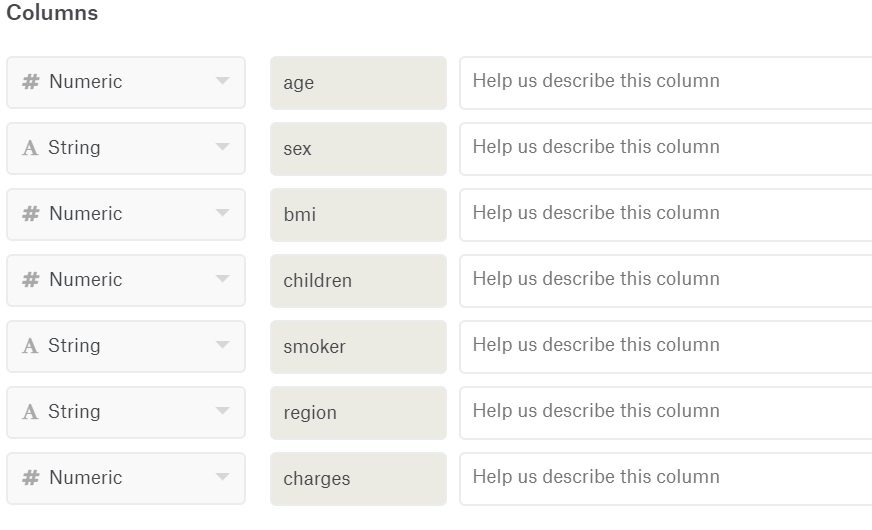
I have created a small POC to create the various dashboards to view the Health Insurance cost predictions based on various parameters like gender(sex), smoker and region.

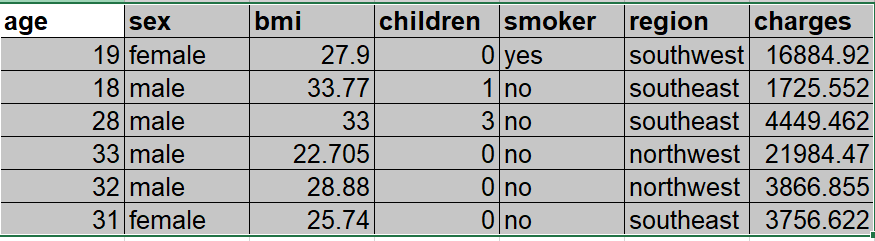
Various graphs are created using Python matplotlib and the data has been derived using the data frames in pandas library.

**About Input Dataset:**

The input data is a csv file containing the below schema.



Below are the sample records from the csv file (insurance.csv)



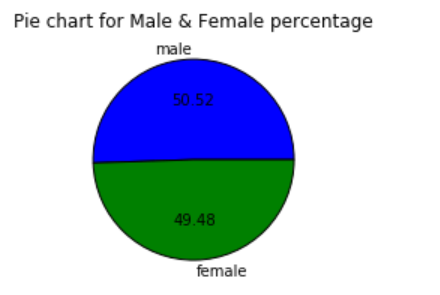
**Analytics:**

For this POC, the data set contains 1338 records and the reference for the same is mentioned at the end.  
Based on this data we are deriving following using Pandas library in python

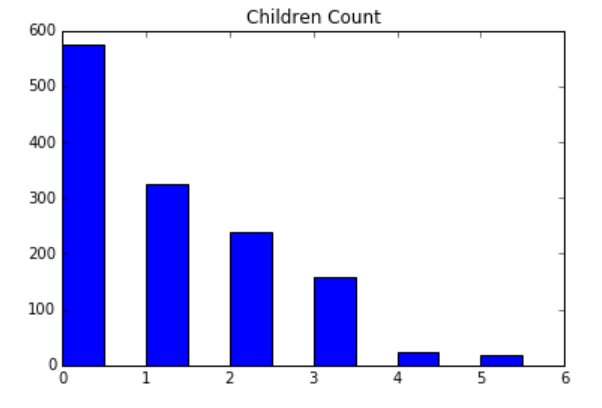
* Calculate the Number of Male and Female percentage on total number of records
* Count the number of children on total number of records
* Calculate the number of smoker and non-smoker percentage on total number of records
* BMI values calculated on dataset
* Age groups are counted on dataset
* Calculate the percentage based on the region values available in dataset
* Calculate Average Charges based on Region, Sex and Smoker

Below are the plots derived using the python library matplotlib

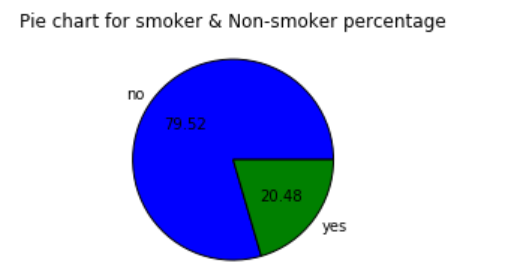
* **Percentage of Male and Female**:



