MASTER NOTES | SESSIONS 1 TO 4 | NOTES, FIGURES, MCQ BANKS AND EXAM CHEAT CODES

Business Decision Making

Master Notes - Sessions 1 to 4: from the first campus class to the Pilgrim Bank decision

ABOUT THIS DOCUMENT

This is the single, complete set of notes for the Business Decision Making module of the Executive Programme in Business Management (EPBM-32, IIM Calcutta, taught by **Prof. Megha Sharma**, Operations Management), covering the April campus immersion and the three online sessions of 9, 16 and 23 August 2026. It merges the four session documents into one story, told in the order the module told it: what a decision is and what kind of world it lives in (Part I); how to decide well under risk, value information and draw a tree (Part II); how to go from a pile of data to an honest estimate (Part III); and how to turn that estimate into a defensible decision with hypothesis tests and regression (Part IV). Part V is the exam cheat sheet - every formula, number, trap and answer template in one place. Nothing taught in class, shown on a slide, written on the board or worked in Excel has been left out; where the same idea returned in a later session, it appears once, where the story needs it, with a pointer back. Every number was recomputed from the actual files. It is written for someone meeting these ideas for the very first time - in plain words, with everyday examples, and with what each idea means on a manager's desk in a product company and in a service company. The exam is one percent of the reason for this document; the other ninety-nine is using these ideas at work for the rest of your career.

PART I	**The Decision and Its Environment**	Campus immersion, April 2026
PART II	**Deciding Under Risk**	Online session 2, 9 August 2026
PART III	**From Data to an Estimate**	Online session 3, 16 August 2026
PART IV	**From Estimate to Decision**	Online session 4, 23 August 2026
PART V	**Exam Cheat Codes and Closing**	Everything to memorise, on a few pages

Abhinav Roy | EPBM-32 | IIM Calcutta **https://www.linkedin.com/in/abhinavroy07**

A How to Use This Document

This is a long document because the module was a rich one. It is built so that you never have to read it all at once: every part stands alone, every section opens with a plain-language lead, and the same five kinds of box mark the same five kinds of content throughout. Learn the boxes once and you can move through a hundred and sixty pages by their colour.

The five kinds of box

KEY INSIGHT (light teal, teal bar)

The idea to carry away from the section - the sentence you would repeat to a colleague. Also used for definitions to memorise, for worked equations, and for the one-breath summary that closes every part.

CONCEPT TRAP (navy)

A mistake that people in the class actually made, or that the professor warned about explicitly. These are the exam's favourite hunting ground: the option that is numerically right for the wrong reason usually comes straight from a concept trap.

MANAGER'S TAKEAWAY and EXAM TIP (dark teal)

Manager's takeaway: what the idea changes about the way you run a team on Monday morning. Exam tip: the answer template for the kind of question this section produces.

"The professor's own words, or a line from a slide, kept because it is sharper than any paraphrase."

QUOTE BLOCK - the attribution line says where it came from

Beyond the boxes: **figures** (every diagram is drawn from the actual numbers and numbered by part, so Figure 2.4 is the fourth figure in Part II); **tables** (used whenever three things are being compared, so you can scan instead of read); **subheads** in navy, which are the map of a section; and at the end of every part a **Product, Service and the Manager's Desk** section, the **manager's checklist**, the **sharpest lines** of the session, a **one-breath summary**, and a **practice MCQ bank** whose options are deliberately close and whose one-line reasons are the actual revision material.

Three reading paths

If you...	Read	Skip for now
are meeting these ideas for the first time	Everything, in order. The parts are a story; each one assumes the one before. Budget about two hours per part, with the MCQs done cold at the end of each.	Nothing - but you may leave the Bayes tables (2.10) and the assignment models (4.11) for a second pass.
have one evening before the exam	Part V first (the cheat codes), then every one-breath summary, every concept trap, and the MCQ banks with their reasons. Then the worked drills: 2.6-2.10 for Thompson Lumber, 4.7-4.12 for Pilgrim Bank.	The stories and the manager's desk sections - read them after the exam; they are the part you keep.
want to use this at work	Every Manager's Takeaway, the manager's checklists (2.11, 3.13, 4.13) and the Product / Service tables. Then the sections behind whichever takeaway you plan to use.	The MCQ banks and the exam-only drills.

Conventions

Convention	What it means here
Numbers	Thompson Lumber is in rupees, as taught; Pilgrim Bank is in US dollars, as in the case. Every Pilgrim number was recomputed from the raw file and matches the professor's screen and the case slides; where the class rounded ("about $6", "about 2%"), the exact figure is given alongside.
Section references	Part.Section - 3.11 is the eleventh section of Part III. "Section 1.8" always means the same page wherever you are.
Class voices	Batchmates are named where the professor named them in the recording; their contributions are part of what was taught.
What is not here yet	Session 5 (optimisation) and the second campus visit (simulation) had not happened when this document was compiled. They will slot in as Parts VI and VII.

Abbreviations, in case one stops you

Short form	Meaning	Short form	Meaning
EMV / EOL	Expected monetary value / expected opportunity loss (2.6)	H0 / H1	Null and alternate hypothesis (4.2)
EVPI / EVwPI / EVSI	Expected value of perfect information / with perfect information / of sample information (2.8, 2.10)	p-value, alpha	Probability of the sample if the null were true; the tolerance you compare it with (4.4, 4.5)
TLC	Thompson Lumber Company, the Part II worked problem	SLR / MLR	Simple (one x) / multiple (several x) linear regression (4.6)
LP / IP	Linear / integer programming - optimisation with continuous / whole-number decisions (1.4)	ANOVA, Significance F	Analysis of variance; the p-value of the test that all slopes are zero (4.8)
VRP / TSP	Vehicle routing problem / travelling salesman problem (1.5)	R Square	Share of the variation the model explains (4.8)
NP-hard	Problems with no known efficient exact solution (1.6)	SD, CI, CLT, SRS	Standard deviation; confidence interval; central limit theorem; simple random sampling (Part III)
EOQ, CPFR, VMI, POS	Economic order quantity; collaborative planning, forecasting and replenishment; vendor-managed inventory; point of sale (1.7)	KYC, HNI, IVR	Know-your-customer paperwork; high-net-worth individual; the automated phone menu
FD, DICGC	Fixed deposit; the deposit-insurance body behind the Rs. 5 lakh cover (1.3)	SOC, MSS, MDR	Security operations centre; managed security services; managed detection and response - the author's own industry, used in the service examples
SLA, MTBF	Service-level agreement; mean time between failures	MTTR, MTTD, CSAT, UEBA	Mean time to respond / to detect; customer-satisfaction score; user-behaviour baselining tools
FMCG, SKU, USP, P&L;	Fast-moving consumer goods; stock-keeping unit (one product line); unique selling proposition; profit-and-loss account	HBS, LMS, EPBM	Harvard Business School (the Pilgrim case); the programme's learning portal; Executive Programme in Business Management

B The Story of the Module in One Page

The module is a single argument, told over four sessions. Here it is, so that you can see where each part sits before you enter it - and so that no idea ever arrives without a reason.

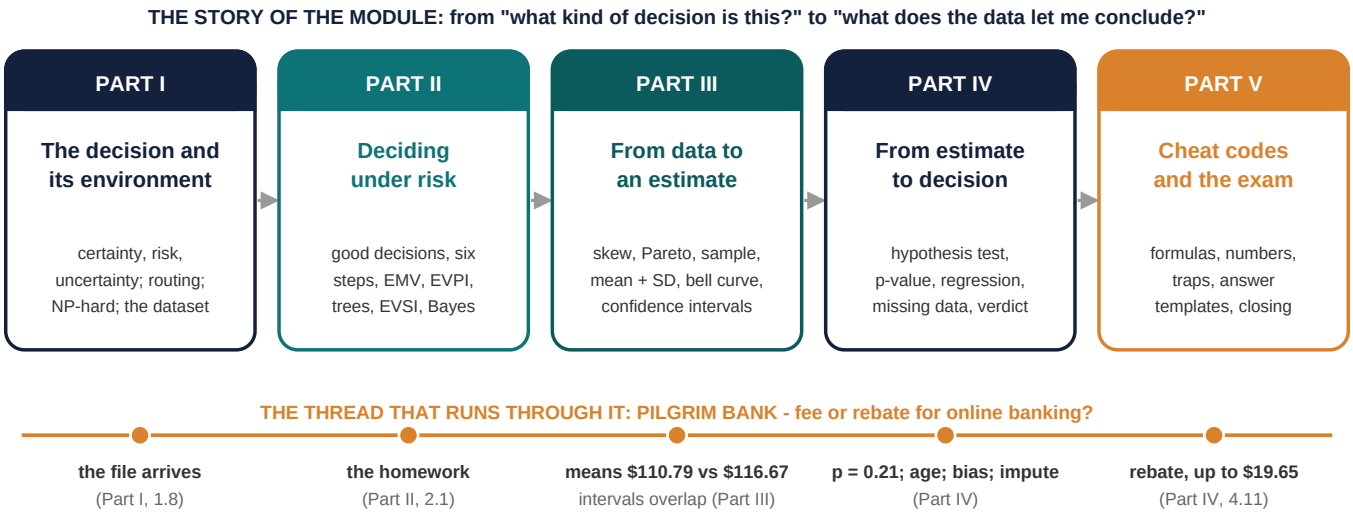


Figure A. The five parts, and the Pilgrim Bank thread that ties them together.

Part I asks what a decision is. A course of action you control meets a state of nature you do not. How much you know about that state of nature sorts every decision into certainty, risk or uncertainty, and the tool follows the environment: optimise, play the long run with EMV, or pick a criterion that matches your temperament. Even "certain" problems turn out to be hard - routing a delivery van is NP-hard - which is why industry runs on heuristics. And a file arrives: 31,634 bank customers and the question "fee or rebate?".

Part II asks what a good decision is, and how to make one under risk. It is judged by its process, not its outcome. Define the problem before you fall for a solution; list every alternative and every state of nature; choose by EMV; check with sensitivity analysis; value information before buying it; and put multi-stage decisions on a tree that reads like a strategy. Bayes shows where the survey probabilities come from. The homework is set: analyse the Pilgrim data and make up your mind.

Part III asks how to trust a number that came from a sample. The bank has five million customers and data on 31,634. A sample must be representative and at least 30; an average must travel with its standard deviation; a confidence interval turns a guess into a range with a stated confidence. The two intervals - online and offline - overlap. The data is inconclusive, and the professor leaves a hint: age is clouding the effect.

Part IV asks how to get from an estimate to a decision. Hypothesis testing - framed so that the thing you want to prove is the alternate; Type I and Type II errors; the p-value as "would this be a miracle?". Regression puts the question in one equation, holds age constant, and finds the online effect. Dropping the customers with missing age biases the sample; imputation with a flag fixes it. The verdict: online customers are $19.65 more profitable like for like. Offer the rebate, cap it there.

Part V is the cheat sheet - and the closing note to the batch.

C Table of Contents

Every part, every section, with its page. Part numbers and section numbers are the same ones used in every cross-reference in the text.

A How to Use This Document 1

B The Story of the Module in One Page 4

C Table of Contents 5

PART I - The Decision and Its Environment (Campus Immersion, April 2026) 7

1.1 The Decisions We Postpone . . . 7
1.2 Course of Action vs State of Nature . . . 8
1.3 The Three Environments: Certainty, Risk, Uncertainty . . . 9
1.4 Certainty at Business Scale . . . 12
1.5 Routing Problems: VRP and TSP . . . 13
1.6 NP-Hard and the Factorial Wall . . . 14
1.7 Also on Campus: Demand, Forecasting, Inventory and the Supply Chain . . . 16
1.8 Decoded: The Pilgrim Bank Dataset . . . 18
1.9 Class Lines, Exam Radar and the Manager's Desk . . . 19
1.10 Practice MCQ Bank with Answers . . . 21

PART II - Deciding Under Risk: Good Decisions, Expected Value, Information and Trees (Online Session 2, 9 August 2026) 32

2.1 Where We Are: The Homework and the Road Ahead . . . 32
2.2 What Makes a Decision Good? . . . 33
2.3 Fall in Love with the Problem . . . 36
2.4 The Six Steps of Decision Making . . . 37
2.5 Environments, Probability and Gut Feeling . . . 39
2.6 EMV: The Thompson Lumber Decision . . . 41
2.7 Sensitivity Analysis . . . 43
2.8 EVPI: The Worth of Perfect Information . . . 45
2.9 Decision Trees . . . 47
2.10 EVSI and Bayes: Is the Survey Worth It? . . . 50
2.11 Product, Service and the Manager's Desk . . . 54
2.12 Practice MCQ Bank with Answers . . . 56

PART III - From Data to an Estimate: Profitability, Sampling, Mean and Standard Deviation, Confidence Intervals (Online Session 3, 16 August 2026) 70

3.1 Where We Are: The Case, the Data and Two Myths . . . 70
3.2 The Profitability Skew: Reading and Rebuilding Exhibit 3 . . . 72

3.3 Letting Go of Customers Without Saying So . . . 74
3.4 Pareto: Customers, Defects and Inventory . . . 76
3.5 Why Retail Is Always on Sale (and What Zara Does Differently) . . . 77
3.6 Population and Sample: Why We Never Have All the Data . . . 79
3.7 Choosing a Sample: Random, Stratified, Convenience - and the Number 30 . . . 80
3.8 What a Sample Is For, and Three Kinds of Data . . . 82
3.9 Mean, Standard Deviation and the Bell Curve . . . 84
3.10 Online vs Offline: The First Attempt . . . 87
3.11 Confidence Intervals: From a Number to a Range . . . 89
3.12 Overlap, Outliers and the Age Cliffhanger . . . 92
3.13 Product, Service and the Manager's Desk . . . 95
3.14 Practice MCQ Bank with Answers . . . 98

PART IV - From Estimate to Decision: Hypothesis Testing, Regression, Missing Data and the Pilgrim Verdict (Online Session 4, 23 August 2026) 118

4.1 Where We Are: Recap and Today's Agenda . . . 118
4.2 Hypothesis Testing: The Delhi Metro Once-Over . . . 119
4.3 Framing the Hypothesis: What You Want to Prove Goes in the Alternate . . . 121
4.4 Two Ways to Be Wrong: Type I, Type II and the Significance Level . . . 122
4.5 The p-value: "Would This Be a Miracle if the Null Were True?" . . . 125
4.6 Regression: A Line Through the Cloud . . . 126
4.7 Running It in Excel, and Reading the Output . . . 128
4.8 Is Age Clouding the Effect? Multiple Linear Regression . . . 131
4.9 The Missing-Data Trap: Testing Whether Dropping Rows Creates Bias . . . 135
4.10 Imputation: Keep Every Row, Add a Flag . . . 136
4.11 The Decision, and the Assignment Worked . . . 139
4.12 Reading Any Regression Output: The Exam Drill . . . 142
4.13 Product, Service and the Manager's Desk . . . 144
4.14 Practice MCQ Bank with Answers . . . 146

PART V - Exam Cheat Codes, and a Note to the Batch 166

5.1 The Cheat Sheet: Every Definition in One Line . . . 166
5.2 Formulas and Numbers to Memorise . . . 168
5.3 The Traps, All in One Place . . . 170
5.4 Answer Templates for Every Kind of Question . . . 172
5.5 The Quick Lookup: If the Question Says X, the Tool Is Y . . . 173
5.6 A Note to the Batch . . . 174

PART I The Decision and Its Environment (Campus Immersion, April 2026)

The module opened on campus, live, with Prof. Megha Sharma (Operations Management, IIM Calcutta). It did not open with mathematics. It opened with a mirror: the decisions we postpone, and why. From there it built the two ideas every later session rests on - the course of action you control and the state of nature you do not - and used them to sort every decision into one of three environments: certainty, risk and uncertainty. It then spent its longest stretch inside the first environment, where the tools are optimisation, and showed why even "certain" problems like routing a delivery van defeat supercomputers. The session ended with a file uploaded to the LMS: the Pilgrim Bank dataset, which becomes the thread of Parts II, III and IV.

By the end of this part you will be able to	Sections in this part
- Explain why we postpone decisions, and separate what you want from what is right - Split any decision into courses of action (yours) and states of nature (the world's) - Classify a decision as certainty, risk or uncertainty, and name the tool for each - Recognise linear and integer programming problems at work - product mix, rosters, SLA windows - Explain VRP, TSP and why NP-hard problems force industry onto heuristics - Tell demand from sales data, and explain the Bass model, EOQ, safety stock, the bullwhip effect and CPFR - Read the Pilgrim Bank file and state its headline numbers	1.1 The Decisions We Postpone 1.2 Course of Action vs State of Nature 1.3 The Three Environments: Certainty, Risk, Uncertainty 1.4 Certainty at Business Scale 1.5 Routing Problems: VRP and TSP 1.6 NP-Hard and the Factorial Wall 1.7 Also on Campus: Demand, Forecasting, Inventory and the Supply Chain 1.8 Decoded: The Pilgrim Bank Dataset 1.9 Class Lines, Exam Radar and the Manager's Desk 1.10 Practice MCQ Bank with Answers

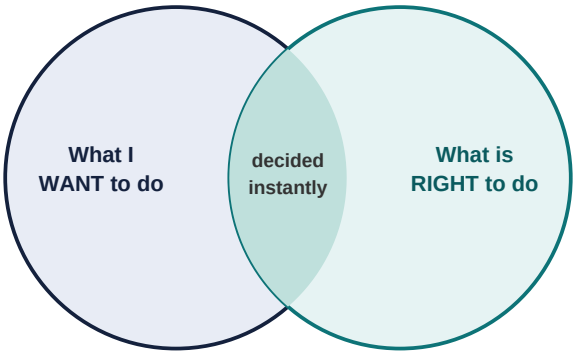
1.1 The Decisions We Postpone

The session did not open with mathematics. It opened with a mirror.

Think of the friend who is a travel enthusiast. The moment a family holiday is mentioned, work is set aside and the tickets are booked - even if the trip is six months away. Why? Because for that person the decision itself is a source of instant gratification. Now flip the character: a couch potato whose family is pushing for the same holiday. The trip keeps getting postponed - "next month", "after this project", "let us see". Same decision, opposite behaviour.

The professor's point was sharp: **we postpone decisions where the right course of action is not the one we want to take.** The family's best alternative is to travel; the couch potato's preference is to stay home. The two are in conflict, so the decision is deferred - in the quiet hope that someday, what I want to do becomes what is right to do. Audit your own pending decisions and most will fit this pattern.

The travel enthusiast books the family holiday at once; the couch potato says "next month". Same decision.



Where the two circles do not overlap is where decisions get postponed - the holiday, the old production line, the loss-making client.

Figure 1.1. The gap between what you want and what is right is where postponed decisions live.

> *"Making a decision and executing it - those are two different things."*
>
> **Opening framing, Session 1**

The same pattern at work. A plant head keeps postponing the shutdown of an old but "emotionally important" production line even though the numbers say retire it. A services delivery head keeps deferring the exit of a loss-making client because the logo looks good on the slide. A product manager delays killing a feature the team loves but users ignore. In every case the analysis is done - the conflict between preference and optimality is what stalls execution.

> **KEY INSIGHT**
>
> Postponement is rarely a data problem - it is a conflict problem. When "what I want" and "what is right" diverge, we delay and hope reality changes. The discipline of decision science is to separate preference from optimality, decide, and then execute. (Part II returns to this with the professor's two-minute writing habit, and the FIFA-match-versus-assignment test.)

1.2 Course of Action vs State of Nature

Before any framework, the everyday picture: a farmer chooses which crop to sow this season. That choice is entirely his. Whether the monsoon arrives on time is entirely not. Every decision problem in this module - however mathematical it becomes - is just these two forces meeting.

Formally, every decision has two building blocks. The **course of action** is your choice, picked from a list of alternatives - fully under your control. The **state of nature** is what the world does after you choose - beyond your control. You decided to invest in TCS stock: that is your course of action. Whether it rises or falls on Monday is the state of nature. The outcome is produced jointly by the two.

The two building blocks, in every kind of job

Context	Course of action (you control)	State of nature (you do not)
Everyday life	Which crop the farmer sows; which stock you buy	Monsoon timing; Monday's market move
Production / plant	Batch size, machine loading, shift plan for the week	Machine breakdown; raw-material quality variation
Operations / SOC	Weekend analyst roster; maintenance window chosen	Actual alert volume; whether an attacker strikes that night
Product company	Launch price and build quantity of the new model	Competitor's counter-launch; actual customer demand
Service company	Airline's overbooking level; bench size an IT firm carries	Number of no-shows; which deals actually convert

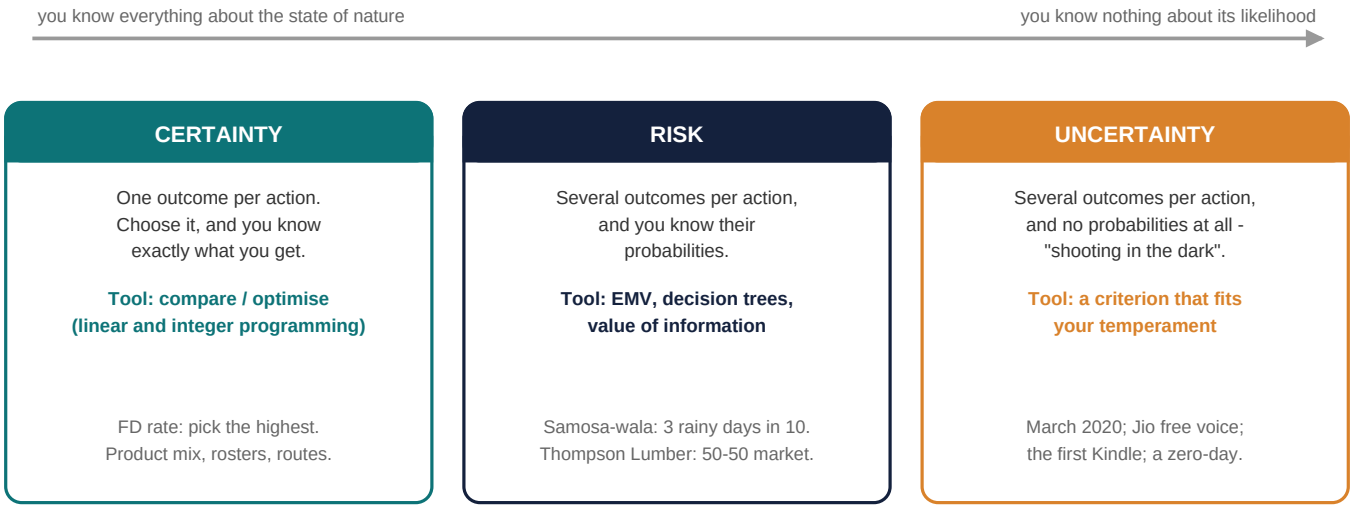
The entire classification of decision environments - certainty, risk, uncertainty - is simply a statement about how much you know about the states of nature before you choose. That is the ladder the next section climbs.

CONCEPT TRAP

Do not judge the quality of a decision by its outcome alone. A sound course of action can meet an unlucky state of nature and still produce a bad result - and a reckless choice can get lucky. Decisions are evaluated on the process and information available at the time of choosing, not on hindsight. Part II makes this the first idea of Session 2.

1.3 The Three Environments: Certainty, Risk, Uncertainty

One question sorts every decision you will ever face: when you pick a course of action, what do you know about the states of nature? Know the outcome exactly - certainty. Know the possible outcomes and their probabilities - risk. Know the possible outcomes but not the probabilities - uncertainty. Session 1 lived in the first environment; the other two are where the module is headed, so all three are built out below with the same discipline: analogy first, then the tool, then where you meet it at work.



Classify the environment FIRST; the tool follows the environment. Most real decisions start on the right and are pushed left with proxies.

Figure 1.2. The three environments on one line. The tool follows the environment.

A. Decision making under certainty - the outcome is guaranteed

Definition: for every course of action there is exactly one state of nature - choose the action and the outcome is known in advance. Nothing is left to chance; the only work is to compare known outcomes and pick the best.

The fixed deposit story from class. You have Rs. 1 lakh. The mandate is rigid: a fixed deposit with a nationalised bank, tenure one year, objective maximise the maturity amount. Open the banks' websites, compare one-year FD rates, pick the highest. Are you guaranteed that amount? Yes. One action, one known outcome - certainty in its purest form.

Class debate 1 - "but inflation eats the real rate". A batchmate argued that even in an FD the real return falls with inflation and tax. The professor's response is the deeper lesson: the objective was fixed as maximise the nominal maturity value of an FD. Within that frame, inflation sits outside the model. Change the objective - maximise real post-tax return across all instruments - and you have changed the entire decision problem, possibly out of certainty altogether. **The objective function is sovereign.**

Class debate 2 - why Rs. 1 lakh and not Rs. 10 lakh? A deliberately chosen number. Indian bank deposits are insured only up to Rs. 5 lakh per depositor per bank (DICGC). At Rs. 1 lakh even a bank failure cannot touch your money - the outcome stays certain. Above Rs. 5 lakh a small probability of loss creeps in, and you have quietly slipped from certainty into risk. Certainty is a fragile assumption; respect its boundary conditions.

Where you meet certainty at work	The decision
Production	Product-mix planning: machine hours, material availability and profit per unit are all known - decide how many tables vs chairs to make this week (classic linear programming)
Operations	Fitting preventive-maintenance windows inside a known downtime budget; assigning 41 warehouses to cities once demand estimates are treated as known
Product-based	Choosing among fixed supplier quotations for a component - prices, lead times and quantities are all on paper before you sign
Service-based	Committed annuity contracts: a managed-services or IT firm with signed multi-year deals knows next quarter's top line from those contracts with near-certainty - a point a batchmate raised in class

B. Decision making under risk - probabilities are known

The everyday picture: the samosa stall. The samosa-wala outside your office does not know whether tomorrow will be sunny or rainy - but ten years of standing there have taught him that it rains about 3 days in 10, and that rain cuts his sales from 200 samosas to 80. He cannot control the weather, but he knows its probabilities. That is decision making under risk: multiple states of nature per action, with probabilities known from history or data.

The tool is **Expected Monetary Value (EMV)**: for each course of action, multiply every outcome by its probability and add them up; choose the action with the best EMV. Watch it work. Each samosa costs Rs. 5 to make and sells at Rs. 10; unsold samosas are waste.

Course of action	Sunny day (p = 0.7)	Rainy day (p = 0.3)	EMV
Fry 200 samosas	Sells 200: profit Rs. 1,000	Sells 80: revenue 800 - cost 1,000 = Rs. -200	0.7 x 1000 + 0.3 x (-200) = Rs. 640
Fry 100 samosas	Sells 100: profit Rs. 500	Sells 80: 800 - 500 = Rs. 300	0.7 x 500 + 0.3 x 300 = Rs. 440

EMV says fry 200 - accept occasional rainy-day losses for the larger average gain. Note what the vendor is really doing: playing the long run. EMV is the right criterion for repeated decisions; for one-off bet-the-company decisions, risk appetite matters too. Decision trees are the same logic drawn as a picture, and EVPI (expected value of perfect information) asks how much you should pay a fortune-teller - i.e., for a better forecast. Part II builds all three on one worked problem.

Where you meet risk at work	The decision
Production	How much safety stock to hold when demand history gives a distribution; whether to schedule overtime given the known win-rate of a pending order
Operations	Spare-parts stocking from known failure rates (MTBF data); sizing the SOC night shift from the historical distribution of alert volumes
Product-based	Diwali build quantity from five years of festive sell-through data; pharma go/no-go using known phase-wise clinical success rates
Service-based	Airline overbooking calibrated to historical no-show rates; insurance pricing itself - the purest risk business there is; a restaurant ordering perishables

C. Decision making under uncertainty - probabilities are unknown

The everyday picture: March 2020. When the first lockdown hit, every business knew the possible states - demand collapses, demand shifts online, recovery in three months or in two years - but nobody had probabilities, because nothing like it had happened in living memory. No history, no distribution, no EMV. That is uncertainty: the states of nature are visible, their likelihoods are not. The same is true for a first-time entrepreneur, or Jio in 2016 deciding to launch voice free - there was no historical data on what incumbents and customers would do to a move nobody had made before.

When probabilities are absent, criteria replace calculations. Consider a manufacturer sizing a plant for a brand-new product with no demand history (figures in Rs. lakh):

Course of action	Demand turns out HIGH	Demand turns out LOW
Build LARGE plant	+100	-40
Build SMALL plant	+40	+10

Four classical criteria, four temperaments:

Criterion	Logic	Working	Verdict
Maximax (optimist)	Best of the best outcomes	Large: 100 vs Small: 40	LARGE
Maximin (pessimist / Wald)	Best of the worst outcomes	Large: -40 vs Small: +10	SMALL
Minimax Regret (Savage)	Minimise the worst "if only I had..." regret	If HIGH, best is 100: regrets are Large 0, Small 60. If LOW, best is 10: regrets are Large 50, Small 0. Worst regrets: Large 50 vs Small 60	LARGE
Laplace (equal likelihood)	No information = treat states as equally likely	Large: (100-40)/2 = 30 vs Small: (40+10)/2 = 25	LARGE

A fifth, the **Hurwicz criterion**, blends the extremes: score = alpha x best outcome + (1 - alpha) x worst outcome, where alpha is your "optimism index" between 0 and 1. Notice that the criteria disagree - maximin says small, the rest say large. That is not a flaw; it is the honest admission that under uncertainty the "right" answer depends on

the decision-maker's temperament and the organisation's capacity to survive the worst case.

Where you meet uncertainty at work	The decision
Production	Capacity for a first-ever product line (EV components in 2019, say) - no demand history exists anywhere
Operations	Entering a brand-new geography with no order data; in security operations, defending against zero-day exploits - known ransomware families come with frequency data (risk), a zero-day comes with none (uncertainty)
Product-based	A category-creating launch: Jio's free voice in 2016, or the first smartwatch - competitor and customer response had no precedent
Service-based	Pricing a never-before service offering; a consulting firm building a practice for a regulation that has not yet been notified

THE KNIGHT DISTINCTION (EXAM GOLD)

Economist Frank Knight (1921) drew the line the module stands on: RISK is measurable randomness - probabilities exist and can be priced, which is why insurance is a business. UNCERTAINTY is unmeasurable - no probabilities, so judgment criteria replace expected-value arithmetic. One line to remember: risk can be insured; uncertainty must be managed.

The spectrum on one slide

Environment	What you know about states of nature	The tool	One-line example
Certainty	Exactly one outcome per action	Compare / optimise (LP, IP)	Pick the highest FD rate
Risk	All outcomes + their probabilities	EMV, decision trees, EVPI	Samosa-wala frying for a 30% chance of rain
Uncertainty	Outcomes visible, probabilities unknown	Maximax, maximin, minimax regret, Laplace, Hurwicz	Plant size for a product with no history

EXAM TIP

For any case question, classify the environment FIRST - the tool follows the environment. Write down (1) the objective function, (2) the courses of action, (3) the states of nature and what is known about them. One-to-one mapping = certainty = optimise. Known probabilities = risk = EMV / decision tree. Unknown probabilities = uncertainty = pick a criterion and justify the temperament.

1.4 Certainty at Business Scale

Do businesses really face decisions under certainty? The class produced three live examples - all from batchmates' own workplaces - and each turns out to be classic optimisation territory.

Example 1: Committed revenue - a known top line

A batchmate made the simplest and sharpest point: with signed contracts in hand, we already know what the top line is going to be. An IT or managed-services firm with multi-year annuity deals can state next quarter's contracted revenue with near-certainty - the states of nature have been pinned down by signatures. Planning delivery capacity against that committed book is decision making under certainty; it is the new-business pipeline that lives under risk.

Example 2: SLA-driven maintenance scheduling

Another batchmate supporting production systems described a contractual SLA: systems must stay up 99.9% of the year. The permitted downtime is therefore a known, fixed budget - roughly 8.8 hours a year. Scheduling preventive-maintenance windows inside that budget is certainty-world optimisation: allowed downtime, maintenance durations and patch cycles are all known in advance, and the task is to find a schedule that fits. This is scheduling via **integer programming**: the decisions are yes/no and whole-number in nature - do the maintenance in this window or not - which is exactly what integer programmes model.

Example 3: Which warehouse serves which city?

A third batchmate runs e-commerce warehouse operations - 41 warehouses across India. How is it decided which warehouse serves which city? Demand first: cities are categorised into tiers (1, 2, 3), order density is mapped area by area, warehouses are networked to stores accordingly, and proper route planning follows. Once demand estimates are treated as known, warehouse-to-city assignment is a deterministic allocation problem - large, but solvable by optimisation, not intuition.

> **KEY INSIGHT**
>
> The Indian analogy for integer programming is the IPL fixture: every team must play every other team, grounds have availability windows, broadcasters have slot constraints - and a match either happens in a slot or it does not. Whole-number, yes/no decisions under known constraints are the signature of integer programming; linear programming is its cousin where quantities may be fractional (litres of blend, tonnes of alloy). Crew rostering at an airline and exam timetabling at a university are the same mathematics wearing different uniforms.

1.5 Routing Problems: VRP and TSP

Before the definitions, the picture every Indian city already knows: the Mumbai dabbawala. Thousands of tiffins, collected from homes, sorted at hubs, delivered to exact desks before lunch, and returned - with near-Six Sigma accuracy and no software. Every morning they are solving, by hand, the two most famous problems in operations research.

The Vehicle Routing Problem (VRP)

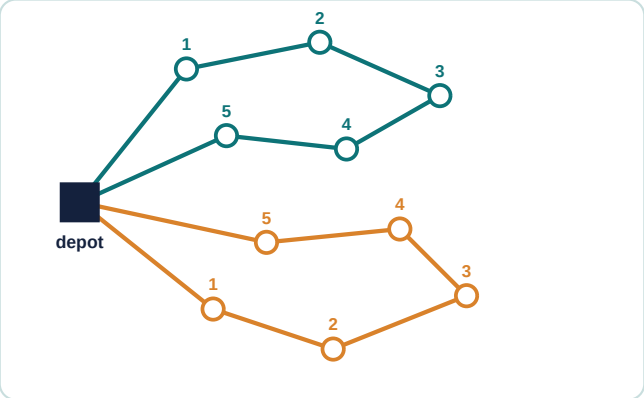
Imagine managing an Amazon delivery station. Every morning, parcels must go out. Delivery associates report to you; each has a capacity constraint - one person cannot carry everything. Two decisions must be made together: (a) **Allocation** - which parcels go to which associate, respecting capacity; and (b) **Sequencing** - the order in which each associate visits the addresses: go to this place first, then this, then this, and return with the receipts.

That combined problem is the Vehicle Routing Problem. The objective is usually to minimise total cost or distance; an equally valid alternative is minimising average delivery time - the sooner people get their parcels, the happier they are. The choice of objective, once again, defines the problem.

The Travelling Salesman Problem (TSP)

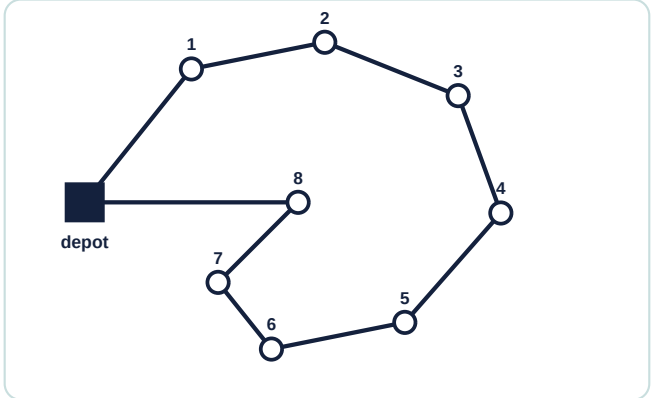
Strip VRP down to a single vehicle and you get the TSP: a salesman starts at a depot or hub, must visit each location exactly once, and return to the depot at the minimum possible cost. It sounds almost trivial. It is not - the next section shows exactly why.

VEHICLE ROUTING PROBLEM: many vehicles, limited capacity



Allocate parcels to associates AND sequence each route.
Objective: least total distance (or fastest average delivery).

TRAVELLING SALESMAN PROBLEM: one vehicle, no capacity limit



Visit every point exactly once and return, at least cost.
TSP is the special case of VRP with a single vehicle.

Figure 1.3. Left: two capacity-limited associates, each with an allocation and a sequence. Right: one vehicle visiting every point once - the TSP.

The same problem in different uniforms

Where you meet it	The routing problem in disguise
Food delivery (service)	Swiggy / Zomato batching three orders onto one rider and sequencing the drops before the biryani goes cold - VRP with time windows
Manufacturing inbound	Maruti-style "milk runs": one truck looping through multiple component suppliers on a fixed circuit instead of each supplier shipping separately - VRP on the supply side
Shop floor (production)	A PCB drilling machine visiting thousands of hole positions in minimum head-travel time - a pure TSP hiding inside a factory
Field service	A Tata Play or Jio technician's day plan: eight installation visits, one van, best sequence - a TSP every single morning
Banking operations	ATM cash-replenishment vans covering machines across a city under security and time constraints - VRP with a very expensive cargo

KEY INSIGHT

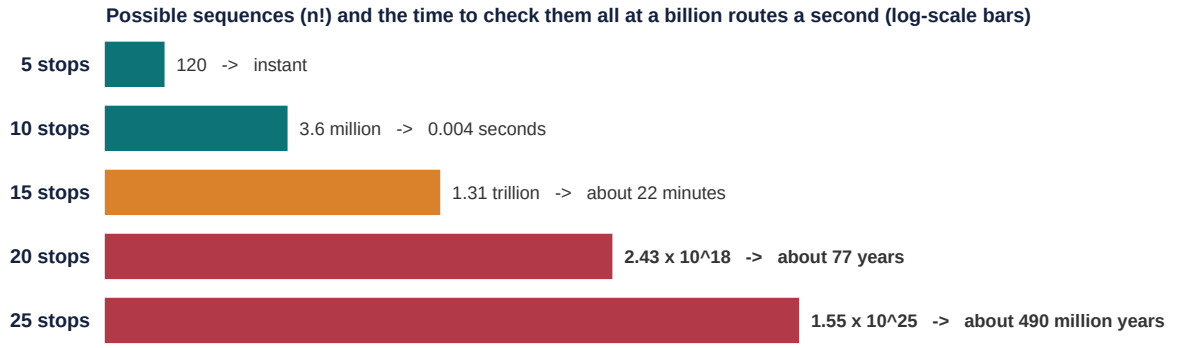
VRP subsumes TSP: the TSP is simply a VRP with one vehicle and no capacity limit. Master the TSP's logic and the VRP is the same logic plus an allocation layer. In exam language: TSP is a special case of VRP.

1.6 NP-Hard and the Factorial Wall

Why is a problem a child can state - "visit these ten houses in the cheapest order" - one that defeats supercomputers? Because the only guaranteed way to find the best route is to enumerate every possible route, and the number of routes explodes factorially.

Factorial refresher: the factorial of a number n, written n!, is the product of all whole numbers from n down to 1. So 5! = 5 x 4 x 3 x 2 x 1 = 120. With 10 delivery stops there are 10! ways to sequence them. For a closed tour that

starts and ends at a fixed depot, the count of distinct circular routes works out to (n-1)! - and half of that again if clockwise and anti-clockwise cost the same. The refinement barely matters: every version grows explosively.



Brute force is the only guaranteed route to the exact optimum, and it hits a wall at about twenty stops. Industry uses heuristics instead.

Figure 1.4. The factorial wall. Five extra stops turn minutes into decades.

Stops (n)	Possible sequences (n!)	Time to check all, at 1 billion routes/second
5	120	Instant
10	36,28,800 (~3.6 million)	0.004 seconds
15	~1.31 x 10^12 (1.31 trillion)	~22 minutes
20	~2.43 x 10^18	~77 years
25	~1.55 x 10^25	~490 million years

Both VRP and TSP belong to the class of problems called **NP-hard**. For such problems, no known method guarantees the optimal answer quickly as size grows - even the fastest computer in the world may effectively take forever, because brute-force enumeration is the only certain route to the true optimum. The family is larger than routing: airline crew pairing, university exam timetabling and factory job-shop scheduling are all NP-hard cousins - which is why your workplace "roster headache" genuinely is mathematically hard.

CONCEPT TRAP

In class shorthand, NP was expanded as "non-polynomial". The formally correct expansion is "non-deterministic polynomial" - NP-hard means at least as hard as the hardest problems in that class. For the exam, the substance is what matters: solution effort explodes with problem size, exact enumeration becomes infeasible, and no efficient exact algorithm is known.

So what does industry do? It stops chasing perfection. Amazon, Delhivery, Swiggy and every logistics player run **heuristics** - clever rules and approximation algorithms that find a very good route in seconds rather than the perfect route in centuries. A near-optimal answer now beats a perfect answer never. That trade - optimality sacrificed for tractability - is one of the deepest managerial lessons of operations research.

MANAGER'S TAKEAWAY

When a planning problem at work is "impossible to get perfect" - the roster, the route plan, the maintenance calendar - it probably is, mathematically. Stop paying for perfection. Buy or build a good heuristic, measure how close it gets, and spend the saved effort on the inputs (demand, capacity, constraints) that the model depends on.

1.7 Also on Campus: Demand, Forecasting, Inventory and the Supply Chain

The campus immersion did not stop at the decision framework. The same days covered the operations side of the module - where the numbers that feed a decision come from, and what goes wrong with them. When the professor asked the Session 2 class to reconstruct the campus days, these were the ideas people remembered. They are collected here in one place, in plain words, because each of them returns later in the story.

Idea	What it means, in plain words	Product company	Service company
Hawthorne effect	People change their behaviour when they know they are being watched. So data collected during a pilot or an audit may not show normal behaviour - it shows best behaviour.	The production line runs faster the week the auditors are on the floor; do not size the plan on that week.	Analysts close tickets faster the week the new dashboard goes live; the real run-rate appears a month later.
Qualitative vs quantitative models	Qualitative models structure judgement (frameworks, checklists, expert panels); quantitative models use data, statistics and optimisation. Both are legitimate, and the module uses both.	A launch go/no-go checklist, and the EMV table behind it.	A client-risk scorecard, and the regression behind the score.
Sales data vs demand data (censored data)	Sales equal demand only until you run out of stock. After a stock-out, sales undercount demand: the customers who wanted the product and found the shelf empty leave no trace. Data cut off like this is called censored. Uncensoring means estimating the demand you never saw - from the selling rate before the stock-out, or from days when stock never ran out.	A Diwali variant sold out on day 3 of a 10-day season; the season's demand was far more than the sales the system recorded.	A fully booked clinic slot or a full training batch shows you capacity, not demand - the people turned away are invisible.
Markdown vs event discount	A markdown is a permanent price cut once demand has fallen short - Ross selling last season's Michael Kors bag cheaply for good; the previous-generation phone repricing permanently when the new one launches. Zara's Black Friday price is an event, not a markdown: it is back to full price on Sunday.	Last season's stock, marked down to clear the shelf (Part III, 3.5).	An off-season hotel rate is an event; a permanently cut list price for a service nobody buys is a markdown.
Forecasting a category that does not exist yet: the Bass model	When a product has no history - the first e-reader, the first smartwatch - you cannot extrapolate sales. The Bass model forecasts adoption from two kinds of buyers: innovators, who adopt on their own (moved by advertising and media; coefficient p), and imitators, who adopt because others have (moved by word of mouth; coefficient q). Together they trace the S-shaped adoption curve. It needs demand data, not censored sales data. Rogers' diffusion categories - innovators, early adopters, early majority, late majority, laggards - describe the same curve in five groups.	Kindle, smartphones, electric two-wheelers: heavy advertising early to reach innovators, then word of mouth carries the growth.	UPI payments, OTT subscriptions, a new managed-security service: the first customers are won by sales effort, the rest by reference.

Idea	What it means, in plain words	Product company	Service company
Why hold inventory - and how much (EOQ and safety stock)	Inventory is a hedge against uncertainty in demand and in supply lead time, and a way to buy or make in economical batches. The economic order quantity (EOQ) is the order size that balances the cost of placing orders (many small orders = high ordering cost) against the cost of holding stock (few large orders = high holding cost); total cost is U-shaped and EOQ sits at its bottom. Safety stock is the extra kept on top for the uncertainty the formula cannot see. Too little means stock-outs and lost sales; too much means capital, space and markdowns.	FMCG holds stock (cheap, steady demand); fashion cannot afford to (markdowns); perishables hold almost none; technology fears obsolescence; luxury keeps stock scarce on purpose; car makers build to order.	The service version of safety stock is surge capacity: the extra analysts, beds or delivery riders kept for the night the alert volume, admissions or orders spike.
The bullwhip effect	A small change in what consumers buy becomes a large swing in what the factory is asked to make, because every tier - retailer, distributor, manufacturer - forecasts on its own, adds its own safety buffer, orders in batches, over-orders during promotions and inflates orders during shortages. It is an information problem disguised as an inventory problem.	COVID: shoppers bought a little more toilet paper once; retailers ordered three times as much, distributors five times, factories ran round the clock, then orders collapsed. Car makers cancelled chip orders in 2020, chip makers moved capacity elsewhere, and 2021 had no chips.	A managed-service client doubles its ticket forecast for one quarter; the provider hires for it; the tickets never come. Capacity planned on amplified signals is the service bullwhip.
CPFR - collaborative planning, forecasting and replenishment	The fix for the bullwhip: retailer and supplier build one shared forecast instead of two independent ones, share point-of-sale data, let replenishment follow actual sales, and shorten the order cycle. Related practices: vendor-managed inventory (the supplier manages the retailer's stock) and real-time data sharing of the kind Walmart does with its suppliers.	A consumer-goods maker and a retail chain planning the Diwali promotion on one shared number.	A managed-security provider and its client agreeing one forecast of monitored assets and alert volume, reviewed quarterly, instead of each guessing the other's next move.
"With chaos comes revenue; with order comes profit"	Disruption - a new technology, a rival failing, a pandemic - creates revenue opportunities for whoever moves fast. Only operational discipline turns that revenue into profit. Chaos is where the opportunity is; order is where the money is kept.	Jio's launch threw the telecom market into chaos and revenue followed; profit needed the operational order underneath.	Food delivery grew on chaos and is still building the order that makes it profitable.
Fear of consequences - why managers freeze	Loss aversion, blame culture and regret aversion make managers postpone or copy precedent. The cure the module offers is to judge the quality of the decision, not the outcome - the first idea of Part II.	The launch nobody wants to own in case it fails.	The client exit nobody wants to sign in case the revenue dip is blamed on them.
Kinds of data	Cross-sectional (many things at one time), time series (one thing over time) and panel (both) - defined fully in Section 3.8, where the Pilgrim file turns out to be cross-sectional.	Sales of every SKU this month; one SKU over five years; every SKU over five years.	Every client's ticket volume this quarter; one client over three years; all clients over three years.

HOW THESE CONNECT TO THE DECISION FRAMEWORK

Every one of these is about the quality of the inputs to a decision. Censored sales data understates the demand state of nature; the Hawthorne effect flatters a pilot; the Bass model turns "no history" (uncertainty) into a forecast you can put a probability on (risk); EOQ and safety stock are the certainty-and-risk halves of one inventory decision; the bullwhip is what happens when every decision-maker in a chain uses a different, distorted estimate of the same state of nature - and CPFR is the agreement to use one. Better data, better collaboration, better judgement: that was the campus day's summary, and it is the module's.

MANAGER'S TAKEAWAY

Before you forecast, ask whether the history you are extrapolating was censored by stock-outs or flattered by observation. Before you hold inventory - or surge capacity - name the uncertainty it is hedging. And before you react to a big swing in orders from the next tier, ask whether the end customer actually changed behaviour or whether you are looking at the whip.

1.8 Decoded: The Pilgrim Bank Dataset

The mystery file Detailed_Data.xls uploaded after the campus session is no mystery: it is the dataset of "Pilgrim Bank (A): Customer Profitability" - the classic Harvard Business School case (Frances Frei and Dennis Campbell) used worldwide to teach data-driven decision making. It is a snapshot of 31,634 retail customers of a US bank for the year 1999 (the "9" prefix in every column = 1999). The decision on the table: online banking has arrived - should the bank charge fees for it, keep it free, or pay customers to adopt it?

Column decode

Column	Meaning	Values in the file	What to notice
ID	Customer identifier	1 to 31,634	One row per customer.
9Profit	Annual profit the bank earns from this customer (USD)	-221 to +2,071	Customer 1 gave $21; customer 2 cost the bank $6.
9Online	Uses online banking? 1 = yes, 0 = no	3,854 ones (12.2%), 27,780 zeros	A binary variable - it takes only two values.
9Age	Age bucket 1 (youngest) to 7 (oldest)	Missing for 8,289 customers	Blanks are people who did not give a date of birth on the bank form, before KYC made it compulsory.
9Inc	Income bucket 1 (lowest) to 9 (highest)	Missing for 8,261 customers	Slabs, not amounts: people answer "5 to 10 lakh" far more readily than an exact figure - so the bank traded precision for completeness, and still got a quarter of the rows blank.
9Tenure	Years as a customer	0 to 40+ (mean 10.2 years)	Complete for everyone, because it comes from the bank's own database.
9District	Geographic district code	1100 / 1200 / 1300	Like a PIN code; only three values.

The full file also has columns prefixed 0 (the year 2000); the professor deleted them for the statistics sessions to keep the sheet clean, and her worked sheets - "Profitability Curve", "Point Estimate", "Online-offline" - were shared after Session 3. Your rows may be sorted differently from hers; the row order changes no result.

The headline numbers (computed from this exact file) - a preview of Parts III and IV

1. Customer profitability is brutally skewed. Mean profit is $111.50 per customer, but the median is just $9. Around 46% of customers are unprofitable, and the top 10% of customers deliver roughly 70% of total profit (about $3.5 million in all). The average customer is a statistical fiction - the numerator-only story hides a wildly unequal denominator. (Section 3.2 draws the curve.)

