

**CIS 5550 BUSINESS ANALYTICS Fall 2023**

Exploring Customer Attrition Patterns and Financial Behavior in Credit Card Usage



Mansi vivekanand dhoke (402702401)

Abhiram Natakarani (402707510)

Vamsi Sai Krishna Reddy (402705235)

**1a. Introduction**

Credit card usage has become an integral part of modern financial transactions, offering individuals and businesses a convenient and flexible means of payment. These plastic cards, issued by financial institutions, enable cardholders to make purchases, access credit, and manage their finances with ease. Credit cards work by providing a line of credit, allowing users to spend up to a predefined limit, typically based on their creditworthiness. They offer benefits such as cashback rewards, points, and travel miles, making them attractive for consumers. Additionally, credit cards offer a grace period during which cardholders can pay off their balances without incurring interest charges. However, prudent management is essential, as excessive credit card debt can lead to financial difficulties. Responsible usage, timely payments, and monitoring of statements are critical to maximizing the advantages of credit cards while avoiding potential pitfalls. Understanding terms and conditions, such as interest rates, fees, and credit limits, is key to making informed financial decisions in the world of credit card usage.

Businesses are motivated to analyze credit card usage data for a variety of reasons. This data can be used to understand customer behavior and preferences, identify fraud, and risk, and improve business operations. For example, businesses can use credit card usage data to identify their most loyal customers and offer them special discounts or promotions. They can also use this data to identify customers who are at risk of defaulting on their payments and offer them debt counseling or other assistance. In addition, credit card usage data can be used to develop new products and services, such as a credit card that offers rewards for travel or dining. Government agencies can also use this data to track economic trends or investigate financial crimes. Overall, credit card usage data is a valuable tool that can be used for a variety of purposes. By analyzing this data, businesses can better understand their customers, improve their products and services, and reduce risk.

By analyzing spending patterns, transaction history, and demographic information, businesses can gain a deeper understanding of their customer base. This knowledge can be used to develop targeted marketing campaigns, personalize product offerings, and improve customer service strategies.

Analyzing customer data can help identify patterns that may indicate fraudulent activity or potential risk. This allows businesses to implement effective fraud detection mechanisms, protect customers' financial information, and minimize losses. By analyzing aggregated customer data, businesses can identify emerging trends in customer preferences, spending habits, and market demands. This information can guide product development, marketing strategies, and overall business decisions.

**(1) How does attrition vary between genders?**

**(2) What are the age groups of credit card holders?**

**(3) What is the distribution of transactions over different card types?**

**(4) What is the Credit Limit Distribution by education level?**

**1b. Background Information about the Dataset**

<https://www.kaggle.com/datasets/sakshigoyal7/credit-card-customers>

**c. Screenshots of Dataset**

**A screenshot of a computer

Description automatically generated**

**d. Data Description**

| **Field Name** | **Data Description** | **Definition** | **Example Value** |
| --- | --- | --- | --- |
| CLIENTNUM | Client number ID | A unique identifier for each customer. | 1234567890 |
| Attrition\_Flag | Customer activity, Attrited or existing | Whether the customer has churned (left the bank's credit card services) or is still an active customer. | Attrited or Existing |
| Customer\_Age | Customer's Age in Years | The customer's age in years. | 30 |
| Gender | Customer's gender, male or female | The customer's gender. | Male or Female |
| Dependent\_count | Number of dependents | The number of people who depend on the customer financially. | 2 |
| Education\_Level | Educational Qualification of the account holder | The customer's highest level of education. | Graduate, High School, Unknown, Uneducated, College, Post-Graduate, Doctorate |
| Marital\_Status | Marital Status of the account holder | The customer's marital status. | Married, Single, Divorced, Unknown |
| Income\_Category | Annual Income Category of the account holder | The customer's annual income category. | Less than $25,000, $25,000-$49,999, $50,000-$74,999, $75,000-$99,999, $100,000 or more |
| Card\_Category | Type of Card (Blue, Silver, Gold, Platinum) | The type of credit card the customer has. | Blue, Silver, Gold, Platinum |
| Months\_on\_book | Period of relationship with bank | The number of months the customer has been a customer of the bank. | 12 |
| Total\_Relationship\_Count | Total number of products held by the customer | The total number of products the customer has with the bank, including credit cards, loans, and deposit accounts. | 3 |
| Months\_Inactive\_12\_mon | Number of months inactive in the last 12 months | The number of months in the last 12 months that the customer has not used their credit card. | 2 |
| Contacts\_Count\_12\_mon | Number of Contacts in the last 12 months | The number of times the customer has contacted the bank in the last 12 months. | 3 |
| Credit\_Limit | Credit Limit on the Credit Card | The maximum amount of credit the customer is allowed to use on their credit card. | $10,000 |
| Total\_Revolving\_Bal | Total Revolving Balance on the Credit Card | The total amount of debt the customer has on their credit card. | $5,000 |
| Avg\_Open\_To\_Buy | Open to Buy Credit Line (Average of last 12 months) | The average amount of credit the customer has available to use on their credit card over the last 12 months. | $5,000 |
| Total\_Amt\_Chng\_Q4\_Q1 | Change in Transaction Amount (Q4 over Q1) | The change in the customer's transaction amount from the fourth quarter of the previous year to the first quarter of the current year. | $100 |
| Total\_Trans\_Amt | Total Transaction Amount (Last 12 months) | The total amount of transactions the customer has made on their credit card in the last 12 months. | $10,000 |

