Text

Description automatically generated

15.4 a)

Diagram

Description automatically generated with medium confidence

As we can see in this picture we can divide the data into two big groups, and if we keep going down, we can divide the data into 6 smaller groups.

15.4 b)

If the data is not standardized, the units of measurement of each variable would be very different and it would affect the measurement of the distance between each cluster, which would cause the elements to be clustered incorrectly.

15.4 c)

As we can see the passengers of group two travel more than the passengers in group one, so we can categorize the passengers in the first group as “passengers who travel occasionally” and passengers in group two as “passengers who travel frequently”

15.4 d)

Chart

Description automatically generated

As we can see in this picture the clusters are stable because if we remove a random sample of 5% the structure of the groups remains the same as when we do the analysis with 100% of the data.

15.4 e)

Graphical user interface, text, application

Description automatically generated with medium confidence

It is very similar but not exactly the same because as we can see in the table, there are 914 observations that the model predicted as passengers of group two and they were group one passengers.

15.4 f)

I would focus the offers on group 1 passengers, I understand that group 2 passengers are less flexible with the price so regardless of the price they will travel, however, group 1 passengers are more flexible with the price and therefore if we focus the offers on them, they will travel more.