****

**JAIPURIA INSTITUTE OF MANAGEMENT, NOIDA**

**PGDM / PGDM (M) / PGDM (SM)**

**FOURTH TRIMESTER (Batch 2022-24)**

**END TERM EXAMINATION, OCTOBER 2023**

|  |  |  |  |
| --- | --- | --- | --- |
| Course Name | Python for Business Analytics (SET – 2) | Course Code | **20822** |
| Max. Time | **2 Hours** | Max. Marks | **40 MM** |

**INSTRUCTIONS:**

1. Attempt all the questions on a single Jupyter Notebook
2. The data for the case is available on Moodle.
3. Write down your Roll no., course name and course code on top of Jupyter Notebook
4. Save your Jupyter notebook with .ipynb extension and as pdf file
5. Upload both the files on Moodle.
6. Label the files as PBA\_roll no (for example: PBA\_23)
7. This is an open book exam. Students may refer to the codes.

**Read the case below and answer the questions given by analyzing the data using Python**.

**Case: CarVar.com**

CarVar.com is a company that deals in buying and selling of old/ used cars. Suppose the company has hired you as an analytics consultant and your manager gave you data having “Price” of used cars and other related information about the car. You are asked to “identify” and “quantify” the factors responsible for estimating the “Price” in a multivariate fashion. Your manager has no knowledge of running a multivariate regression. Help your manager in analyzing the data and extract meaningful insights from it.

The data is listed in the file used\_cars.csv. For each used car, the price and 16 features related to it are given.

**In the capacity of Data Analyst, you have to develop a managerial report. Analyze the data and perform the following tasks:**

1. Perform data preprocessing and cleaning. Also explain the importance of data cleaning. (10 marks)
2. Perform Exploratory Data Analysis (EDA). Discuss and interpret the results you obtained in EDA. (15 marks)
3. Apply multiple regression analysis. Determine the significant predictors of salary offered.

 (10 marks)

1. Validate the regression results. (05 marks)

**Note:** Interpretation of all the outputs should be written by putting comments (or using markdown option) on the Jupyter notebook.