

Dr. Babasaheb Ambedkar Open University
Term End Examination January – 2026

Course	: MSCDS	Date	: 24/02/2026
Subject Code	: MSCDS-202	Time	: 11:30am to 01:45pm
Subject Name	: Data Analysis and Visualization	Duration	: 02.15 Hours
		Max. Marks	: 70

Section A

Answer the following (Attempt any three) (30)

1. Discuss various data cleaning methods
2. Explain various graphical representations used to represent different types of data.
3. Discuss descriptive and inferential statistics and their role in data analysis.
4. Discuss the challenges of high-dimensional data visualization and explain techniques used to visualize such data.
5. Explain in brief about sector-specific data analytics applications in healthcare, finance, education, and governance.

Section B

Answer the following (Attempt any four) (20)

1. Describe any four principles of effective data visualization.
2. Differentiate between data summarization and data aggregation with suitable examples.
3. Discuss the advantages of data visualization over tabular data representation.
4. Compare any two popular data visualization libraries.
5. Describe the difference between supervised and unsupervised learning.
6. How does AI enhance data visualization techniques?

Section C

Part – A (Multiple Choice Questions) (10)

- 1 Which of the following is a qualitative data type?
A Nominal B Interval
C Ratio D Continuous
- 2 Which visualization is commonly used for geographic data?
A Box plots B Heat maps
C Histograms D Line graphs
- 3 High-dimensional data visualization is challenging due to:
A Low storage B Too few variables
C Curse of dimensionality D Simple patterns
- 4 AI in visualization mainly helps in:
A Manual chart creation B Data deletion
C File compression D Automated insight generation
- 5 Which application uses NLP?
A Image compression B Sentiment analysis
C Data normalization D Indexing

- 6 Time series forecasting is used when data is:
A Categorical B Random
C Time-dependent D Nominal
- 7 Which of the following is a supervised learning algorithm?
A K-Means B PCA
C Apriori D Linear Regression
- 8 Inferential statistics is mainly used to:
A Draw conclusions about populations B Describe sample data only
C Clean data D Encode variables
- 9 Effective visualization design should focus on:
A Decoration B Data clarity
C Complex graphics D Maximum colors
- 10 Line charts are most suitable for representing:
A Categorical comparison B Proportions
C Trends over time D Frequency distribution

Part – B (Do as Directed)

(10)

State whether the following statements are true or false

- 1 Data analysis converts raw data into meaningful information for decision-making.
- 2 Data cleaning is performed only after building predictive models.
- 3 EDA is mainly used to encrypt datasets for secure storage.
- 4 Aggregation functions combine multiple data values into a single summary value.
- 5 Visualization is unnecessary when dealing with large datasets.
- 6 Pie charts are best suited for showing trends over time.
- 7 Poor color selection can mislead users while interpreting visualizations.
- 8 Visualization libraries are used exclusively for data cleaning.
- 9 Machine learning techniques cannot be applied in data analysis.
- 10 AI can automatically generate insights from large datasets.
