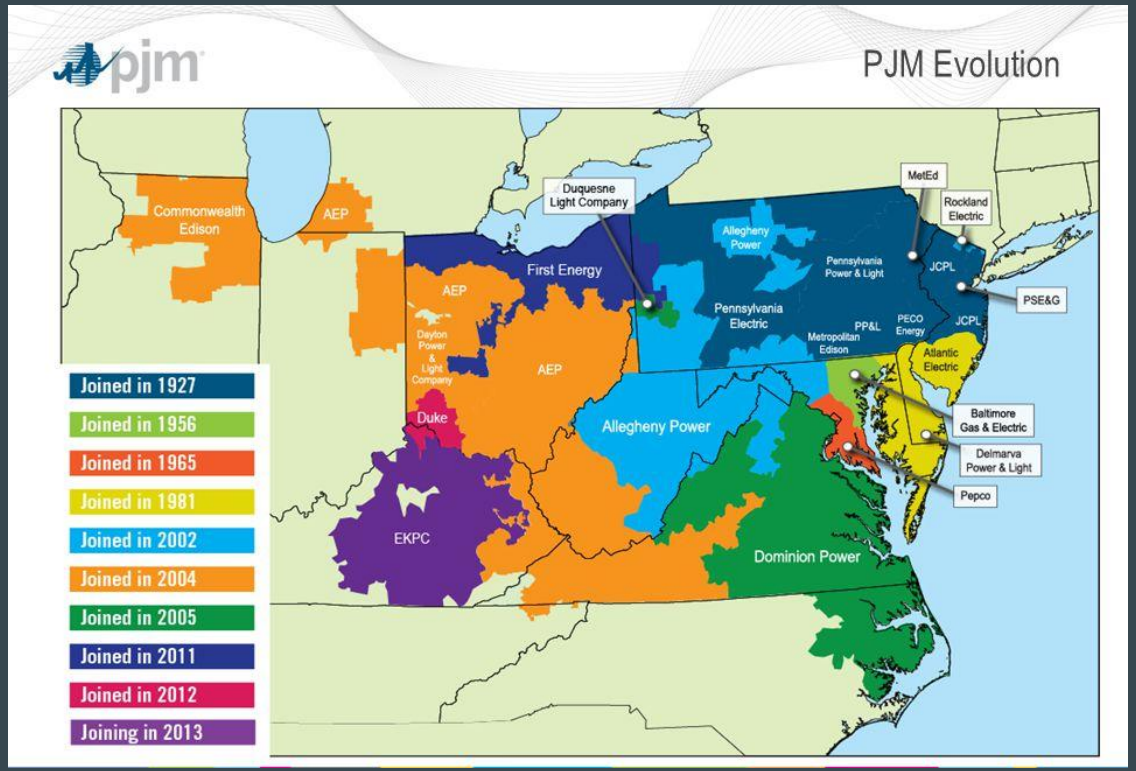
Energy Consumption Forecasting

Κουριδάκης Ανδρέας

Βαγγέλης Ορέστης

Introduction

	Datetime	PJME_MW
0	2002-12-31 01:00:00	26498.0
1	2002-12-31 02:00:00	25147.0
2	2002-12-31 03:00:00	24574.0
3	2002-12-31 04:00:00	24393.0
4	2002-12-31 05:00:00	24860.0



