**#ml5 salesdata**

**import pandas as pd**

**import numpy as np**

**import matplotlib.pyplot as plt**

**from sklearn.cluster import KMeans**

**df = pd.read\_csv("sales\_data\_sample.csv",encoding= 'latin')**

**df.dtypes**

**X = df.iloc[:, [3,4]].values**

**wcss = [] #within cluster sum of square**

**for i in range(1,11):**

**#init argument is the method for initializing the centroid**

**kmeans = KMeans(n\_clusters=i, init="k-means++", random\_state=42)**

**kmeans.fit(X)**

**#we calculate wcss value for each k value**

**wcss.append(kmeans.inertia\_)**

**ks = [1,2,3,4,5,6,7,8,9,10]**

**plt.plot(ks, wcss, 'bx-')**

**plt.title("Elbow method")**

**plt.xlabel("K value")**

**plt.ylabel("WCSS")**

**df.describe()**

**wcss =[]**

**for i in range(1,11):**

**clustering = KMeans(n\_clusters=i, init="k-means++", random\_state=42)**

**clustering.fit(scaled)**

**wcss.append(clustering.inertia\_)**

**ks = [1,2,3,4,5,6,7,8,9,10]**

**plt.plot(ks, wcss, 'bx-')**

**plt.title("Elbow method")**

**plt.xlabel("K value")**

**plt.ylabel("WCSS")**