Lab 03 - IE 6200 – Sec 07 – Omkar Ashok Kalange

Task 1 :

health <- read.csv('Health\_Dataset.csv',header=TRUE, sep=',')

>

> health\_df=data.frame(health)

Output –

Age Avg\_Glucose\_Level BMI Stroke Hypertension Heart\_Diseases

1 18.00 107.38 31.40000 FALSE FALSE FALSE

2 32.00 67.19 31.20000 FALSE FALSE FALSE

3 20.00 83.45 23.60000 FALSE FALSE FALSE

4 60.00 102.77 29.70000 FALSE FALSE FALSE

5 61.00 96.07 24.40000 FALSE FALSE FALSE

6 18.00 87.28 25.70000 FALSE FALSE FALSE

Task 2 :

underwt\_num

[1] 64

> nrow(normal)

[1] 244

> nrow(obese)

[1] 429

> nrow(overwt)

[1] 263

nrow(normal\_hy)

[1] 11

> nrow(overwt\_hy)

[1] 32

Barplot shows that overweight people have high chance of suffering from

Hypertension rather than normal people.

mean\_nor

[1] 22.14508

> median\_nor

[1] 22.45

> range\_nor

[1] 6.4

> int\_quanrange

75%

3.325

> var\_nor

[1] 3.701416

> sd\_nor

[1] 1.923906

Task 3 :

Calculated percentages of each pair of stroke and hypertension. Created a data frame for the same and using cbind() function, added an extra column to the existing pair of stroke and hypertension.

Output :

Stroke Hypertension Percentage

1 TRUE TRUE 0.3

2 TRUE FALSE 1.0

3 FALSE TRUE 8.4

4 FALSE FALSE 90.3

Task 4 :

a.

Coefficient of variance

cov\_gluc

[1] 41.92766%

Skewness

skew\_gluc

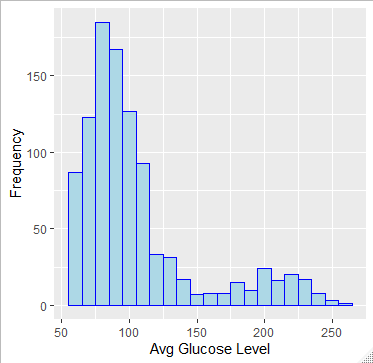
[1] 1.637986

Kurtosis

> kur\_gluc

[1] 1.925254

b. Histogram is shown below :

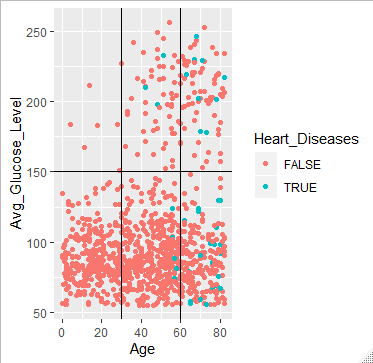


c.

Histogram gives valuable results regarding the skewness of the data. As we obtained skewness from part a) that it is 1.63 which means the curve is positively skewed which suggests that mean lies to the right side of median and mode lies to the left side of median.

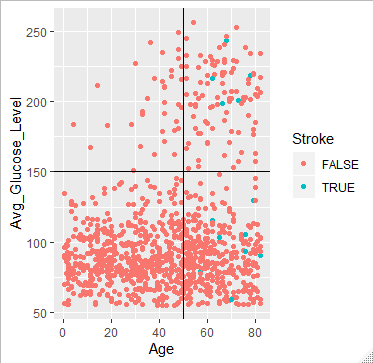
Task 5

We can infer that people having age less than 30 do not suffer from heart disease and people with more than 60 age have high chance of suffering from heart diseases.



Task 6

1.

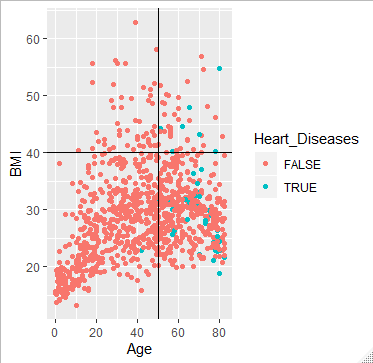


From the analysis of Stroke parameter with Age and Glucose levels, we come up with the conclusion that people with more than 50 age have high chances of getting stroke. Also if their glucose level is below 150, the chances are even more.

2.

Analysis of Age, BMI with the effect of Heart Diseases –

Below is the scatter plot for the same. We can deduce some meaningful observations from the same-



1. People under 50 with BMI more than 40 suffer from no heart diseases.
2. People above 50 with BMI less than 40 suffer a lot from heart diseases.
3. Also people above 50 and BMI greater than 40 have high chances of suffering from heart diseases.
4. People less than 50 seldom have heart diseases if their BMI is less than 40.