Business Report

UL Coded Project

PGPDSBA

Chithira Raj

Table of Contents

List of Tables

List of Figures

1. Context

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

2. Objective

To identify different segments in the existing customers, based on their spending patterns as well as past interaction with the bank, using clustering algorithms, and provide recommendations to the bank on how to better market to and service these customers.

3. Data Dictionary

The data provided is of various customers of a bank and their financial attributes like credit limit, the total number of credit cards the customer has, and different channels through which customers have contacted the bank for any queries (including visiting the bank, online, and through a call center).

S.No.	Variables	Description
1	Sl_No	Primary key of the records
2	Customer Key	Customer identification number
3	Average Credit Limit	Average credit limit of each customer for all credit cards
4	Total credit cards	Total number of credit cards possessed by the customer
5	Total visits bank	Total number of visits that the customer made (yearly) personally to the bank
6	Total visits online	Total number of visits or online logins made by the customer (yearly)
7	Total calls made	Total number of calls made by the customer to the bank or its customer service department (yearly)

Table 1: Data Dictionary

4. Data Overview

4.1. Import libraries and load the data

Sl_No	Customer Key	Avg_Credit_Limit	Total_Credit_Cards	Total_visits_bank	Total_visits_online	Total_calls_made
1	87073	100000	2	1	1	0
2	38414	50000	3	0	10	9
3	17341	50000	7	1	3	4
4	40496	30000	5	1	1	4
5	47437	100000	6	0	12	3

Figure 1: Data Overview

We will drop the Sl_No column as it adds no value to the analysis.

