

**TEAM DPDNS**

# **MOISTURE MINDS**

## **SOIL MOISTURE PREDICTION MODEL**

Team Members:

Dillon Gonsalves

Preeti Vasaikar

Densil Nunes

Nigel Misquitta

Shoydon Alphonso

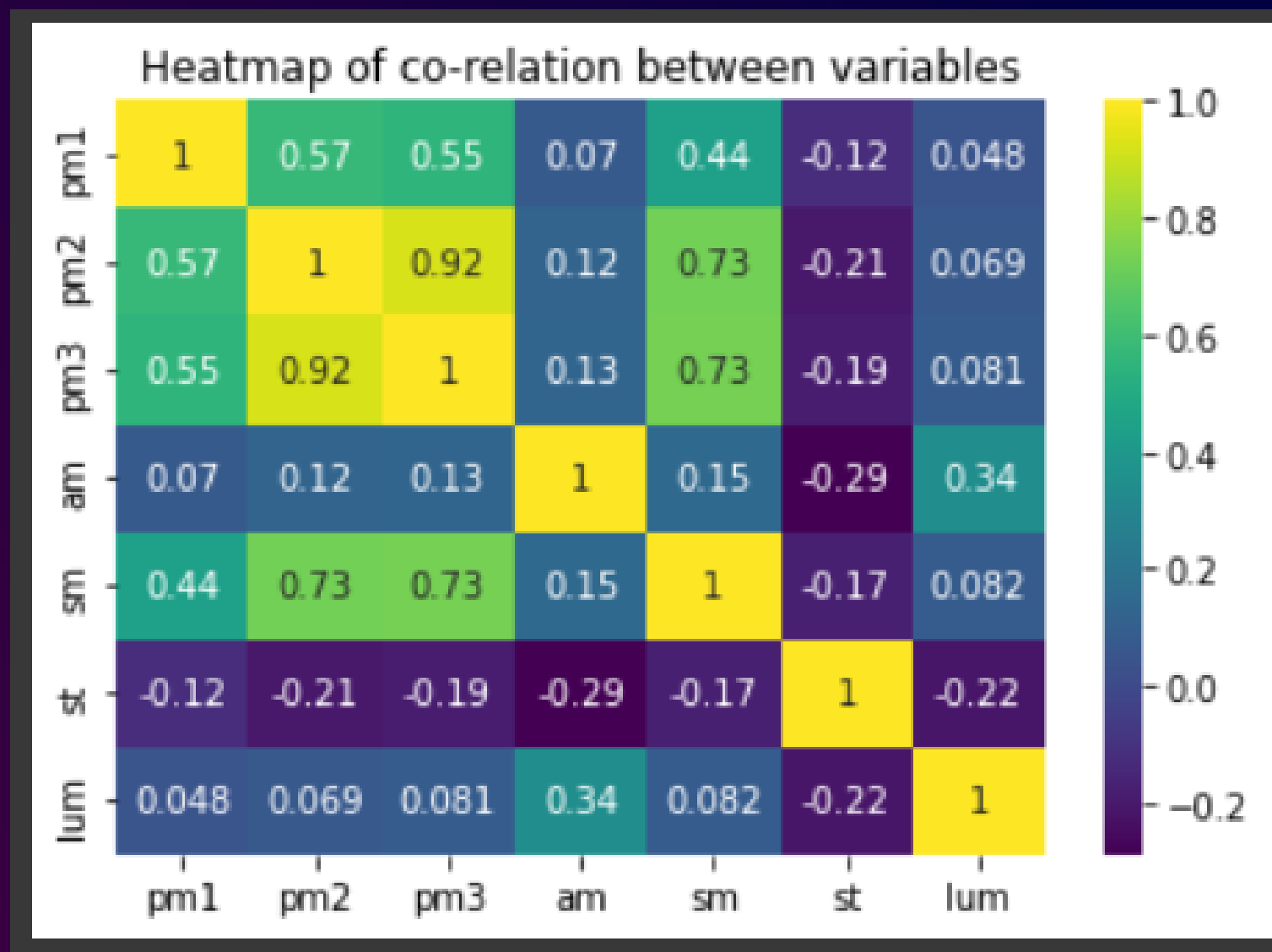
# PROBLEM STATEMENT

To build a machine learning model to predict the soil moisture level for month of March 2023 based on the data of July 2022 to 10th March 2023.

