**Executive Summary**

**Credit Card Attrition**

In the financial world, understanding the relationship between credit card utilization and customer attrition is crucial for maintaining a healthy customer base and financial stability. We want to understand if credit card utilization, typically defined as the ratio of a customer’s current credit card balance to their credit limit, can be a telling indicator of financial behavior.

**Problem Statement**

Are users with a higher credit card utilization ratio more likely to churn (i.e., discontinue their services)?

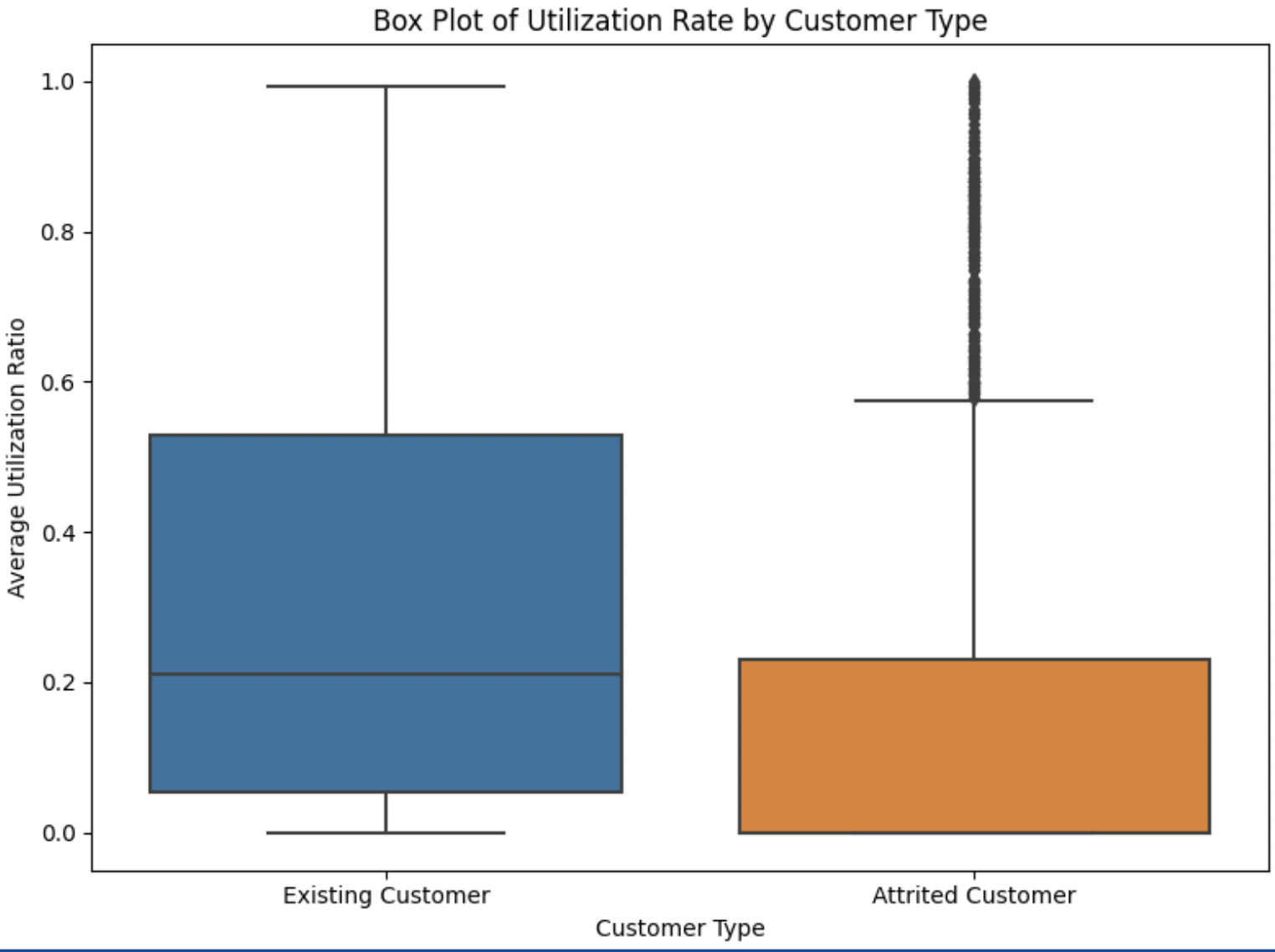
To address this, we formulated a null hypothesis () stating that the utilization ratio is not greater in attrited customers compared to existing customers. Conversely, the alternative hypothesis () posits that the utilization ratio is greater in attrited customers compared to existing customers.

