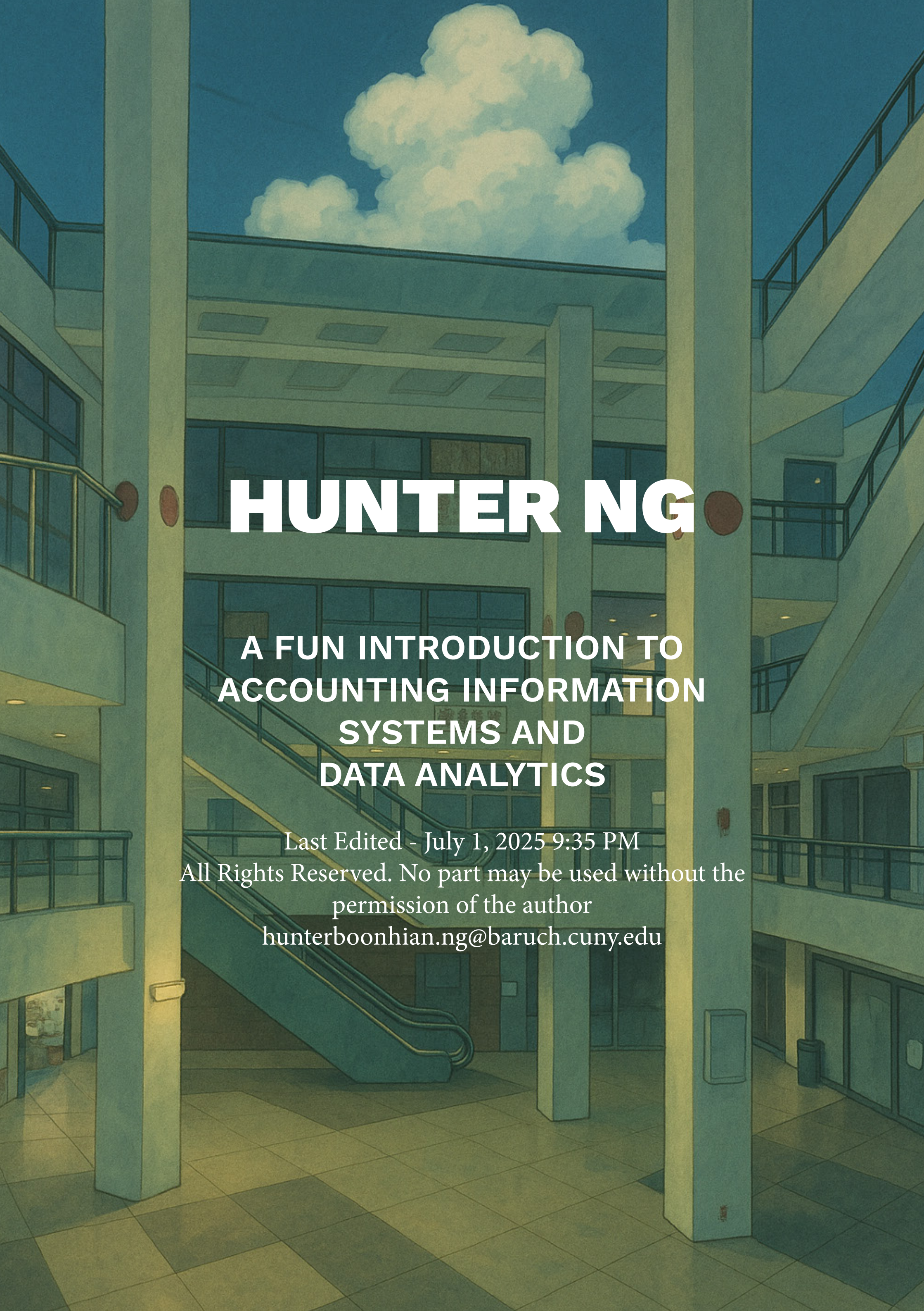


HUNTER NG

A FUN INTRODUCTION TO ACCOUNTING INFORMATION SYSTEMS AND DATA ANALYTICS

Last Edited - July 1, 2025 9:35 PM

hunterboonhian.ng@baruch.cuny.edu

PREFACE

🐇 Hunter Ng. "Expression seeks to impress; perception seeks to understand." For all learners of AIS. hunterng.com

🍩 🍩 🍩

—Chapter 1 -

🍩 🍩 🍩

The Italian Friar 👨

History of Accounting Information Systems

✅ The story of Accounting Information Systems (AIS) is really the story of how we've kept track of money through the ages, and how that process has evolved alongside human progress.

✅ Thousands of years ago, long before computers, people in places like Mesopotamia, Egypt, and Rome were already recording financial information. They didn't have Excel, of course: they used clay tablets, scrolls, and papyrus to log trade, taxes, and government spending. It wasn't fancy, but it got the job done.

✅ Our story starts in medieval Europe, where friars were among the most educated members of society, and for good reason. At the time, the Catholic Church 💒 was not only a spiritual authority but also the main institution responsible for preserving and passing on knowledge. Monasteries and cathedral schools served as early centers of learning, where reading, writing, philosophy, and mathematics were taught. Latin, the language of the Church, was the language of scholarship, and anyone aspiring to understand scripture needed to learn it.

✅ Friars, unlike monks who often lived in isolated communities, were part of active religious orders such as the Franciscans and Dominicans. These friars lived and worked among the public, and their roles often included preaching, teaching, advising, and managing local affairs. Because of this, they required more than theological training; they also needed practical knowledge in arithmetic, logic, and even science.

✅ Luca Pacioli 👨, a Franciscan friar, stands out as a remarkable figure who bridged the worlds of faith and mathematics during the Renaissance. Far from being solely a religious man, Pacioli was a deeply educated scholar whose training empowered him to teach advanced subjects such as algebra and practical bookkeeping to merchants and students across Italy. His most influential contribution came in 1494, when he published Summa de Arithmetica, a comprehensive work that included the first formal description of double-entry bookkeeping. This was a revolutionary system where every transaction is recorded with both a debit and a credit.

✅ Pacioli didn't invent this method, but his ability to organize and explain it clearly gave the system structure, legitimacy, and reach. Thanks to his writing, a practice once confined to the ledgers of Italian merchants became the gold standard of accounting, shaping how businesses around the world manage their finances even today.

💪 Exercise - Single-entry vs double-entry

Meet Hunter Ng, the proud owner of a small surf shop in Maui. He just opened his store, Hunter’s Surf Shack, which sells surfboards, gear, and sunscreen, and also offers surfing lessons to tourists. To keep things simple, Hunter uses a single-entry bookkeeping system. He just records the money coming in and out of his cash register in a notebook.

Below is a summary of Hunter’s cash activity during his first week of business:

- On January 1, Hunter invests $5,000 of his own savings into the business.
- On January 2, he buys surfboards from a wholesaler for $3,000.
- On January 3, he sells surf gear to tourists for $800.
- On January 5, he pays $500 in rent for his small retail space.
- On January 6, he buys suncreen from a supplier for $300
- On January 7, he sells $600 of lessons to tourists for surfboard practice.

Hunter tracks only the cash in and out. He doesn’t write down what he owns (like inventory), what he owes (if any), or his profit. He just looks at how much money is left.

