Electrical & Computer Engineering Department

EE 6013: Topic in Smart Grid

Forecasting The Demand for University of New Brunswick

Project Report

Tolulope Olugbenga (Student ID: 3643581)

November 29th, 2019.

Table of Contents

1 Introduction

Electricity is one of the driving forces of economic development and is essential to our daily life and wellbeing. Power demand forecasting is a difficult task due to the number of the different random variables that needs to be taken into consideration in order to predict human behavior. People often use electricity at any time that suits their lifestyle, and for the most part we all happen to use electricity at the same time. Most people share a similar lifestyle pattern, from when we wake up, to having a shower, making some breakfast, leaving for work, coming back at night, going to bed, doing our laundry on weekends and so on.

Load forecasting is an integral part in the process of the planning and operation of electric utilities; it has played a vital role in the power industry for over a century. In terms of power supply and demand; for the stable supply of electricity, the reserve power must be prepared. Businesses needs of load forecasting includes power systems planning/operations, revenue projection, rate design, energy trading, and so on. Load forecasting is needed by many business entities other than electric utilities, such as regulatory commissions, industrial/commercial companies, banks, trading firms, and insurance companies [1]. The demand pattern is very complex due to the deregulation of energy markets; therefore finding an appropriate forecasting model for a specific electricity network is not a trivial task [2]. Electricity demand is accessed by accumulating the consumption periodically; it can be considered for hourly, daily, weekly, monthly, and yearly periods.

The aim of this project was to predict the demand for the University of New Brunswick (UNB); six different approaches were used namely, simple average, moving average, naive approach, holt-winter method, auto regressive model (AR), and auto regressive integrated moving average (ARIMA). The models were used to predict; one week ahead, two weeks ahead, one month ahead, and the demand for December, 2017. The results were compared using the standard accuracy metrics; mean absolute percent error, mean error, mean absolute error, mean percent error, root mean squared error, and min-max error.

2. Description of the Models

This section gives a brief description of all the models used in this experiment.

2.1 The Simple Average Approach

The simple average approach takes the sum of all occurrences in the data and divides it by the number of the occurrences. It works best when the mean of the data remains the same or close to each other in different time periods. If there are no trends in the data, and the mean remains the same at different time periods; it could be predicted that the next day's demand is similar to the average of the previous days. This forecasting technique is known as the simple average technique; the formula for this approach can be seen below.

$$\text{Hence } \hat{y}_{x+1} = \frac{1}{x} \sum_{i=1}^{x} y_i$$

Equation 1:- The formula for the simple average approach

2.2 The Moving Average Approach

Unlike the simple average approach; the moving average approach takes the average of previous fixed n observations and uses it as a prediction for the next day. This is slight improvement as compared to the simple average approach, as it more specific. Using the moving average model, we can forecast the next day's demand based on the average of a fixed number of previous values; the formula for this approach can be seen below, where i > p. An advancement to this approach will the weighted moving average approach; in this approach the previous values are given different weights in terms of their significance.

$$\hat{y}_i = \frac{1}{p}(y_{i-1} + y_{i-2} + y_{i-3} \ldots\ldots + y_{i-p})$$

Equation 2:- The formula for the moving average approach

2.3 The Naive Approach

The naive approach predicts the current observation as the previously observed value. This approach works best if the previous day's value is similar to the present day. It is sometimes called the similar day approach. The naive approach is not suited for datasets with high variability, it is best suited for stable datasets; the formula for this approach can be seen below.

$$\text{Hence } \hat{y}_{t+1} = y_t.$$

Equation 3:- The formula for the naive approach

Due to weekly similarity of the data; the naive approach implemented in this project takes the previous week's value, e.g. this Tuesday is similar with last Tuesday.

2.4 Holt-Winters Method

Holt-Winters forecasting method is a way to model and predict the behavior of the previous data. This method models three aspects of the data; the average, the trend, and the seasonality (repeating pattern). This method takes into account both the trend and seasonality to forecast future values. The idea behind this method is to apply exponential smoothing to the seasonal components in addition to level and trend. The Holt-Winters seasonal method comprises the forecast equation and three smoothing equations; one for the level L_t, one for trend b_t and one for the seasonal component denoted by S_t, with smoothing parameters α, β and γ. The formulas for this method can be seen below, where s is the length of the seasonal cycle, for $0 \leq \alpha \leq 1$, $0 \leq \beta \leq 1$ and $0 \leq \gamma \leq 1$ [3].

$$\begin{aligned}
\text{level} \quad L_t &= \alpha(y_t - S_{t-s}) + (1-\alpha)(L_{t-1} + b_{t-1}); \\
\text{trend} \quad b_t &= \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1}, \\
\text{seasonal} \quad S_t &= \gamma(y_t - L_t) + (1-\gamma)S_{t-s} \\
\text{forecast} \quad F_{t+k} &= L_t + kb_t + S_{t+k-s},
\end{aligned}$$

Equation 4:- The formulas for the Holt-Winters method

Due to weekly similarity of the data; a seasonality period of 7 was used for most of the horizons.

2.5 The Autoregressive Model (AR)

The autoregressive model is used when there are some correlation between the values in a time series. The process is basically a linear regression of the data in the current series against one or more past values in the same series. An AR(p) model is an autoregressive model where specific lagged values of y_t are used as predictor variables. Lags are where results from one time period affect following periods. The value for "p" is called the order [4]. The formula for this model can be seen below; where p is the order, c is a constant, and e_t is white noise.

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t$$

Equation 5:- The formula for the AR model

2.6 The Autoregressive Integrated Moving Average (ARIMA)

ARIMA is a class of models that explains a given time series based on its own past values, that is, its own lags and the lagged forecast errors, so that equation can be used to forecast future values. An ARIMA model is characterized by p, d, q; where p is the order of the AR term, q is the order of the MA term, and d is the number of differencing required in order to make the time series stationary. An ARIMA model is one where the time series was differenced at least once to make it stationary and you combine the AR and the MA terms; the formula for this model can be seen below [5].

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + .. + \beta_p Y_{t-p} \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + .. + \phi_q \epsilon_{t-q}$$

Equation 6:- The formula for the ARIMA model

The SARIMA takes into account the seasonality of the dataset, seasonal differencing is used in this case. Seasonal differencing is similar to regular differencing, but instead it subtracts the value from the previous season. The model will be represented as SARIMA(p,d,q)x(P,D,Q), where, P, D and Q are SAR, order of seasonal differencing and SMA terms respectively and 'x' is the frequency of the time series. This is best suited for the dataset because it contains some seasonality.

3. Dataset

The time series data was gotten from 'Primary A lagged demand' feeder at UNB. The data contains the power demand at a 15 mins interval starting from '2014-03-20 07:30:00' until '2019-11-05 10:15:00'. The Datetime and the demand were the necessary data needed.

3.1 Filtering the Dataset

I plotted the data, and noticed some unusual spikes and instances where the demand was really low and almost zero. I also noticed there was a huge gap, running into months in the 2014 data; for this reason, only data from 2015 were used to train our models. All the remaining outliers were replaced with their previous week's value. In order to our data simpler; I took the daily mean of the time series. The figure below shows a plot of the final data used in training and testing our models.

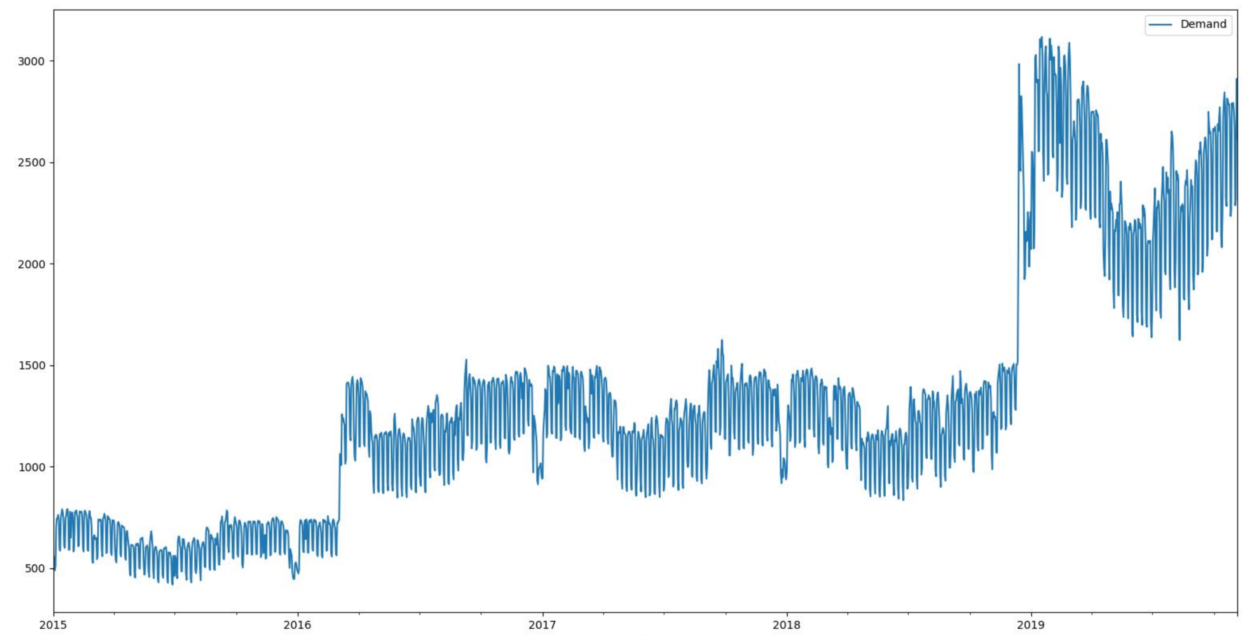

Figure 1:- The figure shows the plot of the time series data used in this experiment

4 Implementation

This section discusses the experiments carried out on the one week, two weeks, one month, and December 2017 periods.

4.1 One Week Forecast

First, I had to split the data into two; one for training, and one for testing. The training data contained the data from 2015 still the last 7 days of the data (data[2015: len(data) - 7]), and the testing data contained the last 7 days of the data.

4.1.1 Simple Average Approach

I took the average of all the data in the time series. The figure below shows a plot of the training data from beginning of 2019, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 50.08%, ME: -1322.25Kw, MAE: 1322.25Kw, MPE: -50.08%, RMSE: 1340.30Kw, and Min-Max Error: 0.50.

