

# Indian Institute of Technology, Mandi

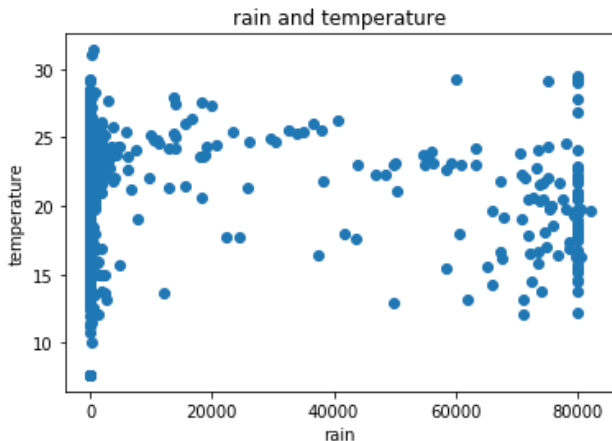
## IC272 – Data Science 3

### Lab Assignment 1

Pooja Patidar (B19255)

Mob no.:- 8278785526

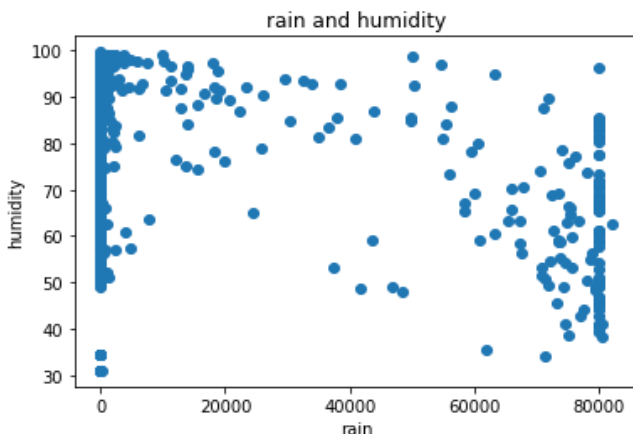
#### 1.) Graph for rain and temperature



**Correlation Coefficient** – ( - 0.109)

**Inference** – Here we can see that correlation is negative which implies that rain and temperature moves in opposite direction that means rain decreases on increasing temperature and value tells that they have negligible connection between them.

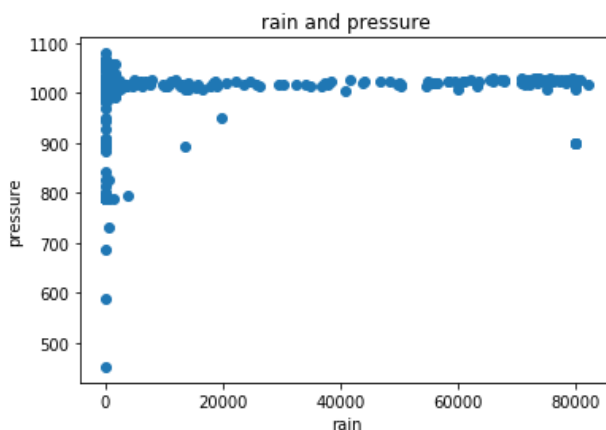
#### 2.) Graph for rain and humidity



**Correlation Coefficient** – ( - 0.435)

**Inference** – Here we can see that correlation is negative which implies that rain and humidity moves in opposite direction that means rain decreases on increasing humidity and value tells that they have less connected with each other.

