

ACC 3202



ACC 3202

Details

- Spring 2026
- Instructor - Hunter Ng hunterboonhian.ng@baruch.cuny.edu (PLEASE NOTE THAT hunterboonhian.ng@baruchmail.cuny.edu [DEFAULT OPTION IN CUNY EMAIL] DOES NOT WORK)
- Teaching Hours – Saturdays 1pm to 2pm or by Appointment



INTRO

Hunter Ng

- Worked in KPMG Digital Trust (Singapore) (Big 4)
- Worked in a local audit firm (Singapore)
- Incorporated company and filed my books
- PhD Candidate in Accountancy



Textbook

S

Prescribed by the Department (Your Connect has access to both texts):

- Accounting Information Systems – Vernon Richardson 3rd Ed
- Introduction to Data Analytics for Accounting – Vernon Richardson, 2nd Ed



Topics Covered

- Accounting Information Systems
 - BPMN
 - UML
 - SQL
 - Databases
 - XBRL
 - Encryption (base 64)
 - Other topics in prescribed textbook
- Data Analytics
 - Data Cleaning
 - Tableau/Data Visualization
 - Basic Python
 - AGILE
 - Basic Game Devt, AI usage
 - Vibe Coding
 - Other topics in prescribed textbook



In the lap of truth



Perception

- AI's information capacity
- Your experience of life
- Mind, body, emotion, energy



Relationships

- Fulfilling your needs
- Going as a beggar
- Outsourcing your problems



Game

- Wanting to get top grades
- But willing to get less



Aliveness

- Young children bursting with life



Love Songs Or Pain Songs

- Intensity of life



Destiny

- Is there such a thing?
- 10 years ago, what seemed like destiny today is your responsibility



Breakups

- Today you feel very sad
- Tomorrow you want to have your third baby



Competition

- If your friend doesn't do well, you feel sad
- If your friend does well, you feel very sad
- Whichever way, you have fixed yourself



Youth Suicide

<https://www.cdc.gov/suicide/facts/data.html>



Why people drink alcohol or take drugs?



Phases of Life

Childhood (0–12)

Focus on physical and cognitive development through play, nutrition, rest, and exploration.

****Goal:**** Healthy growth of body and mind.

Student / Learning Phase (12–24)

A period dedicated to education, skill-building, and character formation, often guided by mentors and teachers.

****Goal:**** Knowledge and discipline.

Working & Family Life (24–48)

Focus on building a career, forming relationships, and raising a family.

****Goal:**** Financial stability and meaningful relationships.

Reflection & Transition (48–60)

Gradually step back from primary responsibilities, mentor others, and shift focus toward personal growth and purpose.

****Goal:**** Wisdom and perspective.

Inner Fulfillment (60+)

Let go of material attachments and focus on deeper meaning, contribution, and inner peace.

****Goal:**** Fulfillment and peace.



Why a physical class and not online?

- To sit here totally as a student
- To experience something which you cannot experience online
- To learn human skills



Data Analyst

How much are data analysts paid?

<https://www.datamation.com/careers/data-analytics-job-openings/>

Can you learn this online?

https://www.youtube.com/watch?v=rGx1QNdYzvs&list=PLUaB-1hjhk8FE_XZ87vPPSfHqb6OcM0cF



**Two parts in
this course**



Attendance and RMP/Reddit/Discord



I'm with you every step



AIS



AIS C1 – AIS and Firm Value

- As an accountant, what are the documents that you need to know by heart?
 - Conceptual Framework
 - US GAAP – ASC (Accounting Standard Codification)
 - Financial Statements (10Ks)



AIS C1 – AIS and Firm Value

Textbook Definition of AIS

- System that records, processes, summarizes, reports and communicates the results of business transactions to provide information for decision making. It is designed to ensure appropriate levels of internal controls for these transactions

Textbook Definition of System

- A system is a set of interrelated components that takes inputs, processes them, and produces outputs (You are a system)



AIS C1 – AIS and Firm Value



Visualize how a POS system works



AIS C1 – AIS and Firm Value

Attributes of Useful information (Conceptual Framework)

Relevance

- Confirmatory (corrects or confirms what is predicted in past)
- Predictive (helps with forecasting)
- Materiality (above a threshold where missing information impacts decisions)

Faithful Representation

- Complete (includes everything)
- Neutral (not biased)
- Free from error (no mistakes)



AIS C1 – AIS and Firm Value

There are many types of data...

Data VS information – data is an input, information is an output

Discretionary Data VS Mandatory Data – what needs to be reported to government, what need not be reported but helps in decision-making (grocery store arrangement, YouTube recommendations)



AIS C1 – AIS and Firm Value

Pathways of AIS

Qualifications	Certifying Body
Certified Information Systems Auditor CISA	ISACA
Certified Information Technology Professional CITP	AICPA
Certified Internal Auditor CIA	IIA



AIS C1 – AIS and Firm Value

Value Chain (obsolete)

- A product gains value at every chain

Example of a Value Chain

Primary Activities

Inbound Logistics → Operations → Outbound Logistics → Marketing

Support Activities

- Procurement
- Information Technology
- Human Resources Management
- Infrastructure



AIS C1 – AIS and Firm Value

ACRONYMS TO KNOW

- Enterprise System / Enterprise Resource Planning (ERP)
- Supply Chain Management Software (SCM)
- Customer Relationship Management (CRM)



AIS C1 – AIS and Firm Value

Supply Chain – flow of materials, information, payments and services from raw materials suppliers, through factories and warehouses

Supply Chain Between Partners

Stage	Entity	Core Activities
Supplier's Supplier	Chinese Firm	Source → Make → Deliver → Return
Supplier	P&G	Source → Make → Deliver → Return
Your Company	Sam's Club	Source → Make → Deliver → Return
Customer	Convenience Store	Source → Make → Deliver → Return
Customer's Customer	End Customer	Source → Make → Deliver → Return



AIS C1 – AIS and Firm Value

Customer Relationship Management



LUXURY WATCHES



AIS C1 – MCQ

Which of the following is a characteristic of relevance?

1. **Materiality**
2. Timeliness
3. Verifiable
4. Cost to generate



AIS C1 – MCQ

Which of the following is mandatory information required by a regulatory body?

1. 10-Q Submission to the Securities and Exchange Commission
2. Cost to produce a Samsung Galaxy
3. Number of organic bananas sold at Whole Foods
4. Cost to build a new Starbucks restaurant



AIS C1 – MCQ

The correct order of events in the value chain is?

1. Inbound logistics -> Operations -> Service
2. Inbound logistics -> Outbound logistics -> Marketing and Sales
3. Inbound logistics -> Operations -> Outbound Logistics
4. Inbound logistics -> Operations -> Marketing and Sales



AIS C1 – AIS and Firm Value

CLASS EXERCISE

- Close-reading of ASC 985 (<https://asc.fasb.org/985/showallinonepageplus>)
- 985-20-25, 985-20-55



AIS C1 – AIS and Firm Value

Q1.

Once technological feasibility is established, *all* subsequent software-related costs must be capitalized.

✘ False

Capitalization **ceases when the product is available for general release**, and maintenance/support costs are always expensed.



AIS C1 – AIS and Firm Value

Q2.

Which of the following *alone* is sufficient to establish technological feasibility?

- A. Completion of a product design approved by management
- B. Completion of coding for low-risk features only
- C. A working model whose completeness and consistency with the product design have been confirmed by testing
- D. Management's expectation that the software will work as intended

Correct answer: C



AIS C1 – AIS and Firm Value

Q3.

A software company incurs the following costs:

- January–March: Concept exploration and architecture design
- April–June: Detailed program design, coding, and testing of high-risk features
- July: Technological feasibility established
- August–October: Final coding, testing, and creation of the product master
- November: Software released to customers

Question:

Indicate which costs are **expensed** and which are **capitalized**, and identify the authoritative basis.

Expected logic:

- Jan–June → Expense (R&D under Subtopic 730)
- Aug–Oct → Capitalize (production costs after feasibility)
- After Nov → Expense (maintenance/support)



AIS C1 – AIS and Firm Value

Q4. A company purchases software code that can only be used in a single software product currently under development. At the purchase date, technological feasibility of the overall product has **not** yet been established.

Question:

How should the purchase be accounted for at acquisition?

Correct answer:

Expense as R&D.

- **No alternative future use**
 - **Feasibility not yet established**
- expense under 985-20-25-9



AIS C1 – AIS and Firm Value

CLASS EXERCISE

Close reading of Conceptual Framework – Chapter 3



AIS C2 – Accountants as Biz Analysts

What does an accountant do?

What an accountant can do?	
Regulatory Compliance	Finance and accounting processes
Tax Returns	Financial Close
Stakeholder assurance	Financial Consolidation, Reporting, Analysis
Investor Relations	Providing Comprehensive Management Information
Raising Capital and loans	People Management
Board Reports	Using IT to make finance and accounting more efficient
Statutory Reporting	



AIS C2 – Accountants as Biz Analysts

Activities + Structures + Rules → Complete understanding of business processes

Why Companies Model Business Processes

- Find and fix ****bottlenecks and inefficiencies**** (save time and cost)
- Design and implement ****systems (ERP, POS, AIS)**** correctly
- Standardize work for ****training, control, and scalability****



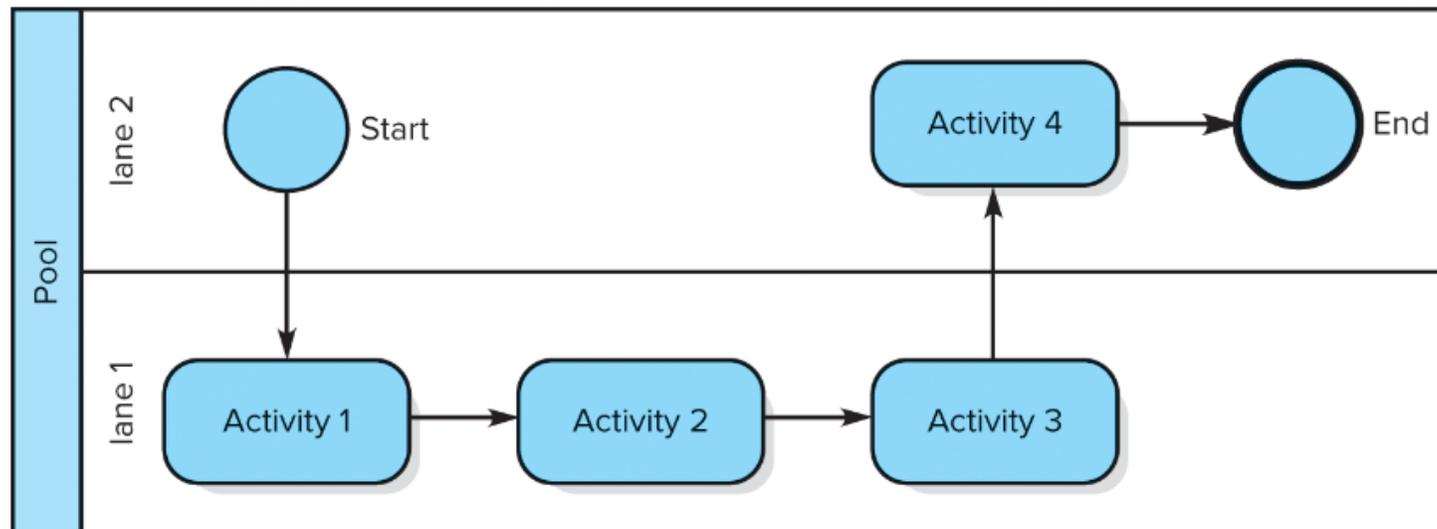
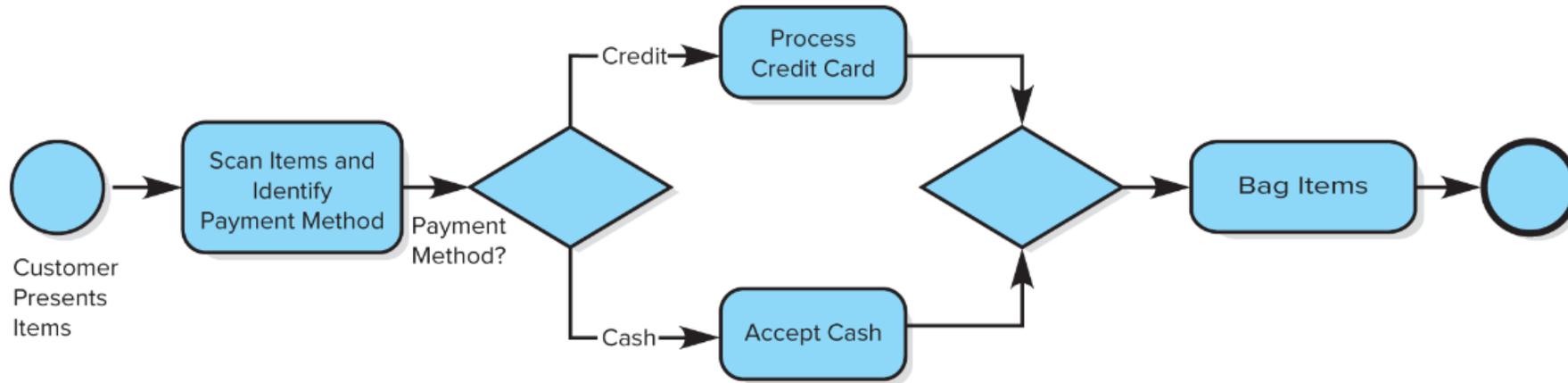
AIS C2 – Accountants as Biz Analysts

