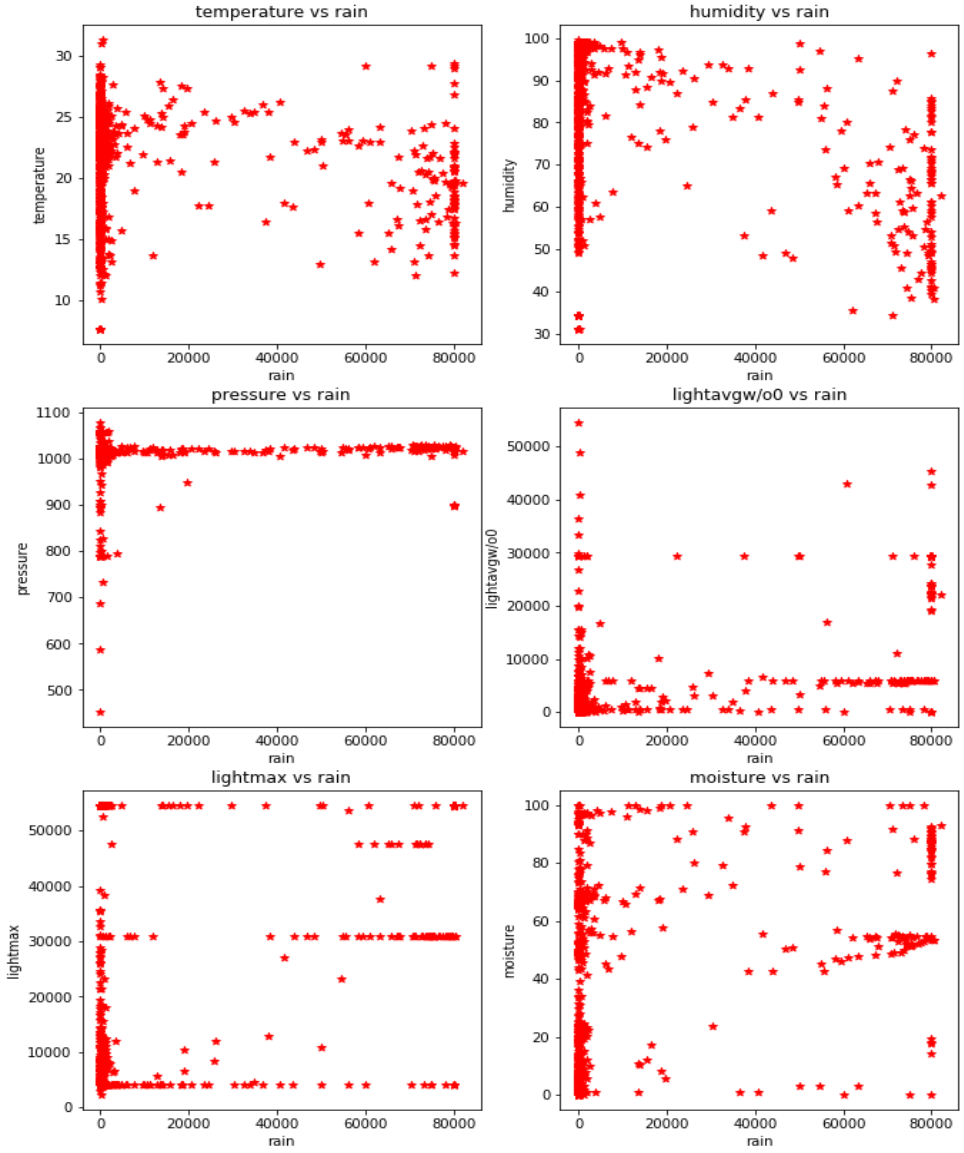
Name: Anooshka Bajaj

1:

	temperature (Celsius)	humidity (%)	pressure (mb)	rain (ml)	lightavgw/o0 (lux)	lightmax (lux)	moisture (%)
mean	21.214888	83.479932	1009.008774	10701.538370	4438.428453	21788.623280	32.386053
median	22.272730	91.380950	1014.677832	18.000000	1656.880000	6634.000000	16.704200
mode	12.72727	99.0	789.392692	0.0	4488.9103	4000	0.0
min	7.672900	31.000000	452.097887	0.000000	0.000000	2259.000000	0.000000
max	31.375000	99.720000	1079.162000	82037.250000	54612.000000	54612.000000	100.000000
std	4.355818	18.210065	46.980477	24852.255288	7573.162806	22064.993089	33.653245

