Names: Albert MANIRAHO
ANDREW ID: maniraha
Kigali, February 28, 2022

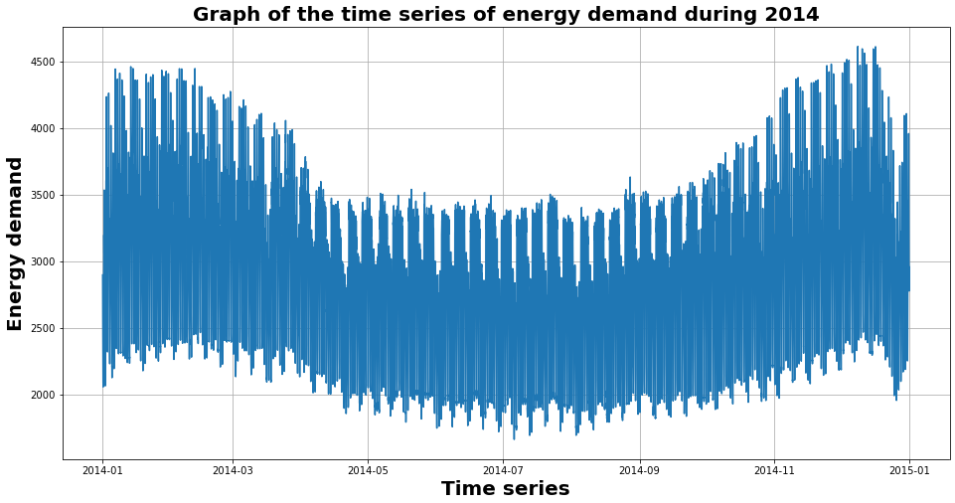
DATA ANALYTICS ASSIGNMENT 2 REPORT

All the libraries and packages used:

```
from scipy import stats
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from statsmodels.graphics.tsaplots import plot_acf
import seaborn as sns
import warnings
import statsmodels.api as sm
from statsmodels.graphics.tsaplots import plot_acf
import scipy.stats as stats
import matplotlib.dates as mdates
from sklearn.metrics import mean_absolute_error as mae
from sklearn.metrics import mean_absolute_percentage_error as mape
from scipy.stats import ttest_ind
warnings.filterwarnings('ignore')
```

QUESTION 1.

The answer to the question has resulted in the following graphic:



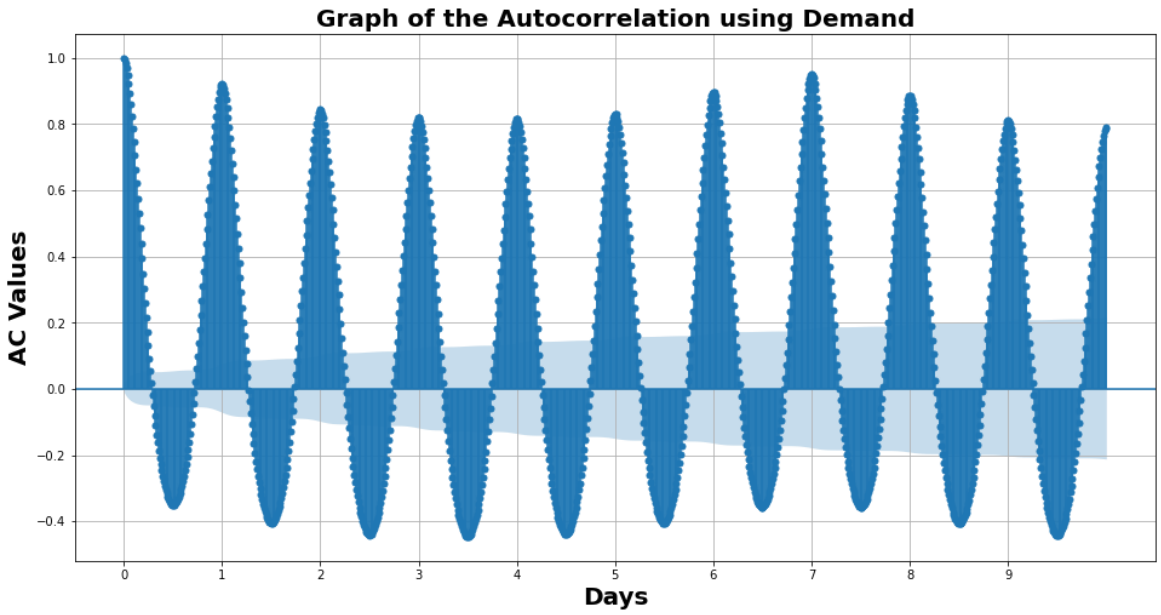
In the first three months from January to March the energy demand was high, and the demand decreased for the subsequent months from April to September and later on it increased in the last three months of the year.

