Unit – 1

1. Define Data science?
2. In Data science project what is the role of Data Analyst and Application architect?
3. In Data science project what is the role of Data scientist?
4. In Data science project what is the role of Application architect?
5. Define AI, ML, DL?
6. Why Data Science is important in this era?
7. Draw Data science workflow and process.
8. In Data science project what is the role of Data scientist, Data engineer and Application architect?
9. List few areas where Data Science can be applicable.
10. What is the relationship between AL and ML.
11. Explain Data science workflow?
12. Difference between Database and Data science?
13. Explain Structured and Unstructured Data?
14. Explain about machine generated data and streaming data?
15. Explain briefly any 4 different types of data?
16. Explain about Data science Venn Diagram?
17. What are the real time applications of Data Science?
18. Explain CAP theorem?
19. List and Explain briefly the usage of any 4 data science toolkits.
20. Explain the Data Science Process in detail, by elaborating each step.
21. What is Data Science? Explain about different data types in Data Science.

Unit – 2

1. Define traditional and Big Data?
2. Techniques involved in traditional data?
3. Define data shuffling
4. What is Univariate and Bi-variate Analysis?
5. Mention few data repositories?
6. Explain any 2 Categories of API.
7. List and Explain 2 types of Secondary data briefly.
8. What is feature engineering?
9. Define Outliers and causes Outliers?
10. What is pre-processing?
11. What is the difference between Data Analyst and Data Engineer?
12. Explain Data pre-processing?
13. Explain about Data repositories?
14. Explain missing value treatments?
15. Explain any 2 methods of collecting Primary Data in detail.
16. What is an API? Explain 3 Basic elements of an API.
17. What are the different methods to explore data?
18. What are the methods to treat missing values?
19. Explain any 4 Steps of Data Exploration and Preparation in detail.
20. What are the different sources of data collection?
21. What is an API? Explain about any social media API for data collection in detail.

Unit – 3

1. Explain the term Machine Learning.
2. What is cost function in linear regression give formula?
3. Mention 3 evaluation metrics formulas in linear regression?
4. Define Gaussian model?
5. Define Bernoulli classifier?
6. Why is Data Analytics important?
7. Define tuning parameters in SVM?
8. Formula for Bayes theorem and mention posterior, prior probability?
9. Formula for kernel linear function and polynomial function
10. Define types of Analytics?
11. Explain about Simple Linear Regression model.
12. Difference between traditional programming and machine learning?
13. Explain different types of machine learning?
14. Explain three components of machine learning algorithm?
15. Explain the Data Analytics workflow.
16. What are the different statistical methods that can be applied on data? Explain briefly.
17. Explain about central limit theorem.
18. Define Bayes law. How does Naïve Bayes solve spam filter problem?
19. Explain Linear regression and Discuss the advantages and disadvantages.
20. Explain SVM and Discuss the advantages and disadvantages.
21. Explain Naive Bayes and Discuss the advantages and disadvantages.

Unit – 4

1. Define Quantitative and Quantitative data? Give examples?
2. Define few lists of ways you can encode data?
3. Mention when to use bar chart in data visualization?
4. What are heat maps?
5. When to not apply, Retinal variables to data?
6. What is Data Visualization? Why use data visualization?
7. What do you mean by Data Encoding?
8. Why is text added to data visualization and mention few texts that can be added?
9. What are the key questions for right visualization?
10. What are the charts used to display comparison variables and relationship variables?
11. What is data visualisation? What are the uses of Data Visualisation?
12. Explain about different visualisation methods in detail.
13. Explain about visual encoding.
14. Explain about Geospatial data visualization?
15. Explain Quantitative data and Qualitative data with suitable examples.
16. Explain any 4 Retinal Variables in detail.
17. What is data encoding?
18. Explain about 4 most common graphs used in data visualization?
19. Explain any 4 Types of Data Visualization Categories in detail.
20. Explain in detail Planar and Retinal encoding.
21. Explain about Distribution, Comparison, Relationship, Composition?

Unit – 5

1. How Audience factor can Influence data Visualization choices.
2. How Content factor can Influence data Visualization choices.
3. How Context factor can Influence data Visualization choices.
4. How Dynamics factor can Influence data Visualization choices.
5. How Purpose factor can Influence data Visualization choices.
6. What are the advantages of Dashboards?
7. Explain the 'Column Data Source' of Bokeh.
8. What is the use of "output_notebook( )" in Bokeh.
9. Explain about "bokeh.models".
10. Explain about "bokeh.plotting".
11. Present your view on Data Science in the domain of Health Care.
12. Present your view on Data Science in the domain of E-commerce.
13. Present your view on Data Science in the domain of fraud detection.
14. Present your view on Data Science in the domain of Manufacturing.
15. Explain any 4 Data Science applications in detail with respect to problem, approach and solution model.
16. What are the categories of visualization technology? Explain briefly.
17. Explain about Bokeh Python visualization library.
18. Explain briefly about data visualization tools used in Data Science.
19. Explain any 2 Data Visualization Tools in detail with pros and cons.
20. What is a Bokeh (Python)? What are the advantages of Bokeh over other packages?
21. Explain the Factors that Influence data Visualization choices in detail.