**Airline Passsengers Segmenting using KMeans Clustering**

**Project Overview**

Dataset Description:

The dataset consists of customer information with various attributes such as balance, miles earned with different credit cards, bonus miles, flight miles, and days since enrolment. The goal is to perform customer segmentation using the KMeans clustering algorithm to identify distinct groups of customers.

Column Descriptions:

ID: Unique ID

Balance: Number of miles eligible for award travel

Qual\_miles: Number of miles counted as qualifying for Topflight status

cc1\_miles, cc2\_miles, cc3\_miles: Miles earned with credit cards (categorized)

Bonus\_miles: Number of miles earned from non-flight bonus transactions

Bonus\_trans: Number of non-flight bonus transactions

Flight\_miles\_12mo: Number of flight miles in the past 12 months

Flight\_trans\_12: Number of flight transactions in the past 12 months

Days\_since\_enroll: Number of days since enrolled in the flyer program

Award: Whether the person had an award flight (free flight) or not

**Methodology**

Data Exploration and Preprocessing:

· Imported necessary packages

· Checked data information (no null values)

· Explored basic statistics of the data using describe

· Dropped the 'ID' column as it is unique and not relevant for analysis

**Exploratory Data Analysis (EDA):**

· Checked the categories of 'cc1\_miles', 'cc2\_miles', 'cc3\_miles' using countplot.

· Explored correlation using a heatmap

**KMeans Clustering:**

· Utilized the Elbow Method to determine the optimal number of clusters (k=5)

· Applied KMeans clustering with k=5

· Added the 'Cluster' column to the dataset

**Visualization:**

Explored the distribution of clusters using a scatter plot (Balance vs. Days\_since\_enroll, Balance vs. Bonus\_miles, Balance vs. Flight\_miles\_12mo)

**Results**

The KMeans clustering algorithm successfully segmented customers into five distinct clusters. The scatter plot revealed patterns in customer segmentation based on balance and days since enrolment. Further analysis and targeted strategies can be implemented for each customer cluster.

**Conclusion**

The application of KMeans clustering provides valuable insights into customer segmentation, allowing businesses to tailor their marketing and service strategies for different customer groups. Understanding customer behaviour and preferences is crucial for enhancing customer satisfaction and optimizing business outcomes.