Capstone Project Submission

Team Member's Name, Email and Contribution:
Sanjay Yadav neer.ping@gmail.com **Contribution :** **1. Exploratory Data Analysis** **2. Data Preprocessing and Normalization** **3. Generating Customers RFM Score** **4. Scaling Data** **5. Optimal Number of Clusters** **6. Building Kmeans and Hierarchical Model**
Please paste the GitHub Repo link.
Github Link:- https://github.com/sanjay2097/Customer-Segmentation
Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words) **Problem Statement :** **Customer segmentation is the process of separating customers into groups on the basis of their shared behavior or other attributes. The groups should be homogeneous within themselves and should also be heterogeneous to each other. The overall aim of this process is to identify high-value customer base i.e. customers that have the highest growth potential or are the most profitable.** **Insights from customer segmentation are used to develop tailor-made marketing campaigns and for designing overall marketing strategy and planning.** **A key consideration for a company would be whether or not to segment its customers and how to do the process of segmentation. This would depend upon the company philosophy and the type of product or services it offers. The type of segmentation criterion followed would create a big difference in the way the business operates and formulates its strategy.**

Customer segmentation has a lot of potential benefits. It helps a company to develop an effective strategy for targeting its customers. This has a direct impact on the entire product development cycle, the budget management practices, and the plan for delivering targeted promotional content to customers. For example, a company can make a high-end product, a budget product, or a cheap alternative product, depending upon whether the product is intended for its most high yield customers, frequent purchasers or for the low-value customer segment. It may also fine-tune the features of the product for fulfilling the specific needs of its customers.

In this project, our task is to identify major customer segments on a transnational data set which contains all the transactions occurring between 01/12/2010 and 09/12/2011 for a UK-based and registered non-store online retail.The company mainly sells unique all-occasion gifts. Many customers of the company are wholesalers.

Approach :

In this project our main aim is to segment the retail customers into proper groups that are defined by their purchase patterns using recency, frequency and monetary value of individual customers.
Segmentation with K-means clustering and Hierarchical clustering : Initially, the data is subject to important stages in an analytics pipeline: exploratory analysis, preprocessing, normalization and standardization. Then, the unsupervised clustering algorithm is used to determine the ideal segments of customers.Various metrics are to generate optimal number of clusters and related cluster visualizations are leveraged to deduce the optimum value of "K" (number of clusters) in the algorithm. The observations from the results are elaborately discussed before reaching the conclusion from the business perspective.

Conclusion :
Using Kmeans we can segment our customer base into 2 major groups and similarly we can segment our customer base through hierarchical clustering.