Activity model – describes sequence of workflow in a business process

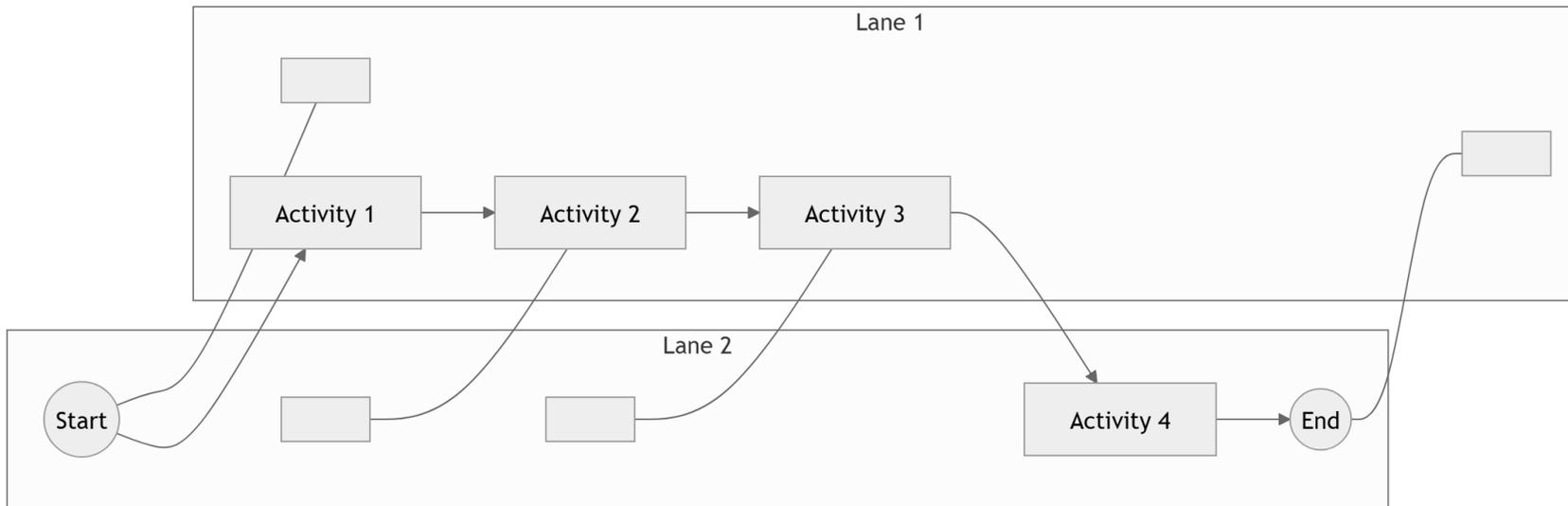
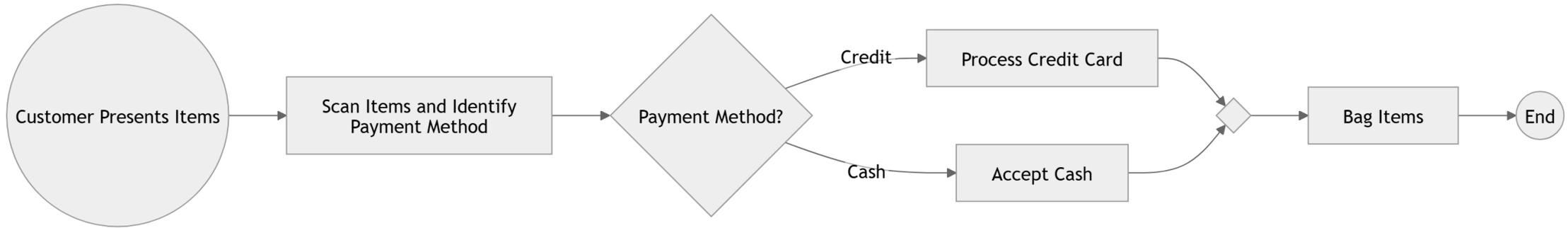
Element	Description	Symbol
Events	Things that happen	Circle
Activities	Where the work takes place, process or subprocess	Rectangle
Sequence Flow	Arrow shows the order of activities	Arrow
Gateways	Gateways control branching and merging of flow paths	Square



AIS C2 - Accountants as Biz Analysts



AIS C2 – Accountants as Biz Analysts



AIS C2 – Accountants as Biz Analysts

Lane (Swimming Pool)

shows the subdivision of responsibilities, Pool – company or organization

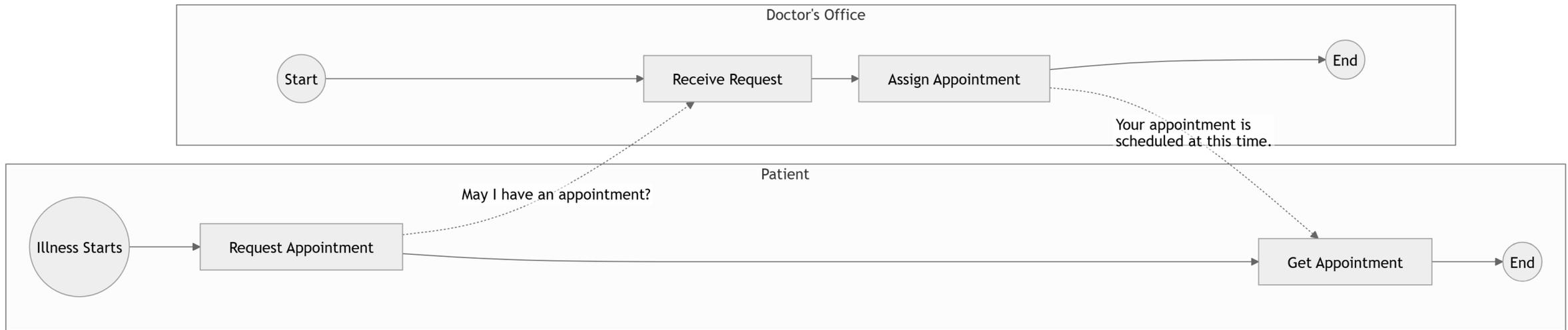
Business Process Modelling Notation BPMN

Good BPMNs are Correct Clear Complete Consistent



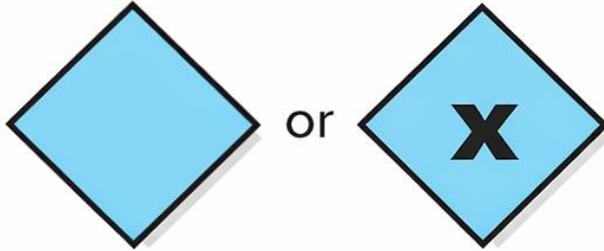
AIS C2 – Accountants as Biz Analysts

Dotted lines to show message flow – start and end events are circles



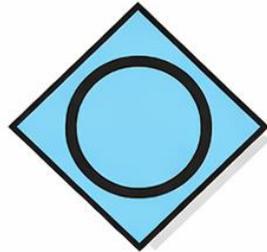
AIS C2 – Accountants as Biz Analysts

Different gateways



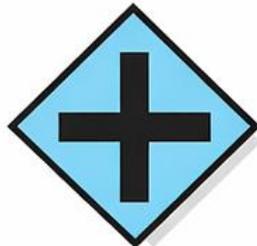
Exclusive Gateway

One path is taken based on a condition.



Inclusive Gateway

One or more paths are taken based on conditions.

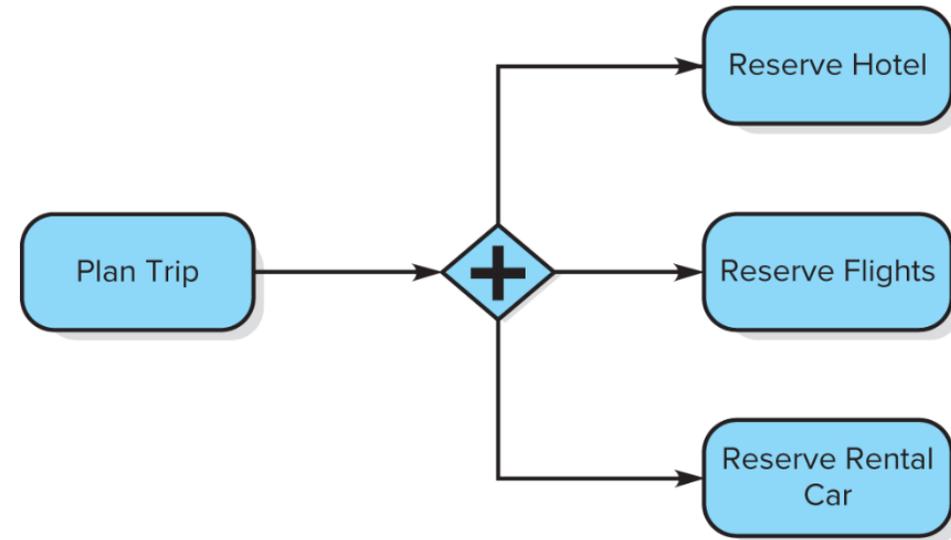
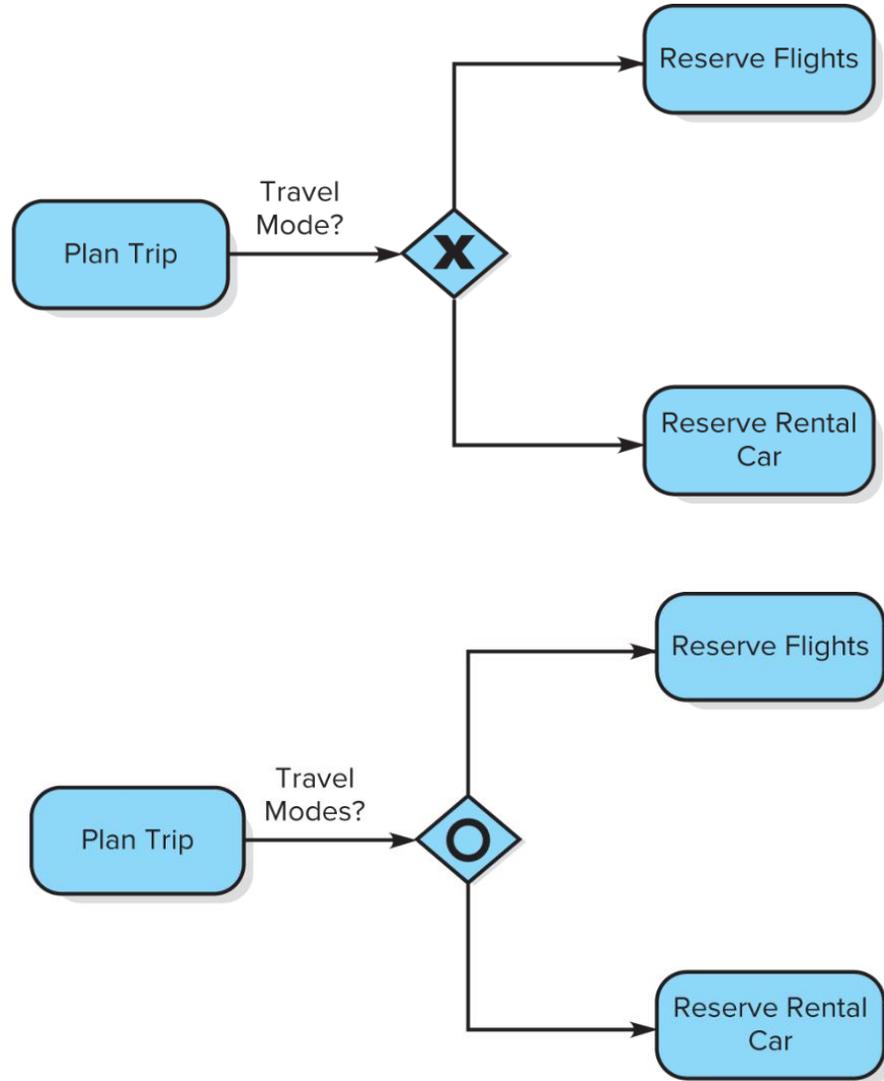


Parallel Gateway

All paths are taken in parallel.



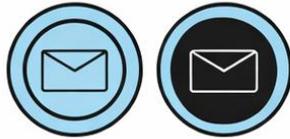
AIS C2 – Accountants as Biz Analysts



AIS C2 – Accountants as Biz Analysts



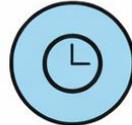
Start message (catching) event. This is used to begin a process based on an incoming message, such as a sales order. Start events can only



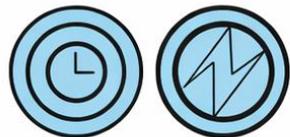
Intermediate message events. The white envelope is the catching event and the black envelope is the throwing event. These indicate that the process sends a message to or receives a message from an



End message (throwing) event. This indicates that the process sends a message when it ends.



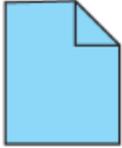
Start timer event. This is used to indicate a process that starts at a particular time or date, such as creating monthly budgets.



Intermediate timer event. This indicates a delay in the process flow until a specific time or date or for a specified period.



AIS C2 – Accountants as Biz Analysts

	A data object. This element represents data that are only available for the duration of a process.
	A datastore. This element represents data that are available across processes.
	An association connecting a data object/store to an activity. The arrowhead shows the direction of data flow when necessary.



AIS C2 – MCQ

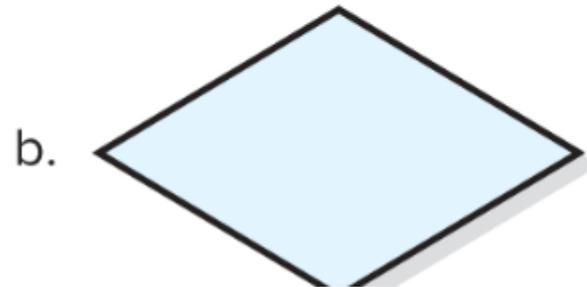
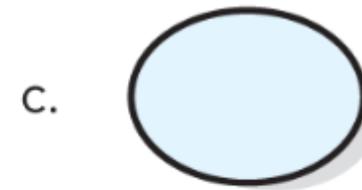
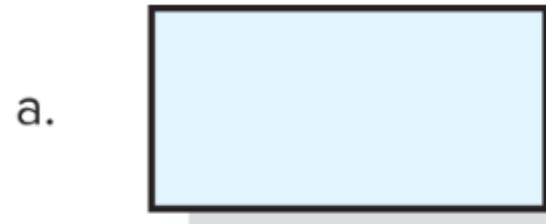
Which of the following is not a role of the accounting function in business?

1. Financial Close
2. Regulatory Compliance
3. Project Management
4. Using IT to make finance and accounting processes more efficient
5. All of the aboves are roles of the accounting function.



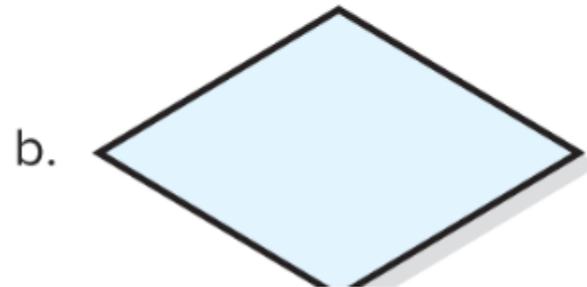
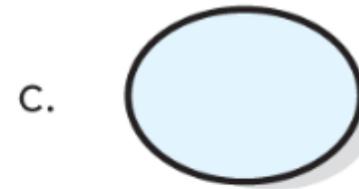
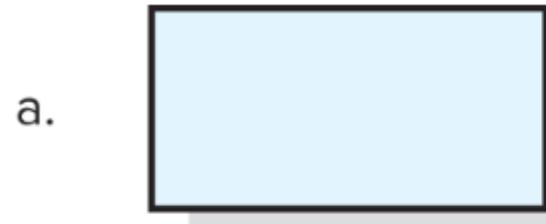
AIS C2 – MCQ

Which of the following is a sequence flow? d



AIS C2 – MCQ

Which of the following is a gateway? b



AIS C2 – Accountants as Biz Analysts

CLASS EXERCISE

Library BPMN Exercise (Draw in class)



AIS C3 – Data Modelling

Unified Modelling Language – UML

UML (Unified Modeling Language) is a standardized visual language used to represent the structure and relationships within systems

- ****Class (Entity / Object Type / Table)****

A distinct category of objects (e.g., Customer, Order)

- ****Association (Relationship / Link / Connection)****

A connection between classes

- ****Multiplicity (Cardinality / Participation)****

The number of instances that can be linked between classes (e.g., 1..1, 0..*)

- ****Attributes (Fields / Columns / Properties)****

Data elements that describe a class (e.g., name, date, amount)



AIS C3 – Data Modelling

Options_for_Multiplicities

0..1 → Zero or one

1..1 → One only

0..* → 0 or more

1..* → 1 or more



AIS C3 – Data Modelling

Database Schema

A schema is the structured blueprint of a database that defines how data is organized: what tables exist, what fields they contain, and how those tables relate to each other.

