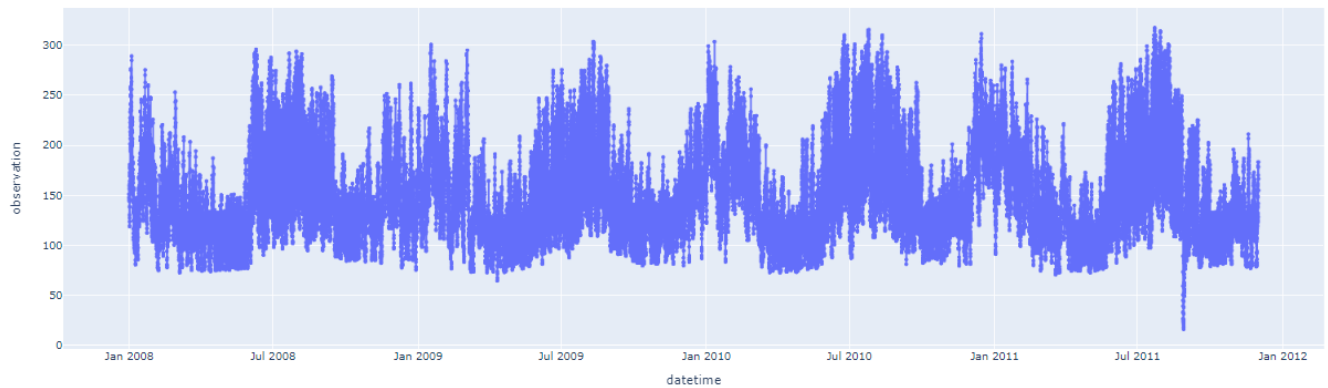# Q1.

- In 1st step **Data cleaning** is done to check for any missing value (nothing was missing).
- Before forecasting the data looked like:

- First two columns are converted to yyyy-mm-dd hh:mm:ss format (for easy handling of the data and the further steps).

- We can see that energy consumption has some pattern. By this it is concluded that the energy consumption has daily, weekly as well as seasonal patterns. So these can be used as features for predicting.

- **Data-Preprocessing**
  First two columns are converted to yyyy-mm-dd hh:mm:ss format later divided in to different columns like date, hour_of_day, day_of_week, quarter_of_year, month_of_year, year, day_of_year, day_of_month, for better insight of the data.

- **Create a model.**

  - A 2-layer LSTM based Neural network is used for the problem.
  - Where X_train features are Load, hour_of_day, day_of_week, quarter_of_year, month_of_year, year, day_of_year, day_of_month, week_of_year.

  - And, Y_train is Load

    e.g. 1st Jan 2008-5th Jan 2008 is used to predict the load for 1st Jan 2009.
    2nd Jan 2008-6th Jan 2008 is used to predict the load for 2nd Jan 2009.

  o Note: since we are trying to predict as per above example. I have removed 29th Feb,2008 from the data, as 2008 is a leap year. This is done so as to make all the 4 years to be non-leap.
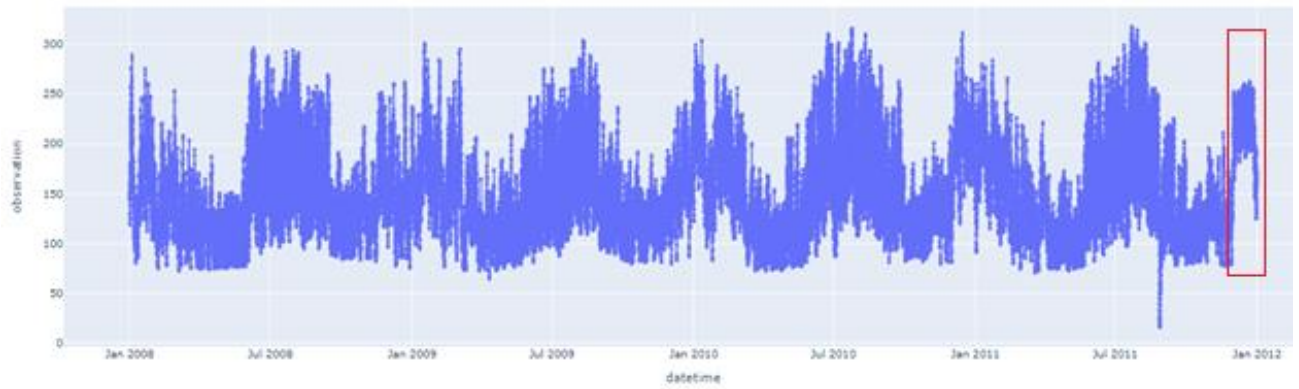
- **The graph below shows the average loads vs different features.**

