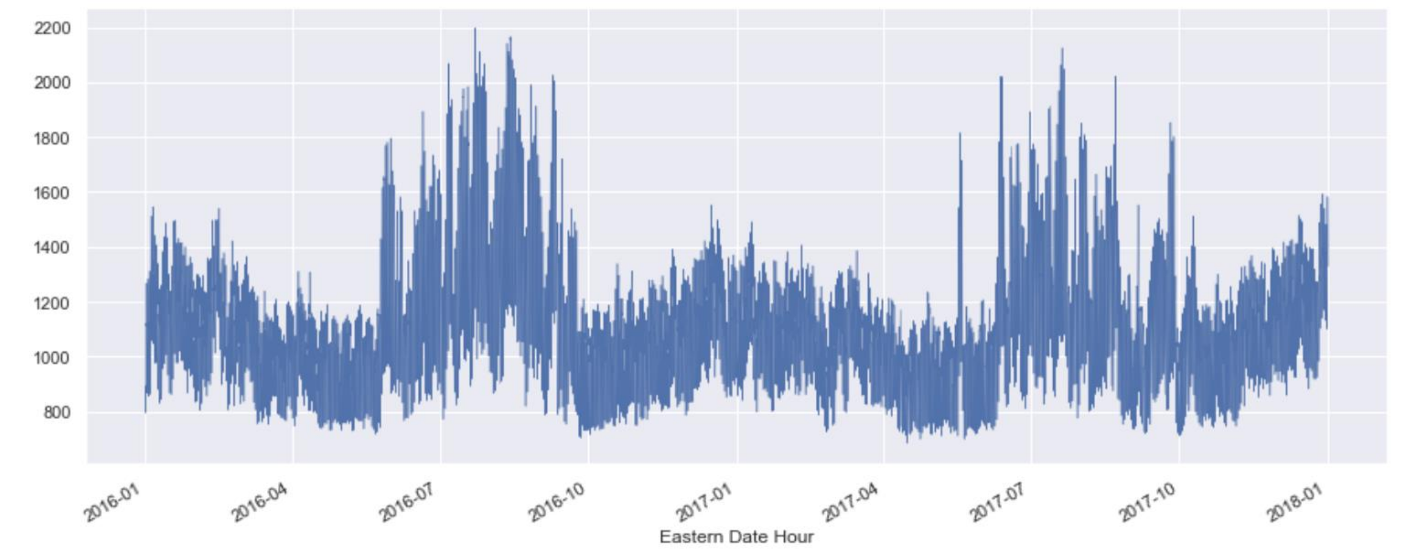
SYSTEM LOAD ANALYSIS

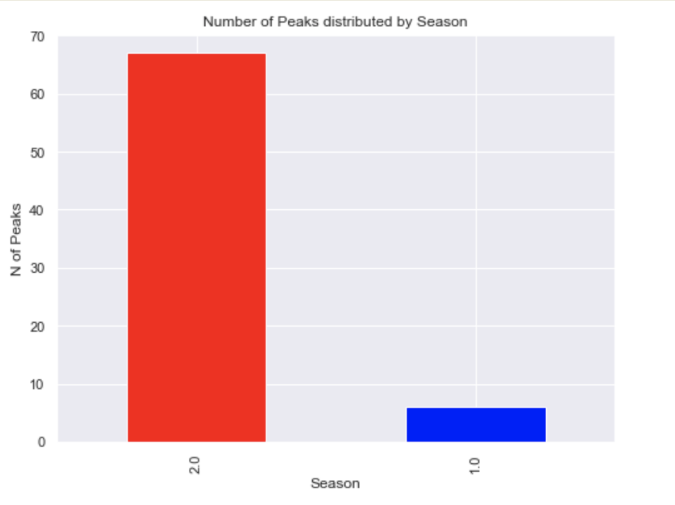
by Santiago M. Novoa
smn405@nyu.edu

Presentation outline:

- Overview of problem- System Load Analysis
- Data:
  - *Data Set Exploration – Visualizations*
  - *Feature Engineering*
  - *Data transformation and adjustments (Normalization)*
- Modelling Approach:
  - *Monthly Averages -> Time Series Patterns*
  - *Hourly residuals ->XGB Regressor*
- Model Evaluation
- Forecasting for efficient infrastructure planning
- Conclusion

