DATA REPORT

1.1 Business Understanding

Topup mama is an organization that digitizes the supply for micro, small and medium-sized restaurants.It is a private company based in Nairobi with an employee size of 51-100 people.

Data source

The data was obtained from the following source: [here](#)

1.2 Business Objectives

The overall objectives are
- To merge the datasets into one.
- Clean the data
- Use machine models to predict customer retention,classify customers and provide product recommendations.

1.3Tools Employed

We used various tools which were essential in carrying out our project such as :

1. [Github Repo](#)
2. [Collaboratory Notebook](#)

2.0 Methodology

The Data Mining approached follows a methodology inspired by the CRISP-DM that follows the following phases;

2.1 Data Exploration

We checked the information of our data set by looking at the data types, the missing values, the column names, descriptions, the number of rows, and the shape of our dataset.

2.2 Data Pre-processing

During this step we were able to;

- Check the missing values in each column.
- Check the duplicate records.
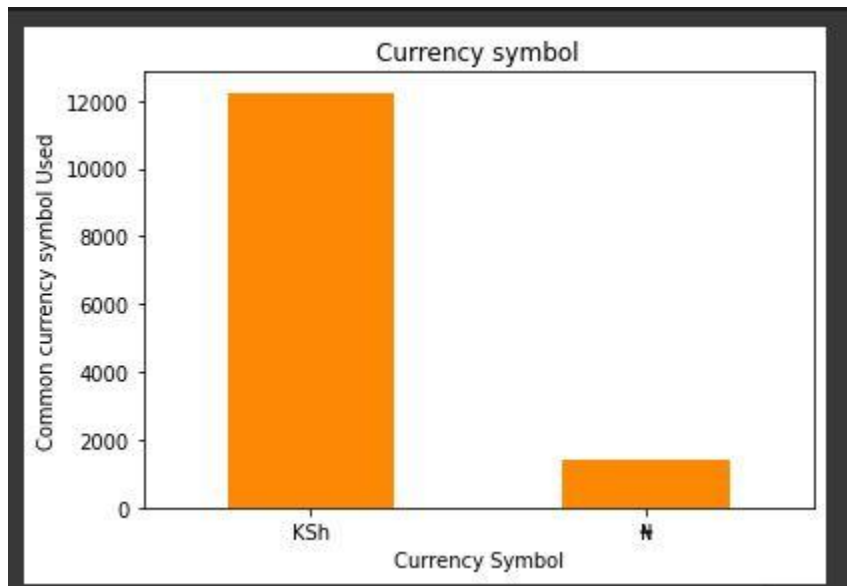- Undersampling the dataset to ensure balance in expected outputs

