Qualetics
DATA MACHINES

Use Case

Data Intelligence in Energy Sector

Time Series Analysis for Analyzing and Predicting Energy Consumption

Introduction

Consumption is an important parameter to determine the energy management and control system of any country. And therefore, an analysis and prediction of the electricity consumption of a particular location becomes crucial and helps the traders and stakeholders to efficiently manage their demand and supply portfolios.

In our study, we tried to analyze and predict the consumption of electricity for the regions served by the Eastern Interconnection power grid of the United States.

Electricity in the United States is supplied to consumers through three large power grids.

1. The Eastern Interconnection - serving the area east of the Rocky Mountains.
2. The Western Interconnection - serving the area west of the Rocky Mountains and Great Plains to the Pacific Coast.
3. The Electric Reliability Council of Texas (ERCOT) - covers most of the state of Texas.

Each of these grids is run by different utility companies that manage the electric supply for the regions.

Problem Statement

The demand for electricity has been continuously increasing. To take the necessary actions for optimizing the supply of energy consumption, finding a model that can visualize, understand, and forecast the energy consumption trends for a period of time is very critical.

In our study, we analyzed the energy consumption data of 20 years and predict the energy consumption levels in the Eastern Interconnection grid for the next year. We also intend to understand the trends of the consumption of energy in two particular seasons - Summer and Winter.

About the Dataset

We considered a total of 12 datasets that have data registered from 1998 to 2018. 11 of these datasets have been registered by the utility companies of the region for the same period of time. The datasets contain the collection of hourly loads of particular locations.

Solution provided

An analysis and prediction of energy consumption tend to provide incisive insights that are fundamental in concerns like micro-grids, energy storage, demand analysis, and energy feedback. Therefore, a time-series analysis was performed on the occurred data.

Time series analysis is performed to forecast future events based on the already occurred events for which the data is collected at regular time intervals. The time series models predict future events by analyzing past trends on the assumption that future trends will hold similar to past trends.

As there are a lot of datasets to be considered, we have to clean the data properly and choose the best time series model to predict the required values.

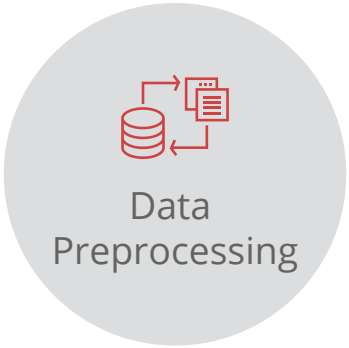
Data cleaning

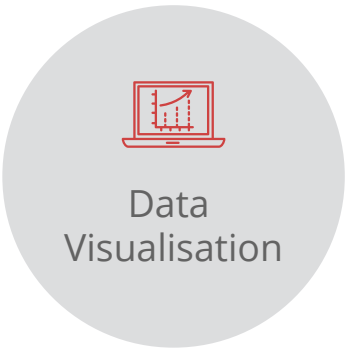
- There is one main dataset that contains columns that have all the columns but has a lot of NAN values.
- There are 11 variables in the main dataset that are defined by the remaining 11 datasets.

- So we have decided to merge the datasets two at a time and have arrived at a final dataset with 11 parameters from the year 2013 to 2018.
- We observed that there are duplicates in the dataset so we have removed all the duplicates.
- Then we have split the date-time column and have retained the date column.
- We have created a new column and calculated the sum of the loads by all the companies.
- We first calculated the daily average of the total loads and then calculated the monthly average loads for all the years.
- We then sorted the values according to the date and started the time series analysis.
- After the detailed cleaning of the data, we have the data from 2013-2018 with all the 11 companies calculating individual loads.

