Customer Segmentation

NAME	ROLL NUMBER
Ch.Laxmi Prashanth Reddy	2011CS020070

Customer Segmentation developed using Python, involves classification of customers in an Online Retail Store and study of their granular purchase behaviour. Helpful in devising strategies which uncovers deeper understanding of purchase decisions taken by the customers, determine appropriate product pricing and recommend as well.

Modules:

Data Import and Process

RFM

Exploratory Data Analysis

K-Means Clustering

Feature scaling

Data Import and Process

It includes

1. Reading the data from csv file
2. Preprocessing the dataset
    1. Removing duplicates value
    2. Removing Duplicate values

RFM analysis

RFM stands for recency, frequency, monetary value. In business analytics, we often use this concept to divide customers into different segments, like high-value customers, medium value customers or low-value customers, and similarly many others.

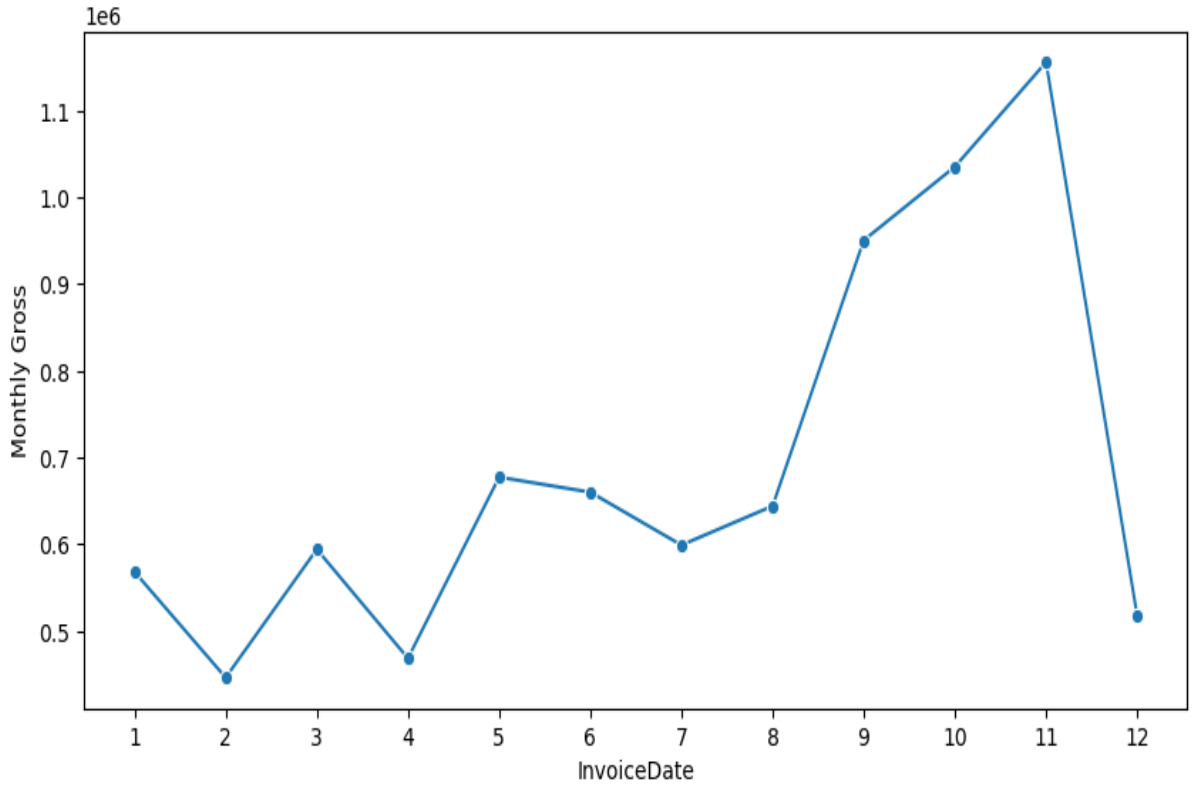
Let's assume we are a company, our company name is geek, let's perform the RFM analysis on our customers

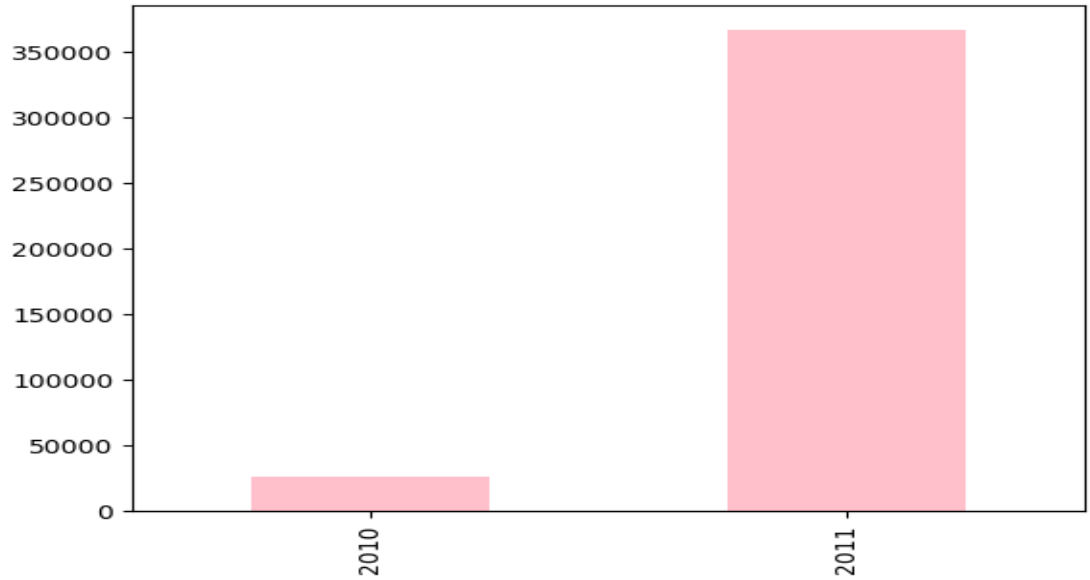
1. Recency: How recently has the customer made a transaction with us
2. Frequency: How frequent is the customer in ordering/buying some product from us
3. Monetary: How much does the customer spend on purchasing products from us.