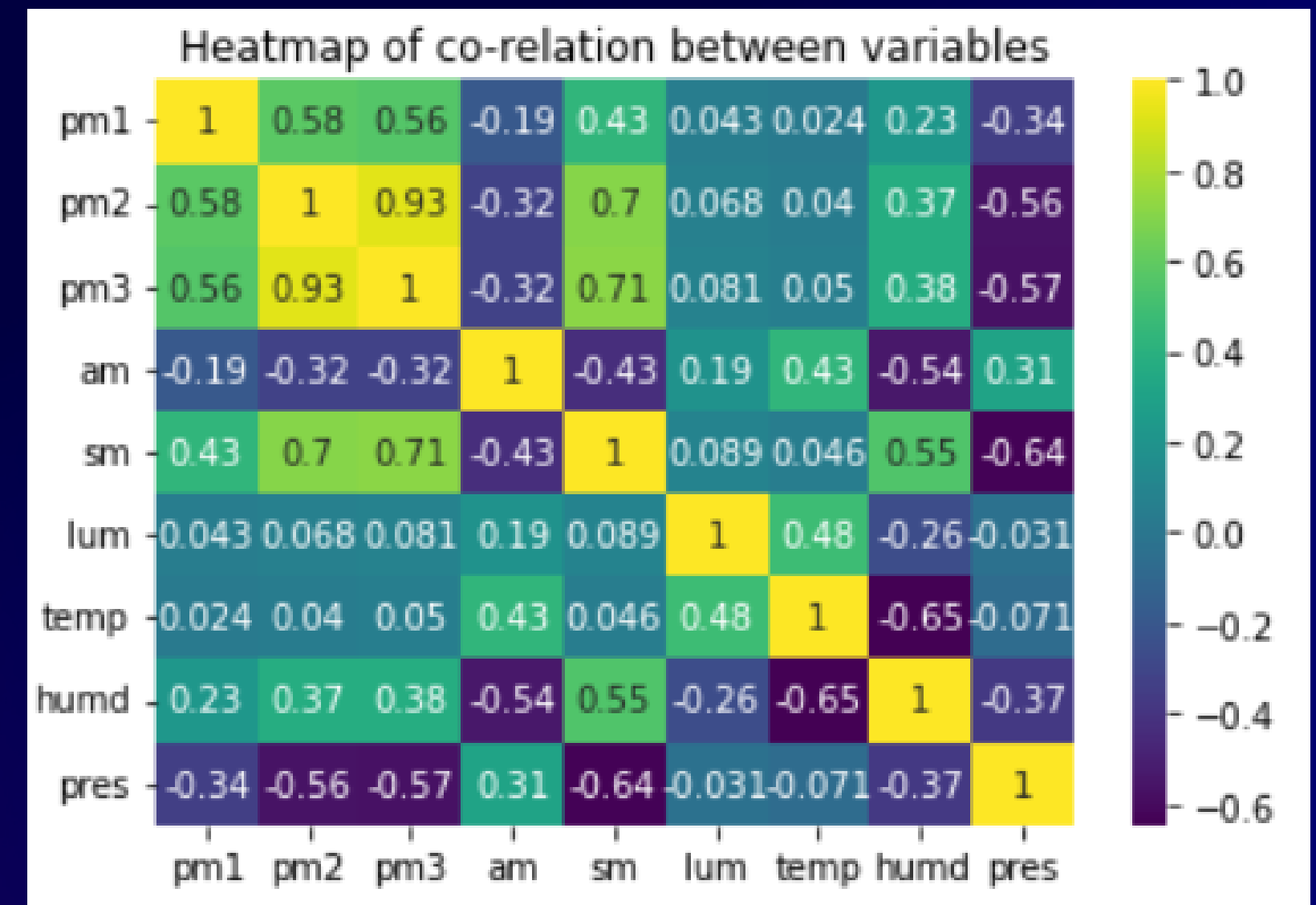


# HEATMAP ON RAW DATA

user\_data1

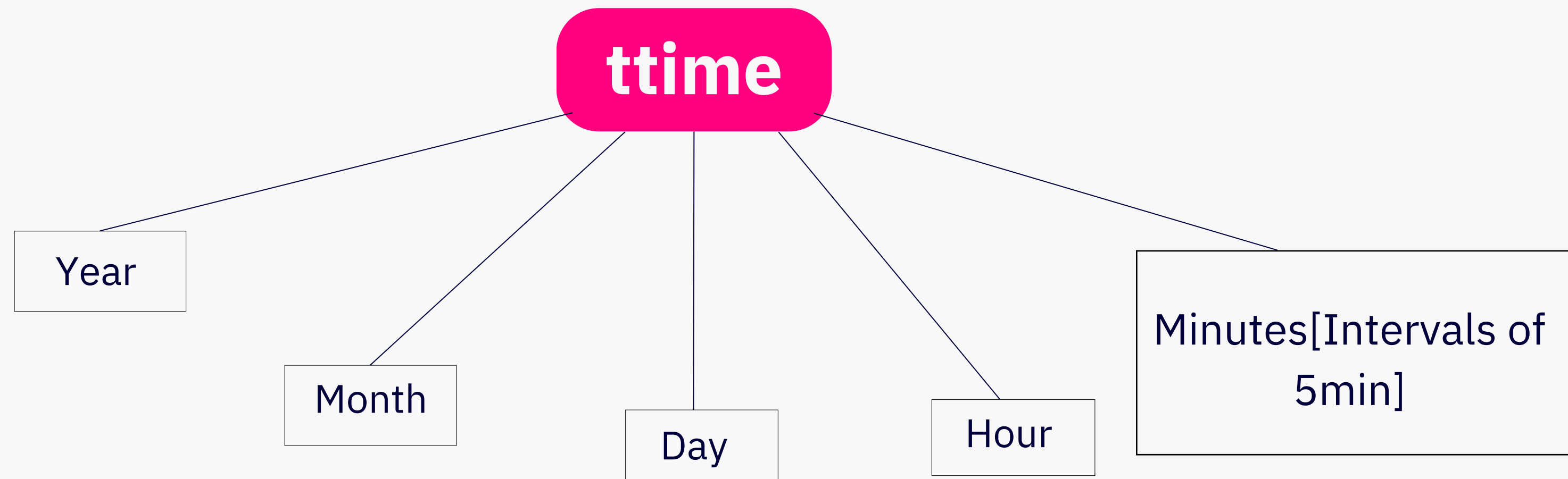


user\_data2



# PREPROCESSING STEP

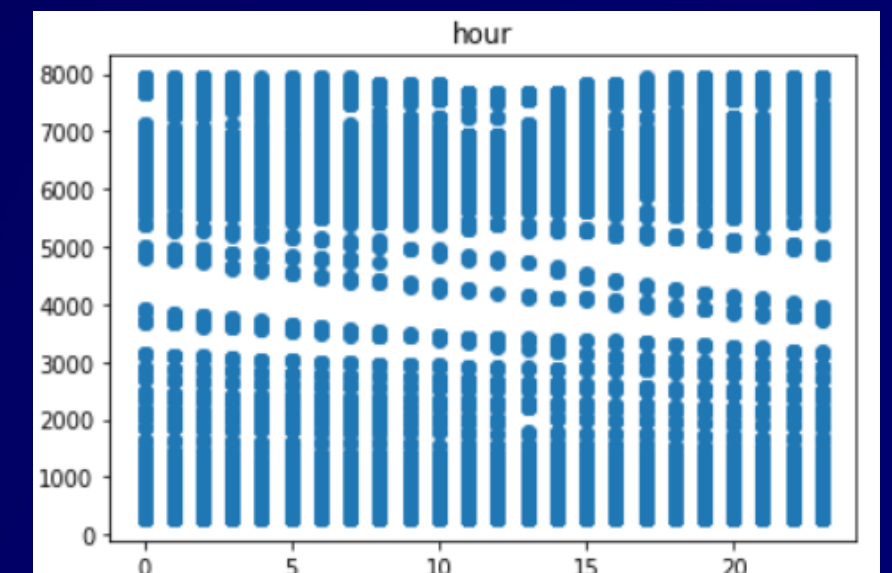
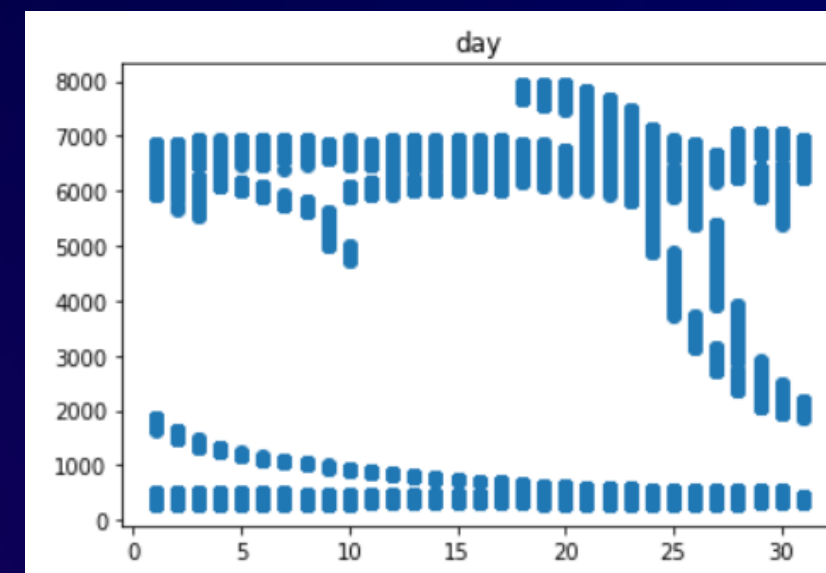
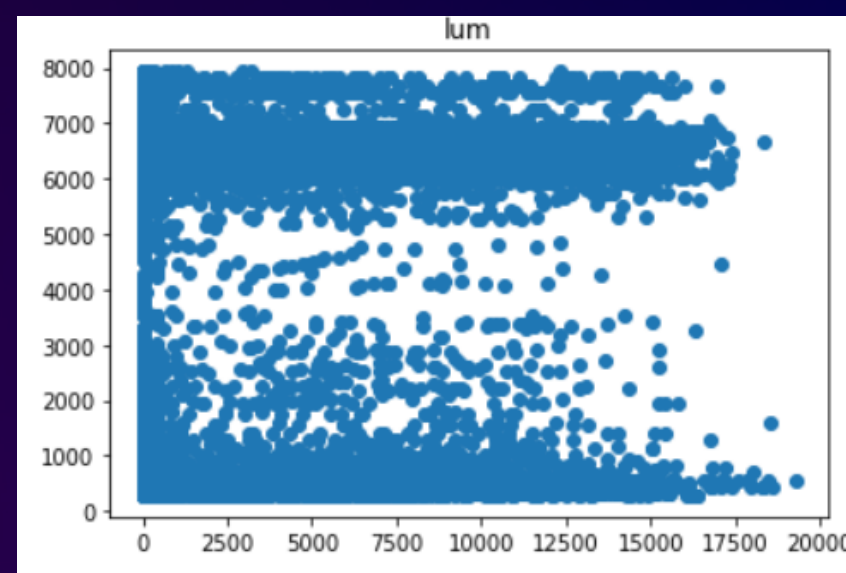
