

AI/ML Data Science

About AnnexIT

AnnexIT is a leading e-learning platform providing live instructor-led interactive online training. We cater to professionals and students across the globe in categories like Big Data & Hadoop, Business Analytics, NoSQL Databases, Java & Mobile Technologies, System Engineering, Project Management and Programming. We have an easy and affordable learning solution that is accessible to millions of learners across the globe.

Python Statistics for Data Science Course

Python Scripting allows programmers to build applications easily and rapidly. This course is an introduction to Python scripting, which focuses on the concepts of Python, it will help you to perform operations on variable types using Pycharm. You will learn the importance of Python in real time environment and will be able to develop applications based on Object Oriented Programming concept. End of this course, you will be able to develop networking applications with suitable GUI

WEEK 2 6 Modules 18 Hours 14 Skills

Course Content

Python Statistics for Data Science Course Curriculum

Understanding the Data

Goal: In this module, you will be introduced to data and its types and accordingly sample data and derive meaningful information from the data in terms different statistical parameters.

Objectives: At the end of this Module, you should be able to:

- Understand various data types
- Learn Various variable types
- List the uses of variable types
- Explain Population and Sample
- Discuss sampling techniques
- Understand Data representation

Topics:

- Introduction to Data Types Numerical parameters to represent data
- Mean
- Mode
- Median
- Sensitivity
- Information Gain
- Entropy
- Statistical parameters to represent data

Hands-On/Demo

- Estimating mean, median and mode using python
- Calculating Information Gain and Entropy

Probability and its uses

Goal: In this module, you should learn about probability, interpret & solve real-life problems using probability. You will get to know the power of probability with Bayesian Inference.

Objectives: At the end of this Module, you should be able to:

- Understand rules of probability
- Learn about dependent and independent events
- Implement conditional, marginal and joint probability using Bayes Theorem
- Discuss probability distribution
- Explain Central Limit Theorem

Topics:

- Uses of probability
- Need of probability
- Bayesian Inference
- Density Concepts
- Normal Distribution Curve

Hands-On/Demo:

- Calculating probability using python
- Conditional, Joint and Marginal Probability using Python
- Plotting a Normal distribution curve

Statistical Inference

Goal: Draw inferences from present data and construct predictive models using different inferential parameters (as a constraint).

Objectives: At the end of this Module, you should be able to:

- Understand the concept of point estimation using confidence margin
- Draw meaningful inferences using margin of error
- Explore hypothesis testing and its different levels

Topics:

- Point Estimation
- Confidence Margin
- Hypothesis Testing
- Levels of Hypothesis Testing

Hands-On/Demo:

- Calculating and generalizing point estimates using python
- Estimation of Confidence Intervals and Margin of Error

Testing the Data

Goal: In this module, you should learn the different methods of testing the alternative hypothesis.

Objectives: At the end of this module, you should be able to:

- Understand Parametric and Non-parametric Testing
- Learn various types of parametric testing
- Discuss experimental designing
- Explain a/b testing

Topics:

- Parametric Test
- Parametric Test Types
- Non- Parametric Test
- Experimental Designing
- A/B testing

Hands-On/Demo:

- Perform p test and t tests in python
- A/B testing in python

Data Clustering

Goal: Get an introduction to Clustering as part of this Module which forms the basis for machine learning.

Objectives: At the end of this module, you should be able to:

- Understand the concept of association and dependence

- Explain causation and correlation
- Learn the concept of covariance
- Discuss Simpson's paradox
- Illustrate Clustering Techniques

Topics:

- Association and Dependence
- Causation and Correlation
- Covariance
- Simpson's Paradox
- Clustering Techniques

Hands-On/Demo:

- Correlation and Covariance in python
- Hierarchical clustering in python
- K means clustering in python

Regression Modelling

Goal: Learn the roots of Regression Modelling using statistics.

Objectives: At the end of this module, you should be able to:

- Understand the concept of Linear Regression
- Explain Logistic Regression
- Implement WOE

- Differentiate between heteroscedasticity and homoscedasticity
- Learn the concept of residual analysis

Topics:

- Logistic and Regression Techniques
- Problem of Collinearity
- WOE and IV
- Residual Analysis
- Heteroscedasticity
- Homoscedasticity

Hands-On/Demo:

- Perform Linear and Logistic Regression in python
- Analyze the residuals using python