**e. Justification for the Choice of the Dataset:**

The choice of the 'Credit Card Customers' dataset for the business analytics project, which aims to explore customer attrition patterns and financial behavior in credit card usage using Tableau, can be justified based on several factors:

* **Relevance to Business Objectives:** The dataset is directly related to credit card customers, aligning with the project's focus on understanding customer attrition patterns and financial behavior in credit card usage. This ensures that the analysis will be highly relevant to the business objectives.
* **Rich Demographic Information:** The dataset includes diverse demographic variables such as customer age, gender, education level, marital status, and income category. This rich demographic information allows for a comprehensive analysis of customer profiles, helping to identify patterns and trends that may influence attrition.
* **Product and Financial Variables:** The inclusion of product-related variables like card category and financial variables such as credit limit, total revolving balance, and transaction amounts provides insights into customers' financial behavior. This enables a thorough exploration of factors affecting customer attrition.
* **Temporal Aspects:** The dataset contains temporal information, including the length of the relationship with the bank (months on book) and the number of months inactive in the last 12 months. Analyzing these temporal aspects can reveal trends over time and help identify critical points in the customer lifecycle.
* **Transaction Metrics:** Key transaction metrics like total transaction amount, total transaction count, and changes in transaction amounts and counts are available. These metrics are crucial for understanding customer behavior and can contribute to predicting attrition based on transactional patterns.
* **Customer Interaction Data:** The dataset includes information on the number of contacts in the last 12 months. Analyzing customer interactions can provide insights into customer engagement and its impact on attrition.
* **Unique Identifier:** The 'CLIENTNUM' column serves as a unique identifier for each customer, facilitating individual-level analysis and tracking customer-specific trends and behaviors over time.
* **Data Quality and Completeness:** The dataset appears to be well-structured with relevant information for the intended analysis. The completeness and quality of the data are critical for the success of any analytics project, and this dataset seems to meet these criteria.
* **Scope for Advanced Analytics:** The dataset's variables offer opportunities for advanced analytics, such as predictive modeling for customer attrition, segmentation analysis, and the creation of data visualizations to communicate insights effectively.

In summary, the 'Credit Card Customers' dataset is a suitable choice for the business analytics project due to its relevance to the objectives, richness of demographic and financial information, inclusion of temporal aspects, and potential for advanced analytics using Tableau.