QUESTION 2.

From the computations done in the codes, the following is the table of the autocorrelation coefficients for the 10 days.

	AutoCorrCoefficients
0	1.000000
1	0.996103
2	0.985689
3	0.969448
4	0.948211
...	...
956	0.746984
957	0.764598
958	0.777886
959	0.786296
960	0.789327

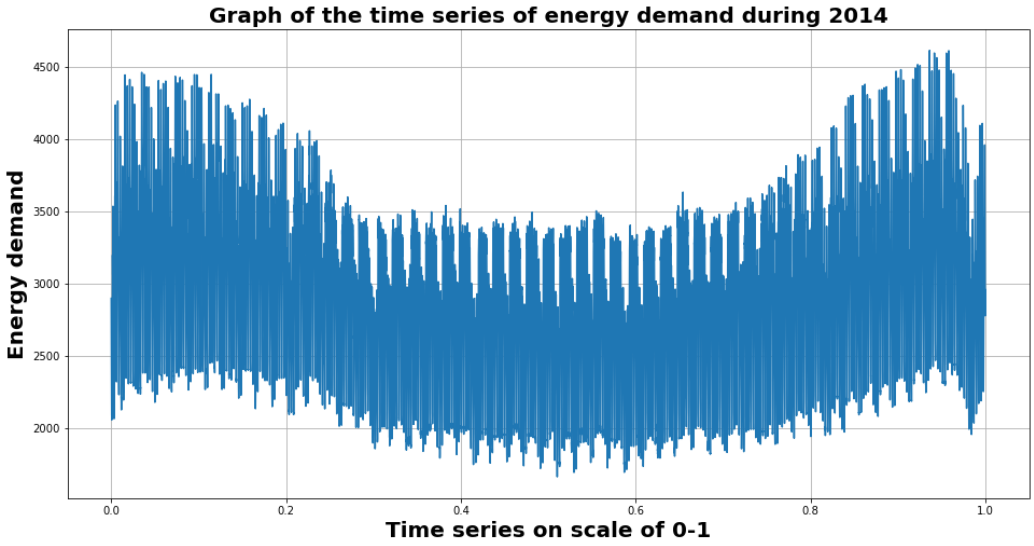
The graph of the autocorrelation using the demand.



There is evidence of seasonality in the graph as the see almost the same pattern on the graph. The most data points are significant as they lie outside the blue area.

QUESTION 3.

