

# Python Functions for Data Science

GoSkills online course syllabus

**Skill level**

Beginner

**Lessons**

25

**Pre-requisites**

No prior experience needed

**Video duration**

1h 35m

**Estimated study time**

1h 35m

**Instructor**

Lavanya Vijayan

## Introduction

---

- 1 Use Python like a data scientist**  
Many people learn Python without knowing which functions matter most in data science.
- 2 Getting started with Python**  
Before diving into the hands-on lessons, it's essential to have your workspace ready for real-time practice.

## Core Python Essentials

---

- 3 Inspect data for validation**  
Inspecting your data is essential for data validation and debugging.
- 4 Handle magnitudes and precision in data**  
Properly handling magnitudes and controlling numeric precision is essential for interpreting data and communicating results.
- 5 Aggregate data with basic functions**  
Computing basic summary statistics is the first step in exploratory data analysis.
- 6 Sort, filter, and transform your data**  
Data in sequences often needs to be ordered or pruned before advanced analysis.

## NumPy and SciPy Fundamentals

---

- 7 Create NumPy arrays in Python**  
NumPy arrays are the performance-optimized backbone of nearly every data science workflow in Python.

**8** **Index and slice NumPy arrays**  
Accurately accessing and extracting data from arrays is crucial for data analysis.

**9** **Reshape NumPy arrays**  
Inspecting dimensions and reorganizing arrays is vital in preparing data for computations and machine learning models.

**10** **Transform and scale NumPy arrays**  
Fast, element-wise numeric transformations are crucial for preprocessing and scaling large datasets.

**11** **Extract key values with NumPy**  
Summarizing data with key statistics is fundamental for understanding distributions and informing modeling decisions.

**12** **Solve matrix-based problems with SciPy**  
Matrix computations, like checking invertibility with determinants and finding inverses, are foundational for tasks such as solving regression systems, performing dimensionality reduction, and transforming feature spaces in data science.

**13** **Run statistical functions with SciPy**  
Working with probability distributions and hypothesis testing are essential for drawing inferences and validating models.

## **Pandas for Data Manipulation**

---

**14** **Create pandas series and dataframes**  
The pandas library makes tabular and time series data easy to load, explore, and manipulate.

**15** **Extract data subsets from pandas objects**  
Being able to extract the exact rows and columns you need is fundamental to data cleaning and analysis.

**16** **Modify pandas objects**  
Updating data in-place keeps your pipeline tidy and prevents errors later on.

**17** **Combine data from pandas objects**  
Combining data from multiple tables or different data sources is essential for creating robust datasets that drive deeper insights.

**18** **Group data from pandas objects**  
Aggregating data by categories helps reveal group-level patterns, which is important in exploratory data analysis.

- 19** Transform data with pandas apply()  
Custom data transformations are essential when built-in methods fall short of specific cleaning or feature engineering needs.

## Visualization Essentials

---

- 20** Create line and scatter plots  
Visualizing quantitative data is essential for spotting trends and relationships between numeric variables in your datasets.
- 21** Display categorical distributions  
Visualizing categorical distributions is essential for comparing groups and understanding proportions in your data.
- 22** Explore numerical distributions  
Understanding the shape, spread, and density of numerical data is essential for exploratory analysis.
- 23** Visualize pairwise relationships  
Mapping pairwise relationships helps you detect correlations, clusters, and anomalies across many variables at a glance.
- 24** Organize your visualizations  
Arranging multiple visualizations in a single figure is essential for comparing insights and crafting coherent dashboards.

## Conclusion

---

- 25** Apply functions to data science  
Congratulations on completing the course!

[Go to GoSkills.com](https://www.goskills.com)