Problem Statement - UL Project – Coded

All Life Bank

Context

All Life 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

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.

Data Description

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).

Data Dictionary

- Sl_No: Primary key of the records
- Customer Key: Customer identification number
- Average Credit Limit: Average credit limit of each customer for all credit cards
- Total credit cards: Total number of credit cards possessed by the customer
- Total visits bank: Total number of visits that the customer made (yearly) personally to the bank
- Total visits online: Total number of visits or online logins made by the customer (yearly)
- Total calls made: Total number of calls made by the customer to the bank or its customer service department (yearly)

We will begin with Univariate analysis by examining the distribution of each individual variable.

We can check for central tendencies (mean, median) and dispersions (variance, standard deviation) to gain insights into each feature.

We'll start by calculating the summary statistics and then proceed to visualizations for each variable.

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

;

	Sl_No	Customer Key	Avg_Credit_Limit	Total_Credit_Cards	Total_visits_bank	Total_visits_online	Total_calls_made
count	660.000000	660.000000	660.000000	660.000000	660.000000	660.000000	660.000000
mean	330.500000	55141.443939	34574.242424	4.706061	2.403030	2.606061	3.583333
std	190.669872	25627.772200	37625.487804	2.167835	1.631813	2.935724	2.865317
min	1.000000	11265.000000	3000.000000	1.000000	0.000000	0.000000	0.000000
25%	165.750000	33825.250000	10000.000000	3.000000	1.000000	1.000000	1.000000
50%	330.500000	53874.500000	18000.000000	5.000000	2.000000	2.000000	3.000000
75%	495.250000	77202.500000	48000.000000	6.000000	4.000000	4.000000	5.000000
max	660.000000	99843.000000	200000.000000	10.000000	5.000000	15.000000	10.000000

Fig: Summary statistics of the dataset

Here are the key insights from the summary statistics of the dataset:

1. **Average Credit Limit**:
   - Mean: $34,574
   - Range: $3,000 to $200,000
   - There is a large spread in credit limits, with the top 25% of customers having credit limits of $48,000 or more.
2. **Total Credit Cards**:
   - Mean: 4.7 cards per customer
   - Range: 1 to 10 cards
   - The median number of credit cards is 5, indicating many customers own multiple cards.
3. **Total Visits to the Bank (Yearly)**:
   - Mean: 2.4 visits
   - Range: 0 to 5 visits
   - Many customers make fewer visits, with 50% of them visiting the bank only twice or less per year.
4. **Total Online Visits (Yearly)**:
   - Mean: 2.6 online logins
   - Range: 0 to 15 logins
   - The distribution of online logins is more spread out, with some customers being more frequent users.
5. **Total Calls Made (Yearly)**:
   - Mean: 3.6 calls per year
   - Range: 0 to 10 calls
   - The middle 50% of customers make 1 to 5 calls per year.

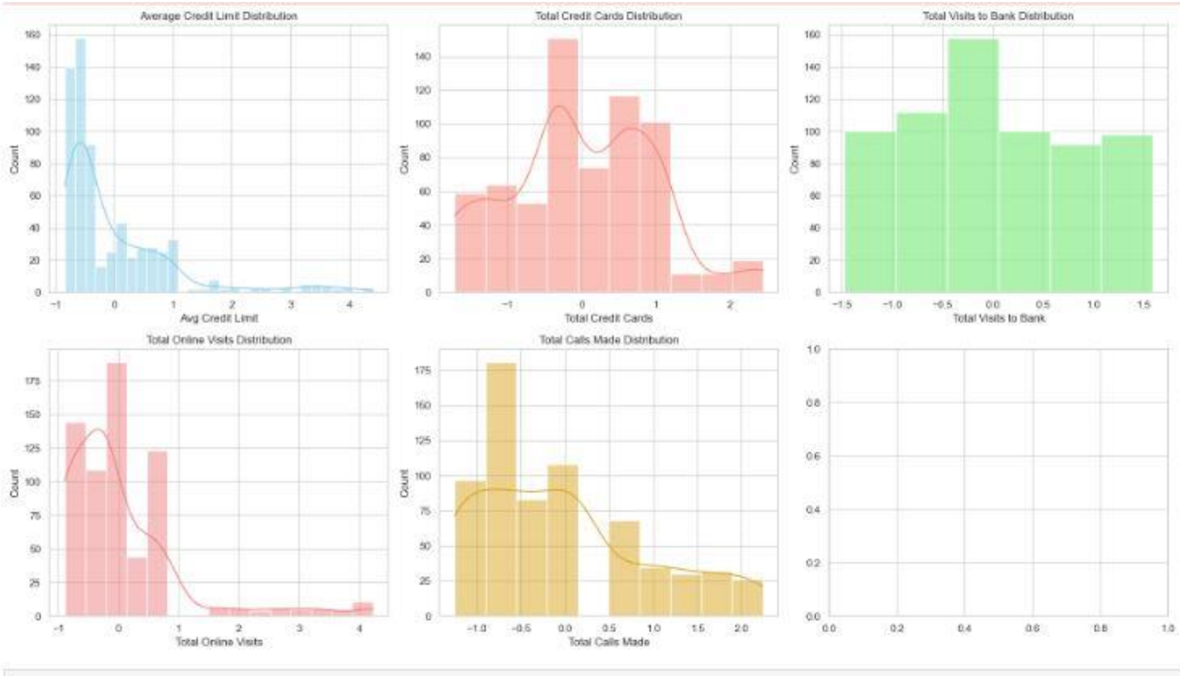


Fig: Univariate visualizations

The Univariate visualizations reveal the following patterns:

1. **Average Credit Limit**:
   - The distribution is skewed to the right, meaning most customers have lower credit limits, with a few having very high limits.
2. **Total Credit Cards**:
   - The number of credit cards is spread somewhat evenly, but most customers hold around 3 to 6 cards.
3. **Total Visits to the Bank**:
   - Many customers visit the bank 2 to 3 times yearly, and very few visit the maximum number of times (5).
4. **Total Online Visits**:
   - There is a right-skewed distribution with many customers having fewer than 5 online logins, although some customers are frequent online users (up to 15 logins).
5. **Total Calls Made**:
   - The distribution is close to normal, with most customers making 1 to 5 calls per year.

For the **Exploratory Data Analysis (EDA)**, we will dive deeper into visualizing and analyse each key variable and relationship to identify actionable insights for the Marketing and Operations teams at AllLife Bank. This will allow us to understand customer behaviour's, spending habits, and interaction patterns with the bank, which is crucial for segmentation.

Here's a structured EDA plan to extract actionable insights:

---

1. Univariate Analysis

This step provides insights into individual variables, helping us understand customer spending and engagement patterns.

- **Average Credit Limit**: Analyse the distribution of credit limits to understand spending power across customers.