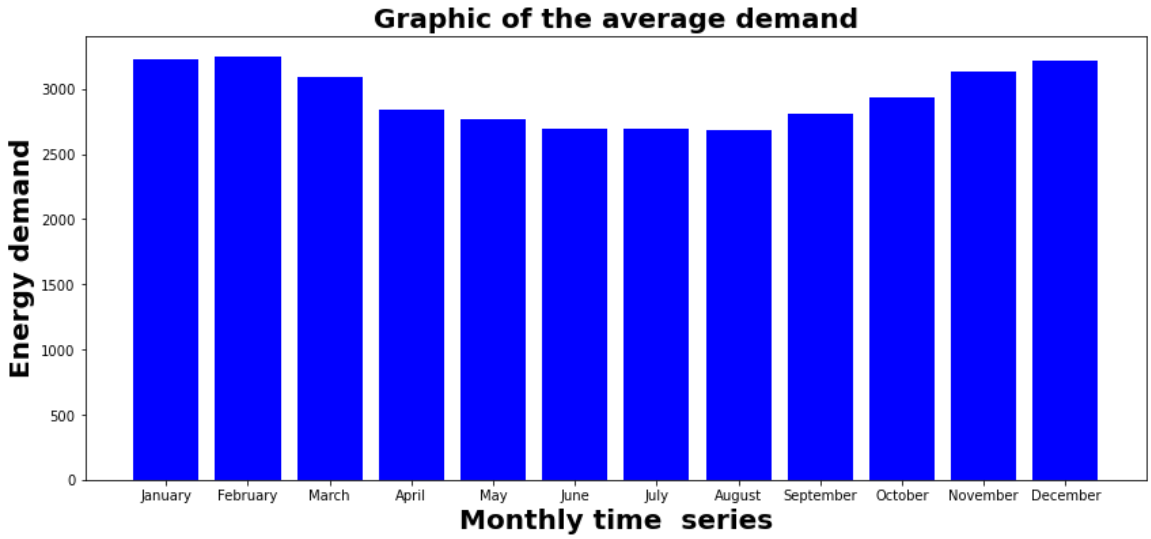
The answer to the question has resulted in the following graphic for the energy demand on a scale of 0-1.



In the first three months from January to March the energy demand was high, and the demand decreased for the subsequent months from April to September and later on it increased in the last three months of the year.

QUESTION 4.

The graphic of the average energy demand in a year/ on a monthly basis.

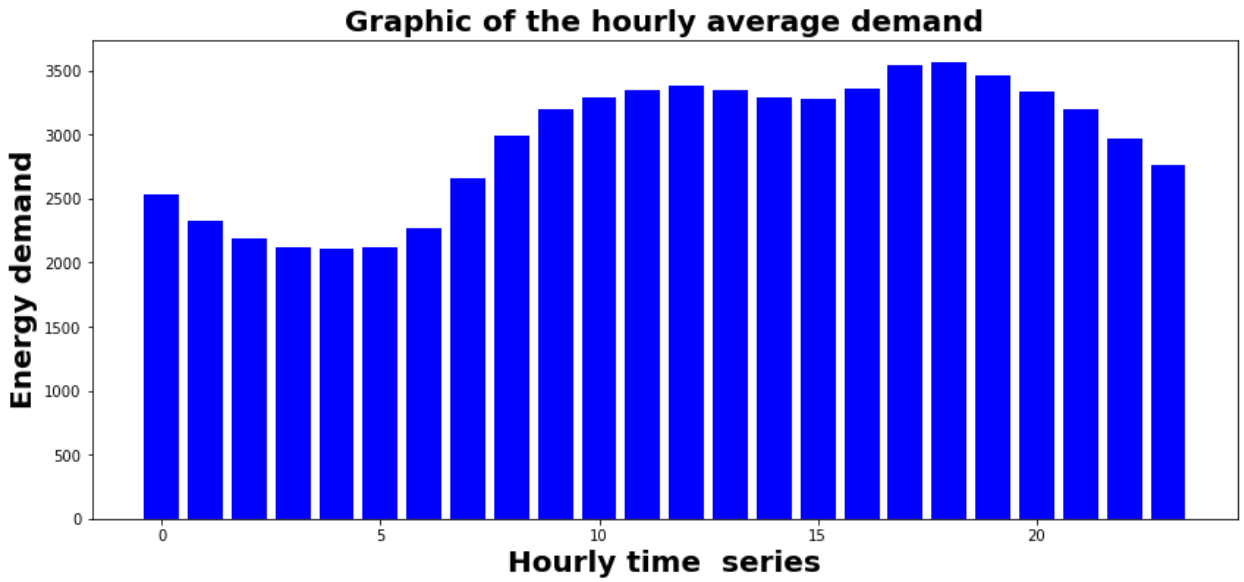


The energy demand in the is higher in the first three months of the year (January-March). After then, the demand slightly decreased (April-August). The demand also increases in the three last months of the year (October-December).

February has shown the highest demand while August has shown the lowest demand.

QUESTION 5.

