# Day-ahead Electricity Demand Forecasting

Boylan Pardosi

# Why day-ahead forecasting is important?

- Optimizing Energy Consumption
- Reducing Operational Costs
- Enhancing Grid Stability

# Forecasting Horizon

- Short term Forecasting (Hourly, Daily)
- Medium term Forecasting (Weekly, Monthly)
- Long term Forecasting (Annual, Multi-year)

# Common Approaches

- Statistical Methods :

  Time Series Analysis (ARIMA, SARIMA)

- Machine Learning :

  Linear Regression, Decision Trees, Random Forrest

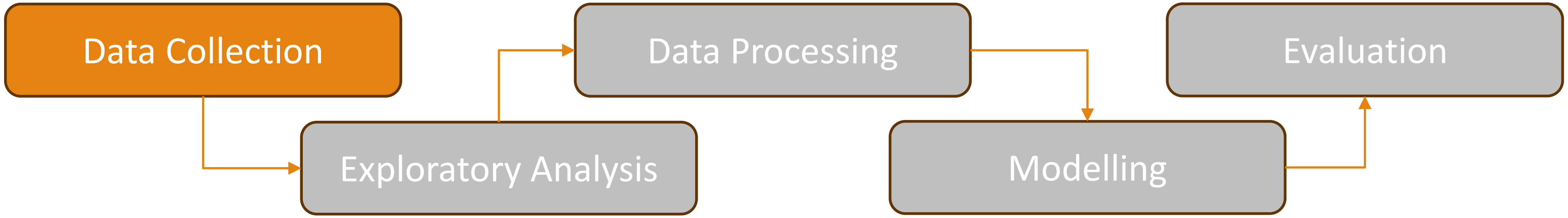
  LSTM, XGBoost

# Why Using XGBoost Model ?

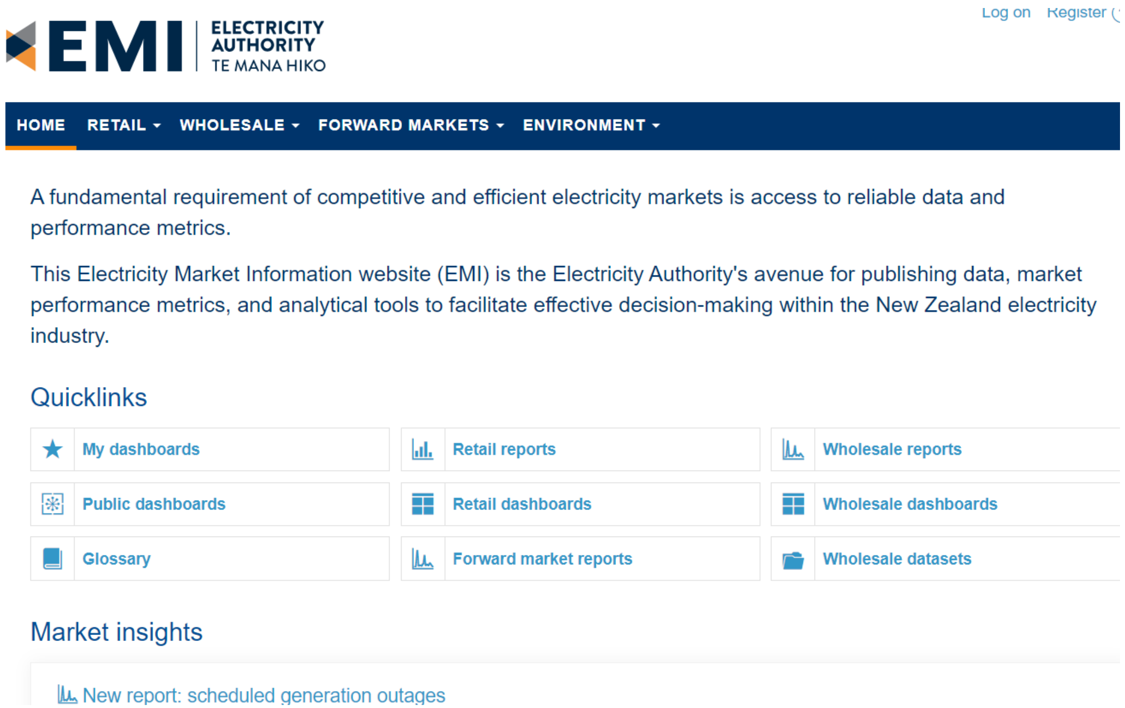
- Handle complex relationship
- High Predictive Performance
- Robustness to overfitting
- Efficiency

# Analytical Workflow

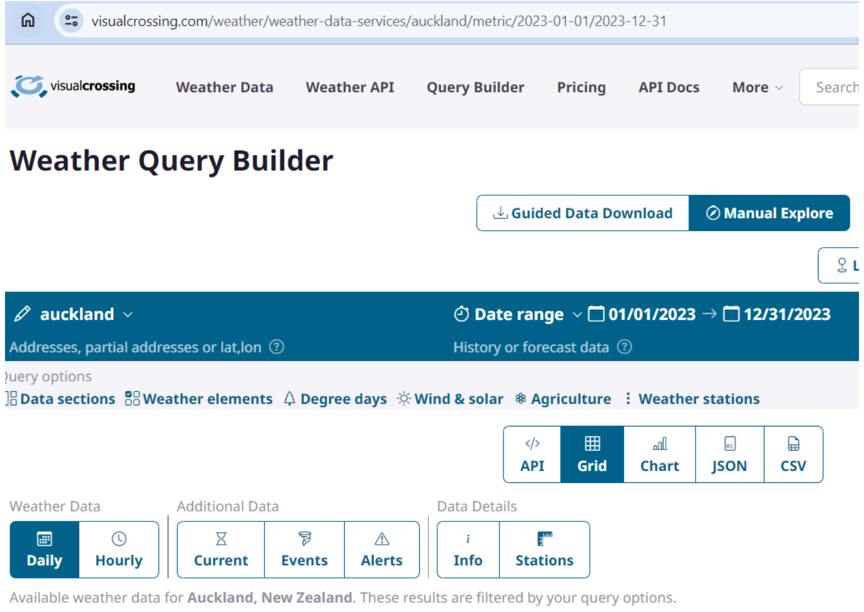
- Define Objective
- Data Collection
- Exploratory Analysis
- Data Processing
- Modelling
- Evaluation