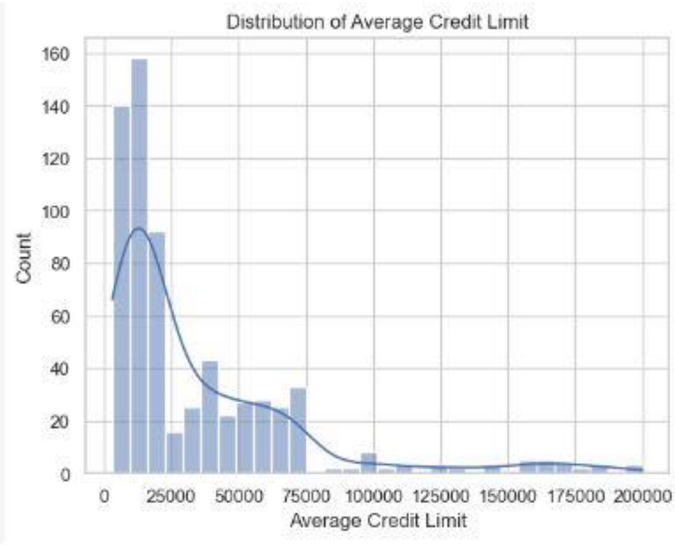


Fig : The graph shows Interaction Metrics between Average Credit Limit

The distribution is skewed to the right, meaning most customers have lower credit limits, with a few having very high limits

- **Bank Visits Distribution (1st Image)**:
  - The number of bank visits seems uniformly distributed between 1 to 5, indicating that customers make varied use of in-person visits.
- **Online Visits Distribution**:
  - A large majority of customers have minimal or no online visits (with a sharp drop after 3 visits), suggesting low engagement through online channels.
- **Calls Made Distribution**:
  - The histogram shows many customers making around 4-6 calls, indicating moderate reliance on customer service calls, but it tapers off sharply for higher call counts.

Insight: There is higher preference for traditional channels like phone calls and in-person visits, with limited engagement online.

- **Insight**: High credit limits may indicate high-value customers. Marketing can target the upper range of this distribution for premium offerings.

- **Total Credit Cards**: Analyse the total credit cards held by customers.

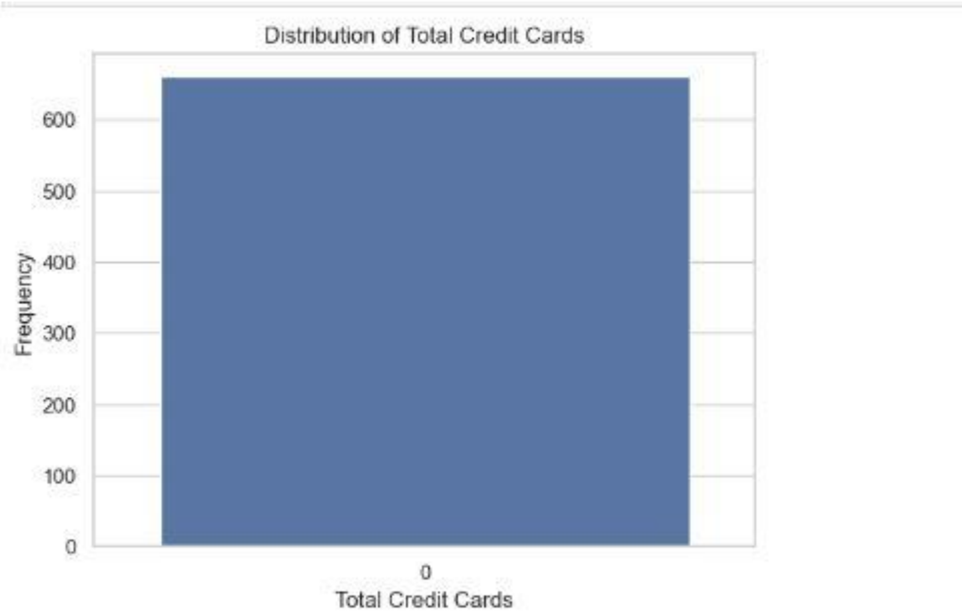


Fig : The graph shows Interaction Metrics Analysis of the total credit cards held by customers

- **Insight**: Customers with multiple credit cards may represent loyal or high-value customers. Marketing can upsell rewards cards or premium cards to this group.

- **Interaction Metrics (Bank Visits, Online Visits, Calls)**:

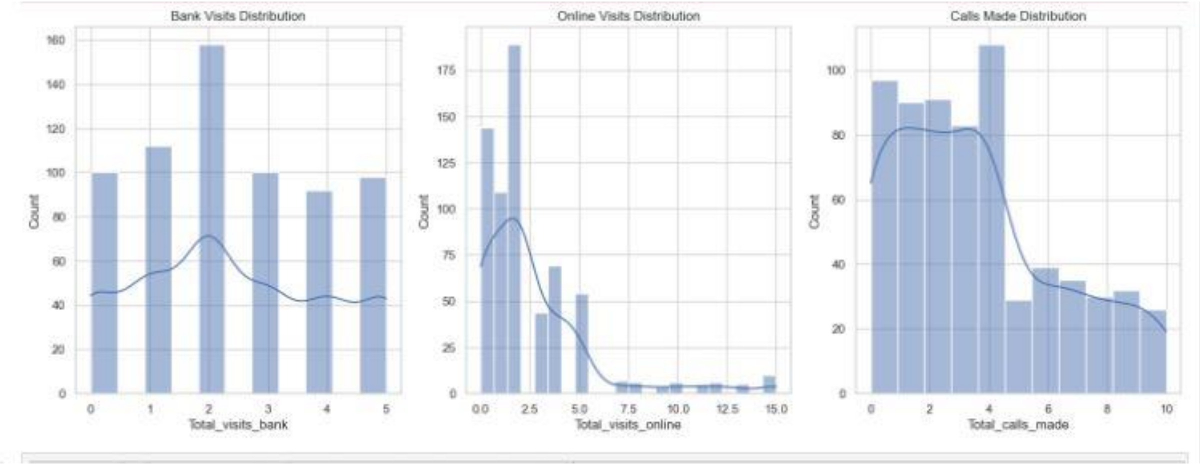


Fig : The graph shows Interaction Metrics between Bank Visits, Online Visits, Calls

Insight: Customers with high call volumes may indicate dissatisfaction or support needs. The Operations team should investigate to streamline customer service.

2. Bivariate Analysis

In this step, we explore relationships between pairs of variables, which can reveal patterns between customer behavior and interaction types.