**f. Evaluation of the Quality and Relevance of the Chosen Dataset:**

* **Data Completeness:** The dataset appears to be relatively complete with no apparent missing values in the provided column descriptions. This is crucial for meaningful analysis and insights, as missing data can introduce bias and affect the validity of the results.
* **Data Consistency:** The dataset seems to be consistent in terms of the format and type of data in each column. Consistency is essential for accurate analysis, and the uniform structure of the dataset contributes to its reliability.
* **Unique Identifier:** The 'CLIENTNUM' column serves as a unique identifier, which is essential for individual-level analysis and tracking customer-specific trends. This unique identifier ensures that each record corresponds to a distinct customer, supporting the integrity of the analysis.
* **Relevance of Variables:** The chosen variables align well with the objectives of exploring customer attrition patterns and financial behavior in credit card usage. Demographic variables (e.g., age, gender, education, marital status) and financial variables (e.g., credit limit, transaction amounts) are relevant for understanding customer behavior and predicting attrition.
* **Temporal Aspects:** The inclusion of temporal aspects such as 'Months\_on\_book' and 'Months\_Inactive\_12\_mon' allows for the analysis of trends over time. This temporal dimension is crucial for identifying patterns in customer behavior and understanding how relationships with the bank evolve.
* **Product and Financial Metrics:** The dataset includes important product and financial metrics like 'Card\_Category,' 'Total\_Relationship\_Count,' and 'Avg\_Utilization\_Ratio.' These metrics provide insights into the types of products customers hold and their financial behavior, contributing to a comprehensive analysis.
* **Event Variable (Attrition\_Flag):** The 'Attrition\_Flag' variable, indicating whether an account is closed or not, is a key event variable for the analysis of customer attrition. Its binary nature (0 or 1) simplifies classification tasks and supports predictive modeling.
* **Diversity in Demographic Information:** The dataset includes a diverse set of demographic variables, such as 'Dependent\_count,' 'Education\_Level,' and 'Income\_Category.' This diversity allows for a nuanced analysis of how different demographic factors may influence customer behavior and attrition.
* **Transaction Metrics :** Transaction-related metrics like 'Total\_Trans\_Amt,' 'Total\_Trans\_Ct,' and 'Total\_Ct\_Chng\_Q4\_Q1' provide valuable information on customer engagement and transactional patterns. These metrics are essential for understanding financial behavior.
* **Potential for Visualization :** The dataset is well-suited for visualization in Tableau, given the variety of variables and their relevance to the business objectives. Visualizations can enhance the interpretation of trends and patterns in customer behavior.

In conclusion, the 'Credit Card Customers' dataset demonstrates good quality and relevance for the business analytics project. Its completeness, consistency, relevant variables, and the presence of a unique identifier make it a suitable choice for exploring customer attrition patterns and financial behavior in credit card usage using Tableau.

**g. Discussion of Data Issues Encountered During the Analysis:**

* **Missing Values:** One common issue encountered during the analysis was the presence of missing values in certain columns. For instance, the 'Education\_Level,' 'Marital\_Status,' and 'Income\_Category' columns may have 'Unknown' values, which could impact the accuracy of demographic analyses. Strategies such as imputation or careful handling of these unknown values may be necessary.
* **Imbalance in Attrition\_Flag:** The 'Attrition\_Flag' variable may exhibit class imbalance, with a disproportionate number of 0s (indicating accounts that are not closed) compared to 1s (indicating closed accounts). Imbalance can affect the performance of predictive models, and techniques such as oversampling or undersampling may be considered to address this issue.
* **Zero Values in Transaction Metrics:** Zero values in columns like 'Total\_Trans\_Amt' and 'Total\_Trans\_Ct' may indicate a lack of transactional activity. Understanding the reasons behind such occurrences is crucial, as it may influence the interpretation of customer behavior and engagement patterns.
* **Potential Data Entry Errors:** While 'CLIENTNUM' serves as a unique identifier, potential data entry errors or duplications should be investigated. Anomalies in this identifier could lead to misinterpretations in individual-level analyses.
* **Limited Granularity in Income Category:** The 'Income\_Category' variable provides income ranges, but the broad categorization may limit the granularity of the analysis. Exploring more detailed income data or creating additional income-related features could enhance the precision of financial behavior analysis.

Addressing these data issues through data cleaning, validation, and, if necessary, collaboration with domain experts is crucial to ensure the reliability and accuracy of the insights derived from the 'Credit Card Customers' dataset during the business analytics project.

**2.Data Collection and Cleaning**

**Accessing the Dataset:**

**SOQL code:**

*SELECT \* FROM BC;*

A screenshot of a computer

Description automatically generated

**Replacing *Female* values with *F* the date column:**

**SOQL code:**

*Update BC*

*Set Gender = REPLACE(Gender,’Female’,’F’);*

A screenshot of a computer

Description automatically generated

**Checking for null values:**

**SOQL Code:**

*SELECT COUNT (\*) FROM Bc WHERE Credit \_Limit IS NULL;*

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**View updated Dataset:**

**SOQL Code:**

*UPDATE BC*

*SET Credit \_Limit = Replace(Credit\_Limit,’NULL’,’0’);*

**A screenshot of a computer

Description automatically generated**

**A screen shot of a computer

Description automatically generated**

**1. Use of SQL Queries for Data Cleaning:** The project utilized SQL queries to address various data cleaning issues, ensuring that the data was properly formatted and prepared for analysis. This approach is effective for handling large datasets and allows for precise and efficient data manipulation.