2. The naive comparison flatters online banking - weakly. Online customers average $116.67 against $110.79 for offline: a difference of $5.88. Tested properly, the p-value is about 0.21 - far above the 0.05 bar. On raw means, you cannot claim online customers are more profitable. (Sections 3.10 and 4.7.)

3. The twist: online users are a different crowd. They are visibly younger (average age bucket 3.35 vs 4.17), slightly higher-income (5.98 vs 5.42) and have shorter tenure (8.65 vs 10.37 years). Youth and short tenure both depress profit - so the raw comparison mixes the effect of the channel with the effect of who chooses the channel. This is **confounding**. (Section 4.8.)

4. Control for the confounders and the story flips. In a regression of profit on online status plus age, income and tenure, the online coefficient becomes significant and positive - about $14 to $20 depending on what is held constant. Yet the model explains under 6% of profit variation: even "significant" effects can be small levers in a noisy world. (Sections 4.10 and 4.11.)

5. The missing-data trap. Age and income are missing for roughly 26% of customers - and not at random: customers with missing age average $73 profit versus $125 for the rest. Silently deleting incomplete rows quietly changes who you are studying. How to handle missingness is itself a decision. (Sections 4.9 and 4.10.)

EXAM TIP

The Pilgrim ladder is the answer template for any "does X drive profit?" question: (1) compare group means, (2) test statistical significance - a difference is not evidence until the p-value says so, (3) control for confounders before concluding, (4) check whether missing data is biasing the sample. Quote the skew line: mean $111.50, median $9 - the median tells the story the mean hides.

WHY THIS FILE, AFTER THAT SESSION

Session 1 built decisions where outcomes are certain and the mathematics is optimisation. Pilgrim Bank is the other half of the module: a decision - price online banking - where outcomes are noisy and the evidence is data. The bridge concept is the state of nature: under certainty you know it; under risk you estimate its probabilities from exactly this kind of data. Part II sets the homework; Parts III and IV answer it.

1.9 Class Lines, Exam Radar and the Manager's Desk

Four ready-to-fire contributions connecting the session to the managed-security world, the concepts most likely to be tested, and what the session means for the way you run a team.

Class lines

#	The line, and the idea behind it
1	**On certainty vs states of nature:** "In managed security we sign 99.9% availability SLAs. Our maintenance calendar is a decision under certainty - but the adversary is a state of nature. Are SLAs essentially promises made under certainty about a world that is uncertain?"
2	**On risk vs uncertainty:** "Our threat landscape splits exactly along today's framework - commodity ransomware comes with frequency data, so defending it is decision-making under risk; a zero-day comes with no probabilities at all, which is decision-making under uncertainty. Should security budgets be split along that line - EMV logic for the known, maximin logic for the unknown?"
3	**On NP-hard routing:** "Alert triage in a Security Operations Centre is a vehicle routing problem - alerts are parcels, analysts are capacity-constrained vehicles, and the sequencing space explodes factorially. Is that why the industry buys orchestration tools - as heuristics for an NP-hard problem we cannot solve exactly?"
4	**On Pilgrim Bank:** "In our data 46% of customers are unprofitable and the top decile funds everyone else. In managed services the same skew exists across client contracts - should churn and retention strategy start from the profit distribution rather than the logo count?"

Product, service and the manager's desk

Concept	Product-based company	Service-based company
Want vs right	Killing the feature the team loves but nobody uses.	Exiting the loss-making client whose logo looks good on the slide.
Course of action vs state of nature	Build quantity (yours) against actual demand (the world's).	Analyst roster (yours) against tonight's alert volume (the world's).
Certainty - optimise	Product mix from known machine hours and margins (LP); which supplier quote to accept.	Maintenance windows inside a known SLA downtime budget (IP); delivery capacity against the contracted book.
Risk - EMV	Diwali build quantity from five years of sell-through data.	Overbooking level from historical no-show rates; night-shift size from the alert-volume distribution.
Uncertainty - criteria	Capacity for a category-creating product; the first smartwatch.	Pricing a never-before service; defending a zero-day.
VRP / TSP / heuristics	PCB drilling head path; milk-run truck circuits.	Field-technician day plans; ATM cash vans; alert triage as routing.
The data file	Which SKUs actually earn the margin - mean vs median.	Which contracts fund the SOC - the profit distribution, not the logo count.

Exam radar

Concept	One-line answer to carry into the exam
Course of action	The alternative you choose; fully under your control
State of nature	What the world does after you choose; beyond your control
Decision under certainty	Each action maps to exactly one state of nature; outcome guaranteed; tool = comparison / LP / IP
Decision under risk	Several states per action with KNOWN probabilities; tool = EMV, decision trees, EVPI

Concept	One-line answer to carry into the exam
EMV	Sum of (payoff x probability) across states; choose the action with the best EMV
Decision under uncertainty	States known, probabilities UNKNOWN; tools = maximax, maximin (Wald), minimax regret (Savage), Laplace, Hurwicz
Knight distinction	Risk is measurable and insurable; uncertainty is unmeasurable and must be managed by judgment
FD example logic	Fixed objective (max maturity), compare known rates, pick highest; Rs. 1 lakh keeps it inside the Rs. 5 lakh deposit-insurance limit
Integer programming	Optimisation with whole-number / yes-no decisions; scheduling, assignment, rostering, IPL fixtures
TSP	Start at depot, visit every location exactly once, return, at minimum cost
VRP	Allocate parcels to capacity-constrained vehicles AND sequence each route; TSP is its special case
NP-hard	No known efficient exact algorithm; brute force is factorial; industry uses heuristics
Censored data	Sales undercount demand after a stock-out; uncensor before forecasting
Bass model	Adoption of a new category = innovators (p, media-driven) + imitators (q, word of mouth); S-curve; needs demand data
EOQ / safety stock	Order size that balances ordering and holding cost (U-shaped total cost); extra stock for the uncertainty the formula cannot see
Bullwhip effect	Small consumer changes amplified up the chain by independent forecasts, batching, promotions and shortage gaming; an information problem
CPFR	One shared retailer-supplier forecast with shared point-of-sale data; the fix for the bullwhip
Hawthorne effect	Observed people behave differently; pilot data can flatter
Pilgrim Bank lesson	Means mislead; test significance, control confounders, respect missing data

ONE-BREATH SUMMARY

Every decision = a course of action you control meeting a state of nature you do not. Know the outcome exactly - certainty: optimise (FD rates, committed revenue, SLA maintenance windows, warehouse assignment, the VRP/TSP routing family whose factorial explosion makes them NP-hard and forces industry onto heuristics). Know the probabilities - risk: play the long run with EMV, like the samosa-wala frying for a 30% chance of rain. Know neither - uncertainty: choose a criterion that matches your temperament, maximax to maximin. And the shared Excel is HBS's Pilgrim Bank: 31,634 customers teaching that averages lie, significance and confounders decide, and missing data is never innocent.

1.10 Practice MCQ Bank with Answers

Fifty-three questions drawn from the campus days, in roughly the order the ideas were taught. The options are meant to be close, so read the one-line reason under each answer even when you got it right.

Q1. According to the professor, we postpone decisions mainly when:

(a) The data is incomplete

(b) The right course of action is not the one we want to take

(c) The deadline is far away

(d) Nobody else is involved

Answer: (b) - The travel enthusiast and the couch potato face the same holiday decision; only the want-vs-right gap differs.

Q2. "Making a decision and executing it are two different things." The stalled step in the plant-head, delivery-head and product-manager examples is:

(a) The analysis

(b) The execution, because preference conflicts with optimality

(c) The data collection

(d) The approval from finance

Answer: (b) - In every case the numbers already point one way; the conflict with what they want to do is what defers action.

Q3. The course of action is:

(a) What the world does after you choose

(b) The alternative you choose, fully under your control

(c) The outcome

(d) The probability of success

Answer: (b) - The farmer picks the crop; the monsoon is the state of nature.

Q4. Buying TCS stock on Friday; the price moves on Monday. The Monday move is:

(a) A course of action

(b) A state of nature

(c) An objective function

(d) A payoff table

Answer: (b) - Beyond your control after you choose.

Q5. For a SOC, which pairing is correct?

(a) Roster = state of nature; alert volume = course of action

(b) Roster = course of action; alert volume = state of nature

(c) Both are courses of action

(d) Both are states of nature

Answer: (b) - You choose the roster; the night's alerts and attackers are the world's move.

Q6. The classification into certainty, risk and uncertainty is a statement about:

(a) How many alternatives you have

(b) How much you know about the states of nature before you choose

(c) The size of the payoff

(d) The number of decision-makers

Answer: (b) - One outcome per action, known probabilities, or unknown probabilities.

Q7. Decision making under certainty means:

(a) The decision is easy

(b) Each course of action leads to exactly one known outcome

(c) Probabilities are known

(d) There is only one alternative

Answer: (b) - Nothing is left to chance; the work is comparing known outcomes.

Q8. In the FD example, the objective was fixed as:

(a) Maximise real post-tax return across all instruments

(b) Maximise the nominal maturity value of a one-year FD at a nationalised bank

(c) Minimise risk

(d) Maximise liquidity

Answer: (b) - Change the objective and you change the whole decision problem - the objective function is sovereign.

Q9. Why did the professor use Rs. 1 lakh and not Rs. 10 lakh in the FD example?

(a) Rounding

(b) Deposit insurance covers up to Rs. 5 lakh per depositor per bank, so Rs. 1 lakh keeps the outcome certain even if the bank fails

(c) Banks do not accept larger FDs

(d) Tax rules change above Rs. 1 lakh

Answer: (b) - Above Rs. 5 lakh a small probability of loss creeps in and the decision slides into risk.

Q10. Planning delivery capacity against a signed multi-year contract book is decision making under:

(a) Certainty

(b) Risk

(c) Uncertainty

(d) None - it is not a decision

Answer: (a) - The signatures pin down the states of nature; the new-business pipeline is what lives under risk.

Q11. The tool for decisions under certainty is:

(a) EMV

(b) Decision trees

(c) Comparison and optimisation - linear and integer programming

(d) Maximin

Answer: (c) - Known outcomes, so the job is to optimise.

Q12. The samosa-wala knows it rains about 3 days in 10 and that rain cuts sales to 80. His decision is under:

(a) Certainty

(b) Risk

(c) Uncertainty

(d) Ignorance

Answer: (b) - Several states per action, with probabilities known from ten years of history.

Q13. Frying 200 samosas at Rs. 5 cost and Rs. 10 price, with p(sunny) = 0.7 and rainy-day sales of 80, gives an EMV of:

(a) Rs. 1,000

(b) Rs. 640

(c) Rs. 440

(d) Rs. 800

Answer: (b) - 0.7 x 1,000 + 0.3 x (-200) = 640. Frying 100 gives 440.

Q14. EMV is the right criterion mainly for:

(a) One-off bet-the-company decisions

(b) Repeated decisions, because it plays the long run

(c) Decisions under certainty

(d) Decisions with no data

Answer: (b) - For one-off decisions risk appetite matters too; the vendor is playing many days.

Q15. March 2020 for most businesses was a decision environment of:

(a) Certainty

(b) Risk

(c) Uncertainty

(d) Optimisation

Answer: (c) - The possible states were visible; nobody had probabilities for them.

Q16. Jio launching free voice in 2016 is an example of:

(a) Risk, from telecom history

(b) Uncertainty - no precedent for how incumbents and customers would react

(c) Certainty

(d) A TSP

Answer: (b) - A category-changing move with no historical data.

Q17. Large plant: +100 / -40; small plant: +40 / +10 (Rs. lakh). The maximin criterion chooses:

(a) Large

(b) Small

(c) Either

(d) Neither

Answer: (b) - Best of the worst outcomes: small's +10 beats large's -40.

Q18. With the same table, maximax chooses:

(a) Large

(b) Small

(c) Either

(d) Neither

Answer: (a) - Best of the best: 100 beats 40.

Q19. With the same table, minimax regret chooses:

(a) Large, with worst regret 50

(b) Small, with worst regret 60

(c) Small, with worst regret 50

(d) Large, with worst regret 60

Answer: (a) - If HIGH the regrets are 0 and 60; if LOW they are 50 and 0; large's worst regret (50) is the smaller.

Q20. The Laplace criterion:

(a) Chooses the best worst case

(b) Treats the states as equally likely when there is no information

(c) Uses an optimism index

(d) Requires known probabilities

Answer: (b) - Large: (100 - 40)/2 = 30 beats small: 25.

Q21. The Hurwicz criterion:

(a) Ignores the worst case

(b) Blends best and worst outcomes with an optimism index alpha between 0 and 1

(c) Is the same as Laplace

(d) Only works with three states

Answer: (b) - Alpha = 1 is maximax, alpha = 0 is maximin.

Q22. The criteria disagreed (maximin said small, the rest said large). The professor's view of this is:

(a) A flaw in the method

(b) An honest admission that under uncertainty the answer depends on temperament and the capacity to survive the worst case

(c) A sign the data is wrong

(d) A reason to use EMV instead

Answer: (b) - There are no probabilities to settle the argument.

Q23. Frank Knight's distinction says:

(a) Risk and uncertainty are synonyms

(b) Risk is measurable and can be priced (insured); uncertainty is unmeasurable and must be managed

(c) Uncertainty can be insured

(d) Risk has no probabilities

Answer: (b) - Insurance is a business precisely because risk is measurable.

Q24. A zero-day exploit, as opposed to a known ransomware family, is a decision environment of:

(a) Risk

(b) Uncertainty

(c) Certainty

(d) Both

Answer: (b) - Known families come with frequency data; a zero-day comes with none.

Q25. The exam tip for any case question is to:

(a) Compute EMV immediately

(b) Classify the environment first - the tool follows the environment

(c) Draw a tree first

(d) List the states of nature last

Answer: (b) - Write the objective, the courses of action, and what is known about the states of nature.

Q26. A 99.9% availability SLA allows roughly how much downtime a year?

(a) 88 hours

(b) 8.8 hours

(c) 0.88 hours

(d) 24 hours

Answer: (b) - 0.1% of 8,760 hours; scheduling maintenance inside it is integer programming.

Q27. Integer programming is distinguished from linear programming by:

(a) Larger numbers

(b) Whole-number or yes/no decision variables

(c) Unknown probabilities

(d) More constraints

Answer: (b) - A match happens in a slot or it does not; litres of a blend may be fractional.

Q28. The IPL fixture was the professor's analogy for:

(a) The TSP

(b) Integer programming

(c) EMV

(d) Uncertainty

Answer: (b) - Every team plays every other, grounds and broadcast slots have constraints, decisions are yes/no.

Q29. Assigning 41 warehouses to cities once demand estimates are treated as known is:

(a) A decision under uncertainty

(b) A deterministic allocation problem solvable by optimisation

(c) A TSP

(d) A hypothesis test

Answer: (b) - Large, but certainty-world once demand is taken as known.

Q30. The two decisions inside a Vehicle Routing Problem are:

(a) Pricing and promotion

(b) Allocation of parcels to capacity-constrained vehicles, and sequencing of each route

(c) Forecasting and inventory

(d) Hiring and training

Answer: (b) - Which parcels go with whom, and in what order.

Q31. A valid alternative objective for a VRP, mentioned in class, is:

(a) Maximise the number of vehicles

(b) Minimise average delivery time

(c) Maximise distance

(d) Minimise the number of parcels

Answer: (b) - The choice of objective defines the problem.

Q32. A PCB drilling machine visiting thousands of hole positions is:

(a) A VRP

(b) A TSP hiding inside a factory

(c) An integer programme only

(d) A decision under uncertainty

Answer: (b) - One head, every hole once, minimum travel.

Q33. The number of ways to sequence 10 stops is 10!, which equals about:

(a) 3,600

(b) 3.6 million

(c) 3.6 billion

(d) 3.6 trillion

Answer: (b) - 36,28,800; 15 stops is about 1.3 trillion.

Q34. Checking every route for 20 stops at a billion routes a second would take about:

(a) 22 minutes

(b) 77 years

(c) 490 million years

(d) 4 seconds

Answer: (b) - 20! is about 2.43 x 10^18; 25 stops is about 490 million years.

Q35. For a closed tour from a fixed depot, the number of distinct circular routes is:

(a) n!

(b) (n - 1)!, halved again if direction does not matter

(c) n squared

(d) 2 to the power n

Answer: (b) - The refinement barely matters - every version grows explosively.

Q36. NP-hard means, in substance:

(a) The problem cannot be solved at all

(b) No known method guarantees the optimum quickly as size grows; brute force explodes

(c) The problem is non-polynomial in name only

(d) The problem has no constraints

Answer: (b) - The formally correct expansion is non-deterministic polynomial; the class shorthand was non-polynomial.

Q37. Because routing is NP-hard, industry:

(a) Buys bigger computers

(b) Uses heuristics that find a very good route in seconds rather than the perfect route in centuries

(c) Stops planning routes

(d) Uses EMV

Answer: (b) - Optimality sacrificed for tractability - a deep managerial lesson.

Q38. Which of these is NOT an NP-hard cousin mentioned in the notes?

(a) Airline crew pairing

(b) University exam timetabling

(c) Factory job-shop scheduling

(d) Comparing FD rates

Answer: (d) - Comparing FD rates is a trivial comparison; the other three explode combinatorially.

Q39. The shared file Detailed_Data.xls is the dataset of:

(a) A Mumbai dabbawala study

(b) Pilgrim Bank (A): Customer Profitability, the HBS case by Frei and Campbell

(c) An IPL fixture

(d) A samosa stall

Answer: (b) - 31,634 US retail bank customers in 1999.

Q40. The "9" prefix in the column names (9Profit, 9Online) means:

(a) Nine variables

(b) The year 1999

(c) Category 9

(d) The ninth sample

Answer: (b) - The full file also has columns prefixed 0 for the year 2000.

Q41. Mean profit is $111.50 and the median is $9. This tells you:

(a) A data error

(b) Profit is heavily skewed - a few very profitable customers pull the mean up

(c) Most customers earn about $111

(d) The SD is small

Answer: (b) - The median tells the story the mean hides.

Q42. 9Tenure is complete for every customer because:

(a) Customers always answer it

(b) It comes from the bank's own database - the customer did not have to supply it

(c) It was imputed

(d) It is a binary variable

Answer: (b) - Age and income came from forms and are missing for about a quarter of customers.

Q43. Roughly what share of Pilgrim customers were unprofitable in 1999?

(a) 12%

(b) 26%

(c) 46%

(d) 70%

Answer: (c) - 14,580 of 31,634; the top 10% deliver about 70% of the profit.

Q44. The four-step Pilgrim ladder for any "does X drive profit?" question is:

(a) Plot, colour, present, decide

(b) Compare group means; test significance; control for confounders; check missing-data bias

(c) Sort, sum, average, chart

(d) Survey, tree, EMV, decide

Answer: (b) - Parts III and IV climb exactly this ladder.

Q45. Sales data equals demand data:

(a) Always

(b) Only until a stock-out; after that sales undercount demand and the data is censored

(c) Only for services

(d) Only when prices are constant

Answer: (b) - The customers who found the shelf empty leave no trace; uncensoring estimates the demand you never saw.

Q46. The Hawthorne effect warns that:

(a) Machines slow down when observed

(b) People behave differently when they know they are being watched, so pilot or audit data can flatter

(c) Surveys always understate demand

(d) Forecasts improve with observation

Answer: (b) - Size the plan on normal weeks, not the week the auditors were on the floor.

Q47. In the Bass model, imitators are moved by:

(a) Advertising and media (coefficient p)

(b) Word of mouth - they adopt because others have (coefficient q)

(c) Price cuts only

(d) Government mandates

Answer: (b) - Innovators respond to external influence (p); imitators to internal influence (q); together they trace the S-curve.

Q48. The Bass model is used when:

(a) Five years of sales history exist

(b) The product category does not yet exist, so there is no history to extrapolate

(c) Demand is certain

(d) The data is a panel

Answer: (b) - Kindle, the first smartwatch - forecast adoption from the two kinds of buyers instead.

Q49. The economic order quantity balances:

(a) Price and quality

(b) Ordering cost against holding cost - total cost is U-shaped and EOQ sits at its bottom

(c) Demand and supply

(d) Revenue and profit

Answer: (b) - Many small orders cost ordering effort; few large orders cost holding space and capital.

Q50. Safety stock exists to:

(a) Reduce the order quantity

(b) Cover the uncertainty in demand and lead time that the EOQ formula cannot see

(c) Increase markdowns

(d) Replace forecasting

Answer: (b) - The service version is surge capacity - extra analysts or riders for the night volume spikes.

Q51. The bullwhip effect is best described as:

(a) A pricing problem

(b) Small changes in consumer demand amplified up the supply chain - an information problem disguised as an inventory problem

(c) A quality problem

(d) A transport problem

Answer: (b) - Independent forecasts, safety buffers, order batching, promotions and shortage gaming each add to the swing.

Q52. CPFR fixes the bullwhip by:

(a) Holding more inventory at every tier

(b) Building one shared retailer-supplier forecast on shared point-of-sale data, with shorter order cycles

(c) Cancelling promotions

(d) Ordering in larger batches

Answer: (b) - Collaborative planning, forecasting and replenishment; vendor-managed inventory is a cousin.

Q53. "With chaos comes revenue; with order comes profit" means:

(a) Disorderly companies are more profitable

(b) Disruption creates the revenue opportunity; operational discipline is what turns it into profit

(c) Profit requires chaos

(d) Order reduces revenue

Answer: (b) - Jio's launch created chaos and revenue; the profit needed operational order underneath.

PART II Deciding Under Risk: Good Decisions, Expected Value, Information and Trees (Online Session 2, 9 August 2026)

Four months after the campus immersion, the module went online. Session 2 picked up where Part I left off and built the analytical heart of the module: what a good decision actually is, why we must define the problem before we fall for a solution, the six-step decision process, Expected Monetary Value (EMV), sensitivity analysis, the value of perfect and imperfect information, decision trees, and how survey probabilities are revised using Bayes' theorem. Everything is taught on one deliberately small problem - Thompson Lumber's backyard sheds - so that the method, not the arithmetic, stays in focus. The Pilgrim Bank file from Part I becomes the homework here, and the last two sections of this part are the tools you will need to answer it.

By the end of this part you will be able to	Sections in this part
- Judge a decision by its process, not its outcome - and explain why the reverse promotes mediocrity - Write a problem statement with no solution hiding inside it, and tell a symptom from a root cause - Run the six steps: problem, alternatives, states of nature, payoffs, model, decide - Compute EMV and EOL, and read a sensitivity range as robust or sensitive - Value information before buying it: EVPI for certainty, EVSI for a survey - Draw a decision tree left to right, solve it right to left, and read the strategy it prescribes - Explain how Bayes' theorem turns a vendor's accuracy into the probabilities on the tree	2.1 Where We Are: The Homework and the Road Ahead 2.2 What Makes a Decision Good? 2.3 Fall in Love with the Problem 2.4 The Six Steps of Decision Making 2.5 Environments, Probability and Gut Feeling 2.6 EMV: The Thompson Lumber Decision 2.7 Sensitivity Analysis 2.8 EVPI: The Worth of Perfect Information 2.9 Decision Trees 2.10 EVSI and Bayes: Is the Survey Worth It? 2.11 Product, Service and the Manager's Desk 2.12 Practice MCQ Bank with Answers

2.1 Where We Are: The Homework and the Road Ahead

Four months had passed since the campus immersion in April, so the professor opened by asking the class to reconstruct what we had done. Part I holds the full picture; this section keeps only what the class added that day - the homework, and the logic of the two extreme options it offered.

The class's own recap named the Hawthorne effect, qualitative versus quantitative models, the Bass model, demand versus sales data and markdowns - the campus ideas collected in Section 1.7 - alongside the decision environments of Section 1.3 and the postponed decisions of Section 1.1.

The road ahead, as the professor drew it: four online sessions - this one on decision analysis, two on statistics, one on simulation - and a second campus visit. Together with the campus immersion, that is her view of "a good enough grasp of the subject". Parts III and IV of this document are the two statistics sessions.

The homework: Pilgrim Bank - charge a fee, or pay a rebate?

The dataset uploaded to the LMS after the campus session (decoded in Section 1.8) is the assignment for the next class. Put yourself in a US bank in the early 2000s. There is almost no internet penetration, laptops are rare, and smartphones do not exist. Online banking is a brand-new service, and the bank has to choose between two deliberately extreme options: **charge customers a fee** for using it, or **pay customers a rebate** for using it.

Why would a bank ever pay you to use a service? Because some new services **reduce the bank's own cost**. Many of us have not walked into our branch for years, and the bank loves that - the physical branch, the tellers, the paper, all cost far more than a server. If online banking cuts cost, the bank should promote it, even pay for migration. But if customers keep visiting the branch anyway, online banking is an **additional facility on top of the old cost** - and the bank should recover that from the people who use it. Two opposite logics; the data decides.

> **WHAT THE PROFESSOR ACTUALLY ASKED FOR**
>
> Analyse the data over the week and make up your mind either way. Uploading it to an AI assistant is fine - you would do that at work too. What matters is that you can support your recommendation, defend it, and show why it beats the other side's argument. "Not just because my bot said so."

2.2 What Makes a Decision Good?

The professor asked the class a simple question - what makes a decision good or bad? - and then spent the next hour dismantling our first answer. That dismantling is the most important idea of the session, so it is worth following step by step.

The class's answers - and the one with a hidden bias

The board filled up quickly: a decision is good if it produces a **desirable outcome**; if it **accounts for the available courses of action**; if it passes an **ethical** test; if it is **effective**; and if it serves the **bigger goal or vision**. Four of these are fine. The first one is a trap, and the professor set it up with a second question.

"How does it feel to be wrong?"

The chat filled with the obvious answers: disappointing, frustrating, bad - especially if you put in real effort - unhappy, ignored, a lesson learned. Then came the twist.

> *"Being wrong feels exactly like being right. Because you do not know you are wrong."*
>
> **Prof. Megha Sharma, Session 2 (an idea made famous by the writer Kathryn Schulz)**

Every feeling the class listed is the feeling of *finding out* that you were wrong. Until that moment, being wrong feels confident, optimistic, settled - identical to being right. That is the catch in every decision under risk or uncertainty: the moment you feel most sure is not evidence that you are right, and the chances of being wrong are often high.

The outcome trap: judging with an answer key you never had

Now look again at "a good decision produces a desirable outcome". When you decide, the outcome has not happened yet. You launch a product; the market reacts weeks or months later. If you later grade the decision by the outcome, you are grading it with **information that was not available at the moment of decision**. That is hindsight, and hindsight is not analysis.

It is worse than unfair - it **promotes mediocrity**. Think about what happens when something unusual lands on your desk. The first instinct is to look for a precedent: has this happened before, and what did we do? We copy it, because it is a safety net - if it fails, "we did what worked last time". Why do we behave this way? Because organisations reward good outcomes and punish bad ones: miss the target and the promotion goes to someone else, or you inherit the territory nobody wants. So everyone plays it safe, nobody reaches for the option that would have realised their full potential, and the organisation settles in the middle - it does not fail, but it never wins big either.

The definition to carry with you

> *"A good decision is based on logic, considers all available data and possible alternatives, and is obtained through rational analysis of these data and alternatives."*
>
> **Session 2 slide 2 - the working definition of the module**

Test	Question to ask about the decision
All courses of action	Were all the possible alternatives considered - including the ones the decision-maker did not like?
All possible outcomes	Were all the states of nature that could follow each alternative listed, not just the comfortable ones?
Rational use of data	Was all the available information used, analysed rationally - or was inconvenient data quietly ignored?

If a decision passes these three tests, it is a good decision **whatever the outcome**. A good decision can meet bad luck and lose money. A careless decision can get lucky and look brilliant. It is still a bad decision - it just happened to work.

CONCEPT TRAP

Outcome and decision quality are two different things joined by chance. Under certainty (Session 1) the link is one-to-one, so judging by outcome is harmless. Under risk and uncertainty, the same good decision can produce different outcomes on different days - so the outcome tells you about luck, not about the decision.

Alcoa, 1987-2000: what rewarding the process looks like

Aluminium was once more precious than gold. Napoleon III served his most honoured guests on aluminium cutlery and the less important ones on gold and silver, because separating the metal from its ore was so difficult that it was almost a craft. The company that changed that was Alcoa - the Aluminum Company of America - founded in 1888 by Charles Martin Hall and his partners after Hall found a practical way to refine aluminium. It is the company behind the modern beverage can.

In 1987 Alcoa hired a new chief executive, **Paul O'Neill**, who came from government and bureaucracy rather than from industry. In those days a new CEO would address the analysts, the analysts would judge the speech - market share, growth, cost-cutting - and then run to the nearest phone booth (no mobiles in 1987) to tell their firms whether to buy or sell. O'Neill's entire speech was about **worker safety**: fewer accidents, fewer injuries. The analysts ran to the phones and told everyone to sell. The stock fell before he had done anything at all - the market had graded him on a first "outcome".

He stayed until 2000. In the class's telling, annual profit went from about USD 200 million in 1987 to about USD 3 billion by the time he retired, and the company's market value multiplied. What did safety have to do with it? In the 1980s American unions resisted almost everything - sharing data, changing how work was done. But no union can refuse when management says "we are changing this process because the old one is hazardous to your members". Safety became the door through which every process in the company was re-examined.

The mechanism was the rule he set: every accident had to reach him within a fixed number of hours - not just the fact of it, but a plan that would ensure the same accident never happened again. One occurrence is a diagnosis; a repeat is unacceptable. He fired a high-performing executive - not because an accident happened on his plant, but

because he **tried to cover it up**. The cover-up is the outcome mindset in action: "I must not be seen as the plant with accidents." Once that mindset takes hold, people fudge the process, and when the process is fudged, luck - not skill - decides your results.

> **KEY INSIGHT**
>
> Organisations that do well over the long run do not reward or punish the outcome of decisions made under risk or uncertainty, because chance played a part. They audit the process: did you look at all the data, all the options, all the outcomes? A robust process does not win every time - but it wins more often than it loses, over the long run. An outcome-driven culture wins or loses by luck.

"But we do live by outcomes" - the captain and the entrepreneur

A batchmate pushed back honestly: a cricket captain choosing the next bowler gets feedback in six balls, but someone who quits a job to start a business may learn only ten years later whether it was right. How do you detach from the outcome in either case? The professor answered in two parts.

Short feedback loops are noisy. Six balls tell you very little. M.S. Dhoni failed in his first matches; Sourav Ganguly kept backing him - not out of favouritism, but because his judgement said this player had potential - and gave him enough opportunities for that judgement to be tested properly. Tune out the noise.

Long feedback loops are mostly regret. Ten years later, what you "know" about the road not taken is speculation. You chose option A and are unhappy; you assume option B would have been greener. You will never know - there is no parallel universe. What you can control is whether you decided well with what you knew. That is why the session introduced a personal tool for exactly this kind of decision - the decision tree, and before it, the power of writing.

The power of writing

Everyone knows the endless argument in the head - with the boss, a subordinate, the household help - where you rehearse what they will say and what you will say back, and win every round. In reality you rarely win those, and you come out more frustrated. The brain is too clever and too cluttered to sort itself out. So the professor's habit, used at home as much as at work, is simple: **write it down**. Not pages - that becomes a burden and you will stop. Two minutes, a few bullet points.

Writing works for a reason. Because of a self-serving bias, you will not write down something that makes you look bad even in your own eyes - so the pettiness and the emotional clutter fall away on their own. What is left is the real issue. Then: try to synthesise a **problem statement** (not a solution); **empathise** - why would this person be behaving this way?; and list the **options you actually have**. Sometimes you discover you have no other option, and that acceptance alone removes the stress. Sometimes you find better options than the one you were stuck on.

And you see why you were postponing. We delay decisions when **what we want to do and what is available or right to do are two different things**. Nobody postpones the FIFA World Cup match at 1 a.m. on a work night. Everybody postpones the assignment due the next morning. Same night, same office the next day - the difference is that in one case want and right were aligned, and in the other they were not. Once you can see that gap on paper, you either do it anyway or consciously decide not to. Either way, it stops living in your head.

MANAGER'S TAKEAWAY

In your next appraisal conversation or project post-mortem, ask "was the process sound with the information available at the time?" before you ask "did it work?". Teams graded on outcomes learn to hide accidents. Teams graded on process learn to report them - and that is the only way the same accident stops recurring.

2.3 Fall in Love with the Problem

What kind of people do we like at work? The ones who talk about solutions - the can-do crowd. The professor agreed, and then pointed at the danger inside that instinct.

Once we have found a solution - built a prototype, done a proof of concept, spent weeks on it - we are **emotionally invested**. When early signs appear that the solution does not fit the problem, we do not drop the solution. We quietly reshape the problem so that the solution fits. On the slide deck the problem now looks different and the solution looks perfect. On the ground, the problem is exactly what it always was, and nothing gets fixed.

For a routine, well-defined problem this does not arise. A gear breaks; replace the gear; the machine runs. That is a certainty scenario, and the analytical decision-making approach of this session applies directly. The trouble is the **unstructured problem** - the one that keeps coming back no matter what you fix. That is usually a problem where you have been treating a symptom rather than the root cause, and for it the professor borrowed the first pillar of **design thinking**.

What design thinking asks of you

Design thinking is not a new buzzword - it has been around for decades as a creative problem-solving method. It began in product and service design, it runs quietly inside systems like Toyota's, IT companies made it popular, and some of its most striking results are in public systems in highly constrained places - parts of Africa, for instance - where solutions transplanted from the West or the East simply do not work. It is usually described as a five or six step process, and in the **first three steps you are not allowed to talk about solutions at all**.

Instead you fall in love with the problem: look at it from every stakeholder's point of view, put yourself in their shoes, without judgement and without the hurry to fix anything. Why does this problem occur? What is causing it? What does it do to the people who live with it? Only after that do you look for solutions - and by then, you are far less likely to be defending a pet solution against the evidence.

The firefighting test

Take a moment and estimate: in an average day, what share of your time goes on the routine, urgent, operational tasks of your job, and what share goes on strategic work - the things that are not urgent but will pay off later? The same question applies at home. Paying the bills is urgent; it has a deadline; it does not move you towards any vision. If the split is 90 percent mundane and 10 percent strategic, your vision keeps receding into the future. If it is 50-50, you are in decent shape.

Firefighting belongs to the mundane column even when the fire is on a strategic project. It is not supposed to be daily. If your systems and processes are sound, a fire should be rare, and when it comes it should teach you something - a risk you had not seen and could not have managed earlier. If fires come every day, one of four things is true: your processes are broken; the processes are right but their implementation is broken; your team is not aligned or motivated; or you keep solving a symptom while the root cause sits untouched and sends the next fire.

HOW TO KNOW WHETHER YOU FIXED THE CAUSE OR THE SYMPTOM

After you solve a problem, do you get a significant period of calm - time you can actually spend on strategy? If yes, you probably treated the cause. If the next fire follows the last one without a pause, you treated a symptom. The calm is the test.

Daniel Kish, and the problem that was not "blindness"

The professor played a short video. Daniel Kish was born with retinoblastoma, a cancer of the eye. His right eye was removed at seven months and his left at thirteen months. The first thing he did on waking from the second surgery was climb out of his cot and wander around the intensive care nursery - "probably looking for the one who did this to me". His parents, he says, were pragmatic rather than poetic: they put his freedom first, so he moved out at eighteen, pays taxes, and rides a mountain bike. Today he teaches blind children to ride bikes and play basketball.

How? **Echolocation** - the same trick bats and dolphins use. He clicks his tongue and reads the echoes; a wall, a tree, a moving car each return the sound differently. In the demonstration he moves a panel towards and away from a listener, and even a sighted person with eyes closed can hear it coming closer. Seeing, it turns out, does not have to happen through the eyes.

Before the video, the professor had asked the class: a child with total vision loss wants to play basketball and go mountain biking - what would you do? The answers were an eye transplant (not possible when the eyes have been removed for cancer) and a guide. Both answers came from defining the problem as "**the child cannot see**". That is the symptom. The real problem is that the child **cannot independently move through an unfamiliar space** - does not know where the furniture is, where the water was spilt, what is moving around them. Define the problem that way and echolocation becomes a solution. Define it as "cannot see" and you are stuck waiting for a miracle.

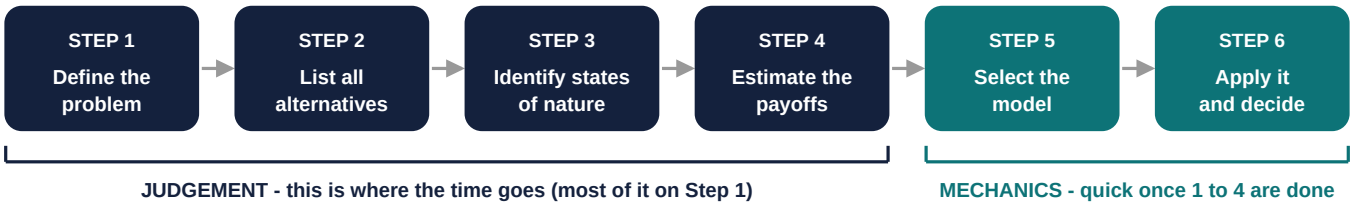
> *"What people say and what the real problem is are often two different things - even when they are trying their best to articulate it."*
>
> **Prof. Megha Sharma, on why step one deserves the most time**

MANAGER'S TAKEAWAY

Before your team builds anything, make them write a problem statement with no solution hiding inside it. "We need a dashboard" is a solution. "Regional heads discover a stock-out three days after it happens" is a problem - and it may have five solutions, of which a dashboard is the most expensive.

2.4 The Six Steps of Decision Making

With the warm-up done, the session turned to the framework that everything else hangs on. Decision making, the professor said, is a simple six-step process - simple to state, and revealing in where the time goes.



Steps 2 and 3 separate what you control (alternatives) from what you do not (states of nature).

Figure 2.1. The six steps. The first four are judgement; the last two are mechanics.

Step	What you do	What the professor stressed
1. Define the problem	State the problem at hand clearly, in writing.	Spend the most time here. Ask: is this the root cause or a symptom? Use the firefighting test. "Solving the right problem roughly beats solving the wrong problem accurately."
2. List the alternatives	Write down every possible course of action - the choices under your control.	List them ALL, including the ones you dislike. Weak options will be eliminated by the process; do not let your bias eliminate them first.
3. Identify the states of nature	List the possible outcomes that could follow - things beyond your control.	Customer reaction could be very bad, mildly bad, neutral, favourable, very favourable. You do not know which; list them anyway.
4. Estimate the payoffs	For each alternative-state combination, estimate the result on your chosen performance measure.	The payoff is whatever the decision is meant to move: profit, revenue, cost, market share. It need not be money. This is project-finance style scenario work.
5. Select the model	Pick a decision theory model.	The right model depends on the environment - certainty, uncertainty or risk (Section 2.5).
6. Apply and decide	Run the model and make the decision.	Mechanical. If steps 1-4 are honest, this takes minutes.

Why "all" alternatives - the rival's product

Nobody likes everybody in their organisation. Suppose the colleague you like least has proposed a new product, and you have been asked to test-market it. You could test it nationally, regionally, or in one small town. Because it is your rival's idea, the national launch will probably never even appear on your list - you are biased before the analysis begins. Now flip it: it is your own idea. The small-town test quietly disappears, because you are sure the idea deserves more. Either way, an alternative has been killed by bias, not by analysis. Listing everything is how you take the bias out. If an option is genuinely weak, the numbers will remove it.

The same skeleton on a shop counter

A kirana store owner has noticed footfall dropping since a quick-commerce dark store opened two streets away. **Problem:** regular customers are shifting part of their basket to app delivery. **Alternatives:** start own home delivery with a delivery boy; list the shop on a delivery platform; do nothing. **States of nature:** the neighbourhood takes to delivery heavily, moderately, or hardly at all. **Payoff:** monthly profit under each combination. That is the entire six-step structure, and it fits on the back of a bill.

CONCEPT TRAP

Alternatives do not have probabilities. A common mistake is to ask "what is the probability of choosing the large plant?". An alternative is your choice - you take it (1) or you do not (0). Only states of nature - the things you cannot control - carry probabilities.

2.5 Environments, Probability and Gut Feeling

Step 5 says "select a model", and the model you can select depends on the environment you are in. The campus session introduced the three environments; this session recapped them and then asked the question every executive quietly wants answered - is gut feeling allowed?

Environment	What you know	Tools	Example
Certainty	Pick an alternative and you know exactly what you get - one outcome per action.	Linear programming, integer programming (operations research).	Rs. 1 lakh in a one-year FD at a public-sector bank: compare the rates, pick the highest. Even evaluating it in hindsight is fair, because action and outcome match one-to-one.
Uncertainty	Several outcomes are possible and you have no probability estimates at all - "shooting in the dark".	The criteria from Session 1 (maximax, maximin, regret-based).	Amazon launching the Kindle when e-readers did not exist; the first smartwatch; Alexa. A pharma molecule may not work, may have side effects, may not get approval, may be beaten by a rival's molecule.
Risk	Several outcomes are possible and you have probability estimates for each.	EMV maximisation or EOL minimisation; decision trees.	Thompson Lumber: the market will be favourable or unfavourable, with a 50-50 estimate.

Where do the probability estimates for a risk decision come from? From the board: **market research** (surveys, focus groups), **expertise** of people in the field, **historical data** and its analysis, and **experience and gut feeling**.

Most decisions start in uncertainty - and we convert them to risk

A batchmate asked the natural question: what about products and situations where there is no data at all? The honest answer is that most of us live in the uncertainty column most of the time - and we convert the problem into risk because risk is far more comfortable to handle. How? With **proxies**. No company launches without some customer testing. Even Google, as cutting-edge as it gets, tests: first within the team that built it, then their family and friends, then all employees, then a beta release. Each stage generates data that turns "we have no idea" into "roughly 60-40". It is still an estimate - but it is a number you can work with.

Is gut feeling allowed?

The class was split - "absolutely not", "it is good", "no, full stop". The professor's answer began with a comparison. Statistics uses a **sample** to draw conclusions about a **population**. Gut feeling does the same thing: the situations you have lived through are a sample of everything that could happen in the world, and from that sample you generalise. The gut of someone on their first day at work and the gut of someone with twenty years in the trade are not the same instrument, because the samples are not the same size.

So using gut feeling in decision making - even analytical decision making - is not bad. What is bad is **stubbornness**: "my gut says so, and I am the boss". Even a boss's gut feeling rests on assumptions, and those assumptions may not hold this time.

Situation	What to do
Gut and evidence agree	The easiest and happiest case. What you thought is right and everything supports it. Proceed.
Gut and evidence disagree - check 1	Ask: is my gut feeling based on an assumption that is not true in this situation? Dig until you can put that assumption on the table. "Yes - I was assuming X, and X does not hold here." Then the evidence makes sense.
Gut and evidence disagree - check 2	You cannot find a faulty assumption. Then ask: am I looking at the right data? The same data can be cut to tell almost any story, some of them contradictory. Which slice are you looking at?

Challenger, 28 January 1986: the right gut, the wrong chart

The Space Shuttle Challenger mission was famous for two reasons: it carried people, and one of them was a civilian schoolteacher. It was also badly delayed, and the delays had become a running joke on American radio - a slow highway "moving like Challenger". Under that pressure, the launch was finally scheduled for a morning with a freezing forecast.

The booster rockets were built by a contractor, Morton Thiokol. Its engineers had argued for months that the seals in the boosters - the O-rings - failed when it was cold, and that the shuttle should not fly at low temperature. Management asked for evidence. The engineers' chart plotted the launches that had suffered O-ring damage against temperature. Damage appeared at low temperatures and at high ones. "This proves nothing," said management. The night before the launch, the engineers again said do not fly. NASA pushed back. The line that has entered management folklore was spoken to the contractor's vice-president: **"Take off your engineering hat and put on your management hat."** The launch was approved. Seventy-three seconds after lift-off the shuttle broke apart, killing all seven on board.

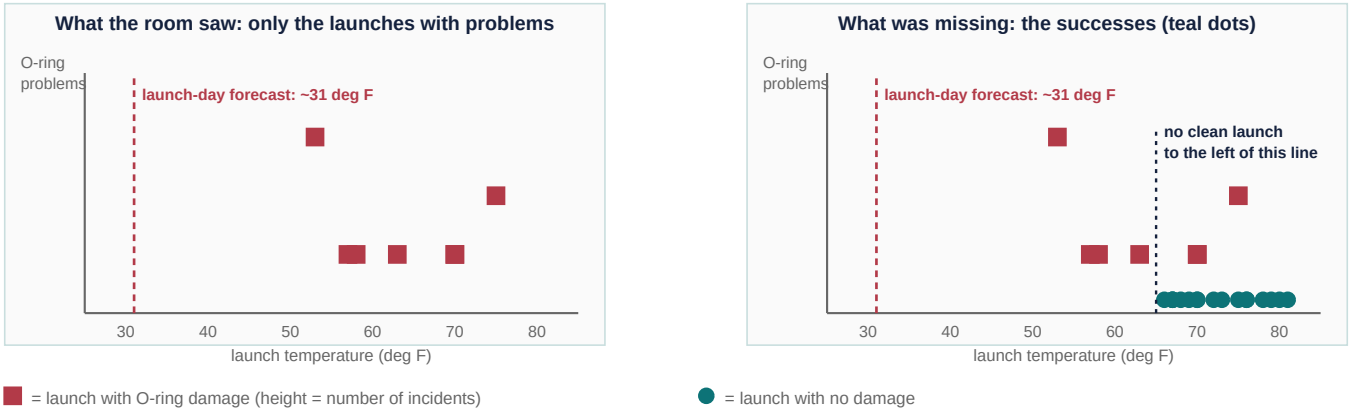


Figure 2.2. The engineers charted failures only. Adding the successful launches shows the pattern: no clean launch had ever happened below the mid-60s, and the forecast was thirty degrees colder.

Who went wrong? The engineers' gut was right, and it was later proved right - the seals had failed in the cold. But they failed the second test from the table above: they never found the data that would **support** their gut. The inquiry found that the chart showed only the failures. It never showed where the **successes** had occurred. Put them on the chart and the story is stark: every clean launch sat at the warm end; below a certain temperature there were no successes at all; and the launch-day forecast was far colder than anything ever attempted. In that zone the probability of failure was effectively one.

This is also the answer to the stories that circulate in every company - "whatever the analysis says, my boss takes the opposite view and turns out right". Those stories, the professor said, are mostly WhatsApp University. What is really happening is that experienced people trust their gut *and ask the right question*: while everyone else looks at

the failures, they ask to see the successes too.

> **KEY INSIGHT**
>
> A chart of only the failures can support any story. Always ask where the successes were. Gut feeling has a real role, especially in senior roles - research supports it - but when it disagrees with the evidence, do the two checks: find the faulty assumption, or find the right data. Never just pull rank.

2.6 EMV: The Thompson Lumber Decision

Every method in the rest of the session is taught on one deliberately small problem, so that the method - not the arithmetic - stays in focus. Meet Thompson Lumber Company (TLC).

TLC has been in the lumber business for over forty years - the professor kept that detail in on purpose, so that nobody dismisses their gut feeling as a fluke. Management is deliberating whether to manufacture and market a new product: backyard sheds. If they launch, they must build capacity, so they have three courses of action: construct a **large plant**, construct a **small plant**, or build **no plant** (do not launch). Whatever they choose, the market will turn out **favourable** (moderate to high demand) or **unfavourable** (low demand, loss-making). The performance measure they chose is **net profit**. A batchmate asked whether the numbers already account for the cost of the plant - yes: every figure below is the net result after costs.

Six steps	TLC
1. Problem	Whether to expand by manufacturing and marketing backyard sheds
2. Alternatives (under TLC's control)	Large plant / small plant / no plant
3. States of nature (beyond control)	Favourable market / unfavourable market
4. Payoffs	Net profit for each of the six combinations - the payoff table below
5. Model	Under risk (once probabilities are attached): EMV
6. Apply	Compute EMV per alternative, pick the best

The payoff table

A payoff table is the standard way to write step 4. **Alternatives go down the rows**, **states of nature go across the columns**, and each cell holds the payoff for that combination. Everybody in the world of decision analysis reads it the same way, which is the point of a convention.

Alternative	Favourable market (Rs.)	Unfavourable market (Rs.)
Construct a large plant	200,000	-180,000
Construct a small plant	100,000	-20,000
Do nothing (no plant)	0	0

As it stands this is a decision under **uncertainty** - nobody has said how likely a favourable market is. TLC's forty years of experience turns it into a decision under **risk**: management believes it is a 50-50 case. Now a model applies.

Expected Monetary Value

EMV sounds technical and is not. For those who do not come from a mathematical background, the professor translated it: **a weighted average of the payoffs, where the weights are the probabilities**. Take each payoff, multiply it by the probability of that state of nature, add them up. Do this for every alternative.

Alternative	Favourable (0.5)	Unfavourable (0.5)	EMV = weighted average
Large plant	200,000	-180,000	0.5 x 200,000 + 0.5 x (-180,000) = **10,000**
Small plant	100,000	-20,000	0.5 x 100,000 + 0.5 x (-20,000) = **40,000**
No plant	0	0	0.5 x 0 + 0.5 x 0 = **0**

The EMV criterion then says: pick the alternative with the **best** EMV. "Best" depends on the payoff. When larger is better - profit, revenue, market share - pick the largest EMV. When smaller is better - cost, risk, defects - pick the smallest. Here the payoff is profit, so TLC should build the **small plant**, with an EMV of Rs. 40,000. (Note the name: "monetary" is just the label. The payoff could be market share; the method is identical.)

"But the small plant never actually earns 40,000!"