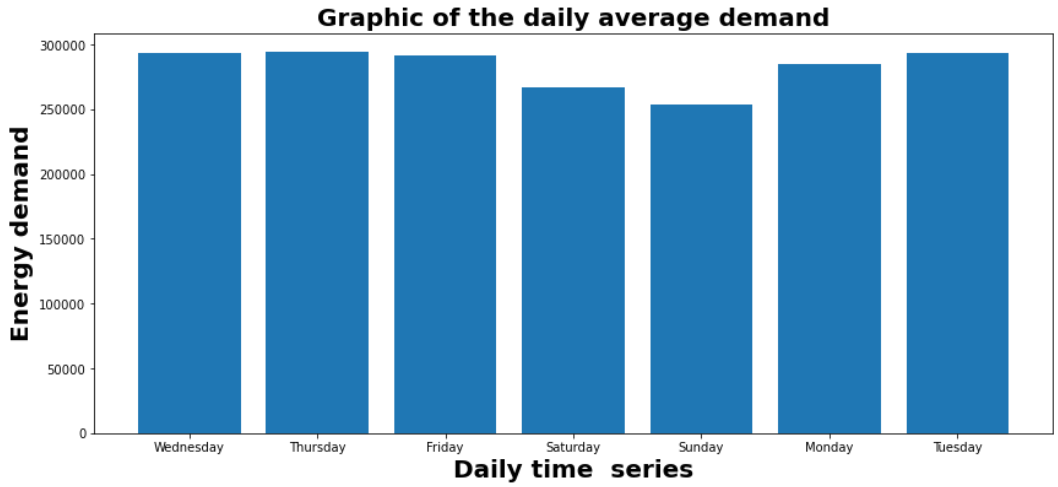
The average demand of the electricity for the 24 hours of the day.



From the graph it is clear that the energy consumption between (00:00pm-05:00am) is slightly low, it is because that the people are asleep and industries activities are operational but not on the full load, so the energy demand is low. After 5:00 am the people are awake and the daily activities are on now, so the demand increases till 20:00pm where the people again go to rest after their activities. So, after this time the energy demand also decreases.

QUESTION 6.

The graphic to show the energy demand across the 7 days of the week.

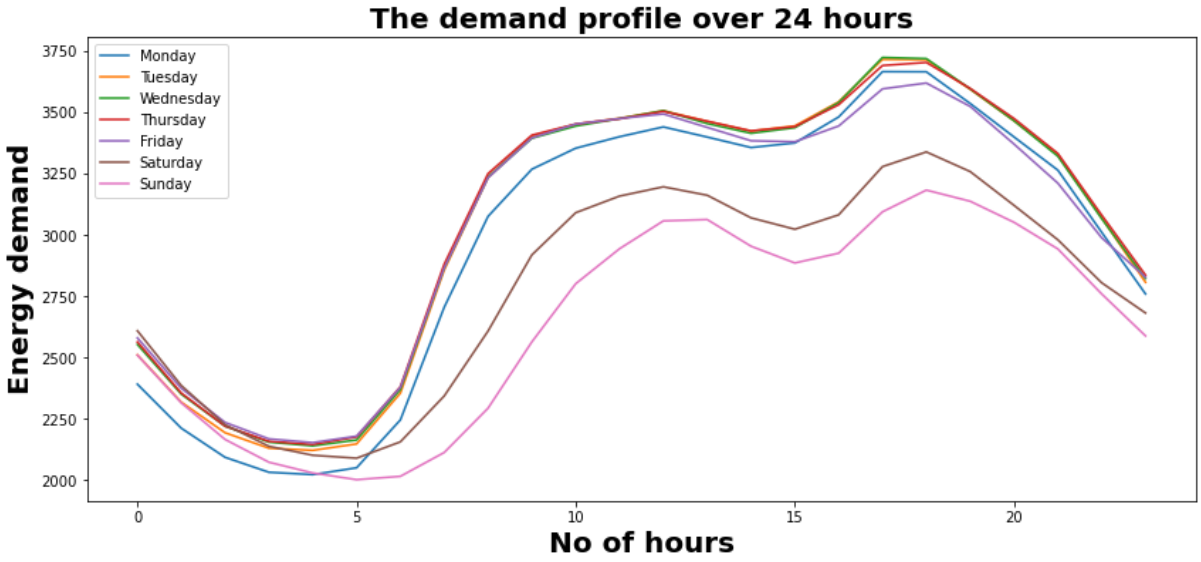


The people are consuming higher in the five working days of the week; therefore, the energy demand/consumption is higher compared to the weekends where the workers are not working. The weekends some activities are shut down like offices and schools, so the energy demand is lower.