- **Credit Limit vs. Total Credit Cards**:

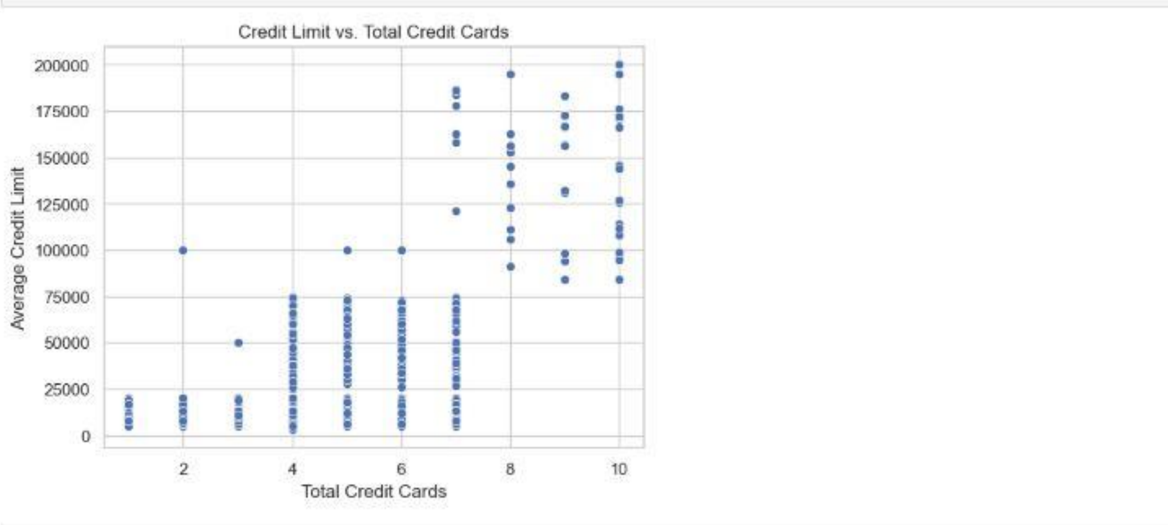


Fig : The graph shows Interaction Metrics between Credit Limit vs. Total Credit Cards

- **Insight**: If higher credit limits correlate with more credit cards, these customers are ideal candidates for premium product upselling.

- **Credit Limit vs. Interaction Metrics**:

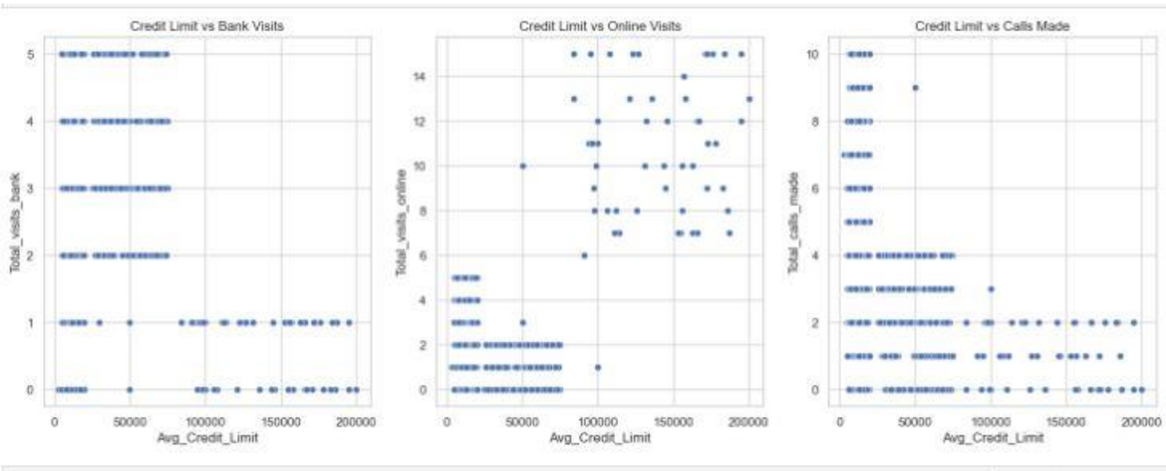


Fig : The graph shows Interaction Metrics between Credit Limit vs. Interaction Metrics:

- **Avg Credit Limit vs Bank Visits**:
  - There is no clear relationship between higher credit limits and the number of bank visits.
  - Customers with both low and high credit limits appear to have a similar distribution of visits.
- **Avg Credit Limit vs Online Visits**:
  - Customers with higher credit limits tend to make more online visits. This suggests that customers with more financial flexibility may prefer online channels.
- **Avg Credit Limit vs Calls Made**:
  - The relationship between credit limits and calls made seems weak. Customers across all credit levels make calls in a similar pattern.

Insight: Online visits may correlate with higher financial activity, while traditional modes (calls and bank visits) are used by customers irrespective of their financial status.

- **Insight**: Customers with high credit limits and high call volumes could indicate high-value customers needing enhanced support services.

For the **bivariate analysis**, we'll explore relationships between pairs of variables, such as how credit limits correlate with the number of credit cards, or how visits and calls are related. I'll start by calculating the correlation matrix and then visualizing key relationships between variables.

Let's begin with the correlation matrix.

	Sl_No	Customer Key	Avg_Credit_Limit	Total_Credit_Cards	Total_visits_bank	Total_visits_online	Total_calls_made	Engagement_Score
Sl_No	1.000000	0.052886	0.677962	0.739329	0.406438	0.033916	-0.684125	-0.351373
Customer Key	0.052886	1.000000	0.068604	-0.010281	-0.000560	0.022506	0.005968	0.024138
Avg_Credit_Limit	0.677962	0.068604	1.000000	0.608860	-0.100312	0.551385	-0.414352	0.078575
Total_Credit_Cards	0.739329	-0.010281	0.608860	1.000000	0.315796	0.167758	-0.651251	-0.251837
Total_visits_bank	0.406438	-0.000560	-0.100312	0.315796	1.000000	-0.551861	-0.506016	-0.422021
Total_visits_online	0.033916	0.022506	0.551385	0.167758	-0.551861	1.000000	0.127299	0.704235
Total_calls_made	-0.684125	0.005968	-0.414352	-0.651251	-0.506016	0.127299	1.000000	0.708157
Engagement_Score	-0.351373	0.024138	0.078575	-0.251837	-0.422021	0.704235	0.708157	1.000000

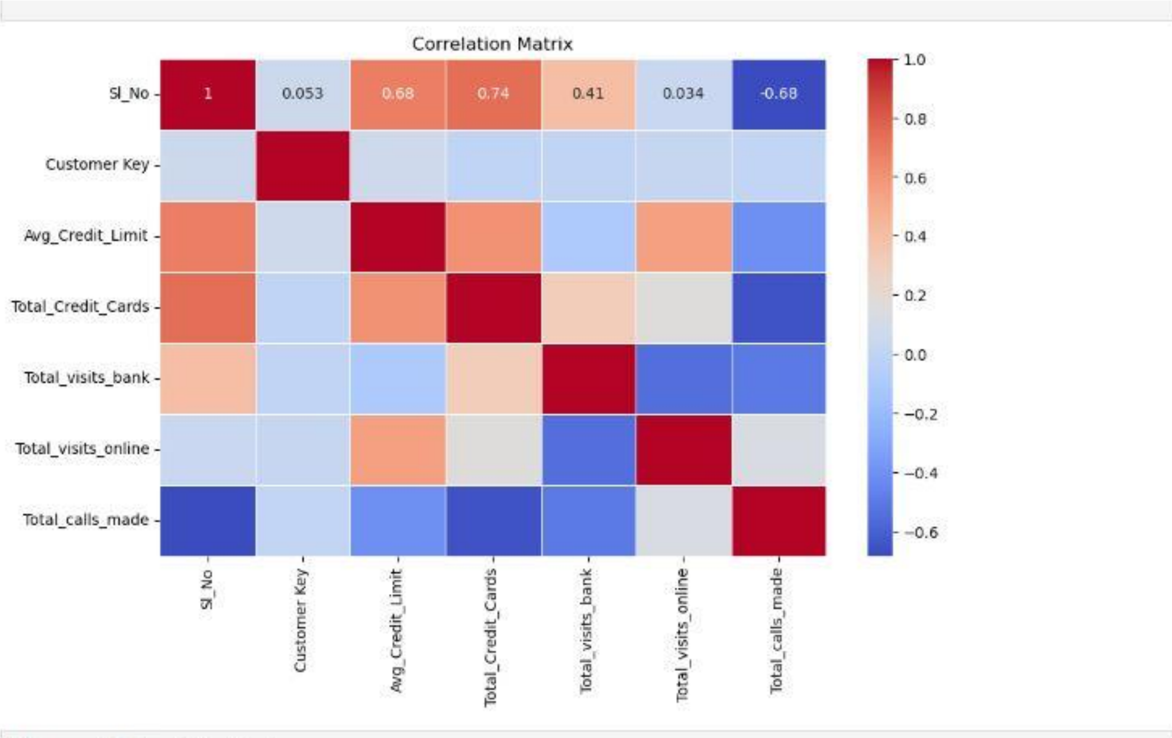
Fig : Calculate the correlation matrix

Performing bivariate analysis

To perform bivariate analysis, here are the steps you can follow:

Step 1: Correlation Matrix

A correlation matrix helps to identify relationships between numerical variables. Here's how to calculate and visualize it:



This will show you the strength and direction of relationships between the variables.

- **Credit Cards & Avg Credit Limit**:
  - A strong positive correlation (~0.74) between the **Total Credit Cards** and **Avg Credit Limit** implies that customers with more cards tend to have higher credit limits.
- **Avg Credit Limit & Bank Visits**:
  - Weak positive correlation (~0.41) suggests a slight tendency for those with higher credit limits to visit banks more often, though not strongly significant.
- **Avg Credit Limit & Calls Made**:
  - There is a negative correlation (~-0.68), indicating that customers with higher credit limits may rely less on calls, potentially because of better service accessibility or online engagement.

Insight: Credit limit tends to correlate more with the number of credit cards but shows mixed or weaker relationships with engagement through visits or calls.

Step 2: Scatter Plots

Scatter plots are useful for visualizing relationships between two continuous variables. For example:

- **Credit Limit vs. Total Credit Cards**:

	Sl_No	Customer Key	Avg_Credit_Limit	Total_Credit_Cards	Total_visits_bank	Total_visits_online	Total_calls_made
count	660.000000	660.000000	660.000000	660.000000	660.000000	660.000000	660.000000
mean	330.500000	55141.443939	34574.242424	4.706061	2.403030	2.606061	3.583333
std	190.669872	25627.772200	37625.487804	2.167835	1.631813	2.935724	2.865317
min	1.000000	11265.000000	3000.000000	1.000000	0.000000	0.000000	0.000000
25%	165.750000	33825.250000	10000.000000	3.000000	1.000000	1.000000	1.000000
50%	330.500000	53874.500000	18000.000000	5.000000	2.000000	2.000000	3.000000
75%	495.250000	77202.500000	48000.000000	6.000000	4.000000	4.000000	5.000000
max	660.000000	99843.000000	200000.000000	10.000000	5.000000	15.000000	10.000000

Fig: Visual relationships between two continuous variables

Total Visits to Bank vs. Total Calls:

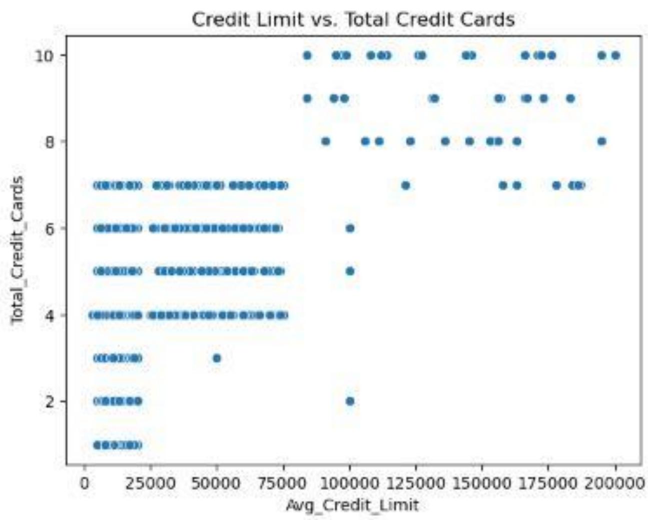


Fig : The graph shows Interaction Metrics between **Total Visits to Bank vs. Total Calls**

Step 3: Box Plots

If you're comparing a categorical variable (e.g., low vs. high spenders), you can use box plots to observe the distribution of one numerical variable across different categories. For example:

- **Distribution of Credit Limit by Total Calls Made**:

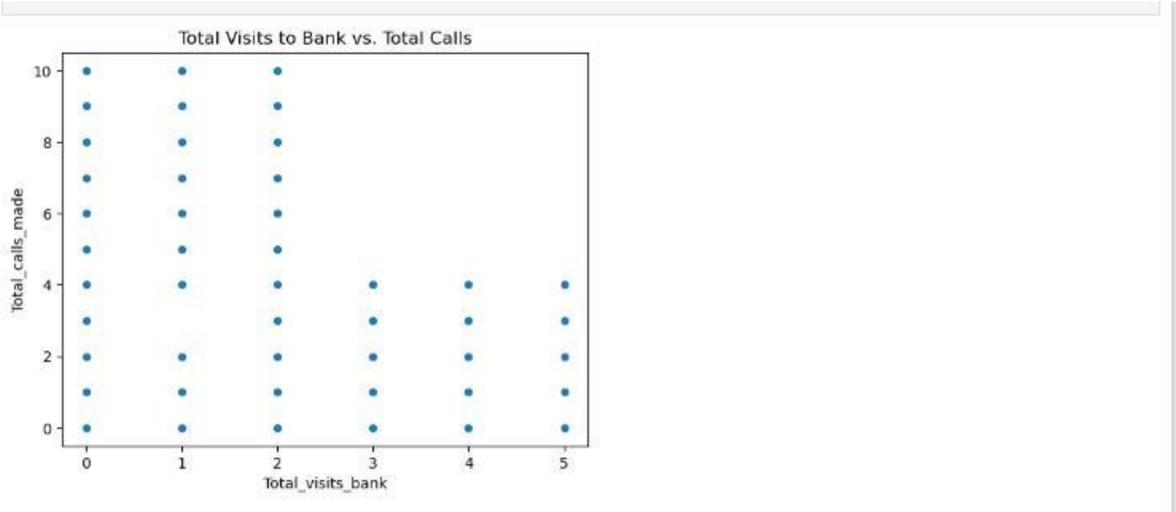


Fig : The graph shows Distribution of Credit Limit by Total Calls Made

Another aspect of the analysis

- **Data Preprocessing**: Handling missing values, outliers, or feature engineering.
- **Clustering (K-means or Hierarchical)**: Segmenting customers based on spending and interactions.
- **Further Visualization**: Any specific pair of variables you'd like to analyse or visualize more.
- **Actionable Insights**: How to derive insights from patterns and relationships observed.

In **Data Preprocessing**, we aim to clean and prepare the data for analysis. Here are the key steps:

1. Missing Value Treatment

First, we need to check for any missing values. Depending on the nature of the missing data, we can either remove records or fill them with appropriate values (mean, median, or a specific value).