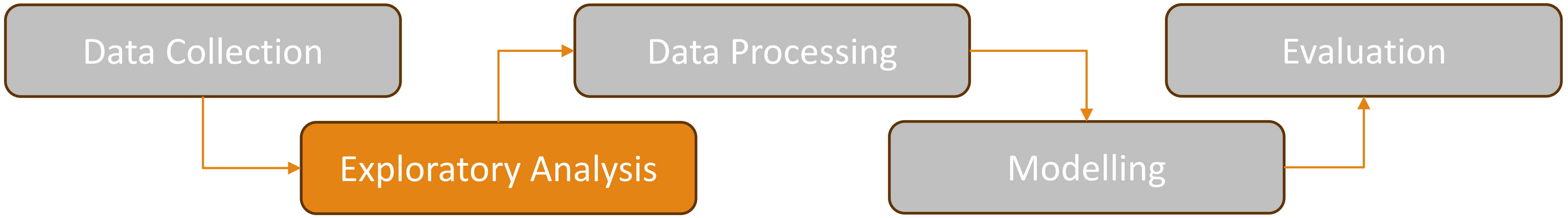

## Electricity Demand Data

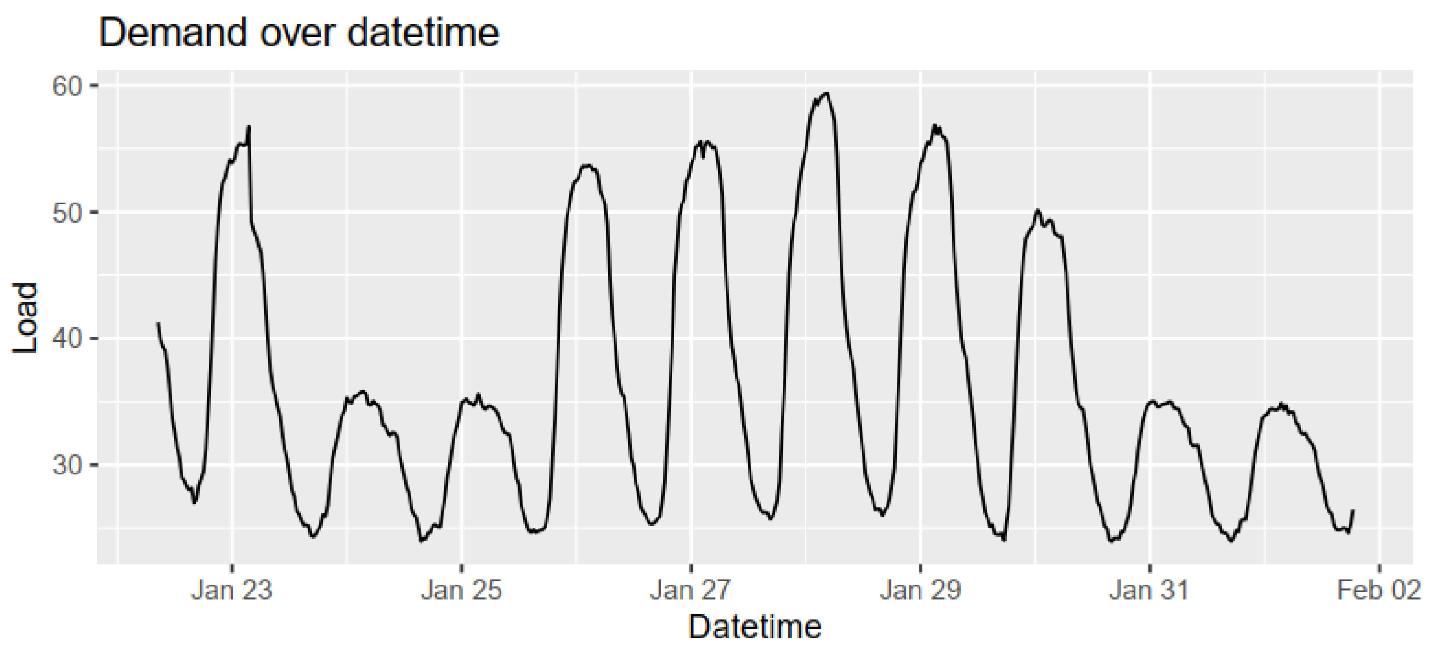


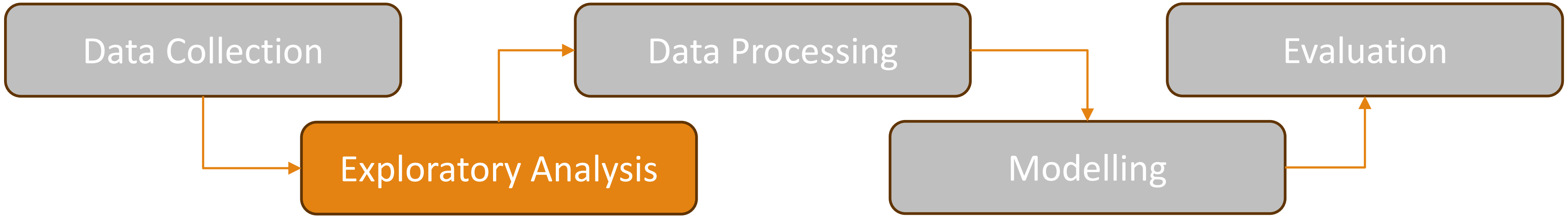

## Weather Data

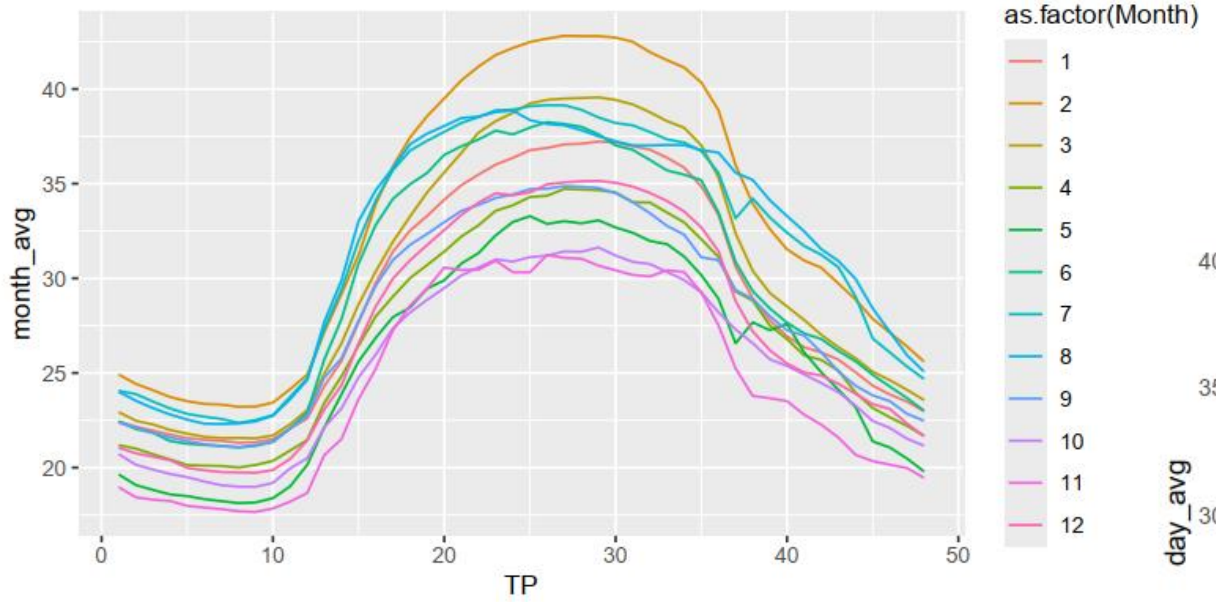


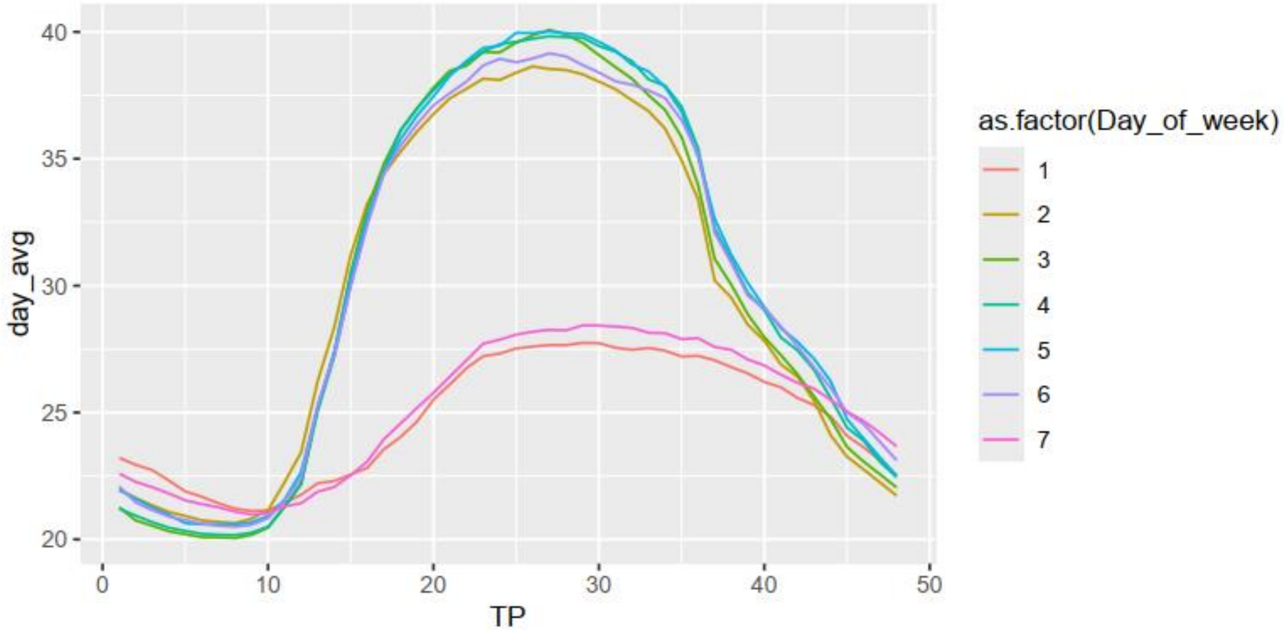

| datetime | tempmax | tempmin | temp | feelslikemax | feelslikemin | feelslike | dew | humidity | precip | precipprob | precipcover |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2023-01-01 | 22.2 | 16.9 | 19.1 | 22.2 | 16.9 | 19.1 | 12.7 | 66.7 | 1.342 | 100 | 8.33 |