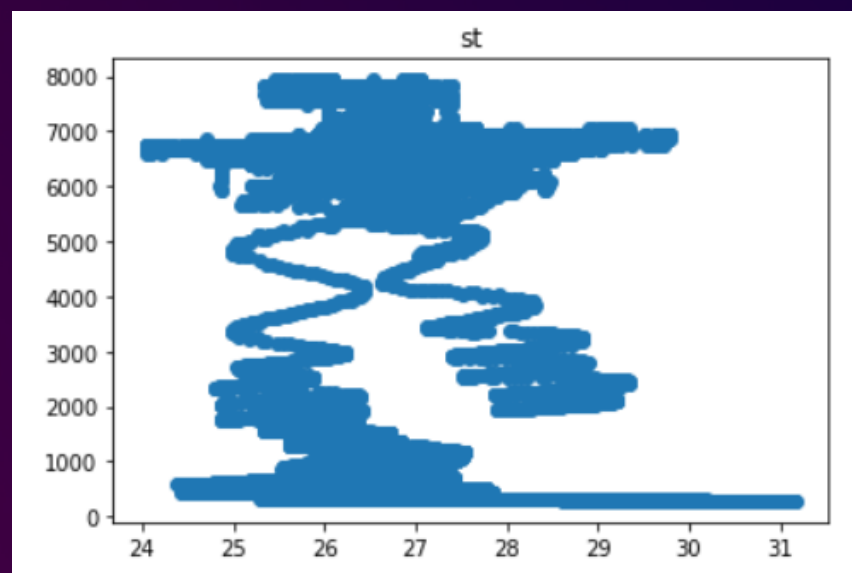
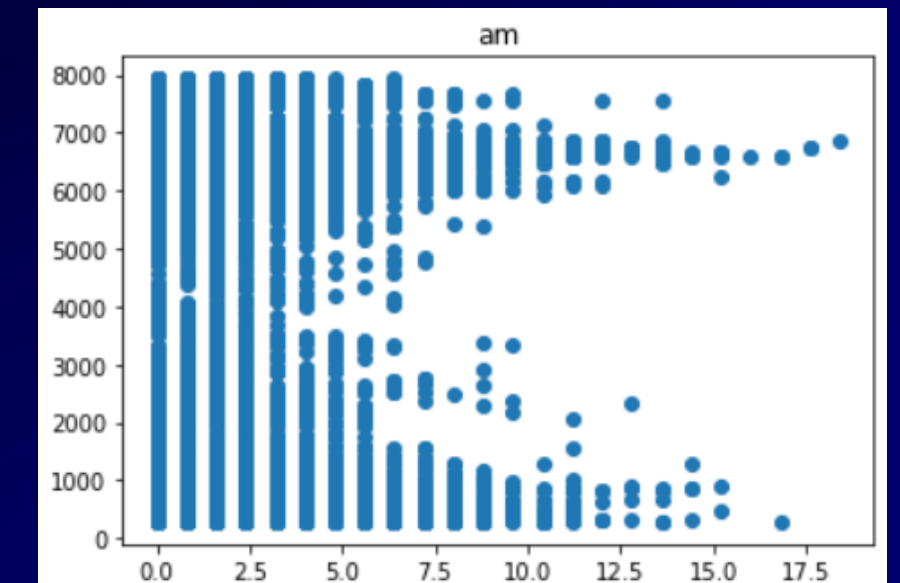
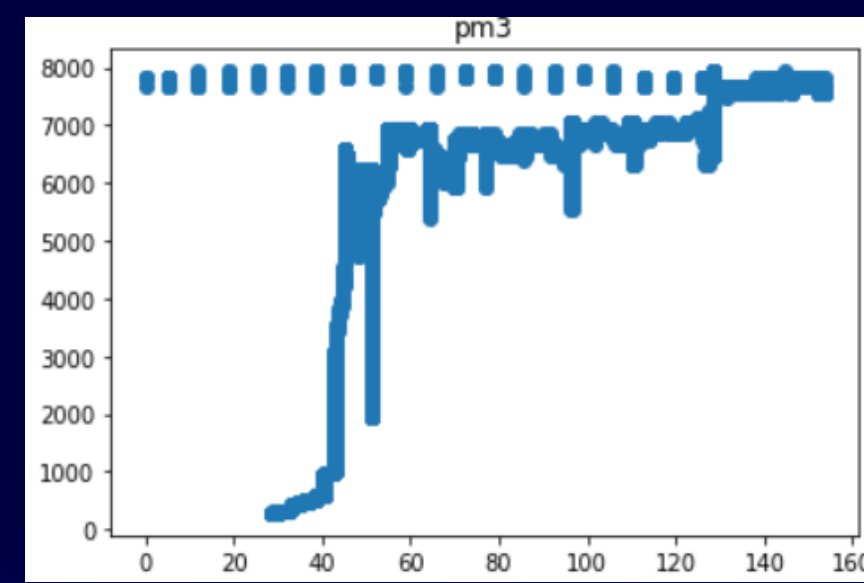
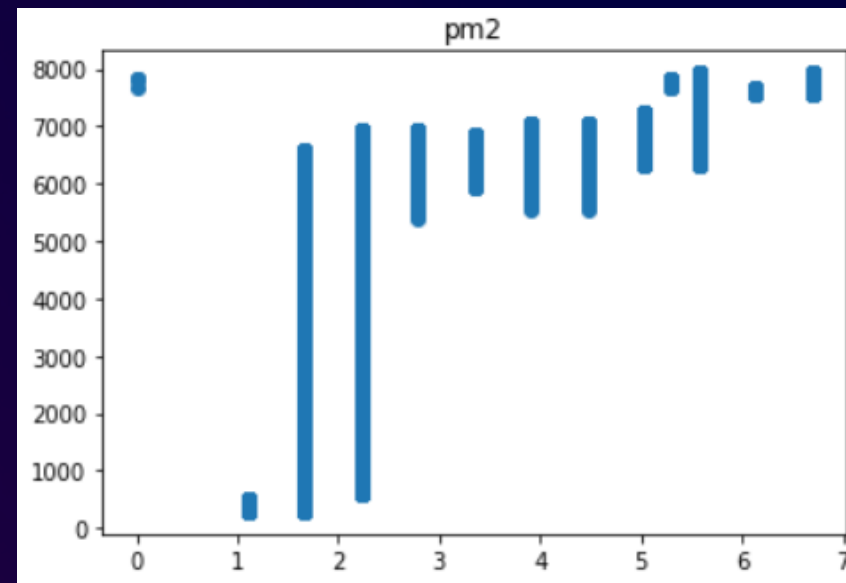
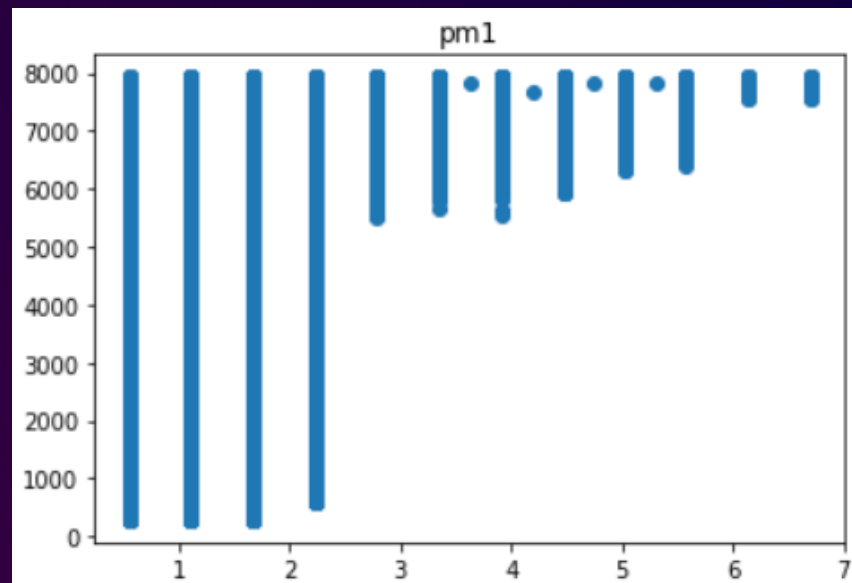
- The ttime data is split into minutes[5 mins], hours, days, month and year in the final dataset.