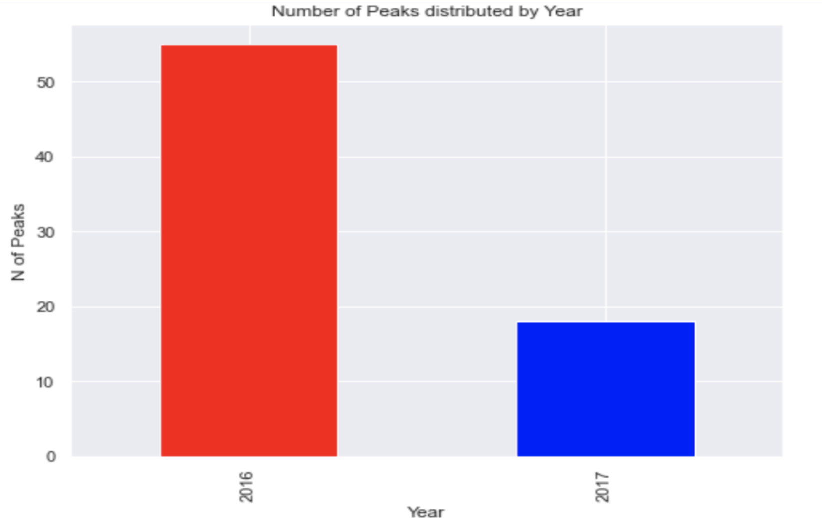
System Load Profile



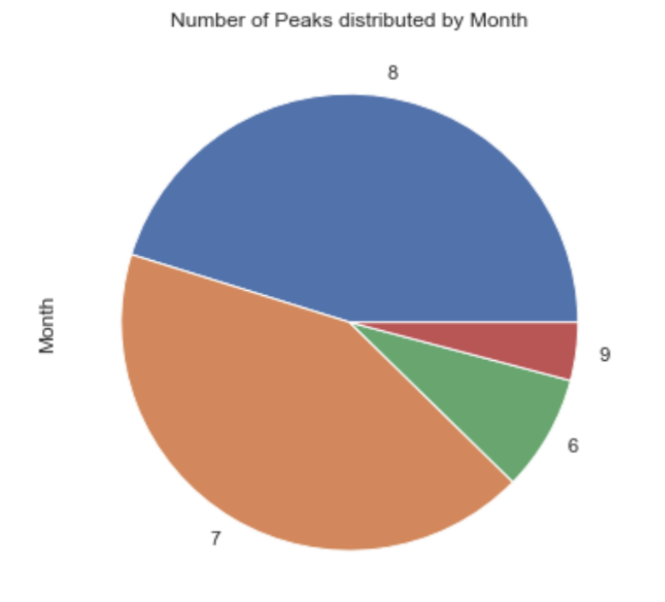


Summer: 67 peaks
Spring: 6 peaks

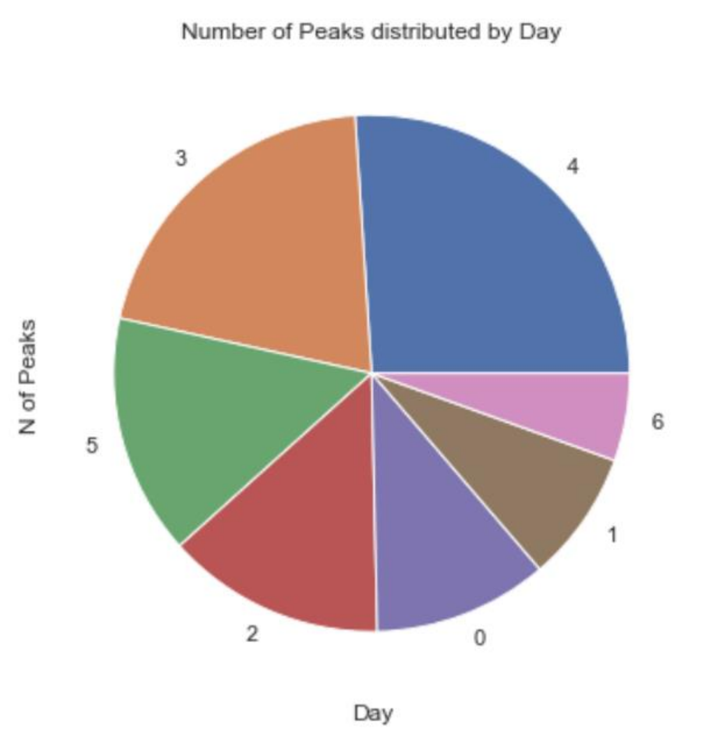
Hourly Data for 2 years (2016, 2017)
Load Profile in MW, with EST timestamp (No data quality issues)
There are 73 hours in which the system load has gone over 2000 MW.
2016: 55 peaks
2017: 18 peaks