Code to check for missing values:

```
Sl_No                  0
Customer Key           0
Avg_Credit_Limit       0
Total_Credit_Cards     0
Total_visits_bank      0
Total_visits_online    0
Total_calls_made       0
dtype: int64
```

Fig. : check for missing values

2. Outlier Detection and Treatment

Outliers can skew your analysis. We use box plots or statistical methods to detect outliers.

Code to detect outliers with box plots:

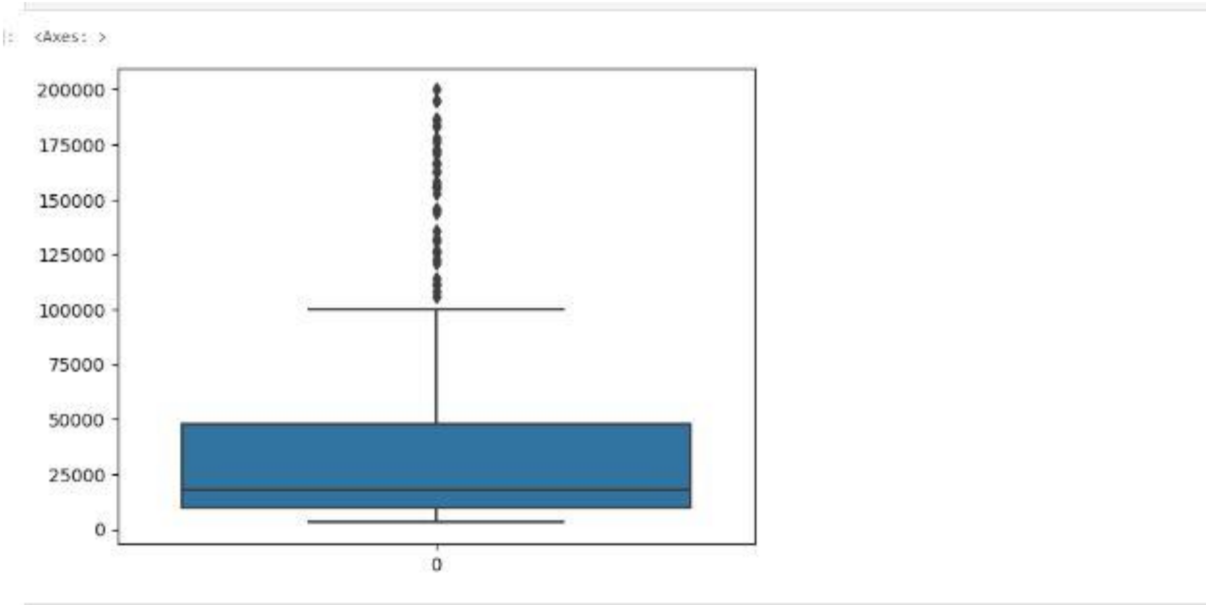


Fig: Box plot to detect outliers with box plots

Treatment options:

- We can remove or cap outliers

	Sl_No	Customer Key	Avg_Credit_Limit	Total_Credit_Cards	Total_visits_bank	Total_visits_online	Total_calls_made	Engagement_Score	Credit_Category
0	1	87073	1.740187	-1.249225	-0.860451	-0.547490	-1.251537	2	Low
1	2	38414	0.410293	-0.787585	-1.473731	2.520519	1.891859	19	Low
2	3	17341	0.410293	1.058973	-0.860451	0.134290	0.145528	8	Low
3	4	40496	-0.121665	0.135694	-0.860451	-0.547490	0.145528	6	NaN
4	5	47437	1.740187	0.597334	-1.473731	3.202298	-0.203739	15	Low

3. Feature Engineering

This step involves creating new variables or transforming existing ones to improve model performance. Examples include:

- **Creating customer engagement score**: Combining visits, calls, and online logins into one metric.

	Avg_Credit_Limit	Credit_Category
0	1.740187	Low
1	0.410293	Low
2	0.410293	Low
3	-0.121665	NaN
4	1.740187	Low

4. Data Scaling

Some machine learning models (like K-means) are sensitive to the scale of the data. Scaling ensures that all features contribute equally to the model.

Code to scale the data:

```
   Sl_No  Customer Key  Avg_Credit_Limit  Total_Credit_Cards  \
0      1         87073          1.740187           -1.249225
1      2         38414          0.410293           -0.787585
2      3         17341          0.410293            1.058973
3      4         40496         -0.121665            0.135694
4      5         47437          1.740187            0.597334

   Total_visits_bank  Total_visits_online  Total_calls_made  Engagement_Score  \
0          -0.860451            -0.547490         -1.251537                 2
1          -1.473731             2.520519          1.891859                19
2          -0.860451             0.134290          0.145528                 8
3          -0.860451            -0.547490          0.145528                 6
4          -1.473731             3.202298         -0.203739                15

  Credit_Category
0             Low
1             Low
2             Low
3             NaN
4             Low
```

Fig: Code to scale the data

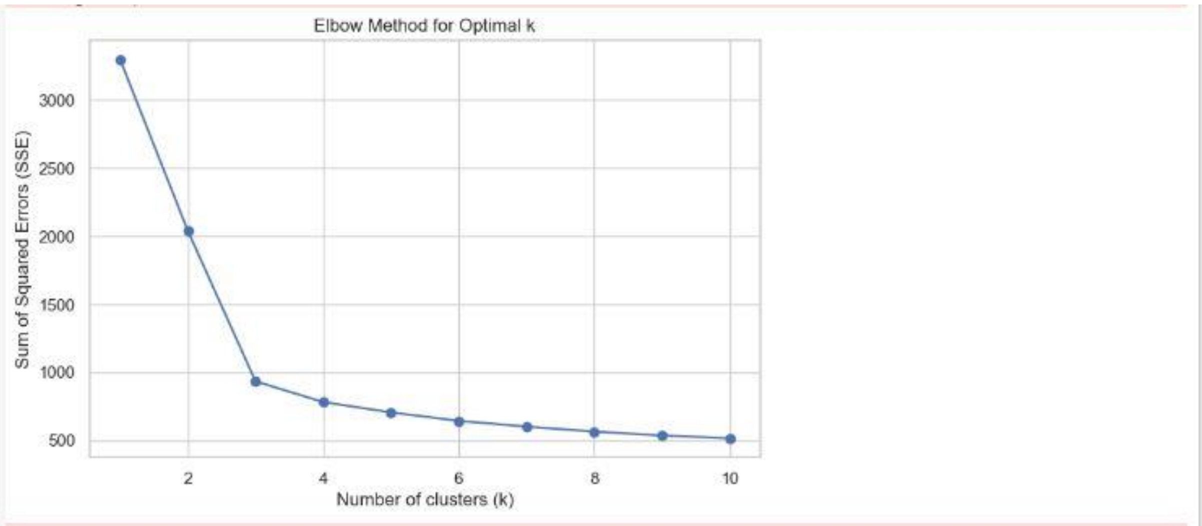
Clustering

For clustering, we can try both **K-means** and **Hierarchical Clustering** to segment customers based on their financial behaviors and interactions with the bank. Here's how we can proceed with each technique.