#### 3.) Graph for rain and pressure

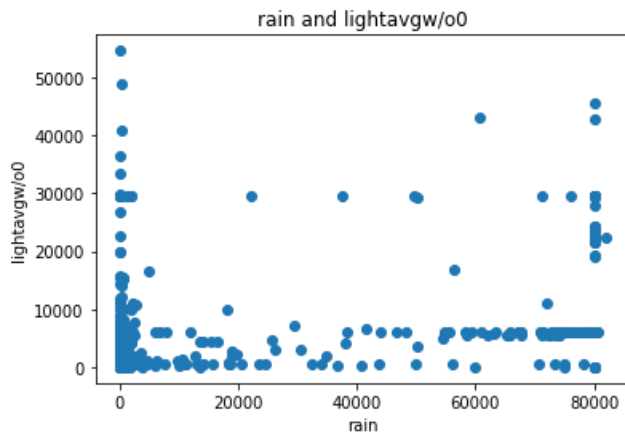


**Correlation Coefficient** – ( 0.071)

**Inference** – Here we can see that correlation is positive which implies that rain and pressure moves in same direction and value tells that they have negligible connection between them.

By observing plot we can say that we can draw regression line and it also termed as a linear relation exist.

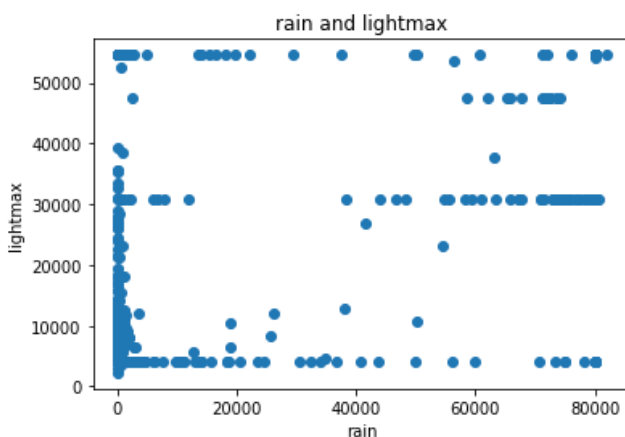
#### 4.) Graph for rain and lightavgw/o0



**Correlation Coefficient – ( 0.527)**

**Inference –** Here we can see that correlation is positive which implies that rain and lightavgw/o0 moves in same direction and value tells that they are moderate correlated/connected.

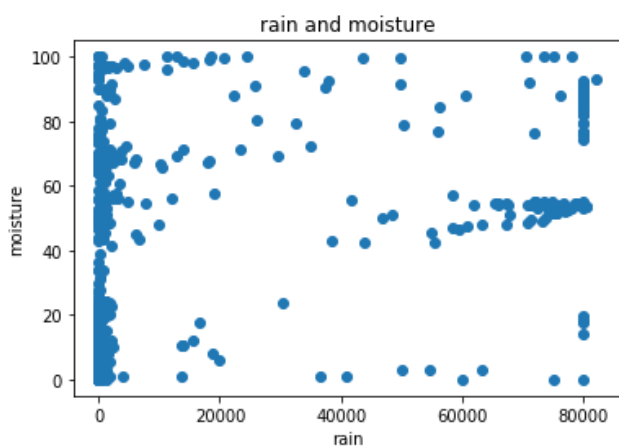
#### 5.) Graph for rain and lightmax



**Correlation Coefficient – ( 0.313)**

**Inference –** Here we can see that correlation is positive which implies that rain and lightmax moves in same direction and value tells that they have low positive correlation.

#### 6.) Graph for rain and moisture

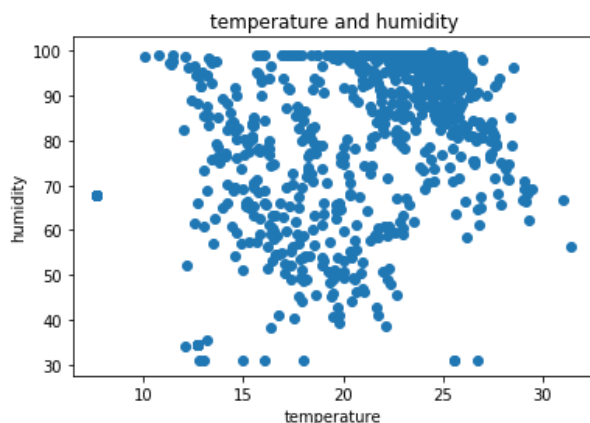


**Correlation Coefficient – ( 0.427)**

**Inference –** Here we can see that correlation is positive which implies that rain and moisture moves in same direction and value tells that they have low positive correlation.

