****

**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 – 1) | 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: Predicting MBA Salary**

Jai Institute of Management, Nagpur (JIMN) is one of the leading private B-schools based in Maharashtra, India. The school was established in the year 1995 and offer courses on Post Graduate Diploma in Management (PGDM). It offers various specializations during the second year of the programme - Marketing, Finance, Operations, HR and Business Analytics. The Institute offers 100 percent placement with an average package of around 9 lakhs per annum. The career development cell (CDC) of the institute has set a target of increasing an average package to 11 lakhs per annum. To this CDC organizes corporate sessions, interaction with industry experts under the integrated development plan (IDP). The cell also facilitates training of the students on various areas like quantitative, verbal and aptitude test. CDC has collected the data of all the placed students of the previous year batch 2021-23 for finding out the pattern and trends in the data. They are also interested in finding out the significant predictors of salary offered to students. In other words, they want to check the impact of their initiatives, like organizing IDP sessions, conducting verbal and aptitude test etc. on the salary offered to students. The data is given in the file mba\_placement.csv.

The Institute has hired you as a data scientist to explore and analyze the placement data and find useful insights from it. The data dictionary is given below:

|  |  |
| --- | --- |
| **Variables** | **Description** |
| Roll\_no | Roll number of a Student |
| SIP\_Score | Summer Internship Score out of 200 |
| Communication\_Skills | Communication skills rating during the IDP |
| Group\_Discussion\_Rating | Group discussion rating rating during the IDP |
| Personal\_Interview\_Rating | Personal interview rating rating during the IDP |
| Quantitative\_Skills\_Score | Score out of 100 on Quantitative skills |
| Verbal\_Skills\_Score | Score out of 100 on Verbal skills |
| Aptitude\_Test\_Score | Score out of 100 on Aptitude test |
| PGDM\_CGPA | CGPA in PGDM |
| Gender | Gender of a student |
| PGDM\_Stream | Total 8 dual specializations |
| SP1: MKT\_HR |
| SP2: MKT\_FIN |
| SP3: MKT\_OPS |
| SP4: MKT\_BA |
| SP5: FIN\_OPS |
| SP6: FIN\_BA |
| SP7: OPS\_HR |
| SP8: OPS\_BA |
| Work\_Experience | Whether student has a Work experience or not |
| Salary\_Offered | Salary Packaged offered to student |

**In the capacity of Data Scientist, 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.