2(a):



Scatter Plots provide a visual representation of correlation between the two variables. If the data show an uphill pattern this indicates a positive relationship and if the data show a downhill pattern this indicates a negative relationship. A line of best fit is a straight line that best represents the data on a scatter plot. The closer the points are to the best fit line more is the correlation coefficient.

Thus,

(i) Temperature and Rain have low negative correlation.
Rainfall is either very low or very high. Average rainfall is not much seen.

(ii) Humidity and Rain have comparatively higher negative correlation.

(iii) Pressure and Rain have very little (close to zero) correlation i.e. pressure and rain are almost independent of each other.

(iv) Lightavgw/o and Rain have positive correlation. For less average light throughout the day, rainfall is usually low but it can be high also.

(v) Lightmax and Rain have very little correlation.

(vi) Moisture and Rain have low positive correlation.

2(b):

(Scatter Plots on next page)

Inference:

(i) Humidity and Temperature have positive correlation.

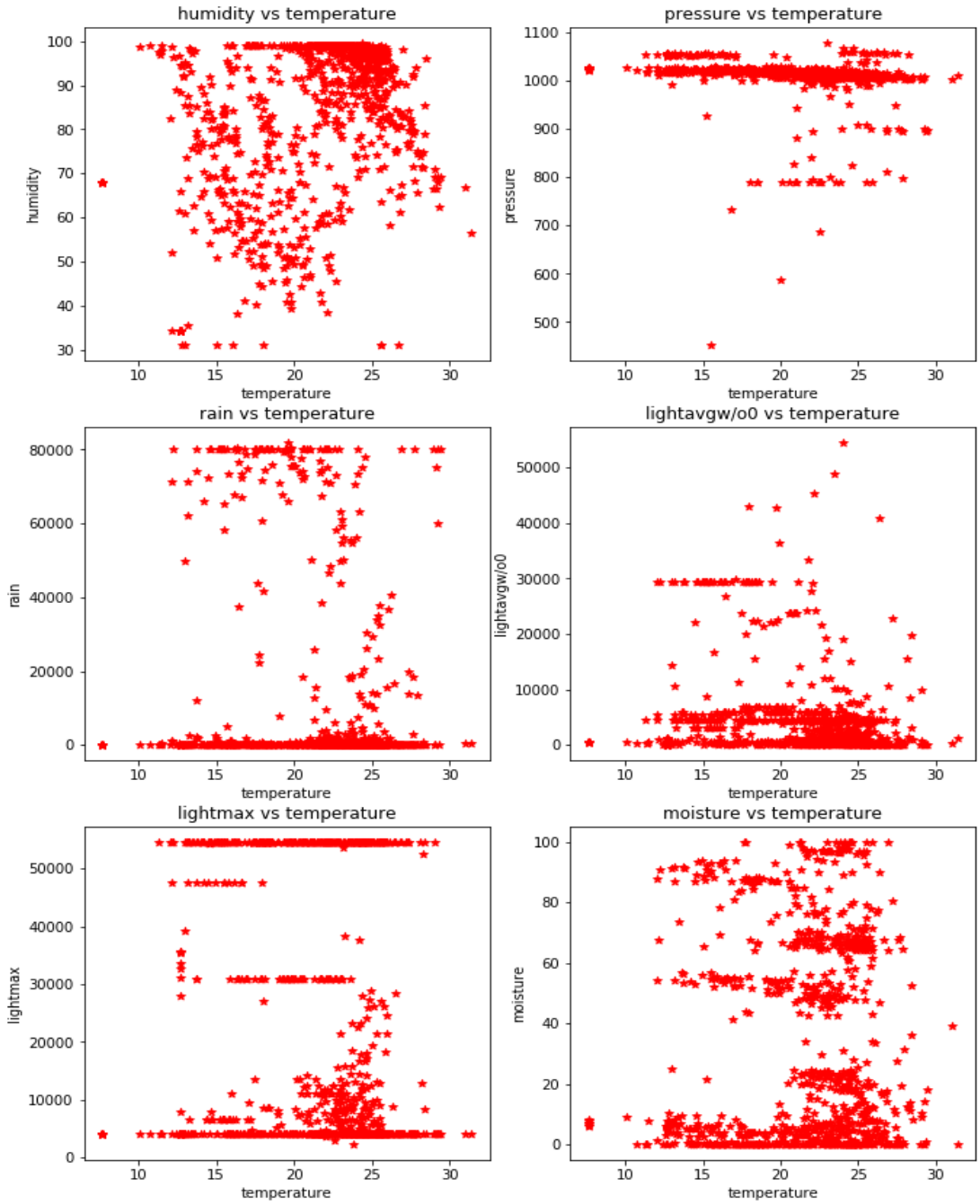
(ii) Pressure and Temperature have very low negative correlation.