Our hypothesis is rooted in the assumption that the reason for high credit card utilization is negligent spending, which could lead to a user being less likely to be able to pay back credit card fees. Over time, such financial strain can lead banks to take action, potentially freezing or closing accounts after extended periods of poor credit practices. Therefore, a higher utilization ratio might be an early indicator of future attrition, making this study not only relevant but essential for strategic financial management and customer retention strategies.

**Our Dataset**

The dataset we chose, titled ‘Data on Credit Card Customer Segmentation' from Kaggle, offers a comprehensive overview of customer behaviors and attributes in the context of credit card usage. It was compiled by aggregating customer-level data over two consecutive quarters and includes 23 columns that feature both quantitative and qualitative data. Key study variables encompass the 'average utilization ratio', ‘Attrition\_fla’, indicating whether a customer is still active, and additional factors such as age, income, marital status, education, and credit limit among others, providing a comprehensive view of customer profiles.

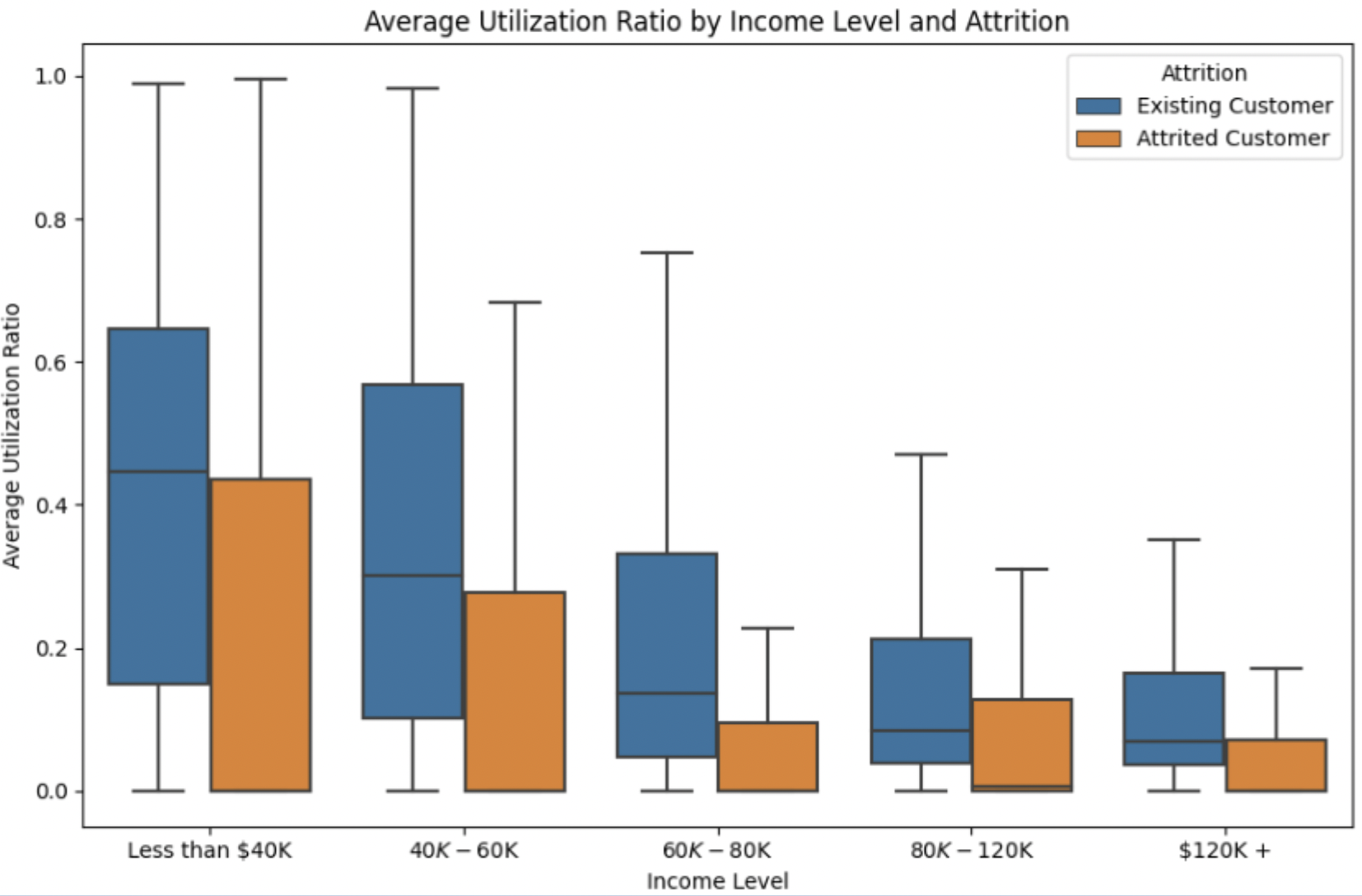
**Confounding Variables**

In analyzing the relationship between credit card utilization and customer attrition, it’s vital to consider key confounding variables that may skew our findings. Income level is a prime factor, as it influences credit limits, and financial stability, affecting both utilization ratios and attrition rates. Credit limit determines utilization ratios, and those with higher limits may stay with the bank. Additionally, education level, which could be correlated with financial literacy and income, can influence banking habits and preferences. Credit usage and customer attrition might be influenced by these underlying factors, rather than a direct effect of the utilization ratio alone. Neglecting these confounders may lead to inaccurate conclusions, so investigating their importance in our inclusion in our analysis can help us better understand the dynamics of customer retention and churn.