QUESTION 7.

The table of the daily demand profile for each day of the week.

dateTimes	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
0	2391.326923	2510.788462	2553.466981	2563.740385	2580.211538	2609.100962	2509.822115
1	2212.317308	2318.504808	2351.047170	2355.471154	2375.052885	2386.254808	2315.302885
2	2093.615385	2193.745192	2220.330189	2220.687500	2235.759615	2226.533654	2166.370192
3	2032.706731	2130.639423	2155.500000	2158.735577	2168.706731	2138.514423	2073.586538
4	2023.168269	2121.923077	2140.292453	2146.048077	2153.413462	2101.947115	2029.841346
5	2050.812500	2148.014423	2164.051887	2176.403846	2179.826923	2089.961538	2002.461538
6	2245.923077	2354.802885	2369.641509	2380.985577	2378.764423	2156.581731	2016.076923
7	2704.875000	2856.302885	2865.608491	2880.221154	2865.163462	2343.307692	2113.182692
8	3075.024038	3233.456731	3235.207547	3249.216346	3232.235577	2607.961538	2293.625000
9	3267.221154	3402.254808	3393.306604	3405.942308	3394.730769	2918.153846	2563.826923
10	3352.697115	3450.586538	3442.990566	3450.956731	3451.197115	3090.004808	2800.668269
11	3399.557692	3474.625000	3473.419811	3472.259615	3473.461538	3157.634615	2943.177885
12	3439.447115	3505.432692	3505.731132	3502.408654	3492.187500	3195.475962	3056.697115
13	3398.163462	3462.009615	3453.150943	3462.730769	3437.528846	3161.110577	3062.100962
14	3355.485577	3422.346154	3413.528302	3423.360577	3383.144231	3069.048077	2953.471154
15	3374.177885	3443.663462	3436.306604	3440.711538	3379.019231	3022.629808	2885.235577
16	3480.057692	3541.889423	3538.556604	3529.865385	3443.206731	3081.336538	2925.028846
17	3664.567308	3714.471154	3722.400943	3689.634615	3594.423077	3277.788462	3093.701923
18	3664.000000	3713.524038	3717.707547	3702.067308	3618.225962	3337.591346	3182.086538
19	3534.865385	3592.793269	3594.566038	3596.245192	3522.908654	3257.543269	3136.673077
20	3398.610577	3464.500000	3464.594340	3473.682692	3369.528846	3119.942308	3050.899038
21	3263.317308	3325.365385	3319.976415	3331.451923	3210.317308	2979.677885	2942.423077
22	3010.355769	3069.552885	3066.033019	3080.225962	2987.932692	2804.649038	2759.177885
23	2759.081731	2806.519231	2822.051887	2835.759615	2827.096154	2681.466346	2588.163462



From the graph, using the hours of the day:

It is seen that the energy consumption between (00:00pm-05:00am) is slightly low, it is because that the people are asleep and industries activities are operational but not on the full load, so the energy demand is low. After 5:00 am the people are awake and the daily activities are on now, so the demand increases till 20:00pm where the people again go to rest after their activities. So, after this time the energy demand also decreases.

From the graph, using days of the week it is seen that:

The people are consuming higher in the five working days of the week; therefore, the energy demand/consumption is higher compared to the weekends where the workers are not working. The weekends some activities are shut down like offices and schools, so the energy demand is lower.

QUESTION 8.

