

# TSF Internal Graded Assessment

**Answer the following questions for the following datasets:**

- 1. Rainfall Dataset**
- 2. Closing Stock Price dataset**
- 3. Tractor Sales dataset**
- 4. Store\_50 dataset**

## SECTION A

### 1. Data preparation (5 marks)

- a. Read the dataset (tab, csv, xls, txt, inbuilt dataset). What are the number of rows and no. of cols & types of variables? (1 MARK)
- b. Convert the data into time series (2 MARK)
- c. Check for defects in the data such as missing values, null, etc. (1 MARK)
- d. Visualize the time series using relevant plots. (1 MARK)

## SECTION B

### 2. Data understanding (15 marks)

- a. Decompose the time series and check for time series components. (4 marks)
- b. Perform dicky fuller test to check the stationarity? What other actions will you take if series is non-stationary? (3+1 marks)
- c. Plot Auto Correlation and Partial Auto Correlation function? What is your inference from these plots? (2+2 marks)
- d. Split dataset into train and test sets. Use last one-month data for testing. (3 marks)

## SECTION C

### 3. Model Building (20 marks)

- a. Fit a base model and observe the residuals, RMSE and MAPE values of the model. Please comment on whether it is good or not. (5 marks)
- c. How would you improve the model? What changes you will make in the base model. Fit the final model. (10 marks)
- d. Analyze the residuals of final model. Feel free to use charts or graphs to explain. (2 marks)
- e. Forecast the births for next 1 months using the final model? (3 marks)