

Data Science Course Content (5 days)

By Dr. Vishwanath Rao

Introduction to Data Science

- Need for Data Scientists
 - Foundation of Data Science
 - What is Business Intelligence
 - What is Data Analysis, Data Mining, and Machine Learning
 - Analytics vs Data Science
 - Value Chain
 - Types of Analytics
 - Lifecycle Probability
 - Analytics Project Lifecycle
-
- Environment set-up
 - Jupyter overview
 - Python Numpy
 - Python Panda
 - Python Matplotlib

Data

- Basis of Data Categorization
- Types of Data
- Data Collection Types
- Forms of Data and Sources
- Data Quality, Changes and Data Quality Issues, Quality Story
- What is Data Architecture
- Components of Data Architecture
- OLTP vs OLAP
- How is Data Stored?

Big Data

- What is Big Data?
- 5VsofBigData
- Big Data Architecture, Technologies, Challenge and Big Data Requirements
- Big Data Distributed Computing and Complexity
- Hadoop
- Map Reduce Framework
- Hadoop Ecosystem

Data Science Deep Dive

- What is Data Science?
- Why are Data Scientists in demand?
- What is a Data Product
- The growing need for Data Science
- Large-Scale Analysis Cost vs Storage
- Data Science Skills
- Data Science Use Cases and Data Science Project Life Cycle & Stages
- Map-Reduce Framework
- Hadoop Ecosystem
- Data Acquisition
- Where to source data
- Techniques
- Evaluating input data
- Data formats, Quantity and Data Quality
- Resolution Techniques
- Data Transformation
- File Format Conversions
- Anonymization

Data Manipulation in R

- Various phases of Data Cleaning
- Functions used in Inspection
- Data Cleaning Techniques
- Uses of functions involved
- Use-cases for Data Cleaning

Data Import Techniques

- Import data from spreadsheets and text files
- Importing data from statistical formats
- Packages installation for database import
- Connecting to RDBMS from R using ODBC and basic SQL queries

- Web Scraping
- Other concepts on Data Import Techniques

Data Visualization

- Storytelling with Data
- Principle tenets
- Elements of Data Visualization
- Infographics vs Data Visualization
- Data Visualization & Graphical functions
- Plotting Graphs
- Customizing Graphical Parameters to improvise the plots
- Various GUIs
- Spatial Analysis
- Other Visualization concepts