(iii) Rain and Temperature have low negative correlation.

(iv) Lightavgw/o and Temperature have low negative correlation.

(v) Lightmax and Temperature have low negative correlation.

(vi) Moisture and Temperature have very low positive correlation (they are almost independent of each other).



3(a):

Correlation Coeffecient of rain with:

temperature -0.108893

humidity -0.434917

pressure 0.070785

lightavgw/o0 0.527490

lightmax 0.312843

moisture 0.426928

3(b):

Correlation Coefficient of temperature with:

humidity 0.401570

pressure -0.181389

rain -0.108893

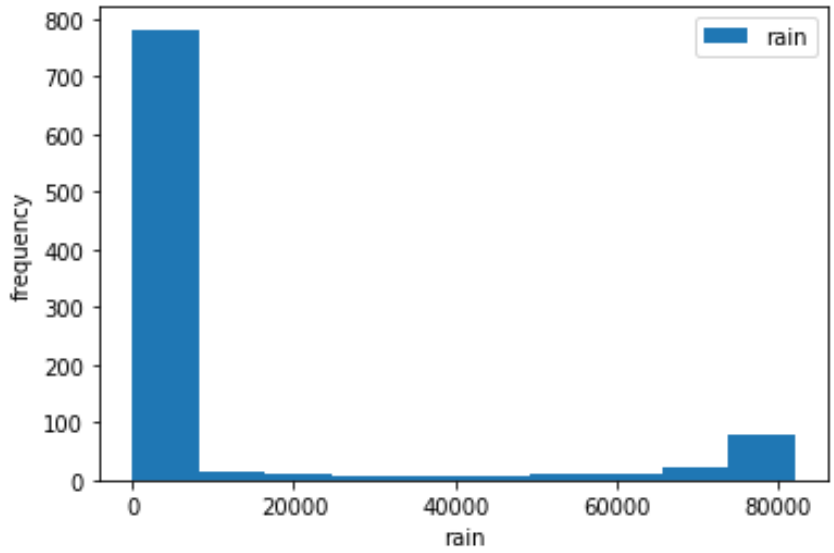
lightavgw/o0 -0.181400

lightmax -0.145884

moisture 0.080660

The inference from scatter plots in ans 2(a) and 2(b) are consistent with the values of correlation coefficient measured in ans 3(a) and 3(b). Positive correlation means that when one variable increases, the other variable increases and vice versa. Negative correlation means that when one variable increases, the other variable decreases and vice versa.

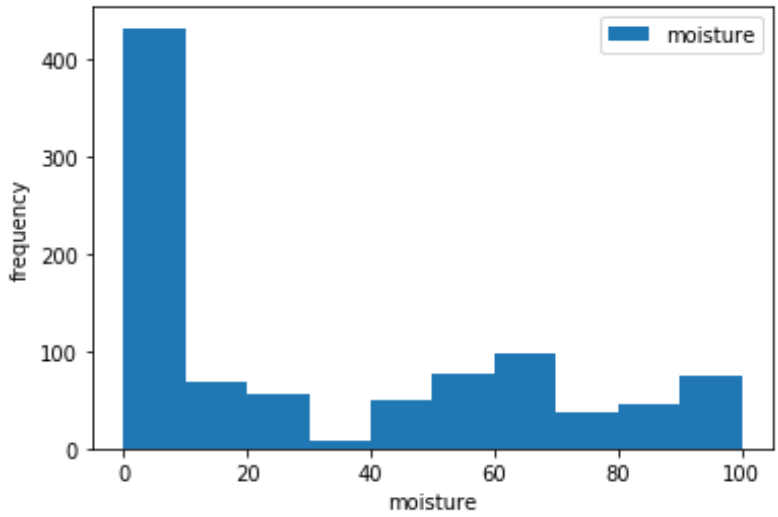
4: Histogram of Rain (measured in ml):