| 2023-01-02 | 21 | 16.1 | 18.3 | 21 | 16.1 | 18.3 | 13.3 | 72.7 | 0.358 | 100 | 8.33 |
| 2023-01-03 | 22.8 | 17 | 19.8 | 22.8 | 17 | 19.8 | 14.1 | 70.1 | 0.008 | 100 | 4.17 |
| 2023-01-04 | 21.1 | 18 | 19 | 21.1 | 18 | 19 | 16.6 | 86.2 | 2.582 | 100 | 20.83 |
| 2023-01-05 | 21.8 | 18.1 | 19.6 | 21.8 | 18.1 | 19.6 | 18.2 | 92 | 26.704 | 100 | 25 |

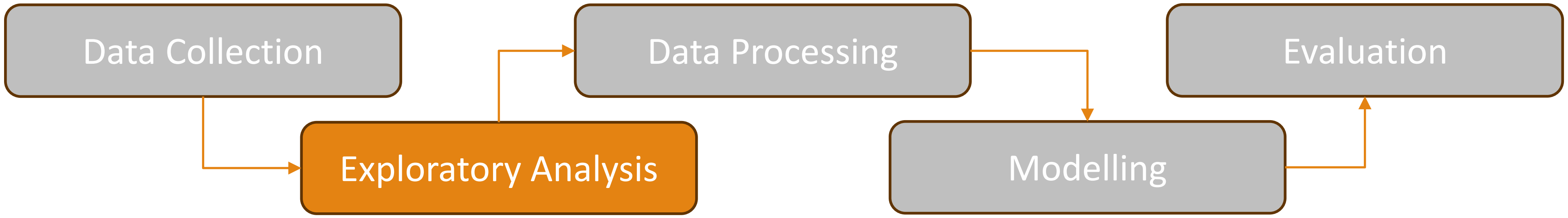

Data Collection
Exploratory Analysis
Data Processing
Modelling
Evaluation

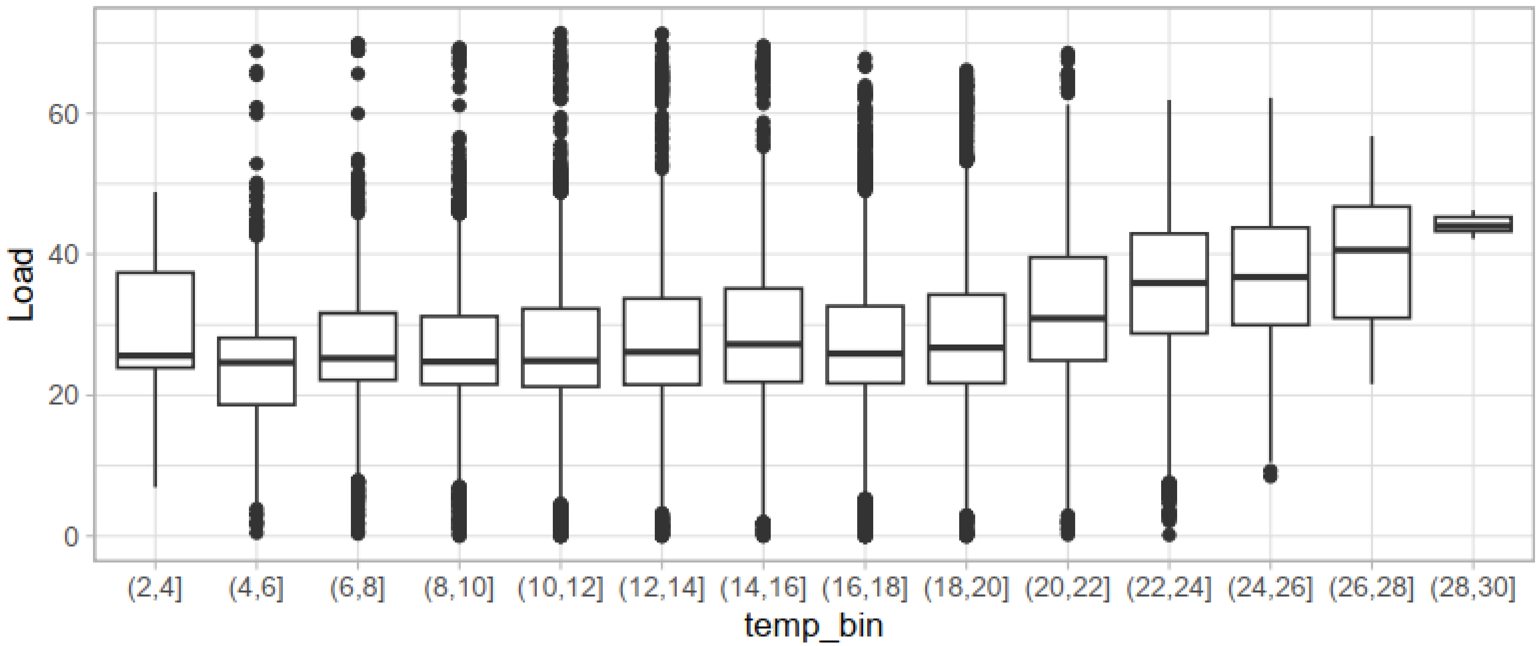

Demand over datetime
60
50
40
30
Load
Jan 23
Jan 25
Jan 27
Jan 29
Jan 31
Feb 02
Datetime