**2. Checking for Missing Values:** A key part of the data cleaning process involved checking each column for missing values. Identifying and addressing missing data is crucial for ensuring the integrity of the analysis, as missing values can lead to biased or inaccurate results.

**3. Focus on Integrity and Relevance of Data:** The overarching objective of these data cleaning and preprocessing steps was to ensure the integrity and relevance of the dataset for the analysis. This highlights the importance of thorough data preparation in any analytical project, as it directly impacts the quality of insights derived from the data

**3.Explanatory Data Analysis**

**Question #1 of Analysis:**

How does attrition vary between genders?

**A screenshot of a computer

Description automatically generated**

**A screenshot of a graph

Description automatically generated**

**Interpretation of Results**

From the visualizations, the attrition rate among credit card customers is similar for both genders, but it is slightly lower for males when compared to females. This means that the rate at which both male and female credit card customers close their accounts or stop using the card is generally comparable, but there is a slightly higher level of customer retention among men.

Several factors can contribute to this gender-specific difference in attrition rates among credit card customers. These factors may include variations in spending habits, financial stability, preferences for certain card features, or even differences in how customer service is perceived by men and women. Understanding these factors can help credit card companies tailor their services and offerings to better meet the needs and preferences of their customers, regardless of gender, to enhance customer loyalty and satisfaction. There are several factors that contribute to the high attrition rate for credit card customers. Some of the most common reasons include:

* **High interest rates:** Credit card interest rates are typically much higher than the interest rates on other types of loans, such as mortgages or car loans. This can make it difficult for customers to pay off their credit card balances, which can lead to delinquency and ultimately churn.
* **Fees:** Credit card companies charge a variety of fees, such as annual fees, late fees, and balance transfer fees. These fees can add up quickly and make it more expensive to use a credit card, which can lead customers to cancel their cards.
* **Poor customer service:** Many credit card customers have negative experiences with customer service representatives. This can lead to frustration and dissatisfaction, which can ultimately lead customers to churn.
* **Lack of rewards:** Credit card companies offer a variety of rewards programs, but many customers do not find these rewards to be valuable enough to keep them using their cards.
* **New card offers:** Credit card companies are constantly offering new cards with attractive sign-up bonuses and rewards programs. This can tempt customers to switch cards, even if they are happy with their current card.

**Business Actionable and Insights:**

1. **Targeted Retention Strategies:**

**Insight:** While the attrition rates are generally similar between genders, there's a slightly higher retention rate among male customers.

**Actionable:** Can implement targeted retention strategies for female customers to bridge the gap. This could include personalized promotions, loyalty programs, or communication campaigns to enhance their overall satisfaction and loyalty.

1. **Understanding Gender-Specific Preferences:**

**Insight:** Various factors contribute to gender-specific differences, such as spending habits and preferences for card features.

**Actionable:** Can conduct surveys or gather additional data to understand the specific preferences of male and female customers. Can use this insight to customize product offerings and marketing strategies that resonate with each gender group.

1. **Addressing High Interest Rates:**

**Insight:** High interest rates contribute to customer churn.

**Actionable:** Can explore options to either reduce interest rates, provide lower interest rate offers for loyal customers, or offer financial education resources to help customers manage their credit card balances more effectively.

By acting on these insights, credit card companies can create a more customer-centric approach, fostering loyalty, satisfaction, and sustainable long-term relationships with their diverse customer base.

**Question #2 of Analysis:**

What are the age groups of credit card holders?

A diagram of a distribution of customer ages

Description automatically generated

**Interpretation of Results:**

The histogram shows the distribution of credit card holders by age group. Each bar's height shows the proportion of credit card holders in that age group. The age range is shown by the x-axis, and the fraction of credit card holders is represented by the y-axis.

The following are key results from the histogram:

* Younger age groups have a higher representation among credit card holders. The bars corresponding to younger age groups (e.g., 20-29, 30-39) are smaller than the bars for older age groups (e.g., 40-60). This indicates that a larger proportion of credit card holders fall within adult age brackets.
* The distribution of credit card holders across age groups is relatively uniform. The bars are consistent in height, suggesting that credit card ownership is prevalent across a wide range of ages.