# DATA INFO

## Preprocessed data

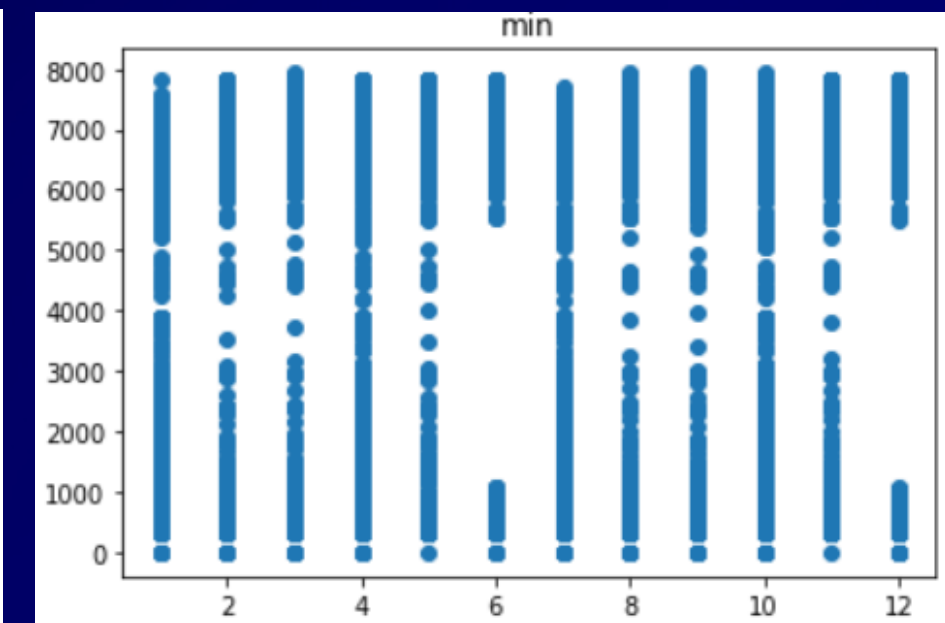
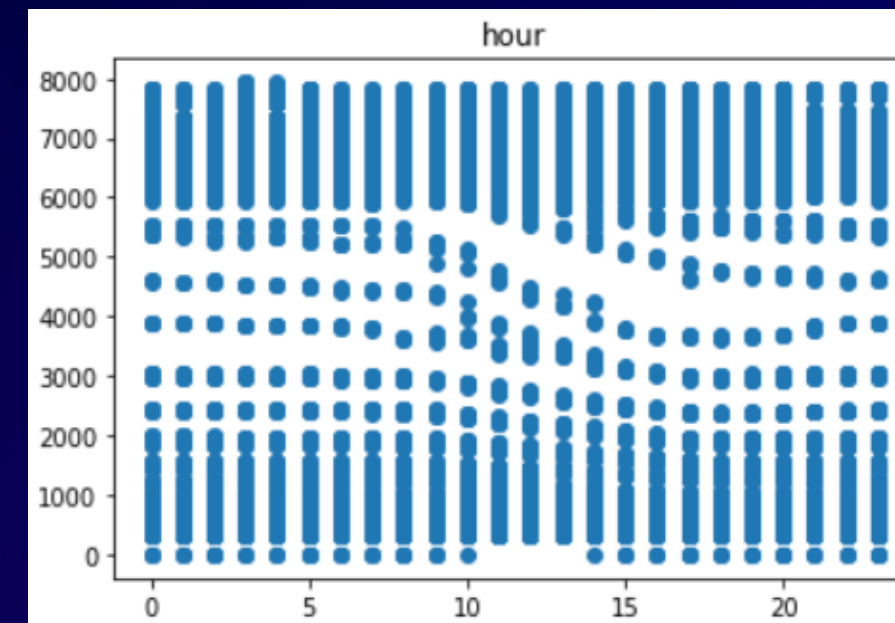
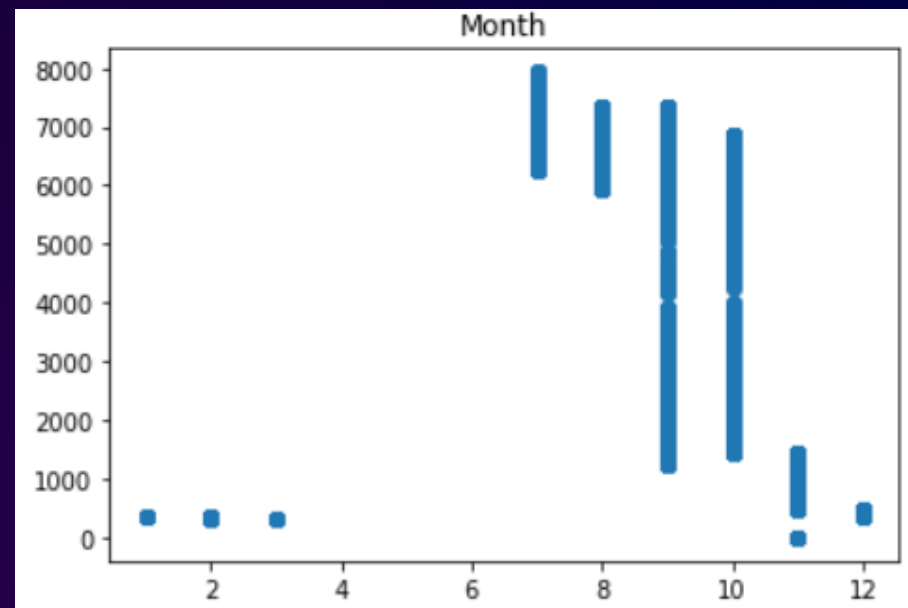
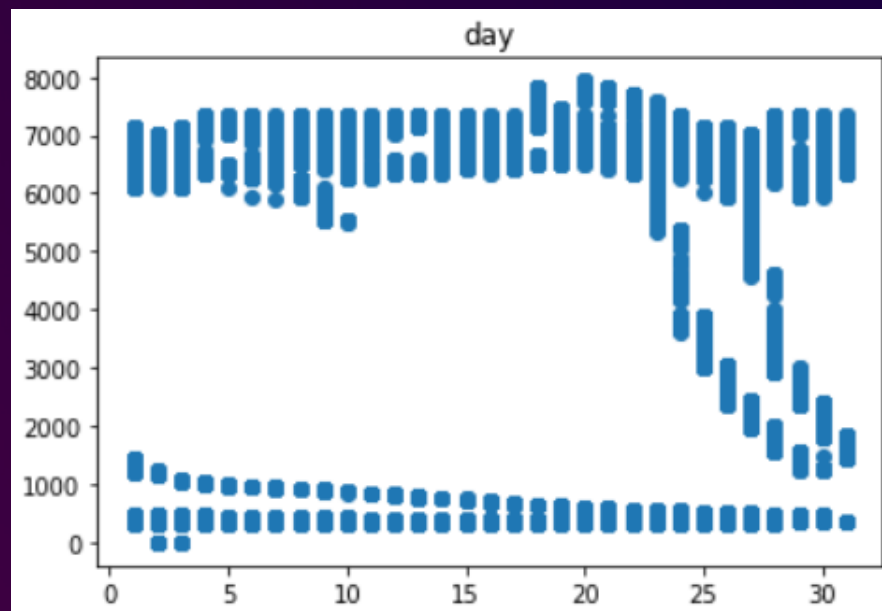
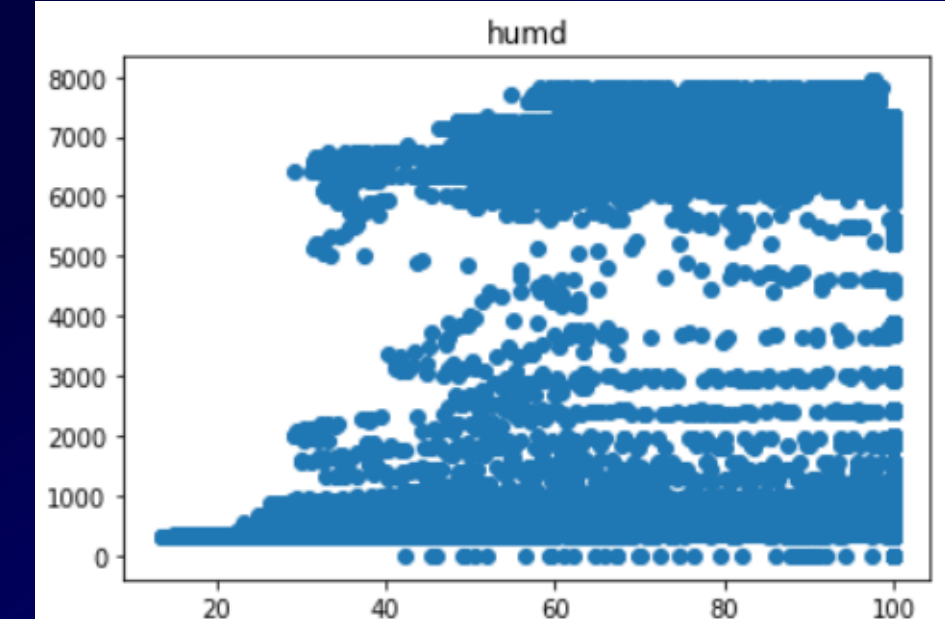
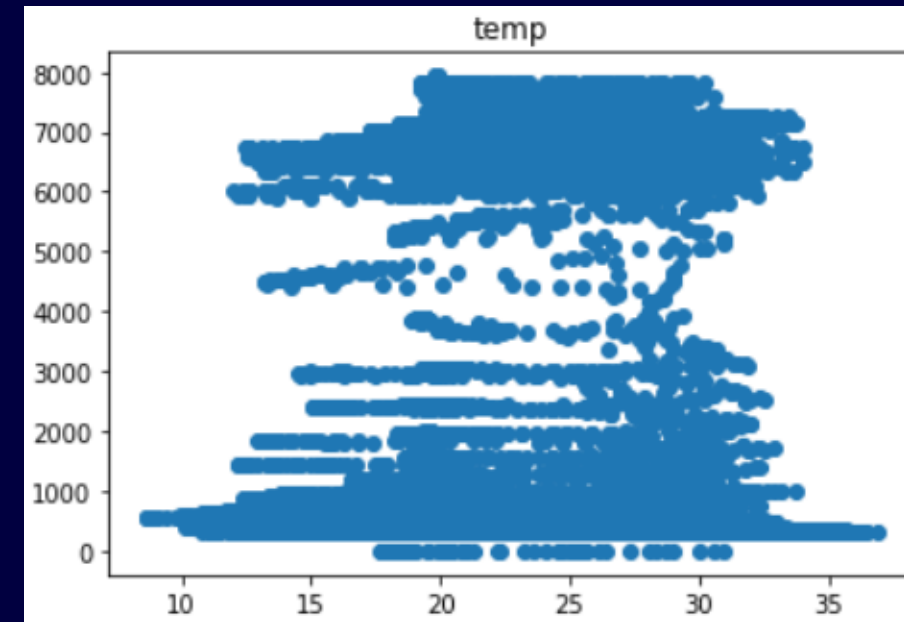
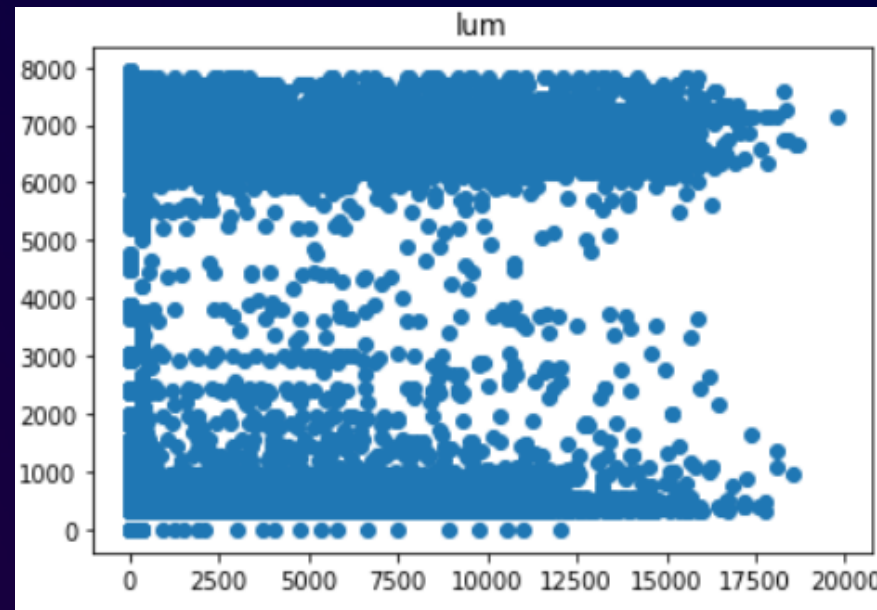
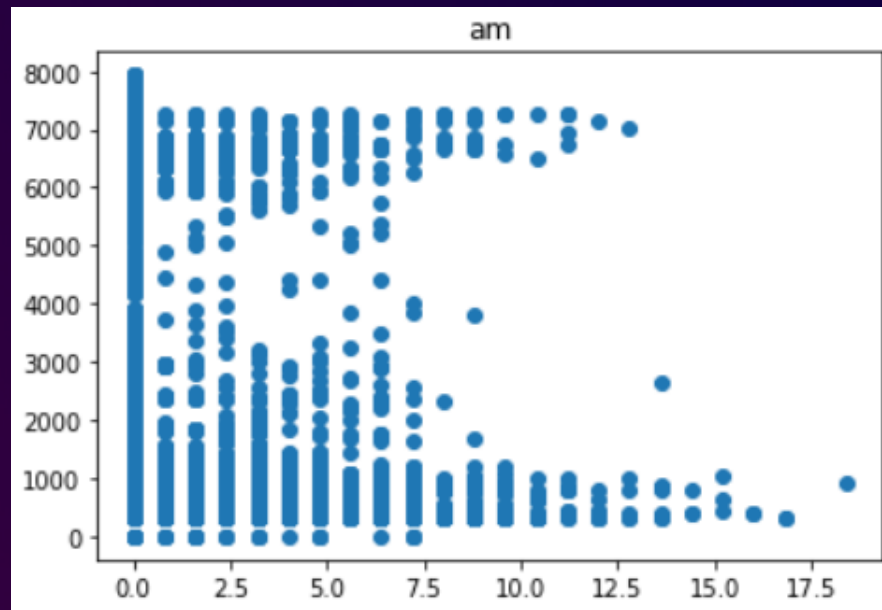
Scatter plots of sm against the other features in the dataset.





