UNIT-I: Introduction to Data Science

1) What does one understand by the term Data Science?
2) What is the difference between data analytics and data science?
3) What is the Role nd responsibilities of DS? Exolain Each?
4) How do you approach solving any data analytics based project?
5) Why is data cleaning crucial? How do you clean the data?
6) Differentiate between box plot and histogram.
7) In your choice of language, write a program that prints the numbers ranging from one to 50
8) Why R is used in Data Visualization?
9) Explain History of R Programming Language?
10) Why do we use the summary function?
11) Which are the important steps of Data Cleaning?
12) Explain Data Science Life Cycle?
13) What are the components of DS
14) What is Vector Explain with example?
15) What are the data structure in R Programming language?
16) Enlist R Data types and Explain each with example?
17) Write code for assignment operator and explain assignment Operator?
18) Enlist any 4-Advantages of DS and explain it?
19) What are the applications of Data Science?
20) What are the Key Components of Data Science?
21) Who is Data Scientists? Role & Responsibility?
22) Explain Rules for Variable Declaration in R Programming?
23) What is Data Frame in R Language?
24) What are the tools used for data analysis?
25) Explain Keyword in R Programming and enlist any eight keywords?
26) Explain matrix in details with example?

UNIT-II: Linear Algebra & Statistical Modeling for Data Science

1) Is linear algebra used in data science?
2) At what conditions does the inverse of a diagonal matrix exist?
3) How do you find eigenvalues of a matrix? Could you provide an example?
4) What's the difference between Cross Product and Dot Product?
5) What is Ax=b? When doees Ax=b has a unique solution?
6) How do you know if a system of two linear equations has one solution, multiple solutions or no solutions?
7) What are scalars, vectors, matrices, and tensors?
8) What is an inverse matrix?
9) What is a random variable?
10) What is a probability distribution?
11) What is a probability mass function?
12) What is a probability density function?
13) What is a joint probability distribution?
14) What are expectation, variance and covariance?
15) Compare covariance and independence.
16) What is the covariance for a vector of random variables?
17) If two random variables are related in a deterministic way, how are the PDFs related?
18) What is population mean and sample mean?
19) What is population standard deviation and sample standard deviation?
20) What is the formula for calculating the s.d. of the sample mean?
21) What is confidence interval?
22) What is standard error?
23) How is missing data handled in statistics?
24) How is the statistical significance of an insight assessed?
25) What is the difference between inferential statistics and descriptive statistics?
26) What is the difference between population and sample in inferential statistics?
27) Most common characteristics used in descriptive statistics?
28) What are observational and experimental data in statistics?
29) What is the Chi-Square test?
30) When to use the Chi-Square test?
31) What should be the data format for the Chi-Square test?
32) How to perform the Chi-square test?

UNIT:III – Optimization for Data Science

1) What are Bayesian Optimization Methods?
2) How do you use the Convex Optimization Approachto Minimize
3) Regret?
4) How do you use the Convex Optimization Approach to Minimize Regret?
5) What is multivariate optimization in data science?
6) What is Optimisation in Machine Learning?