INTERNET of THINGS

Lab 5 Assignment

On

IoT EDA and GIS

HÖGSKOLAN
DALARNA

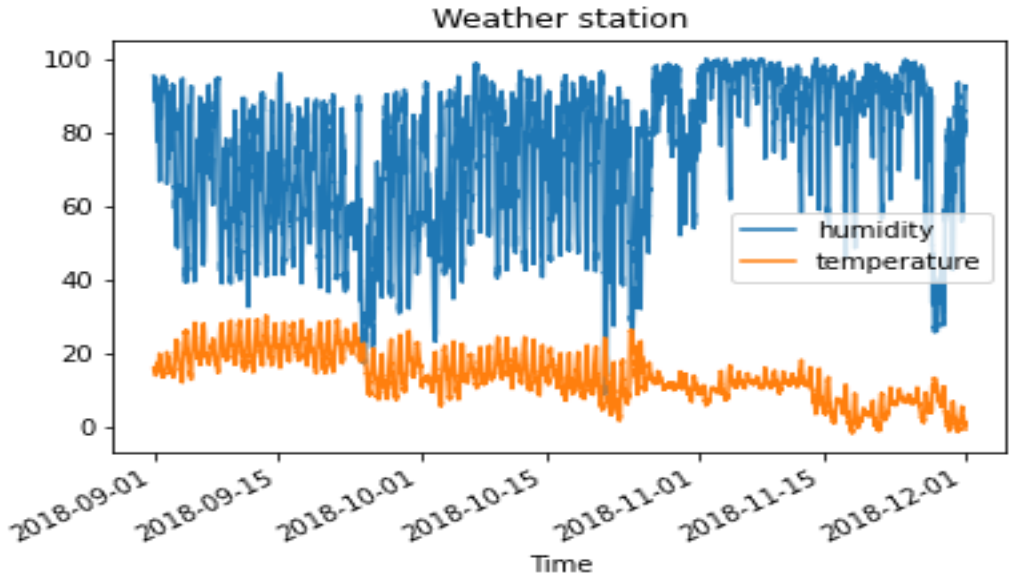
Manthri Bala Kiran

H19baman@du.se

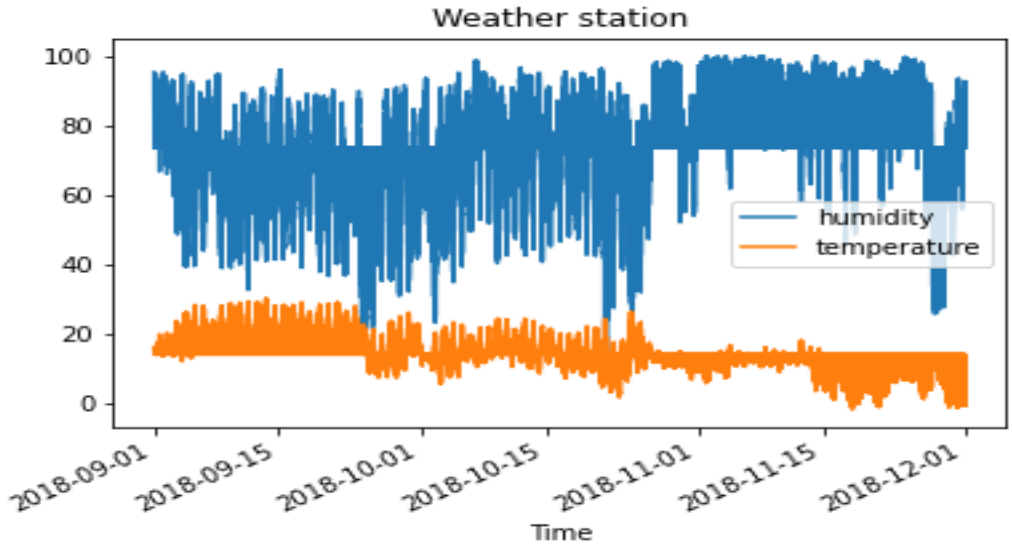
Task 1:

a) The explanation for Weather Station Data:

In the weather station data, I have noticed too many Null Values and I chose to drop them. And I have tried to visualize the temperature and humidity and noticed that over the time frame the temperature started to rise at the starting and continuously dipping down. And Humidity was changing continuously over time and it was high at the end of the time frame.