❓ Questions

Help Hunter reflect on whether this system is enough for running a successful business.

1. Draw out the single-entry accounting record yourself. What is Hunter's cash balance at the end of the week?
2. Can you calculate his profit from these transactions? What challenges do you face in doing so?
3. Can you tell how many surfboards or how much sunscreen Hunter still has?
4. Do you know if Hunter owes money to suppliers or if anyone owes him?
5. What would happen if Hunter forgot to write down a sale or an expense?

🙋‍♀️ Answers:

Date	Description	Cash In	Cash Out
1-Jan	Owner’s investment	$5,000	
2-Jan	Bought surfboards		$3,000
3-Jan	Sold surf gear	$800	
5-Jan	Paid rent		$500
6-Jan	Bought sunscree		$300
7-Jan	Sold lessons	$600	

Hunter’s cash balance is $2,600 at the end of the week.

Q2 You can’t accurately calculate profit because costs like inventory are not matched to sales.

Q3 No, you can’t tell how many surfboards or how much sunscreen he has left since inventory isn’t tracked.

Q4 No, you don’t know if Hunter owes money or is owed money without a record of receivables or payables.

Q5 If he forgets a sale or expense, his records will be inaccurate, potentially leading to poor decisions.

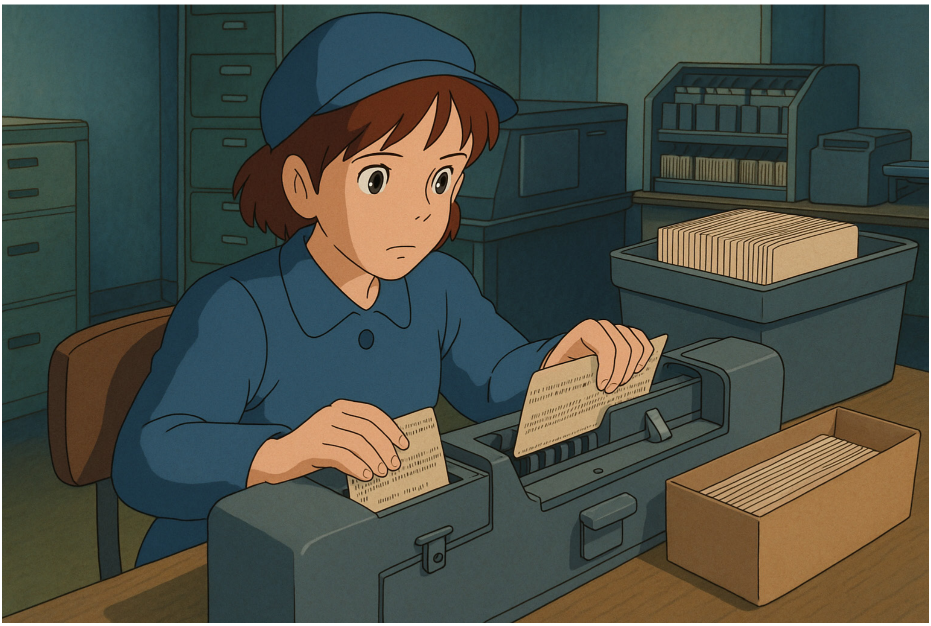
🍩 🍩 🍩

✅ For hundreds of years after that, accounting stayed mostly manual. Businesses used paper ledgers, pens, and a whole lot of patience. Every entry had to be written by hand, and mistakes were easy to make. But then came the early 20th century, and with it, a wave of new technology. Suddenly, offices had mechanical calculators and punch card machines to help with tasks like payroll and inventory. These tools might look ancient now, but at the time, they were a big leap forward, helping reduce errors and speed things up.

✅ The real transformation of accounting kicked off between the 1950s and 1970s, when mainframe computers entered the scene. These massive machines weren't cheap or easy to use, but they allowed big companies and government agencies to process financial data much faster. Tasks like payroll and keeping the general ledger became more efficient and reliable. Still, these early systems weren't perfect as they were expensive, highly customized, and worked in batches, meaning you couldn't see updates in real time.

✅ Then came the personal computer boom 💻 of the 1980s and 1990s, and everything changed. Suddenly, accounting technology wasn't just for big corporations. It was for everyone. Small businesses could now use user-friendly software like QuickBooks or Peachtree to manage their books without needing an IT department. Meanwhile, programs like Microsoft Excel became essential tools for budgeting, forecasting, and number-crunching.

✅ Larger companies took things even further with Enterprise Resource Planning (ERP) systems like SAP and Oracle



Financials. These connected accounting to other parts of the business, like supply chains, HR, and customer management, giving a fuller picture of company performance.

✅ By the 2000s, cloud computing took over. With the rise of high-speed internet, platforms like Xero, NetSuite, and Zoho Books made it possible to check your company's finances from anywhere in the world. These systems didn't just digitize accounting, they made it real-time, collaborative, and smarter. Companies no longer had to install expensive software, they could subscribe online and access updates automatically, thanks to the Software-as-a-Service (SaaS) model.

✅ In the past decade, we've seen AIS 💻 enter an entirely new era. Technologies like artificial intelligence (AI), machine learning, and blockchain are now changing the game. AI can handle repetitive tasks like sorting invoices or spotting unusual transactions, while machine learning can help predict future cash flow or detect fraud. Blockchain, with its secure and transparent recordkeeping, could one day revolutionize audits and compliance.

✅ And it doesn't stop there. Modern AIS platforms now tap into big data analytics to deliver powerful insights, helping businesses make smarter decisions. Accountants aren't just recordkeepers anymore, they're becoming strategic advisors, guiding organizations with real-time financial intelligence.

✅ What started thousands of years ago with handwritten records on clay tablets has turned into a high-tech, cloud-powered system 💻 that helps organizations stay agile, compliant, and competitive in today's fast-paced world.

—Chapter 2 -
🍩 🍩 🍩

The Fraudulent and The Auditor 🦹

How to succeed in AIS as a career

✅ In 2020, Kubient, a digital advertising startup, told investors it had built a powerful AI tool called KAI that could detect and block ad fraud in real time. This technology, they claimed, had been rigorously tested and was already generating serious revenue. In fact, the company reported $1.3 million from a single KAI-related deal, representing nearly all of its quarterly revenue just before going public. But there was one major problem: the KAI tests never happened.

✅ The company's CEO orchestrated a scheme to fabricate internal test results, directing staff to produce fake reports and passing them off to auditors as genuine. The CFO and audit committee chair eventually found ou. Instead of correcting the record, they kept quiet. The company proceeded with its IPO and a follow-on stock offering, raising over $30 million from investors based on this false narrative.

✅ Eventually, the SEC and Department of Justice uncovered the truth. The CEO pled guilty to fraud and was sentenced to prison. The CFO and audit chair were charged with misleading auditors and investors.

✅ This case reveals exactly why AIS 💻 is critical to modern organizations. It's not just about bookkeeping or databases: it's about how financial data flows through systems, how internal controls catch (or fail to catch) manipulation, and how decision-makers rely on these systems to tell the truth. At Kubient, flawed oversight and


$1,3M DEAL!
KAI
REAL-TIME AD FRAUD DETECTION
IPO LAUNCH
This isn't right...
KAI..Test..Results ..FINAL..v3
SEC
DO.
SEC
IPO

willful misuse of AIS led to misleading financial statements, regulatory violations, and the erosion of investor trust.

