



Data Science with Python

simplilearn

Table of Contents:

- > Program Overview
- > Program Features
- > Delivery Mode
- > Prerequisites
- > Target Audience
- > Key Learning Outcomes
- > Certification Details and Criteria
- > Table of Content
- > Course End Projects
- > Customer Reviews
- > About Us

Program Overview:

Establish your mastery of data science and analytics techniques using Python by enrolling in this Data Science with Python course. You'll learn the essential concepts of Python programming and gain in-depth knowledge of data analytics, machine learning, data visualization, web scraping, and natural language processing. Python is a required skill for many data science positions, so jumpstart your career with this interactive, hands-on, Data Science with Python course.

Program Features:

- > 24 hours of Online self-paced learning
- > 44 hours of instructor-led training
- > 4 industry-based course-end projects
- > Interactive learning with Jupyter notebooks integrated labs
- > Dedicated mentoring session from faculty of industry experts

Delivery Mode:

Online Bootcamp - Online self-paced learning and live virtual classroom

Prerequisites:

To best understand the Data Science with Python course, it is recommended that you begin with these courses:

- > Python Basics
- > Math Refresher
- > Data Science in Real Life
- > Statistics Essentials for Data Science

Target Audience:

- > Analytics professionals willing to work with Python
- > Software and IT professionals interested in analytics
- > Anyone with a genuine interest in data science

Key Learning Outcomes:

This Python for Data Science training course will enable you to:

- > Gain an in-depth understanding of data science processes, data wrangling, data exploration, data visualization, hypothesis building, and testing; and the basics of statistics
- > Understand the essential concepts of Python programming such as datatypes, tuples, lists, dicts, basic operators, and functions
- > Perform high-level mathematical computations using the NumPy and SciPy packages and their large library of mathematical functions
- > Perform data analysis and manipulation using data structures and tools provided in the Pandas package
- > Gain an in-depth understanding of supervised learning and unsupervised learning models such as linear regression, logistic regression, clustering, dimensionality reduction, K-NN, and pipeline
- > Use the Scikit-Learn package for natural language processing and matplotlib library of Python for data visualization

Certification Details and **Criteria:**

- > 85 percent of online self-paced completion or attendance of one live virtual classroom
- > A score of at least 75 percent in course-end assessment
- > Successful evaluation in at least one project

Table of **Contents:**

Lesson 00 - Course Overview

- > Course Overview

Lesson 01 - Data Science Overview

- > Introduction to Data Science
- > Different Sectors Using Data Science
- > Purpose and Components of Python
- > Quiz
- > Key Takeaways

Lesson 02 - Data Analytics Overview

- > Data Analytics Process
- > Knowledge Check
- > Exploratory Data Analysis (EDA)
- > Quiz
- > EDA-Quantitative Technique
- > EDA - Graphical Technique
- > Data Analytics Conclusion or Predictions
- > Data Analytics Communication
- > Data Types for Plotting
- > Data Types and Plotting
- > Quiz
- > Key Takeaways
- > Knowledge Check

Lesson 03 - Data Analytics Overview

- › Introduction to Statistics
- › Statistical and Non-statistical Analysis
- › Major Categories of Statistics
- › Statistical Analysis Considerations
- › Population and Sample
- › Statistical Analysis Process
- › Data Distribution
- › Dispersion
- › Knowledge Check
- › Histogram
- › Knowledge Check
- › Testing
- › Knowledge Check
- › Correlation and Inferential Statistics
- › Quiz
- › Key Takeaways

Lesson 04 - Python Environment Setup and Essentials

- › Anaconda
- › Installation of Anaconda Python Distribution (contd.)
- › Data Types with Python
- › Basic Operators and Functions
- › Quiz
- › Key Takeaways

Lesson 05 - Mathematical Computing with Python (NumPy)

- > Introduction to Numpy
- > Activity-Sequence it Right
- > Demo 01-Creating and Printing an ndarray
- > Knowledge Check
- > Class and Attributes of ndarray
- > Basic Operations
- > Activity-Slice It
- > Copy and Views
- > Mathematical Functions of Numpy
- > Assignment 01
- > Assignment 01 Demo
- > Assignment 02
- > Assignment 02 Demo
- > Quiz
- > Key Takeaways

Lesson 06 - Scientific computing with Python (Scipy)

