Customer Segmentation in Banking

Insights and Recommendations for Enhanced Financial Services

Abstract

The project called **"Customer Segmentation in Banking: Insights & Recommendations for Better Financial Services"** digs into customer data. The goal? To find different groups of customers! This helps banks redesign their services to fit what customers really want.

To start off, the study dives deep with Exploratory Data Analysis (EDA). We clean and sort the data, fixing any messiness to make it easy to look at. In fact, a big part of this project is figuring out which attributes matter the most. This way, we can sort customers into clear groups.

Through our analysis, we found five main customer segments: Young Professionals, Established Families, High Net-Worth Individuals (HNWI), Budget-Conscious Customers, & Credit Seekers. Each of these groups shows unique financial habits and likes, all explained in the report.

So, this project shows just how crucial it is for banks to segment their customers. Plus, it gives banks useful ideas on how they can improve their services to boost business.

Table of Contents

Introduction of the Study

Customer Segmentation is a super important practice. It helps financial institutions learn more about their different customers with this knowledge, they can serve them better. By grouping customers based on things like age, spending habits, & preferences, banks can easily adjust their products and services. This way, they can market to each specific group. Doing this often leads to happier customers and also helps the business grow. It lets banks use their resources in smarter ways and run campaigns that work better.

In this study, we will look at different customer groups using data analysis. We want to understand what makes each group unique. These insights will help create strategies that boost customer engagement and overall business performance. The findings of this study will be helpful for both people working in the field & researchers.

Banks collect tons of information from their customers. Think about transaction history, account balances, loans, & even more details! This data helps spot patterns and trends for segmentation. When banks understand how different customer groups behave, they can create special strategies that fit their needs better. This can lead to improved customer satisfaction, loyalty, & even higher profits!

<u>Objectives of the Study</u>

The main goal of this project is to look at customer segmentation in the banking world This helps us engage better with customers develop smarter business strategies. Here's what we'll focus on:

1. **Identify Key Customer Segments**: We will use data to find specific groups of customers in the banking industry. This will include things like demographics, their financial habits, & how they use different products.

2. **Customer Behaviour Analysis**: Next, we'll dive deep into the traits & actions of these customer groups. Understanding what they need, what they like, and how they handle money is super important.

3. **Formulate Focused Strategies**: After looking at the segments, we'll come up with targeted strategies. These will help improve customer engagement, keep them loyal, and boost revenue.

4. **Business Impact Analysis**: Finally, we'll see how these strategies could affect customer satisfaction, loyalty, & overall business success.

By working towards these goals, this study aims to give banks great insights. This way, they can adjust their services to fit different market segments. In doing so, they can really push for more growth & profits!

Literature Review (Background Study)

Customer segregation has always been important for banks. It's how they figure out who needs what. They use this info to match services with customer needs. This idea came from marketing and found its way to banking.

1. **How Bank Customer Segmentation Has Changed:** Over time, banks have really changed how they group customers. At first, they looked at simple things like age, gender, and how much money people made. Now, thanks to data analytics and machine learning, they consider many factors. They also look at things like how you spend your money, your credit scores, and even what you like to do.

2. **Techniques for Data-Driven Segmentation:** Big data has totally changed the game! Advanced methods help banks figure out different types of customers better than before. Some popular techniques now are cluster analysis, decision trees, & neural networks.

3. **Why Customer Segmentation is Key in Banking:** Customer segmentation helps banks find their best customers. This means they can sell more products & services—good for business! Plus, it helps with managing risk better. By understanding the risk level of customers, banks can create smarter strategies for approving loans & credits. This helps keep everything stable.

4. **Challenges Banks Face:** But it's not all smooth sailing. There are challenges too! First off, the quality of data is crucial. If the info is bad or missing, segmentation might not work well—leading to poor strategies. Also, customer behavior changes a lot! That means banks must keep updating their models regularly.