Performing the digital trust audit that uncovered the above case is just one of the possibilities of learning AIS.

💪 Exercise - What is AIS 💻?

1. Which of the following best describes the scope of Accounting Information Systems (AIS)?
A. A tool used only for recording financial transactions
B. A database for storing accounting records
C. A system that combines accounting, technology, and data to support business decisions
D. A basic spreadsheet for calculating profits

2. Which of the following roles might require knowledge of AIS?
A. Financial Analyst
B. ERP Consultant
C. IT Auditor
D. All of the above

3. Which of the following certifications are relevant for professionals pursuing careers in Accounting Information Systems (AIS)?
A. Certified Information Systems Auditor (CISA)
B. Certified Information Technology Professional (CITP)
C. Certified Internal Auditor (CIA)
D. All of the above

4. Which of the following are the two key attributes that make information from an Accounting Information System (AIS) useful to decision makers?
A. Speed and accessibility
B. Relevance and faithful representation
C. Cost-effectiveness and neutrality
D. Completeness and materiality

5. Which document issued by the FASB formally defines the qualitative characteristics of useful financial information, including relevance and faithful representation?
A. GAAP Implementation Guide
B. Statement of Financial Accounting Concepts No. 8
C. Sarbanes-Oxley Act
D. ASC 606 Revenue Recognition Standard

6. Which of the following is considered to be mandatory information required by a regulatory body?
A. Internal management reports for budgeting
B. Sales forecasts used by marketing teams
C. Quarterly financial statements filed with the SEC
D. Trend analyses prepared for executive decision-making

7. In a full-featured ERP system like SAP, which of the following modules would be most directly responsible for handling accounts payable and receivable?
A. Materials Management (MM)
B. Financial Accounting (FI)
C. Sales and Distribution (SD)
D. Human Capital Management (HCM)

8. Which type of software is typically used within AIS to monitor stock levels, track orders, and manage supply chains?
A. CRM (Customer Relationship Management)
B. HCM (Human Capital Management)
C. Inventory Management Systems
D. GRC (Governance, Risk, and Compliance)

🙋 Answers:

1. C - AIS 💻 combines accounting, technology, and data to support decision-making. It's more than just recording transactions.

2. D - Financial analysts, ERP consultants, and IT auditors

all use AIS in different ways. AIS is relevant across finance, tech, and audit roles.

3. D - CISA, CITP, and CIA are all AIS-related certifications. Each reflects a different career path within the AIS field.

4. B - Relevance and faithful representation are the core qualities of useful financial information. They ensure data is meaningful and accurate.

5. B - SFAC No. 8 defines the qualitative characteristics of financial information. It is the key part of FASB's conceptual framework.

6. C - Quarterly filings with the SEC are legally required. They provide transparency to investors and regulators.

7. B - SAP's Financial Accounting (FI) module handles payables and receivables. It's essential for core accounting functions.

8. C - Inventory systems track stock, orders, and supply chains. They support AIS by ensuring accurate cost and inventory records.

🍩🍩🍩

✅ Accounting Information Systems (AIS) 💻 focuses on how financial data is collected, processed, and used in business systems. It combines knowledge of accounting, technology, and process design. AIS professionals are trained not just to use accounting software, but to understand how financial workflows operate and how to improve them.

✅ A key strength of AIS is business process modeling. This involves mapping out how core functions, like purchasing, billing, payroll, or reporting, work in practice. AIS professionals identify where data enters the system, how it flows between departments, and where automation or controls are needed. This skill is critical in modern organizations, especially when configuring enterprise software like SAP or Oracle.

✅ Compared to general software engineers, AIS 💻 professionals bring institutional knowledge of accounting principles, compliance requirements, and internal controls. This allows them to build or adapt systems that align with business needs. For example, they understand how to structure approval hierarchies, maintain audit trails, and comply with standards like GAAP or SOX.

✅ This knowledge opens up a range of roles, such as ERP implementation analyst, business process consultant, financial systems engineer, or IT auditor. These roles are growing rapidly as companies digitize their finance functions and integrate cloud-based platforms. AIS professionals may also specialize in automation, using tools like UiPath or Power Automate to streamline repetitive accounting tasks.

To succeed in AIS, students typically learn process modeling, data analysis (using tools like SQL, Excel, and Power BI), and how accounting concepts apply within financial software. This gives them a clear advantage when working at the intersection of finance, systems, and decision-making.

✅ AIS professionals are increasingly involved in designing and implementing AI tools that handle core accounting tasks. These include classifying invoices, detecting fraud, forecasting cash flows, and generating journal entries based on past data. Unlike generic AI systems, AIS experts ensure these tools are accurate, explainable, and compliant with accounting rules.

✅ As companies adopt AI in finance, they need people who can manage, adjust, and integrate these systems into real workflows. AIS professionals are well-positioned for this because they understand both the financial requirements and how the technology works. They work closely with data scientists and finance teams to make sure the tools meet business needs without creating new risks.

✅ This skillset makes AIS careers more stable in a changing job market. Rather than being replaced by automation, AIS experts are the ones leading its implementation. They set up ERP systems, automate reporting, create dashboards, and apply AI in ways that match how companies actually operate. These are tasks that require domain expertise and are not easily outsourced or automated.

—Chapter 3 -

🍩 🍩 🍩

Factories and Wastage 🏭

Business Process Modelling

✅ A century ago, factories faced a simple but urgent problem: how to get more done with fewer resources. As demand grew and labor costs rose, managers started looking closely at how work actually happened. Essentially, how raw materials moved through a factory, how long each task took, and where time or effort was being wasted. They realized something important: if they could see the full process step by step, they could fix the slow parts, cut out waste, and save money.

✅ That's the basic idea behind business process modeling: a method for mapping out how work flows so you can improve it. Back then, it was about machines and assembly lines. Today, it's about everything from customer service to accounting systems. But the goal hasn't changed: do things better, faster, and cheaper.

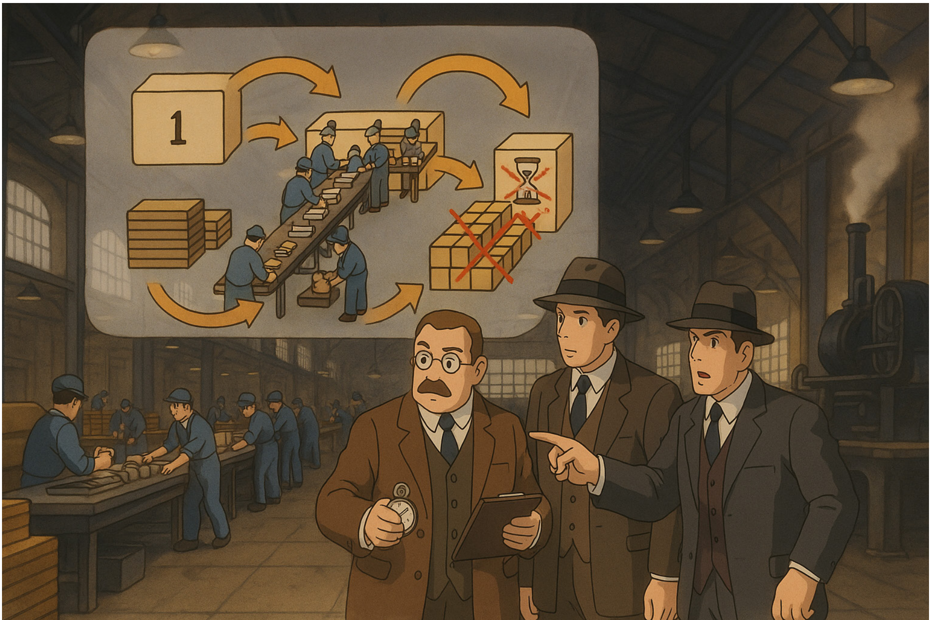
✅ Modern companies face similar pressure. Labor costs are high, customer expectations are rising, and there's more data to manage than ever before. Businesses can't afford inefficient systems or manual tasks that waste time. That's why process modeling is now used in every part of a company to understand how work actually gets done and how to improve it.