Technologies Used

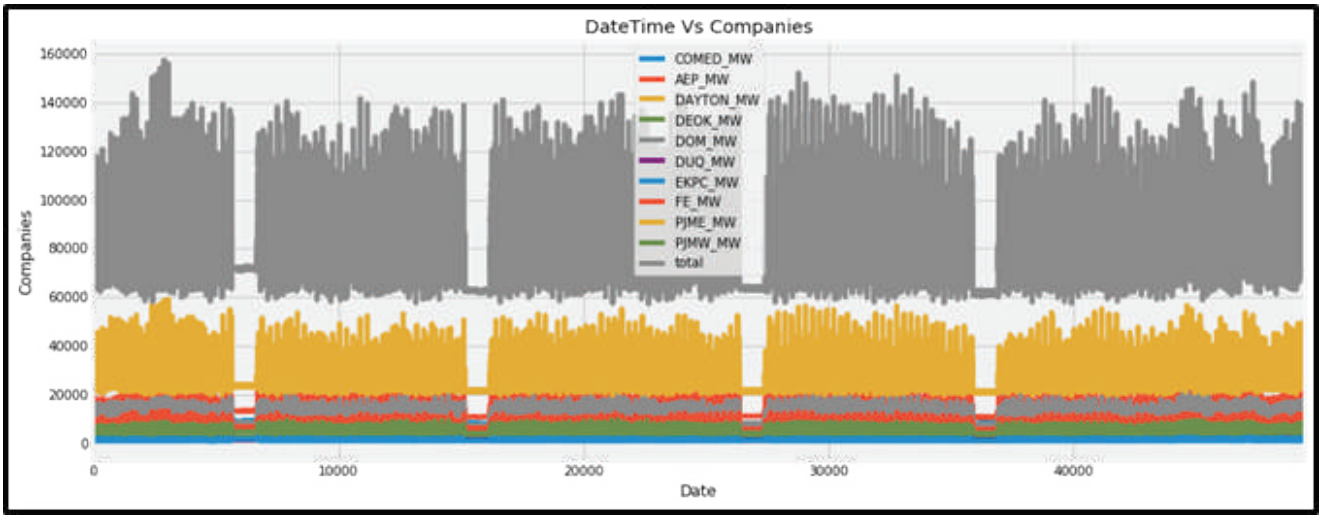
The following technologies were used to perform our analysis.





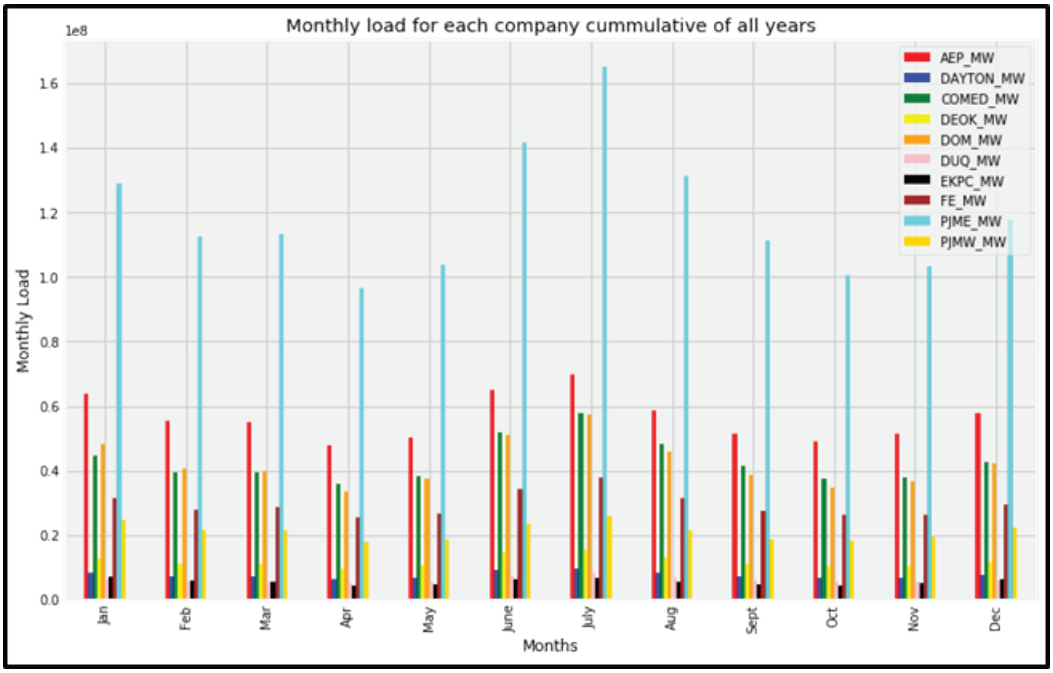
Model Metrics

The Time series model that we used for our study is the ARIMA model. After the data cleansing, we made a line plot on the whole data.



