



AllLife Bank Customer Segmentation

Business Presentation

Contents

Business Problem Overview

Data Overview

Exploratory Data Analysis (EDA)

Model Performance Summary

Business Insights and Recommendations

Business Problem Overview and Solution Approach

AllLife Bank wants to focus on its credit card customer base in the next financial year as the penetration in the market can be improved.

Based on this input, the Marketing team proposes to run personalized campaigns to target new customers as well as upsell to existing customers.

The Operations team wants to upgrade the service delivery model, to ensure that customer queries are resolved faster.

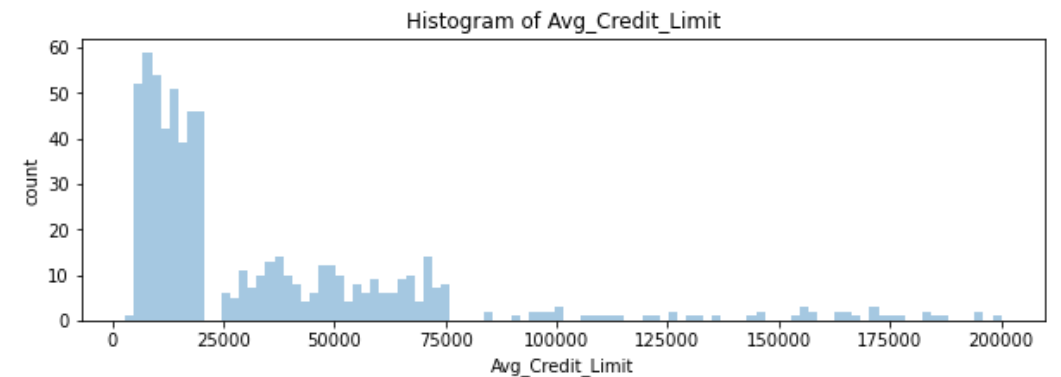
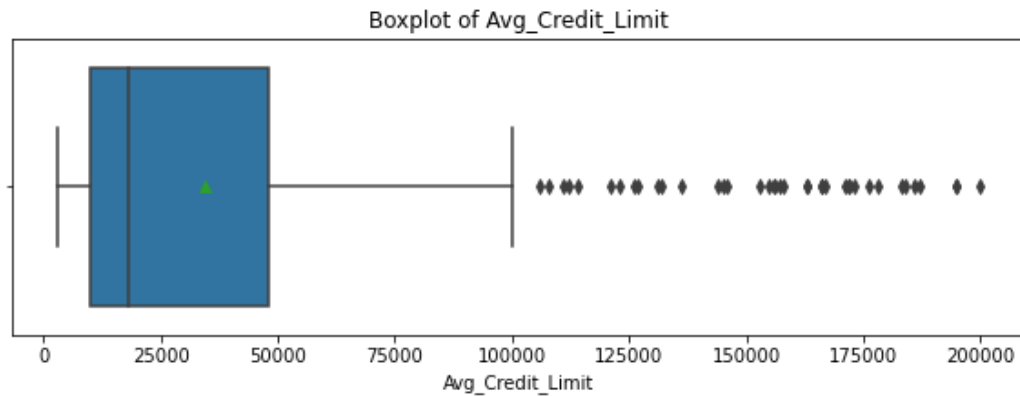
Hence there is a need to identify different customer segments, so that recommendations can be provided to these segments to better market and service them.