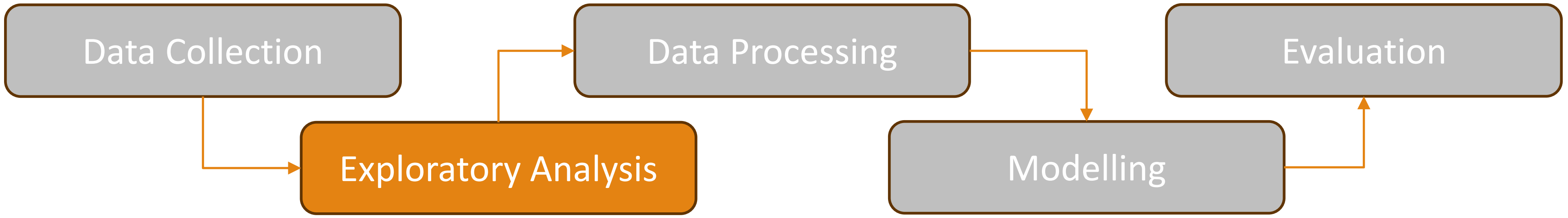

Data Collection
Exploratory Analysis
Data Processing
Modelling
Evaluation

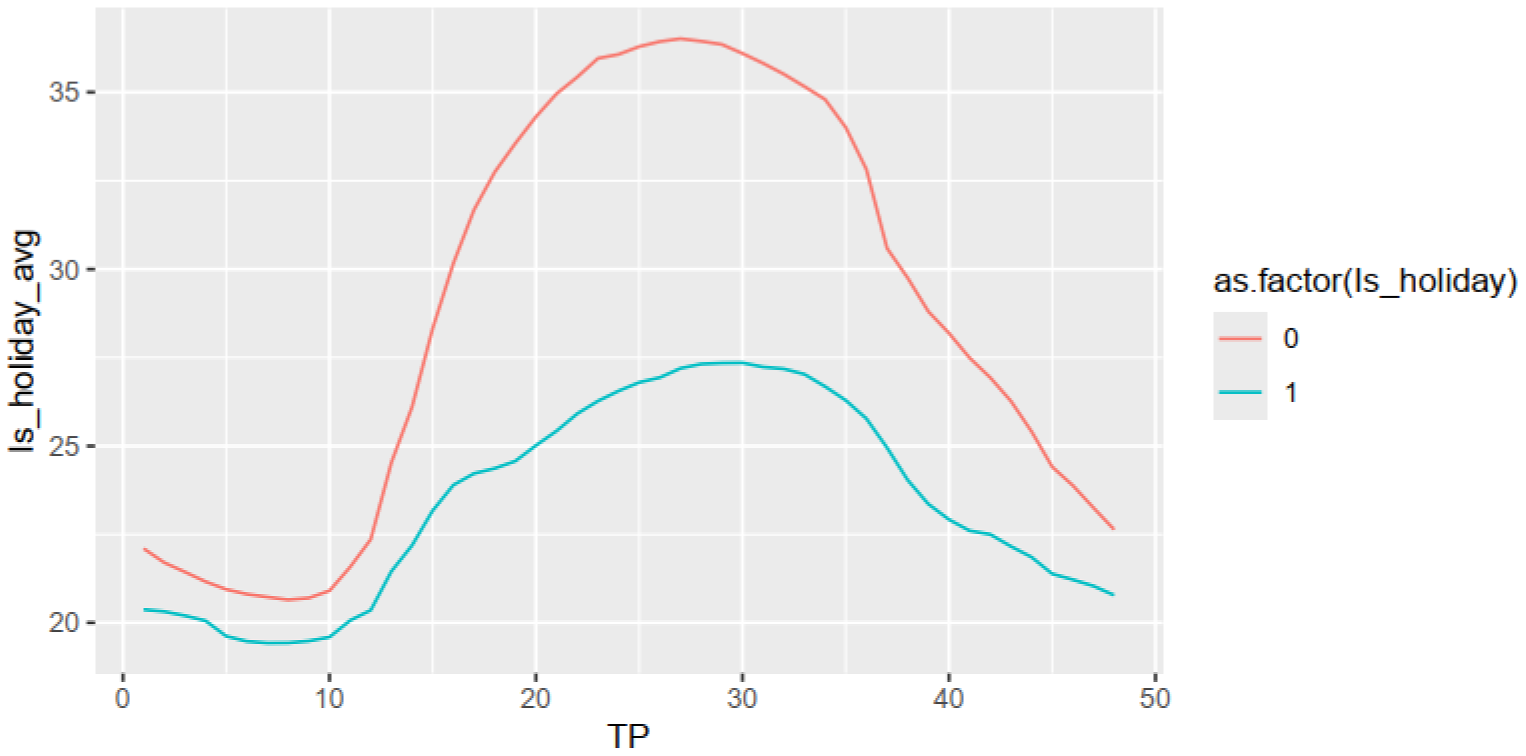

month_avg
TP
as.factor(Month)
1
2
3
4
5
6
7
8
9
10
11
12

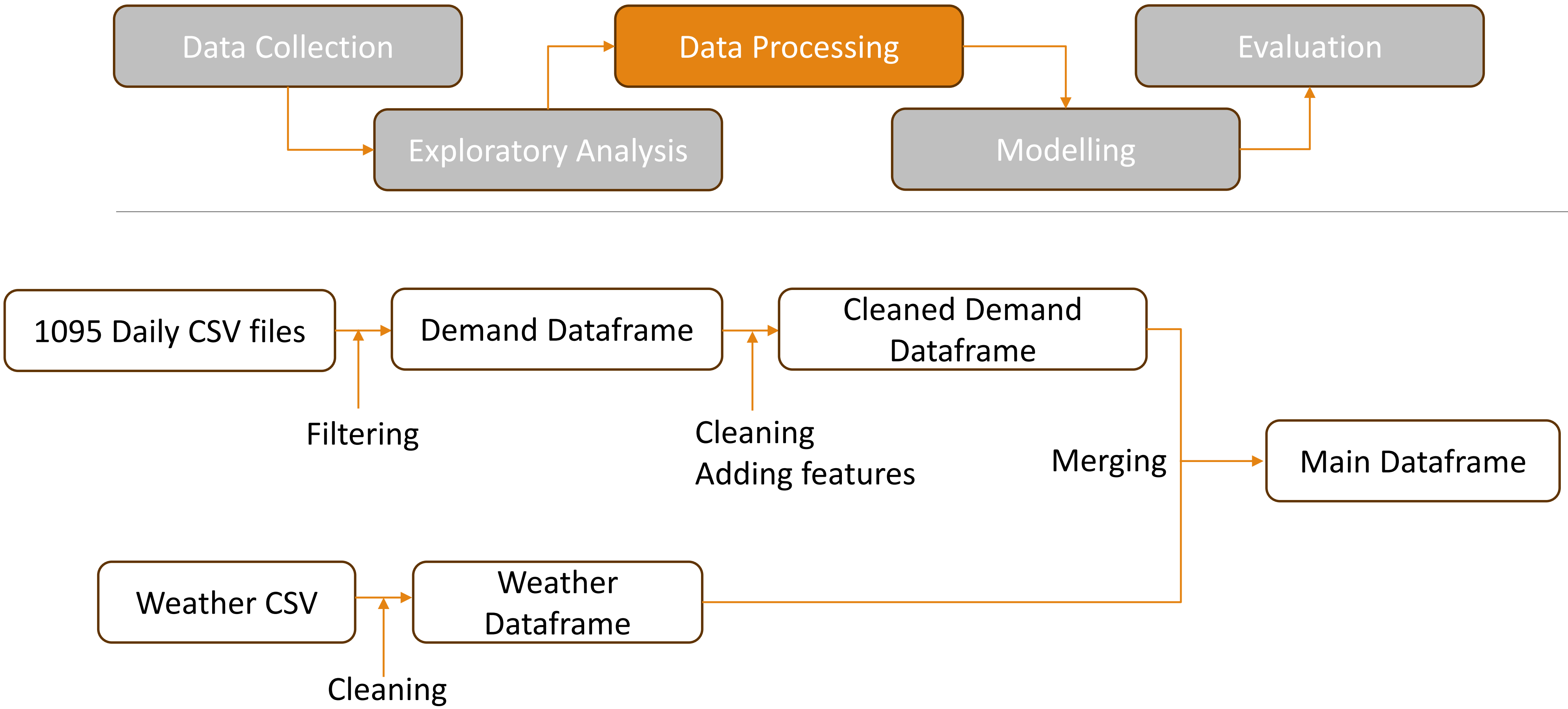

day_avg
TP
as.factor(Day_of_week)
1
2
3
4
5
6
7

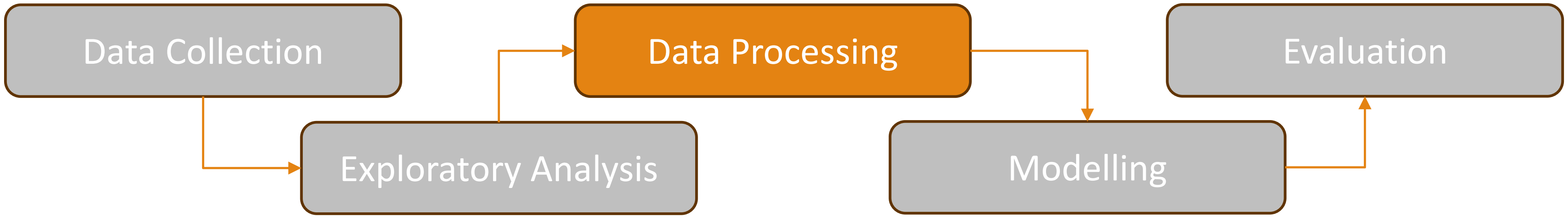

Data Collection
Exploratory Analysis
Data Processing
Modelling
Evaluation


Load
0
20
40
60
(2,4]
(4,6]
(6,8]
(8,10]
(10,12]
(12,14]
(14,16]
(16,18]
(18,20]
(20,22]
(22,24]
(24,26]
(26,28]
(28,30]
temp_bin


Data Collection
Exploratory Analysis