Customers

[Customer_Number (PK), Customer_Name, Customer_City, Customer_State, Customer_Zip, Customer_Phone]

Orders

[Order_Number (PK), Order_Date, Delivery_Date, Order_Amount, Shipping_Cost, Customer_Number (FK)]

Order_Items

[Order_Number + Product_Number (PK), Quantity_Ordered, Price]

Inventory

[Product_Number (PK), Product_Description, Quantity_on_Hand (QOH), Unit_of_Issue, Current_List_Price, Standard_Cost]



AIS C3 – Data Modelling

Primary Key (PK)

- An attribute or combination of attributes that uniquely identifies each instance in a class (table row).

Properties of a Primary Key

- - Uniquely identifies each instance
- - Cannot be null
- - Should not change over time
- - Controlled by the organization that assigns it
- - Sequential values help detect gaps
- - Shorter keys are preferred over longer ones



AIS C3 – Data Modelling

Foreign Key (FK)

is an attribute or combination of attributes that allows tables to be linked together – FK is linked to the PK of another table to support a defined association

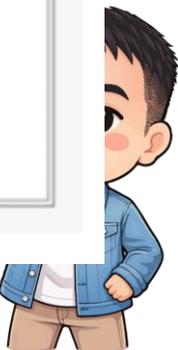
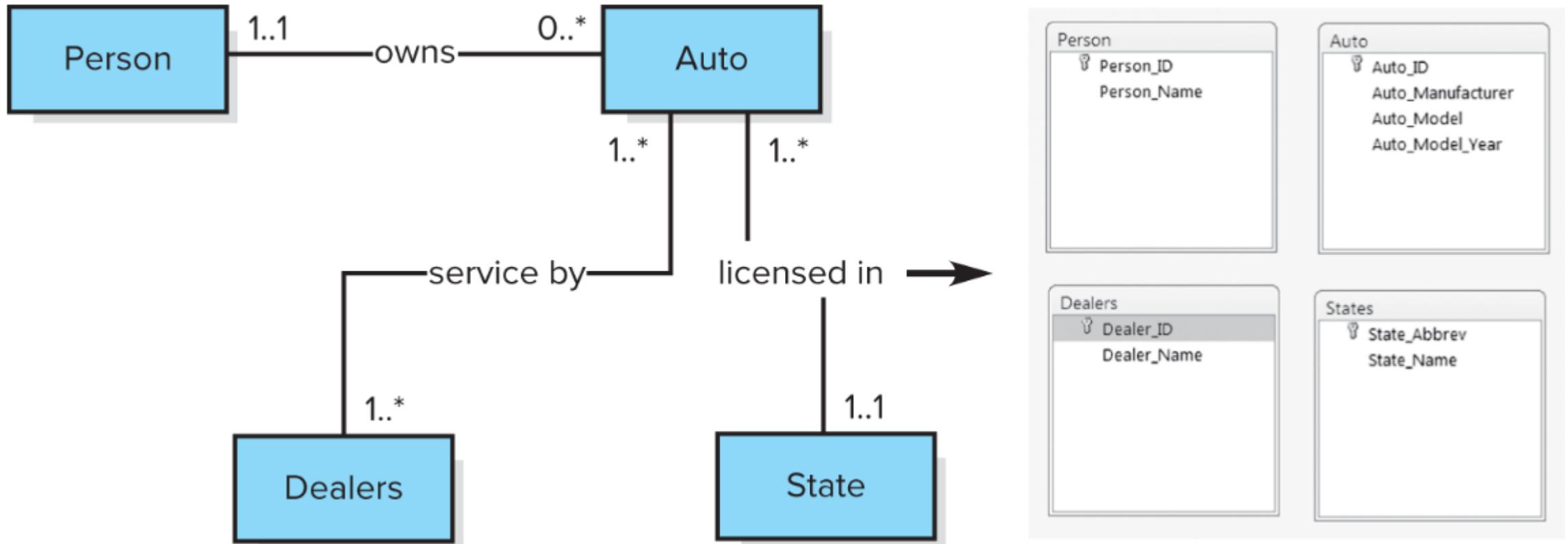
Other relationships

- Generalization (or inheritance)
- Aggregation
- Composition



AIS C3 – Data Modelling

Map a UML class diagram to a Relational Database Schema



AIS C3 – Data Modelling

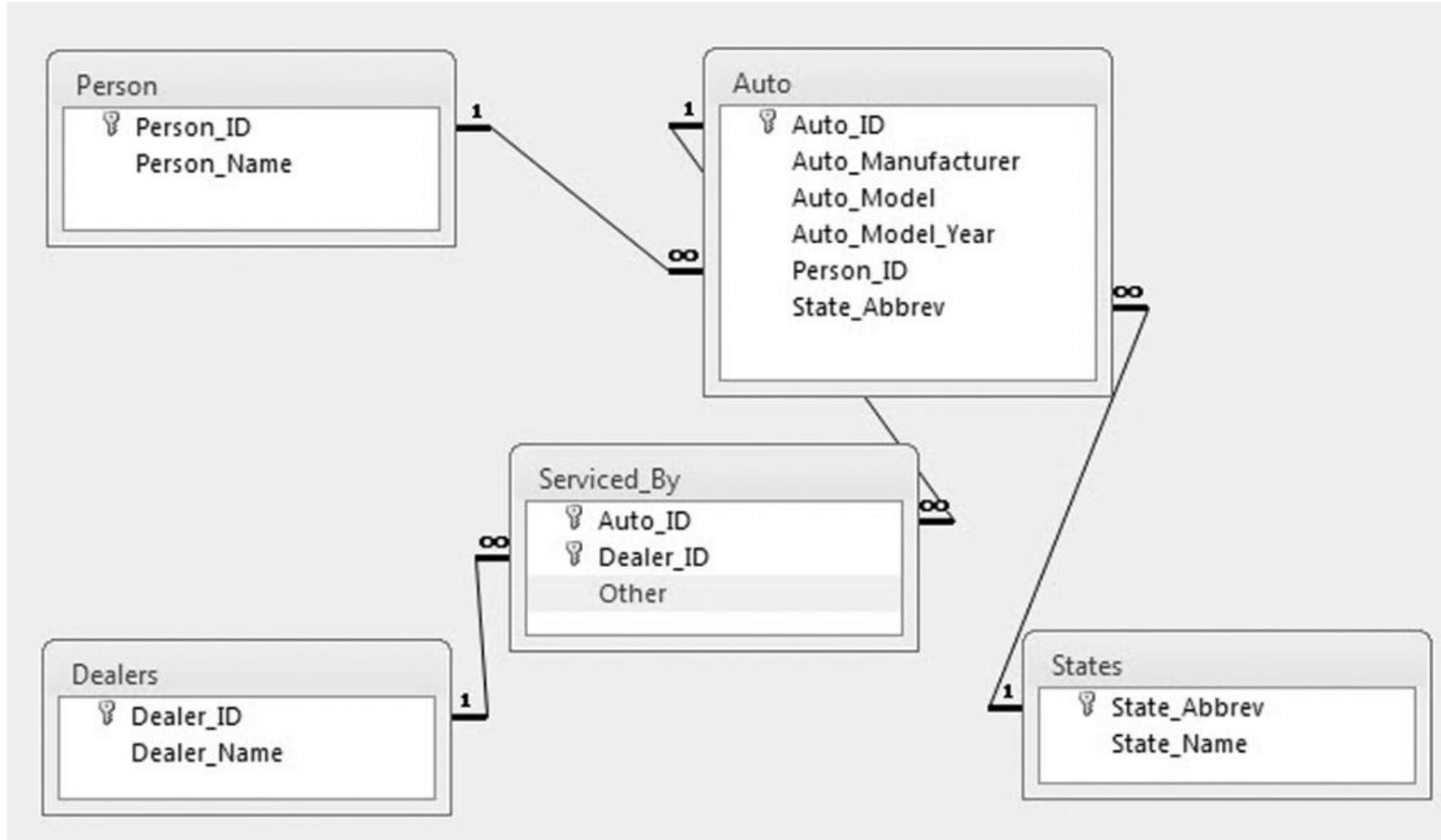
Relational DB schema diagram



LINKING TABLE – reduces repetition in data



AIS C3 – Data Modelling



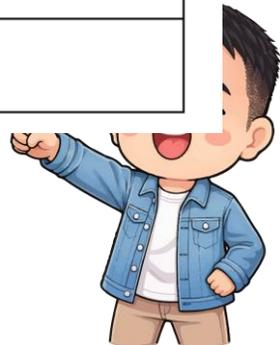
LINKING TABLE – reduces repetition in data



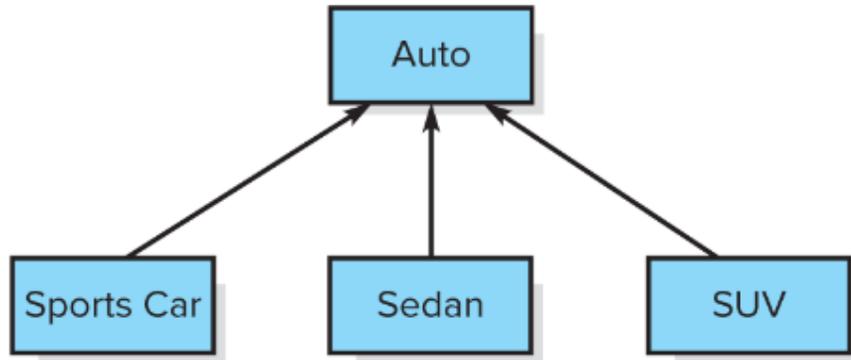
AIS C3 – Data Modelling

- Business rule – succinct statement of a constraint on a business process – represent as a decision table – think of it as an idiot-proof manual – fast-food restaurant

Discounts			
Number	Customer Type	Size of Order	Discount to Offer
1	Business	<10	10%
2	Business	>=10	15%
3	Individual	All	5%



AIS C3 – Data Modelling



Generalization

Groups similar things to reduce redundancy.

Shared features are modeled once, while subtypes have their own unique traits.



Aggregation

A “has-a” relationship where parts can exist independently. (e.g., players belong to a team but can exist on their own)



Composition

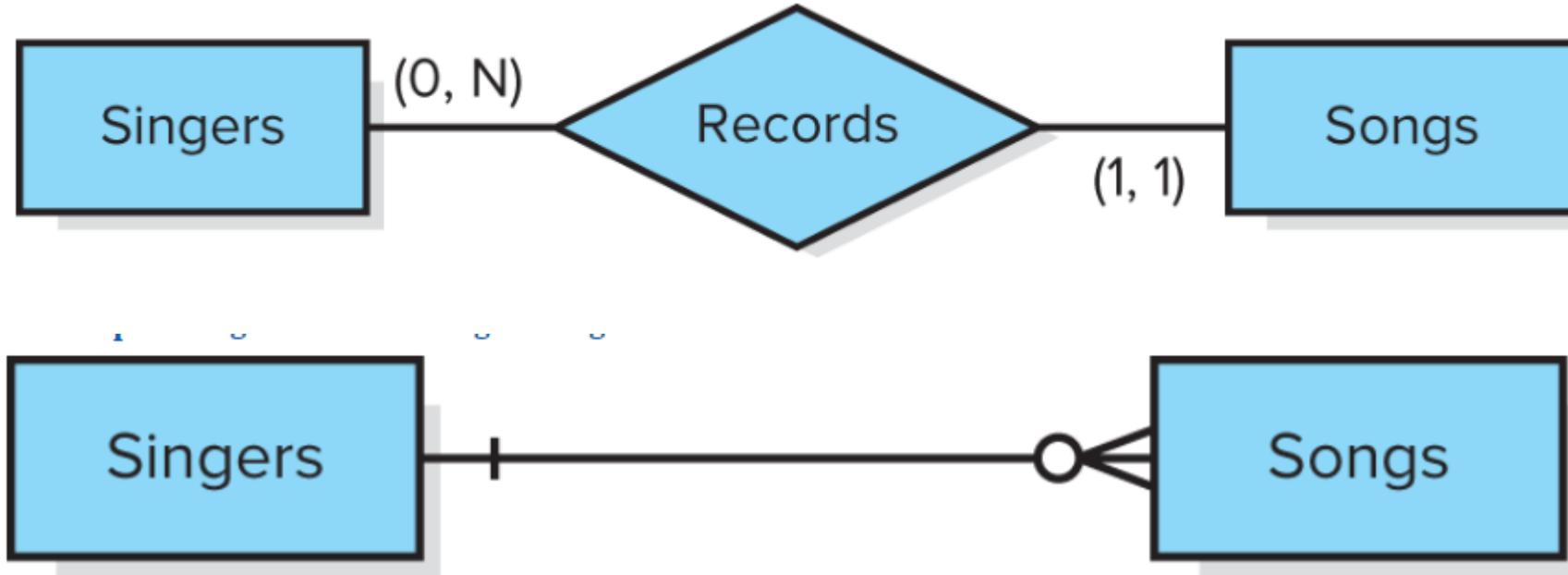
A strong “has-a” relationship where parts cannot exist independently.

(e.g., chapters only exist within a book)



AIS C3 – MCQ

- Business rule – succinct statement of a constraint on a business process – represent as a decision table – think of it as an idiot-proof manual – fast-food restaurant
- Entity Relationship Diagram – ERD – Chen's Notation vs Info Engineering notation



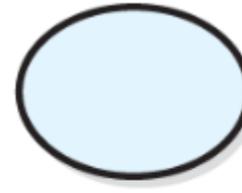
AIS C3 – MCQ

Which of the following represents a class in a UML diagram? A

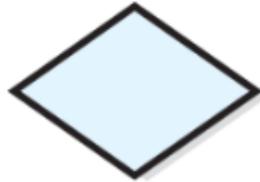
a.



c.



b.

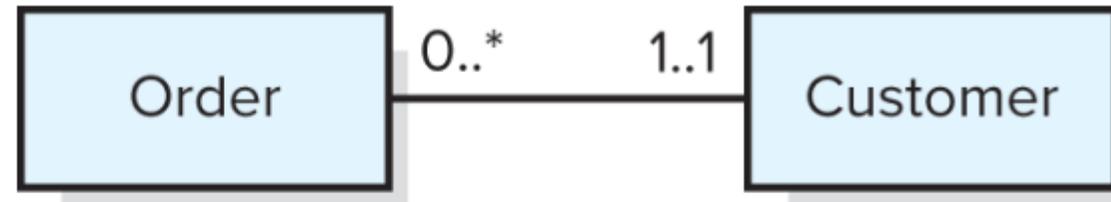


d.



AIS C3 – MCQ

Which of the following statements regarding this class diagram with multiplicities is not true?



- a. An order can involve only one customer.
- b. A customer can place only one order.
- c. A customer can place many orders.
- d. A customer may not have ordered yet.
- e. All of the above are true.



AIS C3 – MCQ

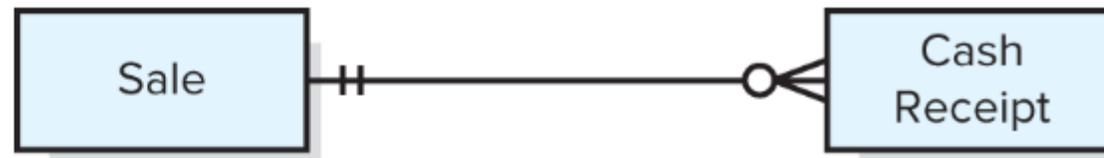
Which of the following is not an enforcement level for a business rule?

