

# Assignment 1

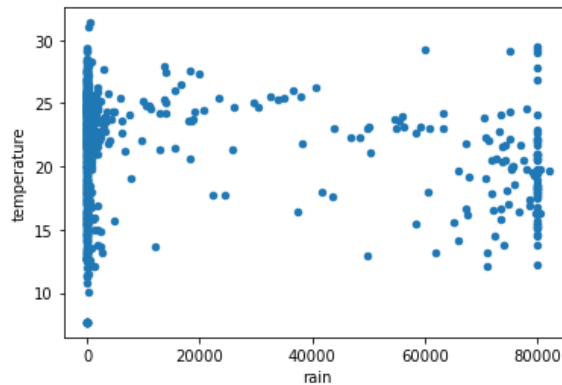
Harsh Arya (B20043)

Ans 1)

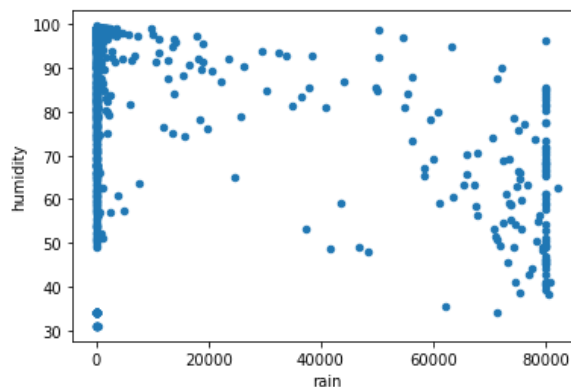
	temperature	humidity	pressure	rain	lightavgw/o0	lightmax	moisture
mean	21.214888	83.479932	1009.008774	10701.538370	4438.428453	21788.623280	32.386053
median	22.272730	91.380950	1014.677832	18.000000	1656.880000	6634.000000	16.704200
mode	12.727270	99.000000	789.392692	0.000000	4488.910300	4000.000000	0.000000
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

Ans 2)

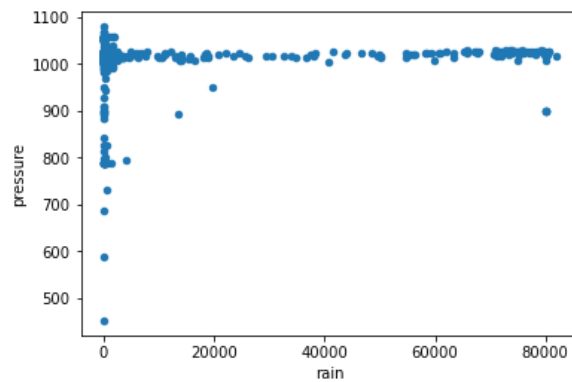
a)



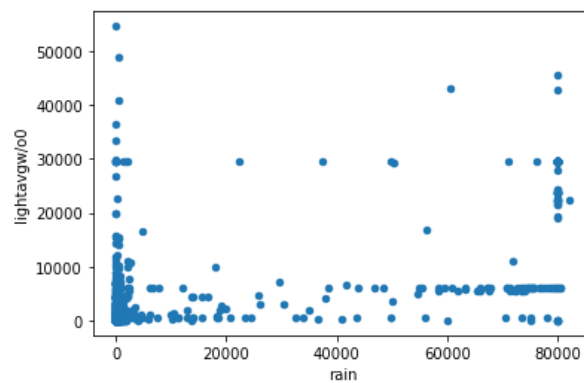
In the above plot there are a lot of outliers and we can see that temperature and rain have a weak negative correlation. Therefore, when we move towards right i.e. more rain, we conclude that temperature decrease slightly.



From the above plot, we can conclude that rain and humidity are negatively correlated, thus when there is increase in rain, humidity decreases.

