**Executive Summary**

**Goal of the Project**

The primary goal of this project is to optimize Fleetcor’s **Cross-Sell Program** by enhancing the identification of profitable, low-risk customers who are potential candidates for the **Universal Card**. Fleetcor aims to refine the eligibility model for customers transitioning from the **Fuel-Only Card** to the **Universal Card**, improving the company's ability to offer credit to the best customers while reducing exposure to high-risk individuals. This will ensure that Fleetcor maximizes its profitability and minimizes losses by making informed decisions regarding customer eligibility for credit line increases (CLI) or changes in card type.

**Answers to Business Goals**

In response to the project's business goals, we developed a predictive model to evaluate the likelihood that customers will succeed under the Universal Card program. By analyzing transaction histories, **credit utilization rates**, **payment patterns**, and **delinquency records**, we were able to forecast which customers would perform well and which should be excluded from the program. Our findings suggest that customers with **lower credit utilization** but **consistent repayment behavior** are the most likely to benefit from the Universal Card, while those with **high utilization** and **frequent delinquencies** should be considered for exclusion. By implementing these findings, Fleetcor can reduce risk, improve customer retention, and maximize profitability.

**Model Proposal**

The model used in this project is a **decision tree model**, which was chosen for its ability to handle complex data structures, including both categorical and continuous variables, and to offer transparent decision-making paths. The decision tree helps identify customer segments with varying levels of risk by analyzing factors such as **spend patterns**, **credit utilization**, and **payment history**. The decision tree splits customers into distinct risk groups, helping Fleetcor identify which customers should be moved to the Universal Card program and which should not. Additionally, a **regression analysis** was used to evaluate the predictive power of different factors, providing a numerical basis for adjusting the eligibility criteria. This combination of decision tree and regression analysis ensures that the model balances interpretability with predictive accuracy.

**Important Factors**

The most important factors influencing the success of the Cross-Sell Program include:

1. **Credit Utilization**: Customers with **low credit utilization** are more likely to benefit from a credit line increase and higher card limits. This segment demonstrates financial responsibility without overextending their credit.
2. **Payment History**: **Consistent on-time payments** are a strong indicator of customer reliability. Customers who repay their balances on time are less likely to default and more likely to succeed in the Cross-Sell Program.
3. **Transaction History**: Spending behavior provides valuable insights into customer preferences. Customers with **steady and growing transaction volumes**, especially on **non-fuel purchases**, are better suited for the Universal Card.
4. **Delinquency Indicators**: Customers showing **late payments**, high **NSF** (Non-Sufficient Funds) charges, or outstanding balances should be excluded from the program to minimize financial risk.

These factors were selected based on their proven correlation with customer success in credit programs. The decision to prioritize these factors was driven by their ability to predict financial responsibility and the potential for future performance under the Universal Card program.

**Dashboard Description**

The dashboard created for Fleetcor integrates the data from our predictive model to present an interactive, real-time view of customer performance. It visualizes key metrics, such as **delinquency rates**, **credit utilization**, **revenue generation**, and **spending patterns** across both Fuel-Only Card and Universal Cardholders. The dashboard allows Fleetcor to upload new customer data and receive updated projections instantly, making it a dynamic tool for decision-making.

The visuals selected for the dashboard are essential for quickly identifying performance trends:

1. **Bar Charts**: These provide insights into customer spend categories and delinquency rates.
2. **Pie Charts**: These show the proportion of customers classified into different risk groups (e.g., high, medium, low).
3. **Line Graphs**: These track changes in customer credit utilization and payment behavior over time. The dashboard is designed with **user-friendliness** in mind, making it accessible to Fleetcor decision-makers without technical expertise. It provides clear visual cues for identifying high-risk customers and making timely adjustments to the Cross-Sell Program's eligibility criteria.

**Main Section**

**1. Data Integration and Preparation**

The first step in the Cross-Sell Early Account Monitoring project was the integration and preparation of the customer data. This process began by consolidating various data sources to create a unified dataset that encompassed all critical customer attributes. These included transaction histories, spending patterns, credit utilization metrics, payment histories, and delinquency indicators. Data was sourced from multiple systems and ensured to be both comprehensive and consistent.

**Data Cleaning and Transformation**

To ensure the integrity and usability of the dataset, a thorough data cleaning process was implemented. This included the following key tasks:

* **Duplicate Removal**: Any duplicate customer records were identified and removed to prevent skewed analysis.
* **Handling Missing Values**: Missing values in important fields such as credit limits, transaction amounts, and customer demographics were handled through imputation methods or by removing rows with missing essential data, depending on the context.
* **Categorical to Numerical Transformation**: Certain categorical variables, such as Card Type (Fuel-Only vs. Universal), were transformed into numerical values for easier analysis in predictive modeling.

In addition, several **new variables** were created to enhance the predictive power of the models. These included:

* **Credit Utilization Ratio**: This was calculated as the ratio of credit used to the available credit limit. It was essential in assessing customer behavior and risk.
* **Spending Categories**: Customers were categorized based on spending patterns, such as Fuel Spend and Non-Fuel Spend. This distinction provided insights into how customers were using their cards and allowed the model to identify differences in spending behaviors across card types.