**Connection between findings and the initial problem or question:**

* "What are the age groups of credit card holders?" was the opening query. The histogram answers this issue clearly by visually displaying the distribution of credit card holders across different age groups. The results of the histogram can be used to inform several elements of credit card marketing, education, and policy, including:
* Tailoring marketing campaigns to specific age groups: By understanding the distribution of credit card holders, marketing efforts can be targeted more effectively towards the most receptive age groups.
* Developing age-specific financial education programs: Financial literacy initiatives can be designed to address the specific needs and behaviors of credit card holders within different age brackets.
* Considering age demographics when formulating credit card policies: Understanding the age distribution of credit card holders can help guide policy decisions related to credit card accessibility, interest rates, and rewards programs.

Overall, the histogram provides valuable insights into the age composition of credit card holders. These insights can be used to inform strategies that enhance financial well-being and responsible credit card usage among different age groups.

**Business Actionable and Insights**

**Insights:**

* Younger generations dominate the credit card market: The histogram clearly shows that a significantly larger portion of credit card holders belong to younger age groups. This indicates a strong affinity towards credit cards among young individuals.
* Credit card ownership is widespread: The relatively consistent bar heights across age brackets suggest that credit cards are not limited to a specific age group, but rather widely used across a large age spectrum.

**Business Actions:**

**Marketing:**

* Target younger age groups: Tailor marketing campaigns towards younger generations (e.g., 20-39), leveraging social media, influencer marketing, and digital platforms frequented by this demographic.
* Develop age-specific messaging: Craft marketing messages that resonate with the specific needs and aspirations of different age groups. Emphasize the benefits relevant to their specific financial goals and lifestyle choices.
* Utilize visuals and content appealing to different age groups: Use visuals and content formats that resonate with the target age group's preferences and online behavior.

**Financial Education:**

* Develop targeted financial literacy programs: Design educational programs tailored to the specific needs and financial literacy gaps of different age groups. Address common financial challenges and responsible credit card usage practices relevant to each age bracket.
* Partner with educational institutions and communities: Collaborate with schools, universities, and community organizations to deliver financial education programs directly to young audiences.
* Utilize digital platforms for financial education: Create engaging online resources and interactive tools available through mobile apps and websites to reach young adults in their preferred digital spaces.

**Policy and Product Development:**

* Consider age-based accessibility: Analyze creditworthiness and risk factors associated with different age groups to determine appropriate credit card accessibility and approval criteria for each demographic.
* Develop age-specific credit card products: Design credit cards with features and rewards programs tailored to the specific needs and spending habits of different age groups.
* Implement age-based interest rate structures: Consider offering competitive interest rates or introductory offers that cater to the financial circumstances of young credit card holders.

By implementing these business actions based on the valuable insights gleaned from the histogram, companies can effectively target their marketing efforts, develop targeted financial education programs, and create credit card products and policies that are more inclusive and responsive to the needs of diverse age groups.

Remember, these are just starting points.

Further analysis and market research are crucial to refine these actions and ensure they are effective in achieving specific business objectives.

**Question #3 of Analysis:**

What is the distribution of transactions over different card types?

**A pie chart of credit cards

Description automatically generated**

**Interpretation of Results:**

The pie chart you sent shows the distribution of the total number of credit card transactions across four different credit card categories: Blue, Gold, Platinum, and Silver. The chart shows that the Blue credit card category has the highest number of transactions (39,870,938), followed by Silver (3,657,718), Gold (891,531), and Platinum (179,995).

This suggests that Blue credit cards are the most popular type of credit card among the customers represented in the chart. This could be due to a number of factors, such as the fact that Blue credit cards typically have lower fees and interest rates than other types of credit cards, or that they are more widely accepted by merchants.

It is also worth noting that the Blue credit card category accounts for over 90% of all transactions. This suggests that a small number of customers are responsible for a large majority of the credit card spending in the chart. This could be due to the fact that these customers are using their credit cards for large purchases, such as cars or homes.

Overall, the pie chart provides a valuable overview of the distribution of credit card transactions across different credit card categories. This information could be used by credit card companies to develop targeted marketing campaigns and product offerings.