Data Processing
Modelling
Evaluation

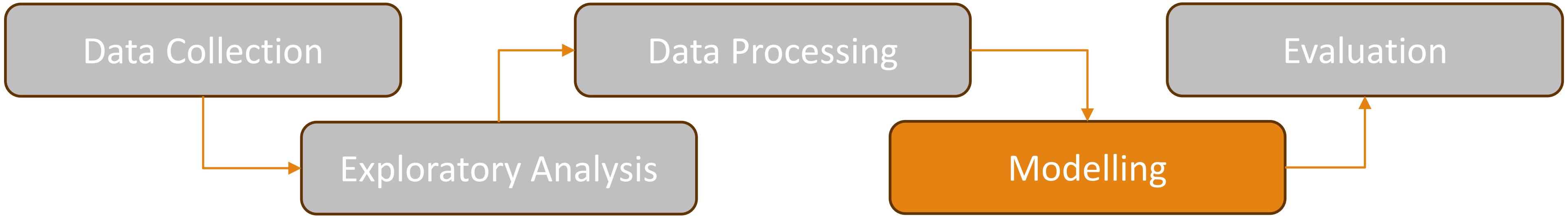

Is_holiday_avg
35
30
25
20
0
10
20
30
40
50
TP
as.factor(Is_holiday)
0
1


Data Collection
Exploratory Analysis
Data Processing
Modelling
Evaluation
1095 Daily CSV files
Filtering
Demand Dataframe
Cleaning
Adding features
Cleaned Demand Dataframe
Merging
Main Dataframe
Weather CSV
Cleaning
Weather Dataframe




Some issues to address:

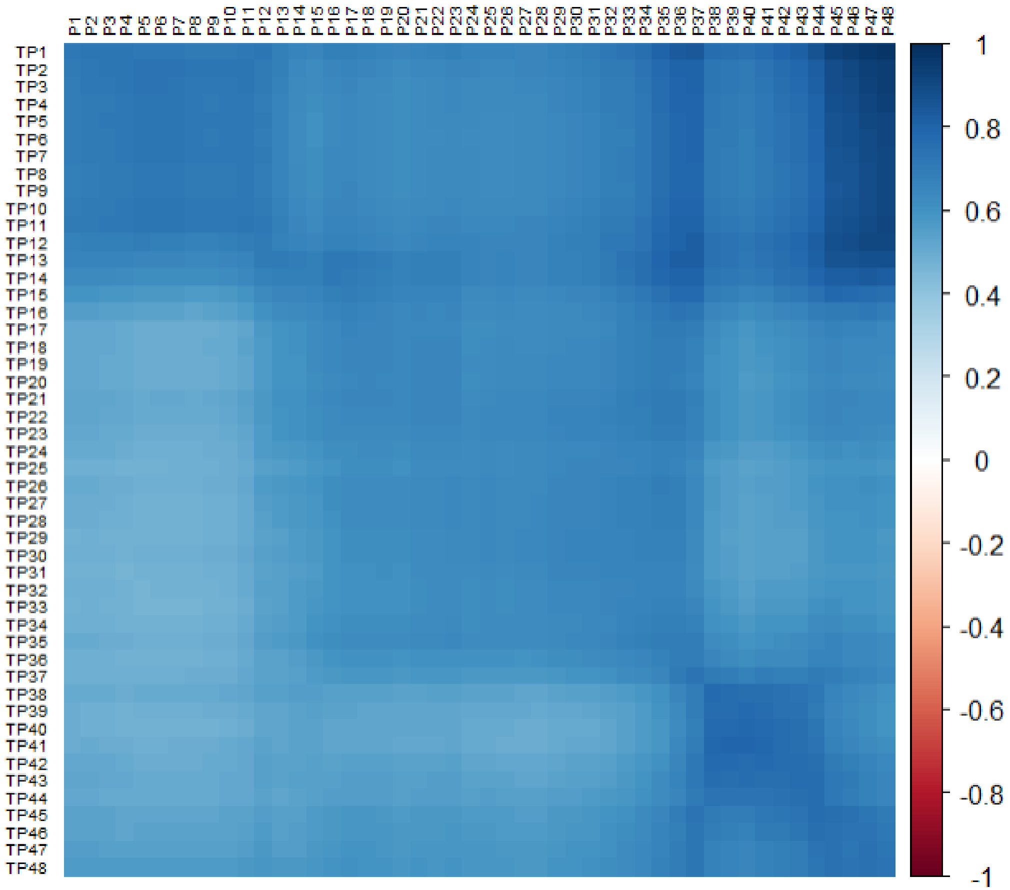
- Demand Null Values → Use the average of the same groups (use zoo)
- Lower time resolution in weather data → insert in between rows and interpolate

