Data Exploratory Analysis for Credit Card Data

Using R Programming

BUSINESS PROBLEM:

In order to effectively produce quality decisions in the modern credit card industry, knowledge must be gained through effective data analysis and modeling. Through the use of dynamic data-driven decision-making tools and procedures, information can be gathered to successfully evaluate all aspects of credit card operations.

PSPD Bank has banking operations in more than 50 countries across the globe.

Mr. Jim Watson, CEO, wants to evaluate areas of bankruptcy, fraud, and collections, respond to customer requests for help with proactive offers and service.

DATA AVAILABLE -

This book has the following sheets:

- Customer Acquisition: At the time of card issuing, company maintains the details of customers.
- Spend (Transaction data): Credit card spend for each customer
- Repayment: Credit card Payment done by customer

Following are some of Watson's questions to a Consultant (like you) to understand the customers spend & repayment behavior.

1. In the above dataset,
   a. Incase age is less than 18, replace it with mean of age values.
   b. Incase spend amount is more than the limit, replace it with 50% of that customer's limit. (customer's limit provided in acquisition table is the per transaction limit on his card)
   c. Incase the repayment amount is more than the limit, replace the repayment with the limit.

2. From the above dataset create the following summaries:
   a. How many distinct customers exist?
   b. How many distinct categories exist?
   c. What is the average monthly spend by customers?
   d. What is the average monthly repayment by customers?
   e. 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)
   f. What are the top 5 product types?
   g. Which city is having maximum spend?
   h. Which age group is spending more money?
   i. Who are the top 10 customers in terms of repayment?

3. Calculate the city wise spend on each product on yearly basis. Also include a graphical representation for the same.

4. Create graphs for
   a. Monthly comparison of total spends, city wise
   b. Comparison of yearly spend on air tickets
   c. Comparison of monthly spend for each product (look for any seasonality that exists in terms of spend)

5. Write user defined R function to perform the following analysis:

You need to find top 10 customers for each city in terms of their repayment amount by different products and by different time periods i.e. year or month. The user should be able to specify the product (Gold/Silver/Platinum) and time period (yearly or monthly) and the function should automatically take these inputs while identifying the top 10 customers.