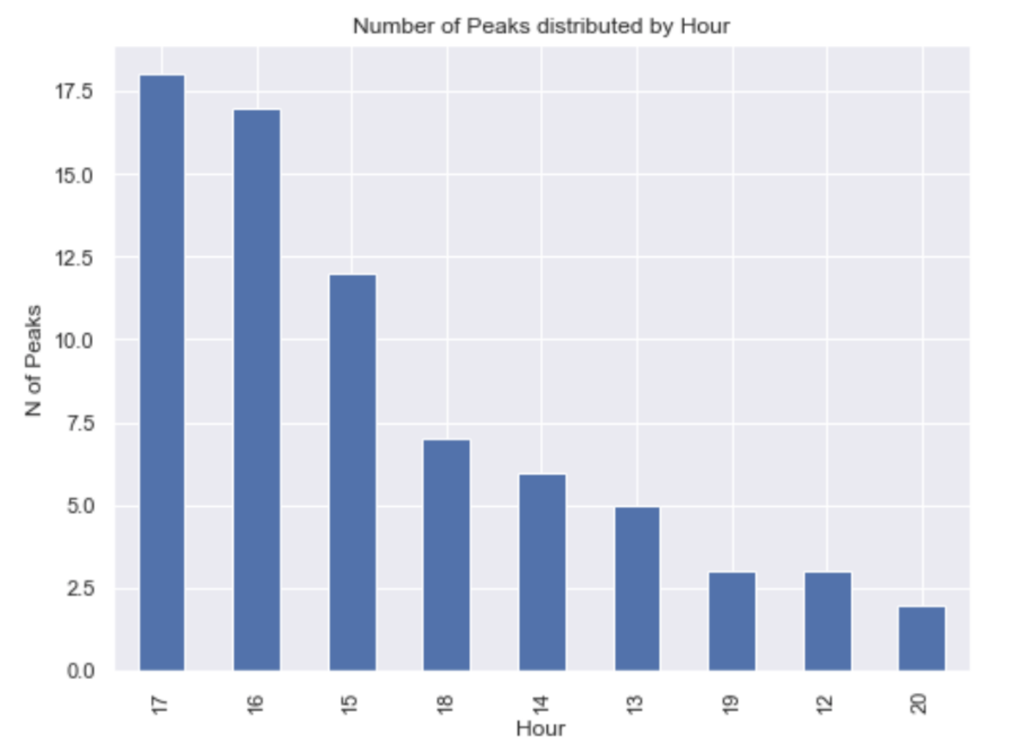
August: 33 peaks
July: 31 peaks
June: 6 peaks
September: 3 peaks



Data Set Exploration – Visualizations



Friday: 19 peaks
Thursday: 15 peaks
Saturday: 11 peaks
Wednesday: 10 peaks
Monday: 8 peaks
Tuesday: 6 peaks
Sunday: 4 peaks



17:00 18 peaks
16:00 17 peaks
15:00 12 peaks
18:00 7 peaks
14:00 6 peaks
13:00 5 peaks
19:00 3 peaks
12:00 3 peaks
20:00 2 peaks

Weather Data

- Fundamental Assumption:

Date time field in weather is expressed in Eastern time too!

Temperature (358 missing values)

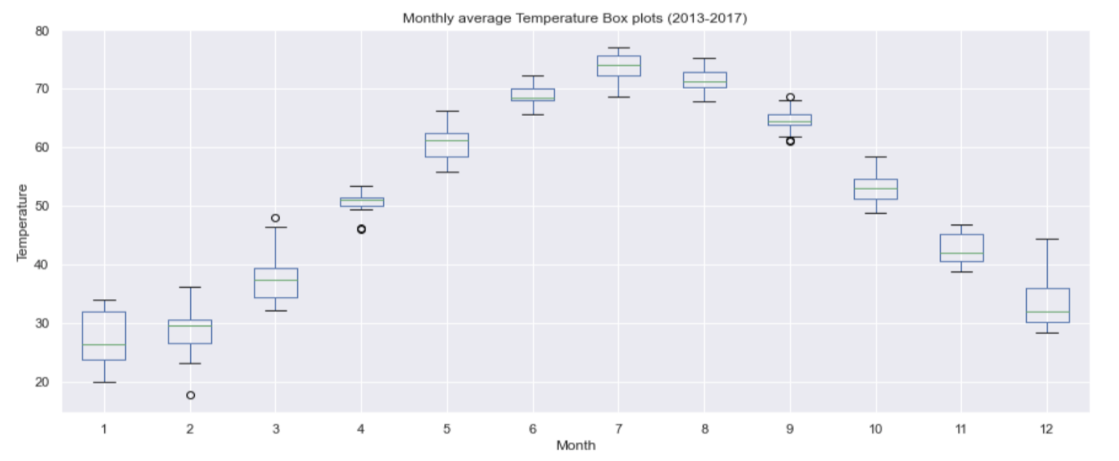
Dew point (938 missing values)