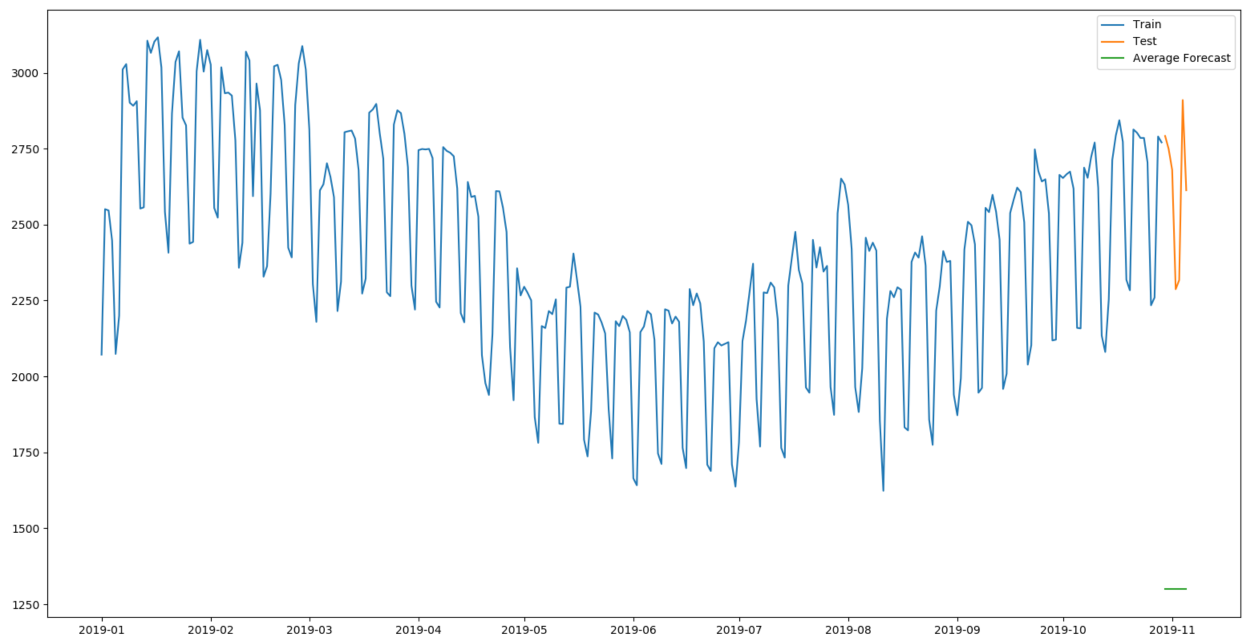

Figure 2:- The figure shows the plot of the training data, the testing data, and the predicted values using the simple average approach.

4.1.2 Moving Average Approach

I took the average of the previous 7 days in the time series. The figure below shows a plot of the training data from beginning of 2019, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 7.28%, ME: -2.69Kw, MAE: 185.05Kw, MPE: -

0.63%, RMSE: 219.25Kw, and Min-Max Error: 0.07. Based on the MAPE and RMSE values; we can observe that there is a slight improvement in the results as compared to the simple average approach.

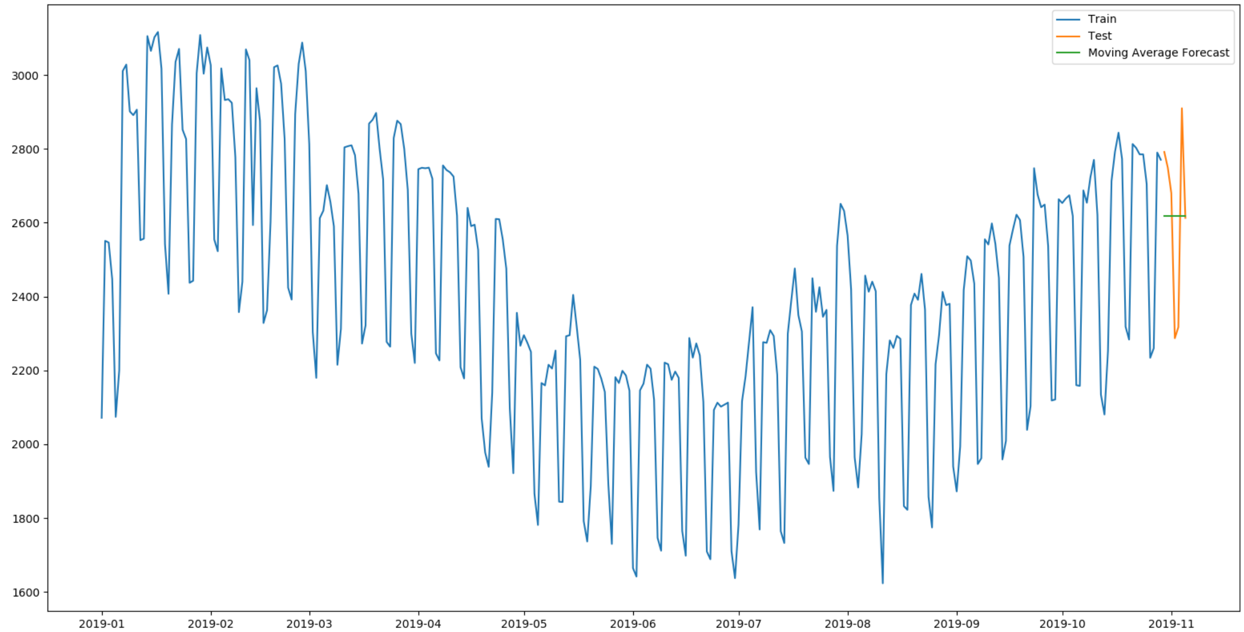

Figure 3:- The figure shows the plot of the training data, the testing data, and the predicted values using the moving average approach.

4.1.3 The Naive Approach

I took the previous week's values as the forecast for the current week, it is sometimes called the similar day approach. The figure below shows a plot of the training data from beginning of 2019, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 2.49%, ME: -2.69Kw, MAE: 65.04Kw, MPE: -0.13%, RMSE: 82.20Kw, and Min-Max Error: 0.02. Based on the MAPE and RMSE values; we can observe some improvement in the results as compared to the moving average approach.

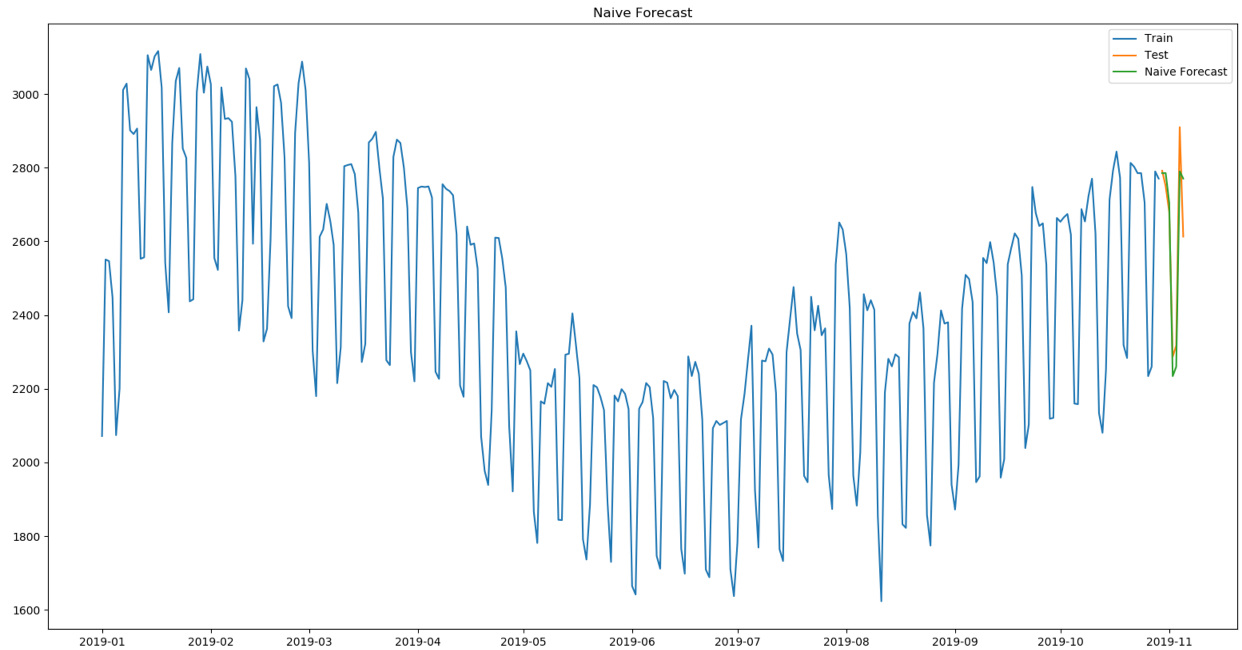

Figure 4:- The figure shows the plot of the training data, the testing data, and the predicted values using the naive approach.

4.1.4 Holt-Winter Method

The Holt-Winter's method was done with a seasonal period of 7. The figure below shows a plot of the training data from beginning of 2019, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 3.37%, ME: 9.74Kw, MAE: 90.14Kw, MPE: 0.38%, RMSE: 112.82Kw, and Min-Max Error: 0.03. Based on the MAPE and RMSE values; this method is currently in second place.

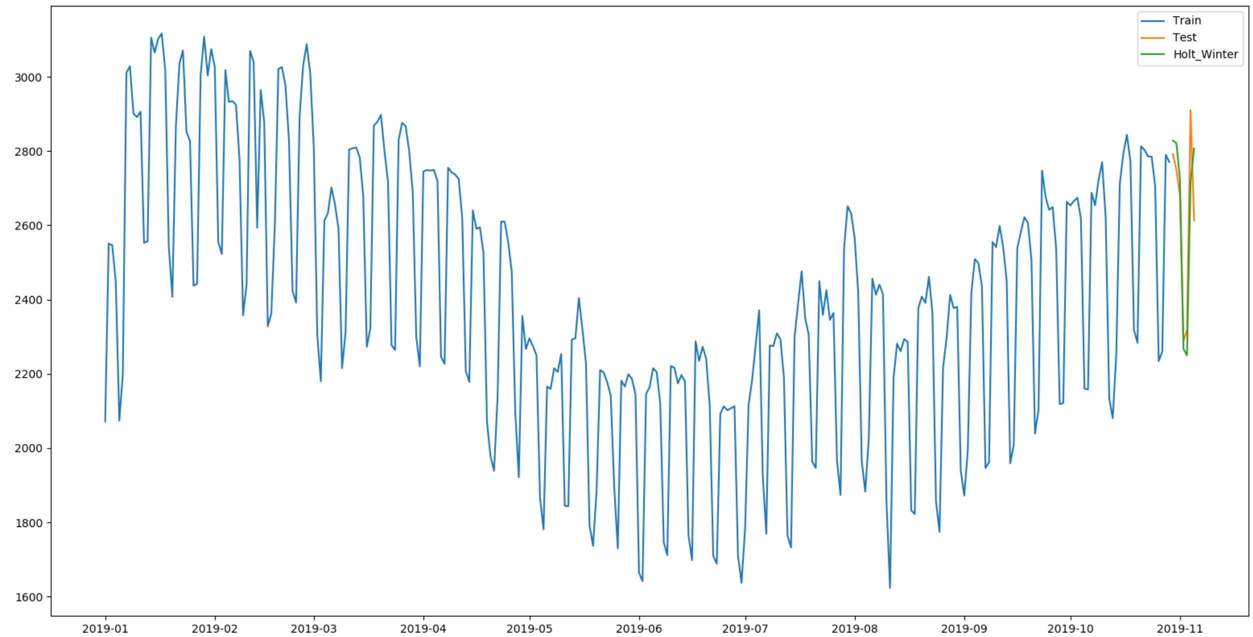

Figure 5:- The figure shows the plot of the training data, the testing data, and the predicted values using the Holt-Winter's method.

4.1.5 Auto Regressive Model (AR)

The figure below shows a plot of the training data from beginning of 2019, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 3.62%, ME: 7.21Kw, MAE: 98.05Kw, MPE: 0.32%, RMSE: 128.86Kw, and Min-Max Error: 0.04. Based on the MAPE and RMSE values; this method is lagging behind the Holt-Winter's method.