Data Preparation

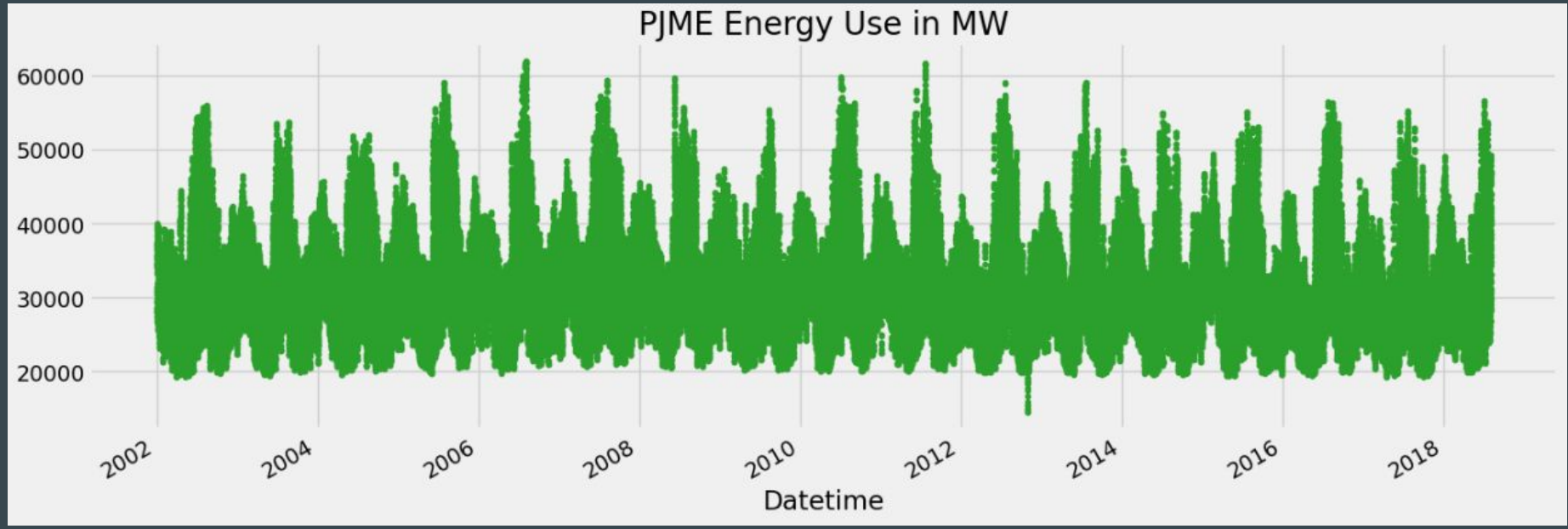
	Datetime	PJME_MW
112487	2014-11-02 02:00:00	23755.0
112488	2014-11-02 02:00:00	22935.0
121223	2015-11-01 02:00:00	21567.0
121224	2015-11-01 02:00:00	21171.0
130127	2016-11-06 02:00:00	20795.0
130128	2016-11-06 02:00:00	21692.0
138863	2017-11-05 02:00:00	21236.0
138864	2017-11-05 02:00:00	20666.0

- No missing values
- Find duplicates rows
- Errors in data collection that do not reflect actual events

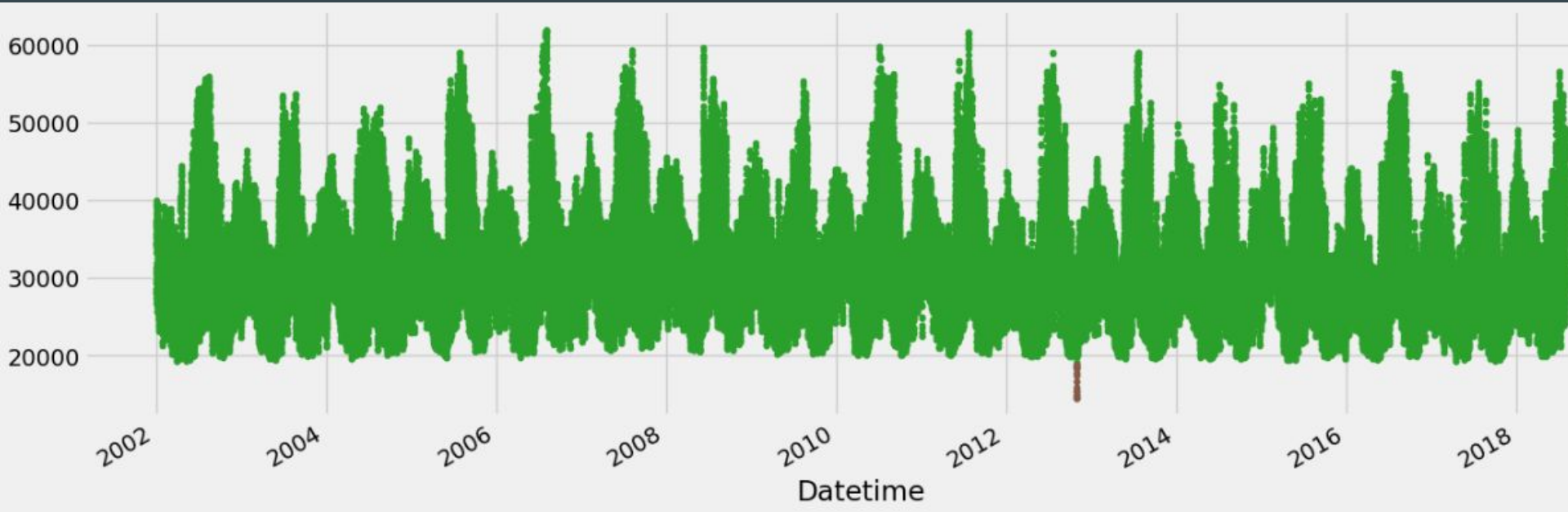
	Datetime	PJME_MW
112482	2014-11-01 21:00:00	29695.0
112483	2014-11-01 22:00:00	28829.0
112484	2014-11-01 23:00:00	27430.0
112485	2014-11-02 00:00:00	25838.0
112486	2014-11-02 01:00:00	23538.0
112487	2014-11-02 02:00:00	23755.0
112488	2014-11-02 02:00:00	22935.0
112489	2014-11-02 03:00:00	22789.0
112490	2014-11-02 04:00:00	22555.0
112491	2014-11-02 05:00:00	23247.0
112492	2014-11-02 06:00:00	23733.0
112493	2014-11-02 07:00:00	24716.0

