Job Oriented Specialisation in Data Science

Course 6: Clustering

Project: Segmenting customers into clusters

Perform Customer Segmentation on the transactional data to build an efficient marketing model. This is a transactional data set that contains all the transactions occurring between 01/12/2010 and 09/12/2011 for a registered non-store online retail.

Invoice No: Invoice number. Nominal, a 6-digit integral number uniquely assigned to each transaction. If this code starts with the letter 'c', it indicates a cancellation.

StockCode: Product (item) code. Nominal, a 5-digit integral number uniquely assigned to each distinct product.

Description: Product (item) name. Nominal.

Quantity: The quantities of each product (item) per transaction. Numeric.

Invoice Date: Invoice Date and time. Numeric, the day and time when each transaction was generated.

UnitPrice: Unit price. Numeric, Product price per unit in sterling.
CustomerID: Customer number. Nominal, a 5-digit integral number uniquely assigned to each customer.

Country: Country name. Nominal, the name of the country where each customer resides.

_____

Guidelines:
    a. The code should be submitted in the form of pdf format of the jupyter notebook. The option could be found in the "*File > Download as*" section in the notebook.
    b. Make sure the pdf format notebook contains all the relevant output of the code cells. Notebooks that have not been run for displaying the outputs will be marked as incomplete.
    c. The code should be well commented on and documented.
    d. All project files should be added in a folder with your name and then compressed. (Right-click on the folder> Send to > Compressed (zipped) folder) and give the file your name. Then you may upload this zipped file on the project submission page.