The sharpest question of the session came here. If TLC builds the small plant, it will either make 100,000 (favourable) or lose 20,000 (unfavourable). Nobody will ever bank exactly 40,000. So where does this number come from, and how can a number that will never occur be the yardstick for the decision?

The professor's answer is the key to understanding every expected value you will ever meet. **EMV is a long-run average.** Think about what "50-50" means: if TLC made this launch ten times, five of the ten markets would be favourable and five unfavourable - the problem is that TLC does not know *which* five. Now suppose TLC goes with the large plant every single time.



Values in thousands of rupees. EMV is simply this per-launch average.

Figure 2.3. Why EMV is a real number even though it never happens on any single launch.

Alternative, repeated 10 times	Five favourable launches	Five unfavourable launches	Net over 10	Per launch
Large plant	5 x 200,000 = 1,000,000	5 x (-180,000) = -900,000	100,000	10,000
Small plant	5 x 100,000 = 500,000	5 x (-20,000) = -100,000	400,000	40,000

The 10,000 and 40,000 are exactly the EMVs from the table. EMV is what you would make *per instance, on average*, if you faced this decision again and again. The small plant does not earn 40,000 on any one launch; it earns 40,000 per launch on average, over many launches.

"But this is a one-time decision - we will never launch the same sheds twice." True. But you will make many *other* decisions, each with its own payoffs and probabilities. If you consistently choose by EMV across all of them, then across a career you come out ahead of someone who does not. It is the casino's logic in reverse: the house loses individual hands all night and wins on volume, because it plays the expected value every time.

EOL: the same answer from the other side

The second tool for decisions under risk is **Expected Opportunity Loss** minimisation. Opportunity loss - regret - is how much you missed out on by not picking the best alternative for the state that actually occurred. For a favourable market the best payoff is 200,000; for an unfavourable market it is 0. Subtract each payoff from the best in its column, weight by probability, and pick the **smallest**.

Alternative	Regret if favourable	Regret if unfavourable	EOL
Large plant	200,000 - 200,000 = 0	0 - (-180,000) = 180,000	0.5 x 0 + 0.5 x 180,000 = 90,000
Small plant	200,000 - 100,000 = 100,000	0 - (-20,000) = 20,000	0.5 x 100,000 + 0.5 x 20,000 = **60,000**
No plant	200,000 - 0 = 200,000	0 - 0 = 0	0.5 x 200,000 + 0.5 x 0 = 100,000

Small plant again, with the lowest expected regret of 60,000. The mathematics guarantees that EMV maximisation and EOL minimisation always select the same alternative, so the class - and the exam - focuses on EMV. Keep one number in mind, though: the winning alternative's EOL, 60,000, is going to reappear in Section 2.8 under another name.

MANAGER'S TAKEAWAY

EMV is a policy, not a prediction. Choose by the long-run average, not by the best case that excites you or the worst case that frightens you - and then, before you present the number, run the check in the next section. A recommendation without its sensitivity range is half a recommendation.

2.7 Sensitivity Analysis

Just before the break, the professor left a piece of food for thought. Several people had pointed out that the whole answer rested on the assumption of 50-50. Fair. So what if it were not 50-50?

Make it extreme. If the probability of a favourable market were 1 - a certain success - would anyone build a small plant? No; you would build large. If it were 0 - certain failure - you would build nothing. And at 0.5 we found the small plant. So somewhere between 0 and 0.5 the best choice switches from no plant to small plant, and somewhere between 0.5 and 1 it switches from small to large. **Where are those thresholds?** That was the homework for the break, and this is the working.

Write each EMV as a formula in P

Let P be the probability of a favourable market, so the probability of an unfavourable market is 1 - P.

Alternative	EMV as a function of P	Simplified
Large plant	200,000 x P + (-180,000) x (1 - P)	380,000 P - 180,000
Small plant	100,000 x P + (-20,000) x (1 - P)	120,000 P - 20,000
No plant	0 x P + 0 x (1 - P)	0

(Why does -180,000 become +180,000 P? Because -180,000 x (1 - P) = -180,000 + 180,000 P; add the 200,000 P and you get 380,000 P - 180,000.)

Find where the lines cross

The small plant stays the best choice as long as its EMV is at least as large as both the others.

Comparison	Condition	Threshold
Small plant vs no plant	120,000 P - 20,000 >= 0	P >= 20,000 / 120,000 = 2/12 = **0.167**
Small plant vs large plant	120,000 P - 20,000 >= 380,000 P - 180,000 180,000 - 20,000 >= 380,000 P - 120,000 P 160,000 >= 260,000 P	P <= 160,000 / 260,000 = 16/26 = **0.615**

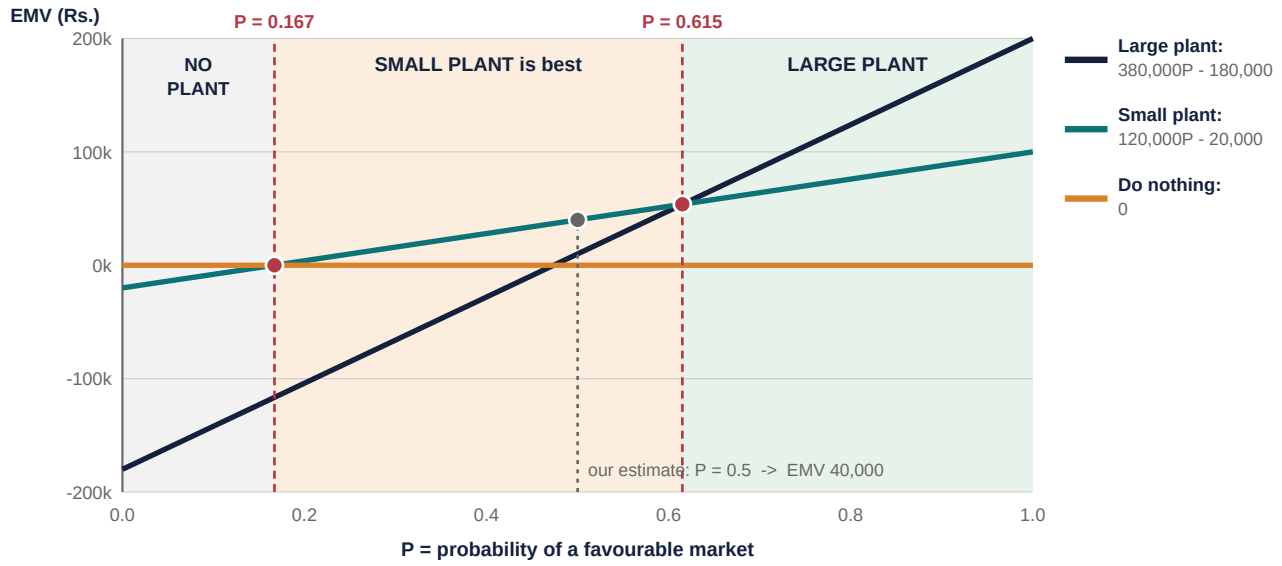


Figure 2.4. EMV of each alternative as P changes. The best choice is the highest line at each P; it switches at 0.167 and 0.615.

Probability of favourable market (P)	Best alternative
Less than 0.167	Do nothing
0.167 to 0.615	Construct a small plant
More than 0.615	Construct a large plant

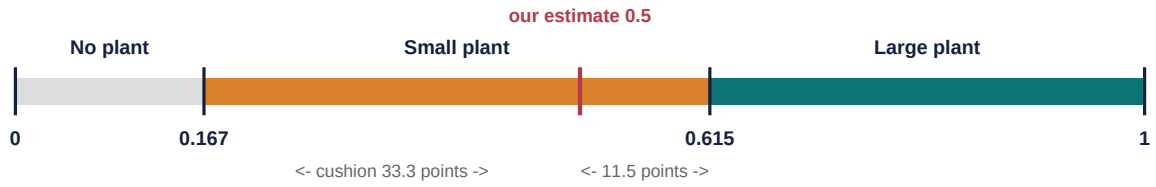


Figure 2.5. The same result on a number line, with TLC's estimate of 0.5 and its cushion on each side.

Robust or sensitive?

TLC's estimate was 0.5. That sits 33.3 percentage points above the lower threshold and 11.5 points below the upper one. Even if management's gut is off by a fair margin, the small plant stays optimal. The decision is **robust**. This is what sensitivity analysis checks: how sensitive - or how robust - your best choice is to the numbers you assumed. If a small change in an input flips the answer, the decision is **sensitive**. If there is a comfortable cushion, it is robust.

Now imagine the upper threshold had come out at 0.515 instead of 0.615. Can anyone's gut feeling tell the difference between a 50 percent and a 51 percent chance? No. Then the decision is sensitive, and the professor's advice is not to freeze but to **prepare for the contingency**: what if I am wrong and the market is stronger than I think? Identify the suppliers you would call, the plant you would lease, the subcontractor who would fill the demand your small plant cannot. Contingency planning - for an upside like this or for a downside risk - always costs something. Risk management is never cheap. But it saves a great deal when the risk actually materialises.

Quick check with other estimates	EMV large	EMV small	Best
P = 0.60	380,000 x 0.6 - 180,000 = 48,000	120,000 x 0.6 - 20,000 = 52,000	Small plant (0.60 is below 0.615)
P = 0.70	380,000 x 0.7 - 180,000 = 86,000	120,000 x 0.7 - 20,000 = 64,000	Large plant (0.70 is above 0.615)
P = 0.15	380,000 x 0.15 - 180,000 = -123,000	120,000 x 0.15 - 20,000 = -2,000	No plant (0.15 is below 0.167)

KEY INSIGHT

Never present an EMV recommendation without its sensitivity range. Robust means the inputs can move a fair bit without changing the answer - decide and move on. Sensitive means a small change flips the answer - decide, and prepare for being wrong. Both are fine. Not knowing which one you are in is not.

2.8 EVPI: The Worth of Perfect Information

A new question. Right now TLC builds the plant first and discovers the market afterwards. Suppose someone could tell you the market condition in advance - and not just tell you, but guarantee it. Perfect information: one hundred percent accurate, no way it can be wrong. Should you want it?

The class hesitated ("no?"), then reasoned its way to yes. Perfect information **changes the sequence of events**. Instead of build-then-discover, you discover-then-build. If the informant says favourable, you build the large plant and make 200,000. If the informant says unfavourable, you build nothing and lose nothing. Notice what happened to the small plant: it vanished. With perfect information you never pick it. And that is the test of whether any information is worth having at all.

WHEN IS INFORMATION WORTH ANYTHING?

Information has value only if it changes what you would do. If, with or without the information, you would take the same course of action, the information is worth nothing to you - however accurate it is - because you were going to do that anyway.

How much is it worth? Value with it, minus value without it

The worth of any information is what you make with it minus what you were making without it - how much better off it leaves you. Without perfect information, the best TLC can do is the small plant with an EMV of 40,000. With perfect information, the informant will say "favourable" half the time (and TLC builds large for 200,000) and "unfavourable" half the time (and TLC builds nothing for 0).

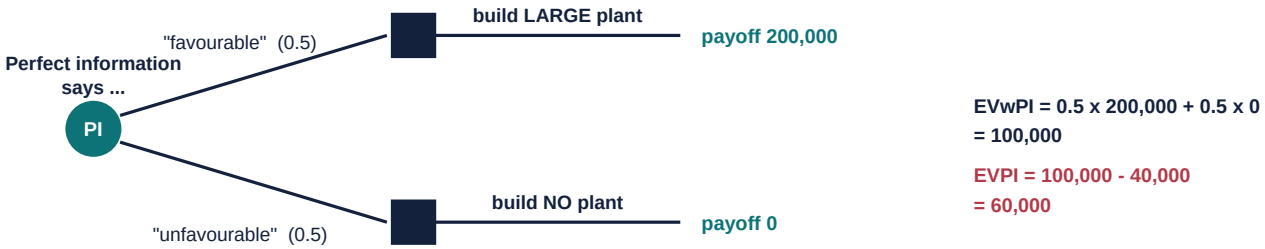


Figure 2.6. With perfect information, TLC always takes the best action for the state that will occur.

Quantity	Working	Value
Expected value WITH perfect information (EVwPI)	Best payoff in each state, weighted by that state's probability: 0.5 x 200,000 + 0.5 x 0	100,000
Expected value WITHOUT perfect information	Best EMV from the payoff table (small plant)	40,000
Expected Value of Perfect Information (EVPI)	EVwPI - best EMV = 100,000 - 40,000	**60,000**

EVPI is the **most** TLC should ever pay for perfect information: at 60,000 it is no worse off than before; below that it gains; above that it loses. How far below 60,000 you can negotiate is up to you - free is best. The slide made the point concrete: a consultant offers certainty about the market for Rs. 65,000. Decline. And notice that 60,000 is exactly the EOL of the small plant from Section 2.6 - the expected regret of your best choice is precisely what perfect information would save you.

Why 0.5 and 0.5 again?

A batchmate asked why the informant's two messages were weighted 50-50 - because there are two messages? No. They are 0.5 and 0.5 because **our estimate of the market** was 0.5 and 0.5. Had we believed 60-40, the informant - who only tells the truth - would say "favourable" 60 percent of the time. The informant has not opened their mouth yet; we are estimating what they will say, and that is the same thing as estimating the market. You cannot hold two different estimates of the same thing. Whatever probabilities you used for EMV, use for EVwPI.

Where does perfect information come from in real life?

Unethically: bribes and inside information. Ethically, the class offered market research (imperfect - Section 2.10), historical profit and loss (that is the past, and some companies' shares rise on a bad year). The professor's answer was a story.

A former student opened a bakery in Calcutta. Her mother had started baking for her as a child, got good at it, took orders by phone, and eventually the family opened a shop. The reviews were excellent, with one constant complaint: everything is delicious, but whatever you ask for, they never have it - you make do with what is on the shelf. Why so little stock? Because the shop's USP was freshness: everything was baked that day, and anything unsold at closing time was thrown away. The daughter, running marketing, wanted to stock more. The mother, watching the bins fill every evening, wanted to stock less. Two stakeholders, exactly opposite views.

Where is the uncertainty coming from? From not knowing **tomorrow's demand, product by product**. Who can remove it? A forecast cannot - it is never spot on. The **customer** can, by **pre-booking**. But why would I tell the bakery today that I want pastries tomorrow? Only for an incentive. And EVPI tells the bakery the ceiling on that incentive: the maximum it should spend to convert walk-ins into pre-bookings. It need not spend all of it - it decides what share of EVPI to pass on as a pre-booking discount. Will every walk-in convert? No; some of us wake up wanting cheesecake with no warning. But every pre-booking shrinks the uncertain part of demand, and the

uncertainty is what was costing money.

Where you have seen this	What is really going on
Newspaper at your door costs the cover price; the same paper at a traffic signal or railway station once cost several times more	The hawker at the signal carries the uncertainty - unsold copies some days, lost sales on others - and prices for it. The subscription vendor has certainty and can afford the cover price.
Annual newspaper subscriptions with a suitcase thrown in (Times of India in the professor's college days)	Committed readers turn uncertain volume into certain volume, so print-run errors shrink. In a competitive market it also ties the reader up for a year.
Rajasthan Patrika vs Dainik Bhaskar	Patrika was the near-monopoly in Rajasthan through the 1990s at a high price and in black and white. Bhaskar entered in the mid-1990s with a lower price, four weekly supplements (kids, women, Bollywood, arts) on glossy paper. Today they split the market roughly 50-50 with identical prices - and annual-payment discounts that exist to stop readers switching, not only to manage uncertainty.
Early-booking discounts on apartments	Partly word of mouth, mostly certainty: a booked flat is demand the builder no longer has to guess about.
OnePlus invite-only launch (2014), Xiaomi flash sales on Flipkart	Nothing to do with uncertainty. Scarcity and privilege as marketing - "I own something not everyone can get". Do not confuse a marketing tactic with an information tactic.

MANAGER'S TAKEAWAY

Before paying for any information - a consultant's study, a proof of concept, a pilot, a data feed - ask two questions. Will it change what I would otherwise do? And is its price below EVPI? If the first answer is no, the information is worth nothing however accurate it is. If the second answer is no, you are paying more than certainty itself is worth.

2.9 Decision Trees

The payoff table has a limitation: it has two dimensions - alternatives and states of nature - and nothing else. The moment a decision has stages (do a survey, read it, then decide, then see what the market does), the table cannot hold it. The tool for that is the decision tree, which the professor called one of the most powerful things you will learn in the whole programme - not because of the arithmetic, but because it provides a structure that takes the clutter out of a problem even when you have no numbers at all.

The grammar of a tree

Element	Symbol	Branches that come out of it	Who controls it
Decision node	Square	One branch per course of action - as many branches as you have choices	You
Outcome (chance) node	Circle	One branch per state of nature, each carrying its probability	Nobody - nature
Payoff	Number at the end of a branch	The result of that complete path	-

Two rules matter more than the symbols. First, the tree is drawn **from left to right**. Second - and this is the important one - it is drawn **in the same sequence in which things would happen in real life**. Take a career decision: you apply for a job (a square - your choice); you get it or you do not (a circle - not your choice); if you get it, you accept or decline (square); they post you to a location you like or do not (circle); your current employer releases you or holds you to the bond (circle); and so on. Drawn in that order, the tree never lets you assume information you would not yet have, and it shows you exactly what you will know at each point. That is the discipline.

TLC as a tree

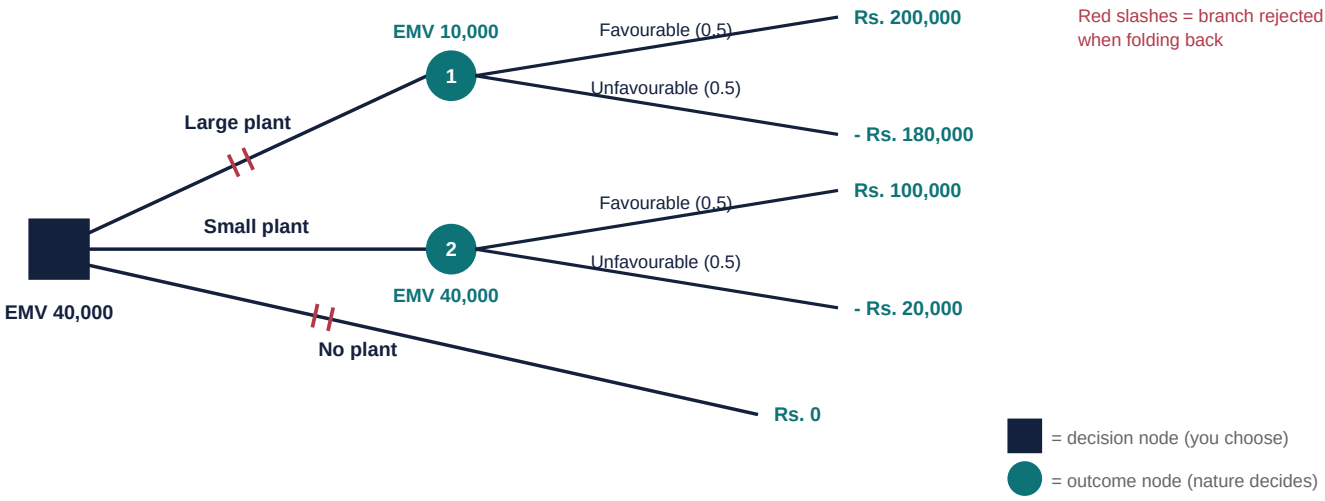


Figure 2.7. The Thompson Lumber problem as a decision tree, drawn left to right and solved right to left.

Read it from the left. TLC chooses (square) among large, small and no plant. Large and small each lead to a circle, because the market decides; each circle has two branches, favourable and unfavourable, with probability 0.5, ending in the payoff. No plant ends immediately at 0 - if you do not launch, you never even find out what the market would have been. That is a subtle and honest feature of the tree.

Solving runs the other way: from right to left, folding the branches back into the nearest node. At a circle you compute the expected value of its branches - 0.5 x 200,000 + 0.5 x (-180,000) = 10,000 at node 1; 40,000 at node 2. At a square, you choose - you always take the best branch - and you mark the rejected branches with two small slashes so that anyone reading the tree can see what was considered and dropped. Back at the root: small plant, EMV 40,000. The tree says exactly what the table said, and now it can grow.

The five steps of decision tree analysis (from the slide)

Step	What you do	Where it happened for TLC
1. Define the problem	Same step 1 as the six-step process.	Launch backyard sheds or not; then, survey or not.
2. Structure or draw the tree	Left to right, in real-life order: squares for choices, circles for nature.	Figure 2.7, then Figure 2.8.
3. Assign probabilities to the states of nature	On every branch out of a circle; from experience, history or Bayes.	0.5 / 0.5 without the survey; 0.45, 0.78, 0.27 with it.
4. Estimate payoffs for each path	At the end of every branch; net of any cost already committed on that path.	200,000 ... 0 without the survey; 190,000 ... -10,000 with it.
5. Solve by computing the EMV at each state-of-nature node	Right to left: average at circles, choose at squares, slash the rejected.	Nodes 2 to 7, then node 1: 49,200.

Two more lines from the slides worth keeping: any problem that fits a decision table can also be drawn as a tree, and trees are most beneficial when a **sequence** of decisions must be made - which is exactly why the survey problem in the next section needs one.

Trees without numbers

A batchmate asked what to do when you have no payoffs and no probabilities. The professor's answer is that people use trees in their personal lives all the time in exactly that state: at the end of each branch you write the pluses and the minuses instead of a number. The tree will not solve the problem for you. What it gives you is a truer picture - everything written down comprehensively in one place - so that when you make the trade-off, you know what you are letting go of, and in favour of what.

A life decision, on paper

An architect with the Central Public Works Department was working on the Central Vista project - prestigious work, a good posting near Delhi, official residence, pension, security. He also had an admission to a master's programme in construction management in Canada, and because he did not yet qualify for study leave, going meant resigning. He did his research, as the session advised. Corporate friends said: are you mad, people dream of a job like yours. People who had built their own careers said: go. Nobody could decide for him, because there is no right answer - and, as the batchmate had said earlier, you never learn what the other branch would have held.

So he drew the tree, in his diary. Two things jumped out. He was just out of a divorce and had no financial responsibilities - no parents to support, no dependants. And he had carried the dream of studying abroad for years. Against that, none of the pluses on the CPWD branch actually excited him; he was fine letting them go even if the other branch turned out badly. He chose Canada. Then, at the last moment, CPWD granted the leave. Then COVID arrived, and he spent the programme alone in a room, studying online - none of the campus culture he had gone for. He came back, later left CPWD anyway, and went on to another course in Canada.

Was it a bad decision? He says no, and he is right by this session's standard: he decided well informed, and COVID was not knowable. Because the tree was in his diary, he could look at it later and see that he had let go of what he had chosen to let go of. No beating himself up - the process was sound; the outcome was chance.

Trees in meetings

Why do meetings run long? Either somebody wrote a casual point on a slide and the whole room went down that road, or nobody did their homework. In many organisations the rule is that the **meeting owner arrives with a tree**. If the problem has already been defined, the discussion can only be about three things: is there a branch missing (an alternative nobody listed)? is there a state of nature missing? are the numbers wrong? A two-hour argument becomes three questions.

A batchmate asked how this differs from the fishbone diagram. A fishbone is for **root-cause analysis** - chains of "why did this happen?" A tree is a **solution mechanic**: it assumes the problem has been defined and lays out the options and what follows from them. Fishbone before step 1; tree after step 4.

MANAGER'S TAKEAWAY

Bring a tree, not a deck. One page that shows what you control, what you do not, what you will know when, and what each path is worth tells the room more than twenty slides - and it makes the disagreement precise: a missing branch, a missing state, or a number.

2.10 EVSI and Bayes: Is the Survey Worth It?

Perfect information at Rs. 65,000 was too expensive. So TLC considers the realistic option: a market survey - imperfect, but it "may help quite a bit". What is imperfect information worth, and how do you even lay out a decision that now has two stages? This is the part of the session where the tree earns its keep.

The survey, described

Fact	Number	What it means
Cost of the survey	Rs. 10,000	Committed the moment TLC decides to survey.
Chance the survey will report a favourable market	45%	So a 55% chance it reports unfavourable. Reading a report is not the same as the market being that way.
If the survey says favourable, chance the market is actually favourable	78%	22% of the time the survey is wrong on the optimistic side.
If the survey says unfavourable, chance the market is actually favourable	27%	73% of the time an unfavourable report is right.

Why 22 percent one way and 27 percent the other - why not symmetric? Because in real life the two kinds of error are rarely equal. The professor's example: the COVID rapid antigen home test. If it came back positive, you were told to trust it and isolate; if it came back negative but you had symptoms, you were told to go for an RT-PCR. Two reasons. The test's false-positive rate and false-negative rate were different. And the **stakes** were different: someone with COVID who believed they were fine could die; someone without COVID who isolated for a week lost very little.

Drawing the survey tree - and the mistake everyone makes

The first node is a square with two branches: conduct the survey, or do not. If you do not, you are back to the tree from Section 2.9 - large, small or no plant at 50-50. If you do, what comes next? This is where the class went wrong: many drew another square with the three plant choices. That is incorrect. **You paid for the survey, so you read it before you decide.** The next node is a circle - the survey result, favourable (0.45) or unfavourable (0.55) - and only then a square for the plant decision, followed by the market circle. One more change: every payoff on the survey side is **reduced by 10,000**, because that money is spent whatever happens. 200,000 becomes 190,000, -180,000 becomes -190,000, and "no plant" is no longer free - it is -10,000.

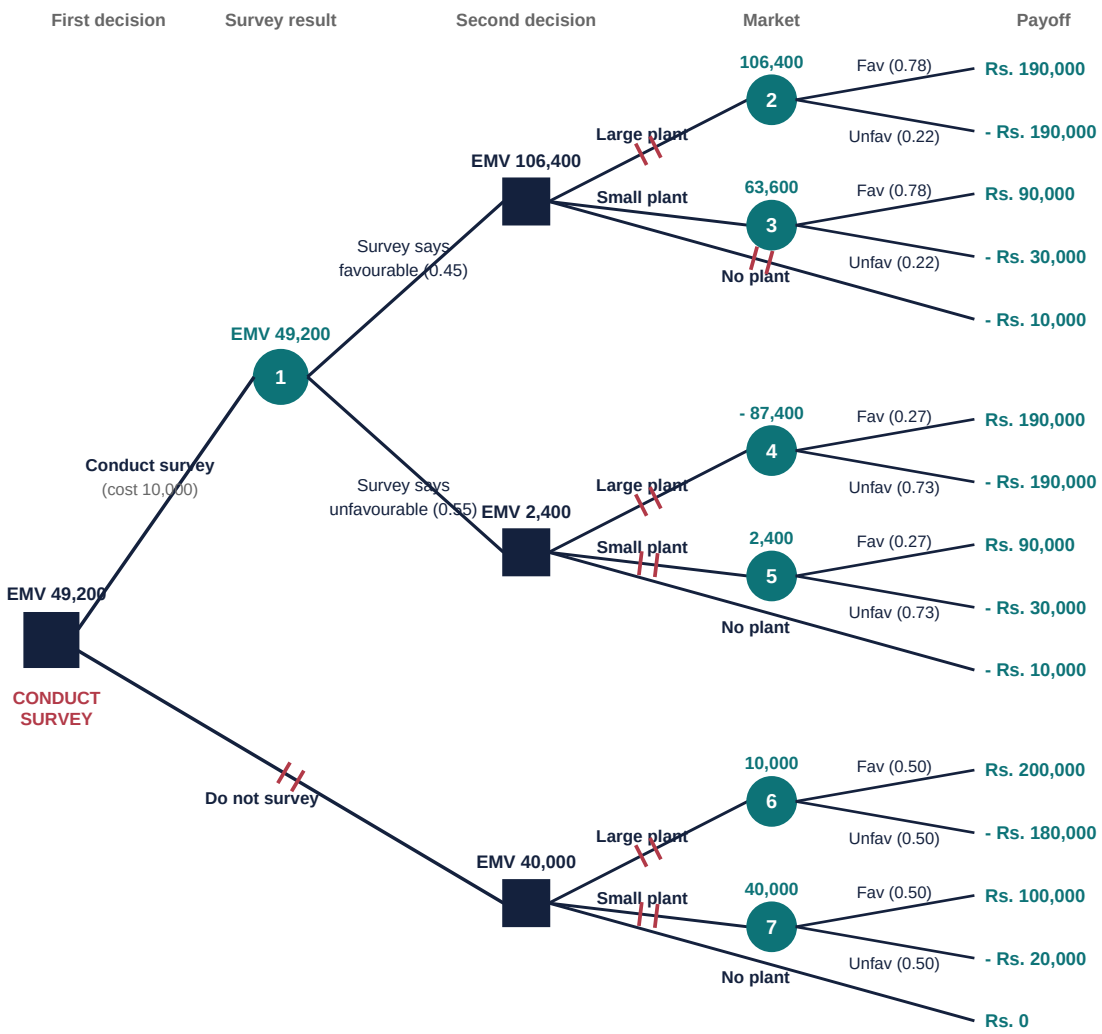


Figure 2.8. The full survey tree, solved. Node numbers match the slides. Slashes mark branches rejected while folding back.

Folding it back

Node	Situation	Working	EMV
2	Large plant, survey said favourable	0.78 x 190,000 + 0.22 x (-190,000) = 148,200 - 41,800	106,400
3	Small plant, survey said favourable	0.78 x 90,000 + 0.22 x (-30,000) = 70,200 - 6,600	63,600
-	No plant, survey said favourable	Survey cost only	-10,000
Square after favourable	Choose the best of 106,400 / 63,600 / -10,000	Large plant	**106,400**
4	Large plant, survey said unfavourable	0.27 x 190,000 + 0.73 x (-190,000) = 51,300 - 138,700	-87,400
5	Small plant, survey said unfavourable	0.27 x 90,000 + 0.73 x (-30,000) = 24,300 - 21,900	2,400

Node	Situation	Working	EMV
-	No plant, survey said unfavourable	Survey cost only	-10,000
Square after unfavourable	Choose the best of -87,400 / 2,400 / -10,000	Small plant	**2,400**
1	Conduct the survey	0.45 x 106,400 + 0.55 x 2,400 = 47,880 + 1,320	**49,200**
6 / 7	Do not survey (from Section 2.6)	Large 10,000 / Small 40,000 / None 0 - choose small	40,000
Root	Survey (49,200) vs no survey (40,000)	Conduct the survey	**49,200**

The tree now reads as a complete strategy, which is something a table could never give you: **conduct the survey; if it comes back favourable, build the large plant; if it comes back unfavourable, build the small plant.** On average TLC makes 49,200 this way, against 40,000 without the survey.

So what is the survey worth? (Three wrong answers and the right one)

The class offered 9,200 (the gain), 190,000 (the best payoff) and 106,400 (the value after a favourable report). The 190,000 is a best case, not a worth; the 106,400 exists only on the branch where the survey says favourable - what if you had paid that much and the survey said no? The 9,200 is closest, and still wrong, for a subtle reason: the 49,200 was earned *after* paying 10,000 for the survey. If the survey had been free, TLC would have made 59,200. The worth of information is always what you make with it, assuming it were free, minus what you make without it.

Quantity	Working	Value
Expected value WITH the survey, assuming it were free	49,200 + 10,000	59,200
Expected value WITHOUT the survey	Best EMV without information (small plant)	40,000
Expected Value of Sample Information (EVSI)	59,200 - 40,000	**19,200**
Cost of the survey	Given	10,000
Net gain from surveying	19,200 - 10,000	9,200

So the survey is worth 19,200; TLC should be willing to pay anything up to that, and at 10,000 it is a good buy. Two sanity checks that were not spelt out in class but follow directly: EVSI can never exceed EVPI, because imperfect information cannot be worth more than perfect information (19,200 against 60,000); and the ratio EVSI / EVPI - here 19,200 / 60,000 = 32 percent - is a handy measure of how efficient a survey is, a way of comparing two research proposals on the same decision.

Where do 0.45, 0.78 and 0.27 come from? Bayes' theorem

The slide put it in three lines: probability values can come from management's experience and intuition, from historical data, or be computed from other data using Bayes' theorem - which incorporates your initial estimates and the known accuracy of the information source, and lets you revise the initial estimates as new information arrives.

The professor flagged this part as **not in the exam**, but insisted we understand what it does, because in real life a market research firm or a consultant will be doing it for us and we will be on the receiving end. Class suggestions for the source of these probabilities: ask the research firm, or use its historical record - out of ten times it said

favourable, the market really was favourable seven times. That record is the right raw material, and using it **directly** would be a serious mistake.

Why? Imagine two products in the same company. One is almost certain to succeed - the COVID vaccine during the pandemic, say 90-10. The other is a genuine 50-50. If you plug the vendor's 70 percent accuracy straight into both trees, both products get identical survey probabilities, which is absurd: a survey on a near-certain product should point far more confidently than a survey on a coin toss. The survey's accuracy has to be **combined with what you already believe about this product** - your prior. That combination is Bayes' theorem: start with the prior, multiply by how often the evidence shows up in each state, and re-scale so the possibilities add to one.

Survey accuracy (from the vendor's history)	If the market is favourable (FM)	If the market is unfavourable (UM)
Survey says favourable (positive)	P(positive | FM) = 0.70	P(positive | UM) = 0.20
Survey says unfavourable (negative)	P(negative | FM) = 0.30	P(negative | UM) = 0.80

TLC's prior is 0.5 favourable, 0.5 unfavourable. Now revise.

If the survey says FAVOURABLE	Prior	x Survey accuracy	= Joint	Posterior = Joint / 0.45
Market favourable	0.50	0.70	0.35	0.35 / 0.45 = **0.78**
Market unfavourable	0.50	0.20	0.10	0.10 / 0.45 = **0.22**
Total = P(survey says favourable)			**0.45**	1.00

If the survey says UNFAVOURABLE	Prior	x Survey accuracy	= Joint	Posterior = Joint / 0.55
Market favourable	0.50	0.30	0.15	0.15 / 0.55 = **0.27**
Market unfavourable	0.50	0.80	0.40	0.40 / 0.55 = **0.73**
Total = P(survey says unfavourable)			**0.55**	1.00

Every number on the survey tree - 0.45, 0.55, 0.78, 0.22, 0.27, 0.73 - is now accounted for. And the "customise to the prior" point becomes visible if you redo the table with a 90-10 prior: the survey then says favourable 65 percent of the time, a favourable report means a 97 percent chance of a favourable market, and even an unfavourable report still leaves a 77 percent chance of a favourable market. Same survey, very different numbers - because the product is different.

BAYES IN ONE BREATH

Prior belief x how often this evidence appears in each world = joint probability; divide by the total to get the revised (posterior) belief. The same logic tells a security analyst how likely an alert is a real incident given the base rate of attacks and the detection tool's accuracy - and it is why a "95% accurate" tool can still be wrong most of the time when real incidents are rare.

MANAGER'S TAKEAWAY

When a vendor quotes you the accuracy of a survey, a test or a detection tool, that number alone is not the probability you need. Ask them to combine it with your prior - the base rate for your product, your market, your environment. If they cannot explain how the prior entered their numbers, they have not done the job.

2.11 Product, Service and the Manager's Desk

Everything in this session was taught on backyard sheds. The ideas do not care what the product is. Here is each of them transplanted into a product business and a service business, followed by what they mean for the way you run a team.

Concept	Product-based company	Service-based company
Define the problem, not the symptom	FMCG: "sales are falling in the North" is the symptom. The root cause turns out to be distributor stock-outs on the top three SKUs - fill rate, not demand.	Airline: "complaints are rising" is the symptom. The root cause is a ground-crew roster that leaves the 6 a.m. bank of flights short-staffed.
Courses of action vs states of nature	Auto maker: build a large EV line, a small line, or none (choices) - against demand turning out high, medium or low (nature).	IT services: staff a dedicated delivery centre for a client, or use flexible staffing (choices) - against the client renewing, shrinking or leaving (nature).
Payoff table and EMV	Consumer electronics: order 1 lakh, 50,000 or 20,000 units of a new model; payoff = contribution under each demand scenario; choose the largest EMV.	Hotel: block 200, 100 or 50 rooms for a conference; payoff = revenue net of unsold-room cost under each attendance scenario.
Sensitivity analysis	Pharma: at what probability of regulatory approval does "build the plant now" beat "wait"? If the threshold is close to the estimate, line up a contract manufacturer as contingency.	Managed security services: at what renewal probability does hiring a dedicated analyst team for one customer beat a shared pool? If it is close, keep the pool and contract surge capacity.
EVPI - pay for certainty only up to its worth	Bakery: the pre-booking discount pool is capped by EVPI. Apparel: pre-season order commitments from retailers reduce the markdown risk that Session 1 discussed.	Consulting: a paid diagnostic before a large engagement. SaaS: an annual plan discount that converts uncertain monthly churn into a certain year.
Decision tree and EVSI	Test-market in one city (the survey), read the result, then choose national, regional or no launch. Pay for the test only if EVSI exceeds its cost.	Bank (Pilgrim): pilot a fee in one district, read adoption and attrition, then decide fee, free or rebate nationwide.
Bayes - revise with new evidence	Quality: the plant's historical defect rate is the prior; the inspection sample updates it before you decide to ship or hold the lot.	SOC: the base rate of real intrusions is the prior; the detection tool's true-positive and false-positive rates update it into "probability this alert is a real incident".
Gut feeling vs evidence	The plant manager's instinct says the new supplier's parts fail in the cold; make him find the data - including the batches that did not fail.	The SOC lead's instinct says a rule is noisy; check the alerts it caught as well as the ones it missed before switching it off.

Concept	Product-based company	Service-based company
Write it down; bring a tree	A one-page pre-mortem for the launch: problem statement, options, what could go wrong.	A post-incident review that asks what was known at the time and which branch was chosen - not who to blame.
Process over outcome	A launch that followed a sound six-step process and still lost to an unforeseeable competitor is a good decision; review the process, not the P&L, in the post-mortem.	A project that delivered on time by skipping the risk review is a bad decision that got lucky; the next one will not be.

The manager's checklist from this session

#	Discipline	What it looks like on Monday morning
1	Grade the process, not the outcome	In reviews and post-mortems, ask what was known at the time and whether all options and outcomes were considered. Reward reporting of failures; punish cover-ups.
2	Fall in love with the problem	Write the problem statement with no solution inside it. Check for symptom vs root cause. Use the firefighting test: did we get calm after the fix?
3	List every alternative, including the ones you dislike	Let the analysis kill weak options; do not let your bias kill them first.
4	Separate what you control from what you do not	Alternatives are choices (no probabilities). States of nature are the world's move (probabilities). Confusing the two breaks every calculation.
5	Decide by EMV, then test with sensitivity	Report the recommendation with its threshold range. Robust: proceed. Sensitive: proceed with a costed contingency plan.
6	Value information before buying it	Would it change what I do? Is the price below EVPI (or EVSI for imperfect information)?
7	Bring a tree to the meeting	Left to right in real-life sequence; solve right to left. The argument becomes: missing branch, missing state, or wrong number.
8	Respect gut feeling, refuse stubbornness	When gut and evidence disagree, find the faulty assumption or find the right data - the successes as well as the failures.
9	Convert uncertainty to risk with proxies	Team test, friends-and-family, employee beta, pilot city: each stage turns "no idea" into an estimate you can put on a tree.
10	Write it down	Two minutes, a few bullets. Problem statement, empathy, options. The clutter and the self-serving bias fall away on paper.

The sharpest lines of the session

> *"Being wrong feels exactly like being right - until you find out."*
>
> **On why confidence is not evidence**

> *"Evaluating a decision on the basis of its outcome promotes mediocrity."*
>
> **On precedent as a safety net**

"Solving the right problem roughly is way better than solving the wrong problem accurately."

On step one of six

"Information is worth something only if it changes what you would do."

On EVPI and EVSI

"Risk management is never cheap, but it saves you a lot of money when the risk materialises."

On contingency plans for sensitive decisions

"Take off your engineering hat and put on your management hat."

The Challenger line - remembered as a warning, not as advice

ONE-BREATH SUMMARY

A good decision is judged by its process - all options, all outcomes, all data, analysed rationally - never by its outcome, because chance sits between the two. Define the problem before you fall for a solution; then list what you control (alternatives), what you do not (states of nature), and the payoff of each pairing. Under risk, choose by EMV - the long-run average - and check with sensitivity analysis whether the choice is robust or needs a contingency. Information is worth the improvement it buys and nothing more: EVPI caps the price of certainty, EVSI the price of a survey. Multi-stage decisions go on a tree - drawn left to right in the order life unfolds, solved right to left - and the probabilities on that tree come from combining what you already believed with how accurate the new evidence is, which is Bayes.

2.12 Practice MCQ Bank with Answers

Sixty-eight questions drawn from this session, in roughly the order the ideas were taught. The options are meant to be close - several are numerically right for the wrong reason - so read the one-line reason under each answer even when you got it right. All numbers refer to the Thompson Lumber problem unless stated otherwise.

Q1. Which statement best captures the module's definition of a good decision?

(a) One that produces the desirable outcome

(b) One that follows the precedent that worked the last time

(c) One based on logic, that considers all available data and alternatives and is arrived at through rational analysis

(d) One approved by the most experienced person in the room

Answer: (c) - Outcome, precedent and seniority are all substitutes for process; the definition is about the process and the information available at the time.

Q2. "Being wrong feels exactly like being right." The point of this line was that:

(a) Wrong decisions are as common as right ones

(b) Until you discover you were wrong, your confidence is identical - so confidence is not evidence of being right

(c) Emotions should be kept out of decision making

(d) Outcomes take time to materialise

Answer: (b) - Every feeling the class listed was the feeling of finding out; before that moment, wrong and right feel the same.

Q3. Why does judging decisions by their outcome "promote mediocrity"?

(a) Outcomes are hard to measure accurately

(b) It makes people take excessive risks to chase results

(c) People copy precedent as a safety net and avoid the full-potential option, because rewards and penalties follow outcomes

(d) It ignores the ethical dimension

Answer: (c) - When outcomes are rewarded and punished, everyone plays safe and the organisation settles in the middle - neither failing nor realising its potential.

Q4. A manager used all available data and considered every option; the market turned bad and the product lost money. By the session's standard the decision was:

(a) Bad, because it lost money

(b) Good, because decision quality is judged on the process and the information available at the time

(c) Neither - too early to tell

(d) Bad, because the loss could have been predicted

Answer: (b) - Under risk, a sound process can meet an unlucky state of nature; the outcome reflects chance, not the decision.

Q5. In the Alcoa story, why was the plant executive fired?

(a) An accident occurred on his plant

(b) He missed his production targets

(c) He tried to cover up an accident - the outcome mindset the CEO was eliminating

(d) He opposed the union

Answer: (c) - Accidents were allowed to happen once; hiding one was not, because cover-ups mean the process is being fudged.

Q6. Why could the unions not resist Paul O'Neill's changes to how work was done at Alcoa?

(a) He offered higher wages in return

(b) Every change was framed as removing a hazard to workers, which no union can oppose

(c) Unions had little power in 1987

(d) The analysts backed him publicly

Answer: (b) - Safety was the door through which every process was re-examined; nobody can argue for keeping a hazardous process.

Q7. The stock price fell immediately after O'Neill's first speech. This illustrates:

(a) That safety programmes destroy shareholder value

(b) The market grading a leader on an early "outcome" before any process had been built

(c) That analysts had perfect information

(d) Sensitivity analysis

Answer: (b) - The analysts judged the outcome of one speech; the process he built took thirteen years to show in profits.

Q8. Why does design thinking forbid talking about solutions in its first steps?

(a) Solutions are expensive to prototype

(b) Customers cannot evaluate solutions

(c) Once invested in a solution, we reshape the problem so that the solution fits it

(d) Problems are more interesting than solutions

Answer: (c) - Emotional investment in a solution creates the bias of moulding the problem to fit; the ban protects step one.

Q9. For the blind child who wants to play basketball, the session's "real problem" was:

(a) The child cannot see

(b) No eye transplant is possible

(c) The child cannot independently navigate an unfamiliar space

(d) Basketball is unsafe for the child

Answer: (c) - "Cannot see" is the symptom; the navigation problem is the one echolocation actually solves.

Q10. Which is the professor's test that you treated a symptom rather than the root cause?

(a) The fix cost more than the budget

(b) The next fire arrives without a significant period of calm

(c) The team disagreed with the fix

(d) The fix took longer than a week

Answer: (b) - Calm after a fix is the evidence that the cause was treated; back-to-back fires mean the cause is still there.

Q11. If firefighting happens daily, which of these was NOT listed as a possible reason?

(a) The processes are broken

(b) The processes are fine but their implementation is broken

(c) The team is not aligned or motivated

(d) The organisation has too many strategic projects

Answer: (d) - The four listed reasons were broken processes, broken implementation, an unaligned team, and treating symptoms - not too much strategy.

Q12. Which step of the six-step process should receive the most time?

(a) Step 1: define the problem

(b) Step 4: estimate the payoffs

(c) Step 5: select the model

(d) Step 6: apply the model

Answer: (a) - "Solving the right problem roughly beats solving the wrong problem accurately" - steps 5 and 6 are mechanical.

Q13. Which of these is a state of nature rather than a course of action?

(a) Constructing a small plant

(b) Conducting a market survey

(c) The market turning out unfavourable

(d) Charging a fee for online banking

Answer: (c) - The other three are choices under the decision-maker's control; the market's reaction is not.

Q14. "What is the probability that TLC chooses the large plant?" The correct response is:

(a) One in three, since there are three alternatives

(b) 0.5, from the payoff table

(c) The question is malformed - alternatives are chosen (1 or 0); only states of nature carry probabilities

(d) It equals the EMV divided by the payoff

Answer: (c) - This was the professor's explicit concept trap: you pick alternatives, you do not roll dice on them.

Q15. Why does the six-step process insist on listing ALL alternatives?

(a) To make the payoff table larger

(b) To neutralise the bias that quietly drops options we dislike or over-favours our own idea before the analysis begins

(c) Because the model cannot run with fewer than three

(d) To satisfy auditors

Answer: (b) - The rival's-product example: national launch vanishes from a rival's list, the small-town test vanishes from your own.

Q16. A payoff in step 4 need NOT be:

(a) Estimated for every alternative-state combination

(b) Monetary

(c) Tied to the performance measure of the decision

(d) Comparable across alternatives

Answer: (b) - Market share is a perfectly valid payoff; "monetary" in EMV is only a label.

Q17. Decision making under uncertainty differs from decision making under risk in that:

(a) Under uncertainty there is only one state of nature

(b) Under uncertainty the probabilities of the states are unknown; under risk they are estimated

(c) Under risk the outcome of each action is certain

(d) Uncertainty uses EMV while risk uses linear programming

Answer: (b) - Both have several states of nature; the difference is whether you can attach probabilities.

Q18. Amazon launching the Kindle before e-readers existed as a category is an example of decision making under:

(a) Certainty

(b) Risk

(c) Uncertainty

(d) Sensitivity

Answer: (c) - No historical data and no probability estimates for how the market would react - "shooting in the dark".

Q19. Tools for decision making under certainty include:

(a) EMV maximisation and EOL minimisation

(b) Decision trees

(c) Linear and integer programming

(d) Bayes' theorem

Answer: (c) - Under certainty each action has one known outcome, so the job is optimisation - operations research.

Q20. Per the session, using gut feeling in analytical decision making is:

(a) To be avoided altogether

(b) Acceptable, because it is inference from a sample of experience - what is unacceptable is stubbornness when the evidence disagrees

(c) Only acceptable for the most senior person

(d) A substitute for market research

Answer: (b) - Gut feeling works like statistics - a sample generalised to the whole; the sin is refusing to examine it.

Q21. Your gut feeling and the data disagree. Which of these is NOT one of the two checks the professor prescribed?

(a) Find the assumption in your gut feeling that may not hold in this situation

(b) Check whether you are looking at the right data

(c) Overrule the data, since experience outranks analysis

(d) Both (a) and (b) are the prescribed checks

Answer: (c) - The two checks are the faulty assumption and the wrong data; pulling rank is exactly the stubbornness the session warned against.

Q22. The Challenger engineers' chart failed to convince because:

(a) It used the wrong temperature units

(b) It showed only launches with O-ring problems, omitting the successful launches that all sat at warmer temperatures

(c) It was drawn by managers rather than engineers

(d) It had too few data points to plot

Answer: (b) - Failures alone appeared at all temperatures; adding the successes showed that no clean launch had happened below the threshold.

Q23. The lesson the professor drew from Challenger was:

(a) Gut feeling should be ignored in engineering

(b) Managers should never overrule engineers

(c) The engineers' gut was right but they never found the data to support it - look at the successes, not only the failures

(d) Launch schedules should never be delayed

Answer: (c) - Their instinct was correct and later proved; what failed was the second check - finding the right data.

Q24. The EMV of an alternative is:

(a) Its most likely payoff

(b) Its best-case payoff

(c) The probability-weighted average of its payoffs across the states of nature

(d) Its payoff minus its opportunity loss

Answer: (c) - A weighted average where the weights are the probabilities - the professor's translation for non-mathematicians.

Q25. At P = 0.5, EMV of the large plant is:

(a) Rs. 200,000

(b) Rs. 10,000

(c) Rs. 40,000

(d) Rs. 20,000

Answer: (b) - 0.5 x 200,000 + 0.5 x (-180,000) = 100,000 - 90,000 = 10,000.

Q26. The EMV criterion selects the SMALLEST EMV when:

(a) Never - it always selects the largest

(b) The payoff is a cost or risk measure, where smaller is better

(c) The probabilities are below 0.5

(d) There are more than two states of nature

Answer: (b) - "Best" depends on the payoff: largest for profit, revenue or share; smallest for cost, risk or defects.