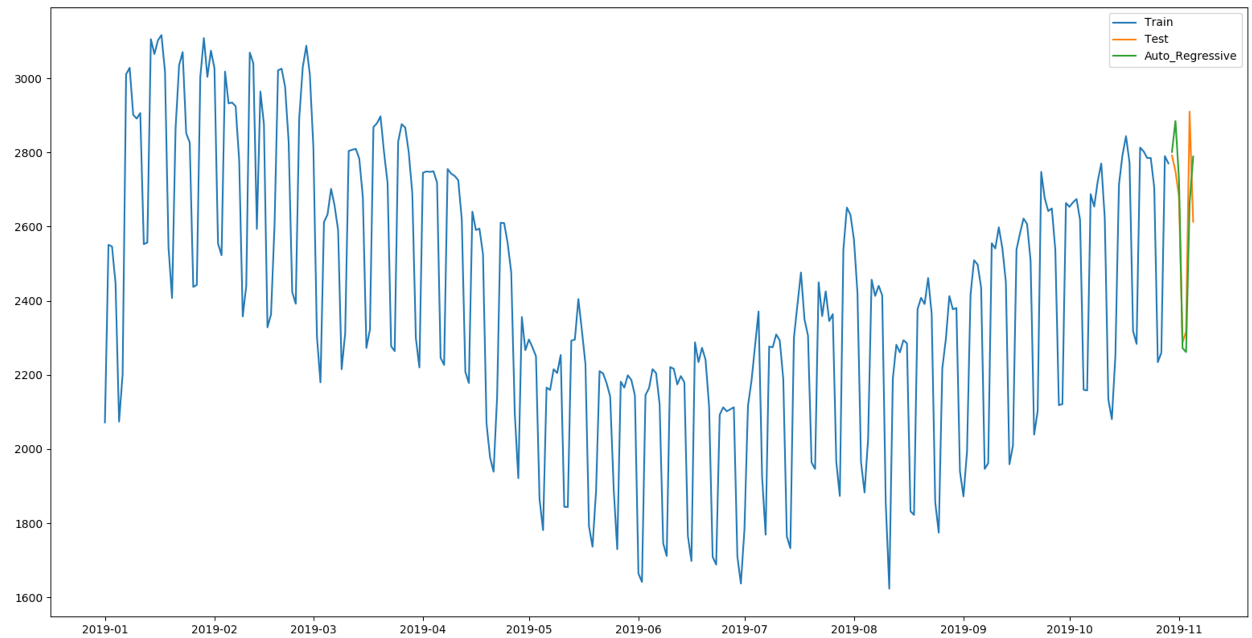

Figure 6:- The figure shows the plot of the training data, the testing data, and the predicted values using the AR model.

4.1.6 Auto Regressive Integrated Moving Average (ARIMA)

The ARIMA model was trained with an order of (0, 1, 2), and a seasonal order of (1, 0, 1, 7). The figure below shows a plot of the training data from beginning of 2019, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 3.02%, ME: 15.48Kw, MAE: 81.15Kw, MPE: 0.64%, RMSE: 107.95Kw, and Min-Max Error: 0.03. Based on the MAPE and RMSE values; this method performed better than the Holt-Winter's method, while the naive approach is in first place.

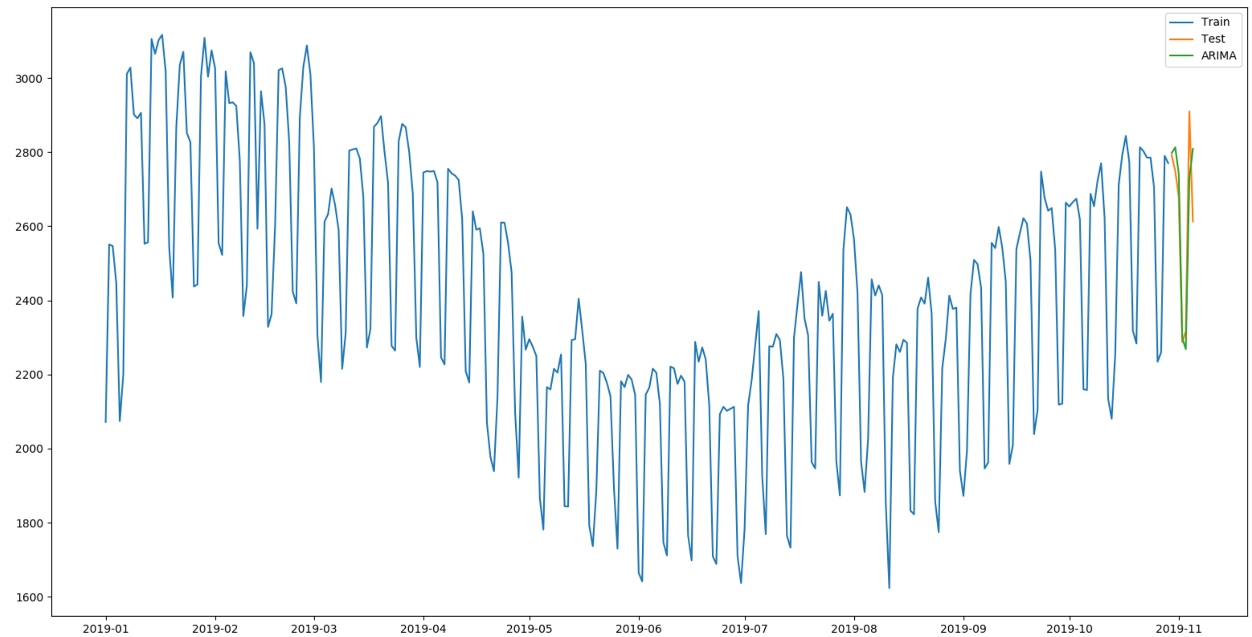

Figure 7:- The figure shows the plot of the training data, the testing data, and the predicted values using the ARIMA model.

4.2 Two Weeks Forecast

First, I had to split the data into two; one for training, and one for testing. The training data contained the data from 2015 still the last 14 days of the data (data[2015: len(data) - 14]), and the testing data contained the last 14 days of the data.

4.2.1 Simple Average Approach

I took the average of all the data in the time series. The figure below shows a plot of the training data from beginning of 2019, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 50.22%, ME: -1326.17Kw, MAE: 1326.17Kw, MPE: -50.22%, RMSE: 1345.64Kw, and Min-Max Error: 0.50.

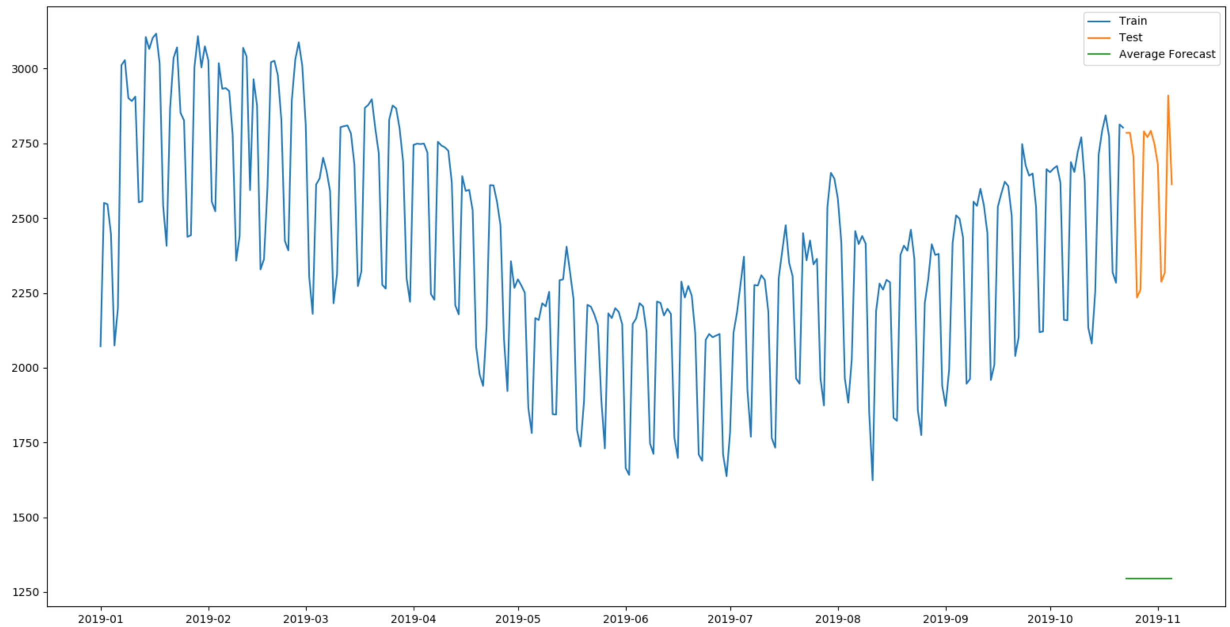

Figure 8:- The figure shows the plot of the training data, the testing data, and the predicted values using the simple average approach.

4.2.2 Moving Average Approach

I took the average of the previous 7 days in the time series. The figure below shows a plot of the training data from beginning of 2019, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 7.58%, ME: 41.06Kw, MAE: 186.57Kw, MPE: 2.40%, RMSE: 231.76Kw, and Min-Max Error: 0.07. Based on the MAPE and RMSE values; we can observe that there is a slight improvement in the results as compared to the simple average approach.

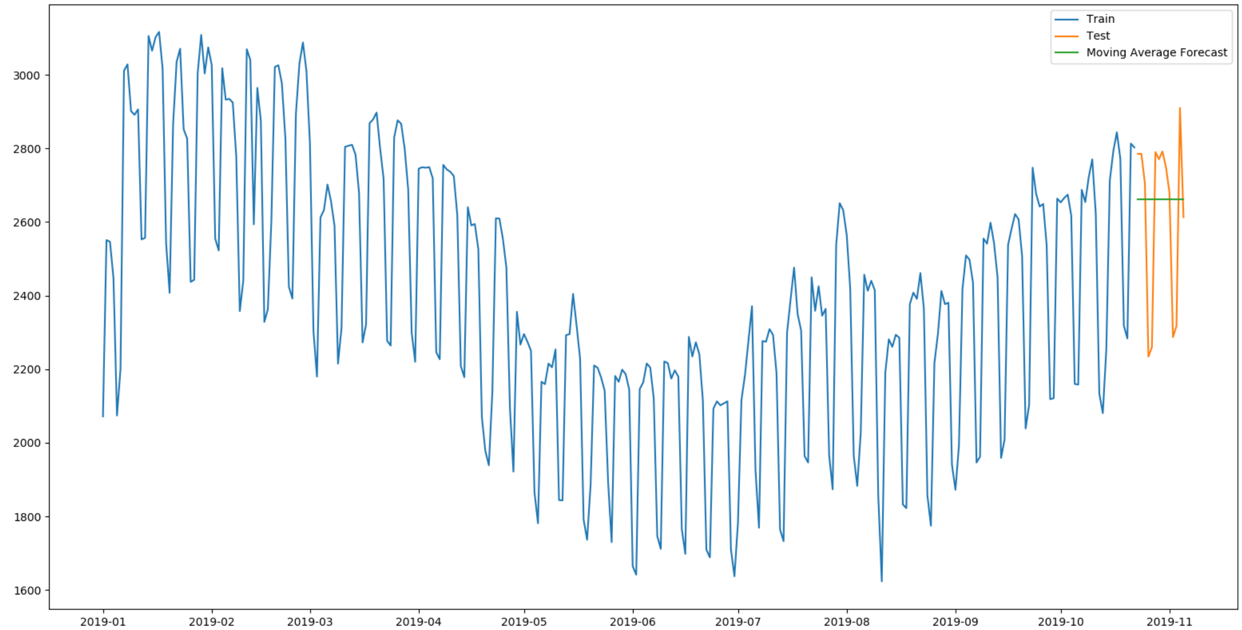

Figure 9:- The figure shows the plot of the training data, the testing data, and the predicted values using the moving average approach.

