AllLife Bank Customer Segmentation

Business Presentation

Contents

Business Problem Overview and Solution Approach

- AllLife Bank wants to focus on its credit card customer base in the next financial year. They have been advised by their marketing research team that penetration in the market can be improved. The Marketing team proposes to run personalized campaigns to target new customers as well as upsell to existing customers.

- Another insight from the market research was that the customers perceive the support services of the back end poorly. The Operations team wants to upgrade the service delivery model, to ensure that customer queries are resolved faster. Head of Marketing and Head of Delivery both decide to reach out to the Data Science team for help.

Business Problem Overview and Solution Approach

- The objective of the model is to:
  - Identify different segments in the existing customer, based on their spending patterns and past interaction with the bank using clustering algorithms.
  - Recommend to the bank on how to better market to and service these customers.

Data Overview

- The data contains information about 660 customers and their characteristics.
- The characteristics include Sl_No primary key of the records, Customer key, Average credit limit, Total number of credit cards, Total annual visits to the bank, Total annual visits to the bank online assets, Total annual calls made to the bank or its customer service.
- Sl_No primary key and Customer key variables were dropped from the data as they are unique for almost every to all customers and will not add value to the analysis.
- Outliers in Average_Credit_Limit and Total_visits_online variables were treated and capped.

EDA – Correlation Matrix

Observations

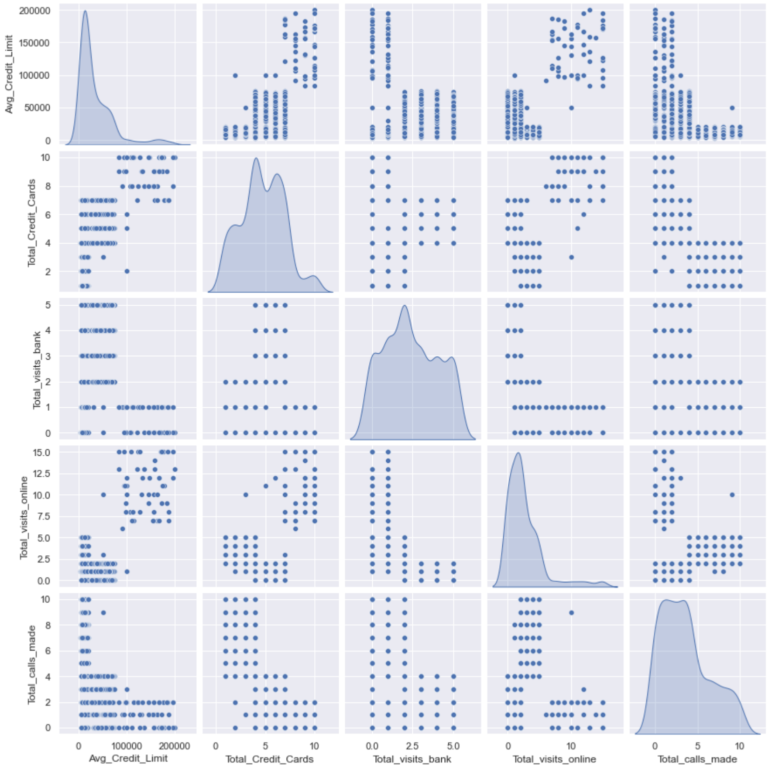
- Total_Credit_Cards and Avg_Credit_Limit are correlated (0.61) suggesting that customers who have higher number of credit cards are able to secure a higher average credit limit based on their higher income bracket and perhaps credit rating.

- Total_visits_online and Avg_Credit_Limit are correlated (0.55) suggesting those with higher average credit limit accessed the bank services more online.

- Total_calls_made is negatively correlated to Total_Credit_Cards (-0.65), Total_visits_bank (-0.51) and Avg_Credit_Limit (-0.41) suggesting that the more calls made, the less total number of credit cards customers will have with the bank, less visits perhaps to sign up or utilize more personal services and gain less average credit limit perhaps due to less credit cards with the bank.

- Total_visits_online and Total_visits_bank seem negatively correlated (-0.55) which is perhaps logical given that customers who can use the banks services online will less likely visit the bank personally and vice versa.