By observing graph we can say that regression line have some positive slop.

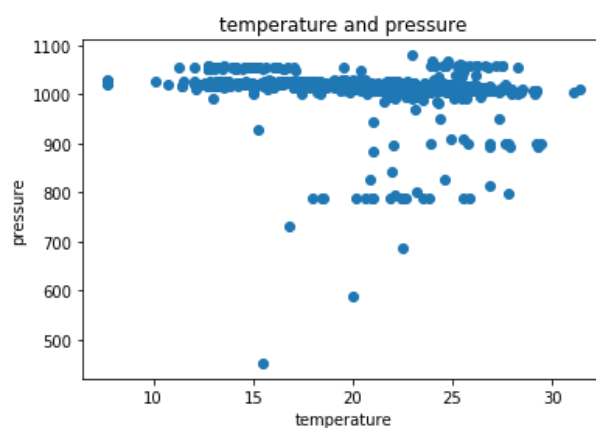
## 7.) Graph for temperature and humidity



**Correlation Coefficient – ( 0.402)**

**Inference –** Here we can see that correlation is positive which implies that temperature and humidity moves in same direction and value tells that they have low positive correlation.

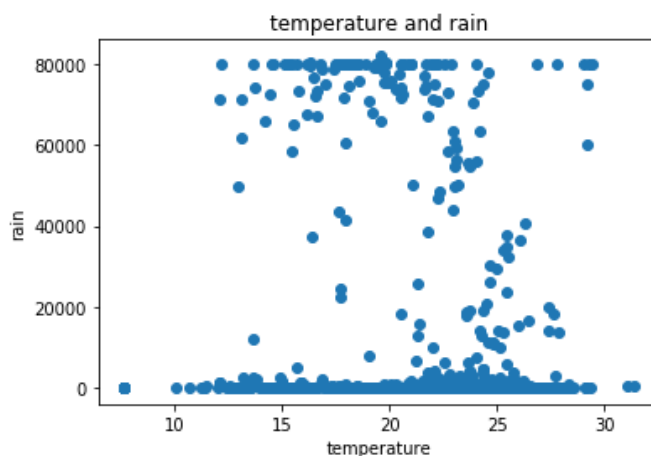
## 8.) Graph for temperature and pressure



**Correlation Coefficient – ( -0.181)**

**Inference –** Here we can see that correlation is negative which implies that temperature and pressure moves in opposite direction and value tells that they have negligible connection between them.

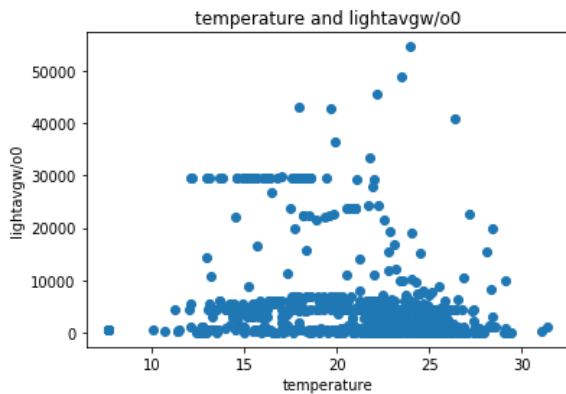
## 9.) Graph for temperature and rain



**Correlation Coefficient – ( -0.109)**

**Inference –** Here we can see that correlation is negative which implies that temperature and rain moves in opposite direction that means rain decreases on increasing temperature and value tells that they have negligible connection between them.

