

IC 272: DATA SCIENCE - III  
LAB ASSIGNMENT – VI  
Auto-regression

---

Student's Name:

Mobile No:

Roll Number:

Branch:

---

1 a.

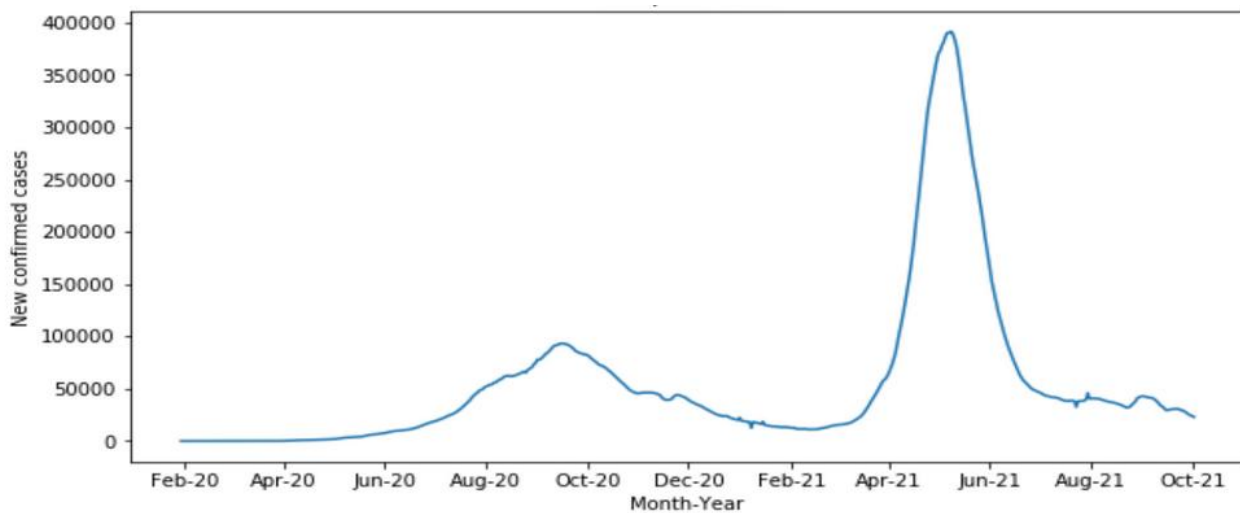


Figure 1 No. of COVID-19 cases vs. days

**Inferences:**

1. Infer from the plot whether the days one after the other have similar power consumption?
2. State the reason behind inference 1.
3. Infer from the plot the duration of first and second waves.

Note: The plot above is for illustration purposes. Replace it with the plot obtained by you. Suitably rename x-axis and y-axis legends.

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

**Inferences:**

1. From the value of Pearson's correlation coefficient, what do you infer about the degree of correlation between the two time sequences?

IC 272: DATA SCIENCE - III  
LAB ASSIGNMENT – VI  
Auto-regression

2. We generally expect observations (here number of COVID-19 cases) on days one after the other to be similar. To what extent does it hold? Answer with respect to the value of Pearson's correlation coefficient.
3. State the reason behind Inferences 1 and 2.

c.

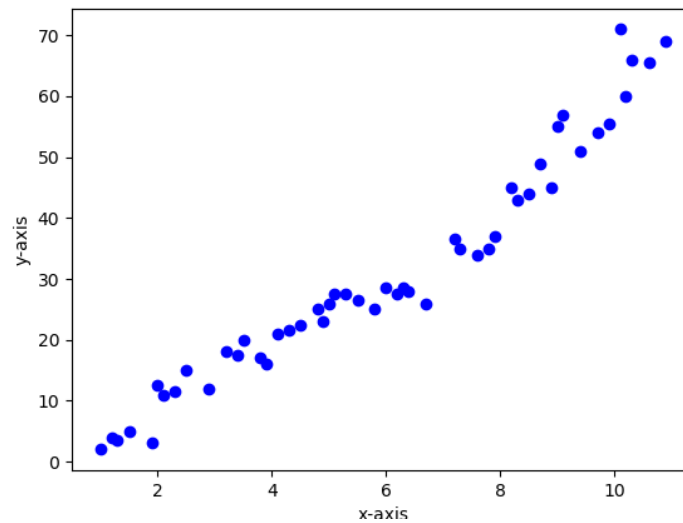


Figure 2 Scatter plot one day lagged sequence vs. given time sequence

**Inferences:**

1. From the nature of the spread of data points, what do you infer about the nature of correlation between the two sequences?
2. Does the scatter plot seem to obey the nature reflected by Pearson's correlation coefficient calculated in 1.b?
3. State reason behind inference 2.

Note: The scatter plot above is for illustration purposes. Replace it with the scatter plot obtained by you. Suitably rename x-axis and y-axis legends.

IC 272: DATA SCIENCE - III  
LAB ASSIGNMENT – VI  
Auto-regression

---

d.

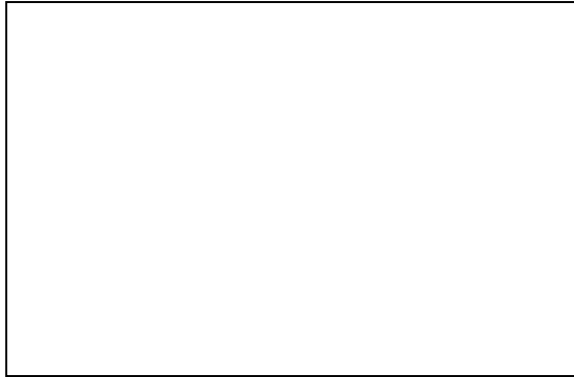


Figure 3 Correlation coefficient vs. lags in given sequence

**Inferences:**

1. Infer the trend of correlation coefficient value with respect to increase in lags in time sequence.
2. Explain the reason behind the observed trend.

e.



