

(2 ½ Hours)

[Total Marks: 75]

- N.B. 1) All questions are compulsory.
2) Figures to the right indicate marks.
3) Illustrations, in-depth answers and diagrams will be appreciated.
4) Mixing of sub-questions is not allowed.

Q. 1 Attempt ANY FOUR from the following: (20M)

- Define Data Science. Discuss in detail applications of Data Science.
- Explain Data Warehousing (DW) and Data Mining (DM) in detail.
- Explain in detail different Data Sources.
- Describe in detail Data Transformation.
- Define Data Wrangling. Discuss Data Wrangling Techniques.
- Discuss about Feature Engineering and Time Series Data.

Q. 2 Attempt ANY FOUR from the following: (20M)

- Explain Data Visualization techniques in detail.
- Define Descriptive Statistics. Explain Mean, Median, Mode and Standard Deviation in detail.
- Explain in detail Classification and Regression analysis.
- Define bias, variance and discuss about bias-variance tradeoff.
- Discuss different techniques for evaluating model performance.
- Explain Ensemble Learning with Bagging and Boosting.

Q. 3 Attempt ANY FOUR from the following: (20M)

- Explain storytelling in analysis in detail.
- Discuss visualization tools in detail.
- List and Discuss Data Management Activities.
- Elaborate the concept of Data Governance.
- Illustrate Extraction, Transformation and Load (ETL) in detail.
- Give the importance of Data Quality.

Q. 4 Attempt ANY FIVE from the following: (15M)

- Differentiate between structured and unstructured data.
- Explain Hyperparameter Tuning.
- Define Accuracy and Explain in brief.
- Discuss any three libraries of Data Science.
- Differentiate between underfitting and overfitting.
- Define Precision, Recall and F1-Score.