4.2.3 The Naive Approach

I took the previous week's values as the forecast for the test weeks, it is sometimes called the similar day approach. The figure below shows a plot of the training data from beginning of 2019, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 2.29%, ME: 41.06Kw, MAE: 59.73Kw, MPE: 1.60%, RMSE: 76.62Kw, and Min-Max Error: 0.02. Based on the MAPE and RMSE values; we can observe some improvement in the results as compared to the moving average approach.

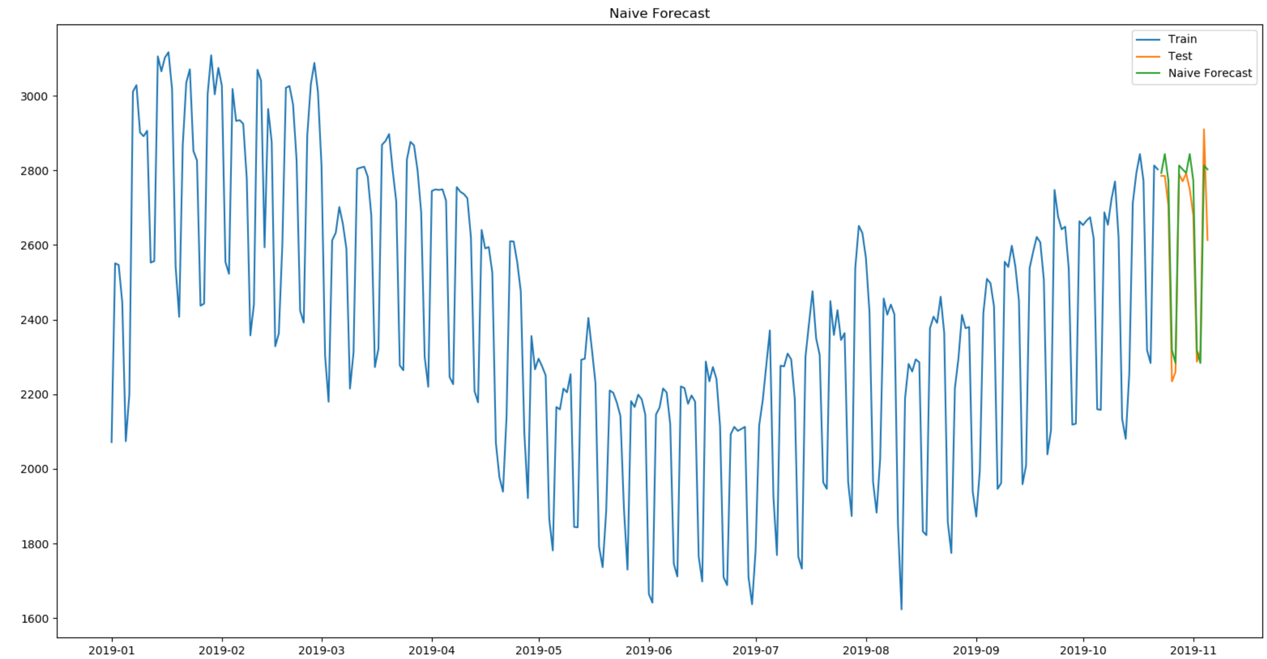

Figure 10:- The figure shows the plot of the training data, the testing data, and the predicted values using the naive approach.

4.2.4 Holt-Winter Method

The Holt-Winter's method was done with a seasonal period of 7. The figure below shows a plot of the training data from beginning of 2019, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 2.95%, ME: 41.24Kw, MAE: 79.04Kw, MPE: 1.59%, RMSE: 97.99Kw, and Min-Max Error: 0.03. Based on the MAPE and RMSE values; this method is currently in second place.

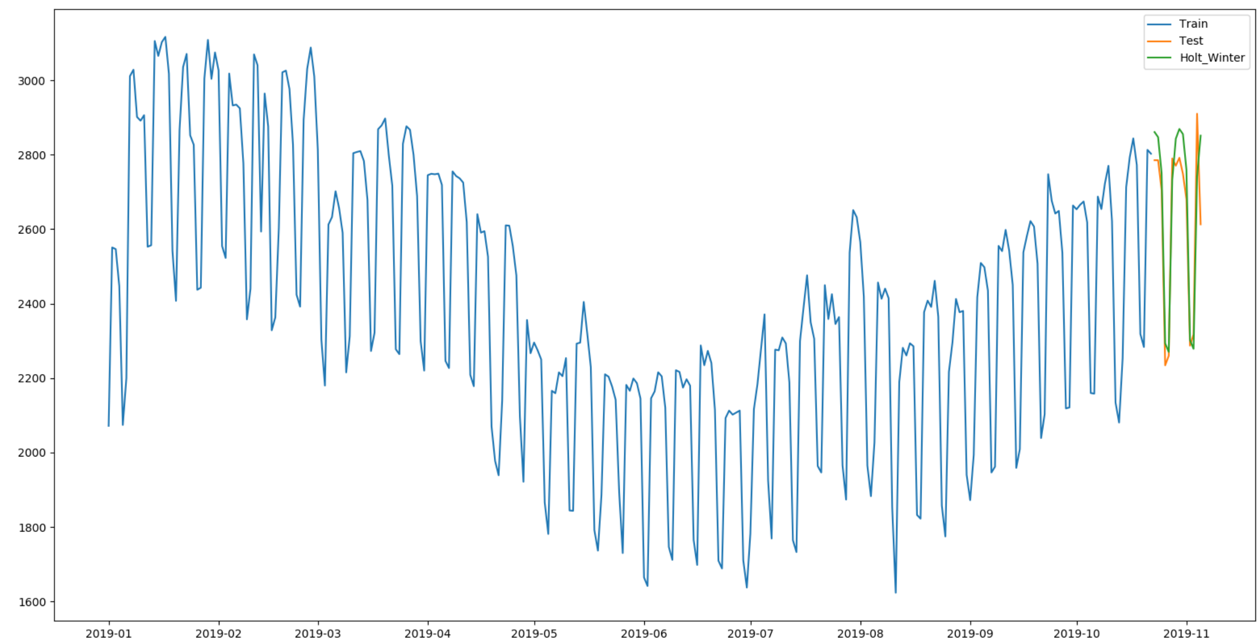

Figure 11:- The figure shows the plot of the training data, the testing data, and the predicted values using the Holt-Winter's method.

4.2.5 Auto Regressive Model (AR)

The figure below shows a plot of the training data from beginning of 2019, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 3.28%, ME: 13.46Kw, MAE: 88.27Kw, MPE: 0.59%, RMSE: 119.33Kw, and Min-Max Error: 0.03. Based on the MAPE and RMSE values; this method is lagging behind the Holt-Winter's method.

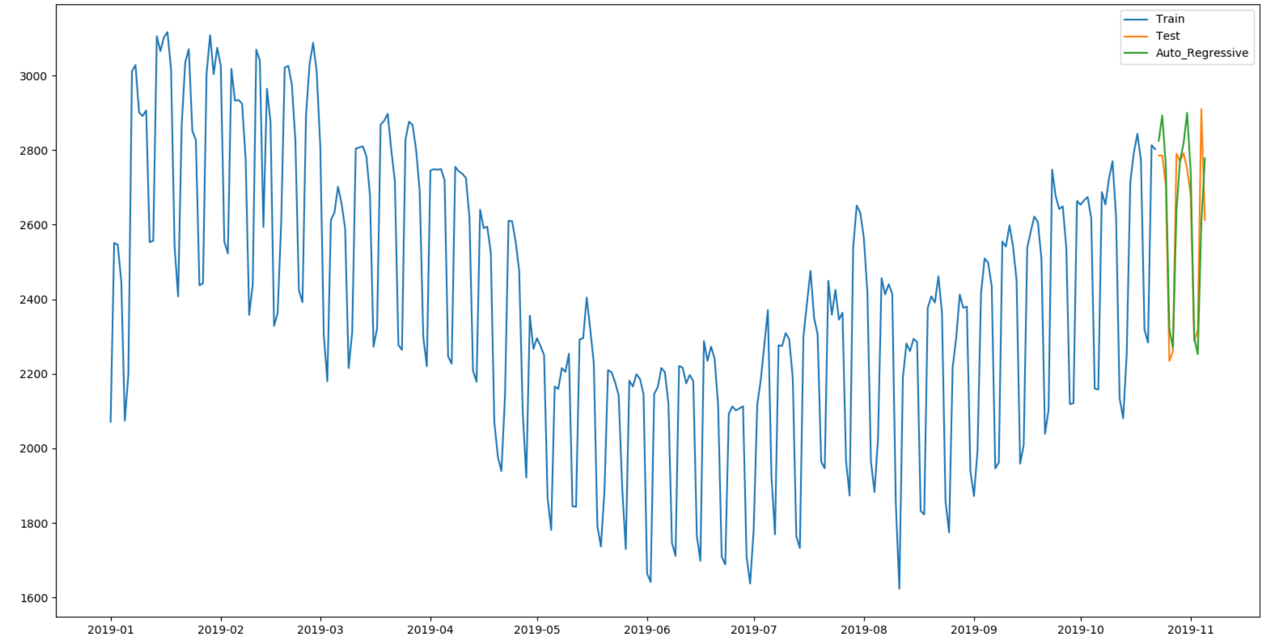

Figure 12:- The figure shows the plot of the training data, the testing data, and the predicted values using the AR model.

4.2.6 Auto Regressive Integrated Moving Average (ARIMA)

The ARIMA model was trained with an order of (1, 1, 7), and a seasonal order of (1, 0, 2, 7). The figure below shows a plot of the training data from beginning of 2019, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 2.58%, ME: 32.54Kw, MAE: 67.98Kw, MPE: 1.21%, RMSE: 83.32Kw, and Min-Max Error: 0.03. Based on the MAPE and RMSE values; this method performed better than the Holt-Winter's method, while the naive approach is in first place.

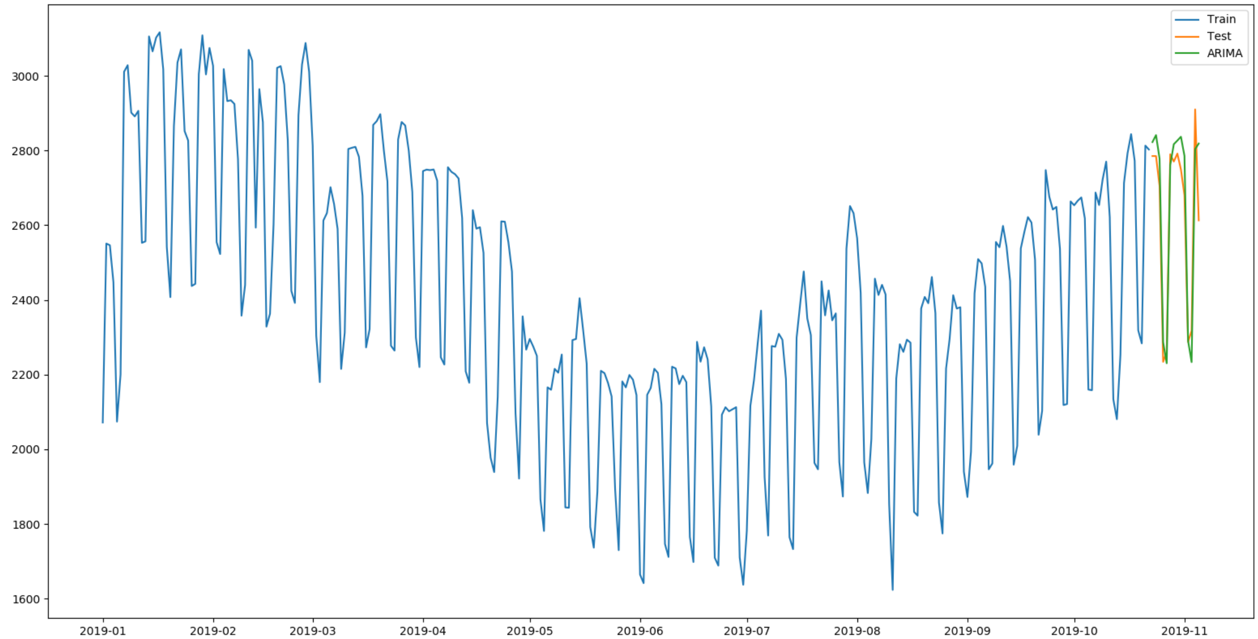

Figure 13:- The figure shows the plot of the training data, the testing data, and the predicted values using the ARIMA model.