EDA – Pairplot



Observations

- There appears to be distinct groups of customers with higher average credit limits using the bank online services more while those with lower average credit limits use online services less.
- Customers with higher number of credit cards also use more bank online services and vice versa forming 2 distinct groups.
- Customers also made less calls when they use more bank online services.
- Customers with higher average credit limit have higher numbers of credit cards.
- From the kde plots with the number of peaks shown, it is estimated that at least 2 to 3 clusters are needed.

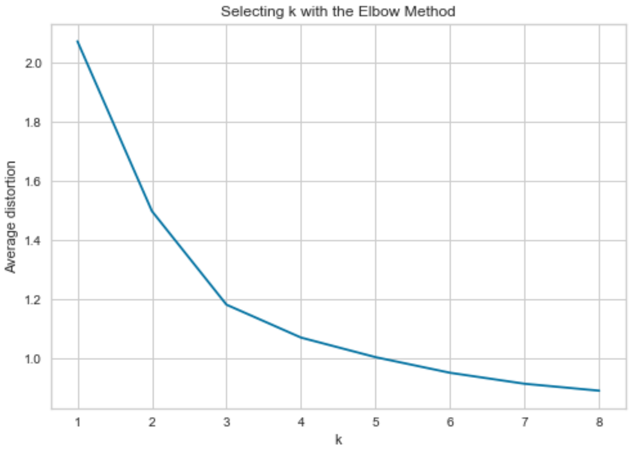
Model Performance Summary – Approach

- Customer data was subjected to standard scaling
- Apply K-Means Clustering to scaled customer data
  - Apply Elbow Method and Silhouette Scoring to determine number of clusters
  - Fit K-Means algorithm on scaled data
  - Profile customers in K-Means clusters and generate insights

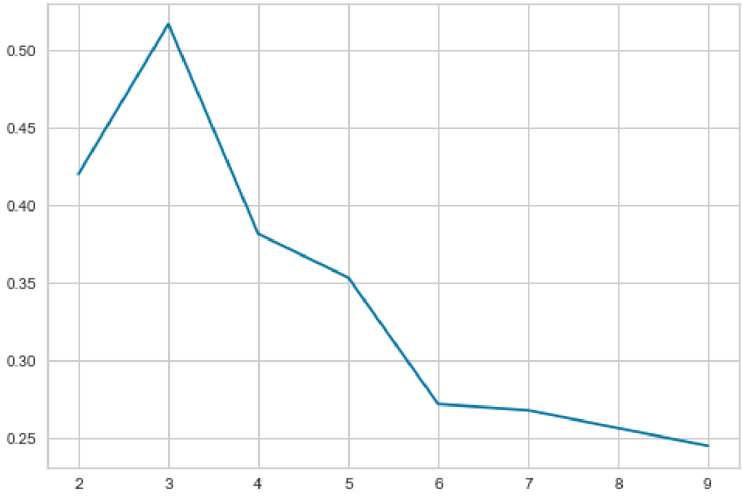
Model Performance Summary – Approach

- Apply Hierarchical Clustering to scaled customer data
  - Apply Cophenetic Correlation to find the best distance metric and linkage method to determine clusters
  - Build dendrograms for best distance metric and find the best combination of linkage method and distinct clusters to determine number of clusters
  - Fit Hierarchical Agglomerative Clustering algorithm on scaled data
  - Profile customers in Hierarchical Clustering clusters and generate insights
- Compare both clustering generated customer segment clusters and consolidate insights

Model Performance Summary – K-Means Evaluation



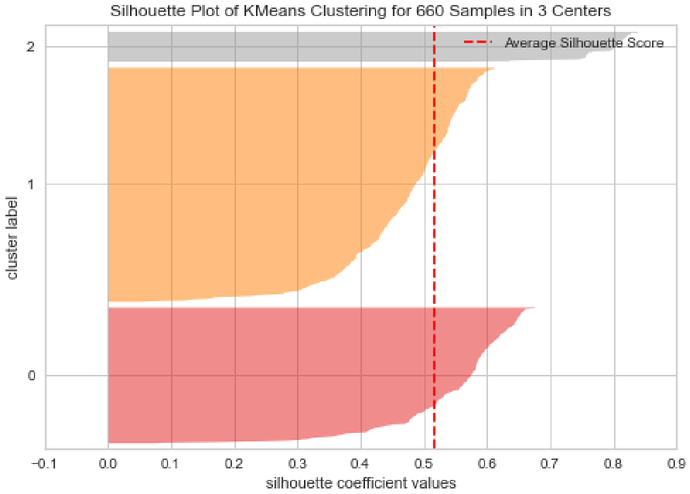
- K-Means Elbow Method
  - 2 – 3 clusters will be ideal for K-Means Clustering segmentation