✅ Over time, process modeling has evolved. In the early days, people used flowcharts to show how paper moved through an office. By the 1970s and 80s, companies started modeling how data moved through computer systems, especially as software became central to operations. In the 1990s, tools like Business Process Model and Notation (BPMN) gave organizations a standardized way to document, analyze, and redesign even the most complex workflows, especially when adopting big systems like SAP or Oracle.

✅ Today, process models are everywhere. If you've ever drawn a workflow to explain how invoices are approved or how customer orders are handled, you've done process modeling. Tools like draw.io and even whiteboards are used every day to map and improve how businesses work.

✅ Understanding a flowchart is a lot like understanding how a program works. Both rely on breaking down a process into a clear series of steps that move from a starting point to a specific outcome. You need to know what happens first, what comes next, and what decisions or conditions change the path along the way.

✅ This is the same logic used in programming. In code 💻, you give instructions in a specific order. You might include conditions (like "if this, then do that") or loops (like



“keep doing this until you’re done”). A flowchart works the same way, just in visual form.

✅ For AIS 💻 professionals, this way of thinking is especially useful. It helps you design automations, build rules into accounting software, or explain what a system should do without needing to write the code yourself. If you can map the process, you can help build the solution.

🍩 🍩 🍩

Business process modeling (BPM) is the practice of visually representing how work is done within an organization. It allows people to document, analyze, and improve the flow of tasks, decisions, and data through a series of structured steps. These visual models help stakeholders understand how a business process operates from start to finish. There are 5 components.

1. Events
Events are things that happen during a process. They mark key points like the beginning, end, or a change that affects

the flow of activities. Example: A process might begin with "Invoice Received" or end with "Payment Sent."

2. Activities

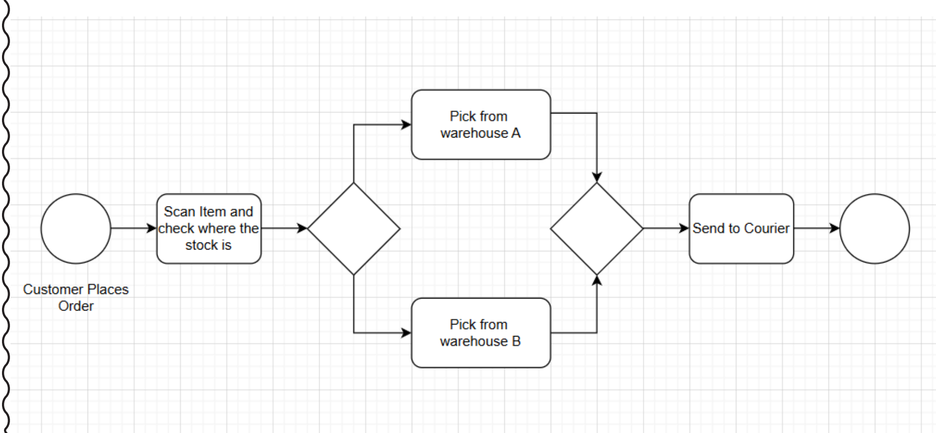
Activities are where work takes place. They represent actions, tasks, or sub-processes and are usually shown as rounded rectangles in a diagram. Each activity uses a short verb phrase to describe the task, such as "Process Payment" or "Bill Customer."

3. Sequence Flows

Sequence flows are arrows that connect events, activities, and decisions. They show the order in which tasks occur from start to finish. In diagrams, the direction typically flows left to right or top to bottom, making it easy to follow the path of the process.

4. Gateways

Gateways are decision points in the process. They are usually shown as diamonds and represent places where the process splits or merges depending on conditions. Example: A gateway might ask, "Is the invoice over $10,000?" If yes, the process follows an approval path. If no, it moves directly to payment.



5. Annotations
Annotations provide additional explanations or notes that help clarify parts of the diagram. They are connected with dashed lines and used to add context without crowding the diagram with too much detail.
Example: An annotation might say, “Manager approval required for amounts over $5,000.”
💪 Exercise:

1. Which of the following is not an example of business process documentation?
A. Business process models
B. Training manuals
C. Organization charts
D. Internal audit

2. In BPMN, what does a circle typically represent?
A. Task
B. Start event
C. Gateway
D. Data store

3. What is the primary purpose of using BPMN in an organization?
A. To improve employee morale
B. To diagram software architecture
C. To visually model business processes for analysis and communication
D. To document financial transactions

4. Which of the following BPMN elements represents a decision point?
A. Event
B. Pool
C. Gateway
D. Sub-process

5. In a BPMN diagram, which element is used to group and separate different participants in a process?
A. Lane
B. Data object
C. Annotation
D. Message flow

🙋‍♀️ Answers:

1. Answer: D. Internal audit

2. Answer: B. Start event

3. Answer: C. To visually model business processes for analysis and communication

4. Answer: C. Gateway

5. Answer: A. Lane
Lanes divide a BPMN pool to show roles or departments

—Chapter 4 -
🍩 🍩 🍩

Grocery Mayhem 🛒

Data Modelling

✅ In 2013, one of America's most successful retailers made a bold move. Target opened over 100 stores across Canada in less than a year, hoping to replicate its U.S. success. What followed wasn't a story of growth, but one of collapse. Within two years, Target Canada shut down completely after losing over $2.5 billion. Stores were half-stocked. Inventory sat in warehouses. Customers walked out empty-handed. At the core of this failure was not a lack of demand or branding: it was bad data.

✅ Target's systems, particularly its new SAP platform 💻, were flooded with incorrect, incomplete, or inconsistent

product information. Dimensions were in the wrong units. Currency fields were missing. Descriptions were vague or duplicated. Vendors supplied unreliable data, and the staff tasked with entering it had little training or experience. There were no safeguards or validations in place. At one point, Target estimated that only 30% of its product data was accurate. In the U.S., that number was closer to 99%.

✅ What Target overlooked was a basic truth of technology: systems 💻 are only as smart as the structure of their data. They rushed into operations without building a solid data foundation: one that connects products, prices, packages, and people into a usable, reliable system. What they needed was a clear, standardized method for designing and managing this information before launch.

