Reg. No. \_\_\_\_\_\_\_\_\_\_\_\_\_



**End Semester Examination – Nov / Dec – 2019**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| **Code :** | **17CS2067** | **Duration :** | **3hrs** |
| **Sub. Name :** | **INTRODUCTION TO DATA ANALYTICS** | **Max. Marks :** | **100** |

**ANSWER ALL QUESTIONS (5 x 20 = 100 Marks)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Q. No.** | **Sub Div.** | **Questions** | **Course**  **Outcome** | **Marks** |
| 1. | a. | Describe in detail three options for building an in-house data science team. | CO1 | 8 |
| b. | Differentiate between data analysis and data engineering. | CO1 | 6 |
| c. | Describe on applying data science to our subject area. | CO1 | 6 |
| **(OR)** | | | | |
| 2. | a. | Explain four V’s of big data. | CO2 | 8 |
| b. | Illustrate MapReduce Programming architecture. | CO2 | 12 |
|  |  |  |  |  |
| 3. | a. | Identify the types of data analytics. | CO3 | 8 |
| b. | Describe simple linear regression models with example. | CO3 | 12 |
| **(OR)** | | | | |
| 4. | a. | Describe data preprocessing of raw data to actionable insights. | CO3 | 10 |
| b. | Explain the skillsets that are useful in business centric data science. | CO3 | 10 |
|  |  |  |  |  |
| 5. |  | Demonstrate data classification technique using average nearest neighbor and k-nearest neighbour algorithms. | CO3 | 20 |
| **(OR)** | | | | |
| 6. |  | Explain different types of statistical plots used in data visualizations for an audience of non-analytical people. | CO4 | 20 |
|  |  |  |  |  |
| 7. |  | Describe most commonly used python libraries in data science. | CO4 | 20 |
| **(OR)** | | | | |
| 8. | a. | Demonstrate visualizing, mapping, and graphing in R. | CO5 | 10 |
| b. | Ouline the various types of joins used in SQL. | CO5 | 10 |
|  | | **Compulsory**: |  |  |
| 9. |  | Illustrate how data science is used to describe and predict criminal activity. | CO6 | 20 |