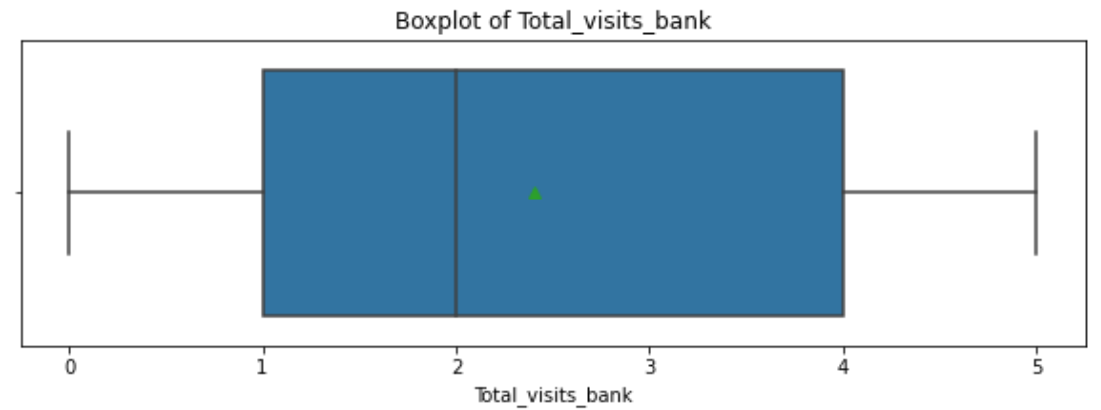
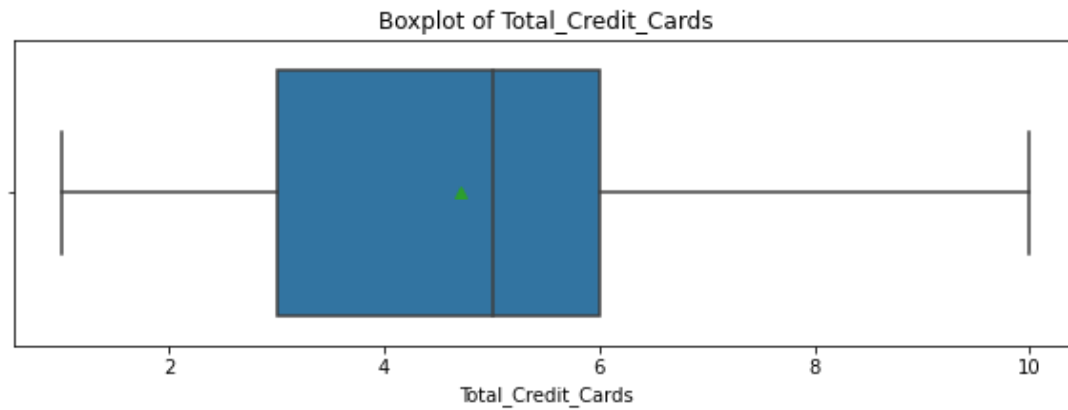
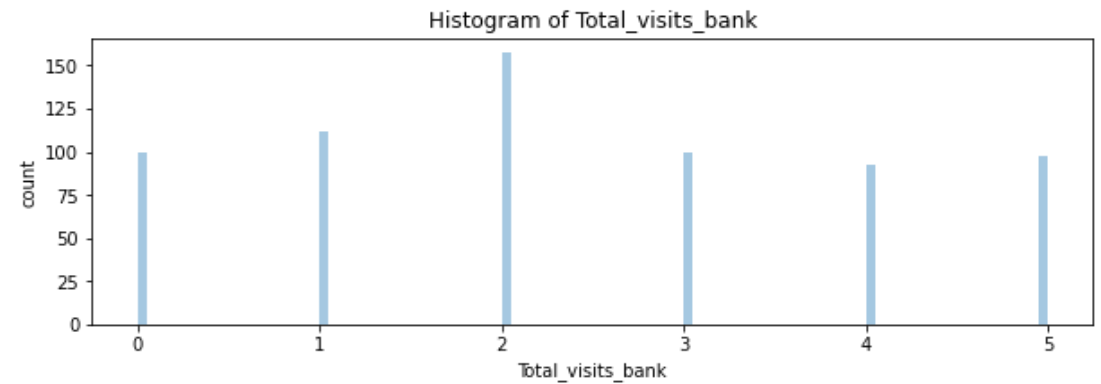
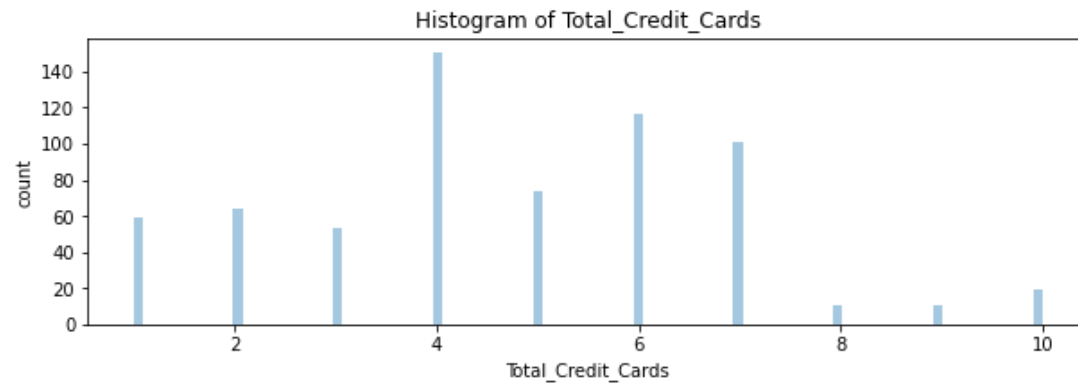
Data Overview

