ITA Associates

Applied Data Science with Python

Master applied data science with Python and unleash the power of data-driven insights

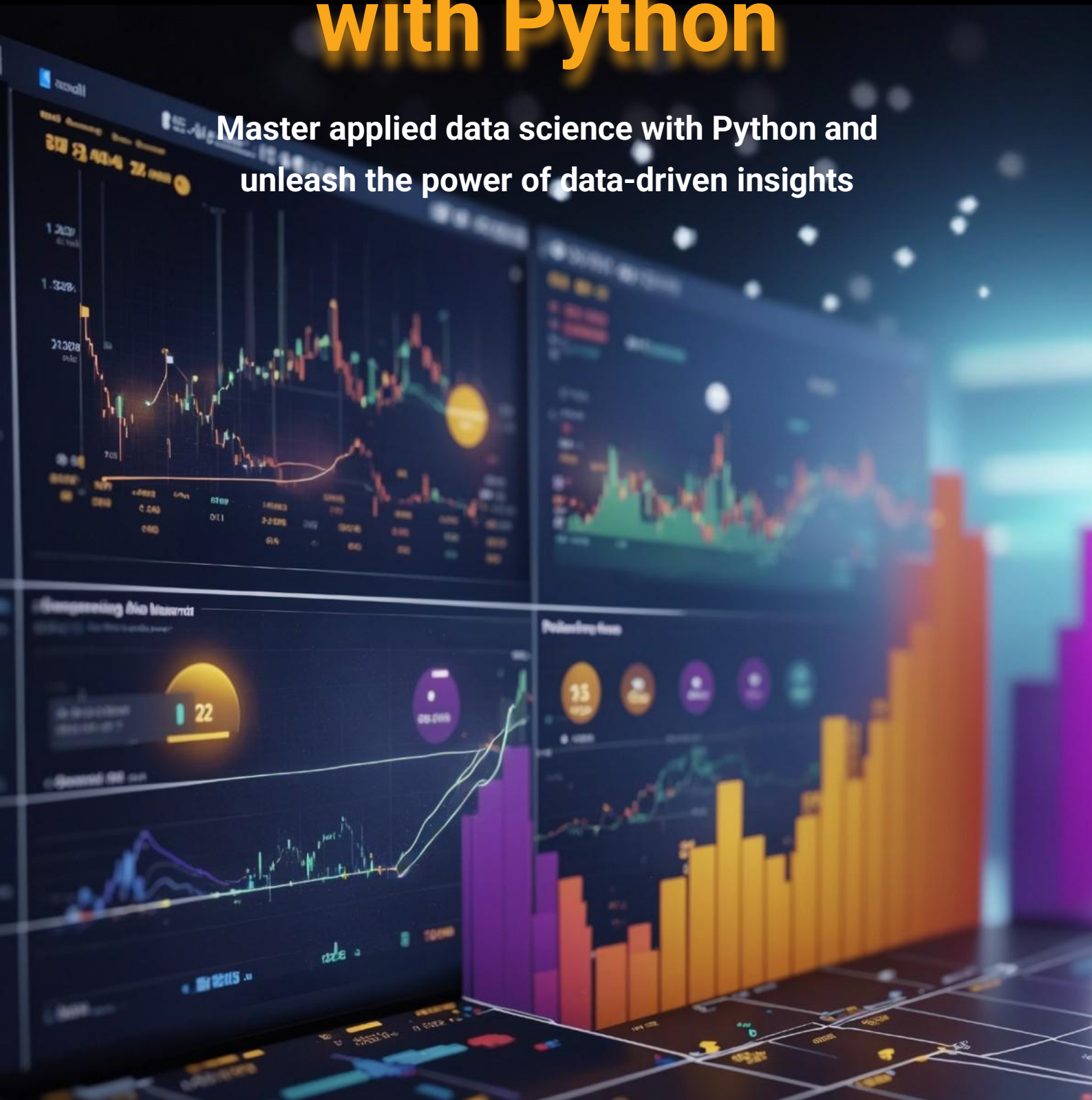


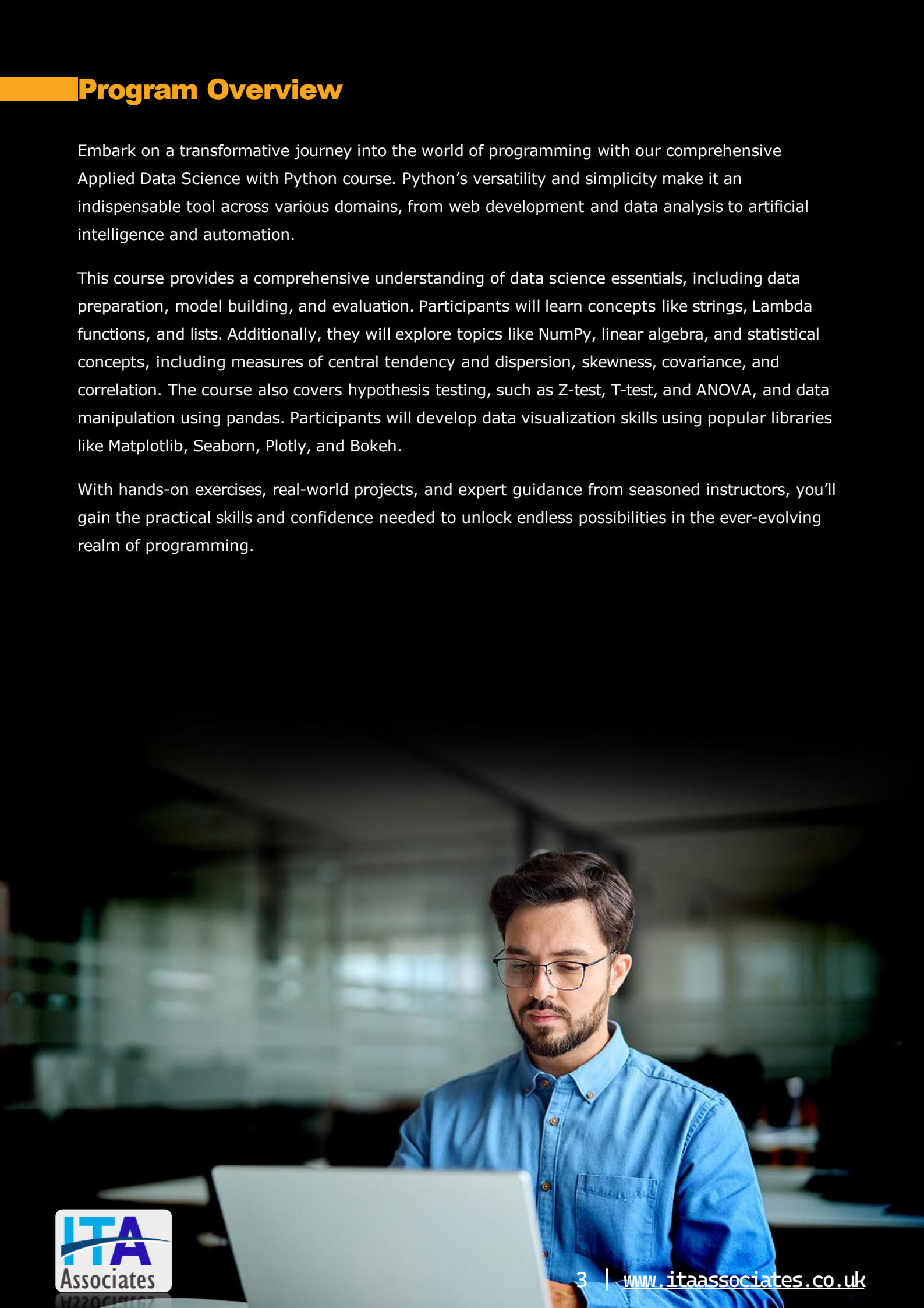
Table of Contents

Program Overview

Embark on a transformative journey into the world of programming with our comprehensive Applied Data Science with Python course. Python's versatility and simplicity make it an indispensable tool across various domains, from web development and data analysis to artificial intelligence and automation.

