Task 3: SQL for Data Analysis

Dataset: [Bike Store Relational Database](https://www.kaggle.com/datasets/dillonmyrick/bike-store-sample-database?select=products.csv)

Tool: MySQL

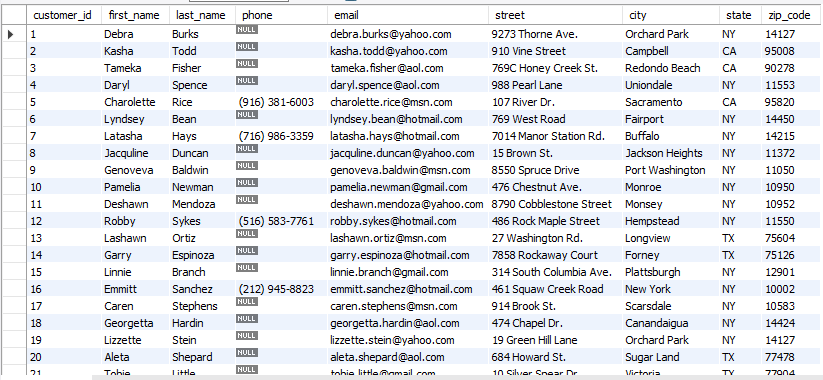
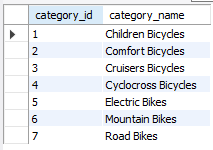
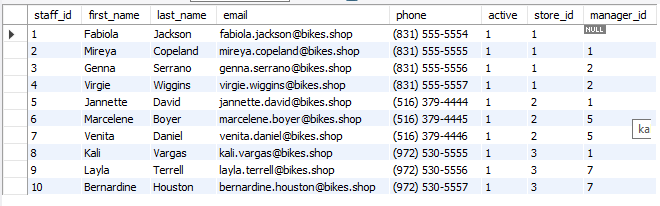
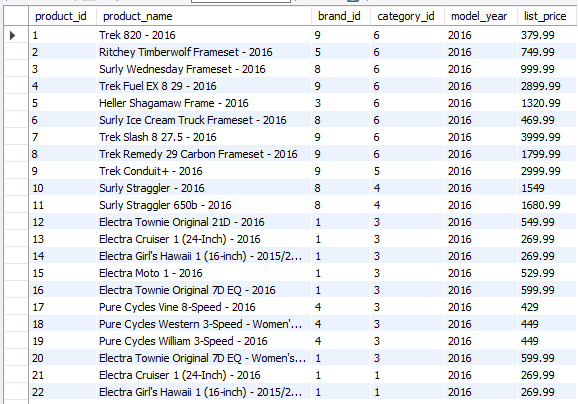
Name: Sasindu Chanaka Piyumal.

Email ID: [sasinduha@gmail.com](mailto:sasinduha@gmail.com)

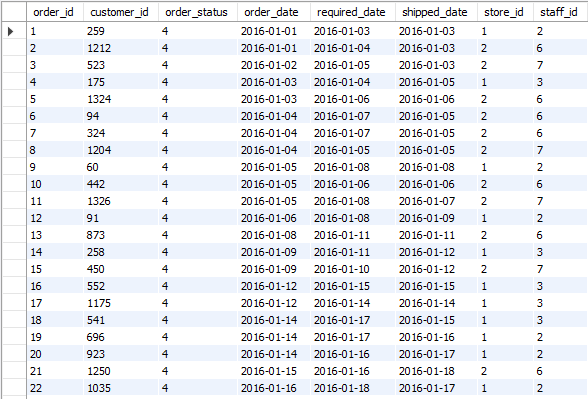
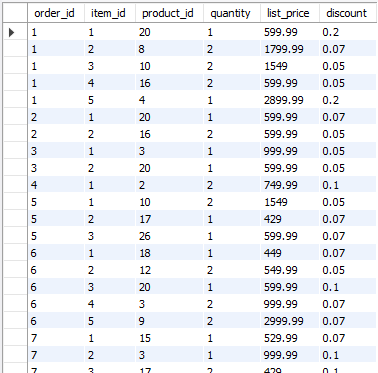
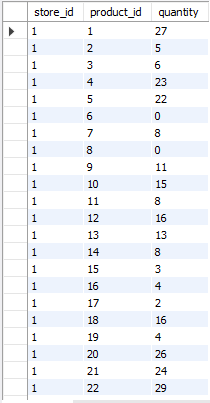
**Introduction**