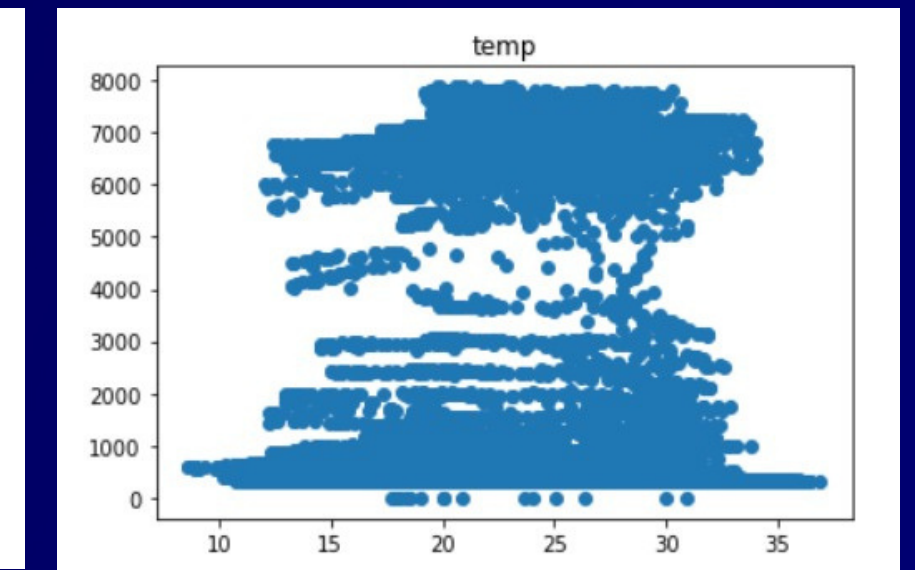
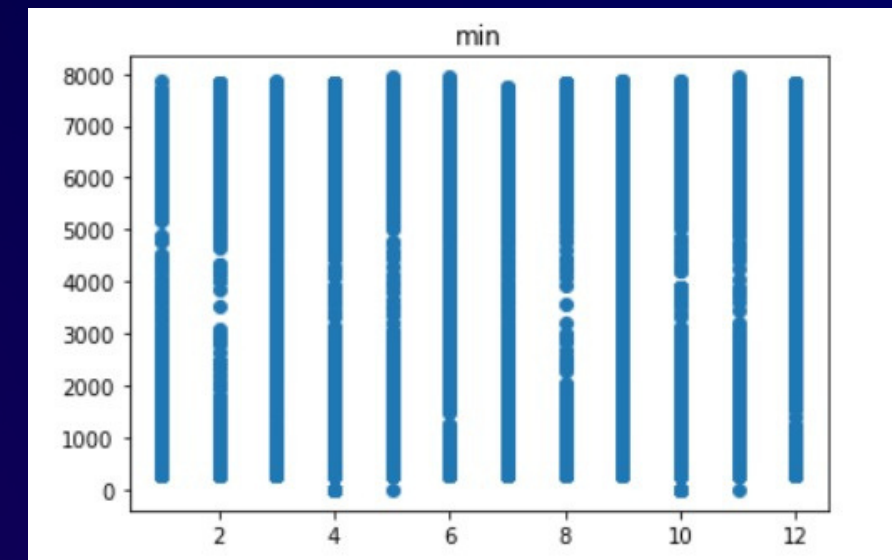
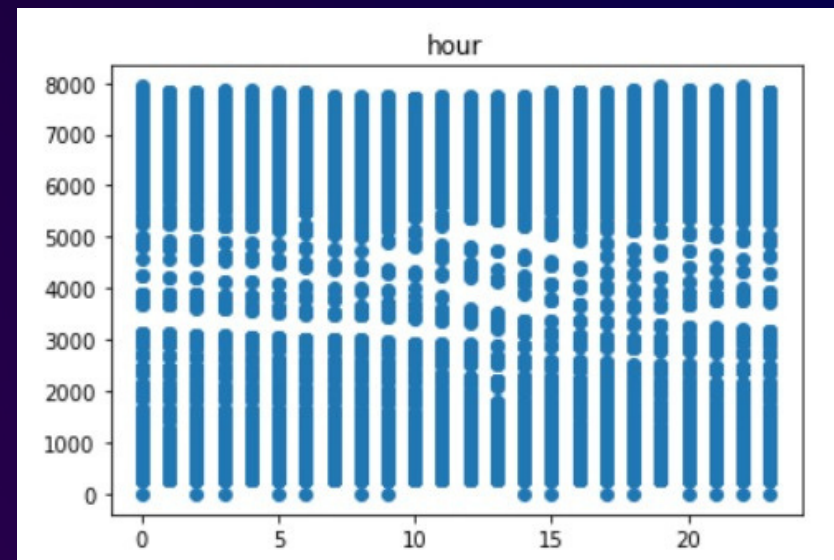
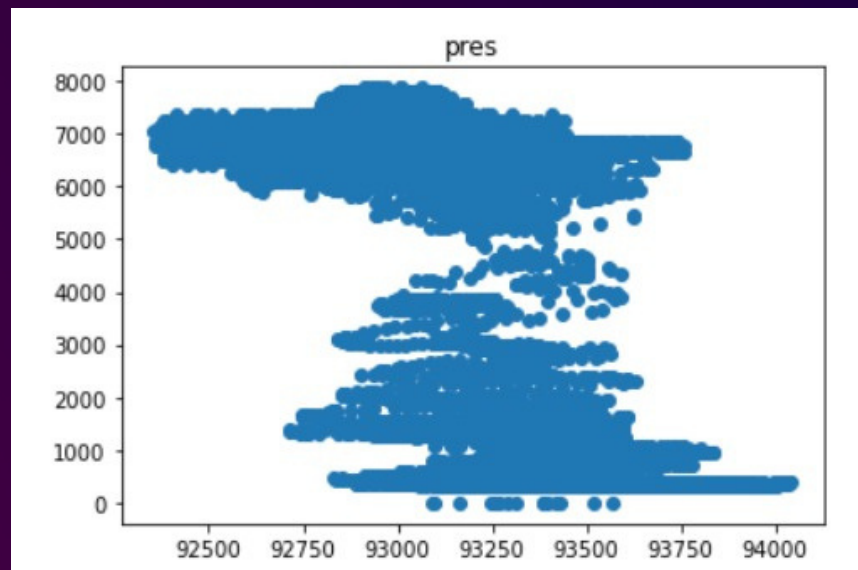
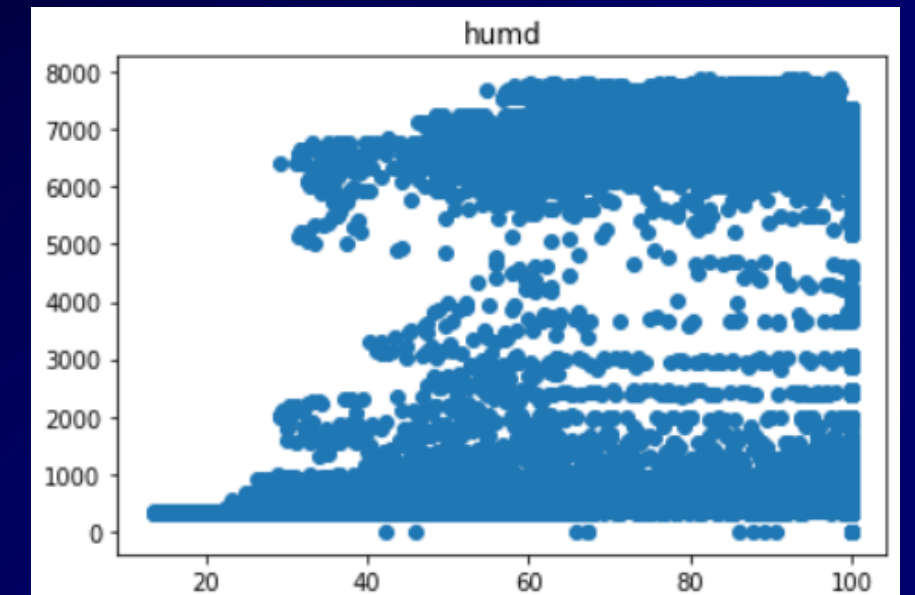
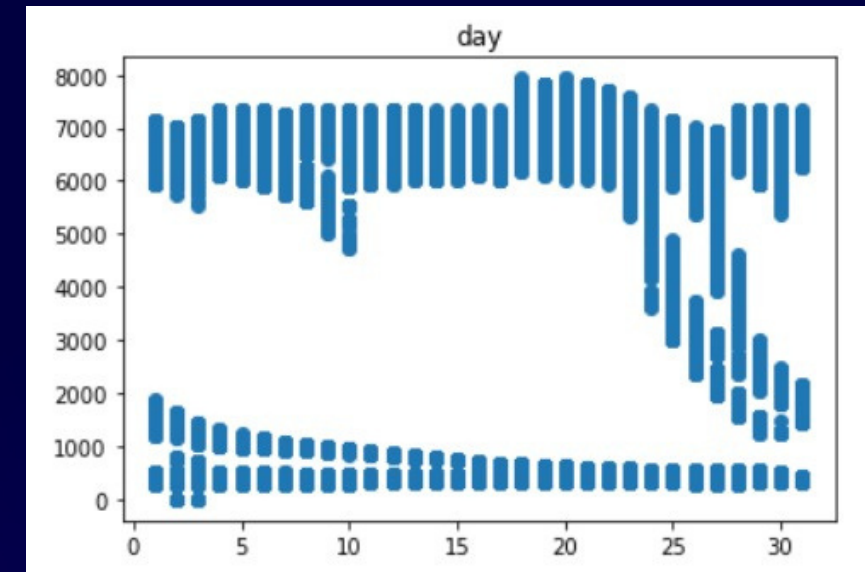
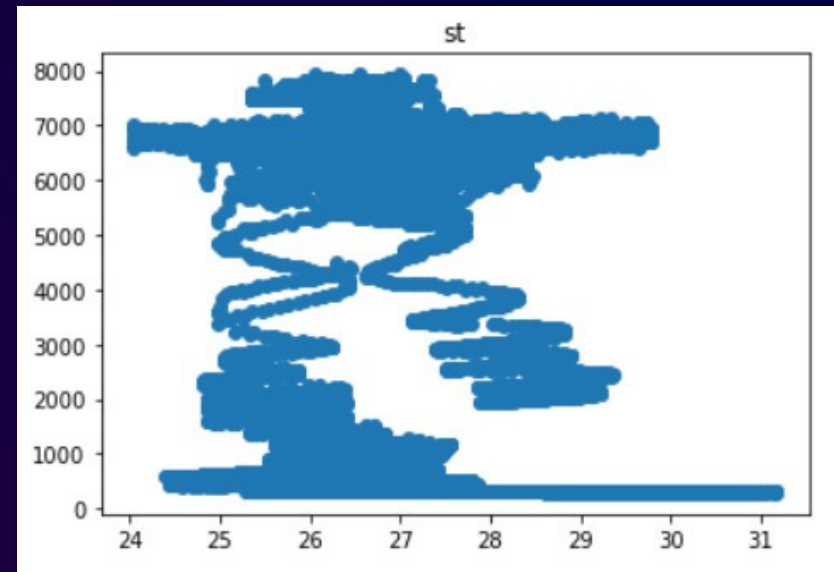
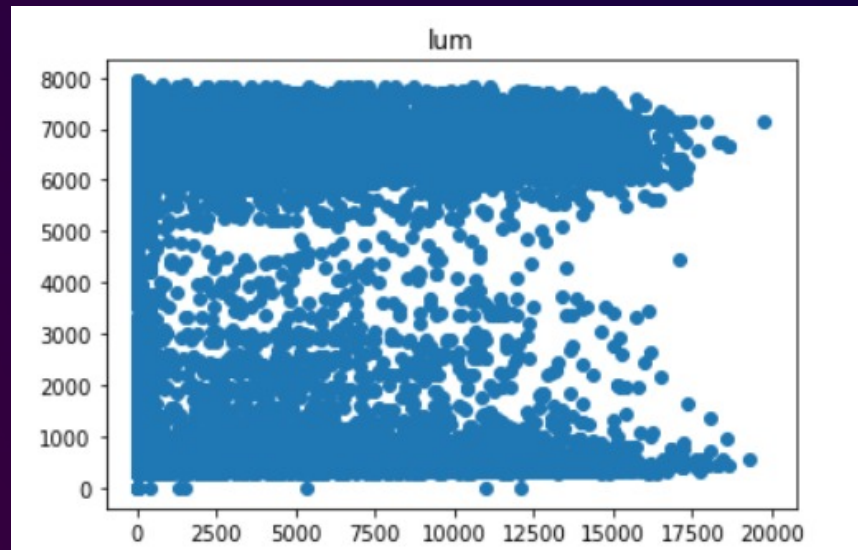
# Preprocessed data\_2

