**UNDERSTANDING OF THE BRAND AND MARKET SITUATION**

**Code-AssignEDA.R**

The given data had several columns such as mentioned in Fig 1 except the column other brand and unbranded. The values of other brand and unbranded are calculated first to get an understanding in the data and cleaned the data to only required parts.

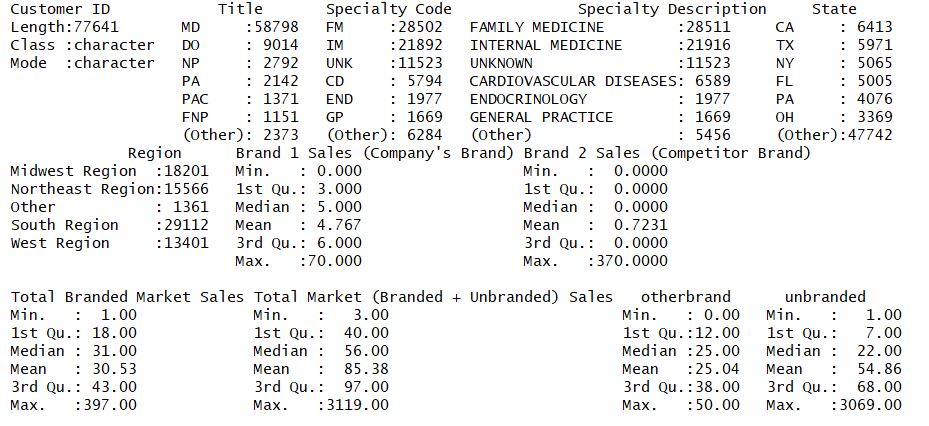


Figure 1: the data used for plotting graphs

Once those two columns where generated. Then calculated the sum of all the required columns such as brand1, brand2, other brands and unbranded and then represented in the form of pie chart and bar graphs

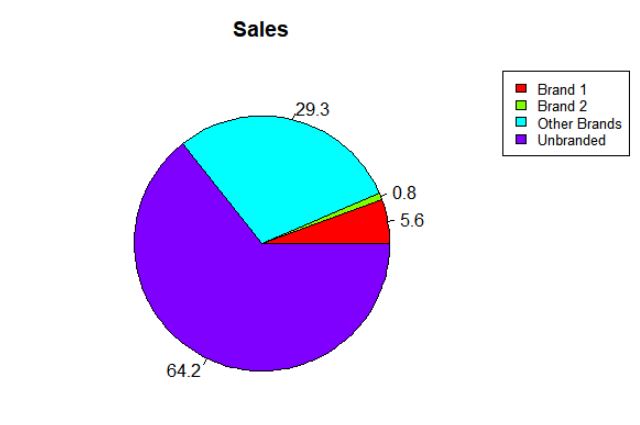


Figure 2: Pie Chart

Figure 2 shows the pie chart of the given sales data. In this they show percentage of each category of items sold and we can do a overall comparison based on the percentages.

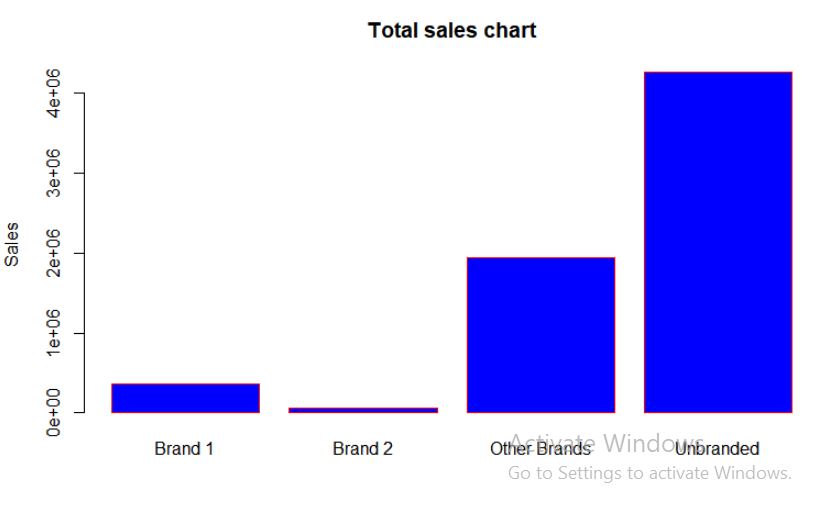


Figure 3: Bar graph

Figure 3 shows the Bar graph of the given sales data. In this they represent each category as a bar and we can easily compare the values based on height of bars.

In the given scenario it is noted that the Brand 1 is having more sales than brand 2(Brand 1 compitator) by 4.8% of the total sale but it is noticed that it is having very less compared to other brands ie more than 20% of the sale. It is also noticed than the unbranded sales is more than branded sale by 14.2% of the total sale. The same understanding with the values can be understood from the bar grapgh

**EXPLANTION ABOUT THE PREDICTIVE MODEL BUILD**

**Code-predict.py**

The target variable for the given model is Brand 1 sales so it is continuous. So the model can either be Regression or Neural network.  
Here it model is build using linear regression with multiple variants.  
The model is build using sklearn library in python.

Since the data set contains categorical data it is converted to numerical categories using get\_dummy function.

The total given data is spitted into two parts. ie training and testing.

Training is 75% and Testing is 25%.

The model is trained using the Training data and then tested using the testing data using score function and produce the result as accuracy.

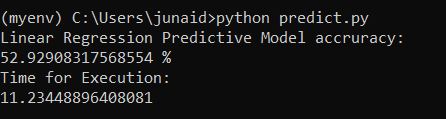


Figure 4:The accuracy and time taken for the predictive model build

Figure 4 shows the accuracy obtained in the model after the predictions are made and also time for execution.