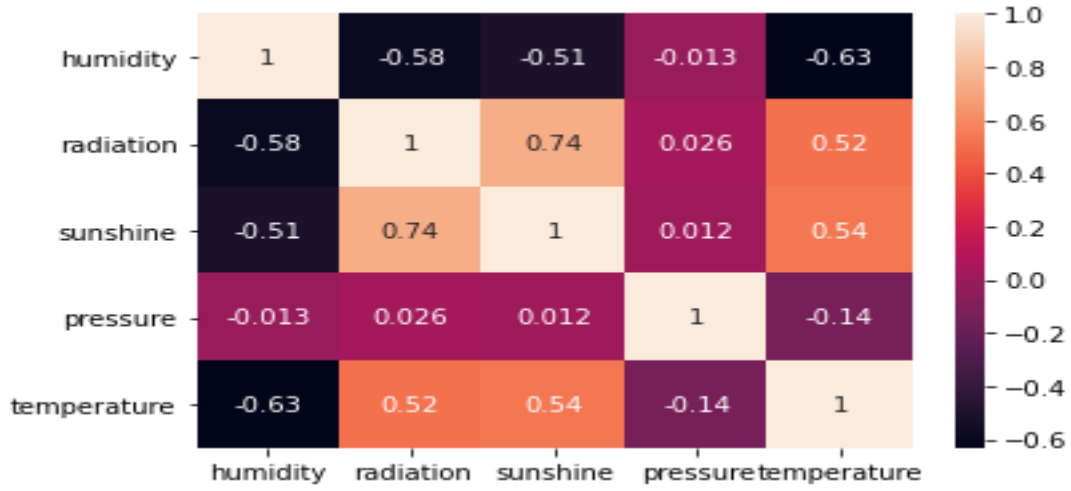


When I tried to fill the Na with the FFILL. We can see the same observation but as they were many missing values In the data frame and f-fill has given more data points.



In the next phase, I have tried to see the correlation of the features in the data. And noticed that the humidity and the temperature were highly negatively correlated. And Radiation and sunshine were

highest positively correlated. And we can also say that if the high amount of sunshine and temperature would be high enough and less humid.



The data was interesting to speak about if we have more amount of data and fewer missing values over time we can have more insights into the behavior. The data was collected and stored for every 5 mins, but I feel we can change that to 30 or 1hour.

The main characteristics from the data set would be Humidity, Sunshine, Radiation, and Temperature.

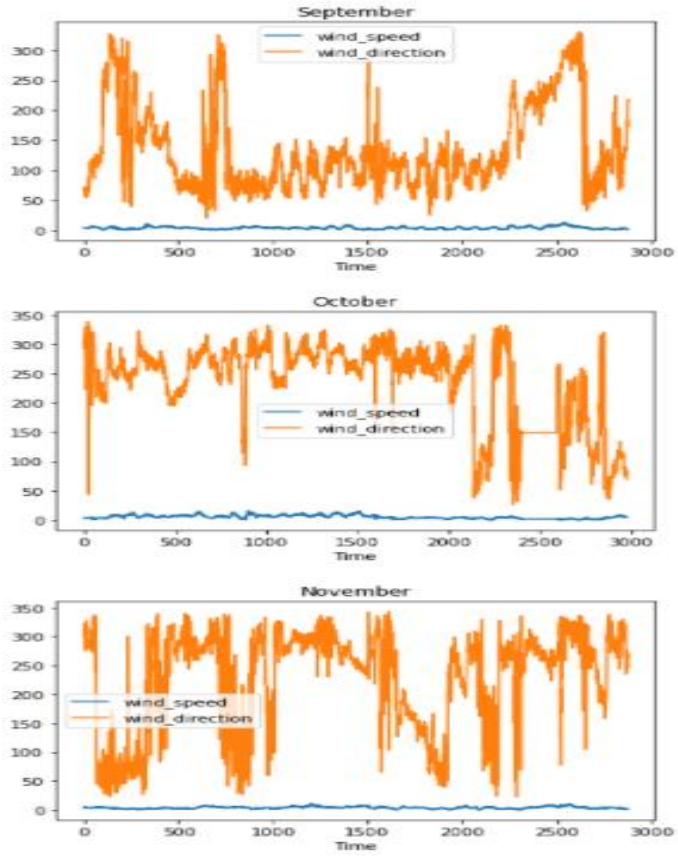
b) Wind Data:

In the three wind data files, I have noticed that October has missing values so I tried to handle the missing values with f-fill and drop missing values.

```
sept
 wind_speed          3.815301
wind_direction     132.810118
dtype: float64
oct fill na
 wind_speed          5.205481
wind_direction     235.194922
dtype: float64
oct drop na
 wind_speed          5.446737
wind_direction     241.164227
dtype: float64
nov
 wind_speed          3.757534
wind_direction     227.598412
dtype: float64
wind_speed           4.269835
wind_direction     198.937346
dtype: float64
```

We can see that October has the highest wind speed compare with the other months and we can also see that when we drop the NA values the average was quite higher it might be either case like

the forward fill might have less value in the wind speed making the average less and the count of the observation can also be considered.



From the above figure, we can see the wind directions across the months. In September the wind direction was continuously changing but settled mostly between 50-200 degrees and in October the wind direction was mostly above 200 degrees and in November the wind direction was continuously changing over time. And from the correlation plot, I have noticed that the wind speed and direction were slightly positively correlated.

For the wind data set, I feel we can store the data for every one hour, or 30 mins if we require more data points.

c) Lab Feedback:

The lab material was relevant and the lab was interesting in my point of view.