PCA is used to reduce high-dimensional data to a lower-dimensional space in a way that preserves variability.

VISUALIZE 2D PROJECTION

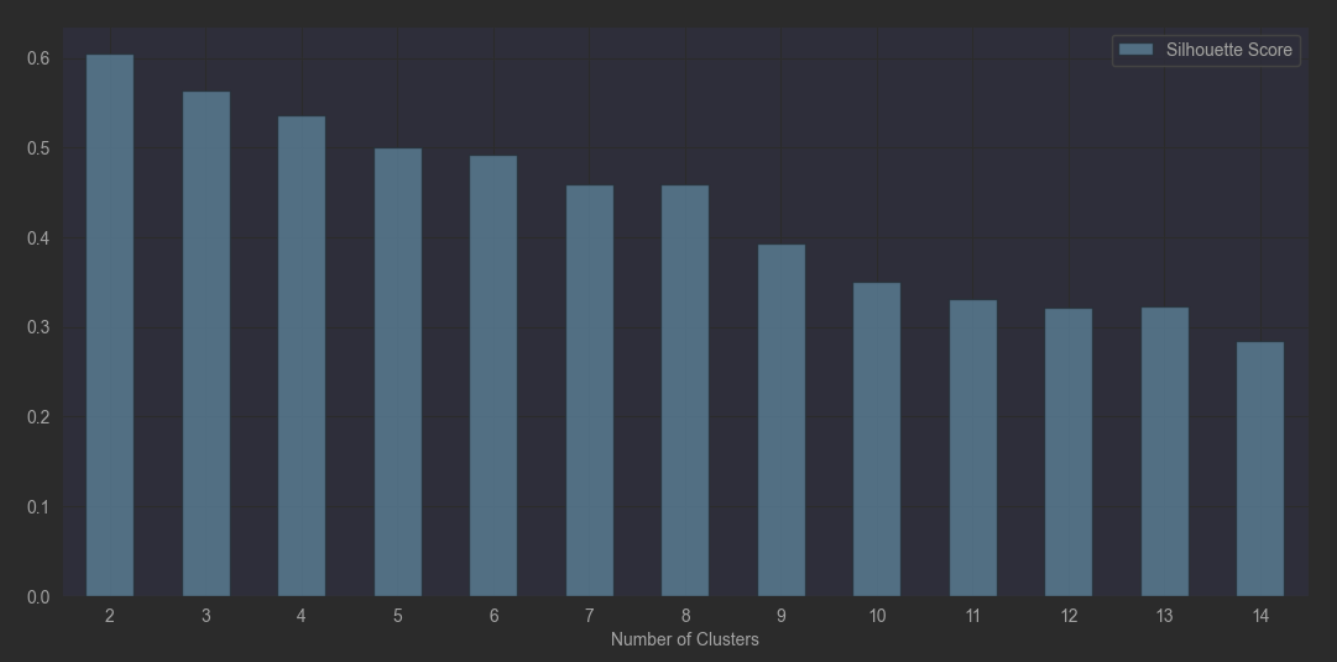


From the above plot we observe the two groups or clusters identified by the Churn and Principal Components

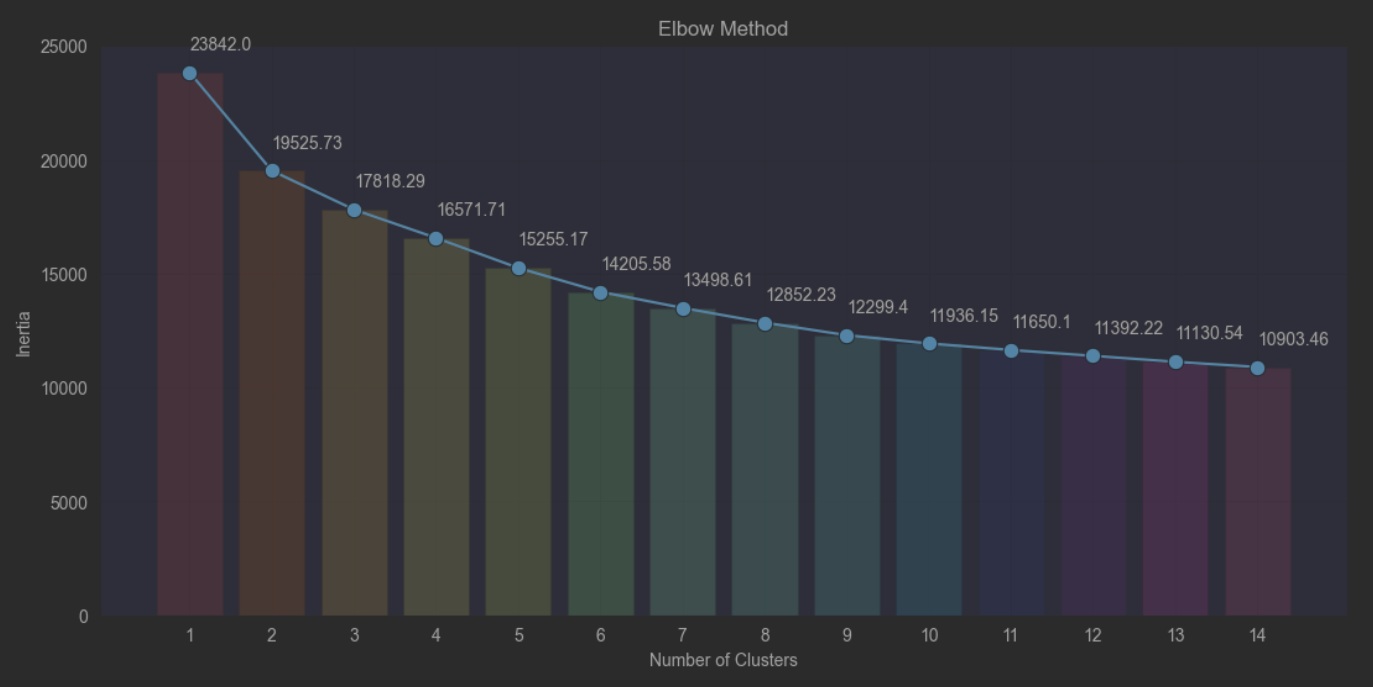
To select K, which is the number of clusters, you run an ensemble of k-means over a range of different values of K.