1. K-means Clustering

Step 1: Finding the Optimal Number of Clusters

We use the **Elbow Method** and **Silhouette Scores** to find the optimal number of clusters for K-means.




```
[[-0.02106178  0.37368962  0.6663945  -0.55367163 -0.55300488]
 [-0.59579625 -1.05962278 -0.9015185   0.32299678  1.14810882]
 [ 2.83176409  1.86222621 -1.10576269  2.82731942 -0.87432983]]
```

Fig : Elbow Method for optimal number of clusters for K-means.

Step 2: Apply K-means Clustering

Once we determine the optimal number of clusters (from the Elbow curve), apply K-means:

```
[[-0.02106178  0.37368962  0.6663945  -0.55367163 -0.55300488]
 [-0.59579625 -1.05962278 -0.9015185   0.32299678  1.14810882]
 [ 2.83176409  1.86222621 -1.10576269  2.82731942 -0.87432983]]
```

Step 3: Cluster Profiling

Profile each cluster to understand the behavior of customers in each group.

```
         Avg_Credit_Limit  Total_Credit_Cards  Total_visits_bank  \
Cluster
0               -0.021062            0.373690           0.666395
1               -0.595796           -1.059623          -0.901518
2                2.831764            1.862226          -1.105763

         Total_visits_online  Total_calls_made
Cluster
0                  -0.553672         -0.553005
1                   0.322997          1.148109
2                   2.827319         -0.874330
```

2. Hierarchical Clustering

Step 1: Applying Hierarchical Clustering with Dendrograms

We use dendrograms to visualize the merging of clusters and help decide the number of clusters.

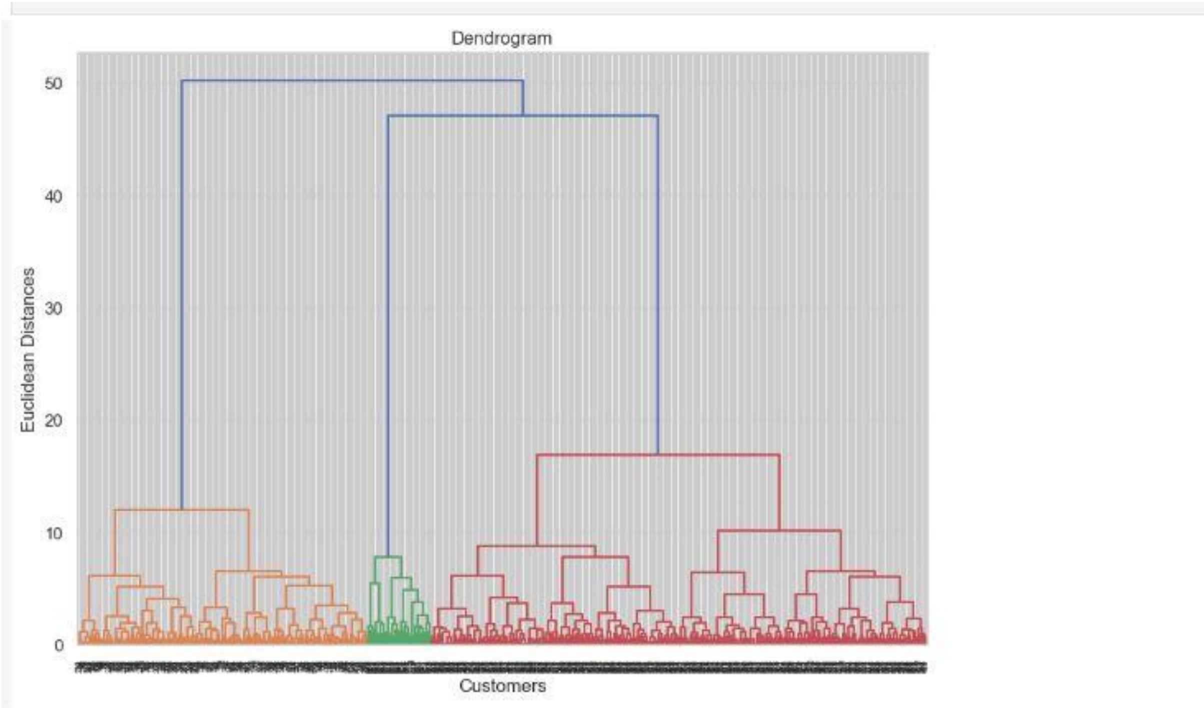


Fig : Dendrogram to visualize the merging

- The **dendrogram** suggests that there are clear sub-clusters of customer behaviors.
- The largest separation occurs at the top (blue and green clusters), indicating distinct customer profiles, potentially driven by credit activity or engagement channel preferences.

Insight: This clustering may point to 3-4 key customer segments. For example:

- **High Credit Activity, Online Engagement** (one cluster)
- **Low Credit Activity, Traditional Engagement** (another cluster)
- **Moderate Credit Activity, Balanced Usage of Channels**.

Step 2: Choosing the Number of Clusters

Once you inspect the dendrogram, you can choose a threshold to cut the dendrogram and determine the number of clusters:

```
HCluster
3    385
1    225
2     50
Name: count, dtype: int64
```

Step 3: Cluster Profiling

Just like in K-means, you can profile clusters created through hierarchical clustering:

```
          Avg_Credit_Limit  Total_Credit_Cards  Total_visits_bank  \
HCluster
1                -0.596408           -1.054310          -0.898610
2                 2.831764            1.862226          -1.105763
3                -0.019212            0.374308           0.668767

          Total_visits_online  Total_calls_made
HCluster
1                    0.320643          1.143653
2                    2.827319         -0.874330
3                   -0.554573         -0.554819
```

Fig : Profile clusters created through hierarchical clustering

3. Comparing K-means and Hierarchical Clustering

You can compare the clusters from both methods to check if the groupings are consistent or offer different insights. For example:

```
Cluster  HCluster
0        1             1
         3           385
1        1           224
2        2            50
dtype: int64
```

Fig: comparing the clusters from both methods

Conclusion

Once you have the clusters, analyse the customer segments (spending patterns, interaction with the bank, etc.) and derive actionable insights for the bank's marketing and operations teams.

Analyzing the customer segments

To analyse the customer segments, we need to focus on profiling each cluster (group) identified through either **K-means** or **Hierarchical Clustering**. This involves understanding the characteristics of each segment based on their spending behavior (e.g., credit limit) and their interaction with the bank (e.g., visits, calls).

Here are the key steps for customer segmentation analysis:

Step 1: Cluster Profiling

After running the clustering algorithm, you'll have a `Cluster` or `HCluster` label for each customer. The next step is to examine the **mean** and **distribution** of key variables for each cluster to derive insights.

