

SKILLSOURCE TECHSOFT

Data Science with PYTHON

Syllabus

Call us on +91 8830322917 / +91 7709636236
hr.skillsourcetek@gmail.com www.techsoftskillsource.com

Introduction to Data Science and Analytics

- What is Data Science?
- What's the need & why it's in demand?
- What is Data Analytics?
- Components of Data Science.
- Real-life examples & applications.
- Introduction to different programming languages used for Data Science.

Python Programming 1: Basics

- Why Python for Data Science?.
- Installing Python.
- Python IDEs.
- Python Basic Syntax & Data types.
- Lists.
- String Manipulation.
- Conditional Statements.
- Looping Statements.
- Dictionaries.
- Tuples.
- Functions.
- Array.

Python Programming 2: Libraries

- Pandas.
- Numpy.
- Sci-kit Learn.
- Matplot library.
- Seaborn.
-

Data Pre-processing & Exploration

- Extracting data from different sources.
- Reading XLSX, CSV etc. files.
- Handling Missing Values.
- Handling Outliers.
- Different Data Mugging Techniques.
- One Hot Encoder & Feature Scaling.

- **Statistics**

- Mean Mode, Median.
- Random Variable.
- Probability, Probability Distribution of Random Variables.
- Type of Random Variables - Based on Scale of Measurement.
- Normal, Binomial, Poisson Distribution.
- Standard Normal Distribution and Z-Score.
- Sampling & Sampling Distribution.
- Central Limit Theorem.
- Simulation.
- Hypothesis and hypothesis Testing.
- Hypothesis Testing using z-test, t-test.

Machine Learning

- What is Machine Learning ?.
- Overview & Terminologies.
- Difference between AI & ML.
- Supervised & Unsupervised ML
- Supervised ML Models.
- Linear & Logistic Regression.
- Regression methods, Classification.
- Sampling & Sampling Distribution.
- K Nearest Neighbours KNN.
- Decision Tree, Random Forest.
- Unsupervised ML Models.
- K Means Clustering.
- Under fitting and Overfitting.
- Confusion Metrix.
- K-Fold Cross Validation.
- Regression Evaluation Metrics.
- Time Series Analysis.
- Support Vector Machine (SVM).
- Na'ive Bayes.