- The data contains information about 660 customers and 7 variables.
- The data include customer information such as customer identification number, average credit limit, number of credit cards, total visits to bank, total visits online and total calls made to the bank.
- The record primary key and customer identification number from the data will be dropped as it is unique for each customer/record and will not add value to our analysis.
- There are no missing values and no outliers from the dataset.
- There are few duplicate values of customer identification number in the dataset, and it was decided that all the data should be recorded as the values of the data with the same customer identification number are completely different.
- The data will be scaled as a certain variable (Avg_Credit_Limit) is very large compared to other variables.

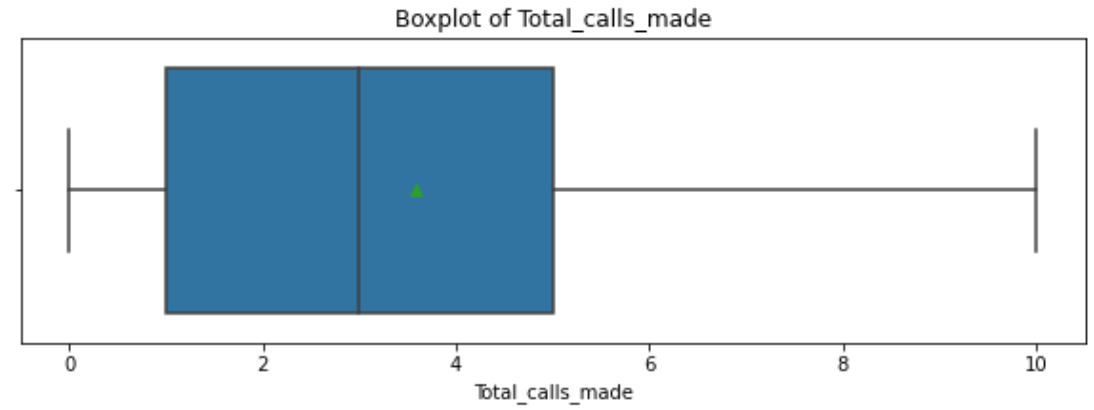
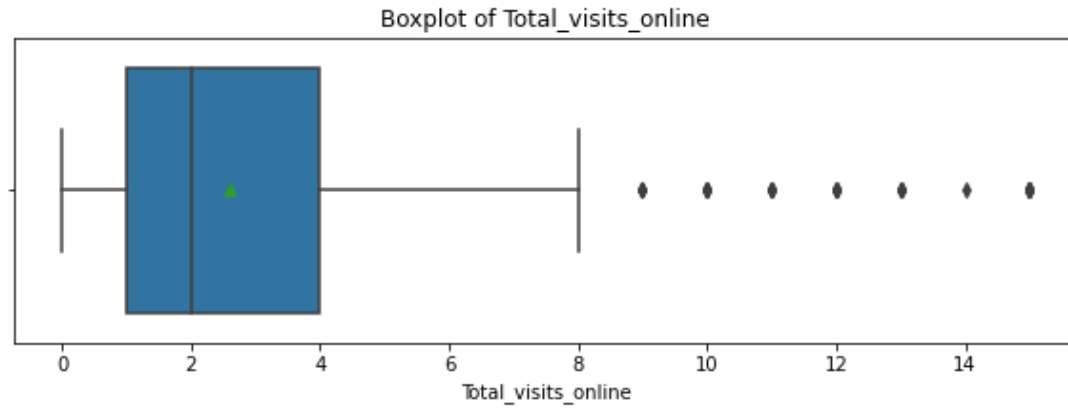
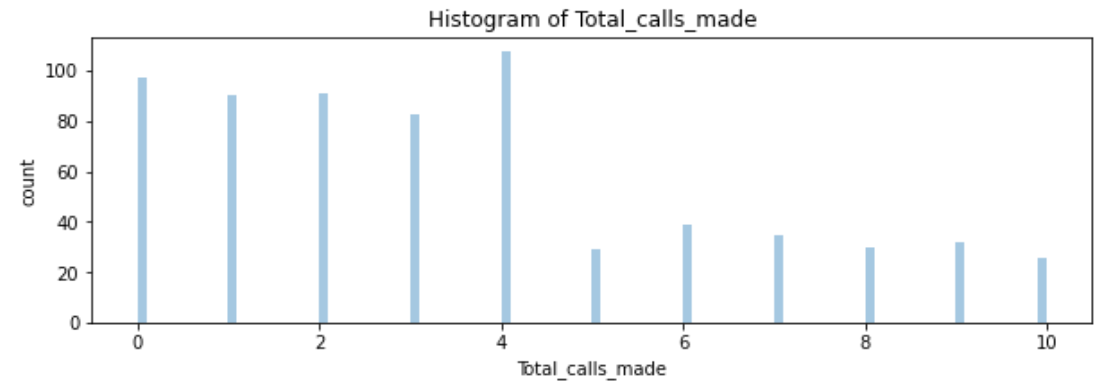
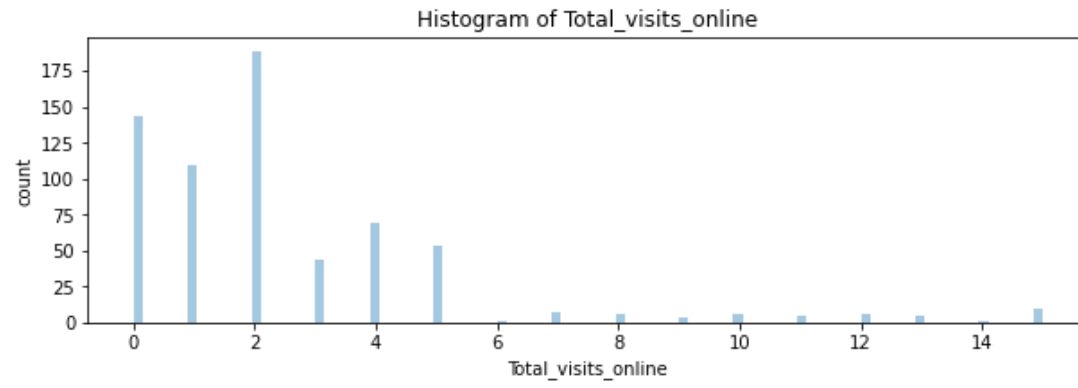
Exploratory Data Analysis

