

# Introduction to Data Science Course

## Table of Contents

### **Module 1 : Foundation for Data Science**

- Introduction to Data Science
- Analytics Landscape
- Life Cycle of a Data Science Projects
- Data Science Tools & Technologies

### **Module 2 : Probability & Statistics**

- Measures of Central Tendency
- Measures of Dispersion
- Descriptive Statistics
- Probability Basics
- Marginal Probability
- Bayes Theorem
- Probability Distributions
- Hypothesis Testing

### **Module 3 : Basics of Python**

- Install Anaconda
- Data Types & Variables
- String & Regular Expressions

### **Module 4 : Python Built-in Data Structures**

- Python list
- Python dictionaries
- Python set
- Python tuple
- Comprehensions

### **Module 5 : Control & Loop Statements in Python**

- For Loop
- While Loop
- Break Statement
- Next Statements
- Repeat Statement
- if, if...else Statements
- Switch Statement

## **Module 6 : Functions & Classes in Python**

- Writing your own functions (UDF)
- Calling Python Functions
- Functions with Arguments
- Calling Python Functions by passing Arguments
- Lambda Functions
- Classes & Objects

## **Module 7 : Working with Data**

- Reading files with Python
- Writing files from Python
- Reading files using Pandas library
- Saving Data using Pandas library

## **Module 8 : Analyze Data using Pandas**

- Clean & Prepare Datasets
- Manipulate DataFrame
- Summarize Data
- Churn Insights from Data

## **Module 9 : Visualize Data**

- Charts using Matplotlib
- Charts using Seaborn
- Charts using ggplot

## **Module 10 : Advanced Statistics & Predictive Modeling**

- ANOVA
- Linear Regression (OLS)
- Case Study: Linear Regression
- Principal Component Analysis
- Factor Analysis
- Case Study: PCA/FA
- Logistic Regression (MLE)
- Case Study: Logistic Regression
- K-Nearest Neighbor Algorithm
- Case Study: K-Nearest Neighbor Algorithm
- Decision Tree
- Case Study: Decision Tree

## **Module 11 : Time Series Forecasting**

- Understand Time Series Data
- Visualizing Time Series Components
- Exponential Smoothing
- Holt's Model
- Holt-Winter's Model
- ARIMA
- Case Study: Time Series Modeling on Stock Price

## **Module 12 : Introduction to Machine Learning**

- What is Machine Learning?
- Supervised Learning
- Unsupervised Learning
- Using Scikit-learn
- Scikit-learn classes
- Case Study: Machine Learning Algorithm