The Bike Store dataset is a comprehensive relational database designed to simulate the operations of a bicycle retail business. It encompasses various key aspects of a real-world bike store, including product inventory, customer details, sales transactions, store locations, and staff information. This dataset is ideal for practicing and demonstrating SQL querying skills, data analysis, and business intelligence reporting.

The data is organized into multiple related tables, including:

* Customers – Contains information such as customer names, contact details, and location.
* Categories– Metadata about categories
* Staff – Includes details of employees like salespersons and managers.
* Products – Holds data about bicycles including brand, category, model, and price
* Brands – Provide metadata for product classification.



* Orders – Records of customer purchases including dates, order status, and store/salesperson information.
* Order\_Items – Line items in each order, specifying products, quantities, and prices.
* Stock – Information on product availability in each store.

This dataset allows users to explore important business questions such as revenue trends, inventory management, customer behavior, and employee performance. It's widely used for learning SQL operations like JOIN, GROUP BY, AGGREGATE FUNCTIONS, and subqueries, making it a powerful resource for students and professionals in data-related fields.

**Methodology**

The following methodology was followed to analyze the Bike Store dataset and extract valuable business insights using MySQL:

1. Data Acquisition

The Bike Store dataset was downloaded from Kaggle, which provides a realistic simulation of a retail bicycle business, including information on customers, stores, staff, products, orders, and inventory.

1. Database Setup

A new database was created in MySQL to store and manage the dataset. This setup allowed efficient querying and relational data management.

1. Data Import

The individual dataset files (typically in .csv format) were imported into the MySQL database. Each file was imported as a separate table representing different entities such as customers, products, orders, order\_items, staff, stores, brands, and categories

1. Data Exploration

Initial exploratory queries were run to understand the structure, relationships, and contents of each table. This included checking column names, data types, null values, foreign key connections, and sample data.

1. Scenario Analysis and Question Formulation

Based on business scenarios and use cases, a set of 25 meaningful and relevant analytical questions were designed. These questions cover various aspects such as sales trends, customer behavior, product performance, and staff efficiency.

1. Query Execution

The questions were answered by writing and executing SQL queries using MySQL. Queries made use of SQL concepts such as SELECT, JOIN, GROUP BY, ORDER BY, AGGREGATE FUNCTIONS, subqueries, and date functions.

1. Result Interpretation

The output of each query was analyzed to derive insights and support decision-making. Results were documented and, where applicable, visualized using basic tabular formats to enhance clarity.

**Results**

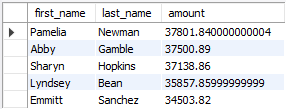
1. What are the total sales made by the store?



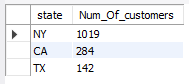
1. What is the total number of orders placed?



1. Who are the top 5 customers by total purchase amount?



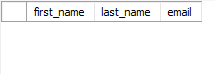
1. How many customers are there in each state?



1. What is the total quantity of bikes sold by each store?



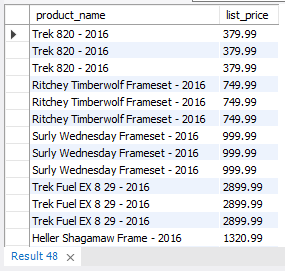
1. Which customers have never placed an order?



1. What is the average order value?



1. What are the names and prices of all the bikes in stock?



1. What are the top 10 most expensive bikes?