Monthly total loads

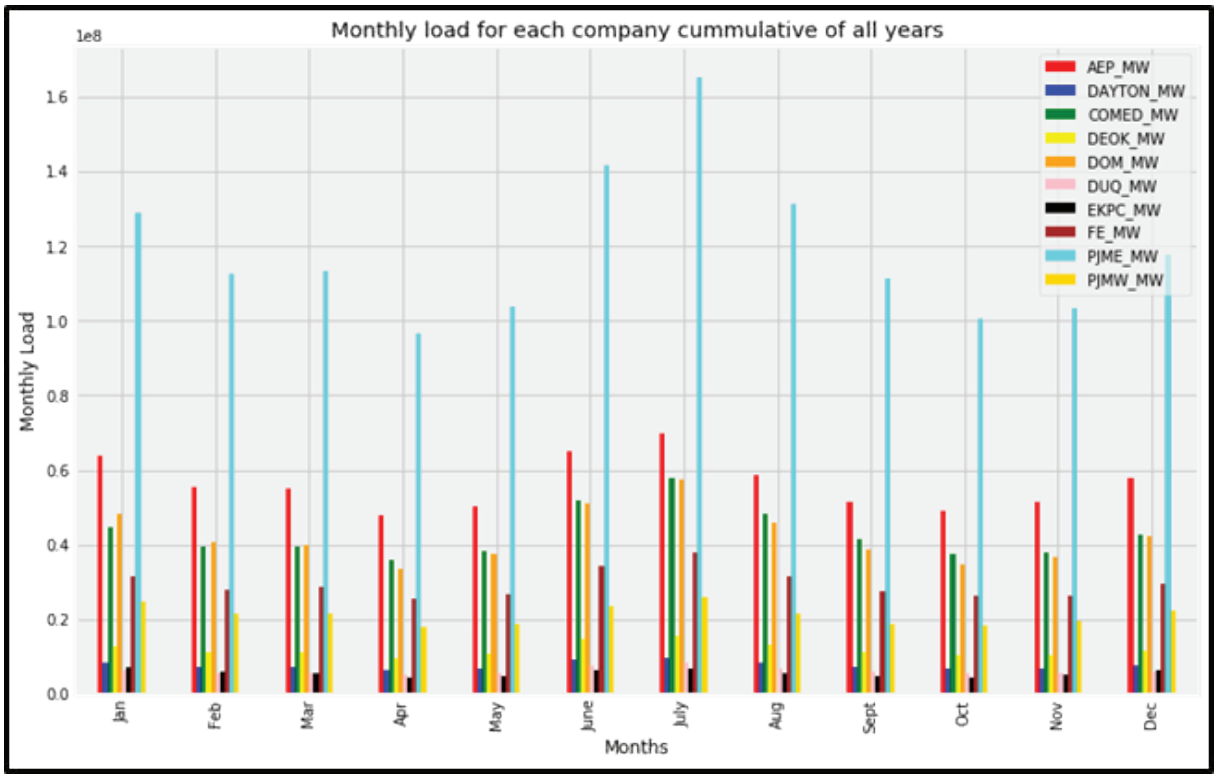
We calculated the individual total loads of all the companies from 2013 to 2018 (Month-wise).



Inference: From the above plot, we can infer that the PJME-MW is having the highest load calculated for the entire duration of our study.

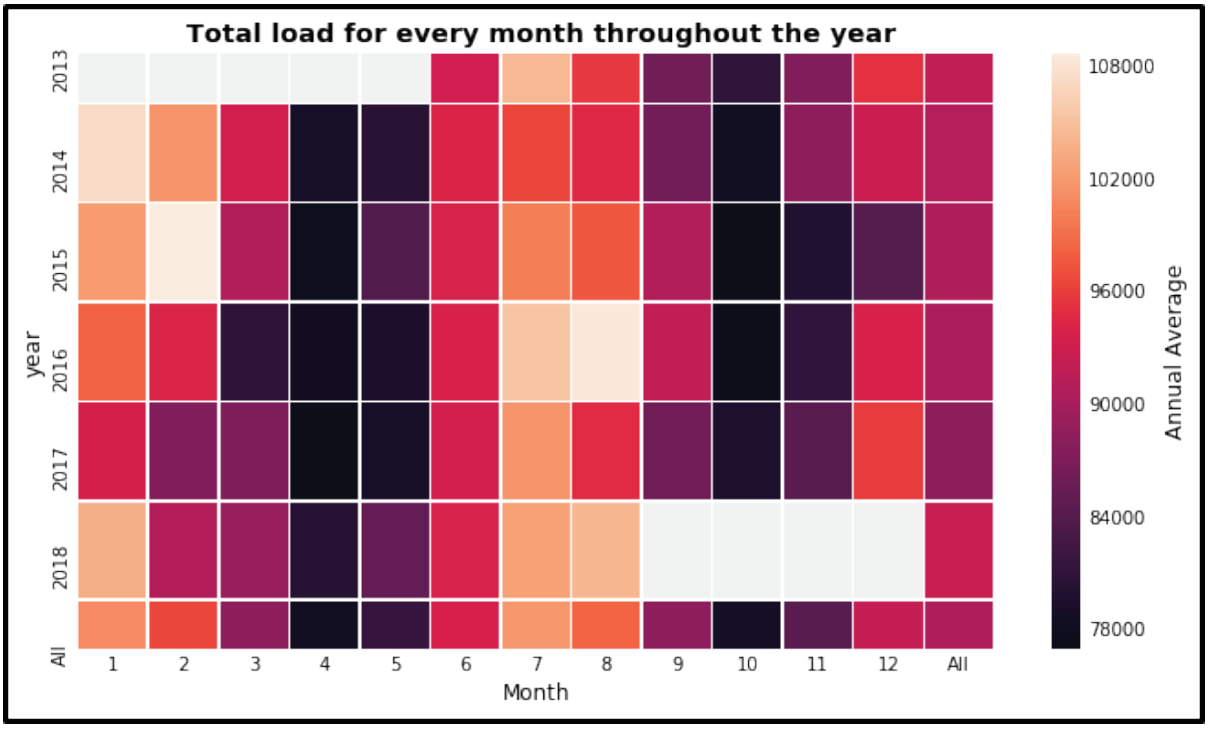
Monthly total loads

We calculated the individual total loads of all the companies from 2013 to 2018 (Month-wise).