Here are some additional insights that can be drawn from the data:

The Blue and Silver credit card categories together account for over 99% of all transactions. This suggests that the vast majority of customers in the chart are using credit cards that are primarily focused on affordability and convenience.

The Gold and Platinum credit card categories account for a very small percentage of all transactions, but they still have a significant impact on the overall spending. This suggests that a small number of customers are using their credit cards for high-value purchases.

The Platinum credit card category has the highest average transaction amount of all four categories. This suggests that Platinum cardholders are the most likely to use their credit cards for large purchases.

Overall, the pie chart suggests that the credit card market is dominated by Blue and Silver credit cards, which are primarily focused on affordability and convenience. However, there is still a significant demand for Gold and Platinum credit cards among a small number of customers who are looking for more premium features and benefits.

**Business Actionable and Insights: Abhiram**

**Question #4 of Analysis:**

What is the Credit Limit Distribution by education level?

A blue and green color scheme

Description automatically generated with medium confidence

**Interpretation of Results:**

The tree map depicts the distribution of credit limit total by education level. The size of each square represents the total credit limit for a specific academic level. The color of each rectangle represents the average credit use ratio for that academic level.

The tree map reveals several key findings:

* Higher education levels are associated with higher credit limits. The size of the rectangles for higher education levels (e.g., Graduate degree, High School diploma) is significantly larger than the size of the rectangles for other education levels (e.g., Post-Graduate, Doctorate). This suggests that people with higher education levels are more likely to be approved for credit cards and have least credit limits.
* Average credit utilization ratios are higher for people with lower education levels. The color of the rectangles for lower education levels is generally darker than the color of the rectangles for higher education levels. This indicates that people with lower education levels tend to use a higher percentage of their available credit.

**Clear connection between findings and the initial problem or question:**

The initial question was "What is the Credit Limit Distribution by education level." The tree map provides a clear and concise answer to this question by visually illustrating the distribution of credit limit sum by education level. The findings from the tree map can be used to address several problems related to credit access and utilization, such as:

* Identifying populations that may be at risk of over-indebtedness.
* Developing targeted financial education programs to help people with lower education levels make informed decisions about credit.
* Advocating for policies that promote equitable access to credit for all borrowers.

Overall, the tree map provides a valuable tool for understanding the relationship between education level and credit limit. The findings from the tree map can be used to address a number of important problems related to credit access and utilization.

**Business Actionable and Insights**

**Insights:**

* Higher education levels are associated with higher credit limits: This suggests that individuals with higher education levels have better access to credit and potentially higher borrowing capacity.
* Average credit utilization ratios are higher for people with lower education levels: This indicates that individuals with lower education levels may be more likely to utilize a higher percentage of their available credit, potentially putting them at risk of over-indebtedness.

**Business Actions:**

**Financial Education:**

* Develop targeted programs for lower education levels: Design educational programs specifically tailored to individuals with lower education levels, focusing on responsible credit card usage, budgeting, and debt management.
* Partner with community organizations and educational institutions: Collaborate with organizations and institutions serving lower-income communities to provide financial education workshops and resources readily accessible to their target audience.
* Utilize diverse educational formats: Offer a variety of educational materials, including interactive online modules, videos, and infographics, to cater to different learning styles and preferences.

**Credit Access and Policy:**

* Advocate for fair lending practices: Support policies and regulations that promote equitable access to credit for individuals of all educational backgrounds.
* Offer alternative credit products: Develop credit card products with lower starting limits and tailored features to encourage responsible credit building among individuals with lower education levels.
* Implement alternative credit scoring models: Explore the use of alternative data sources beyond traditional credit scores to assess creditworthiness and offer fairer access to credit for individuals with limited credit history.

**Debt Management:**

* Develop personalized debt management tools: Offer online tools and resources that help individuals track their credit utilization, manage their debt effectively, and develop personalized financial plans.
* Partner with credit counseling agencies: Collaborate with credit counseling agencies to offer free or low-cost debt counseling services to individuals struggling with high credit utilization or over-indebtedness.
* Provide early warning systems: Implement systems that notify individuals when their credit utilization reaches high levels, allowing them to take proactive measures to avoid further debt accumulation.

