Marwadi University	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Machine Learning (01CT1519)	**Aim: KMeans Clustering**	
Assignment 6	**Date:30-09-2025**	**Enrolment No:92301733054**

Code:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import random

dataset = pd.read_csv("/content/E-commerce Customer Behavior - Sheet1.csv")
dataset.head()

features = ['Total Spend', 'Items Purchased', 'Days Since Last Purchase', 'Average Rating']
X = np.array(dataset[features])

def calculate_distance(a, b):
    return np.sqrt(np.sum((a - b) ** 2))

def assign_clusters(centroids, X):
    assigned_cluster = []
    for i in X:
        distances = [calculate_distance(i, c) for c in centroids]
        assigned_cluster.append(np.argmin(distances))
    return assigned_cluster

def update_centroids(clusters, X):
    new_centroids = []
    df_temp = pd.concat([pd.DataFrame(X), pd.Series(clusters, name='cluster')], axis=1)
    for c in set(df_temp['cluster']):
        current_cluster = df_temp[df_temp['cluster'] == c][df_temp.columns[:-1]]
        cluster_mean = current_cluster.mean(axis=0)
        new_centroids.append(cluster_mean)
    return np.array(new_centroids)

k = 3
init_indices = random.sample(range(len(X)), k)
centroids = np.array([X[i] for i in init_indices])
print("Initial Centroids:\n", centroids)

epochs = 10
for i in range(epochs):
    clusters = assign_clusters(centroids, X)
    centroids = update_centroids(clusters, X)

    if i == 0 or i == epochs - 1:
        plt.figure(figsize=(10,7))
        plt.scatter(X[:,0], X[:,1], c=clusters, alpha=0.3)
        plt.scatter(centroids[:,0], centroids[:,1], color='black', marker='X', s=200)
        plt.title(f'K-Means Clustering - Iteration {i+1}')
```

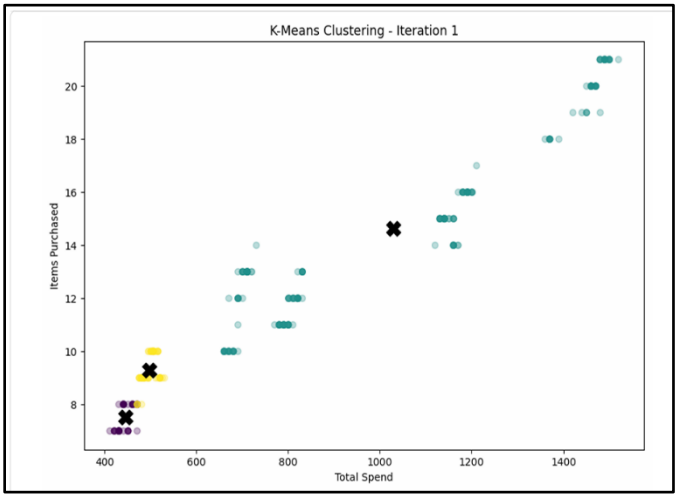
Marwadi University	Marwadi University Faculty of Technology Department of Information and Communication Technology	
Subject: Machine Learning (01CT1519)	Aim: KMeans Clustering	
Assignment 6	Date:30-09-2025	Enrolment No:92301733054

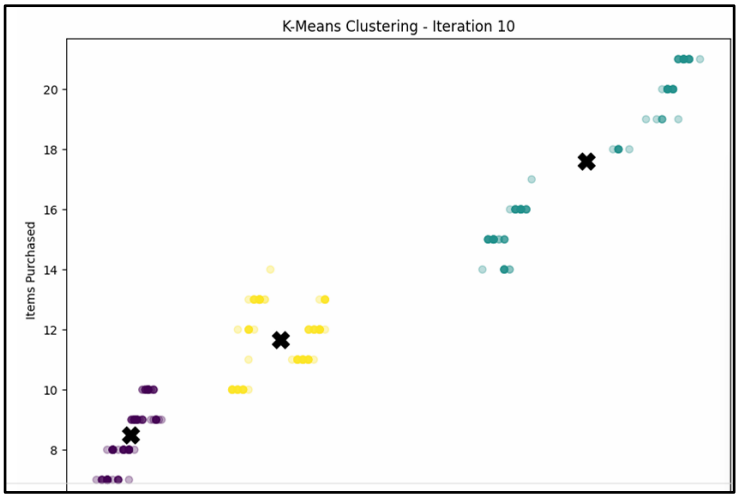
```
    plt.xlabel(features[0])
    plt.ylabel(features[1])
    plt.show()

dataset['Cluster'] = clusters

for i in range(k):
  print(f"\nCluster {i} Summary:")
  print(dataset[dataset['Cluster']==i][features].describe())
```

Results:





Conclusion: In this assignment, we learned how to use K-Means clustering to group customers based on their shopping behavior. It helped us identify loyal buyers, moderate shoppers, and low-engagement customers, showing how data can guide better business decisions.