

WALMART SALES ANALYSIS

ABOUT

The aim of this project is to analyse the sales data of Walmart in Myanmar southeast Asia. This data was downloaded from kaggle.com and aims to understand the top performing branches, cities and products.

PURPOSE

The major aim of this project is to gain insight into the sales data of Walmart to understand the different factors that affect sales in the different branches.

ABOUT THE DATA

This dataset contains sales transactions from the three different branches of Walmart, respectively located in Mandalay, Yangon and Naypyitaw in Myanmar. My data contains 12 columns and 1000 rows but was truncated when imported into the MySQL database to 995 due to the elimination of duplicate entries.

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 the product. The result of this can help us 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 customer segments, purchase trends and the profitability of each customer segment.

APPROACH

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.

- 1) Create a database
- 2) Create table and insert the data.
- 3) 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.

Feature Engineering: This will help us generate some new columns from existing ones. For my analysis, I created **FACT COLUMNS** using simple mathematical formulas and **LOOK UP COLUMNS** using the date and time columns.

Exploratory Data Analysis (EDA): Exploratory data analysis is done to answer the listed questions and aims of this project.

BUSINESS QUESTIONS

Generic Question

1. How many unique cities are in MYANMAR?
2. In which city is each branch?

Product Questions

1. How many unique product lines does the sales data have?
2. What is the most common payment method?
3. What is the most selling product line?
4. What is the total revenue by month?
5. What month had the largest COGS?
6. What product line had the largest revenue?
7. What is the city with the largest revenue?
8. What product line had the largest VAT?
9. Fetch each product line and add a column to those product line showing "Good", "Bad". Good if its greater than average profit.
10. Which branch sold more products quantity than the average product quantity sold?
11. What is the most common product line by gender?
12. What is the average rating of each product line?

Sales Questions

1. Number of sales made in each time of the day per weekday
2. Which of the customer types brings the most revenue?
3. Which city pays the highest tax percent/ VAT (Value Added Tax)?
4. Which customer type pays the most in VAT?

Customer

1. How many unique customer types does the data have?
2. How many unique payment methods does the data have?
3. What is the most common customer type?
4. Which customer type buys the most?
5. What is the gender of most of the customers?
6. What is the gender distribution per branch?
7. Which time of the day do customers give most ratings?
8. Which time of the day do customers give most ratings per branch?
9. Which day of the week has the best average and total ratings?
10. Which day of the week has the best average and total ratings per branch?

MATHEMATICAL CALCULATIONS

VAT = 5% of COGS

Net profit = COGS – VAT

Gross profit = COGS + VAT

COGS = unit price * quantity sold