✅ Data modeling is the practice of designing how information is structured so that systems can store, retrieve, and use it effectively. It goes beyond simply naming database fields or creating spreadsheets. At its core, data modeling defines what kinds of information exist in a system: what details each of these contain, like names, prices, or dimensions, and how they relate to each other. It is the blueprint of a system's data layer, shaping everything from daily operations to executive decisions.

✅ Historically, the challenge wasn't just designing good data. It was communicating how that data should behave across large teams. In the early days of system development, each team used its own method for sketching out how a system worked. Some used flowcharts, others relied on inconsistent diagrams or text-based specifications. As projects grew more complex, this led to confusion, duplication, and costly rework. It was clear the industry needed a shared language for describing system structure.

✅ Officially adopted as a global standard in 1997, Unified Modeling Language (UML) introduced a visual vocabulary that teams across industries could use to describe systems in a consistent and logical way.

✅ UML (Unified Modeling Language) emerged from the rise of object-oriented programming (OOP) as a way to visually represent the structure and behavior of complex software systems. As OOP 💻 gained popularity through languages like C++ and Java, developers needed a clearer way to communicate how classes, objects, and their relationships worked. UML was created to fill that gap: it took the core ideas of OOP, such as encapsulation, inheritance, and association, and translated them into standardized diagrams like class diagrams and sequence diagrams. This visual approach made it easier for teams to design, document, and understand software systems, bridging the gap between technical developers and business stakeholders.

✅ One of the most powerful elements of UML is the class diagram, which plays a central role in data modeling. In a class diagram, each entity in a system is represented as a class. These classes contain attributes, like price or weight, and define relationships with other classes. For example, one customer might place many orders, or one supplier might provide multiple products. Class diagrams also allow designers to show how many items can relate to one another and whether some categories share the same core structure.

—Chapter 5 -

🍩 🍩 🍩

Disease and Life 🦠

Introduction to Data

✅ In 1854, a deadly cholera outbreak swept through Soho in London, killing over 600 people within a matter of days. At the time, most doctors believed the disease was caused by “bad air” from rotting waste. But Dr. John Snow thought otherwise. He believed the illness was being spread through contaminated water.

✅ To investigate, Snow did something groundbreaking. He collected data on where the victims lived and created a detailed map of the cases. He noticed that most of the sick people got their water from the same public pump on Broad Street. Some who lived farther away but still drank from that pump also became ill. Snow shared this information with local officials and persuaded them to remove the pump's handle. After that, the number of new cholera cases dropped sharply.

✅ This was one of the first times data was used to solve a public health crisis. Snow's method of mapping and analyzing information became a model for how we use data today. In this course, you will learn to work with data in the same spirit. Whether it is sales numbers or inventory logs, data helps us see patterns, understand problems, and make smarter decisions based on facts rather than guesses.

✅ Turning data into information is a lot like how our brains make decisions without us even realizing it. Imagine walking into a store and seeing a shirt for $200. You don't need to do math: your brain instantly tells you it's too expensive. Why? Because it remembers what shirts usually cost. That's your brain using past data to make a quick call.

✅ Businesses do the same thing. They collect tons of raw details: prices, dates, product names, customer orders. On their own, these facts don't mean much. But once you organize them, spot patterns, and ask the right questions, they start to tell a story. Suddenly, you can see what's selling, what's not, or when to restock. Just like your instincts help you decide what to buy or avoid, good information helps businesses make smarter choices.

✅ Data on its own is just raw facts. For example, "42," "2025-07-01," or "Widget A." These pieces don't mean much until you organize and interpret them. When you put data into a structure, it becomes information. It now answers a question and supports a decision. I will introduce several types of data but take note that everybody has their own typology. The below is an introduction to how I

see data.

1. Structured Data 📒

This is the neat and tidy kind—the kind that fits perfectly into rows and columns, like a spreadsheet or a database table. Each row is usually a record (like a customer), and each column is a field (like name, age, or purchase amount).

Example:
Imagine a table of online store orders. One row might say:

- Customer: Jane Doe
- Age: 32
- Item: Running Shoes
- Price: $89.99
- Date: June 15, 2025

Structured data is easy to search, filter, and analyze using tools like Excel, SQL, or Python's Pandas library. It's the foundation of most traditional business reporting and dashboarding.

2. Unstructured Data 📁

Unstructured data is messier. It doesn't fit neatly into tables, and it often comes in the form of text, images, audio, or video.

Example:

- A product review: "The shoes are comfy, but the delivery was slow."
- A customer service phone call recording
- A photo posted on social media

Even though it doesn't come in neat boxes, unstructured data is incredibly valuable. Data scientists use natural language processing (NLP) to analyze text (like reviews or tweets), or computer vision to make sense of images and videos. It's harder to process, but it offers rich insights that numbers alone can't provide.

3. Semi-Structured Data 📁

This type of data is a mix between structured and unstructured. It doesn't fit perfectly into tables, but it has some organizational structure: usually through tags or formatting that machines can recognize. AI can help with this now.

Example:

- JSON or XML files used in APIs
- Log files from a website showing user activity

- A form submission with labeled fields

It’s not as rigid as a spreadsheet, but it’s not totally raw either. Semi-structured data is often used in web development, data integration, and modern applications that talk to each other through APIs.

4. Time-Series Data 📁

This is data collected over time, usually at regular intervals. It’s essential in areas like finance, IoT (Internet of Things), and forecasting. Timestamps also affect data – for example, UTC.

Example:

- Hourly stock prices for Apple
- Daily temperature readings from a weather station
- Website traffic logs recorded every minute

What makes time-series data unique is that time is a built-in dimension, so analysis often involves identifying trends, seasonality, or anomalies.

5. Geospatial Data 📁

Geospatial data deals with locations and spatial relationships. It tells you where something happened.

Example:

- GPS coordinates from a delivery truck
- Heatmaps showing foot traffic in a mall
- Satellite imagery used in agriculture or environmental science

This kind of data is often used in mapping, logistics, urban planning, and even climate research. It's analyzed using tools like GIS (Geographic Information Systems) and visualized with maps.

6. Numerical Data

Numerical data can be measured and used in calculations. It comes in two flavors:

Integers

These are whole numbers (no decimals).

Examples:

- Number of items in stock: 120
- Age in years: 34

You can add, subtract, multiply, and use them in statistical models directly.

Floats (also called real numbers or doubles)

Floats are numbers with decimal points, used when precision matters.

Examples:

- Price: $4.99
- Temperature: 98.6°F
- Distance: 3.14 kilometers

Float values are extremely common in financial, scientific, and engineering data. They're also the default for many statistical functions that return averages, variances, or rates.

7. Categorical Data (also called nominal data)

This type of data represents groups, categories, or labels, with no natural order between them.

Examples:

- Colors: “Red”, “Blue”, “Green”
- Countries: “USA”, “Canada”, “Japan”
- Payment methods: “Cash”, “Card”, “PayPal”

You can count how often each category appears, but you can’t meaningfully say one is greater than another. You wouldn’t say “Visa” is greater than “Mastercard”—they’re just different.

8. Ordinal Data 🗂️

Ordinal data is a special type of categorical data that does have a clear order, but the differences between values aren’t necessarily equal or measurable.