Example for **K-means** profiling:

This gives you the average values for each key variable in each cluster. You can interpret them as follows:

1. **Avg_Credit_Limit**: Does the cluster have high or low spending power?
2. **Total_Credit_Cards**: Do customers in this segment own multiple credit cards, indicating higher engagement or spending potential?
3. **Total_visits_bank, Total_visits_online, Total_calls_made**: What are their interaction patterns with the bank? For example, if a cluster has a high number of calls, it might indicate dissatisfaction or need for customer service.

```
              Sl_No  Customer Key  Avg_Credit_Limit  Total_Credit_Cards  \
Cluster
0        417.528497  54881.329016         -0.021062            0.373690
1        117.857143  55239.830357         -0.595796           -1.059623
2        611.280000  56708.760000          2.831764            1.862226

         Total_visits_bank  Total_visits_online  Total_calls_made  \
Cluster
0                 0.666395            -0.553672         -0.553005
1                -0.901518             0.322997          1.148109
2                -1.105763             2.827319         -0.874330

         Engagement_Score  Cluster
Cluster
0                6.471503      0.0
1               11.357143      1.0
2               12.580000      2.0
```


```
              Sl_No  Customer Key  Avg_Credit_Limit  Total_Credit_Cards  \
Cluster
0        417.528497  54881.329016         -0.021062            0.373690
1        117.857143  55239.830357         -0.595796           -1.059623
2        611.280000  56708.760000          2.831764            1.862226

         Total_visits_bank  Total_visits_online  Total_calls_made  \
Cluster
0                 0.666395            -0.553672         -0.553005
1                -0.901518             0.322997          1.148109
2                -1.105763             2.827319         -0.874330

         Engagement_Score  Cluster
Cluster
0                6.471503      0.0
1               11.357143      1.0
2               12.580000      2.0
```


Fig : **K-means** profiling

Step 2: Visualizing Cluster Characteristics

You can visualize these findings to get a better understanding of the differences between clusters. For example:

- **Average Credit Limit per Cluster**:

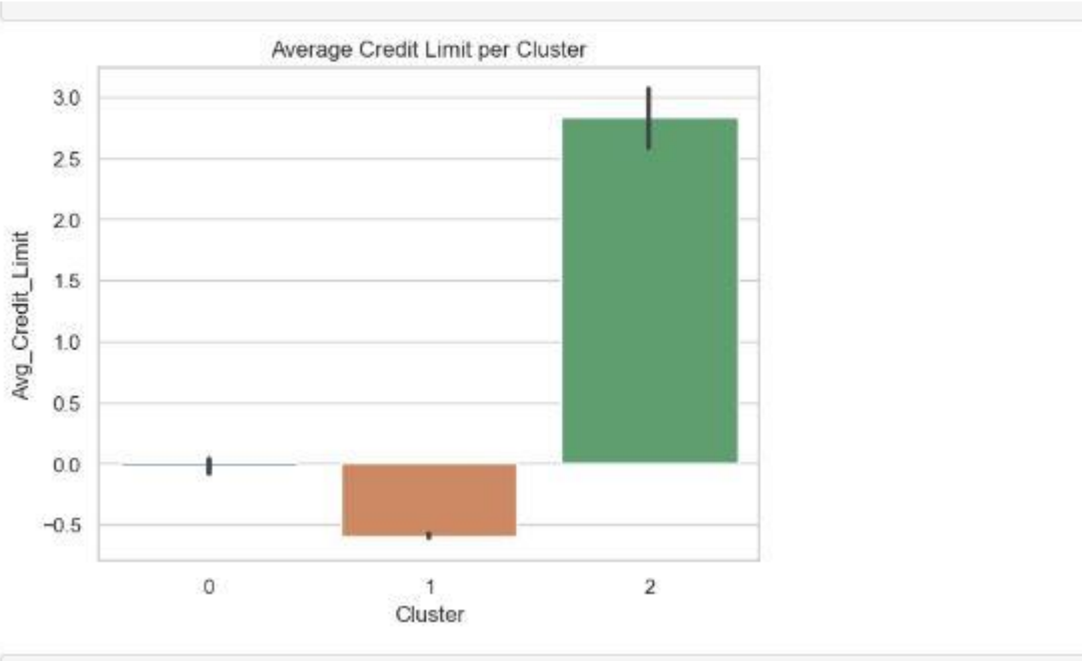


Fig : The graph shows Distribution of Average Credit Limit per Cluster

- The boxplot shows the **average credit limit across clusters**.
  - **Cluster 0 and 1** have similar, lower average credit limits (possibly basic or low-value customers).
  - **Cluster 2** has significantly higher average credit limits, suggesting a premium segment.

Insight: Cluster 2 could represent high-value customers who might benefit from personalized services and focused online engagement.

Customer Interactions (Visits, Calls) per Cluster:

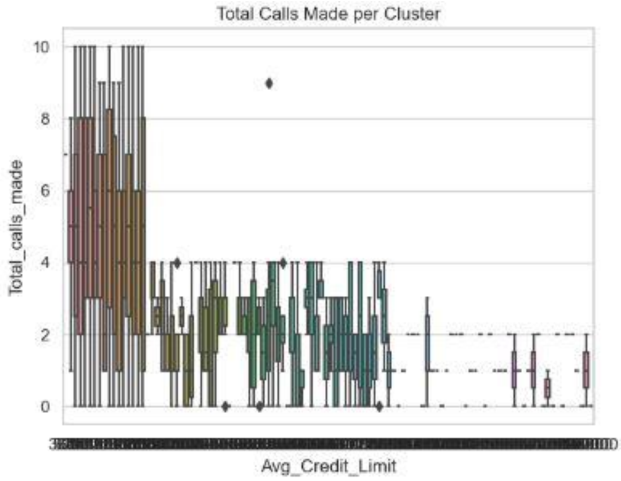


Fig : Customer Interactions (Visits, Calls) per Cluster:

```
Index(['Sl_No', 'Customer Key', 'Avg_Credit_Limit', 'Total_Credit_Cards',
       'Total_visits_bank', 'Total_visits_online', 'Total_calls_made'],
      dtype='object')
[100000  50000  30000  20000  15000   5000   3000  10000  13000  11000
   9000   6000   8000  19000  16000  18000  17000  14000  12000   7000
  73000  49000  67000  61000  75000  48000  56000  72000  70000  51000
  69000  40000  44000  31000  37000  65000  46000  74000  58000  39000
  52000  33000  47000  71000  41000  59000  64000  45000  54000  66000
  27000  43000  36000  25000  57000  26000  38000  35000  34000  28000
  63000  29000  68000  42000  62000  32000  60000  55000 157000  94000
 163000 131000  96000 136000 121000 158000 108000 166000 176000 178000
  91000 156000 146000  84000 155000 200000 195000 187000 106000 114000
 126000 173000 153000 184000 123000 144000  97000  98000 127000 171000
 186000 183000 111000 112000 132000  95000 172000  99000 145000 167000]
```

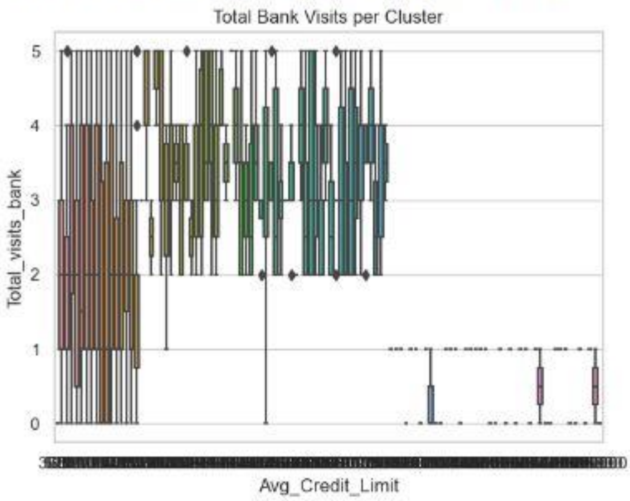


Fig: Total bank visits vs Avg credit limit

Step 3: Segment Interpretation

Once we have the mean values and visualizations, we can interpret each segment:

1. **Cluster 1 (High Credit, Low Interaction)**:
   - Customers in this cluster may have high credit limits and multiple credit cards, but interact less with the bank (fewer visits, calls, or online logins).
   - **Actionable Insight**: These are high-value customers who might be satisfied with the bank's services. Focus on upselling or premium services.