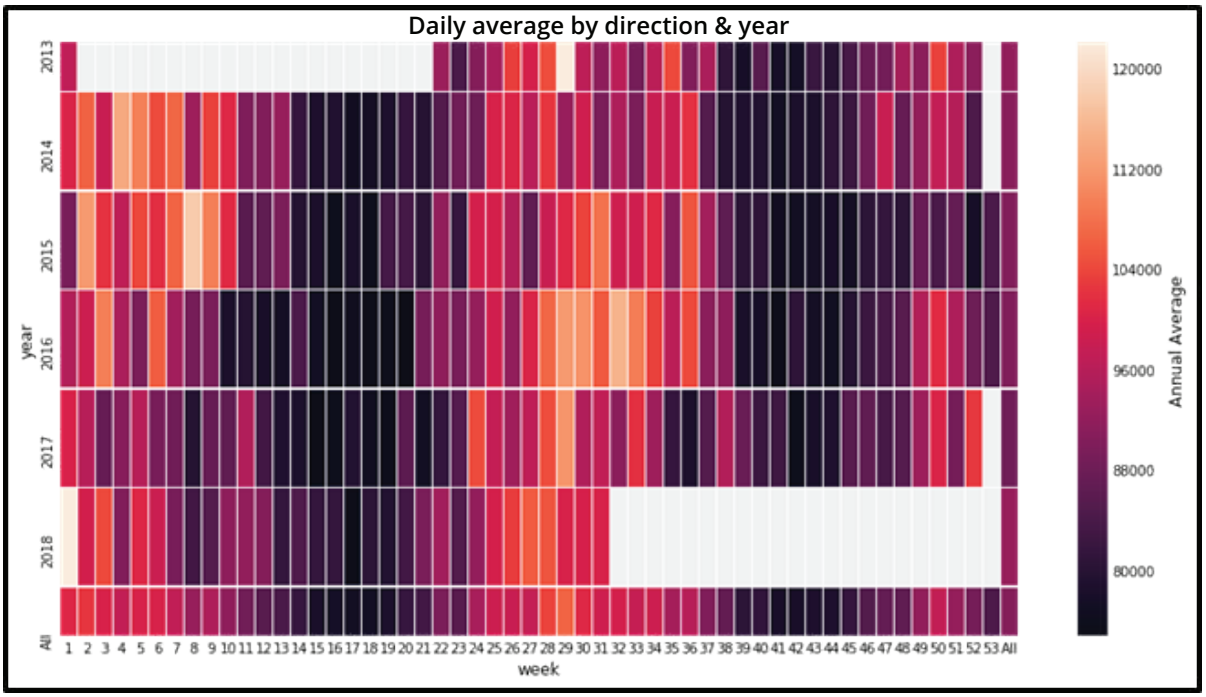
Inference: From the above plot, we can infer that the PJME-MW is having the highest load calculated for the entire duration of our study.