4.3 One Month Forecast

First, I had to split the data into two; one for training, and one for testing. The training data contained the data from 2015 still the last 31 days of the data (data[2015: len(data) - 31]), and the testing data contained the last 31 days of the data.

4.3.1 Simple Average Approach

I took the average of all the data in the time series. The figure below shows a plot of the training data from beginning of 2019, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 49.89%, ME: -1302.46Kw, MAE: 1302.46Kw, MPE: -49.89%, RMSE: 1345.64Kw, and Min-Max Error: 0.50.

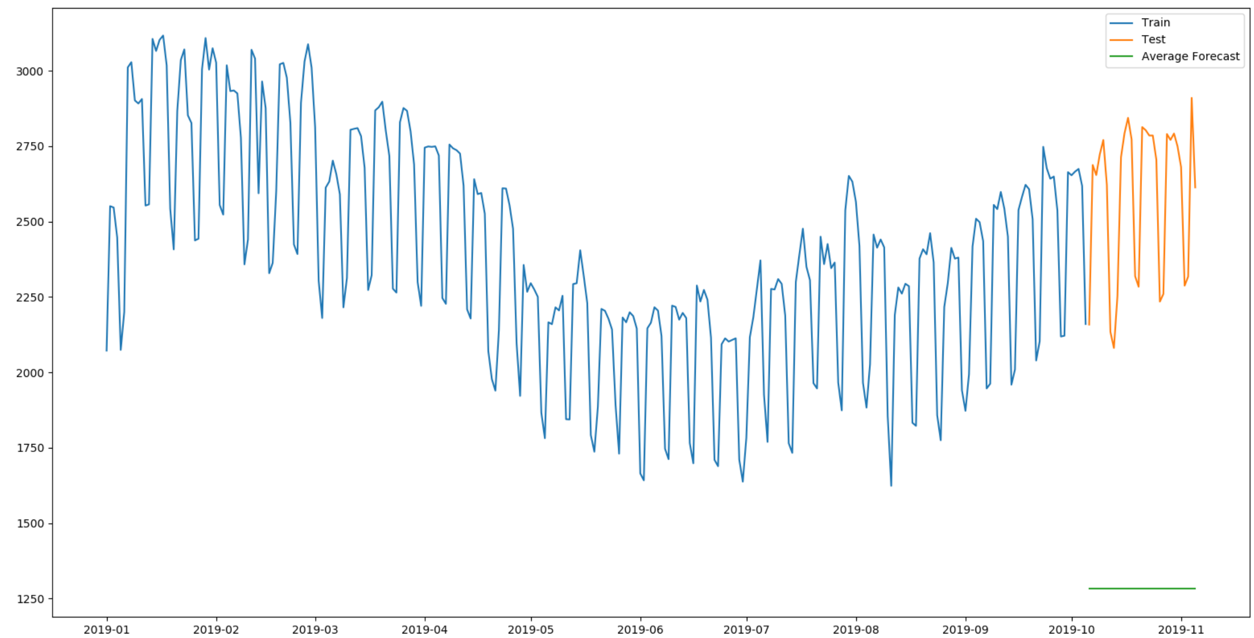

Figure 14:- The figure shows the plot of the training data, the testing data, and the predicted values using the simple average approach.

4.3.2 Moving Average Approach

I took the average of the previous 7 days in the time series. The figure below shows a plot of the training data from beginning of 2019, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 9.97%, ME: -75.89Kw, MAE: 253.51Kw, MPE: -1.92%, RMSE: 264.37Kw, and Min-Max Error: 0.09. Based on the MAPE and RMSE values; we can observe that there is a slight improvement in the results as compared to the simple average approach.

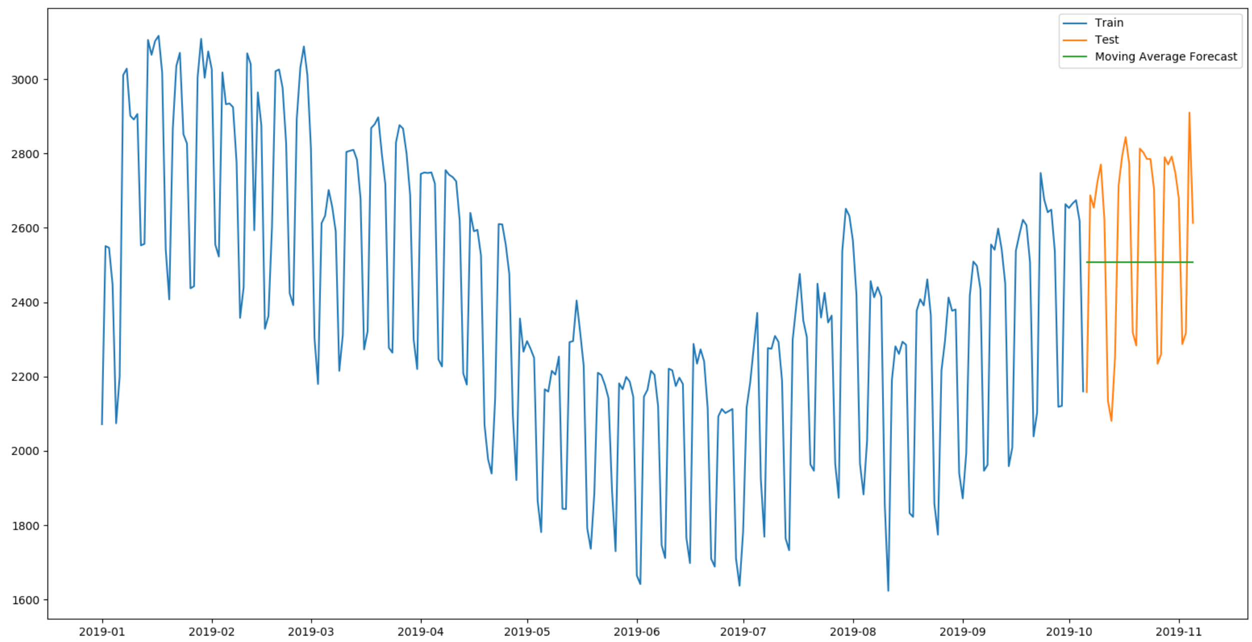

Figure 15:- The figure shows the plot of the training data, the testing data, and the predicted values using the moving average approach.

4.3.3 The Naive Approach

I took the previous week's values as the forecast for the test weeks, it is sometimes called the similar day approach. The figure below shows a plot of the training data from beginning of 2019, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 4.38%, ME: -78.66Kw, MAE: 112.01Kw, MPE: -2.90%, RMSE: 137.33Kw, and Min-Max Error: 0.04. Based on the MAPE and RMSE values; we can observe some improvement in the results as compared to the moving average approach.

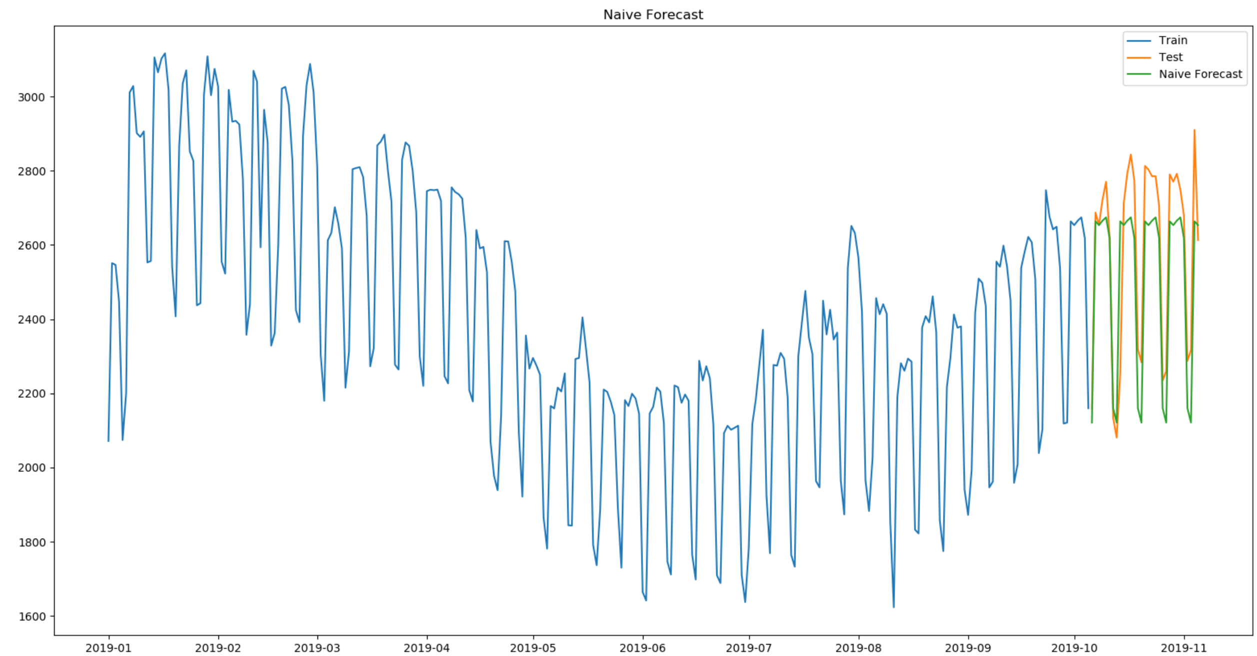

Figure 16:- The figure shows the plot of the training data, the testing data, and the predicted values using the naive approach.

<u>4.3.4 Holt-Winter Method</u>

The Holt-Winter's method was done with a seasonal period of 7. The figure below shows a plot of the training data from beginning of 2019, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 3.97%, ME: -65.64Kw, MAE: 101.31Kw, MPE: -2.41%, RMSE: 127.41Kw, and Min-Max Error: 0.04. Based on the MAPE and RMSE values; this methods did better than the Naive approach.

