* **Given the e-commerce data in cust\_data.xlsx, use k-means clustering algorithm to cluster customers with similar interest.**
* import pandas as pd
* import numpy as np
* import matplotlib.pyplot as plt
* import seaborn as sns
* import warnings
* warnings.filterwarnings("ignore")
* df = pd.read\_csv("Mall\_Customers.csv")
* df
* df.head()
* df.info()
* x = df.iloc[:,3:]
* x
* from sklearn.cluster import KMeans
* wccs = []
* for i in range(1,30):
* km = KMeans(n\_clusters = i, init = "k-means++", n\_init = 10, max\_iter = 300, random\_state = 1)
* km.fit(x)
* wccs.append(km.inertia\_)
* Wccs
* plt.plot(range(1,30), wccs, marker = "\*")
* plt.title("Elbow Method")
* plt.xlabel("K Value")
* plt.ylabel("Inertia")
* plt.grid()
* plt.show()
* km1 = KMeans(n\_clusters = 5, init = "k-means++", n\_init = 10, max\_iter = 300, random\_state = 1)
* labels = km1.fit\_predict(x)
* km1.inertia\_
* km1.cluster\_centers\_
* labels
* df["c\_labels"] = labels
* df
* centroid\_df = pd.DataFrame(km1.cluster\_centers\_, columns = ["X","Y"])
* centroid\_df
* sns.scatterplot(data = df, x = "Annual Income (k$)", y = "Spending Score (1-100)", hue = "c\_labels")
* plt.scatter(centroid\_df["X"], centroid\_df["Y"], s = 40, color = "yellow")
* plt.show()