This course provides a comprehensive understanding of data science essentials, including data preparation, model building, and evaluation. Participants will learn concepts like strings, Lambda functions, and lists. Additionally, they will explore topics like NumPy, linear algebra, and statistical concepts, including measures of central tendency and dispersion, skewness, covariance, and correlation. The course also covers hypothesis testing, such as Z-test, T-test, and ANOVA, and data manipulation using pandas. Participants will develop data visualization skills using popular libraries like Matplotlib, Seaborn, Plotly, and Bokeh.

With hands-on exercises, real-world projects, and expert guidance from seasoned instructors, you'll gain the practical skills and confidence needed to unlock endless possibilities in the ever-evolving realm of programming.



Key Features of the Program

- Industry-based projects for experiential learning
- 40+ assisted practices and lesson-wise knowledge checks
- Interactive learning with Jupyter notebooks labs
- Lifetime access to self-paced learning content
- Practical skills and hands-on experience in applying Python to address data science challenges
- Dedicated live sessions by faculty of industry experts
- 60+ hours of blended learning

Delivery Mode

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

Who Should Enroll in this Program

This program caters to professionals from various industries and backgrounds, and the diversity of our students adds richness to class discussions and interactions. Exposure to any programming language, even at a beginner level, can expedite learning. However, Python's simplicity and readability make it accessible to beginners with little to no prior programming experience. With dedication, practice, and the right resources, anyone can grasp Python programming and unlock its vast potential in various fields, including web development, data analysis, artificial intelligence, and more. We have summarized the same into below 3 categories:

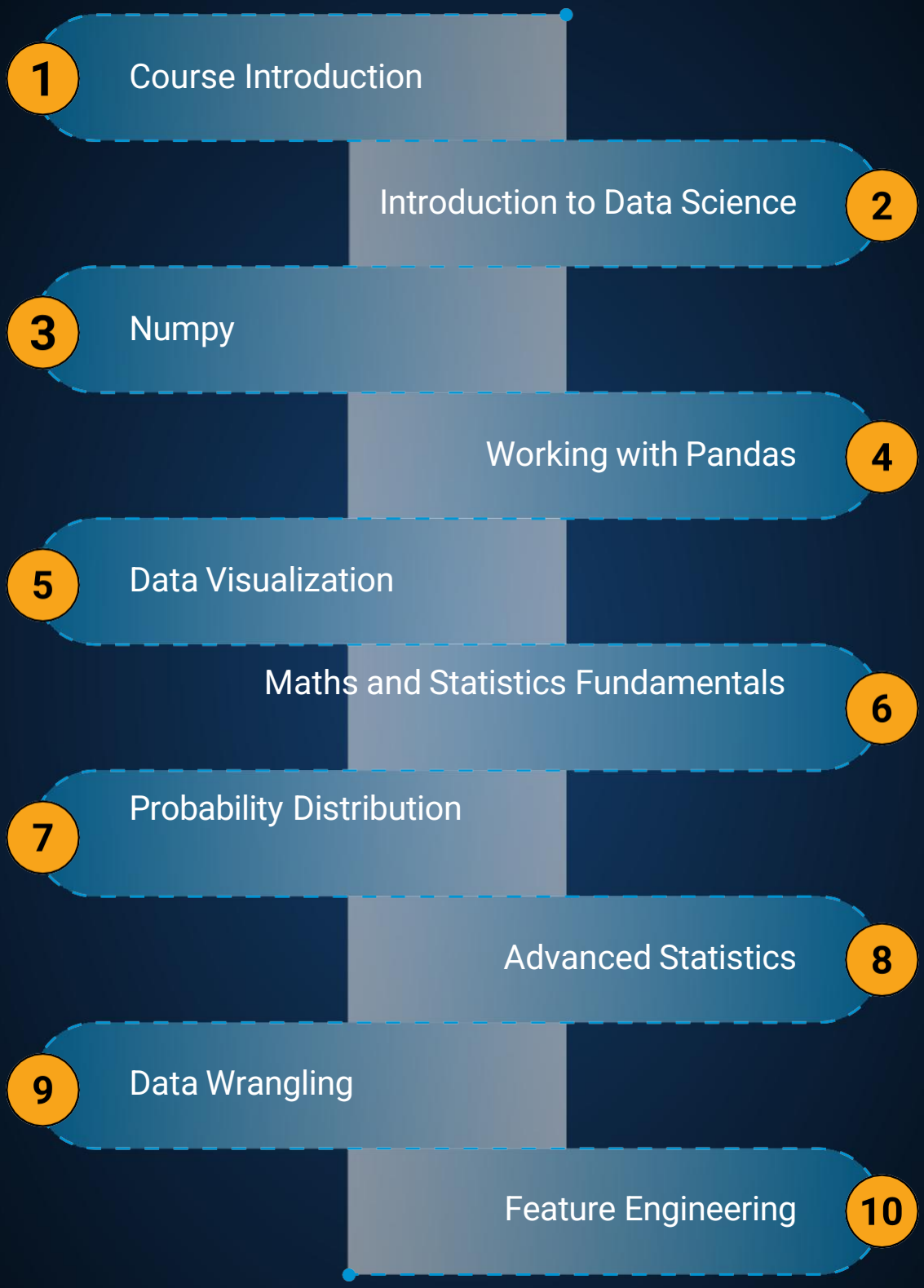
- Analytics professionals willing to work with Python
- Software and IT professionals interested in analytics
- Anyone with a genuine interest in data science

