**Description: Data Source[**[**Link**](https://www.kaggle.com/c/walmart-recruiting-store-sales-forecasting)**]**

The project involves analysing a dataset containing sales transactions from different branches of Walmart. It includes details such as branch information, customer details, product information, sales quantities, prices, taxes, and ratings. The dataset is used to perform various analyses, including product analysis, sales analysis, and customer analysis. Feature engineering is used to extract additional information such as time of day and weekday from the data. The project aims to answer various business questions related to product performance, sales trends, customer behaviour, and revenue analysis.

**Purpose Of Project:**

The main aim of this project is to gain insights into the sales data of Walmart to understand the different factors that affect sales across branches. It involves analysing the sales trends of different products, customer behaviour, and the effectiveness of sales strategies.

**About Data:**

| **Column Name** | **Description** | **Data Type** |
| --- | --- | --- |
| Invoice id | Invoice ID of the sales transaction | VARCHAR (30) |
| branch | Branch at which sales were made | VARCHAR (5) |
| city | Location of the branch | VARCHAR (30) |
| Customer type | Type of the customer | VARCHAR (30) |
| gender | Gender of the customer making purchase | VARCHAR (10) |
| Product line | Product line of the product sold | VARCHAR (100) |
| Unit price | Price of each product | DECIMAL (10,2) |
| quantity | Amount of the product sold | INT |
| VAT | Amount of tax on the purchase | FLOAT |
| Total cost | Total cost of the purchase | DECIMAL (10,2) |
| date | Date of the purchase | DATE |
| time | Time of the purchase | TIMESTAMP |
| Payment mode | Payment method | VARCHAR (15) |
| cogs | Cost Of Goods sold | DECIMAL (10,2) |
| Gross margin percent | Gross margin percentage | FLOAT |
| Gross income | Gross Income | DECIMAL (10,2) |
| rating | Rating | FLOAT |

**Analysis List**

Product Analysis

Conduct analysis on the data to understand the different product lines, the products lines performing best and the product lines that need to be improved.

Sales Analysis

This analysis aims to answer the question of the sales trends of product. The result of this can help use measure the effectiveness of each sales strategy the business applies and what modifications are needed to gain more sales.

Customer Analysis

This analysis aims to uncover the different customers segments, purchase trends and the profitability of each customer segment.

**Approach Used**

1. **Data Wrangling:** This is the first step where inspection of data is done to make sure **NULL** values and missing values are detected and data replacement methods are used to replace, missing or **NULL** values.
2. Build a database.
3. Create table and insert the data.
4. Select columns with null values in them. There are no null values in our database as in creating the tables, we set **NOT NULL** for each field, hence null values are filtered out.
5. **Feature Engineering:** This will help use generate some new columns from existing ones.
6. Add a new column named `**time of day**` to give insight of sales in the Morning, Afternoon and Evening. This will help answer the question on which part of the day most sales are made.
7. Add a new column named `**day name**` that contains the extracted days of the week on which the given transaction took place (Mon, Tue, Wed, Thru, Fri). This will help answer the question on which week of the day each branch is busiest.
8. Add a new column named **`month name`** that contains the extracted months of the year on which the given transaction took place (Jan, Feb, Mar). Help determine which month of the year has the most sales and profit.
9. **Exploratory Data Analysis (EDA):** Exploratory data analysis is done to answer the listed questions and aims of this project.

**Recommendations for Sales Analysis:**

1. **Optimize Product Mix:** Analyse the most selling product lines and adjust inventory to meet demand, ensuring popular products are always available.
2. **Improve Marketing Strategies:** Identify peak sales times and tailor marketing campaigns to maximize impact during these periods.
3. **Enhance Customer Segmentation:** Use customer analysis to tailor promotions and offerings to different customer types, increasing customer loyalty and satisfaction.
4. **Implement Pricing Strategies:** Analyse pricing data to adjust prices for different products or in different locations to maximize revenue and profit margins.
5. **Streamline Operations:** Use sales trend analysis to optimize staffing levels and inventory management, reducing costs and improving efficiency.
6. **Monitor Competitor Performance:** Analyse competitor sales data to identify market trends and adjust strategies to stay competitive.
7. **Leverage Technology:** Use advanced analytics and machine learning to gain deeper insights into sales data and identify new opportunities for growth.

**Note:**

Hello and welcome to this project! As you explore the data and queries, I've created, I hope you find them insightful and useful for your own analysis. Feel free to suggest any improvements or modifications to the queries based on your needs. Your feedback will be invaluable in enhancing the effectiveness and efficiency of this project for future users. Enjoy exploring the data and happy analyzing!