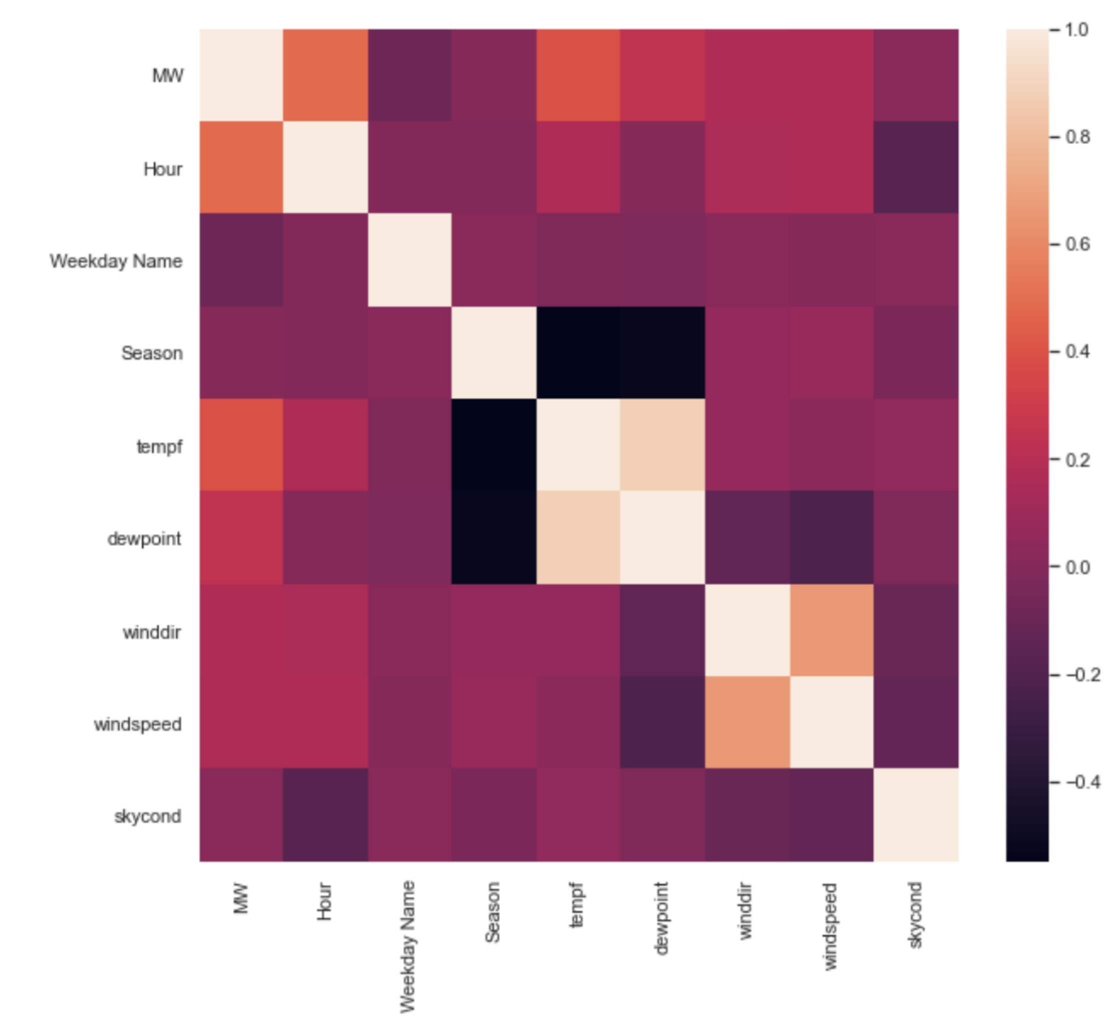
Wind direction (6772 missing values)