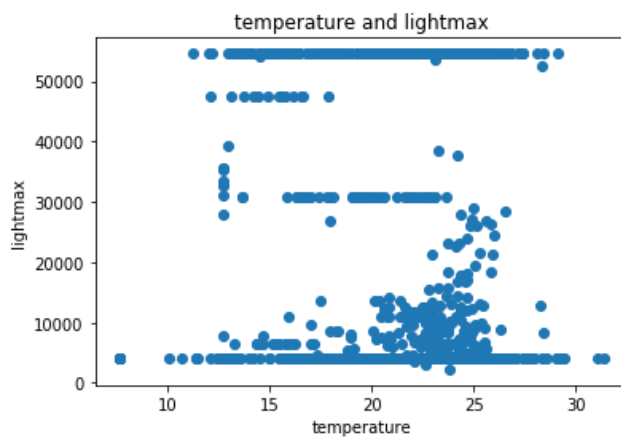
### 10.) Graph for temperature and lightavgw/o0



**Correlation Coefficient** – (- 0.181)

**Inference** – Here we can see that correlation is negative which implies that temperature and lightavgw/o0 moves in opposite direction and value tells that they have negligible connection between them.

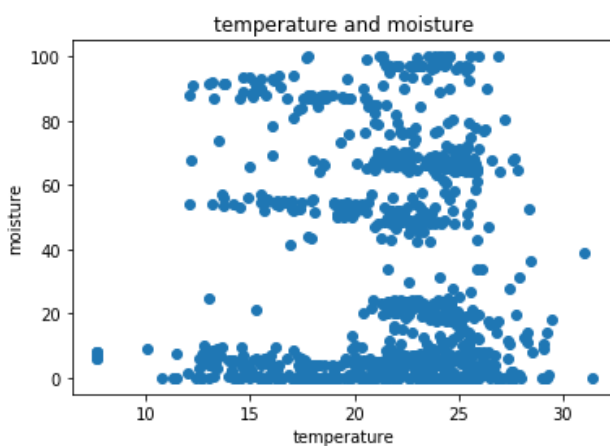
### 11.) Graph for temperature and lightmax



**Correlation Coefficient** – (- 0.146)

**Inference** – Here we can see that correlation is negative which implies that temperature and lightmax moves in opposite direction and value tells that they have negligible connection between them.

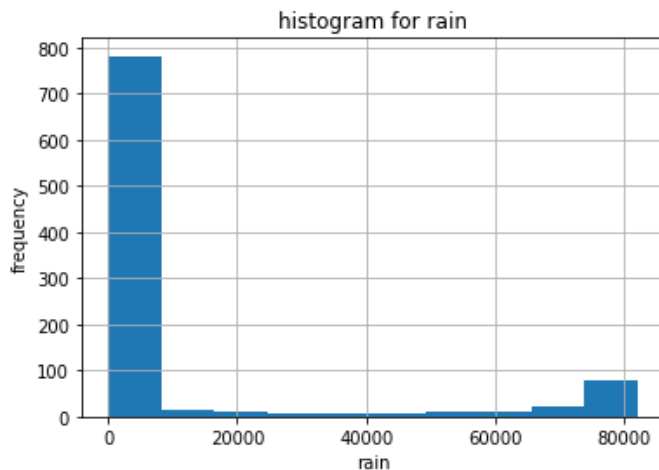
### 12.) Graph for temperature and moisture



**Correlation Coefficient** – (- 0.081)

**Inference** – Here we can see that correlation is negative which implies that temperature and moisture moves in opposite direction and value tells that they have negligible connection between them.

### 13.) Histogram for Rain



mean : 10701.538  
median : 18.0  
mode : 0.0  
minimum : 0.0  
maximum : 82037.25