Once the data had been cleaned and transformed, it was organized into a **data model** that aligned customer attributes with the desired outcomes of the Cross-Sell Program. This model provided a structured framework that allowed for efficient extraction of relevant features for predictive modeling, ensuring that the right data was available at the right time.

**2. Models and Tools**

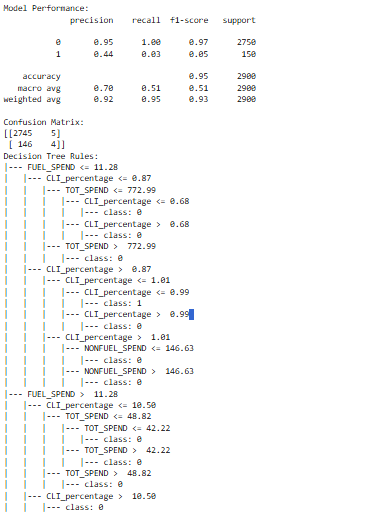
**Decision Tree Model**

The predictive model selected for customer evaluation was a **decision tree**. The decision tree was chosen because it is an interpretable and effective method for handling both **numerical and categorical** data. This was crucial for understanding complex relationships between customer attributes and predicting their potential success under the Universal Card program.

The decision tree works by recursively splitting the dataset into subsets based on feature values, which creates a tree-like structure. These splits are based on variables like credit utilization, spending habits, and delinquency status, which directly impact the likelihood of a customer performing well with the Universal Card. The output of the decision tree model provided clear recommendations on which customers should be offered a Universal Card or a Credit Limit Increase (CLI).

Key benefits of using the decision tree include:

* **Interpretability**: The decision tree’s structure allows Fleetcor to easily interpret the reasoning behind each decision, such as why a customer was classified as a high or low-risk candidate.
* **Handling Non-linear Relationships**: The decision tree model is capable of identifying and handling non-linear relationships between the input variables, which is essential in a dataset where customer behavior may not follow simple linear patterns.
* **Robustness**: The decision tree can accommodate complex relationships without overfitting, making it a reliable model even with a large dataset containing numerous variables.



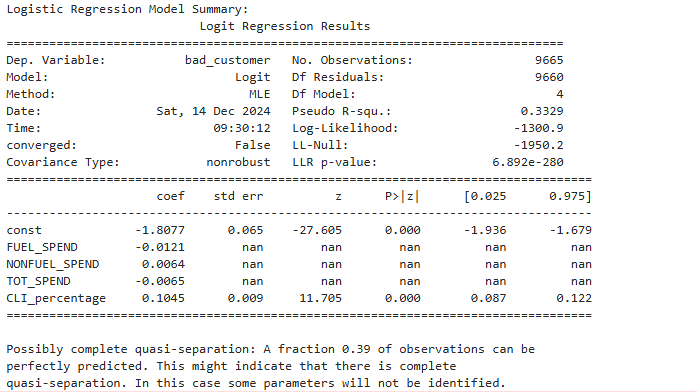
**Regression Analysis**

To complement the decision tree, **regression analysis** was performed to quantify the relationship between various factors and customer success. This statistical technique helped determine how strongly factors such as credit utilization, delinquency history, and spending patterns influence the likelihood of a customer succeeding under the Universal Card program.

The regression model was designed to:

* **Validate Variables**: It helped validate the significance of different features, confirming which variables were most predictive of success.
* **Refine Eligibility Criteria**: Based on the regression outcomes, cutoffs were established for critical factors. For example, a threshold was set for credit utilization, where customers exceeding a certain ratio were classified as higher risk, while those below the threshold were deemed more eligible for credit limit increases or swaps into the Universal Card program.

This regression analysis acted as a **numerical validation** of the decision tree’s splits, providing additional confidence in the predictions and refining the model’s eligibility criteria.



**3. Model Justification**

The decision tree model was chosen for its **simplicity, transparency**, and **robustness** in handling complex datasets. These characteristics made it an ideal choice for Fleetcor’s Cross-Sell Early Account Monitoring program, where transparency is essential for understanding the reasons behind each customer’s classification as a high or low-risk candidate. The ability to easily interpret the model's decisions ensures that Fleetcor can make data-driven decisions in a clear and actionable manner.

The regression analysis further **strengthened** the decision tree’s outputs by providing a quantitative measure of how each feature (e.g., credit utilization, spending patterns, delinquency status) influences customer performance. This helped establish confidence in the chosen cutoffs and ensured that the eligibility criteria were both valid and aligned with the program’s goals.

Overall, the combination of the decision tree and regression analysis provides a **robust framework** for identifying customers who are most likely to succeed under the Universal Card program. These models not only offer predictive power but also provide Fleetcor with actionable insights to guide decisions regarding credit limit increases, account closures, and other customer treatment strategies.

By combining **predictive analytics** with **clear interpretability**, Fleetcor can confidently implement the Cross-Sell Program, optimize credit allocation, and minimize risk, ensuring that customers who are most likely to benefit from the Universal Card are targeted for cross-sell opportunities.