Observations:

Rainfall less than 10000ml is most favourable and high rainfall of around 80000ml also has a possibility of happening. Rainfall around 10000ml to 70000ml happens rarely.

Histogram of Moisture (measured between 0 to 100 percent)



Observations:

Moisture content is mostly less than 10%. Moisture content around 30% to 40% is very rare.

5:

Observations:

(i) For stationoid t10: less rainfall is more likely to happen. Most time rainfall is less than 250ml. Rainfall upto 2000ml happens rarely.

(ii) For stationoid t11: rainfall less than 2500ml is more likely.

(iii) For stationoid t12: similar to stationoid t11. Rainfall less than 2500ml is more likely to happen. Heavy rainfall upto 15000ml happens rarely.

(iv) For stationoid t13: rainfall less than 10000ml most likely. Heavy rainfall upto 80000ml happens rarely.

(v) For stationoid t14: similar to stationoid t13. Rainfall less than 10000ml is most likely to happen. Heavy rainfall upto 80000ml can also happen but rarely.

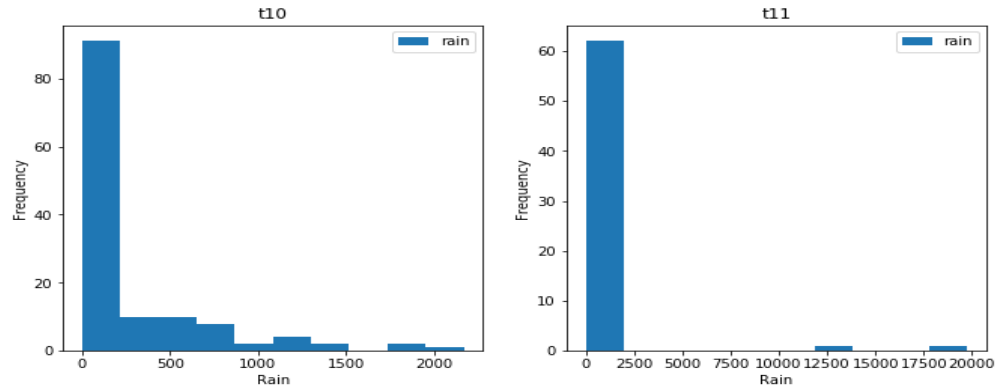
(vi) For stationoid t15: rainfall less than 250ml is most likely. Rainfall around 500ml is also possible but rainfall upto 1750ml happens rarely.

(vii) For stationoid t6: rainfall less than 250ml most likely. Rainfall upto 1750ml happens rarely.

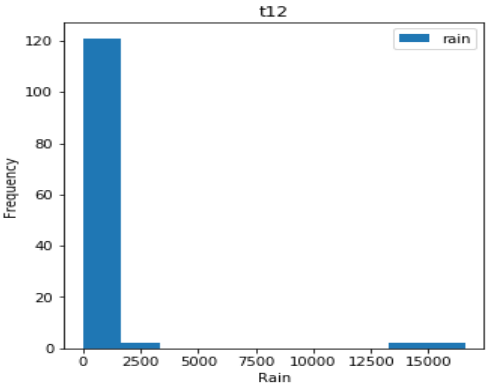
(viii) For stationoid t7: rainfall less than 8000ml most likely. Heavy rainfall upto 60000ml happens rarely.