4.2. Check the structure of data

Shape of the dataset: 660 rows and 6 columns

4.3. Check the types of the data

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 660 entries, 0 to 659
Data columns (total 6 columns):
 #   Column               Non-Null Count  Dtype
---  ------               --------------  -----
 0   Customer Key         660 non-null    int64
 1   Avg_Credit_Limit     660 non-null    int64
 2   Total_Credit_Cards   660 non-null    int64
 3   Total_visits_bank    660 non-null    int64
 4   Total_visits_online  660 non-null    int64
 5   Total_calls_made     660 non-null    int64
dtypes: int64(6)
memory usage: 31.1 KB
```

Figure 2: Datatypes

4.4. Check for and treat (if needed) missing values

	0
Customer Key	0
Avg_Credit_Limit	0
Total_Credit_Cards	0
Total_visits_bank	0
Total_visits_online	0
Total_calls_made	0

Figure 3: Missing values check

4.5. Data Duplicates

There are no duplicate rows.

4.6. Statistical Summary

	count	mean	std	min	25%	50%	75%	max
Customer Key	660.0	55141.443939	25627.772200	11265.0	33825.25	53874.5	77202.5	99843.0
Avg_Credit_Limit	660.0	34574.242424	37625.487804	3000.0	10000.00	18000.0	48000.0	200000.0
Total_Credit_Cards	660.0	4.706061	2.167835	1.0	3.00	5.0	6.0	10.0
Total_visits_bank	660.0	2.403030	1.631813	0.0	1.00	2.0	4.0	5.0
Total_visits_online	660.0	2.606061	2.935724	0.0	1.00	2.0	4.0	15.0
Total_calls_made	660.0	3.583333	2.865317	0.0	1.00	3.0	5.0	10.0

Figure 4: Statistical Summary - Numeric

4.7. Insights

- Serial Number and Customer identification number are unique, hence will not add value to the modelling.
- The average credit limit is 34,574, but the maximum reaches 200,000. The wide range (3,000 to 200,000) and high standard deviation (37,625) indicate that there are customers with very high credit limits compared to others. This might imply the presence of both high-value customers and potentially risky customers within the dataset.
- Customers hold an average of approximately 4.7 credit cards, with most customers having between 3 to 6 cards. The maximum of 10 cards suggests that a segment of customers may be over-leveraged, which could be a target for risk assessment.
- On average, customers visit the bank about 2.4 times, with a range from 0 to 5 visits. A substantial proportion of customers may prefer online interactions, as indicated by the low average number of bank visits.
- The average number of online visits is slightly higher at 2.6, with a maximum of 15 visits. This suggests a growing trend toward online banking, which may influence customer engagement strategies.
