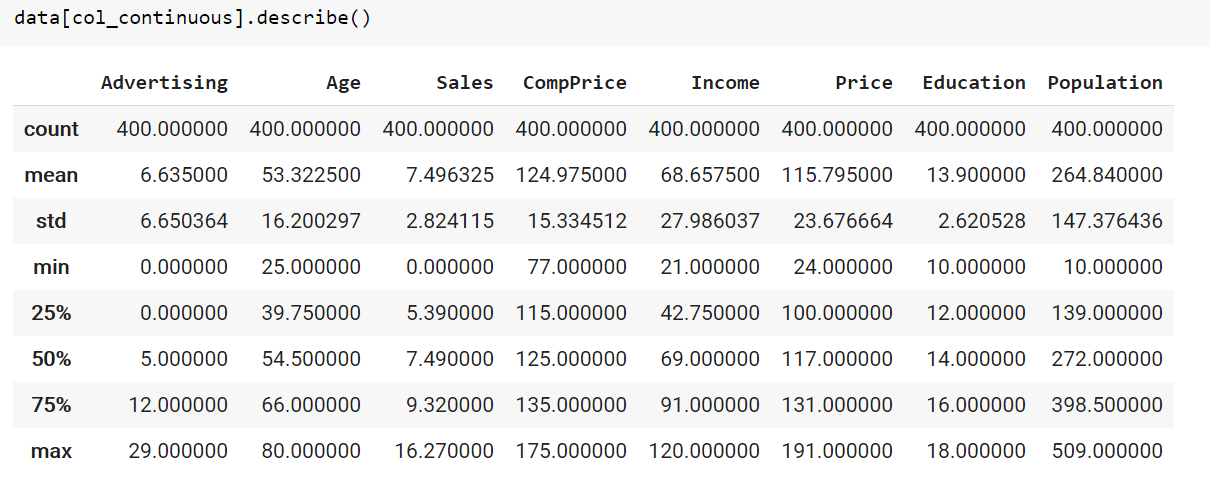
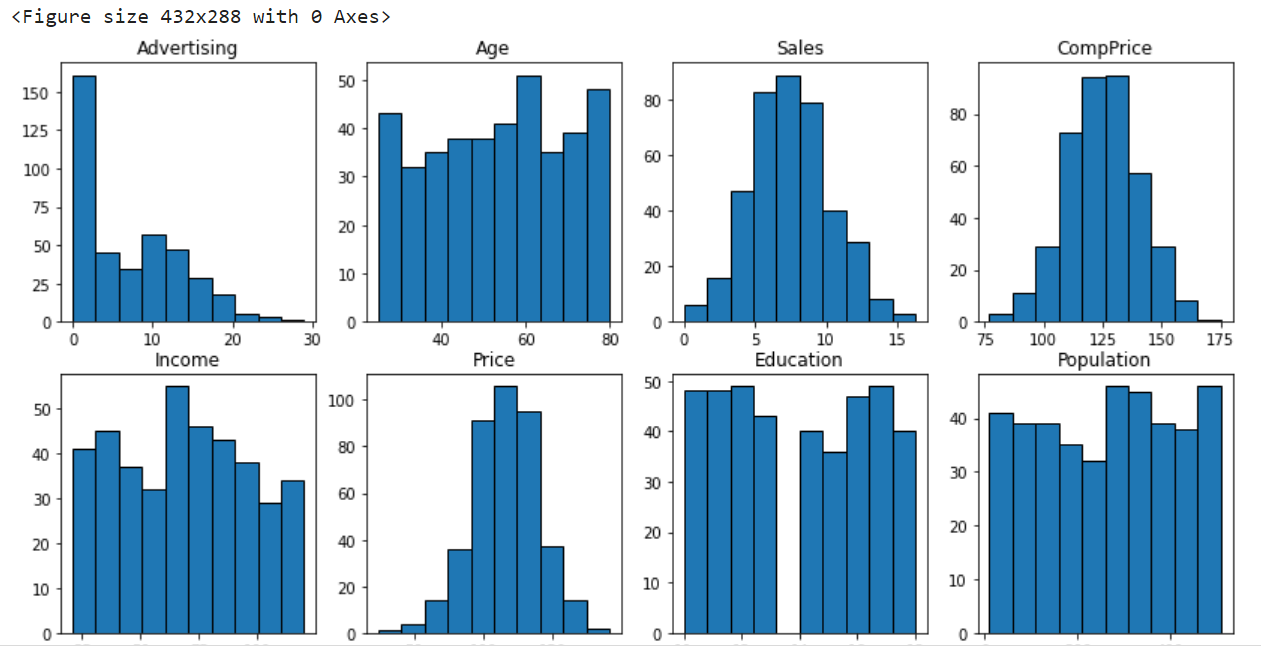
**Data statistics:**

In this project, first we analyze this data using statistics method like the variable and then we try to visualize the statistics in histogram.