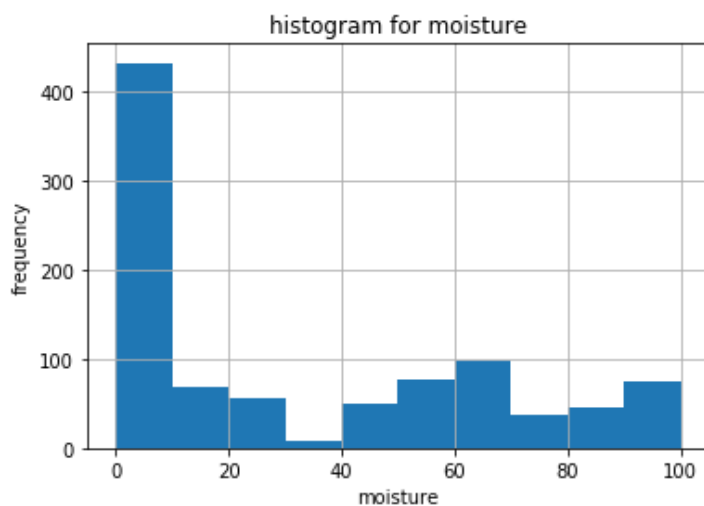
**Inference** –At 0, frequency is highest so we can conclude that 0 is mode for given data. Similarly, for mean and median. From x-axis we find range of data, minimum is

at 0 and maximum is above 80000 (82037.25 on calculation).

Mode < Median < Mean, we conclude that data is positively skewed.

We also conclude that there are so many outliers in this, because we can see the unusual value (here values from 10000 to 70000).

### 14.) Histogram for moisture



mean : 32.386  
median : 16.7042  
mode : 0.0  
minimum : 0.0  
maximum : 100

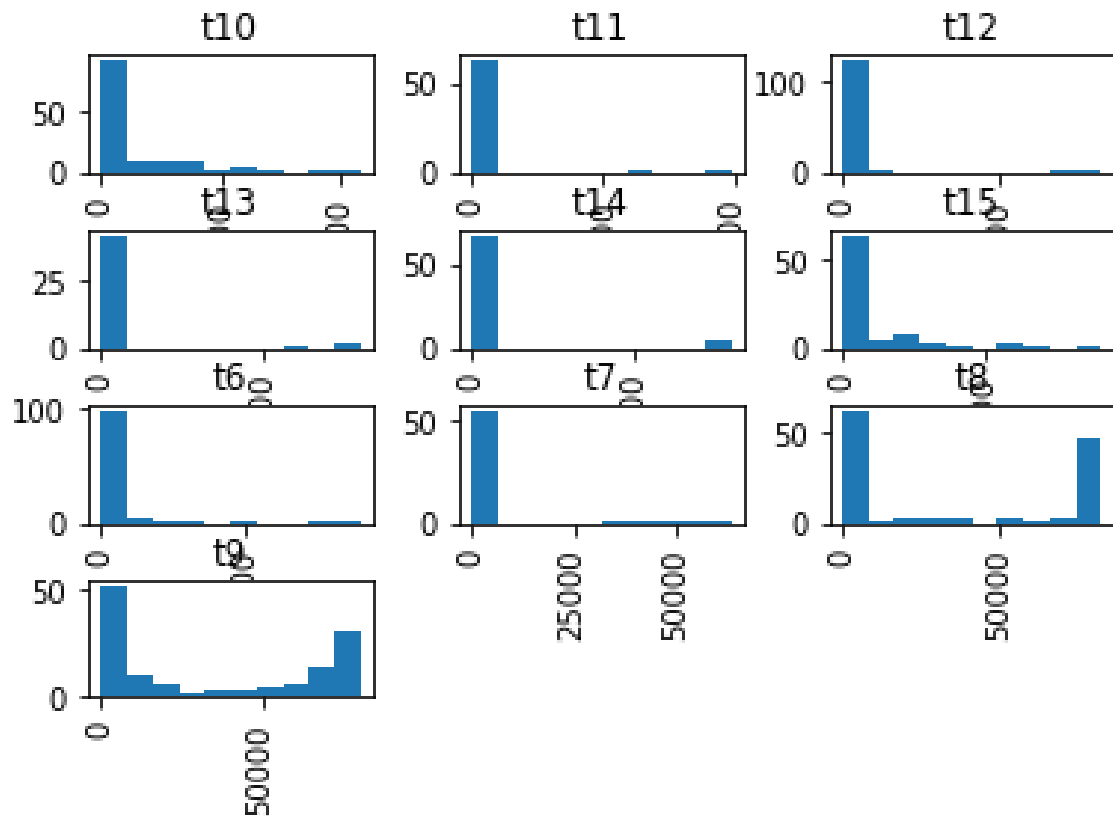
**Inference** –At 0 frequency is highest so we can conclude that 0 is mode for given data. Similarly, for mean and median. From x-axis we find range of data, minimum is at 0 and

maximum is at 100.

