

Detailed Curriculum

Module - 1 Python for Data Science

1. Basic building blocks
2. Conditional statements
3. Loop statements
4. String
5. List
6. Dictionary
7. Tuple
8. File handling
9. Function

Module - 2 R for Data Science

1. Introduction
2. Variables
3. Data Types
4. Numbers
5. Characters
6. Logical and Operators
7. List
8. Array
9. Matrices
10. Data Frame

Module - 3 Data Science Library and data visualization Using Python

1. Numpy
2. Pandas
3. Matplotlib
4. Seaborn
5. Plotly Bokeh
6. Sklearn

Module - 4 Maths Behind Data Science

A. Descriptive Statistics

1. Types of Data- Quantitative
Qualitative
2. Types of Scale- Nominal,
Ordinal, Ratio, Interval
3. Measure of Central Tendency
 - ✓ Mean
 - ✓ Mode
 - ✓ Median
4. Measure of Dispersion (Spread)
 - ✓ Range
 - ✓ IQR
 - ✓ Variance
 - ✓ Standard Deviation
5. Covariance and Correlation
6. Exploratory Data Analysis (EDA)
7. Univariate Analysis
8. Bivariate Analysis
9. Multivariate Analysis

B. Inferential Statistics

1. Sample & Population Terminology
2. Sampling Techniques
3. Sampling Errors
4. Continuous & Discrete Random Variables
5. Probability Distribution Functions
Continuous & Discrete
6. Cumulative Distribution Functions
7. Normal Distribution
8. Uniform Distribution
9. Exponential Distribution
10. Bernoulli / Binomial Distribution
11. Poisson Distribution
12. Chi-Square etc...
13. Goal of Statistical Inference
14. Inferences about population parameter from sample statistic using standard normal distribution
15. Sampling Distribution
16. Central Limit Theorem (CLT)
17. Z-score, Confidence level, Confidence Interval, Significance Level

C. Hypothesis Testing

1. What is Hypothesis
2. Null Hypothesis & Alternate Hypothesis
3. Why Hypothesis Testing
4. Types of Hypothesis Tests
5. One sample t-test
6. Paired t-test
7. Z-test
8. Chi-square test
9. One sample proportion test
10. Anova test etc...
11. Types of errors- Type 1 & Type 2

Module - 5 Exploratory data analysis /Data Cleaning Techniques/ Data Preparation Techniques

1. Identification of variables and data types
2. Univariate, bivariate, multivariate analysis
3. Variable transformations
4. Binning/ Discretization/ Binarization
5. Missing value treatment
6. Outlier treatment
7. Categorical to Numerical
8. Correlation Analysis
9. Feature Selection

Module - 6 Big Data Engineering: SQL/Hadoop/Spark

1. Introduction to Database
2. Introduction to SQL
3. Basic Operations of SQL
 - Joins
 - Groupby
 - Orderby
 - Where and having
4. Introduction to Hadoop and spark
 - Partition by
 - Windows Function-lead,lag
 - Sub queries

Module - 7 Tableau

1. Introduction to TABLEAU
2. Different types of charts
3. Dashboard Making
4. Case studies

Module - 8 Advance Excel

1. Introduction
2. Different methods of Excel

Module - 9 Overview of Machine Learning

1. Types of machine Learning
2. Case studies