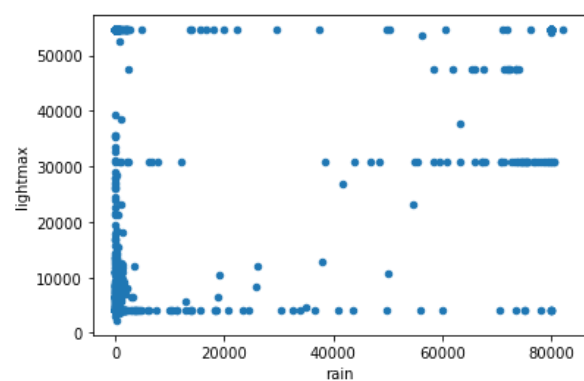


Pressure and rain have almost no correlation, therefore pressure does not change appreciably with the increase in the amount of rain.

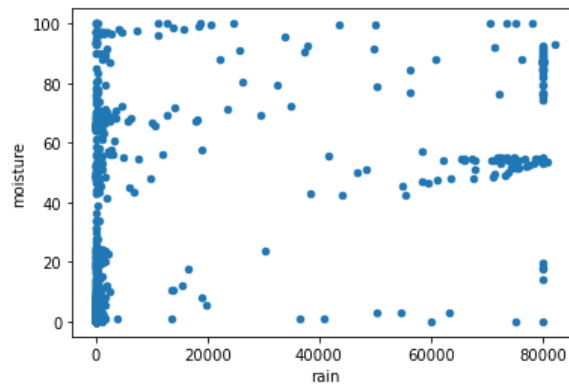


Average light and rain correlation which light increases slightly amount of rain.

have weak positive implies that average with the increase in

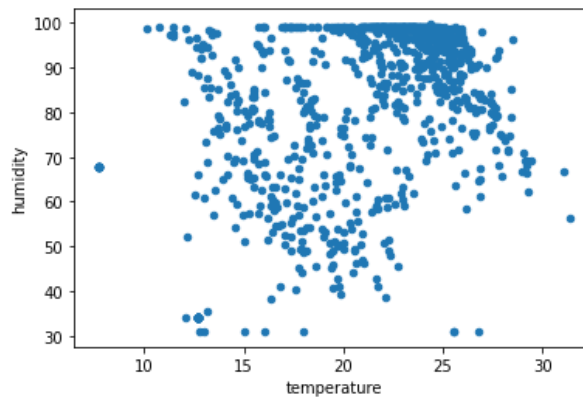


The maximum lux count by sensor and rain have a moderately positive correlation, thus increasing with the increase in the amount of rain.

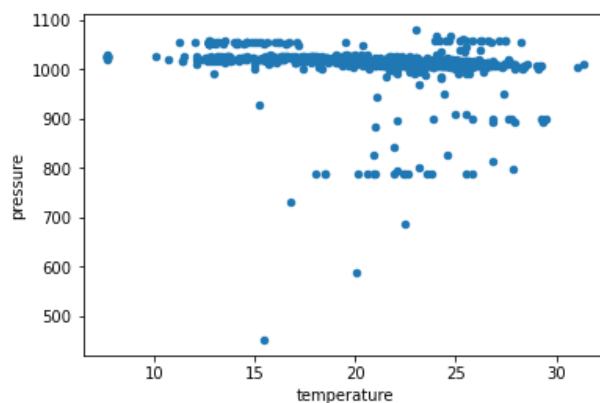


Moisture and rain are also positively correlated; therefore, moisture increases with the increase in the amount of rain.

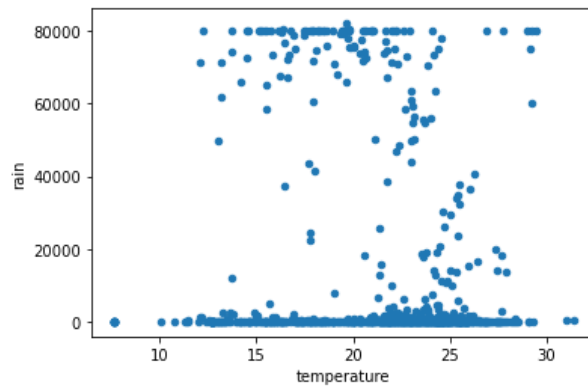
b)



