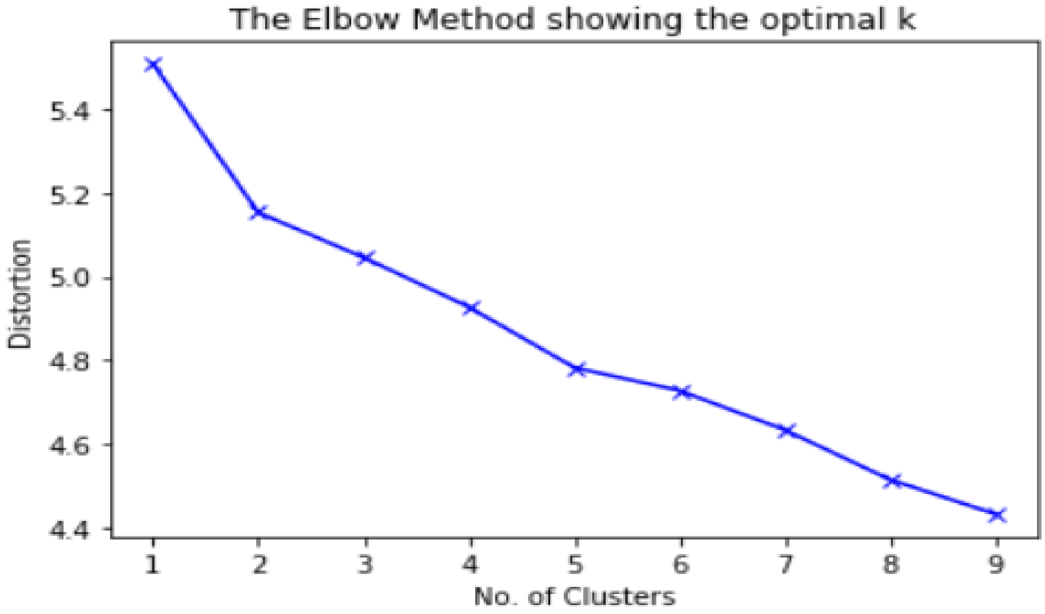
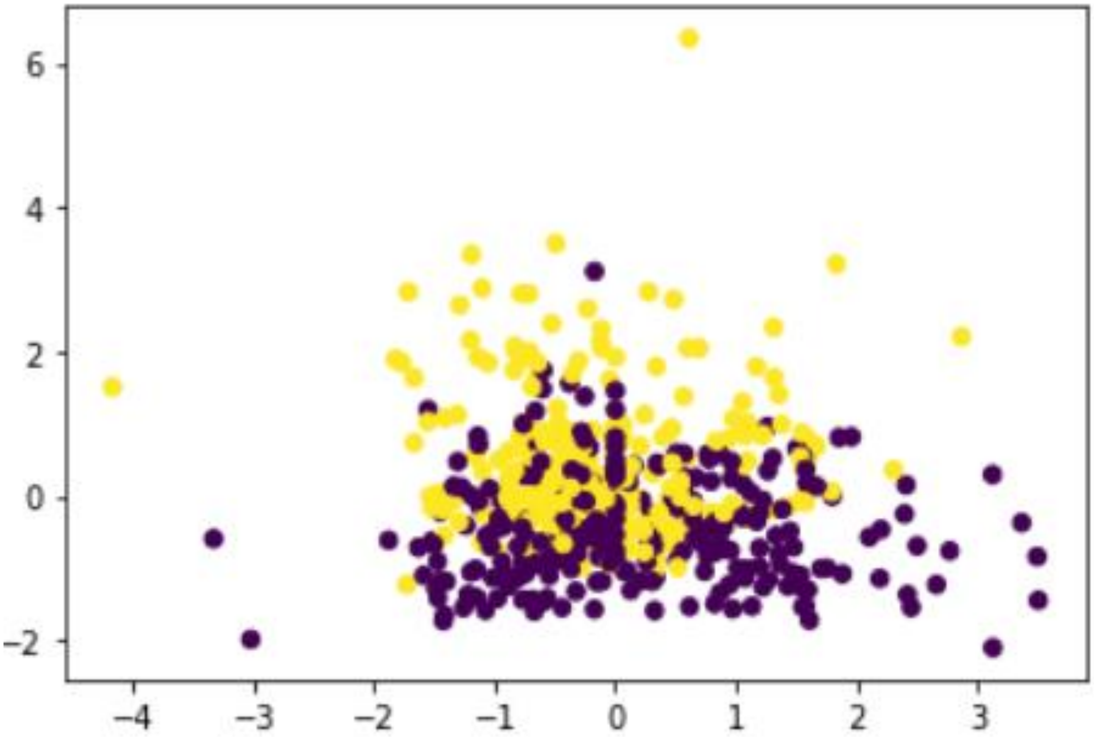
1. For filling up missing values, I replaced them with the mean given in the question. And for Normalizing all values I used *StandardScaler's* fit method.