Heat map showing the monthly average load of all the years

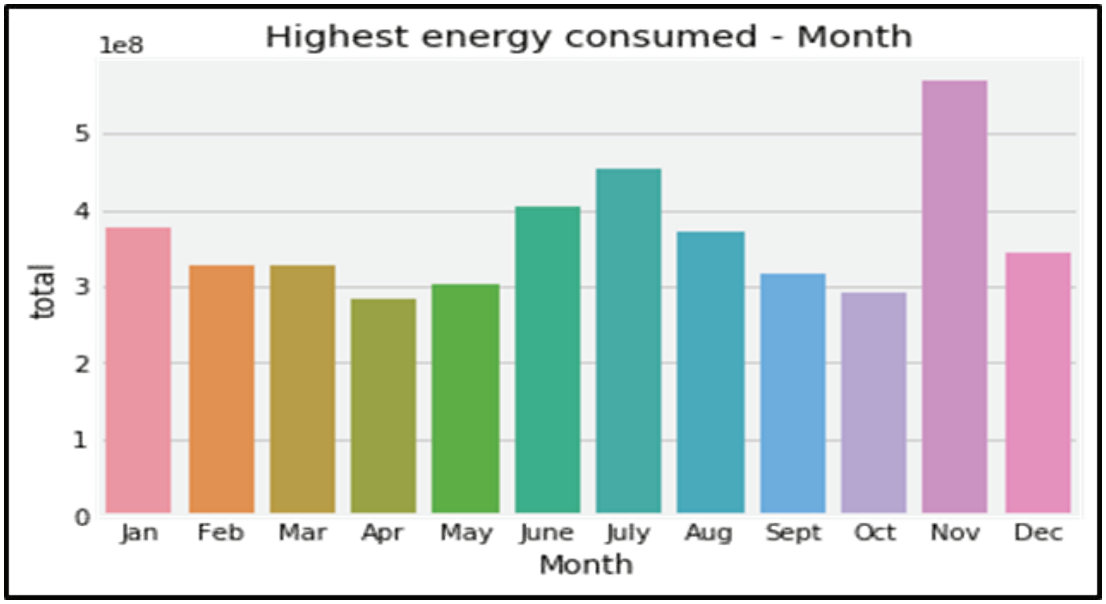


Inference: From the above heatmap, we can infer that the months January to May of 2013 and the months September to December of 2018 have seen the highest average monthly loads.

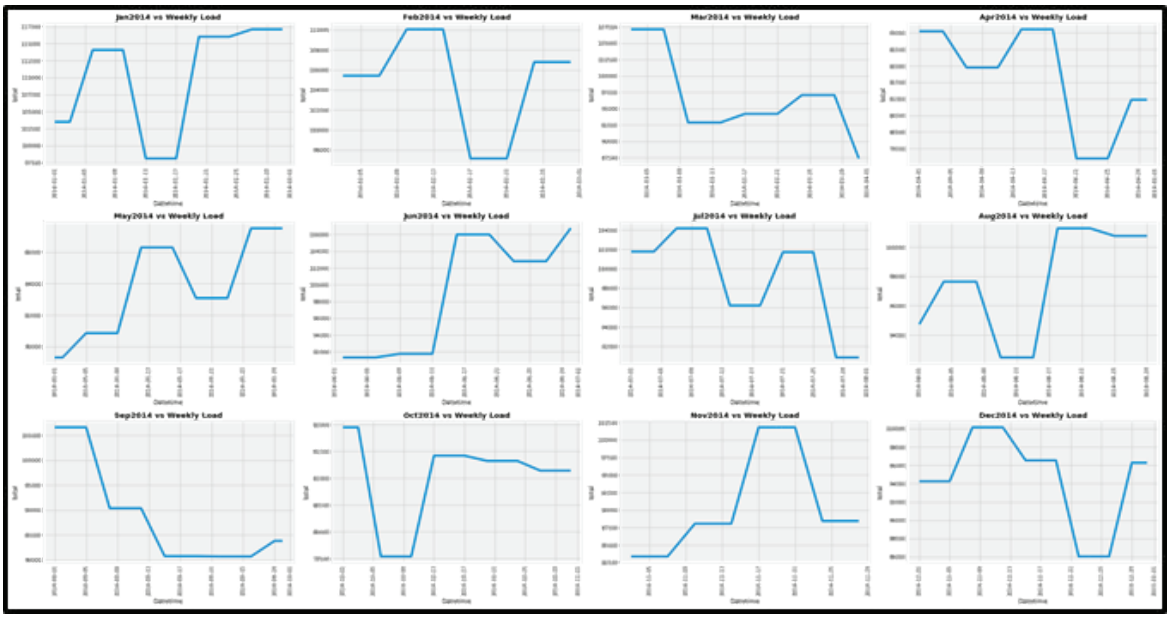
Heatmap showing weekly load consumed



We calculated the total loads to analyze the month in which the highest load was consumed.