- Customers make an average of 3.6 calls, with most calls falling between 1 to 5. The significant standard deviation indicates some customers may rely heavily on customer service, possibly reflecting dissatisfaction or a need for assistance.

5. Exploratory Data Analysis

5.1. Univariate Analysis

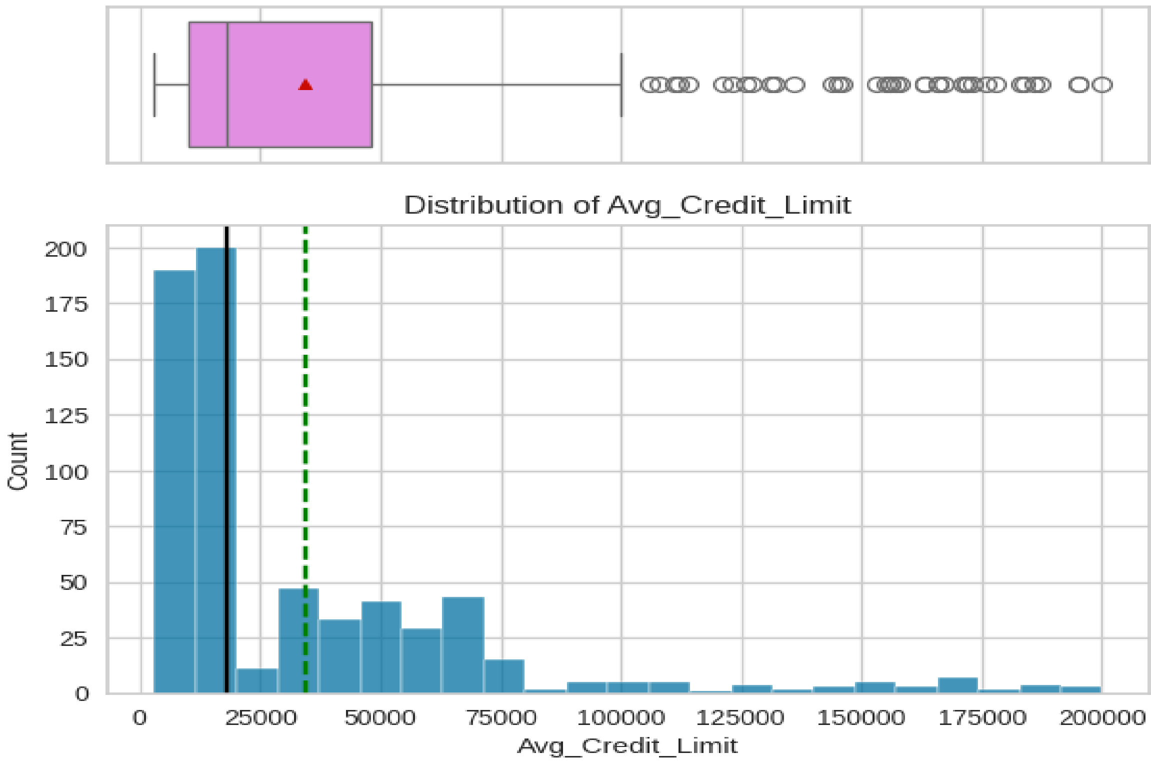

Figure 5: Avg_credit_Limit

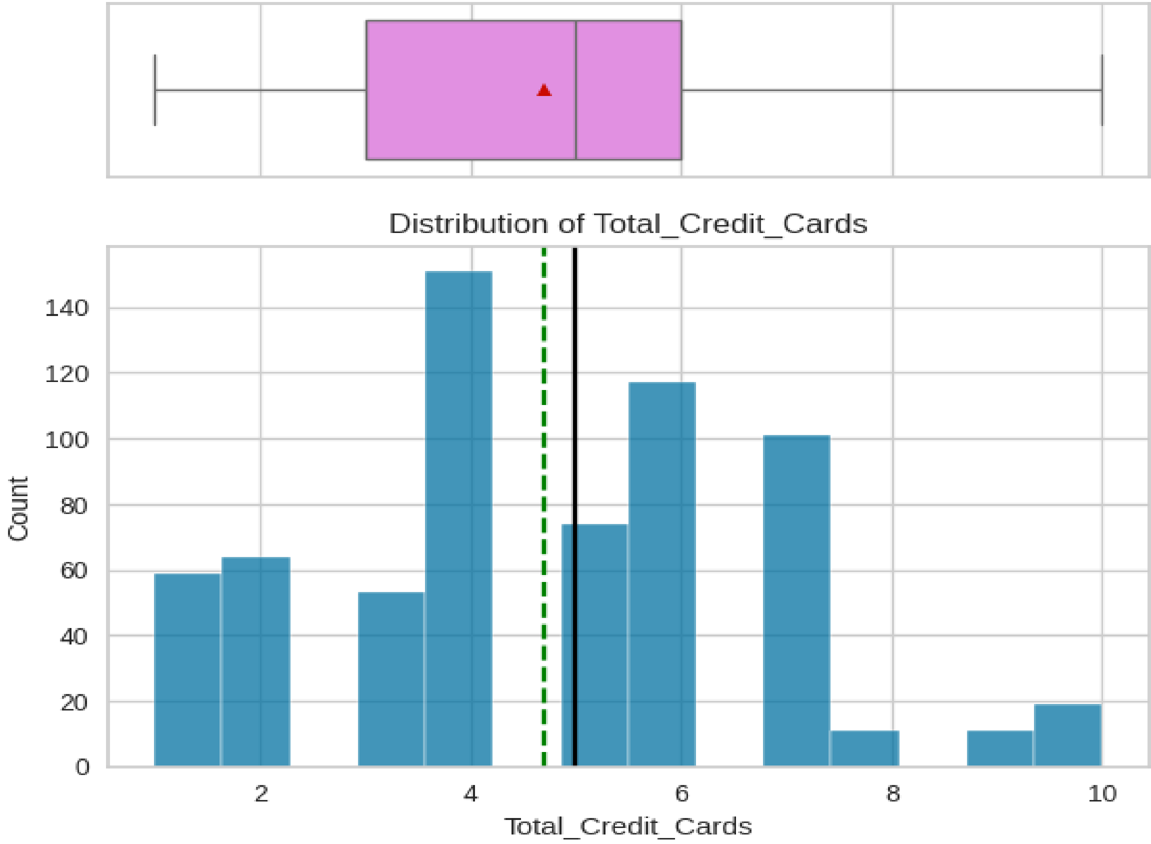

Figure 6: Total_Credit_Cards

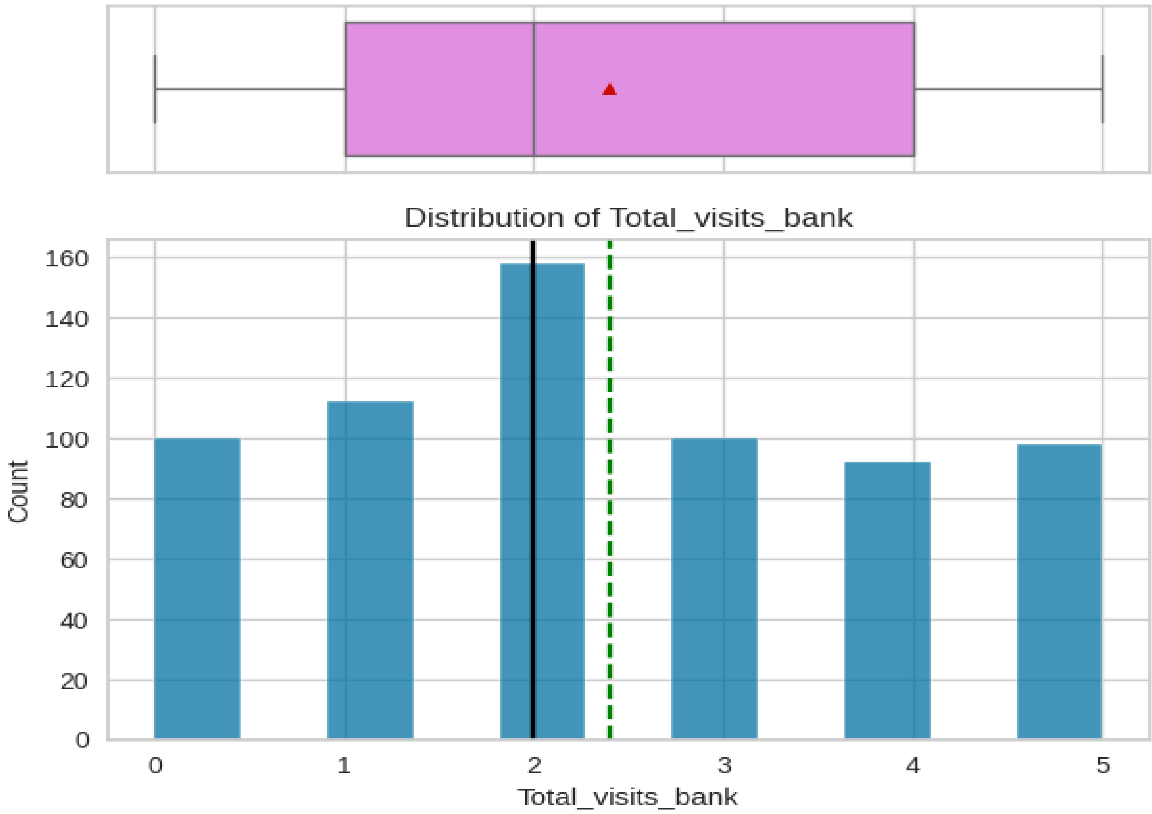

Figure 7:Total_visits_bank

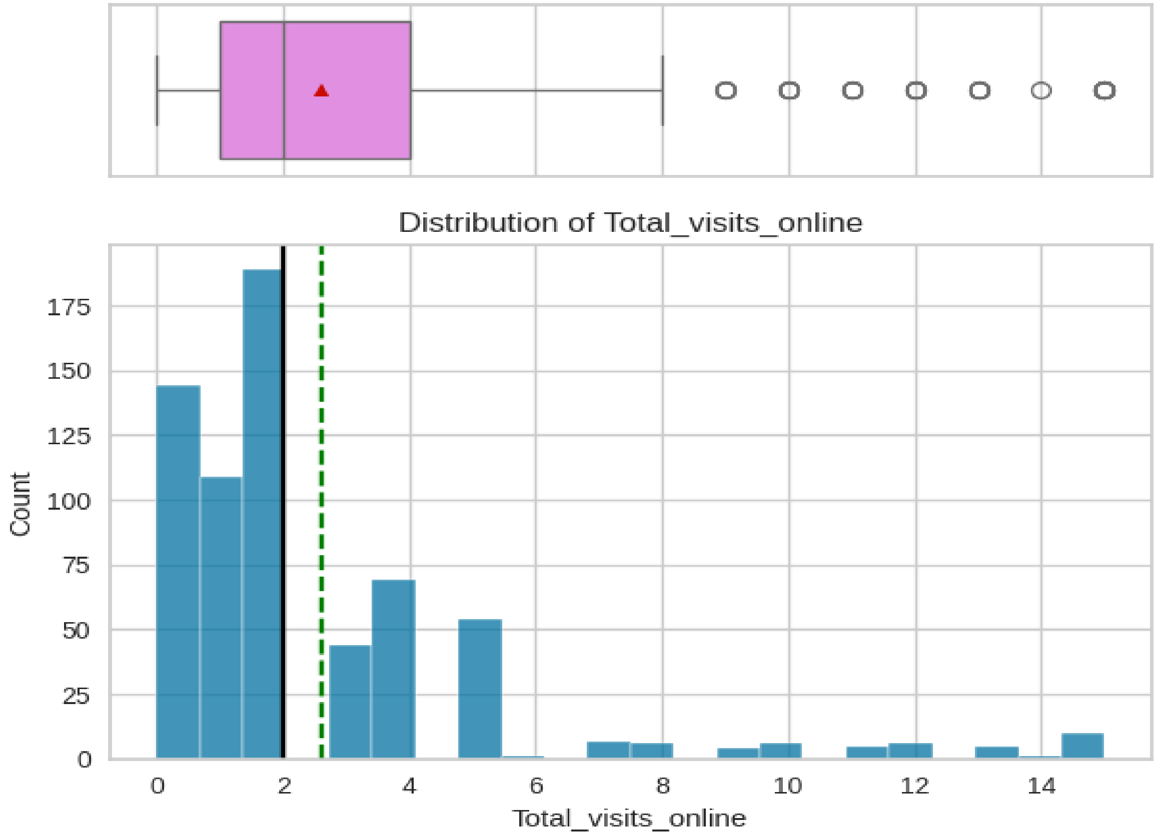

Figure 8: Total_visits_online

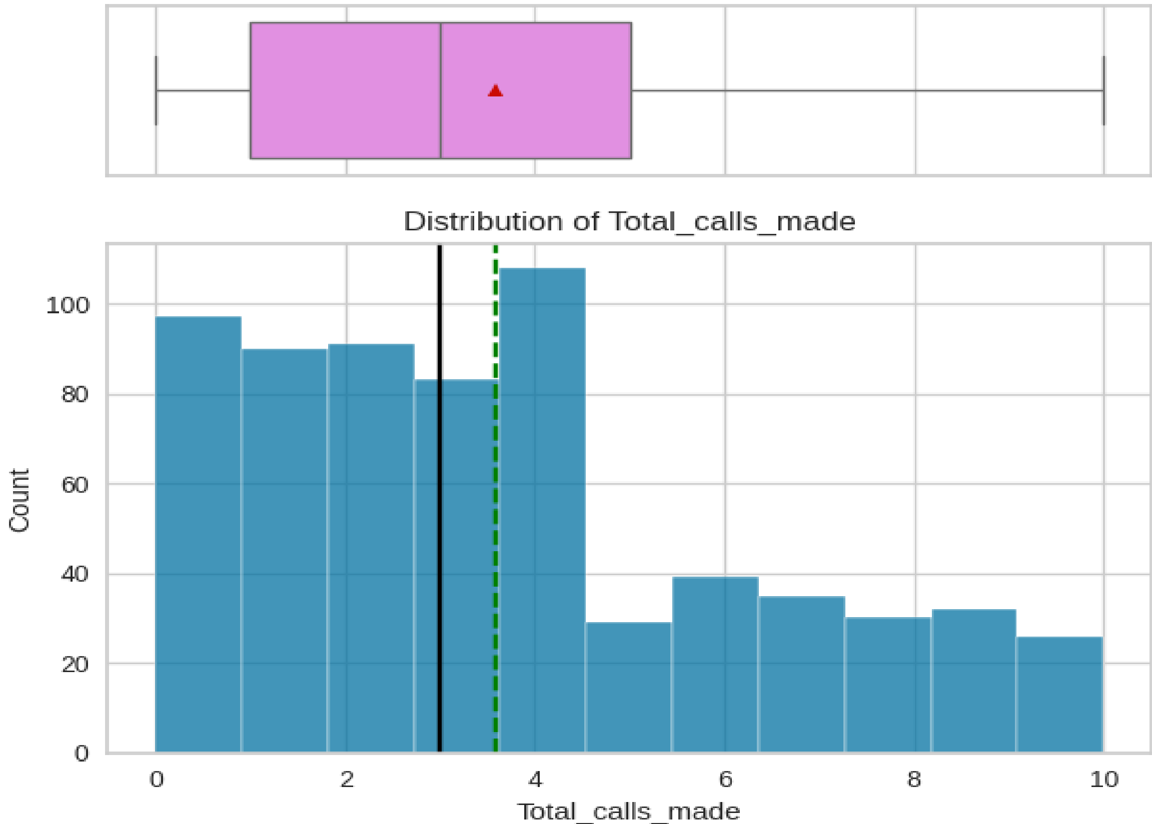

Figure 9: Total_calls_made

Insights

- The data shows a wide range of average credit limits among customers, with a significant portion having lower limits and a smaller group having much higher limits.
- The box plot suggests a right-skewed distribution, meaning there are a few customers with very high credit limits, pulling the tail to the right.
- The data shows a concentration of customers with a moderate number of credit cards.
- The histogram confirms the slight right-skewness observed in the box plot. The majority of customers have between 4 and 6 credit cards, with a decreasing frequency as the number of credit cards increases.
- The histogram confirms the slight right-skewness observed in the box plot. The majority of customers visit the bank between 2 and 3 times, with a decreasing frequency as the number of visits increases.
- The box plot suggests a slight right-skewness, meaning there are a few customers with a larger number of online visits, pulling the tail to the right.
- The median number of total visits online is around 2. This suggests that 50% of customers visit online two times or fewer.
- The histogram confirms the slight right-skewness observed in the box plot. The majority of customers make between 3 and 5 calls, with a decreasing frequency as the number of calls increases.

5.2. Bivariate Analysis

Correlation Check

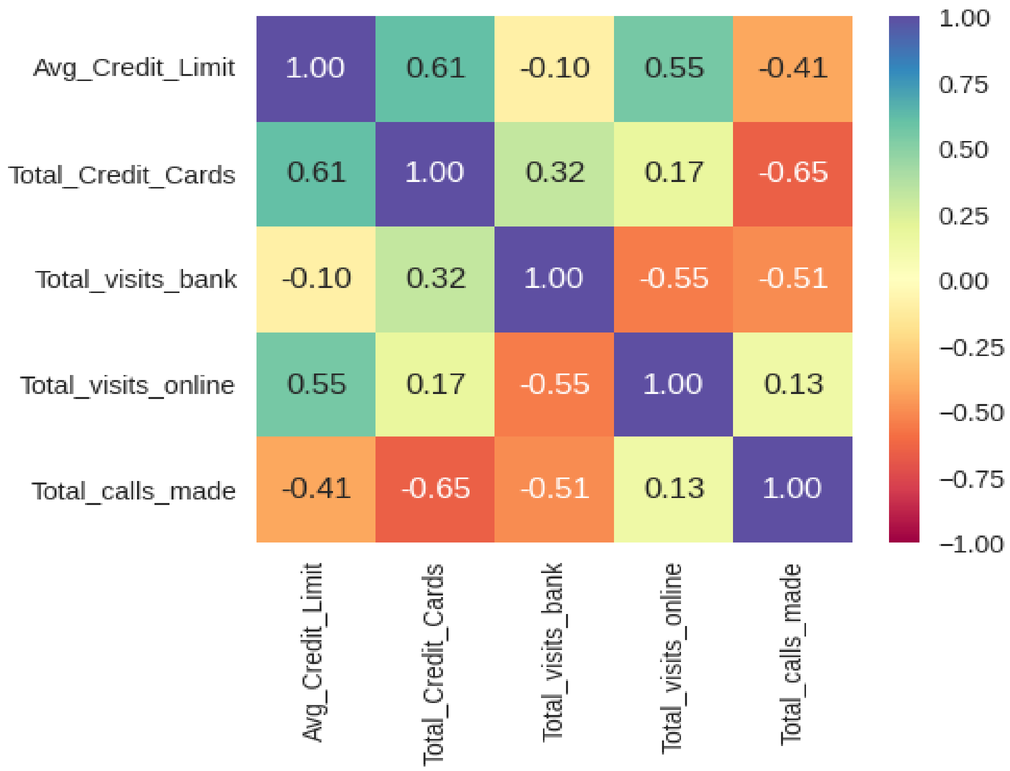

Figure 10: Heatmap

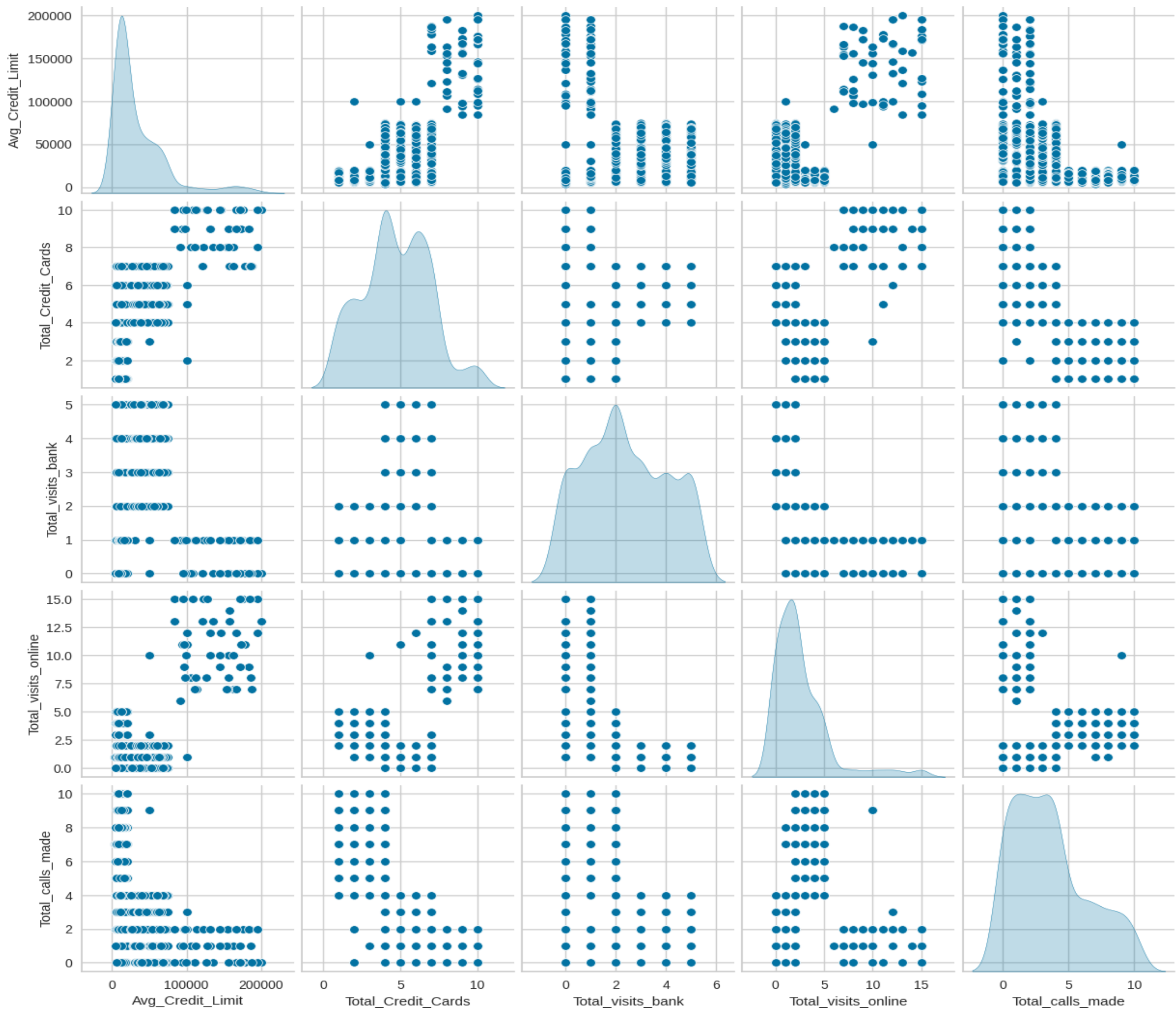

Figure 11: Pair Plot

Insights

Positive Correlations:

- Avg_Credit_Limit and Total_Credit_Cards: There's a strong positive correlation between average credit limit and the number of credit cards, suggesting that customers with higher credit limits tend to have more credit cards.
- Total_visits_online and Total_calls_made: A moderate positive correlation exists between online visits and calls made, indicating that customers who visit online more frequently are also more likely to make calls.