Before-perdiction	After-perdiction
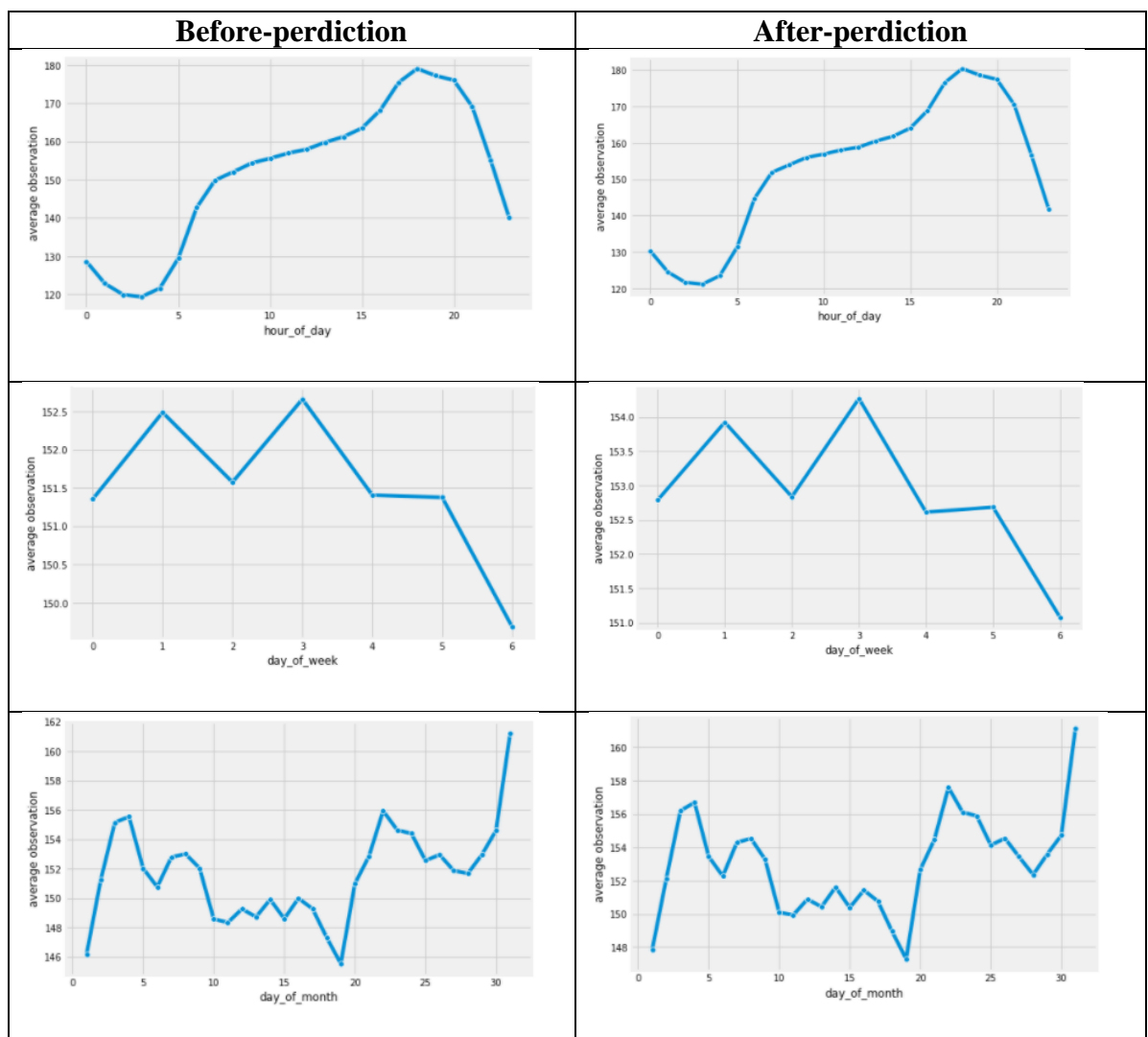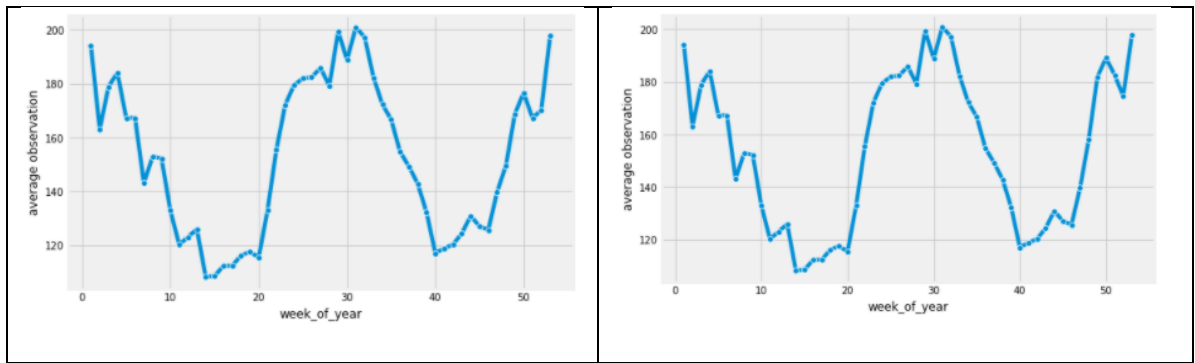
- As can be seen from the graphs after prediction, the trend is not changing, this indicates that using these as features will help in good predictions.

- Predicted values are stored in the submission.csv file.

/end