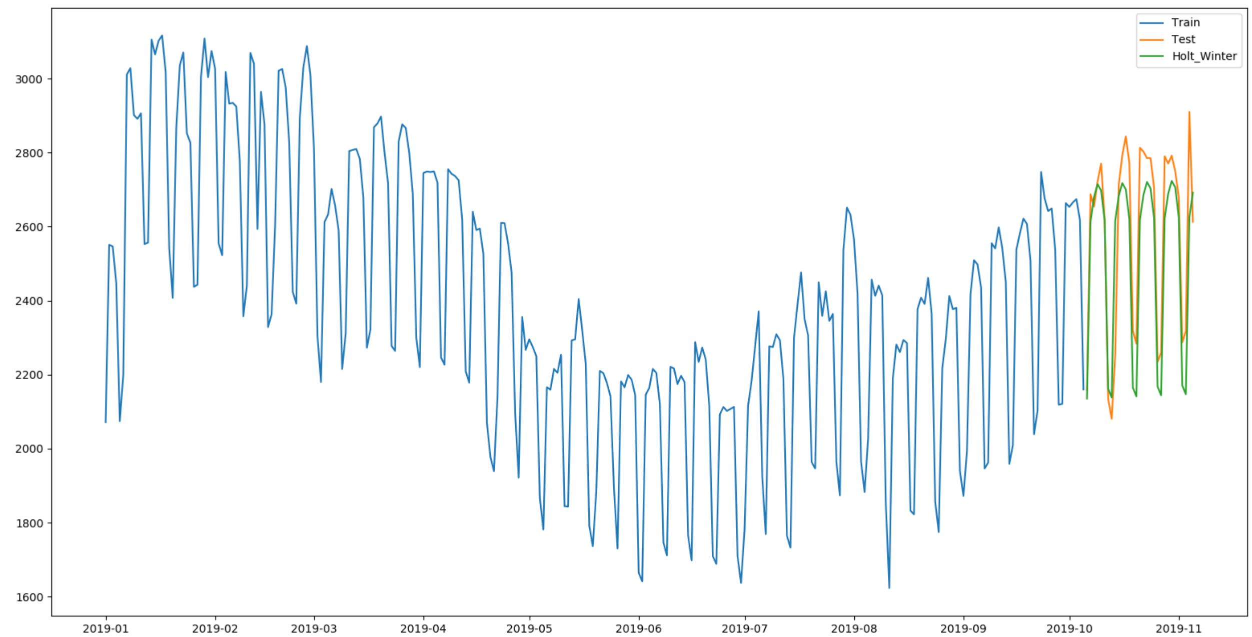

Figure 17:- The figure shows the plot of the training data, the testing data, and the predicted values using the Holt-Winter's method.

4.3.5 Auto Regressive Model (AR)

The figure below shows a plot of the training data from beginning of 2019, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 4.24%, ME: -58.14Kw, MAE: 108.85Kw, MPE: -2.03%, RMSE: 135.66Kw, and Min-Max Error: 0.04. Based on the MAPE and RMSE values; this method is lagging behind the Holt-Winter's method.

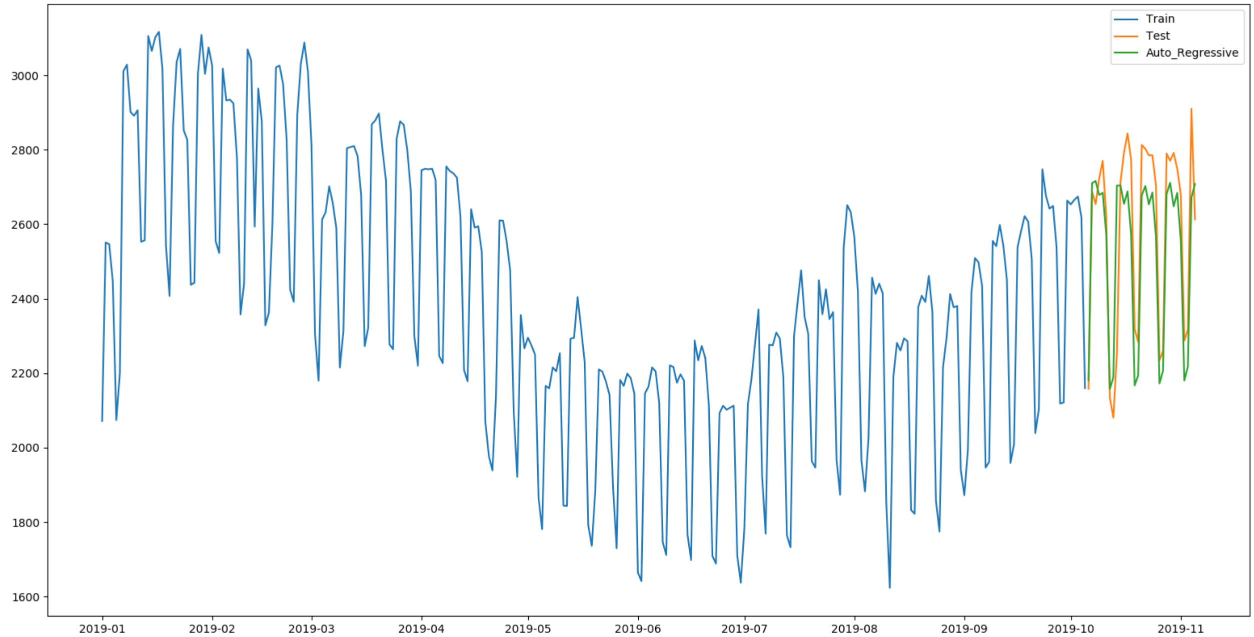

Figure 18:- The figure shows the plot of the training data, the testing data, and the predicted values using the AR model.

4.3.6 Auto Regressive Integrated Moving Average (ARIMA)

The ARIMA model was trained with an order of (1, 1, 2), and a seasonal order of (1, 0, 2, 7). The figure below shows a plot of the training data from beginning of 2019, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 3.12%, ME: -35.87Kw, MAE: 79.48Kw, MPE: -1.23%, RMSE: 109.16Kw, and Min-Max Error: 0.03. Based on the MAPE and RMSE values; this method had the best performance in the horizon.

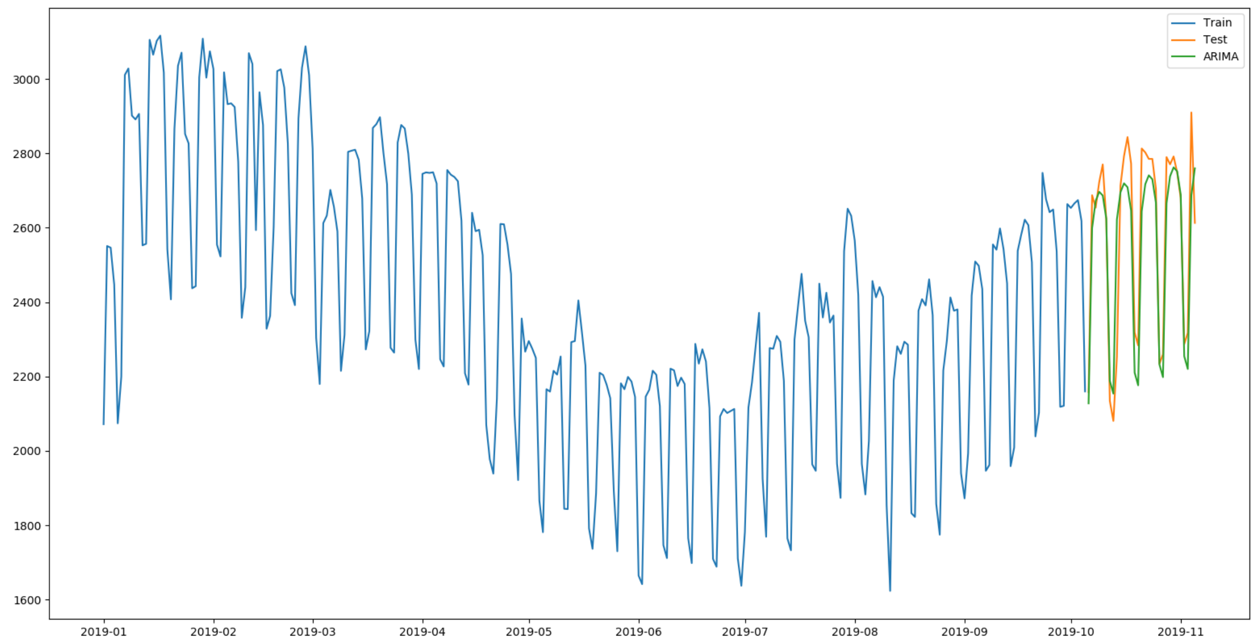

Figure 19:- The figure shows the plot of the training data, the testing data, and the predicted values using the ARIMA model.

4.4 December, 2017's Forecast

First, I had to split the data into two; one for training, and one for testing. The training data contained the data from 2015 still the end of November 2017, and the testing data contained December, 2017's data.

4.4.1 Simple Average Approach

I took the average of all the data in the time series. The figure below shows a plot of the training data from beginning of 2017, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 17.43%, ME: -211.95Kw, MAE: 227.37Kw, MPE: -15.79%, RMSE: 274.12Kw, and Min-Max Error: 0.17.

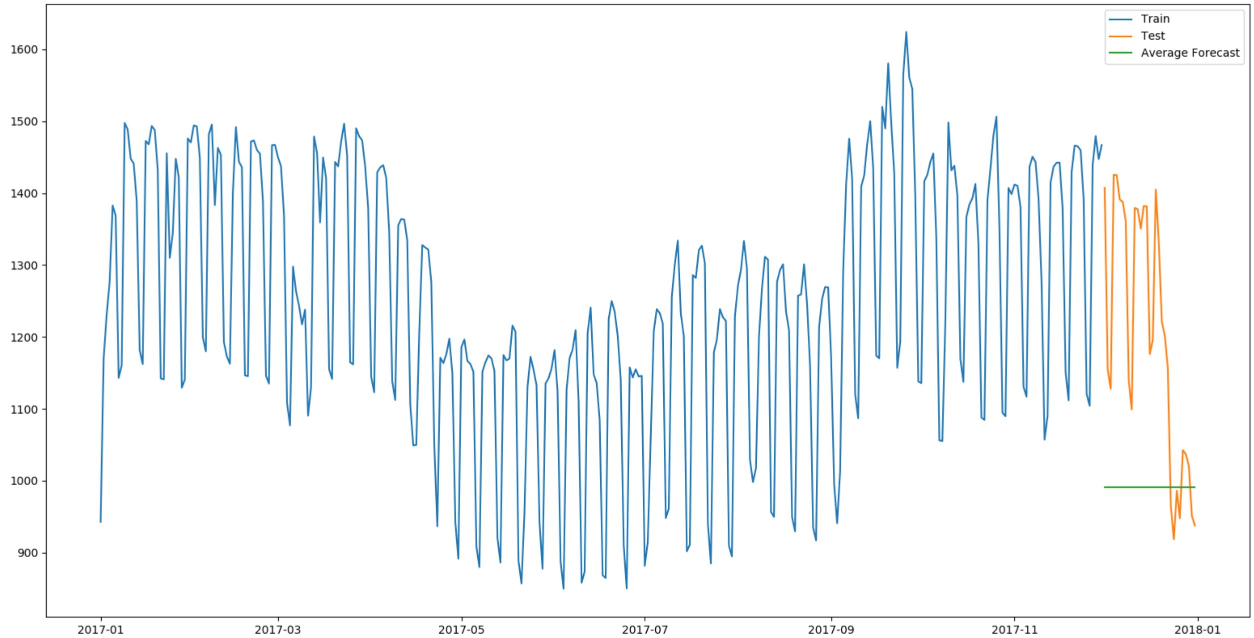

Figure 20:- The figure shows the plot of the training data, the testing data, and the predicted values using the simple average approach.

4.4.2 Moving Average Approach

I took the average of the previous 7 days in the time series. The figure below shows a plot of the training data from beginning of 2017, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 14.24%, ME: 41.09Kw, MAE: 159.60Kw, MPE: -5.71%, RMSE: 178.63Kw, and Min-Max Error: 0.12. Based on the MAPE and RMSE values; we can observe that there is a slight improvement in the results as compared to the simple average approach.