Negative Correlations:

- Total_visits_bank and Total_visits_online: There's a strong negative correlation between visits to the bank and online visits, suggesting that customers who visit the bank frequently tend to visit online less often.
- Total_visits_bank and Total_calls_made: A moderate negative correlation exists between bank visits and calls made, indicating that customers who visit the bank frequently are less likely to make calls.

No Significant Correlations: There are no significant correlations between Total_Credit_Cards and Total_visits_bank, Total_Credit_Cards and Total_visits_online, and Total_Credit_Cards and Total_calls_made.

6. Data Preprocessing

6.1. Missing Value treatment

There are no missing values.

6.2. Duplicate value check

There are no duplicate rows.

6.3. Outlier Detection

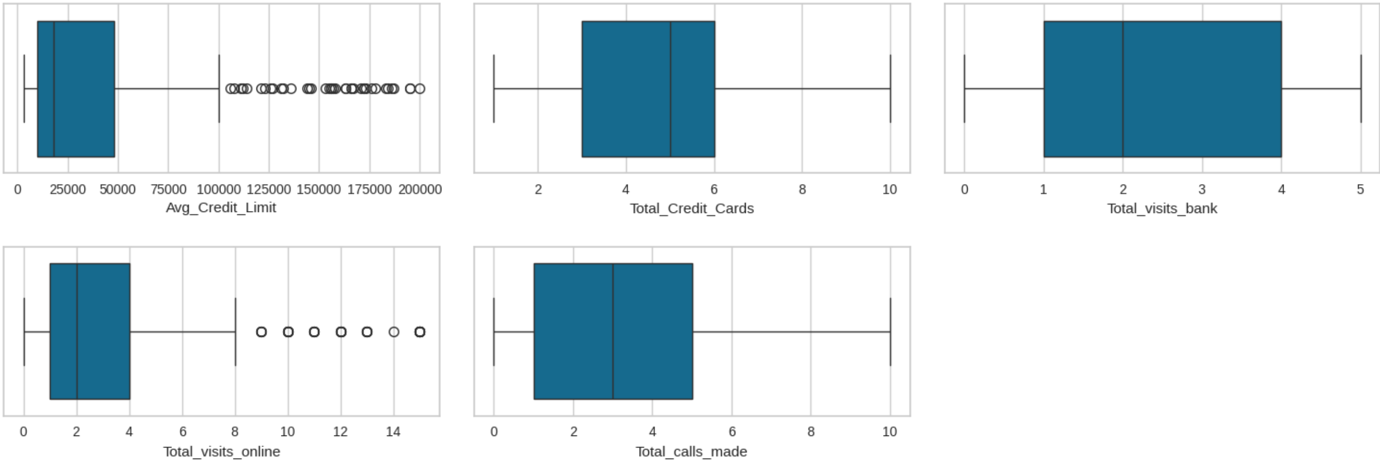

Figure 12: Outliers

There are outliers in the few columns. We have a few options for handling these outliers:

- Use the IQR (Interquartile Range) to determine the lower and upper bounds of the column and either replace or remove the outliers.
- These customers have average credit limits that are significantly higher than the rest of the population. However, this could reflect their genuine circumstances, so we shouldn't remove the outliers. It ultimately depends on each individual's financial and professional situation.

6.4. Feature Engineering

We are excluding the Customer Key from the DataFrame, as it does not provide significant value for the clustering analysis.

6.5. Scaling

Algorithms like K-means or hierarchical clustering calculate the distance between data points. If features are on different scales, the algorithm may produce biased clusters. Scaling brings features to a similar range, improving the accuracy of clustering. Features with larger scales (e.g., Credit limit versus total credit cards) can dominate distance metrics, leading the clustering algorithm to prioritize them. Scaling ensures all features have equal importance.

Avg_Credit_Limit	Total_Credit_Cards	Total_visits_bank	Total_visits_online	Total_calls_made
1.740187	-1.249225	-0.860451	-0.547490	-1.251537
0.410293	-0.787585	-1.473731	2.520519	1.891859
0.410293	1.058973	-0.860451	0.134290	0.145528
-0.121665	0.135694	-0.860451	-0.547490	0.145528
1.740187	0.597334	-1.473731	3.202298	-0.203739

Figure 13: Scaling

7. K-means Clustering

7.1. Plot the elbow curve