Wind speed (1000 missing values)

Sky condition (284 missing values)

Missing Data was filled using Linear extrapolations

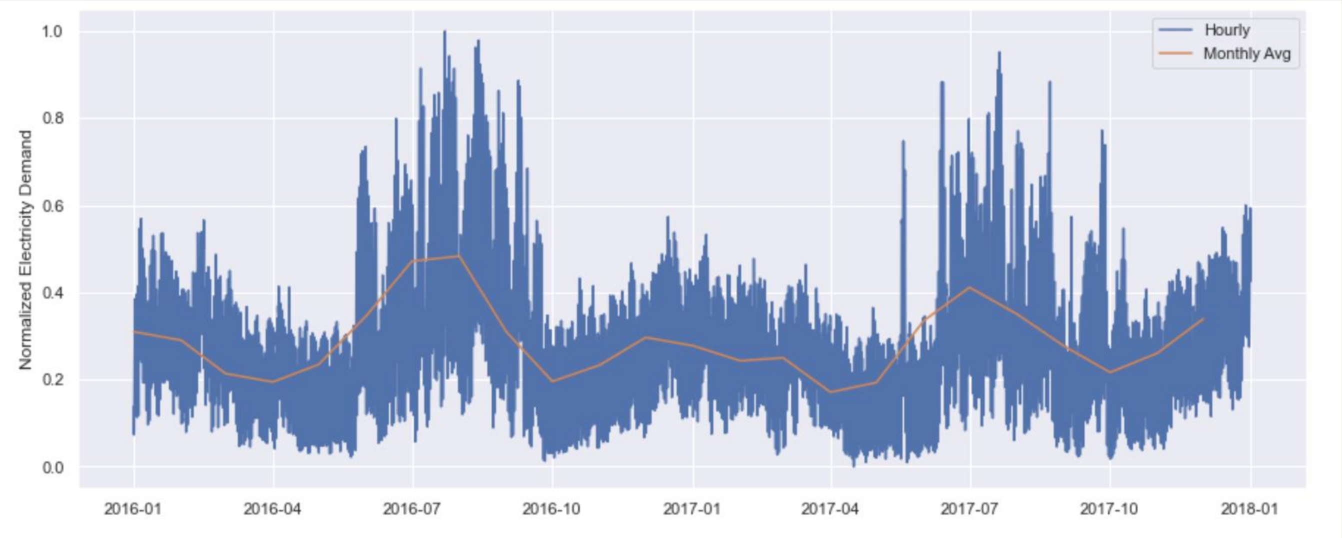




Modelling Approach:

How well can peak demand be explained by weather, day of week, season and hour?

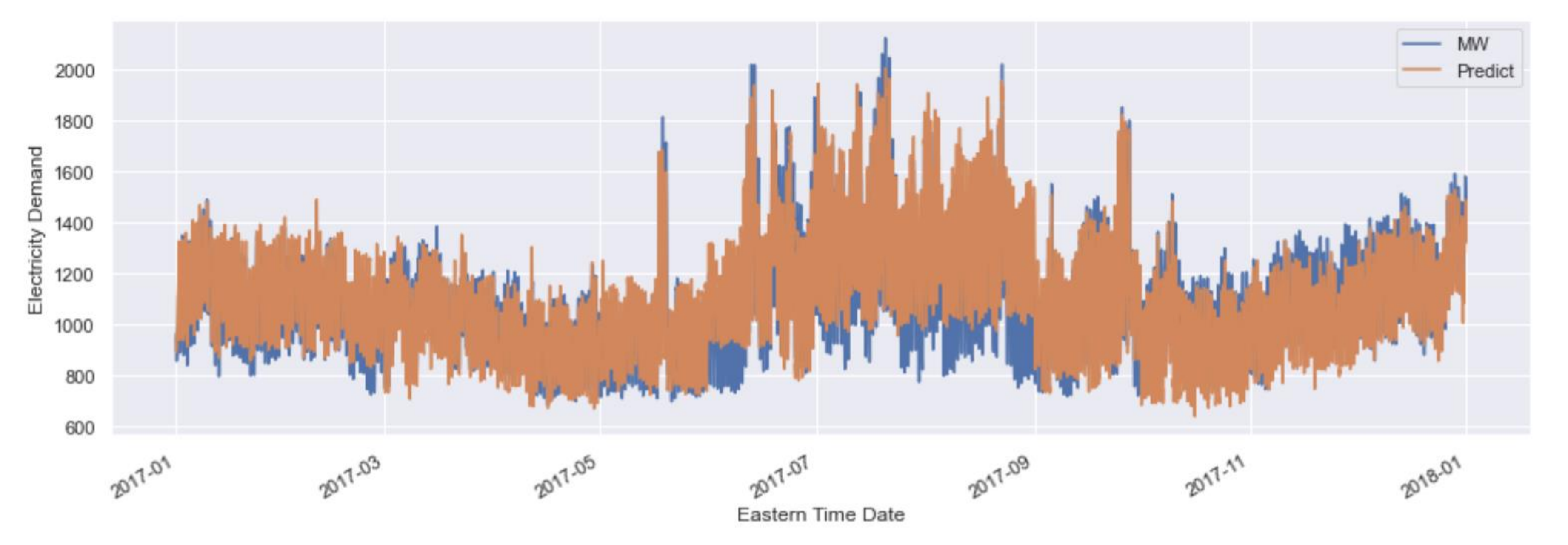
First Component; Monthly average electricity demand, which is modeled using traditional timeseries methods.