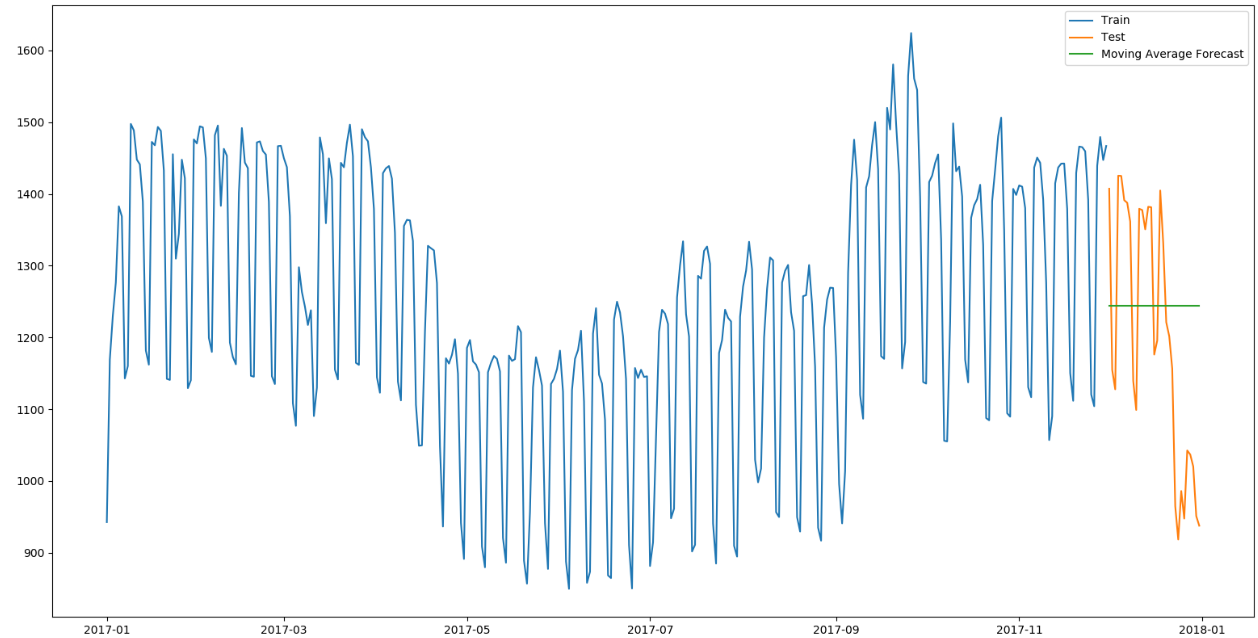

Figure 21:- The figure shows the plot of the training data, the testing data, and the predicted values using the moving average approach.

4.4.3 The Naive Approach

I took the previous week's values as the forecast for the test weeks, it is sometimes called the similar day approach. The figure below shows a plot of the training data from beginning of 2017, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 13.82%, ME: 133.26Kw, MAE: 148.71Kw, MPE: 12.52%, RMSE: 208.55Kw, and Min-Max Error: 0.11.

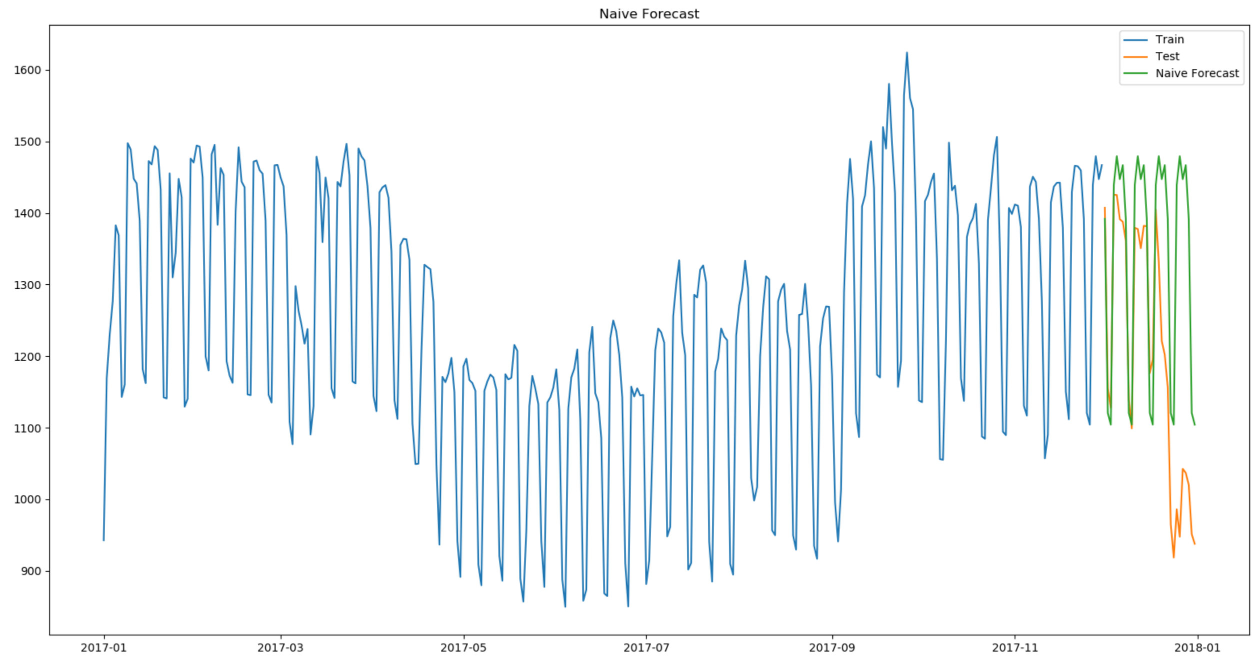

Figure 22:- The figure shows the plot of the training data, the testing data, and the predicted values using the naive approach.

4.4.4 Holt-Winter Method

The Holt-Winter's method was normally done with a seasonal period of 7, but there was an improvement when a seasonal period of 365 was used. The figure below shows a plot of the training data from beginning of 2017, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 11.81%, ME: 93.69Kw, MAE: 134.11Kw, MPE: 8.94%, RMSE: 155.35Kw, and Min-Max Error: 0.10. Based on the MAPE and RMSE values; this method has the best performance so far.

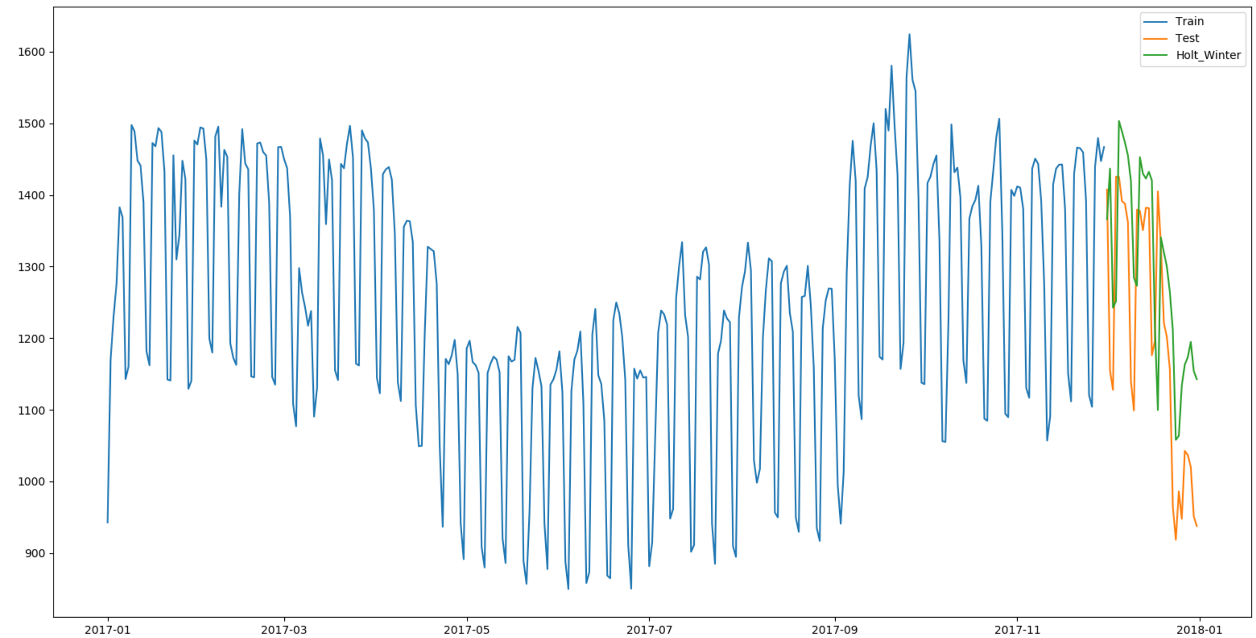

Figure 23:- The figure shows the plot of the training data, the testing data, and the predicted values using the Holt-Winter's method.

4.4.5 Auto Regressive Model (AR)

The figure below shows a plot of the training data from beginning of 2017, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 12.81%, ME: -121.94Kw, MAE: 136.43Kw, MPE: 11.63%, RMSE: 195.06Kw, and Min-Max Error: 0.10.

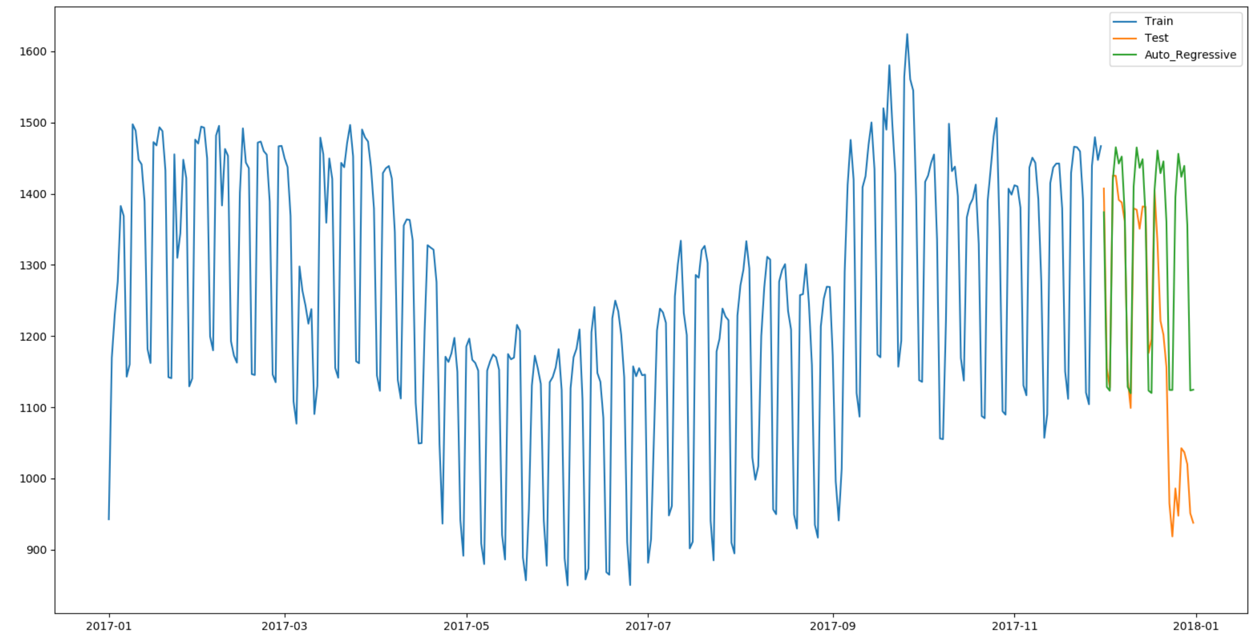

Figure 24:- The figure shows the plot of the training data, the testing data, and the predicted values using the AR model.

