**3\_Python Statistics for Data Science**

3. Python Statistics For Data Science



| **Module** | **Time** (**Mins**) |
| --- | --- |
| **Module 1: Understanding the Data** |  |
| **Module 2: Probability and its uses** |  |
| **Module 3: Statistical Inference** |  |
| **Module 4: Data Clustering** |  |
| **Module 5: Testing the Data** |  |
| **Module 6: Regression Modelling** |  |

# Module 1: Understanding the Data

# 

## Introduction to Data Types

## Numerical parameters to represent data.

### Mean

### Mode

### Median

### Sensitivity

### Information Gain

### Entropy

## Statistical parameters to represent data.

# Module 2: Probability and its uses

# 

## Uses of probability

## Need of probability

## Bayesian Inference

## Density Concepts

## Normal Distribution Curve

# Module 3: Statistical Inference

# 

## Point Estimation

## Confidence Margin

## Hypothesis Testing

## Levels of Hypothesis Testing

## Levels of Hypothesis Testing

# 

# 

# 

# Module 4: Data Clustering

# 

## Association and Dependence

## Causation and Correlation

## Covariance

## Simpson’s Paradox

## Clustering Techniques

# Module 5: Testing the Data

# 

## Parametric Test

## Parametric Test Types

## Non- Parametric Test

## Experimental Designing

## A/B testing

## Chi-Square testing

# Module 6: Regression Modelling

# 

## Logistic and Regression Techniques

## Problem of Collinearity

## WOE and IV

## Residual Analysis

## Heteroscedasticity

## Homoscedasticity