- > Introduction to SciPy
- > SciPy Sub Package - Integration and Optimization
- > Knowledge Check
- > SciPy sub package
- > Demo - Calculate Eigenvalues and Eigenvector
- > Knowledge Check
- > SciPy Sub Package - Statistics, Weave and IO
- > Assignment 01
- > Assignment 01 Demo
- > Assignment 02
- > Assignment 02 Demo
- > Quiz
- > Key Takeaways

Lesson 07 - Data Manipulation with Pandas

- > Introduction to Pandas
- > Knowledge Check
- > Understanding DataFrame
- > View and Select Data Demo
- > Missing Values
- > Data Operations
- > Knowledge Check
- > File Read and Write Support
- > Knowledge Check-Sequence it Right
- > Pandas Sql Operation
- > Assignment 01
- > Assignment 01 Demo
- > Assignment 02
- > Assignment 02 Demo
- > Quiz
- > Key Takeaways

Lesson 08 - Machine Learning with Scikit-Learn

- > Machine Learning Approach
- > Understand data sets and extract its features
- > Identifying problem type and learning model
- > How it Works
- > Train, test and optimizing the model
- > Supervised Learning Model Considerations
- > Knowledge Check
- > Scikit-Learn
- > Knowledge Check
- > Supervised Learning Models - Linear Regression
- > Supervised Learning Models - Logistic Regression
- > Unsupervised Learning Models
- > Pipeline
- > Model Persistence and Evaluation
- > Assignment 01
- > Knowledge Check
- > Assignment 01
- > Assignment 02
- > Assignment 02
- > Quiz
- > Key Takeaways

Lesson 09 - Natural Language Processing with Scikit Learn

- > NLP Overview
- > NLP Applications
- > Knowledge Check
- > NLP Libraries-Scikit
- > Extraction Considerations
- > Scikit Learn-Model Training and Grid Search
- > Assignment 01
- > Demo Assignment 01
- > Assignment 02
- > Demo Assignment 02
- > Quiz
- > Key Takeaway

Lesson 10 - Data Visualization in Python using matplotlib

- > Introduction to Data Visualization
- > Knowledge Check
- > Line Properties
- > (x,y) Plot and Subplots
- > Knowledge Check
- > Types of Plots
- > Assignment 01
- > Assignment 01 Demo
- > Assignment 02
- > Assignment 02 Demo
- > Quiz
- > Key Takeaways

Lesson 11 - Web Scraping with BeautifulSoup

- > Web Scraping and Parsing
- > Knowledge Check
- > Understanding and Searching the Tree
- > Navigating options
- > Demo3 Navigating a Tree
- > Knowledge Check
- > Modifying the Tree
- > Parsing and Printing the Document
- > Assignment 01
- > Assignment 01 Demo
- > Assignment 02
- > Assignment 02 demo
- > Quiz
- > Key takeaways

Lesson 12 - Python integration with Hadoop MapReduce and Spark

- > Why Big Data Solutions are Provided for Python0
- > Hadoop Core Components
- > Python Integration with HDFS using Hadoop Streaming
- > Demo 01 - Using Hadoop Streaming for Calculating Word Count
- > Knowledge Check
- > Python Integration with Spark using PySpark
- > Demo 02 - Using PySpark to Determine Word Count
- > Knowledge Check
- > Assignment 01
- > Assignment 01 Demo
- > Assignment 02
- > Assignment 02 Demo
- > Quiz
- > Key takeaways

Course End **Projects:**

The course includes four real-world, industry-based projects. Successful evaluation of one of the following projects is a part of the certification eligibility criteria:

Project 1: Products rating prediction for Amazon

Amazon, one of the leading US-based e-commerce companies, recommends products within the same category to customers based on their activity and reviews of similar products. Amazon would like to improve this recommendation engine by predicting ratings for the non-rated products and add them to recommendations accordingly.

Domain: E-commerce

Project 2: Demand Forecasting for Walmart

Predict accurate sales for 45 stores of Walmart, one of the US-based leading retail stores, considering the impact of promotional markdown events. Check if macroeconomic factors, such as CPI and unemployment rate, have an impact on sales.

Domain: Retail

Project 3: Improving Customer Experience for Comcast

Comcast, one of the largest US-based global telecommunication companies wants to improve customer experience by identifying and acting on problem areas that lower customer satisfaction. The company is also looking for key recommendations that can be implemented to deliver the best customer experience.