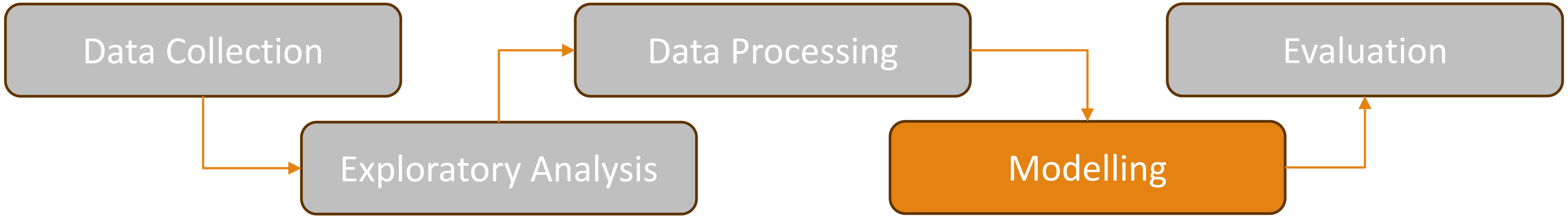



# Feature Selection (1) :

- Lag features

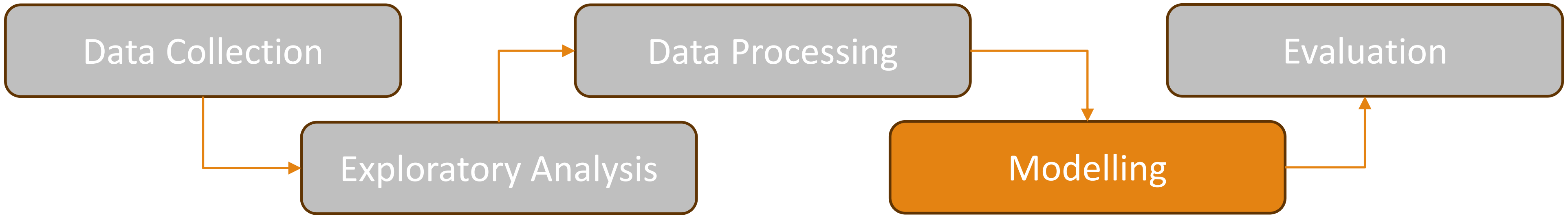
Use correlation matrix





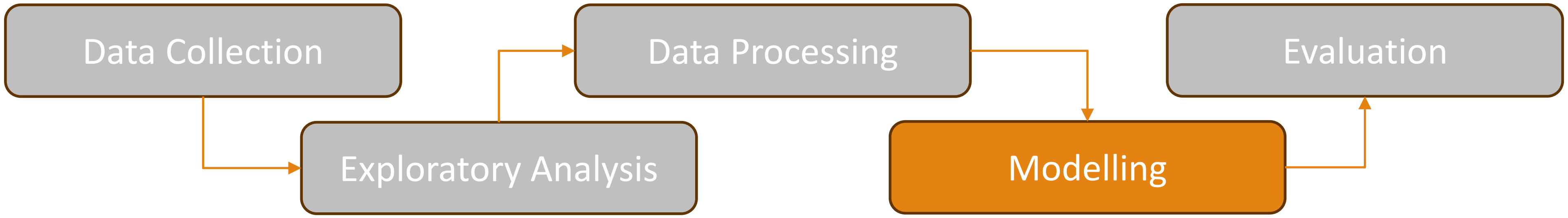

## Feature Selection (2) :

- Day of week
- Month
- Holiday flag
- Temperature
- Humidity



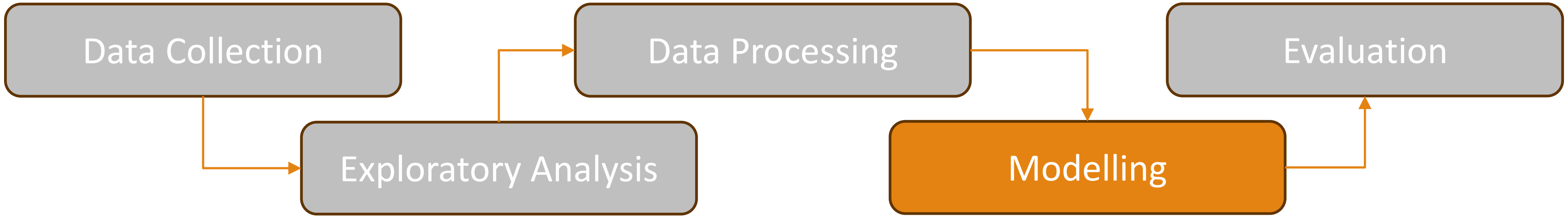

XGBoost Parameter selection :

- Use hyper parameter tuning with gridsearch cross validation
- Different parameters for 48 prediction target

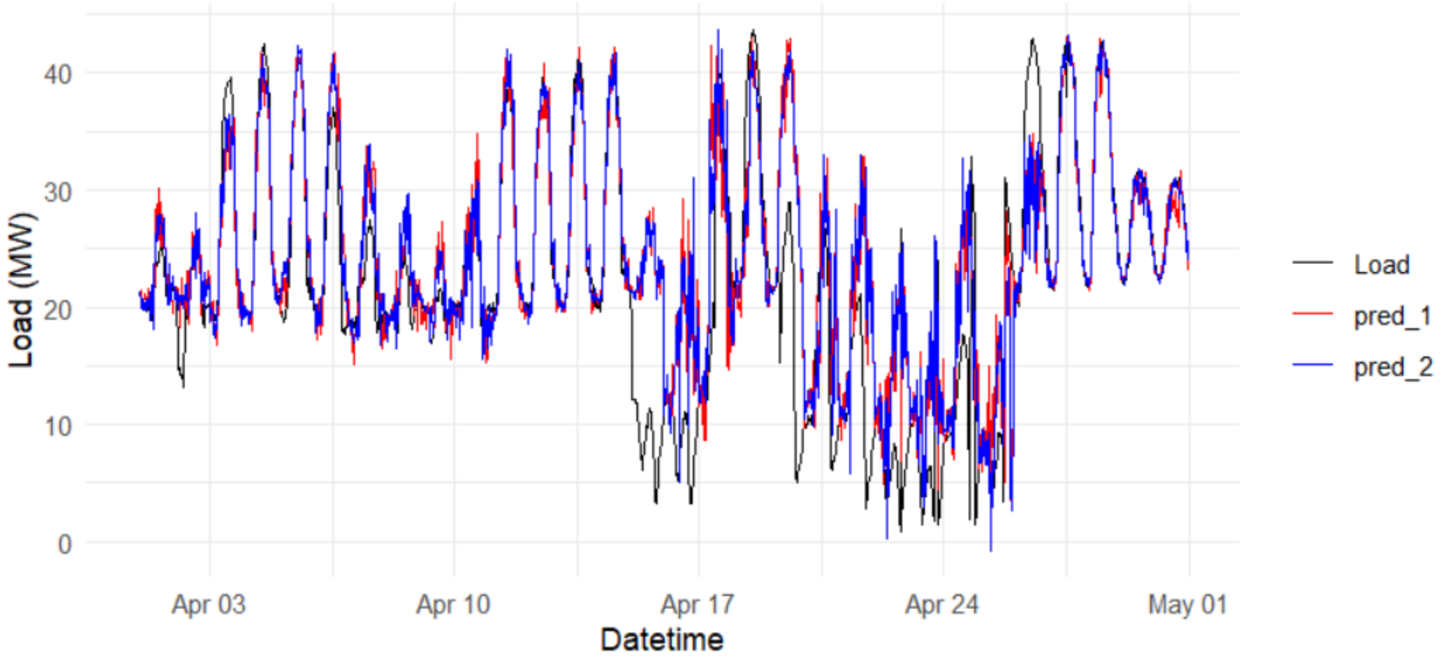


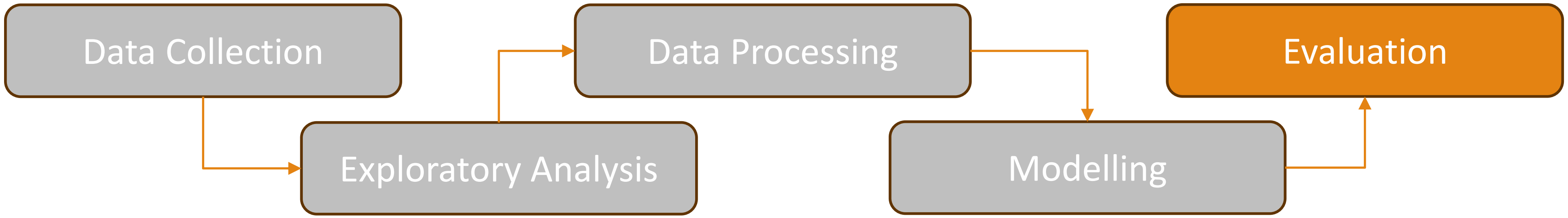

Dut to potential inaccuracy of the weather data,
we consider two models :