2. **Cluster 2 (Low Credit, High Interaction)**:
   - Customers in this cluster may have low to medium credit limits but frequent interactions with the bank, indicating they require more support or have issues with services.
   - **Actionable Insight**: Improve customer service efficiency for this segment and offer tailored credit products to increase engagement.

3. **Cluster 3 (Moderate Credit, High Engagement)**:
   - Customers have moderate credit limits and are actively using both bank branches and online services.
   - **Actionable Insight**: Offer personalized campaigns to these customers for product cross-selling or loyalty rewards based on their frequent usage.

- The boxplot shows **Total Bank Visits per Cluster** in relation to the credit limit.
  - Customers with high bank visits (left section) are spread across credit limits.
  - However, the right side (customers with fewer visits) tends to be associated with higher credit limits.

Insight: Higher credit limit customers prefer fewer in-person visits, hinting at a preference for online channels or self-service options.

Step 4: Compare Segments with Business Goals

- **Marketing Team Focus**: Look for clusters where customers have a moderate-to-high credit limit and multiple credit cards. These are good targets for upselling premium services or credit limit increases.
- **Operations Team Focus**: Focus on clusters with high bank visits or calls, indicating areas where customer service improvements can be made. These customers likely need faster query resolution.

Step 5: Recommendations

Based on the profiles and insights from clustering, here are some actionable recommendations:

1. **Cluster 1 (High Credit, Low Interaction)**:
   - **Recommendation**: Upsell premium credit products and personalized offers to retain high-value customers.
2. **Cluster 2 (Low Credit, High Interaction)**:
   - **Recommendation**: Improve customer service for these customers to reduce the need for frequent calls or visits. You could introduce self-service options or online chat support.
3. **Cluster 3 (Moderate Credit, High Engagement)**:
   - **Recommendation**: Run targeted campaigns that encourage product bundling or loyalty programs, rewarding frequent engagement.

Conclusion

This analysis helps you identify different customer segments, interpret their behavior, and generate actionable insights for marketing and operations. You can now advise AllLife Bank on how to tailor its marketing and service strategies to better address the needs of each customer segment.

Based on the customer segmentation, here are tailored **actionable insights** and **strategic recommendations** for AllLife Bank's Marketing and Operations teams:

1. Cluster 1: High Credit, Low Interaction

- **Profile**: These customers have high credit limits, multiple credit cards, but minimal interactions with the bank (fewer visits, calls, or online logins).
- **Insights**:
  - Likely satisfied with the bank's services since they rarely reach out.
  - Represent high-value customers with substantial credit potential.
- **Recommendations**:
  - **Upsell Premium Services**: Offer exclusive products like premium credit cards with added benefits, personalized loan offerings, or financial advisory services.
  - **Enhance Engagement with Loyalty Programs**: Since these customers don't interact much, introduce loyalty programs or rewards for using their existing credit cards more frequently.
  - **Send Personalized Communications**: Use targeted marketing to inform them about services that add value (e.g., higher credit limits or lower interest rates on loans).

2. Cluster 2: Low Credit, High Interaction

- **Profile**: Customers in this segment have lower credit limits but frequently interact with the bank via calls and visits.

- **Insights**:
  - High interaction frequency might indicate a need for support, potentially due to dissatisfaction or complex issues.
  - Lower credit limit could imply a lower income bracket or higher risk profile, requiring careful engagement.

Recommendations:

  - **Improve Customer Support Accessibility**: Streamline service channels for this segment by introducing dedicated customer support lines or self-service options in online banking to reduce dependency on in-person visits and calls.
  - **Introduce Entry-Level Credit Products**: Consider offering lower-tier credit products or microloans with manageable limits. Emphasize secure usage practices to encourage responsible credit behavior.
  - **Financial Literacy Campaigns**: Educate these customers on managing credit through informative emails, app notifications, or webinars, potentially increasing their creditworthiness and satisfaction over time.

3. Cluster 3: Moderate Credit, High Engagement

- **Profile**: Customers in this group hold moderate credit limits and frequently interact with the bank through various channels.
- **Insights**:
  - Actively engaged and likely open to bank promotions, potentially receptive to personalized offerings.
  - Likely have a mix of credit and banking needs, including cross-services.
- **Recommendations**:
  - **Cross-Sell Banking Products**: Market complementary products, such as savings accounts, investment plans, or insurance options, aligning with their engagement level.
  - **Introduce Tiered Rewards System**: Encourage continued interaction and spending by rewarding online transactions, in-branch visits, and regular usage of credit cards.
  - **Personalized Offers Based on Usage Patterns**: Use their frequent interactions to understand individual preferences and tailor offers such as cashback on specific spending categories or discounts on bank services.

Actionable Insights for Marketing and Operations

Based on the EDA, here are specific, data-driven insights:

- **For Marketing**:
    - **High Credit Limit Customers**: Target with premium product offers and upsell opportunities (e.g., rewards cards or concierge services).
    - **High Interaction Customers**: Offer personalized services that focus on building loyalty (e.g., cross-selling savings or investment products).
- **For Operations**:
    - **Customers with High Call Volumes**: Likely experiencing issues or dissatisfaction. Improve support efficiency by introducing self-service options or dedicated support lines.
    - **Frequent Bank Visitors**: Convert in-person visits to online interactions by promoting digital banking solutions and providing incentives for using online channels.

General Recommendations

1. **For the Marketing Team**:
    - **Targeted Campaigns**: Use the segmentation analysis to design personalized marketing messages that appeal to each cluster's unique needs and behaviors. For instance, sending premium card offers to high-value customers and cross-product offers to actively engaged users.
    - **Product Awareness**: Run awareness campaigns to introduce new digital banking features, especially for high-interaction segments that might benefit from online services to reduce branch dependency.
2. **For the Operations Team**:
    - **Enhanced Service Model**: Streamline customer service operations by directing high-frequency callers and visitors toward automated or online channels where possible, reducing operational strain.
    - **Customer Feedback Loops**: Periodically survey high-interaction customers (Cluster 2) to gain insights on their pain points, helping to refine the service delivery model further.

Summary

These recommendations enable AllLife Bank to:

- Boost customer satisfaction by addressing service inefficiencies.
- Drive revenue through targeted product offerings and personalized marketing.
- Retain high-value customers with exclusive offerings and loyalty incentives.

By leveraging these insights, AllLife Bank can optimize both its marketing and service delivery strategies, building stronger, more valuable customer relationships.