Examples:

- Customer satisfaction: “Very Unsatisfied”, “Unsatisfied”, “Neutral”, “Satisfied”, “Very Satisfied”
- Education level: “High School”, “Bachelor’s”, “Master’s”, “PhD”
- Hotel star ratings: 1 to 5 stars

With ordinal data, you know the ranking, but not the exact “distance” between categories. For instance, the difference between “Neutral” and “Satisfied” isn’t necessarily the same as between “Satisfied” and “Very Satisfied.”

Why it matters:
Ordinal data can be used in ranking or sorting, but you need to be careful with statistical operations—taking the average of ordinal categories usually doesn’t make sense.

9. Factors (especially in R) 🗂️

In R (and some other statistical languages), factors are a data type used to store categorical variables efficiently. They can be:

- Nominal factors: no order (like types of fruit)

- Ordinal factors: with order (like survey ratings)

Why use them?
Factors help statistical models know how to treat the variable. For example, logistic regression in R treats factors differently than continuous numbers.

—Chapter 6 -
🍩 🍩 🍩

Next stop: New York City 🗽

Introduction to Databases

✅ In the 1950s, booking a flight on American Airlines was anything but fast. If you wanted to reserve a seat, an agent had to call a central office where a group of operators manually flipped through a giant rotating file of index cards: one for each flight. To check if seats were available, they looked at marks on the side of the card. It worked, but only barely. As air travel exploded, the system couldn't keep up. It took up to three hours to book a single ticket.

✅ Then came a chance meeting in 1953. The president of American Airlines happened to sit next to an IBM executive on a flight. They got talking, and soon realized that IBM's work with the U.S. military, using big computers to process radar data in real-time, could also work for airline reservations.

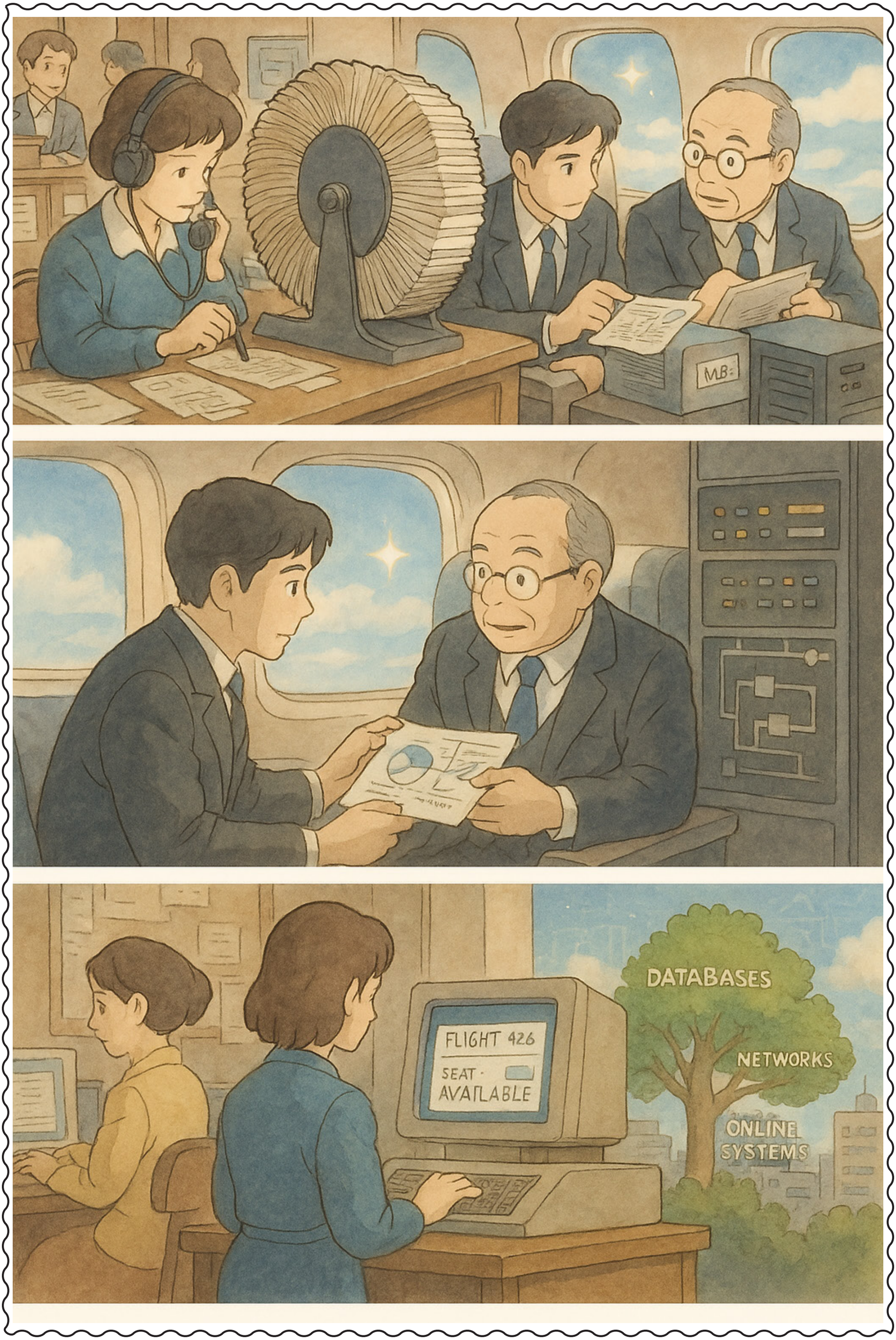
✅ From that conversation, SABRE was born: the Semi-Automated Business Research Environment, the world's first computerized airline reservation system. Instead of index cards and phone calls, SABRE used computers and teleprinters. Agents could now send booking requests electronically and get real-time seat availability.

By 1964, SABRE handled all of American Airlines' reservations. It could process over 80,000 transactions a day, which was unthinkable for a paper-based system. And the real breakthrough? SABRE didn't just make booking faster. It mapped the entire reservation process into a structured, digital system in what we now call a database.

✅ From there, the technology spread. IBM adapted it for other airlines. Eventually, it became the backbone for everything from travel agents to online booking sites. And the logic behind SABRE was how to structure data so different parts of a system talk to each other. This helped shape how modern databases and software systems work today. Before computers, businesses kept records in filing cabinets. If someone wanted to check which customer bought what, they had to sort through stacks of paper. It was slow, messy, and easy to lose things. As computers became more common in the 1960s and 70s, people realized that information could be stored and organized digitally. This gave rise to the database, which is a smarter, faster way to keep track of information.

✅ To keep these databases running smoothly, businesses rely on people called Database Administrators, or DBAs. You can think of a DBA like a digital librarian. They make sure the information is well organized, easy to find, and protected from mistakes or hackers. They also make backups, fix problems, and help systems run quickly and correctly.

✅ Today, many databases live in the cloud. That means they don't sit on computers in the office. Instead, they are hosted by companies like Amazon, Google, or Microsoft. This makes it easier for businesses to set up a database without buying expensive equipment. They can log into a website, choose what kind of database they want, and start using it right away. The cloud also helps with back-



MB:
FLIGHT 426
SEAT ·
AVAILABLE
DATABASES
NETWORKS
ONLINE
SYSTEMS

ups, security, and scaling up if more people start using the system.

