# HACKATHON ONLINE PM2.5 DATA OF THAILAND

300232-ยงสุข

# DATA

**Data Explorer**

20.1 kB

- sample_submission.csv
- train.csv

**TRAIN --> 1095**

Jan 1st, 2013 to Dec 31st, 2015.

**SUBMISSION --> 366**

Jan 1st, 2016 to Dec 31st, 2016

# MISSING DATA

```
df.isnull().sum()
```

```
Date      0
PM2.5     7
dtype: int64
```



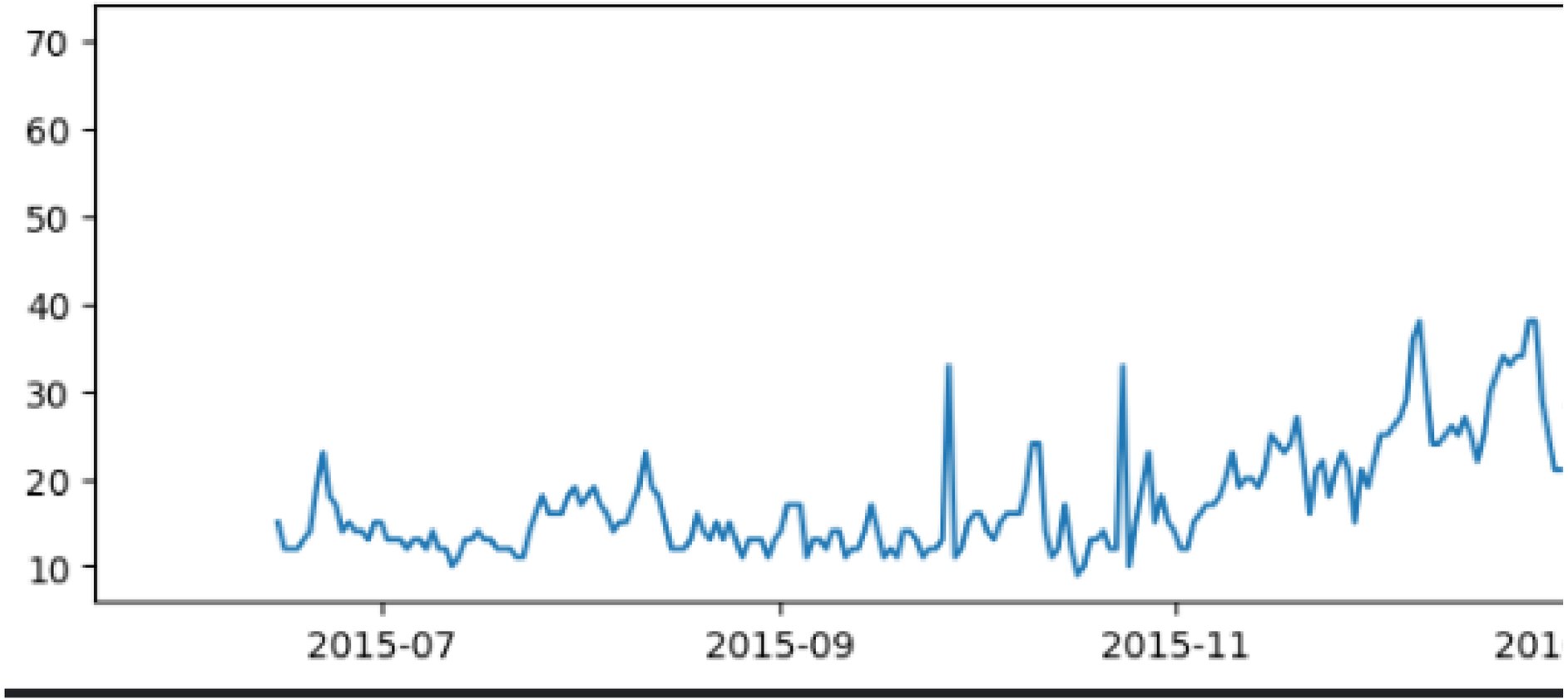
```
# fill missing values with the mean
mean_value = df['PM2.5'].mean()
print(f'mean_value : {mean_value}')

df = df.fillna({'PM2.5': mean_value})

# Check again if there are still NaN values in the data
print(df.isnull().sum())
```

```
mean_value : 32.869485294117645
Date      0
PM2.5     0
dtype: int64
```

# GRAPH



# SETTING DATA

```
# Create a TimeSeriesDataFrame
train_data = TimeSeriesDataFrame.from_data_frame(
    df_format,
    id_column=0,
    timestamp_column='Date'
)
```

| | | PM2.5 |
|---|---|---|
| item_id | timestamp | |
| PM2.5 | 2013-01-01 | 26.0 |
| | 2013-01-02 | 21.0 |
| | 2013-01-03 | 21.0 |
| | 2013-01-04 | 20.0 |
| | 2013-01-05 | 22.0 |
| | ... | ... |
| | 2015-12-27 | 38.0 |
| | 2015-12-28 | 29.0 |
| | 2015-12-29 | 25.0 |
| | 2015-12-30 | 21.0 |
| | 2015-12-31 | 21.0 |

1095 rows × 1 columns


```
predictor = TimeSeriesPredictor(
    prediction_length=366,
    path="Model",
    target="PM2.5",
    eval_metric="RMSE"
)

predictor.fit(
    train_data,
    presets = "best_quality"
)
```

