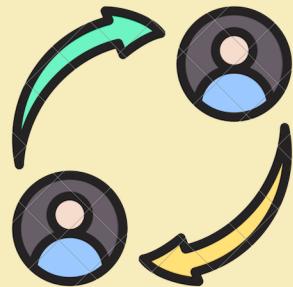


# INTERNSHIP PRESENTATION

Tasks completed



**Name: Pinak Mahapatra**  
**Adani Digital labs**  
**Analytics Team**



## Tasks Completed:

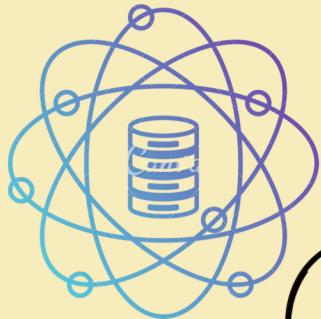
- Designed and developed interactive dashboards using Google Analytics and Looker Studio at Adani One. These dashboards effectively tracked key performance indicators (KPIs) and identified trends, while also communicating insights to non-technical stakeholders.
- Designed and implemented a content generation model using the GPT-4 API, seamlessly integrating LLMs through the Langchain framework. The model featured three hierarchical levels of content, each building upon the other. Additionally, optimized the code to produce unique descriptions with an ideal balance of burstiness and complexity.
- Successfully implemented the multimodal CLIP model for image classification and further refined an aesthetic scoring system to curate and select optimal images for the Adani One website.





# (SQL)Structured Query Language!



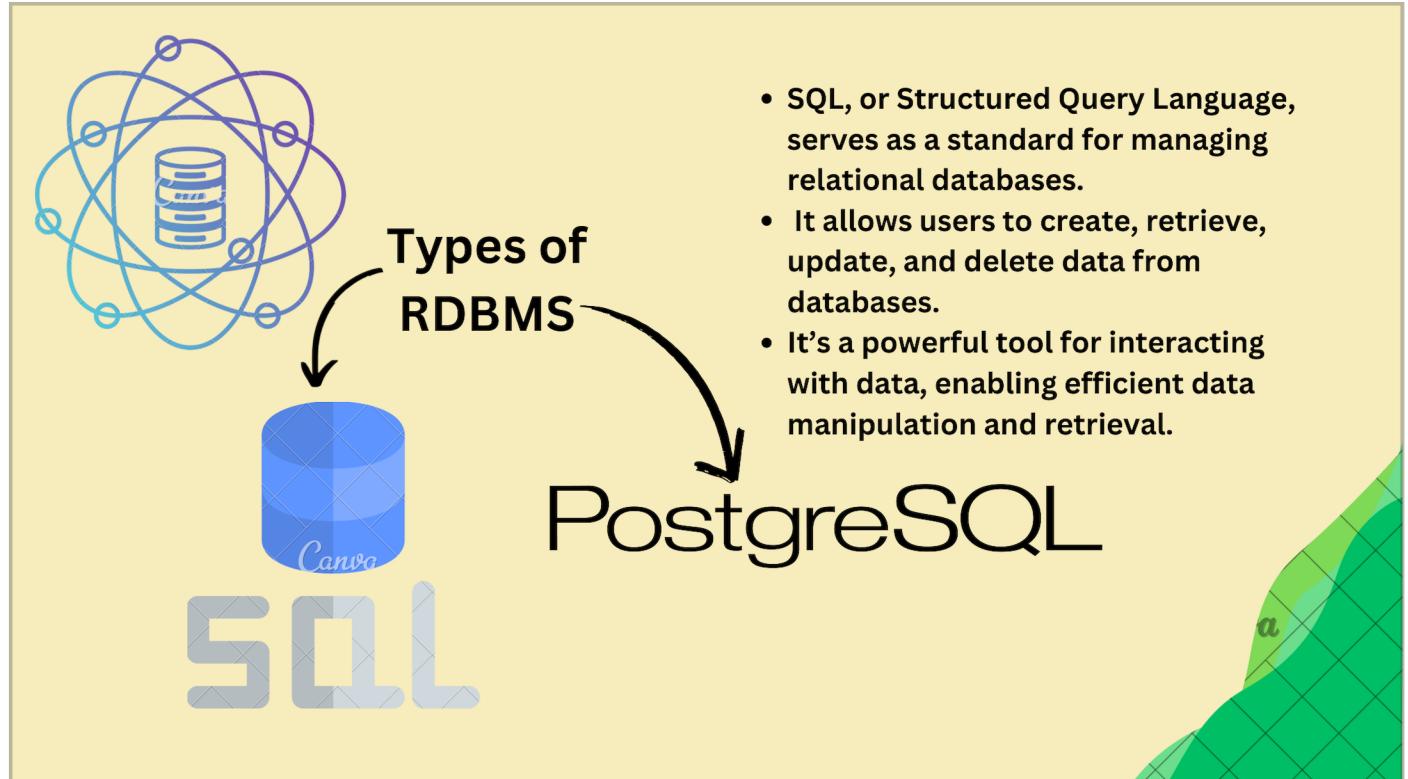


## Types of RDBMS

 **SQL**

- SQL, or Structured Query Language, serves as a standard for managing relational databases.
- It allows users to create, retrieve, update, and delete data from databases.
- It's a powerful tool for interacting with data, enabling efficient data manipulation and retrieval.