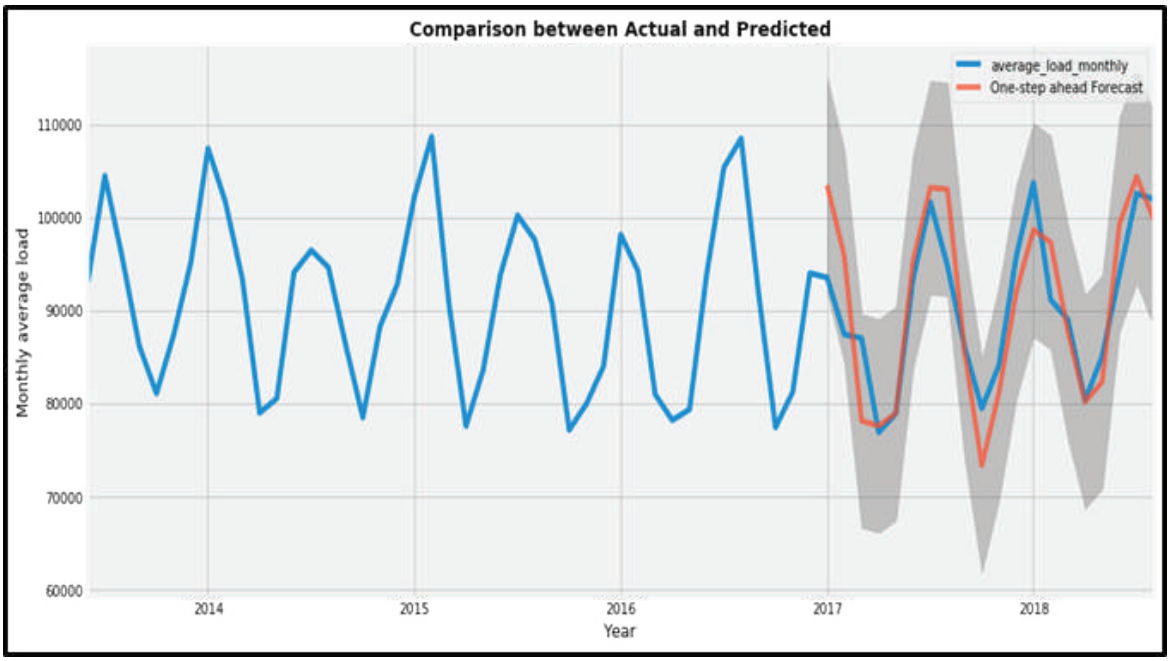


Inference: We can infer that November month is having the highest load and so the power grid has to generate more electricity to meet the demand.

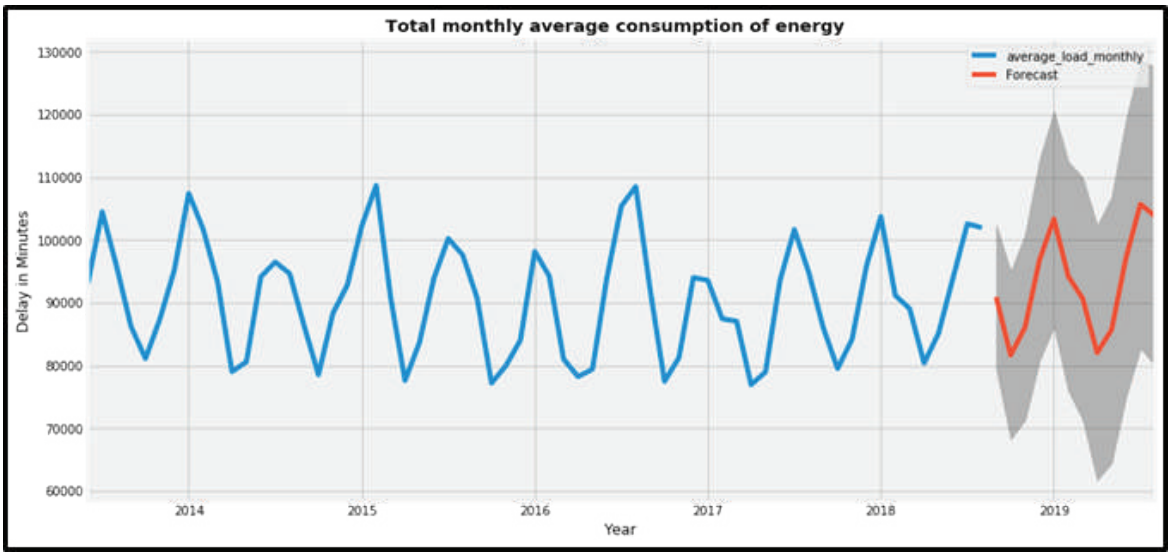


Monthly average load prediction

We predicted the average monthly load for the next year. The blue line in the below plot is the actual average monthly load and the predicted values are shown in the red line.



Predicted values

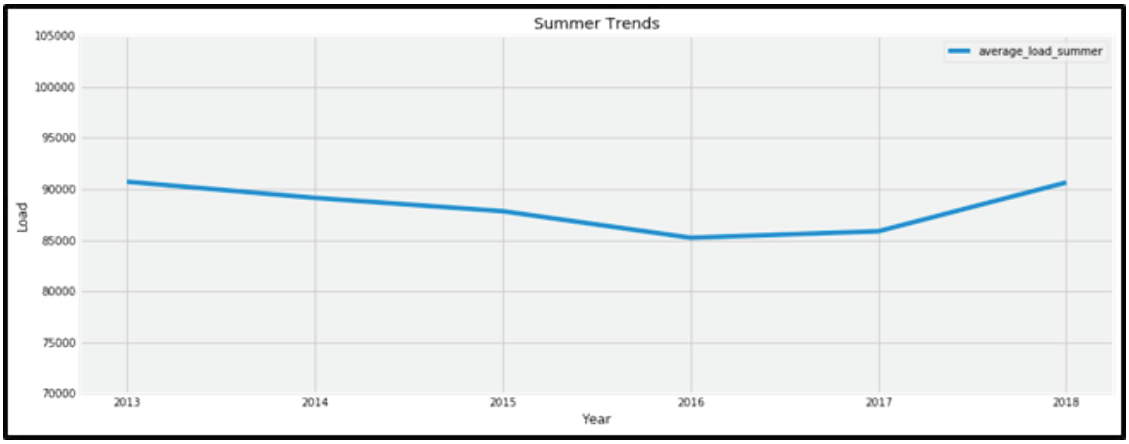


The predicted monthly average of total load values are given below:

