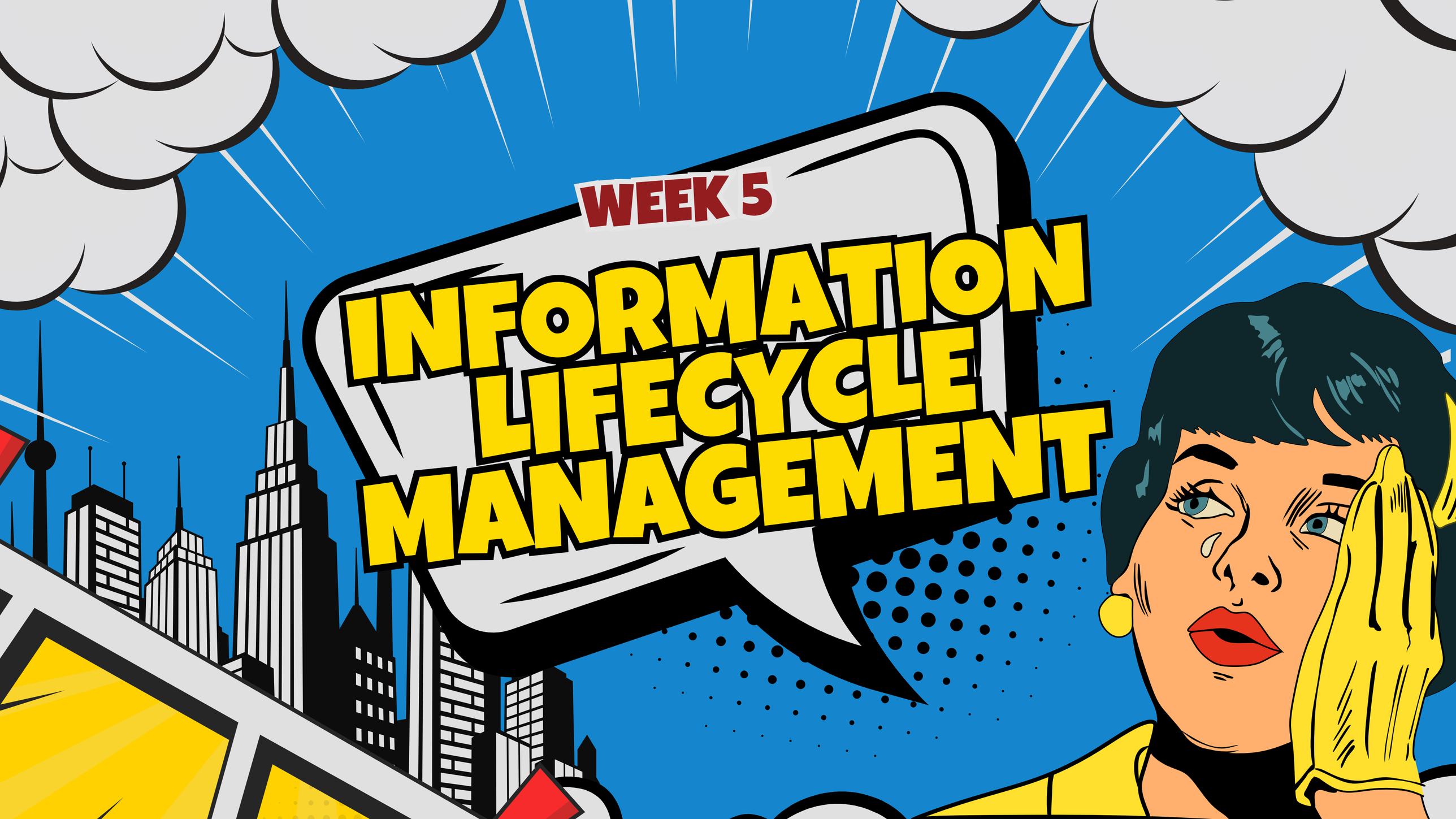
WEEK 5

INFORMATION LIFECYCLE MANAGEMENT



INTRODUCTION

What is Information Lifecycle Management?

- Information Lifecycle Management is a comprehensive approach to managing data and information throughout its entire existence - from creation to final disposal. It's a strategic methodology that ensures information is properly handled, stored, protected, and disposed of according to its value and regulatory requirements.

STAGES OF THE INFORMATION LIFECYCLE

- CREATION
- STORAGE
- USE
- ARCHIVAL
- DISPOSAL

STAGES OF THE INFORMATION LIFECYCLE

Creation

Where data is initially generated from various sources such as employees, customers, machines, or systems. This can include structured data like spreadsheets and databases, or unstructured data like emails, images, and documents.

Example: A student writes a research paper or a company creates a financial report.

STAGES OF THE INFORMATION LIFECYCLE

Storage

Once created, the data moves into the Storage stage, where it is securely kept for future access. Storage may be physical, such as paper files in filing cabinets, or digital, such as databases, servers, or cloud systems. Proper storage ensures data is protected, organized, and available when needed.

Example: Saving the research paper on a computer hard drive, cloud storage, or filing it in a cabinet.

STAGES OF THE INFORMATION LIFECYCLE

Use

The next stage is Use, where data is actively accessed, processed, or shared to support daily operations and decision-making. At this stage, raw data is transformed into meaningful information through reporting, analysis, or direct application in tasks.

Example: The teacher reads the paper to grade it, or employees use the financial report for decision-making.

STAGES OF THE INFORMATION LIFECYCLE

Archival

As data ages and is no longer needed for daily operations, it enters the Archival stage. Here, it is moved to long-term storage, where it remains available for legal, regulatory, or historical purposes but does not burden active systems. Archived data is often stored in cheaper storage solutions optimized for long-term retention.

Example: After the semester, the research paper is stored in the school's database for record-keeping, or the company files the report for auditing purposes.

STAGES OF THE INFORMATION LIFECYCLE

Disposal

Finally, information reaches the Disposal stage, where it is securely destroyed once it is no longer useful or legally required. Secure disposal prevents unauthorized access, data breaches, and compliance violations. This may include shredding physical documents, wiping or encrypting digital files, or physically destroying storage devices.

Example: Deleting old drafts from the computer, shredding outdated printed reports, or permanently erasing digital files no longer needed.

DATA RETENTION AND ARCHIVING POLICIES

Legal Requirements

Governments require specific retention periods (e.g., tax records for 7 years).