Scatter plots of sm against the other features in the dataset.

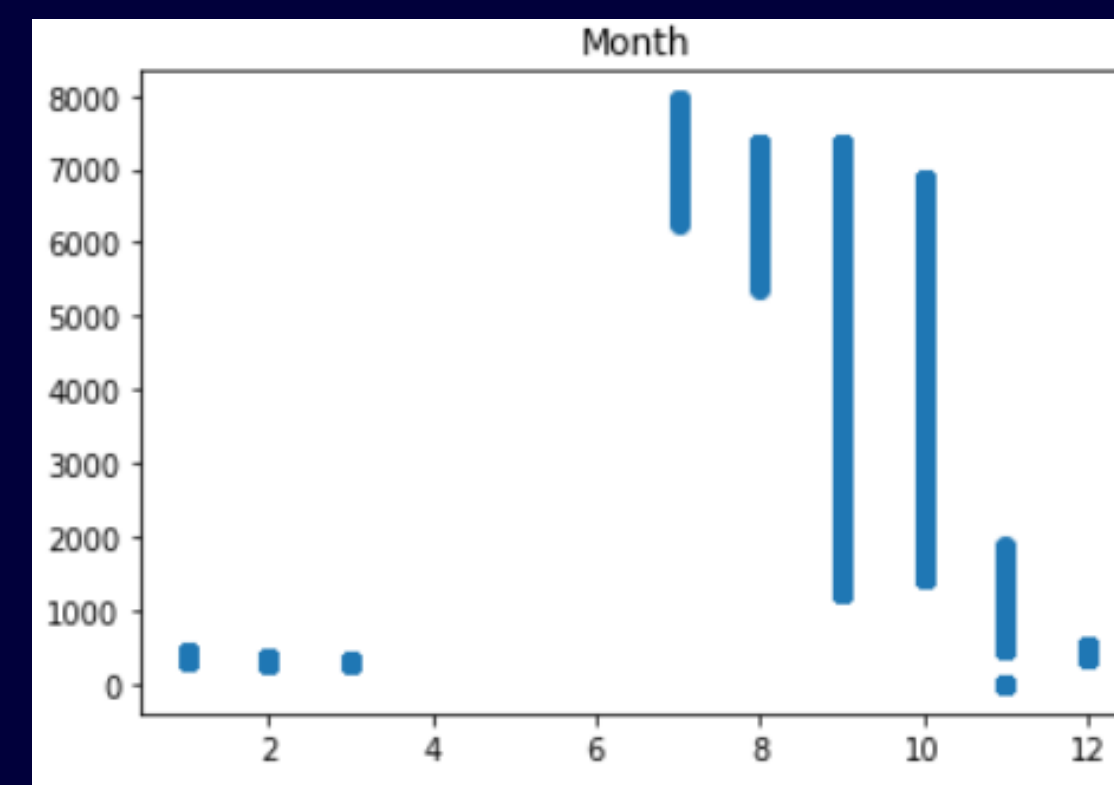
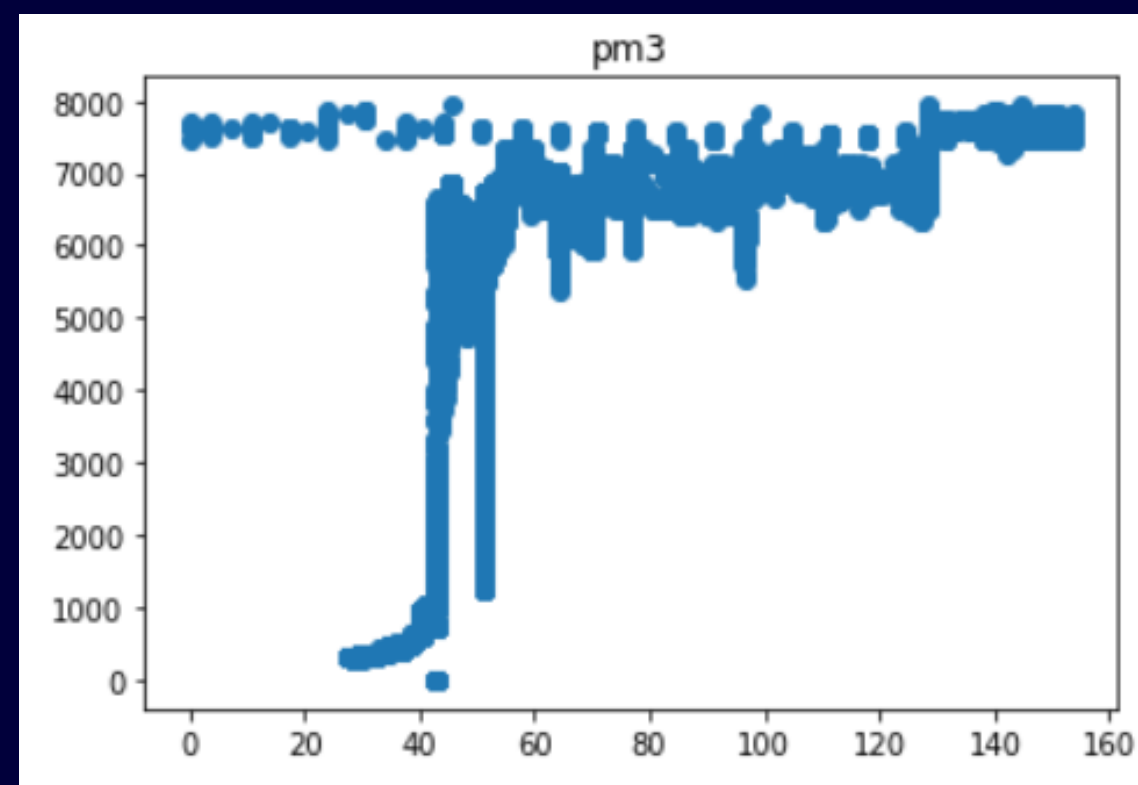
