Student's Name: Aryan Garg

Mobile No: 8219383122

Roll Number: B19153

Branch:EE

1 a.
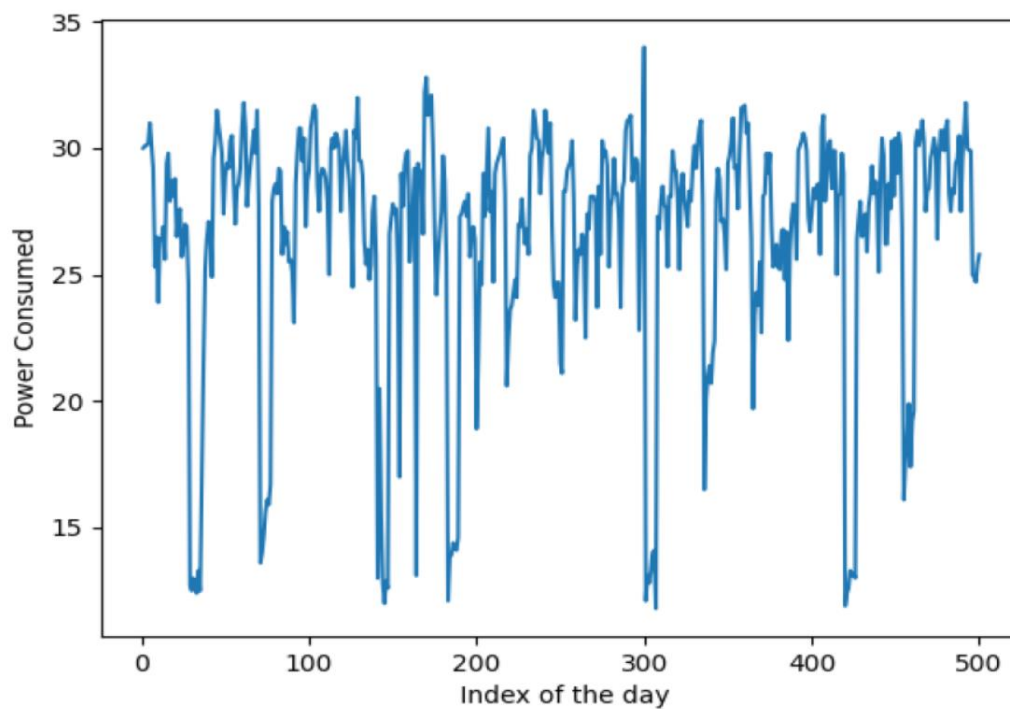
Figure 1 Power consumed (in MW) vs. days

Inferences:

1. The days don't have similar power consumption
2. Because not every day is same for any person, and so is the data.

b. The value of the Pearson's correlation coefficient is 0.768

Inferences:

1. Strong correlation exists between the two-time sequences
2. The relation between the sequences is nearly linear and increasing.
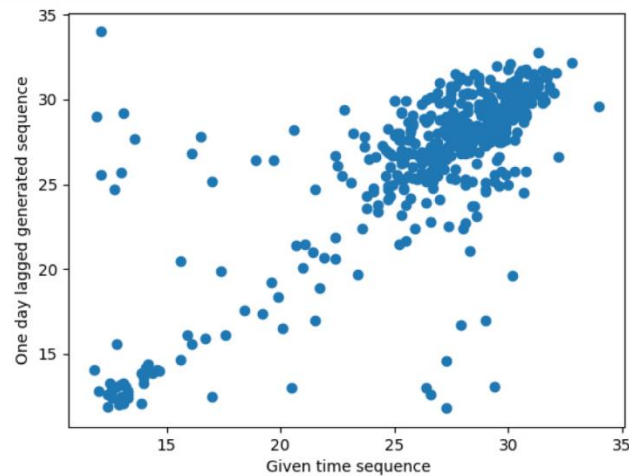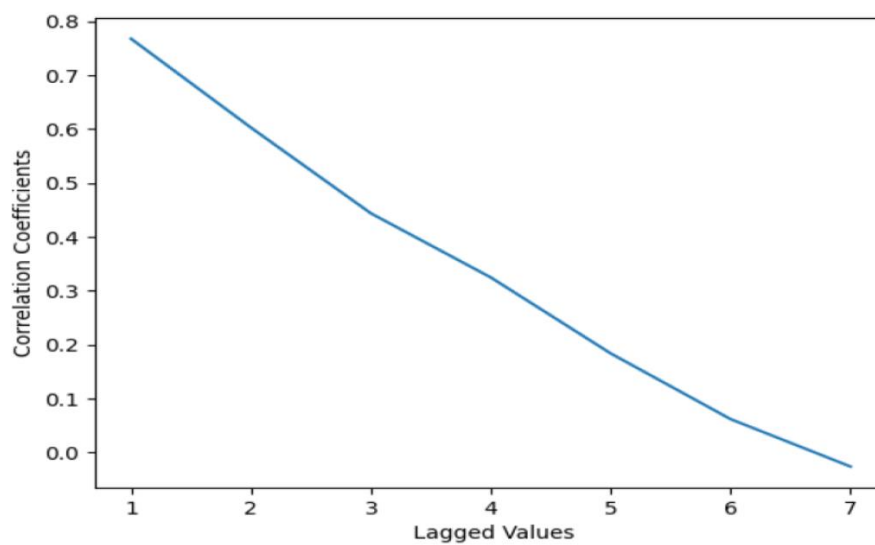3. Due to high correlation point 2 was observed.