- Handle duplicates rows (mean)

Plotting the data



Outlier Analysis



On October 29–30, Hurricane Sandy brought high winds and coastal flooding to a large portion of the eastern United States, leaving an estimated 8 million customers without power. "https://en.wikipedia.org/wiki/List_of_major_power_outages#2012

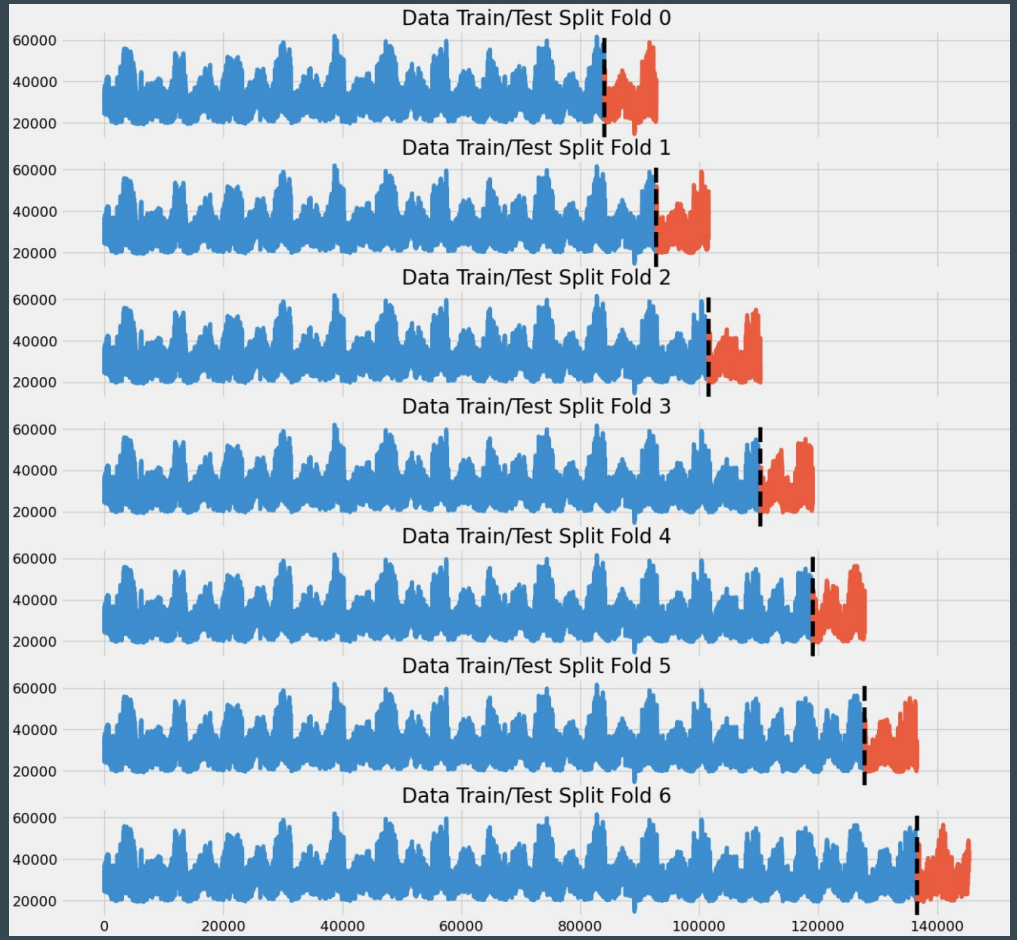
Features Creation

Datetime	PJME_MW	hour	dayofweek	quarter	month	year	dayofyear	dayofmonth	weekofyear
2002-01-01 01:00:00	30393.0	1	1	1	1	2002	1	1	1
2002-01-01 02:00:00	29265.0	2	1	1	1	2002	1	1	1
2002-01-01 03:00:00	28357.0	3	1	1	1	2002	1	1	1
2002-01-01 04:00:00	27899.0	4	1	1	1	2002	1	1	1
2002-01-01 05:00:00	28057.0	5	1	1	1	2002	1	1	1
...	...	...	...	...	...	...	...	...	...
2018-08-02 20:00:00	44057.0	20	3	3	8	2018	214	2	31
2018-08-02 21:00:00	43256.0	21	3	3	8	2018	214	2	31
2018-08-02 22:00:00	41552.0	22	3	3	8	2018	214	2	31
2018-08-02 23:00:00	38500.0	23	3	3	8	2018	214	2	31
2018-08-03 00:00:00	35486.0	0	4	3	8	2018	215	3	31