- K-Means Silhouette Score
  - 3 clusters will be the most ideal for K-Means Clustering segmentation

Model Performance Summary – K-Means Evaluation

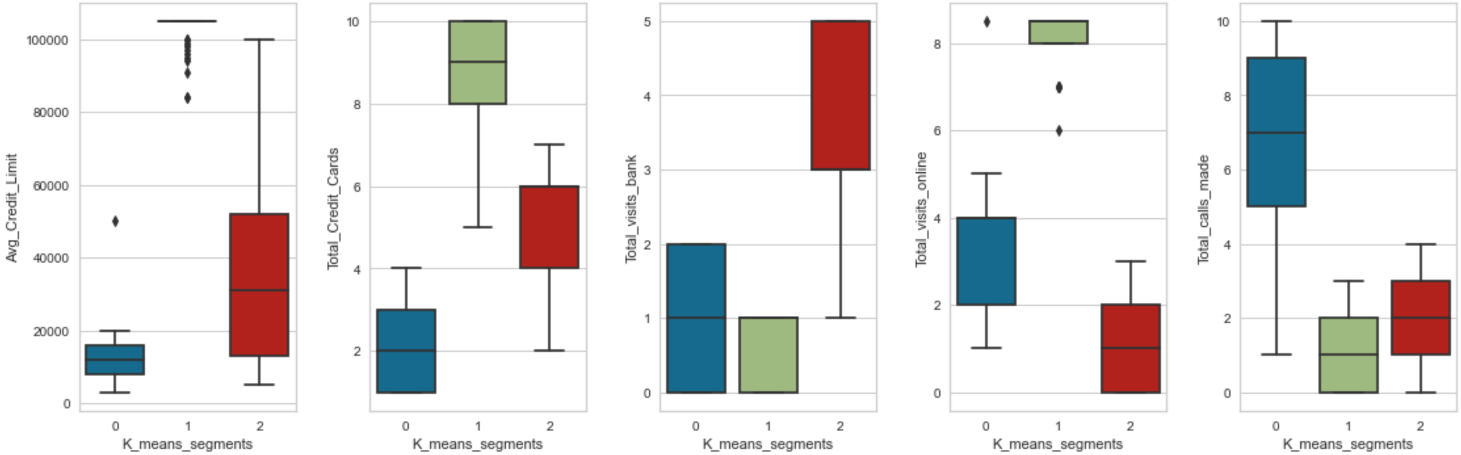
- Silhouette Visualizer for K-Means 2 clusters



- Silhouette Visualizer for K-Means 3 clusters

Model Performance Summary – K-Means Customer Profile

Boxplot of original numerical variables for each cluster



K_means_segments	Avg_Credit_Limit	Total_Credit_Cards	Total_visits_bank	Total_visits_online	Total_calls_made	count_in_each_segment
0	12174.11	2.41	0.93	3.55	6.87	224
1	102660.00	8.74	0.60	8.18	1.08	50
2	33782.38	5.52	3.49	0.98	2.00	386