c.

Figure 2 Scatter plot one day lagged sequence vs. given time sequence

Inferences:

1. High correlation is observed
2. The scatter plot and the correlation coefficient are in good agreement.
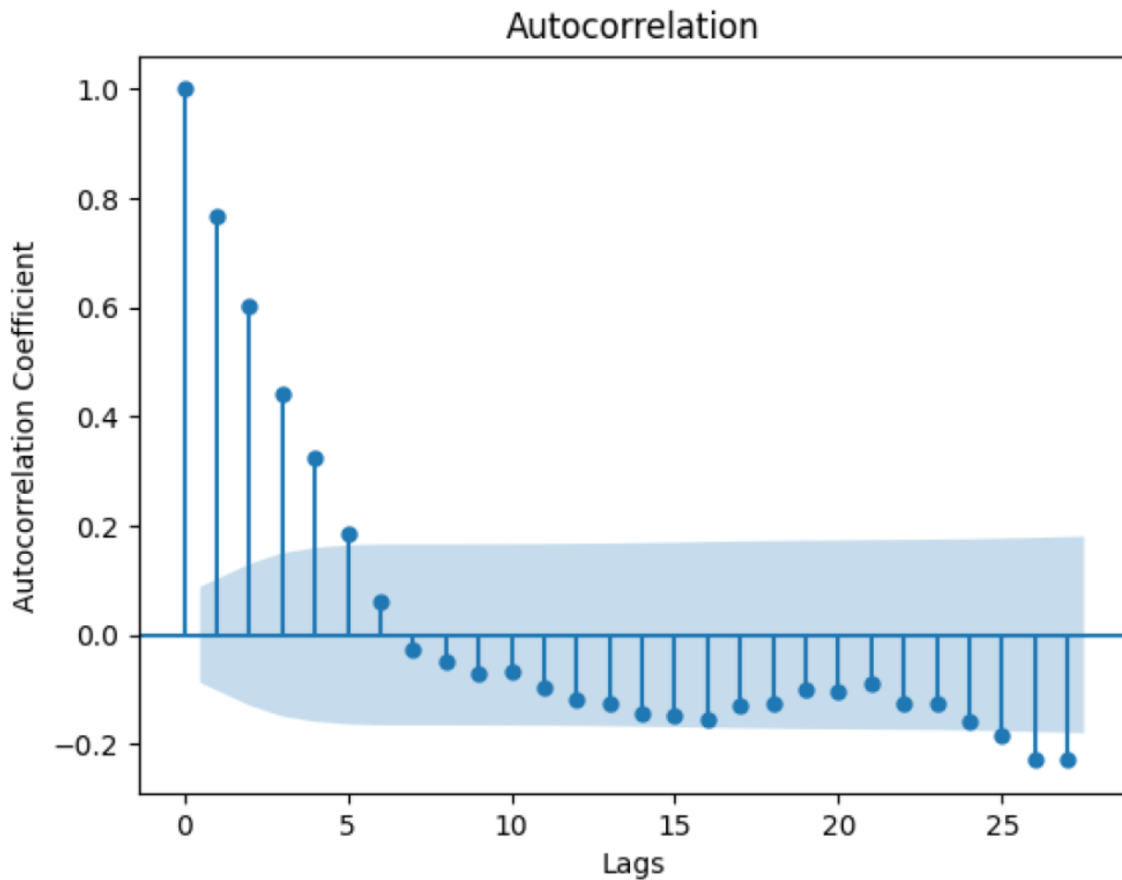
d.

