MAR6668 Individual Assignment 1 (Due on Sept 20 before class)

This assignment must be answered on an individual basis without consulting with the classmates' write-ups. Please submit the write-up of all your answers in a word/PDF file. Please also include your Python codes in separate files as supplement materials.

1. A credit card company wants to divide its users into different segments. The dataset ("Q1 CreditCardBehavior.csv") contains the credit card usage and payment data of 1,500 consumers in the past six months. Below is the description of the variables.

   CUST_ID: Identification of Credit Card holder

   BALANCE: Account balance at the end of the six months

   PURCHASES: Amount of purchases made from account in the past six months

   PURCHASES_TRX: Number of purchase transactions made in the past six months

   CREDIT_LIMIT: Credit limit for user

   PAYMENTS: Amount of card payment made by user in the past six months

   PRC_FULL_PAYMENT: Percent of full payment paid by user in the past six months

   Use clustering analysis to divide the consumers into **four** segments.

   (Github link:

   https://raw.githubusercontent.com/zoutianxin1992/MarketingAnalyticsPython/main/Marketing%20Analytics%20in%20Python/2023/Assignment%201/2023Q1_CreditCardBehavior.csv)

   (1) Explain why data normalization is needed before running the clustering analysis.

   (2) How many consumers are there in each segment? In which segment did the segment's average consumer make the most purchase in dollars with the credit card in the past six months? Which segment *in total* made the most purchase in dollars with the credit card in the past six months?

   (3) How many dollars does an average consumer in each segment spend *per transaction*?

   (4) Describe in word the unique characteristics of each segment that distinguishes it from other segments. Can you give a name for each segment (e.g., online shoppers are labelled as image idolizers, review researchers , etc.)?

(5) Is dividing consumers into four segments a reasonable choice? Why?

2. Roblox is a popular online gaming platform, where users can create their own games and play those created by other peer users. Dataset "Q2-RobloxTotalUsers.csv" records the total number of Roblox users (in millions) since its launch in 2016. The data are bi-annual and have 11 periods.
   (Github link: https://raw.githubusercontent.com/zoutianxin1992/MarketingAnalyticsPython/main/Marketing%20Analytics%20in%20Python/2023/Assignment%201/2023Q2-RobloxTotalUsers.csv)

   (1) Use Bass model to predict the total number of Roblox users in second half of 2023. [**Hint:** Total number of users is A(t) in Bass model, but we need N(t) to estimate Bass model.]
   (2) Compare the values of p and q for your Bass model. What do you find? Explain how the product nature of Roblox may explain this finding.
   (3) Plot new adoptions of Roblox, N(t), as a function of time for this new product for 10 periods (5 years). Report the total number of adopters after five years.
   (4) When will the total adoptions of Roblox surpass 60% of the total market size?