Model Performance Summary – K-Means Customer Profile

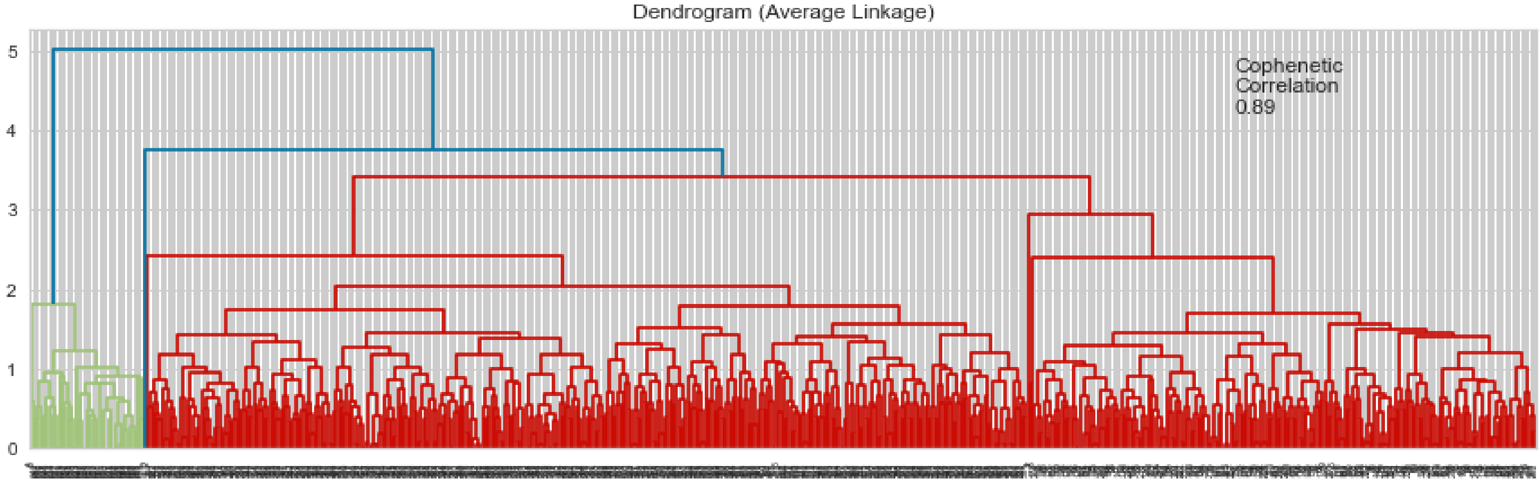
- Cluster 0: The Dissatisfied Budget Customer
  - This cluster has the lowest Average Credit Limit.
  - They also have the lowest number of credit cards.
  - They visit the bank annually ranging from 0 to 2 times and averaging 1 time.
  - 50% of them visit or do online logins 2 to 4 times to the bank's website yearly.
  - They clock the highest number of calls to the banks with 50% of them making 5 to 9 calls yearly.

- Cluster 1: The Tech Savvy Premium Customer
  - This cluster has the highest Average Credit Limit.
  - They also have the highest number of credit cards.
  - They visit the bank the lowest amount of times a year but visit or do online logins the most averaging 8 times yearly.
  - They make the lowest number of calls to the bank so it may mean they are pretty satisfied with the services.