- Tax Records: Must be kept for 7 years in the U.S. (IRS rule).
- Employment Records: Some countries require employee contracts to be stored for 6 years after employment ends.
- Contracts: Legal agreements may need retention for a statutory limitation period (e.g., 6 years in the UK).

Regulatory Requirements

- Healthcare: HIPAA requires patient records to be retained for a defined period.
- Finance: SEC mandates retention of trading records.

Pharmaceutical Industry (FDA, U.S.): Clinical trial records must be retained for at least 2 years after a drug is approved.

Business Requirements

Keeping data for performance reviews, strategic planning, and audits.

- A company keeps old sales data to check what products sell best.
- A company keeps customer feedback to improve future products.

MANAGING DATA THROUGH ITS LIFECYCLE

Document Management Systems (DMS)

- Organize, track, and manage documents in digital form.
- Provide version control and easy retrieval of files.
- Reduce reliance on paper records.

Cloud Storage with Lifecycle Management Rules

- Automatically moves inactive files to cheaper, long-term storage.
- Ensures important data remains accessible while lowering costs.
- Offers scalability and flexibility for growing organizations.

MANAGING DATA THROUGH ITS LIFECYCLE

Document Management Systems (DMS)

EXAMPLE:

- A school uses a DMS to store students' grades, assignments, and attendance records digitally instead of keeping piles of paper folders.
- Employees in a law firm can quickly search and retrieve contracts without digging through filing cabinets.

Cloud Storage with Lifecycle Management Rules

EXAMPLE:

- A company stores all its project files in Google Drive or AWS. Files not opened in 1 year automatically move to a cheaper "archive storage" but can still be retrieved when needed.
- A photography studio keeps recent client photos in fast-access cloud storage, but after 2 years, old photos are shifted to low-cost cloud storage to save money.

MANAGING DATA THROUGH ITS LIFECYCLE

Encryption and Access Controls

- Protect sensitive information from unauthorized access.
- Encryption ensures data remains unreadable without the correct key.
- Access controls (like passwords, roles, and multi-factor authentication) allow only authorized staff to view or edit data.

Backup and Disaster Recovery Systems

- Create regular copies of critical data to avoid loss.
- Ensure quick recovery after system failures, cyberattacks, or natural disasters.
- Provide business continuity and reduce downtime.

MANAGING DATA THROUGH ITS LIFECYCLE

Encryption and Access Controls

EXAMPLE:

- A bank encrypts customer account numbers, so even if hackers steal the database, the information looks like random symbols without the decryption key.
- At a school, only teachers with a password and 2-step verification can access students' grades.

Backup and Disaster Recovery Systems

EXAMPLE:

- A hospital creates daily backups of patient medical records. If a cyberattack or system crash happens, they can quickly restore the files and continue treating patients.
- A company stores copies of its business files in the cloud. If the office computer servers are destroyed by fire, they can still recover all data from the cloud and continue operations.

CHALLENGES IN MANAGING THE INFORMATION LIFECYCLE

Complexity

Volume and variety of data are growing rapidly. Multiple data sources and formats increase management difficulties.

EXAMPLE:

A student has photos, videos, notes, and assignments saved on their phone, laptop, and Google Drive. Since the files are in different places and formats, it's hard to keep everything organized.

Cost

Managing the information lifecycle requires significant investment in storage infrastructure, advanced data management tools, compliance software, and skilled IT professionals.

EXAMPLE:

The student runs out of free Google Drive space and has to start paying for extra storage every month. That's an added cost just to manage their files.

CHALLENGES IN MANAGING THE INFORMATION LIFECYCLE

Security Concerns

Data is constantly under threat from cyberattacks such as ransomware, phishing, and hacking attempts, which can result in severe financial and reputational damage.

Example: If someone hacks a school's system and leaks student records, the school loses trust.

Compliance Risks

Failing to comply can result in legal penalties, reputational damage, and loss of customer trust.

Example: If a hospital doesn't keep patient records properly, it can face big legal problems.


THANK
YOU

Information Management

Home Content Contact

DATA QUALITY

MANAGEMENT

Start

Information Management

Home Content Contact

Definition of Data Quality Management (DQM)

Data Quality Management is the process of making sure that data is correct, complete, consistent, and updated so it can be trusted and used effectively. It focuses on keeping information reliable by setting rules, checking for errors, and improving data over time.

Ex: A bank uses DQM to ensure customer account details (like names, balances, and transactions) are accurate and up to date.

DIMENSIONS OF DATA QUALITY

Accuracy – Data should correctly represent real-world values.

Completeness – No missing or incomplete values.

Consistency – Data should be uniform across systems and databases.

Timeliness – Data must be up-to-date and available when needed.

Data Quality Frameworks and Standards

ISO 8000

- International Organization for Standardization. ISO 8000 is a global standard for data quality.

Focus: Ensures data is accurate, complete, and exchangeable across systems.

Data Quality Frameworks and Standards

DAMA-DMBOK

- DAMA = Data Management Association International; DMBOK = Data Management Body of Knowledge.

Purpose: Helps organizations set governance rules, improve quality, and ensure secure and consistent data use.

Assessing and Measuring Data Quality

Techniques:

- **Validation Checks** – Ensure inputs meet rules (e.g., email format).
- **Data Profiling** – Analyzing datasets to spot inconsistencies. Audits – Independent reviews to check compliance.

Improving Data Quality

Techniques & Tools:

- **Data Cleaning –** Removing errors, duplicates, or wrong entries.

Example: A university cleans student records to merge duplicate IDs.

- **Standardization** – Converting data into a uniform format.

Example: An airline standardizes date formats (MM/DD/YYYY) across all booking systems.

Improving Data Quality

Techniques & Tools:

- **Deduplication –** Detecting and removing duplicate data.

Example: Amazon merges duplicate product listings to avoid confusion for buyers.

- **Data Enrichment –** Adding external data to improve value.

Example: Banks enrich customer profiles with updated credit scores from external agencies.

Governance Strategies

Strategies:

- **Clear Policies –** Rules for data handling and usage.

Example: Facebook sets strict policies on user data access for employees.

- **Regular Monitoring –** Continuous checks on data quality.

Example: Banks monitor transaction data daily to detect unusual activity.

Governance Strategies

Strategies:

- **Accountability Structures –** Assigning responsibility to teams/roles.

Example: In government agencies, specific offices are made accountable for maintaining citizen databases.