**PostgreSQL**



- Descriptive statistics summarize and organize characteristics of a data set.

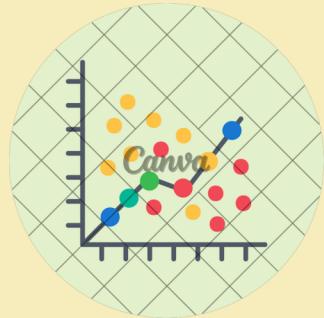
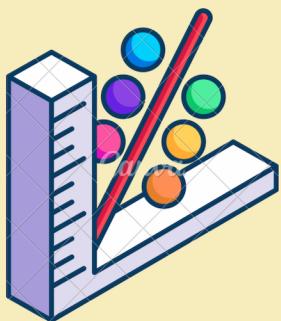
- Measures of central tendency help you find the middle, or the average, of a dataset. The three most common measures are Mean, Median and Mode

## Descriptive Analytics!

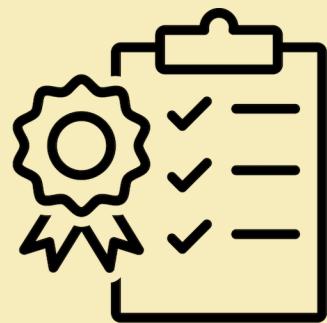


- Correlation analysis is a statistical technique used to quantify the degree to which two variables relate.
- It helps researchers evaluate the correlation coefficient, which indicates how much one variable changes when the other changes.

## Correlation Analysis



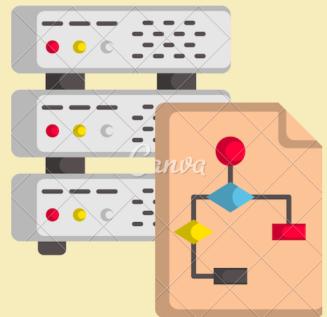
- A correlation coefficient quantifies the strength and direction of a relationship between variables. It ranges from -1 to 1: close to 1 indicates a strong positive correlation, close to -1 signifies a strong negative correlation, and close to 0 implies a weak or no linear relationship.



# Underlying Principles!

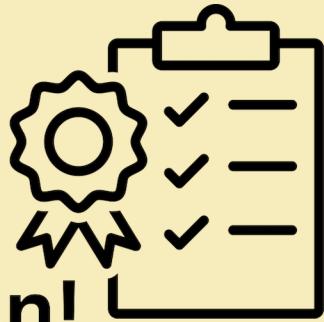


- **Invention and Purpose:** SQL (Structured Query Language) was developed by IBM in the 1980s. Its primary purpose is to interact with databases. Despite its strict syntax, SQL remains human-readable and widely used.
- **Functionality:** SQL allows users to: Execute queries against databases. Retrieve data from tables. Insert, update, and delete records. Create new databases, tables, stored procedures, and views.
- **Relational Databases (RDBMS):** SQL operates within the context of Relational Database Management Systems (RDBMS). RDBMS stores data in tables with columns and rows. SQL's readability makes it accessible even to non-programmers.





# Data Normalization!

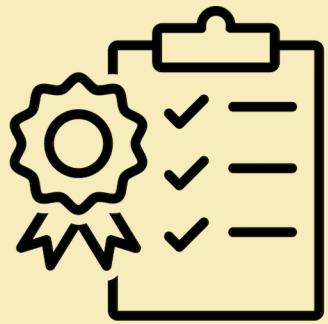


Database normalization is a crucial database design principle that organizes data consistently and reduces redundancy. It involves dividing data into multiple tables linked by relationships.



## Normal Forms of Tables

First Normal Form (1NF)      Second Normal Form (2NF)      Third Normal Form (3NF)



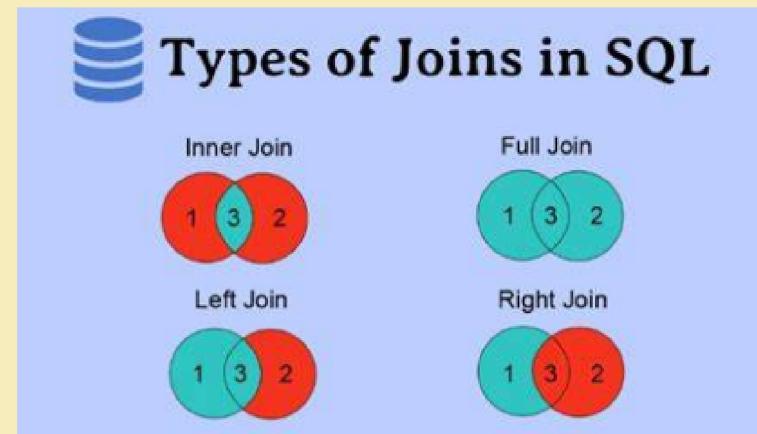
# SQL Joins



MySQL

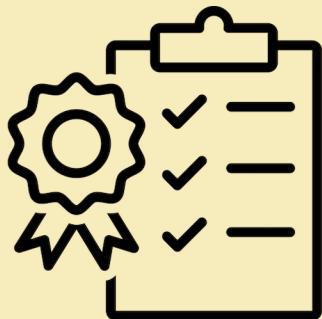


- **Right Join:** Returns all records when there's a match in either the left or right table.
- Includes unmatched records from both tables.





