



K. K. Wagh Institute of Engineering Education & Research, Nashik
(An Autonomous Institute From A.Y. 2022-23)

| WINTER-2025 | |
|---|------------------------|
| Exam Seat No.: | |
| Academic Year:2025-2026 | Semester:III |
| Class:PG-II | Program:MCA |
| Branch Code:M.C.A. | Pattern:2022 |
| Name of Course:Elective II:C: Industry Elective | Course Code:MCA223004C |
| Max. Marks:60 | Duration:2.30 Hrs. |

Instructions: Candidates should read carefully the instructions printed on the Question Paper and on the cover page of the Answer Book, which is provided for their use.

1. This question paper contains 2 page(s).
2. Answer to each new question is to be started on a new page.
3. Assume suitable data wherever required, but justify it.
4. Draw the neat labelled diagrams, wherever necessary.
5. The last columns indicates the Course Outcome and level of Blooms Taxonomy of the Question/sub-question.

Marks CO

Question No. 1

- 1a) Explain the four V of Big Data Volume, Variety, Velocity, Veracity with simple examples. (6) 1

Question No. 2

- 2a) Demonstrate how you would perform slicing and filtering operations to extract specific records with an example. (6) 2

Question No. 3

- 3a) Apply the concept of data cleaning to a small dataset and demonstrate how you would identify dirty data using any two appropriate methods. (8) 3

OR

- 3b) Use suitable techniques to fill missing data in dataset and demonstrate the process with a simple example. (8) 3

- 3c) Demonstrate the steps involved in exploring a dataset after importing it using a Python library. (8) 3

OR

- 3d) Illustrate concept of hierarchical indexing with an example. (8) 3

Question No. 4

- 4a) Illustrate why data visualization is important in decision making with a simple scenario. (8) 4

OR

- 4b) Apply your understanding of conventional data visualization tools by selecting any two tools and demonstrating how they can be used to visualize data for a given purpose. (8) 4

- 4c) Illustrate any four types of data visualization with examples. (8) 4

OR

- 4d) What challenges may occur during data visualization? Explain any four. (8) 4

Question No. 5

5a) Differentiate between Apache Spark and Power BI. (8) 5

OR

5b) Illustrate the Spark unified stack with a labelled diagram. (8) 5

5c) Illustrate how Power BI generates quick insights from uploaded data. (8) 5

OR

5d) Analyze the computational trade-offs involved in a Spark application by examining any four key factors. (8) 5

..... End of question paper.....