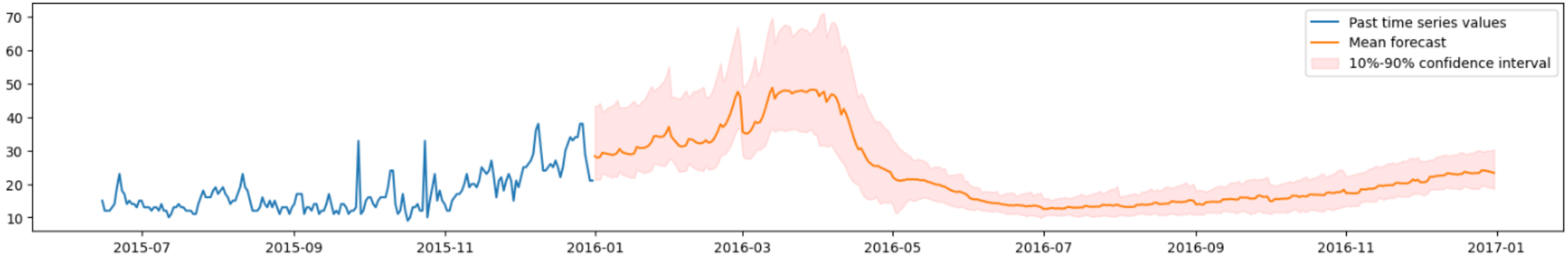
```
Training complete. Models trained: ['Naive', 'SeasonalNaive', 'ETS', 'Theta/T1', 'Theta/T
2', 'ARIMA', 'AutoETS', 'DynamicOptimizedTheta', 'DeepAR/T1', 'DeepAR/T2', 'DeepAR/T3', 'D
eepAR/T4', 'DeepAR/T5', 'DeepAR/T6', 'DeepAR/T7', 'DeepAR/T8', 'DeepAR/T9', 'DeepAR/T10',
'AutoARIMA', 'TemporalFusionTransformer', 'WeightedEnsemble']
Total runtime: 3615.13 s
Best model: WeightedEnsemble
Best model score: -13.2056
```

In [7]:
```
predictions
```

Out[7]:

| | | mean | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| item_id | timestamp | | | | | | | | | | |
| PM2.5 | 2016-01-01 | 28.308779 | 21.331375 | 22.185627 | 22.431164 | 27.162949 | 28.319559 | 31.228170 | 30.229418 | 34.767056 | 43.442341 |
| | 2016-01-02 | 27.825802 | 21.324696 | 22.230389 | 23.533859 | 26.839954 | 27.801414 | 31.115444 | 30.565639 | 33.853561 | 43.226280 |
| | 2016-01-03 | 28.040714 | 21.325983 | 22.481062 | 24.116966 | 26.783602 | 27.978806 | 31.495663 | 31.228693 | 34.465359 | 44.232681 |
| | 2016-01-04 | 29.318140 | 23.297709 | 23.821404 | 25.607523 | 27.908997 | 29.221718 | 31.967928 | 31.506733 | 33.537914 | 41.607315 |
| | 2016-01-05 | 29.184425 | 22.908043 | 23.525404 | 25.514139 | 27.659130 | 29.211117 | 32.105625 | 31.693974 | 33.672344 | 42.103333 |
| | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| | 2016-12-27 | 23.986422 | 19.414381 | 19.962297 | 21.464231 | 22.696068 | 24.037848 | 24.532393 | 26.384960 | 28.901962 | 29.671486 |
| | 2016-12-28 | 23.911253 | 19.089859 | 19.709999 | 21.209757 | 22.518881 | 23.932920 | 24.582596 | 26.602961 | 29.222471 | 29.964825 |
| | 2016-12-29 | 23.697933 | 18.919994 | 19.706202 | 21.200985 | 22.513699 | 23.718616 | 24.395830 | 26.495195 | 29.209251 | 29.903389 |
| | 2016-12-30 | 23.511806 | 18.841398 | 19.554466 | 21.037495 | 22.245239 | 23.379288 | 24.285988 | 26.281944 | 29.304993 | 30.047525 |
| | 2016-12-31 | 23.306200 | 18.458420 | 19.476740 | 20.721758 | 22.019705 | 23.303688 | 24.098824 | 26.118814 | 29.480566 | 30.288113 |

366 rows × 10 columns


70
60
50
40
30
20
10
2015-07
2015-09
2015-11
2016-01
2016-03
2016-05
2016-07
2016-09
2016-11
2017-01
Past time series values
Mean forecast
10%-90% confidence interval

# RMSE

| | Private Score | Public Score |
|---|---|---|
| 8BestW-En-0.4-raw.csv<br>Complete · 5h ago | 8.67892 | 8.77495 |
| BestW-En-0.5.csv<br>Complete · 1d ago | 8.76138 | 8.5838 |
| PM2.5-data-of-thailand_02 - Version 3<br>Complete · 1d ago | 8.76177 | 8.58056 |
| W-En-0.6.csv<br>Complete · 1d ago | 14.88159 | 12.94713 |
| W-En-0.5.csv<br>Complete · 1d ago | 12.65734 | 10.88871 |
| W-En-TimeSeries.csv<br>Complete · 1d ago | 13.3641 | 11.4692 |