The following methods will be used for identifying the number of clusters

* Silhouette score

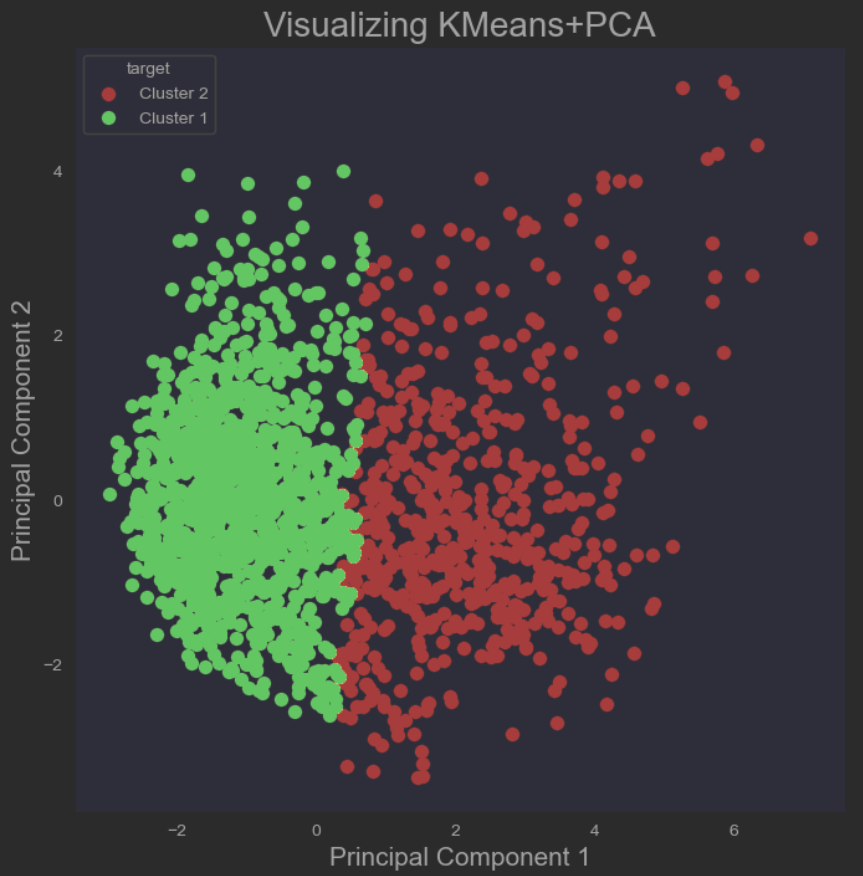


* Elbow method



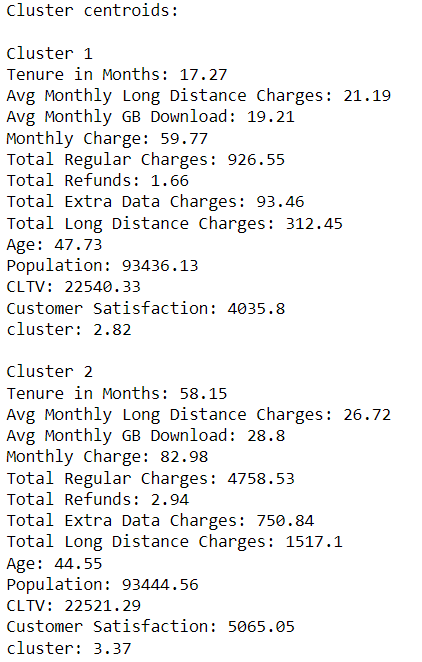
From the above clustering methods, we can see the optimal number of clusters is two.