**Our Data Analysis and Visualizations**

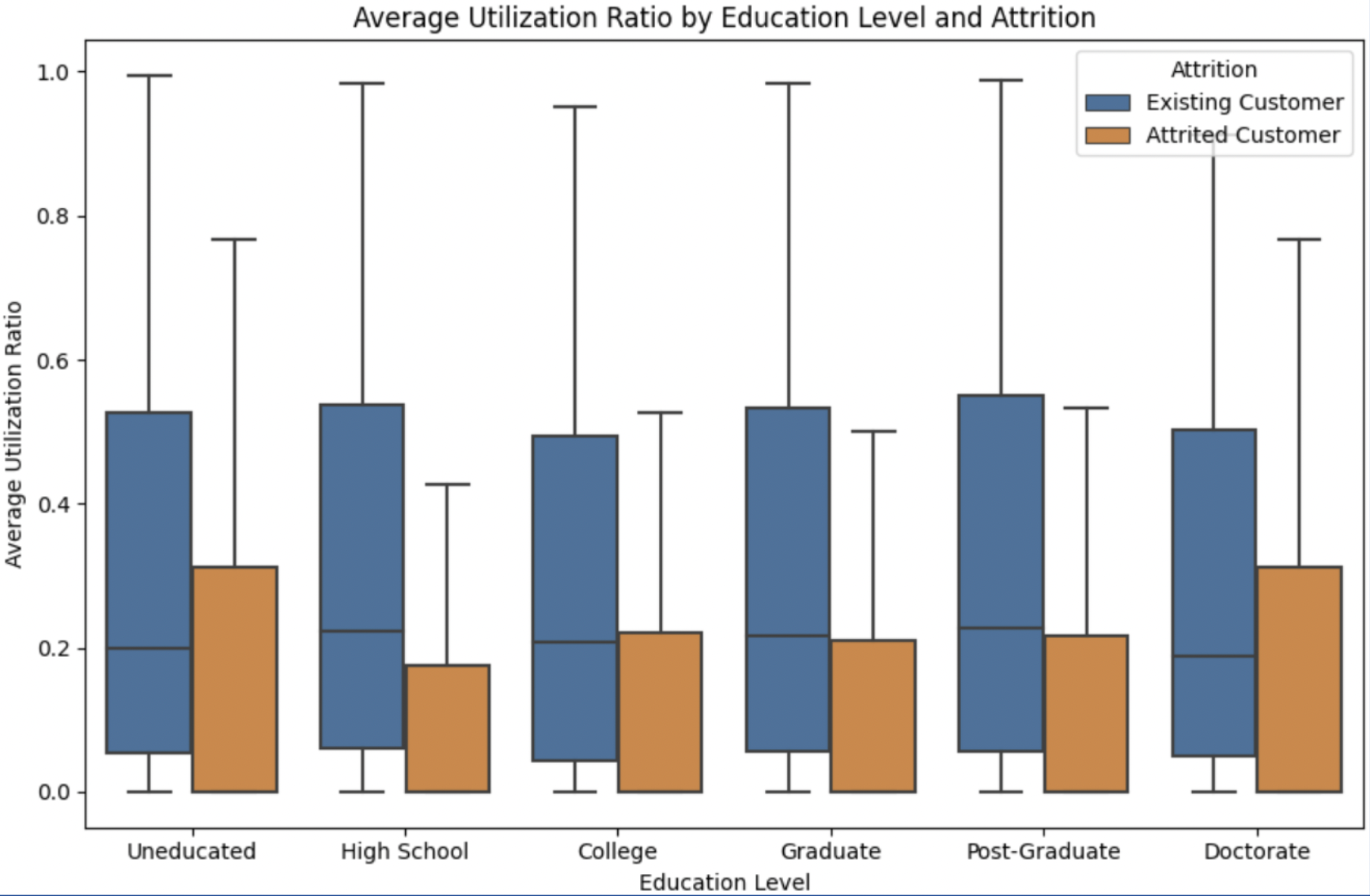
In the following section, we analyze the existing and attrited customer segments with our confounding variables: income, education, and credit limit.

***Customer Type and Utilization Boxplot:***

Existing customers exhibit a notably higher average utilization ratio compared to attrited customers. The spread of utilization is more pronounced among existing customers than attrited customers.

