**Unsupervised Learning – Clustering**

**K-Means and Hierarchical Clustering**

**Problem Statement** – Apply clustering on online retail transnational data which contains all the transactions occurring between 01/12/2010 and 09/12/2011 for a UK-based and registered non-store online retail.

**Building Model steps** –

1. Read and visualize the data
2. Clean the data
   1. Handle missing values
   2. Prepare the data for modelling
      1. Perform RFM analysis
         1. Recency
         2. Frequency
         3. Monetary
      2. Performed standardization and normalization
3. Built clustering model
   1. Identified optimal number of Clusters using Silhouette coefficient
   2. Run the model using optimal number of Clusters

Also performed hierarchical clustering on the same dataset. In hierarchical method calculated distance between datapoints using linkage method. Plotted dendrogram to visualize the clusters. And at the end cut the dendrogram to get the clusters.