Inferences:

1. Correlation coefficient decreases with higher lag values.
2. As we increase lag, data error reduces, hence the correlation decreases as the data stops having a linear/increasing relation.

e.

Figure 4 Correlation coefficient vs. lags in given sequence generated using 'plot_acf' function
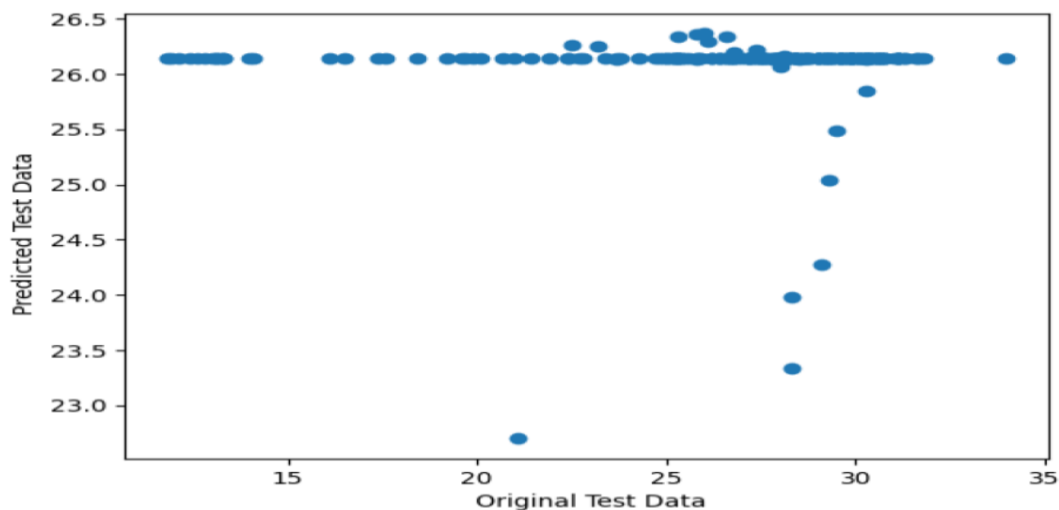
Inferences:

1. The correlation coefficient value decreases w.r.t lags in time sequence till lag ~= 15 and then increases for a while before decreasing again.
2. As we decrease lag, the data starts becoming less related to each other and hence the correlation co-efficient decreases

2. The RMSE between predicted power consumed for test data and original values for test data is 3.192.

Inference:

1. Low error suggests high accuracy.

3 a.

Figure 5 Predicted test data time sequence vs. original test data sequence

The RMSE between predicted power consumed for test data and original values for test data is 4.515.

Inferences:

1. Decent accuracy observed.
2. Point 1 was observed due to low error value.
3. Clearly, the model is reliable.
4. This model is less reliable than the model above

b.

Table 1 RMSE between predicted and original data values wrt lags in time sequence

Lag value	RMSE
1	4.539
5	4.538
10	4.532
15	4.560
25	4.515

Inferences:

1. RMSE decreases till 15 lag and then increases for a while till decreasing again.
2. In general, the accuracy increases with increase in lag hence the RMSE decreases.

c. The heuristic value for optimal number of lags 0.126.

The RMSE value between test data time sequence and original test data sequence is 4.538

Inferences:
1. Heuristics didn't decrease (improve) the error.
2. RMSE by heuristics is more than the one we get by using brute force.

d.

The optimal number of lags without using heuristics for calculating optimal lag is 25

The optimal number of lags using heuristics for calculating optimal lag is 6

Inferences:
1. Prediction accuracies obtained without heuristics for calculating optimal lag is higher than the one when we calculate the heuristic.
2. RMSE value is less when we don't calculate the heuristic value.