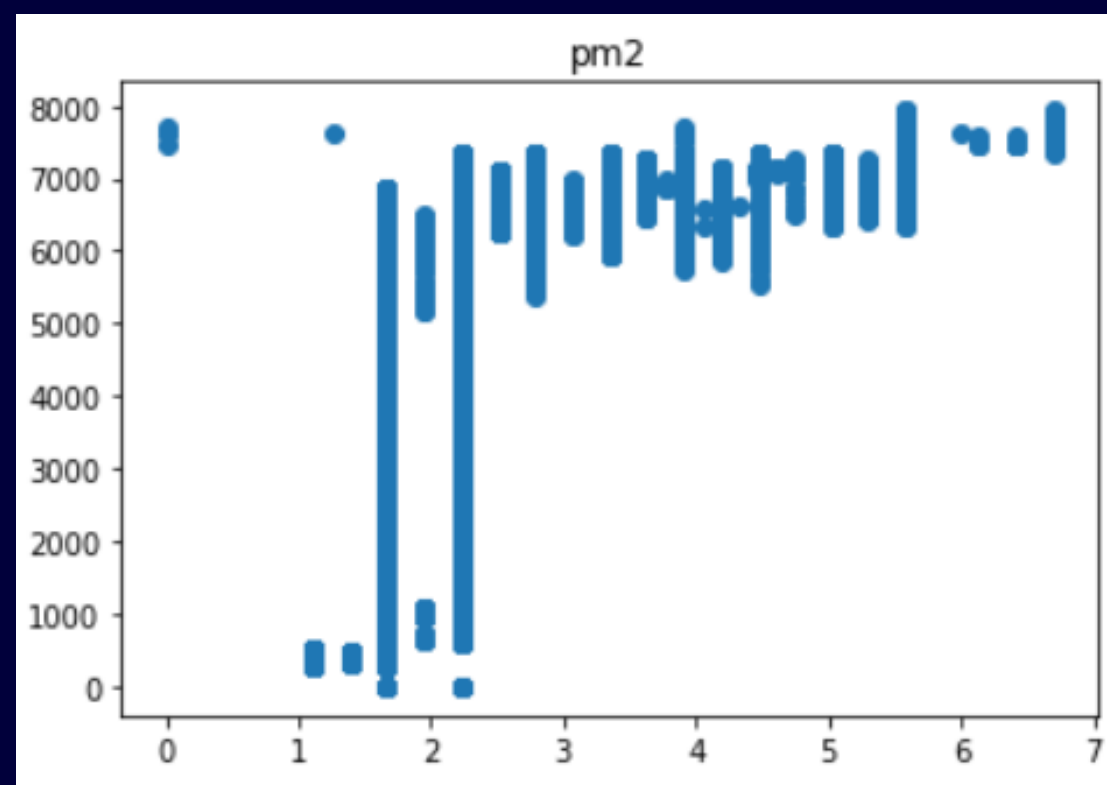
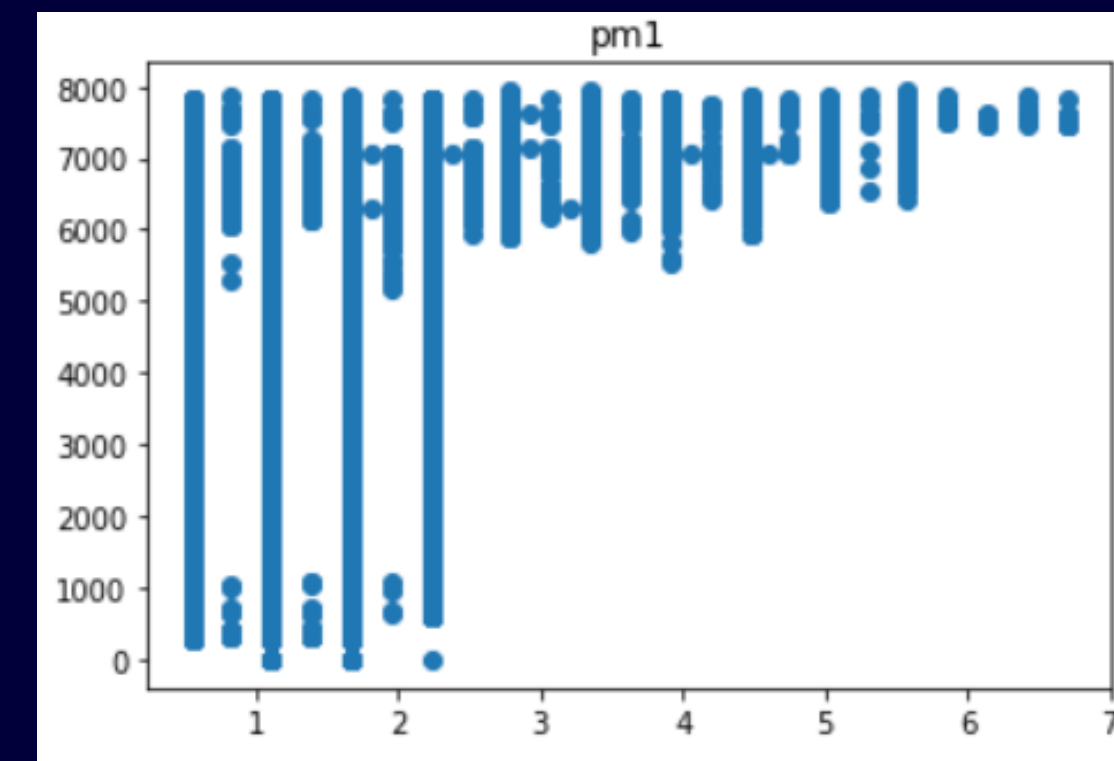
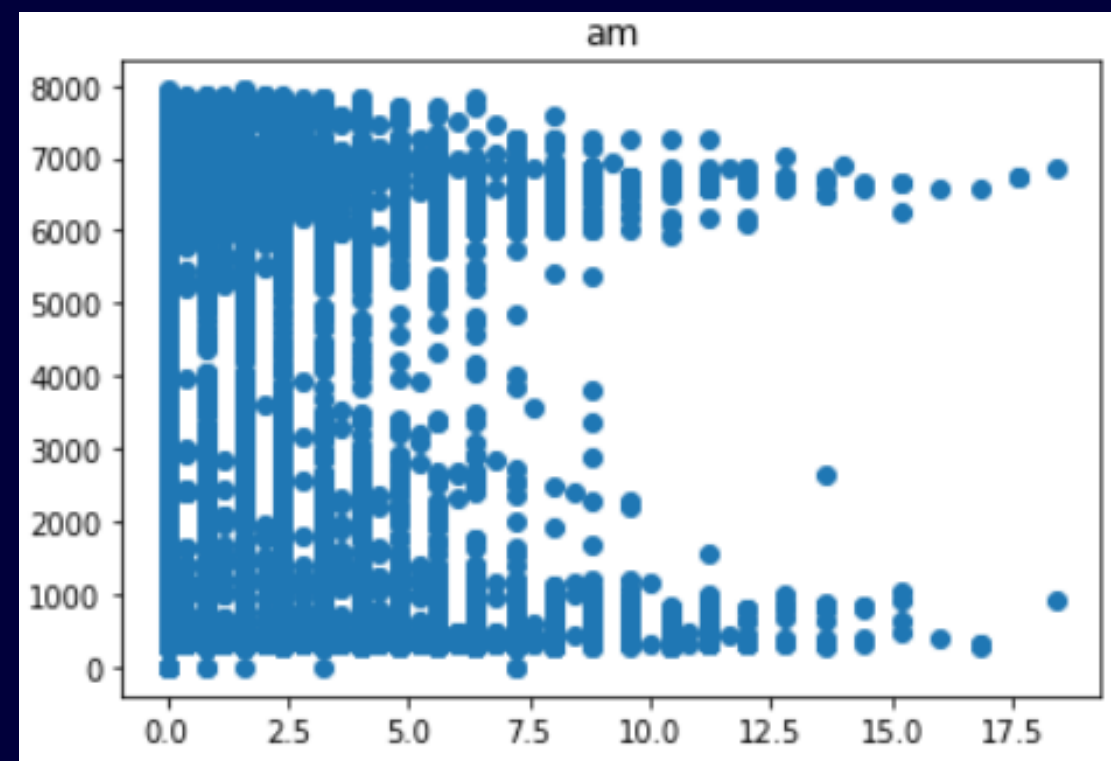
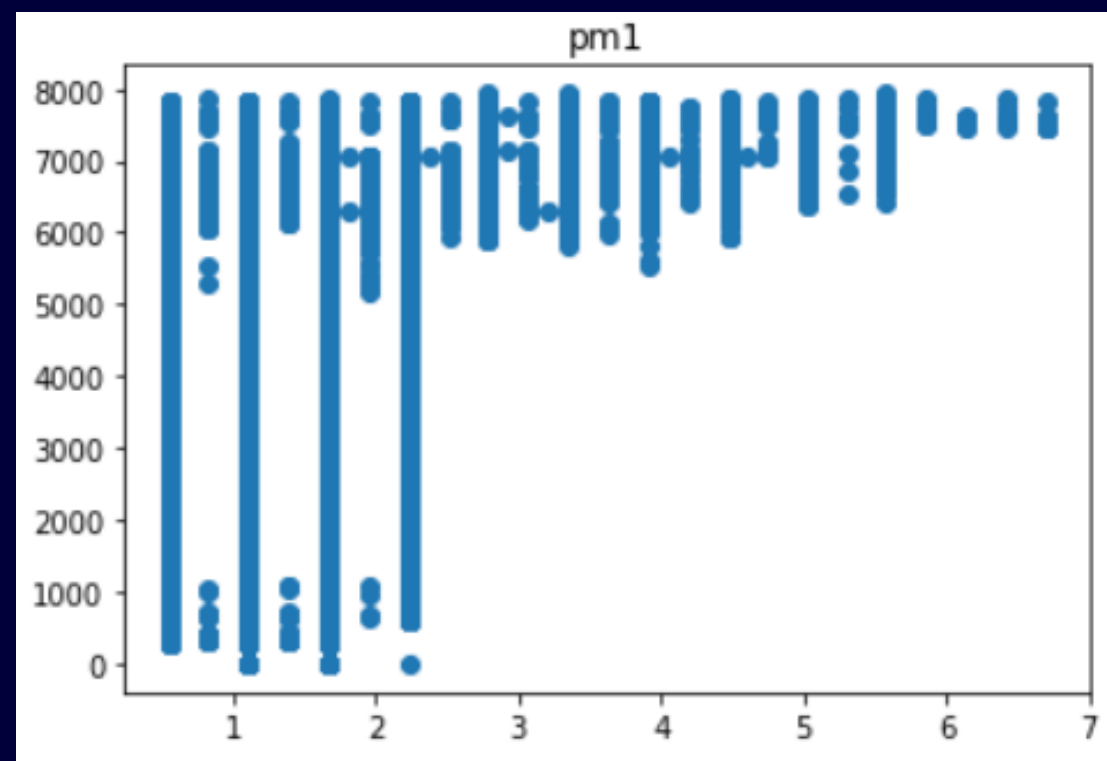