- Avg_Credit_Limit is right skewed.
- We decided to keep the outliers from the boxplot as they could form a certain cluster.





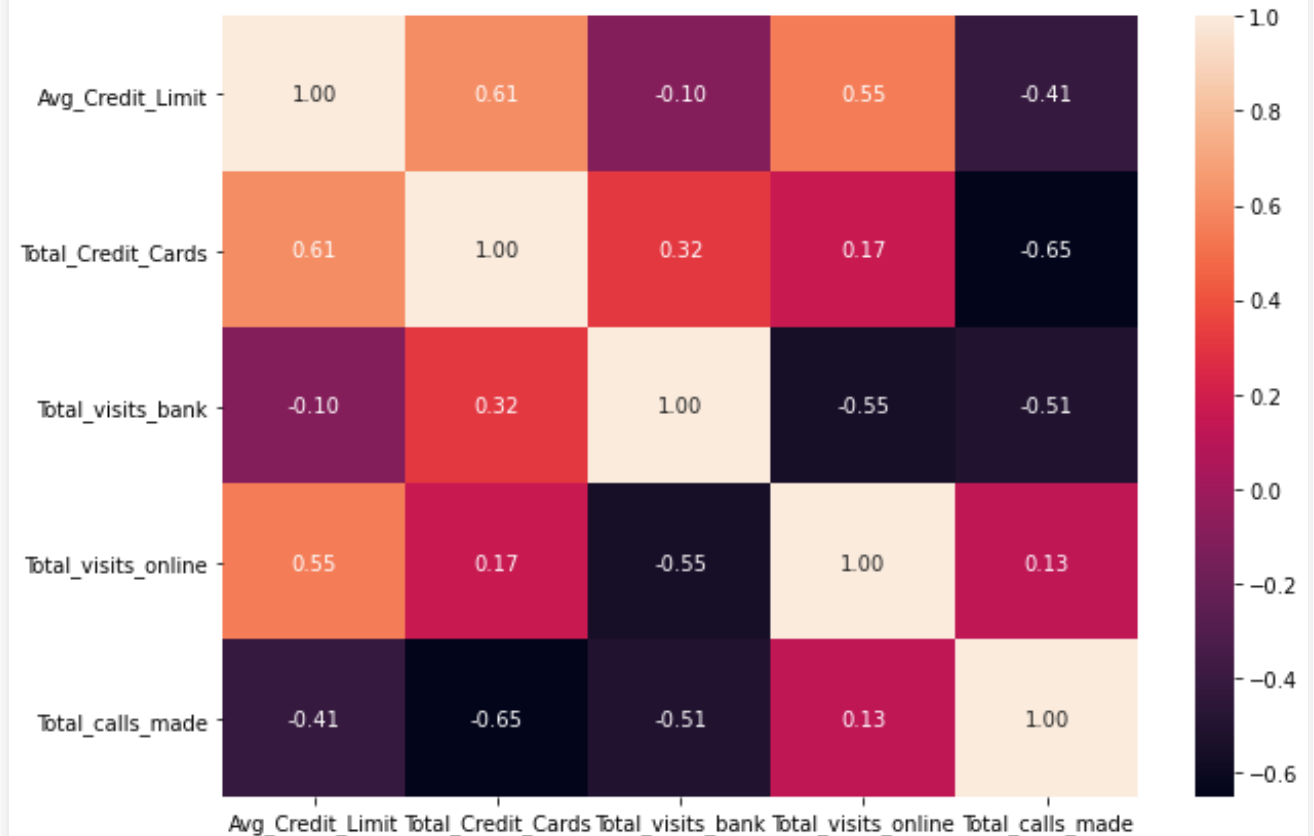
Exploratory Data Analysis



Exploratory Data Analysis

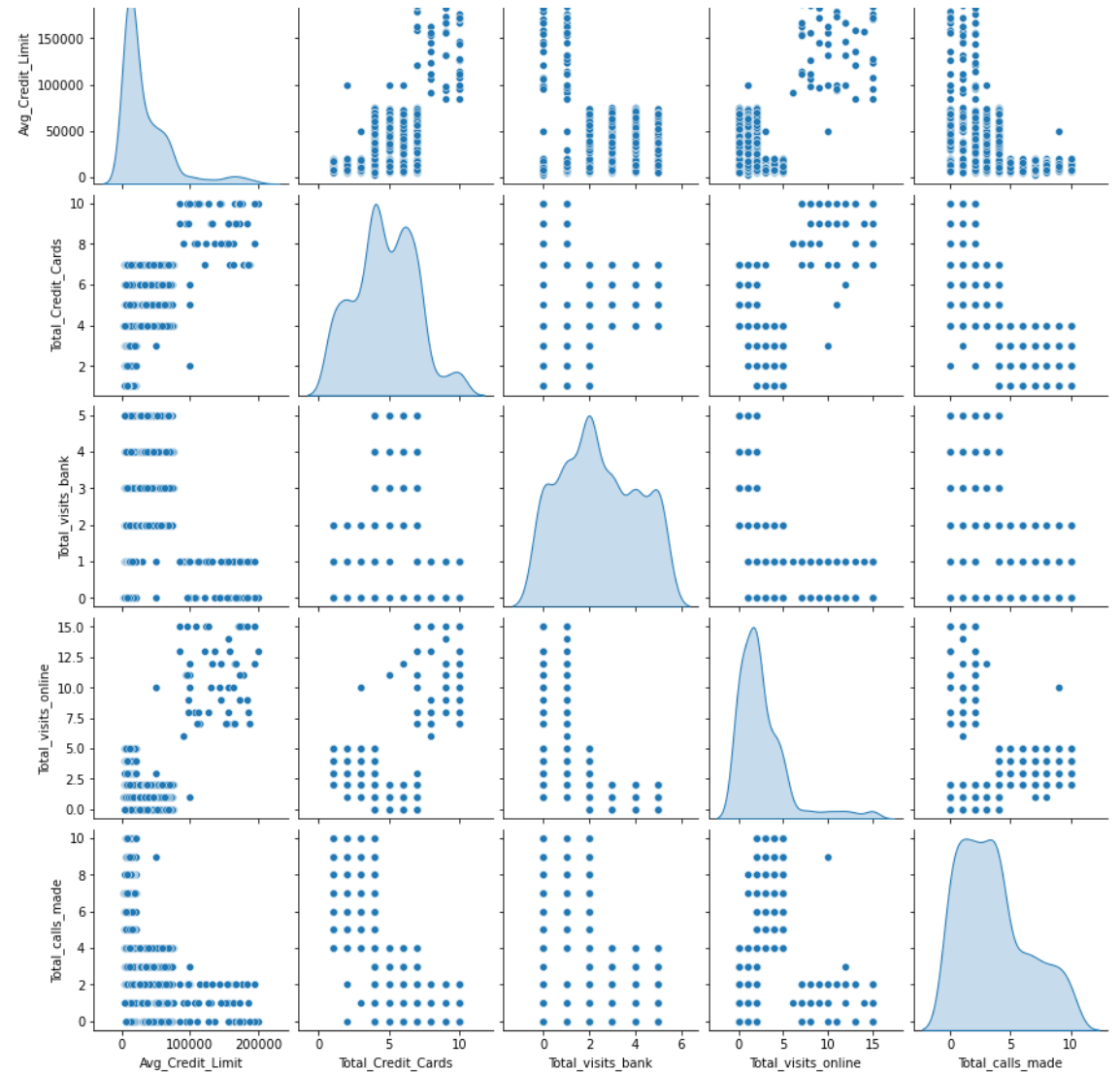
