import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

from sklearn.cluster import KMeans

mall\_data = pd.read\_csv("F:\exam\exp 4\Mall\_Customers.csv")

mall\_data.head()

mall\_data.info()

plt.figure(figsize=(10,10))

sns.countplot(x="Genre", data=mall\_data)

plt.figure(figsize=(16,10))

sns.countplot(x="Age", data=mall\_data)

plt.figure(figsize=(20,8))

sns.barplot(x='Annual Income (k$)',y='Spending Score (1-100)',data=mall\_data)

X = mall\_data.iloc[:,[2,3,4]].values

X

wcss = []

for i in range(1,11):

kmeans = KMeans(n\_clusters=i, init='k-means++', random\_state=50)

kmeans.fit(X)

wcss.append(kmeans.inertia\_)

wcss

sns.set()

plt.plot(range(1,11),wcss)

plt.xlabel("Number of clusters")

plt.ylabel("WCSS value")

plt.show()

kmeans = KMeans(n\_clusters = 5, init = 'k-means++',random\_state = 0)

y = kmeans.fit\_predict(X)

fig = plt.figure(figsize = (10,10))

ax = fig.add\_subplot(111, projection='3d')

ax.scatter(X[y == 0,0],X[y == 0,1],X[y == 0,2], s = 40 , color = 'red', label = "cluster 1")

ax.scatter(X[y == 1,0],X[y == 1,1],X[y == 1,2], s = 40 , color = 'blue', label = "cluster 2")

ax.scatter(X[y == 2,0],X[y == 2,1],X[y == 2,2], s = 40 , color = 'green', label = "cluster 3")

ax.scatter(X[y == 3,0],X[y == 3,1],X[y == 3,2], s = 40 , color = 'yellow', label = "cluster 4")

ax.scatter(X[y == 4,0],X[y == 4,1],X[y == 4,2], s = 40 , color = 'purple', label = "cluster 5")

ax.set\_xlabel('Age of a customer-->')

ax.set\_ylabel('Anual Income-->')

ax.set\_zlabel('Spending Score-->')

ax.legend()

plt.show()