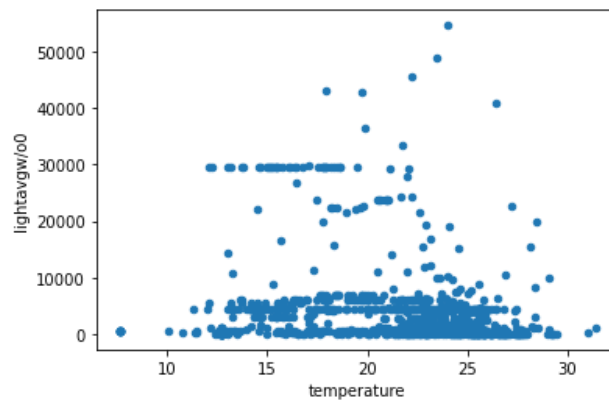
Temperature and humidity are positively correlated with each other. Increase in the temperature implies increase in humidity and vice versa. And the humidity has its peak value in the temperature range of 20 to 30.



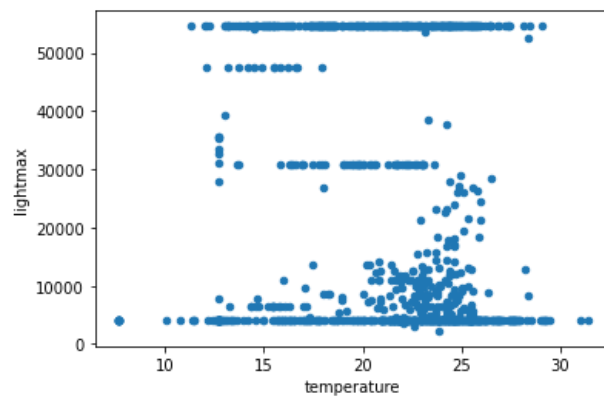
From the above plot we conclude that temperature and pressure have weak correlation. And pressure decreases slowly with the increase in temperature.



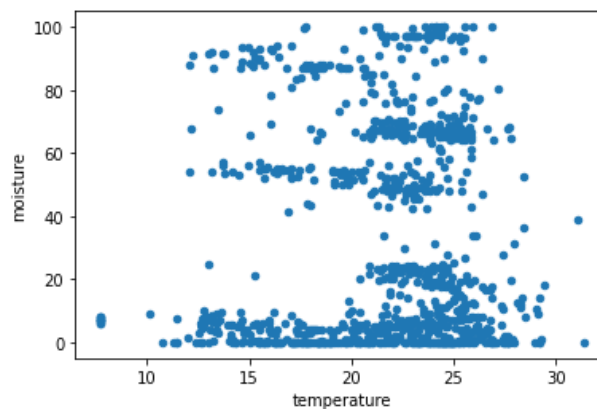
Rain and temperature are weakly correlated, and amount of rain decreases with the increase in temperature.



Average light and temperature are weakly correlated and average light seems to decrease with the increase in temperature.



The maximum lux count by sensor and temperature are weakly correlated and lightmax seems to decrease with the increase in temperature.



Temperature and moisture seem to be uncorrelated with a lot of outlier data points.

### Ans 3)

a)

The correlation of rain with temperature is  $-0.108893$ . Thus, rain and temperature have a weak negative correlation.

The correlation of rain with humidity is  $-0.434917$ . Thus, rain and humidity have strong negative correlation.

The correlation of rain with pressure is  $0.070785$ . Thus, rain and pressure have weak positive correlation.

The correlation of rain with lightavgw/o0 is  $0.527490$ . Thus, rain and lightavgw/o0 has strong positive correlation.

The correlation of rain with lightmax is  $0.312843$ . Thus, rain and lightmax have moderately positive correlation.

The correlation of rain with moisture is  $0.426928$ . Thus, rain and moisture have strong positive correlation.

b)

The correlation of temperature with humidity is  $0.401570$ . Thus, temperature and humidity have strong positive correlation.

The correlation of temperature with pressure is  $-0.181389$ . Thus, temperature and pressure have weak negative correlation.