Exploratory Data Analysis - Heatmap

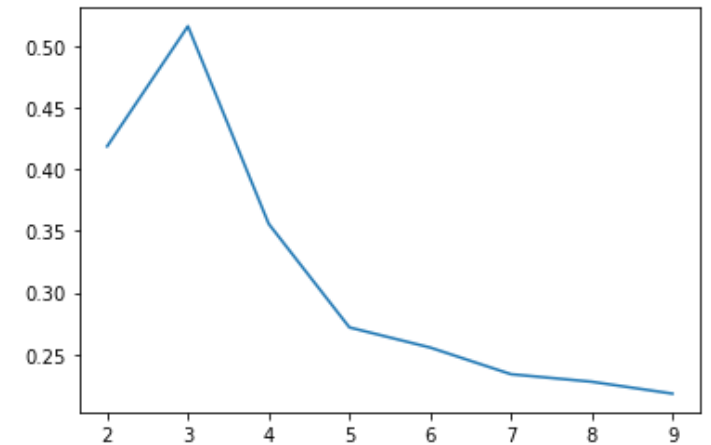
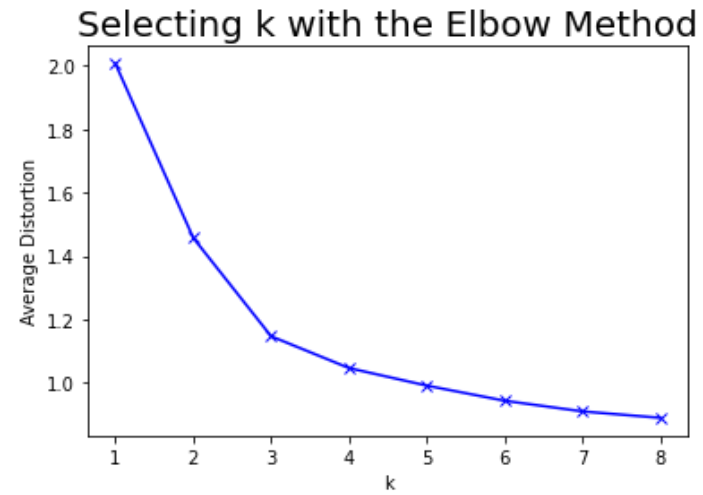
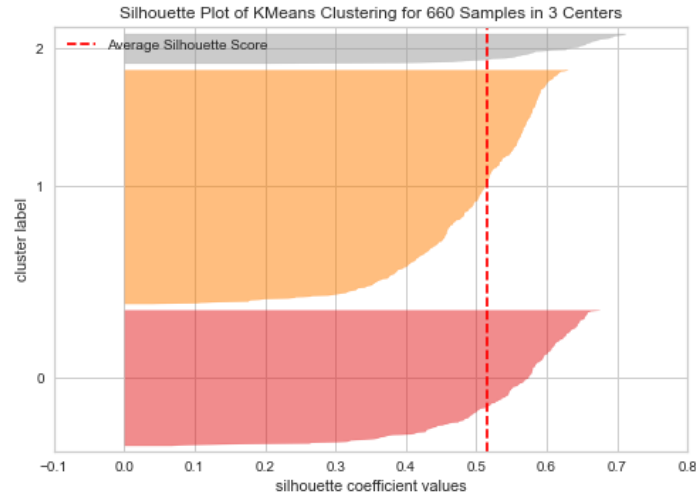
- No multicollinearity is detected in the data.
- Positive correlation:
 - Total_Credit_Cards and Avg_Credit_Limit
 - Total_visits_online and Avg_Credit_Limit
- Negative correlation:
 - Total_calls_made and Total_Credit_Cards
 - Total_visits_online and Total_visits_bank
 - Total_calls_made and Total_visits_bank