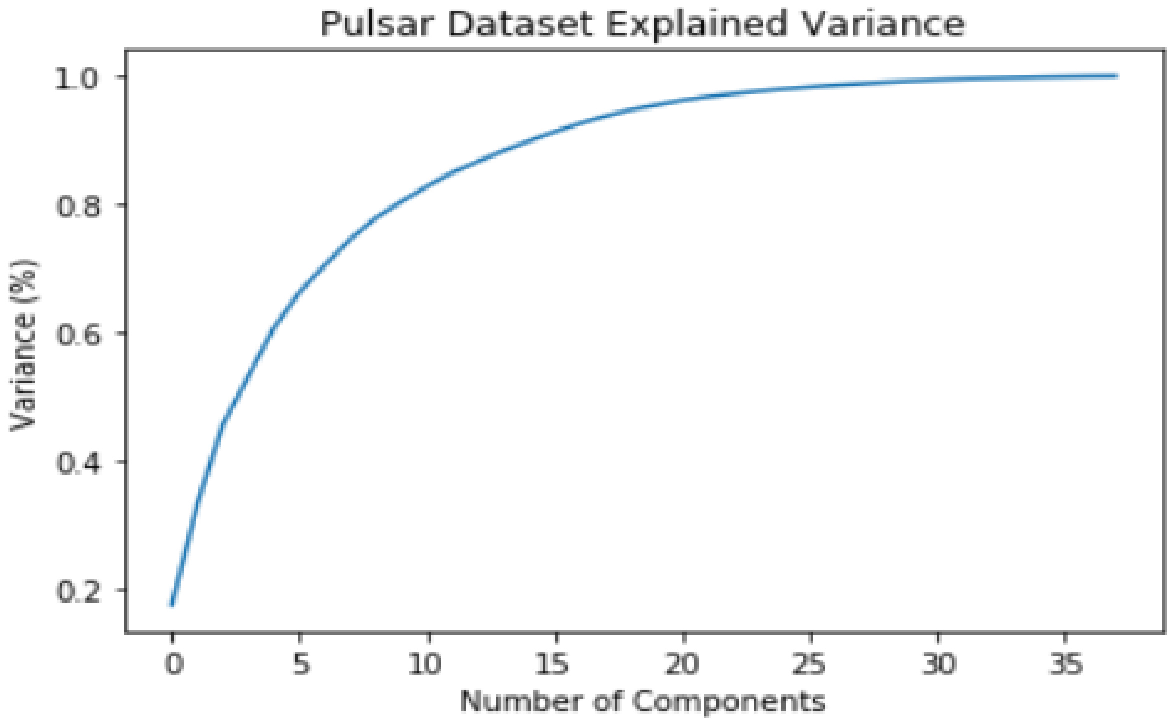
2. To determine value of k in K-Means there are multiple *methods*. I have used *Elbow* method. In which when we see an elbow in the graph then the number at which the elbow is generated is taken as the value of k. So, the elbow was generated at 2. Thus, value of k was chosen as 2.

