PROJECT 1: DATABASE DESIGN FOR SYMPHONY WAVE MUSIC APP

Project Overview*:*

The SymphonyWave Music App is an exciting project that aims to create a robust relational database to power a music-playing application. This database will enable users to discover, listen to, and manage their music library seamlessly. The project's primary objective is to design and implement a data structure that supports essential functionalities such as music playback, personalised recommendations, and user-generated playlists.

Project Scope:
The scope of this project encompasses the entire database design process, from conceptualization to implementation. It includes the creation of tables, defining relationships, and ensuring data integrity. Below, we outline the key steps and deliverables for the SymphonyWave Music App database design.

Project Steps:
1. Conceptual Model (ER Diagram)
    - Create a conceptual model of the database using an Entity-Relationship (ER) Diagram, depicting entities and their relationships.
    - Identify and define the main entities in the database: Users, Artists, Albums, Tracks, Playlists, and Followers.
    - Specify the attributes for each entity, capturing essential data such as user profiles, artist information, album details, track metadata, playlist information, and follower relationships.
    - Define the cardinality and type of relationships between entities (e.g., one-to-many, many-to-many).

2. Relational Model
    - Based on the conceptual model, design the relational model by converting entities into database tables.
    - Define primary keys, foreign keys, and attributes for each table.
    - Ensure that referential integrity is maintained by enforcing foreign key constraints.

3. Physical Model (PostgreSQL Implementation)
    - Implement the designed database using the PostgreSQL Relational Database Management System.
    - Create the necessary tables, specifying data types, constraints, and relationships.
    - Populate tables with sample data to test the database's functionality.

4. Testing and Validation
    - Validate the database by performing data integrity checks.

- Ensure that relationships between tables function as expected.
- Execute sample queries to retrieve and manipulate data to verify correctness.

Deliverables:

Conceptual Model (ER Diagram):

- A visual representation of the conceptual database model, illustrating entities and relationships. Use tools like draw.io for this purpose.

Relational Model:

- A structured document outlining the database schema, including table definitions, attributes, primary keys, and foreign keys.

Physical Model (PostgreSQL Implementation):

- The actual implementation of the database using PostgreSQL, including SQL scripts to create tables, constraints, and relationships.

By completing this project, we aim to deliver a well-designed and functional database that forms the foundation of the SymphonyWave Music App, enhancing the user experience and enabling a seamless music listening experience.

PROJECT 2: Data Cleaning for Improved Analytics

Project Overview:
In today's data-driven world, the quality of data is paramount for making informed decisions and deriving accurate insights. Data cleaning, also known as data preprocessing, is a crucial step in the data analysis pipeline. This project focuses on cleaning and preparing a real-world dataset to enhance its quality, accuracy, and usability for downstream analytics.

Project Objectives:
- Identify and rectify data quality issues in the dataset.
- Standardise data formats and values to ensure consistency.
- Handle missing data appropriately.
- Remove duplicates and outliers.
- Create a cleaned dataset ready for analysis.

Project Steps:
1. Data Assessment:
    - Obtain and import the raw dataset.
    - Examine the data's structure, columns, and overall quality.
    - Document data quality issues, inconsistencies, and missing values.

2. Data Cleaning:
    - Address Missing Data:
        - Decide on appropriate strategies for handling missing values (e.g., imputation, removal).
        - Implement chosen strategies to fill or remove missing data.
    - Standardize Data:
        - Standardize data formats (e.g., date formats, naming conventions).
        - Normalize categorical data values (e.g., capitalization, spelling errors).
    - Remove Duplicates: Identify and remove duplicate records if necessary.
    - Outlier Handling: Identify and handle outliers appropriately, considering the impact on analysis.

3. Data Transformation:
    - Feature Engineering (if applicable): Create new features that may be useful for analysis.

4. Documentation: Maintain clear documentation of all data cleaning steps, transformations, and decisions made.

5. Validation:
    - Verify the data's quality and integrity after cleaning.
    - Conduct basic summary statistics and visualizations to ensure the data behaves as expected.

6. Data Export:
   - Export the cleaned dataset to a structured format (e.g., CSV) for analysis.

7. Reporting:
   - Summarize the data cleaning process, highlighting the improvements made.
   - Document any challenges encountered and lessons learned.

Deliverables:
- A cleaned and well-documented dataset.
- A report detailing the data cleaning process, transformations, and validation.
- Any code/scripts used for data cleaning.

Conclusion: Data cleaning is a critical step in the data analysis workflow, ensuring that the data used for analytics is accurate and reliable. This project equips you with valuable skills in data preprocessing, enabling you to handle real-world data quality challenges effectively. The cleaned dataset can then be used for further analysis and insights generation.

DATASET

PROJECT 3: Pizza Sales Analysis Project

Project Overview:
Your Data Won't Speak Unless You Ask It The Right Data Analysis Questions
Background: As a Data Analyst, your role extends beyond merely gathering a company's existing data; it involves the strategic preparation of business data to uncover the most valuable insights. Regardless of an organisation's technological sophistication, data alone won't offer a ready-made solution. The key lies in asking precise questions related to data analysis.

Hence, effectively converting data into actionable business decisions starts with recognizing the specific challenges you aim to tackle. Even before data collection, craft a set of targeted questions that align with your goals and the intended audience. These questions will function as a guideline as you navigate through the data analysis process, streamlining the extraction of useful insights that align with your analysis objectives.

Project Overview
This data analysis project aims to analyse the sales data of a pizza company to uncover insights and trends that can help improve business strategies, optimise menu offerings, and enhance overall sales performance.

Objectives

1. Collect and preprocess sales data, including order details, time of purchase, and menu items
2. Perform exploratory data analysis to identify patterns in sales, popular menu items, peak ordering hours, and customer preferences using visually appealing visuals to further explain findings

Analytical Questions
The current data analysis project seeks to examine pizza sales patterns with the goal of identifying trends, factors, and opportunities that impact sales performance. Key Performance Index (KPI):

1. What is the Total Revenue made on Sales
2. What is Average Revenue
3. What is the Total number of Order made
4. What is the Average Order Value
5. What is the Total Quantity sold Patterns and Trends
6. Identifying the pizza size that generated the highest sales in the first quarter of the year
7. Determining the pizza category that recorded the highest sales in December
8. Analyze sales trends across different days of the week, months, days, and time periods.
9. Identifying which pizza size exhibits the highest turnover rate in terms of quantity ordered
10. Visualize the trajectory of total income over time and assessing how prices change as time progresses

DATASET