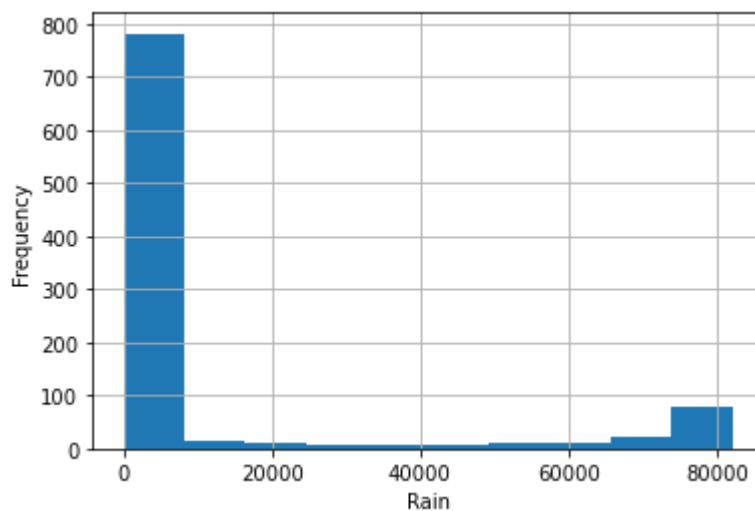
The correlation of temperature with rain is  $-0.108893$ . Thus, temperature and rain have weak negative correlation.

The correlation of temperature with lightavgw/o0 is  $-0.181400$ . Thus, temperature and lightavgw/o0 have weak negative correlation.

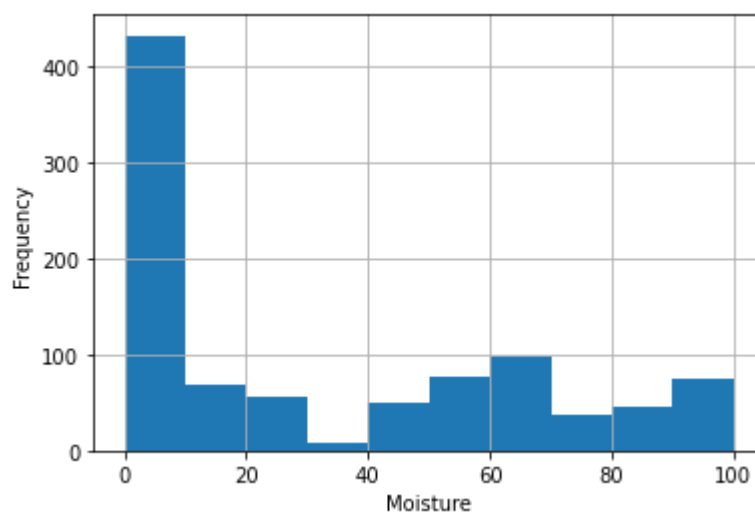
The correlation of temperature with lightmax is -0.145884. Thus, temperature and lightmax have weak negative correlation.

The correlation of temperature with moisture is 0.080660. Thus, temperature and moisture have weak positive correlation.

**Ans 4)**

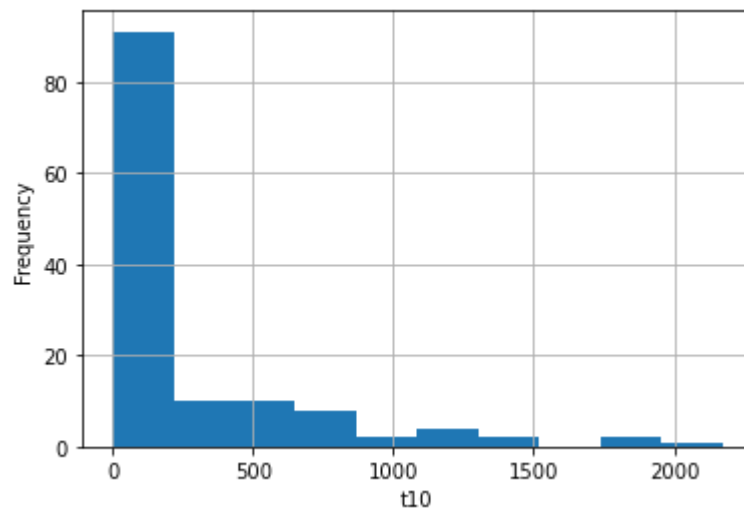


In the above histogram, most of the data points lie in the bin 0 – 8200 ml. Thus, there is a high probability that the amount of rain on a day is between 0 – 8200 ml and there is very less frequency of rain amount 20000-60000ml.

