The attached R script conducts K-means clustering analysis on two datasets, becuase I did not have a data set with 3000 rows. So two fulfil that requirment I used two different data sets.

It begins by installing and loading necessary packages, then defines a function for plotting to determine the optimal number of clusters. After processing the data, it performs K-means clustering, visualizes the results, and interprets the clustering outcomes.

Cluster Analysis 1

A screenshot of a computer

Description automatically generated

In the cluster plot below, there is a slight overlap of the clusters, but after analysing the results cluster centre, except "Z\_CostContact" all other variables have different cluster centres, thus the cluster analysis was a success.

A screenshot of a computer

Description automatically generated

Cluster Analysis 2

A screenshot of a computer

Description automatically generated

At first glance at the cluster diagram, we can see there is a slight overlap of the clusters in the middle, but after analysing the results of the cluster centre, none of the variable have the same centre thus the cluster analysis was a success.

A screenshot of a computer screen

Description automatically generated