	PES University, Bengaluru (Established under Karnataka Act No. 16 of 2013)	**UE20CS934**
	Oct 2024 : END SEMESTER ASSESSMENT (ESA) **M TECH DATA SCIENCE AND MACHINE LEARNING_ SEMESTER II** **UE20CS934 – Time Series Forecasting**	
Time: 3 Hrs	Answer All Questions	Max Marks: 100

INSTRUCTIONS
- All questions are compulsory.   - Section A should be handwritten in the answer script provided   - Section B and C are coding questions which have to be answered in the system.
Section A (20 marks)

1	a)	Explain the concept of seasonality in time series analysis. How can seasonality be detected and accounted for in a time series model? Provide examples where applicable	5
	b)	What are the key differences between the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) in time series analysis? Explain their respective roles in identifying patterns within a time series?	7
	c)	What is the difference between additive and multiplicative time series models? In what situations would you choose one over the other? Provide examples to illustrate your answer.	8

Section B (40 marks)	
2	Your task is to analyse the dataset **Title**: Retail Sales Data for a Store(retail_sales.csv)	

- **Description**: A time series dataset representing the daily total sales for a retail store in New York, USA, during the year 1992 to 2016. The dataset includes information on the date and total sales made each day. There are 293 records in total monthly data.
- **Columns**:
    - **ds(date)**: The date of the sales record in "yyyy-dd-mm" format.
    - **y(sales)**: The total sales amount in USD for the corresponding date (numeric).

Forecast sales for the next 30 days using the final model?

a)	Read the dataset (tab, csv, xls, txt, inbuilt dataset). What are the number of rows and no. of cols & types of variables? (2 mark) Convert the data into time series (3 marks) Check for defects in the data such as missing values, null, etc. (2 mark) Visualize the time series using relevant plots. (3 mark)	10
b)	Decompose the time series and check for components of time series. (4 marks) Perform dicky fuller test to check the stationarity? What other actions will you take if series is non-stationary? (3+2 marks) Plot Auto Correlation and Partial Auto Correlation function? What is your inference from these plots? (3+3 marks)	15
c)	Split dataset into 80-20train and test sets.. (5 marks) Fit ARIMA model and observe the RMSE and MAPE values of the model for test data.(10 marks)	15

Section C (40)

3	a)	Fit exponential smoothing model and observe the residuals, RMSE and MAPE values of the model for test data. (15 marks) marks	15
	b)	How would you improve the exponential smoothing model? Make the changes and fit the final exponential smoothing model. (10 marks) Analyze the residuals of this final model. Feel free to use charts or graphs to explain. (5 marks)	15
	c)	Forecast sales for the next month using the final model?	10