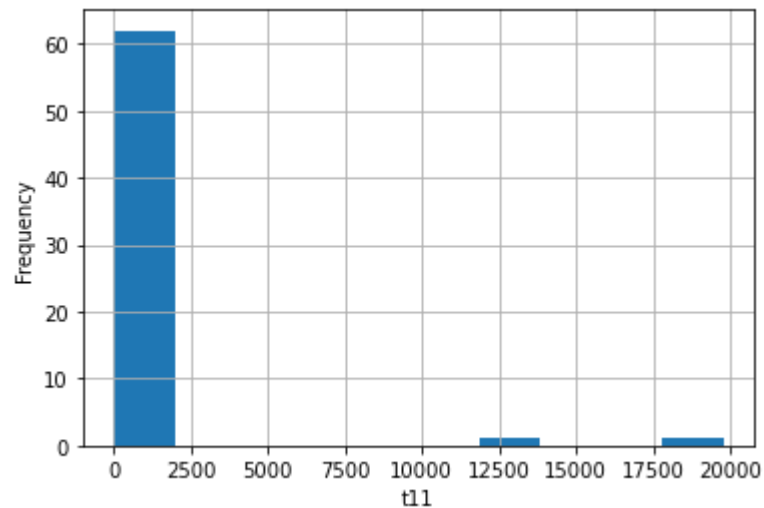


Moisture content has a good probability of being in the range 0-10 percent.

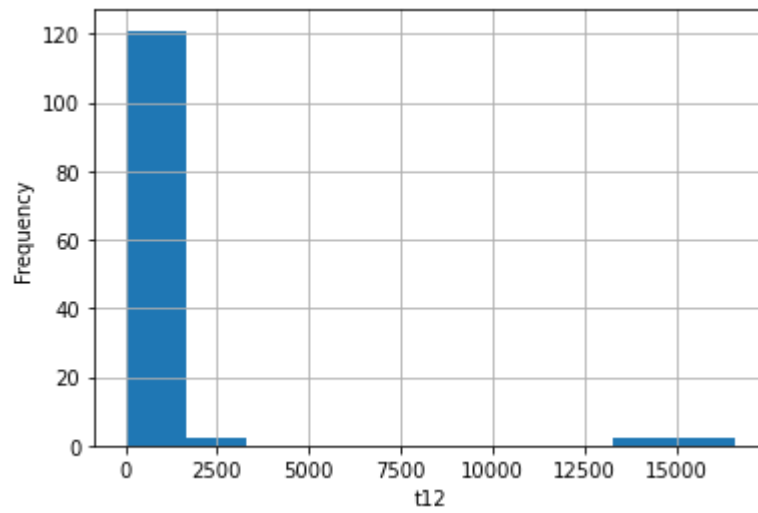
**Ans 5)**



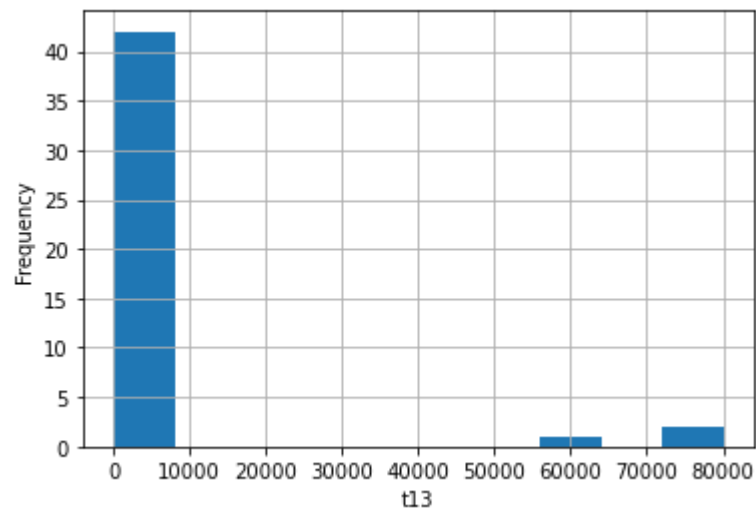
The amount of rain near t10 station has a good probability of being in the range 0 – 217ml i.e. we have less rain at this station and the probability decreases as the amount of rain increase.



