

Practical Data Science with Python

Learn via: **Classroom / Virtual Classroom / Online**

Duration: **5 Day**

<https://bilginc.com/en/training/practical-data-science-with-python-379-training/>

Overview

An introduction to Statistics, Python, Analytics, Data Science and Machine Learning. Sets up practitioners with working knowledge of whole field of data science, along with immediate practical knowledge of key analytical tasks.

This 5-day course is hands-on, practical and workshop based. It is the start of an experienced developer's journey towards becoming a Data Scientist. If you are a software engineer, in business intelligence, or you are an SQL specialist, this is the course for you.

By attending this course you will learn how to become a professional Data Scientist. You're going to be able to demystify and understand the language around data science and understand the core concepts of analytics and automation. You'll also develop practical, hands-on, advanced skills in Python, targeted towards data analysis and Machine Learning so you can create sophisticated statistical models.

Target Audience

For fledging data science practitioners, and for IT professionals who wish to move to the exciting world of data analytics and machine learning.

Prerequisites

- GCSE level mathematics or above. Alternatively, familiar and comfortable with logical and mathematical thinking
- Familiar with basic knowledge of programming: variables, scope, functions

What You Will Learn

At the end of the course attendees will know:

- Fundamental concepts of Data Science
- Methodologies used in Machine Learning
- Summary statistics and how to use statistical inference to analyse data
- Hands on Python programming language for numerical analysis
- Most used simple machine learning algorithms

At the end of the course attendees will be able to:

- Speak the language of data scientists
- Write Python programs to analyse data
- Understand a Python program in the context of data analytics
- Explore and visualise data using Python
- Build working machine learning models

Outline

01 Introduction to Data Science

- Understanding Big Data challenges for storage and analytics
- Identifying potential Big Data projects
- Designing successful Data Science projects

02 Introduction to Machine Learning

- Identify types of machine learning:
 - Supervised

- Unsupervised
- Reinforcement Learning
- Identify use cases

03 Jupyter Notebook

- Identify Anaconda and Jupyter
- Work with Jupyter Notebooks
- Practical Lab Activity

04 Python Fundamentals Review

- Review of Python techniques:
 - Data Types and Assignment
 - Lists, Tuples, Strings, Sets and Dictionaries – and how to address from them
 - Selection and Iteration structures
 - Subroutine definitions
- Practical Lab Activity

05 Introduction to Pandas and Numpy

- Dataframes and how to address from them
- Read from a CSV and exploration of documentation for reading from other sources
- Dataframe methods and using the documentation
- Practical Lab Activity

06 Exploratory Data Analysis

- In the context of identifying appropriate Machine Learning methods:
 - Interpreting descriptive statistics using Pandas
 - Interpreting correlations and associations
- Practical Lab Activity

07 Data Visualisation

- In the context of identifying appropriate Machine Learning methods:
 - Interpreting visualisations using Pandas, Seaborn, and Matplotlib
 - Interpreting correlations and associations
 - Interpreting visualisations for EDA
- Practical Lab Activity

08 Data Preparation

- Data Preparation techniques in the context of selected Machine Learning methods:
 - Checking the quality of the data and understanding it's source
 - Identifying when to remove, replace, or retain missing data
 - Handling Imbalanced Data
 - Scaling and Normalisation requirements
- Practical Lab Activity

09 Linear Regression

- Identify appropriate situations for using Linear Regression
- Use the Supervised Machine Learning workflow to create, evaluate, tune, and visualise a linear regression model
- Practical Lab Activity

10 Logistic Regression

- Identify appropriate situations for using Logistic Regression
- Use the Supervised Machine Learning workflow to create, evaluate, and tune a logistic regression model
- Practical Lab Activity

11 Decision Trees and Random Forests

- Identify appropriate situations for using Decision Trees
- Use the Supervised Machine Learning workflow to create, evaluate, tune, and visualise a Decision Tree model
- Identify alternative classification models such as Random Forests and how to compare them
- Practical Lab Activity

12 Clustering with K-means

- Understanding and implementing the k-means clustering algorithm
- Evaluate the model performance and select k
- Practical Lab Activity