Dynamic web training
The Ultimate Training Experience

**Phone:** 1300 888 724
**Email:** info@dynamicwebtraining.com.au

# Python for Data Analysis

## Python

## About The Course

Starting from Python programming basics, this 3-day course will build up your Python skills so that you will write code with confidence and clarity in the end. This course will teach you how to use Python as a scripting and modular programming language.

Starting from understanding the syntax, formatting, indenting, and documenting code with comments and docstrings, you will learn Python programming basics such as using data types, operators, and control flow. Furthermore, you will learn how to define and use your own functions and classes; create and use collection types such as tuples, strings, ranges, lists, dictionaries, and sets; handle problems using exceptions; manipulate and generate collections using comprehensions, iterators, and generators. Students will also learn how to investigate classes, instances, polymorphism, and inheritance in Python, managing files and other resources, and finally, unit testing and debugging.

**Duration:** 3 days

**Class size:** 10 students max

**Times:** 9:00am - 5:00pm

**Price:** *Refer to our website for current course and package pricing*

### After the course?

Each student will receive:

- Certificate of completion
- Training manual
- 12 months FREE email support
- FREE class re-sit (if necessary)

## Who Should Do This Course?

This course is ideal for data analysts, business analysts, and reporting professionals who want to use Python to work with real business data. It is also suitable for Power BI or SQL practitioners who want to extend their analytics capability using Python.

## Prerequisites

Participants must have completed the Python Beginner course or have equivalent knowledge of Python fundamentals including data types, control flow, functions, and basic scripting. Familiarity with tabular data concepts such as rows, columns, and basic filtering will be beneficial but is not required. No prior experience with data analytics libraries is assumed.

# Content

## Unit 1: Python Setup and Getting Started

- Understand Why Python Is Used for Data Analytics
- Compare Python with SQL Server and Power BI
- Install Essential Python Libraries for Data Analytics
- Understand What a Python Script Is and How It Works

## Unit 2: Overview of the Data Analysis Module Library

- Identify the Core Python Modules Used in Data Analytics
- Understand the Role of Each Module in the Analytics Workflow
- Recognise the Basic Syntax for Importing and Using Modules
- Import and Verify All Core Analytics Modules

**Phone:** 1300 888 724
**Email:** info@dynamicwebtraining.com.au

## Unit 3: Introduction to IPython

- Start and Use IPython for Interactive Coding
- Use Magic Commands Including %time, %timeit, %who, and %history
- Run Python Scripts from Within IPython Using %run
- Reset the Session and Suppress Output
- Get Help and Documentation Using the ? Operator

## Unit 4: Introduction to NumPy

- Create and Manipulate NumPy Arrays
- Reshape, Index, and Slice Arrays
- Perform Vectorized Math and Broadcasting
- Use Aggregation Functions Including sum, mean, and axis
- Generate Random Numbers for Sampling and Testing
- Compare NumPy Arrays to Python Lists

## Unit 5: Introduction to Pandas

- Create and Manipulate Series and Data-Frames
- Select, Filter, Sort, and Reindex Data
- Apply Functions and Mapping to Columns
- Perform Basic Descriptive Statistics and Summarisation
- Inspect DataFrames Using .info() and .shape

## Unit 6: Data Loading and File Formats

- Read and Write Data in CSV and JSON Formats
- Read and Write Data in Excel Format
- Load the Sales Dataset from sales_data.xlsx for Analysis
- Retrieve Data from Public URLs and Web Tables
- Retrieve Data from Public APIs
- Understand How to Read Database Data

## Unit 7: Data Cleaning and Preparation

- Remove and Add Columns and Rows to Focus Analysis
- Split Columns and Add Calculated Columns Including Revenue
- Rename Columns and Set Data Types for Consistency
- Fix Spelling Mistakes, Abbreviations, and For-matting Errors
- Apply Forward Fill to Resolve Blank Cells
- Combine Multiple Cleaning Steps Using Method Chaining

## Unit 8: Regular Expressions (Regex) in Pandas

- Understand What Regular Expressions Are and Why They Are Useful
- Use Regex Patterns to Match, Extract, and Replace Text in Pandas
- Apply Regex with Common Pandas String Functions
- Recognise and Use Common Regex Wildcards and Patterns

## Unit 9: Analytical Functions and Grouping

- Group and Aggregate Data Using groupby and agg
- Apply Multiple and Custom Aggregation Functions
- Understand the Difference Between agg and transform
- Iterate Over Groups and Select Subsets
- Filter Groups Using Advanced Conditions
- Group by Multiple Columns for Deeper Insights
- Rank Values and Find Top N Within Groups

## Unit 10: Data Modelling

- Distinguish Between Analytical and Transactional Queries
- Identify Dimensions and Measures in Your Data
- Understand the Difference Between Measures and Calculated Columns
- Recognise Star and Snowflake Schema Patterns
- Apply Best Practices for Analytical Data Models

## Unit 11: Creating a Dates Dimension Table

- Understand the Purpose and Benefits of a Dates Dimension Table
- Create a Comprehensive Dates Table in Pandas
- Add Financial Year and Financial Month Columns
- Join the Dates Table to Sales Data for Time-Based Analysis
- Aggregate Revenue by Financial Year

## Unit 12: Joins and Combining Data

- Use Inner, Left, Right, and Outer Joins to Combine Tables
- Debug Joins Using the indicator Parameter
- Build a Star Schema Model by Joining Fact and Dimension Tables

**Phone:** 1300 888 724
**Email:** info@dynamicwebtraining.com.au

- Add Calculated Columns Including Total Cost, Status, and Profit
- Concatenate DataFrames Vertically and Horizontally
- Reshape Data Using Pivot Tables, Melt, Stack, and Unstack

## Unit 13: Advanced Analytics

- Calculate Running Totals and Cumulative Metrics
- Apply Rolling and Expanding Window Functions
- Understand the Difference Between Rolling and Expanding Windows
- Rank Values Within Groups Using dense and first Methods
- Calculate Lag, Day-to-Day Change, and Percentage Change
- Normalise Values and Calculate Z-Scores Within Groups
- Compute Percentile Ranks Within Groups

## Unit 14: Data Visualisation with Matplotlib

- Choose the Right Chart Type for Your Data and Analytical Question
- Create Line Charts, Bar Charts, Scatter Plots, and Histograms
- Understand the Difference Between Bar Charts and Histograms
- Customise Plots with Titles, Labels, Colours, and Styles
- Create Subplots to Display Multiple Charts Together
- Save Charts to File for Use in Reports and Presentations

## Unit 15: Data Visualisation with Seaborn

- Understand What Seaborn Is and How It Builds on Matplotlib
- Apply Seaborn Themes and Styles to Produce Polished Charts
- Create Categorical Plots Including Bar, Box, and Count Plots
- Create Distribution Plots Including Histograms with KDE and KDE Plots
- Create Regression Plots with Trend Lines
- Create a Correlation Heatmap to Explore Relationships Between Variables
- Create a Pair Plot to Visualise All Numeric Relationships in One View
- Create Facet Grids to Compare Distributions Across Multiple Categories

## Looking for course dates?

**To view a full list of course dates, please visit our website at www.dynamicwebtraining.com.au**
**Alternatively please contact our office on 1300 888 724**