

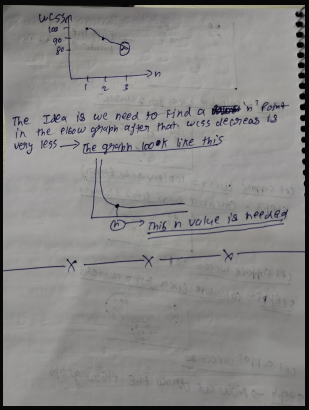
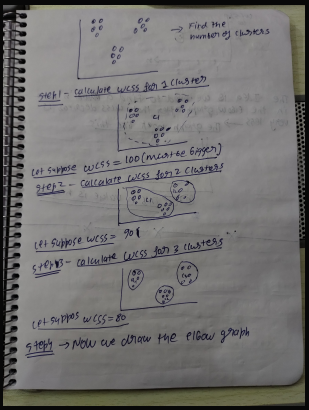
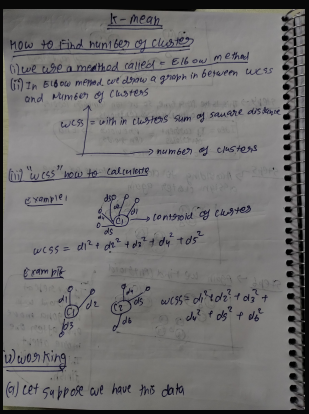
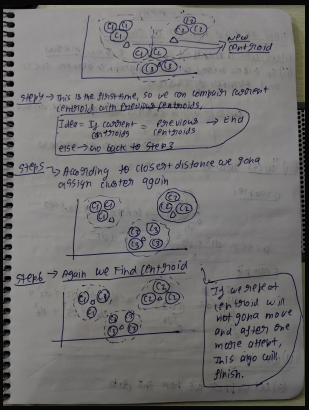
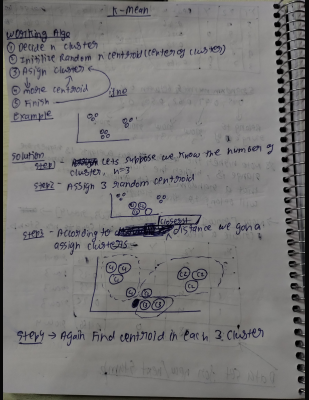
K-Mean Clustering

Geometric Intuition

video

Here k mean = Number of cluster

working



How to find number of clusters

generally the number of cluster is less than 20

code

video

github

Elbow method

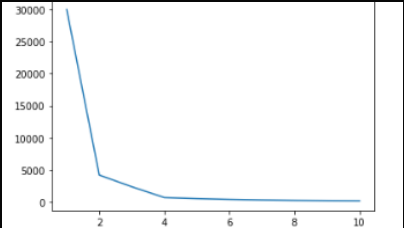
from sklearn.cluster import KMeans

wcss = []

```
for i in range(1,11):
    km = KMeans(n_clusters=i)
    km.fit_predict(df)
    wcss.append(km.inertia_)
```

```
[29957.898287999997,
4184.141269999999,
2362.7133490000015,
681.9696600000001,
514.1616803171117,
388.8524026875981,
295.4391895943191,
234.48687343199217,
199.99120032567834,
171.40590752168507]
```

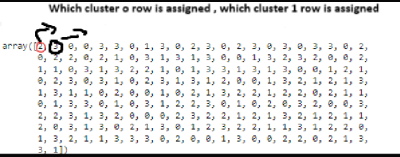
wcss



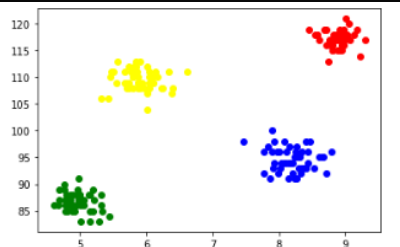
plt.plot(range(1,11),wcss)

Now we know Number of cluster = 4

```
X = df.iloc[:,].values
km = KMeans(n_clusters=4)
y_means = km.fit_predict(X)
```



y_means



```
plt.scatter(X[y_means == 0,0],X[y_means == 0,1],color='blue')
plt.scatter(X[y_means == 1,0],X[y_means == 1,1],color='red')
plt.scatter(X[y_means == 2,0],X[y_means == 2,1],color='green')
plt.scatter(X[y_means == 3,0],X[y_means == 3,1],color='yellow')
```