Customer Segmentation Analysis in Python

Steps to solve the problem :

1.Importing Libraries

Numpy, Pandas, Matplotlib, Seaborn, Scikit-Learn, Kneed, and Scipy

1. Pandas – for loading data as a dataframe and wrangling the data.
2. Numpy and Scipy – for performing some basic mathematical computations.
3. Scikit-Learn – for building our Customer Segmentation Model.
4. Seaborn, Matplotlib and Plotly Express – for data visualization

2.Exploration of data

The first step in understanding our data is to conduct exploratory data analysis. This process involves examining the basic properties of the dataset, such as viewing the initial rows, calculating summary statistics, and identifying any missing

values. By performing EDA, we can gain insights into the structure and distribution of the data.

3.Data Visualization

Way to communicate your customer segments is to visualize them and their relationships using charts, graphs, or maps. Visualization can help you highlight the similarities and differences between your clusters, as well as their patterns and trends over time or across dimensions.

4.Clustering using K-Means

K-Means clustering is an unsupervised machine learning algorithm that divides the given data into the given number of clusters. Here, the "K" is the given number of predefined clusters, that need to be created.

It is a centroid based algorithm in which each cluster is associated with a centroid. The main idea is to reduce the distance between the data points and their respective cluster centroid.

The algorithm takes raw unlabelled data as an input and divides the dataset into clusters and the process is repeated until the best clusters are found.

K-Means is very easy and simple to implement. It is highly scalable, can be applied to both small and large datasets. There is, however, a problem with choosing the number of clusters or K. Also, with the increase in dimensions, stability decreases. But, overall K Means is a simple and robust algorithm that makes clustering very easy.

5.Selection of Clusters

In the context of customer segmentation, customer clustering analysis is the use of a mathematical model to discover groups of similar customers based on finding the smallest variations among customers within each group.

6.Conclusion

Customer segmentation is a unique strategy that can help to find your target audience, improve client retention, and overall improve your number of clients and sources of revenue. Thus it plays an important role.