

Credit Card Case Study

Exploratory Data Analysis for Credit Card Data

Analytics in Credit Card Industry:

Analytics has penetrated every industry owing to the various technology platforms that collect information, allowing service providers to understand exactly what customers want. The Credit Card industry is no exception. Within credit card payment processing, there is a significant amount of data available that can be beneficial in countless ways.

Understanding the customer behavior: The data available from a credit card processor identifies the types of consumers and their business spending behaviors. This helps in developing marketing campaigns to directly address these behaviors, resulting in greater sales.

Personalize offering based on data results: Data reveals specific interests and needs in individual customers that a company can leverage, addressing their needs more efficiently. Specific promotions can be sent out related to where these customers are located, quickly building sales.

Use trends and patterns to get new customers: The transactions and activities of existing customers reflect larger trends. This information provides a strategy to target potential customers.

Uncover suspicious activity: Data from credit card processing is crucial for fighting fraud. When combined with artificial intelligence, this data can be analyzed quickly to uncover suspicious purchase activity.

Reduce chargebacks: Detecting suspicious activity and patterns in data can assess whether a transaction might result in a chargeback. Using analytics to track each

transaction can reveal anomalies, helping to reject suspicious transactions and save the business from chargebacks.

Business Problem:

To produce quality decisions in the modern credit card industry, knowledge must be gained through effective data analysis and modeling. Using dynamic data-driven decision-making tools and procedures, information can be gathered to evaluate all aspects of credit card operations. For example, a bank with operations in more than 50 countries needs to evaluate areas of bankruptcy, fraud, and collections, and respond to customer requests with proactive offers and services.

Data Available:

The dataset includes the following sheets:

- **Customer Acquisition:** Details of customers at the time of card issuing.
- **Spend (Transaction data):** Credit card spend for each customer.
- **Repayment:** Credit card payments done by customers.

Questions to address customer spend and repayment behavior:

Q1. Data Cleaning:

- If age is less than 18, replace it with the mean of age values.
- If age is less than 18, replace it with the mean of age values.
- If the spend amount is more than the limit, replace it with 50% of that customer's limit.
- If the repayment amount is more than the limit, replace it with the limit.

Q2. Summary Statistics:

- How many distinct customers exist?
- How many distinct categories exist?

- What is the average monthly spend by customers?
- What is the average monthly repayment by customers?
- If the monthly rate of interest is 2.9%, what is the profit for the bank for each month? (Profit is defined as interest earned on monthly profit. Monthly profit = Monthly repayment – Monthly spend. Interest is earned only on positive profits and not on negative amounts.)
- What are the top 5 product types?
- Which city has the maximum spend?
- Which age group is spending more money?
- Who are the top 10 customers in terms of repayment?

Q3. City-wise Analysis:

- Calculate the city-wise spend on each product on a yearly basis and include graphical representations.

Q4. Graphs:

- Monthly comparison of total spends, city-wise.
- Comparison of yearly spend on air tickets.
- Comparison of monthly spend for each product, identifying any seasonality in spending.

Q5. Custom Python Function:

- Write a user-defined Python function to find the top 10 customers for each city in terms of repayment amount by different products and time periods (yearly or monthly). The function should allow specification of the product (Gold/Silver/Platinum) and the time period for analysis.