Q27. "The small plant never actually earns 40,000." The right interpretation of EMV = 40,000 is:

(a) It is the guaranteed profit

(b) It is the median profit

(c) It is the long-run average per instance if the same decision were repeated many times under the same probabilities

(d) It is a rounding artefact of the table

Answer: (c) - Over ten launches, five favourable and five unfavourable, the small plant nets 400,000 - 40,000 per launch.

Q28. Over ten repetitions with five favourable and five unfavourable markets, the large plant nets in total:

(a) Rs. 1,000,000

(b) Rs. 100,000

(c) Rs. 10,000

(d) Rs. -900,000

Answer: (b) - 5 x 200,000 - 5 x 180,000 = 1,000,000 - 900,000 = 100,000 over ten instances, i.e. 10,000 per instance.

Q29. EOL minimisation compared with EMV maximisation:

(a) Can recommend a different alternative

(b) Always recommends the same alternative

(c) Is used only under uncertainty

(d) Needs no probabilities

Answer: (b) - The mathematics guarantees the same choice; that is why the class focuses on EMV.

Q30. For TLC at P = 0.5, the EOL of the small plant is:

(a) Rs. 40,000

(b) Rs. 60,000

(c) Rs. 90,000

(d) Rs. 100,000

Answer: (b) - Regrets are 100,000 (favourable) and 20,000 (unfavourable): 0.5 x 100,000 + 0.5 x 20,000 = 60,000 - which also equals EVPI.

Q31. The range of P (probability of a favourable market) over which the small plant is the best choice is:

(a) 0 to 0.167

(b) 0.167 to 0.615

(c) 0.615 to 1

(d) 0.5 to 1

Answer: (b) - Below 0.167 do nothing wins; above 0.615 the large plant wins; the small plant holds the middle.

Q32. If P = 0.70, EMV recommends:

(a) Small plant, EMV 64,000

(b) Large plant, EMV 86,000

(c) Do nothing

(d) Indifferent between small and large

Answer: (b) - 380,000 x 0.7 - 180,000 = 86,000 beats 120,000 x 0.7 - 20,000 = 64,000; 0.70 is above the 0.615 threshold.

Q33. If P = 0.60, EMV recommends:

(a) Large plant, because 0.60 is above 0.5

(b) Small plant, EMV 52,000

(c) Large plant, EMV 48,000

(d) Do nothing

Answer: (b) - 0.60 is still below the 0.615 threshold: small = 52,000 beats large = 48,000. The 0.5 estimate is irrelevant; the threshold is what matters.

Q34. A decision is called "sensitive" when:

(a) Its EMV is negative

(b) A small change in an input flips the recommended alternative

(c) The probabilities are subjective

(d) The payoff table has more than three rows

Answer: (b) - Sensitivity analysis asks how much the inputs can move before the answer changes; a thin cushion means sensitive.

Q35. When sensitivity analysis shows a very thin cushion, the professor's advice is to:

(a) Abandon the decision until better data arrives

(b) Choose the alternative with the smallest possible loss instead

(c) Proceed, but prepare a contingency plan - accepting that risk management has a cost

(d) Buy perfect information at any price

Answer: (c) - If the upper threshold were 0.515 nobody's gut can tell 50% from 51%; so line up suppliers, a leased plant or a subcontractor in advance.

Q36. EVPI equals:

(a) The EMV of the best alternative

(b) The expected value with perfect information minus the best EMV without it

(c) The cost of a market survey

(d) The best payoff in the table

Answer: (b) - Worth of information = what you make with it minus what you make without it.

Q37. EVwPI for TLC is 100,000 because:

(a) It is the best payoff in each state weighted by that state's probability: 0.5 x 200,000 + 0.5 x 0

(b) It is the large-plant payoff halved

(c) It is the sum of the positive payoffs divided by three

(d) It is EMV(large) + EMV(small) + 50,000

Answer: (a) - Options (c) and (d) also produce 100,000 by coincidence; only (a) is the actual logic - with perfect information you always pick the best action for the state that will occur.

Q38. Information is worth paying for only if:

(a) It is 100 percent accurate

(b) It would change the course of action you would otherwise take

(c) It comes from a reputed agency

(d) Its cost is below the best payoff

Answer: (b) - If you would do the same thing with or without it, it is worth nothing however accurate it is.

Q39. A consultant offers TLC perfect information about the market for Rs. 65,000. TLC should:

(a) Accept, since 65,000 exceeds the small plant's EMV

(b) Decline, since EVPI of 60,000 is the ceiling on what the information is worth

(c) Accept only if the probability of a favourable market is 0.5

(d) Counter-offer at 62,000

Answer: (b) - Any price above 60,000 leaves TLC worse off than deciding without the information.

Q40. In computing EVwPI, the informant's "favourable" and "unfavourable" messages are weighted 0.5 and 0.5 because:

(a) The informant has exactly two possible messages

(b) They are our own estimates of the market, and we cannot hold two different estimates of the same thing

(c) Perfect information is always a 50-50 proposition

(d) EMV happened to be computed at 0.5

Answer: (b) - Had the prior been 60-40, a truthful informant would say "favourable" 60 percent of the time.

Q41. The Calcutta bakery could reduce its demand uncertainty most directly by:

(a) Stocking more of every product

(b) Offering a pre-booking discount funded from EVPI, converting some walk-ins into committed orders

(c) Discarding fewer items at closing time

(d) Raising prices on popular items

Answer: (b) - The customer is the one who can remove tomorrow's demand uncertainty, and needs an incentive to do so; EVPI caps that incentive.

Q42. Per the session, OnePlus's invite-only launch and Xiaomi's flash sales were:

(a) Ways to obtain perfect demand information

(b) Marketing tactics creating scarcity and exclusivity - not uncertainty-management devices

(c) EVSI exercises

(d) Markdown strategies

Answer: (b) - They created the feeling of owning something not everyone can get; inventory was not the constraint.

Q43. In a decision tree:

(a) Squares are chance nodes and circles are decision nodes

(b) Squares are decision nodes whose branches are courses of action; circles are outcome nodes whose branches are states of nature with probabilities

(c) Both node types are drawn from right to left

(d) Branches out of a square carry probabilities

Answer: (b) - Square = you choose; circle = nature decides. Only the circle's branches have probabilities.

Q44. A decision tree is drawn from left to right in the sequence in which events unfold so that:

(a) It fits on a landscape slide

(b) You never assume information that will not yet be available, and you see what you will know at each point

(c) The probabilities sum to one

(d) The payoffs appear last

Answer: (b) - The sequence is the discipline: read the survey before you choose the plant, not after.

Q45. Solving ("folding back") a decision tree means:

(a) Working left to right, computing EMV at squares

(b) Working right to left: computing expected value at circles, choosing the best branch at squares and marking the rejected ones

(c) Working right to left, choosing the best branch at circles

(d) Evaluating only the leftmost node

Answer: (b) - Branches fold into their nearest node: average at circles, choose at squares, slash what you drop.

Q46. In the survey tree, what immediately follows the "conduct survey" branch?

(a) A decision node with large, small and no plant

(b) A chance node for the survey result - favourable (0.45) or unfavourable (0.55) - because you read the report before deciding

(c) The market outcome node

(d) A payoff of -10,000

Answer: (b) - Drawing a square straight after the survey was the mistake most of the class made.

Q47. Why do the payoffs on the survey side read 190,000 and -190,000 instead of 200,000 and -180,000?

(a) The survey changes the market

(b) The Rs. 10,000 survey cost is committed the moment you decide to survey, so every payoff on that side is reduced by 10,000

(c) Because of a five percent tax

(d) Because the probabilities changed to 0.78 and 0.22

Answer: (b) - Even "no plant" becomes -10,000 on the survey side; the money is spent regardless of what follows.

Q48. Given a favourable survey result, the best action and its EMV are:

(a) Small plant, 63,600

(b) Large plant, 106,400

(c) Large plant, 200,000

(d) No plant, -10,000

Answer: (b) - 0.78 x 190,000 + 0.22 x (-190,000) = 106,400 beats 63,600 and -10,000.

Q49. Given an unfavourable survey result, the best action and its EMV are:

(a) No plant, -10,000

(b) Large plant, -87,400

(c) Small plant, 2,400

(d) Small plant, 40,000

Answer: (c) - 0.27 x 90,000 + 0.73 x (-30,000) = 2,400 - small, positive, and better than -10,000 or -87,400.

Q50. The EMV of "conduct the survey" (node 1) is:

(a) 49,200

(b) 59,200

(c) 40,000

(d) 108,800

Answer: (a) - 0.45 x 106,400 + 0.55 x 2,400 = 47,880 + 1,320 = 49,200 - already net of the survey cost.

Q51. EVSI for the TLC survey is:

(a) Rs. 9,200

(b) Rs. 19,200

(c) Rs. 49,200

(d) Rs. 60,000

Answer: (b) - Value with the survey if it were free (49,200 + 10,000 = 59,200) minus 40,000 without it; 9,200 is the net gain after paying.

Q52. TLC's optimal strategy is:

(a) Skip the survey and build the small plant

(b) Survey; build large if favourable, small if unfavourable

(c) Survey; build large if favourable, nothing if unfavourable

(d) Survey; build small in both cases

Answer: (b) - The folded-back tree prescribes a complete policy, not a single action - that is what a table cannot do.

Q53. Which of the following is always true?

(a) EVSI is greater than or equal to EVPI

(b) EVSI is less than or equal to EVPI

(c) EVSI equals EVPI minus the survey cost

(d) EVSI is zero whenever the survey costs money

Answer: (b) - Imperfect information can never be worth more than perfect information: 19,200 against 60,000 here.

Q54. P(favourable market | survey favourable) = 0.78 and P(unfavourable market | survey unfavourable) = 0.73 are not symmetric because:

(a) The slide contains a typo

(b) A test's false-positive and false-negative error rates need not be equal - as with the COVID rapid antigen test

(c) 0.45 + 0.55 must equal one

(d) The survey costs money

Answer: (b) - Different error rates, and different stakes for each error, are the norm rather than the exception.

Q55. Why should a survey vendor's historical accuracy (say 70 percent) NOT be plugged straight into the tree?

(a) Historical data is always unreliable

(b) The probabilities must be customised to the current prior - a near-certain product and a 50-50 product cannot get identical survey probabilities

(c) Vendors always overstate accuracy

(d) Trees accept only round numbers

Answer: (b) - The COVID-vaccine example: a survey on a 90-10 product must point more confidently than the same survey on a coin toss.

Q56. With a prior of 0.5/0.5, P(survey positive | favourable) = 0.70 and P(survey positive | unfavourable) = 0.20, the posterior P(favourable | survey positive) is:

(a) 0.70

(b) 0.35

(c) 0.78

(d) 0.45

Answer: (c) - Joint probabilities 0.35 and 0.10 sum to 0.45; 0.35 / 0.45 = 0.78.

Q57. P(survey positive) = 0.45 was obtained as:

(a) 0.70 x 0.50 only

(b) 0.70 x 0.50 + 0.20 x 0.50

(c) 0.70 x 0.78

(d) 1 - 0.78 + 0.22

Answer: (b) - The total of the joint probabilities across both states of nature - the denominator in Bayes' theorem.

Q58. The Pilgrim Bank assignment asks whether to charge a fee or offer a rebate for online banking. The rebate makes sense when:

(a) Online customers are wealthier

(b) Online banking lowers the bank's servicing cost, so migrating customers online is worth paying for

(c) Customers refuse to pay fees

(d) Interest rates are falling

Answer: (b) - The fee logic applies when online banking is an additional facility on top of continued branch cost.

Q59. Sales data equals demand data only when:

(a) Prices were constant

(b) There was no stock-out

(c) Marketing spend was zero

(d) The Bass model was used

Answer: (b) - After a stock-out, sales undercount demand - the campus-session point the class recalled.

Q60. Ross selling last season's bag at a permanently lower price is:

(a) A markdown - a permanent reduction once demand falls short

(b) An event discount like Zara on Black Friday

(c) An EVPI exercise

(d) A stock-out

Answer: (a) - A true markdown never returns to the original price; Zara's Black Friday price does.

Q61. The professor's distinction between a fishbone diagram and a decision tree is:

(a) Fishbone is for root-cause analysis (chains of why); the tree structures the solution once the problem is defined

(b) Both solve the same problem in different notation

(c) Fishbone is for numbers; the tree is for pros and cons

(d) The tree is for root cause; the fishbone for options

Answer: (a) - Fishbone belongs before step one; the tree belongs after step four.

Q62. The CPWD architect's Canada decision illustrates:

(a) The value of perfect information

(b) A well-informed decision with a poor outcome (COVID) still being a good decision - process over outcome

(c) That personal decisions cannot use trees

(d) Sensitivity analysis

Answer: (b) - The tree in his diary showed he had let go of what he had chosen to let go of; COVID was not knowable.

Q63. The slide's five steps of decision tree analysis are:

(a) Draw, colour, present, argue, decide

(b) Define the problem; draw the tree; assign probabilities to the states of nature; estimate payoffs; solve by computing EMVs at the state-of-nature nodes

(c) List alternatives; compute EOL; buy information; survey; launch

(d) Sort, sum, average, chart, decide

Answer: (b) - Steps 3 and 4 put numbers on the tree; step 5 folds it back from the right.

Q64. Decision trees are most beneficial when:

(a) There is only one alternative

(b) A sequence of decisions must be made, such as survey first and then choose the plant

(c) Probabilities are unknown

(d) The payoff table has two columns

Answer: (b) - Any decision table can be drawn as a tree; a sequence of decisions cannot be held in a table at all.

Q65. According to the slide, probability values for a tree can come from:

(a) Only market surveys

(b) Management's experience and intuition, historical data, or computation from other data using Bayes' theorem

(c) Only historical data

(d) The payoff table

Answer: (b) - Bayes incorporates the initial estimate and the accuracy of the information source, and revises the estimate as new information arrives.

Q66. EVwPI is defined as:

(a) The best single payoff in the table

(b) The long-run average return if perfect information were available before each decision: the best payoff in each state, weighted by that state's probability

(c) The EMV of the best alternative

(d) The survey cost

Answer: (b) - For TLC: 0.5 x 200,000 + 0.5 x 0 = 100,000.

Q67. Google testing a new product first inside the team, then with family and friends, then with all employees, then in a beta is an example of:

(a) Decision making under certainty

(b) Converting uncertainty into risk with proxies - each stage produces an estimate you can put on a tree

(c) Sensitivity analysis

(d) Perfect information

Answer: (b) - Most decisions start in uncertainty; proxies move them into the risk column.

Q68. The kirana store owner facing a new dark store two streets away lists: start own delivery, list on a platform, or do nothing. Those three are:

(a) States of nature

(b) Courses of action

(c) Payoffs

(d) Probabilities

Answer: (b) - How heavily the neighbourhood takes to delivery is the state of nature; monthly profit is the payoff.

PART III From Data to an Estimate: Profitability, Sampling, Mean and Standard Deviation, Confidence Intervals (Online Session 3, 16 August 2026)

Session 3 was the first of the two statistics sessions, and it was taught entirely on the Pilgrim Bank dataset that had been our homework: should the bank charge a fee for online banking, or pay customers a rebate to use it? On the way to that question the class rebuilt the case's profitability curve in Excel, met the Pareto principle in customers, defects and inventory, learned why every business decision is about a population but is made with a sample, what makes a sample trustworthy, why an average is dangerous without its standard deviation, what the bell curve says, and how to turn a single number into a confidence interval - and then discovered that the two intervals overlap, so the data cannot yet say who is more profitable. Every number in this part was recomputed from the actual 31,634 rows.

By the end of this part you will be able to	Sections in this part
- Read and rebuild a profitability skew, and say who the "desirable" customers are - Explain the Pareto principle in customers, defects and inventory - and why retail is always on sale - Tell a population from a sample, and a representative sample from a convenient one - Say why 30 is "large", and what a larger sample buys - Report a mean with its standard deviation and read the bell curve's 68-95-99.7 rule - Build a confidence interval, choose its confidence level, and interpret overlap correctly	3.1 Where We Are: The Case, the Data and Two Myths 3.2 The Profitability Skew (Exhibit 3) 3.3 Letting Go of Customers, Politely 3.4 Pareto: Customers, Defects, Inventory 3.5 Why Retail Is Always on Sale 3.6 Population and Sample 3.7 Choosing a Sample, and the Number 30 3.8 What a Sample Is For; Kinds of Data 3.9 Mean, Standard Deviation, Bell Curve 3.10 Online vs Offline: The First Attempt 3.11 Confidence Intervals 3.12 Overlap, Outliers and the Age Cliffhanger 3.13 Product, Service and the Manager's Desk 3.14 Practice MCQ Bank with Answers

3.1 Where We Are: The Case, the Data and Two Myths

Session 2 ended with a homework: analyse the Pilgrim Bank dataset and make up your mind - fee or rebate for online banking. Session 3 opened by asking what we had found, and the first answer set up the whole class.

Part II's ideas travel with us, in a new costume. A good decision is still judged by its process - but this week the process is statistical: was the sample representative, was the average read with its spread, does the interval actually separate the two groups? The problem is still not "fee or rebate" but "are online customers more profitable than offline customers for the whole bank?" - and even that will turn out to hide an age effect. Fee and rebate are our alternatives; whether online customers are truly more profitable in the population is the state of nature we are trying to read from data. And information is still worth only what it changes: the professor's closing point this week is that the data we already have has more to say if we look at it differently.

How the class opened: "this data is not concrete enough to decide"

Venkatesh went first. His group had looked for a relationship between the columns and found none; the average profit per customer was about $111 but the standard deviation was about $272, which meant the customers ranged from heavily loss-making to very profitable; and the age and income columns had thousands of blanks. His conclusion: the data is too discrete and too incomplete to support a decision. Prachi gave a thumbs up - she had not made a decision either.

The professor agreed with the description and disagreed with the conclusion. She had chosen this case precisely because it is **rich and messy**: an old dataset, but a very real one. Missing values are the normal condition of business data. If you run a survey, do you get answers to every question? Some people go a step further and make every question mandatory - and then people either abandon the survey or type garbage. You want information; the other side is not obliged to give it. And you still have to decide.

> **NOT DECIDING IS USUALLY NOT ONE OF THE OPTIONS**
>
> Whatever the data looks like, a decision eventually gets taken - and if you postpone it, it often gets taken for you. Options expire. If a competitor starts offering a rebate for online banking, "charge a fee" quietly stops being an option, whatever your spreadsheet says. This is even more true in services, where there are no patents to protect a head start. Waiting for clean data is itself a decision, and usually a bad one.

Myth 1: "retaining customers is the goal"

Before touching the data the professor busted two myths that had crept into the homework discussion. The first was that a bank should offer the rebate because it retains customers and builds relationships. Retention and relationships are **means**, not ends. The end comes from the company's vision and mission and, in the medium run, its strategy. For a start-up whose valuation depends on market share, buying customers with giveaways can be a perfectly good idea. Beyond that phase, a for-profit company has to bring in profit; a permanently red bottom line is not a strategy. Many start-ups that do wonderfully in the growth phase show cracks a couple of years later - often after the founders have exited.

The deeper point is about what loyalty is. Customers seek value. You cannot satisfy them with freebies for long, because freebies are not sustainable and the customer who came for a freebie will leave for someone else's freebie. Solid loyalty comes from value delivered over a long time. It does not come from a rebate.

Myth 2: "the customer is God"

Growing up, the professor saw signs in shops saying customers are God and should be treated as such. They are not. **Customers are the means of the company's success** - which is exactly why some customers are worth keeping, some are worth changing, and some are worth letting go. The profitability curve in the next section makes that uncomfortable idea impossible to avoid.

The dataset, in one breath

Section 1.8 decodes the file column by column; the professor walked through it again in class. The short version: every column except the first is prefixed with 9, which means "the year 1999"; 9Profit is the dollar profit or loss each customer gave the bank; 9Online is a binary variable (1 = used online banking, 0 = did not - 3,854 ones and 27,780 zeros); 9Age (seven categories, 8,289 blanks) and 9Inc (nine slabs, 8,261 blanks) came from forms filled before KYC made such details compulsory; 9Tenure is complete because it comes from the bank's own database; 9District has three values. Why categories rather than precise ages and incomes? Think of a survey that insists on your exact income. Response rates collapse. Offer slabs - up to 5 lakh, 5 to 10, 10 to 15, 15 to 20, above 20 - and people answer. The bank traded precision for completeness, and it still ended up with a quarter of the rows blank.

> **THE PROFESSOR'S OWN WORKBOOK**
>
> The Excel file shared after the class contains her worked sheets: "Profitability Curve", "Point Estimate" and "Online-offline". Every number in this part was recomputed from the raw "Data" sheet and matches those sheets exactly. If your own numbers differ in the third decimal, you probably sorted or copied a column differently; if they differ in the first digit, check your cell references.

3.2 The Profitability Skew: Reading and Rebuilding Exhibit 3

Exhibit 3 of the case is a single curve that says more about the bank than any table. The professor asked everyone to reproduce it from the raw data before discussing it, so this section does the same: first how to read it, then how to build it.

How to read the curve

Sort the 31,634 customers from the most profitable to the least. On the horizontal axis put the **cumulative percentage of customers** - 10% means "the top 10% most profitable customers", 20% means the top 20%, and so on. On the vertical axis put the **cumulative percentage of the year's profit** that those customers brought in. The curve starts at zero (no customers, no profit) and must end at 100% (all customers, all the profit). What happens in between is the story.

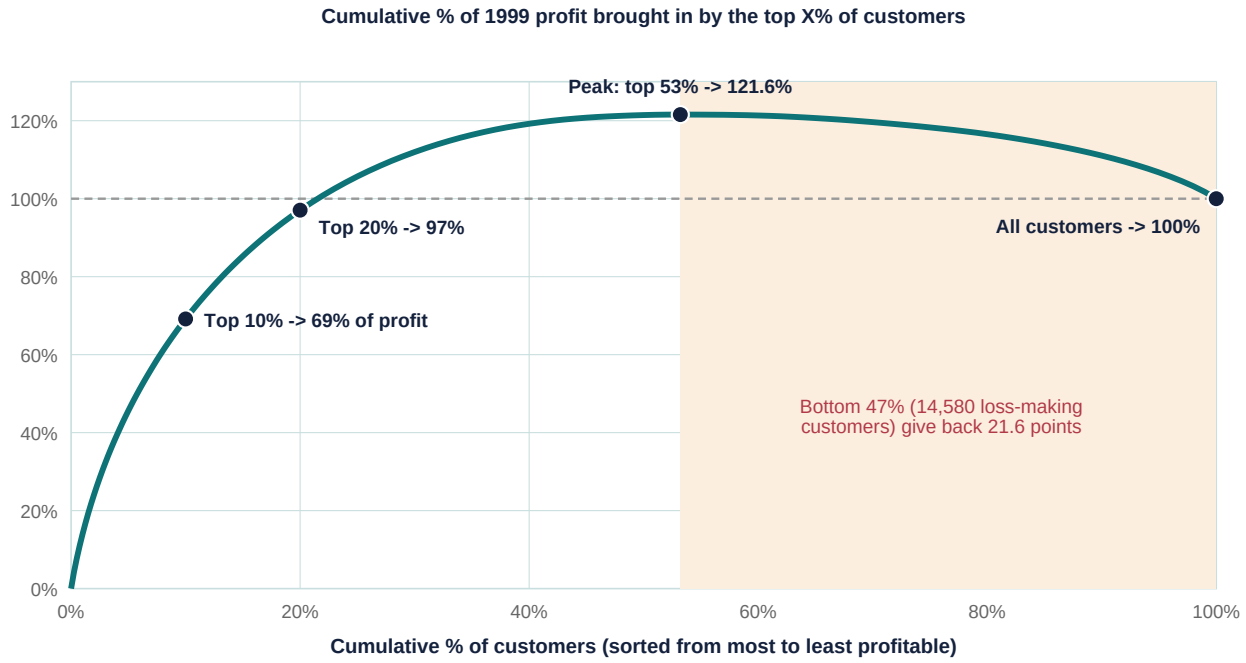


Figure 3.1. The profitability skew, rebuilt from the 31,634 rows. The curve rises fast, peaks at 121.6%, and then falls back to 100% as the loss-making customers are added.

Top X% of customers	Cumulative share of 1999 profit	In words
1%	14.0%	Three hundred customers bring one dollar in seven.
5%	45.6%	Fewer than 1,600 people bring almost half the profit.
10%	69.2%	The professor's "about 70%": one customer in ten pays for most of the bank.
20%	97.1%	Serve only the top fifth and you keep essentially all of the profit.
30%	112.0%	Beyond 100% - more profit than the bank actually made. Keep reading.
53% (the peak)	121.6%	The 16,832 customers with a positive profit. Everyone after this point costs money.
100%	100.0%	All customers: the real profit of $3,527,276.

CONCEPT TRAP: "HOW CAN PROFIT BE 121% IN THE SAME YEAR?"

It cannot - not as a total. The curve is cumulative and the customers are sorted by profit, so at the peak you have added up every profitable customer ($4,288,303) and none of the loss-makers. The 14,580 customers who follow lose the bank $761,027 between them, which drags the curve back down to the real total of $3,527,276. In other words, 21.6 points of the profit earned by the good customers went into cancelling the losses caused by the others. A curve that goes above 100% is not an error; it is the signature of a customer base with a loss-making tail.

This is the professor's definition of a **desirable customer**: anyone up to the point where the curve stops rising. As long as the curve climbs or stays flat, serving the next customer adds to profit. The moment it starts dipping, the next customer is taking profit away - and she would rather not serve that customer than keep them for the sake of "building a relationship". Whether a bank can act on that is the subject of Section 3.3.

From the case slides	Profitable customers (16,832)	Loss-making customers (14,580)
Average profit	$251.45	-$52.19
Standard deviation	$306.38	$44.77
Extreme	Maximum $2,071	Minimum -$221
What it tells you	The profitable customers vary enormously - a few giants and many modest earners.	The losses are small and similar: no single customer sinks the bank; the number of them does.

The slides read the same numbers two ways: the profit from profitable customers alone ($4,288,303) is a 21% increase over the actual total, or equivalently about 17% of the profit the good customers earn is spent on the loss-making ones ($761,027 of $4,288,303). And notice the spread: **profitable customers have much higher variation** than the loss-makers, which is why Section 3.9 will insist that no average travels alone.

A curious coincidence worth remembering: the top 20% of customers bring in $3,423,486, which is 79.8% of the $4,288,303 earned by all the profitable customers. The 80/20 rule of Section 3.4 is sitting right there in the data.

How to build it in Excel (the recipe from the class)

You need only the profit column. Put it in a new sheet so the raw data stays intact - a habit the professor insisted on. Then follow the steps below; the formulas assume the sorted profit is in column B starting at row 2, the serial number in column A, the cumulative profit in column C, and the two totals in cells F1 and E1.

Step	What to do	Formula or setting
1	Copy 9Profit into a new sheet and sort it from largest to smallest	Data > Sort, largest to smallest. The top value is 2,071; the bottom is -221.
2	Add a serial number column: 1 for the most profitable customer, 2 for the next, and so on	A2 = 1, A3 = A2 + 1, drag down to 31,634.
3	Add a cumulative profit column: the profit of the top 1, top 2, top 3 ... customers together	C2 = B2; C3 = C2 + B3; drag down. The value in row 10 is the profit of the top 10 customers.
4	Get the two totals	Total customers: =COUNT(B:B) = 31,634. Total profit: =SUM(B:B) = 3,527,276 - the same as the last cumulative value.
5	Convert to percentages, fixing the total with dollar signs so it does not move when you drag	%Customer: =A2/$E$1. %Profit: =C2/$F$1. Format both as percentages.

Step	What to do	Formula or setting
6	Plot	Select the two percentage columns; Insert > Scatter (markers only, no lines - lines make 31,634 points a mess).
7	Clean the axes	Show 20%, 40%, 60% ... rather than 0.2000. A chart is for a reader, so keep it readable.

The last row should read 100% customers and 100% profit; if it does not, a reference has slipped. The peak sits at about 53% of customers and 121.6% of profit - which the professor summarised as: **if the bank could let go of the bottom half of its customers, it would have made about 20% more profit last year.** That is not always possible, and it is never possible bluntly, so the next question is how to do it politely.

MANAGER'S TAKEAWAY

Draw this curve for whatever you manage - customers, contracts, products, tickets, alert sources. If it peaks well above 100% you are carrying a loss-making tail that your profitable core is quietly paying for. The curve tells you how big the tail is; the next sections tell you what you can and cannot do about it.

3.3 Letting Go of Customers Without Saying So

Can a bank pick and choose which customers to serve? Half the chat said yes, half said no, and both halves were right - it depends on how you do it.

Explicitly, no. Implicitly, all the time.

In India a bank cannot simply say "I will not serve you"; in the Western world refusing a customer would be treated as discrimination and could cost the bank its licence. So a bank cannot look at last year's losses and show a customer the door. But as Senthil pointed out, banks discriminate between customers every day through **preferential treatment**. The class's first examples - multiple channels, cheaper prices for online use, surprises for top-tier customers - were all about how to *keep* the good ones. The professor wanted the other half: how do you nudge the loss-makers out? Another student got there: charge them more for what they use, and they will either stop using it or leave.

The credit card: a fee that is really a filter

Around 2014-15 many banks were handing out credit cards with no annual fee. They then discovered that a lot of undesirable customers were signing up precisely because the card was free. So the premium cards acquired a fee of about Rs. 10,000 a year - waived if you spend above a threshold. The fee is not there to earn Rs. 10,000; it is there to keep the undesirable customers away. The waiver is there to make sure the profitable customers, who would earn the bank far more than the fee, never feel it. One rule for everyone, two very different effects.

The 27-minute wait: a story the professor is still sore about

Jet Airways ran a loyalty programme called JetPrivilege (later spun off as InterMiles). Her husband flew Jet constantly and had accrued a pile of miles; she rarely flew them. One day she booked her child's ticket with his miles, had to reschedule, and called the JetPrivilege helpline.

> *"Your expected wait time is about 27 minutes. You could also reach us by email, WhatsApp or the chatbot."*
>
> **What the IVR told the professor, a Blue-tier member**

She hung up, annoyed, and said so. Her husband picked up the phone, dialled the same number, and was talking to an agent in under a minute. Had the entire queue vanished the moment she hung up? Obviously not. He was a Gold or Silver tier member - a desired customer. She was an entry-level Blue - an undesired one. The airline was deliberately **disincentivising** her from using the precious call-centre agents and steering her to the cheap channels, so that the agents stayed free for the customers who mattered. The alternative would have been to hire many more agents, or to let her call degrade her husband's experience. They made the trade-off. That is how companies shed customers subtly.

What the bankers in the class added

Who	What they described	The mechanism
Subarna (banking)	Every customer gets a profitability score and is placed in a programme tier. If the score is not maintained, the customer is not thrown out but downgraded to the next tier, where the bank's cost of serving is lower and the facilities are fewer.	Graduated, not abrupt: the customer is weaned off expensive service rather than fired.
Vikram (a customer)	Basic, priority, prestige and so on - different account types with different facilities. Card applications are judged on the total relationship value (deposits, FDs, everything), not on one balance.	The bank prices service to the whole relationship, not to the transaction in front of it.
Gopal (insurance)	Two car-insurance claims in a year, together above Rs. 1 lakh - and the insurer refused to renew the policy the next year, with no explanation.	The explicit version. Legal in insurance, unthinkable for a bank account.

The professor added a caution to Gopal's story. These are the practices a *well-run* organisation would think through. Most organisations are not that deliberate; many of their processes were never aligned with what they are trying to achieve. It is common, in companies and in people, to believe you are working towards point A while your actions are carrying you towards point B, far away. The insurer may have had a policy - or just a faulty system.

Lever	How it reshapes the customer base	Where you have seen it
Fee with a spend waiver	Filters out low-value entrants; profitable customers never pay.	Premium credit cards; "no fee if you spend Rs. 3 lakh".
Minimum balance with a charge below it	Prices the cost of a small, active account back to the customer.	Savings accounts; monthly maintenance charges.
Tiered service queues	Scarce, expensive service goes to the customers who justify it.	JetPrivilege call routing; bank relationship managers.
Channel steering	Pushes low-tier customers to cheap channels (email, chat, app).	The IVR that offers WhatsApp before an agent.
Charging for what they use	Heavy users of costly touchpoints either change behaviour or leave.	Fees for branch transactions, printed statements, cheque books.
Downgrade rather than exit	Keeps the relationship while cutting the cost of serving it.	Subarna's tier downgrade.

Lever	How it reshapes the customer base	Where you have seen it
Non-renewal	The explicit exit - only where the law allows it.	Gopal's car insurance.

MANAGER'S TAKEAWAY

You cannot discriminate between customers by name, but you can design rules that apply to everyone and bite only where you want them to. A fee with a waiver, a tier with a threshold, a queue with a priority: each is one rule with two effects. Decide the trade-off consciously - who gets the scarce, expensive service and who gets steered elsewhere - or your systems will decide it for you, and not in the direction you intended.

3.4 Pareto: Customers, Defects and Inventory

The profitability skew is one instance of a pattern that shows up almost everywhere a manager looks: a small fraction of the things accounts for most of whatever you are measuring. It is named after Vilfredo Pareto, an Italian economist who noticed a century ago that about 20% of the people owned about 80% of the land.

The exact split is never 80/20 - in this bank it is closer to 20/97 - but the shape is the same, and so is the management response: do not spread your effort evenly across everything. Find the vital few, fix or protect them, and you get most of the gain for a fraction of the work.

Where the curve appears	The vital 20%	The trivial 80%	What you do about it
Customers and profit (service management)	The top fifth of customers bring 97% of Pilgrim Bank's profit.	The rest add little, and the bottom 47% actually subtract.	Protect and reward the top; steer or shed the tail (Section 3.3).
Quality - total quality management (TQM)	A handful of root causes.	The many symptoms, defects and unhappy customers they produce.	Fix the few root causes; most of the defects disappear with them.
Inventory	The 20% of SKUs that account for 80% of the inventory holding cost.	Thousands of items that cost little to hold.	Manage the expensive few tightly instead of trying to optimise every SKU.
A security operations centre (service example)	A few detection rules, log sources or assets generate most of the alert volume and most of the confirmed incidents.	Hundreds of rules that fire rarely and matter rarely.	Tune the noisy few; patch the exposed few; do not treat 400 rules as equals.
A product company's warranty returns	A few failure modes on a few components.	A long list of one-off complaints.	Redesign the components behind the few failure modes.

Inventory is where the professor spent longest, because inventory holding cost is one of the biggest costs in most companies and the biggest by far in fashion retail. To see why, she took a long detour through why your favourite store is always on sale. That detour is the next section; it is worth it because it ends with the idea of **alignment** that runs through the whole module.

MANAGER'S TAKEAWAY

Before you launch an improvement programme, draw the Pareto curve of whatever hurts - cost, defects, alerts, complaints, churn. Then work only on the part of the curve that is steep. Spreading effort evenly across a long tail feels thorough and achieves almost nothing.

3.5 Why Retail Is Always on Sale (and What Zara Does Differently)

Think about how often you saw a discount sale as a child. The professor's generation saw one a year: an off-season sale for woollens, because India's winter lasts two months and the stock is dead for the other ten. Now look around.

Marks and Spencer near her home has been on some offer for more than 200 of the last 300 days. Pantaloons, Shoppers Stop, Lifestyle, Max - permanently on sale. Myntra calls one of its sales the "End of Reason Sale", which is at least honest. Her grandmother had a theory: there is no real sale, they jack the price up and then cut it. Check it and it is not true - the same pair of jeans genuinely costs less on sale than off it. So why does a full-price retailer keep giving its margin away?

Two kinds of retailer, two reasons to discount

Retailer	What the sale is	Example
Discount as the business model	The bargain *is* the value proposition. The customer comes because it feels like a steal, and the sale is permanent by design.	Koutons ran 50% + 40% off (which is 70% off, not 90% - the 40% comes off the remaining half). Cantabil went to 50% + 49%, roughly 75% off.
Full-price retailer on sale	A symptom of poorly managed demand: stock that did not sell has to be cleared at a price often below its cost - a markdown, in Session 1's vocabulary.	Shoppers Stop, Pantaloons, Lifestyle, Marks and Spencer.

Why managing demand got so much harder

Who decides which men's formal shirts Shoppers Stop will carry next season? The merchandising and buying team, using two inputs the class named at once: past data and future trends (fashion weeks, what is happening in the world). Forecasting demand is old. What changed over the last four or five decades is the business around it.

Shift	What happened	Why it fills the store with unwanted stock
Consolidation of retail	Standalone mom-and-pop stores became chains run by corporates.	The owner of a small store knew his customers personally and bought for them. A corporate buyer - a highly paid fashion-school graduate buying for a value store like Style Bazaar - is not the target customer and may not understand them at all. Merchandisers in the US and India keep a sign on the wall: **"Remember, you are buying to sell, not to wear yourself."** If the customer likes bright yellow and you hate it, buy bright yellow.
Consolidation and offshoring of manufacturing	Apparel that was made in the West for the West in the 1960s-80s moved to China, Vietnam, Bangladesh, India and Pakistan because it was far cheaper.	You cannot order weekly from China for a European store - the logistics cost is prohibitive and the lead time is long (one student: orders go in at least six months ahead). So you commit early and cannot respond when a trend shows up. Result: your size or your shade is out of stock, while racks of what nobody wanted wait for the sale.

Zara: the same problem, a business built around solving it

Zara faces the same trends, the same lead times, the same uncertainty. Its USP is **fast fashion**, and everything is structured around that. Stores are replenished roughly every 15 days, mostly from centralised European production (the centralisation point came up in Session 1). Walk into a Zara: it is not crammed. There are few pieces on the floor, a back room attached to the store, fulfilment centres within the city and larger ones behind those. The spacious floor is deliberate - they want you to browse, because the more you browse the more likely you are to buy. Every piece of the system points the same way. That is what alignment looks like, and it is why Zara discounts less than the store next door.

There is an ethical shadow. Zara's buyers are said to roam the boutique districts of Paris and elsewhere looking for pieces that are selling well; a boutique designer can reach perhaps 12,000 customers, Zara can put a copy in front of millions within weeks. The supply chain that makes the model work also makes the copying pay.

"Reduce variety" - yes and no

Senthil wrote in the chat that the way to cut inventory is to cut variety. The professor's answer was "yes and no", and she asked the class for cases on each side.

Situation	Can you cut variety?	Why
Departmental store (Padmanabhan)	No	Drop the product a customer came for and they go to the store that has it - along with everything else they would have bought from you.
Dell computers	Yes, through delayed differentiation	Dell offered enormous variety - RAM, processor, screen, all your choice - and still won on inventory, because it stocked components, not finished laptops. The same component goes into many configurations; the product is differentiated at the last moment, to order.
Asian Paints	Yes, the same trick	Paint used to ship in every colour. Now the store keeps a colourless base and mixes pigments to order. Far fewer SKUs, and if anything goes past its expiry it is only a little pigment, not tins of paint.
DMart (Mahadev KT)	Partly	Fewer brands per category, concentrating on the ones that sell - plus the bargaining power of a big buyer, so suppliers work on DMart's terms, as with Costco and Walmart.

Spencer's and Jiffy: starting a channel is easy, aligning it is not

Spencer's, the RPG (Sanjiv Goenka) group's supermarket chain, launched a quick-delivery service called Jiffy only about two years ago - long after BigBasket and Grofers had built the category - and is still struggling, even in its home market of Kolkata. Many companies bolted an online channel onto their physical business during and after COVID; some shut it down, some deliver poor service. Starting a new channel is not the problem. Making sure every process behind it is aligned with the promise is where most businesses fail.

MANAGER'S TAKEAWAY

When your business keeps "going on sale" - discounting, over-delivering, firefighting - it is usually not a pricing problem but a demand-management and alignment problem. Ask whether the people making the buying, staffing or capacity decisions actually know the customer, how long your lead times are, and whether every process points at the same promise. Zara wins on alignment, not on cheaper cloth.

3.6 Population and Sample: Why We Never Have All the Data

How many customers are in the Excel sheet? 31,634. How many customers does Pilgrim Bank have? The case says five million. So the decision we are about to make - fee or rebate, for everyone - rests on data about fewer than 1% of the people it will affect. Is that allowed? That question is the whole of statistics.

Two words that carry the rest of the module

Term	Definition	In the Pilgrim case	In other studies
Population	The set of **all** the elements that are **of interest to us**. Both halves matter: all of them, and the ones this study cares about.	The bank's five million customers - the people the fee or rebate will apply to.	For an inventory study the elements are SKUs; for a defects study they are the possible root causes. What is the population for one study is not the population for another.
Sample	A subset of the population that we actually study - survey, test, observe or experiment on.	The 31,634 customers in the file: about 0.6% of the population.	The 1,000 people in a shampoo survey; the 200 bulbs burned to failure; the 70 EPBM-32 participants if the question is about executive students in general.

THE ONE-LINE TEST

Everything we do - every average, every interval, every decision - is **about the population**. The sample is only a stepping stone, used because we cannot reach the population. Ask of any statistic: which population is this a guess about, and how was the sample that produced it chosen?

Why not just get data on everyone?

Ideally you would. If it is within budget and within the time you have, the whole population beats any sample. Three things stop you.

Reason	What it means	The professor's examples
It is expensive	Reaching every element costs more than the decision is worth.	Procter and Gamble wants to know what customers think of one shampoo. Some bought it in a supermarket, some from a corner shop, some used a sachet in a hotel. Tracking every user down is possible in principle - but if the change you hope to make is worth a 5% gain and the data costs 10%, it will never pay.
It takes too long	By the time the full picture arrives it is out of date.	Ideally a census would run every year, because things change every year. It does not, because even if you spent the money the results would arrive with such a lag that they would be of little use.
It is not possible	Some elements will not, or cannot, give you data.	Customers can simply refuse to talk. Those who talk may say one thing, feel another and do a third. And in **destructive testing** the act of measuring destroys the unit - see below.

Destructive testing: when measuring the population would leave nothing to sell

The late actor Irrfan Khan fronted an advertisement for Syska LED bulbs claiming they would run for 3,000 hours before failing. Tyre makers claim a tyre will not fail before 5,000 kilometres. How would a company know? It lights a bulb and leaves it on until it dies, noting the hours. That bulb is now useless. Test every bulb and you have nothing left to sell - so you test a sample and infer.

The claim itself is a decision under uncertainty. Claim 3,500 hours and most bulbs fail earlier: the free replacements you promised will bankrupt you. Claim 2,500 hours when the bulbs really last 3,000: you have left an opportunity on the table. The sample tells you where to draw the line - and, as Section 3.8 will show, gives you the evidence to defend it in court.

Her favourite example is Honda's Activa, launched in 2001. Until then the two-wheeler market was neatly segmented: scooters (Bajaj Chetak, Bajaj Super, LML Vespa) for the family man; scooterettes (TVS Scooty, Bajaj Sunny, Hero Puch, Kinetic Honda) for school and college girls; and bikes (Bajaj Caliber, Hero Honda Passion, the Pulsar) for young men who would not be seen on a scooterette. The Activa looked like a scooter, was priced like a bike - above Rs. 50,000 when scooters cost Rs. 34,000 to 38,000 - and did not even carry a spare wheel. Why would anyone buy it? Honda turned the missing spare wheel into the talking point: our tyres will not fail before about 5,000 kilometres, so you do not need one. A claim like that can only come from destructively testing a sample of tyres. The Activa went on to erase the old gender lines of the segment entirely.

A less glamorous case: Johnson and Johnson's large facility in Tamil Nadu makes sutures - the thread that closes a body after surgery. To know how much strain a suture can take, you pull it until it breaks. The tested sutures are gone; the untested ones ship, with the confidence the sample bought.

MANAGER'S TAKEAWAY

Before you trust a number, name the population it claims to describe and ask why the sample was used instead. "Expensive", "slow" and "impossible" are all legitimate reasons - but each of them means the number in front of you is an inference, not a fact, and the next two sections are about how good an inference it is.

3.7 Choosing a Sample: Random, Stratified, Convenience - and the Number 30

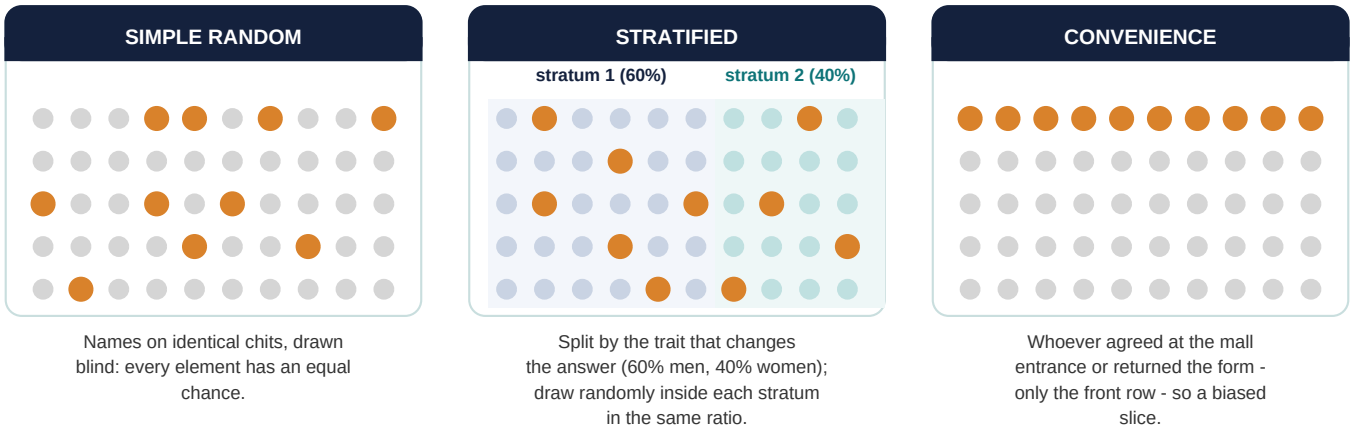
A sample is a subset of the population. Which subset? Any subset gives you numbers; only a good subset gives you numbers you can trust. The first and most important condition is that the sample be representative - and the class discovered that it had the wrong idea of what that word means.

Representative does not mean "the same mix"

The professor set a trap. A population has 10,000 people: 5,000 men and 5,000 women. You take a sample of 100. Must it contain 50 men and 50 women? Most of the class said yes, or "preferably". The answer is: **in general, no**. Omkar had it right - it depends on the question. If men and women are likely to answer differently - a study of gender bias at the workplace - then yes, balance the sample the way the population is balanced. If the question has nothing to do with gender, forcing a 50/50 split actually introduces a bias.

Why? Because gender is only one of infinitely many characteristics people have. Why balance for gender and not for eye colour, hair texture, skin colour, or whether they prefer sweets to savouries? You cannot balance a sample on infinitely many things. So statisticians define representativeness by the *method* instead of the composition:

> *"A sample is representative of the population if it was chosen by a method such that every element of the population had an equal chance of being included in it."*
>
> **The definition to memorise - it is about the method, not the mix**



Orange dots = the 10 elements that end up in the sample; grey dots = the rest of the population.

Figure 3.2. Three ways to pick ten elements from a population of fifty. Only the first two give every element a fair chance; the third takes whoever was easy to reach.

Method	How it works	When to use it	Watch out for
Simple random sampling (SRS)	Every name on an identical chit, folded, into a fishbowl, drawn blind - the party game from childhood. With five million customers you let software draw the 31,000.	The default. Whenever nothing about the question demands otherwise.	Nothing, as long as the draw is truly blind. It is the standard the other methods are measured against.
Stratified sampling	Split the population into strata by the trait that changes the answer (men and women; urban and rural; corporate and retail customers). Draw a simple random sample *inside each stratum*, keeping the same ratio as the population: if the population is 60/40, the sample is 60/40.	When different strata are likely to respond differently to the question you are asking.	Stratifying on a trait that does not matter for the question creates the bias you were trying to avoid.
Convenience sampling	You reached out to a random list, but not everyone replied. Or you stood at the mall entrance with a questionnaire and most people walked past. You use whatever data you managed to get.	You do not choose it - you end up with it. Most market research is convenience sampling.	Bias. The people who responded may differ systematically from those who did not. Use with great care and larger error margins.
Rare-event sampling	Special methods for phenomena that occur, say, once in a million - a rare disease. A random draw would simply never find the cases.	Only when the thing you study is extremely rare. Most of us will never need it.	Mentioned for completeness; the formulas of this module assume the ordinary methods above.

What bias looks like: the Rs. 75-lakh alumni

A school organised the 25th reunion of the batch that graduated in 2000 and wrote to everyone asking, among other things, how much they earned. It then announced that its alumni earned Rs. 75 lakh a year on average, twenty-five years out of school. Would you believe it? The class did not. Put yourself in the shoes of an alumnus whose career has not gone well and who knows the others are doing better: do you write back with your salary? Now imagine two classmates who founded start-ups that became unicorns: they do reply, and their income alone drags the average into the stratosphere. One segment of the population was far more likely to respond than another. That is **bias**: the sample no longer gives a general feel of the population it was meant to describe, because a small slice crowds out the picture.

How big must a sample be? The number is 30, not 30%

Back to the elephant in the room. 31,634 is 0.6% of five million - not even 1%. Can we use it to decide for everyone? What would be the minimum? Several people said 10% - "that is what our organisation uses". The answer surprised most of the class: **a sample of size 30 or more is, statistically, a large sample**, and the formulas of this module - estimation, hypothesis testing and the rest - apply to it. Thirty is an absolute number, not a percentage, and it does not depend on the size of the population.

Why 30? Because of the **central limit theorem**. Take a sample and compute its average - say the average profit per customer. Take another sample and compute its average again. Do this many times and the averages you get start to follow the bell-shaped curve (Section 3.9) once each sample has about 30 or more elements, almost regardless of what the underlying data looks like. That regularity is what the formulas lean on, which is why 30 is the threshold across industries and geographies, not a convention of one company.