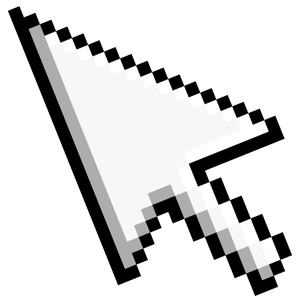
- **Training & Awareness –** Educating staff on the importance of quality.

Example: Airlines train staff to input passenger data correctly to avoid boarding delays.

Information Management

Home Content Contact

THANK YOU!



MASTER DATA MANAGEMENT

&

REFERENCE DATA MANAGEMENT

GROUP 7

WHAT IS DATA?

Data

raw facts and once it is organize or processed it called information

Problem

Data can be duplicated, inconsistent, or incomplete.

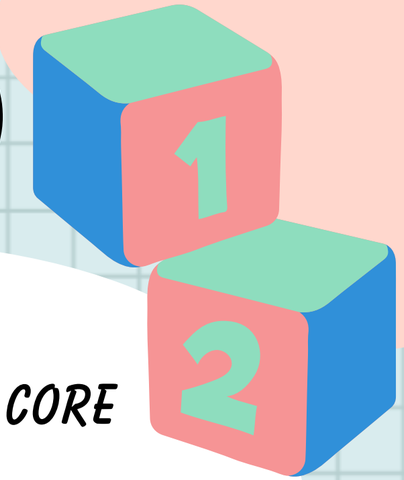
Solution

MDM

and

RDM

MASTER DATA MANAGEMENT (MDM)



MDM IS THE PRACTICE OF CREATING ONE TRUSTED VERSION OF THE TRUTH FOR CORE BUSINESS DATA.

EXAMPLE:

IF THE SAME STUDENT'S NAME APPEARS DIFFERENTLY IN EACH SYSTEM (E.G., JOHN PAUL CRUZ VS. J. P. CRUZ), MDM FIXES THIS.

MDM CREATES ONE SINGLE, CORRECT VERSION OF THAT STUDENT RECORD ACROSS ALL SYSTEMS.

IMPORTANCE OF MDM

- Removes duplication.
- Improves reporting and analytics.
- Boosts efficiency.

REFERENCE DATA MANAGEMENT (RDM)

RDM MANAGES LISTS, CODES, AND CATEGORIES TO ENSURE ALL SYSTEMS USE THE SAME STANDARD.

EXAMPLE:

- ADMISSIONS → PHP
- CANTEEN → PH PESO
- ACCOUNTING → ₱
- 👉 RDM STANDARDIZES INTO PHP (PHILIPPINE PESO).

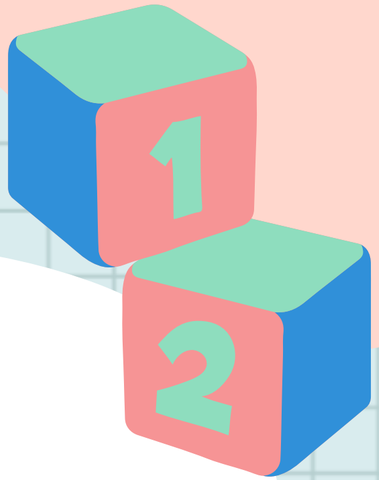
ROLE OF RDM IN INTEGRATION

RDM ensures that when systems exchange data, they are aligned and not "lost in translation."

1. STANDARDIZATION ACROSS SYSTEMS

* Ensures that codes, categories, and values mean the same thing in every system.
* Example: "M" = Male, "F" = Female across HR, CRM, and Finance.

ROLE OF RDM IN INTEGRATION



2. DATA VALIDATION IN INTEGRATION

* Prevents wrong values from entering systems during integration.
* Example: If a new vendor has currency = "PHL," RDM rejects it since "PHL" is not a valid currency code.

3. ENABLES INTEROPERABILITY

* Different applications can talk to each other smoothly because they use the same reference values.

ROLE OF RDM IN REPORTING

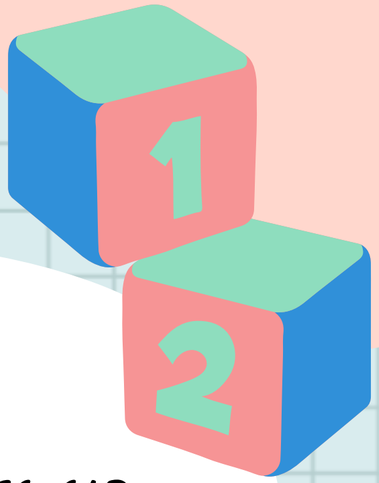
1. ACCURATE AGGREGATION

* Combines data from multiple sources without duplication or mismatch.
* Example: "US," "USA," and "United States" are all reported as one country.

2. IMPROVED DECISION-MAKING

* Executives rely on trusted, standardized data → not confusing or conflicting numbers.

ROLE OF RDM IN REPORTING



3. HISTORICAL CONSISTENCY

* Even when categories change (like product lines), RDM maintains a history so old reports remain consistent.

* Example

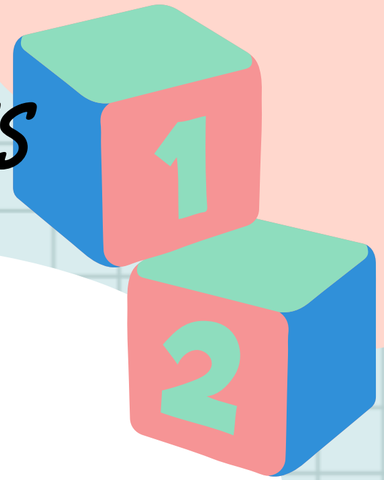
Category_ID: 101

Name: Smartphones

Aliases: Mobile Phones, Cell Phones

MDM ARCHITECTURE & PROCESSES

KEY COMPONENTS OF MDM ARCHITECTURE & PROCESSES



1. DATA SOURCES

Definition: The original systems where data is created and stored.

- Examples: ERP (Enterprise Resource Planning), CRM (Customer Relationship Management), HR, Finance, spreadsheets.

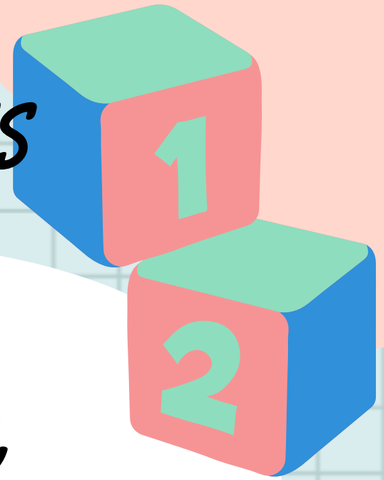
KEY COMPONENTS OF MDM ARCHITECTURE & PROCESSES

2.INTEGRATION LAYER

Definition: Tools and processes (like ETL – Extract, Transform, Load) that move data from source systems to the MDM hub.

