

# Examination Question Paper

PG-DBDA September 2021

Module Name: Advanced Analytics Using  
Statistics

Exam Type: Main

DATE: 03-03-2022

Duration: 2hrs

Max. Marks: 40

## Instructions for Uploading:

- All questions are compulsory
- Make sure to attempt all sub questions
- Your analysis should have justifications or explanation (in English)
- Before creating Zip file of your .ipynb files make sure that you have closed all .ipynb files
- Zip File should be renamed full PRN followed by exam name. (For example: 210940125001-AA.zip)
- Uploading is allowed only once

Q1. Take following hotel booking data.(20 marks)

- Find the outliers in 2 features using IQR method(7 marks)
- Select features for prediction of Average daily rate using any 2 methods (LASSO, RFE and PCA)( 5 marks for each method max 10 marks)
- Apply Linear Regression for prediction of average daily rate. (8 marks)

[Note: Keep data cleaning to minimum.]

Link : hotel\_bookings\_2015.zip

Q2. Given Time Series Data.

Plot after Decomposing the time series (2 marks)

Plot time series after resampling on day, week, & month (3 marks)

Apply ARIMA / Linear Regression approach for forecasting. Work on forecasting at least 2 month data. (10 marks)

Calculate RMSE, & MAE values. (5 marks)

[Note: Keep data cleaning to minimum. Forecast can be fine-tuned as per time availability.]

Data Link : power\_forecast\_302\_2010.zip

\*\*\*\*\*All the best\*\*\*\*\*