Even with the cloud, DBAs are still very important. They help pick the right tools, watch for issues, set up safety rules, and make sure data follows privacy laws. Their job has changed over time, but their goal is still the same: keep information safe, fast, and reliable.

—Chapter 7 -
🍩 🍩 🍩

Tag and Win 📦

Introduction to XBRL

✅ The idea behind XBRL was born in 1998, when an American accountant named Charlie Hoffman recognized a big problem: while the internet was booming and companies were sharing tons of data online, financial reports were still stuck in static formats. He wondered: "What if financial data could be tagged and structured in a way that both people and machines could understand?"

✅ Drawing inspiration from XML (eXtensible Markup Language), which was already being used to structure web data, Hoffman and a small group of forward-thinkers began developing what would become XBRL. The goal was to digitally tag every line item in a financial report—like revenue, net income, or total assets—so software could read and compare it automatically.

✅ In 2000, the first working version of XBRL was released. That same year, the XBRL International Consortium was formed, bringing together accounting firms, technology providers, regulators, and companies from around the world to collaborate on the project. Since then, XBRL has continued to grow in scope and adoption.

By the mid-2000s, governments and regulators started to see the potential of XBRL. It wasn't just a tech upgrade: t could transform how financial information is shared and analyzed.

✅ Some major milestones:

- U.S. SEC (Securities and Exchange Commission): In 2009, the SEC began requiring public companies to submit financial reports in XBRL format.
- European Union: The European Banking Authority adopted XBRL for regulatory reporting (COREP/FINREP).
- Japan and China: Both countries made XBRL mandatory for stock exchange filings.
- Singapore, South Korea, Australia, and others also rolled out XBRL reporting standards for public and private entities.

✅ Now, XBRL is used in over 60 countries, helping create a global standard for business and financial reporting. Today, XBRL powers:

- Financial reports from public companies
- Banking and insurance supervision
- Tax filings
- Sustainability and ESG disclosures (via newer standards like iXBRL and XBRL-JSON)

✅ It makes it possible for:

- Investors to quickly compare company performance
- Regulators to spot red flags or trends in industries
- Companies to save time and reduce errors in reporting
- Software tools to automate audits and analysis


<Revenue>
<NetIncome
<Assets>
=
<Llabllities>

✅ XBRL has evolved beyond just tagging financial statements. New versions, like Inline XBRL (iXBRL), allow companies to embed tags directly into human-readable documents like HTML or PDFs. It looks like a normal report to us, but machines see the tags underneath.

✅ There's also a push toward real-time and high-frequency reporting, as well as integration with ESG (Environmental, Social, and Governance) data. That means in the future, not just financial health, but also climate risks or social impact could be tracked using standardized, machine-readable data.

—Chapter 8 -

🍩 🍩 🍩

The Phone Miracle 👼

Data Analytics I

✅ In the early 2000s, Rwanda faced a quiet crisis. Across its hills and rural villages, children were dying—often from preventable illnesses like malaria, pneumonia, and diarrhea. The country's child mortality rate was among the highest in the world. Clinics were understaffed, underfunded, and largely disconnected from one another. Health officials suspected where the problems were, but they lacked one crucial thing: timely, accurate data.

✅ Then something remarkable happened. In partnership with NGOs and data scientists, the Rwandan government introduced a simple but powerful tool—basic mobile phones. These weren't smartphones or advanced systems, just ordinary devices placed in the hands of community health workers. These workers, often volunteers living in the villages they served, began using the phones to report vital events: a child falling ill, a birth, a vaccination, a death. Each report was sent directly to a centralized database.

✅ For the first time, the Ministry of Health could see the country's health picture in near real time. If a cluster of malaria cases emerged in one district, they saw it. If a clinic was running low on antibiotics, they knew. And instead of reacting to annual surveys or incomplete reports, they could respond within days or even hours.

✅ The impact was profound. Between 2000 and 2015, Rwanda's under-five child mortality rate dropped by over 60 percent. Vaccination coverage improved dramatically. Medicine and equipment reached clinics faster. And instead of relying on blanket policies, officials began tailoring interventions to the specific needs of each region.

✅ As factories and businesses grew during the Industrial Revolution, companies started using punch cards, time-and-motion studies, and control charts to improve efficiency. Then came computers, and everything changed. Suddenly, we could store, search, and process huge amounts of data.

✅ Spreadsheets like Excel became office staples. Databases let us organize information with precision. In the 1990s, Business Intelligence tools helped companies generate reports and make better decisions. Then the internet exploded, and with it came massive amounts of data—from websites, phones, sensors, and social media.

✅ That's when "big data" took off. And today? We've moved into the era of smart analytics. With AI and machine learning, systems can detect fraud, recommend songs, or adjust prices in real time. Tools that once required a whole IT team are now available to anyone in the cloud.

D 🅿️ The 2Ds and 2Ps

✅ In traditional data analytics, especially before the rise of big data and AI, one of the most widely used frameworks to understand and classify analytics tasks was based on 2Ds and 2Ps:

1. Descriptive Analytics – What happened?
   This is the most basic form of analytics. It summarizes raw data to help you understand the past. Think charts, summaries, averages, and totals.
   Example: A sales report showing total revenue for last quarter.

2. Diagnostic Analytics – Why did it happen?
   This digs a bit deeper to identify root causes, patterns, and relationships. It often involves comparing segments, analyzing trends, or using regression techniques.
   Example: Analyzing why sales dipped in one region compared to others.

3. Predictive Analytics – What might happen next?
   This uses historical data to forecast future outcomes. It often involves machine learning, statistical modeling, and algorithms.
   Example: Predicting customer churn or next month's sales.

4. Prescriptive Analytics – What should we do about it?
   This takes predictions and goes a step further to recommend actions. It combines business rules, optimization models, and decision engines.
   Example: Suggesting the best product to promote next week based on predicted demand and inventory.

As technology and data evolve, the 2Ds and 2Ps framework is no longer the full picture. Today's analytics world is being reshaped by new types of data, real-time processing, and advanced AI capabilities.

1. Real-Time / Streaming Analytics – What's happening right now? 💾

Thanks to IoT devices, sensors, and constant data flow from the web, many companies now monitor and respond to events as they happen. Example: A logistics company tracking vehicle routes and adjusting in real time to avoid traffic.

2. Cognitive / Adaptive Analytics – How should the system learn and evolve? 💾

This refers to AI-powered systems that not only analyze data but learn and adapt over time without constant human input. Think machine learning pipelines that evolve on their own. Example: A recommendation engine that gets better as more users interact with it—like Netflix or TikTok.

3. Conversational and Natural Language Analytics – Can we just ask the data? 💾

With tools like ChatGPT, Power BI Q&A, or Google Looker, users can ask data questions in plain English—and get answers without writing queries or code. Example: Typing "Show me this year's revenue by product" and instantly getting a chart.

4. Automated Analytics (AutoML and No-Code Platforms) – Can machines build models for us? 💾

AutoML tools like Google AutoML, DataRobot, or Amazon SageMaker help users create predictive models without being data scientists. No-code platforms let business users analyze data visually. Example: A marketing manager builds a churn prediction model without writing a single line of Python.

5. Ethical, Responsible, and Explainable AI (XAI) – Can we trust what the model is doing? 💾

With AI making more decisions, organizations are focusing on transparency, bias detection, and fairness. Example: A credit scoring model must explain why someone was denied a loan—not just predict default risk.