Key Learning Outcomes

This Applied Data Science with Python course will enable you to:

- Explain the fundamentals of data science and its practical applications.
- Explore the processes of data preparation, model building, and evaluation.
- Apply Python concepts like strings and comprehensively understand Lambda functions and lists.
- Develop a solid understanding of the fundamentals of NumPy.
- Explore array indexing and slicing techniques.
- Apply principles of linear algebra in data analysis.
- Understand the application of calculus in linear algebra.
- Calculate measures of central tendency and dispersion.
- Gain a clear understanding of statistical concepts such as skewness, covariance, and correlation.
- Describe the null hypothesis and alternative hypothesis.
- Examine different hypothesis tests, including Z-test and T-test.
- Understand the concept of ANOVA.
- Work with pandas' two primary data structures: Series and DataFrame.
- Utilize pandas for tasks such as data loading, indexing, reindexing, and data merging.
- Prepare, format, normalize, and standardize data using data binning techniques.
- Create visualizations with Matplotlib, Seaborn, Plotly, and Bokeh.

Learning Path



Learning Path

Lesson 1: Course Introduction

Get started with this program by understanding the course components and the topics covered. This will help you to be prepared for the upcoming sessions.

Topics covered

- Learning Path
- Program components

Lesson 2: Introduction to Data Science

Embark on a comprehensive journey through the data science process, starting with an introduction to its fundamental concepts. Delve into Python's role in data science, exploring essential packages and tools used for data manipulation, analysis, and visualization. By understanding the types of plots commonly used in data visualization, along with practical examples, you will acquire the skills necessary to effectively analyze and communicate insights from diverse datasets.

Topics covered

- Introduction
- Data Science Process
- Python for Data Science
- Python Packages for Data Science
- Types of Plots with Examples

Lesson 3: Numpy

In this module, you will comprehensively understand NumPy, a fundamental library for numerical computing in Python. Explore the array object and its attributes, mastering essential array functions, arithmetic operations, and statistical functions for efficient data manipulation and analysis. Additionally, you will delve into advanced topics such as string manipulation, array indexing, and slicing, equipping them with the necessary skills to work effectively with NumPy arrays in various data science applications.