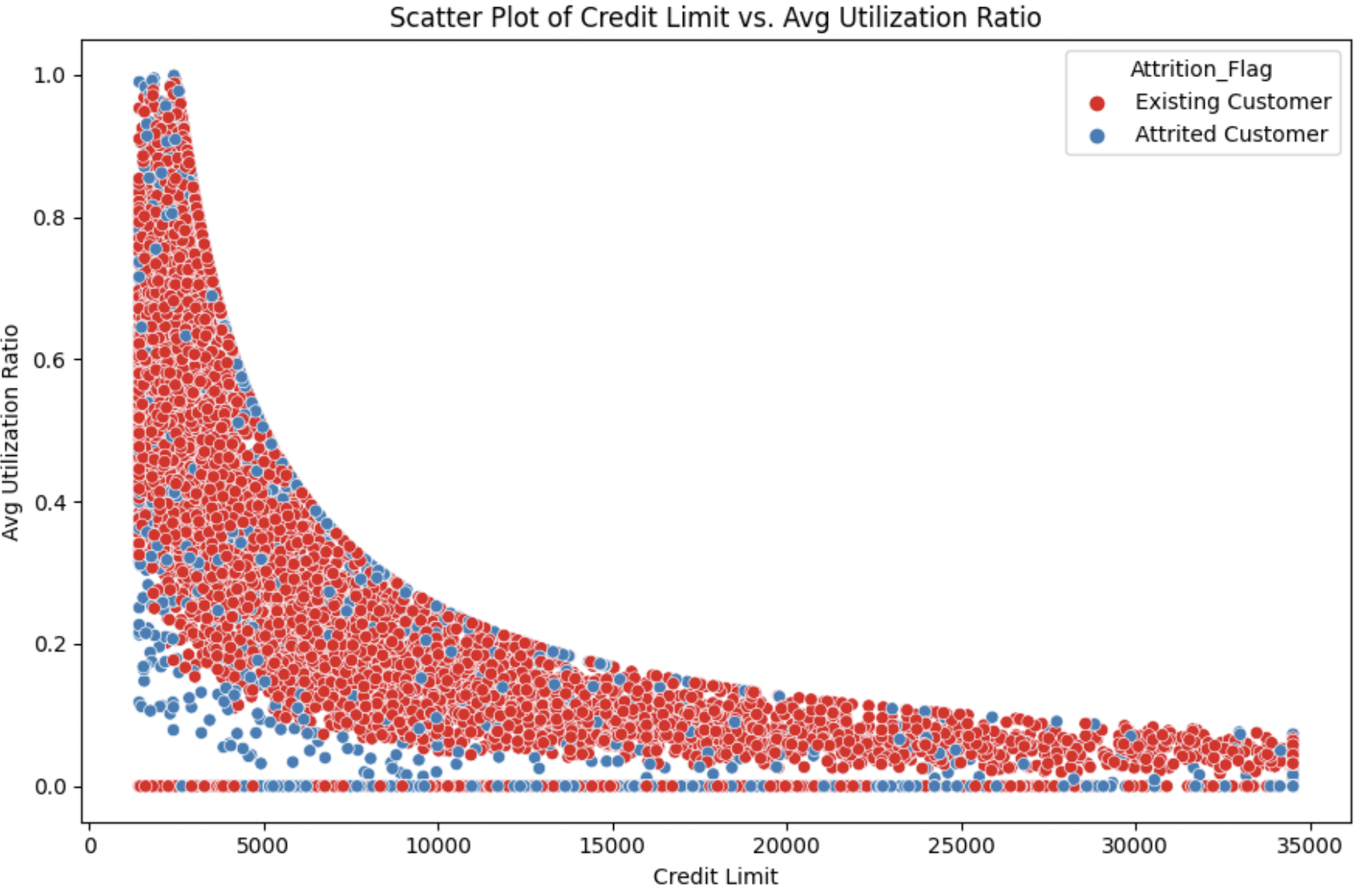
***Income Level and Utilization Boxplot:***

Investigating the correlation between utilization and income levels by attrition status, there is a significant dispersion of average utilization among individuals with lower income levels.

Across all income levels, existing customers consistently demonstrate a higher spread in utilization compared to their attrited counterparts.

***Education Level and Utilization Boxplot:***

While no significant differences are observed in average utilization among different education levels, existing customers maintain a higher utilization ratio than attrited customers at each education level.