CONCEPT TRAP: "LARGE SAMPLE" MEANS n >= 30, NOT 30% OF THE POPULATION

The class jumped to percentages because that is how most companies talk. Statistics does not. With 31,634 rows we are a thousand times past the threshold. What a larger sample buys is not permission but precision: the cushion around every estimate shrinks as the sample grows (Section 3.11). And if you are stuck below 30 - a rare event, a tiny pilot - you are not helpless; there is a separate body of small-sample statistics for exactly that.

> *"You can always go wrong, whatever the size of the sample. That is why it is statistics and not mathematics."*
>
> **Prof. Megha Sharma, on why a sample never gives certainty**

MANAGER'S TAKEAWAY

When someone brings you a survey, a pilot result or a customer study, ask two questions before you look at the numbers: how were the participants chosen, and who did not respond? A convenience sample of the people who bothered to reply - the loudest complainers, the happiest fans, the unicorn founders - can be worse than no data at all, because it looks like evidence.

3.8 What a Sample Is For, and Three Kinds of Data

We take a sample for exactly three purposes, and the whole of business statistics in this module is those three. Each one is about the population; the sample is just how we get there.

Purpose	The question it answers	Examples from the class
1. Estimation	What is my best guess of an unknown number about the population - its **average**, its **proportion**, or its **standard deviation**?	The average profit per customer across all five million (this session). The average monthly grocery spend of a bank's customers - which a retailer like Lifestyle would love to know, because it tells them the size of the wallet they are competing for. The proportion of voters in favour of a policy: nobody asks every voter, so we run opinion polls and exit polls.
2. Hypothesis testing	Can I **support a claim** I want to make, or **reject a claim** somebody else has made, with evidence from a sample?	Syska claims its LED lasts 3,000 hours; a consumer advocacy group takes them to court because a bulb lasted 2,500. Syska never claimed every bulb would; it needs sample data that supports the claim. That is why advertisements now carry the fine print - "tested on 1,000 subjects under laboratory conditions" - from Horlicks (taller, stronger, sharper) to hair-fall shampoos and Kesh King.
3. Relationships (regression)	How does one variable move with another, so that I can work backwards from a target?	You have a first-quarter sales target for a new product and want to know how much pre-launch promotion to spend. If past data says sales are such-and-such a function of promotion, you can solve backwards: a target of Rs. 50 crore might imply about Rs. 11 crore of promotion. A film's first-weekend box office is driven by pre-release promotion - trailers, music launch, the cast on talk shows - and the same logic sets the promotion budget.

This session did estimation. Session 4 does hypothesis testing on the same data. Regression comes later in the module.

Three kinds of data, and why the Pilgrim data is the weaker kind

The question we really want answered is causal: if a customer moves online, does the bank's profit from that customer go up? The ideal way to answer it is an **experiment**. Keep one group offline as a control group, ask another group to bank online, and track both groups over time. That would give you the strongest kind of evidence - and it would take five to ten years, by which time the technology itself would be obsolete. So the case does what most businesses do: it uses a snapshot as a proxy, comparing the customers who happen to be online today with those who are not. Understanding the three kinds of data tells you what that proxy can and cannot prove.

Kind of data	What it is	Examples	What it can tell you
Cross-sectional	A snapshot: many elements observed at **one** point in time.	The Pilgrim file - 31,634 customers, all in 1999. The 2025 GDP of 200 countries. Units sold of Maggi, Yippee and other noodles on 31 December 2025.	How elements differ from each other right now. It cannot by itself show what happens to one element over time.
Time series	The **same variable** observed at many points in time.	India's GDP from 1991 to 2025. Maggi packets sold in 1999, 2000, 2001 ... - yearly, quarterly or daily. The demand history your company uses to forecast.	How one thing changes over time - trend, seasonality, the effect of an intervention on the same customer.
Panel	Both at once: many elements, each followed over time.	The GDP of all 200 countries, every year from 1991 to 2025. The same customer IDs observed in 1999, 2000, 2001 ...	The most powerful of the three: for each element a time series, for each year a cross-section, and the whole picture together.

THE HYPOTHESIS, IN ONE LINE

Because we are using cross-sectional data as a proxy, the question changes shape. We cannot check whether customers *became* more profitable by going online. We can check whether the customers who are online are **more profitable than the customers who are offline**. If they are, the rebate is worth paying; if they are not, the bank should recover its cost with a fee. Everything from Section 3.10 onwards is an attempt to answer that one comparison honestly.

Setting	Cross-sectional use	Time-series use	Panel use
Product company	Defect rate by supplier this quarter.	Defect rate of one supplier, month by month.	Every supplier, every month - which one improved after the audit?
Managed security services	Mean time to detect, by customer, this month.	One customer's detection time over 24 months.	All customers over 24 months - did the new tooling help everyone or only the large ones?

3.9 Mean, Standard Deviation and the Bell Curve

After the break the class copied the profit column into a fresh sheet and did the simplest possible piece of estimation: what is our best guess of the average profit per customer across all five million customers?

The point estimate: $111.50, and why it is the "best" guess

We know nothing about the five million except the 31,634 in front of us. So the best guess of the population average is the sample average: **=AVERAGE(profit column) = $111.50**. Not because it is certainly right, but because it is the only information we have - it beats pulling a number out of the air. A single number offered as the guess for an unknown population quantity is called a **point estimate**.

Parameters are fixed; statistics wobble

The professor then asked a question that sounds pedantic and is not. The average profit per customer for the entire population in 1999 - is that a fixed number or a random variable? It is fixed. If we had all five million rows we would compute it and everyone would get the same answer. We do not know it, but it exists and it does not move. Quantities like that are called **population parameters**: mean, standard deviation, proportion. Unknown, but constant.

The sample average, on the other hand, **is a random variable**. Think of EPBM-32 as a population of roughly 70 people and suppose the professor wants the batch's average work experience. The true answer is fixed as of today. But if she is in a hurry and takes the first few people who appear on her screen, her sample average depends on who turned up. If the most senior people happen to appear, the sample average is high; if the youngest do, it is low. Because the draw is random, the sample average has a probability attached to it. Once a particular sample is drawn its average is a fixed number - but which sample you drew was chance.

	Population parameter	Sample statistic
Examples	Population mean, population SD, population proportion	Sample mean, sample SD, sample proportion, sample minimum and maximum
Nature	A constant. Unknown to us, but fixed.	A random variable: it changes with who lands in the sample.

	Population parameter	Sample statistic
Symbols used in class	Mean: the Greek letter mu; SD: sigma	Mean: x-bar (x with a bar over it); size: n
Pilgrim, 1999	The true average profit of all five million customers - never observed	$111.50 from these 31,634 customers; another random 31,634 would give a slightly different number

Never leave an average alone

Somebody had said it at the very start: an average by itself can be very misleading. Here is the professor's way of seeing why. Take two batches. One is EPBM-32, where work experience runs from early-career to seasoned; the other is a niche programme that admits only people with three to five years of experience. Take a sample of the same size from each and compute the average. Which average is more likely to be close to its batch's true figure? The niche batch, obviously: everyone is between three and five, so the scope for going wrong is tiny. In EPBM-32 the sample could land on the veterans or on the rookies.

That property - how alike or how spread out the values are, how homogeneous or heterogeneous the group is - is what the **standard deviation** (SD) measures. The higher the SD, the less trust you can place in the average, because the true value could be far from it.

How the standard deviation is computed

Take the five numbers 1, 2, 3, 4, 5. Their average is 3. To measure spread, find how far each number is from the average: -2, -1, 0, +1, +2. Averaging those distances would be the natural next step, except that they always add up to zero - the average is the balance point, the centre of gravity of the data, so the negatives and positives cancel exactly. The fix is to **square** the distances so that everything becomes positive.

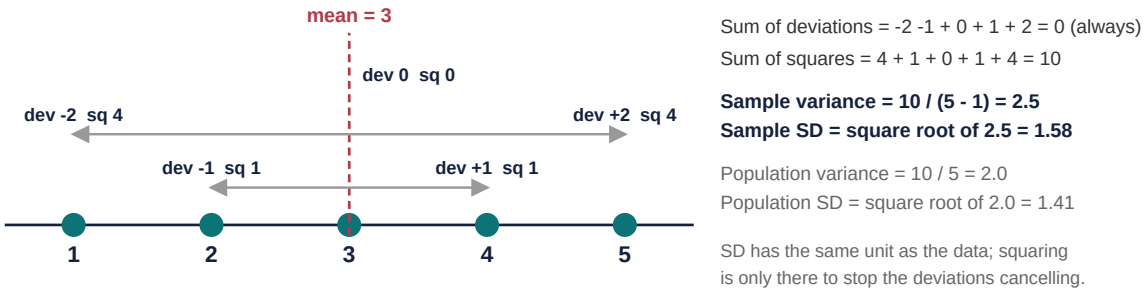


Figure 3.3. Variance and standard deviation for the numbers 1 to 5. Squaring stops the deviations cancelling; the square root brings the answer back into the units of the data.

Quantity	Recipe	Excel	Unit
Mean (x-bar)	Add all n values and divide by n.	=AVERAGE(range)	Same as the data ($)
Variance	Take each value minus the mean, square it, add them all up - then divide by **n** if the data is the whole population, or by **n - 1** if it is a sample.	=VAR.S(range) or =VAR.P(range)	Squared units ($ squared) - awkward
Standard deviation	The square root of the variance.	=STDEV.S(range) for a sample; =STDEV.P(range) for a population	Back to the units of the data ($)

In Excel, STDEV.S is the one you will almost always want - S for sample. It divides by 31,633 for our data; STDEV.P would divide by 31,634. The difference is negligible with 31,634 rows and matters with 8. For the profit

column: **=STDEV.S = $272.84**.

> **WHY n - 1 FOR A SAMPLE? (A QUESTION FROM THE CLASS)**
>
> Two answers, one short and one intuitive. The short one: dividing by n - 1 makes the sample variance an **unbiased** estimate of the population variance - on average, across all possible samples, it lands on the true value; dividing by n systematically underestimates it. The intuitive one is about **degrees of freedom**. Suppose n = 10 and I tell you the sample mean is 32. Now I show you nine of the ten values. You can work out the tenth yourself - it is whatever makes the mean 32. So once the mean has been used, only nine of the ten numbers are free to vary; you have lost one degree of freedom, and you divide by 9. For a population there is no estimated mean standing in for the true one, so nothing is lost.

When is a standard deviation "too high"? Only relative to the mean

Is $272.84 a large standard deviation? On its own the number means nothing. If the average profit per customer were $50,000, an SD of $273 would be trivial - the customers would all be very alike. Against an average of $111.50 it is enormous. To decide where "large" begins, the professor turned to the shape that most data in nature follows.

The bell curve

Measure the height of every five-year-old banyan tree in India. They differ, but plot how many trees are 3 feet tall, how many 3.5 feet, and so on, and you get a hump: very few very short trees, very few very tall ones, most somewhere in the middle. IQ behaves the same way. So do countless other things that occur in nature - which is why this shape is called the **normal distribution** (also the Gaussian distribution, after Gauss, and informally the bell curve). It has two properties the class needs.

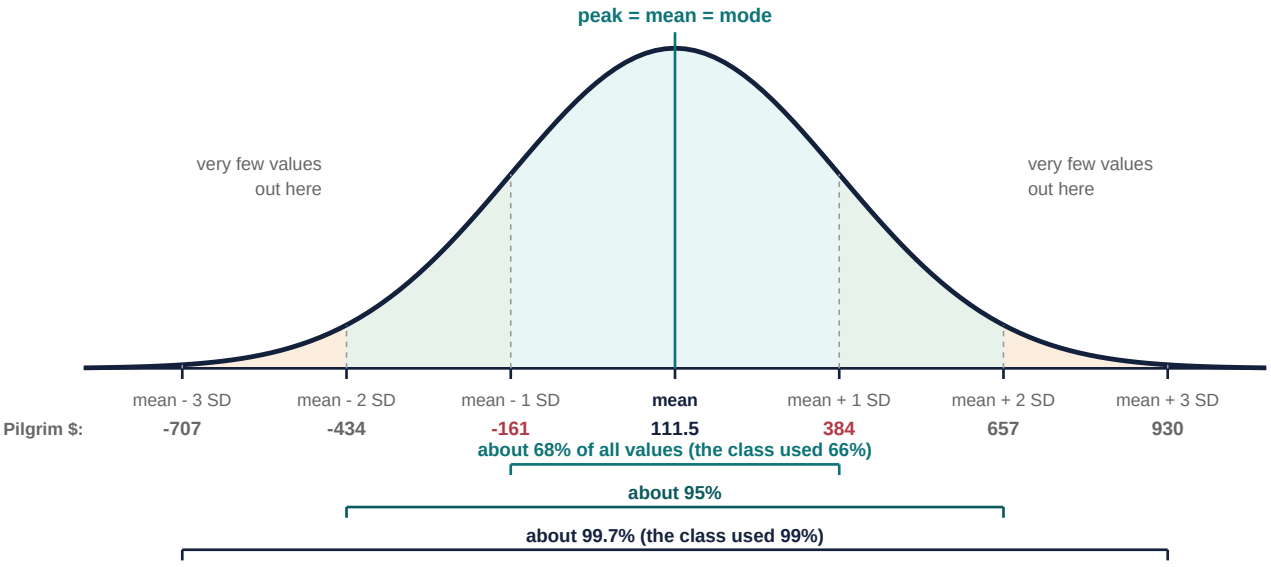


Figure 3.4. The normal distribution, with the Pilgrim profit numbers ($111.50 mean, $272.84 SD) written along the bottom. Even the middle band runs from a $161 loss to a $384 profit.

Property	What it says	Applied to Pilgrim profit ($111.50, SD $272.84)
The peak is at the mean	The most frequent value (the mode) is also the average.	The typical customer sits near $111 - if the data were bell-shaped.

Property	What it says	Applied to Pilgrim profit ($111.50, SD $272.84)
Mean +/- 1 SD	Holds about 68% of all values (the professor used 66% - close enough for a rule of thumb).	-$161 to +$384: the middle two-thirds of customers lie somewhere between a loss of $161 and a profit of $384.
Mean +/- 2 SD	Holds about 95% of all values.	-$434 to +$657
Mean +/- 3 SD	Holds about 99.7% (the professor said 99%).	-$707 to +$930

Now the rule of thumb. Assume, roughly, that your data follows the normal curve and compute mean plus and minus one SD. If that range is too wide to support the decision you want to make, the SD is high and the average cannot be trusted much. Here the range is -$161 to +$384. Forget fee versus rebate - with that spread you cannot even say whether the bank's customers are, on average, profitable. Now imagine the same mean with an SD of $20: the range would be $91.50 to $131.50. You would know your customers are profitable and roughly by how much, and you could decide with some confidence.

> *"Never, ever give anybody an average without the standard deviation. If somebody submits a report to you with an average, ask them for the standard deviation - otherwise the number could be quite misleading."*
>
> **Prof. Megha Sharma - the rule of the session**

> **A SMALL HONESTY NOTE ABOUT THE RULE OF THUMB**
>
> The profit data is not actually bell-shaped: the median profit is $9 while the mean is $111.50, which tells you the distribution is skewed by a tail of very profitable customers (Section 3.2). The professor was explicit that the +/- 1 SD range is "just a quick rule of thumb" to sense whether the spread is uncomfortably large. It is - and that is all the rule was asked to tell us. The central limit theorem is what makes the confidence intervals of Section 3.11 valid even when the data itself is not normal, because they are about the sample *average*, which is.

> **MANAGER'S TAKEAWAY**
>
> **An average is a claim about the typical case; the standard deviation says how typical the typical case is. Report both, always: "MTTR is 42 minutes, SD 65" is a completely different message from "MTTR is 42 minutes, SD 4". The first says your response times are all over the place and the average hides a long tail; the second says the process is under control.**

3.10 Online vs Offline: The First Attempt

Now the real question. Copy the profit column and the 9Online column into a new sheet, sort by 9Online so that all the zeros come first, and insert a blank row where the ones begin - at row 27,780 or thereabouts. Above the blank row are the offline customers; below it the online ones. There are smarter ways in Excel, but this one works for everybody.

Step	What to do	Excel
1	Copy 9Profit and 9Online into a new sheet	Two columns, side by side.
2	Sort both columns together by 9Online, smallest to largest	Data > Sort by 9Online. Zeros on top, ones at the bottom.

Step	What to do	Excel
3	Insert a blank row where 0 turns into 1	Scroll to about row 27,780; Insert > Row.
4	Select the offline profits in one go	Click the first offline profit, then Ctrl + Shift + Down Arrow: the selection stops at the blank row.
5	Compute the three numbers for each group	=AVERAGE(...), =STDEV.S(...), =COUNT(...) for the offline block, then the same for the online block.

Group	Customers (n)	Average profit	Standard deviation	Mean - 1 SD to mean + 1 SD
Offline (9Online = 0)	27,780	$110.79	$271.30	-$160.5 to +$382.1
Online (9Online = 1)	3,854	$116.67	$283.66	-$167.0 to +$400.3
Everyone	31,634	$111.50	$272.84	-$161.3 to +$384.3
Difference, online minus offline		$5.88 (about 5.3% higher)	Both around $270-280	The two ranges almost coincide

"Why not drop the loss-making customers first?" - a question from the class

A student suggested keeping only the customers with positive profit, on the grounds that the loss-makers are a liability and would distort the comparison. The professor pushed back on two levels. First, those customers do not leave the portfolio because you left them out of a spreadsheet. If the bank offers a rebate they will use it too; if it charges a fee they will pay it or leave. Either way they are part of the population the decision applies to, so they belong in the sample. Second, you cannot make one rule for the profitable and another for the rest - "you gave me a loss, so you pay a fee" is exactly the discrimination Section 3.3 ruled out. What you can do is set one rule with a threshold - a Rs. 5,000 card fee waived at Rs. 3 lakh of spend, a minimum balance below which a charge applies - which is legitimate precisely because it is the same rule for everybody.

Why a customer makes a loss	What it looks like	What online banking might do
Too little revenue to cover the cost of serving them	A small balance, and a trip to the ATM to withdraw Rs. 100 - the professor's hostel days, when nobody withdrew more because a visible note meant paying for everyone's dinner.	Little - the revenue problem is the balance, not the channel.
Plenty of balance, but a high cost to serve	Uses online banking and the ATM and still visits the branch weekly for a printed statement, occupies the teller, and calls the contact centre.	A lot - every visit and call that moves to the app is cost the bank no longer incurs.

Which leads to a point the fee-or-rebate debate often misses: a rebate does not have to turn everyone profitable to be worth paying. If a customer who costs the bank $50 offline costs it $30 online, going online is good for the bank even though the customer is still a loss. **Reducing losses counts.** Keep that in mind for the cliffhanger in Section 3.12.

So - are online customers more profitable? Only in the sample

Several people typed "yes, online, $116 against $110 - offer the rebate". The professor's answer: that is true of the sample, and we are not deciding for the sample. Two problems stand between the table and a decision.

Problem	Why it blocks the conclusion
The averages are close	$116.67 against $110.79 is a gap of under 6%. Small gaps are easily produced by chance.
The standard deviations are huge	Around $270-280 in both groups, against averages of about $110-117. With that much spread, the sample averages could easily have landed the other way round.

Remember that sample means are random variables. It is entirely possible that, by the luck of who ended up in this sample, online customers look more profitable here while in the real population of five million they are less profitable. If that is the case and the bank pays a rebate, it pays money to customers who were already costing it more - and profitability falls twice over. If the SDs had been $2 or $3, or if the online average had been $298 instead of $116, the gap would have been far too large to blame on sampling luck and the decision would have been easy. With these numbers the honest verdict is: not yet.

CONCEPT TRAP: "THE SAMPLE SAYS ONLINE WINS, SO ONLINE WINS"

A difference in sample averages is evidence, not proof. Whether it is enough evidence depends on how big the gap is relative to the spread and the sample sizes. The tool that makes that judgement explicit is the confidence interval - the subject of the next section, introduced with a pen and a notional gun.

MANAGER'S TAKEAWAY

When two groups differ by a few percent on a metric with a large spread - team A closes tickets 6% faster, region B converts 5% better - do not act on the difference until you have asked whether it would survive a different sample of the same size. Most small differences in noisy data are noise.

3.11 Confidence Intervals: From a Number to a Range

A point estimate is a single best guess. The class had just seen why a single number is not enough to decide. The professor introduced the alternative with the strangest exercise of the module: she asked everyone to switch on their cameras.

The pen at gunpoint

Guess the price of the pen she was using on the smartboard. One catch: you are at gunpoint - a notional gun, she is a non-violent person, but assume the threat is credible. Guess right and you walk free. She is kind, so you may give either a single number or a range. Type it in the chat. Every single person typed a range. Why? Because a range gives you a far higher chance of being right - a single number is almost certain to be wrong. And if you were thinking clearly, why not the widest range of all, minus infinity to plus infinity? That makes the probability of being wrong exactly zero.

Two lessons were hiding in the joke. First, when the stakes are high you demand high accuracy - and you buy that accuracy by widening the range. Second, a range wide enough to be certainly right is useless: "the profit is somewhere between minus infinity and plus infinity" helps nobody decide anything. Everything about confidence intervals is the trade-off between those two.

Point estimate plus or minus a cushion

Look out of the window and estimate the chance of rain today. "About 70%, give or take 5%." That is a confidence interval: a point estimate (70%) and a cushion (5%) added on each side, giving a range from 65% to 75%. Formally, a **confidence interval** (also called an **interval estimate**) is a range within which we believe the unknown population parameter lies, stated together with how confident we are - the level of accuracy we chose. Both words matter: it is an *interval*, with a lower and an upper limit, and it carries a *confidence*, chosen by us.

Confidence level	The multiplier z (from the standard normal table)	When to use it
90%	1.645	When the cost of being wrong is modest and you want a tight range.
95%	1.96	The default in most business analysis. Use it unless you have a reason not to.
99%	2.576	When the cost of going wrong is very high - the range widens to buy safety.
100%	(infinite)	Never. The interval becomes minus infinity to plus infinity - the pen answer that keeps you alive and tells you nothing.

The level you choose is a business decision, not a statistical one: the more it would cost to be wrong, the more confidence you buy. At gunpoint the class bought 100% and paid for it with a useless answer. In the Pilgrim case a wrong rebate costs money but not lives, so 95% is the usual compromise.

What sets the size of the cushion

Driver	Effect on the cushion	The professor's illustration
The confidence you want	More confidence, bigger cushion.	At gunpoint you went for the widest range you could.
The standard deviation of the data	More spread (heterogeneity), bigger cushion.	Estimating average work experience: the niche batch (everyone between 3 and 5 years) needs a small cushion; EPBM-32, with rookies and veterans, needs a big one.
The sample size	More data, smaller cushion - and the population itself needs none.	Sample 3 of the 70 in the batch and you might land on the three youngest or the three most senior - a gross under- or overestimate. Sample 65 of 70 and the extremes balance out. Sample all 70 and you have the true number: no cushion at all.

THE FORMULA (FOR A POPULATION MEAN, LARGE SAMPLE)

Confidence interval = x-bar +/- z x sigma / square root of n

x-bar is the sample mean (the point estimate); z is 1.645, 1.96 or 2.576 for 90%, 95% or 99%; sigma is the standard deviation (in practice the sample SD from STDEV.S); n is the sample size. The whole term after the +/- is the cushion. Subtract it for the lower limit, add it for the upper limit. Notice where each driver sits: z and sigma are on top (bigger means wider), n is underneath (bigger means narrower), and it is the square root of n, which is why the cushion shrinks slowly.

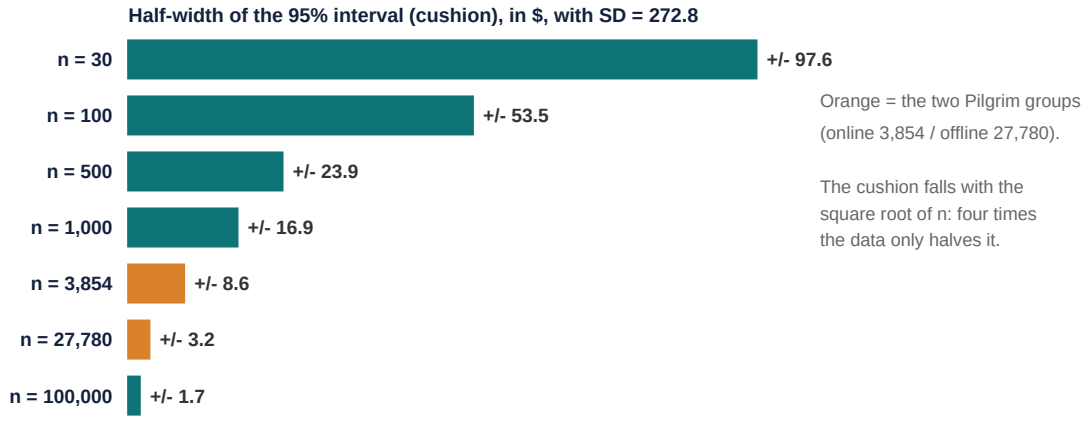


Figure 3.5. The cushion for a 95% interval at different sample sizes, using the Pilgrim SD. Going from 3,854 customers to 27,780 cuts it from $8.6 to $3.2, not to nothing.

TWO LINES FROM THE CASE SLIDES ON READING AN INTERVAL

Confidence is the probability that the interval contains the real population mean - the higher, the better. The **interval** is the range - the narrower, the better, because a wide range gives little help in deciding. And the test of whether an interval is narrow enough: **if your decision would differ at the lower limit and at the upper limit, the interval is too wide.** With $108.50 to $114.51 for the average customer, the bank would act the same way at either end - so that interval is narrow enough for its question.

Computing the two intervals in Excel

Quantity	Offline customers	Online customers	Excel
Sample mean (x-bar)	$110.786	$116.667	=AVERAGE(block)
Sample SD (sigma)	$271.301	$283.665	=STDEV.S(block)
Sample size (n)	27,780	3,854	=COUNT(block)
Cushion at 95% = 1.96 x SD / SQRT(n)	$3.190	$8.956	=1.96*SD/SQRT(n), or =CONFIDENCE.NORM(0.05, SD, n)
Lower limit = x-bar - cushion	$107.596	$107.711	=mean - cushion
Upper limit = x-bar + cushion	$113.977	$125.623	=mean + cushion
Width of the interval	$6.38	$17.91	upper - lower

(The 1.96 itself comes from the normal table: =NORM.S.INV(0.975) returns 1.95996. The 0.975 is because 5% of the curve is left out, split as 2.5% in each tail.)

How to say it: **with 95% confidence, the average profit that Pilgrim Bank earns from an offline customer across its whole population lies between $107.60 and $113.98; for an online customer, between $107.71 and $125.62.** That is a statement about the five million, made from the 31,634, with a stated accuracy. It is exactly what the point estimate could not give us.

The homework, worked (so you can check your own sheet)

The professor asked everyone to build the 90%, 95% and 99% intervals for all customers together, and for the online and offline groups separately. Here they are, computed from the raw data with the multipliers above.

Group	n	Mean	SD	90% (z = 1.645)	95% (z = 1.96)	99% (z = 2.576)
All customers	31,634	111.50	272.84	108.98 to 114.03	108.50 to 114.51	107.55 to 115.45
Offline	27,780	110.79	271.30	108.11 to 113.46	107.60 to 113.98	106.59 to 114.98
Online	3,854	116.67	283.66	109.15 to 124.18	107.71 to 125.62	104.90 to 128.44

Two things to notice before the next section. Every interval widens as the confidence rises - that is the cushion doing its job. And at every level, the online interval reaches into the offline one. Widening the confidence never separates them; it only makes the overlap bigger.

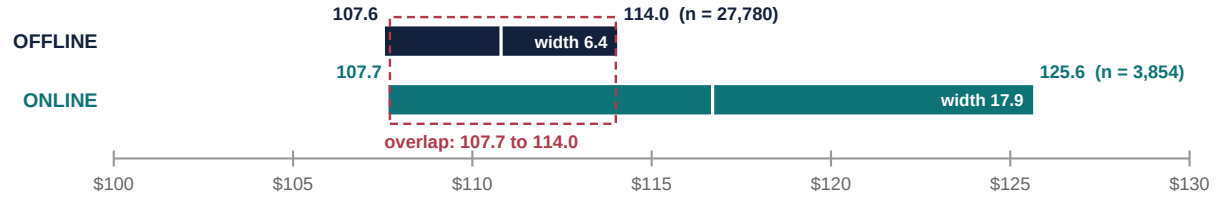
MANAGER'S TAKEAWAY

Replace "the number is X" with "the number is between A and B, at 95% confidence" in anything that will be used to decide. It forces three honest inputs into the open - how sure you need to be, how noisy the data is, and how much of it you have - and it stops a point estimate from pretending to a precision it does not have.

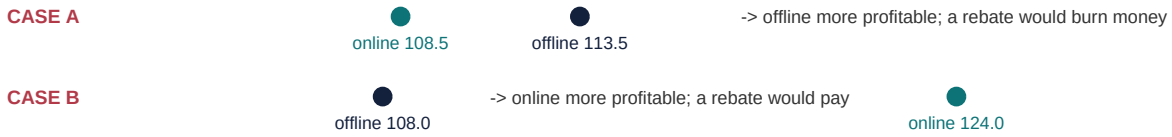
3.12 Overlap, Outliers and the Age Cliffhanger

With the two intervals on the board the professor asked the class to conclude. Somebody obliged: online is more profitable, because its interval goes up to $125. She said that was exactly what she had wanted to hear - because it is wrong, and it is the most common mistake people make with confidence intervals.

A. What the data gave us (95% confidence)



Both of these population pictures are consistent with the intervals above:



B. What would have allowed a conclusion (illustrative numbers)

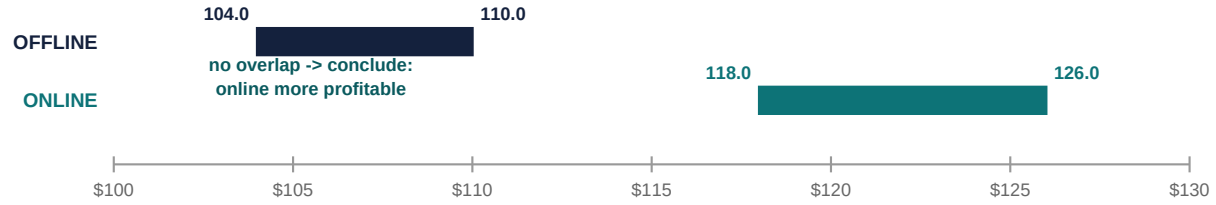


Figure 3.6. The two 95% intervals share the stretch from $107.7 to $114.0 (panel A), so both Case A and Case B are possible. Only non-overlapping intervals (panel B) would settle it.

Why "it goes up to 125" proves nothing

A confidence interval does not say the true value is near its upper end, or near its lower end, or in the middle. It says the true value could be **anywhere inside it**. So look at the two intervals together. The offline average could be $113.5 while the online average is $108.5 - both numbers sit comfortably inside their intervals, and offline is more profitable (Case A). Or the offline average could be $108 and the online $124, and online is more profitable (Case B). Both pictures are consistent with everything we know. The data is **inconclusive**.

> *"If the two confidence intervals overlap, you cannot conclude anything."*
>
> **The rule of the session - overlap means "not yet", not "roughly equal"**

Had the intervals come out like panel B - offline entirely below online, no shared ground - the conclusion would have been immediate: online customers are more profitable, offer the rebate. Madhav asked whether that would also need the sample sizes to match. No. Even 3,000 against 28,000 is fine, as long as the intervals do not overlap. In real life the group receiving a new intervention is almost always small at the start; if unequal sizes disqualified the comparison, we could never evaluate anything new. And if a sample were genuinely too small, you would know from the count before doing any analysis at all.

Why the online interval is three times as wide

The offline interval is about $6 wide; the online one about $18. The standard deviations are similar ($271 against $284), so the spread explains almost none of it. The difference is the sample size sitting under the square root: 27,780 against 3,854. A smaller denominator makes a bigger cushion. This is the practical face of "the larger the sample, the better": the online group is the one where the bank knows least, and its interval says so.

CONCEPT TRAP: "THEN LET US COLLECT MORE ONLINE DATA"

It is the reflex answer, and the professor discouraged it. Collecting more data is expensive and slow, and it is often not required. The data in front of you usually has much more to say if you look at it with a fresh pair of eyes. Session 4 will use the same 31,634 rows - not a single new one - and reach a conclusion. The cliffhanger below is the first clue how.

Two good questions from the class

Prachi: are there data sets where mean and SD should not be used? Yes. For a binary variable like 9Online, an average is nearly meaningless - it just tells you the proportion of ones, and the SD adds nothing. With many categories, a median is often the better summary. Her second case was income data where 99% of people earn under Rs. 20 lakh and 1% are billionaires. That is not a reason to abandon the SD; it is the question of **outliers** - values starkly different from the rest.

Should an outlier be excluded? The reflex is "yes, he is not like everybody else". The professor's answer: it depends on what the data is for. Suppose Mr Ambani buys a house in your neighbourhood. If you are estimating how much income tax the locality will contribute, keep him in - his 30% of a very large income is real tax. If you are deciding whether the locality should be charged as a high-net-worth area, take him out - he is one person pulling the average up, and the other residents are not rich. The same data point is an outlier for one question and the whole point for another.

Madhav KT: should we sample people who fall outside the confidence interval? Two answers. As a way of building a sample, no - never collect data *after* the analysis to fill in what you wish you had seen; that is a shortcut to bias. Collect first, analyse second. But as a way of *flagging* individual cases, yes: that is **anomaly detection**. If 90% of your monthly credit-card spend sits between Rs. 22,000 and Rs. 30,000 and this month shows a few lakh,

the card company raises a trigger because the value is beyond your usual interval. It is the same mathematics, pointed at a different question.

Question	Short answer
Mean and SD for a 0/1 variable?	No - use the proportion (12.2% of customers were online). For many categories, consider the median.
Exclude an outlier?	Decide by the question: keep it when the total matters (tax revenue), drop it when the typical case matters (is this locality rich?).
Add data that lies outside the interval?	Never, if you are building a sample: collect first, analyse later. Yes, if you are flagging anomalies: a value outside its usual interval is the alert.
Collect more data when intervals overlap?	Usually unnecessary and expensive. Look harder at the data you have (Session 4).

The cliffhanger: is age clouding the effect?

The professor closed with a hint for Session 4, insisting the focus there will be the methodology rather than the answer. Make a grid: online and offline customers across the top, young and older customers down the side. Who is more likely to be profitable - young customers or old? Old, obviously; think of your own bank balance in college. Who is more likely to be online? The young. So the online group is full of young, low-profit customers and the offline group is full of older, high-profit ones. Could that mix be hiding the true effect? Her numbers were hypothetical, and they make the point brutally.

The hint for Session 4: the two groups have different age mixes, and age drives profit.

	ONLINE	OFFLINE	
YOUNG	**-$30 per person** 80 people	**-$50 per person** 20 people	online wins by $20
OLDER	**+$350 per person** 20 people	**+$300 per person** 80 people	online wins by $50
	(80 x -30 + 20 x 350) / 100 **= $46 average**	(20 x -50 + 80 x 300) / 100 **= $230 average**	

Column by column, OFFLINE looks $184 better. Row by row, ONLINE is better for everyone.

Figure 3.7. The professor's hypothetical grid. Online is better for young customers (-$30 against -$50) and for older customers (+$350 against +$300), yet the online average is far lower - because 80% of the online group is young and 80% of the offline group is old.

Within every age group, moving online is better for the bank - by $20 for the young and $50 for the old. Notice also that the young customers stay loss-making online; the rebate does not make them profitable, it makes them *less* loss-making, which Section 3.10 said counts. But average the columns and offline wins by $184, because the two groups do not contain the same kind of customers. Compare the raw column averages and you conclude the opposite of the truth. (Statisticians call this reversal Simpson's paradox; the professor simply called it age clouding the effect.)

Is the hint just a story? The real data whispers the same thing. In the file, the share of customers who bank online falls steadily with age, and average profit rises steadily with age:

9Age category	1 (youngest)	2	3	4	5	6	7 (oldest)
Customers with an age on file	710	3,650	5,390	5,376	3,236	2,290	2,693
Share who are online	19.3%	21.5%	15.5%	13.4%	9.5%	5.0%	3.6%
Average profit	$3.5	$58.5	$115.1	$135.7	$145.8	$160.4	$192.3

The average online customer is in age category 3.3; the average offline customer is in 4.2. The two groups being compared are not the same population sliced by channel - they are two different populations. What to do about that, and how to say "online is more profitable" with a stated confidence, is Session 4. Come with the grid drawn.

BEFORE SESSION 4: THE HOMEWORK

1. If you did not draw the profitability curve yourself in class, build it from the raw data (Section 3.2) - not from the professor's sheet.
2. Compute the 90%, 95% and 99% confidence intervals for all customers, for online customers and for offline customers (the answers are in Section 3.11 - check, do not copy).
3. Think about the age grid: what would you compare, and what would you need to hold constant, to see the real effect of going online?

MANAGER'S TAKEAWAY

Before you crown a winner between two groups, ask three questions in order. Do the intervals overlap? If they do, you have no verdict, however tempting the point estimates look. Are the groups the same kind of thing? A team that looks faster may simply get easier tickets; a channel that looks worse may simply attract younger customers. And is the outlier part of the question or noise in it? The answer changes with the decision, not with the data.

3.13 Product, Service and the Manager's Desk

Everything in this session was taught on a bank's customer file. The ideas do not care what the business is. Here is each of them transplanted into a product business and a service business, followed by what they mean for the way you run a team.

Concept	Product-based company	Service-based company
Profitability skew	FMCG distributor: a fifth of the SKUs earn most of the margin; a long tail of slow movers eats shelf space and working capital.	Managed security services: a few large contracts fund the SOC; a tail of small contracts with heavy ticket volumes cost more to serve than they pay.
Letting go politely	Minimum order quantities and freight-paid thresholds price tiny orders out without refusing anyone.	Tiered support: premium customers reach an engineer at once; entry-tier customers are routed to the portal and chat first.
Pareto	Warranty returns cluster on a few components in a few failure modes - redesign those, not the whole product.	A handful of detection rules and log sources generate most of the alert noise - tune those before touching the other 400.
Population vs sample	1,000 units from a batch of 200,000 are tested; the pass/fail decision is about the batch.	400 tickets sampled from 60,000 closed this quarter; the quality score is a guess about all 60,000.

Concept	Product-based company	Service-based company
Representative sampling	Sample from every shift and every machine, not only the day shift you happen to visit.	Sample tickets across customers, severities and shifts - not only the escalations everyone remembers.
Convenience bias	Product reviews on your website come from the delighted and the furious; nobody in the middle writes.	A CSAT survey answered by 8% of customers measures the mood of the 8%, which is rarely the mood of the 100%.
Destructive testing	Crash tests, burn-in, tensile tests: you learn the limit by breaking the sample.	A penetration test or red-team exercise against a sample of assets - you cannot break every system to learn what breaks.
Mean with its SD	Fill weight 500 g with SD 2 g is a controlled line; 500 g with SD 20 g is a problem hiding inside a good average.	Mean time to detect 12 minutes with SD 3 is a process; with SD 40 it is an average covering a few disasters.
Confidence interval	"The defect rate is between 1.2% and 1.8% at 95%" - not "1.5%".	"Analyst utilisation is between 62% and 71% at 95%" - the range is the honest number.
Overlap rule	Two suppliers' defect-rate intervals overlap: no evidence to switch, whatever the point estimates say.	Two shifts' response-time intervals overlap: no verdict on which team is faster.
Anomaly detection	Control-chart limits on a machine: a reading beyond them stops the line.	Behavioural baselines: a user who normally logs in 9 to 6 from Bangalore appears at 3 a.m. from elsewhere - outside the interval, so an alert.
The lurking variable (age)	A variant that "sells better" only because it is stocked in bigger stores.	A team that "closes faster" only because it receives the easier ticket mix.

The manager's checklist from this session

#	Discipline	What it looks like on Monday morning
1	Draw the curve	Sort whatever you manage by contribution and plot the cumulative share. If it peaks above 100%, you have a tail your core is paying for.
2	Define the desirable customer	Up to where the curve stops rising. Then choose deliberately who gets scarce, expensive service and who is steered elsewhere.
3	One rule, two effects	Never a rule by name. A fee with a waiver, a tier with a threshold, a queue with a priority - identical for everyone, biting only where you intend.
4	Pareto before programmes	Find the vital few causes, SKUs or rules before launching an improvement drive across everything.
5	Name the population	For every number: which population is this about, and why is it a sample? Expensive, slow or impossible are fine reasons - but they make the number an inference.
6	Ask how the sample was chosen	Random or convenient? Who did not respond? A biased sample that looks like evidence is worse than no data.
7	Never an average without its SD	Report both, demand both. Then compute mean +/- 1 SD and ask whether that range would support the decision.
8	State a range and a confidence	Decide how sure you need to be from the cost of being wrong; let the SD and the sample size set the width.

#	Discipline	What it looks like on Monday morning
9	Check overlap before declaring a winner	Overlapping intervals mean "no verdict yet" - not "roughly equal" and not "the higher upper limit wins".
10	Look for the lurking variable	Before comparing two groups, ask whether they are the same kind of thing. Slice by the variable that drives the outcome (age, ticket mix, store size) and compare like with like.

The sharpest lines of the session

> *"Not making a decision is often not a choice. Options expire, and the decision gets taken for you."*
>
> **On messy data and postponed decisions**

> *"Customers are not God. They are the means of the company's success."*
>
> **On why some customers are worth letting go**

> *"Remember, you are buying to sell, not to wear yourself."*
>
> **The sign on the merchandiser's wall - on knowing the customer you are not**

> *"You can always go wrong, whatever the size of the sample. That is why it is statistics and not mathematics."*
>
> **On the number 30**

> *"Never, ever give anybody an average without the standard deviation."*
>
> **The rule of the session**

> *"If the two confidence intervals overlap, you cannot conclude anything. Your data has much more value to offer - look at it with a fresh pair of eyes."*
>
> **The verdict, and the hint for Session 4**

ONE-BREATH SUMMARY

A small share of customers brings most of the profit and a large share brings losses, so the profitability curve peaks above 100% and tells you who is worth keeping; you cannot fire the rest by name, but one rule with a threshold shapes the base for you. Every decision is about a population, made with a sample because the population is expensive, slow or impossible to reach; the sample is trustworthy if every element had an equal chance of being drawn, and large enough at 30. A sample average is a random variable, so report it with its standard deviation and, better, as a confidence interval: point estimate plus or minus a cushion that grows with the confidence and the spread and shrinks with the square root of the sample size. When two intervals overlap, the data has not spoken - and before collecting more, check whether a lurking variable like age is making two different populations look like one.

3.14 Practice MCQ Bank with Answers

One hundred and two questions drawn from this session, in roughly the order the ideas were taught. The options are meant to be close - several are numerically right for the wrong reason - so read the one-line reason under each answer even when you got it right. All numbers refer to the Pilgrim Bank 1999 data unless stated otherwise.

Q1. When a student said the Pilgrim data was too messy and incomplete to support a decision, the professor's response was that:

(a) More data should be collected before deciding

(b) Real business data is always messy, and not deciding is usually not one of the options

(c) Only the rows with complete age and income should be used

(d) The decision should wait for the year-2000 columns

Answer: (b) - Missing values are the normal condition of data; postponing a decision usually means it gets taken for you when options expire.

Q2. Making every survey question mandatory tends to:

(a) Eliminate missing values with no side effects

(b) Improve the quality of the answers

(c) Make people abandon the survey or type garbage

(d) Raise the response rate

Answer: (c) - The information you want is not the respondent's obligation; forcing it produces exits or junk.

Q3. In the professor's framing, retaining customers and building relationships are:

(a) The end goal of every for-profit company

(b) Means to an end that comes from the vision, mission and strategy

(c) Irrelevant for start-ups

(d) The same thing as loyalty

Answer: (b) - Retention is a means; the end is profit for a for-profit firm (or market share for a start-up in its valuation phase).

Q4. Loyalty bought with freebies is weak because:

(a) Freebies reduce market share

(b) Customers seek value, and someone who came for a freebie leaves for the next one

(c) Regulators prohibit them

(d) Only large companies can afford them

Answer: (b) - Solid loyalty comes from value delivered over a long time, not from a rebate.

Q5. On the profitability skew (Exhibit 3), the horizontal axis is:

(a) Customers in ID order

(b) Cumulative % of customers, sorted from most to least profitable

(c) Profit in ascending order

(d) Tenure with the bank

Answer: (b) - Sorting descending by profit is what makes "the top 10%" mean the ten percent most profitable customers.

Q6. The top 10% most profitable customers bring roughly what share of the year's profit?

(a) About 50%

(b) About 70%

(c) About 97%

(d) About 121%

Answer: (b) - The computed figure is 69.2%; 97% is the top 20% and 121% is the peak at about 53%.

Q7. The top 20% most profitable customers bring roughly:

(a) 70%

(b) 80%

(c) 97%

(d) Exactly 100%

Answer: (c) - 97.1% - which the class rounded to "about 100%". 80% is the classic Pareto figure, not this bank's.

Q8. The curve rises above 100% because:

(a) Excel double-counted some customers

(b) The cumulative profit of the profitable customers exceeds the net total, and the loss-makers then pull it back down

(c) Profit grew 21% during the year

(d) The vertical axis was mislabelled

Answer: (b) - At the peak every profitable customer has been added ($4.29 million) and no loss-maker yet.

Q9. The peak of the curve occurs at:

(a) 20% of customers

(b) The last customer with a positive profit, about 53% of customers

(c) The median customer

(d) 100% of customers

Answer: (b) - Customer 16,832 is the last one with profit above zero; everyone after subtracts.

Q10. The professor read the 121.6% peak as:

(a) The bank overstated its profit by 21.6%

(b) Letting go of roughly the bottom half of customers would have raised last year's profit by about 20%

(c) 21.6% of customers are online

(d) Online customers bring 21.6% more profit

Answer: (b) - The loss-making tail returned 21.6 points of the profit the good customers had earned.

Q11. How many customers in the file were loss-making in 1999?

(a) 3,854

(b) 14,580

(c) 16,832

(d) 27,780

Answer: (b) - 14,580 customers (46.1%) had negative profit; 16,832 were profitable; 3,854 and 27,780 are the online and offline counts.

Q12. In the recipe =A2/$E$1 for the percentage of customers, the dollar signs are there to:

(a) Format the result as currency

(b) Keep the total-customers cell fixed when the formula is dragged down

(c) Round the result

(d) Convert the result to a percentage

Answer: (b) - Without the absolute reference, dragging would move the denominator down a row each time.

Q13. The chart type the professor used for the skew, and why:

(a) Line chart, because the points are ordered

(b) Scatter with markers only, because lines through 31,634 points make a cluttered mess

(c) Pie chart, because the shares add to 100%

(d) Stacked bar, to show losses separately

Answer: (b) - Markers only; and tidy the axes to 20%, 40% ... rather than 0.2000 for the reader's sake.

Q14. Can a bank explicitly refuse to serve a customer because they were loss-making last year?

(a) Yes, anywhere

(b) Generally no - in India it cannot refuse service and in the West it would be treated as discrimination - so it uses implicit means

(c) Only in India

(d) Yes, after two loss-making years

Answer: (b) - Refusal risks the licence; preferential treatment achieves the same end quietly.

Q15. The professor's 27-minute JetPrivilege wait, against her husband's one-minute connection, illustrates:

(a) A call-centre system failure

(b) Deliberate steering of low-tier customers to cheap channels so agents stay free for high-tier customers

(c) Understaffing during a peak

(d) A loyalty programme error

Answer: (b) - The queue did not vanish when she hung up; Gold and Silver tiers are routed ahead of Blue.

Q16. A Rs. 10,000 annual card fee that is waived above a spending threshold is designed to:

(a) Earn Rs. 10,000 per customer

(b) Keep undesirable customers out while the profitable ones never pay it

(c) Comply with regulation

(d) Reward every customer equally

Answer: (b) - One rule for everyone, two effects: the fee filters entrants, the waiver retains the profitable.

Q17. In Subarna's bank, a customer who stops meeting the profitability score for a tier is:

(a) Asked to close the account

(b) Downgraded to the next tier, where the cost to serve and the facilities are both lower

(c) Charged a penalty

(d) Offered a rebate to stay

Answer: (b) - Graduated, not abrupt: the customer is weaned off expensive service rather than fired.

Q18. Gopal's car insurer refusing to renew after two claims above Rs. 1 lakh is an example of:

(a) Implicit channel steering

(b) Explicit refusal - possible in insurance, unthinkable for a bank account

(c) A Pareto effect

(d) Stratified sampling

Answer: (b) - The professor added that many organisations do this through faulty systems rather than deliberate policy.

Q19. The Pareto principle says:

(a) Effects always split exactly 80/20

(b) A small share of the causes or elements accounts for most of the effect

(c) 20% of the profit comes from 80% of customers

(d) It applies only to inventory

Answer: (b) - The exact split varies (this bank is nearer 20/97); the shape and the management response are the same.

Q20. In total quality management, the Pareto idea implies that:

(a) 80% of defects cannot be fixed

(b) A handful of root causes produce most of the defects, so fix those first

(c) Inspect 20% of the output

(d) 20% of workers cause the defects

Answer: (b) - Fix the vital few causes and most of the symptoms disappear with them.

Q21. In this data, the top 20% of customers bring what share of the profit earned by the profitable customers (before losses are netted)?