Exploratory Data Analysis - Pairplot

- There seems to be at least 2 clusters visible from the pairplot.
- However, pairplot can only compare 2 dimensions at a time, and there is a possibility that more than 2 clusters are formed.

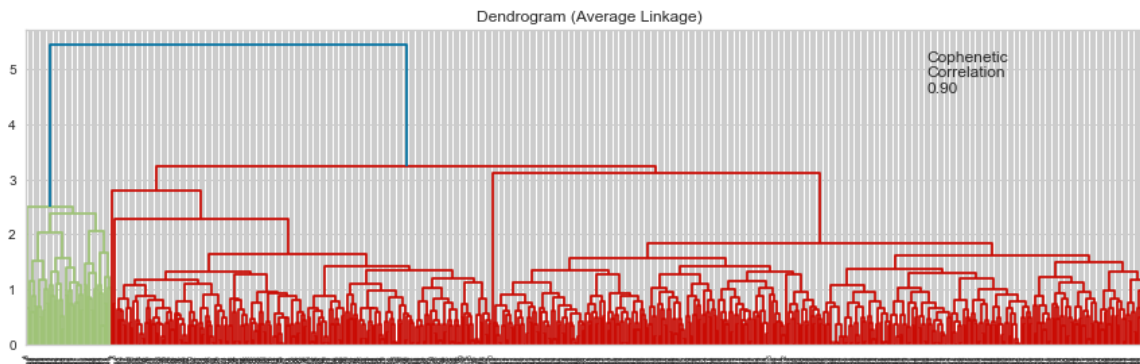




K-Means Clustering

- Using elbow method, it seems that k value of 3 or 4 is good.
- Seeing silhouette scores, it seems that k value of 3 is good.
- Hence k value of 3 is selected for the model.

Hierarchical Clustering

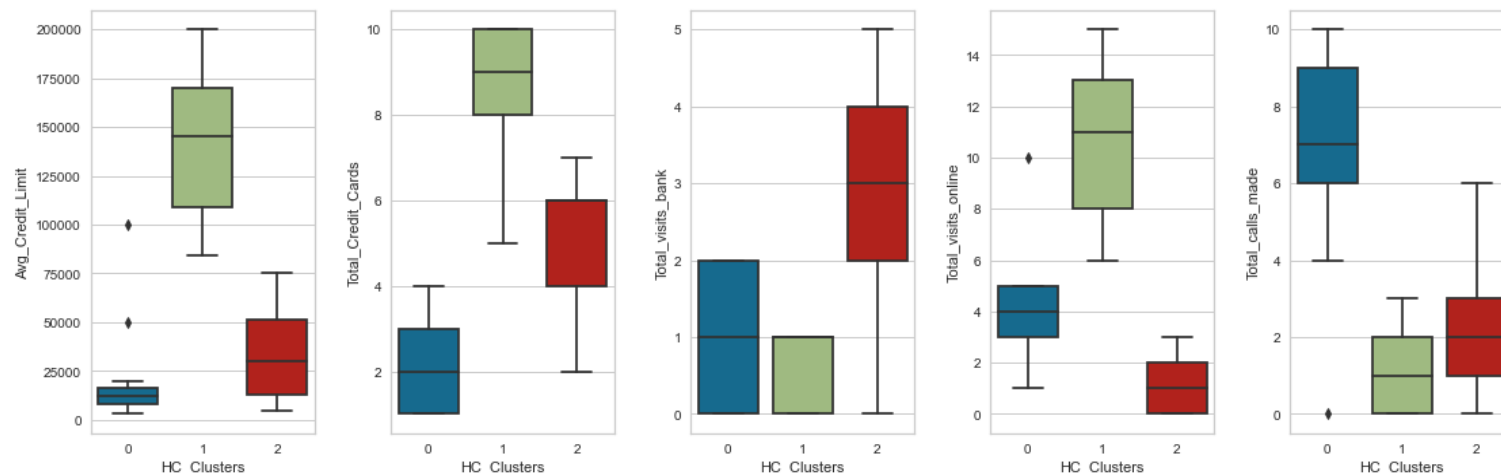


- Using the cophenetic correlation for various distance metrics and linkage methods, it was found that the highest cophenetic correlation score is the highest with Euclidean distance and Average linkage.
- Upon creation of hierarchical clustering models using various linkage methods, it was found that the data tends to gather around 3 clusters.