1. Strict enforcement
2. Optional enforcement
3. Pre-override enforcement
4. Post-override enforcement
5. **Guideline**



AIS C3 – MCQ

Which of the following best describes the meaning of this diagram?



1 and only 1

Each sale can
have zero or
many cash
receipts



- a. Each sale can result in many subsequent cash receipts.
- b. Each cash receipt can apply to many sales.
- c. All sales must have cash receipts.
- d. Each sale can result in only one cash receipt.
- e. None of the above



AIS C3 – MCQ

Dr. Franklin runs a small medical clinic specializing in family practice. She has many patients. When the patients visit the clinic, she may perform several tests to diagnose their conditions. She bills the patient one amount for the visit plus additional amounts for each test.



AIS C3 – MCQ

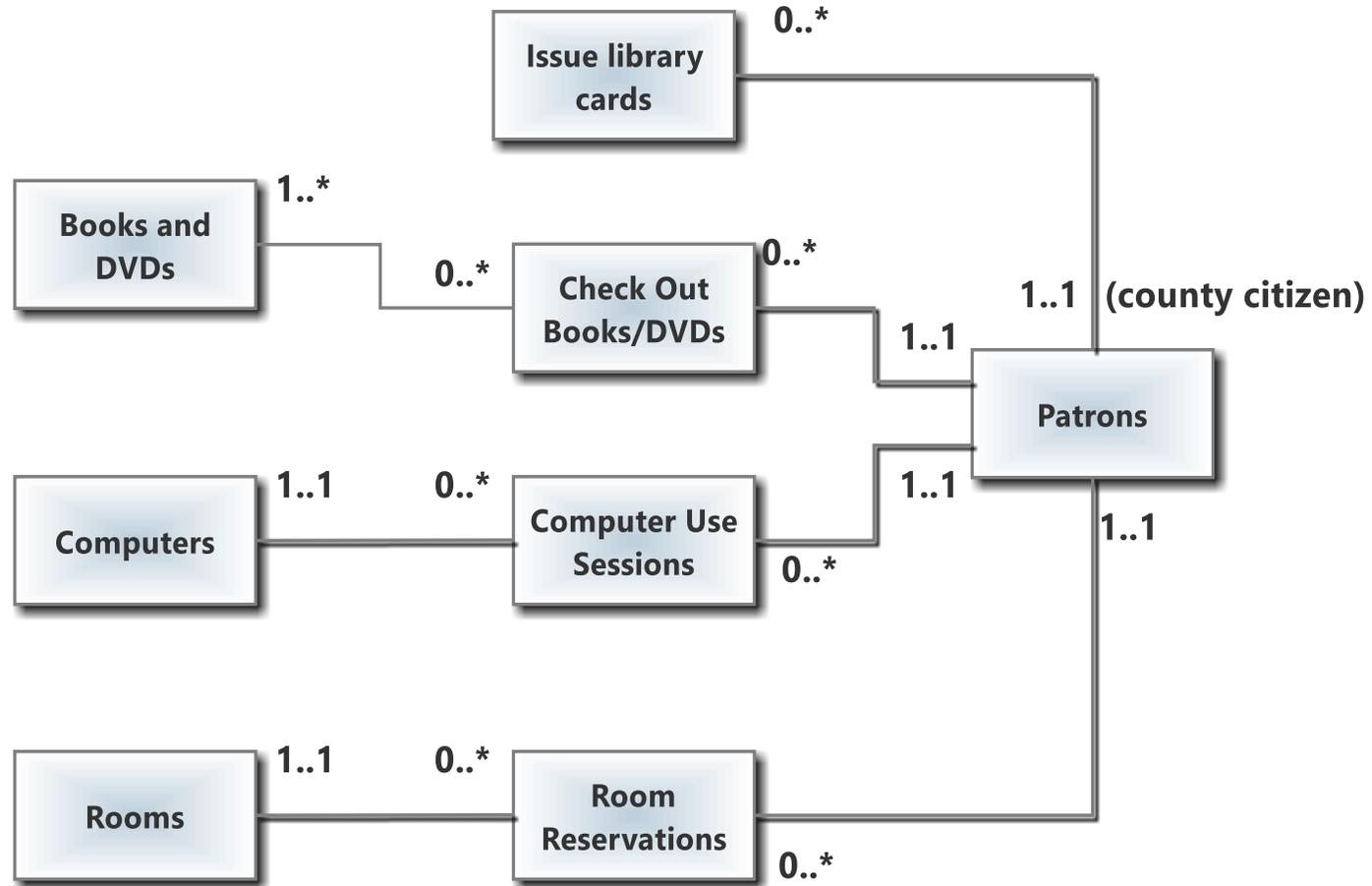
5. Develop a UML class diagram with classes, associations, and multiplicities based on the following narrative:  LO 3-2,
 LO 3-3



The Multnomah County Library provides a variety of services to citizens of the county. First, the library offers a number of traditional books and movies that may be checked out by patrons. Each patron may check out up to five books and two movies at one time. The library also offers several computers for patron use. The library tracks computer use sessions. Each patron may use the computer for up to 30 minutes per session. The library also maintains a number of small meeting rooms that patrons may reserve for classes or events during normal operating hours. Although library services are offered to all county citizens, each patron must obtain a library card to check out books, use the computers, or reserve a room.



AIS C3 – MCQ



AIS C3 – Data Modelling

- Business Process Modelling Notation (BPMN)
- Unified Modelling Language (UML)



AIS C4 – Relational DB and Enterprise Systems

A database is a collection of related data for various uses

- Hierarchical data models – widely used in mainframe DBMS – organize into tree-like structure using parent/child relationships
- Network Data model – flexible model representing objects and relationships
- Relational Data Model – stores information in the form of related 2-D tables (flexibility and scalability, simplicity, reduced information redundancy)

How would you store data?



AIS C4 – Relational DB and Enterprise Systems

DBMS – computer program that creates, modifies, and queries database

Data dictionary – describes the data fields in each database record

Database administrator – database administrator is responsible for the design, implementation, repair and security of a firm's database



AIS C4 – Relational DB and Enterprise Systems

Requirements of a Table

- The primary key must always have a value
- A foreign key must either be empty or match a value in the related table
- Each attribute (column) must have a unique name
- Values in a column must be of the same data type
- Each field must contain only a single value (no multiple values in one cell)
- Non-key attributes should describe the entity (the table) and not something else



AIS C4 – Relational DB and Enterprise Systems

Entities and Attributes

- A class can be grouped into resources, events and agents
- Attributes – characteristics, properties or adjectives that describe each class
- Rows in a table are called records/tuples/instance
- PK is an attribute or combination of attributes that uniquely identifies a specific row in a table
- FK is an attribute in one table that is a primary key in another table



AIS C4 – Relational DB and Enterprise Systems

Demonstration of how tables link

Product No.	Description	Price	Amount
1001	Apple iPhone13	799.00	799.00
Order Total			799.00

Customer Table			
Customer ID	Customer Last Name Name	Customer First Name	Customer Address
301-555-0145	John	Davis	169 Apple Rd., Rockville, MD 20852
718-555-9981	Emily	Patel	2500 Campanile Dr., NY, NY 10001

Primary Key



Sales Table			
Invoice No.	Customer ID	Date	Product No.
24731	301-555-0145	6/12/2022	1001
26845	301-555-0145	8/28/2022	1104
31259	718-555-9981	11/18/2022	1001
32567	718-555-9981	12/1/2022	2152
34041	718-555-9981	12/15/2022	3165

Primary Key



Product Table		
Product No.	Customer Name	Product Price
1001	Apple iPhone 13	799.00
1104	Samsung Galaxy S21	899.00
1001	AirPods	149.00
2152	iPhone case	39.00
3165	iPhone Charger	25.00

Primary Key



AIS C4 – Relational DB and Enterprise Systems

Microsoft Access

Simple DBMS that can be used to run databases for individuals and small firms

Access DEMO



AIS C4 – Relational DB and Enterprise Systems

Class Exercise

Your team has been hired to design the **database** for a new mobile game called **Campus Royale**. It's a multiplayer game played **only by college students**/

Players earn points by:

- Attending class on time
- Resisting TikTok during lectures
- Completing side quests like *“Submit an assignment before 11:59 PM”*

The university wants **aggregate insights** (not individual surveillance... allegedly 🙄).

Q: What data would you capture in the game's database to make the game work well and generate useful insights?



AIS C4 – Relational DB and Enterprise Systems

SAP Modules in an ERP System

Core (Financial & Logistics)

- **FI (Financial Accounting)** – General ledger, accounts payable/receivable, financial reporting
- **CO (Controlling)** – Cost accounting, budgeting, internal performance analysis
- **MM (Materials Management)** – Procurement, inventory management
- **SD (Sales and Distribution)** – Order processing, billing, shipping
- **PP (Production Planning)** – Manufacturing planning and scheduling

Human Resources

- **HCM / HR (Human Capital Management)** – Employee data, payroll, recruitment

Supply Chain & Operations

- **WM (Warehouse Management)** – Warehouse operations and storage
- **EWM (Extended Warehouse Management)** – Advanced warehouse processes
- **QM (Quality Management)** – Quality control and compliance

Customer & Vendor Management

- **CRM (Customer Relationship Management)** – Customer interactions and sales tracking
- **SRM (Supplier Relationship Management)** – Supplier sourcing and procurement



AIS C4 – Relational DB and Enterprise Systems

1 Hierarchical data model

Structure: Tree (parent → child)

Rule: Each child has **one and only one parent**

How it works

- Data is organized like a family tree
- You navigate **top-down**
- Relationships are **one-to-many only**

Problems

- What if a student takes **multiple courses**?
- What if a course is cross-listed across departments?
- You must **duplicate data** or redesign the tree

University

└─ College

└─ Department

└─ Course

└─ Student



AIS C4 – Relational DB and Enterprise Systems

2 Network data model

Structure: Graph (records connected by links)

Rule: A child can have **multiple parents**

How it works

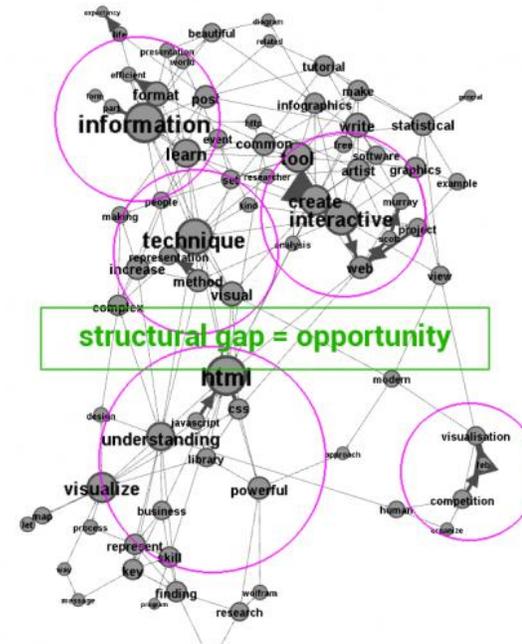
- More flexible than hierarchical
- Supports **many-to-many** relationships
- Relationships are stored as explicit **pointers**

Problems

- Very **complex to design**
- Queries depend on **knowing paths in advance**
- Changes to structure can break applications
- Hard to teach, audit, and maintain

University system (network)

- Student ↔ Course
- Course ↔ Instructor
- Instructor ↔ Department



AIS C4 – Relational DB and Enterprise Systems

③ Relational data model (what we use today)

Structure: Tables (rows & columns)

Rule: Relationships handled using **keys**, not pointers

Core idea

- Data is stored in **simple tables**
- Tables are connected logically using **primary keys and foreign keys**
- You don't "navigate" — you **query**

Why the relational model won (and why we teach it)

☑ 1. Flexibility without chaos

- Supports **one-to-one, one-to-many, many-to-many**
- No rigid tree or fixed navigation paths

☑ 2. Data independence

- You can change **how data is stored** without rewriting all applications
- Users query **what they want**, not *how to traverse the structure*



AIS C4 – SQL

Structured Query Language

Different implementations: *MYSQL*, *MONGODB* – think of it as countries and languages.

English is spoken in USA, Britain, Hong Kong, India, Singapore. Each country is a SQL program, but they all speak SQL

SELECT Name FROM Customer



AIS C4 – SQL

W3schools SQL command tutorial



AIS C4 – SQL

CRUD

RESTFUL



AIS C4 – MCQ

In the hierarchical model, mapping from parent to child is

1. One to one
2. One to many
3. Many to many
4. Many to one



AIS C4 – MCQ

Advantages of relational data models

1. **Reduced information redundancy**
2. Low cost
3. Ease of implementation
4. efficiency



AIS C4 – MCQ

Which statements about enterprise systems is correct?

1. Most enterprise systems are designed mainly for accounting and finance functions
2. SAP, Oracle and Microsoft all offer products for enterprise systems
3. Most enterprise systems are designed for the service industry
4. Small companies do not use enterprise systems at all



AIS C4 – MCQ

The FROM Clause to the SELECT statement used in SQL indicates

1. Name of the tables from which to retrieve data
2. Name of columns from which to retrieve data
3. Name of databases from which to retrieve data
4. Name of query from which to retrieve data



AIS C4 – MCQ

Explain differences among hierarchical, network and relational data models. What makes the relational data model the most popular data model in use today?

Hierarchical data models organize data into a tree-like structure. In a hierarchical data model, data elements are related to each other using one-to-many relationships. A network data model is a flexible model representing objects and their relationships. It allows many-to-many relationships. The relational data model is a data model that stores information in the form of related two-dimensional tables. While hierarchical and network data models require relationships to be formed at the database creation, relational data models can be made up as needed.

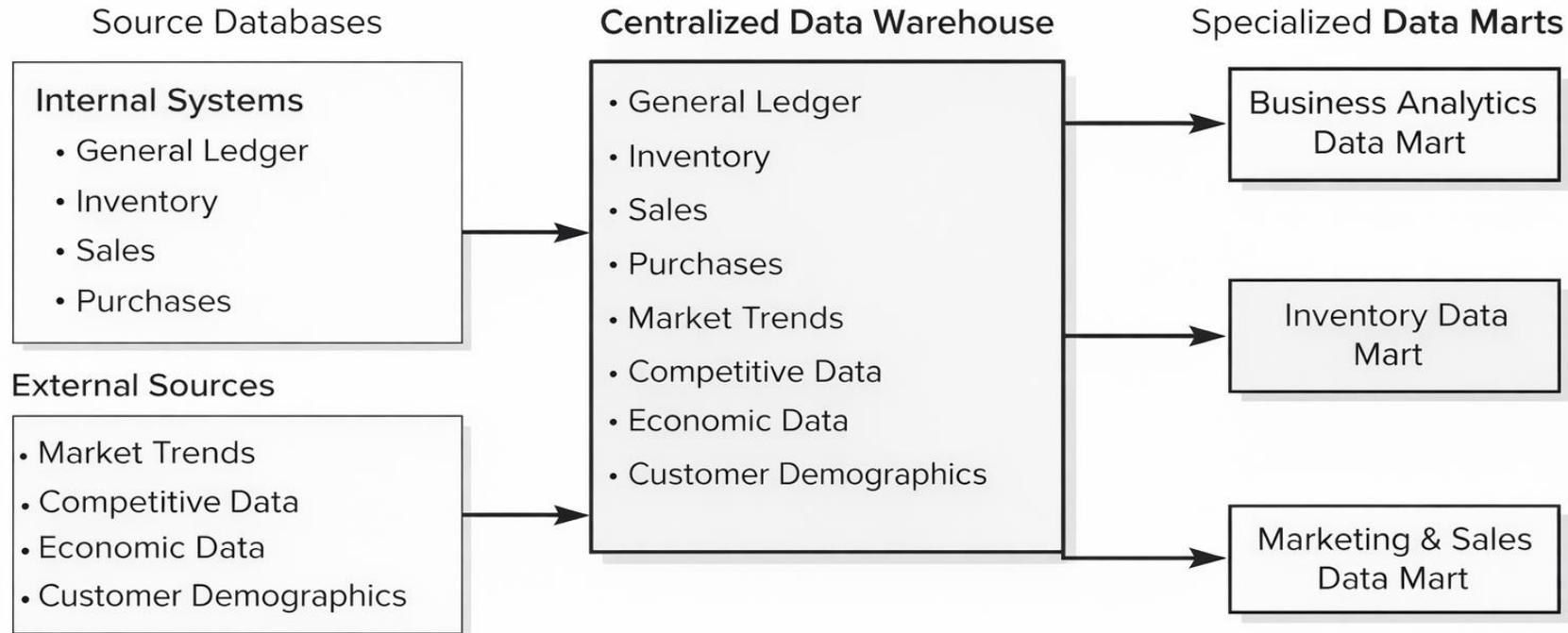
The relational database is the most popular data model in use today because it has the following advantages:

- flexibility and scalability
- simplicity
- reduced information redundancy.