- *Why Important: Different systems store data in different formats. The integration layer converts them into a common format.
- Analogy: Like a translator that makes sure all "languages" from HR, Finance, and CRM can be understood in one place.

KEY COMPONENTS OF MDM ARCHITECTURE & PROCESSES



3. MASTER DATA HUB

Definition: The central database (the "brain") that stores the clean, consolidated Golden Records.

- *Why Important: It's the single source of truth for the organization.
- example: Like the registrar's office in a school that keeps the official student list—no duplicates, no errors.

KEY COMPONENTS OF MDM ARCHITECTURE & PROCESSES

4. DATA QUALITY TOOLS

Definition: Automated tools that clean, validate, and standardize data.

- Functions: Fix spelling mistakes, Remove duplicates, Standardize formats (e.g., "PH" vs. "Philippines")
- Why Important: Ensures the Golden Record is accurate and trustworthy.

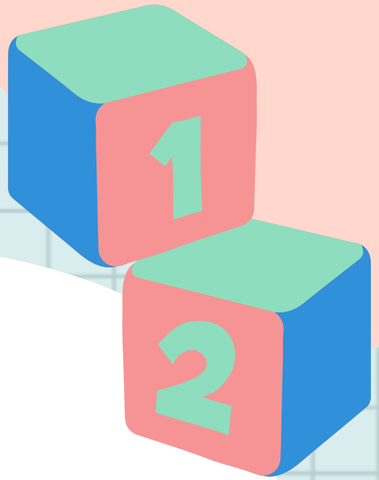
KEY COMPONENTS OF MDM ARCHITECTURE & PROCESSES

5. CONSUMERS

Definition: Systems, applications, and reports that use the clean master data.

- Examples: Business applications (ERP, CRM), BI dashboards, analytics, financial reports.
- Why Important: They rely on MDM to get consistent data for decision-making.

MDM PROCESSES



DATA COLLECTION

- Gathering raw data from multiple source systems.

DATA CONSOLIDATION

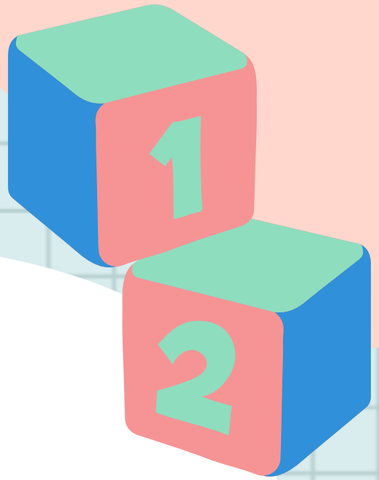
- Merging duplicate records, resolving conflicts.
- Example: "JP Cruz" + "J Paul Cruz" → John Paul Cruz

MDM PROCESSES

GOVERNANCE & VALIDATION

- Applying business rules and checking accuracy.
- Example: Date of Birth must be valid, email must follow standard format.

MDM PROCESSES



DATA DISTRIBUTION

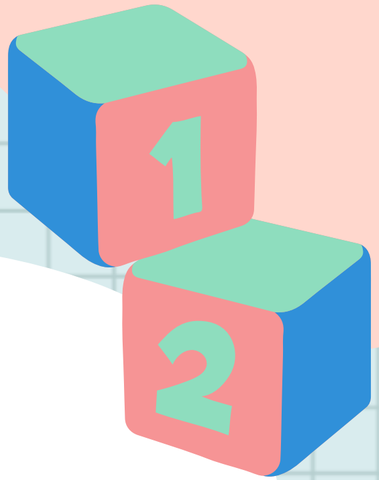
- Sharing clean, Golden Records to other systems.
- Example: Customer info sent to Sales, Finance, and Marketing

  → all see the same record.

MONITORING & MAINTENANCE

- Continuous checking, updating, and improving data quality.
- Ensures MDM doesn't become outdated or inconsistent again.

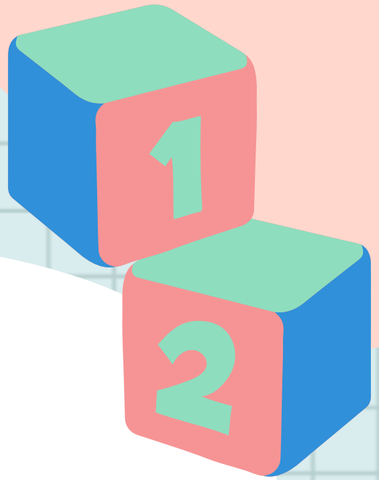
DATA QUALITY TOOLS



POPULAR TOOLS:

- Informatica MDM
- IBM InfoSphere
- SAP MDG
- Oracle MDM
- Talend

DATA QUALITY TOOLS



POPULAR TOOLS:

- Cloud-based MDM:

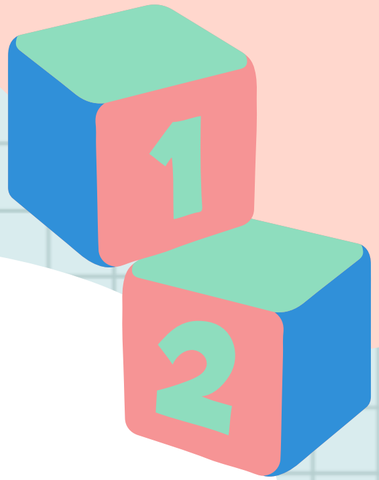
Definition: Hosting MDM in the cloud instead of on-premises.

Benefits: Flexible, scalable, cost-efficient.

Example: Instead of buying servers, companies can use cloud MDM that grows as their data grows.

