



LEARN DATA SCIENCE USING PYTHON

LEARN

DATA SCIENCE

using Python

Data Science is a study of data. We believe Data is everything in today's world that can be treated as 'Information' about different aspects of Business.

In this course, we will cover all the components of Data Science like Data Management, Programming, Visualization, and Analytics.

All the components of Data Science mentioned above are equally important to successfully become a Data Scientist.

This course focuses mainly on Python tools. It starts with programming using Python followed by analytical and visualization techniques. It includes multiple hands-on exercises and project work in the domains of marketing, banking, finance, risk management, retail, clinical, manufacturing/inventory management, etc.

The ultimate purpose of this course is to cover all the basics and advance predictive modeling techniques which are a must in today's competitive edge.

This course will cover all the programming, visualization and analytical models to be implemented in Python.

KEY FEATURES:

- Online/Classroom Training
- Flexible Schedule
- 35 Days (2 Hrs. per day)
- Certification and Job Assistance
- Project work & Exercises
- Study Materials (Software / Reference Book / Interview Preparation / Project Data / Python Codes)



FACULTY PROFILE:

Our Faculty has 9+ years of experience in the field of Data and Analytics, Data Science, Training, and Analytical application development.

Our Faculty is a certified Predictive Modeler and Programmer from SAS Institute Inc.

Our Faculty worked on SAS (All analytical modules), Python (Data Science and Data Management libraries), and Big Data (Hadoop and Hive) during his 9+ years of tenure.

He has successfully implemented analytics and developed analytical applications in the field of AML, Risk Management, Healthcare Fraud, Retail, and Market Research.

A former employee of TCS and SAS Institute Inc.



COURSE CONTENT

1. INTRODUCTION OF DATA SCIENCE

- Welcome/General Discussion about the expectation from course
- Definition of Data
- Difference between data management and data analytics
- Data Science components

2. PROGRAMMING USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Python Overview
- Python Data Types
- Python operations using Numbers, String, Logical, Arithmetic and so on
- Python Strings
- Python Lists
- Python Tuple
- Python Dictionary
- FOR and WHILE loops
- IF/THEN/ELSE in Python
- Data Manipulation using Numpy and Pandas

3. STATISTICS USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Levels of Measurement and Variable types
- Descriptive Statistics and Picturing Distributions
- Confidence Interval for the Mean

4. HYPOTHESIS TESTING AND ANOVA USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- One-Sample T-Test of Comparing Means
- Two-Sample T-Test of Comparing Means
- One Way ANOVA
- Assumptions of ANOVA Modeling
- N-Way ANOVA
- ANOVA Post Hoc Studies



5. EXPLORATORY DATA ANALYSIS USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Data Exploration by using Scatter Plots
- Pearson and Spearman Correlations

6. LINEAR REGRESSION USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Fit Simple Linear Regression Model
- Assumptions of Linear Regression Model
- Analyze the output of the Linear Regression
- Producing Predicted Values
- Difference between Simple Linear Regression and Multiple Linear Regression Models
- Fit Multiple Linear Regression Model
- Stepwise Regression/Model Selection Techniques

7. REGRESSION DIAGNOSTICS USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Residual Analysis
- Influential Observation
- Difference between Influential Observation and Outliers
- Collinearity Diagnostics

Model Building Process using SAS and Python

8. CATEGORICAL DATA ANALYSIS USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Examining Distributions
- Test of Associations by using the chi-square test
- Fisher's Exact p-values for Pearson Chi-square test

9. LOGISTIC REGRESSION USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Odds and Odds Ratio
- Simple Logistic Regression
- Multiple Logistic Regression with categorical predictors
- Analyze the output of Logistic Regression



10. MEASURE MODEL PERFORMANCE USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Apply the principles of honest assessment to model performance measurement
- Rare event adjustments
- Assess classifier performance using the confusion matrix
- Model selection and validation using training and validation data
- Create and interpret graphs (ROC, lift, and gains charts) for model comparison and selection
- Establish effective decision cut-off values for scoring

11. DECISION TREE MODELING USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Introduction to Decision Tree Modeling
- Model essential for Decision Tree Models
- Decision Tree Model Development by using CHAID, Entropy/Information Gain, and Gini
- Decision Tree Model Tuning

12. GRADIENT BOOSTING (XGBOOST) USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Introduction to Boosting
- Example of Boosting
- Regression Decision Tree
- Gradient Boosted Trees Regression.