```
Number of Clusters: 1   Average Distortion: 2.006922226250361
Number of Clusters: 2   Average Distortion: 1.4571553548514269
Number of Clusters: 3   Average Distortion: 1.1466276549150365
Number of Clusters: 4   Average Distortion: 1.0463825294774463
Number of Clusters: 5   Average Distortion: 0.9984086474611271
Number of Clusters: 6   Average Distortion: 0.9651175868007105
Number of Clusters: 7   Average Distortion: 0.9259135376350698
Number of Clusters: 8   Average Distortion: 0.9211835468499175
Number of Clusters: 9   Average Distortion: 0.8752234180939921
Number of Clusters: 10  Average Distortion: 0.8505796034595803
Number of Clusters: 11  Average Distortion: 0.8350921089299926
```

Figure 14: Distortion

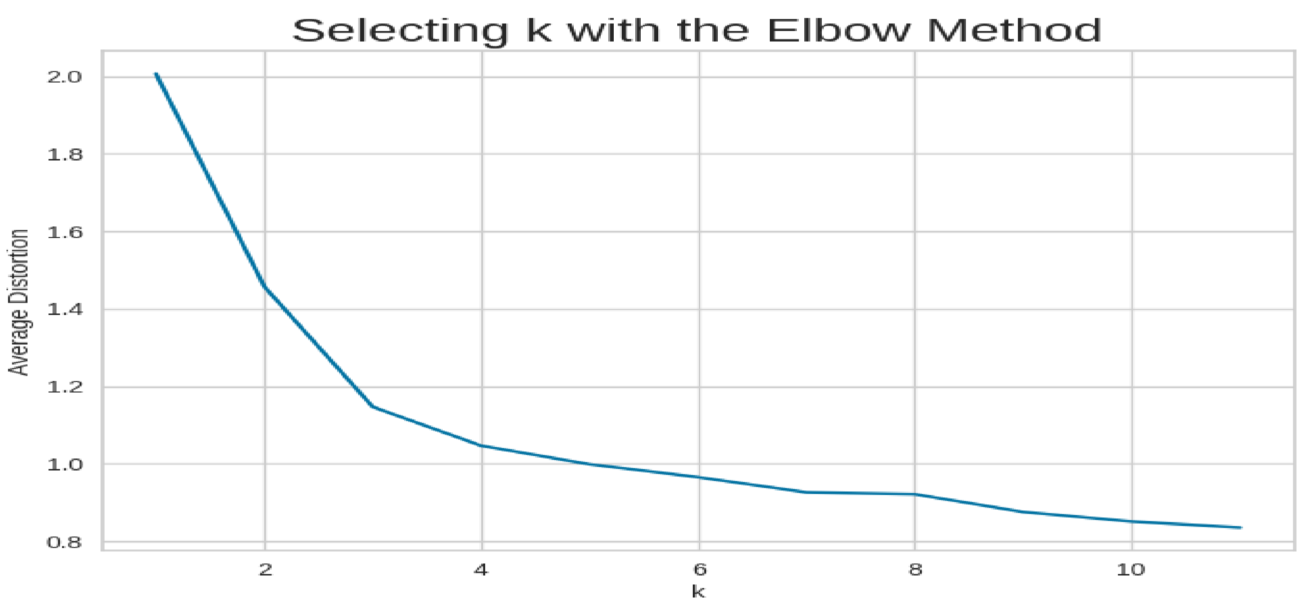

Figure 15: Elbow curve

The appropriate value of k from the elbow curve seems to be 3 or 4.

7.2. Let's check the silhouette scores

