

Python Pandas

This Python Pandas training course will teach you all about using Pandas for data analysis, from the beginning to creating one- and two-dimensional indexed data structures, indexing and slicing-and-dicing that data to derive results, loading data from local and Internet-based resources, and finally creating effective visualizations to form quick insights.

Prerequisites / Further Training

Have a look at our [Python Bootcamp](#)

Intended Audience

- Analysts who wants to more on data analysis and reporting
- Programmers who wants to performing data exploration and analysis on Python using pandas

DAY 1

Pandas Foundations

- Introduction
- Dissecting the anatomy of a DataFrame
- Accessing the main DataFrame components
- Understanding data types
- Selecting a single column of data as a Series
- Calling Series methods
- Working with operators on a Series
- Chaining Series methods together
- Making the index meaningful
- Renaming row and column names
- Creating and deleting columns

Essential DataFrame Operations

- Selecting multiple DataFrame columns

- Selecting columns with methods
- Ordering column names sensibly
- Operating on the entire DataFrame
- Chaining DataFrame methods together
- Working with operators on a DataFrame
- Comparing missing values
- Transposing the direction of a DataFrame operation
- Determining college campus diversity

DAY 2

Beginning Data Analysis

- Developing a data analysis routine
- Reducing memory by changing data types
- Selecting the smallest of the largest
- Selecting the largest of each group by sorting
- Replicating nlargest with sort_values
- Calculating a trailing stop order price

Selecting Subsets of Data

- Selecting Series data
- Selecting DataFrame rows
- Selecting DataFrame rows and columns simultaneously
- Selecting data with both integers and labels
- Speeding up scalar selection
- Slicing rows lazily
- Slicing lexicographically

DAY 3

Boolean Indexing

- Calculating boolean statistics
- Constructing multiple boolean conditions
- Filtering with boolean indexing
- Replicating boolean indexing with index selection
- Selecting with unique and sorted indexes
- Gaining perspective on stock prices

- Translating SQL WHERE clauses
- Determining the normality of stock market returns
- Improving readability of boolean indexing with the query method
- Preserving Series with the where method
- Masking DataFrame rows
- Selecting with booleans, integer location, and labels

Index Alignment

- Examining the Index object
- Producing Cartesian products
- Exploding indexes
- Filling values with unequal indexes
- Appending columns from different DataFrames
- Highlighting the maximum value from each column
- Replicating idxmax with method chaining
- Finding the most common maximum

DAY 4

Grouping for Aggregation, Filtration, and Transformation

- Defining an aggregation
- Grouping and aggregating with multiple columns and functions
- Removing the MultiIndex after grouping
- Customizing an aggregation function
- Customizing aggregating functions with *args and **kwargs
- Examining the groupby object
- Filtering for states with a minority majority
- Transforming through a weight loss bet
- Calculating weighted mean SAT scores per state with apply
- Grouping by continuous variables
- Counting the total number of flights between cities
- Finding the longest streak of on-time flights

Restructuring Data into a Tidy Form

- Tidying variable values as column names with stack
- Tidying variable values as column names with melt
- Stacking multiple groups of variables simultaneously
- Inverting stacked data
- Unstacking after a groupby aggregation
- Replicating pivot_table with a groupby aggregation
- Renaming axis levels for easy reshaping
- Tidying when multiple variables are stored as column values
- Tidying when two or more values are stored in the same cell
- Tidying when variables are stored in column names and values
- Tidying when multiple observational units are stored in the same table

DAY 5

Combining Pandas Objects

- Appending new rows to DataFrames
- Concatenating multiple DataFrames together
- Comparing President Trump's and Obama's approval ratings
- Understanding the differences between concat, join, and merge
- Connecting to SQL databases

Time Series Analysis

- Understanding the difference between Python and pandas date tools
- Slicing time series intelligently
- Using methods that only work with a DatetimeIndex
- Counting the number of weekly crimes
- Aggregating weekly crime and traffic accidents separately
- Measuring crime by weekday and year

- Grouping with anonymous functions with a DatetimeIndex
- Grouping by a Timestamp and another column
- Finding the last time crime was 20% lower with merge_asof

Visualization with Matplotlib, Pandas, and Seaborn

- Getting started with matplotlib
- Object-oriented guide to matplotlib
- Visualizing data with matplotlib
- Plotting basics with pandas
- Visualizing the flights dataset
- Stacking area charts to discover emerging trends
- Understanding the differences between seaborn and pandas
- Doing multivariate analysis with seaborn Grids
- Uncovering Simpson's paradox in the diamonds dataset with Seaborn

Duration and pricing

In [Price Group A](#)

Certificate

Please read [here about our certificates](#)

Schedule

On the calendar on this page below.

If your browser doesn't display the calendar below, please click on [this link](#) or try using [Google Chrome](#), alternatively please enquire via our '[Contact Us](#)' page.

Bookings

Please click [click here](#) or send us an [email](#).

Questions

Please [email us](#)