   Another big issue is about ethics—especially when it comes to privacy and security. Banks hold tons of customer data & must follow strict rules to keep that information safe from breaches.

5. **Trends for the Future:** Looking ahead, exciting trends in customer segmentation are on the rise! Banks are starting to use AI and machine learning more for these processes. This will help create more accurate & flexible models. With digital banking and fintech becoming popular, new factors like online behavior & social media usage are starting to matter too. These will be key in future segmentation strategies!

Research Methodology

The study's method is built around key steps. Each plays a role in reaching the main goal: create insights that boost customer engagement & business plans.

1. **Data Collection:** The study used data from Kaggle. This dataset includes things like age, marital status, balance, loan status, and default history. It's got about 80MB of info! Also, we gathered more data by surveying around 50 people with a structured questionnaire (you can see it in Annexure D).

2. **Data Preprocessing:** Before diving into analysis, we needed to prep the dataset well to make sure of its quality & consistency. Here's what we did:

   - **Handling Missing Values:** We took care of missing values by filling them in or removing incomplete records – whatever kept the study's quality high.

   - **Data Cleanup:** We fixed the inconsistent date formats and removed some columns like 'contact', 'duration', 'previous', & 'poutcome'.

   - **Feature Engineering:** We created new features by combining 'day', 'month', & 'year' into one neat 'date' column in DD-MM-YYYY format. We also added extra features like customer lifetime value & transaction frequency.

3. **Feature Selection:** We did an exploratory data analysis (EDA) to find key features for sorting customers. Features like age, loan status, balance, marital status, & job type stand out for understanding customer behavior.

4. **Clustering Algorithms:** In our quest to segment customers better, we tried out several clustering algorithms – K-Means, Hierarchical Clustering, and Density-Based Spatial Clustering of Applications with Noise.

5. **Data Analysis & Interpretation:** Next, we looked at patterns in each segment based on their features. Important attribute distributions across segments were shown using fun visualizations like histograms, pie charts, bar charts, & treemaps. We compared segments to spot big differences in financial habits, product use, and risk profiles.

6. **Tooling and Software:** This was all done using Python! We used Pandas for data handling and Scikit-learn for running the clustering algorithms. For visualizations, we relied on Matplotlib & Seaborn. The entire analysis happened in Google Colab for easy teamwork and efficient processing.

7. **Ethical Considerations:** Since this research involved sensitive customer data, we kept ethical issues at the forefront. We varied our methods for data anonymization to protect customer identities. Plus, we made sure our analysis followed all rules related to data privacy by not using any personal identifiers in our study.

Overall, we aimed for a careful approach throughout this research!

Data Analysis and Interpretation

Data analysis for this study was all about finding clear customer groups in banking data. We looked what makes each group unique—things like their habits and characteristics. This helped us get useful insights. These insights can aid in crafting focused customer engagement & strategies for business growth.

1. **Finding Customer Groups:** To break down the customers into different groups, we used the K-Means clustering algorithm. This method worked on both demographic & financial information. After running it with the best value of clusters, we found five distinct customer segments.

2. **Characteristics of Each Segment:** Every segment stands for a specific type of customer who shares similar traits & behaviors.

a) **Segment 1: Young Professionals**

- **Demographics:** Mostly younger folks aged 25 to 35, earning steady incomes.

- **Financial Behavior:** They have moderate account balances and are active users of online banking. Plus, they like personal loans.

- **Insights:** This group is tech-savvy! They respond nicely to digital banking solutions. Marketing can focus on digital services & personalized loan offers.

b) **Segment 2: Established Families**

- **Demographics:** Aged between 35 and 50; usually married with kids.

- **Financial Behavior:** They have higher account balances and use multiple products like mortgages & savings accounts.

- **Insights:** These customers value stability and long-term planning. Banks can market family-friendly products like insurance plans & retirement packages.

c) **Segment 3: High Net-Worth Individuals (HNWI)**

- **Demographics:** Typically older customers—usually over 50—with significant wealth.