MDM GOVERNANCE & STEWARDSHIP

MDM GOVERNANCE & STEWARDSHIP



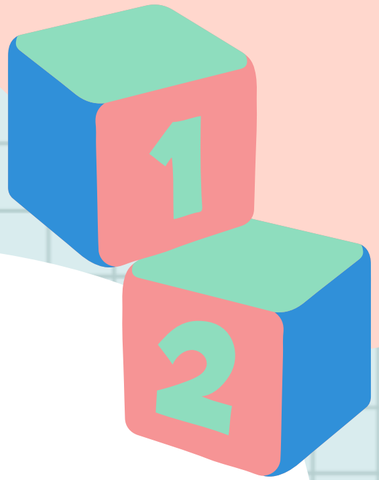
1. DATA GOVERNANCE

Definition: the policies, rules, and decision-making framework for managing data in MDM.

2. DATA STEWARDSHIP

Definition: the people and roles responsible for maintaining data quality, security, and compliance.

KEY ROLES IN GOVERNANCE & STEWARDSHIP



1. DATA OWNERS

- Senior leaders (like department heads).
- Define what the data means and who can access it.

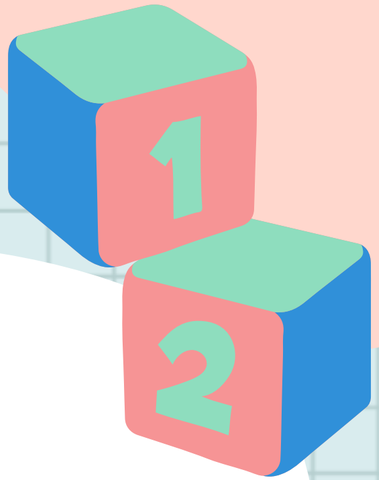
Example: HR Head decides how "Employee Records" are structured.

2. DATA STEWARDS

- Day-to-day caretakers of data.
- Ensure accuracy, fix errors, and monitor data quality.

Example: A Finance Data Steward ensures vendor bank accounts are correct.

KEY ROLES IN GOVERNANCE & STEWARDSHIP



3. DATA ADMINISTRATORS

- Manage system rules, permissions, and integrations.

Example: They make sure HR only gets Employee data, not Customer data.

GOVERNANCE PROCESSES IN MDM

1. DATA POLICIES

- Rules for how data is created, stored, shared, and secured.

Example: "All customer records must include email + phone number."

2. ACCESS CONTROL (RBAC – ROLE-BASED ACCESS CONTROL)

- Each system/role only gets the data it needs.

Example: Finance can see vendor bank info not employee medical records.

GOVERNANCE PROCESSES IN MDM

3. DATA QUALITY MONITORING

- Continuous checks for duplicates, missing fields, or invalid codes.

Example: If "PHL" is entered as currency, RDM + governance rules reject it.

4. AUDIT & COMPLIANCE

- Logs who accessed what, when, and how.
- Ensures compliance with laws like GDPR, HIPAA, or local data privacy laws.

WHY GOVERNANCE & STEWARDSHIP MATTER

- PREVENTS DATA CHAOS → WITHOUT RULES, EVERY SYSTEM WOULD DEFINE ALL DEPARTMENT AS "CUSTOMER".
- PROTECTS SENSITIVE DATA → PREVENTS FINANCE FROM SEEING INDIVIDUAL EPMLOYEE SALARY RECORDS.
- SUPPORTS TRUST → PEOPLE ONLY TRUST DASHBOARDS AND REPORTS IF THEY BELIEVE THE DATA IS ACCURATE AND CONSISTENT.
- REGULATORY COMPLIANCE → AVOIDS LEGAL ISSUES (E.G., DATA PRIVACY VIOLATIONS).


THE END
Thanks for listening!

DATA INTEGRATION

&

DATA PROVENANCE

GROUP 8

DATA INTEGRATION

Data integration involves combining data from disparate sources into a unified view for analysis, reporting, and other business intelligence purposes

Example:

- Cooking : Source(Recipe , Procedure, Equipment) output: (Adobo)

PRINCIPLES OF DATA INTEGRATION

DATA QUALITY AND CONSISTENCY

DATA GOVERNANCE

SCALABILITY AND PERFORMANCE

FLEXIBILITY AND ADAPTABILITY

METADATA MANAGEMENT

DATA QUALITY AND CONSISTENCY

Ensuring accuracy, completeness, and uniformity of data across all sources and targets.

Example : Students Record : Name (Jam, Age : 20)
Class Record : Name (Jam, Age : 90).

DATA GOVERNANCE

Establishing policies and procedures for managing data assets, including security, privacy, and compliance.

Example: Data Privacy (G-cash)

SCALABILITY AND PERFORMANCE

Designing integration solutions that can handle increasing data volumes and processing demands efficiently.

Example: Order Monday (100) - Order Friday (1000)

FLEXIBILITY AND ADAPTABILITY

Building systems that can easily integrate new data sources and adapt to evolving business needs.

Example: Student Record

METADATA MANAGEMENT

Metadata management refers to organizing, optimizing and using metadata to improve the accessibility and quality of an organization’s data.

Methods of Data Integration (ETL,ETL, DATA PIPELINE)

METHOD OF DATA INTEGRATION

ETL (Extract, Transform, Load):

- Extract: Data is extracted from source systems.
- Transform: Data is cleansed, standardized, and transformed into a suitable format in a staging area before being loaded.

METHOD OF DATA INTEGRATION

ETL (Extract, Transform, Load):

- Load: The processed data is loaded into a target system, typically a data warehouse.

Example: Hospital Different source:(Patient Admission, Lab test result)

Transform : Remove duplicate name, Standard date Format

Load : Sql Server, Snowflake

METHOD OF DATA INTEGRATION

ELT (Extract, Load, Transform):

Extract: Raw data is extracted from source systems.
Load: The raw data is loaded directly into the target system (e.g., cloud data warehouse or data lake).

METHOD OF DATA INTEGRATION

ELT (Extract, Load, Transform):

Transform: Transformations are performed within the target system, leveraging its processing power.

METHOD OF DATA INTEGRATION

Data Pipelines:

A data pipeline is a series of automated processes that move and transform data from its source to a destination.
It can encompass various integration methods like ETL, ELT, or real-time streaming, depending on the specific requirements

METHOD OF DATA INTEGRATION

Data Pipelines:

Data pipelines facilitate continuous data flow, enabling timely insights and supporting various analytical and operational use cases.
They are crucial for managing the entire data lifecycle, from ingestion to consumption.

METHOD OF DATA INTEGRATION

Types of Data Pipelines

- Batch Processing Pipelines: Process large volumes of data at scheduled intervals, ideal for tasks like monthly reporting.

Example: Business(Daily, Monthly, Yearly) Report .