Topics covered

- Fundamentals of NumPy
- NumPy: Array Object
- Attributes of NumPy Arrays
- NumPy Array Functions
- Arithmetic Operations using NumPy
- Statistical Function in Numpy
- String Function in Numpy
- NumPy Array Indexing
- NumPy Array Slicing

Lesson 4: Working with Pandas

Through these topics, you will gain a comprehensive understanding of pandas, a powerful library for data manipulation and analysis in Python. Explore fundamental data structures such as Series and DataFrame, mastering essential statistical operations and handling techniques for dates, times, categorical data, and text data. Additionally, delve into advanced functionalities, including iteration, sorting, and plotting with Pandas, equipping them with the skills needed to process and analyze diverse datasets efficiently.

Topics covered

- Fundamentals of pandas
- Data Structures
- Introduction to Series
- Introduction to pandas DataFrame
- Introduction to Statistical Operations in pandas
- Date and TimeDelta in pandas
- Date Handling in pandas
- Timedelta in pandas
- Categorical Data Handling
- Text Data in pandas
- Iteration
- Sorting
- Plotting with pandas

Lesson 5: Data Visualization

Through these topics, you will gain proficiency in data visualization using Matplotlib and Seaborn, two powerful libraries in Python. You will learn to create various types of plots, including line plots, scatter plots, bar charts, box plots, radar charts, area plots, polar plots, tree maps, and pie charts using Matplotlib. Additionally, using Seaborn, you will explore advanced visualization techniques such as 3D visualization, violin plots, pair plots, heatmaps, joint plots, swarm plots, and 3D graphs with multiple columns.

Topics covered

- Introduction
- Introduction to Matplotlib
- Line Plot
- Scatter Plot
- Bar Chart
- Box Plot
- Radar Chart (Spider chart)
- Area Plot
- Polar Plot
- Tree Map
- Pie Chart
- Matplotlib for 3D Visualization
- Introduction to Seaborn
- Plotting Graphs Using Seaborn
- Violin Plot
- Pair Plot
- Heatmap
- Joint Plot
- Swarm Plot
- Plotting 3D Graphs for Multiple Columns Using Seaborn

Lesson 6: Maths and Statistics Fundamentals

This comprehensively explores linear algebra, calculus, and statistics—the foundational pillars of data science. Grasp essential concepts such as scalars, vectors, matrices, and their operations, along with understanding norms, ranks, determinants, inverses, eigenvalues, and eigenvectors. Furthermore,delve into the application of calculus within linear algebra, establishing a solid mathematical framework for data analysis. Additionally, uncover the importance of statistics in data science, mastering various types of data and crucial statistical measures, including central tendency, dispersion, shape, covariance, and correlation. By mastering these concepts, you will be able to manipulate and analyze complex datasets, extract meaningful insights, and make informed decisions in data-driven environments.

Topics covered

- Linear Algebra
- Scalars and Vectors
- Vector Operation
- Norm of a Vector
- Matrix and Matrix Operations
- Rank of Matrix
- Determinant of Matrix
- Inverse of Matrix
- Eigenvalues and Eigenvectors
- Calculus in Linear Algebra
- Importance of Statistics for Data Science
- Types of Data
- Measures of Central Tendency
- Measures of Dispersion
- Measures of Shape
- Covariance and Correlation

Lesson 7: Probability Distribution

In this module, you will explore the core principles of probability theory essential for data science. Understand random variables, probability distributions (both discrete and continuous), and key concepts like probability density functions and cumulative distribution functions. Additionally, delve into crucial theorems like the Central Limit Theorem and Bayes' Theorem, along with estimation theory, equipping them to make informed statistical inferences and extract valuable insights from data.

Topics covered

- Probability and Its Importance
- Random Variable
- Probability Distribution
- Discrete Probability Distribution
- Continuous Probability Distribution
- Probability Density Function and Mass Function
- Cumulative Distribution Function
- Central Limit Theoram
- Bayes' Theorem
- Estimation Theory

Lesson 8: Advanced Statistics

In this module, you will master hypothesis testing methods essential for data analysis. You will understand concepts like null and alternative hypotheses, confidence intervals, margin of error, and confidence levels. Additionally, you will explore distributions, including the standard normal distribution (Z-distribution), t-distribution, and chi-square distribution, along with associated tests like the t-test, z-test, and f-test. By understanding these techniques, you can make statistically sound decisions, analyze variance, and draw reliable conclusions from data.