- **Financial Behavior:** Very high account balances & frequent users of wealth management services.

- **Insights:** They may benefit from personalized wealth management services, special financial products, and priority banking.

d) **Segment 4: Price-Sensitive Customers**

- **Demographics:** A mix of ages; middle to low-income customers.

- **Financial Behavior:** Lower account balances and frugal spending, often using just basic checking accounts.

- **Insights:** This group might be looking for savings products, budget-friendly tools, and no-fee banking options.

e) Segment 5: Credit Seekers

- **Demographics:** Younger to middle-aged people; often with variable incomes.

- **Financial Behavior:** Moderate to high loan balances with frequent credit applications; they might sometimes default on payments.

- **Insights:** For this group, managing risk is super important! Banks could offer credit education programs, debt consolidation options, and tailored repayment plans.

3. **Visualizing the Segments:** To understand each segment better, we created some visuals:

- **Histograms:** These showed account balances, ages, & loan amounts across segments—really helpful for spotting patterns! Like which segment has higher balances or takes out more loans.

- **Pie Charts:** They displayed the portion of customers in each segment so it was easier to see the size and importance of each group to the bank.

- **Bar Charts:** These compared average balances, loan amounts, & product usage among segments—highlighting key financial behavior differences.

4. **Understanding the Findings:** Each segment is quite different in needs & preferences! For example, HNWI customers want services from wealth management experts while price-sensitive folks likely prefer simple & cheap banking options. The segmentation also shows paths for targeted marketing and product development. By understanding what each group needs specifically, banks can offer great services that lead to happier customers who are more loyal!

5. **Implications for Business Strategy:** What we found in this study has some important takeaways for business strategy:

- **Personalized Marketing:** We can create tailored campaigns that focus on exactly what each segment needs regarding products & services.

- **Product Development:** New tools can be designed to cater to specific groups—for instance, digital solutions for young professionals or retirement help for established families.

- **Risk Management:** Knowing about segments like credit seekers helps banks manage risks better and reduce chances of defaults along with financial losses.

This study gives banks great ideas on how to connect with their customers more effectively!

Recommendations and Conclusions

Customer segmentation research has helped the bank understand its customers better. It shows the diverse needs & behaviors have. Based on this, we can suggest a few recommendations and conclusions:

Recommendations

1. **Personalized Banking Solutions• Tailored Products:** The bank should create products that fit the specific needs of different customer groups For instance, young professionals could benefit from digital banking options and custom loan offers. On the other hand, High Net-Worth Individuals (HNWI) might appreciate exclusive investments & wealth management services.• **Segment-Specific Marketing:** Marketing strategies should speak directly to each group's unique traits. For example, ads promoting low-cost banking may not attract Established Families as much as they would Budget-Conscious Customers. These customers likely care more about long-term plans & family-friendly products.

2. **Enhanced Digital Experience• Digital Engagement:** Technology keeps growing in banking, especially for younger folks. A smart move for the bank would be to improve its digital platforms even more! With personalized dashboards, AI-driven financial advice, and smooth mobile experiences, customer satisfaction could really take off!• **Data-Driven Decision Making:** This focus will let the bank make informed choices across all areas—marketing, product development, and customer service too! Predictive analytics can help guess what a customer might need or want next. This way, the bank stays a step ahead!

3. **Risk Management & Credit Education• Targeted Risk Mitigation:** Customers like Credit Seekers need special support when it comes to managing risks since they often struggle with loans. The bank could offer credit education programs, debt consolidation help, & personalized repayment plans to lower risk levels.• **Credit Monitoring:** Keeping a close eye on high-risk sectors allows the bank to act quickly when it comes to managing loans, preventing potential issues before they arise.