- Streaming Pipelines: Handle real-time data, processing events as they occur, such as user interactions or sensor data.

Example: Facebook (Like, Comment, Share) Real-time.

Data Warehousing and Integration Challenges

DATA WAREHOUSING AND INTEGRATION CHALLENGES

Data Warehousing is the process of collecting, storing, and managing large amounts of data from different sources in one central place (called a data warehouse) for analysis and decision-making.

DATA WAREHOUSING AND INTEGRATION CHALLENGES

While data warehouses provide centralized repositories for structured data, several challenges exist in integrating heterogeneous sources:

- Data Variety: Managing structured, semi-structured, and unstructured data.

DATA WAREHOUSING AND INTEGRATION CHALLENGES

- Data Quality: Ensuring accuracy, completeness, and consistency.
- Scalability: Handling large and growing volumes of data.
- Latency: Balancing real-time and batch data integration needs.

Addressing these challenges requires advanced integration strategies and quality assurance mechanisms.

DATA PROVENANCE

DATA PROVENANCE

Data provenance, also known as data lineage, refers to tracking the history and flow of data.

DATA PROVENANCE

Key aspects include:

- Tracking Data Flow: Identifying where data originates and how it moves across systems.
- Monitoring Transformations: Recording how data is cleaned, aggregated, or transformed.
- Ensuring Accountability: Providing transparency for auditing, compliance, and troubleshooting.

DATA PROVENANCE

By maintaining lineage, organizations can increase trust in their data and ensure regulatory compliance.

THANK YOU

DATA WAREHOUSING & ONLINE ANALYTICAL PROCESSING (OLAP)

DATA WAREHOUSING

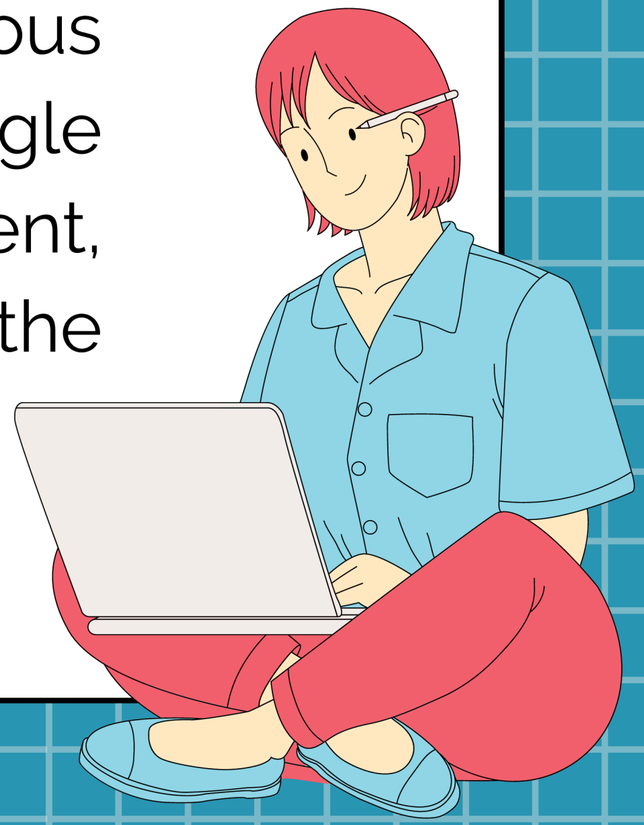
The process of gathering, keeping, and managing data from multiple sources in order to produce insightful business information.

ONLINE ANALYTICAL PROCESSING

Is a technology that uses the multi-dimensional data model to enable users to evaluate data from several database systems simultaneously.

DATA WAREHOUSING ARCHITECTURE

Is a system that helps businesses make better decisions by combining data from various sources and organizing it under a single architecture. It simplifies data management, reporting, and storage, increasing the effectiveness of analysis.



Data staging is a process in which data goes through various transformations, data cleaning procedures, and organization to be loaded into the data warehouse facility.

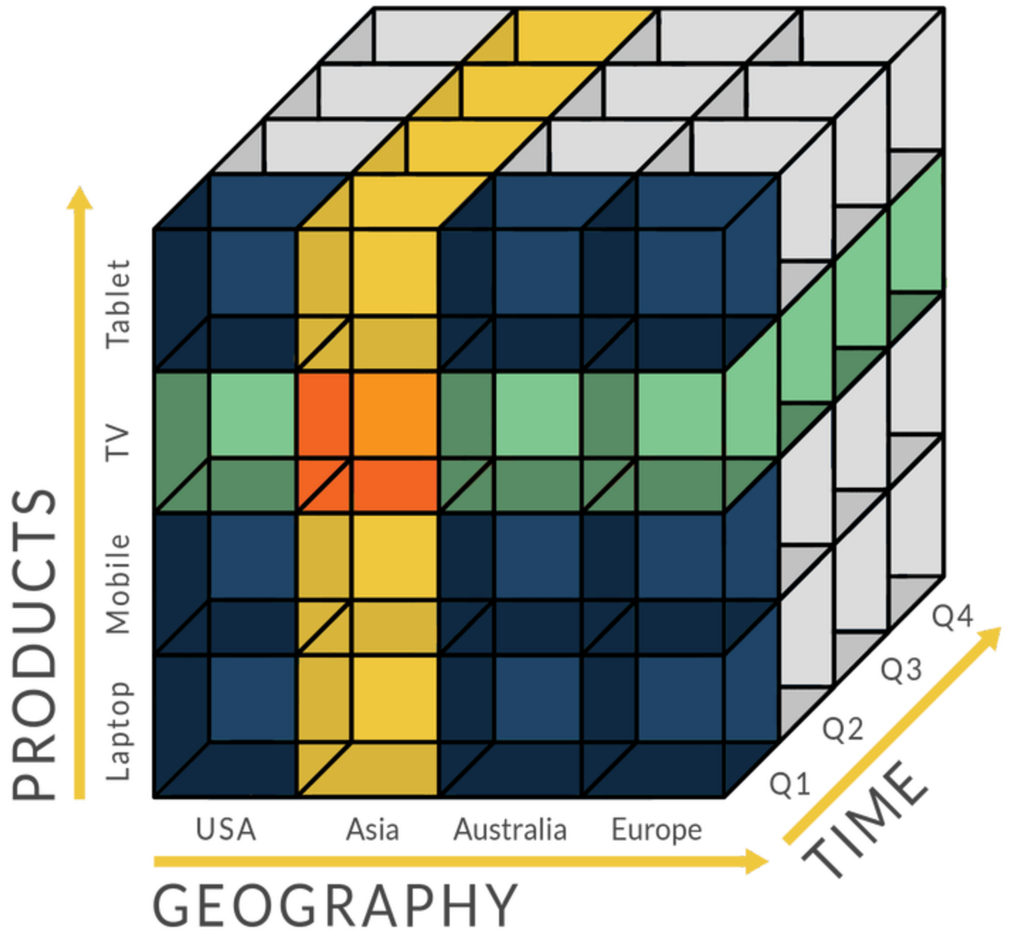
ETL tools manage this process:

Extract: Pulls raw data from sources.
Transform: Standardizes and formats the data.
Load: Moves the data into the data warehouse

Data Marts is the smaller, specialized areas within the data warehouse that serve certain groups by giving them direct access to the data they need.

