The data set is given there providing different parameters and the price in different places on different dates. The fact that price of homes is increasing over time. The average value of the price according to the past data was calculated in the code. We can plot the data using different tools in python. We can see that the maximum average value is in CA(California). The average value here crossed 600000. It can be seen that not only CA but some different states also crossed the value. And the price will increase in the future.

The RegionID column is removed as we can identify it by seeing the StateName only. Similarly, the less important columns were removed already to make the csv more clear to understand. We can predict the future value by using the previous data. For that reason, we have to use some machine-learning models here.

To make a better data set we must clear the NaN values. This work was also done there.

➢ **State Name and average value plot**
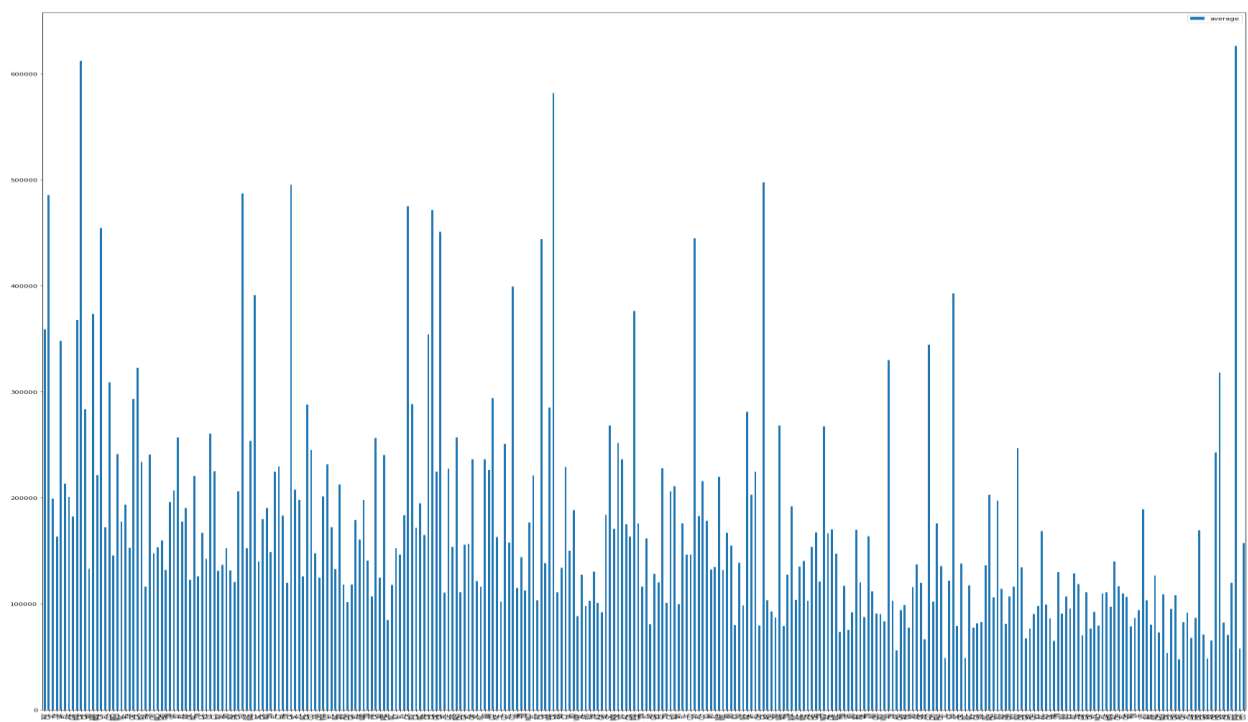


• For a clear view we used a bar plot here. As there are so many stats names exist so the x-axis tricks are difficult to read.

➢ **Scatter plot**