(a) About 20%

(b) About 69%

(c) About 80%

(d) About 97%

Answer: (c) - $3,423,486 of $4,288,303 is 79.8% - the 80/20 rule almost exactly; 97% is the share of the net total.

Q22. Koutons' "50% + 40% off" is a total discount of:

(a) 90%

(b) 70%

(c) 45%

(d) 80%

Answer: (b) - The 40% comes off the remaining half: 50% + 40% of 50% = 70%. Cantabil's 50% + 49% is about 75%.

Q23. A full-price retailer that is on sale most days of the year is showing:

(a) A successful discount business model

(b) Poorly managed demand - stock that did not sell has to be marked down

(c) Price jacking, as the professor's grandmother believed

(d) Very low costs

Answer: (b) - Koutons discounts by design; Shoppers Stop or Marks and Spencer discount because forecasting and lead times failed them.

Q24. "Remember, you are buying to sell, not to wear yourself" warns merchandisers against:

(a) Using employee discounts

(b) Substituting their own taste for the target customer's

(c) Overbuying before a season

(d) Copying boutique designs

Answer: (b) - Consolidation put buying in the hands of people who are not the customer - a NIFT graduate buying for a value store.

Q25. Offshoring apparel manufacturing to China, Vietnam and Bangladesh worsened discounting because:

(a) Quality fell

(b) Long lead times and logistics costs stop retailers responding to trends, so the wrong stock piles up

(c) Import tariffs rose

(d) Labour became scarce

Answer: (b) - Orders go in months ahead; your size sells out while the unwanted shades wait for the sale.

Q26. Zara replenishes its stores roughly every:

(a) Six months

(b) 15 days

(c) Six weeks

(d) Season

Answer: (b) - Frequent replenishment from centralised production, small displays and city fulfilment centres are all aligned around fast fashion.

Q27. Delayed differentiation means:

(a) Launching a product after competitors

(b) Keeping generic components or a base product and customising at the last step, as Dell and Asian Paints do

(c) Postponing markdowns to the end of the season

(d) Cutting the number of brands stocked

Answer: (b) - Stock components, not finished variety: the same component serves many configurations; pigments are mixed to order.

Q28. Where can variety NOT easily be reduced, according to the class?

(a) A configure-to-order laptop maker

(b) A departmental store - drop the item a customer came for and they take all their shopping elsewhere

(c) A paint company

(d) A value supermarket with strong bargaining power

Answer: (b) - Padmanabhan's example; Dell, Asian Paints and DMart were the cases where it works.

Q29. The lesson of Spencer's late and struggling Jiffy quick-delivery service is that:

(a) Quick commerce cannot work for supermarkets

(b) Starting a new channel is easy; aligning every process behind it is where businesses fail

(c) Kolkata is a difficult market

(d) Online channels should not be launched after COVID

Answer: (b) - Many post-COVID online channels died for the same reason.

Q30. A population is best defined as:

(a) Everyone in the country

(b) The set of all elements that are of interest to the study

(c) The 31,634 customers in the file

(d) The sample plus the people who did not respond

Answer: (b) - Both halves matter: all the elements, and only those this study cares about - so it changes from study to study.

Q31. For an inventory-management study, the "elements" of the population would be:

(a) Customers

(b) SKUs

(c) Suppliers

(d) Store managers

Answer: (b) - For a defects study they would be the possible root causes; for Pilgrim they are customers.

Q32. Which of these is NOT one of the reasons the professor gave for using a sample instead of the population?

(a) Collecting population data is expensive

(b) It is time-consuming

(c) It is sometimes impossible, as in destructive testing

(d) A sample is more accurate than population data

Answer: (d) - The population is always better if you can afford it in money and time; a sample is the fallback.

Q33. Which is an example of destructive testing?

(a) An exit poll

(b) Lighting an LED bulb and leaving it on until it fails, noting the hours

(c) Counting ATM withdrawals

(d) A customer satisfaction survey

Answer: (b) - The measured unit cannot be sold afterwards, so testing the whole population would leave nothing to sell.

Q34. Honda's decision not to give the 2001 Activa a spare wheel was:

(a) Pure cost cutting

(b) A talking point backed by a tyre-life claim that could only come from testing a sample destructively

(c) A regulatory requirement

(d) A mistake later corrected

Answer: (b) - Priced like a bike, shaped like a scooter, no stepney - and a claim of about 5,000 trouble-free kilometres.

Q35. A sample is representative of the population when:

(a) It has the same demographic mix as the population

(b) It was chosen by a method that gave every element an equal chance of inclusion

(c) It is at least 10% of the population

(d) It is balanced on gender

Answer: (b) - The definition is about the method, not the mix - you cannot balance on infinitely many characteristics.

Q36. A population of 5,000 men and 5,000 women; a sample of 100. Must it contain 50 men and 50 women?

(a) Always

(b) Only if the question is one where men and women are likely to answer differently

(c) Never

(d) Only if the sample is smaller than 30

Answer: (b) - Balance for a trait only when it drives the answer; otherwise forcing the balance introduces bias.

Q37. In stratified sampling, the strata in the sample should be in:

(a) Equal proportions

(b) The same ratio as in the population, e.g. 60/40 if the population is 60/40

(c) Inverse proportion to their size

(d) Any convenient proportion

Answer: (b) - And within each stratum the elements are drawn by simple random sampling.

Q38. Within each stratum of a stratified sample, elements are chosen by:

(a) Convenience

(b) Simple random sampling

(c) Expert judgement

(d) Taking the first names on the list

Answer: (b) - Stratification fixes the ratio; randomness inside each stratum keeps the sample fair.

Q39. The main danger of convenience sampling is:

(a) It is too expensive

(b) Bias - the people who responded may differ systematically from those who did not

(c) The sample is too large

(d) It requires special software

Answer: (b) - You did not choose the sample; the respondents did.

Q40. A school reports that its alumni earn Rs. 75 lakh a year on average, based on replies to a reunion survey. The most likely problem is:

(a) Inflation

(b) Non-response bias: those doing badly stayed silent while a few unicorn founders replied

(c) The sample was stratified

(d) The sample was smaller than 30

Answer: (b) - A segment of the population was far more likely to respond and dragged the average up.

Q41. Rare-event sampling is needed when:

(a) The population is small

(b) The phenomenon is so rare that a random draw would never find the cases

(c) Testing is destructive

(d) The data is a time series

Answer: (b) - A one-in-a-million disease will not appear in a simple random sample; most managers never need this method.

Q42. Statistically, a "large sample" is one with:

(a) At least 10% of the population

(b) At least 30 elements, regardless of the population size

(c) At least 1,000 elements

(d) At least 30% of the population

Answer: (b) - An absolute number, not a percentage - the class's 10% answer was the corporate convention, not the statistics.

Q43. The reason 30 is the threshold is:

(a) Convention in one industry

(b) The central limit theorem: sample averages start to follow the bell curve once samples have about 30 or more elements

(c) An Excel limitation

(d) The Pareto principle

Answer: (b) - That regularity is what the estimation and testing formulas rely on.

Q44. Beyond 30, a larger sample buys you:

(a) Permission to use the formulas

(b) Precision - a smaller cushion around every estimate

(c) Representativeness

(d) Nothing further

Answer: (b) - Permission arrives at 30; accuracy keeps improving with the square root of n.

Q45. The three things a sample is used for are:

(a) Estimation, hypothesis testing and identifying relationships (regression)

(b) Mean, median and mode

(c) Sorting, summing and charting

(d) Simple random, stratified and convenience sampling

Answer: (a) - This session did estimation; Session 4 does hypothesis testing; regression comes later.

Q46. An exit poll is an example of:

(a) Estimating a population proportion

(b) Hypothesis testing

(c) Regression

(d) Destructive testing

Answer: (a) - Nobody asks every voter; a sample estimates the share in favour.

Q47. The fine print "tested on 1,000 subjects under laboratory conditions" exists because:

(a) It is a legal formality with no statistical content

(b) A claim like "lasts 3,000 hours" or "makes you taller" must rest on sample evidence that can be defended in a hypothesis test

(c) It reports a regression

(d) It describes stratified sampling

Answer: (b) - Syska, Horlicks and hair-fall shampoos have all been challenged by consumer groups; the sample is the defence.

Q48. Working out how much pre-launch promotion to spend to hit a first-quarter sales target is:

(a) Estimation

(b) Hypothesis testing

(c) Regression - identifying the relationship and solving backwards

(d) Sampling

Answer: (c) - If sales are a known function of promotion, a Rs. 50 crore target can be turned into a promotion budget.

Q49. The Pilgrim 1999 file is:

(a) Time-series data

(b) Cross-sectional data - many customers at one point in time

(c) Panel data

(d) Experimental data

Answer: (b) - It is a snapshot; following the same customers into 2000 and 2001 would make it a panel.

Q50. India's GDP from 1991 to 2025 is:

(a) Cross-sectional data

(b) Time-series data

(c) Panel data

(d) Convenience data

Answer: (b) - One variable, many points in time; the 2025 GDP of 200 countries would be cross-sectional.

Q51. Panel data is:

(a) Many elements at one point in time

(b) One element over time

(c) Many elements, each followed over time - the most powerful of the three

(d) Data from a consumer panel survey

Answer: (c) - For each country a time series, for each year a cross-section, and the whole picture together.

Q52. Why did the case use cross-sectional data as a proxy instead of an experiment with a control group?

(a) Experiments on customers are illegal

(b) Tracking two groups for five to ten years is too slow - the technology would be obsolete by then

(c) Cross-sectional data is more accurate

(d) Control groups need more than 30 customers

Answer: (b) - So the question becomes "are online customers more profitable than offline customers?" rather than "does going online make a customer more profitable?".

Q53. A population parameter, such as the true average profit of all five million customers, is:

(a) A random variable

(b) A fixed but unknown constant

(c) Equal to the sample mean

(d) Undefined

Answer: (b) - If we had every row we would all compute the same number; we just do not have every row.

Q54. A sample mean is a random variable because:

(a) Excel rounds it

(b) Its value depends on who happens to land in the sample

(c) The population keeps changing

(d) It has a standard deviation

Answer: (b) - Sample the batch's work experience from the three most senior people and the mean is high; from the three youngest and it is low.

Q55. The deviations of the values from their mean always:

(a) Sum to zero

(b) Sum to the variance

(c) Sum to the standard deviation

(d) Sum to n - 1

Answer: (a) - The mean is the centre of gravity of the data, so negatives and positives cancel - hence the squaring.

Q56. The difference between STDEV.S and STDEV.P in Excel is that:

(a) S divides the sum of squares by n - 1; P divides by n

(b) S divides by n; P divides by n - 1

(c) They are identical

(d) S is for small samples only

Answer: (a) - S for sample (31,633), P for population (31,634); negligible here, material with a handful of values.

Q57. Why n - 1 for a sample variance?

(a) It is a safety margin

(b) Because the sample mean has already used up one degree of freedom, and dividing by n - 1 makes the estimate unbiased

(c) An Excel convention

(d) To correct for rounding

Answer: (b) - Given the mean and nine of ten values, the tenth is determined - only nine are free to vary.

Q58. For the sample 1, 2, 3, 4, 5, the sample variance is:

(a) 2.0

(b) 2.5

(c) 10

(d) 1.58

Answer: (b) - Sum of squares 10, divided by n - 1 = 4. Dividing by 5 gives the population variance 2.0; 1.58 is the SD.

Q59. The standard deviation is expressed in:

(a) Squared units of the data

(b) The same units as the data

(c) Percent

(d) No units

Answer: (b) - Variance is in squared units; the square root brings it back - $272.84 is in dollars.

Q60. Whether a standard deviation is "large" is judged:

(a) Against a fixed cut-off of 100

(b) Relative to the mean - $273 is trivial against a $50,000 mean and enormous against $111.50

(c) Against a cut-off of 30

(d) Against the median

Answer: (b) - Then check mean +/- 1 SD and ask whether that range could support the decision.

Q61. For a normal distribution, mean +/- 1 SD contains about:

(a) 50% of values

(b) 68% of values (the professor used 66%)

(c) 95% of values

(d) 99.7% of values

Answer: (b) - 95% is +/- 2 SD; 99.7% is +/- 3 SD.

Q62. Mean +/- 1 SD for the Pilgrim profit data is:

(a) $91.5 to $131.5

(b) -$161 to +$384

(c) $107.6 to $114.0

(d) $0 to $223

Answer: (b) - 111.50 +/- 272.84. The first option is what the range would be if the SD were 20; the third is the offline confidence interval.

Q63. If the mean were $111.50 but the SD were $20, mean +/- 1 SD would be:

(a) -$161 to +$384

(b) $91.5 to $131.5

(c) $0 to $223

(d) $111.5 to $131.5

Answer: (b) - Now you could say the average customer is profitable, and by roughly how much - the point of the comparison.

Q64. The median profit is $9 while the mean is $111.50. This indicates:

(a) A symmetric bell-shaped distribution

(b) A distribution skewed by a tail of very profitable customers

(c) A data error

(d) That the SD must be small

Answer: (b) - Half the customers earn the bank $9 or less; a few earn up to $2,071 and pull the mean up.

Q65. The offline customers' mean profit and standard deviation are:

(a) $110.79 and $271.30

(b) $116.67 and $283.66

(c) $111.50 and $272.84

(d) $107.60 and $113.98

Answer: (a) - The second pair is online; the third is everyone; the fourth is the offline confidence interval.

Q66. The number of online customers in the file is:

(a) 27,780

(b) 3,854

(c) 31,634

(d) 3,142

Answer: (b) - 27,780 are offline; 3,142 is the count for district 1100.

Q67. A student proposed dropping the loss-making customers before comparing online and offline. The professor said:

(a) Yes - losses distort the average

(b) No - those customers stay in the population the fee or rebate will apply to, and rules must be identical for everyone

(c) Only drop the offline loss-makers

(d) Only drop customers with missing age

Answer: (b) - A threshold rule (fee waived above Rs. 3 lakh of spend) is legitimate because it is the same rule for all.

Q68. A rebate can be worth paying even when:

(a) Online customers are less profitable in every group

(b) It only reduces a customer's loss, say from $50 offline to $30 online

(c) The customer never uses online banking

(d) The fee would raise more

Answer: (b) - Reducing losses counts; a rebate does not have to make everyone profitable.

Q69. Why could the class not conclude that online customers are more profitable from $116.67 versus $110.79?

(a) The sample is smaller than 30

(b) The gap is under 6% while the SDs are around $270-280, so a different sample could easily reverse it

(c) The online sample is too large

(d) Age was missing for some customers

Answer: (b) - Sample means are random variables; with that much spread the population could be the other way round.

Q70. A point estimate is:

(a) A range with an upper and lower limit

(b) A single number offered as the best guess of a population parameter

(c) The true population value

(d) The most frequent value

Answer: (b) - The sample mean of $111.50 is the point estimate of the population mean.

Q71. At "gunpoint", everyone in the class gave a range for the pen's price because:

(a) They were told to

(b) A range has a far higher chance of being right than a single number

(c) Point estimates are not allowed

(d) The standard deviation was unknown

Answer: (b) - High stakes make you demand accuracy, and you buy accuracy by widening the range.

Q72. Minus infinity to plus infinity is a 100% confidence interval. It is:

(a) The most useful interval

(b) Certainly right and completely useless - which is why 100% confidence is never used

(c) Statistically invalid

(d) The default in Excel

Answer: (b) - Confidence and usefulness trade off; 90%, 95% and 99% are the working compromises.

Q73. A confidence interval is built as:

(a) Point estimate plus or minus a cushion

(b) Mean plus or minus one standard deviation

(c) Minimum to maximum of the sample

(d) Median plus or minus the SD

Answer: (a) - The cushion is z times the SD divided by the square root of n; mean +/- 1 SD describes the data, not the estimate.

Q74. You should choose 99% rather than 90% confidence when:

(a) The sample is larger

(b) The cost of going wrong is very high

(c) The SD is small

(d) Always - more confidence is better

Answer: (b) - The level is a business choice driven by the cost of error; higher confidence costs you a wider interval.

Q75. The multiplier z for a 95% confidence interval is:

(a) 1.645

(b) 1.96

(c) 2.576

(d) 1.5

Answer: (b) - 1.645 is 90%, 2.576 is 99%. NORM.S.INV(0.975) returns 1.96.

Q76. The multiplier z for a 99% confidence interval is:

(a) 1.645

(b) 1.96

(c) 2.576

(d) 3.0

Answer: (c) - Wider than 95%: more confidence means a bigger cushion.

Q77. The cushion of a confidence interval grows with:

(a) The confidence level and the SD, and shrinks with the sample size

(b) The sample size

(c) The mean

(d) The tenure of the customers

Answer: (a) - z and sigma sit on top of the formula; the square root of n sits underneath.

Q78. Doubling the sample size reduces the cushion by:

(a) Half

(b) A factor of the square root of 2, about 29%

(c) Nothing

(d) Three quarters

Answer: (b) - The square root is why you need four times the data to halve the cushion.

Q79. The 95% confidence interval for offline customers' average profit is:

(a) $107.60 to $113.98

(b) $107.71 to $125.62

(c) $108.50 to $114.51

(d) -$160 to +$382

Answer: (a) - The second is online; the third is all customers; the fourth is mean +/- 1 SD.

Q80. The online interval ($17.9 wide) is about three times the offline interval ($6.4) mainly because:

(a) The online SD is much higher

(b) The online sample is far smaller (3,854 against 27,780), and n sits in the denominator

(c) The online mean is higher

(d) Online customers have more losses

Answer: (b) - The SDs are similar ($284 against $271); the sample size does almost all the work.

Q81. The correct way to state the offline result is:

(a) The average offline customer earns exactly $110.79

(b) With 95% confidence, the population average profit from an offline customer lies between $107.60 and $113.98

(c) Offline profit will be $113.98 next year

(d) 95% of offline customers earn between $107.60 and $113.98

Answer: (b) - The interval is about the population mean, not about individual customers.

Q82. The two 95% intervals overlap between $107.7 and $114.0. Therefore:

(a) Online is more profitable because its upper limit is higher

(b) Nothing can be concluded - the true means could sit either way round

(c) Offline is more profitable because its interval is tighter

(d) The two groups are equally profitable

Answer: (b) - Case A (offline 113.5, online 108.5) and Case B (offline 108, online 124) are both consistent with the data.

Q83. "Online must be more profitable - its interval goes up to $125.6." This reasoning is wrong because:

(a) The interval should have been 99%

(b) The true value can be anywhere in its interval, not preferentially near the upper end

(c) The online sample is too small to count

(d) $125.6 is an outlier

Answer: (b) - The professor called this the most common mistake made with confidence intervals.

Q84. Non-overlapping intervals would allow a conclusion even if:

(a) The sample sizes were 3,000 and 28,000

(b) Only if the sample sizes were equal

(c) Only if the SDs were equal

(d) Never - overlap is not the test

Answer: (a) - What matters is overlap, not matching sizes; new interventions always start with a small group.

Q85. At 90% confidence the intervals are offline $108.11 to $113.46 and online $109.15 to $124.18. They:

(a) No longer overlap, so online wins

(b) Still overlap, so the verdict is unchanged

(c) Overlap only because the SD changed

(d) Cannot be compared at different z values

Answer: (b) - Tightening the confidence narrows both intervals but never separates them; at 99% the overlap is larger still.

Q86. When intervals overlap, the professor's advice on collecting more online data was:

(a) Always collect more

(b) Usually unnecessary - the existing data has more to say if you look at it differently, which is Session 4

(c) Collect more offline data instead

(d) Reduce the confidence level until they separate

Answer: (b) - Session 4 uses the same 31,634 rows and reaches a conclusion.

Q87. For the binary variable 9Online, the professor said mean and SD:

(a) Are the right summary

(b) Do not make much sense - the average just reports the proportion of ones; for many categories use the median

(c) Should be replaced by the mode

(d) Should be computed with STDEV.P

Answer: (b) - 12.2% of customers are online is the sensible statement, not "mean 0.122, SD 0.327".

Q88. Mr Ambani buys a house in your neighbourhood. Should his income be excluded from the locality's average?

(a) Always - outliers are always excluded

(b) It depends on the question: keep it when estimating total tax revenue, drop it when asking whether the typical resident is wealthy

(c) Never - every data point counts

(d) Replace it with the median

Answer: (b) - The same value is an outlier for one question and the whole point for another.

Q89. Collecting additional data points after seeing the analysis, to fill in what you wish you had observed, is:

(a) Good practice

(b) A shortcut to a biased result - collect first, analyse later

(c) Stratified sampling

(d) Required for anomaly detection

Answer: (b) - Data gathered to match a conclusion cannot test that conclusion.

Q90. Flagging a credit-card month of several lakh when 90% of months fall between Rs. 22,000 and Rs. 30,000 is:

(a) Estimation

(b) Anomaly detection - a value outside its usual interval triggers an alert

(c) Stratification

(d) Regression

Answer: (b) - The same confidence-interval mathematics pointed at a different question - the one a SOC baseline asks.

Q91. The professor's closing hint was that, in the Pilgrim data, younger customers are:

(a) More likely to be online and less profitable

(b) More likely to be online and more profitable

(c) Less likely to be online

(d) No different from older customers

Answer: (a) - The real file agrees: about 20% of the youngest category is online against under 4% of the oldest, and profit rises with age.

Q92. In the hint grid, within each age group, online beats offline by:

(a) $20 for the young and $50 for the old

(b) $184 for everyone

(c) $0 - they are equal

(d) $30 for the old only

Answer: (a) - -$30 vs -$50 for the young; +$350 vs +$300 for the old.

Q93. Yet the overall column averages in that grid come out as:

(a) Online $46, offline $230

(b) Online $230, offline $46

(c) Both $138

(d) Online $116, offline $110

Answer: (a) - (80 x -30 + 20 x 350)/100 = 46; (20 x -50 + 80 x 300)/100 = 230.

Q94. The reversal in the grid - online wins in every row but loses overall - is caused by:

(a) Sampling error

(b) The two groups having different age mixes, and age driving profit

(c) A high standard deviation

(d) An Excel sorting error

Answer: (b) - Compare like with like (young with young, old with old) and the true effect reappears - the plan for Session 4.

Q95. The Excel function that returns the cushion of a 95% interval directly is:

(a) =CONFIDENCE.NORM(0.05, SD, n)

(b) =STDEV.S(range)

(c) =NORM.S.INV(0.95)

(d) =AVERAGE(range)/SQRT(n)

Answer: (a) - It equals 1.96 x SD / SQRT(n); NORM.S.INV(0.975), not 0.95, gives the 1.96.

Q96. A SOC manager reviews the tickets that analysts remember as difficult to judge quality. This is:

(a) Simple random sampling

(b) Convenience sampling with a bias towards memorable cases

(c) Stratified sampling

(d) Rare-event sampling

Answer: (b) - Draw tickets at random across customers, severities and shifts instead.

Q97. The decision to charge a fee or pay a rebate applies to:

(a) The 31,634 customers in the sample

(b) The population of about five million customers

(c) The 3,854 online customers

(d) The profitable customers only

Answer: (b) - That is why the whole exercise is about inferring the population from the sample.

Q98. Which statement is consistent with the professor's view of customers?

(a) Customers are God and must always be served

(b) Customers are the means of the company's success, so some are worth keeping, some changing and some letting go

(c) Every customer is equally valuable

(d) Loss-making customers should be refused service

Answer: (b) - The profitability curve makes the idea unavoidable; Section 3.3 shows how it is done without refusing anyone.

Q99. From the case slides, the standard deviation of profit among profitable customers ($306) against loss-making customers ($45) tells you:

(a) Loss-makers are more varied

(b) Profitable customers vary enormously while losses are small and similar - the number of loss-makers, not any one of them, hurts the bank

(c) The data has errors

(d) The mean is reliable

Answer: (b) - Average profitable customer $251.45; average loss-maker -$52.19; maximum $2,071; minimum -$221.

Q100. "The profit of the profitable customers is 21% more than the actual total" and "17% of the good customers' profit goes to loss-makers" are:

(a) Contradictory

(b) Two readings of the same numbers: $4,288,303 against $3,527,276, and $761,027 of $4,288,303

(c) Both wrong

(d) About different years

Answer: (b) - The peak of the curve is 121.6%; the losses are 17.7% of the gross.

Q101. The confidence level of an interval is:

(a) The width of the interval

(b) The probability that the interval contains the real population mean

(c) The sample size

(d) The standard deviation

Answer: (b) - Higher confidence is better; a narrower interval is better; the two pull against each other.

Q102. According to the case slides, an interval is too wide when:

(a) It exceeds one standard deviation

(b) Your decision would differ at its lower limit and at its upper limit

(c) It contains zero

(d) The sample is below 30

Answer: (b) - If the bank would act the same way at $108.50 and at $114.51, that interval is narrow enough for its question.

PART IV From Estimate to Decision: Hypothesis Testing, Regression, Missing Data and the Pilgrim Verdict (Online Session 4, 23 August 2026)

Part III ended in a deadlock: the confidence intervals for online and offline customers overlapped, so the Pilgrim Bank data could not say who was more profitable. Session 4 broke the deadlock without collecting a single new row. It introduced hypothesis testing through a woman with a baby on the Delhi Metro and a broken toy helicopter, explained Type I and Type II errors through two legal systems, defined the p-value as "would this be a miracle if the null were true", built simple and multiple linear regressions in Excel, showed that dropping customers with missing age data biases the sample, fixed it with imputation and a flag, and arrived at an answer: within comparable customers, online customers are about $19.65 more profitable, so offer a rebate of up to that amount. Every regression in this part was recomputed from the 31,634 rows, the assignment is worked, and Section 4.12 is an exam-style drill in reading regression output.

By the end of this part you will be able to	Sections in this part
- Write a null and an alternate hypothesis the right way round, and choose alpha from the cost of error - Name Type I and Type II errors in any setting, and say which one a test controls - Define the p-value in one sentence and use it: below alpha, reject - Read a regression output in the right order - Significance F, p-values, R Square, coefficients - Interpret a slope "holding the other variables constant", including a 0/1 dummy - Test whether dropping missing rows creates bias, and fix it with imputation plus a flag - Size an incentive to the honest effect, not the first significant number	4.1 Where We Are: Recap and Agenda 4.2 Hypothesis Testing: The Delhi Metro 4.3 Framing: What You Want to Prove 4.4 Type I, Type II and Alpha 4.5 The p-value: Would This Be a Miracle? 4.6 Regression: A Line Through the Cloud 4.7 Running It in Excel; Reading the Output 4.8 Is Age Clouding the Effect? MLR 4.9 The Missing-Data Trap: Bias 4.10 Imputation: Keep Every Row, Add a Flag 4.11 The Decision, and the Assignment 4.12 Reading Any Output: The Exam Drill 4.13 Product, Service and the Manager's Desk 4.14 Practice MCQ Bank with Answers

4.1 Where We Are: Recap and Today's Agenda

The professor opened with a ten-minute reconstruction of Session 3, because everything in this session stands on it. Put yourself back in early 2000, at Pilgrim Bank, without the hindsight we have today, trying to decide whether to pay a rebate or charge a fee for online banking - with data on fewer than 1% of your five million customers.

Part III idea	Where Part III left it	What this part does with it
Population vs sample	Decisions are for the population (5 million); we only have a sample (31,634), because population data costs time and money or is impossible to get (destructive testing).	Every regression coefficient is an estimate for the population, and every p-value is a statement about it.
Sample size and representativeness	n >= 30 is a large sample; larger is more accurate; the sample must be chosen without bias (simple random, stratified, rare-event; convenience sampling is the risky one).	Dropping the 8,289 customers without an age turns a representative sample into a biased one - Section 4.9.

Part III idea	Where Part III left it	What this part does with it
Estimation: point and interval	The sample mean is the best single guess; the confidence interval is point estimate +/- a cushion that depends on the confidence level, the spread and the sample size.	Regression output carries the same intervals - for the intercept and for every slope.
Comparing with intervals	Offline $107.60 to $113.98 and online $107.71 to $125.62 overlap, so the data was **inconclusive**: Case A (offline higher) and Case B (online higher) were both possible.	Same verdict from the first regression (p = 0.21), then a way past it.
Inconclusive - now what?	Do not rush to collect more data: expensive, slow, and sometimes the old data has to be scrapped and redone. Look harder at what you have.	The whole session uses the same 31,634 rows and ends with a decision.
Kinds of data	Cross-sectional (Pilgrim), longitudinal or time series, panel.	Cross-sectional data cannot prove that going online *makes* a customer more profitable; it can show that online customers *are* more profitable, like for like.

The three things a sample is for - and where we are on the list

Session 3 said a sample is used for exactly three purposes: estimation, hypothesis testing and identifying relationships (regression). Estimation is done. Today's agenda is the other two - but with a twist the professor announced up front: **we will do the hypothesis testing implicitly, through regression**. Every regression output contains a hypothesis test, so learning to read the output is learning to test hypotheses. First the theory, in stories; then Excel.

One more session follows this one - optimisation, with less mathematics and more interpretation, the professor's promise to those who found the statistics heavy - and then the second campus visit, on simulation. Those are not in this document yet.

> **A NOTE ON THE NUMBERS IN THIS PART**
>
> Every regression in this part was re-run on the raw "Data" sheet of the Pilgrim Bank workbook and matches the professor's screen and the case slides to the second decimal (110.786 and 5.881 for the first regression, 17.08 / 27.19 / 25.86 for the second, 72.96 / 52.22 for the bias check, 70.93 / 19.65 / 25.60 / -51.85 for the final model). Where the class rounded - "about $6", "about 2%" - the exact figures are given alongside.

4.2 Hypothesis Testing: The Delhi Metro Once-Over

"Anybody who has never done hypothesis testing in their career, or in life?" Quite a few hands went up in the chat. The professor's claim: all of us do it, day in and day out - we just never wrote H0 and H1. Then she told a story to prove it.

The Manhattan once-over, Delhi edition

In Manhattan there is a habit with a name: the *Manhattan once-over*. A stranger walking towards you looks you up and down in a fraction of a second, judges your clothes, your accessories, what sort of person you are, and moves on. The ladies' compartment of the Delhi Metro has the same custom. Whoever boards is looked over by everyone, once, and then everyone goes back to their phones.

One day a woman boarded who was not just poorly dressed but shabbily dressed - what she wore looked like rags. She was carrying a baby, and the baby was dressed for a festival: fancy clothes and what looked like gold bangles. The once-over did not end. Everybody kept staring, for a long time. Why? The class got it immediately: **it did not add up**. A nanny? Nannies are not sent alone on the Metro with a baby in gold. Somebody typed "kidnapped". (A mother in the class offered a gentler reading - a well-dressed child usually means a mother who has stopped dressing for herself - and the professor accepted it as a perspective, though torn clothes are still a stretch.)

Whatever each passenger concluded, they had just run a hypothesis test. Two statements were on the table, **the baby is hers** and **the baby is not hers**. Nobody knew the truth. Everybody compared what they saw with a sample - a lifetime of observing what usually goes with what - and found a mismatch. If one of us walked in with a baby like that, nobody would give a second look, because our appearance and the baby's would match.

The formal version

Element	Definition	On the Metro
Null hypothesis (H0)	One of two statements **about the population**. The test is built around this one.	The baby is hers.
Alternate hypothesis (H1, also Ha)	The other statement about the population.	The baby is not hers.
Mutually exclusive	The two cannot both be true. If one is true the other is false.	Either she is the legitimate carer or she is not.
Collectively exhaustive	Together they cover every possibility. One of them **must** be true.	There is no third case.
The sample	The statements are about the population; their truth is checked with sample data.	What we have seen over a lifetime of babies in gold bangles and the adults who carry them.
The population / the truth	What we want to know but never observe directly.	Whether the baby is really in the right custody.

Statistics is used this way far beyond business - in law and in biostatistics especially - whenever a sample is used to infer something about a particular case or about a population in general.

Only two verdicts, and only one of them is a conclusion

After the test you do one of two things: **reject H0** or **do not reject H0**. That is all. Both verdicts are worded around the null, and they are not symmetric.

Verdict	What it means	What you are entitled to say
Reject H0	The sample is too unlike what H0 would produce. You side with H1.	You **accept H1**. This is the only conclusive outcome, and H1 is the only hypothesis you can ever accept.
Do not reject H0	The sample is not strong enough evidence against H0.	You are **inconclusive**. You do *not* accept H0; you say the data does not let you decide. "Not rejecting is not the same as accepting."

CONCEPT TRAP: "WE DID NOT REJECT THE NULL, SO THE NULL IS TRUE"

No. Failing to reject means the sample could not tell the two stories apart - perhaps because the effect is small, perhaps because the sample is noisy. The Session 3 intervals overlapped; that did not prove online and offline customers are equally profitable, it proved we could not yet tell. A student asked the professor to repeat this point twice, and she did: hypothesis testing is all about the null, the only decisive verdict is "reject", and the only thing you ever accept is the alternate.

4.3 Framing the Hypothesis: What You Want to Prove Goes in the Alternate

If "reject H0" is the only conclusive verdict, then which statement you call H0 decides whether the test can ever help you. This is the single most important practical rule of the session, and the professor proved it with a shampoo.

The hair-fall shampoo, framed two ways

An FMCG company has a new shampoo formula it believes reduces hair fall. It cannot test every bottle - it wants to sell most of them - so it tests a sample on a focus group. Two ways to write the hypotheses:

	Case 1	Case 2
H0	This formula reduces hair fall	This formula does **not** reduce hair fall (it is just as good as the old one)
H1	This formula does not reduce hair fall	This formula **reduces** hair fall
If H0 is rejected	You have established that it does *not* work. Nothing to advertise.	You have established that it works. Advertise - and defend the claim in court with the data.
If H0 is not rejected	Inconclusive. You cannot advertise, because not rejecting is not accepting.	Inconclusive. Perhaps the effect is too weak to show; no claim yet.
Useful to the FMCG company?	**No.** Whatever the outcome, it cannot do the one thing it wanted.	**Yes.** One of the two outcomes lets it act.
Useful to whom, then?	A competitor or a consumer advocacy group that wants to prove the claim false: rejecting H0 is exactly their win.	The company making the claim.

> *"Whatever you want to establish, put it as your alternate hypothesis. The negation of it becomes the null."*
>
> **The framing rule - same data, same sample, but the stakeholder decides the frame**

The same problem and even the same sample are framed differently by different stakeholders, because each wants to establish a different thing. That is not cheating; it is the point. The test can only ever hand you the alternate, so the alternate must be the statement you need.

Framing Pilgrim Bank - and why the frame is a business decision

The class's first attempts were "H0: charge a fee" and "H0: do not charge a fee". Wrong kind of statement. Hypotheses are about **what the data shows** - a fact about the population - not about the action you will take. The fact here is who is more profitable. Two legitimate frames:

	Frame A	Frame B
H0	Average profit (offline) >= average profit (online)	Average profit (offline) <= average profit (online)
H1	Average profit (offline) < average profit (online)	Average profit (offline) > average profit (online)
Reject H0 means	Online customers are more profitable -> **offer the rebate**	Offline customers are more profitable -> **charge the fee**
Do not reject H0 means	Inconclusive -> you fall back to **charging the fee**	Inconclusive -> you fall back to **offering the rebate**
The risk in the fallback	A blanket fee even when the two groups may be just as good: possible customer pushback, but no increase in the bank's cost.	A rebate for everyone even when online may not be more profitable: many customers move online, cost rises, no extra profit to show for it.
Who chooses it	A bank more worried about spending money it may not recover.	A bank more worried about annoying customers, willing to take a calculated hit.

Notice two things. The **equality sign always sits in the null** - because the test starts by assuming the null is true and asking how likely our sample would then be, and "equal" is the natural place to start. And there is **no statistically right choice between A and B**. Abhishek preferred A because rejecting gives you the rebate; the professor pointed out that A's fallback charges everyone a fee on thin evidence. The frame encodes which mistake you would rather make. That is a management call, and it has to be made before the data is looked at.

MANAGER'S TAKEAWAY

Before anyone runs a test, write down what you want to be able to say afterwards and make that the alternate. Then write down what you will do if the test comes back inconclusive - because it often will - and check that you can live with that fallback. Frame A and Frame B use identical data and lead to opposite default actions; the difference is not statistics, it is which risk you chose.

4.4 Two Ways to Be Wrong: Type I, Type II and the Significance Level

Back on the Metro. You have decided the baby is probably not hers. What do you do? Watch the baby's behaviour - is it sleeping peacefully? Report to the police? Or nothing? Whatever you choose, the professor said, you will not be at peace that day.

Report her, and imagine the police being rough with her while she says: it was my childhood dream to give my child what I never had; I have been saving since I was thirteen for these bangles. You may not believe her. But India has had people who carved a road through a mountain with a hammer, alone, over decades; is she paying the price of an unusual dream? Do nothing, and the thought comes back on the way to the office: what about that baby, what about its parents? **Whatever you do, you will never know the truth** - short of a DNA test, which was not available in the compartment. And that is exactly the situation of every hypothesis test: you decide from a sample, so both verdicts carry a probability of error, and neither probability is zero.

	YOUR VERDICT FROM THE SAMPLE	
THE TRUTH (never known)	**Reject H0**	**Do not reject H0**
H0 is actually TRUE (the baby is hers)	**TYPE I ERROR** You reported an innocent mother. **Controlled at alpha (5%)**	**CORRECT** You left her alone.
H0 is actually FALSE (the baby is not hers)	**CORRECT** You reported a kidnapper.	**TYPE II ERROR** A kidnapper walked away. **Reduced by a bigger sample**

Push the Type I risk down and the Type II risk goes up: the test is deliberately skewed to protect the null.

Figure 4.1. The four outcomes of a test. Two are correct; the other two are named errors, and the test is designed to keep the first within a chosen tolerance.

Error	Definition	On the Metro	How it is handled
Type I	Rejecting H0 when H0 is actually true.	The baby was hers and you reported her: an innocent woman goes through the ordeal.	The methodology keeps its probability within a tolerance you choose: the **level of significance, alpha**.
Type II	Not rejecting H0 when H0 is actually false.	The baby was not hers and you did nothing: you were party to jeopardising a child's future.	The test does not control it directly. The way to shrink it is a **larger sample**.

Alpha: how much Type I error you can live with

Level of significance (alpha)	When it is used	The professor's examples
10%	The cost of going wrong is minor.	A small promotion: if it flops, no great harm.
5%	The most commonly used level - the default.	Almost every market-research report you will read.
1%	The cost of going wrong is high.	An automation vendor claims its robots will raise your productivity. Adopting means redesigning processes, letting people go, hiring others. You want the pilot to show a *significant* improvement before you commit.
0.1% or 0.01%	The cost of going wrong is catastrophic.	Decisions in a nuclear power plant.

Alpha can never be zero. Zero error would require the whole population, and then you would not be testing, you would be counting. The test is deliberately **skewed**: it guards the Type I error tightly and lets the Type II error be whatever the sample size makes it. Which is why the framing of Section 4.3 matters so much - the error you guard is the error of wrongly rejecting the statement you put in the null.

Two legal systems, two nulls

A student asked how this plays out in law, and the professor used the contrast between India and the systems of several Arab countries (the film *Dor*, 2006, turns on the difference). India works on "innocent until proven guilty": the null is that the accused is innocent, and only if the evidence rejects it is the person declared guilty - with exceptions in a few sections, such as domestic violence and dowry, where the burden shifts. In the other system the null is that the accused is guilty, and the accused must prove innocence.

	Indian system	The other system
H0	The accused is innocent	The accused is guilty
H1	The accused is guilty	The accused is innocent
Type I error (reject a true H0)	Punishing an innocent person	Setting a guilty person free
Type II error (fail to reject a false H0)	Setting a guilty person free	Punishing an innocent person
What the system minimises	Punishing the innocent - so, unavoidably, more of the guilty walk.	Letting the guilty walk, because it "sets an example" - so, unavoidably, more innocents are punished.
Side effect	Layers of appeal - district court, high court, Supreme Court, and for capital punishment the President - to push the error near zero. Justice is slow.	The onus is on the accused; if they cannot prove innocence the case ends. Justice is fast.

Both systems face the same two mistakes; they only differ in which one they refuse to make. Push the probability of one error down and the other rises. Who decides? Not the statistics - the value system, or in a company the strategy. As the professor put it, ask which mistake you are more worried about, and set up the hypothesis accordingly. Where the repercussions are severe, the tolerance goes to 1% or lower and you add layers of checking; where they are trivial, 10% is fine.

The toy helicopter: verdicts are not written in stone

You are five and you have begged for a year for a remote-control helicopter that really flies. On your sixth birthday you get it. On Monday you leave it on your study table and go to school; your sibling stays home with a stomach ache. You come back, run to your room, and see it lying in two pieces. What do you do? The chat was unanimous: accuse, cry, chase the sibling. The sibling says: I did not break it - the wind blew the curtain, it fell, I only put it back. Does anyone believe them? Nobody. Given a lifetime of that sibling borrowing your toys, your stationery and your geometry box, the probability that they broke it is very high, and you do not reject the null that they did.

Thirty years later, in Goa over a beer, the sibling brings it up: I really had not broken it. Now most people believe them. Why? The sibling has no incentive to lie today, and you have three decades more data about them. Same incident, same hypothesis - but the data set grew, and the verdict flipped. **Hypothesis testing is not written in stone.** More data can validate a verdict or overturn it. If you remember only two things from this section, remember the Delhi Metro and the sibling's birthday; the rest follows.

MANAGER'S TAKEAWAY

Set alpha from the cost of being wrong, not from habit: 10% for a cheap experiment, 5% by default, 1% or lower when the commitment is expensive to reverse. Then remember that the test only guards one of the two errors, and that a verdict reached on today's sample is provisional - a pilot that passed at 5% can still fail at scale, and a rejected idea can come back with better data.

4.5 The p-value: "Would This Be a Miracle if the Null Were True?"

There are two ways to test a hypothesis. One you already know: build confidence intervals and see whether they overlap. The other is what every market-research report and every statistics package gives you: the p-value.

> *"The p-value is the probability of getting a sample like the one we have got, if the null hypothesis is true."*
>
> **The definition to memorise**

It sounds convoluted, so back to the Metro. If the baby really is hers, what is the probability that she would be dressed the way she is? Very small. That small probability is why the compartment suspected her. If one of us walks in with a baby in gold bangles, the probability of our appearance given "the baby is hers" is high, so nobody looks twice. **In probability we always suspect a miracle.** When the null hypothesis would make our sample a near-impossible event, we would rather believe the null is wrong than that a miracle happened. The sibling story works the same way: given past behaviour, the probability that they broke the helicopter was high, so their innocence looked like a miracle and you did not believe it.

Step	What happens	Who does it
1	State H0 and H1 (Section 4.3) and choose alpha (Section 4.4).	You
2	Enter the data and the hypothesis into software - Python, R, SPSS, or plain Excel.	You
3	The software computes the p-value: the probability of a sample like yours if H0 were true.	The software
4	Compare: if **p-value < alpha**, reject H0 (and accept H1). If p-value >= alpha, do not reject - such a sample is something that could commonly occur under H0, so you do not doubt it.	You

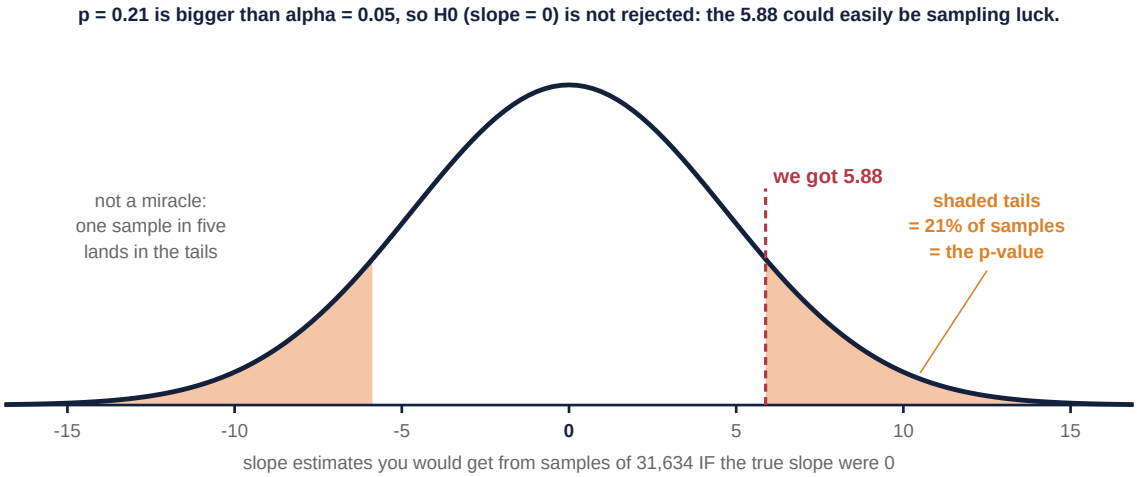


Figure 4.2. The p-value of the first Pilgrim regression, drawn. If the true slope were zero, samples of this size would produce a slope at least as far from zero as 5.88 about 21% of the time. That is not rare enough to reject.

TWO ROADS TO THE SAME VERDICT

Confidence intervals and p-values are two ways of asking one question. Overlapping intervals said "inconclusive" in Session 3; the first regression of this session says the same thing with p = 0.21. The professor made the class walk both roads on purpose: if two valid techniques are applied to the same data, they must agree. A fancier name does not buy a different answer.

Where you will meet hypothesis tests at work

Setting	The claim being tested	The frame
Marketing	"Our new formula reduces hair fall." "Our bulb lasts 3,000 hours."	H1 is the claim; the fine print "tested on 1,000 subjects" is the sample that lets it be defended.
Challenging a claim	A competitor or consumer group says the claim is false.	The mirror image: their H1 is "it does not work".
Warranty design	How long should the warranty be - six months? a year? Urban Company offers 30 days on some services.	Test claims about product life on sample data so that you neither bleed on replacements nor forgo the customer confidence a warranty buys.
Operations	"The new roster cuts response time." "The automation vendor's robots raise productivity."	H1 is the improvement; alpha is set by how expensive the change is to reverse.

MANAGER'S TAKEAWAY

Whenever a report shows a p-value, a hypothesis test is going on - find out what its null was before you read the number. Then the rule is one line: p below alpha, reject the null and act; p above alpha, the result could easily be luck, so do not act on it however attractive the point estimate looks.

4.6 Regression: A Line Through the Cloud

"Anybody who has never done regression?" Many hands. The professor's reassurance: since COVID, everyone has been on the receiving end of it. The newspaper graphs projecting how many cases there would be next week, how many hospital beds would be needed, what the mortality would be - regression was one of the tools behind them.

The idea, with a film release

You are releasing a film. It may be a hit or a miss, but you know one thing: the promotion before release drives the first weekend's box-office collection. (After the first weekend, reviews and word of mouth take over; the first weekend is hype.) The same holds for a new product: first-quarter sales track pre-launch promotion. So you want an equation - first-weekend collection as a function of the pre-release promotion budget - because then you can plug in the collection you are targeting and read off the budget to spend. Regression is the tool that finds that equation from past data.

Role	Other names	Symbol	How to spot it	Pilgrim
Independent variable	Input variable, predictor	x	The one whose value is known or decided first.	9Online: the customer decides to bank online or not.

Role	Other names	Symbol	How to spot it	Pilgrim
Dependent variable	Output variable, response	y	The one that happens afterwards and that you want to predict.	9Profit: the profit the customer then gives the bank.

Houses, areas and prices

Suppose you have data on 1,000 houses: the built-up area of each and its selling price. A 1,000 sq ft flat sold for Rs. 80 lakh; a 1,500 sq ft one for Rs. 89 lakh; an 800 sq ft one for Rs. 1 crore; and so on. You want selling price as a function of area - area is x, because it exists first; price is y. The first thing you do in any regression is **plot the scatter chart**. The points will not all sit on one curve: two houses of the same area sell for different prices because of locality, city, floor, direction, interiors, and a dozen other things. But if the cloud looks roughly like a straight band, you fit a **linear** regression. If it looks like the COVID curve - flat, then flat, then suddenly steep - you would fit a non-linear one.

Two 1,500 sq ft houses sell for different prices: area explains part of the price, other factors the rest (that is R-squared).

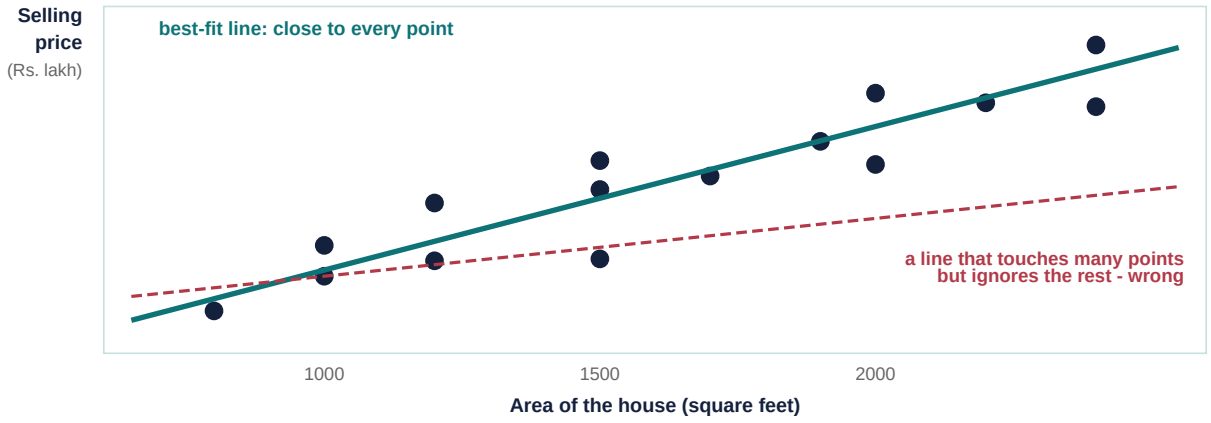


Figure 4.3. A scatter of house prices against area, the best-fit line, and the mistake schoolchildren make on graph paper: drawing the line through the most points instead of close to all of them.