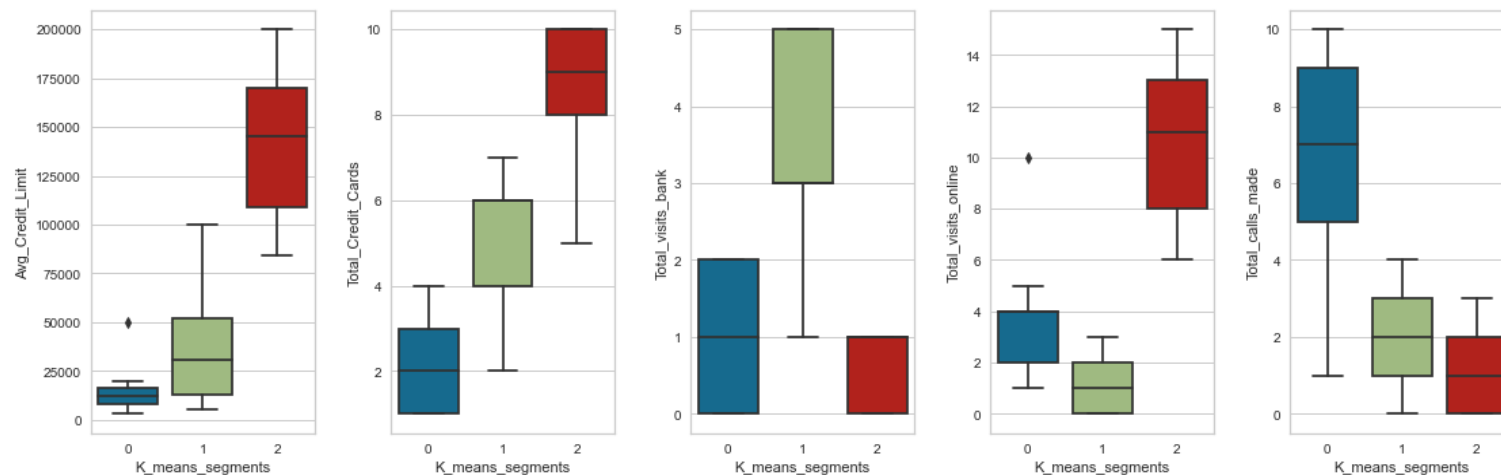
Cluster Profiling

The cluster profiles for the clusters made using k-means clustering and hierarchical clustering are similar.

Boxplot of numerical variables for each cluster



Boxplot of numerical variables for each cluster



Cluster Profiling

Clusters 0:

- Has the lowest Avg_Credit_Limit
- Has the lowest Total_Credit_Cards
- Has a somewhat low Total_Visits_bank
- Has a somewhat low Total_visits_online
- Has the highest Total_calls_made

Cluster 1:

- Has the highest Avg_Credit_Limit
- Has the highest Total_Credit_Cards
- Has the lowest Total_Visits_bank
- Has the highest Total_visits_online
- Has the lowest Total_calls_made

Cluster 2:

- Avg_Credit_Limit higher than Cluster 0 but far lower than Cluster 2
- Has medium number of Total_Credit_Cards
- Has the highest Total_Visits_bank
- Has the lowest Total_visits_online
- Has a somewhat low Total_calls_made

Business Insights & Recommendations

Cluster 0 – Non-breadwinners/new customers

- Maintaining a high level of customer satisfaction from customer service calls.
- As they are comfortable with calls, it is also possible to contact them via calls to promote some bank products to ensure customer loyalty to the bank.

Cluster 1 – Busy yet loyal breadwinners

- Improving the UI for online banking, whether by computer or phone. It is important to make sure these customers are comfortable using the online services.
- Providing loyalty benefits to the customers, such as mileage points and offer them premium membership.

Cluster 2 – Bank visitors

- Provide excellent customer service on retail banks
- Promote products that ensure savings for retirement
- If there are lots of customers from this segment, it's possible to gain data on location and build bank branches that are close to their location.