13. TIME SERIES FORECASTING MODELS USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Introduction to Time Series Forecasting
- Component Factors affecting Time Series
- Moving Average (MA)
- Exponential Smoothing
- Trend Fitting Models (Linear trend, Quadratic trend, and Exponential trend)
- Autoregressive Integrated Moving Average (ARIMA) Model
- Vector Autoregression (VAR) Model
- Autoregressive Conditional Heteroskedasticity (ARCH) Model
- Generalized Autoregressive Conditional Heteroskedasticity (GARCH) Model
- Long Short-Term Memory (LSTM) Model



PROJECT CONTENT

1. PORTFOLIO DATA ANALYSIS

Domain Risk Management

Problem Statement As an analyst, you need to advise your client to decide which mutual fund risk category should invest in.

Topic Descriptive Analytics, Distributions, and Visualization

2. EFFECTIVENESS OF PRODUCTION PROCESS

Domain Manufacturing/Inventory Management

Problem Statement As a manager/supervisor of a company, you need to measure the effectiveness of the production of cereal boxes. The aim is to analyze whether or not the cereal boxes' weight is as per company specifications.

Topic Hypothesis Testing (One-Sample tests)

3. PRODUCT ASSORTMENT STRATEGIES IN RETAIL STORE

Domain Marketing/Retail

Problem Statement As a regional sales manager of a company, you need to analyze the mean sales comparison between two types of displays of products in the retail store. The aim is to decide whether or not the Promotional display of the product is more effective than the Normal display of the product. This helps management to decide the display location of the product in a store that will maximize sales

Topic Hypothesis Testing (Two-Sample tests)



4. DRUG ANALYSIS

Domain

Clinical

Problem Statement

Before you launch the new drug in the market, you need to analyze the effect of new drug and its different doses on the blood pressure of the human body

Topic

Analysis of Variance (ANOVA Models)

5. ANALYZING CONSUMPTION OF OXYGEN IN THE HUMAN BODY DURING RUNNING

Domain

Physiology

Problem Statement

In exercise physiology, an objective measure of aerobic fitness is how effectively the body can absorb and use oxygen during their 1.5 miles run. Factors affecting oxygen consumption are runtime, age, and gender, run pulse, rest pulse, and so on. The aim is to identify the key factors affecting oxygen consumption during a run.

Topic

Analysis of Variance (EDA and Linear Regression Models)

6. TITANIC EVENT ANALYSIS

Domain

Event Analysis

Problem Statement

On the 14th of April, the Titanic hit an iceberg and sank. There were 1517 fatalities from different age groups, class (1, 2, and 3), and gender. The objective is to measure how all these factors are associated with the survival status of passengers.

Topic

Odds, Odds Ratio, Chi-Square tests, Ordinal associations, and Logistic Regression Model



7. MARKETING CAMPAIGN FOR A BANK

Domain

Marketing

Problem Statement

A target marketing campaign for a bank was undertaken to identify a segment of customers who are likely to respond to an insurance product. Here, the target variable is whether or not the customers bought insurance product and it depends on factors like Product usage in three months, demographics, transaction patterns as like deposit amount, checking account, a branch of the bank, Residential information (like urban, rural) and so on.

Topic

Classification, Categorical Data Analysis, Logistic regression, Decision Tree and Gradient Boosting (XGBOOST)



COURSE CERTIFICATION:

The whole Data Science course content is planned by industry experts for you to land the best positions in top MNCs.

Toward the finish of each topic, there will be a project associated with that topic. The project contains live data and codes from Python.

Techinfoplace course fruition authentication will be granted to you when you complete the task work.

After completion of the projects from each topic, your skills will be equal to 6 months of rigorous industry experience in each topic.

FAQ:

1. Why should I learn Data Science from Techinfoplace?

- Techinfoplace offers exclusive Data Science online/classroom courses for professional/fresher who wants to expand their knowledge base and start a career in this field. There are many reasons for choosing Techinfoplace.
- A personal mentor to track your progress
- Techinfoplace online/classroom sessions conducted by experienced professionals in this field
- Real-time exercises, assignments, and projects
- Study materials, reference books for every topic.
- 24/7 learning support
- The large community of like-minded learners

2. What are the different modes of training that Techinfoplace provides?

- Online Session
- Classroom Session



3. Who can attend this course?

- Anyone who wants to play with Data can enter into this field.

4. What kind of projects are included as part of the training?

- This course is known for the projects involved in it. This is purely Job Oriented training.
- You will work on highly exciting projects in the domains of high technology, Retail, Banking, Marketing, Clinical, Manufacturing, and so on.
- After completing the projects successfully, your skills will be equal to 6 months of rigorous industry experience in each topic.
- You will also encounter a different interview question for each topic which will help in cracking Data Science interviews

5. Is there any discount for this course?

- Yes. The fees mentioned for this course are for one candidate. If you come with at least 3 enrollments, a discount of 10% on the base price would be given to that group.