4. **Customer Retention & Loyalty Programs• Loyalty Programs:** These programs should reward customers for sticking around and making big transactions. They're crucial for keeping customers happy—especially Young Professionals & HNWI! Benefits like better interest rates or exclusive events could be part of it!• **Feedback Mechanisms:** Regular feedback is key! It helps the bank understand how happy customers are and where improvements are needed. This way, offerings can be refined to meet customer demands effectively.

Conclusion:

Through segmentation analysis, we've found that the bank's customers are quite diverse—each with their own preferences & behaviors! By understanding these differences better, the bank can create smarter strategies that directly connect with them and provide suitable products and services.

Key features from this study include:

• Five distinct customer segments: Young Professionals, Established Families, High Net-Worth Individuals (HNWI), Budget-Conscious Customers, and Credit Seekers exist.

• There's a need for customized banking solutions along with targeted marketing campaigns.

• Strategies must focus on improving digital experiences while making informed decisions based on data—this is important to stay competitive in a changing market!

• Appropriate risk management policies should aim to cut down default rates & boost financial stability.

The insights from this study make a case for revamping the bank's customer engagement efforts by maximizing product offerings and enhancing overall satisfaction. By applying these suggestions, bank is set to strengthen its market position and grow steadily over the next few years!

Bibliography and References

1. *Kaggle (https://www.kaggle.com/datasets):*

   bank_cust_segment

   bank_transactions

   cust_category

   [retrieved from Kaggle Datasets]

2. *Survey Participants: This project benefited from the responses of 48 individuals who contributed valuable insights through the survey.*

3. *Special Thanks: I would like to express my gratitude to my brother and friends for their assistance and support throughout this project.*

<u>Appendix</u>

Appendix A: Dataset Description

Customer Segmentation Dataset

- **Source**: Kaggle

- **Description**: Contains customer demographic and financial information used for segmentation analysis.

- **Key Columns**:

    o age: Age of the customer

    o job: Type of job

    o marital: Marital status

    o balance: Account balance

    o loan: Whether the customer has a loan

    o default: Whether the customer has credit default

Appendix B: Data Cleaning Scripts Extract of Python Script for Data Cleaning:

```
# Data Cleaning for Transactions
# Interpolate missing values for CustAccountBalance
1.  df_transactions['CustAccountBalance'] =
    df_transactions['CustAccountBalance'].interpolate(method='linear')
2.
3.  # Function to parse dates in various formats
4.  def parse_dates(date_str):
5.    for fmt in ('%d-%m-%Y', '%m/%d/%Y', '%Y-%m-%d'):
6.      try:
7.        return pd.to_datetime(date_str, format=fmt)
8.      except ValueError:
9.        continue
10.   return pd.to_datetime(date_str, errors='coerce')
11.
12. # Apply date parsing function and fill missing values with median date
13. df_transactions['CustomerDOB'] =
    df_transactions['CustomerDOB'].apply(parse_dates)
14. median_dob = df_transactions['CustomerDOB'].median()
```

```
15. df_transactions['CustomerDOB'] =
    df_transactions['CustomerDOB'].fillna(median_dob)
16.
17. # Calculate age and categorize into age groups
18. df_transactions['CustomerAge'] = (pd.to_datetime('today') -
    df_transactions['CustomerDOB']).dt.days // 365
19. bins = [0, 18, 25, 35, 45, 55, 65, 100]
20. labels = ['0-18', '19-25', '26-35', '36-45', '46-55', '56-65', '66+']
21. df_transactions['AgeGroup'] = pd.cut(df_transactions['CustomerAge'],
    bins=bins, labels=labels, right=False)
22.
23. # Data Cleaning for Segments
24. df_segments = df_segments.drop(columns=['contact', 'duration',
    'previous', 'poutcome'])
25. df_segments['Date'] = df_segments['day'].astype(str).str.zfill(2) + '-' +
    df_segments['month']
26. df_segments = df_segments.drop(columns=['day', 'month'])
27.
28. # Save cleaned datasets
29. df_transactions.to_csv('cleaned_transactions.csv', index=False)
30. df_segments.to_csv('cleaned_segments.csv', index=False)
```

2. Explanation of Cleaning Steps

- **Dropped Columns**: Columns not relevant for the analysis were removed to streamline the dataset.