# Merged\_Dataset(Final dataset)

After combining clean data1 and data2 and averaging out the similar values to get the final dataset with all parameters we used to build the model to predict the soil moisture values.





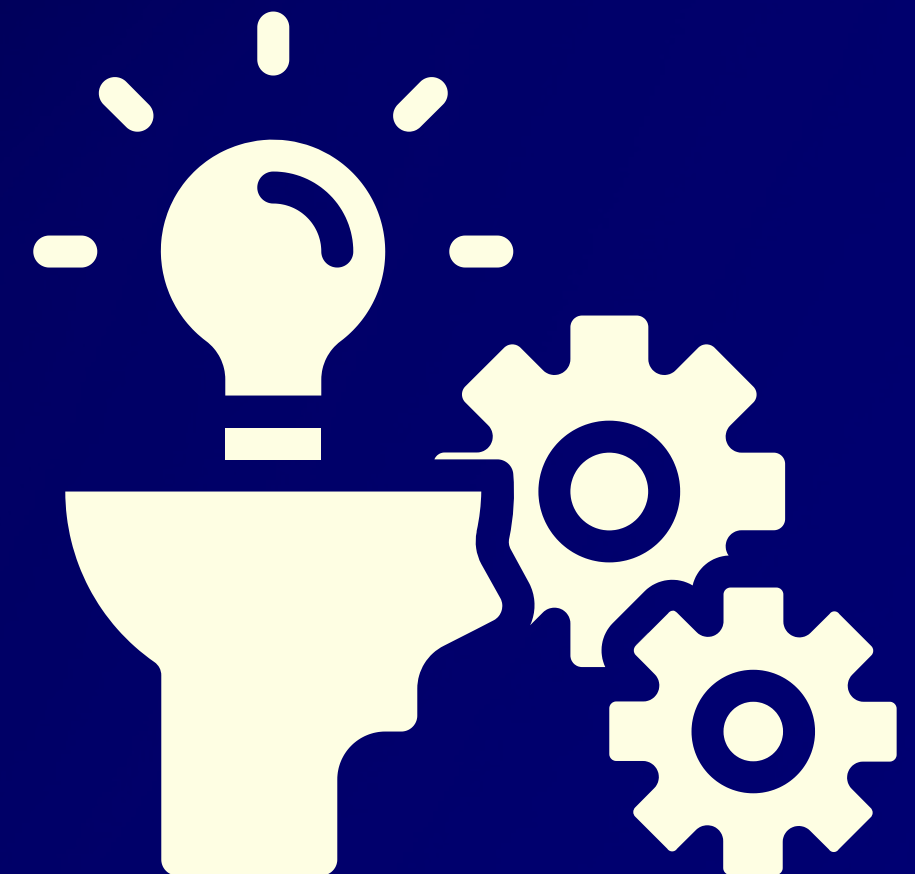




# PREPROCESSING OF FINAL DATA

```
st      13697
day      0
Month    0
year     0
hour     0
min      0
temp    12872
humd     12872
pres     12872
pm1      0
pm2      0
pm3      0
am        0
sm        0
lum       0
dtype: int64
```

- These are the final features in the final dataset with null values.
- These null values are imputed in the model using mean strategy



# MODEL SELECTION

## Random Forest Regressor

### RFR IMPUTATION

MAE : 32.106408898305084  
score: 0.9995056557239824

### RFR AN EXTENSION TO IMPUTATION

MAE : 31.790357067191284  
score: 0.99951847418523

## RFR Pipeline

### RFR PIPELINE IMPUTATION

MAE: 32.094662076271185  
score: 0.9995066681120112

### RFR PIPELINE AN EXTENSION TO IMPUTATION

MAE: 31.739522397094426  
score: 0.9995207551616186

## XG Boost

### XG BOOST IMPUTATION

MAE: 40.77477760407307  
score: 0.9992572415266866

### XG BOOST AN EXTENSION TO IMPUTATION

MAE: 43.111823557652805  
score: 0.9992541126453306

# SELECTED MODEL (Best Model)

**RFR Pipeline An Extension  
to Imputation**

**SimpleImputer**  
strategy="mean"

**RandomForestRegressor**  
n\_estimators=1000,  
random\_state=32

**Imputation  
with  
classification**

**MAE:** 31.739522397094426  
**Score:** 0.9995207551616186

# RESULTS

## Proposed Solution

Our model is capable of predicting soil moisture from 11th March 2023 to 31st March 2023.





The background features a series of overlapping, wavy lines that create a sense of motion and depth. The lines transition in color from a vibrant pink on the left, through a deep purple in the center, to a bright blue on the right. The lines are thin and numerous, creating a complex, layered effect.

**THANK  
YOU**