Describing a line: slope and intercept

For one independent variable, the linear equation is **y = intercept + slope x x** - the y = mx + c of school mathematics. The professor's way of explaining the two numbers: imagine you must describe a line to someone who cannot see it. First you say how steep it is and which way it tilts - that is the **slope** (m). But many parallel lines share a slope, so you also say where it cuts the y-axis - that is the **intercept** (c). Regression finds the best slope and the best intercept for your cloud of points.

"Best" has a precise meaning. Think of the physics practical where you plotted V against I to verify Ohm's law. The points were never perfectly collinear, and the tempting mistake was to draw the line through the maximum number of points. The correct line - the **best-fit line** - passes close to *all* the points, and may pass through none of them. Excel or any other package finds it; what it hands you is the slope and the intercept, and your job is to interpret them.

Term	Meaning
Simple linear regression (SLR)	One independent variable: y = intercept + slope x x. "Simple" means one x; "linear" means the scatter looks like a straight band.
Multiple linear regression (MLR)	More than one independent variable: y = intercept + slope1 x x1 + slope2 x x2 + ... For example, selling price as a function of area *and* the age of the building.

Term	Meaning
The error term	The case slide writes y = intercept + slope x x + error. The error is everything the line does not capture - all the other reasons two 1,500 sq ft houses sell for different prices.

4.7 Running It in Excel, and Reading the Output

The first regression of the session - probably the first regression of many people's lives - was the simplest possible one: 9Profit as the dependent variable, 9Online as the independent variable. Here is the recipe exactly as the class followed it, including the error one student hit.

Step	What to do	Watch out for
1	Insert a new sheet (the professor named it SLR). Copy 9Profit and 9Online from the Data sheet, side by side.	No blank rows in between - Session 3's blank row must go. You need all 31,634 contiguous rows.
2	Data tab > Data Analysis (right-hand end of the ribbon).	Not there? File > Options (may be under More) > Add-ins > Manage: Excel Add-ins > Go > tick Analysis ToolPak > OK. Not the VBA one.
3	Data Analysis > scroll to Regression > OK.	The dialog looks slightly different on Windows and Mac; the fields are the same.
4	Input Y Range: 9Profit, including its label. Input X Range: 9Online, including its label. Tick **Labels**.	Y is what happens later (profit); X is what is decided first (online).
5	Select ranges by clicking the label cell, then Ctrl + Shift + Down Arrow.	Do not select the whole column: the blank cells below row 31,635 make Excel throw an error - the mistake a student shared on screen. Ctrl + Up Arrow gets you back to the top.
6	OK. The output appears on a new sheet just before the one you were on.	Tick Labels so the variable is named 9Online in the output; otherwise it is "X Variable 1", which is confusing with several variables.
7	Optional: set the Confidence Level box to 90 to get 90% limits in addition to the 95% ones.	Left blank, Excel prints the 95% limits twice.

The three tables, and the order to read them

Every regression output, whatever the software, has three tables: a summary at the top, an ANOVA table in the middle (analysis of variance - a statistical test the class returned to for multiple regression), and the coefficients table at the bottom. The class started at the bottom.

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.00705
R Square	0.00005
Adjusted R Square	0.00002
Standard Error	272.84
Observations	31634

3 **R Square: share of the** variation explained (0.005%)

ANOVA	df	SS	MS	F	Significance F
Regression	1	117,000	117,000	1.572	0.210
Residual	31632	2,354,800,000	74,440		
Total	31633	2,354,900,000			

1 **Significance F: read FIRST** in a multiple regression

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	110.786	1.637	67.68	0.000	107.58	113.99
9Online	5.881	4.690	1.254	0.210	-3.31	15.07

2 **p-value of each slope:** above 0.05 -> do not use

4 **Coefficients -> equation:** 110.79 + 5.88 x 9Online

Excel's Data Analysis > Regression output for 9Profit on 9Online (SS values rounded). The orange cells are the two places the verdict lives.

Figure 4.4. Excel's output for the first regression, with the reading order the professor taught. For a simple regression the slope's p-value is the verdict; for a multiple regression read Significance F first.

From coefficients to equation to prediction

The bottom table has one row for the intercept and one for 9Online. The intercept is **110.786**; the slope for 9Online is **5.881**. Not the intercept - the professor was precise - but the *estimate* of the intercept, because the true value belongs to the five-million population and this is the best guess from the sample. So the equation the analysis suggests is:

THE FIRST REGRESSION EQUATION

9Profit = 110.786 + 5.881 x 9Online

For an offline customer, 9Online = 0: predicted profit = 110.786 + 5.881 x 0 = **$110.79**.

For an online customer, 9Online = 1: predicted profit = 110.786 + 5.881 x 1 = **$116.67**.

Online looks more profitable, by about $6. And yes - these are exactly the two averages from Session 3. Same data, same result: a regression on a 0/1 variable is a difference of two means.



Figure 4.5. What the "line" means when the predictor can only be 0 or 1: the intercept is the offline average, the slope is how much higher the online average is.

The check before you use any regression: the p-value

In Session 3 the same two averages made us suspicious once we saw the standard deviations, and the confidence intervals overlapped. Does the regression also say "inconclusive"? Yes. For a simple regression, the check is the **p-value of the slope**. Here it is 0.210 - 21%. Our level of significance is 5% (or 10%, or 1%; none of them is 21%). Since the p-value is larger than alpha, **do not use this regression**. Stop right there: the same verdict as the overlapping intervals, reached by a different road.

What hypothesis was tested to get that 0.21? Every regression tests the same thing for every slope: **H0: the slope is 0** against **H1: the slope is not 0**. A slope of zero means the line is flat - profit equals the intercept whatever 9Online is, area does not move the selling price, the scatter is a shapeless ball with no tilt. In words: the independent variable and the dependent variable are not related. The p-value is then the probability that, if the true slope were zero, a sample of 31,634 customers would still throw up a slope as far from zero as 5.88. Twenty-one percent. Not a miracle - one sample in five. So we cannot rule out that the true slope is zero, that being online has no effect on profit, and we can recommend neither a fee nor a rebate.

The class asked the obvious question: if you knew it was inconclusive, why make us do it? Two reasons. This simple model was a stepping stone to the multiple regression that would settle it. And the class needed to see that two valid techniques on the same data agree - a fancier method does not manufacture a different answer.

The other columns, decoded

Column	What it is	In the first regression
Coefficients	The point estimates of the intercept and the slope(s), for the population.	110.786 and 5.881
Standard Error	The standard deviation of that estimate. Estimates are random variables - a different sample of 31,634 would give a slightly different slope - and this is their spread (the batch work-experience example again).	1.637 for the intercept, 4.690 for the slope
t Stat	The estimate divided by its standard error. Another way to run the same test: compare it with a threshold from the t-distribution table (a relative of the normal distribution, like the F-distribution you will meet next).	67.68 and 1.254

Column	What it is	In the first regression
P-value	The probability computed from the t Stat. The reason p-values are more popular than t Stats: you compare them directly with alpha instead of looking up a table.	0.000 and 0.210
Lower / Upper 95%	The confidence interval for each coefficient - point estimate +/- cushion, exactly as in Session 3. A 90% interval (from the dialog box) is narrower.	Slope: -3.31 to 15.07 at 95%; -1.83 to 13.60 at 90%. Note that the 95% interval contains zero - the same message as the p-value.
Multiple R	The correlation coefficient (Section 4.8).	0.00705
R Square	The coefficient of determination (Section 4.8).	0.00005 - "almost 0%"

CONCEPT TRAP: "A SLOPE OF 1 MEANS PERFECT CORRELATION"

A student mixed up the slope with the correlation coefficient. They are different animals. The correlation coefficient r lives between -1 and +1 and measures how tightly the points hug a straight line: +1 perfectly positive, -1 perfectly negative, 0 no linear relationship. The slope depends on the units - it could be 5.88 dollars per unit of 9Online here, or 30,000 rupees per square foot for house prices - and says how much y changes per unit of x. Only if all the points lie exactly on the line y = x do both happen to equal 1.

MANAGER'S TAKEAWAY

Reading a regression output is a fixed drill. For a simple regression: p-value of the slope below alpha? If not, stop - do not quote the coefficient, do not build a business case on it. If yes, write the equation, plug in the two or three cases that matter to the decision, and state the difference in units your audience understands: "$5.88 per online customer" not "the coefficient is significant".

4.8 Is Age Clouding the Effect? Multiple Linear Regression

After the break the professor returned to the question she had left hanging at the end of Session 3. The data is inconclusive on online versus offline. But our understanding of customers says something the averages ignore: the share of online customers is very unlikely to be the same in every age group.

Younger customers are more likely to be online; older customers are more likely to be offline. And older customers are more profitable - think of your own bank balance at twenty. So the online group is crowded with young, low-profit customers and the offline group with older, high-profit ones. The professor "cooked up" numbers to show what that can do - different numbers from the Session 3 hint, same lesson.

The professor's made-up numbers for 23 August (the 16 August version used -30 / -50 / 350 / 300). Same lesson.

	ONLINE	OFFLINE	
YOUNGER	**-$20 per person** 80 people	**-$70 per person** 20 people	**online wins by $50**
OLDER	**+$400 per person** 20 people	**+$300 per person** 80 people	**online wins by $100**
	(80 x -20 + 20 x 400) / 100 **= $64 average**	(20 x -70 + 80 x 300) / 100 **= $226 average**	

Column by column, OFFLINE looks $162 better. Row by row, ONLINE is better for both ages.

Figure 4.6. Two groups of 100 customers. Within each age, moving online raises profit ($50 for the young, $100 for the old). Across the whole file, offline "wins" by $162 - because the groups have opposite age mixes.

When an older customer moves from offline to online, profit rises from $300 to $400. When a younger customer moves online, the loss shrinks from $70 to $20. Within each age, online is better. Average the columns and online gives $64 against offline's $226 - offline looks *way* more profitable. The composition of the two groups is drowning the effect. The professor called it age **clouding out** the online effect. Does it happen in our data? That is what multiple regression checks.

Why regression rather than seven confidence intervals

One way to check would be to build a separate confidence interval, or run a separate regression, for each of the seven age categories - seven pieces of analysis, each on a smaller sample. The other way is one equation that holds age constant while it looks at the online effect:

> **THE MULTIPLE LINEAR REGRESSION**
>
> **9Profit = intercept + slope1 x 9Online + slope2 x 9Age**
> "Multiple" because there is more than one independent variable. The professor made a new sheet (MLR1) with columns B, C and D of the Data sheet - 9Profit, 9Online and 9Age - and immediately hit the problem somebody had raised in the very first session: 8,289 of the 9Age cells are blank. Regression cannot take a blank cell; it simply errors out.

First attempt: drop the customers without an age

There are two ways to deal with missing values: let go of those rows, or fill them in (imputation - Section 4.10). The class began with the primitive one. Select all the data, sort it by 9Age in descending order so that the blanks collect at the bottom, and run the regression on the rows above them. The good news: 23,345 customers with complete data. The bad news: the sample of 31,634 has shrunk by 8,289 - a quarter of the information thrown away. The regression itself is the same dialog as before, with Input X now spanning both 9Online and 9Age, down to row 23,346.

	Coefficient	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	17.080	5.059	3.376	0.0007	7.17	27.00
9Online	27.188	5.523	4.923	8.59E-07	16.36	38.01
9Age	25.857	1.125	22.992	1.09E-115	23.65	28.06
Multiple R 0.149 | R Square 0.0222 | Observations 23,345 | Significance F 1.67E-114						

THE SECOND REGRESSION EQUATION (customers with an age on file)

9Profit = 17.08 + 27.19 x 9Online + 25.86 x 9Age

The professor's interpretation drill

Before she filled in the values she asked a question the exam will ask: two customers, both online, one in age category 3 and one in category 4 - who is more profitable, and by how much? Work it through. The intercept 17.08 is the same for both. The online term 27.19 x 1 is the same for both. The only difference is the age term: 25.86 x 3 for the first customer and 25.86 x 4 for the second. So the category-4 customer is more profitable by **exactly $25.86** - the slope of age. Writing both out: $121.85 against $147.71.

Coefficient	How to read it	In one sentence
Slope of 9Age = 25.86	Holding the channel fixed - within online customers, or within offline customers - each step up in age category adds $25.86 of profit.	Older customers are more profitable, about $26 per category, whether they are online or not.
Slope of 9Online = 27.19	Offline is coded 0 and online 1, so moving from offline to online within the same age category adds $27.19. Equivalently: two customers of the same age, one online and one offline - the online one is $27.19 more profitable on average.	Like for like, online customers are worth about $27 more. (Madhav: so every offline older customer is a missed opportunity? Every offline customer of any age is.)
Intercept = 17.08	The predicted profit when both variables are zero - an offline customer in a non-existent age category 0. Not meaningful on its own; it anchors the line.	Do not interpret it as a customer type.

The checks: Significance F first, then the p-values

In a simple regression you went straight to the slope's p-value. In a multiple regression you do not. The first thing to look at is **Significance F** in the ANOVA table. It is a p-value too, but for a different test - an F-test, based on the F-distribution - whose null hypothesis is that **all the slopes are simultaneously zero**: slope1 = slope2 = 0. In three dimensions, imagine the cloud of points as a sphere - profit unrelated to 9Online and to 9Age alike. The alternate is that at least one slope is not zero, so the regression has something to capture.

Check	Rule	This regression
1. Significance F	Greater than your alpha (0.05)? Stop; the regression makes no sense. Smaller? Go on.	1.67E-114, which is a decimal point, 113 zeros and then 167 - effectively 0. Go on.

Check	Rule	This regression
2. Each slope's p-value	Any slope with a p-value above alpha? Do not use that variable as it stands; fix the model. All below? Use the equation.	9Online: 8.59E-07 = 0.000000859. 9Age: 1.09E-115. Both far below 0.05. (The intercept's p-value is not part of the check.)
3. Then R Square	How much of the variation the model explains - context decides what counts as good (below).	0.0222, about 2%.
4. Only then the coefficients	Write the equation; interpret each slope holding the others constant.	As above.

CONCEPT TRAP: READING "8.59E-07" AS 8.59

E-07 means "times ten to the power minus seven". 8.59E-07 is 0.000000859 - six zeros after the decimal point, then 859. It is enormously smaller than 0.05. Excel writes tiny p-values this way, and reading the mantissa without the exponent turns a significant result into an apparently hopeless one. When a p-value is shown as 0.000 or 0, Excel has run out of decimals: it is effectively zero.

The top table: Multiple R and R Square

Now the summary table. **Multiple R** is a measure of correlation. Correlation asks: when two variables are measured on the same subjects, do larger values of one go with larger values of the other, or with smaller ones? Steps walked per day and fitness level move together - positive correlation. Hours spent in a chair per day and fitness move in opposite directions - negative correlation. It is measured by the correlation coefficient r, between -1 and +1: every point on an upward straight line gives +1, on a downward one -1, a shapeless cloud gives 0, and real data lands in between.

R Square is the one that matters. Its formal name is the **coefficient of determination**, and its formal definition - the percentage of the variation in the dependent variable that is explained by the independent variables - drew a blank from the class ("Neither did I," said the professor, and explained it properly).

Imagine you were the god of house prices and could set every price yourself. Would area be the only thing you looked at? No: locality, city, builder, age of the building, distance from amenities, direction. You might say area decides 50% of a house's price and the rest is everything else. That weight is R Square. It is why two houses of the same 1,500 sq ft can sell for Rs. 80 lakh and Rs. 65 lakh: area explains part of why prices differ across houses, and the other factors explain the rest. Here, 9Online and 9Age together explain about 2% of why one customer's profit differs from another's; the other 98% is tenure, income, PIN code and things nobody measured.

Context	What R Square would count as	Why
Verifying Ohm's law in a physics practical	95% could earn you a B grade	It is a universal law; the missing 5% is your experimental error.
Which music to play in a Shoppers Stop store to lift purchase intention	10% would be an excellent result	Purchase intention depends on dozens of things; if music alone moves 10% of it, that is a big lever.
Pilgrim: profit on 9Online alone	0.005% - "almost 0%"	Being online, by itself, explains nothing about why profits differ.
Pilgrim: profit on 9Online and 9Age	2.2%	Adding age lifted it from nothing to 2%. Small - and enough for the question we are asking, which is the online effect, not "what drives profit".

A student asked whether 2% means we have the wrong variables. It depends on the question, the professor said. If the question is "what are the main drivers of customer profitability", then yes, these two are not them, and you would go on to add income, tenure and PIN code. If the question is "does the online channel matter, like for like", 2% is not bad at all - and the p-values say the effect is real. But before adding anything, there was a problem of our own making to fix: we had deleted 8,000 customers.

MANAGER'S TAKEAWAY

When two groups differ in composition - age, ticket mix, store size, contract size - never compare their raw averages. Put the confounding variable into the model and read the effect you care about "holding the rest constant". And judge R Square against the question: a small R Square with tiny p-values means a real but modest lever, not a failed analysis.

4.9 The Missing-Data Trap: Testing Whether Dropping Rows Creates Bias

Deleting the 8,289 customers without an age was technically fine - the regression ran - and possibly disastrous. Session 3 said a sample must be representative. Were the customers we removed like the ones we kept, or different in a way that matters?

Think about who does not bother to give a bank their age. Perhaps the casual customers: people who keep their main account elsewhere, opened this one for small things, and did not fill in the form carefully. If those are systematically the less profitable customers, then deleting them means deleting loss-makers, and every effect we estimate on the remaining 23,345 will be tilted - some overstated, some understated. That is what **bias** means in this context: the rows we dropped were not a random slice of the file.

The test: are the two groups equally profitable?

The professor asked the class to think like a statistician: what result would say "bias" and what result would say "no bias", before running anything? The logic is simple. If the missing-age customers were removed at random, their average profit should be roughly the same as everyone else's. If it is significantly higher or lower, dropping them changes the sample. The tool is a regression of profit on a **dummy variable**:

Step	What to do	Detail
1	Copy 9Profit and 9Age to a new sheet (the professor named it "check for bias").	Keep sheets clean and named; you will want to refer back.
2	Add a column **HasAge**: 1 if the age cell is filled, 0 if it is blank.	Formula: =IF(C2<>"", 1, 0), or after sorting simply type 1 down to row 23,346 and 0 below.
3	Regress 9Profit (Y) on HasAge (X), all 31,634 rows.	Remove any blank separator row first.
4	Look at the p-value first, then the slope.	A p-value above 0.05 would mean the slope may well be zero - no difference, no bias. A small p-value with a large slope means bias.

	Coefficient	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	72.963	2.986	24.433	1.2E-130	67.11	78.82
HasAge	52.225	3.476	15.024	7.7E-51	45.41	59.04
R Square 0.0071 | Observations 31,634 | Significance F 7.7E-51						

THE BIAS-CHECK EQUATION

9Profit = 72.96 + 52.22 x HasAge

A customer without age data: 72.96 + 0 = **$72.96** on average. A customer with age data: 72.96 + 52.22 = **$125.19**. The p-value of 7.7E-51 is effectively zero, so the difference is real: the customers we deleted are, on average, $52 less profitable than the ones we kept. That is a lot of bias. Dropping them was not a good idea - and we cannot go back and ask 8,289 people for their age.

What the bias did to the answer

Look back at the second regression's online slope: $27.19. A student asked at the end why we could not simply have stopped there, since online was already significant. Because that number was **overestimated**: by removing the loss-making customers who had not given an age, we compared a rosier subset. A bank that acted on $27 might offer a rebate of up to $27 - and would be paying more than the effect is worth. Section 4.10 shows the corrected figure.

CONCEPT TRAP: "THE REGRESSION RAN, SO THE SAMPLE IS FINE"

Software will happily fit a model to whatever rows survive your cleaning. Whether those rows still represent the population is your problem, not Excel's. Every time you drop rows for missing values, test whether the dropped rows differ from the kept rows on the outcome you care about. If they do not, you lose only some precision; if they do, you have quietly changed the population you are studying.

MANAGER'S TAKEAWAY

Missing data is rarely missing at random. The customers who skip the form, the tickets nobody categorised, the assets that never reported in - they tend to be different. Before deleting them, run the one-variable check: does "is missing" predict the outcome? If it does, keep the rows and use the method in the next section.

4.10 Imputation: Keep Every Row, Add a Flag

If you cannot drop the rows and cannot recover the values, you fill them in. Somebody had already suggested it. The professor's method is deliberately simple, and it works because of one extra column.

The rule

Rule	Why
Fill every missing 9Age with a value of your choice - 0, the average, the median, 5, 100, anything.	The regression needs a number in every cell.
Use the same value for every missing cell.	So that all the "plugged" customers are treated identically.
Keep an indicator column - HasAge: 1 for a real age, 0 for a plugged one - and put it in the regression.	This is what makes the choice of fill value irrelevant: the model can tell a real 0 from a plugged 0, and it estimates separately what "no age on file" is worth.
Then run the regression on all 31,634 rows with three independent variables: 9Online, 9Age (filled) and HasAge.	Every customer is back in the sample; the bias of Section 4.9 is gone.

The professor filled the blanks with 0. As she said, you could have used 1, 5 or 100 - with the indicator in place, the online and age slopes come out the same and the predictions for every customer are identical. Only the intercept and the indicator's coefficient rearrange themselves (see the box on average-imputation below).

	Coefficient	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	70.926	3.000	23.643	1.6E-122	65.05	76.81
9Online	19.649	4.685	4.194	2.75E-05	10.47	28.83
9Age (0 where missing)	25.603	1.086	23.582	6.7E-122	23.48	27.73
HasAge	-51.849	5.598	-9.263	2.1E-20	-62.82	-40.88
Multiple R 0.156 | R Square 0.0243 | Observations 31,634 | Significance F 4.8E-168						

Checks first: Significance F is 4.8E-168 - as close to zero as it gets - so the regression makes sense. Every slope's p-value is far below 0.05 (the intercept's is not checked). R Square is 2.4%. Now the equation.

THE FINAL REGRESSION EQUATION (all 31,634 customers)

9Profit = 70.93 + 19.65 x 9Online + 25.60 x 9Age - 51.85 x HasAge

where 9Age is 0 for customers with no age on file, and HasAge is 1 if the age is real and 0 if it was plugged.

Compare it with the biased version. The online slope has fallen from $27.19 to **$19.65** - "reduced considerably", as the student who ran it first reported. The age slope barely moved, 25.86 to 25.60. And a new number appears: HasAge at -51.85, which the professor turned into a set of questions.

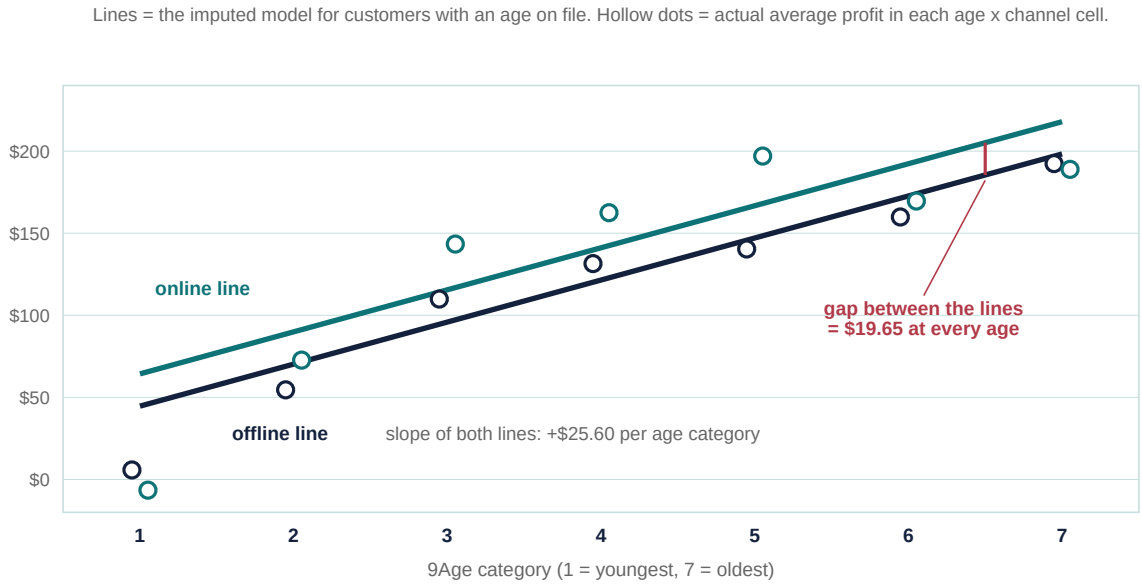


Figure 4.7. The final model for customers with an age on file: two parallel lines, $19.65 apart, rising $25.60 per age category. The hollow dots are the actual averages in each age-and-channel cell, so you can see how well the lines track the data.

The interpretation drill, question by question

The question	The working	The answer
Two customers, both without age data; one online, one offline. Who is more profitable, by how much?	For both, 9Age = 0 and HasAge = 0, so those terms vanish. Both get 70.93. The online one also gets 19.65 x 1.	The online customer, by **$19.65**: $90.58 against $70.93. "Among the casual customers, online is better by 19.64."
Two customers, both online; one in age category 7, one without age data. Who is more profitable, by how much?	Intercept and online term are common. Category 7: 25.60 x 7 - 51.85 = 179.20 - 51.85 = 127.35. No age: 0 - 0 = 0.	The category-7 customer, by about **$127** ($218.09 against $90.58).
Two customers, both online; one in age category 1, one without age data.	Category 1: 25.60 x 1 - 51.85 = -26.25. No age: 0.	The customer **without** age data, by about **$26**.
So, with age data or without - who is more profitable?	It depends on the age category: the with-age customer is worth 25.60 x category - 51.85 relative to a no-age customer. Category 1: -26.25. Category 2: -0.65. Category 3: +24.95.	Category **3 and above**: the customer with age data is more profitable. Category 2 and below: the customer without age data is.
Two offline customers, categories 2 and 5.	Only the age term differs: 3 categories x 25.60.	The category-5 customer, by **$76.80**.
Predicted profit of an online customer in category 4.	70.93 + 19.65 + 25.60 x 4 - 51.85	**$141.13**.

WHAT DOES -51.85 MEAN? (AND WHY IT IS NOT A CONTRADICTION)

HasAge is not "customers who gave their age are $52 less profitable". It works together with the age term. For a real age category k, the pair of terms is 25.60k - 51.85: negative for k = 1, about zero for k = 2, positive from k = 3. The indicator is simply re-anchoring the age line so that a plugged 0 is not read as "a customer of age zero". Read HasAge and 9Age together, always.

CONCEPT TRAP: THE AVERAGE-IMPUTATION VERSION ON THE CASE SLIDE

The slide shows the same model with blanks filled by the average age (4.05) instead of 0: 9Profit = -32.66 + 19.64 x 9Online + 25.60 x AgeAvg -51.73 x AgeExist. Re-running it gives an intercept of -32.66 and an AgeExist coefficient of +51.74, not -51.73 - the slide's minus sign is a typo. Check it yourself: a no-age customer must predict the same $70.93 in both versions, and -32.66 + 25.60 x 4.046 = 70.93. The online slope (19.65) and the age slope (25.60) are identical in both versions, which is exactly the point: the fill value does not matter once the flag is in.

MANAGER'S TAKEAWAY

Never delete a record because one field is blank. Fill the field with any constant, add a flag that says "this was filled", and put the flag in the model. You keep the whole population, you lose nothing, and the flag itself often tells you something - here, that the customers who do not fill in forms are a different kind of customer.

4.11 The Decision, and the Assignment Worked

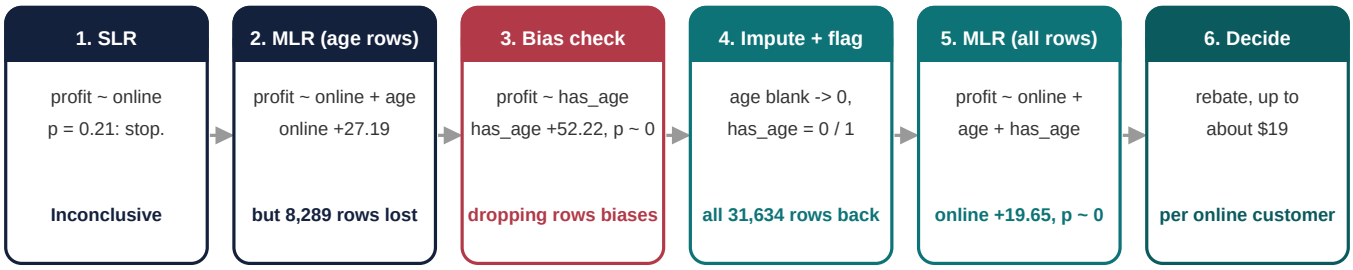
Twelve months of module, four sessions of Pilgrim Bank, and here is the answer - together with why the first answer we could have given would have been wrong, what the case itself wants a manager to learn, and the assignment the professor left for the week, worked so that you can check your own sheet.

The verdict: offer the rebate, and it tells you how much

A student asked whether every answer from a regression has to be hedged with conditions - "if online, and in age category 3 or above, and has declared age, then...". For understanding the output, yes, the professor said; those questions were exercises. For the Pilgrim decision, no. The bank wanted to know whether to charge a fee or offer a rebate. The final model says that **within any comparable cohort - same age category, or both without an age on file - online customers are on average $19.65 more profitable**. That answers the question directly, and it also sizes the answer: the rebate can go up to about $19 per online customer, and not beyond, because beyond that the bank would be paying more than the channel is worth.

If the bank had stopped at...	It would have concluded	What was wrong with it
The two averages ($110.79 vs $116.67)	Online is better by $6 - rebate	A sample difference with p = 0.21: as likely to be luck as not. Overlapping intervals, same message.
The regression on the 23,345 customers with an age	Online is better by $27.19 - rebate up to $27	The 8,289 deleted customers were $52 less profitable than the rest, so the sample was biased and the effect overstated. A $27 rebate would overpay.
The regression on all 31,634 with imputation and a flag	Online is better by $19.65, like for like - rebate up to about $19	The recommendation. Significance F and every p-value pass; the whole population is represented.

Notice how the two frames of Section 4.3 now converge. Under Frame A the null (offline at least as profitable) is rejected, so the rebate follows. Under Frame B the null (online at least as profitable) is not rejected, so the fallback is the rebate as well. When the evidence is strong, the frame stops mattering; the frame matters precisely when the evidence is weak, which is when a manager most needs to have chosen it deliberately.



Same 31,634 rows throughout - not one new customer was surveyed. The conclusion came from looking at the data differently.

Check order for any multiple regression: Significance F first, then each p-value, then R-squared, and only then read the coefficients.

Figure 4.8. The route from deadlock to decision, in six steps, on one unchanged data set.

What the case itself wants you to learn

The case slides close with a lesson that goes beyond the statistics. Managers design service models around beliefs about how customers behave - that online customers cost less to serve, that they stay longer. Too often those beliefs are treated as facts and left unchallenged for years, so that important changes in behaviour go unnoticed. The service manager's job is to **surface the implicit assumptions, convert them into testable hypotheses, test them, and update the model**. The slides call this managerial research design, and they give the Pilgrim version in five steps.

Step	In general	At Pilgrim Bank
1. Identify the business problem or question	Pose a question that forces a strong opinion, then reverse-engineer the assumption behind it.	"Should Pilgrim charge a fee or offer a rebate?" surfaces the assumption: the impact of the online channel on customer profitability.
2. Refine it into clear, researchable hypotheses	Convert assumptions into true/false statements. "Online customers have a lower cost to serve after adoption." "Online customers stay with the bank longer."	Online adoption causes increased customer profitability.
3. Identify the data requirements	What would you need to observe to test that?	Customer profitability, online status, demographics.
4. Identify and perform the analyses	Choose a test that fits the time, resources and data you have. With only cross-sectional data, modify the hypothesis to what it can test.	"Online customers are more profitable than offline customers" - the comparison of averages, then the regressions.
5. Interpret and communicate the results	Update the assumption, and the service model if required.	Online adoption is associated with higher profitability, like for like - about $19.65.

WHAT THE DATA CAN AND CANNOT SAY

The hypothesis in step 2 says "causes". The data is cross-sectional, so step 4 quietly changes it to "are more profitable". Session 3 explained why: without following the same customers over time, we cannot prove that moving online *makes* a customer more profitable - only that, holding age constant, the online ones *are*. The rebate decision is still sound: it is the best evidence available at a reasonable cost, and not deciding is not an option.

The assignment: do it again with income - and beyond

The homework is to repeat the whole route with income in place of age, as if 9Age did not exist: 9Profit on 9Online first (the answer will not change); add 9Inc, dropping the customers without income data; test whether dropping them creates bias; if it does, create HasInc, impute, and compare. In the long run the professor wants a comprehensive model with age, income, tenure and, if possible, the district. Using an AI assistant to run the regressions is fine - "even ChatGPT would do it for you"; what she wants is that you can **interpret the output**. Her own solution goes on the LMS around Thursday. The numbers below are worked from the raw data for you to check against - not to copy.

Model	n	Online slope	Other slopes	Checks	What it says
9Profit on 9Online, 9Inc - customers with income only	23,373	-4.91 (p = 0.37)	9Inc +17.77 (p ~ 0)	Significance F ~ 0; online fails its p-value	Holding income constant, no significant online effect in this subset. Income adds about $17.77 per slab.
Bias check: 9Profit on HasInc	31,634	-	HasInc +54.32 (p ~ 0); intercept 71.37	Passes	Customers without income data are $54 less profitable - dropping them biases the sample, exactly as with age.
9Profit on 9Online, 9Inc (0 where missing), HasInc	31,634	-4.72 (p = 0.31)	9Inc +17.77; HasInc -42.53	Significance F ~ 0; online fails its p-value	Still no significant online effect once income is held constant - a genuinely different picture from the age model.
9Profit on 9Online, 9Age, HasAge, 9Inc, HasInc (both imputed)	31,634	+11.87 (p = 0.011)	9Age +26.89; HasAge -94.31; 9Inc +18.77; HasInc -63.55	All pass; R Square 4.4%	Holding both age and income constant, online customers are about $12 more profitable.
9Profit on 9Online, 9Tenure (the case slide's practice model)	31,634	+16.55 (p = 0.0003)	9Tenure +6.21 per year	All pass; R Square 3.7%	Same tenure, online is worth about $16.55 more; same channel, each extra year of tenure adds $6.21. (The slide quotes $14.67 and $6.21.)
Comprehensive: online, age, income (both imputed with flags), tenure, district (base 1200)	31,634	+13.82 (p = 0.003)	9Age +16.67; 9Inc +16.85; 9Tenure +4.75; district 1100 -21.19; 1300 -13.20	All pass; R Square 6.2%	Like for like on everything measured, online customers are about $14 more profitable - the number a careful bank would size the rebate to.

Two things worth noticing before Thursday. First, 7,728 customers are missing *both* age and income (561 are missing only age, 533 only income), which is why the two bias checks look alike: it is largely the same group of casual customers. Second, the online effect is not one number - it is $19.65 holding age constant, insignificant holding only income constant, $11.87 holding both, $13.82 holding everything measured. That is not the statistics misbehaving; it is what "like for like" means, and the honest recommendation sizes the rebate to the most complete comparison you can defend. If the professor's solution frames it differently, follow hers for the exam and understand why.

HOW THE EXAM WILL TEST THIS

No Excel in the exam - "your Excel may not work, you might make a small mistake". She will give you an output like the ones in this part and ask questions like the ones she asked: two customers, one in this category and one in that, who is more profitable and by how much; who should get the rebate; what should the condition be for upgrading a customer's credit card. Section 4.12 is the drill.

MANAGER'S TAKEAWAY

Size the incentive to the effect, not to the enthusiasm. The first significant number you find ($27) is not the number; the number is the one that survives the bias check and holds the confounders constant ($19.65, or $14 with everything in). And write the assumption you are testing on the wall - "online customers cost less to serve" - because the day it stops being true, your service model stops making money.

4.12 Reading Any Regression Output: The Exam Drill

Everything the professor asked in class follows one routine. Do it in the same order every time and the questions answer themselves.

#	Look at	Ask	If the answer is no
1	Significance F (ANOVA table) - multiple regression only	Is it below alpha (0.05 unless told otherwise)?	Stop. The regression makes no sense; all slopes may be zero.
2	P-value of each slope (not the intercept)	Is every one below alpha?	Do not use the equation as it stands; the offending variable needs fixing or removing.
3	R Square	How much of the variation does the model explain, and is that reasonable for this kind of question?	Low is not a failure if the p-values pass and the question is about one lever; it is a failure if the question was "what drives y".
4	Coefficients	Write y = intercept + slope1 x x1 + slope2 x x2 ...	-
5	The question's two customers	Write each customer's equation; strike out the terms that are identical; the difference is what remains.	-
6	The sentence	"Holding the other variables constant, a one-unit increase in x changes y by (slope)". For a 0/1 variable: "compared with the 0 group, the 1 group is (slope) higher".	-

Drill 1: the tenure model

Output (all 31,634 customers): Significance F = 3.8E-259; 9Online coefficient 16.553, p = 0.0003; 9Tenure coefficient 6.211, p ~ 0; intercept 46.364; R Square 0.037. Equation: **9Profit = 46.36 + 16.55 x 9Online + 6.21 x 9Tenure**.

Question	Working	Answer
Do the checks pass?	Significance F far below 0.05; both slopes far below 0.05.	Yes - use the equation.
Two customers with the same tenure, one online and one offline: who is more profitable, by how much?	The tenure terms cancel; the online one has 16.55 x 1 extra.	The online customer, by $16.55.
Two offline customers, one with 3 years and one with 8 years of tenure.	Five extra years x 6.21.	The 8-year customer, by $31.05.
An online customer with 10 years against an offline customer with 12 years.	Online advantage 16.55; tenure disadvantage 2 x 6.21 = 12.42.	The online 10-year customer, by $4.13.
Predicted profit of an offline customer with 5 years.	46.36 + 0 + 6.21 x 5	$77.41.
Predicted profit of an online customer with 20 years.	46.36 + 16.55 + 124.20	$187.11.

Drill 2: the final age model - decision questions

Equation: **9Profit = 70.93 + 19.65 x 9Online + 25.60 x 9Age - 51.85 x HasAge**, with 9Age = 0 and HasAge = 0 for customers without an age on file.

Question	Working	Answer
Who should the bank offer the rebate to?	The online slope is +19.65 with p = 2.75E-05, and it holds within every age category and for customers without an age.	Every customer who uses the online channel - the effect is not confined to a cohort. Size it up to about $19.
Should the bank push its older offline customers online?	An offline category-7 customer predicts 70.93 + 179.20 - 51.85 = $198.28; online the same customer predicts $217.93.	Yes - each one is about $19.65 of missed opportunity, and they are the most valuable customers to begin with.
A card-upgrade rule: which customers predict above $150?	Solve 70.93 + 19.65 x online + 25.60 x age - 51.85 > 150. Offline with age: 25.60 x age > 130.92, so age >= 6. Online with age: 25.60 x age > 111.27, so age >= 5.	Offline customers in categories 6-7; online customers in categories 5-7. No-age customers never cross it (max $90.58).
Two customers, both offline, one in category 4 and one without age data.	25.60 x 4 - 51.85 = 50.55 against 0.	The category-4 customer, by $50.55.
What is the profit difference between an online category-2 customer and an offline category-3 customer?	Online cat 2: 70.93 + 19.65 + 51.20 - 51.85 = 89.93. Offline cat 3: 70.93 + 76.80 - 51.85 = 95.88.	The offline category-3 customer, by $5.95 - one age step (25.60) outweighs the online effect (19.65).

Exam trap	Why people fall for it	The fix
Reading a slope's size before its p-value	Big numbers look important.	Order: Significance F, p-values, then coefficients. A slope with p = 0.21 does not exist for decision purposes.
Interpreting the intercept as a customer	"The average customer earns $70.93."	The intercept is the prediction when every x is zero - often a customer that cannot exist. It anchors the line; that is all.
Reading HasAge on its own	-51.85 looks like "giving your age costs $52".	Read it with the age term: 25.60 x category - 51.85.
Comparing raw group averages when the groups differ in mix	It is the obvious first thing to do.	Put the mix variable in the model; interpret the effect holding it constant.
Treating 8.59E-07 as 8.59	The exponent is small print.	E-07 means seven places to the left of the decimal point: 0.000000859.
Believing "not rejected" means "no effect"	It sounds like a verdict.	It is the absence of a verdict. More data or a better model may still reject.

4.13 Product, Service and the Manager's Desk

Everything in this session was taught on a bank's customer file, a shampoo and a baby on a train. The ideas transplant. Here is each of them in a product business and a service business, followed by what they mean for the way you run a team.

Concept	Product-based company	Service-based company
Frame the alternate as what you want to prove	New formula: H1 "reduces hair fall". Only rejecting the null lets you advertise.	New SOC playbook: H1 "cuts mean time to respond". Only rejecting the null justifies the rollout.
Choose alpha by the cost of a wrong rejection	Adopting a vendor's automation line: 1%, because reversing it means re-hiring and re-tooling.	A new ticket-routing rule that can be switched off tomorrow: 10% is fine.
Type I vs Type II	Recalling a good batch (Type I) vs shipping a bad one (Type II). Which one your QA is tuned to avoid is a values decision.	Escalating a benign alert (Type I) vs missing a real intrusion (Type II). A SOC tunes for Type II and pays with analyst hours on false positives.
p-value	"If the new line made no difference, how often would a pilot show this much improvement?"	"If the two shifts were equally fast, how often would 400 sampled tickets show this gap?"
Regression with a 0/1 variable	Sales per store on "new packaging (1) / old (0)": the slope is the packaging effect.	Resolution time on "senior analyst (1) / junior (0)": the slope is the seniority effect.
Confounding - the clouded effect	New packaging "sells worse" because it was piloted in small stores. Add store size to the model.	Night shift "resolves faster" because it gets fewer high-severity tickets. Add severity to the model.
Missing data and bias	Warranty cards returned by 30% of buyers - the ones who bother are not the average buyer. Flag "returned card" rather than dropping the rest.	CSAT answered by 8% of customers. Model the outcome with a "responded" flag before believing the score.
Imputation with a flag	Blank "age of machine" fields set to 0 with HasAge = 1/0 in the model.	Blank "contract size" fields set to 0 with HasSize = 1/0.

Concept	Product-based company	Service-based company
R Square in context	Fill-weight variation explained by line speed at 90%: tight process. Sales explained by ad spend at 15%: still a useful lever.	MTTR explained by ticket severity at 40%; by analyst tenure at 5% - small but real if the p-value passes.
Update the service model	The product spec that "customers want more features" - test it yearly.	The service design that assumes online customers cost less to serve - test it, because behaviour changes.

The manager's checklist from this session

#	Discipline	What it looks like on Monday morning
1	Write the alternate first	What do you want to be able to say after the test? That is H1. Its negation is H0.
2	Decide the fallback before the data	If the test is inconclusive, what will you do? Make sure you can live with it - that is the real choice between Frame A and Frame B.
3	Set alpha from the cost of reversal	10% for a cheap, reversible experiment; 5% by default; 1% or lower when the commitment is hard to undo.
4	Name both errors out loud	Which mistake is your organisation tuned to avoid - the false alarm or the missed case? Say it; the test only guards one.
5	p below alpha, act; above, wait	A point estimate with p = 0.21 is an anecdote, however attractive. Do not build a case on it.
6	Read the output in order	Significance F, p-values, R Square, then the equation. Never the coefficient first.
7	Never compare raw averages of unlike groups	Put the mix variable in the model and read the effect holding it constant.
8	Check what you deleted	Before dropping rows for missing values, test whether "is missing" predicts the outcome. If it does, impute with a flag.
9	Size incentives to the honest effect	The rebate is $19.65, not $27, and possibly $14 once everything measured is held constant.
10	Keep the verdict provisional	More data can flip a hypothesis test - the helicopter. Revisit the assumptions your service model rests on.

The sharpest lines of the session

> *"All of us do hypothesis testing - not just once, but day in, day out."*
>
> **On the Delhi Metro once-over**

> *"Whatever you want to establish, put it as your alternate hypothesis."*
>
> **The framing rule**

> *"Not rejecting is not the same as accepting. The only hypothesis you can ever accept is the alternate."*
>
> **On the two verdicts**

> *"In probability we always suspect a miracle. We would rather believe our assumption is wrong than that the miracle happened."*
>
> **On the p-value**

> *"Hypothesis testing is not written in stone. The more data you get, it may change."*
>
> **The toy helicopter, thirty years on**

> *"Whatever tool you use, if you have the same data it should give you the same result. Just because a fancier name comes in does not mean it gives you a different answer."*
>
> **On confidence intervals versus regression**

> *"The prompts come from understanding the business and not being scared of it."*
>
> **On using AI to run the regressions - and why you still need to understand them**

> **ONE-BREATH SUMMARY**
>
> A hypothesis test is two statements about the population - a null and an alternate, mutually exclusive and collectively exhaustive - judged from a sample; you can only ever reject the null and accept the alternate, so the thing you want to prove goes in the alternate. Two errors are possible: rejecting a true null (guarded by alpha, chosen from the cost of being wrong) and keeping a false null (shrunk only by more data). The p-value is how rare your sample would be if the null were true; below alpha, reject. Regression fits the best line through a cloud and tests, for every slope, whether it could be zero; read Significance F, then the p-values, then R Square, then the equation. Groups that differ in mix must be compared holding the mix constant, and rows with missing values must be kept - imputed with a flag - because dropping them biased the Pilgrim sample by $52 a customer. Done properly, online customers are $19.65 more profitable like for like: offer a rebate, and cap it there.

4.14 Practice MCQ Bank with Answers

One hundred and six questions drawn from this session, in roughly the order the ideas were taught. The options are meant to be close - several are numerically right for the wrong reason - so read the one-line reason under each answer even when you got it right. All numbers refer to the Pilgrim Bank 1999 data and the regressions in this part unless stated otherwise.

Q1. A hypothesis test always involves:

(a) One statement about the sample

(b) Two statements about the population, tested with sample data

(c) Two statements about the sample, tested with population data

(d) As many statements as there are variables

Answer: (b) - H0 and H1 are claims about the population; the sample is only the evidence.

Q2. The null and alternate hypotheses must be:

(a) Mutually exclusive and collectively exhaustive

(b) Both true at once

(c) Both testable with a census

(d) Independent of each other

Answer: (a) - They cannot both be true, and one of them must be - together they cover every possibility.

Q3. On the Delhi Metro, the "sample" the passengers used was:

(a) The other passengers in the compartment

(b) A lifetime of observing what usually goes with what

(c) The baby's birth certificate

(d) The police record

Answer: (b) - Nobody knew the truth; everyone compared what they saw with their experience and found a mismatch.

Q4. The only two possible outcomes of a hypothesis test are:

(a) Accept H0 or accept H1

(b) Reject H0 or do not reject H0

(c) Prove H1 or prove H0

(d) Significant or insignificant sample

Answer: (b) - Both verdicts are worded around the null.

Q5. "Do not reject H0" means:

(a) H0 has been proven

(b) The evidence is not strong enough to decide - inconclusive

(c) H1 has been proven false

(d) The sample was too small

Answer: (b) - Not rejecting is not accepting; the professor repeated it twice on request.

Q6. The only hypothesis you can ever accept is:

(a) The null

(b) The alternate

(c) Whichever has the higher p-value

(d) Both, with enough data

Answer: (b) - Rejecting H0 is the only conclusive verdict, and it means accepting H1.

Q7. The rule for framing a hypothesis is:

(a) Put the status quo in the alternate

(b) Put whatever you want to establish in the alternate; its negation becomes the null

(c) Put the larger group in the null

(d) Put the action you will take in the null

Answer: (b) - Only "reject H0" is decisive, so the statement you need must be the one that gets accepted.

Q8. Case 1 for the shampoo (H0: the formula reduces hair fall) is useful to:

(a) The FMCG company that wants to advertise

(b) A competitor or consumer advocacy group that wants to prove the claim false

(c) Nobody

(d) The regulator only

Answer: (b) - Rejecting that null establishes that the formula does not work - their win, not the company's.

Q9. Under Case 2 (H0: does not reduce; H1: reduces), rejecting H0 lets the company:

(a) Advertise, and defend the claim in court with the data

(b) Only run more tests

(c) Withdraw the product

(d) Accept that the formula is no better

Answer: (a) - That is why the company frames it this way: one of the two outcomes lets it act.

Q10. For Pilgrim Bank, "H0: charge a fee" is a poor hypothesis because:

(a) Fees are unpopular

(b) Hypotheses must be about what the data shows (profit), not the action you will take

(c) It has no alternate

(d) The fee amount is unknown

Answer: (b) - The fact being tested is who is more profitable; fee or rebate is what you do with the verdict.

Q11. In the Pilgrim hypotheses, the equality sign always appears in:

(a) The alternate

(b) The null

(c) Neither

(d) Whichever is convenient

Answer: (b) - The test begins by assuming the null is true and asking how likely the sample is; "equal" is the natural starting point.

Q12. Frame A: H0 average offline >= average online. If H0 is rejected, the bank:

(a) Charges a fee

(b) Offers the rebate

(c) Does nothing

(d) Collects more data

Answer: (b) - Rejecting establishes that online customers are more profitable.

Q13. Frame B: H0 average offline <= average online. If H0 is not rejected, the bank:

(a) Charges a fee

(b) Offers the rebate as its fallback

(c) Does nothing

(d) Switches to Frame A

Answer: (b) - And that fallback is a calculated hit: costs may rise if online is not really more profitable.