Null hypothesis: There is **no** statistically significant difference between demand during the weekend (Saturday and Sunday) and during the working week (Monday through Friday)

Alternative hypothesis: There a statistically significant difference between demand during the weekend (Saturday and Sunday) and during the working week (Monday through Friday)

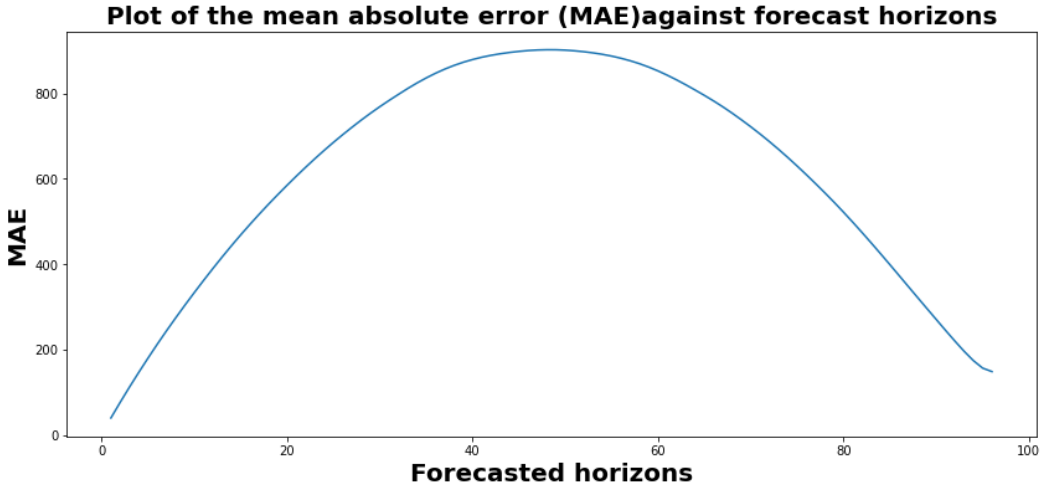
Since the p-value is less than 0.05, the null hypothesis is rejected then the alternative hypothesis is kept so There a statistically significant difference between demand during the weekend (Saturday and Sunday) and during the working week (Monday through Friday)

```
The p-value is : 0.0

The t-statistics is: 51.363399969318884
```

QUESTION 9.

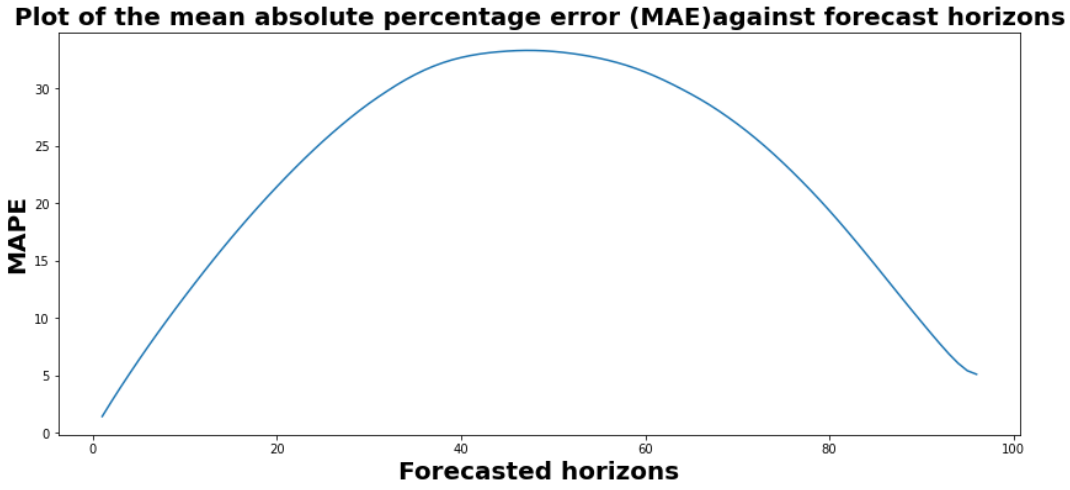
The results of the codes give the following graphic:



From the graph, it is seen that when the value of k increases (forecasted horizons) the mean absolute error increases (MAE) up to the horizon and MAE starts to reduce as the value of k (forecasted horizons) increases.

QUESTION 10.

The results of the codes give the following graphic:



From the graph, it is seen that when the value of k increases (forecasted horizons) the mean absolute error increases (MAPE) up to the horizon and MAPE starts to reduce as the value of k (forecasted horizons) increases.