***Credit Limit and Utilization Scatterplot:***

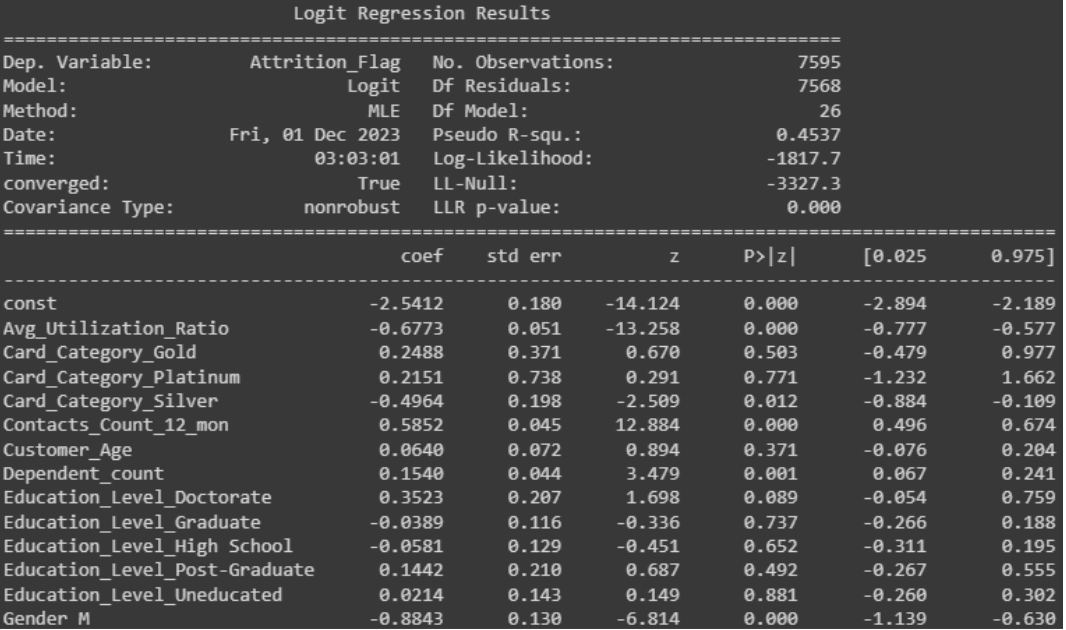
Examining the relationship between credit limit and average utilization ratio reveals that attrited customers are concentrated in lower credit limit and utilization ratio ranges. A notable presence of attrited customers with a 0 utilization ratio is also observed. These findings differ from the broader and more scattered distribution observed among existing customers across both variables.

**Ideal Experiment**

Acquiring credit card data poses significant challenges, and ethically, it is inappropriate to encourage individuals to augment their spending for experimental purposes in the real world. In an ideal setting devoid of resources and real-world constraints, authentic credit card data could be obtained from a bank, ensuring anonymization as per the Payment Card Industry Data Security Standards. Employing a random assignment and double-blind methodology would mitigate biases by preventing contact between the two groups. The control group would maintain regular utilization, while the treatment group would undergo a defined increase in utilization for a specified duration at a predetermined rate. Data, including demographics and confounding variables data, would be collected from both groups. Churn rates could then be compared using significance tests. To ensure accuracy and minimize bias, the experiment should be repeated multiple times, and the average results should be scrutinized.