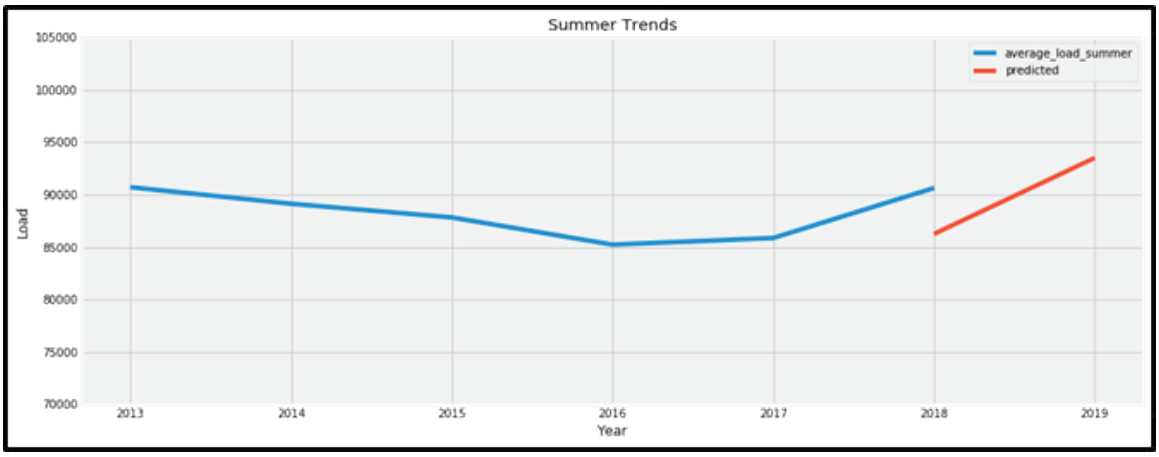
	Date	Total Load
0	01-01-2019	93045.20053
1	01-02-2019	89052.74152
2	01-03-2019	90055.5566
3	01-04-2019	82540.00763
4	01-05-2019	89348.48167
5	01-06-2019	97399.96464
6	01-07-2019	98487.40238
7	01-08-2019	93535.70465
8	01-09-2019	93801.94334
9	01-10-2019	89680.98736
10	01-11-2019	88388.12899
11	01-12-2019	90944.06113

Season-wise analysis

Summer: We divided the data based on months. We considered only March, April, May, and June for the summer-season analysis, studied the trends and predicted the values for the next summer.



Predicted



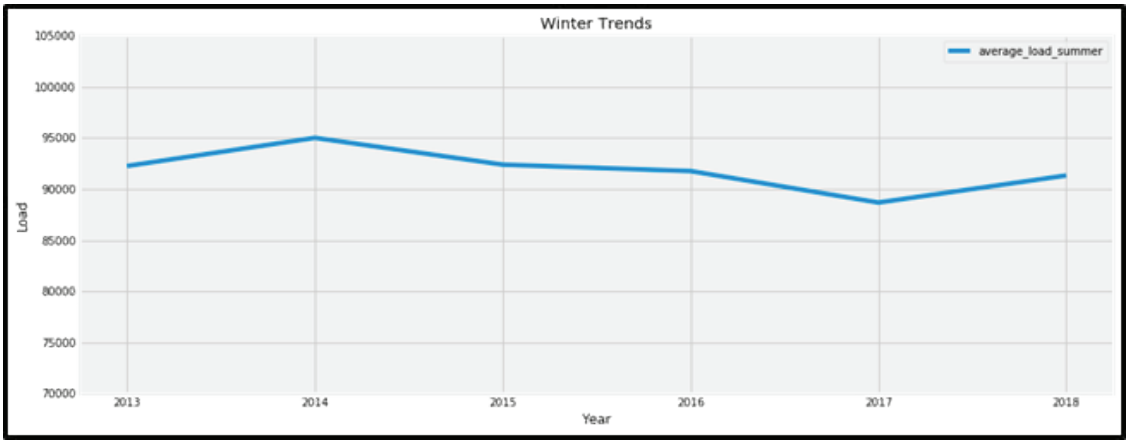
ipsum

Inference:

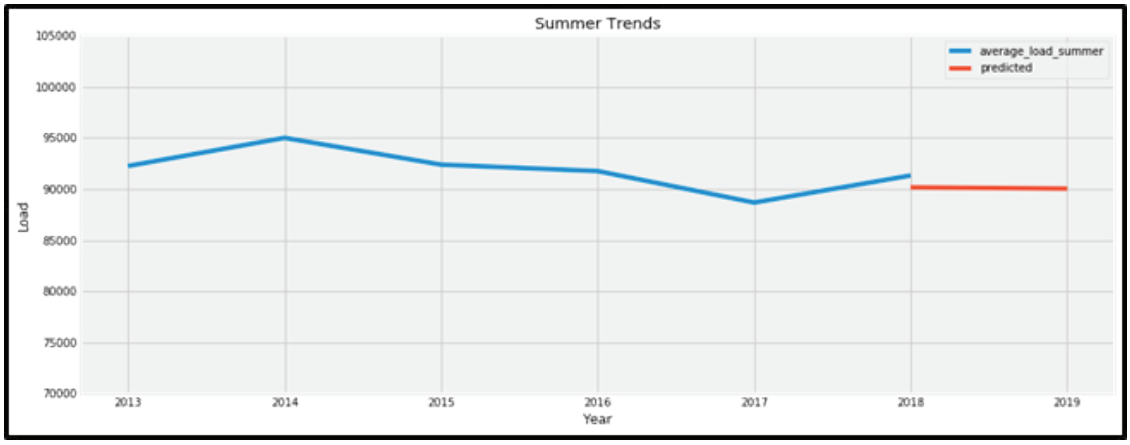
The actual monthly average load for the summer of 2018 is: 90634.872677

The predicted monthly average load for the summer of 2019 is: 93505.855114

Winter: We divided the data based on the months. We considered only November, December, January, February for our winter-season analysis, studied the trends and predicted the values for the next winter.



Predicted



Inference:

The actual monthly average load for the winter of 2018 is: 91321.152043

The predicted monthly average load for the winter of 2019 is: 90037.439149

Qualetics as a Solution

Energy efficiency yields benefits like economical savings, reduced greenhouse gas emissions, and alleviates energy poverty. Analyzing the energy consumption data will provide actionable insights on the means of saving energy, proposing actions, and quantifying the gains that are obtained as a result of the proposed actions.

We studied the trends of the hourly loads and completed our analysis with some predictions of the average monthly load for the next year. We predicted the average total load in that region for the summer and winter of 2019. Our analysis showed that the ARIMA model was the best suitable model for predicting energy consumption.

Building predictive models involve dealing with large volumes of data and complex algorithms. Therefore, it becomes necessary to have an efficient computing environment with a team of expert data scientists to provide the utility companies, power grids, governments, and other stakeholders with accurate forecasts (actionable insights) that can save operating and maintenance costs, enhance the reliability of power supply and delivery system, and better decision making in the future. This is where the Qualetics platform can be trusted.

Data Extraction and Processing Pipelines

Our proprietary architecture built using some of the industry-standard protocols and tools like MQTT, Kafka, NoSQL data stores allows us to set up a data pipeline from the source to the analytical data store with relative ease. This allows us to apply pre-designed learning models to data as it is being generated.

Data Analysis

In addition to pre-designed and pre-trained models, having a continuous data stream that can pick up new data points or variations allows us to refine and tweak the models over time or develop new models altogether. Our expert Data Science team can periodically monitor the efficiency and performance of the models and update them over time

Visualization and Integration Platform

The results of the analysis are delivered through our proprietary Visualization and Integration platform. The platform can be integrated into your Products, Systems, and Processes by Single Sign-On (SSO) or APIs.

Along with features such as continuous monitoring and automated alerts, we can automate an otherwise manual and tedious process such as data analysis and help you build intelligent frameworks from your most valuable resource, your data.

About Qualetics Data Machines

Qualetics Data Machines is a venture-funded startup established in July 2018, in the state of New Jersey in the USA. The company was founded and is led by the CEO, Sumanth Vakada, a software industry veteran with 20-years of experience in building modern software applications across various business domains. Qualetics currently has its core team based in the USA and an extended Application Development and Data Science team based in Hyderabad, India.

Our mission is to enable organizations to make an easier transition into the fields of Data Analytics and Artificial Intelligence. We aim to achieve this more efficiently than an organization might expect, were it to invest in the manpower and technology to build such complex platforms. For a universal need such as Data Science and AI, a dedicated focus is absolutely critical and Qualetics empowers our clients with the knowledge that allows them to apply in their core business offerings.

Key Features

Data Ingetion Platform

Visual Insight Delivery

Lisence Pre-Existing Models

API based Intelligence Delivery

Integration with Apps

Develop Custom Models & Insights

For Inquiries, please contact:

Mike Fowler

Chief Commercial Officer

mike@qualetics.com

630.715.4540

www.qualetics.com