Step-7 Describing Segments

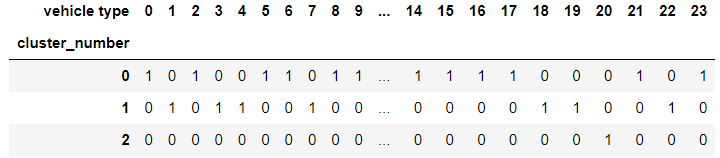
Describing segments is similar to the previous step i.e., Profiling segments. But in step-7, we have to gain information from features that have not been used in the segmentation process. From step-5, we can tell that the following features have been used for segmentation: **max sales volume, market share, Public EV Charging stations** and **Average gross income**. The rest of the features which have not been used for segmentation are: **vehicle type** and **state.**

The variables that will be used for describing segments are called descriptor variables. Let us consider **vehicle type and state** features as **descriptor variables** for this step. Both of the descriptor variables are categorical variables. Hence, we can use mosaic plot to visualize information on these variables.

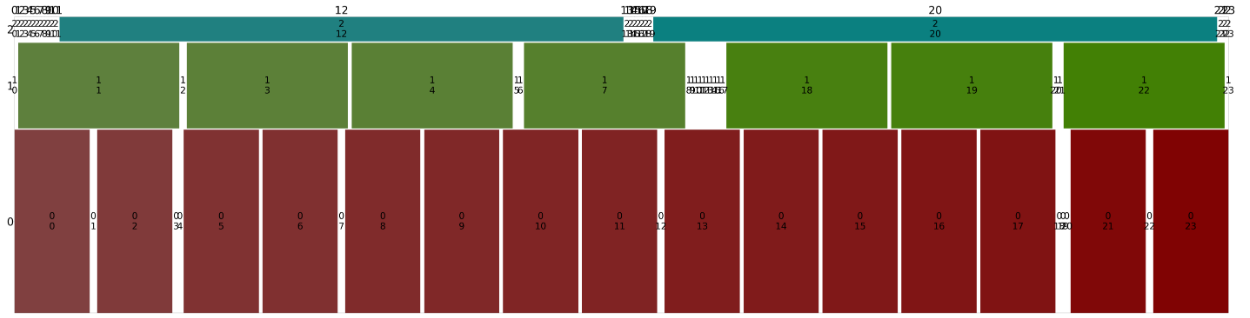
We are using **K-means algorithm** to group our data into clusters. After plotting the WCSS v/s number of clusters graph, we can understand that from 3 number of clusters, the graph declines to a straight line. We will use crosstab and mosaic functions to compare the descriptor variables and cluster number.

\*The crosstab and plots are displayed in the next page\*

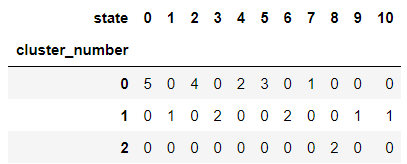
Cross-Tab for Cluster Number and Vehicle Type



Mosaic Plot for Cluster Number and Vehicle Type



Cross-Tab for Cluster Number and State



Mosaic Plot for Cluster Number and State