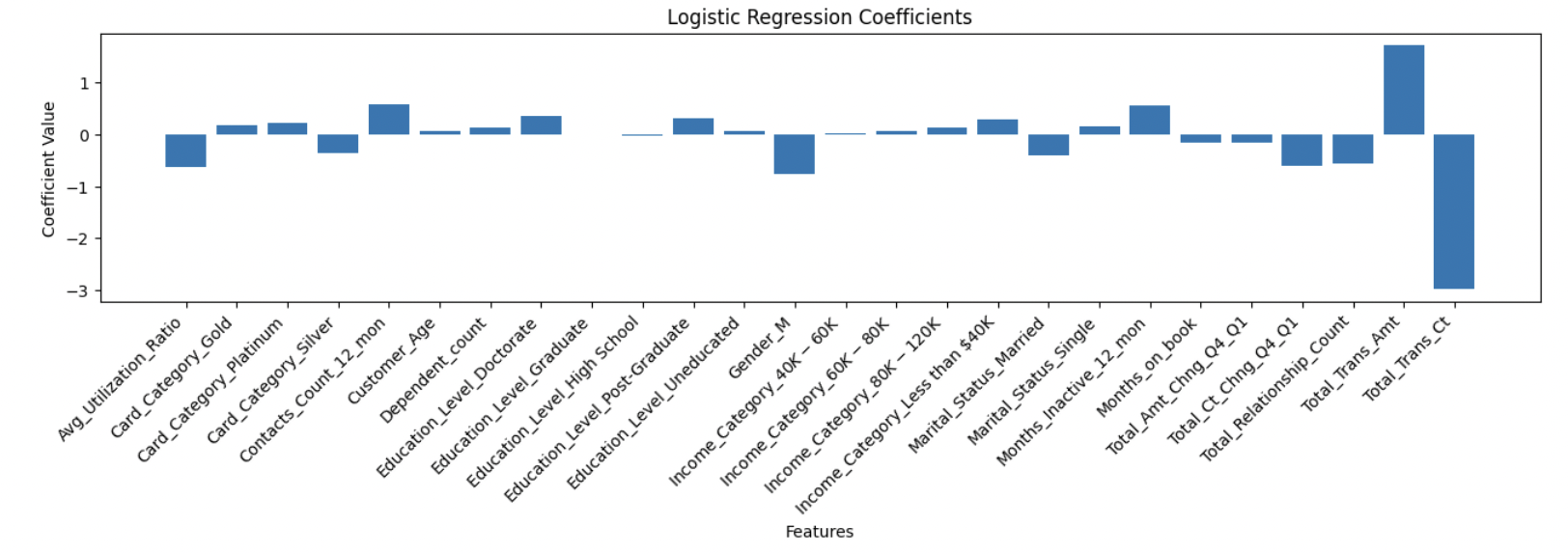
**Significance Testing Using T-Tests**

We conducted a one-tailed T-test to determine the accuracy of our initial hypothesis. Our hypothesis claimed that the utilization ratio would not be higher in attrited customers compared to existing customers. We assessed this by comparing the means of the two groups. Our alternative hypothesis asserted that the utilization ratio would indeed be higher in attrited customers. However, upon testing, we failed to reject the null hypothesis at various Alpha levels (1%, 5%, and 10%, suggesting that we could not achieve statistical significance in supporting this claim. On the other hand, we decided to reverse the test to examine whether the utilization ratio would be lower in shared customers compared to existing customers. In this case, we were able to attain statistical significance and reject the null hypothesis. This implies that our initial hypothesis of higher credit utilization being linked to negligent spending, where people exceed their spending limits and struggle to repay their fees, was simply false. Instead, the loss of customers is more likely attributed to inactivity or low account usage.

**Logistic Regression** 

Once we understood the direction of statistical significance that the average utilization ratio had with customer attrition, we wanted to test the variables' usefulness in predicting customer attrition, assuming all other variables were kept constant. We would need to perform a logistic regression using various parameters as logistic regression predicts a binary classification variable. We would be predicting the log odds for a one-unit change in the corresponding independent variable, assuming all other variables are held constant. We used a variance inflation factor to identify and remove variables with high collinearity. We could have also tried other classification models like a decision tree.