Second Component; Hourly differentials based on our features , modeled using extreme gradient boosted regression.

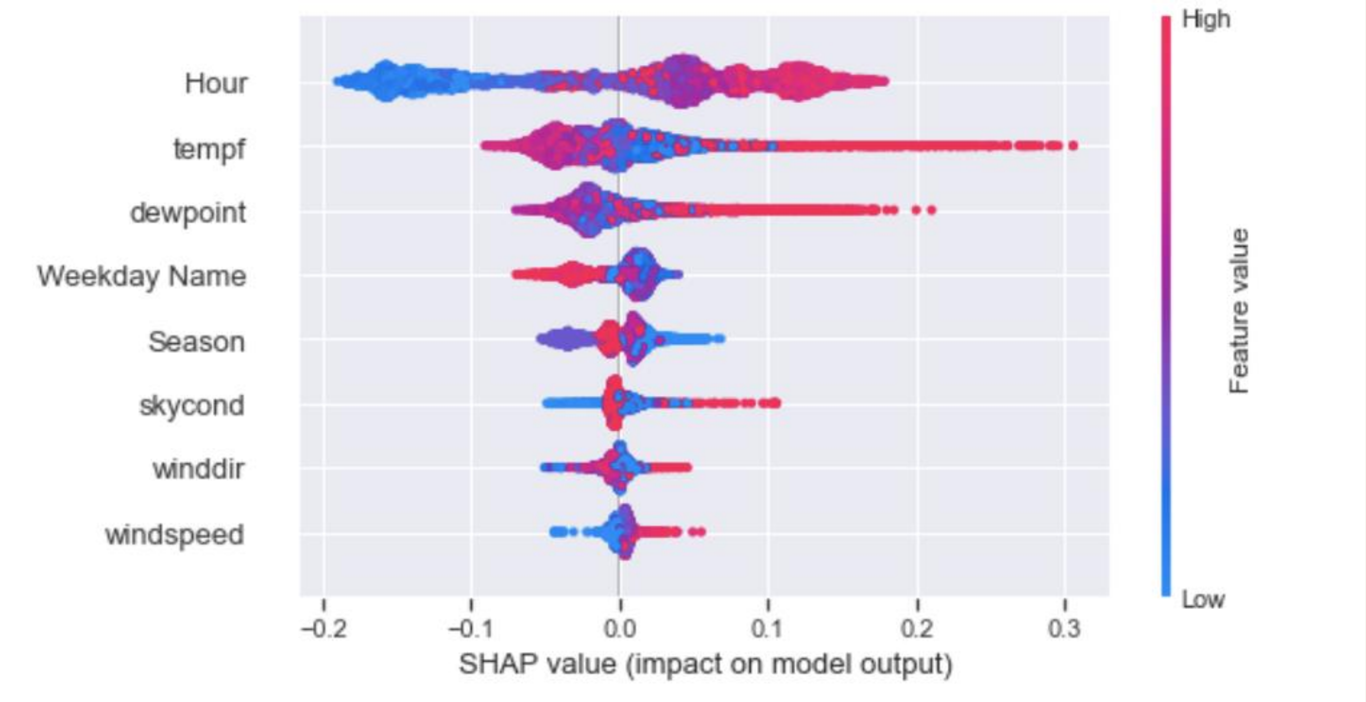
Model Performance

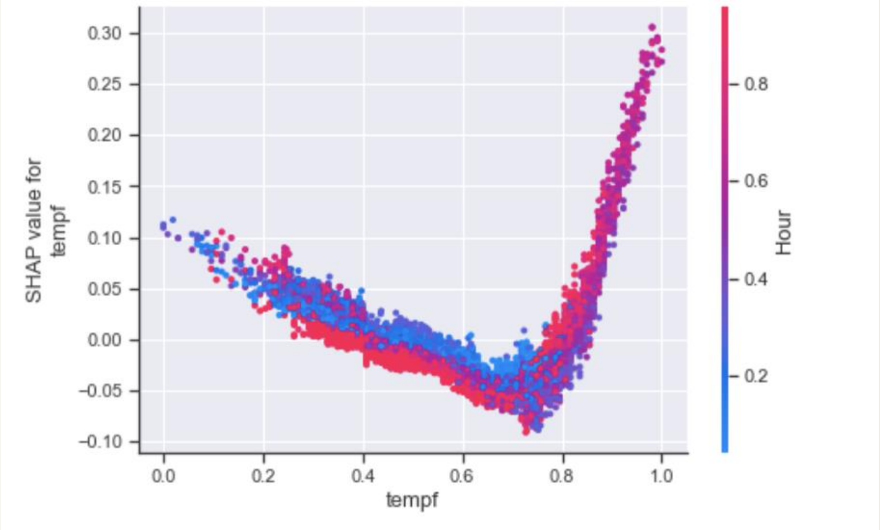
Once both models were trained and fitted with our training data (Year 2016), we proceeded to add their outputs together, in order to get a prediction