# Table Constraints



MySQL

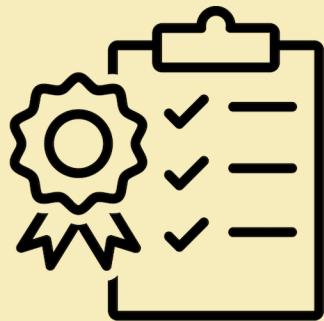
**Table constraints in SQL** are rules that ensure the correctness and integrity of data within a table. They define conditions that restrict the values allowed in specific columns.



- **PRIMARY KEY:** Ensures each row in a table has a unique identifier.
- **FOREIGN KEY:** Ensures the referential integrity of the data in one table to match values in another table.
- **NOT NULL:** Ensures that a column cannot have a NULL value.
- **UNIQUE:** Ensures all values in a column are different.
- **CHECK:** Ensures the value in a column meets a specific condition.
- **DEFAULT:** Sets a default value for a column if no value is specified.



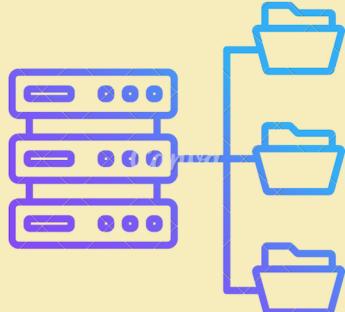
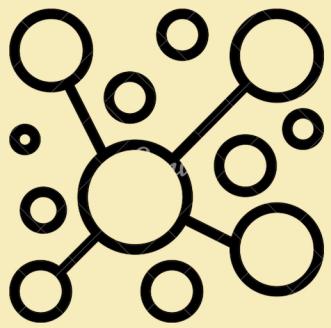
## Clustering and Indexing

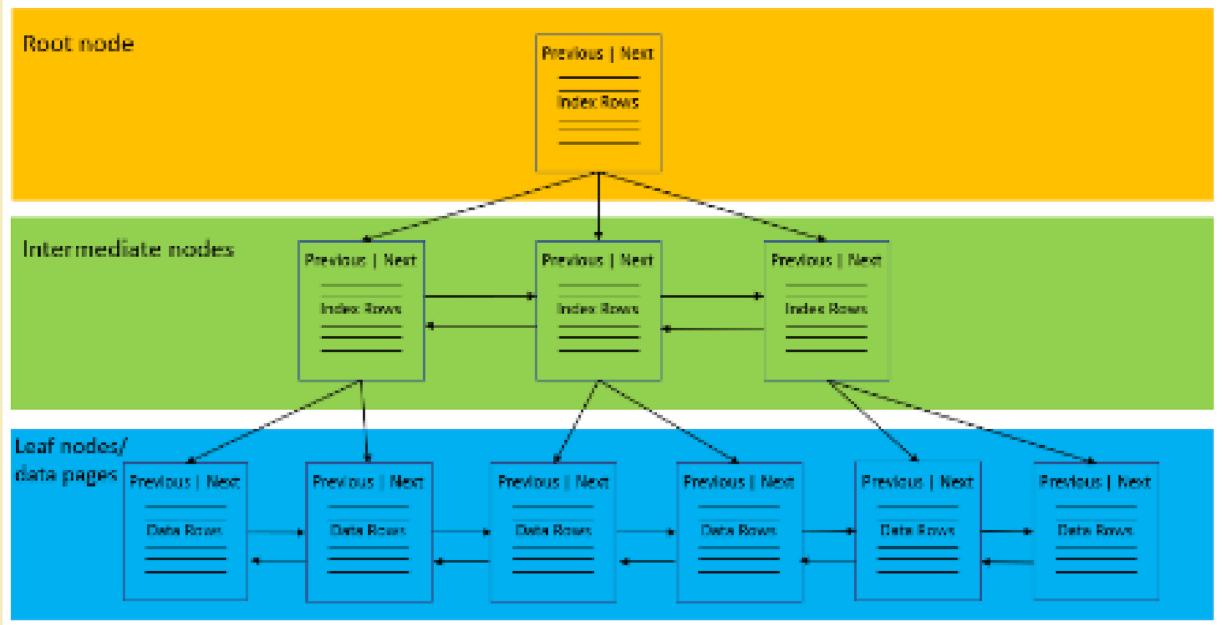


**MySQL**

# Clustering

- Clustering in the context of databases refers to the physical arrangement of data within a table based on a clustered index key.
- Rows are stored on disk in the same order as this key, which improves query performance and reduces disk space usage.
- However, it comes with complexities and limitations, such as reduced update performance and the restriction of one clustered index per table
- It works best for tables where data can be logically organised





Thank  
you!