BCIS 5110

Programming Languages for Business Analytics

Project 1

Project 1 includes 3 steps as follows. For each step follow the description and requirements of the step and provide your answers/results in the report.

- The report should not be longer than 10 pages.
- Make sure to document every step you take and explain what you have done.

Objectives & Deliverables

In this project you will design an algorithm that can simulate (generate) a cohort of 20,000 customers along with their accounts' information and activity for 12 months.

Given the information in the next steps, design an algorithm that does the followings:

1. Generates 20,000 customers using the given characteristics.
2. Generates accounts information and activity for the customers using the information given.
3. Stores customers' and accounts information and activity in the right format for further analysis.

In your report, you need to provide the following information:

1. Description of the project/problem (less than a page) **[2 points]**
2. How you plan to solve the problem and your design of the algorithm **[3 points]**
   - Describe how you plan to approach the problem. Explain about your design of the algorithm, how you are breaking down the problem into smaller parts, and how you approach each one.
3. Your Algorithm represented by a Flowchart **[5 points]**
   - Note that your Flowchart needs to be self-explanatory, so please add comments to explain how it works.
4. Your Algorithm represented by a Pseudocode **[5 points]**
   - Note that your Pseudocode needs to be self-explanatory, so please add comments to explain how it works.

Please follow these steps when you are writing your report:

- Make sure to document every step in your report.
- Keep your report organized and clean.
- Use a software to draw the Flowchart.
- Make sure your Pseudocode looks clean and organized.

Please note that our algorithm will be tested later in the second project to see if it functions as described or not. Points will be deducted if your algorithm does not function as described in this project.

Please note that the performance of your algorithm matters. You need to keep track of the memory and time when it comes to implementing your algorithm in the next project.

Introduction

The FarziBank (FB) provides financial services to its customers including providing credit cards. Customers can have one or more credit cards (accounts) with the bank each with its own credit line. FarziBank stores customers' credit card activity to analyze and understand their behavior and make better decisions.

Before you get into the main steps of this project, you need to get familiar with the terminology around the customers and accounts' information.

Customer Information	
Customer ID	An identification number uniquely assigned to each customer of the bank
Gender	Gender of a customer (Male, Female)
Age	Age of the customer
Marital Status	Marital status of the customer (Single, Married)
Number of Children	Number of children of the customer
Education Level	Education level of the customer (No education, High school, Bachelor's degree, Master's degree, Ph.D.)
Number of Accounts	Number of accounts (credit cards) the customer has with the bank
Annual Income	Annual income of the customer (in dollars)
Total Credit Line	Total available credit line from all the accounts customer has (in dollars)
Account Information	
Account Number	A number uniquely assigned to each customer's account
Date Opened	The date on which the account was opened
Credit Line	Total available credit line of the account
Annual Fee	An amount in dollars customer has to pay annually for the account
APR	Annual interest rate is a yearly rate in percentage that is applied to the balance
Account's Activity	
Purchase Amount	Amount of purchase customer uses the card to make (in dollars)
Cash Advance Amount	Amount of cash customer takes out from the card (in dollars)
Balance Amount	Amount of credit line customer has used (in dollars)
Payment Amount	Amount of debt that customer pay back (in dollars)
Payment Period	From 1st of each month to 10th of the month when customers should pay at least the Minimum Amount Due of the previous month
Minimum Amount Due	Amount of the balance that customer should pay in the payment period (before the due date)
Past Due Amount	Minimum Amount Due that customer was supposed to pay and is now past due
Interest Amount charged	Amount of interest in dollars charged to the customer
Late Payment Fee	Amount in dollars charged to the customer if they do not pay the Minimum due amount by the due date
Delinquency Counter	A number indicating how many consecutive months customer has missed the Minimum Due payment

Please note the following information:

- Each customer can own one or more credit cards (accounts) with the bank.
- Customers can use their cards (accounts) for purchasing or taking cash out.
- Every time customer purchases something, the amount of the purchase (Purchase Amount) will be added to the balance of their card (account).
- Every time customer takes cash out, the amount of the cash (Cash Advance Amount) will be added to the balance of their card (account).
- Balance of an account is basically the amount that customer owes the bank (because of the purchases or cash advances they made)
- Interest would be applied to the balance amount every day based on the interest rate.
- Customers are supposed to pay back their debts (balances).
- Every time customer makes a payment (to pay their debts) the amount (Payment Amount) will be deducted from the balance.
- At the end of each month (last day of the month) we have the closing balance which is the balance at the end of the last day of the month.
- Customers are supposed to pay at least the Minimum Amount Due (which is a proportion of their closing balance) by the due date.
- The due date for each closing balance is the end of the 10th day of the next month.
- If customers do not pay at least the Minimum Amount Due by the end of 10th day, the amount which is due will become past due (Past Due Amount), they will be charged a Late Payment Fee, and their Delinquency Counter will be incremented by 1.
- As soon as the customer pays their Past Due Amount, their Delinquency Counter will be reset to 0 along with their Past Due Amount.
- If customers miss their Minimum Amount Due payments for 3 consecutive months, their account will be closed by the bank. I.e., they cannot use it to make any purchases or cash advances.
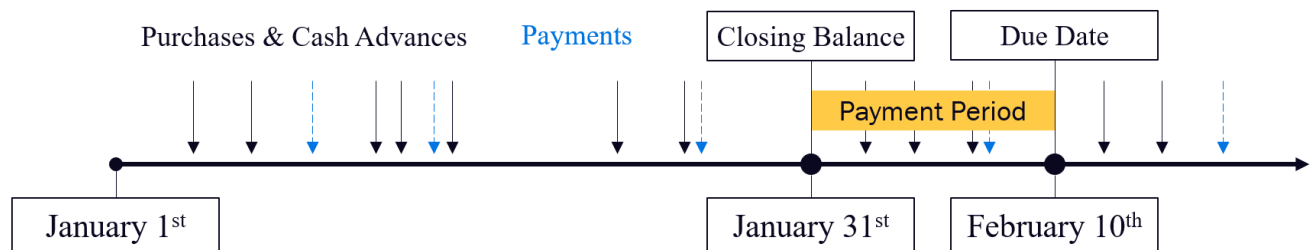
Note: for random number generation in the following steps of the project, use Uniform distribution unless it is explicitly mentioned that distribution you should use.

Step 1 – Generating Customers

You need to generate 20,000 customers with the following characteristics:

1. **Customer ID** should be a **7-digit** number randomly generated for each customer (unique).
   - *Example*: 6522489
1. **Age** of each customer should be a random number uniformly distributed in **[20, 80]**.
2. **Gender** of each customer should be randomly selected from '**Male**' or '**Female**' each with equal probability of **0.5**.
3. **Marital Status** for each customer should be randomly selected from either '**Single**' or '**Married**' using the following distribution:

Age	Single	Married
[20, 30]	0.75	0.25
(30, 60]	0.25	0.75
(60, 80]	0.5	0.5

If you ever need to convert Marital Status to numeric, 'Single' can be represented by 0 and 'Married' by 1.

4. **Number of Children** for each customer should be a random integer number in **[0, 4]** selected based on the following probabilities:

Age	0	1	2	3	4
[20, 40]	0.4	0.3	0.2	0.1	0
(40, 80]	0.1	0.3	0.3	0.2	0.1

5. **Education Level** for each customer should be randomly selected from the following options using the given probabilities:

Age	No Education	High School	Bachelor's Degree	Master's Degree	Ph.D.
[20, 25]	0.1	0.5	0.3	0.1	0
(25, 35]	0.1	0.5	0.3	0.05	0.05
(35, 80]	0.1	0.5	0.25	0.1	0.05

If you ever need to convert Education Level to numeric, use the following mapping:

   - 0 for 'No Education'
   - 1 for 'High School'
   - 2 for 'Bachelor's Degree'
   - 3 for 'Master's Degree'
   - 4 for 'Ph.D.'

6. **Annual Income** for each customer should be generated using the following formula:

$$\text{Annual Income} = 40 \times 52 \times (\$15 + EducationLevel \times \$10 + (Age/10) \times \$2)$$

   - *Example: for a customer who has a Bachelor's Degree and is 43 years old the annual salary would be* $= 40 \times 52 \times (\$15 + 2 \times \$10 + (43/10) \times \$2) = \$90,688$

7. **Number of Accounts** for each customer should be generated using the following formula:

$$\text{Number of Accounts} = MaritalStatus + NumberOfChildren + 1$$

   - *Example: a customer who is single and has 2 children, has* $= 0 + 2 + 1 = 3$ *accounts.*

8. **Total Credit Line** for each customer should be generated using the following formula:

$$\text{Total Credit Line} = NumberOfAccounts * (AnnualIncome/10)$$

   - *Example: for a customer with 3 accounts and an annual income of $90,688, Total Credit Line* $= 3 * (\$90,688/10) = \$27206.40$

Step 2 – Generating Accounts

For each customer, you need to generate their accounts' information based on the following characteristics:

1. **Date Opened** for each account should be a random date selected before **January 1st, 2022**.
   - Cannot be older than customer's age,
   - Should not be before customer became 20 years old,
2. **Account Number** for each account should be an **8-digit** number generated based on the following rule:
   - The first 7-digits are Customer ID.
   - Last digit should be a number in **[1, 9]** assigned to each account in the order the accounts were opened.
   - *Example: for a customer with Customer ID 6522489 who has 3 accounts with the following dates opened:*

Date Opened	Account ID
February 23rd 2011	*6522489**1***
June 13th 2014	*6522489**2***
September 2nd 2017	*6522489**3***

3. **Account Credit Line** for each account should be a random proportion of the customer's Total Credit Line (from Step 1.9)
   - *Example: for the customer with 3 accounts and Total Credit Line of $27206.40 each account's credit line could be:*

Account ID	Credit Line
*6522489**1***	*0.1 × $27206.40 = $2,720.64*
*6522489**2***	*0.3 × $27206.40 = $8,161.92*
*6522489**3***	*0.6 × $27206.40 = $16,323.84*

4. **Annual Fee** for each account should be generated using the following formula:

$$\text{Annual Fee} = AccountCreditLine \times 0.01$$

5. **Annual Interest Rate** for each account should be randomly selected in **[15%, 30%]**.

Step 4 – Generating Account Activity

For each account of the customer, you need to generate the account's activity for **12 months** (from January 1st 2022 to December 31st 2022) based on the following instructions.

1. Customers will use each card every **d** number of days where **d** is a random integer number in **[0, 7]**
   - Note that **d** is not fixed for every customer or account and should change every time.
2. Every time customer uses their card, they either make a purchase (with probability of **0.95**) or take cash out (with probability of **0.05**).
3. If a customer purchases something, the Purchase Amount would be a random number in **(0, Available Credit Line]**, where
   - Available Credit Line = Credit Line – Current Balance
4. If a customer takes out cash, the Cash Advance Amount would be a random number in **(0, Available Cash]**, where
   - Available Cash = 10% × Credit Line, if 10% × Credit Line ≤ Available Credit Line
   - Available Cash = Available Credit Line, if 10% × Credit Line > Available Credit Line

   In other words,

$$\text{Available Cash} = \min(\text{Available Credit Line}, 10\% \times \text{Credit Line})$$

5. As mentioned before, every time customer makes a purchase or cash advance, the amount would be added to their account's balance.
6. At the end of every day, the bank will apply interest to the remaining balance of the account by multiplying the balance by the daily interest rate, thus

$$\text{Current Balance} \leftarrow \text{Current Balance} \times (1 + \text{Daily Interest Rate})$$

where

$$\text{Daily Interest Rate} = \text{Annual Interest Rate (APR)} / 365$$

   - *Example: if Annual Interest Rate is 22% for the account, Daily Interest Rate would be 22%/365 = 0.06%, thus the balance should be multiplied by 1.0006 at the end of each day.*
7. Customers pay back their debts (balances) for each of their accounts based on the following scenarios:
   - 10% of customers pay the exact amount they spent every day on the same day (they pay their entire balance at the end of each day)
   - 30% of customers pay the entire balance every **d** number of days where **d** is random integer number in **[1, 7]**
   - 30% of customers pay only a proportion **p** of the entire balance every **d** number of days **p** is a random number in **[0, 1]** and **d** is a random integer number in **[1, 7]**
   - 30% of customers act in the Payment Period according to the following:
       i. 10% of these customers pay the entire balance within the Payment Period (the day should be randomly selected from **[1, 10]**)
       ii. 15% of these customers pay only the Minimum Amount Due within the Payment Period (the day should be randomly selected from **[1, 10]**)
       iii. 75% of these customers do not pay at all and miss the due date.
8. At the end of each month the following amounts will be calculated:
   - Closing Balance = current/remaining balance
   - Minimum Amount Due = 10% × Closing Balance
   - Total Purchases of the month = sum of all the purchases made during the month
   - Total Cash Advances of the month = sum of all the cash advances made during the month
   - Total Payments of the month = sum of all the payments made during the month
   - Total Interests charged in the months = sum of all interest amounts charged in the month
9. After the 10th of each month, if the customer has not paid at least the Minimum Amount Due of the previous month before the due date:
   - Past Due Amount ← Minimum Amount Due
   - Late Payment Fee of $30 will be added to the balance.
   - Delinquency Counter will be incremented by 1.
   - Note that Delinquency Counter and Past Due Amount will be reset to 0 as soon as the customer pays the Past Due Amount.
10. If the Delinquency Counter for an account becomes 3 (i.e., customer misses paying the Minimum Amount Due for 3 consecutive months)
   - The account will be closed by the bank.
   - Customer cannot use the card to make any purchases or cash advances.