**Marketing:**

* Develop targeted marketing campaigns: Tailor marketing messages and channels to resonate with individuals of different educational backgrounds. Focus on promoting responsible credit usage and highlighting the benefits of financial management.
* Offer educational content in marketing materials: Include informative content within marketing campaigns that educates consumers about credit basics, responsible credit usage, and available financial resources.
* Partner with financial influencers: Collaborate with financial influencers who resonate with specific target audiences to reach individuals with educational content and promote responsible financial habits.

By taking these actionable and leveraging the insights gained from analyzing the relationship between education levels and credit limit, businesses can play a significant role in promoting financial literacy, ensuring responsible credit use, and ultimately contributing to a more equitable and financially stable society.

**4. Methodologies Adopted**

**1. Explanation of the Analysis Methods Used:**

* + **Treemap for Credit Limit Distribution by Education Level:** Created a treemap in Tableau, where each rectangle represents a different education level. The size of each rectangle corresponds to the proportion of customers in that education category. Implemented color gradients or patterns within each rectangle to represent the credit limit distribution. Utilized tooltips to display additional information, such as the average credit limit for each education level.
  + **Pie Chart for Distribution of Transactions Over Different Card Types:** Constructed a pie chart in Tableau to represent the distribution of transaction amounts among various card categories. Applied color-coding to distinguish between card types. Included data labels or percentages to highlight the contribution of each card category to the total transaction volume. Enhanced interactivity by enabling users to drill down into specific card types.
  + **Bar Chart for Attrition Variation Between Genders:** Developed a bar chart in Tableau with two bars representing attrition and non-attrition for each gender. Applied distinct colors to differentiate between closed and open accounts. Added labels to display the attrition percentages for both genders. Implemented interactivity for users to compare attrition rates easily.
  + **Histogram for Age Groups of Credit Card Holders:** Created a histogram in Tableau to illustrate the frequency distribution of customer ages. Defined appropriate age bins to capture meaningful age ranges. Applied color coding or overlays to differentiate attrition and non-attrition cases within the age distribution. - Integrated tooltips to display additional information, such as the count of customers in each age group.

**Additional Considerations:** Consistent Color Coding: Maintained consistent color coding across different visualizations to facilitate easy interpretation and comparison.Tooltips and Labels: Leveraged tooltips and labels in each visualization to provide users with additional context and details.These visualization techniques contribute to a holistic understanding of customer attrition patterns and financial behavior, providing stakeholders with actionable insights through an interactive and user-friendly interface in Tableau.

**2. Justification for Chosen Statistical Techniques:**

* **Visual Analysis for Demographics (Bar Charts, Histograms):**

Chosen for their simplicity and effectiveness in showcasing categorical data. These techniques allow for easy comparison between groups, crucial for understanding demographic trends in customer behavior.

* **Heatmaps for Multi-Variable Comparison:**

Effective in comparing large sets of data across different categories. Heatmaps can highlight trends and variances across categories like income groups and card types, offering a comprehensive view of customer preference.

**Statistical User Categorization Technique:**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer code

Description automatically generated**

**What the Function Does:**

This function calculates the mean, median, and standard deviation of the 'Credit\_Limit' variable from the given dataset. This function takes a dataset as an argument and returns a list containing these descriptive statistics.

1. **Mean Credit Limit:**

This provides the average credit limit across all customers. It indicates the central tendency of the credit limits in the dataset.

1. **Median Credit Limit:**

This represents the middle value of the credit limits when ordered from smallest to largest. It is less sensitive to extreme values and provides a robust measure of central tendency.

1. **Standard Deviation Credit Limit:**

This measures the amount of variation or dispersion in the credit limits. A higher standard deviation suggests greater variability in credit limits among customers.

In the context of the business analytics project, understanding the central tendency and variability of credit limits can help in identifying trends, potential outliers, and patterns that may be relevant to customer behavior, attrition, or financial decision-making.

**Where This Function Can Be Used:**

**1. Benchmarking and Overview**

The mean credit limit provides a benchmark or average value, offering a quickoverview of the typical credit limit customers have.

**2. Identifying Central Tendency**

The median credit limit gives an indication of the central tendency of credit limits, which is particularly useful when dealing with skewed distributions or outliers.

**3. Variability and Risk Assessment**

The standard deviation of credit limits reveals the extent of variability. A higher standard deviation may indicate a wider range of credit limits, suggesting greater diversity among customers.

**4. Segmentation Insights**

Analyzing these statistics across different customer segments (e.g., by income category, card category) can reveal patterns and differences in credit limits, aiding in targeted marketing or product development.