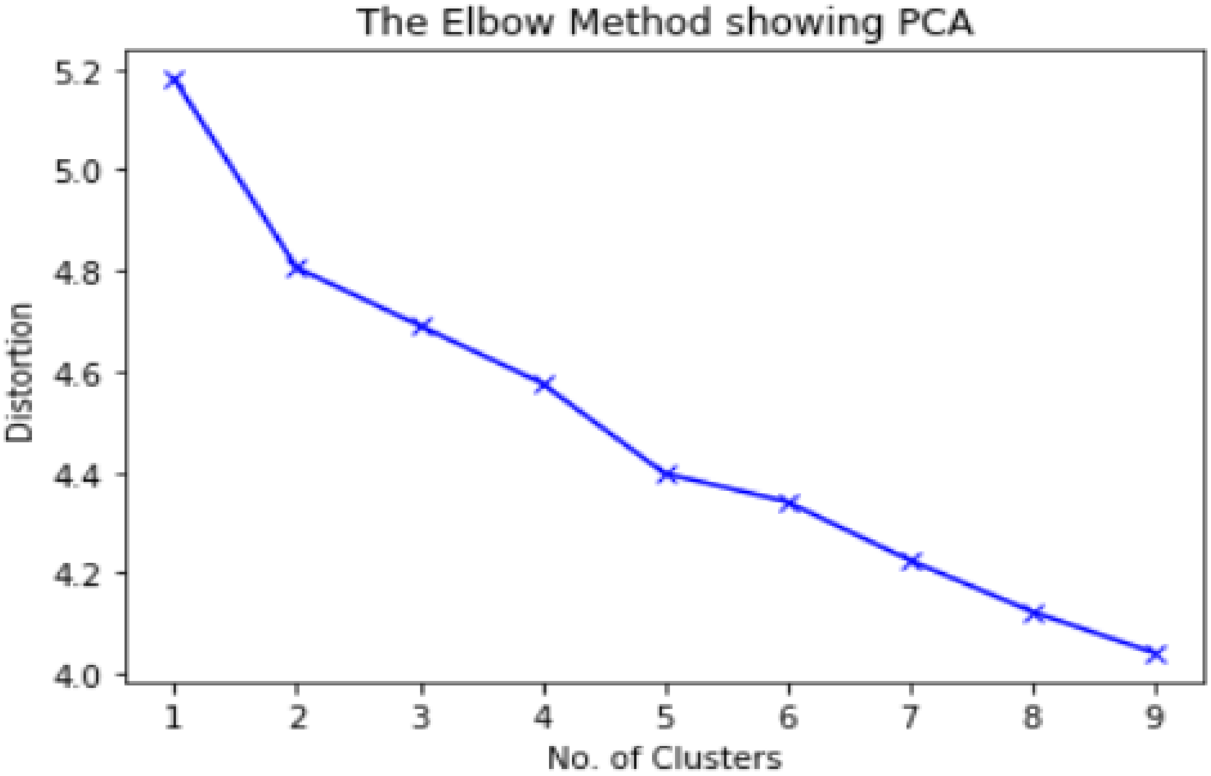


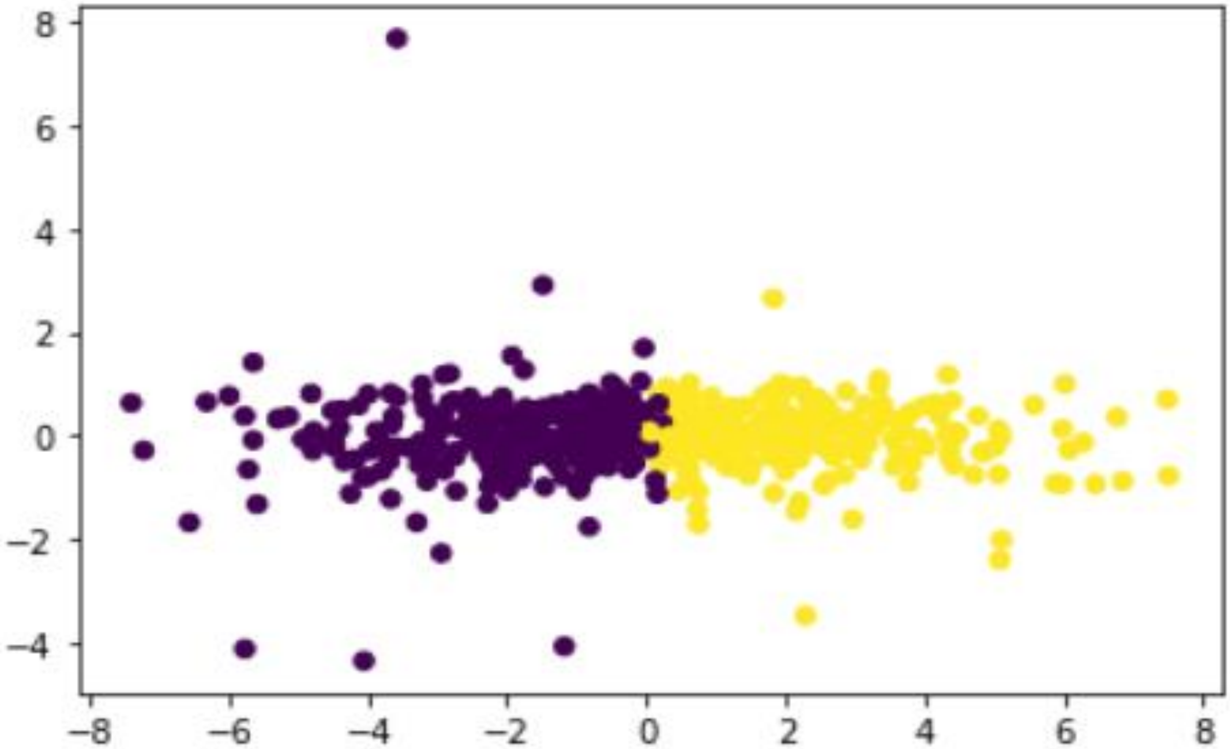
3. As mentioned above elbow was created at 2 and taking value of k=2 for *K-Means*. And then plotting scattered plot I got 2 different clusters. But the clusters were not clearly visible as a difference.