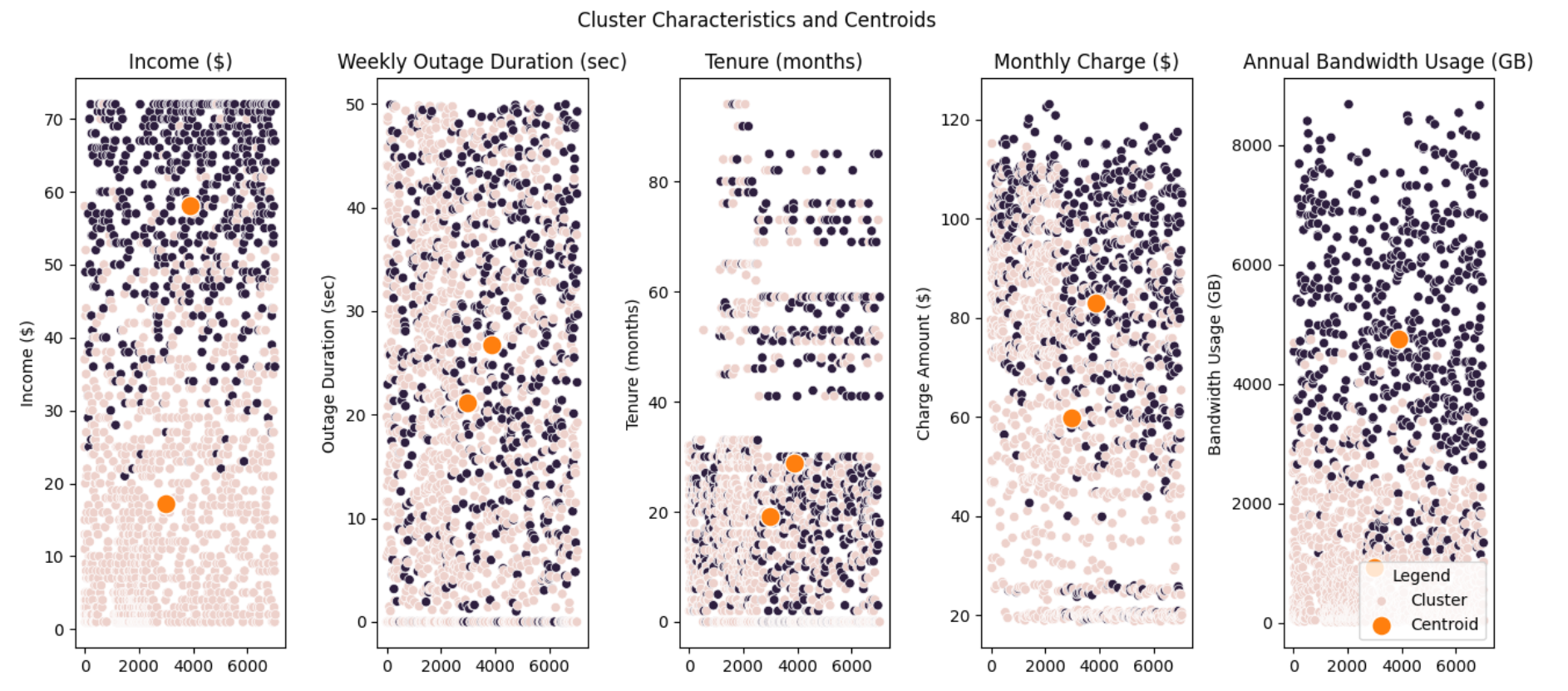
One method of clustering data in Python, is using scikit-learn, clusters can be identified using K-means.



From the above plot we can clearly observe two groups identified by the target (Clusters)

**K-means** uses random initialization and sets the centroids to existing datapoints at random. An advantage of k-means is that it is able to predict, or assign a class to, a new customer.





The results of this analysis have several important implications for the telecommunications company. First, the company can use customer segmentation to allocate resources more efficiently and develop marketing campaigns considering the specific characteristics of each cluster. Second, the company can focus on efforts to retain customers, especially customers from cluster 1 who have shorter tenure.

Considering the centroids of each cluster, it is recommended that the telecommunications company focus on extending the stay of customers in cluster 1 through targeted marketing campaigns and incentives. For Cluster 2 clients with longer average tenure, the company must ensure that they feel valued through reward programs or targeted discounts for additional services.