

## Data Science with Python Online Training Course

Our Data Science with Python course empowers learners with essential data analysis and machine learning skills using Python. Designed for beginners and experienced professionals, the course covers data manipulation, visualization, statistical analysis, and key machine learning algorithms. With practical projects, hands-on exercises, and expert guidance, students gain real-world, industry-ready skills.

Course Duration: 35 To 45 Hours

Flexible Timing: Weekdays & Weekend

Minimum Batch Size

100 % Placement Assistance

### Key Features:

Instructor-led Live Online Training

Hands-on Learning

Flexible Schedules

Experienced Trainers

100% Job Assistance

Real-Time Projects

Study Materials

Interview Preparation Guidelines

Global Certification Support

### Data Science with Python Course Prerequisites:

No advanced technical background is needed, but familiarity with basic math concepts, especially statistics and probability, is useful. Knowledge of Excel and programming basics (Python is a plus) can help. Curiosity about data analysis, problem-solving, and hands-on learning will make the course smoother and more rewarding.

## Data Science with Python Training Syllabus

### Module 1: Introduction to Data Science

- Overview of Data Science, Machine Learning, Deep Learning, and Artificial Intelligence
- Types of Data Analytics: Descriptive, Diagnostic, Predictive, and Prescriptive
- Applications of Data Science in different industries

### Module 2: Introduction to Python

- Overview of Python and its applications in data science
- Installing Python, setting up IDEs (Jupyter Notebook, VS Code)
- Jupyter Notebook basics and useful shortcuts

### Module 3: Python Basics

- Core data types: strings, integers, floats, and booleans
- Data structures: Lists, Dictionaries, Tuples
- Basic operations: slicing, IF statements, loops (for, while)
- Functions and arrays; selecting data by position and labels

### Module 4: Essential Python Packages

- Overview and usage of key packages:
  - Pandas for data manipulation
  - Numpy for numerical operations
  - Matplotlib and Seaborn for data visualization
  - Sci-kit Learn for machine learning basics

# ARIYALAM

## **Module 5: Importing and Exporting Data**

- Reading data from CSV, Excel, and SQL databases
- Saving and loading Python data objects
- Writing data to CSV, Excel, and JSON formats

## **Module 6: Data Manipulation Techniques**

- Selecting rows and columns, renaming, and filtering data
- Sorting and grouping data
- Merging and joining datasets
- Aggregation and data munging for cleaning and preparation

## **Module 7: Statistics and Probability Essentials**

- Measures of Central Tendency (Mean, Median, Mode) and Dispersion (Variance, Standard Deviation)
- Probability concepts and types (conditional probability, Bayes' theorem)
- Skewness, Normal Distribution, and Bias-Variance Tradeoff
- Analyzing data deviation, outliers, and plotting techniques (box plot, scatter plot)

## **Module 8: Data Preprocessing and Feature Engineering**

- Handling missing values (imputation techniques, e.g., KNN)
- Categorical encoding (one-hot encoding, label encoding)
- Data normalization and scaling
- Outlier analysis and treatment
- Correlation analysis and feature selection

## **Module 9: Introduction to Machine Learning Basics**

- Supervised vs. Unsupervised Learning
- Error Metrics for Classification: Confusion Matrix, Precision, Recall, Specificity, F1 Score
- Error Metrics for Regression: MSE, RMSE, MAE, MAPE

## **Module 10: Supervised Learning Algorithms**

- Linear Regression: slope, intercept, R-squared value
- Logistic Regression: odds ratio, probability of success/failure, ROC curve
- Bias-Variance tradeoff and model evaluation

## **Module 11: Unsupervised Learning Algorithms**

- K-Means Clustering and K-Means++ for grouping data
- Hierarchical Clustering: understanding dendrograms and distance measures

## **Module 12: Advanced Machine Learning Algorithms**

- K-Nearest Neighbors for classification
- Naive Bayes for probability-based classification
- Decision Trees (CART, C5.0) and Random Forest for predictive modeling
- Support Vector Machines (SVM) and Ensemble Techniques

## FAQ:

### 1. What is data science, and why is it important?

Data science is the field of studying data to find trends and useful information. It's important because it helps companies and people make better decisions based on data instead of guesses.

### 2. Do I need a coding background to learn data science?

No, you don't need a coding background. Many data science courses start with the basics, and you'll learn essential programming skills in Python during the course.

### 3. What skills will I learn in a data science course?

You'll learn Python programming, data analysis, statistics, data visualization, and machine learning. These skills help you analyze data, find insights, and build predictive models.

### 4. What kinds of jobs can I get with data science skills?

With data science skills, you can pursue jobs like data analyst, data scientist, business analyst, or machine learning engineer. Many industries hire data science professionals.

### 5. Is mathematics important for Data Science?

Yes, math—especially statistics, probability, and linear algebra—is crucial in Data Science for developing algorithms, performing analyses, and understanding data patterns. Courses often include these basics for better understanding.

### 6. What top companies have Jobs in Data Science?

Accenture, Google, Microsoft, IBM, Deloitte, Wipro, and More.