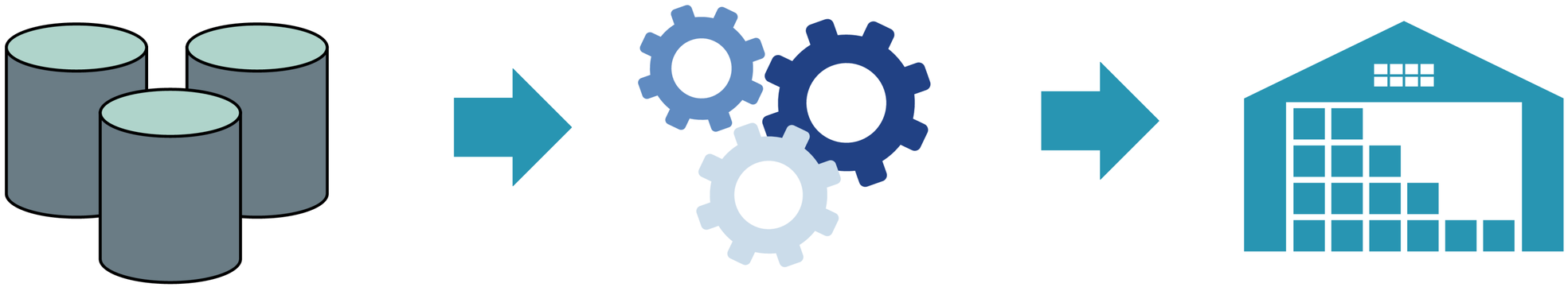
Ex. Sales, Marketing

An **OLAP cube** is a special way of storing data in a multi-dimensional structure that makes it easier and faster to analyze. Instead of looking at data in flat tables,


PRODUCTS
Laptop
Mobile
TV
Tablet
USA
Asia
Australia
Europe
GEOGRAPHY
Q1
Q2
Q3
Q4
TIME

EXTRACT, TRANSFORM, LOAD

ETL is the process of gathering data from different sources and integrating it into a central storage system called a data warehouse. During this process, raw data is cleaned, structured, and organized using business rules so that it's ready for storage, analysis.



EXTRACT

- **What happens:** Data is collected from different sources such as databases, spreadsheets, CRM systems, IoT devices, or web apps.
- **Goal:** Gather raw data, no matter the format (structured, semi-structured, unstructured).
- **Technologies/tools:** SQL, APIs, data connectors, tools like Apache Sqoop, Talend, Informatica, Fivetran.

TRANSFORM

- **What happens:** The extracted data is cleaned, validated, and reformatted according to business rules. This step handles removing duplicates, correcting errors, standardizing formats (e.g., date/time), and sometimes enriching the data.
- **Goal:** Convert messy raw data into consistent, high-quality information ready for analysis.
- **Technologies/tools:** Apache Spark, AWS Glue, Microsoft SSIS, dbt, Talend.

LOAD

- **What happens:** The transformed data is stored in the data warehouse (like Amazon Redshift, Google BigQuery, Snowflake, or Microsoft SQL Server).
- **Goal:** Make the data accessible for analytics, dashboards, reporting, and machine learning.
- **Technologies/tools:** ETL pipelines, batch/streaming loaders, cloud services (Airbyte, Informatica, Azure Data Factory).

OLAP VS. OLTP

Online Analytical Processing (OLAP)

- refers to software tools used for the analysis of data in business decision-making processes. It generally allow users to extract and view data from various perspectives.

Online Transaction Processing (OLTP)

- is a data processing approach emphasizing real-time execution of transactions. The majority of OLTP systems are meant to manage numerous short atomic operations that keep databases in line.

OLAP VS. OLTP

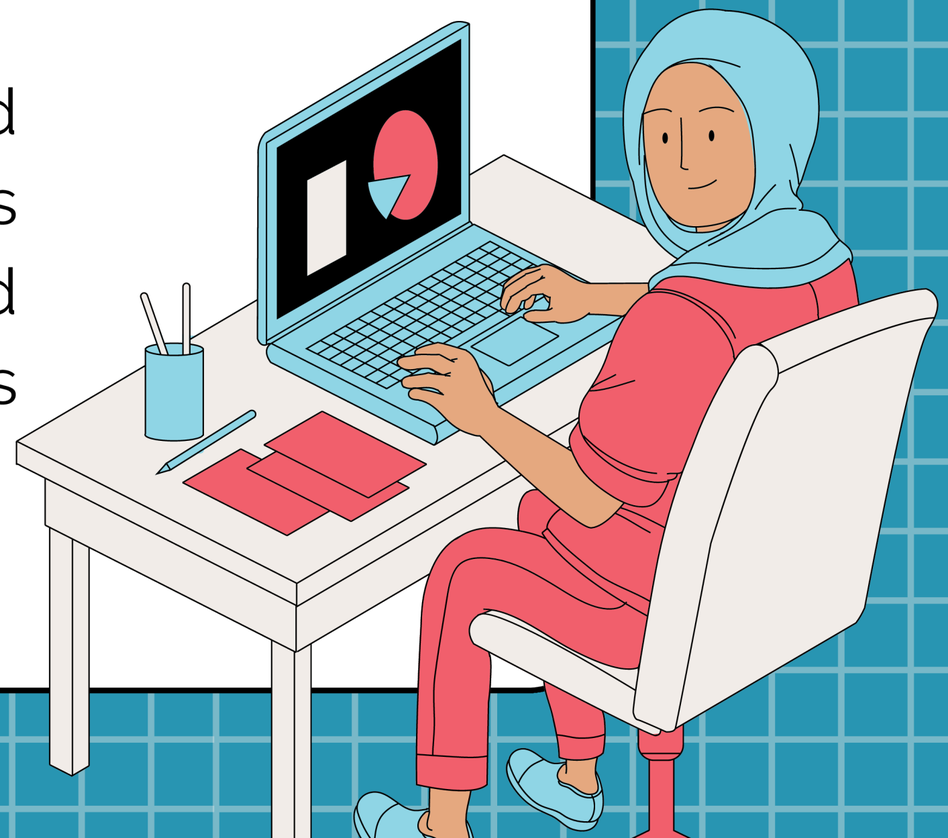
	OLAP (Online Analytical Processing)	**OLTP (Online Transaction Processing)**
Purpose	Used for **analysis** and **decision-making**	Used for **day-to-day transactions**
Use Cases	Forecasting, business reporting, dashboards, market trend analysis	Banking transactions, e-commerce purchases, airline booking systems
Applications/Tools	Microsoft SSAS, Oracle OLAP, SAP BW, Tableau	MySQL, PostgreSQL, SQL Server, Oracle DB