AIS C9 - XBRL

Data Architecture



AIS C9 – XBRL

Class Exercise

Draw a data storage schema for Starbucks



AIS C9 – XBRL

XBRL stands for extensible Business Reporting language

- **Apr 1998** — Charles Hoffman (CPA, Knight, Vale & Gregory, Tacoma, WA) proposes using XML for business reporting—the seed of **XBRL** (then XFRML).
- **Jan 1999** — Prototype presented to the **AICPA**; the AICPA endorses and backs further development.
- **Jul 2000** — First **XBRL specification** released for financial statements of commercial/industrial companies.
- **Jun 2001** — Initiative launched to apply XBRL to internal company transactions (**XBRL GL / General Ledger**).
- **Oct 2001** — XBRL acknowledges country-specific reporting needs and organizes work into jurisdictional taxonomies.
- **Feb 2002** — **APRA** (Australia) announces plans to collect supervisory returns using XBRL.
- **Dec 2003** — **XBRL 2.1** specification released (becomes the stable baseline).
- **May 2004** — **XBRL-UK** releases a taxonomy for **UK GAAP**.
- **Jul 2005** — **XBRL International** releases an updated **GL taxonomy**, improving capture of detailed accounting entries/business events.
- **2008–early 2009** — The **U.S. SEC** mandates public companies submit financial statements as **XBRL (interactive data)** to EDGAR and post them on corporate websites, with phased adoption beginning in 2009.



AIS C9 – XBRL

XBRL is XML-based markup for business and financial reporting.

It tags financial facts (such as revenue, assets, and net income) so computers can read, compare, and analyze them.

```
<xbrli:context id="CurrentYear">
  <xbrli:entity>
    <xbrli:identifier scheme="http://www.sec.gov/CIK">0000123456</xbrli:identifier>
  </xbrli:entity>
  <xbrli:period>
    <xbrli:startDate>2025-01-01</xbrli:startDate>
    <xbrli:endDate>2025-12-31</xbrli:endDate>
  </xbrli:period>
</xbrli:context>

<!-- Unit -->
<xbrli:unit id="USD">
  <xbrli:measure>iso4217:USD</xbrli:measure>
</xbrli:unit>

<!-- Company Information -->
<dei:EntityRegistrantName contextRef="YearEnd">Example Corporation</dei:EntityRegistrantName>
<dei:TradingSymbol contextRef="YearEnd">EXMP</dei:TradingSymbol>

<!-- Financial Facts -->
<us-gaap:Assets contextRef="YearEnd" unitRef="USD" decimals="-3">2500000</us-gaap:Assets>
<us-gaap:Liabilities contextRef="YearEnd" unitRef="USD" decimals="-3">900000</us-gaap:Liabilities>
<us-gaap:StockholdersEquity contextRef="YearEnd" unitRef="USD" decimals="-3">1600000</us-gaap:StockholdersEquity>

<us-gaap:Revenues contextRef="CurrentYear" unitRef="USD" decimals="-3">4200000</us-gaap:Revenues>
<us-gaap:NetIncomeLoss contextRef="CurrentYear" unitRef="USD" decimals="-3">580000</us-gaap:NetIncomeLoss>
```



AIS C9 – XBRL

Class Practice

Go to EDGAR, Find Apple 10K for 2023, Find the statements of operations (income statement), Find the total cost of sales and list down all the related tag information

Go to EDGAR, Find Apple 10K for 2023, Find the statements of comprehensive income, Find the change in FV of derivative instruments, and click the tags at the top to change to custom, and list down all the related tag information



AIS C13 – Accounting Information Systems and Internal Controls

Firms choose a code of ethics: Corporate governance – set of processes and policies in managing an organization with sound ethics and internal and external control mechanisms to safe guard interests of stakeholders

Preventive/detective/corrective controls – different types of controls

COSO Internal control framework

COBIT – control objectives for information and related technology (COBIT)



AIS C13 – Accounting Information Systems and Internal Controls

Sarbanes-Oxley Act – SOX – requires companies registered with SEC and their auditors to annually assess and report on the design and effectiveness of internal control over financial reporting



AIS C13 – AIS and Internal Controls

COSO

Framework

- ```
Control Environment
1. Commitment to integrity and ethical values
2. Oversight by the board/management
3. Clear structure, authority, and responsibility
4. Commitment to competence
5. Accountability enforcement

Risk Assessment
6. Set clear objectives
7. Identify and analyze risks
8. Assess fraud risk
9. Identify and analyze significant changes

Control Activities
10. Develop control activities
11. Implement IT general controls
12. Establish policies and procedures

Information and Communication
13. Use relevant information
14. Communicate internally
15. Communicate externally

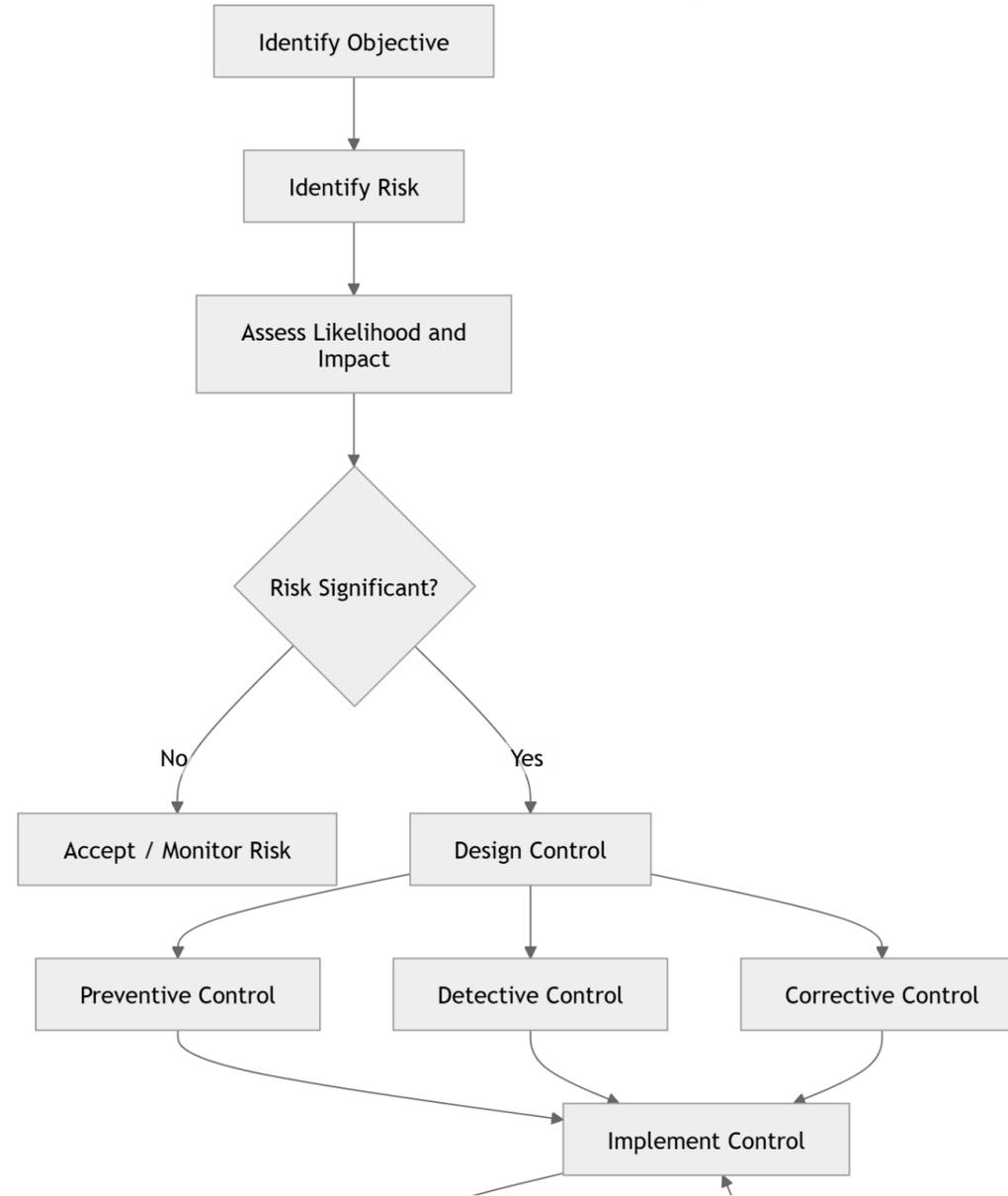
Monitoring Activities
16. Perform ongoing or separate evaluations
17. Report and address deficiencies
```



# AIS C13 – Accounting Information Systems and Internal Controls

AIS C13 Controls

COSO IC Framework



# AIS C13 – Accounting Information Systems and Internal Controls

**Control objectives** – internationally accepted set of best IT and security and control practices

**COSO IC Framework** – control environment, risk assessment, control activities, information and communication, monitoring

**COSO ERM Framework** – Internal environment, objective-setting, event identification, risk assessment, risk response, control activities, information and communication, monitoring



# AIS C13 – Accounting Information Systems and Internal Controls

## COBIT Framework

Supporting tool set that bridges the gap among IT Control requirements



# AIS C13 – Accounting Information Systems and Internal Controls

## ISO 27000

- Internationally accepted sets of information security framework



# AIS C13 – MCQ

Based on SOX, which of the following sections is about internal controls?

1. 302
2. 401
3. 404
4. 906



# AIS C13 – MCQ

SOX requires companies to use COSO or COSO ERM as framework in evaluating IC

1. True

2. False



# AIS C13 – MCQ

A field check is a

1. Preventive control
2. Detective control
3. Corrective control
4. General control
5. Output total



# AIS C13 – MCQ

Which is not an example of a batch total?

1. Record count
2. Financial total
3. Hash total (calculate this instead of the file)
4. Exception total



# AIS C13 – MCQ

The computer sums the four digits of a customer number to calculate value of fifth digit and compares calculation to the number typed during data entry. This is an example of?

1. Field check
2. Parity check
3. Check digit verification
4. Batch Total



# AIS C13 – MCQ

The ISO 27000 series are a framework for?

1. IT Governance
2. IT general controls
3. Data management
4. **Information security management**



# AIS C13 – MCQ

Segregation of duty is

1. Preventive control
2. Detective control
3. Corrective control
4. Personnel general control



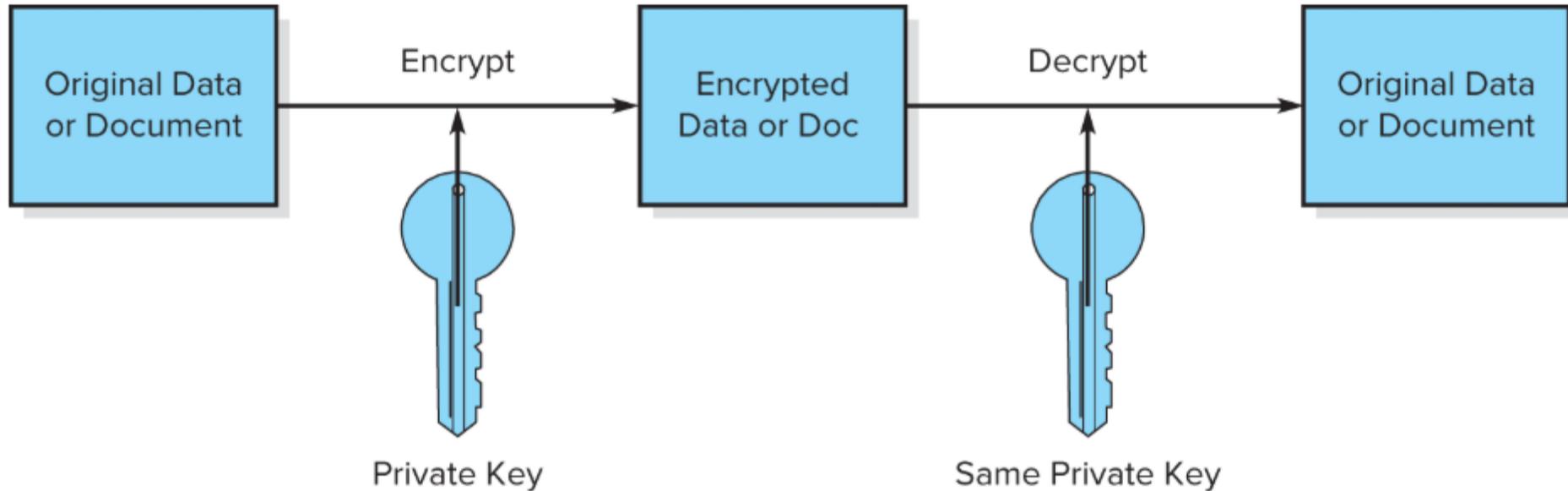
# AIS C14 – IS and Computer Fraud

- **Virus** — a self-replicating program that runs and spreads by modifying other programs or files.
- **Worm** — a self-replicating, self-propagating, self-contained program that uses networking mechanisms to spread itself.
- **Trojan horse** — a non-self-replicating program that seems to have a useful purpose in appearance, but in reality has a different, malicious purpose.
- **Spam** — sending unsolicited bulk information.
- **Botnet (bot)** — a collection of software robots that overruns computers to act automatically in response to the bot-herder's control inputs through the Internet.
- **Denial-of-service (DoS)** — the prevention of authorized access to resources (such as servers) or the delaying of time-critical operations.
- **Spyware** — software that is secretly installed into an information system to gather information on individuals or organizations without their knowledge; a type of malicious code.
- **Spoofing** — sending a network packet that appears to come from a source other than its actual source.
- **Social engineering** — manipulating someone to take certain action that may not be in that person's best interest, such as revealing confidential information or granting access to physical assets, networks, or information.