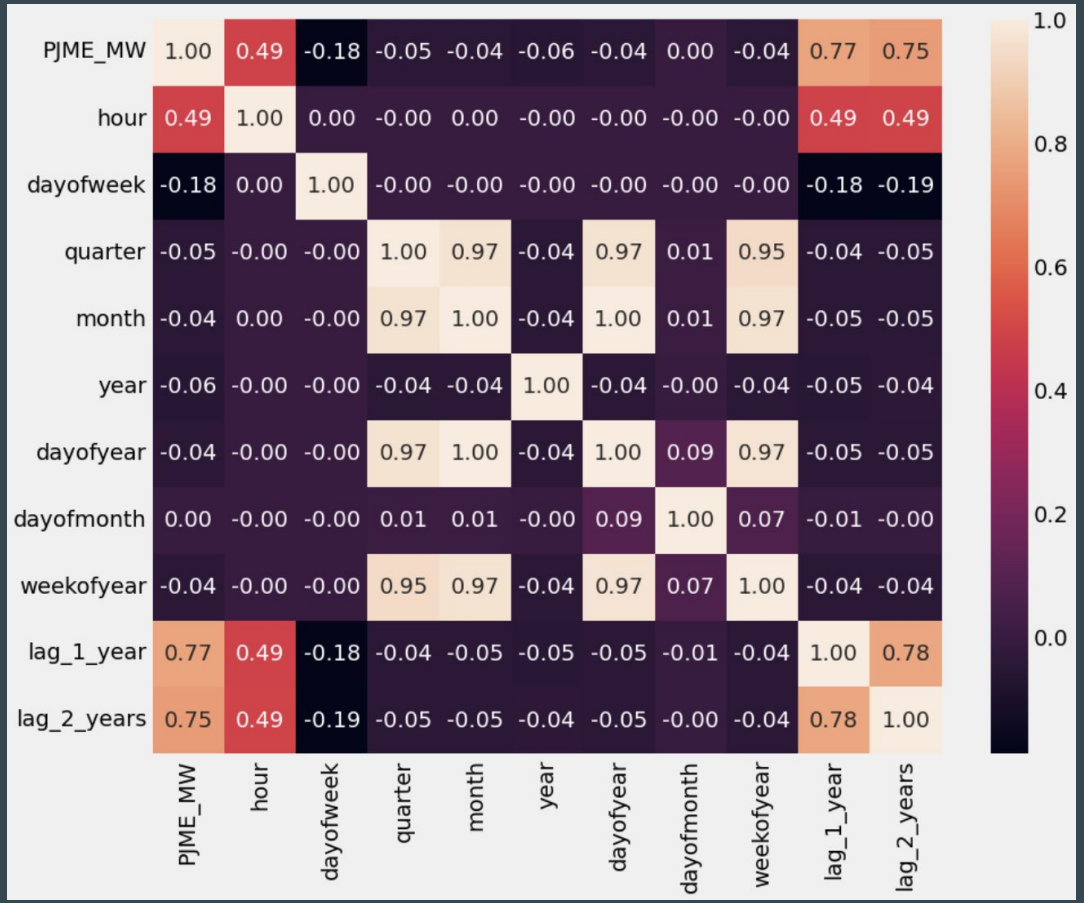
Key findings

- More energy consumption during summer months
- More energy consumption during weeks 26-31
- More energy consumption during the afternoon hours (13:00 - 19:00)
- All years have about the same total consumption