—Chapter 9 -
🍩 🍩 🍩

Cane the Chicken 🍗

Data Analytics II - Accounting Perspective

✅ When most fast-food chains want to grow, they turn to franchising. It's quick, widespread, and gets the brand in front of more people. But Raising Cane's chose a different path. They kept nearly all of their restaurants company-owned. Not because they were cautious, but because they were strategic.

✅ Instead of letting others decide where to build or how to run a store, Cane's built a tightly controlled operation powered by data. They partnered with advanced analytics firms to understand not just where people live, but how they move, shop, and eat. Using tools like SiteZeus and Spatial.ai, they analyzed traffic flows, mobile device activity, social signals, and neighborhood trends to predict exactly where their next restaurant should go.

✅ This wasn't guesswork. It was science. These tools created what Cane's calls a "brand thumbprint"—a detailed profile of the kind of places where Cane's thrives. That could be a college town with a strong late-night scene, a busy retail corridor in the suburbs, or a city block with high foot traffic and drive-thru potential. If a location matched the data, they moved in. If not, they waited.

✅ And when they did open a new store, they didn't franchise it. They built it themselves. They hired the team, trained the staff, managed the operations, and kept the standards tight. Every Raising Cane's location, from Louisiana to California, follows the same playbook, run by the same core leadership.

✅ This level of control allows them to grow with confidence. Once they establish a presence in a new city, they don't scatter. They zoom in and open more stores nearby, creating a network of locations that reinforce each other. It's called market infill, and it keeps delivery times short, supply chains efficient, and customer loyalty strong.

✅ Raising Cane's isn't just expanding. It's scaling with precision. Every store is placed with intention. Every decision is based on data. That's how a small chicken-finger chain from Baton Rouge became a national phenomenon—one carefully chosen corner at a time.

✅ Accounting isn't just about keeping records anymore. Today, it's about using data to answer important questions and guide smart decisions. Whether you're an investor checking a company's health or a manager trying to improve performance, accounting analytics helps make sense of the numbers.

✅ It starts with asking the right questions. Is the company making a profit? Can it pay its debts? Is its stock worth the price? To answer these, analysts use data from reports, markets, and forecasts. They calculate key metrics like profit margins, debt levels, and returns to get a clear picture.

✅ Then they dig deeper. If profits dropped, was it because sales fell or costs rose? Comparing data over time

or across companies helps reveal the real reasons behind changes.

✅ Next comes looking ahead. Analysts use past data to forecast future earnings or cash flow. They also test different scenarios: what if revenue slows, or interest rates rise? These tools help businesses and investors prepare, not just react.

Once the analysis is done, it’s all about communication. Good analysts turn complex numbers into simple reports, charts, or dashboards that help decision-makers act fast.

✅ Inside companies, managers use accounting analytics to track what’s working and what’s not. They monitor things like inventory, budgets, and production costs. If something looks off, they dig into the details using tools like Excel or Power BI to find and fix problems.

✅ Managers also plan for the future. They forecast demand, estimate costs, and decide how much to produce, using both data and experience. The goal is to run more efficiently and stay ahead of surprises.

—Chapter 10 -
🍩 🍩 🍩

Putting to Practice 🤺

Data Analytics III - Practical Training

✅ Learning data analytics doesn't require diving straight into complex algorithms or coding. Many of the most powerful insights come from simple, practical techniques that anyone can learn with a bit of guidance. I will introduce several steps that you can use interchangeably as a guide to your data analytics journey.

1. Descriptive Statistics 💾

This is where most data analytics begins: understanding what's in your dataset.

Techniques to teach:

- Mean, median, mode (average, middle, most common)
- Minimum, maximum, and range
- Standard deviation and variance (how spread out the data is)
- Frequency counts

Practical example: Summarize customer ages or transaction values to understand your audience.

2. Data Visualization 💾

People understand visuals faster than tables of numbers. Teach learners to create simple but powerful charts.

Tools and skills:

- Bar charts, pie charts, and line graphs
- Histograms (for distributions)
- Scatter plots (to spot relationships)
- Using Excel, Google Sheets, or beginner-friendly platforms like Tableau Public or Power BI

Practical example: Plot daily website visitors over time or compare product sales across regions.

3. Data Cleaning 💾

Real-world data is never perfect. Teaching beginners how to clean and prepare data is one of the most important skills in analytics.

Techniques to cover:

- Handling missing values (drop, fill, or impute)
- Removing duplicates
- Correcting typos or inconsistent entries (e.g., “NY” vs “New York”)
- Standardizing formats (e.g., dates, currency)

Practical example: Clean up a customer list before doing a marketing campaign.

4. Sorting, Filtering, and Grouping 💾

These techniques help learners slice and dice their data to reveal insights.

Teach using:

- Pivot tables (in Excel or Google Sheets)
- GROUP BY and WHERE clauses in SQL
- Filtering by category, region, or time period

Practical example: Use a pivot table to compare average order values by product category.

5. Basic Exploratory Data Analysis (EDA) 💾

EDA involves asking open-ended questions to explore patterns and relationships.

Concepts to introduce:

- Checking distributions (e.g., are sales normally distributed?)
- Looking for correlations (e.g., does higher marketing spend = more sales?)
- Visualizing outliers

Practical example: Explore trends in student grades or sales performance by employee.

6. Intro to Correlation and Causation 💾

Help learners understand that correlation doesn't always mean causation, but spotting strong relationships is a starting point for deeper analysis.

Tools:

- Excel's CORREL function
- Scatter plots with trend lines

- Introduction to correlation matrices

Practical example: See if customer satisfaction is linked to repeat purchases.

7. Basic Predictive Modeling 💾

Once students are comfortable with analysis, introduce simple predictive models.

Techniques to explore:

- Linear regression (predicting one variable based on another)

- Logistic regression (predicting yes/no outcomes)
- Decision trees (classifying data into categories)

Tools: Excel (Data Analysis Toolpak), Google Sheets, or beginner Python tools like scikit-learn.

Practical example: Predict how much a customer might spend based on past behavior.

8. Storytelling with Data 💾

Analysis is only useful if it's communicated clearly. Teach students to present insights in plain language, supported by visuals.

Practice areas:

- Writing a summary of findings
- Designing simple dashboards
- Choosing the right chart for the right messagePractical example: Create a 1-page report showing how different stores performed last month.

EPILOGUE

🍩 🍩 🍩

🐇 Hunter Ng. For all learns of AIS and DA. Please direct all comments and feedback to hng@gc.cuny.edu.

The online guide will feature exercises for each Chapter for students to practice. Thank you.

ABOUT THE AUTHOR

🍩 🍩 🍩

Hunter Ng is a Singaporean-born researcher and PhD candidate in Accountancy at Baruch College in New York City. With a background in cloud computing, fintech, and cybersecurity from roles at Fujitsu, the Monetary Authority of Singapore, and KPMG, he brings real-world experience to his academic work. His current research focuses on large language models and their applications in accounting disclosures. Beyond academia, Hunter has founded a floral school, written a novel, and held leadership roles in retail management, showcasing a unique blend of analytical rigor and creative entrepreneurship. More details can be found at hunterng.com