The regression output showed us that Avg\_Utilization\_Ratio was statistically significant (p-value = ~ 0) and it had the 4th largest coefficient impact on the prediction of the log-odds of a customer attriting. For each one-unit increase in the average utilization ratio, the odds of customer attrition decrease by approximately 49.2%.



**Conclusion and Recommendations**

In summary, our variable relationship analysis reveals that a lower utilization ratio significantly correlates with an elevated chance of churning. It is the fourth most influential factor in predicting customer attrition within this bank's dataset. Leveraging the low utilization ratio as a predictive indicator enables the identification of at-risk customers based on their business objectives, allowing targeted marketing efforts for retention.

Our primary recommendation for mitigating churn involves raising the minimum card utilization limit and establishing a higher baseline average payment requirement for each customer. Furthermore, incentivizing low-utilizing customers with attractive schemes, such as 12 to 18 months of interest-free periods or zero-fee balance transfers, can contribute to customer retention. Notably, Cash Back Rewards emerge as a particularly impactful strategy, eliciting a more pronounced short-term behavioral impact on customers. To enhance the effectiveness of these recommendations, we propose additional measures, including personalized communication campaigns and proactive customer engagement initiatives, tailored to individual usage patterns. Implementing these strategies collectively will likely yield a comprehensive approach to address and reduce low credit card utilization, ultimately fostering customer loyalty and minimizing churn.