Graded Assignment: Credit Card Customer Segmentation

This homework deals with the following topics:
- Data Preprocessing for Clustering
- K-Means Clustering
- Clustering Analysis and Interpretation

The Assignment

In this graded assignment, we will practice applying the K-Means clustering model to conduct and analyze customer segmentation. This assignment will cover loading datasets, data preprocessing, K-Means clustering, and clustering analysis.

To begin this assignment, we have provided you with an .ipynb file in Codio, which contains the instructions as inline comments. We will be using the Kaggle Credit Card dataset to complete this practice assignment. Details about the dataset are described in the following sections.

For each question, there are clear instructions in each cell. Follow those instructions and write the code after each block of:

```
# YOUR CODE HERE
raise NotImplementedError()
```

Make sure to delete the line raising an error! Please use the exact variable name if it is specified in the comment.

We'll run a Python test script against your program to test whether the code in each cell is correct.

Dataset Overview:

The Kaggle Credit Card Dataset summarizes the usage behavior of about 9000 active credit card holders with 18 customer level behavioral variables:
- **CUST_ID** : Identification of Credit Card holder
- **BALANCE** : Balance amount left in their account to make purchases
- **BALANCE_FREQUENCY** : How frequently the Balance is updated, score between 0 and 1 (1 = frequently updated, 0 = not frequently updated)
- **PURCHASES** : Amount of purchases made from account
- **ONEOFF_PURCHASES** : Maximum purchase amount done in one-go
- **INSTALLMENTS_PURCHASES** : Amount of purchase done in installment

- **CASH_ADVANCE** : Cash in advance given by the user
- **PURCHASES_FREQUENCY** : How frequently the Purchases are being made, score between 0 and 1 (1 = frequently purchased, 0 = not frequently purchased)
- **ONEOFFPURCHASESFREQUENCY** : How frequently Purchases are happening in one-go (1 = frequently purchased, 0 = not frequently purchased)
- **PURCHASESINSTALLMENTSFREQUENCY** : How frequently purchases in installments are being done (1 = frequently done, 0 = not frequently done)
- **CASHADVANCEFREQUENCY** : How frequently the cash in advance being paid
- **CASHADVANCETRX** : Number of Transactions made with "Cash in Advanced"
- **PURCHASES_TRX** : Number of purchase transactions made
- **CREDIT_LIMIT** : Limit of Credit Card for user
- **PAYMENTS** : Amount of Payment done by user
- **MINIMUM_PAYMENTS** : Minimum amount of payments made by user
- **PRCFULLPAYMENT** : Percent of full payment paid by user
- **TENURE** : Tenure of credit card service for user

We will be loading the dataset from the file: `data.csv`

What to submit

Before submitting, click "Validate" to check if your answers are right. When you are satisfied with your answers, in Codio go to the "Education" menu and select "Mark as Completed."