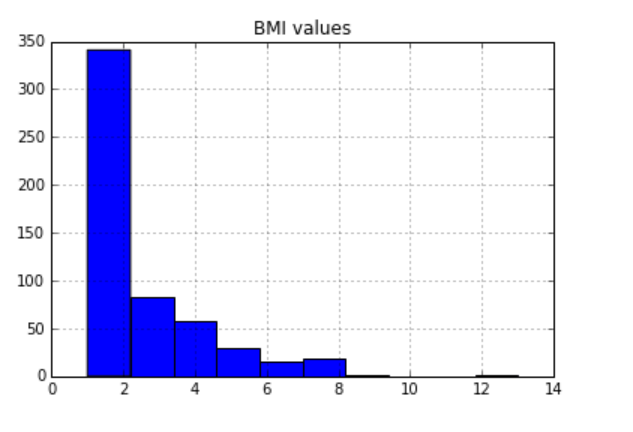
* **Number of Children**



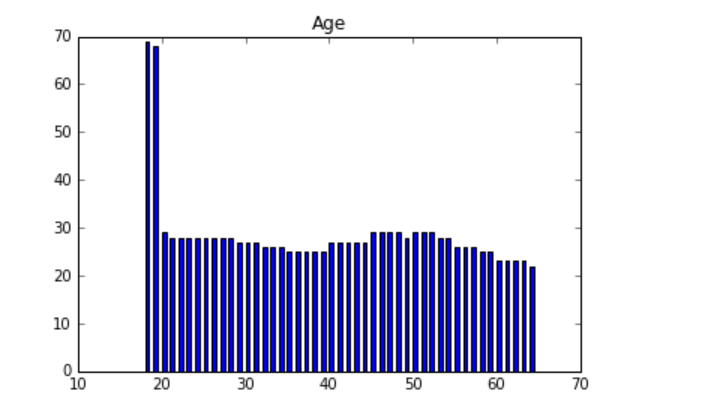
* **Percentage of Smoker and Non-smoker**



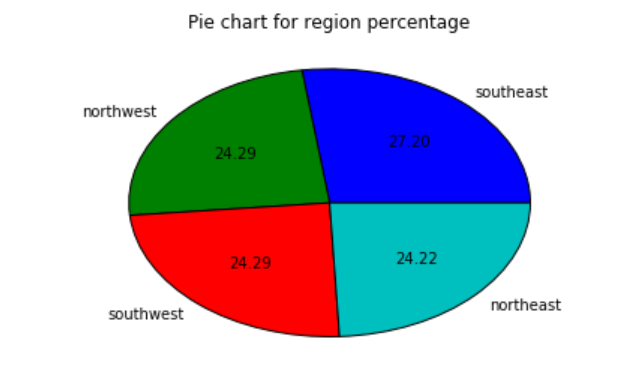
* **Histogram for BMI values**



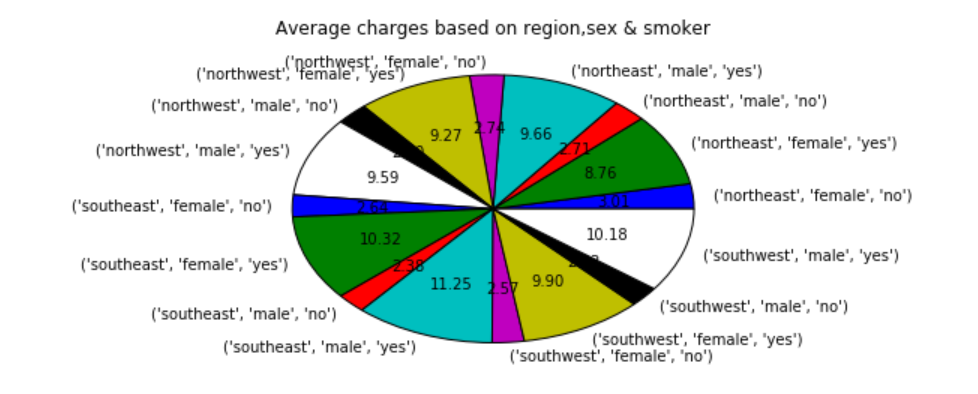
* **Customer Age chart**



* **Percentage of each region**



* **Average Charges based on Region, Sex and Smoker**



References:

Data set used: <https://www.kaggle.com/annetxu/health-insurance-cost-prediction/>