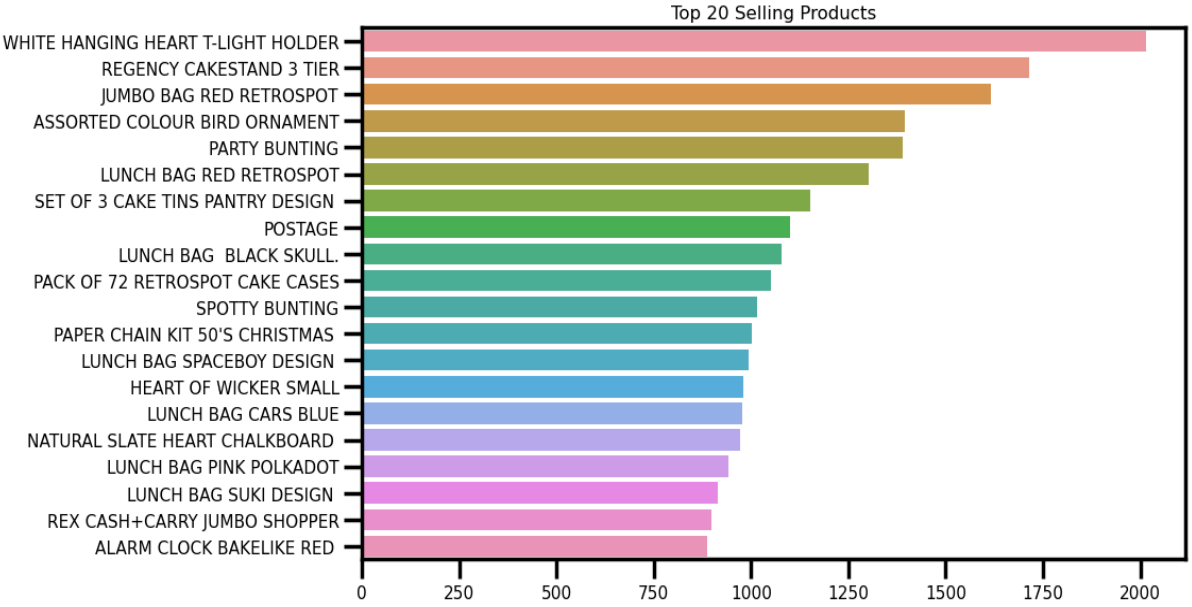
Exploratory Data Analysis

Performing various visualization plots like:

i. Barplot
ii. Lineplot
iii. countplot





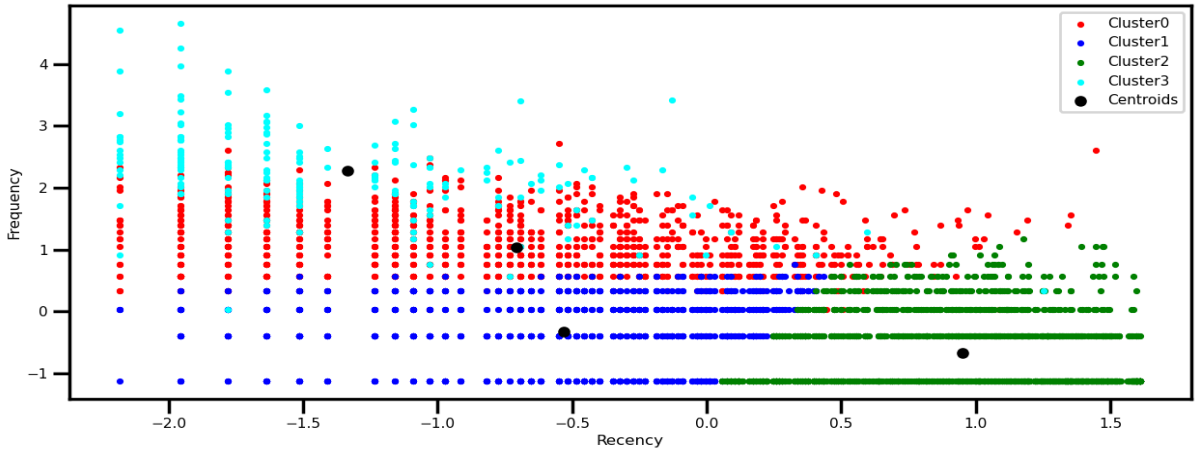


Algorithem Used

Kmeans algorithem

K-Means Clustering is an Unsupervised Learning Algorithem, which groups the unlabeled dataset into different clusters. Here K defines the number of pre-defined clusters that need to be created in the process, as if K=2, there will be two clusters, and for K=3, there will be three clusters, and so on

`



Feature Extraction

Using this dataset to find frequent itemset

Applying various unsupervised algorithems