```
For n_clusters = 2, silhouette score is 0.41842496663215445
For n_clusters = 3, silhouette score is 0.5157182558881063
For n_clusters = 4, silhouette score is 0.3556670619372605
For n_clusters = 5, silhouette score is 0.3284672118706534
For n_clusters = 6, silhouette score is 0.24806123278744782
For n_clusters = 7, silhouette score is 0.24633384025835273
For n_clusters = 8, silhouette score is 0.23241103950361397
For n_clusters = 9, silhouette score is 0.2237028865804097
For n_clusters = 10, silhouette score is 0.21150786170622632
For n clusters = 11, silhouette score is 0.20800747719844545
```

Figure 16: Silhouette score

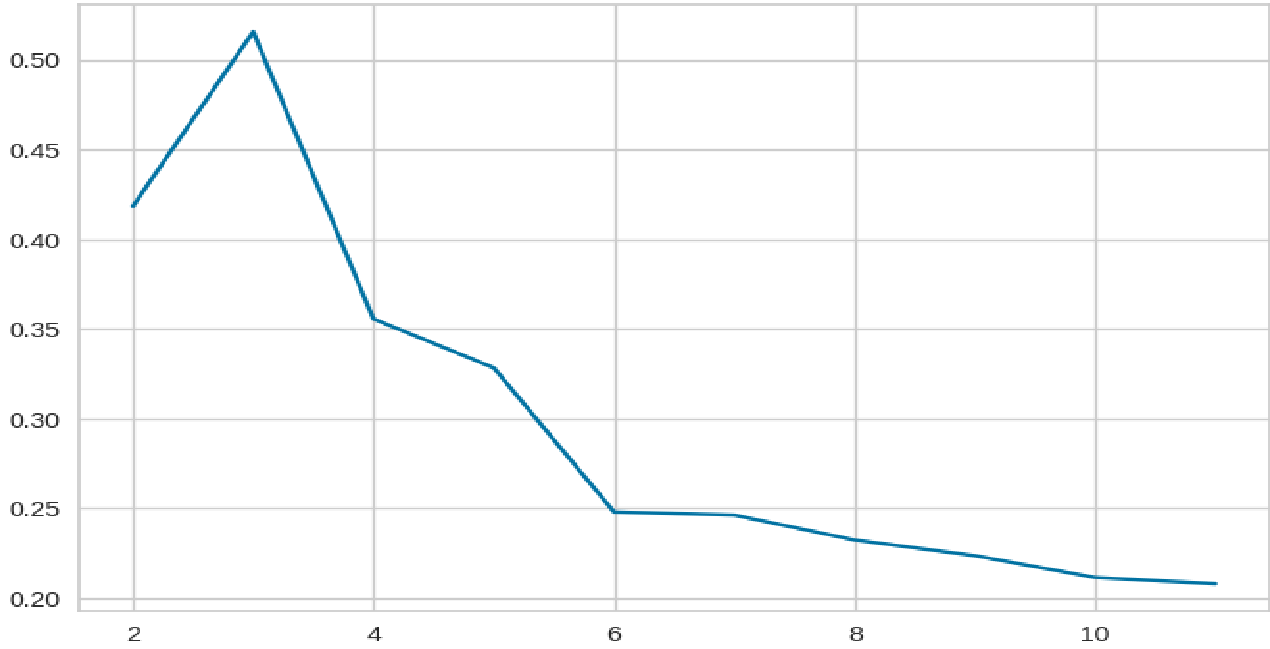

Figure 17: Silhouette Score plot

From the silhouette scores, it seems that 3 is a good value of k.

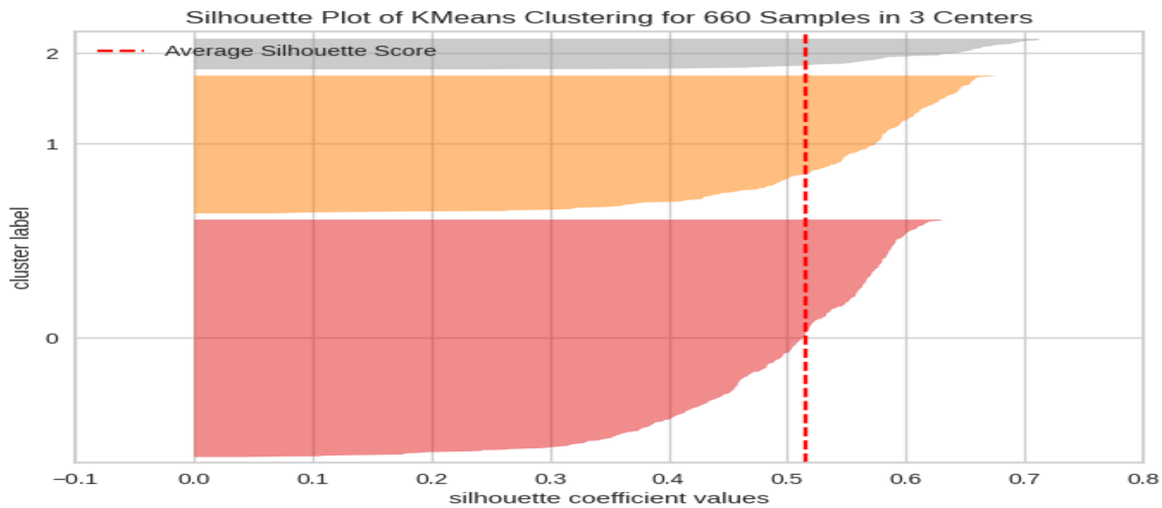

Figure 18: K-means - Silhouette Plot for 3 centers

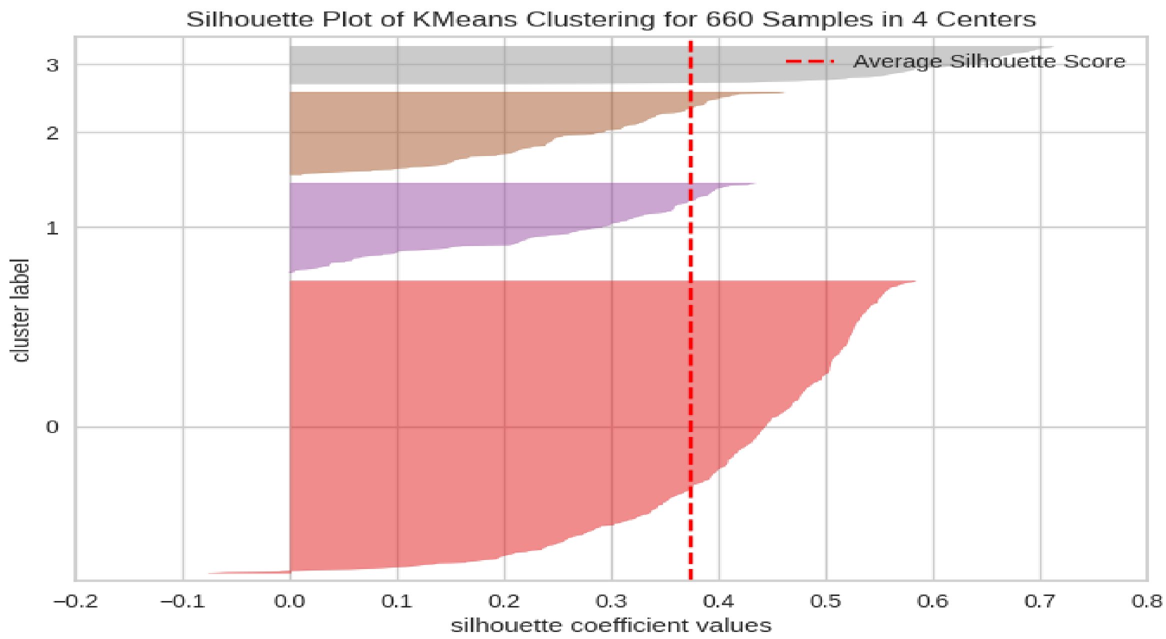

Figure 19: K-means - Silhouette Plot for 4 centers

Cluster-Specific Insights

For 3 Clusters:

- Cluster 2 has the highest silhouette scores, indicating a well-separated cluster.
- Clusters 0 and 1 show slightly more variation in silhouette values, implying moderate overlap or less-defined separation.

For 4 Clusters:

- Cluster 0 (red) seems to be the best-defined cluster, similar to the 3-cluster scenario.
- The other clusters, especially cluster 1 (purple), have lower average silhouette scores, indicating some overlap or poorly separated clusters.

7.3. Selecting final model

By combining the insights from both methods—a noticeable bend in the elbow curve and a high silhouette score at 3 clusters—we concluded that 3 is the optimal number of clusters for this dataset. This choice strikes a balance between model simplicity and cluster quality, ensuring that the clusters are well-defined and meaningful for further analysis.

7.3. Cluster Profiling

K_means_segments	count
0	386
1	224
2	50

Figure 20: Cluster Counts

K_means_segments	Avg_Credit_Limit	Total_Credit_Cards	Total_visits_bank	Total_visits_online	Total_calls_made	count_in_each_segment
0	33782.383420	5.515544	3.489637	0.981865	2.000000	386
1	12174.107143	2.410714	0.933036	3.553571	6.870536	224
2	141040.000000	8.740000	0.600000	10.900000	1.080000	50

Figure 21: K-means Cluster Profile

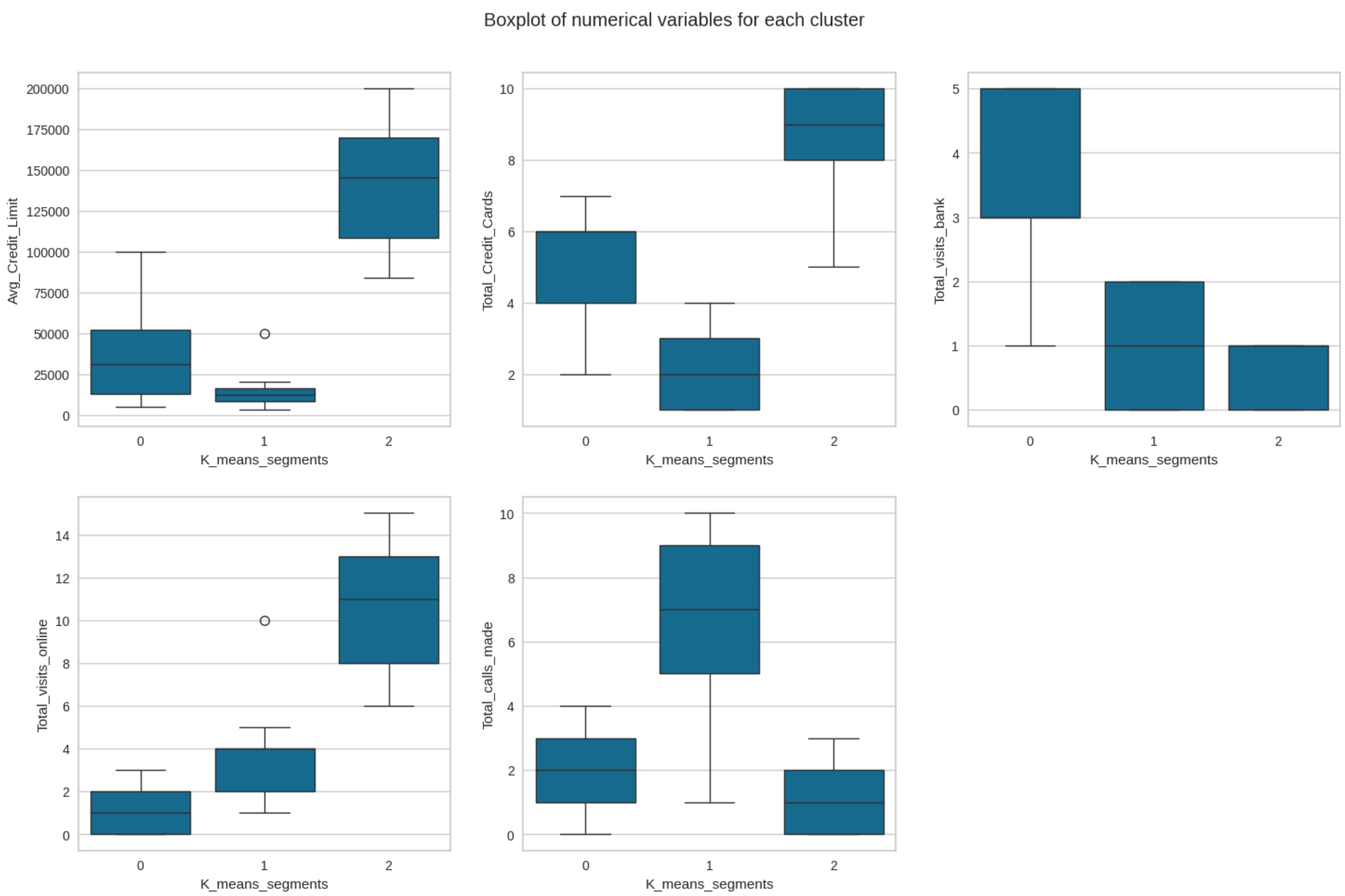

Figure 22: Boxplot for numerical variables in each K-means cluster

7.4. Insights

Cluster 0: Moderate Engagement

- Avg_Credit_Limit: Moderate credit limit, generally between 50,000 and 100,000.
- Total_Credit_Cards: Holds around 4–6 credit cards.
- Total_visits_bank: Frequently visits the bank, typically around 4–5 visits.
- Total_visits_online: Makes around 2–4 online visits, with low online engagement.
- Total_calls_made: Moderately engaged via calls, making 2–4 calls.

Summary: This cluster likely represents users with moderate engagement across all channels. They have a mid-range credit limit, visit the bank frequently, and show balanced online activity and call frequency.

Cluster 1: Low Engagement

- Avg_Credit_Limit: The lowest credit limit, ranging between 10,000 and 25,000.
- Total_Credit_Cards: Holds only 1–2 credit cards, suggesting minimal involvement with credit services.
- Total_visits_bank: Rarely visits the bank, with just 1 visit.
- Total_visits_online: Makes slightly more online visits than bank visits, around 3–4 times, but still relatively low.
- Total_calls_made: Rarely makes calls, with around 1–2 calls.

Summary: This cluster likely represents low-engagement customers with minimal credit limit, fewer credit cards, and limited interaction through both bank visits and calls. They may need targeted marketing to increase their engagement.

Cluster 2: High Engagement

- Avg_Credit_Limit: The highest credit limit, often above 150,000 and up to 200,000.
- Total_Credit_Cards: Holds around 8–10 credit cards, reflecting extensive use of credit products.
- Total_visits_bank: Very few bank visits (1 or fewer), indicating a preference for other channels of interaction.
- Total_visits_online: Extremely active online, with frequent visits (around 10–14).
- Total_calls_made: High engagement through calls, making 7–9 calls regularly.

Summary: This cluster represents high-value customers with significant credit limits, numerous credit cards, and a preference for online and phone interactions over bank visits. They are highly engaged and represent a prime target for premium products or services.

8. Hierarchical Clustering

8.1. Computing Cophenetic Correlation for each linkage method