Mode < Median < Mean, we conclude that data is positively skewed.

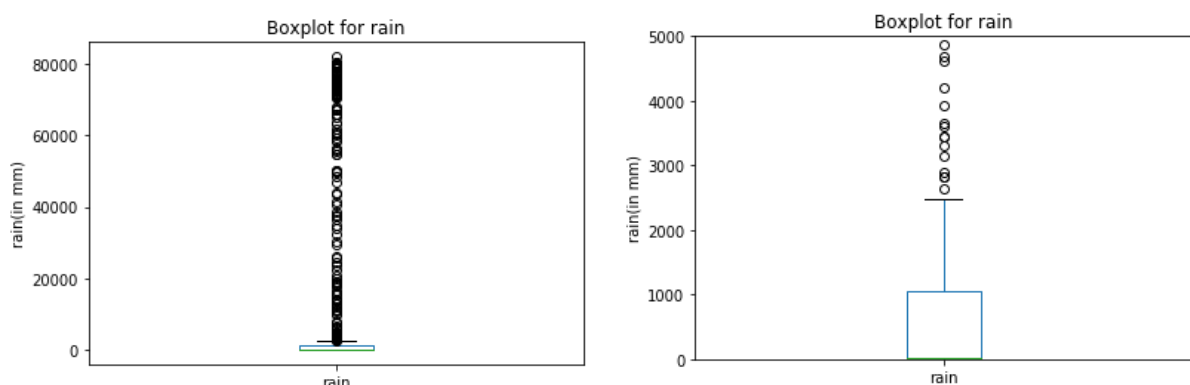
We also conclude that there are no outliers.

### 15.) Histogram for each station



Each histogram is positively skewed and contains many outliers. Mode of each station is 0, mean and median also lies in the range of the highest bar in graph. Minimum of each station is 0.

### 16.) Boxplot for rain

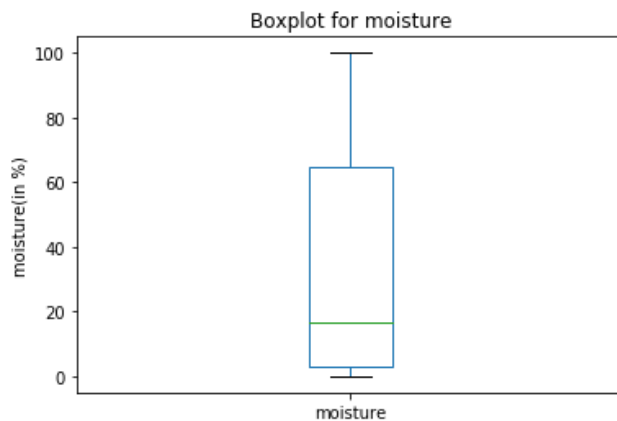


graph by reducing y limit or noise.

It contains so many outliers that implies data is noisy. This condition we observed in histogram of rain too.

It also tells about lower extreme (i.e. 0), lower quartile (i.e. 0), upper quartile (i.e. 1000), upper extreme (i.e. 2500) and median (i.e. 18). All this values I find by reducing the limit of y axis by using `plt.ylim( )`.

### 17.) Boxplot for moisture



It contains no outliers. Here lower extreme (i.e. 0), lower quartile (i.e. 0), upper quartile (i.e. 60), upper extreme (i.e. 100) and median (i.e. 16.7042).