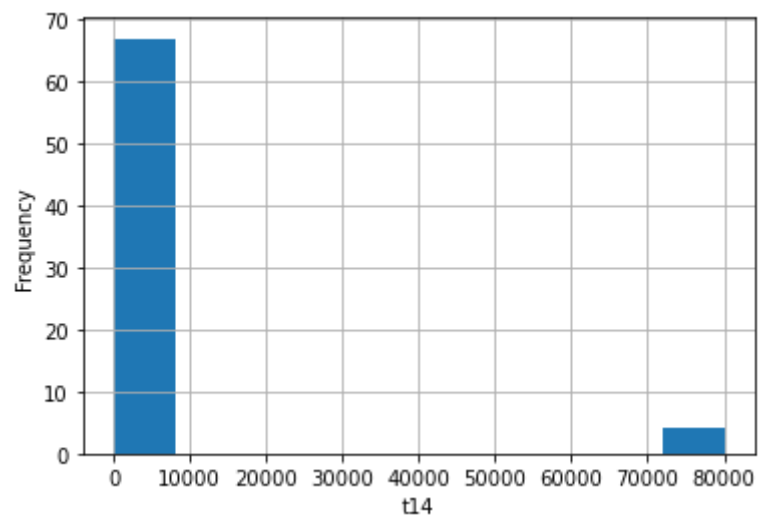
The amount of rain near station t11 is almost lies in the range 0 – 2000 ml and has almost zero rain frequency after this bin.



The amount of rain near station t12 has a high probability of being in the range 0 – 1657 ml.

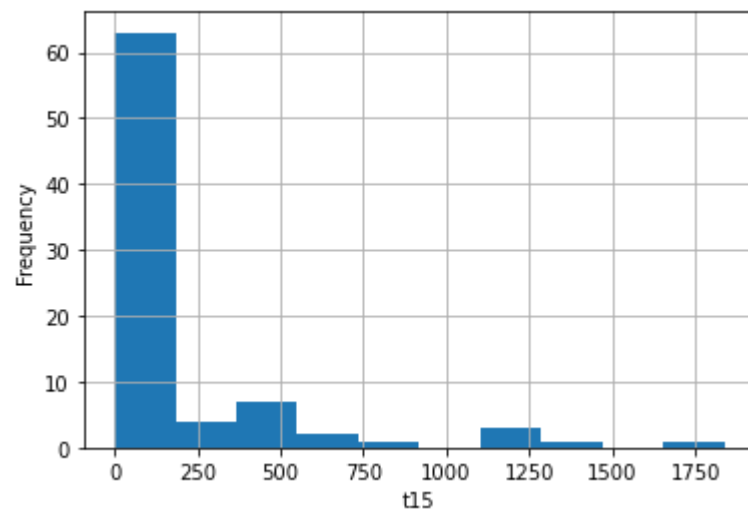


The amount of rain near station no. t13 has a high probability of being in the range 0 – 8000 ml and the probability decreases as the amount of rain increase.

