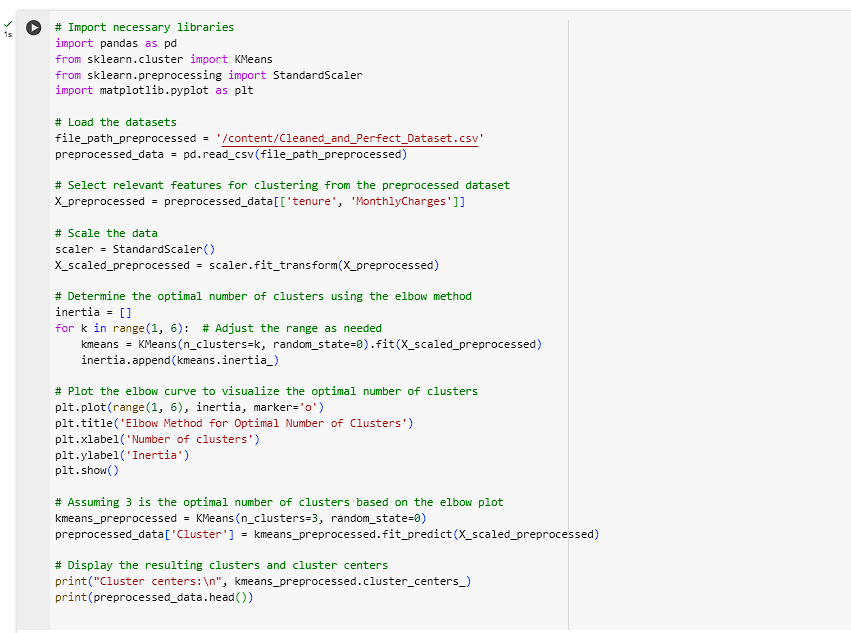
# K-mean Clustering

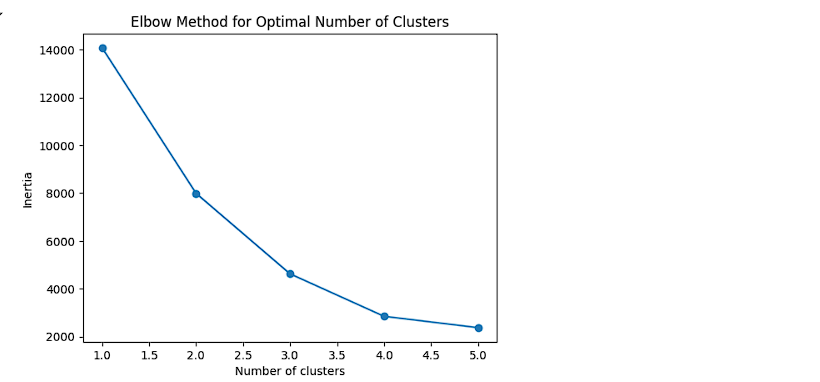
Input:



**Explanation for Your Video:**

In this K-Means clustering process, I used the elbow method to determine the optimal number of clusters, which helps in finding the appropriate number of groupings for the data. I applied K-Means clustering to segment the data into three distinct clusters, visualized the results, and analyzed the patterns based on key features like tenure and monthly charges.

Clusters:



A screenshot of a computer code

Description automatically generated

Being a statistical simulation model analysis report, this paper focuses on the K-Means clustering method applied to concrete compressive strength data.

To find significant variations within the customers database, this paper used the K-Means clustering algorithm on a preprocessed data. The clustering was performed based on two features: Tenure (how long the customer has been associated with the service) and MonthlyCharges (that the customer is charged every month). The objective was to partition the customers into three homogeneous groups since this information would be useful in decision making, especially in explaining customer behavior and likely churn status.

Key Steps Involved

Data Preprocessing:

The dataset was cleaned by dealing with missing values and followed by the process of encoding the categorical variables. Tenure and MonthlyCharges were adopted as desirable metrics to form clusters for what interested analysis.

After that, StandardScaler was used to scale the features since it is important to give all of them an equal chance at the clustering stage by avoiding situations where a feature with a much larger value will sway the clustering output.

Optimal Number of Clusters:

Using the elbow method it was possible to identify the right number of clusters to be used in the study. This enables one arrive at the number of clusters that will improve the performance of the model instead of encouraging the creation of new clusters hence over complicating the model.

For further analysis, 3 clusters were selected to the customers because the inertia of elbow curve starts decreasing at a diminishing rate at the third cluster.

Clustering Results:

The K-Means algorithm was run with 3 clusters, and the resulting cluster centers are:

Cluster 0: In fact, this group is defined by the least tenure among the customers as well as the least MonthlyCharges. It may be argued that they are likely more recent users of the services, whereby have cheaper tariffs.

Cluster 1: This group also consist of customers with higher MonthlyCharges and their tenure is also slightly higher. It means that these customers are probably the true ‘customers’, which are more likely to be the loyal subscribers, paying for more numerous and expensive services.

Cluster 2: This particular cluster of customers has a relatively moderate customer tenure and MonthlyCharges. They are an average audience, perhaps, those who have worked with the service for some time and utilize basic service offers.

Cluster Interpretation:

Cluster 0 could be perhaps associated with a higher churn rate because of the low tenure and lower MonthlyCharges; the clients are possibly dissatisfied or do not consider themselves committed to the service.

It was possible that Cluster 1 could have been the most loyal group which therefore comes with low likelihood of churning. These consumers have been using the service for a longer time and they are charged more for the same, therefore they could be using other services or subscribing to the superior plans.

It is possible that Cluster 2 might be a middle ground; a middle group. These are not new customers nor are they the oldest, and their expenditure lies in the middle level.

Implications of Findings

This clustering analysis provides actionable insights for targeted marketing strategies:

Cluster 0 is suitable for retention campaigns because the churned customers can be contacted with a special offer or offered better terms so that they become satisfied again.

Due to the high likelihood of churn, cluster one should be targeted for loyalty programs to ensure that they don’t leave their loyalty programs. Another solution that might be useful for this group is that cable providers should try to sell more high-margin products and services.

Understanding the above findings, the following strategies may be recommended to push the current Cluster 2 closer to Cluster 1: As the usage increases with more services demanded and/or better packages taken over time, the number of months charged will likewise increase.

Therefore, using the K-Means clustering algorithm we have effectively used the customer data and were able to identify a clustering model which classifies customers into 3 different clusters. They can help inform customer-service decisions, as well as the formulation of customer loyalty and retention strategies and approaches for increasing up-selling, which in turn results in satisfying customers and minimizing churn.