**5. Attrition Risk Assessment**

Comparing the credit limits of customers who have churned (attrition) versus those who haven't can provide insights into whether credit limits play a role in customer retention.

**6. Customer Experience and Satisfaction**

Understanding the distribution of credit limits contributes to assessing customer satisfaction and experience. For instance, if a significant proportion of customers has low credit limits, it might impact their satisfaction and financial flexibility.

**7. Financial Planning and Product Development**

The analysis of credit limits can inform financial planning and product development strategies. For example, it may guide decisions on introducing new card categories or adjusting credit limit offerings.

**Acknowledgment of Limitations in the Analysis:**

* **Data Representativeness**

The dataset may not fully represent the broader population due to potential biases in sample selection or data collection methods. As such, findings might not be universally applicable across different demographics or geographic regions.

* **Static Time Frame**

The analysis is based on a static dataset, which limits the ability to observe changes over time or capture dynamic trends in customer behavior and credit usage.

* **Lack of External Factors:**

The analysis primarily focuses on internal data points and does not account for external economic or social factors that could significantly impact credit behavior, such as market fluctuations or regulatory change.

**2. Suggestions for Future Research or Improvements:**

* **Enhanced Predictive Modeling**

Explore the use of advanced predictive modeling techniques beyond basic logistic regression, such as random forests, gradient boosting, or neural networks. Evaluate the performance of these models in predicting customer attrition and financial behavior.

* **Customer Segmentation Refinement**

Further refine customer segmentation by considering additional variables or combining existing ones. This can lead to more granular insights into specific customer groups and their distinct attrition patterns.

* **Behavioral Analysis Over Time**

Conduct a longitudinal analysis to understand how customer behaviors, such as transaction counts and credit card utilization, evolve over time. This can reveal trends, seasonality, or changes in financial behavior patterns.

* **Feature Engineering**

Experiment with feature engineering techniques to create new variables that might better capture customer behavior. For example, derive new features based on transaction frequency, clustering techniques, or interaction patterns with different card categories.

* **Social and Economic External Factors**

Integrate external datasets related to social and economic factors, such as unemployment rates, interest rates, or economic indicators. Analyzing the correlation between these external factors and customer attrition can provide a broader context for decision-making.

* **Customer Engagement Metrics**

Explore the inclusion of additional customer engagement metrics beyond transaction counts, such as online interactions, customer service calls, or participation in loyalty programs. Understanding the holistic customer journey can offer insights into attrition drivers.

* **Sentiment Analysis**

Incorporate sentiment analysis on customer feedback, reviews, or comments to gauge customer satisfaction and identify potential triggers for attrition. This qualitative data can complement quantitative analyses and provide a more comprehensive view.

* **Temporal Anomalies Detection**

Implement anomaly detection techniques to identify unusual temporal patterns in customer behavior. This can help in flagging potential fraud, irregularities, or significant changes in financial behavior that may contribute to attrition.

* **Dynamic Dashboards with What-If Analysis**

Develop dynamic dashboards in Tableau that allow for what-if analysis. Users should be able to manipulate key variables (e.g., credit limit, transaction frequency) and observe the projected impact on customer attrition in real-time.

* **Customer Retention Strategies Analysis**

Investigate the effectiveness of various customer retention strategies. Analyze historical data to understand which strategies have been more successful in mitigating attrition and identify areas for improvement.

* **Privacy-Preserving Techniques**

Explore privacy-preserving techniques, such as differential privacy or secure multiparty computation, to protect sensitive customer information while still allowing for meaningful analysis.

* **Geospatial Analysis**

If applicable, consider incorporating geospatial data to explore regional variations in customer behavior and attrition patterns. This can be particularly relevant for understanding the impact of local economic conditions.

* **Cross-Validation and Model Interpretability**

Implement robust cross-validation techniques to ensure model generalizability and interpretability. Understanding the factors contributing to model predictions is crucial for building trust in the analytical results.

* **User Feedback and Collaboration**

Actively seek feedback from end-users and stakeholders throughout the analysis process. Collaborate with business experts to validate findings, identify additional variables of interest, and ensure the alignment of insights with business objectives.

By addressing these suggestions, future research and improvements can elevate the depth and applicability of the business analytics project, leading to more actionable insights and informed decision-making.