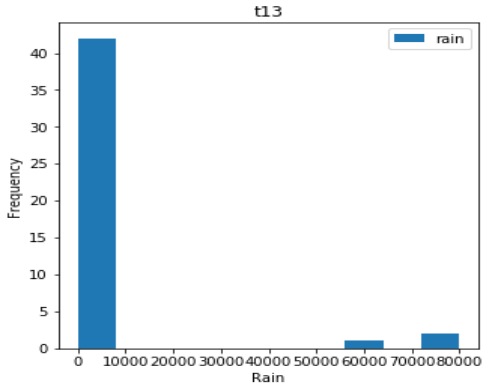
(ix) For stationoid t8: rainfall less than 8000ml most likely to happen. Heavy rainfall around 80000ml also favourable.

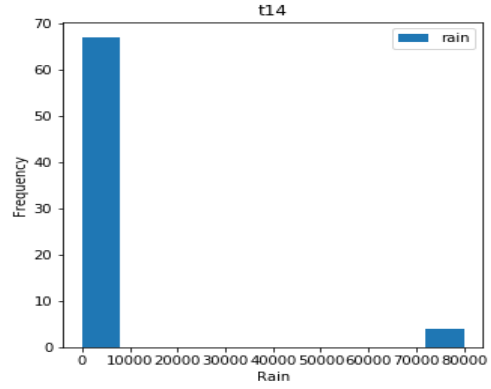
(x) For stationoid t9: rainfall less than 8000ml most favourable. Heavy rainfall of around 72000ml to 80000ml also favourable. Rainfall of range around 8000ml to 72000ml has lower possibility of happening.

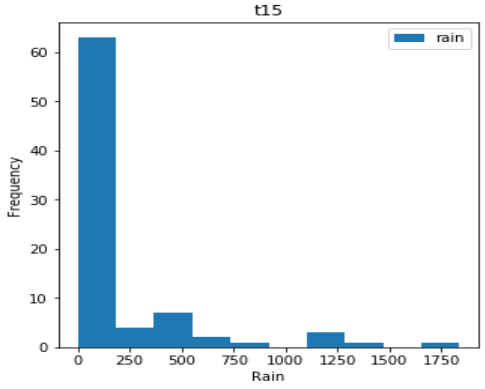

t10
rain
Frequency
Rain
0 20 40 60 80
0 500 1000 1500 2000

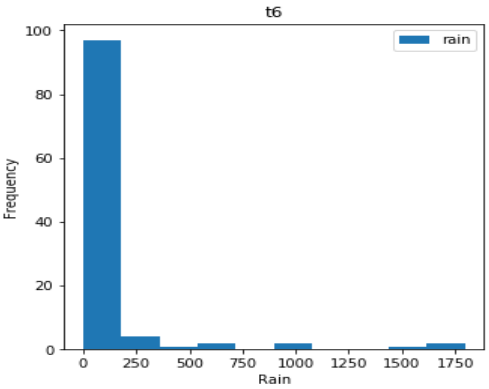
t11
rain
Frequency
Rain
0 10 20 30 40 50 60
0 2500 5000 7500 10000 12500 15000 17500 20000

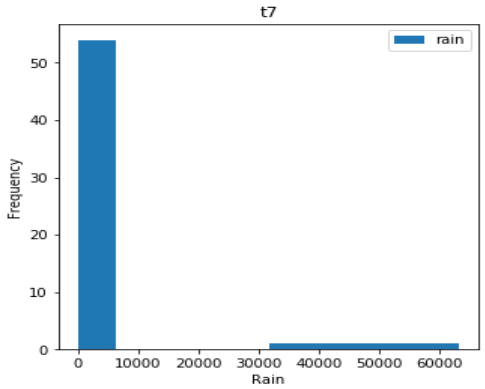

t12
rain
Frequency
Rain
0 20 40 60 80 100 120
0 2500 5000 7500 10000 12500 15000

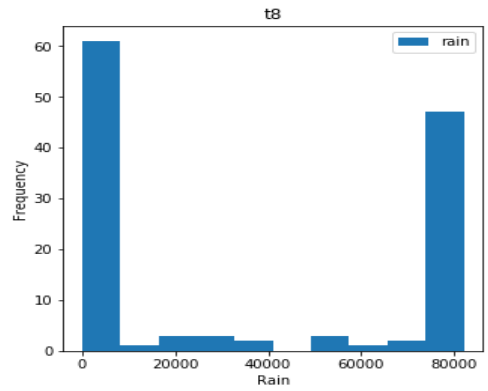

t13
rain
Frequency
Rain
0 5 10 15 20 25 30 35 40
0 10000 20000 30000 40000 50000 60000 70000 80000