- **Date Combination**: The 'day' and 'month' columns were combined to create a unified date format for easier analysis.

Appendix C: Data Analysis Outputs

1. Sample Output of Clustering Analysis

- **Cluster Canters**: Displays the centroid values for each cluster.

- **Cluster Assignments**: Shows which cluster each customer has been assigned to.

2. Visualizations

- **Histograms**: Distribution of key variables (e.g., age, balance).

- **Bar Charts**: Comparison of average balances and loan amounts across segments.

Appendix D: Survey Questionnaire

- **Survey Link**: https://forms.gle/Ctwe44BQnsuZ7X267

- **Survey Questions**:

    Name

    Age Group

    - *18 - 24*

    - *25 - 34*

    - *35 - 44*

    - *45 - 54*

    - *55 - 64*

    - *65 +*

    Gender

    - *Male*

    - *Female*

    - *Non-Binary*

    - *Prefer not to say*

    Marital Status

    - *Single*

    - *Married*

    - *Divorced*

    - *Widowed*

Where are you currently located in India?

- *North*
- *South*
- *East*
- *West*

City

State

Do you have a Bank Account?

- *Yes*
- *No*

If Yes, Where is your Bank Branch at?

- *Same as my location mentioned above*
- *N/A*
- *Other…*

Do you currently have a loan with any bank or otherwise?

- *Yes*
- *No*
- *Prefer not to say*

If yes, what type of loan(s) do you have?

- *Personal Loan*
- *Home Loan*
- *Car Loan*
- *Other*

Which types of transactions do you perform most frequently?

- *Withdrawals*
- *Deposits*
- *Transfers*
- *Bill Payments*
- *Investment Transactions*
- *Other…*

What is your usual range of transaction?

- *Less than ₹1,000*
- *₹1,000 - ₹5,000*
- *₹5,000 - ₹10,000*
- *₹10,000 - ₹50,000*
- *More than ₹50,000*

