Online_Retail_Customer_Segmentation

Unsupervised ML Clustering Analysis

The need of customer segmentation:

The differences in customers' behaviour, demographics, geographies, etc. help in classifying them in groups. Learning about different groups in the customer can help with following:

Target Marketing Client understanding Optimal product placement Searching for new customers Revenue growth

Recency-Frequency-Monetary (RFM) model to determine customer value:

The RFM model is quite useful model in retail customer segmentation where only the data of customer transaction is available. RFM stands for the three dimensions:

Recency – How recently did the customer purchase? Frequency – How often do they purchase? Monetary Value – How much do they spend? A combination of these three attributes can be defined to assign a quantitative value to customers. e.g. A customer who recently bought high value products and transacts regularly is a high value customer.

Segmentation with K-means clustering:

Initially, the data is subject to important stages in an analytics pipeline: exploratory analysis, preprocessing, feature engineering and standardizaton. Then, the unsupervised classification technique, K-means clustering algorithm, is used to determine the ideal segments of customers. Silhouette analysis 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

• K-Means Clustering with Silhouette gives the highest score for number of clusters 3.

• Sales has been increased from 2010 to 2011.

- RFM for Cluster ID box plots tells well about Cluster detail.