```
Cophenetic correlation for Euclidean distance and single linkage is 0.7391220243806552.
Cophenetic correlation for Euclidean distance and complete linkage is 0.8599730607972423.
Cophenetic correlation for Euclidean distance and average linkage is 0.8977080867389372.
Cophenetic correlation for Euclidean distance and weighted linkage is 0.8861746814895477.
Cophenetic correlation for Chebyshev distance and single linkage is 0.7382354769296767.
Cophenetic correlation for Chebyshev distance and complete linkage is 0.8533474836336782.
Cophenetic correlation for Chebyshev distance and average linkage is 0.8974159511838106.
Cophenetic correlation for Chebyshev distance and weighted linkage is 0.8913624010768603.
Cophenetic correlation for Mahalanobis distance and single linkage is 0.7058064784553605.
Cophenetic correlation for Mahalanobis distance and complete linkage is 0.6663534463875359.
Cophenetic correlation for Mahalanobis distance and average linkage is 0.8326994115042136.
Cophenetic correlation for Mahalanobis distance and weighted linkage is 0.7805990615142518.
Cophenetic correlation for Cityblock distance and single linkage is 0.7252379350252723.
Cophenetic correlation for Cityblock distance and complete linkage is 0.8731477899179829.
Cophenetic correlation for Cityblock distance and average linkage is 0.896329431104133.
Cophenetic correlation for Cityblock distance and weighted linkage is 0.8825520731498188.
```

Figure 23: Cophenetic Correlation

Highest cophenetic correlation is 0.8977080867389372, which is obtained with Euclidean distance and average linkage.

8.2. Dendrograms

We observe that the cophenetic correlation is highest with Euclidean distance and average linkage.

Next, let's examine the dendrograms for the various linkage methods.

Unlike KMeans, where the number of clusters must be specified beforehand, hierarchical clustering (via a dendrogram) doesn't require predefining the number of clusters. The appropriate number of clusters can be visually determined by cutting the dendrogram at a chosen level.

- Cluster Structure: Dendrograms visually depict which points or clusters are merged at different levels, giving insights into the nested structure of the data.
- Choosing the Number of Clusters: By selecting a threshold along the y-axis (height), the dendrogram helps determine an appropriate number of clusters. The number of vertical lines cut by this threshold represents the number of clusters.

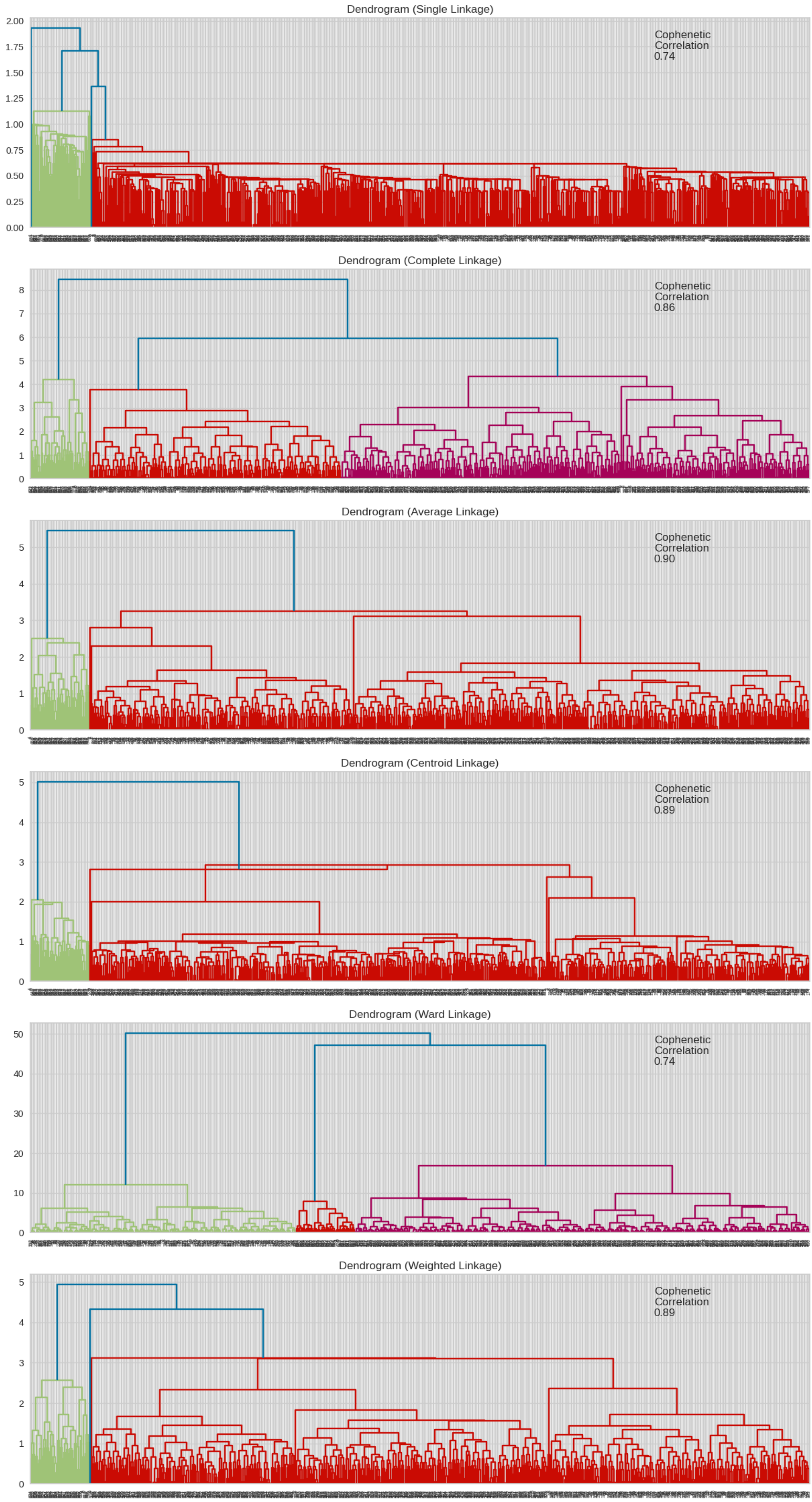

Figure 24: Dendrograms

8.3. Final model

Based on the dendrograms, the Ward linkage shows the best cluster separation, even though it has a lower cophenetic correlation compared to other linkage methods. Opting for 4 clusters appears to be a suitable choice.

8.4. Cluster Profiling

HC_segments	count
0	225
3	194
1	191
2	50

Figure 25: Hierarchical Clustering counts

HC_segments	Avg_Credit_Limit	Total_Credit_Cards	Total_visits_bank	Total_visits_online	Total_calls_made	count_in_each_segment
0	12151.111111	2.422222	0.937778	3.546667	6.857778	225
1	38298.429319	5.670157	2.523560	0.947644	2.099476	191
2	141040.000000	8.740000	0.600000	10.900000	1.080000	50
3	29474.226804	5.365979	4.448454	1.010309	1.891753	194

Figure 26: Hierarchical Cluster profile

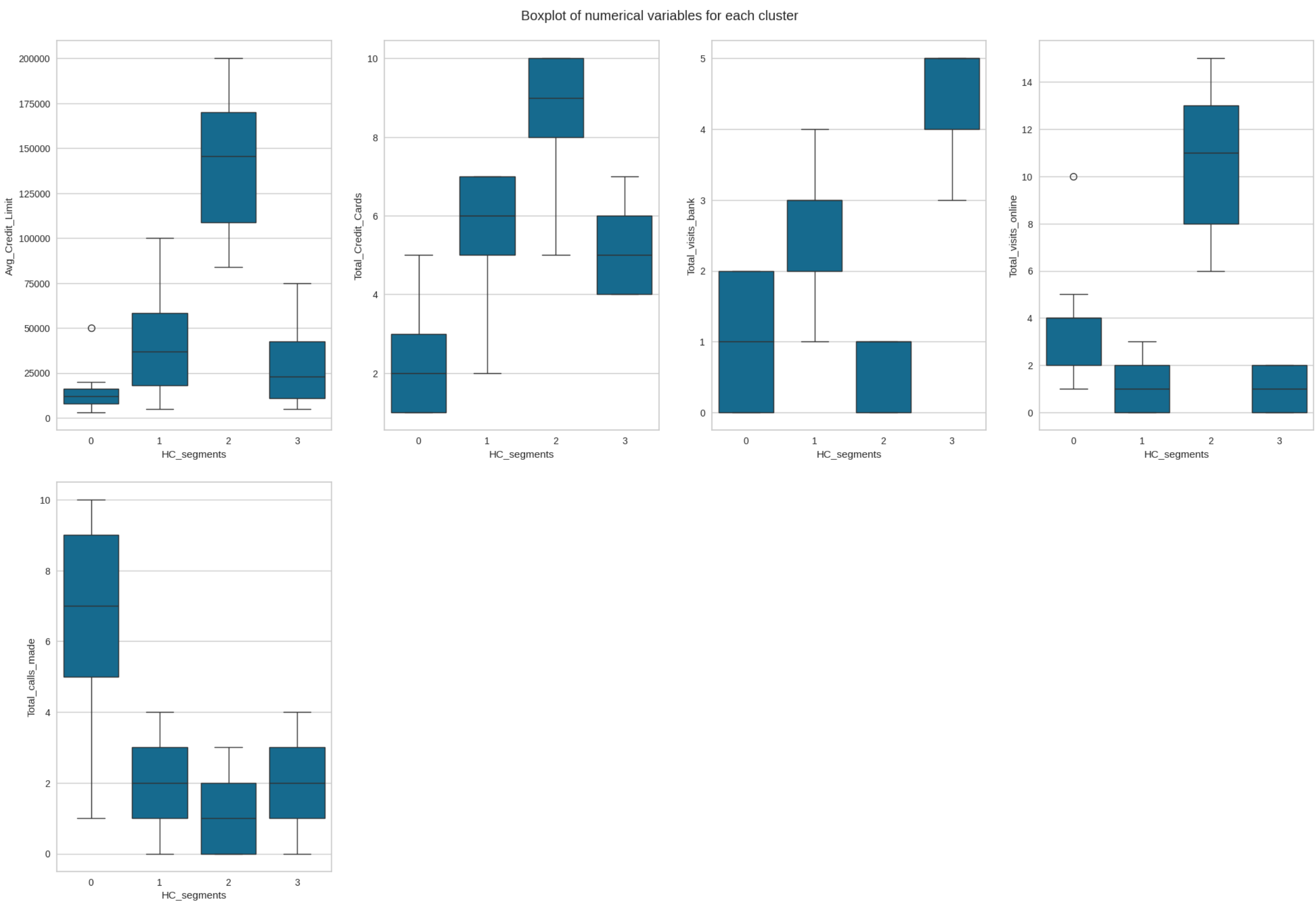

Figure 27: Boxplot for each hierarchical cluster

8.5. Insights

Cluster 0

- Avg_Credit_Limit: Customers in this segment have the lowest average credit limits, typically between 25,000 and 50,000.
- Total_Credit_Cards: This group holds around 4-6 credit cards on average.
- Total_visits_bank: Customers in this cluster make the most visits to the bank, with a median of 4 visits. There's notable variability, indicating diverse banking habits.
- Total_visits_online: They have the fewest online banking interactions, averaging about 2-4 visits.
- Total_calls_made: This group tends to make a higher number of calls, averaging around 8 calls.

General Insight: Cluster 0 represents customers with lower credit limits who rely heavily on phone communication and traditional banking, visiting banks frequently but utilizing online services minimally.

Cluster 1

- Avg_Credit_Limit: Customers in this cluster have similar low credit limits, generally around $25,000.
- Total_Credit_Cards: They typically possess 2-3 credit cards, which is fewer compared to other clusters.
- Total_visits_bank: This group makes the least bank visits, averaging 1 visit.
- Total_visits_online: Online visits are also low, usually around 1-2 visits.
- Total_calls_made: Customers here make the fewest calls, averaging 2-4 calls.

General Insight: Cluster 1 consists of customers with the lowest engagement overall, showcasing minimal credit card usage and low interaction with both banking channels.

Cluster 2

- Avg_Credit_Limit: This cluster includes customers with significantly higher average credit limits, generally ranging from 75,000to150,000.
- Total_Credit_Cards: Customers in this segment possess a higher number of credit cards, typically between 6-8 cards.
- Total_visits_bank: Their bank visits are lower, averaging 1 visit.
- Total_visits_online: They are more active in online banking, averaging about 10-12 visits.
- Total_calls_made: This group makes fewer phone calls, typically around 2-4 calls.

General Insight: Cluster 2 is characterized by affluent customers with higher credit limits and card usage. They prefer online banking over physical visits and phone interactions.

Cluster 3

- Avg_Credit_Limit: Customers in this cluster have moderate credit limits, generally between 50,000to75,000.
- Total_Credit_Cards: They typically have around 4-5 credit cards.
- Total_visits_bank: Their banking behavior is moderate, with about 1-2 visits.
- Total_visits_online: Online visits are also moderate, averaging 2-4 visits.
- Total_calls_made: Phone calls are average, typically around 2-4 calls.

General Insight: Cluster 3 represents a balanced group of customers with moderate credit limits and interaction with both banking and online services.

Summary of Customer Clusters:

