| **INSTRUCTIONS:** | |
| --- | --- |
| 1. | Figures to the right indicate maximum marks for that question. |
| 2. | The symbols used carry their usual meanings. |
| 3. | Assume suitable data, if required & mention them clearly. |
| 4. | Draw neat sketches wherever necessary. |

| Examination | : Sessional I | Seat No. | : |
| --- | --- | --- | --- |
| Date | : 11/03/2022 | Day | : Friday |
| Time | : 9:00 a.m. to 10:15 a.m. | Max. Marks | : 36 |

| Q.1 | Do as directed: |  |
| --- | --- | --- |
| (a) | Fill in the blanks:  (i) Data science involves multiple disciplines: \_\_\_\_ , \_\_\_\_\_ and \_\_\_\_.  (ii) The full form of data format CSV is \_\_\_\_\_\_\_\_\_\_ and the full form of data format JSON is \_\_\_\_\_\_. | [02] |
| (b) | State true or false.  (i) Big data analytics has nothing to do with security and privacy.  (ii) Big data analytics has nothing to do with data modeling. | [02] |
| (c) | Describe Brewer’s CAP theorem. | [02] |
| (d) | Match the following:   | Column A | Column B | | --- | --- | | What happened? | Predictive Analytics | | Why did it happen? | Descriptive Analytics | | What will happen? | Prescriptive Analytics | | How can we make it happen? | Diagnostic Analytics | | [02] |
| (e) | Find mean, median and mode for below dataset:  5, 9 , 12, 12, 15, 22, 22, 22, 27 | [02] |
| (f) | What is ‘Tukey five number summaries’? How to find the same in R language? | [02] |
|  |  |  |
| Q.2 | Attempt the following: |  |
| (a) | What is an outlier? How to identify a few outliers quickly? How should outliers be handled? | [03] |
| (b) | Describe steps involved in Knowledge Discovery for Data (KDD). | [04] |
| (c) | Explain apriori algorithm to perform Market Basket Analytics. | [05] |
| Q.3 | Attempt the following: |  |
| (a) | How does data stream analytics differ from regular data analytics? | [03] |
| (b) | Describe characteristics of Big Data. (any four) | [04] |
| (c) | Provide schemas for multidimensional models for data about data-warehouses generally. How do fact/s and dimensions connect in RDBMS? | [05] |