4. For applying PCA we need to know the number of components. For that I used *Pulsar Dataset Explained Variance* method. From that taking 90% of variance I got value of components to be taken as 15. After applying *PCA* to the dataset, I again checked for value of k using Elbow method. I got value of k = 2. After that taking value of k=2 in K-Means and plotting scattered plot.

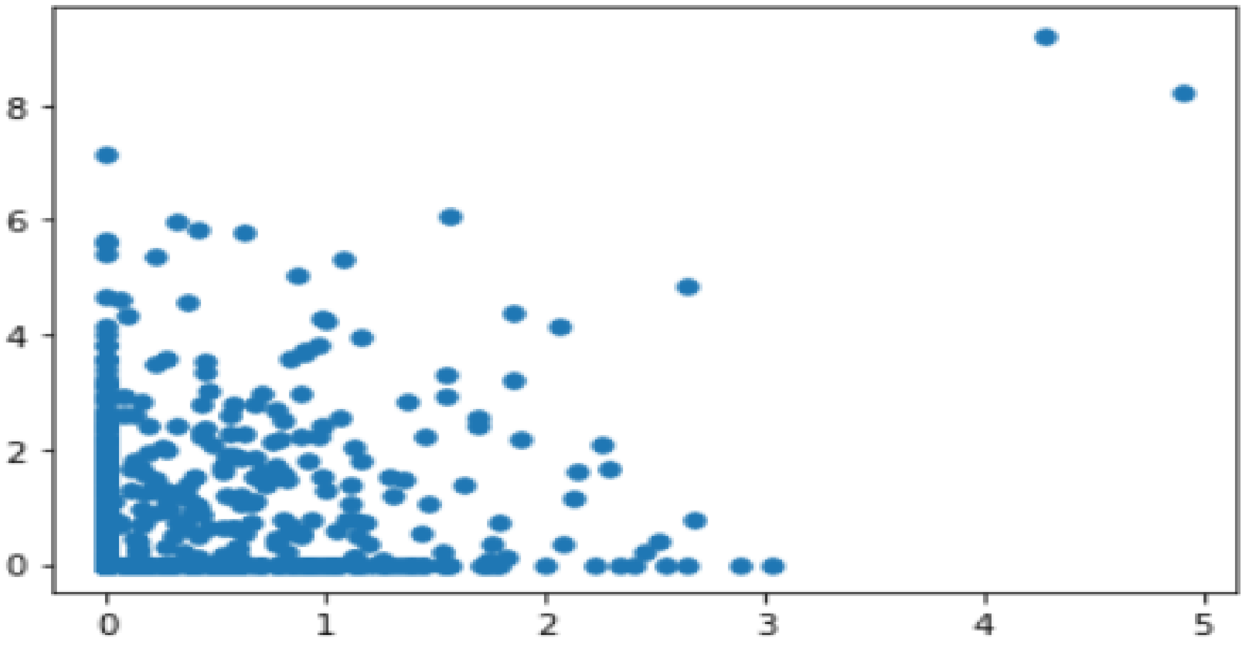



The Elbow Method showing PCA
Distortion
5.2
5.0
4.8
4.6
4.4
4.2
4.0
1
2
3
4
5
6
7
8
9
No. of Clusters


8
6
4
2
0
−2
−4
−8
−6
−4
−2
0
2
4
6
8

5. In the first cluster (without PCA) the data cannot be differentiated but after applying PCA we can differentiate the data and 2 clear clusters are formed. This difference is generated as we have only done data normalizing in first scatter plot and in the second scattered plot we have done data normalization & PCA as well. So, data is normalized & also feature reduction is done. Thus, there is a vast difference in both the scattered plot.

6. As seen in the scatter plot of *Autoencoder*, we can see that all the data is related to 0. The autoencoder tends to perform better when k is small when compared to PCA, meaning the same accuracy can be achieved with less components and hence a smaller data set. This is important when dealing with very large data sets. Autoencoders require more computation than PCA. Although, for very large data sets that can't be stored in memory, PCA will not be able to be performed. The autoencoder construction using keras can easily be batched resolving memory limitations.


8
6
4
2
0
0
1
2
3
4
5