

# Statistics For Data Science Using Python

Book Contents:

Preface:

Why we write this book?

Whom this for?

How the book is organized?

Requirements:

A Quick note on anaconda and the required libraries like scipy, pandas, matplotlib, etc.

## 1. What is Data Science

— We can explain the basics of data science, its life cycles process, etc and end with what is the role and importance of statistics in Data Science

## 2. Measures of Central Tendency

— We can talk about Mean, median and mode and their respective use cases and the relation between them.

## 3. Variance and Standardization

— We can discuss on variances, standardization, CLT, etc

## 4. Sampling distributions

— We can explain on PMF, CDF, etc and their important distributions like normal, gaussian, exponential etc.

## 5. Estimations

— We can introduce on Pop mean vs sample mean, Confidence Intervals, Approx margin of error, etc.

## 6. Hypothesis Testing

— Z- test, T-Test, F-Test, ANOVA, Chi-Square

## 7. Correlations

— We can talk on residual plots, pearson coefficients, etc.

## 8. Regression Analysis

— We can discuss on basic linear regression and logistic regression.