Topics covered

- Hypothesis Testing and Mechanism
- Null and Alternative Hypothesis
- Confidence Interval
- Margin of Error
- Confidence Levels
- Z-Distribution (Standard Normal Distribution)
- T-Distribution
- T-Test
- Z-Test
- Choosing Between T-test and Z-test
- P-Value
- Chi-square Distribution
- Analysis of Variance or ANOVA
- F-Distribution
- F-Test

Lesson 9: Data Wrangling

Through these topics, you will acquire essential data preparation and manipulation skills, crucial steps in the data analysis pipeline. Learn the importance of thorough data collection and inspection, techniques to handle duplicates, and strategies for cleaning messy datasets. Additionally, delve into data transformation, binning, and outlier detection methods to ensure data quality and reliability.

Topics covered

- Introduction
- Data Collection
- Data Inspection
- Dealing with Duplicates
- Data Cleaning
- Data Transformation
- Data Binning
- Handling Outliers
- Merging and Joining Data
- Aggregating Data
- Reshaping Data

Lesson 10: Feature Engineering

In this module, learners will explore the fundamentals of feature engineering, a critical aspect of data preprocessing in machine learning. They will learn various methods for transforming variables, including feature scaling, label encoding, one-hot encoding, and hashing, essential for preparing categorical and numerical data for model training. Additionally, learners will delve into grouping operations, enabling them to aggregate and summarize data efficiently. By mastering these techniques, learners will be equipped to engineer informative features from raw data, enhancing machine learning models' predictive power and performance.

Topics covered

- Introduction
- Feature Engineering Methods
- Transforming Variables
- Features Scaling
- Label Encoding
- One Hot Encoding
- Hashing
- Grouping Operations

Projects

Sales Analysis for Business Growth

Analyze the sales data of a retail clothing company and support management in formulating their sales and growth strategy.

Marketing Campaign Analysis

Perform exploratory data analysis and hypothesis testing to better understand the various factors contributing to customer acquisition.

Real Estate Data Visualization

Analyze the housing dataset using various types of plots to gain insights into the data.

Housing Price Analysis

Analyze housing data to uncover insights into house prices, comprehend the elements influencing them, and understand the impact of various house features on their price.

Customer Behaviour Analysis

Utilize various probability distributions to analyze customer behaviors and store performance metrics using a custom dataset.

Certificate

Upon completing this Python course, you will receive the certificates from Simplilearn. This certificate will testify to your skills as an expert in Python.

Customer Reviews

Prachi

Sr Manager - Digitalization & Innovation

The course was well structured. My instructor, Tim, was efficient and interactive. He ensured that all the queries got addressed without a miss—overall, it was an excellent learning experience.

Jyothish Chandran

Manager

A very well-experienced trainer, I enjoyed Tim's sessions. The way he teaches and progresses in each class is simply superb. Classes are blended with realistic and easily understandable examples. Thanks, Tim, for all your efforts to keep us informed well and for sharing your expertise

About Simplilearn

Simplilearn is the world's #1 online bootcamp provider, enabling learners around the globe with rigorous and highly specialized training offered in partnership with world-renowned universities and leading corporations. We focus on emerging technologies and skills transforming the global economy, such as artificial intelligence, data science, cloud computing, programming, and more. Our hands-on and immersive training includes live virtual classes, integrated labs and projects, 24x7 support, and a collaborative learning environment. Over two million professionals and 2000 corporate training organizations across 150 countries have harnessed our award-winning programs to achieve their career and business goals.

For more information, please visit our website: Applied Data Science with Python

United Kingdom

UBC 1st Floor

The Mille

1000 Great West Road

TW89DW

Phone: +44 7397 538 969

E-mail: info@itaassociates.co.uk

www.itaassociates.co.uk