Customer Purchase Analysis

Project Description:

Given a dataset of customer purchase history for an online retail company containing various attributes.

Attributes given:

- Transaction ID
- Customer ID
- Customer Name
- Product ID
- Product Name
- Product Category
- Purchase Quantity
- Purchase Price
- Purchase Date
- Country

Project Objective:

Create a behaviour analysis based on the purchase trends of customers using SQL, Python and Power BI.

Using SQL to setup Database and Data Transformation

Using Python to Extract Data from SQL Data Base and find trends in Data

Using Power BI to Create a Dashboard of the above discovered Insights.

Methodology:

SQL Workbench:

Created Database to work on the project data

Checked for the missing values and duplicated

Created new tables to bifurcate the data into different tables

- Customers Table
- Products Table
- Orders table
- Orderdetails Table
- Category Table

Python:

Imported the above created tables from SQL Workbench into Python

Used function such as merge, groupby to join and aggregate tables and data to answer the asked questions

Libraries used:

- Pandas
- Matplotlib
- Pymysql
- Datetime
- Csv

Csv libraray was used to create csv files of the tables imported from SQL Workbench.

Pandas library was used to do analysis on the tables by merging and aggregating.

DateTime library was used to convert certain columns into Date format and further run date related operations on it.

Pymysql was used to connect MySQL Workbench to Python

Matplotlib was used for plotting and presentation of data analysis

Power BI:

Visualization of data using the tables created in SQL Workbench

Connected Workbench to Power BI and extracted the tables.

Built interactive visualization and insights.

Analysis:

Varius Customer Purchase Trends were discovered over year, quarter and month.

The dataset had two categories Home Appliances and Electronics, it was observed that home appliances were sold more but it was also observed that Electronics contributed more towards the revenue.

Based on the count of products sole, it was observed that home appliances topped the list with the number of units sold.

Conclusion:

Various trends and patterns were discovered in the Customer Purchase Data. Power BI dashboard highlights them all in a very concise manner. Since this data was only about 1000 customers, a larger data set could have provided some more exciting trends in the data and some new customer purchase behaviours would have been discovered

Project Link:
https://drive.google.com/drive/folders/1QwY_9KOG1G7zRxv9WKpXiWiy5Q8Knf3R?usp=sharing