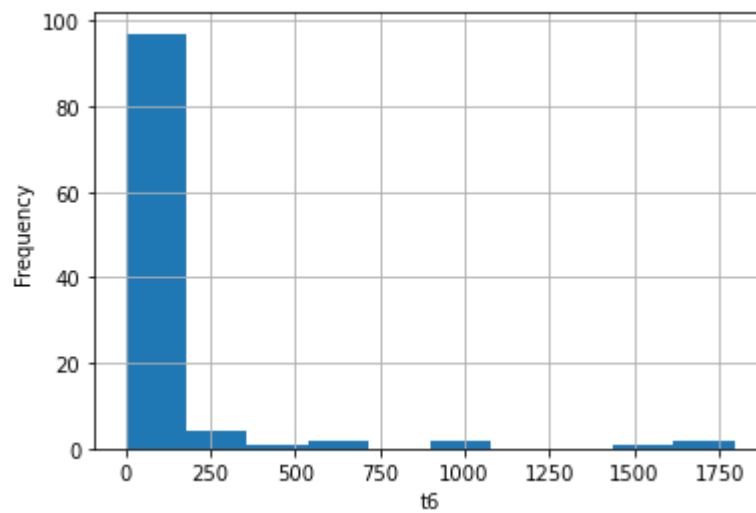




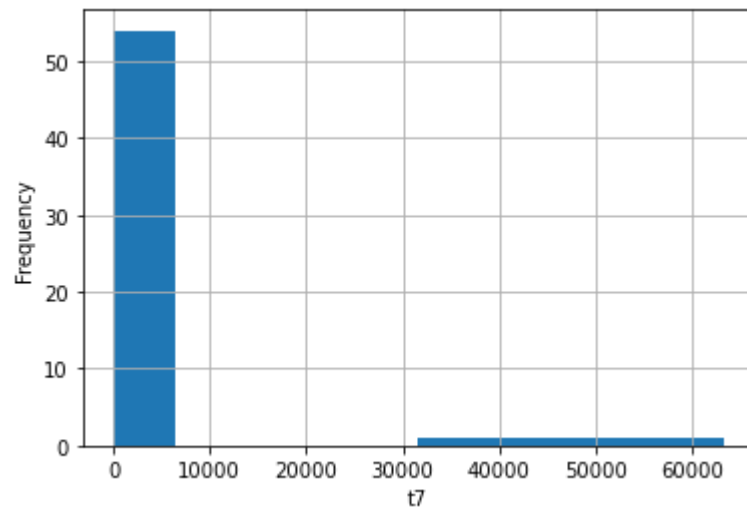
The amount of rain near station t14 has a high probability of being in the range 0 – 8000 ml and the probability decreases as the amount of rain increase.



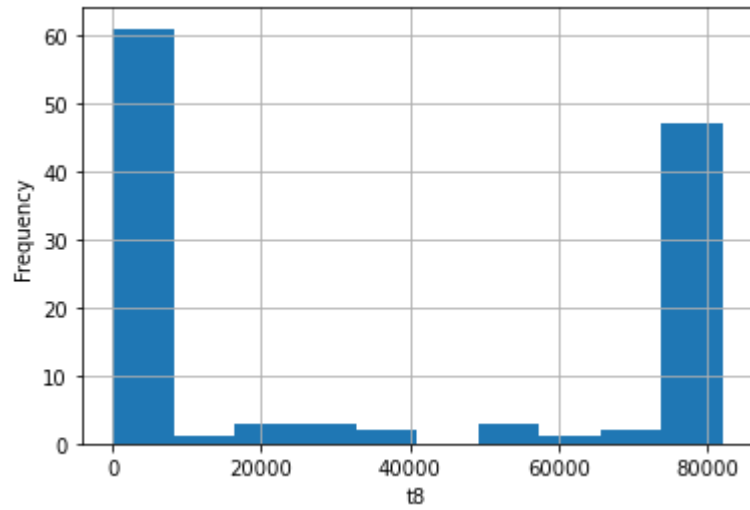
The amount of rain near station t15 has a high probability of being in the range 0 – 1836 ml.