- Cluster 0: Customers with low credit limits and high engagement with bank visits and calls.
- Cluster 1: Customers showing minimal banking activity and engagement.
- Cluster 2: Affluent customers favoring online banking with high credit limits and multiple cards.
- Cluster 3: Customers with moderate profiles, showcasing balanced engagement across banking channels.

9. K-means vs Hierarchical Clustering

K_means_segments	Avg_Credit_Limit	Total_Credit_Cards	Total_visits_bank	Total_visits_online	Total_calls_made	count_in_each_segment
0	33782.383420	5.515544	3.489637	0.981865	2.000000	386
1	12174.107143	2.410714	0.933036	3.553571	6.870536	224
2	141040.000000	8.740000	0.600000	10.900000	1.080000	50

Figure 28: K-means Cluster Profile

HC_segments	Avg_Credit_Limit	Total_Credit_Cards	Total_visits_bank	Total_visits_online	Total_calls_made	count_in_each_segment
0	12151.111111	2.422222	0.937778	3.546667	6.857778	225
1	38298.429319	5.670157	2.523560	0.947644	2.099476	191
2	141040.000000	8.740000	0.600000	10.900000	1.080000	50
3	29474.226804	5.365979	4.448454	1.010309	1.891753	194

Figure 29: Hierarchical Cluster profile

9.1. Insights

K-means Clustering:

- Cluster 0: 384 customers
- Cluster 1: 224 customers
- Cluster 2: 50 customers

Hierarchical Clustering:

- Cluster 0: 225 customers
- Cluster 1: 191 customers
- Cluster 2: 50 customers
- Cluster 3: 193 customers

Cluster Size:

- K-Means produces three clusters, while Hierarchical produces four, indicating more granularity in the latter.
- K-Means shows a large concentration in Cluster 0 (386 customers), whereas Hierarchical has its largest cluster (Cluster 0) at 225 customers, indicating a more balanced distribution across clusters.

Customer Characteristics:

- Cluster 0 (K-Means) vs. Cluster 1 (Hierarchical):

  K-Means Cluster 0 has a significantly lower average credit limit (33,782.38) compared to Hierarchical Cluster 1 (38,298.43). However, both have similar total credit card counts and customer behavior patterns.

- Cluster 1 (K-Means) vs. Cluster 0 (Hierarchical):

  K-Means Cluster 1 exhibits low credit limits (12,174.11) and high call volume (6.87), whereas Hierarchical Cluster 0 has a similar credit limit (12,151.11) and call volume (6.86). This indicates consistency in customer behavior across both models.

- Cluster 2 (K-Means) vs. Cluster 2 (Hierarchical):

  Both clusters exhibit similar characteristics, indicating a strong group of high-spending customers with significant engagement online.

- Hierarchical Cluster 3:

  Introduced by Hierarchical clustering, this cluster (with 194 customers) has an average credit limit of 29,474.23 and a higher number of bank visits (4.45). This segment can be further explored, as it represents customers who may have different needs and behaviors than those identified by K-Means.

12. Actionable Insights and Recommendations

Cluster Characteristics:

K-means Clustering:

- Cluster 0: Contains the majority of customers (386) with an average credit limit of approximately 33,782 and a significant number of credit cards (5.5). This group makes frequent visits to the bank (3.49) but fewer online visits and calls.
- Cluster 1: Contains 224 customers, with a lower average credit limit (12,174) and fewer credit cards (2.4). This group shows a higher tendency for online interactions (3.55 calls made).
- Cluster 2: A small group (50 customers) with the highest average credit limit (141,040) and the most credit cards (8.74). They are very active online (10.9 visits).

Hierarchical Clustering:

- Cluster 0: Similar to K-means Cluster 1, it has 225 customers with a low average credit limit (12,151) and more online interactions.
- Cluster 1: Similar to K-means Cluster 0 but contains 191 customers with a higher average credit limit (38,298).
- Cluster 2: Matches K-means Cluster 2 (50 customers, 141,040).
- Cluster 3: Contains 194 customers with an average credit limit of 29,474, moderate banking activity, and slightly fewer online interactions compared to Cluster 1.

Recommendations:

Targeted Marketing Strategies:

- Cluster 0 (High Credit, High Bank Visits): Implement a loyalty program targeting this group, encouraging them to engage with both online and offline services. Offer incentives for online transactions to increase engagement.
- Cluster 1 (Low Credit, High Online Engagement): Focus on educational content and promotions to increase their credit limits and encourage more offline engagement. Consider offering financial literacy workshops that explain the benefits of having multiple credit cards.
- Cluster 2 (Ultra High Credit): Develop exclusive products tailored for high-net-worth individuals, such as premium credit cards with enhanced benefits. Consider personalized financial advisory services.
- Cluster 3 (Moderate Credit with Mixed Behavior): Encourage this group to increase their credit card usage and online banking through targeted campaigns that highlight convenience and potential rewards. Customer Engagement:

Use insights from both clusters to create personalized communication strategies. For example, tailor emails and promotions based on the preferred interaction channel of each cluster (online vs. in-person). Increase engagement through feedback surveys to understand their needs better and adapt products or services accordingly.

Cross-Selling Opportunities:

Identify customers who have low credit limits but high engagement (Cluster 1 and Cluster 0 in K-means). Offer tailored products that can help them increase their limits and credit score, such as secured credit cards or credit-building loans.

Data-Driven Decision Making:

Continuously monitor customer behavior and update clustering models regularly to identify emerging trends. This will ensure that strategies remain relevant and effective. Utilize additional variables in future clustering to enhance customer segmentation, such as customer demographics or transaction behaviors.

Risk Management:

For high-credit customers in Cluster 2, ensure risk assessment measures are in place, considering their high transaction volumes and engagement levels.