for the electricity demand at our desired hourly frequency.

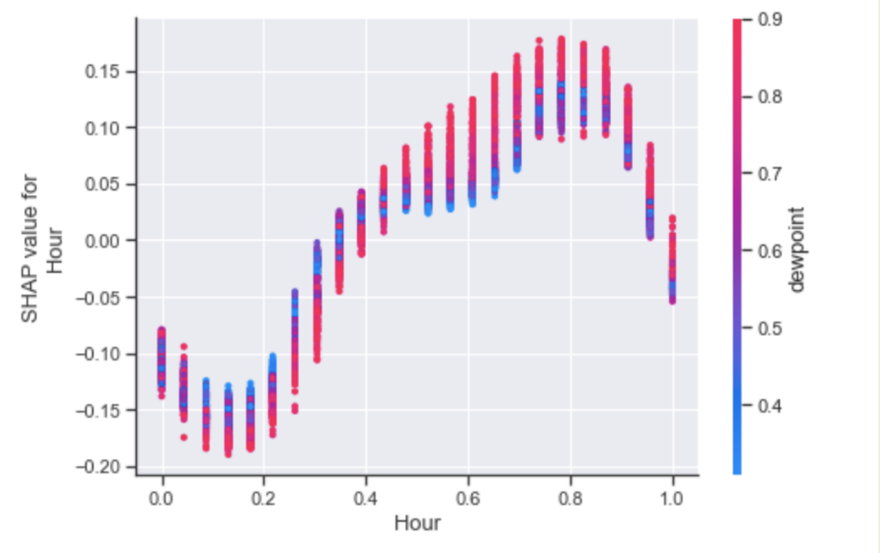


Model’s Forecast RMSE: 96.00337
Model’s Forecast MAPE: 6.85235 %

SHAP summary plot (Model interpretability)







The summary plot combines feature importance with feature effects.
Each point on the summary plot is a Shapley value for a feature and an instance.