Model Performance Summary – K-Means Customer Profile

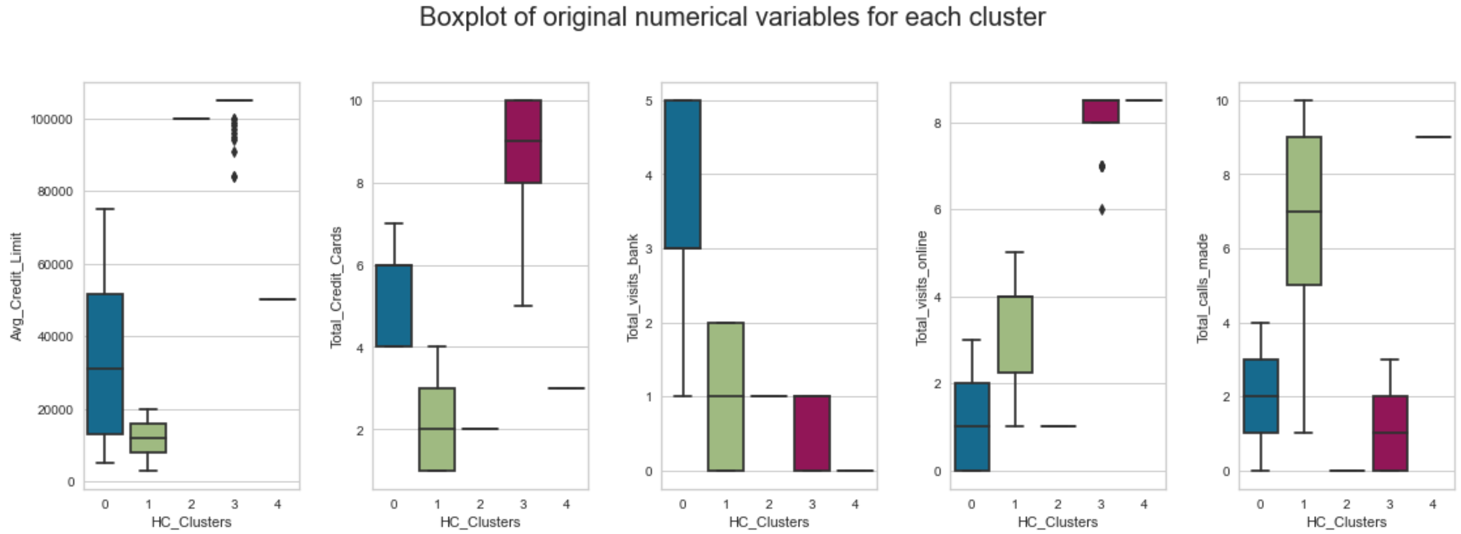
- Cluster 2: The Personable Mass Affluent Customer
  - This cluster has the widest ranging average credit limit customers.
  - 50% of them have 4 to 6 credit cards.
  - They visit the bank the most per year at 3 to 5 times but visit or do online logins the least averaging once a year and mostly do online up to 2 times.
  - They appear to be satisfied with the services going by the low average number of calls a year to the bank.

Model Performance Summary – Hierachical Evaluation



- The Cophenetic Correlation is highest for average, centroid and weighted linkage methods with Euclidean distance metric.
- We will move ahead with average linkage.
- 5 appears to be the appropriate number of clusters from the dendrogram for average linkage.

Model Performance Summary – Hierachical Customer Profile



HC_Clusters	Avg_Credit_Limit	Total_Credit_Cards	Total_visits_bank	Total_visits_online	Total_calls_made	count_in_each_segments
0	33541.45	5.52	3.49	0.98	2.01	386
1	12027.03	2.40	0.93	3.53	6.87	222
2	100000.00	2.00	1.00	1.00	0.00	1
3	102660.00	8.74	0.60	8.18	1.08	50
4	50000.00	3.00	0.00	8.50	9.00	1

Model Performance Summary – Hierachical Customer Profile

- Cluster 0: The Personable Mass Affluent Customer
  - This cluster has the widest ranging average credit limit customers.
  - 50% of them have 4 to 6 credit cards.
  - They visit the bank the most per year at 3 to 5 times but visit or do online logins the least averaging once a year and mostly do online up to 2 times.
  - They appear to be satisfied with the services going by the low average number of calls a year to the bank.

- Cluster 1: The Dissatisfied Budget Customer
  - This cluster has the lowest Average Credit Limit.
  - They also have the lowest number of credit cards.
  - They visit the bank annually ranging from 0 to 2 times and averaging 1 time.
  - 50% of them visit or do online logins 2 to 4 times to the bank's website yearly.
  - They clock the highest number of calls to the banks with 50% of them making 5 to 9 calls yearly.

Model Performance Summary – Hierachical Customer Profile

- Cluster 3: The Tech Savvy Premium Customer
  - This cluster has the highest Average Credit Limit.
  - They also have the highest number of credit cards.
  - They visit the bank the lowest amount of times a year but visit or do online logins the most averaging 8 times yearly.
  - They make the lowest number of calls to the bank so it may mean they are pretty satisfied with the services.

- Clusters 2 and 4 have only 1 customer each and will not be considered for analysis.