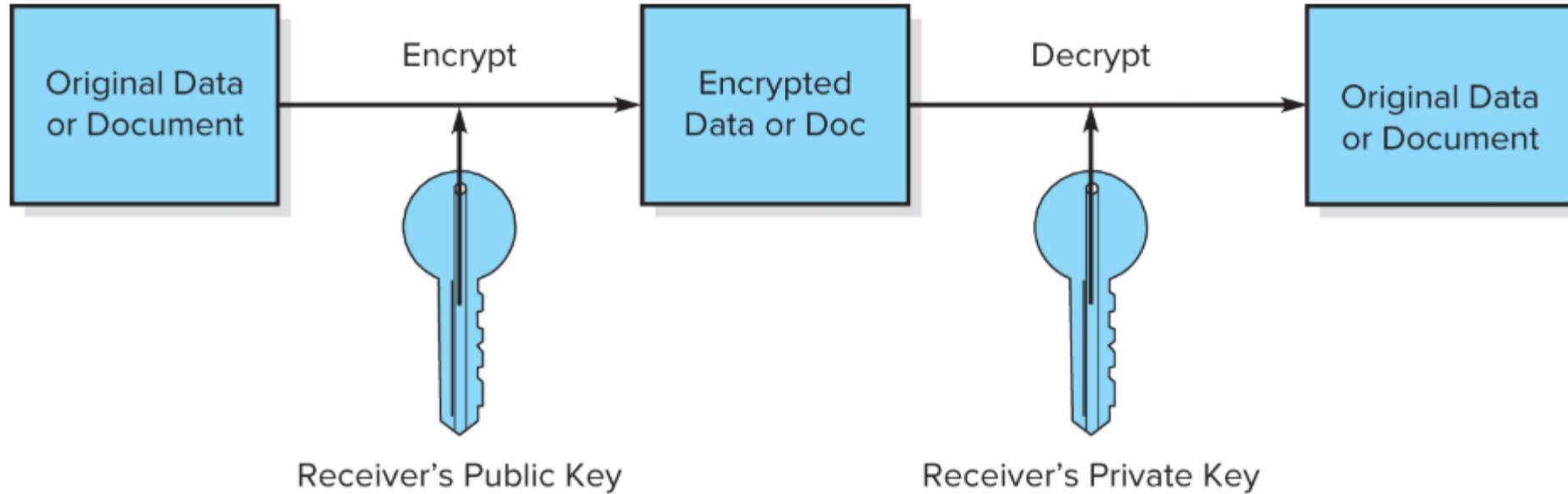
# AIS C14 – Encryption

Encryption Methods – Symmetric, Asymmetric,



# AIS C14 – Encryption

Encryption Methods – Symmetric, Asymmetric,



# AIS C14 – Encryption

## What happens during Chrome Browsing

### The thing that “remembers you” (cookies/tokens)

- “Remember me for 14 days” is almost always a **cookie** (or a similar token like a JWT in storage). It’s a long, random **bearer token** the site recognizes later.
- That cookie is saved **on disk** in Chrome’s profile and sent back to the site on future visits. The site sees the token and logs you in automatically.

### Typical cookie storage locations (Chrome):

- **Windows:** %LOCALAPPDATA%\Google\Chrome\User Data\Default\Network\Cookies
- **macOS:** ~/Library/Application Support/Google/Chrome/Default/Cookies
- **Linux:** ~/.config/google-chrome/Default/Cookies

These are SQLite databases; Chrome encrypts secrets at rest using the OS keystore (e.g., DPAPI on Windows, Keychain on macOS).

### Is that a “private key”?

- **No.** A private key (in the crypto sense) is used for asymmetric cryptography and identity proofs. Normal browsing **does not use a client private key**.
- The 14-day cookie is **not** a private key; it’s just a random secret the website issued to you. If someone steals it, they can impersonate you until it expires—hence protections like Secure



# AIS C14 – Encryption

Cookies Demo



# AIS C14 – Encryption

## Symmetric Key Encryption Demo

- Call two students up
- Reveal the way to do a symmetric encryption and decryption such as a cipher shift
- Get student B to leave classroom, then student A and class comes up with a word.
- Both students A and B know how to encrypt and decrypt

## Asymmetric Key Encryption Demo

- Call two students up
- Reveal the way to do a asymmetric encryption and decryption where both closing and opening are different
- Get student B to leave classroom, then student A and class comes up with a word.
- Student B can decrypt
- Then get student A to leave the room and student B encrypts



# AIS C14 – Encryption

Demonstrate some mathematical operations used to encrypt



# AIS C15 – Monitoring AIS

## Operating System

**OS role:** Core system software that makes the computer run

### **Five control objectives for a reliable OS:**

**Protect itself from users** — prevent user programs from taking over or damaging the OS.

**Protect users from each other** — isolate users so one can't access/corrupt another's data or programs.

**Protect users from themselves** — isolate a user's own modules/memory so one faulty part can't break another.

**Protect itself from itself** — modular OS components shouldn't be able to corrupt/destroy one another.

**Protect from the environment** — handle power failures/disasters with controlled shutdown and recovery.



# AIS C15 – Monitoring AIS

## CAAT approaches

- **Auditing around the computer (black-box):** Auditor predicts results from inputs, compares to system outputs
- **Auditing through the computer (white-box)**
- Common packages: **ACL** and **IDEA**; well-suited for large data files to surface items needing further scrutiny.



# AIS C16 – Balanced Scorecard (not tested)

**Why it matters:** Many firms say IT spending shows little/negative ROI because projects aren't tied to strategy

**What the Balanced Scorecard (BSC) is:** A performance measurement framework that links mission/strategy to concrete objectives and metrics. It's widely used across business, government, and nonprofits.



# AIS C16 – Balanced Scorecard (not tested)

**\*\*Plan & Translate Strategy\*\***: BI, analytics, and dashboards help set strategy and convert it into measurable objectives and budgets

**\*\*Execute & Monitor Operations\*\***: ERP and transaction systems run daily activities while dashboards track performance (KPIs)

**\*\*Learn & Adapt\*\***: BI and analytics provide feedback to refine strategy, supported by collaboration tools across teams



# AIS C18 – System Devt Lifecycle (not tested)

**SDLC (System Development Life Cycle)** is a structured process used to plan, build, test, and maintain information system. It is the **step-by-step method** organizations use to develop systems from start to finish

## ### System Development Life Cycle (SDLC)

- **Planning** – Define problem, objectives, and feasibility
- **Analysis** – Gather requirements and understand user needs
- **Design** – Specify system architecture, data, and processes
- **Development** – Build and configure the system
- **Testing** – Verify the system works correctly and meets requirements
- **Implementation** – Deploy the system and train users
- **Maintenance** – Monitor, fix issues, and update the system over time



# AIS C18 – System Devt Lifecycle (not tested)

## CONTROL, CONTROL CONTROL

Two real excerpts (what went wrong)

### •Landauer Inc. (2014 10-K):

- Lacked processes to **restrict access** to key financial systems and programs.
- **Segregation of duties** was inadequate; certain personnel had access to financial apps and data “beyond” their job without independent monitoring.
- Implication: AIS **design failure**—had access/SOD been designed correctly, the misstatement likely wouldn’t have occurred.

### •WaferGen Bio-Systems Inc. (8-K):

- A **spreadsheet error** led to the **wrong basic/diluted loss per share**, using an incorrect share count (709,639 vs. 1,226,570).
- Implication: **Ineffective spreadsheet controls** within the AIS environment.



# AIS C18 – System Devt Lifecycle (not tested)

## Project Management

**What a project is:** A defined sequence of tasks that produces a specific outcome.

**Project management (PM):** Planning, organizing, supervising, and directing an IT project end-to-end. The **project manager** coordinates all development phases, owns day-to-day success, and works from a **project charter** that states objectives and requirements.



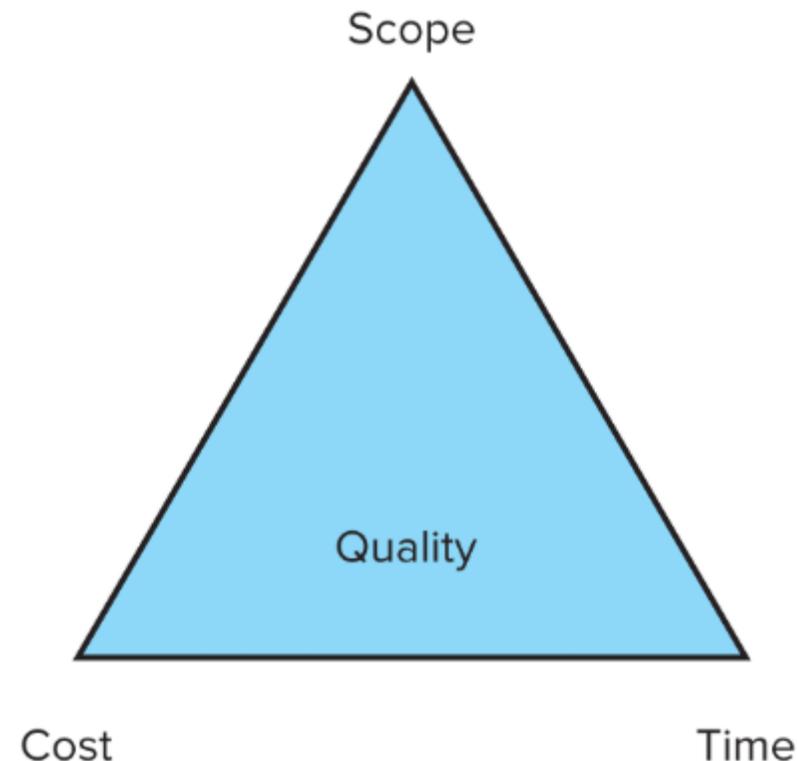
# AIS C18 – System Devt Lifecycle (not tested)

## Project Management Triangle

- Scope (what the project delivers)
- Cost (budget/resources available)
- Time (schedule/deadline)

You cannot change one without affecting the others

**Scope Creep**

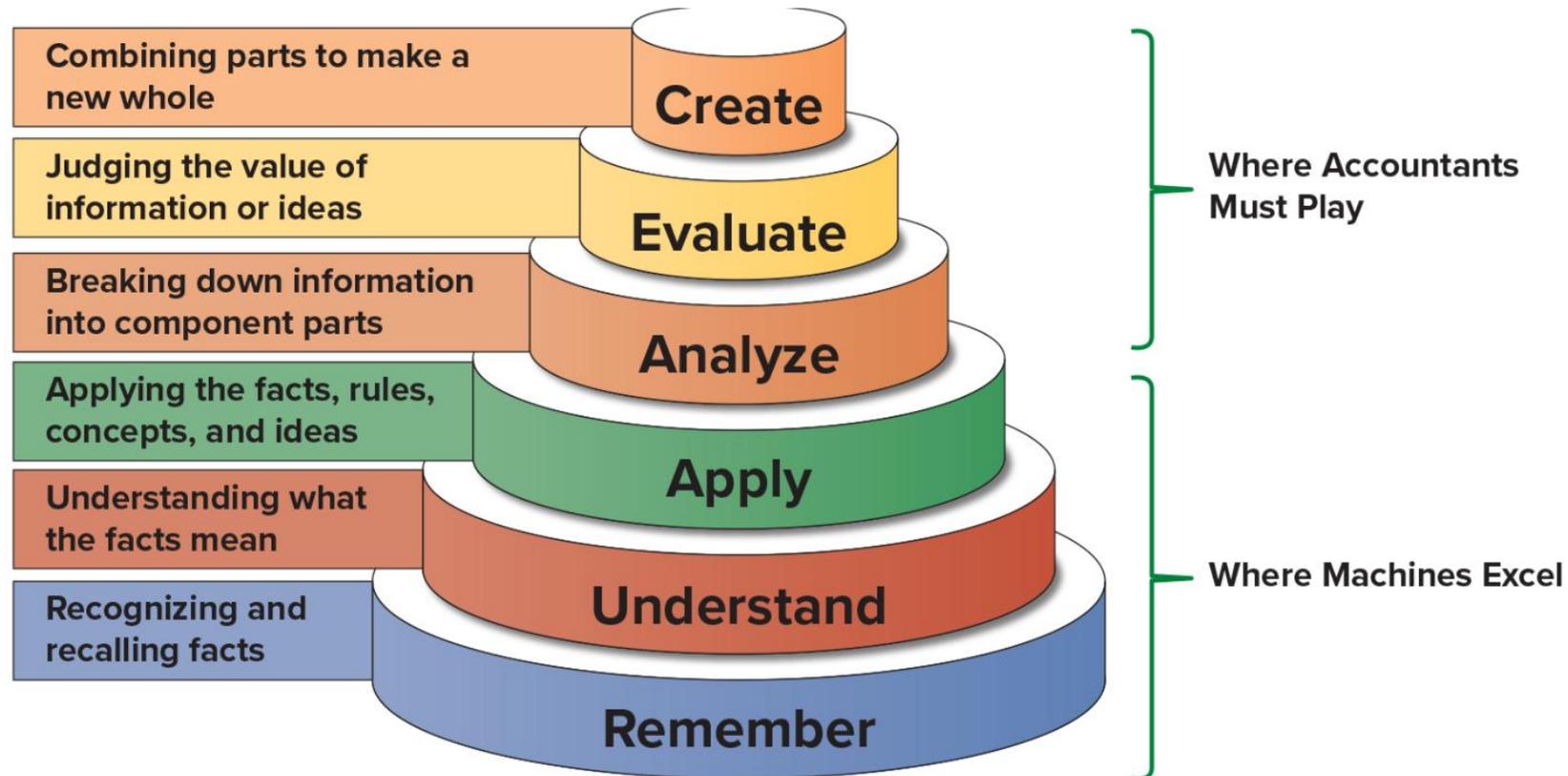


**DA**



# DA C1 – Using DA to answer acting questions

## Bloom's Taxonomy



# DA C1 – Using DA to answer acting questions

- AMPS – ask the right questions | extract, transform and load relevant data | apply appropriate data analytic techniques | interpret and share results with stake holders
  - Ask the question
  - Master the data
  - Perform the analysis
  - Share the story



# DA C1 – Using DA to answer acting questions

## Ask the question

- Did we make a profit last year?
- Why did sales, general and administrative expenses increase relative to industry?

## Master the data

- Does the data have high data integrity?

## Perform the analysis

- Data visualization, regression, pivot table, cash flow analysis, renford's law, what-if/goal seek analytics

## Share the story

- Dashboard or visualization



# Many ways to represent data



# DA C1 – Using DA to answer acting questions

## Data Acquisition and Preparation Tools

- Excel, SQL, Alteryx

## Code Tools

- SAS, SPSS, R

## Data Visualization Tools

- Excel, tableau, powerBI, pivotable



# DA C1 – MCQ

What is the appropriate ordering of critical thinking skills in Bloom's Taxonomy, where the higher order skill is to the left of the ">" symbol?

1. Remember > Apply
2. Apply > Analyze
3. Analyze > Evaluate
4. Create > Analyze



# DA C1 – MCQ

What visualization type is most appropriate for evaluating the relationships between values?

