Capstone Project Submission

Instructions:

i) Please fill in all the required information.

ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:

1) **Mohd Danish:** **Email:** mdanish63364@gmail.com
   1) Data Cleaning:
      - Handling null and missing values.
      - Handling cancelled orders.
   2) Feature Engineering:
      - Introducing new variables with DateTime
      - Introduced Total amount with Quantity and UnitPrice
   3) EDA(Exploratory Data Analysis):
      - Barplots(Top 5 Products, Top 5 Months, Top 5 Customer IDs)
      - Boxplot(handling RFM Outliers)
      - Distplot(All numerical features)
      - Correlation map between all features.

   4) Data Transformation:
      - Introducing RFM table
      - Log Transformation applied on Recency, Frequency and Moetary
   5) Machine Learning Clustering Algorithms
      - K-Means
      - K-Means with Silhouette score
      - K-Means with Elbow method
      - Hierarchical Clustering
   6) Group Colab
2) **Abdul Rahman Talha:** **Email:** rahman88talha@gmail.com
   1) Feature Engineering:
      - Data Preprocessing
      - Introducing new features
   2) EDA(Exploratory Data Analysis):
      - Barplots(Top 10 countries, Customers over years)
      - Distplot
      - Correlation map
   3) Data Transformation:
      - Log Transformation
      - Splitting into quantiles
   4) Machine Learning Clustering algorithms
      - K-Means Clustering
      - K-Means with Silhouette score and Calinski Harabasz
      - Hierarchical Clustering
   5) PPT

Please paste the GitHub Repo link.

Github Link:- Rahman88talha/Capstone-Customer_Segmentation_Clustering

Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)

Businesses all over the world are growing every day. With the help of technology, they have access to a wider market and hence, a large customer base. Customer segmentation refers to categorizing customers into different groups with similar characteristics. Customer segmentation can help businesses focus on each customer group in a different way, in order to maximize benefits for customers as well as the business. This project mainly deals in segmenting customers of an online business store in the UK.

In this project, we performed one of the most essential applications of machine learning, Customer Segmentation by using K-Means Clustering Algorithm. In this project, we implemented customer segmentation in Python. Whenever you need to find your best customer, customer segmentation is the ideal methodology. we have seen descriptive analysis of our data and then implemented several versions of the K-means algorithm to identify the optimum number of clusters, we used the elbow method and silhouette analysis.

Conclusion:

-Various steps to perform customer segmentation. We started with data wrangling in which we tried to handle null values to feature modifications.
Next, we did some EDA and tried to draw observations from the features we had in the dataset.

-We saw how we can segment our customer depending on our business requirements

-RFM (Recency, Frequency and Monetary) data frame ease our problem to solve in a particular order, it makes easy to recommend and display new launched products to few customers. RFM analysis can help in answering many questions with respect to their customers and this can help companies to make marketing strategies for their customers based on their interest.

-Used cluster profiling the average of recency, frequency and monetary values for each customer segment was identified.

-Used the K-means algorithm to segment our customer in various clusters having similar similarity.