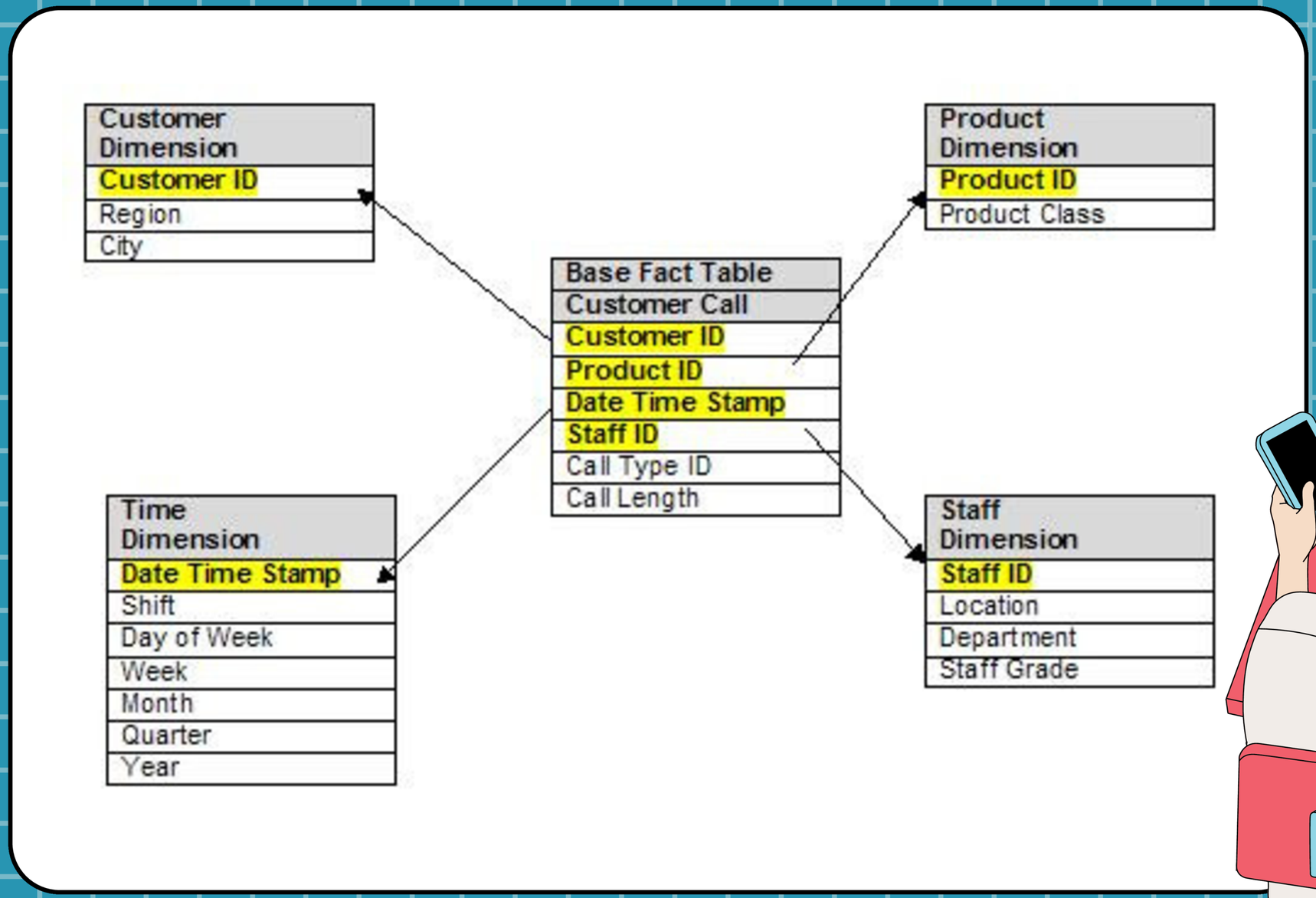
DESIGNING A STAR SCHEMA FOR BUSINESS INTELLIGENCE

A **star schema** is a type of multidimensional data model that structures information in a way that makes it simple to interpret and analyze. A star schema is important in Business Intelligence (BI) because it provides the foundation for storing and organizing data in a way that makes reporting and analysis fast and efficient.

Fact tables: holds the key business metrics like transactions in amounts and quantities.
Dimension tables: provide the context like time and product

The **fact table** is the center (measures) and the **dimension tables** are the sides (descriptions), and they are connected through primary key–foreign key relationships to make data meaningful.




Customer Dimension
Customer ID
Region
City
Product Dimension
Product ID
Product Class
Base Fact Table
Customer Call
Customer ID
Product ID
Date Time Stamp
Staff ID
Call Type ID
Call Length
Time Dimension
Date Time Stamp
Shift
Day of Week
Week
Month
Quarter
Year
Staff Dimension
Staff ID
Location
Department
Staff Grade

THANK YOU

References:

https://www.tutorialspoint.com/overview-of-data-warehousing-and-olap
https://www.geeksforgeeks.org/dbms/data-warehouse-architecture/
http://geeksforgeeks.org/software-testing/what-is-a-data-staging-area-in-data-warehouse/
https://aws.amazon.com/what-is/etl/
https://www.geeksforgeeks.org/dbms/difference-between-olap-and-oltp-in-dbms/
https://www.databricks.com/glossary/star-schema