1. Bar Chart
2. Pie Chart
3. Histogram
4. Scatterplot



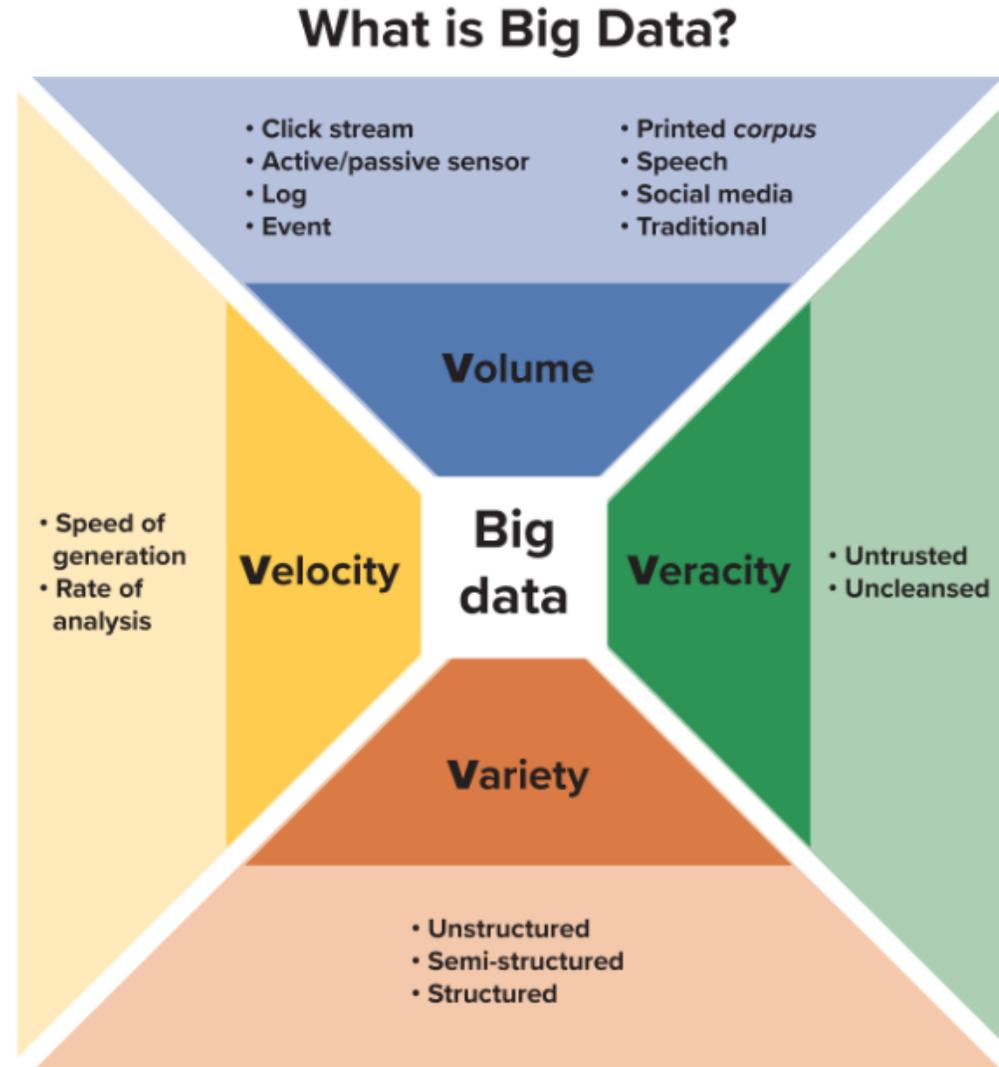
# DA C1 – MCQ

TABLEAU DEMO WITH VIDEO



# DA C2 – Master the data – introduction to accounting data

What is big data?



# DA C2 – Master the data – introduction to accounting data

## What accounting data is available?

- Financial accounting data
- Financial Accounting Related Data – risk factors, MDNA, Financial Statements and supplemental data
  - XBRL – extensible business reporting language
    - Format for presenting financial statements
- Managerial Accounting Data
- Tax Data



# DA C2 – Master the data – introduction to accounting data

## Non-accounting data sources

- Economic data – CPI, GDP, unemployment
- Current and Historical stock prices
- Social Media
- General Chat Sites
- Analyst Research Reports



# DA C2 – Master the data – introduction to accounting data

## Data Ethics

Moral responsibility associated with gathering and protecting data, and an understanding of how much responsibility affects individuals

Gathering data – does company send privacy notice?

Protecting data – is customer's credit card number protected?



# DA C2 – Master the data – introduction to accounting data

## PivotTable

- Tool that allows reorganization of certain data using cross tabulations without changing the data



# DA C2 – MCQ

What would most likely be disclosed in an 8-K filing to the SEC?

1. Financial statements
2. Change in auditor
3. Management discussion and analysis
4. Overview of a company's main operations



# DA C2 – MCQ

What is computer based standard to define and exchange financial information from financial statements?

1. SEC edgar
2. XBRL
3. Earnings call
4. Customer relationship management



# DA C3 – Master the data – types used in accounting

## HOW TO CATEGORIZE DATA



# DA C3 – Master the data – types used in accounting

- **Categorical (Qualitative)** – Groups items into categories
  - \*Examples:\* Gender (male, female, nonbinary), Product type (electronics, clothing), Payment method (cash, credit)
- **Nominal Data** – Categories with no order (just labels/counting)
  - \*Examples:\* Customer ID, Zip code, Brand name, Product category
- **Ordinal Data** – Categories with a meaningful order (ranking)
  - \*Examples:\* Customer satisfaction (poor, fair, good, excellent), Credit rating (AAA, AA, A), Class rank
- **Numerical (Quantitative)** – Data expressed as numbers
  - \*Examples:\* Price, Quantity sold, Revenue, Age
- **Interval Data** – Numerical data with equal spacing, but no true zero
  - \*Examples:\* Temperature (°C, °F), Calendar years, Time of day
- **Ratio Data** – Numerical data with a true zero (ratios make sense)
  - \*Examples:\* Sales revenue, Number of units sold, Weight, Distance, Inventory quantity

## ### Data Type Pairings

- **Categorical**
  - Nominal
  - Ordinal
- **Numerical**
  - Interval
  - Ratio



# DA C3 – Master the data – types used in accounting

## ### Additional Ways to Categorize Data

- **String / Alphanumeric** – Text data (letters, numbers, symbols)  
**Examples:** Customer name, Product ID, Email address
- **Date** – Time-based data  
**Examples:** Order date, Invoice date, Timestamp
- **Number** – Numeric values  
**Examples:** Price, Quantity, Revenue
- **Geographic** – Location-based data  
**Examples:** City, State, Country, ZIP code
- **Dimension** – Descriptive fields used to categorize data  
**Examples:** Product category, Customer segment, Region
- **Measure** – Quantitative values used for analysis  
**Examples:** Sales amount, Profit, Units sold



# DA C3 – Master the data – types used in accounting



# DA C3 – Master the data – types used in accounting

## Data Dictionary

- Centralized repository of information about a set of data

| Field            | DataType   | Definition                                                             | Example  |
|------------------|------------|------------------------------------------------------------------------|----------|
| Transaction_ID   | Short Text | Unique identifier for each transaction                                 | 1        |
| Date             | Date       | The date the transaction was initiated                                 | 1/1/2025 |
| Transaction_Type | Short Text | Indicates which type of transaction was made (Return or Sale)          | Return   |
| Amount           | Number     | Entire transaction amount, negative for returns and positive for sales | -27.32   |



# DA C3 – MCQ

Data considering the text in blogs would generally be considered

1. Structured
2. Interval
3. Unstructured
4. Ratio



# DA C3 – MCQ

Gold, silver and bronze medals in the Olympics would be a form of

1. Ratio data
2. Ordinal data
3. Numerical data
4. Interval data



# DA C3 – MCQ

If there are 10 return transactions and 40 sales transactions, what is proportion of return transactions?

1. 10/50
2. 10/40
3. 40/10
4. 50/10



# DA C4 - Master the data - Prepping data

## Golden Mantra - Extract Transform Load (ETL)

| <u>Problem (Without ETL)</u> | <u>Solution (With ETL)</u>       |
|------------------------------|----------------------------------|
| Data entry errors            | Preventive controls & validation |
| Unauthorized changes         | Access control & security        |
| Duplicate/inconsistent data  | Centralized database             |
| Multiple versions of files   | Version control & single source  |



# DA C4 – Master the data – Prepping data

What are the differences among DB, Excel and Tableau

## Tool

Database (e.g., SQL, Oracle)

Excel

Tableau / Power BI

## What it is

A system for storing structured data

A spreadsheet tool

Data visualization tools

## Primary Purpose

Store, manage, and query large datasets

Analyze, manipulate, and do calculations on data

Create dashboards and visual insights



# DA C4 – Master the data – Prepping data

## What is the format of a DB? Relational Database

- Organizes data into separate tables
- Key components:
  - Tables – collections of related data
  - Fields – columns (data attributes)
  - Records – rows (individual entries)
- Supported by:
  - Data dictionary (defines data elements)
  - Database diagram (shows relationships between tables)
  - Data Integrity (Truth in Data)



# DA C5 – Perform Analysis – Types of DA

## Philosophy 2 - 4 types of accounting data analytics

- Descriptive analytics – what happened?
- Diagnostic analytics – why did it happen?
- Predictive analytics – will it happen in the future?
- Prescriptive analytics – what should we do?



# DA C5 – Perform Analysis – Types of DA

What is DA? Another definition: finding previously unknown linkages

- Anomaly Detection
- Benford's law (do the digits class exercise – get 20 students to propose 4 digit combinations)



# DA C5 – Perform Analysis – Types of DA

## Techniques for Identifying Anomalies/Outliers

| Technique                                                                                             | Example of Accounting Question It Addresses                                                                                                   | Specific Application of Technique                                                                                                |
|-------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------|
| Internal controls testing                                                                             | Why were some large cash disbursements not approved by senior management, and do they represent a violation of a company's internal controls? | Evaluate internal controls for cash disbursements and check for appropriate authorization.                                       |
| Fuzzy matching (test looking for similar but not exact matches, to be explained further in Chapter 7) | Why are some addresses of approved vendors similar to the addresses of those employees that work at the company?                              | Check addresses of employees and vendors and see if they are similar.                                                            |
| Sequence checks and sequence analysis                                                                 | Are any checks or purchase orders missing? Does it signify errors or fraud or can they be explained?                                          | Check to see if any checks or purchase orders are missing by running a sequence check to make sure all checks are accounted for. |
| Duplicate transactions                                                                                | Why are there duplicates of some transactions in the financial reporting records? Are they fraud or just errors?                              | Check all amounts, dates, and vendors to see if duplicate transactions were erroneously recorded.                                |
| Benford's Law                                                                                         | Is the distribution of the first digit of a dataset different than expected?                                                                  | Evaluate the first digit of refunds given to see if refunds conform with Benford's law, or require further investigation.        |
| Variance analysis (typically performed in management accounting)                                      | Why is the labor rate and labor use variance for direct labor at the manufacturing plant unfavorable?                                         | Using standard costs, evaluate the labor use and labor rate and see if it is different than we predicted.                        |



# DA C5 – Perform Analysis – Types of DA

| Predictive Analytics Technique                                                                               | Example of Type of Question It Addresses                                         | Type of Test                                                                                                          |
|--------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------|
| <b>Classification:</b> separate or classify a sample (or population) into two or more groups of classes.     | Which companies do we predict will go bankrupt versus those we predict will not? | Run classification analysis to predict which companies will go bankrupt and which will not.                           |
| <b>Regression:</b> predict a specific dependent variable outcome value based on independent variable inputs. | Can we predict the interest rate given to prospective loan borrowers at a bank?  | Run regression analysis to predict the interest rate given based on prospective loan borrower's characteristics.      |
| <b>Time series analysis:</b> predict future values based on past values of the same variable.                | Can we predict next year's sales (or earnings)?                                  | Use time series analysis incorporating past values of sales (or earnings) to predict next year's sales (or earnings). |
| <b>Neural Network:</b> predict any value that we want, through a black box                                   |                                                                                  |                                                                                                                       |



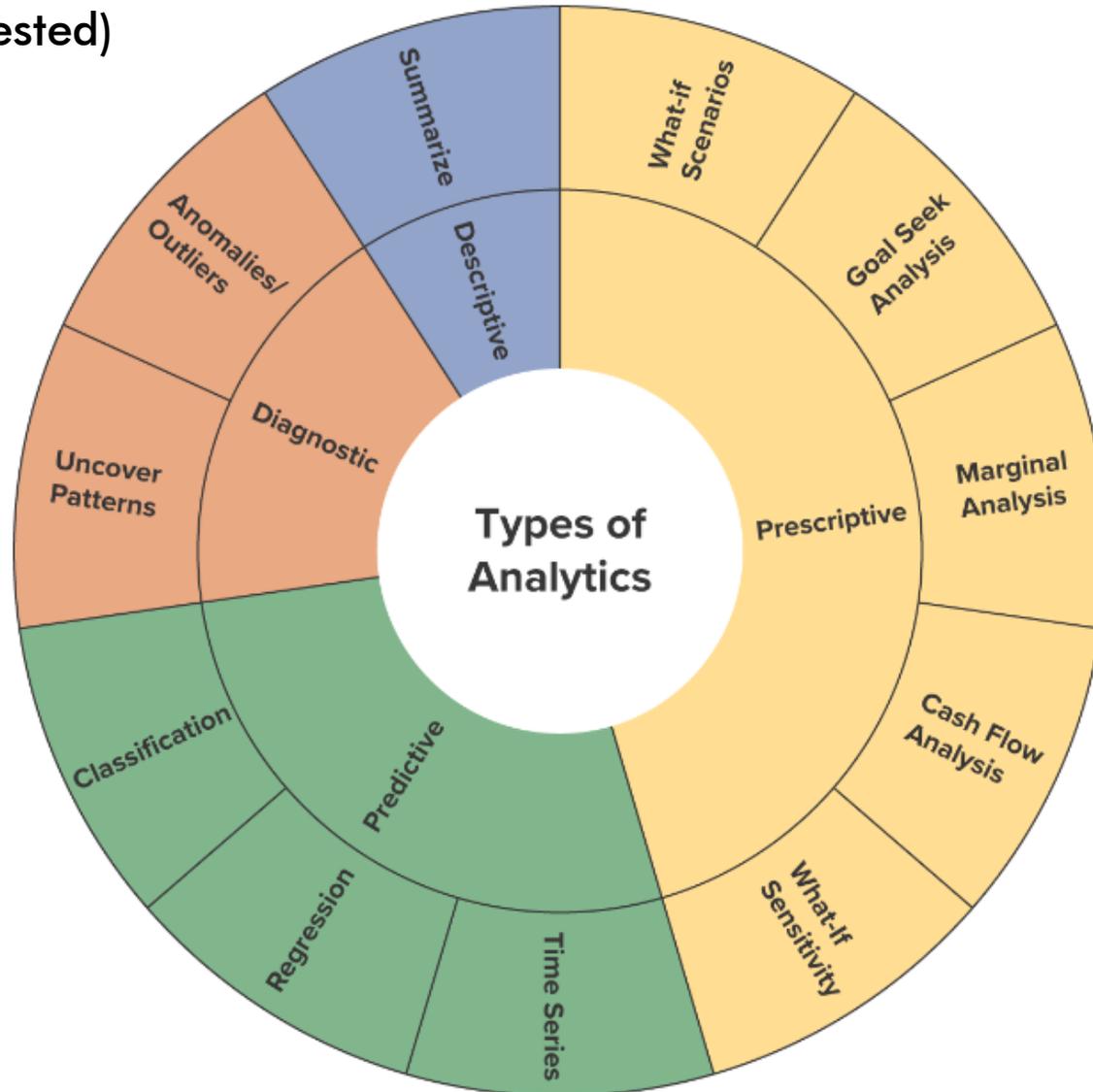
# DA C5 – Perform Analysis – Types of DA