Domain: Telecom

Project 4: Attrition Analysis for IBM

IBM, one of the leading US-based IT companies, would like to identify the factors that influence attrition of employees. Based on the parameters identified, the company would also like to build a logistics regression model that can help predict if an employee will churn or not.

Domain: Workforce Analytics

Project 5: NYC 311 Service Request Analysis

Perform a service request data analysis of New York City 311 calls. You will focus on data wrangling techniques to understand patterns in the data and visualize the major complaint types.

Domain: Telecommunication

Project 6: MovieLens Dataset Analysis

The GroupLens Research Project is a research group in the Department of Computer Science and Engineering at the University of Minnesota. The researchers of this group are involved in several research projects in the fields of information filtering, collaborative filtering, and recommender systems. Here, we ask you to perform an analysis using the exploratory data analysis (EDA) technique for user datasets.

Domain: Engineering

Project 7: Stock Market Data Analysis

As a part of this project, you will import data using Yahoo DataReader from the following companies: Yahoo, Apple, Amazon, Microsoft, and Google. You will perform fundamental analytics, including plotting, closing price, plotting stock trade by volume, performing daily return analysis, and using pair plot to show the correlation between all of the stocks.

Domain: Stock Market

Project 8: Titanic Dataset Analysis

On April 15, 1912, the Titanic sank after colliding with an iceberg, killing 1502 out of 2224 passengers and crew. This tragedy shocked the world and led to better safety regulations for ships. Here, we ask you to perform an analysis using the EDA technique, in particular applying machine learning tools to predict which passengers survived the tragedy.

Domain: Hazard

Customer Reviews:



C Muthu Raman

Technical Project Leader - Mahindra Truck & Bus

Simplilearn facilitates a brilliant platform to acquire new and relevant skills with ease. Well laid-out course content and expert faculty ensure an excellent learning experience.



Mukesh Pandey

Technical Lead|Python|MS SQL Server|SSIS|Power BI|T-SQL

Simplilearn is an excellent platform for online learning. Their course curriculum is comprehensive and up to date. We get lifetime access to the recorded sessions in case we need to refresh our understanding. If you are looking to upskill, I suggest you sign up with Simplilearn. They offer classes in almost all disciplines.



Mukesh Pandey

Sr. Software engineer at Coupa Software

Incredible mentorship and amazing, unique lectures. Simplilearn provides a great way to learn with self-paced videos and recordings of online sessions. Thanks, Simplilearn, for providing quality education.



Surendaran Baskaran

I took the Data Science with Python course with Simplilearn. The instructor is knowledgeable and shares their skills and knowledge. My learning experience has been outstanding with Simplilearn. The practice labs and materials are helpful for better learning. Thank you, Simplilearn!

About Us:

Simplilearn is the world's #1 online Bootcamp provider that enables learners through rigorous and highly specialized training. We focus on emerging technologies and processes that are transforming the digital world, at a fraction of the cost and time as traditional approaches. Over one million professionals and 2000 corporate training organizations have harnessed our award-winning programs to achieve their career and business goals.

For more information, please visit our website:

<https://www.simplilearn.com/big-data-and-analytics/python-for-data-science-training>



simplilearn.com

Simplilearn is the world's #1 online bootcamp for digital economy skills training focused on helping people acquire the skills they need to thrive in the digital economy. Simplilearn provides outcome-based online training across technologies and applications in Data Science, AI and Machine Learning, Cloud Computing, Cyber Security, Digital Marketing, DevOps, Project Management, and other critical digital disciplines.

Through individual courses, comprehensive certification programs, and partnerships with world-renowned universities, Simplilearn provides millions of professionals and thousands of corporate training organizations with the work-ready skills they need to excel in their careers. Based in San Francisco, CA, and Bangalore, India, Simplilearn has helped more than one million professionals and 2,000 companies across 150 countries get trained, acquire certifications, and reach their business and career goals. With over 1,000 live classes each month, real-world projects, and more, professionals learn by doing at Simplilearn. Ongoing industry recognition for the company includes the 2020 Aegis Graham Bell Award for Innovation in EdTech and the 2020 Stevie® Gold Award for Customer Service Success.

India - United States - Singapore

© 2009-2020 - Simplilearn Solutions. All Rights Reserved.

The certification names are the trademarks of their respective owners.