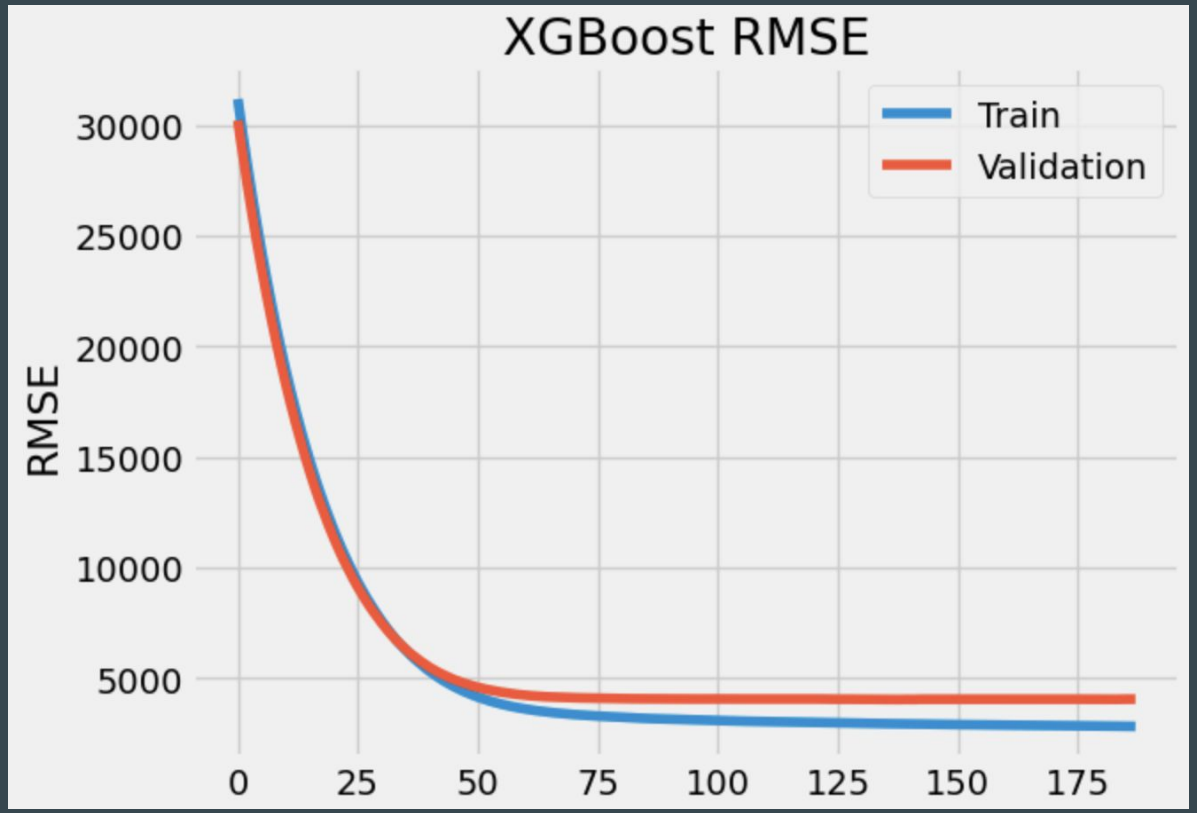
Cross Validation



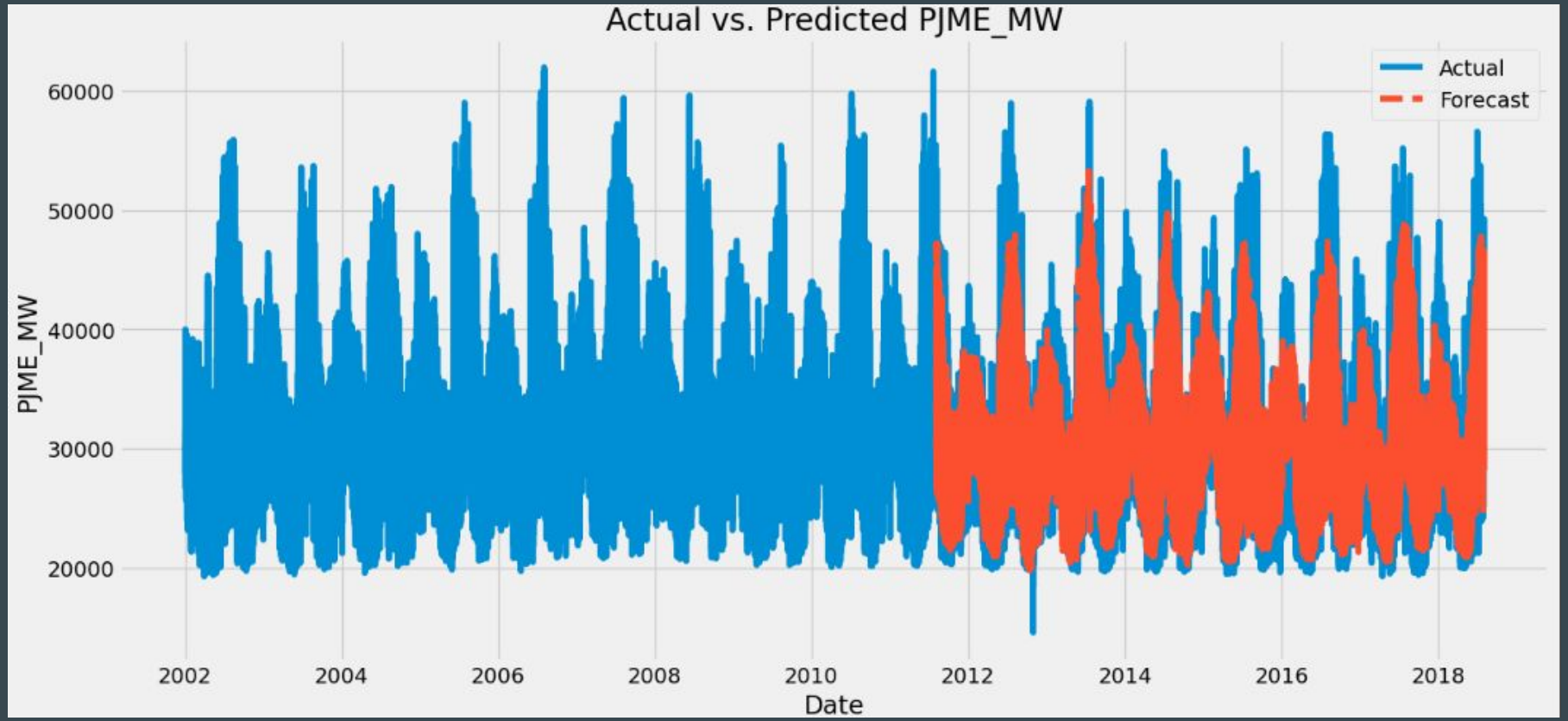
Correlation matrix



XGBoost



- MAPE score across all folds: 7.87
- RMSE score across all folds: 3479


Actual vs. Predicted PJME_MW
Actual
Forecast
PJME_MW
60000
50000
40000
30000
20000
2002
2004
2006
2008
2010
2012
2014
2016
2018
Date

Future work

- Predict the future
- Consider the summer months