| Prescriptive Analytics Techniques                                                                                                                                                                                | Example of Accounting Question Technique Addresses                                                              | Specific Application of Technique                                                                                                                                                               |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Marginal/incremental analysis:</b> What-if analysis that determines the change in profit associated typically with the cost or benefit of the next (or marginal) unit.                                        | Should Tesla make or outsource the production of its batteries?                                                 | Estimate the implications of two scenarios: make batteries or buy batteries from another party as input to making the decision on whether to make or buy batteries.                             |
| <b>Cash flow analysis (sometimes called capital budgeting):</b> Evaluates future cash flows for potential investments or expenditures, typically using net present value (NPV) or internal rate of return (IRR). | Which assets should we invest in at Boeing, a new piece of equipment or a new computer system?                  | Estimate the amounts, timing, and risk associated with the predicted cash flows associated with investment in the new equipment using NPV and IRR capital budgeting techniques.                 |
| <b>Scenario analysis:</b> What-if analysis that examines the effect of changing multiple inputs at the same time on an outcome variable.                                                                         | How will Walmart respond to various scenarios of increased tariffs with China?                                  | Analyze all of the potential scenarios to arrive at an overall potential impact of the future events.                                                                                           |
| <b>Goal-seek analysis:</b> What-if analysis that determines the required input value(s) to get a desired output (i.e., backwards analysis).                                                                      | For McGraw-Hill, what are the needed sales to break even on sales of this text?                                 | Perform breakeven analysis to figure out the needed level of sales.                                                                                                                             |
| <b>Sensitivity analysis:</b> What-if analysis that examines the effect of changing one input on an outcome variable.                                                                                             | What happens to Bank of America profits if the interest rates change in the future from 3% to 4% or 5% or more? | Test to see how sensitive Bank of America's profits are to interest rate changes based on our predictions.                                                                                      |
| <b>Optimization:</b> Analysis of the best feasible solution based on an objective function, decision variables, and business constraints. (See Chapter 12 for more discussion).                                  | How does Netflix price their streaming service to maximize profits?                                             | Analysis of how profitability changes if the price of streaming service changes. As price changes, some customers will change streaming service providers or drop streaming service altogether. |



# DA C5 – Perform Analysis – Types of DA

A taxonomy wheel (not tested)



# DA C5 – Perform Analysis – Types of DA

Many of the technique we will cover are intertwined.



# DA C6 – Perform the analysis – Descriptive

Descriptive Analytics – What happened?

- Summary Statistics
- Data Aggregation
- Data Visualization

Dupont analysis / ROE Analysis

$$\text{ROE} = \frac{\text{Net Income}}{\text{Sales}} \times \frac{\text{Sales}}{\text{Average Total Assets}} \times \frac{\text{Average Total Assets}}{\text{Average Shareholders' Equity}}$$

Which simplifies to:

$$\text{ROE} = \frac{\text{Net Income}}{\text{Average Shareholders' Equity}}$$



# DA C7 – Perform the Analysis - Diagnostic

Diagnostic Analytics – performed to investigate the underlying reasons for past results that cannot be answered by simply looking at the descriptive data

- Hypothesis Testing
- Sequence Checks
- Drill-down Analysis



# DA C8 – Perform the Analysis - Predictive

Predictive Analytics – identifying patterns in historical data and using them to predict future outcomes.

- Classification
- Regression
- Time Series
- Hypothesis Testing
- Machine Learning



# DA C9 – Perform the Analysis - Prescriptive

Prescriptive analytics – identify the best possible options given constraints or changing conditions

- Constraints
- Changing conditions
- What-if analysis
- Marginal analysis
- Net present value
- Internal Rate of Return
- Goal-seek analysis
- sensitivity analysis
- Optimization
- Break-even Analysis



# DA C10 – Share the story

Most important part and this boils down to UI, Presentation Skills

## SUMMARY

- Data analytics techniques are methods used to turn raw data into useful information for decision-making.
- Bread-and-butter techniques
  - Regression
  - Summary Statistics
  - Classification/Neural Network/Machine Learning



# DA (Practical)



# Excel Functions we are going to cover

- VLOOKUP
- LEFT/RIGHT/CONCAT
- IF CONDITION
- YEAR/DAY/MONTH
- ISBLANK/ISNUMBER/ISTEXT
- NPV/IRR/PMT/RATE

WHAT-IF SCENARIO



# DA Techniques

Regression

Cluster Analysis

CART

Data Mining

Neural Network

Association Rules

And many more...



# DA Techniques that we will cover

Regression

CART

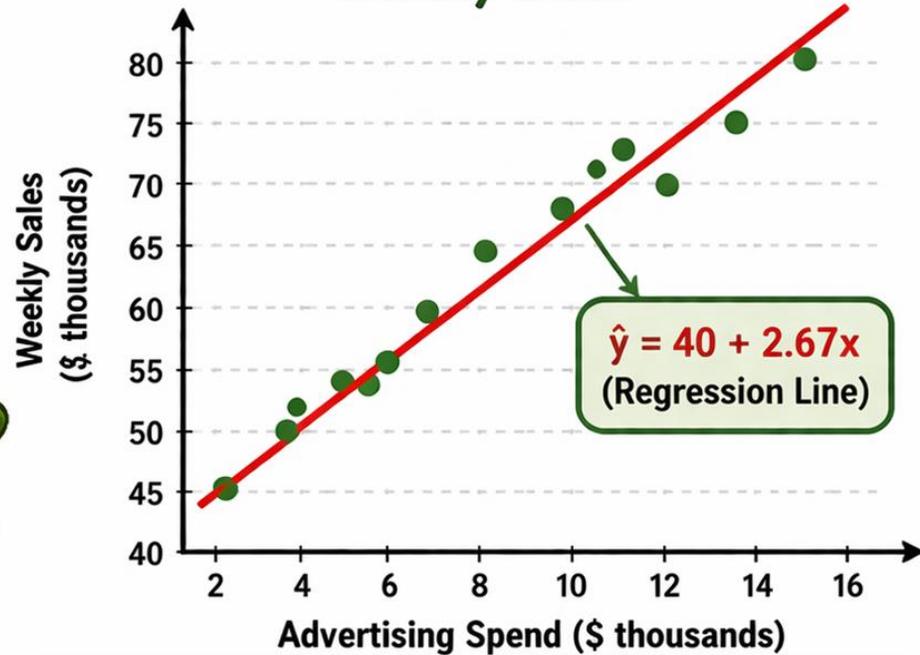
Neural Network



**REGRESSION:**  
Does more advertising lead to higher burger sales?



### Burger Ad Spend vs. Weekly Sales

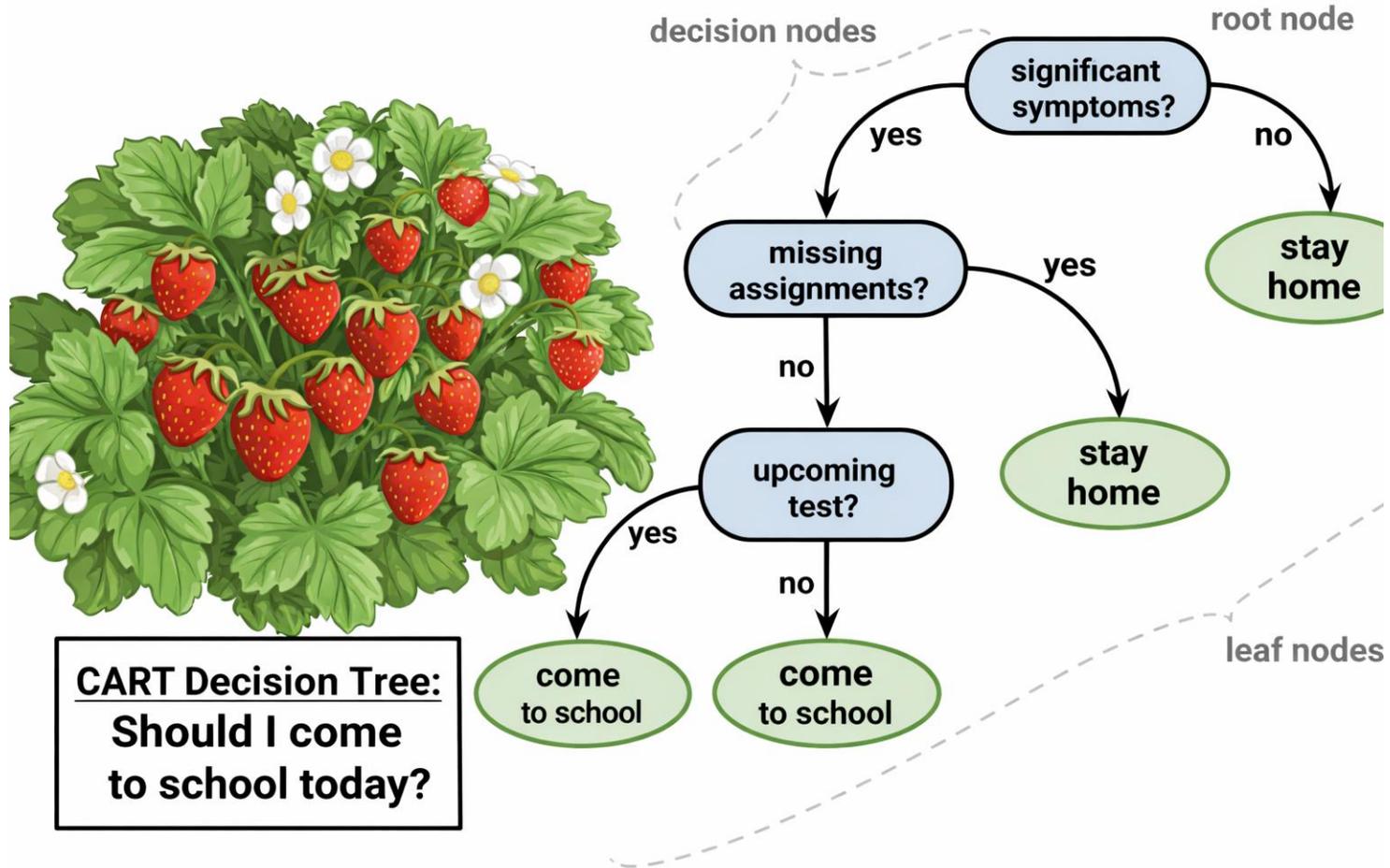


#### WHAT THE REGRESSION SHOWS:

- Slope (**2.67**): For every additional \$1,000 spent on ads, weekly sales increase by about **\$2,670**.
- Intercept (**40**): Estimated sales are **\$40,000** when ad spend is \$0.
- Strong positive relationship.



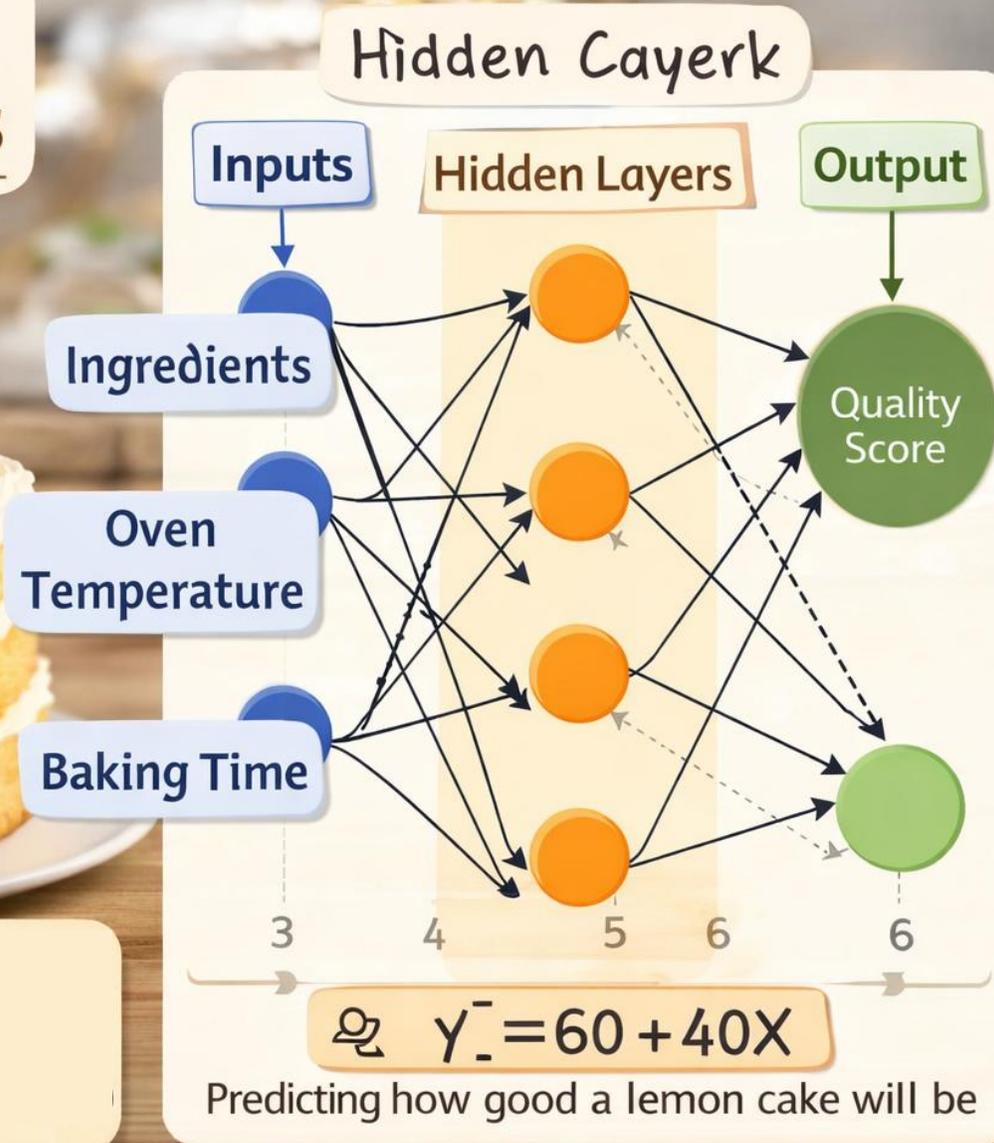
# CART



# Lemon Cake & Neural Networks



Lemon Cake Quality



# DA Techniques that we will cover

- Use *regression* when you want to predict a number.
- Use *classification* when you want to assign something to a group.
- Use a *neural network* when the pattern is too complex for simpler methods (however, for smaller datasets, may be overfitting)



# DA Techniques that we will cover

- Introduction to R (datasets – iris, beer and diapers)
- Introduction to Python



# DA Techniques that we will cover

| Burger ID | Tomato Sauce (grams) | Price (\$) | Burgers Sold |
|-----------|----------------------|------------|--------------|
| 1         | 10                   | 5          | 120          |
| 2         | 15                   | 6          | 140          |
| 3         | 20                   | 7          | 160          |
| 4         | 25                   | 8          | 180          |
| 5         | 30                   | 9          | 200          |
| 6         | 12                   | 5          | 130          |
| 7         | 18                   | 6          | 150          |
| 8         | 22                   | 7          | 170          |

Burger Dataset