3.0 Modeling

I implemented unsupervised learning model only due to limited time,
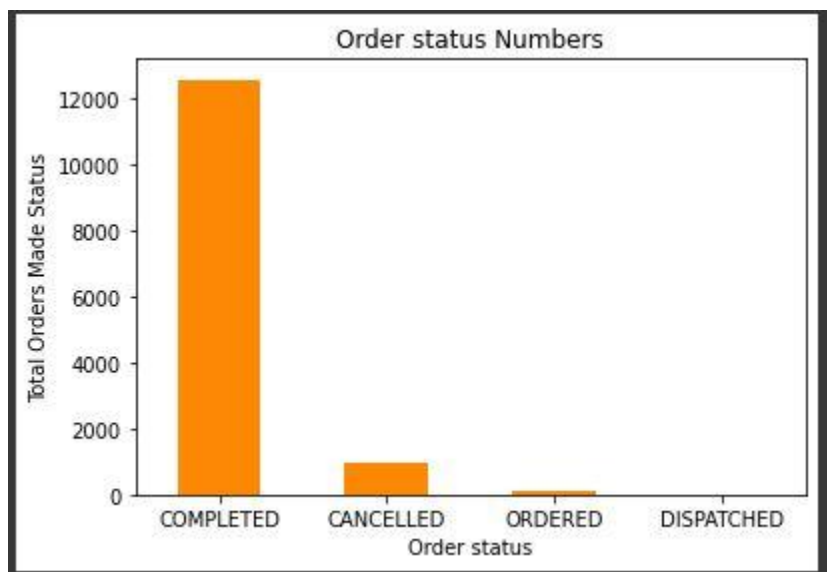
- KMeans Clustering:
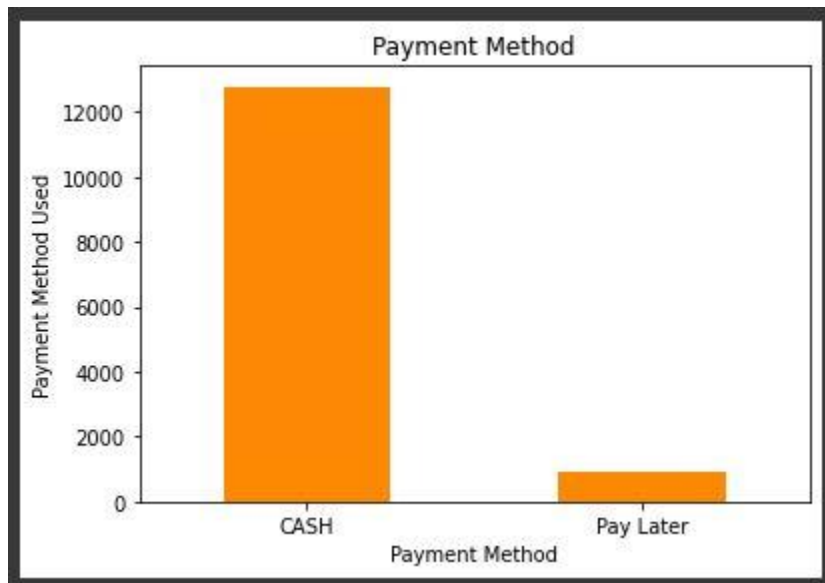
  This was to classify customers

4.0 Conclusions

1.The most commonly used currency symbol is KSH

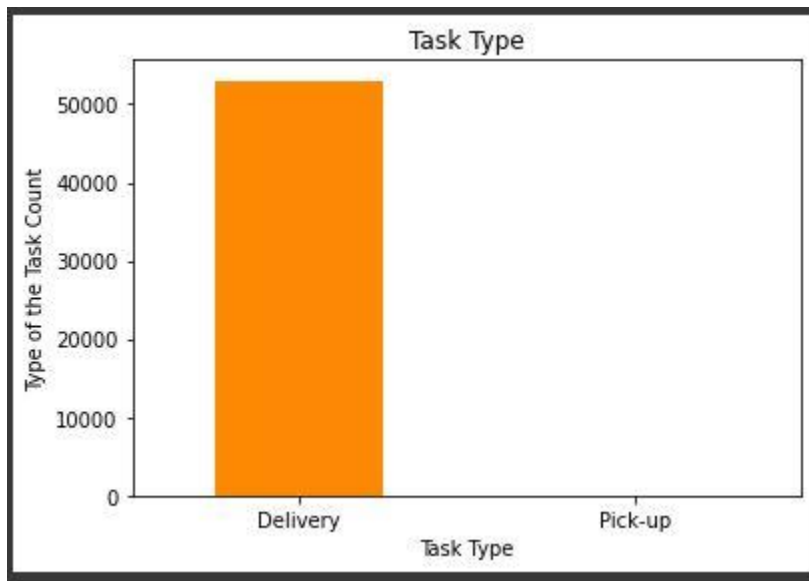2.Most orders have been completed and very few have been dispatched.

3.The most preferred payment method is cash and not Pay Later.

3. Most tasks were delivery and few were pickup.

5.The most ordered items are flour and sugar and least is wine.

category_name	
Flour & Sugar	4539
Cooking Fat & Oil	3091
Salt & Seasoning	1985
Cleaning & Hygiene	956
Fruits & Vegetables	922
Packaging	731
Rice & Cereals	485
Meat & Fish	461
Beverages	304
Spreads	70
Noodles & Spaghetti	61
Vegetables	56
Beer & Bitters	3
Wine	1