Model Performance Summary – Cluster Profiling Comparison

K_means_segments	Avg_Credit_Limit	Total_Credit_Cards	Total_visits_bank	Total_visits_online	Total_calls_made	count_in_each_segment
0	12174.11	2.41	0.93	3.55	6.87	224
1	102660.00	8.74	0.60	8.18	1.08	50
2	33782.38	5.52	3.49	0.98	2.00	386

HC_Clusters	Avg_Credit_Limit	Total_Credit_Cards	Total_visits_bank	Total_visits_online	Total_calls_made	count_in_each_segments
0	33541.45	5.52	3.49	0.98	2.01	386
1	12027.03	2.40	0.93	3.53	6.87	222
2	100000.00	2.00	1.00	1.00	0.00	1
3	102660.00	8.74	0.60	8.18	1.08	50
4	50000.00	3.00	0.00	8.50	9.00	1

- On close observation, clusters 0, 1 and 2 from K-Means Clustering can be mapped directly to clusters 1, 3 and 0 from Hierarchical Clustering.
- Apart from the mean of Avg_Credit_Limit having minute value differences, other variable values are similar or very close to each other between mapped clusters from K-Means Clustering and Hierarchical Clustering.

Model Performance Summary – Overall Customer Profile

- **The Dissatisfied Budget Customer**
  - This cluster has the lowest Average Credit Limit.
  - They also have the lowest number of credit cards.
  - They visit the bank annually ranging from 0 to 2 times and averaging 1 time.
  - 50% of them visit or do online logins 2 to 4 times to the bank's website yearly.
  - They clock the highest number of calls to the banks with 50% of them making 5 to 9 calls yearly.

- **The Personable Mass Affluent Customer**
  - This cluster has the widest ranging average credit limit customers.
  - 50% of them have 4 to 6 credit cards.
  - They visit the bank the most per year at 3 to 5 times but visit or do online logins the least averaging once a year and mostly do online up to 2 times.
  - They appear to be satisfied with the services going by the low average number of calls a year to the bank.

Model Performance Summary – Overall Customer Profile

- **The Tech Savvy Premium Customer**
  - This cluster has the highest Average Credit Limit.
  - They also have the highest number of credit cards.
  - They visit the bank the lowest amount of times a year but visit or do online logins the most averaging 8 times yearly.
  - They make the lowest number of calls to the bank so it may mean they are pretty satisfied with the services.

Business Insights and Recommendations

- **The Dissatisfied Budget Customer**
  - This cluster contains budget customers with the lowest number of credit cards.
  - They registered the highest amount of calls to the bank which may suggest dissatisfaction with the bank services in this customer segment.
  - Numbering 1/3 of all customers, attention needs to be paid to this group of budget customers on improving quality of services.

Business Insights and Recommendations

- **The Personable Mass Affluent Customer**
    - This cluster has the mass affluent band of customers constituting slightly over half of all customers.
    - This band of customers with the average credit limit of 75 percentile and above can be upsell to premium customers as their credit limit rivals that of the premium customers.
    - They can also be persuaded to have more credit cards with the bank and use more online banking services.
    - The opportunity to market it to them can be when they visit the banks as most of them in this profile prefer to and a more personable sales pitch can be conducted.

Business Insights and Recommendations

- **The Tech Savvy Premium Customer**
    - This cluster appears to contain the premium customers with the highest average credit limit and number of credit cards.
    - However their numbers are low at less than 10% of all customers.
    - Since they may bring more profit margin to the bank, their numbers have to be increased and the best opportunity comes from upselling to the Personable Mass Affluent Customer.

Business Insights and Recommendations

- Comments on additional data sources for model improvement
  - Additional data can be obtained from dissatisfied customers' feedback in the budget range to understand their pain points and design intervention methods to address their concerns.
  - Demographics of the customers can be obtained to help in targeted marketing techniques. Chances are the Personable Mass Affluent Customer segment has a high number of older customers prefer to visit the bank for their banking needs thus more effort can be spent on encouraging them to move online.

Business Insights and Recommendations

- Model implementation in real world and potential business benefits from model
  - The model implemented in the real world will help in finding out the customer segment to target with our customized upsell marketing efforts and generate more profit margin.
  - In addition, the concerns of the Dissatisfied Budget Customer can be uncovered and addressed by identifying them and gathering feedback from the right customers to improve service offerings and reduce customer churn.

Business Insights and Recommendations

- Other Recommendations
    - Identify current Tech Savvy Premium Customer who were upsold in the past vs. unsuccessful upsold customers can help in building a supervised learning predictive model to more accurately identify customers with high potential for a successful upsell effort.

greatlearning
Power Ahead

Happy Learning !