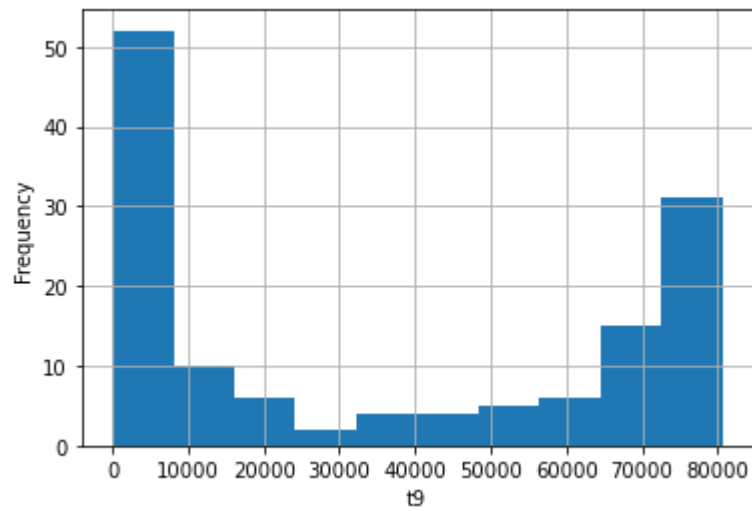
The amount of rain near station t6 has a high probability of being in the range 0 – 180 ml.



The amount of rain near station t7 has a high probability of being in the range 0 – 6326 ml.

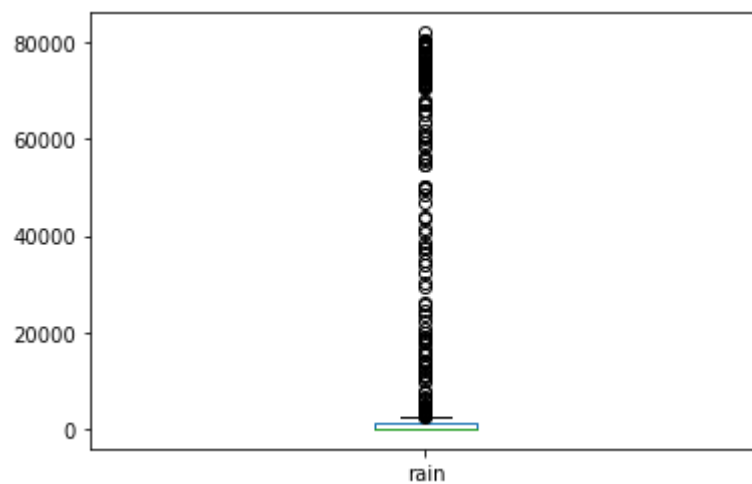


The amount of rain near station t8 generally lie between 0 – 8204 ml and 73836 – 82040 ml.

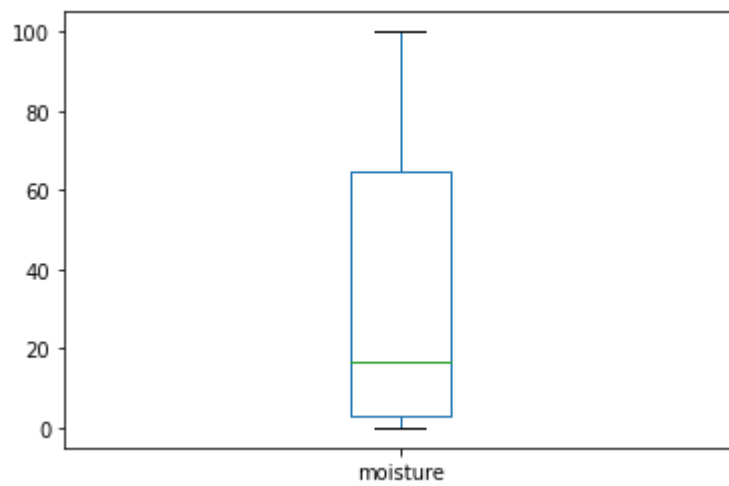


The probability of the amount of rain near station t9 decreases as the amount of rain increases and then the probability starts to increase as it goes beyond 32233 ml.

**Ans 6)**



In the boxplot of rain most of the data points are concentrated in the range 0 – 1000 ml. Inter Quartile range of this plot is spread over very small range and it has lot of outliers which are spread across a really large range.



The median of moisture is around 20% and Inter quartile range of this plot is spread out in the good range, largest observation (top whisker) touches 100 and the bottom whisker is near zero. Also, the box plot of the moisture has no outliers.