4.4.6 Auto Regressive Integrated Moving Average (ARIMA)

The ARIMA model was trained with an order of (1, 1, 2), and a seasonal order of (2, 0, 1, 7). The figure below shows a plot of the training data from beginning of 2017, the testing data, and the predicted values. The accuracy metrics function calculated values of; MAPE: 13.55%, ME: 133.83Kw, MAE: 144.65Kw, MPE: 12.66%, RMSE: 206.81Kw, and Min-Max Error: 0.11.

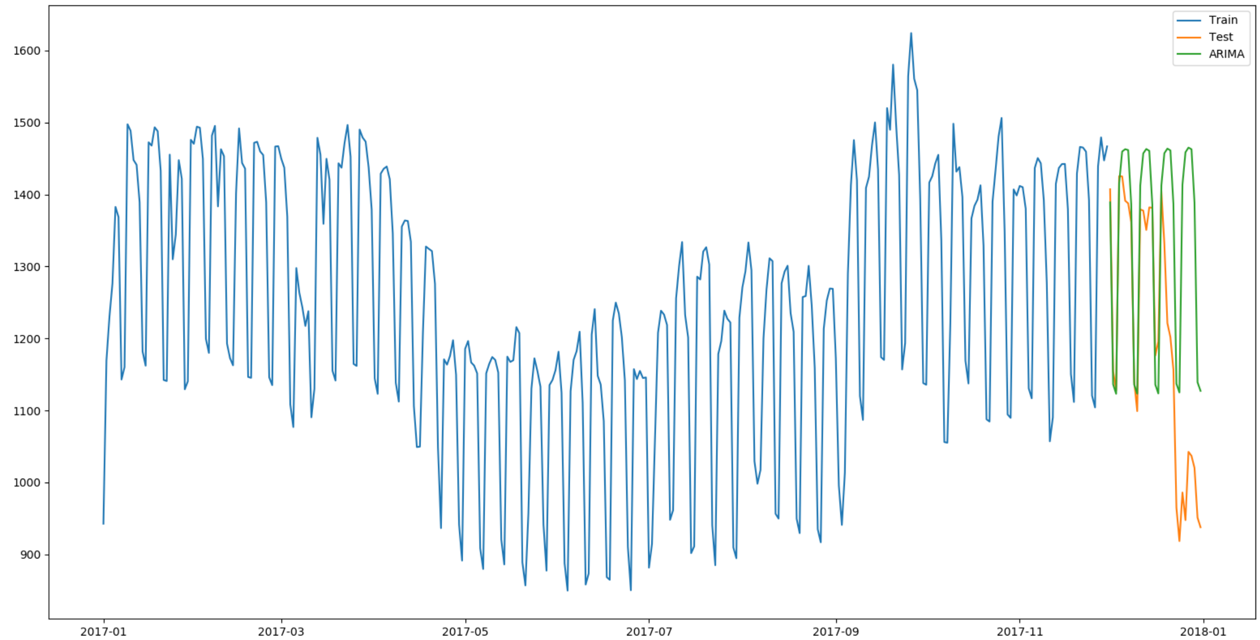

Figure 25:- The figure shows the plot of the training data, the testing data, and the predicted values using the ARIMA model.

5 Evaluation

5.1 Comparison of the results on a week's forecast

The table below shows a summary of the accuracy metrics gotten from the six methods based on how they performed.

	Simple Average	Moving Average	Naive Approach	Holt-Winter	AR model	ARIMA model
MAPE (%)	50.08	7.28	2.49	3.37	3.62	3.02
ME (Kw)	-1322.25	-2.69	-2.69	9.74	7.21	15.48
MAE (Kw)	1322.25	185.05	65.04	90.14	98.05	81.15
MPE (%)	-50.08	0.63	-0.13	0.38	0.32	0.64
RMSE (Kw)	1340.30	219.25	82.20	112.82	128.86	107.95
Min-Max Error	0.50	0.07	0.02	0.03	0.04	0.03

Table 1:- The table shows a summary of the accuracy metrics of all 6 methods on a week's forecast

In the one week's forecast; the naive approach had the best performance, followed by the ARIMA model, and the Holt-Winter's method. The naive method performed best because there is a high similarity between the test week's demand and the previous week's demand.

5.2 Comparison of the results on two weeks forecast

The table below shows a summary of the accuracy metrics gotten from the six methods based on how they performed.

	Simple Average	Moving Average	Naive Approach	Holt-Winter	AR model	ARIMA model
MAPE (%)	50.22	7.58	2.29	2.95	3.28	2.58
ME (Kw)	-1326.17	-41.06	41.06	41.24	13.46	32.54
MAE (Kw)	1326.17	186.57	59.73	79.04	88.27	67.98
MPE (%)	-50.22	2.40	1.60	1.59	0.59	1.21
RMSE (Kw)	1345.64	231.76	76.62	97.99	119.33	83.32
Min-Max Error	0.50	0.07	0.02	0.03	0.03	0.03

Table 2:- The table shows a summary of the accuracy metrics of all 6 methods based on two weeks forecast

From the table above; we can see that the Naive approach had the best performance for the second time in a row, with ARIMA and Holts-Winter method staying as the second and third best in a row respectively. The naive method still performs best because there is a high similarity between the test week's demand and the previous week's demand in UNB. As the horizon gets wider, the naive forecast may not be efficient anymore.

5.3 Comparison of the results on a month's forecast

The table below shows a summary of the accuracy metrics gotten from the six methods based on how they performed.

	Simple Average	Moving Average	Naive Approach	Holt-Winter	AR model	ARIMA model
MAPE (%)	49.89	9.97	4.38	3.97	4.24	3.12
ME (Kw)	-1302.46	-75.89	-78.66	-65.64	-58.14	-35.87
MAE (Kw)	1302.46	253.51	112.01	101.31	108.85	79.48
MPE (%)	-49.89	-1.92	-2.90	-2.41	-2.03	-1.23
RMSE (Kw)	1326.85	264.37	137.33	127.41	135.66	109.16
Min-Max Error	0.50	0.09	0.04	0.04	0.04	0.03

Table 3:- The table shows a summary of the accuracy metrics of all 6 methods based on a month's forecast

From the table above; we can see that the Naive approach wasn't the best or the second best in this case. This is because of the wider horizon of a month. The ARIMA model was the best in the case, followed by the Holt-Winter's method, the AR model, and then the Naive approach.

5.4 Comparison of the results on December, 2017's forecast

The table below shows a summary of the accuracy metrics gotten from the six methods based on how they performed.

	Simple Average	Moving Average	Naive Approach	Holt-Winter	AR model	ARIMA model
MAPE (%)	17.43	14.24	13.82	11.81	12.81	13.55
ME (Kw)	-211.95	41.09	133.26	93.69	121.94	133.83
MAE (Kw)	227.37	159.60	148.71	134.11	136.43	144.65
MPE (%)	-15.79	5.71	12.52	8.94	11.63	12.66
RMSE (Kw)	274.12	178.63	208.55	155.35	195.06	206.81
Min-Max Error	0.17	0.12	0.11	0.10	0.10	0.11

Table 4:- The table shows a summary of the accuracy metrics of all 6 methods based on a month's forecast

Due to the sudden change in demand in UNB as Christmas holidays approaches; I decided to test the models on December, 2017 prediction. The results gotten was expected because it is a pretty difficult to predict this time period accurately. The Holt-Winter model had the best performance in this case, followed by the Moving Average, AR and the ARIMA model.

6 Challenges / Observations

The first challenge of this experiment was to get enough data that would help to train the model much better, and then removing outliers in the data. Another challenge is to find an efficient model for predicting different horizons, as different models perform better with different dataset. The simple average model works better if the average across the dataset remains the same. The moving average model could be improved by adding weights to the previous values being used for prediction, which could be time consuming. The naive approach works best if the previous week was similar of close with the tested week; but if there are trends in the data and stationarity is lost, the naive approach become inefficient as the horizon gets wider. The major challenge with the ARIMA model was to find the right parameters to use in training the model, as different horizons and dataset work best with different parameters. For future work; I aim at adding other variables that affect demand like Temperature, and Dew Point. I believe this would improve the accuracy of the models.

7 Conclusions

The aim of this project was to test different models and see how well they perform on different horizons; one week, two weeks, one month, and December, 2017 forecasts. In the one and two weeks forecast; the naive approach had the best performance, this is due to the high similarity in the data on a weekly basis. Due to the non-stationarity in the data, the naive approach becomes inefficient with a wider horizon. In the one month forecast; the ARIMA model had the best performance, which is expected because ARIMA combines three methods (Auto regression, Integration, and Moving Average). In December, 2017's forecast; the Holt-Winter model had the best performance after I changed the seasonality period to 365 days.

References

[1] T. Hong and S. Fan, “Probabilistic electric load forecasting: A tutorial review,” *Int. J. Forecast.*, vol. 32, no. 3, pp. 914–938, Jul. 2016.

[2] E. Almeshaiei and H. Soltan, “A methodology for Electric Power Load Forecasting,” *Alexandria Eng. J.*, vol. 50, no. 2, pp. 137–144, 2011.

[3] G. Singh, “7 methods to perform Time Series forecasting (with Python codes),” *Analytics Vidhya*, 2018. [Online]. Available: https://www.analyticsvidhya.com/blog/2018/02/time-series-forecasting-methods/.

[4] “Autoregressive Model: Definition & The AR Process - Statistics How To.” [Online]. Available: https://www.statisticshowto.datasciencecentral.com/autoregressive-model/. [Accessed: 01-Dec-2019].

[5] “ARIMA Model - Complete Guide to Time Series Forecasting in Python | ML+.” [Online]. Available: https://www.machinelearningplus.com/time-series/arima-model-time-series-forecasting-python/. [Accessed: 09-Nov-2019].

[6] J. Brownlee, “Time Series Forecast Case Study with Python: Monthly Armed Robberies in Boston,” *Machine Learning Mastery*, 2019. [Online]. Available: https://machinelearningmastery.com/time-series-forecast-case-study-python-monthly-armed-robberies-boston/.

[7] “Analytics Vidhya.” [Online]. Available: https://courses.analyticsvidhya.com/courses/take/creating-time-series-forecast-using-python/texts/6142791-parameter-tuning-for-arima-model. [Accessed: 12-Nov-2019].

[8] “Build High Performance Time Series Models using Auto ARIMA in Python and R.” [Online]. Available: https://www.analyticsvidhya.com/blog/2018/08/auto-arima-time-series-modeling-python-r/. [Accessed: 19-Nov-2019].

[9] R. Weron, *Modeling and forecasting electricity loads and prices: A statistical approach*. wiley, 2006.

[10] H. K. Alfares and M. Nazeeruddin, "Electric load forecasting: Literature survey and classification of methods," *Int. J. Syst. Sci.*, vol. 33, no. 1, pp. 23–34, Jan. 2002.

[11] H. Hahn, S. Meyer-Nieberg, and S. Pickl, "Electric load forecasting methods: Tools for decision making," *Eur. J. Oper. Res.*, vol. 199, no. 3, pp. 902–907, Dec. 2009.

[12] W. C. Hong, "Electric load forecasting by support vector model," *Appl. Math. Model.*, vol. 33, no. 5, pp. 2444–2454, May 2009.

[13] G. Gross and F. D. Galiana, "Short-Term Load Forecasting," *Proc. IEEE*, vol. 75, no. 12, pp. 1558–1573, Dec. 1987.

[14] I. Moghram and S. Rahman, "Analysis and evaluation of five short-term load forecasting techniques," *IEEE Trans. Power Syst.*, vol. 4, no. 4, pp. 1484–1491, 1989.

[15] M. T. Hagan and S. M. Behr, "The Time Series Approach to Short Term Load Forecasting," *IEEE Trans. Power Syst.*, vol. 2, no. 3, pp. 785–791, 1987.