- With temperature data
- Without weather data




# Prediction result






# Error measurements

## Model 1 (with weather features)

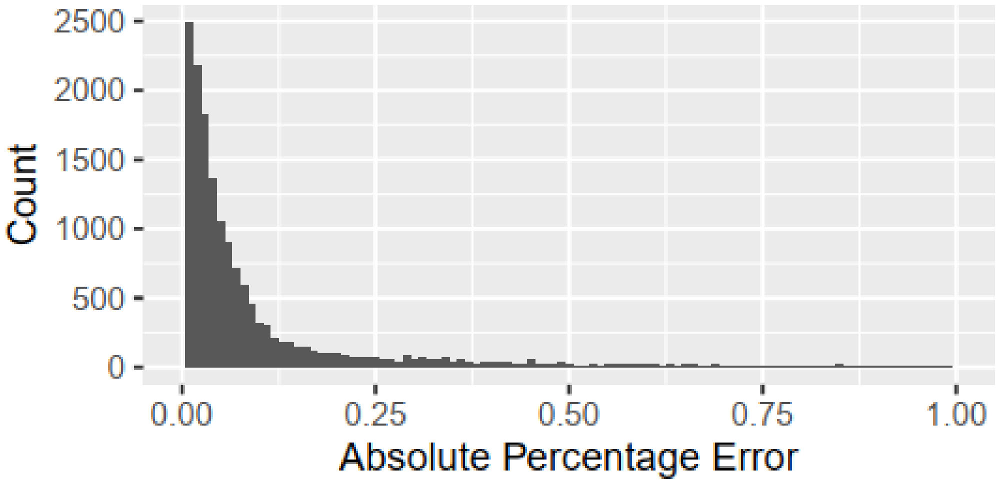
| ape_1 |
|---|
| Min. : 0.00 |
| 1st Qu.: 0.02 |
| Median : 0.05 |
| Mean : 0.23 |
| 3rd Qu.: 0.11 |
| Max. :59.38 |
| NA's :35040 |

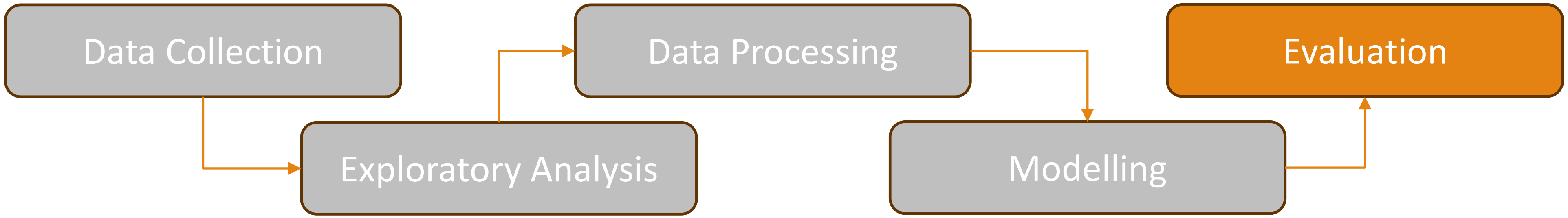
| se_1 |
|---|
| Min. : 0.00 |
| 1st Qu.: 0.26 |
| Median : 1.66 |
| Mean : 29.02 |
| 3rd Qu.: 9.17 |
| Max. :1605.97 |
| NA's :35040 |

## Model 2 (without weather features)

| ape_2 |
|---|
| Min. : 0.00 |
| 1st Qu.: 0.02 |
| Median : 0.04 |
| Mean : 0.23 |
| 3rd Qu.: 0.10 |
| Max. :60.61 |
| NA's :35040 |

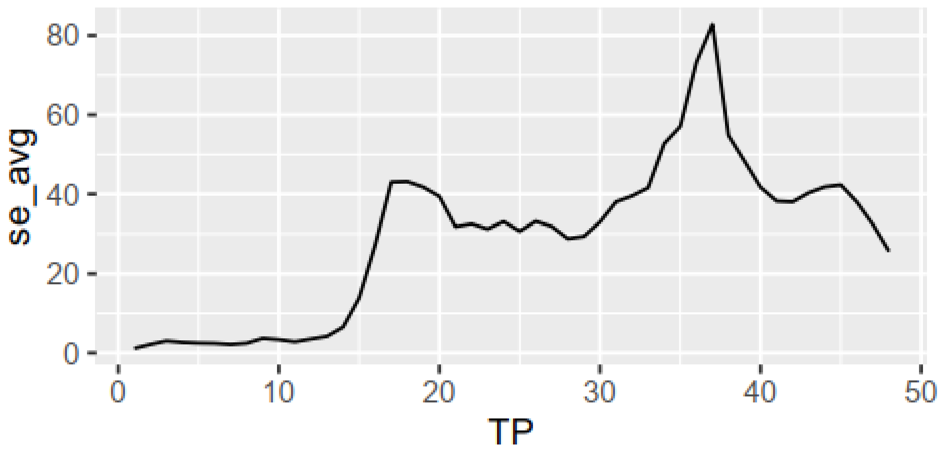
| se_2 |
|---|
| Min. : 0.00 |
| 1st Qu.: 0.23 |
| Median : 1.40 |
| Mean : 29.31 |
| 3rd Qu.: 8.25 |
| Max. :1759.06 |
| NA's :35040 |



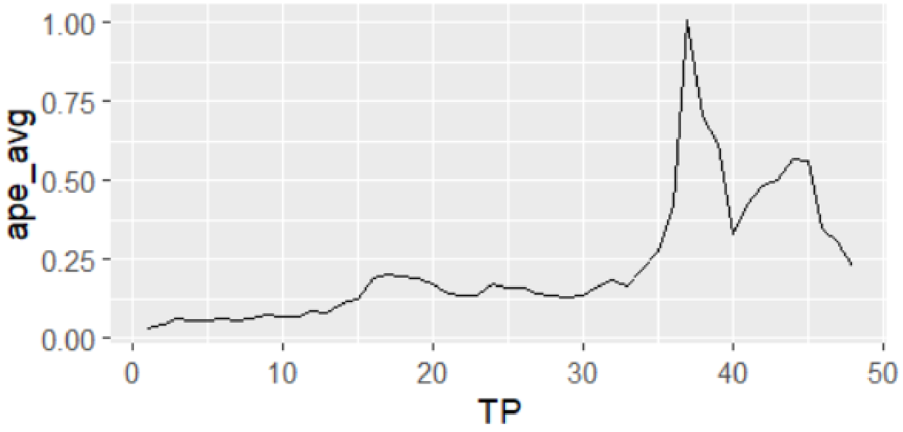


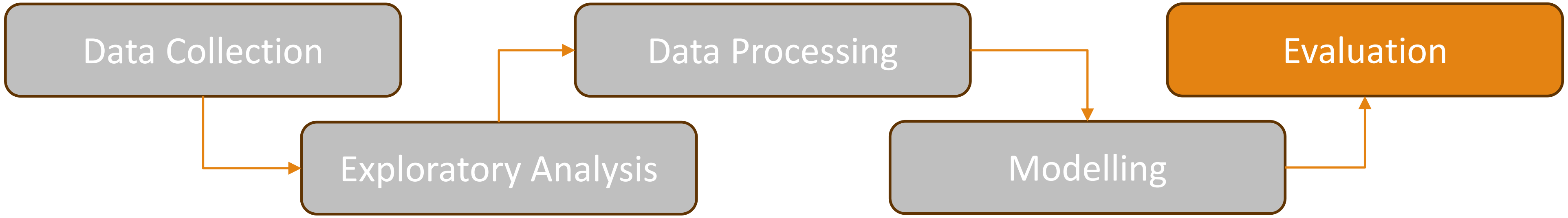

## Error measurements vs TP

### Squared Error




### Absolute Percentage Error





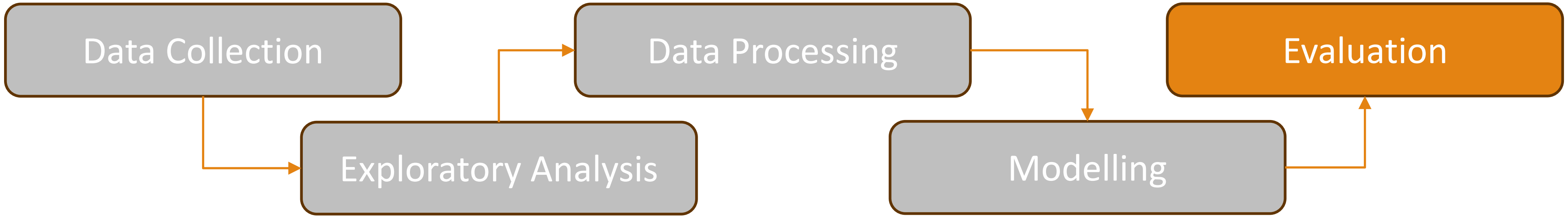

# Feature Importance

TP =1

| Feature | Gain | Cover | Frequency |
|---|---|---|---|
| time_lag_1 | 0.5716362 | 0.3978664 | 0.3421053 |
| time_lag_2 | 0.3508768 | 0.3933362 | 0.2960526 |
| Month | 0.0602796 | 0.0673301 | 0.1052632 |
| temp | 0.0067493 | 0.0485289 | 0.1019737 |
| humidity | 0.0063271 | 0.0708627 | 0.0855263 |
| Day_of_week | 0.0036240 | 0.0219231 | 0.0625000 |
| Is_holiday | 0.0005069 | 0.0001526 | 0.0065789 |