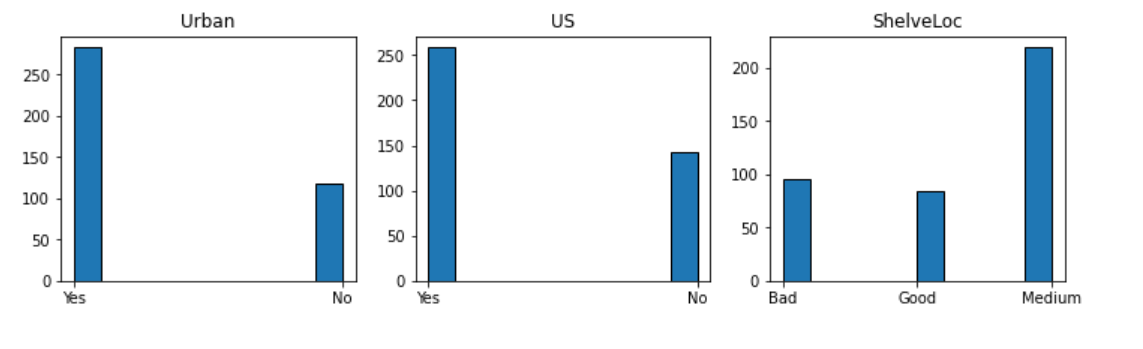


First, we have a brief impression of this data.

Then we get the histogram.



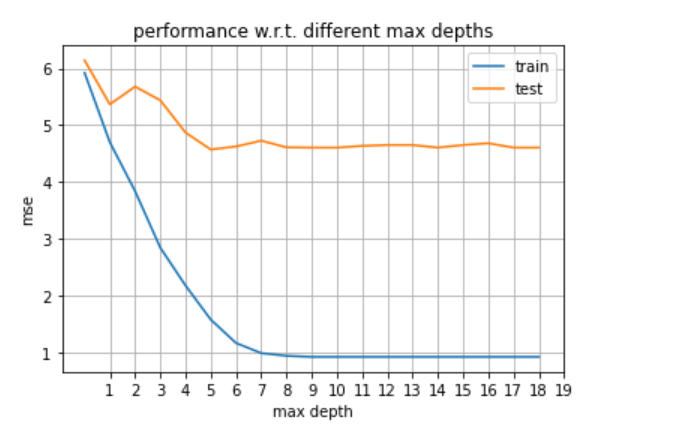
Because the following column is string type, so we plot it separately and change it into integer.



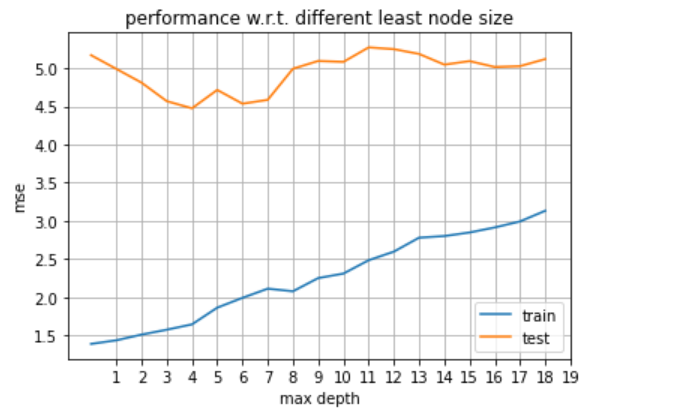
**Decision tree:**

Then we start to analyze these data using Decision Tree.

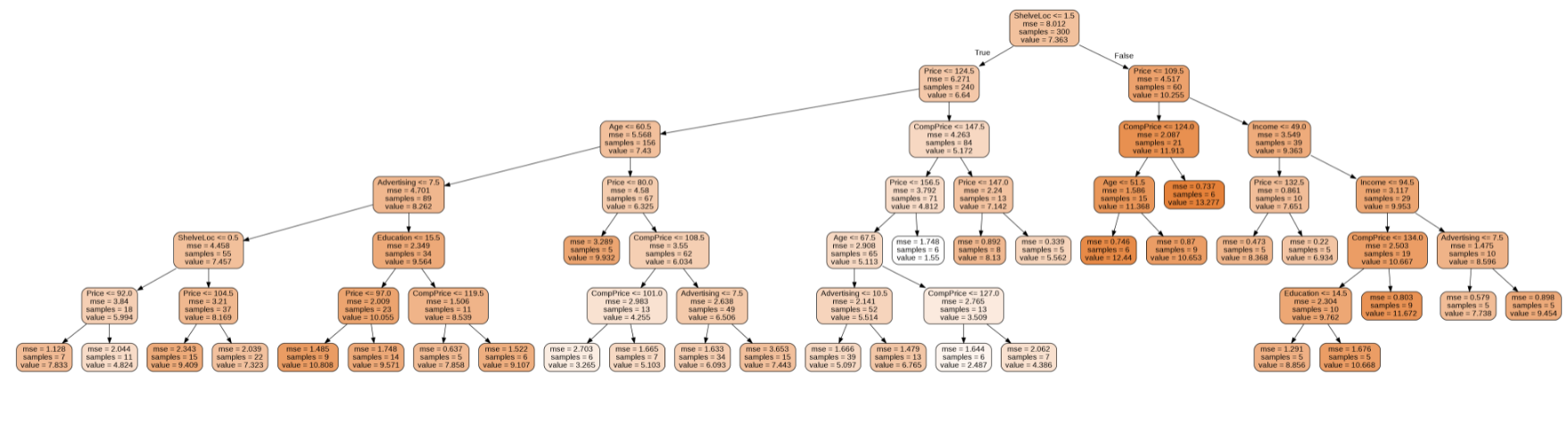
First, we train to find the best max depth. And we can see it is 6.