Figure 4 Correlation coefficient vs. lags in given sequence generated using 'plot\_acf' function

**Inferences:**

1. Infer the trend of correlation coefficient value with respect to lags in time sequence.
2. Explain the reason behind the observed trend.

IC 272: DATA SCIENCE - III  
LAB ASSIGNMENT – VI  
Auto-regression

2

a. The coefficients obtained from the AR model are;

b. i.

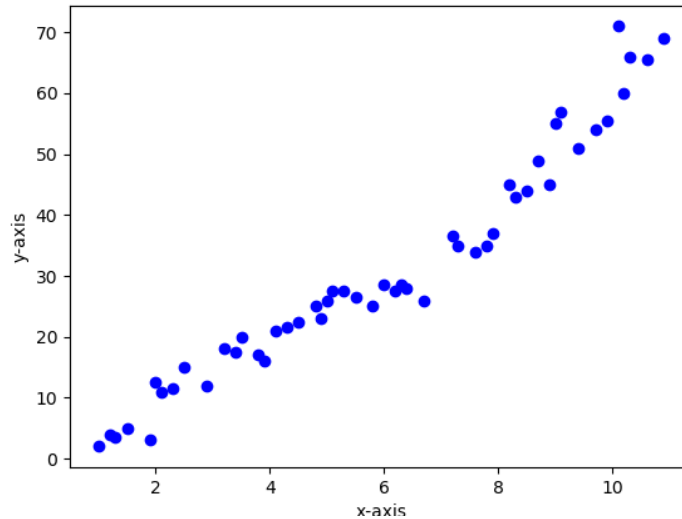


Figure 5 Scatter plot actual vs. predicted values

**Inferences:**

1. From the nature of the spread of data points, what do you infer about the nature of the correlation between the two sequences?
2. Does the scatter plot seem to obey the nature reflected by Pearson's correlation coefficient calculated in 1.b?
3. State reason behind inference 2.

Note: The scatter plot above is for illustration purposes. Replace it with the scatter plot obtained by you. Suitably rename x-axis and y-axis legends.

ii.



Figure 6 Predicted test data time sequence vs. original test data sequence

IC 272: DATA SCIENCE - III  
LAB ASSIGNMENT – VI  
Auto-regression

**Inferences:**

1. From the plot of predicted test data time sequence vs. original test data sequence comment how reliable is the model for future predictions with suitable reasons.

iii.

The RMSE(\%) and MAPE between predicted power consumed for test data and original values for test data are .

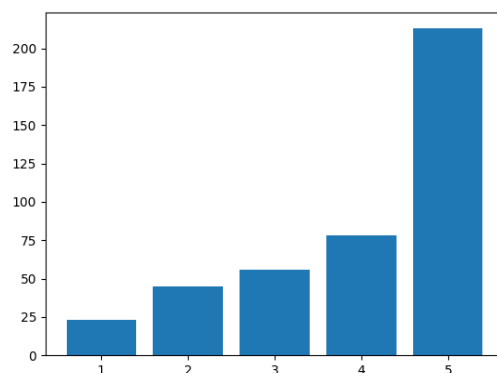
**Inferences:**

1. From the value of RMSE(\%) and MAPE value comment how accurate is the model for the given time series is.
2. State the reason behind Inference 1.

3

**Table 1 RMSE (%) and MAPE between predicted and original data values wrt lags in time sequence**

Lag value	RMSE (%)	MAPE



**Figure 7 RMSE(%) vs. time lag**

## IC 272: DATA SCIENCE - III

### LAB ASSIGNMENT – VI

#### Auto-regression

---

#### Inferences:

1. Infer the trend of RMSE(%) with respect to increase in lags in time sequence.
  2. State the reason behind Inference 1.
- Note: The bar chart above is for illustration purposes. Replace it with the bar chart obtained by you.  
Suitably rename x-axis and y-axis legends.

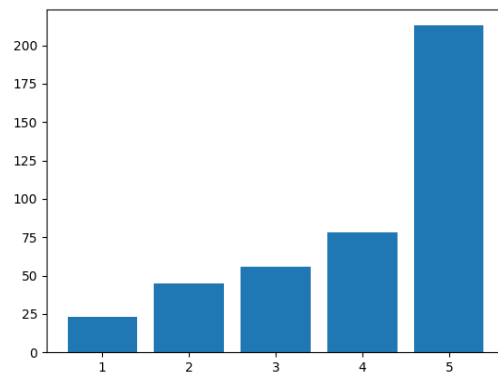


Figure 8 MAPE vs. time lag

#### Inferences:

1. Infer the trend of MAPE with respect to increase in lags in time sequence.
  2. State the reason behind Inference
- Note: The bar chart above is for illustration purposes. Replace it with the bar chart obtained by you.  
Suitably rename x-axis and y-axis legends.

#### 4

The heuristic value for the optimal number of lags is

The RMSE(%) and MAPE value between test data time sequence and original test data sequence are

#### Inferences:

1. Based upon the RMSE(%) and MAPE value, did heuristics for calculating the optimal number of lags improve the prediction accuracy of the model?
2. State the reason behind Inference 1



IC 272: DATA SCIENCE - III  
LAB ASSIGNMENT – VI  
Auto-regression

---

3. Compare the prediction accuracies obtained without and with the heuristic for calculating optimal lag with respect to RMSE(%) and MAPE values.

**Guidelines for Report (Delete this while you submit the report):**

- The plot/graph/figure/table should be centre justified with sequence number and caption.
- Inferences should be written as a numbered list.
- Use specific and technical terms to write inferences.
- Values observed/calculated should be rounded off to three decimal places.
- The quantities which have units should be written with units.