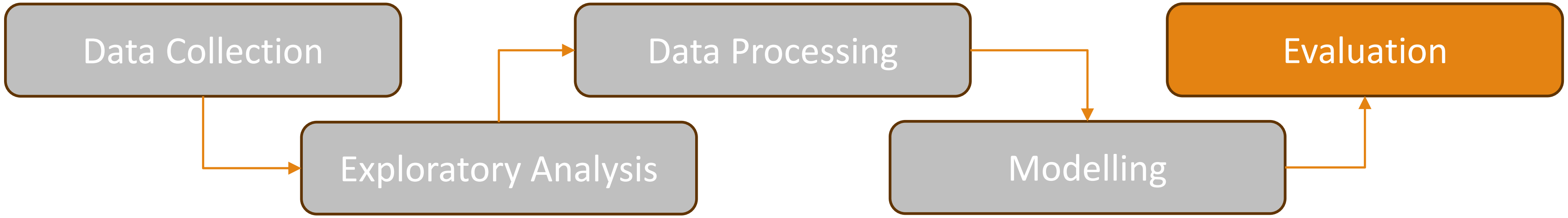
TP =30

| Feature | Gain | Cover | Frequency |
|---|---|---|---|
| time_lag_2 | 0.5356285 | 0.2661932 | 0.2452830 |
| Day_of_week | 0.2016390 | 0.1385776 | 0.1415094 |
| time_lag_1 | 0.1551237 | 0.2885125 | 0.2232704 |
| temp | 0.0406069 | 0.1275923 | 0.1509434 |
| humidity | 0.0306791 | 0.1007975 | 0.1446541 |
| Month | 0.0227920 | 0.0344904 | 0.0628931 |
| Is_holiday | 0.0106729 | 0.0307937 | 0.0220126 |
| Is_holiday_prev | 0.0028579 | 0.0130429 | 0.0094340 |

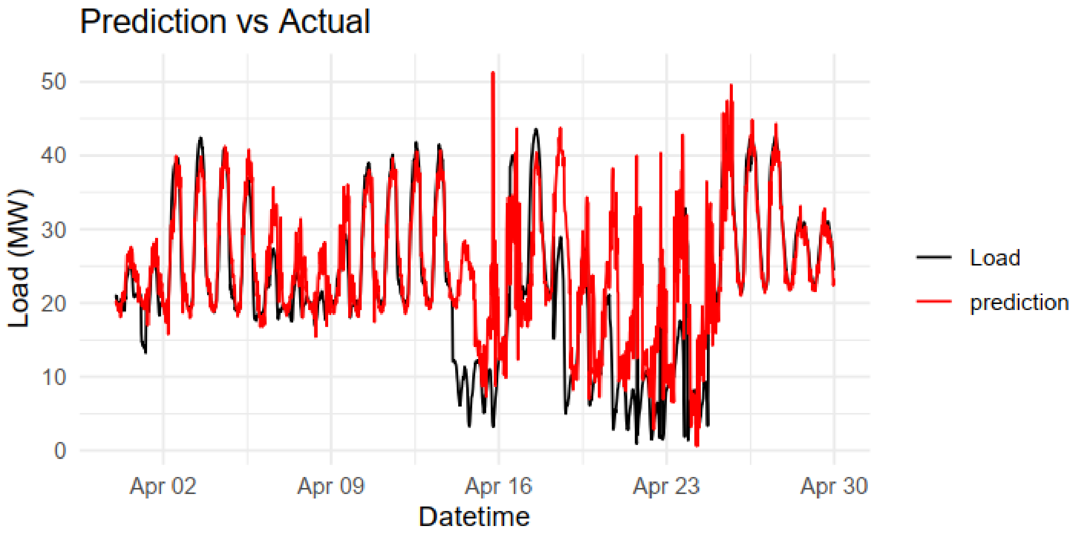


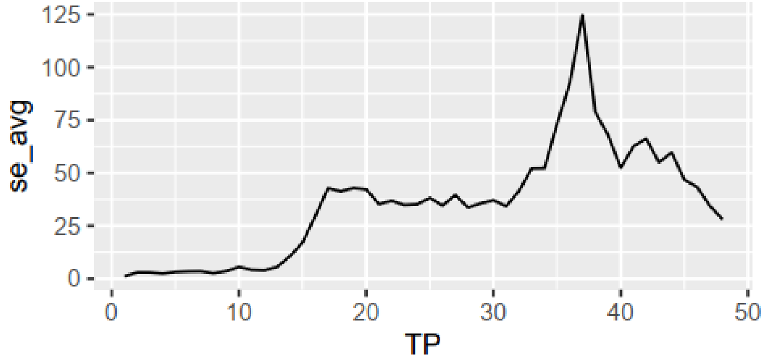

Compare with other algorithm: LSTM

- Similar step with additional step: data transformation (one hot encoding and minmax scaling)
- Use 50 unit layer and ‘relu’ activation
- Use 50 epochs with 32 batch size

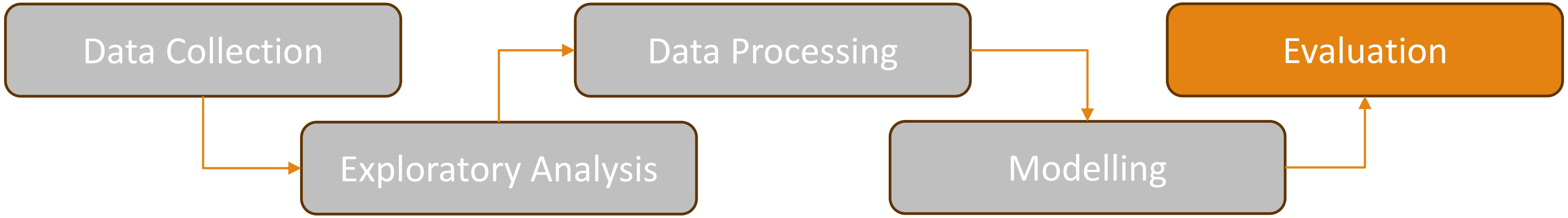



# Result: LSTM




| se | ape |
|---|---|
| Min. : 0.0000 | Min. : 0.00000 |
| 1st Qu.: 0.4022 | 1st Qu.: 0.02418 |
| Median : 2.5594 | Median : 0.05599 |
| Mean : 35.3817 | Mean : 0.26275 |
| 3rd Qu.: 14.5196 | 3rd Qu.: 0.12859 |
| Max. :2304.5917 | Max. :56.49364 |




Conclusion :

- We can drop the weather features as it do not significantly improve performance
- Most of the time prediction is quite accurate but at very low occurrence error can drastically increased, increasing the error mean
- Comparable result with LSTM with slightly better performance