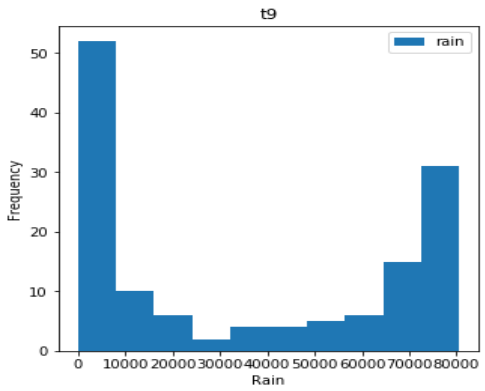

t14
rain
Frequency
Rain
0 10 20 30 40 50 60 70
0 10000 20000 30000 40000 50000 60000 70000 80000


t15
rain
Frequency
Rain
0 10 20 30 40 50 60
0 250 500 750 1000 1250 1500 1750

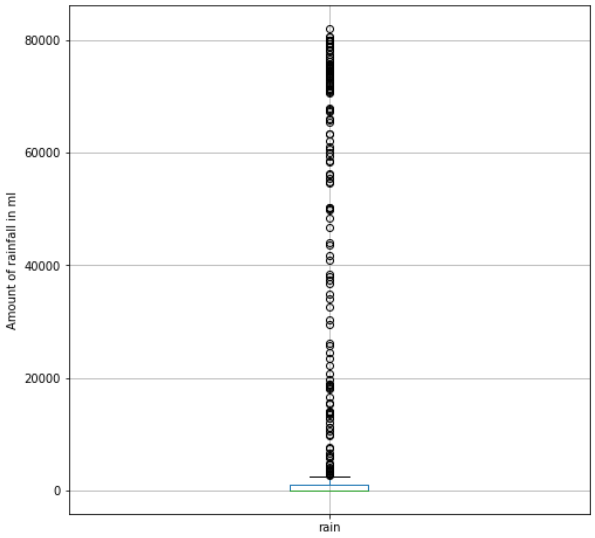

t6
rain
Frequency
Rain
0 20 40 60 80 100
0 250 500 750 1000 1250 1500 1750


t7
rain
Frequency
Rain
0 10 20 30 40 50
0 10000 20000 30000 40000 50000 60000


t8
rain
Frequency
Rain
0 10 20 30 40 50 60
0 20000 40000 60000 80000


t9
rain
Frequency
Rain
0 10 20 30 40 50
0 10000 20000 30000 40000 50000 60000 70000 80000

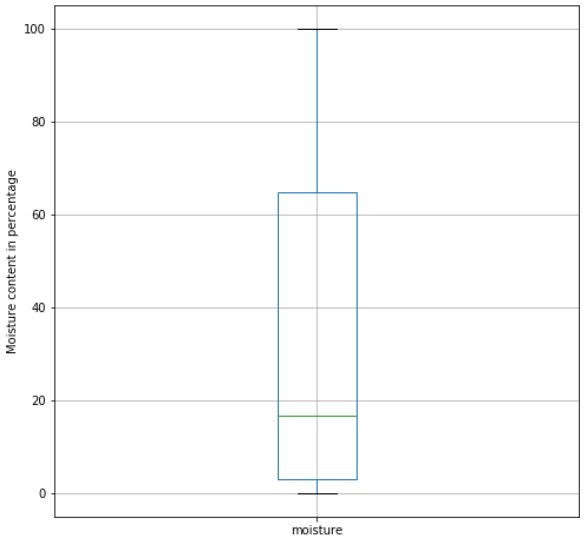
6: Boxplot for rain:



Observations:

- Median is very close to the first quartile
- It has many outliers. Thus, the data collection is not steady.
- There is huge difference between the maximum value and the average value.

Boxplot for moisture:



Observations:

- Median is close to the first quartile.
- It has no outliers. Thus, the data collection is steady i.e. moisture contents are in the interquartile range.