Next, we are going to find best least node size and we can see from following plot, it is 5.



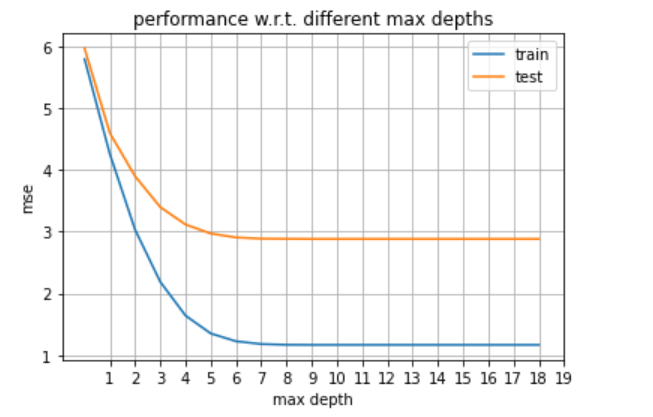
In the end, we can plot the tree:



**Bagging of trees:**

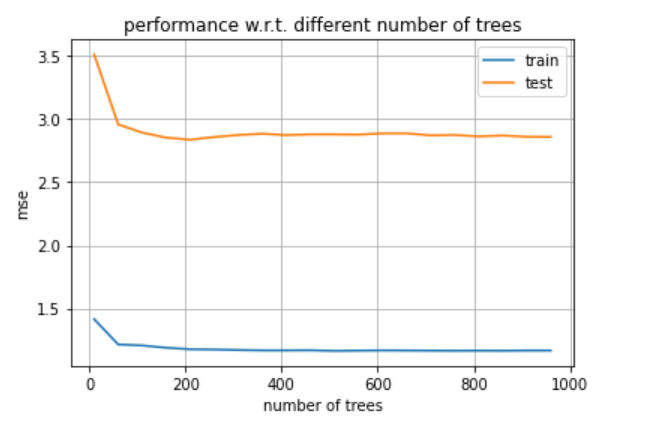
In bagging, we just need to find best max depth and best number of trees.

First:



The best depth is 10.

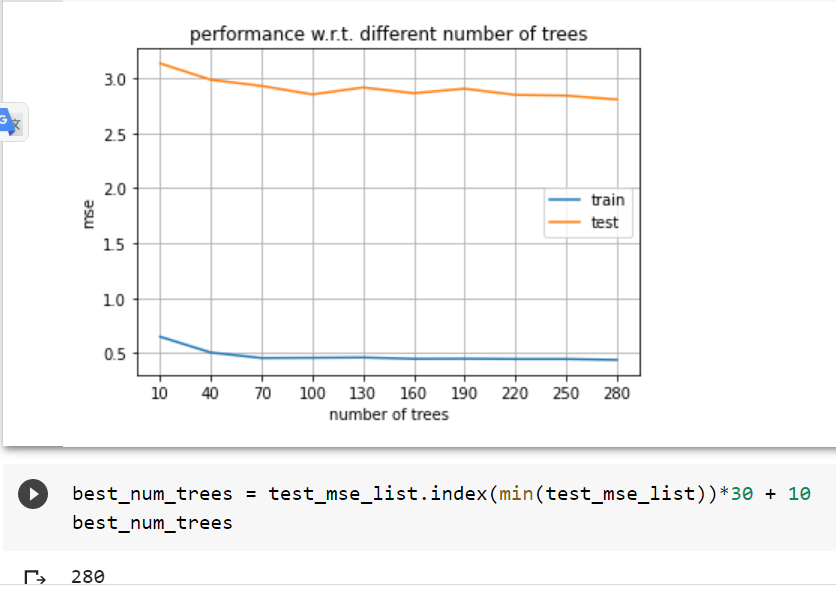
Then:



The best number of trees is 210.

**Random forests:**

Find best number of trees: 280

****

**AdaBoost:**

**Analyze:**