Q14. The risk of Frame A's fallback (fee when inconclusive) is:

(a) Higher servicing cost

(b) Customer pushback from a blanket fee on thin evidence, though no increase in the bank's cost

(c) Regulatory penalty

(d) Losing the online channel

Answer: (b) - Frame B's fallback risks cost; Frame A's risks customers.

Q15. Choosing between Frame A and Frame B is:

(a) A statistical calculation

(b) A management decision about which mistake you would rather risk

(c) Determined by the sample size

(d) Determined by the p-value

Answer: (b) - Same data, opposite default actions; the frame encodes the risk appetite.

Q16. A Type I error is:

(a) Not rejecting a false null

(b) Rejecting a true null

(c) Using too small a sample

(d) Choosing the wrong alpha

Answer: (b) - On the Metro: reporting a mother whose baby really is hers.

Q17. A Type II error is:

(a) Rejecting a true null

(b) Not rejecting a false null

(c) Rejecting the alternate

(d) A calculation error

Answer: (b) - On the Metro: doing nothing while a kidnapper walks away.

Q18. The level of significance, alpha, is:

(a) The probability of a Type II error

(b) The tolerance for a Type I error - the maximum probability you accept of rejecting a true null

(c) The confidence level

(d) The correlation coefficient

Answer: (b) - The methodology keeps the Type I probability within alpha; it does not directly control Type II.

Q19. The most commonly used alpha is:

(a) 10%

(b) 5%

(c) 1%

(d) 0.1%

Answer: (b) - 10% for low-stakes decisions, 1% for high stakes, 0.1% or 0.01% for nuclear plants.

Q20. A vendor claims its automation will raise your productivity; adopting means redesigning processes and re-staffing. The appropriate alpha is:

(a) 10%, to give the vendor a fair chance

(b) 1%, because the cost of going wrong is very high

(c) 50%

(d) It does not matter

Answer: (b) - You want the pilot to show a significant improvement before committing.

Q21. The way to reduce the probability of a Type II error is:

(a) Lower alpha

(b) A larger sample

(c) A higher alpha

(d) Reframe the null

Answer: (b) - The test guards Type I; Type II shrinks with sample size.

Q22. If you push the probability of Type I error down, the probability of Type II error:

(a) Also falls

(b) Rises

(c) Stays the same

(d) Becomes zero

Answer: (b) - The two trade off - the legal-systems example.

Q23. In the Indian legal system, H0 is that the accused is innocent. A Type I error is therefore:

(a) Setting a guilty person free

(b) Punishing an innocent person

(c) Delaying the verdict

(d) Refusing to hear the case

Answer: (b) - Rejecting the true null "innocent" convicts an innocent.

Q24. In a system where H0 is that the accused is guilty, a Type I error is:

(a) Punishing an innocent person

(b) Setting a guilty person free

(c) Both

(d) Neither

Answer: (b) - Rejecting the true null "guilty" frees the guilty; the Type II error there is punishing an innocent.

Q25. India's layers of appeal up to the President for capital punishment exist to:

(a) Speed up justice

(b) Push the probability of punishing an innocent (Type I) towards zero - at the cost of slow justice

(c) Reduce Type II error

(d) Raise alpha

Answer: (b) - Each layer is another check against the error the system refuses to make.

Q26. Why did nobody believe the sibling at age six?

(a) Siblings always lie

(b) Given past behaviour, the probability that they broke it was high - their innocence would have been a miracle

(c) There was CCTV

(d) The wind never blows indoors

Answer: (b) - The same logic as the p-value: suspect the explanation that requires a miracle.

Q27. Thirty years later most people believe the sibling. The lesson is:

(a) Memories fade

(b) More data - and no remaining incentive to lie - can change the verdict; hypothesis testing is not written in stone

(c) Beer improves honesty

(d) The null was wrong from the start

Answer: (b) - Same hypothesis, larger data set, different decision.

Q28. The p-value is:

(a) The probability that H0 is true

(b) The probability of getting a sample like ours if H0 is true

(c) The probability that H1 is true

(d) The level of significance

Answer: (b) - It is about the sample given the null, not about the null given the sample.

Q29. You reject H0 when:

(a) The p-value is greater than alpha

(b) The p-value is less than alpha

(c) The p-value equals the confidence level

(d) The slope is positive

Answer: (b) - A sample that would be a near-miracle under H0 is evidence against H0.

Q30. A p-value of 0.21 with alpha = 0.05 means:

(a) Reject H0

(b) Do not reject H0 - a sample like this occurs one time in five under H0

(c) H0 is proven

(d) The test is invalid

Answer: (b) - Not a miracle, so the null stands unrefuted.

Q31. The two ways of testing a hypothesis taught in the module are:

(a) Regression and correlation

(b) Confidence-interval overlap and the p-value

(c) Mean and median

(d) t Stat and R Square

Answer: (b) - Both must agree on the same data - and did (inconclusive, then p = 0.21).

Q32. In a regression, the independent variable is:

(a) The one you want to predict

(b) The one whose value is known or decided first; the input or predictor, x

(c) Always a 0/1 variable

(d) The one with the larger mean

Answer: (b) - The dependent variable is the response, y, which happens afterwards.

Q33. In the Pilgrim regression, 9Profit is:

(a) The independent variable

(b) The dependent (response) variable

(c) A dummy variable

(d) The intercept

Answer: (b) - The customer first decides to go online; the profit follows.

Q34. The first thing to do in any regression is:

(a) Compute the p-value

(b) Plot the scatter chart

(c) Delete missing rows

(d) Compute R Square

Answer: (b) - The scatter tells you whether a linear model is even sensible.

Q35. The best-fit line is the line that:

(a) Passes through the maximum number of points

(b) Is close to all the points, even if it passes through none

(c) Starts at the origin

(d) Has slope 1

Answer: (b) - The school-practical mistake is drawing through the most points.

Q36. To describe a straight line you need:

(a) Only its slope

(b) Its slope and its intercept

(c) Two p-values

(d) Its R Square

Answer: (b) - Many parallel lines share a slope; the intercept says where the line cuts the y-axis.

Q37. "Simple" in simple linear regression means:

(a) A small sample

(b) One independent variable

(c) No intercept

(d) A 0/1 dependent variable

Answer: (b) - Multiple regression has more than one independent variable; linear refers to the shape.

Q38. A COVID case curve that stays flat and then shoots up would call for:

(a) Simple linear regression

(b) Non-linear regression

(c) No regression

(d) A dummy variable

Answer: (b) - The scatter does not look like a straight band.

Q39. Excel's regression tool is found under:

(a) Insert > Chart

(b) Data > Data Analysis (after enabling the Analysis ToolPak add-in)

(c) Formulas > Statistical

(d) Review > Analyse

Answer: (b) - File > Options > Add-ins > Manage Excel Add-ins > Go > Analysis ToolPak.

Q40. A student's regression threw an error because:

(a) The data had negative profits

(b) The whole column was selected, including the blank cells below the data

(c) Labels were ticked

(d) The sample exceeded 30

Answer: (b) - Select from the label cell with Ctrl + Shift + Down Arrow instead.

Q41. Ticking "Labels" in the regression dialog:

(a) Adds a chart

(b) Names the variables in the output (9Online instead of X Variable 1)

(c) Changes the confidence level

(d) Removes the intercept

Answer: (b) - With several variables, unnamed outputs become confusing.

Q42. Every regression output has three tables. They are:

(a) Mean, SD, count

(b) Summary statistics, ANOVA, coefficients

(c) Slope, intercept, residuals

(d) Input, output, chart

Answer: (b) - Whatever the software, the same three appear.

Q43. The first Pilgrim regression equation is:

(a) 9Profit = 5.88 + 110.79 x 9Online

(b) 9Profit = 110.79 + 5.88 x 9Online

(c) 9Profit = 116.67 + 5.88 x 9Online

(d) 9Profit = 110.79 - 5.88 x 9Online

Answer: (b) - Intercept 110.786, slope 5.881.

Q44. The predicted profits from that equation for offline and online customers are:

(a) $110.79 and $116.67 - the two sample averages

(b) $5.88 and $116.67

(c) $110.79 and $121.85

(d) $116.67 and $110.79

Answer: (a) - A regression on a 0/1 variable reproduces the group means; the slope is their difference.

Q45. The p-value of 9Online in the first regression is 0.21. Therefore:

(a) Use the equation: online is $5.88 better

(b) Do not use the equation - the same verdict as the overlapping intervals

(c) Charge a fee

(d) The intercept is wrong

Answer: (b) - 21% is above every usual alpha; the slope could well be zero in the population.

Q46. The hypothesis behind every slope's p-value is:

(a) H0: slope = 1

(b) H0: slope = 0 (no relationship) against H1: slope not 0

(c) H0: intercept = 0

(d) H0: R Square = 0

Answer: (b) - A zero slope is a flat line - the independent variable does not move the dependent one.

Q47. The "Standard Error" next to a coefficient is:

(a) The SD of the profit data

(b) The standard deviation of the estimate - how much the coefficient would vary across samples

(c) The p-value

(d) The residual

Answer: (b) - Estimates are random variables; this is their spread.

Q48. Why are p-values more popular than t Stats?

(a) They are more accurate

(b) They can be compared directly with alpha, without looking up a t-table

(c) They are always smaller

(d) Excel does not print t Stats

Answer: (b) - Both test the same hypothesis; the p-value skips the table lookup.

Q49. If the regression dialog's confidence level is left blank, Excel:

(a) Prints no intervals

(b) Prints the 95% limits twice

(c) Uses 90%

(d) Uses 99%

Answer: (b) - Enter 90 and you get the 90% limits alongside the 95% ones.

Q50. The 95% interval for the 9Online slope is -3.31 to 15.07. This tells you:

(a) The slope is certainly positive

(b) The interval contains zero - consistent with p = 0.21, the slope may be zero

(c) Online is better by $15

(d) The sample is too small

Answer: (b) - A coefficient whose interval straddles zero is not significant at that level.

Q51. The correlation coefficient r lies between:

(a) 0 and 1

(b) -1 and +1

(c) 0 and infinity

(d) -infinity and +infinity

Answer: (b) - +1 perfectly positive, -1 perfectly negative, 0 no linear relationship.

Q52. A slope of 1 means perfect correlation. This statement is:

(a) True

(b) False - the slope depends on units, and only on the line y = x do both equal 1

(c) True for positive slopes

(d) True for large samples

Answer: (b) - Slope and correlation coefficient are different measures.

Q53. In the professor's clouding-out grid (23 August), the average profits were:

(a) Online $64, offline $226

(b) Online $226, offline $64

(c) Online $46, offline $230

(d) Both $145

Answer: (a) - (80 x -20 + 20 x 400)/100 = 64; (20 x -70 + 80 x 300)/100 = 226.

Q54. In that grid, within each age group online was better by:

(a) $50 for the young and $100 for the old

(b) $162 for both

(c) $20 and $70

(d) Nothing

Answer: (a) - -20 vs -70 and 400 vs 300; the reversal comes from the opposite age mixes.

Q55. Instead of seven separate confidence intervals by age category, the professor used:

(a) A pivot table

(b) One multiple linear regression with 9Online and 9Age

(c) Seven scatter charts

(d) A t-table

Answer: (b) - The equation holds age constant while it reads the online effect.

Q56. The regression on the 23,345 customers with an age gave:

(a) 9Profit = 17.08 + 27.19 x 9Online + 25.86 x 9Age

(b) 9Profit = 70.93 + 19.65 x 9Online + 25.60 x 9Age

(c) 9Profit = 110.79 + 5.88 x 9Online

(d) 9Profit = 72.96 + 52.22 x 9Age

Answer: (a) - The second is the imputed model (without its HasAge term); the fourth is the bias check.

Q57. Two online customers, age categories 3 and 4, under that equation. The category-4 customer is more profitable by:

(a) $27.19

(b) $25.86

(c) $17.08

(d) $51.72

Answer: (b) - Only the age term differs: one category step x 25.86.

Q58. The 9Online slope of 27.19 in that model means:

(a) Online customers are 27% of the file

(b) Within the same age category, online customers are $27.19 more profitable on average than offline

(c) Online customers are $27.19 more profitable in total

(d) Each age step adds $27.19

Answer: (b) - Holding age constant is what "within the same age category" means.

Q59. How many customers remained after dropping those without an age?

(a) 31,634

(b) 23,345

(c) 8,289

(d) 27,780

Answer: (b) - 8,289 rows were lost; 31,634 - 8,289 = 23,345.

Q60. In a multiple regression, the first thing to check is:

(a) The intercept

(b) Significance F in the ANOVA table

(c) R Square

(d) The largest coefficient

Answer: (b) - Only if it is below alpha do you move on to the individual p-values.

Q61. The null hypothesis of the F-test behind Significance F is:

(a) The intercept is zero

(b) All the slopes are simultaneously zero

(c) R Square is one

(d) The sample is biased

Answer: (b) - The alternate is that at least one slope is not zero, so the regression captures something.

Q62. A Significance F of 1.67E-114 means:

(a) Stop - the regression is invalid

(b) Effectively zero - go on to the p-values

(c) The slopes are 1.67

(d) Alpha should be raised

Answer: (b) - A decimal point, 113 zeros, then 167.

Q63. 8.59E-07 equals:

(a) 8.59

(b) 0.000000859

(c) 859

(d) 0.0859

Answer: (b) - E-07 shifts the decimal seven places to the left.

Q64. Which p-value is NOT part of the check before using a regression?

(a) The p-value of 9Online

(b) The p-value of 9Age

(c) The p-value of the intercept

(d) Significance F

Answer: (c) - Only the slopes are checked; the intercept is an anchor.

Q65. Multiple R in the summary table is:

(a) The coefficient of determination

(b) A measure of correlation

(c) The sample size

(d) The p-value

Answer: (b) - R Square is the coefficient of determination.

Q66. R Square is defined as:

(a) The slope squared

(b) The percentage of variation in the dependent variable explained by the independent variables

(c) The probability the model is correct

(d) The correlation between the two independent variables

Answer: (b) - The "god of house prices" weight that area gets versus everything else.

Q67. An R Square of 95% when verifying Ohm's law would be:

(a) Excellent

(b) Poor - it is a universal law, so the missing 5% is experimental error

(c) Impossible

(d) Typical for social science

Answer: (b) - Context decides what counts as good.

Q68. An R Square of 10% for in-store music predicting purchase intention would be:

(a) A failed study

(b) Very good - purchase intention depends on dozens of things, and music alone moves a tenth of it

(c) Exactly average

(d) Meaningless without a larger sample

Answer: (b) - Small R Square with significant p-values means a real but modest lever.

Q69. The R Square of the first regression (9Profit on 9Online) was:

(a) About 2%

(b) About 0.005% - effectively zero

(c) 21%

(d) 50%

Answer: (b) - 0.00005; adding age raised it to about 2.2%.

Q70. Dropping the customers without an age could bias the sample because:

(a) It reduces n below 30

(b) Those may be systematically the casual, less profitable customers - not a random slice

(c) Excel requires equal groups

(d) Age is the dependent variable

Answer: (b) - Representativeness was the Session 3 requirement, and deletion can destroy it.

Q71. The professor's test for bias was:

(a) Compare the two age slopes

(b) Regress 9Profit on a dummy HasAge (1 if age present, 0 if blank) and look at the p-value and slope

(c) Draw two histograms

(d) Compute the median age

Answer: (b) - If the dropped and kept customers had equal average profit, the slope would be near zero with a large p-value.

Q72. The bias-check regression gave:

(a) 9Profit = 72.96 + 52.22 x HasAge, p ~ 0: bias

(b) 9Profit = 72.96 + 5.88 x HasAge, p = 0.21: no bias

(c) 9Profit = 110.79 + 52.22 x HasAge

(d) 9Profit = 52.22 + 72.96 x HasAge

Answer: (a) - Customers with age data average $125.19; without, $72.96.

Q73. Because of that result, the $27.19 online slope from the complete-case model was:

(a) Correct

(b) Overestimated - the deleted loss-makers made the remaining sample rosier

(c) Underestimated

(d) Irrelevant

Answer: (b) - A rebate sized at $27 would overpay; the corrected figure is $19.65.

Q74. The professor's imputation rule is:

(a) Fill missing values with the mean only

(b) Fill with any constant, the same for every missing cell, and add an indicator that marks plugged values

(c) Fill with random values

(d) Delete the column

Answer: (b) - With the flag in the model, the choice of fill value does not affect the slopes or the predictions.

Q75. The final regression equation on all 31,634 customers is:

(a) 9Profit = 70.93 + 19.65 x 9Online + 25.60 x 9Age - 51.85 x HasAge

(b) 9Profit = 17.08 + 27.19 x 9Online + 25.86 x 9Age

(c) 9Profit = 70.93 + 25.60 x 9Online + 19.65 x 9Age - 51.85 x HasAge

(d) 9Profit = -32.66 + 19.65 x 9Online + 25.60 x 9Age

Answer: (a) - Online 19.65, age 25.60, HasAge -51.85 - check the assignments of the numbers.

Q76. Two customers, both without age data, one online and one offline. The online one is more profitable by:

(a) $25.60

(b) $19.65

(c) $51.85

(d) $70.93

Answer: (b) - Age and HasAge are zero for both; only the online term differs.

Q77. Two customers, both online, one in category 7 and one without age data. Who is more profitable, by how much?

(a) The no-age customer, by $51.85

(b) The category-7 customer, by about $127

(c) The category-7 customer, by $179.20

(d) Equal

Answer: (b) - 25.60 x 7 - 51.85 = 127.35.

Q78. Two customers, both online, one in category 1 and one without age data. Who is more profitable?

(a) The category-1 customer, by $25.60

(b) The customer without age data, by about $26

(c) Equal

(d) The category-1 customer, by $51.85

Answer: (b) - 25.60 x 1 - 51.85 = -26.25: the with-age customer predicts lower.

Q79. From which age category onwards is a customer with age data more profitable than one without (same channel)?

(a) Category 1

(b) Category 2

(c) Category 3

(d) Category 5

Answer: (c) - 25.60 x 3 = 76.80 exceeds 51.85; at category 2 the terms almost cancel (-0.65).

Q80. Two offline customers, categories 2 and 5. The difference in predicted profit is:

(a) $25.60

(b) $76.80

(c) $51.85

(d) $19.65

Answer: (b) - Three category steps x 25.60.

Q81. The predicted profit of an online customer in age category 4 is:

(a) $141.13

(b) $192.98

(c) $121.85

(d) $90.58

Answer: (a) - 70.93 + 19.65 + 102.40 - 51.85.

Q82. The predicted profit of an offline customer without age data is:

(a) $0

(b) $70.93

(c) $72.96

(d) $110.79

Answer: (b) - Only the intercept survives. ($72.96 is the bias-check intercept - a different model.)

Q83. HasAge = -51.85 means:

(a) Giving your age costs $52 of profit

(b) Read with the age term: a real category k is worth 25.60k - 51.85 relative to a plugged zero - the indicator re-anchors the age line

(c) Customers without age data are more profitable in every category

(d) The model failed

Answer: (b) - Negative for k = 1, about zero for k = 2, positive from k = 3.

Q84. If the blanks are filled with the average age instead of zero, which coefficients change?

(a) All of them

(b) Only the intercept and the HasAge coefficient; the online and age slopes and every prediction stay identical

(c) Only the online slope

(d) None

Answer: (b) - Intercept becomes -32.66 and HasAge +51.74 (the case slide's minus sign is a typo).

Q85. The professor's final recommendation for Pilgrim Bank is:

(a) Charge a fee

(b) Offer a rebate of up to about $19 per online customer

(c) Offer a rebate of up to $27

(d) Collect more data

Answer: (b) - Within comparable cohorts, online customers are $19.65 more profitable; the rebate is capped there.

Q86. The R Square of the final model is about:

(a) 0.005%

(b) 2.4%

(c) 24%

(d) 95%

Answer: (b) - Small, and enough for the question, which is the online effect rather than what drives profit.

Q87. Repeating the exercise with income (all 31,634 rows, income imputed with a flag), the online slope is:

(a) +19.65, significant

(b) -4.72 with p = 0.31 - not significant

(c) +27.19

(d) +11.87

Answer: (b) - Holding income alone constant, the data shows no significant online effect.

Q88. Holding both age and income constant (both imputed with flags), the online slope is:

(a) +19.65

(b) +11.87 with p = 0.011 - significant

(c) -4.72

(d) +52.22

Answer: (b) - Different sets of controls give different like-for-like comparisons.

Q89. In the model 9Profit = 46.36 + 16.55 x 9Online + 6.21 x 9Tenure, two customers with the same tenure, one online:

(a) The online customer is $6.21 more profitable

(b) The online customer is $16.55 more profitable

(c) The offline customer is $16.55 more profitable

(d) Equal

Answer: (b) - Same tenure cancels; the online term remains. (The case slide quotes $14.67 for its version.)

Q90. In the same model, two offline customers whose tenures differ by one year:

(a) Differ by $16.55

(b) Differ by $6.21, the longer-tenured one being higher

(c) Differ by $46.36

(d) Do not differ

Answer: (b) - The tenure slope is per year of tenure.

Q91. Three districts (1100, 1200, 1300) enter a regression as:

(a) One variable coded 1100, 1200, 1300

(b) Two 0/1 dummy variables, with one district as the base

(c) Three dummies

(d) They cannot be used

Answer: (b) - With 1200 as the base, the 1100 dummy is -21.19 and the 1300 dummy is -13.20 in the comprehensive model.

Q92. In the comprehensive model (online, age, income, tenure, district), the online slope is about:

(a) $5.88

(b) $13.82, p = 0.003

(c) $19.65

(d) $27.19

Answer: (b) - Like for like on everything measured; a careful bank would size the rebate to this.

Q93. How many customers are missing both age and income?

(a) 8,289

(b) 7,728

(c) 561

(d) 533

Answer: (b) - 561 are missing only age and 533 only income - largely the same casual customers behind both bias checks.

Q94. The case slides' lesson for service managers is to:

(a) Trust assumptions about customer behaviour built into the service model

(b) Surface implicit assumptions, convert them into testable hypotheses, test them, and update the model

(c) Avoid statistics

(d) Always offer rebates

Answer: (b) - Assumptions treated as facts hide important changes in behaviour.

Q95. With only cross-sectional data, the hypothesis "online adoption causes increased profitability" is modified to:

(a) "Online adoption will be discontinued"

(b) "Online customers are more profitable than offline customers"

(c) "Profit causes online adoption"

(d) It cannot be tested at all

Answer: (b) - A snapshot can show association like for like, not causation over time.

Q96. In the exam, the professor said she will:

(a) Ask you to run regressions in Excel

(b) Give you a regression output and ask interpretation questions - who is more profitable, by how much, who gets the rebate

(c) Ask for the derivation of the F-test

(d) Ask you to draw the Metro

Answer: (b) - No Excel in the exam; interpretation is the skill.

Q97. If one slope in a multiple regression has a p-value above alpha while Significance F passes, you should:

(a) Use the equation as it is

(b) Not use it as it stands - fix the model, for example by removing or replacing that variable

(c) Raise alpha

(d) Ignore R Square

Answer: (b) - The check is: Significance F, then every slope's p-value, and only then the equation.

Q98. Lowering alpha from 5% to 1% makes it:

(a) Easier to reject H0 and reduces Type II error

(b) Harder to reject H0, reducing Type I error but raising Type II error

(c) Impossible to reject H0

(d) Irrelevant to either error

Answer: (b) - A stricter tolerance for false alarms means more missed effects.

Q99. Hypothesis testing helps set the length of a product or service warranty by:

(a) Guessing

(b) Testing claims about product life on sample data so that the firm neither bleeds on replacements nor forgoes the confidence a warranty buys

(c) Copying competitors

(d) Using R Square

Answer: (b) - Urban Company's 30-day warranties and a manufacturer's six-month ones are the same problem.

Q100. "Holding the other variables constant" in interpreting a slope means:

(a) Setting them to zero

(b) Comparing two customers who are identical on every other variable in the model

(c) Removing them from the model

(d) Averaging them

Answer: (b) - That is why the online slope is read "within the same age category".

Q101. The professor's view on using ChatGPT to run the regressions was:

(a) Forbidden

(b) Fine - but you must be able to interpret the output, and the prompts come from understanding the business

(c) Required

(d) Only for the assignment

Answer: (b) - "The good part is GenAI can do it for you; you just have to prompt it."

Q102. Which statement is correct about the same data analysed with confidence intervals and with regression?

(a) They can give opposite verdicts

(b) If both techniques are valid, they must give the same result - a fancier name does not buy a different answer

(c) Regression is always more conclusive

(d) Confidence intervals are obsolete

Answer: (b) - Inconclusive by intervals, p = 0.21 by regression: the same verdict.

Q103. The case slides warn that managers often build beliefs about customer behaviour into service designs as:

(a) Testable hypotheses reviewed every year

(b) Factual claims left unchallenged for a long time, so important changes in behaviour are overlooked

(c) Regression coefficients

(d) Confidence intervals

Answer: (b) - Surface the assumption, convert it into a hypothesis, test it, and update the model.

Q104. The slides suggest surfacing an implicit assumption by:

(a) Reading the annual report

(b) Posing a question that depends on the assumption to elicit a strong opinion, then reverse-engineering the assumption behind the opinion

(c) Running a regression first

(d) Asking the regulator

Answer: (b) - "Should Pilgrim charge a fee or offer a rebate?" flushes out the belief about what online customers cost to serve.

Q105. In the managerial research design, the step after "refine the question into researchable hypotheses" is:

(a) Communicate results

(b) Identify the data requirements

(c) Update the service model

(d) Choose alpha

Answer: (b) - Problem, hypotheses, data requirements, analyses, interpretation and communication - five steps.

Q106. After a hypothesis test, the slides say you should:

(a) File the result

(b) Update your assumptions and, if required, the service model built on them

(c) Collect more data regardless

(d) Change alpha

Answer: (b) - The test exists to change what you believe and what you design.

PART V Exam Cheat Codes, and a Note to the Batch

Everything in the four parts before this one, compressed to what you must be able to reproduce cold: the definitions in one line, the formulas, the numbers of the two worked problems, every trap the professor set, and a template for each kind of question the exam can ask. Read it the evening before; read it again in the morning. Then close it, because the point of the module was never the exam.

By the end of this part you will be able to	Sections in this part
- Reproduce every definition and formula from memory - Reconstruct the Thompson Lumber and Pilgrim Bank numbers without the file - Spot the trap inside a tempting option - Answer any question type with the right template in the right order	5.1 The Cheat Sheet: Every Definition in One Line 5.2 Formulas and Numbers to Memorise 5.3 The Traps, All in One Place 5.4 Answer Templates for Every Kind of Question 5.5 The Quick Lookup: If the Question Says X, the Tool Is Y 5.6 A Note to the Batch

5.1 The Cheat Sheet: Every Definition in One Line

One line each, in the order the module taught them. If you can say every line without looking, you know the module.

Term	One line	Where
Course of action	The alternative you choose - fully under your control. Alternatives never carry probabilities.	1.2, 2.4
State of nature	What the world does after you choose - beyond your control. Only states carry probabilities.	1.2, 2.4
Certainty	One known outcome per action. Tool: compare / optimise (LP, IP).	1.3
Risk	Several outcomes per action with known probabilities. Tool: EMV, EOL, trees, EVPI. Insurable (Knight).	1.3
Uncertainty	Several outcomes per action, probabilities unknown. Tool: maximax, maximin, minimax regret, Laplace, Hurwicz. Must be managed (Knight).	1.3
Linear / integer programming	Optimisation under known constraints; integer when decisions are whole-number or yes/no (rosters, fixtures, SLA windows).	1.4
VRP / TSP	Allocate parcels to capacity-limited vehicles and sequence each route / one vehicle, every point once, back to the depot at least cost. TSP is a special case of VRP.	1.5
NP-hard	No known efficient exact algorithm; brute force is factorial; industry uses heuristics.	1.6
Censored (sales) data	Sales equal demand only until a stock-out; after it, sales undercount demand. Uncensor before you forecast.	1.7
Bass model	Adoption of a new category = innovators (p, media-driven) + imitators (q, word of mouth); the S-curve; needs demand data.	1.7
EOQ / safety stock	The order size that balances ordering and holding cost; the extra stock kept for the uncertainty the formula cannot see.	1.7

Term	One line	Where
Bullwhip / CPFR	Small consumer changes amplified up the chain by independent forecasts, batching, promotions and shortage gaming; fixed by one shared forecast on shared data.	1.7
Good decision	Based on logic, considers all available data and alternatives, obtained by rational analysis - judged by process, never by outcome.	2.2
Six steps	Define the problem; list all alternatives; identify states of nature; estimate payoffs; select the model; apply and decide. Spend most time on step 1.	2.4
Payoff table	Alternatives down the rows, states across the columns, the payoff in each cell.	2.6
EMV	Probability-weighted average of an alternative's payoffs; the long-run average per instance; pick the best (largest for profit, smallest for cost).	2.6
EOL	Expected opportunity loss (regret); minimising it picks the same alternative as maximising EMV; the best alternative's EOL equals EVPI.	2.6
Sensitivity analysis	How far the inputs can move before the best choice changes. Robust: cushion is wide. Sensitive: prepare a contingency.	2.7
EVPI	Expected value WITH perfect information minus best EMV without it; the most you should pay for certainty; information is worth only what it changes.	2.8
Decision tree	Squares = your choices; circles = nature's branches with probabilities; drawn left to right in real-life sequence, solved right to left; it prescribes a strategy.	2.9
EVSI	Expected value with the survey (as if free) minus EMV without it; never exceeds EVPI; EVSI / EVPI is the survey's efficiency.	2.10
Bayes' theorem	Prior x how often the evidence appears in each state = joint; divide by the total = posterior. Turns a vendor's accuracy into the tree's probabilities.	2.10
Profitability skew	Customers sorted by profit; cumulative % of profit vs cumulative % of customers; peaks above 100% when there is a loss-making tail.	3.2
Pareto principle	A small share of causes or elements accounts for most of the effect (customers, defects, SKUs).	3.4
Population / sample	All the elements of interest to the study / the subset you actually observe. Every decision is about the population.	3.6
Representative sample	Chosen by a method that gives every element an equal chance - not "the same demographic mix".	3.7
Large sample	n >= 30, an absolute number, from the central limit theorem: sample means become bell-shaped.	3.7
Point estimate	A single best guess for an unknown population parameter (e.g. the sample mean).	3.9
Parameter vs statistic	A parameter is a fixed unknown constant of the population; a statistic is a random variable that depends on who landed in the sample.	3.9
Standard deviation	Square root of the variance (sum of squared deviations / n or n - 1); same unit as the data; large only relative to the mean; never report a mean without it.	3.9
Normal distribution	Bell curve: peak at the mean (= mode); 68% within 1 SD, 95% within 2, 99.7% within 3.	3.9

Term	One line	Where
Confidence interval	Point estimate +/- cushion; cushion = z x SD / square root of n; z = 1.645 / 1.96 / 2.576 for 90 / 95 / 99%. Overlap = inconclusive. Too wide if your decision would differ at the two limits.	3.11
Null / alternate hypothesis	Two statements about the population, mutually exclusive and collectively exhaustive; what you want to establish goes in the alternate; equality sits in the null.	4.2, 4.3
Reject / do not reject	The only verdicts. Reject H0 = accept H1 (conclusive). Do not reject = inconclusive; you never accept H0.	4.2
Type I / Type II error	Rejecting a true null / not rejecting a false null. Alpha is the tolerance for Type I; a bigger sample shrinks Type II.	4.4
p-value	The probability of a sample like yours if the null were true. Below alpha, reject.	4.5
Regression	Best-fit line y = intercept + slope x x (simple) or with several x's (multiple); the line close to all points, not through the most.	4.6
Reading order	Significance F (multiple regression), then each slope's p-value, then R Square, then the coefficients.	4.7, 4.8
R Square	Coefficient of determination: share of the variation in y explained by the x's; judged against the question, not a fixed bar.	4.8
Dummy variable	A 0/1 variable; its slope is the difference between the 1 group and the 0 group, holding the rest constant.	4.7, 4.9
Confounding / clouding	Groups with different mixes (age) make raw averages lie; put the mix variable in the model.	3.12, 4.8
Imputation with a flag	Fill missing values with any constant, the same everywhere, plus a 1/0 indicator in the model; never drop rows without testing for bias first.	4.9, 4.10

5.2 Formulas and Numbers to Memorise

The formulas first, then the two worked problems the exam will draw on. Every number below is worked in full in the section shown.

Formula	Statement	Where
EMV	EMV(alternative) = sum over states of (payoff x probability)	2.6
EOL	Regret = best payoff in that state - your payoff; EOL = sum of (regret x probability)	2.6
Sensitivity thresholds	Write each EMV as a function of P; set two EMVs equal; solve for P	2.7
EVwPI	Sum over states of (best payoff in that state x probability of the state)	2.8
EVPI	EVwPI - best EMV without information (= EOL of the best alternative)	2.8
EVSI	(EMV with the survey + survey cost) - best EMV without the survey; net gain = EVSI - cost	2.10
Bayes	P(state | evidence) = P(evidence | state) x P(state) / P(evidence); P(evidence) = sum of the joints	2.10
Mean	Sum of values / n	3.9

Formula	Statement	Where
Variance / SD	Variance = sum of (value - mean)^2 / (n - 1) for a sample (n for a population); SD = square root	3.9
Empirical rule	mean +/- 1 SD holds ~68%; +/- 2 SD ~95%; +/- 3 SD ~99.7%	3.9
Confidence interval	x-bar +/- z x sigma / square root of n; z = 1.645 (90%), 1.96 (95%), 2.576 (99%); Excel: CONFIDENCE.NORM(0.05, SD, n)	3.11
Cushion and n	The cushion falls with the square root of n: four times the data halves it	3.11
Decision rule	Reject H0 if p-value < alpha (alpha = 0.10, 0.05, 0.01 by cost of error)	4.5
Regression line	y = intercept + slope1 x x1 + slope2 x x2 + ...; predict by plugging in; compare two cases by striking common terms	4.6-4.10

Thompson Lumber (Part II) - the numbers

Item	Value
Payoffs (favourable / unfavourable)	Large plant 200,000 / -180,000; small plant 100,000 / -20,000; no plant 0 / 0. Prior 0.5 / 0.5.
EMV	Large 10,000; small 40,000; none 0 -> build the small plant.
EOL	Large 90,000; small 60,000; none 100,000 -> small plant again; 60,000 = EVPI.
Sensitivity	EMVs: large 380,000P - 180,000; small 120,000P - 20,000; none 0. Thresholds P = 0.167 and P = 0.615. At 0.5 the small plant is robust (cushions 33.3 and 11.5 points).
Perfect information	EVwPI = 0.5 x 200,000 + 0.5 x 0 = 100,000; EVPI = 100,000 - 40,000 = 60,000. Decline the Rs. 65,000 consultant.
Survey	Cost 10,000; P(says favourable) = 0.45; P(FM | fav) = 0.78; P(FM | unfav) = 0.27. Payoffs on the survey side reduced by 10,000.
Survey tree	Node 2 (large | fav) 106,400; node 3 (small | fav) 63,600; node 4 (large | unfav) -87,400; node 5 (small | unfav) 2,400; node 1 (survey) = 0.45 x 106,400 + 0.55 x 2,400 = 49,200 vs 40,000 without.
Strategy	Survey; large plant if favourable; small plant if unfavourable.
EVSI	(49,200 + 10,000) - 40,000 = 19,200; net gain 9,200; efficiency 19,200 / 60,000 = 32%.
Bayes (not in the exam, but know the mechanics)	P(pos | FM) = 0.70, P(pos | UM) = 0.20; joints 0.35 and 0.10 -> 0.45; posteriors 0.78 / 0.22. Negative: 0.15 and 0.40 -> 0.55; posteriors 0.27 / 0.73.

The samosa stall and the plant under uncertainty (Part I) - the numbers

Item	Value
Samosa EMV	Fry 200: 0.7 x 1,000 + 0.3 x (-200) = 640. Fry 100: 0.7 x 500 + 0.3 x 300 = 440. Fry 200.
Plant under uncertainty (Rs. lakh)	Large +100 / -40; small +40 / +10. Maximax: large. Maximin: small. Minimax regret: large (worst regret 50 vs 60). Laplace: large (30 vs 25). Hurwicz: alpha x best + (1 - alpha) x worst.
Factorial wall	10 stops = 3.6 million sequences; 15 = 1.3 trillion (22 minutes); 20 = 2.4 x 10^18 (77 years); 25 = 1.55 x 10^25 (490 million years) at a billion routes a second.

Pilgrim Bank (Parts I, III and IV) - the numbers

Item	Value
The file	31,634 customers (a 0.63% sample of about 5 million); 1999 data (prefix 9); profit -221 to +2,071; 3,854 online (12.2%), 27,780 offline; 8,289 missing age, 8,261 missing income, 7,728 missing both; tenure and district complete.
Profit	Total 3,527,276; mean 111.50; median 9; SD 272.84 (STDEV.S). Profitable 16,832 (53%), loss-making 14,580 (46%), zero 222. Profitable customers earned 4,288,303; loss-makers lost 761,027.
Skew	Top 1% -> 14% of profit; top 10% -> 69%; top 20% -> 97% (and 80% of the profitable customers' profit); peak 121.6% at 53% of customers; ends at 100%.
Groups	Offline: n 27,780, mean 110.79, SD 271.30. Online: n 3,854, mean 116.67, SD 283.66. Difference 5.88 (5.3%).
95% intervals	Offline 107.60 to 113.98 (width 6.4); online 107.71 to 125.62 (width 17.9). Overlap 107.7 to 114.0 -> inconclusive at 90%, 95% and 99%. All customers: 108.50 to 114.51.
Mean +/- 1 SD	-161 to +384 (all); the rule-of-thumb sign that the SD is too high to trust the mean alone.
SLR	9Profit = 110.79 + 5.88 x 9Online; p(9Online) = 0.21 -> do not use; R Square 0.00005; 95% CI of the slope -3.31 to 15.07 (contains zero).
MLR, rows with age (n 23,345)	9Profit = 17.08 + 27.19 x 9Online + 25.86 x 9Age; Significance F 1.7E-114; p-values 8.6E-07 and 1.1E-115; R Square 2.2%.
Bias check	9Profit = 72.96 + 52.22 x HasAge; p 7.7E-51 -> customers without age data are $52 less profitable; dropping them biases the sample.
Final model (n 31,634)	9Profit = 70.93 + 19.65 x 9Online + 25.60 x 9Age - 51.85 x HasAge (age = 0 and HasAge = 0 where missing); Significance F 4.8E-168; all p-values ~0; R Square 2.4%. Mean-imputed version: intercept -32.66, HasAge +51.74, same slopes.
Verdict	Like for like, online customers are $19.65 more profitable -> rebate, up to about $19 (not the biased $27).
Assignment models	Income only: online -4.72, p 0.31 (not significant). Age + income: online +11.87, p 0.011. Tenure: 46.36 + 16.55 x online + 6.21 x tenure. Comprehensive (with district, base 1200): online +13.82, p 0.003; 1100 -21.19; 1300 -13.20; R Square 6.2%.
Age pattern	Share online falls from ~20% (youngest) to 3.6% (oldest); average profit rises from $3.5 to $192. Average online customer is in age category 3.3, offline 4.2.

5.3 The Traps, All in One Place

Every concept trap from the four parts, in one table. In the exam, the wrong option that looks most right is usually one of these.

The tempting error	The correction	Where
Judging a decision by its outcome	Under risk, chance sits between decision and outcome; grade the process with the information available at the time.	1.2, 2.2
Giving an alternative a probability	You choose alternatives (1 or 0); only states of nature carry probabilities.	2.4
"NP" = non-polynomial	Non-deterministic polynomial; the substance is that exact enumeration explodes.	1.6

The tempting error	The correction	Where
"The small plant never earns 40,000, so EMV is meaningless"	EMV is the long-run average per instance; choose by it consistently across many decisions.	2.6
Weighting the informant's messages 50-50 "because there are two"	They carry your own prior; a 60-40 belief gives a 60-40 informant.	2.8
Drawing a decision square straight after "conduct survey"	A circle comes first: you read the survey before you decide.	2.10
EVSI = 9,200	That is the net gain; EVSI = 19,200 is the worth of the survey if it were free.	2.10
Plugging a vendor's 70% accuracy straight into the tree	Combine it with the prior through Bayes; a 90-10 product and a 50-50 product cannot share survey probabilities.	2.10
"How can profit be 121%?"	The cumulative curve adds the profitable customers first; the loss-makers pull it back to 100%.	3.2
Representative = same demographic mix	Representative = every element had an equal chance; balance a trait only when it changes the answer.	3.7
"Large sample" = 30% of the population	n >= 30, an absolute number, from the central limit theorem.	3.7
Reporting an average alone	Always with its SD; check mean +/- 1 SD against the decision you want to make.	3.9
"The sample says online wins, so online wins"	A difference in sample means is evidence, not proof; check the spread and the sample sizes.	3.10
"Online is more profitable - its interval goes up to 125"	The truth can be anywhere in an interval; overlapping intervals are inconclusive.	3.12
"Collect more data" as the reflex	Expensive and slow; look at the same data differently first (Part IV did).	3.12, 4.8
"We did not reject the null, so the null is true"	Not rejecting is inconclusive; you only ever accept the alternate.	4.2
Putting the action ("charge a fee") in the hypothesis	Hypotheses are about what the data shows (who is more profitable).	4.3
"A slope of 1 means perfect correlation"	Slope depends on units; the correlation coefficient lives between -1 and +1.	4.7
Reading the coefficient before the p-value	Order: Significance F, p-values, R Square, then coefficients.	4.7, 4.8
Reading 8.59E-07 as 8.59	It is 0.000000859 - far below 0.05.	4.8
Comparing raw averages of groups with different mixes	Put the mix variable in the model; read the effect holding it constant.	4.8
"The regression ran, so the sample is fine"	Test whether the dropped rows differ on the outcome; if they do, impute with a flag.	4.9
Reading HasAge = -51.85 on its own	Read it with the age term: 25.60 x category - 51.85.	4.10
The case slide's mean-imputation sign	AgeExist is +51.74 with mean imputation, not -51.73; the slopes and predictions do not change.	4.10

The tempting error	The correction	Where
Sizing the rebate to $27	That estimate came from a biased sample; the honest figure is $19.65 (about $14 with everything controlled).	4.11

5.4 Answer Templates for Every Kind of Question

The exam rewards order. For each kind of question, do the steps in this order and write them down in this order; the marks are in the steps as much as in the final number.

If the question asks you to...	Do this, in this order	Where
Classify a decision	1. Objective. 2. Courses of action. 3. States of nature and what is known about them. One outcome per action = certainty (optimise); known probabilities = risk (EMV / tree); unknown = uncertainty (criterion + temperament).	1.3
Choose under uncertainty	Build the payoff table; apply maximax, maximin, minimax regret, Laplace (and Hurwicz if an alpha is given); state that they may disagree and which temperament each represents.	1.3
Choose under risk	Payoff table; EMV per alternative = sum of payoff x probability; pick the best (largest for profit, smallest for cost); optionally confirm with EOL.	2.6
Do a sensitivity analysis	Write each EMV as a function of P; find where the lines cross; state the range for each alternative; say whether the estimate sits comfortably inside (robust) or near an edge (sensitive -> contingency).	2.7
Value perfect information	EVwPI = best payoff per state x its probability, summed; EVPI = EVwPI - best EMV; compare with the price offered; say whether the information would change the action.	2.8
Draw and solve a tree	Left to right in real-life order: square (choice), circle (nature), payoffs; every payoff on a survey branch net of the survey cost; solve right to left: average at circles, choose at squares, slash the rejected; read off the strategy.	2.9, 2.10
Value a survey	EMV with survey (from the tree) + cost - EMV without = EVSI; net gain = EVSI - cost; check EVSI <= EVPI; efficiency = EVSI / EVPI.	2.10
Revise probabilities (Bayes)	Prior x P(evidence | state) = joint; sum the joints = P(evidence); joint / P(evidence) = posterior; repeat for the other evidence.	2.10
Comment on a sample	Population vs sample; method of selection (random, stratified, convenience) and bias; n >= 30; larger = more precise, not more legitimate.	3.6, 3.7
Summarise a data set	Mean and SD together; mean +/- 1 SD; median if skewed; say whether the SD is large relative to the mean.	3.9
Build a confidence interval	x-bar, SD, n; z for the level; cushion = z x SD / square root of n; state lower and upper; interpret as "with X% confidence the population mean lies between".	3.11
Compare two groups with intervals	Build both; if they overlap, inconclusive (both orderings possible); if not, conclude; sample sizes need not match.	3.12
Frame a hypothesis	What do you want to establish? That is H1; its negation with the equality sign is H0. State what you will do on "reject" and on "do not reject".	4.3
Name the errors	Type I = reject a true null; Type II = keep a false null; alpha controls Type I; a bigger sample shrinks Type II; say which the setting is tuned to avoid.	4.4

If the question asks you to...	Do this, in this order	Where
Use a p-value	State the null being tested (slope = 0, or means equal); p < alpha -> reject and act; p >= alpha -> inconclusive, do not act.	4.5, 4.7
Read a regression output	Significance F (multiple regression) -> each slope's p-value -> R Square in context -> write the equation -> interpret each slope holding the rest constant.	4.7-4.12
Compare two customers from an equation	Write each one's equation; strike out identical terms; the difference is what remains; state who is higher and by how much, with units.	4.8, 4.10, 4.12
Handle missing data	Do not drop silently; regress the outcome on a 1/0 "has data" dummy; if significant, impute with a constant plus the flag and re-run on all rows.	4.9, 4.10
Recommend for Pilgrim Bank	Means -> intervals overlap -> SLR p = 0.21 -> MLR with age -> bias check -> impute -> online +19.65, like for like -> rebate up to about $19.	4.11

5.5 The Quick Lookup: If the Question Says X, the Tool Is Y

A last-minute index from the words a question uses to the tool it wants.

If the question mentions...	It wants...	Where
"one outcome per action", FD rates, quotations, a signed contract book	Certainty: compare or optimise; linear / integer programming	1.3, 1.4
probabilities from history, "3 days in 10", success rates	Risk: EMV (or EOL), decision tree, EVPI	1.3, 2.6
"no history", a first-ever product, a pandemic, a zero-day	Uncertainty: maximax / maximin / minimax regret / Laplace / Hurwicz; Knight	1.3
routes, parcels, depots, sequencing, rosters, fixtures	VRP / TSP; integer programming; NP-hard; heuristics	1.4-1.6
stock-outs, "sales were", a new category with no history, inventory, a chain of orders	Censored data and uncensoring; the Bass model; EOQ and safety stock; the bullwhip and CPFR	1.7
"was it a good decision?", hindsight, an unlucky result	Process over outcome; Alcoa; the three tests	2.2
"the problem keeps coming back", firefighting, a pet solution	Fall in love with the problem; symptom vs root cause; design thinking; the calm test	2.3
"gut feeling", "my boss was right", a chart of failures	Gut = sample inference; two checks; Challenger - show the successes	2.5
"what if the probability were different?"	Sensitivity analysis: thresholds, robust vs sensitive, contingency	2.7
"a consultant offers certainty for Rs. X"	EVPI; worth only if it changes the action; cap = EVPI	2.8
a survey, a test market, a pilot, "read the result then decide"	Decision tree with a survey stage; EVSI; Bayes for the probabilities	2.9, 2.10
"top 10% of customers", a cumulative curve, "let go of customers"	Profitability skew; Pareto; one rule with a threshold (fee with waiver, tiers)	3.2-3.4

If the question mentions...	It wants...	Where
"is 31,634 enough?", "how big a sample?", "who responded?"	Population vs sample; n >= 30; representative = equal chance; convenience bias	3.6, 3.7
"the average is 111", "is that reliable?"	SD; mean +/- 1 SD; the bell curve; parameter vs statistic	3.9
"give a range", "how confident?", "90% / 95% / 99%"	Confidence interval; z; cushion drivers; overlap rule	3.11, 3.12
"is the baby hers?", a claim to prove, a claim to challenge	H0 / H1 framing; reject / do not reject; Type I / II; alpha	4.2-4.4
"p = 0.21", "significant?", "Significance F"	p-value rule; F-test of all slopes; reading order	4.5, 4.7, 4.8
"intercept", "slope", "coefficient", "R Square", "two customers"	Regression equation; plug in; strike common terms; R Square in context	4.6-4.12
"8,289 blanks", "drop the rows?", "impute"	Bias check with a dummy; imputation with a flag; the corrected online effect	4.9-4.11
"older customers are offline", "the averages reverse"	Clouding / confounding; hold age constant in the model	3.12, 4.8

5.6 A Note to the Batch

A few words before you close this, from one batchmate to the rest of EPBM-32.

Four sessions ago most of us thought this module would be about formulas. It turned out to be about honesty: knowing what we know before we choose, grading the process and not the luck, never letting an average travel without its standard deviation, writing the hypothesis the right way round, and passing up the first significant number for the one that survives the bias check. Those are not exam skills. They are the habits of managers who are right often rather than right loudly.

The samosa-wala, Thompson Lumber, the bakery in Calcutta, the baby in gold bangles, the broken helicopter and 31,634 bank customers were all the same lesson in different clothes: decide with every option and all the data you can honestly get, value information before you buy it, know which mistake you would rather make - and then decide, because not deciding is rarely one of the options.

Thank you, EPBM-32 - for the questions that made the sessions better, for the shared screens when Excel refused to cooperate, and for the company. All the very best for the exam, and for the far more interesting decisions that come after it. Decide well, and may your states of nature be kind.

> *"Solving the right problem roughly is way better than solving the wrong problem accurately."*
>
> **Prof. Megha Sharma - the line to keep, from all four sessions**

Abhinav Roy | EPBM-32 | IIM Calcutta **https://www.linkedin.com/in/abhinavroy07**