Estimating future peak demands - Planning

Eastern Date Hour	MW	Year	Month	Day	Hour	Weekday Name	Season	Predict
2017-07-20 15:00:00	2109.7	2017	7	20	15	3	2.0	2007.877382
2017-07-20 16:00:00	2125.1	2017	7	20	16	3	2.0	1977.701430
2017-07-21 17:00:00	2048.2	2017	7	21	17	4	2.0	1965.844816
2017-07-21 16:00:00	2044.7	2017	7	21	16	4	2.0	1965.100082
2017-08-22 14:00:00	1965.2	2017	8	22	14	1	2.0	1958.212826
2017-07-01 16:00:00	1691.4	2017	7	1	16	5	2.0	1946.465223
2017-08-22 15:00:00	1997.8	2017	8	22	15	1	2.0	1943.846927
2017-07-12 16:00:00	1904.1	2017	7	12	16	2	2.0	1943.813364
2017-06-13 14:00:00	1988.7	2017	6	13	14	1	1.0	1941.047299
2017-08-22 18:00:00	1957.1	2017	8	22	18	1	2.0	1935.589279
2017-07-20 18:00:00	1974.0	2017	7	20	18	3	2.0	1925.281633
2017-06-13 13:00:00	1937.2	2017	6	13	13	1	1.0	1925.241624
2017-06-18 16:00:00	1697.0	2017	6	18	16	6	1.0	1919.576260
2017-08-22 16:00:00	2022.1	2017	8	22	16	1	2.0	1917.607263
2017-07-20 14:00:00	2061.7	2017	7	20	14	3	2.0	1913.734177
2017-08-01 18:00:00	1824.6	2017	8	1	18	1	2.0	1910.269599
2017-07-01 17:00:00	1754.5	2017	7	1	17	5	2.0	1908.427617
2017-07-18 16:00:00	1960.7	2017	7	18	16	1	2.0	1907.453248
2017-08-01 16:00:00	1851.0	2017	8	1	16	1	2.0	1903.477318
2017-07-18 15:00:00	1925.3	2017	7	18	15	1	2.0	1902.420327

For this Analysis we decided to run our forecasting model in a hypothetical scenario where the Temperature of 2017 goes up by 10 degrees.

Being that our model is highly susceptible to changes in temperature forecasts, we observe significant changes in predicted system Load.

While our original 2017 forecast peaked at 2007 MW, by changing the temperature 10 degrees higher our new peak would be 2242 MW.

Also worth noting that the original 2017 forecast only passed the 2000 MW Threshold once, while the artificial ‘higher temperature forecast crosses the 2000 MW limit in all of the top 20 peaks

Eastern Date Hour	MW	Year	Month	Day	Hour	Weekday Name	Season	Predict
2017-07-01 16:00:00	1691.4	2017	7	1	16	5	2.0	2242.517765
2017-08-18 18:00:00	1609.2	2017	8	18	18	4	2.0	2235.855901
2017-07-20 18:00:00	1974.0	2017	7	20	18	3	2.0	2235.244099
2017-08-18 15:00:00	1638.9	2017	8	18	15	4	2.0	2228.481589
2017-08-03 15:00:00	1771.4	2017	8	3	15	3	2.0	2207.231158
2017-08-01 18:00:00	1824.6	2017	8	1	18	1	2.0	2202.987517
2017-07-12 16:00:00	1904.1	2017	7	12	16	2	2.0	2198.282365
2017-07-18 20:00:00	1799.2	2017	7	18	20	1	2.0	2189.108434
2017-08-04 14:00:00	1751.8	2017	8	4	14	4	2.0	2186.578012
2017-07-13 15:00:00	1807.0	2017	7	13	15	3	2.0	2186.172195
2017-07-01 17:00:00	1754.5	2017	7	1	17	5	2.0	2184.280054
2017-08-18 17:00:00	1648.2	2017	8	18	17	4	2.0	2183.975935
2017-08-21 19:00:00	1683.7	2017	8	21	19	0	2.0	2182.136419
2017-08-03 18:00:00	1762.6	2017	8	3	18	3	2.0	2181.413213
2017-08-03 19:00:00	1688.7	2017	8	3	19	3	2.0	2181.237828
2017-07-01 15:00:00	1659.9	2017	7	1	15	5	2.0	2177.905250
2017-08-18 19:00:00	1548.5	2017	8	18	19	4	2.0	2175.192943
2017-08-03 16:00:00	1800.3	2017	8	3	16	3	2.0	2173.714759
2017-08-22 15:00:00	1997.8	2017	8	22	15	1	2.0	2172.358433
2017-08-01 17:00:00	1847.8	2017	8	1	17	1	2.0	2171.832728

Conclusions

- We can observe clear seasoned patterns that emerge from visualizing historical peak electricity demand. While some of our predictors are clearly correlated (Summer -> High Temperatures), it is obvious that Peaks are largely a seasonal event.
- Our additive forecasting method proved to be reasonably capable of predicting demand by weather, day of week, season and hour at a system level.
- Our Feature importance analysis allowed us to quantify the relationship’s we were able to infer from our preliminary data visualizations (Temperature, Season and hour)
- Forecasting is an important tool for many stakeholders in the Power industry; in this case we were able to roughly estimate what could serve as boundaries for infrastructure maintenance and planning needs.