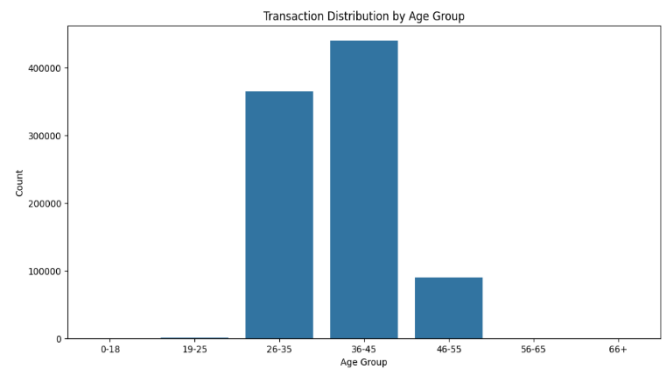### *Frequency of Bank Visits*

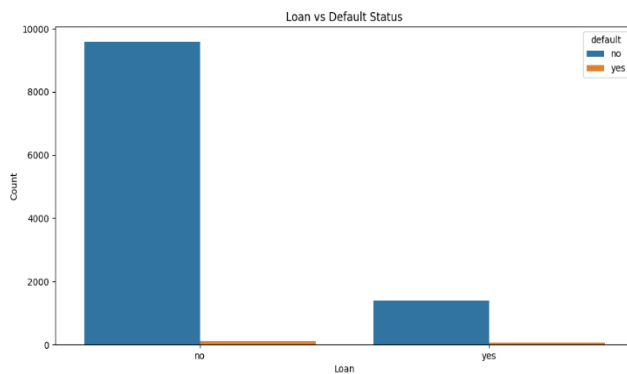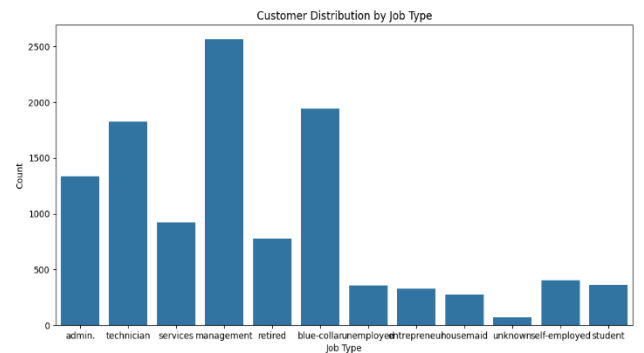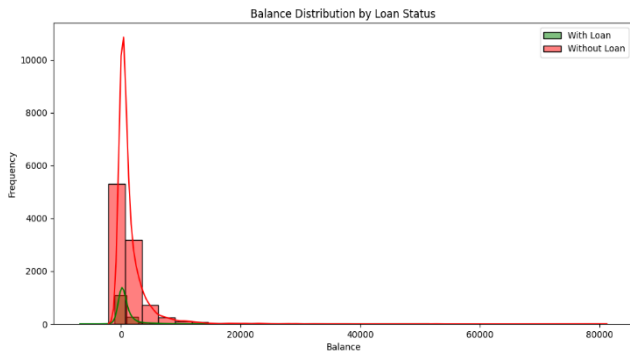- *Daily*
- *Weekly*
- *Monthly*
- *Rarely*
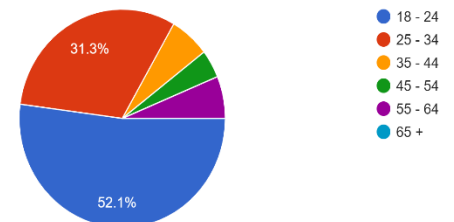
Do you own multiple banks?

- *Yes*
- *No*

How satisfied are you with your bank's overall services? *(if you have 2 or more banks, give an average rating)*

- *1 - Very Dissatisfied*
- *2*
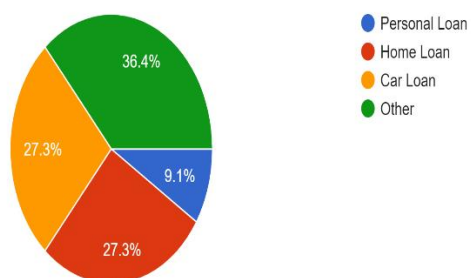- *3*
- *4*
- *5 - Very Satisfied*

Appendix E: Reports (Visualizations)

Balance Distribution by Loan Status

Customer Distribution by Job Type

Loan vs Default Status

Transaction Distribution by Age Group

Customer Distribution by Marital Status
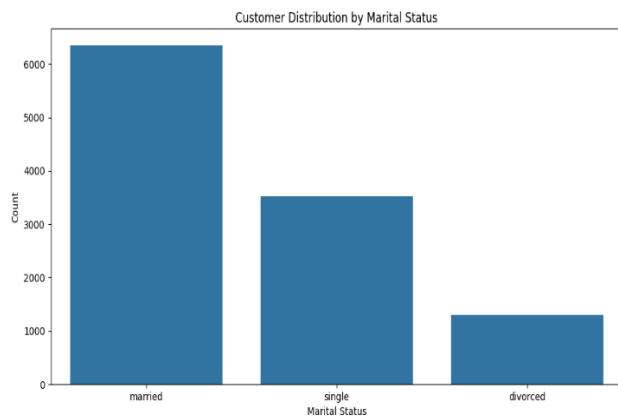
Age Group
48 responses

If yes, what type of loan(s) do you have?
11 responses

Which types of transactions do you perform most frequently?
48 responses