

Hands-On Data Analysis with Panda - TTPS4878

Quick Start to Using the Pandas Library to Reshape, Clean, Aggregate, Analyze & Visualize Your Data

Duration: 3 Days

Skill Level: Introductory

Available Format: Instructor-Led Online; Instructor-Led, Onsite In Person ; Blended; On Public Schedule

Data analysis has become a necessary skill in a variety of domains where knowing how to work with data and extract insights can generate significant value. Geared for data team members with incoming Python scripting experience, **Hands-On Data Analysis with Pandas** will show you how to analyze your data, get started with machine learning, and work effectively with Python libraries often used for data science, such as pandas, NumPy, matplotlib, seaborn, and scikit-learn.

What You'll Learn

Overview

Data analysis has become a necessary skill in a variety of domains where knowing how to work with data and extract insights can generate significant value. Geared for data team members with incoming Python scripting experience, **Hands-On Data Analysis with Pandas** will show you how to analyze your data, get started with machine learning, and work effectively with Python libraries often used for data science, such as pandas, NumPy, matplotlib, seaborn, and scikit-learn.

Using real-world datasets, you will learn how to use the powerful pandas library to perform data wrangling to reshape, clean, and aggregate your data. Then, you will be able to conduct exploratory data analysis by calculating summary statistics and visualizing the data to find patterns. In the concluding lessons, you will explore some applications of anomaly detection, regression, clustering, and classification using scikit-

learn to make predictions based on past data. You'll leave the course armed with the skills required to use pandas to ensure the veracity of their data, visualize it for effective decision-making, and reliably reproduce analyses across multiple datasets.

A core component of our **Python Journey** series, **Hands-on Data Analysis with Pandas**, is a highly rated training course that has helped hundreds of students get up and running with solid, practical Python skills, enabling them to leverage Python for day to day analytics on the job.

Objectives

This course is approximately **50% hands-on**, combining expert lecture, real-world demonstrations and group discussions with machine-based practical labs and exercises. Our engaging instructors and mentors are highly experienced practitioners who bring years of current "on-the-job" experience into every classroom.

Working in a hands-on learning environment, guided by our expert team, attendees will learn to:

- Understand how data analysts and scientists gather and analyze data
- Perform data analysis and data wrangling using Python
- Combine, group, and aggregate data from multiple sources
- Create data visualizations with pandas, matplotlib, and seaborn
- Apply machine learning (ML) algorithms to identify patterns and make predictions
- Use Python data science libraries to analyze real-world datasets
- Use pandas to solve common data representation and analysis problems
- Build Python scripts, modules, and packages for reusable analysis code
- Perform efficient data analysis and manipulation tasks using pandas
- Apply pandas to different real-world domains with the help of step-by-step demonstrations
- Get accustomed to using pandas as an effective data exploration tool.
- Optional: Master the integration of AI analytics to transform and elevate data analysis, encompassing predictive modeling and decision-making enhancement techniques.
- Optional: Acquire practical skills in applying machine learning and deep learning to automate data wrangling and generate dynamic, predictive data visualizations.

Need different skills or topics? If your team requires different topics or tools, additional skills or custom approach, this course may be further adjusted to accommodate. We offer additional python, data science, AI / machine learning and other related topics that may be blended with this course for a track that best suits your needs. Our team will collaborate with you to understand your needs and will target the course to focus on your specific learning objectives and goals.

Audience

This course is geared for Python-experienced attendees who wish to be equipped with the skills you need to use pandas to ensure the veracity of your data, visualize it for effective decision-making, and reliably reproduce analyses across multiple datasets. Ideal attendees include experienced data analysts, developers, engineers or anyone tasked with utilizing Python for data analytics or eventual machine learning tasks. **Attending students are required to have a background in basic Python for data science.**

Pre-Requisites

Attending students are required to have a background in basic Python for data science.

Take Before: Students should have incoming practical skills aligned with those in the course(s) below, or should have attended the following course(s) as a pre-requisite:

- **TTDS6000** Understanding Data Science | A Technical Overview - 1 day (helpful but not required)
- **TTPS4800** Introduction to Python Programming Basics (3 days)

TTPS4800	Introduction to Python Programming Basics
TTPS4820	Mastering Python Programming Boot Camp
TTPS4824	Python Essentials for Networking & Systems Administration
TTPS4872	Quick Start to Python for Data Science Primer: A Hands-on Technical Overview
TTPS4873	Fast Track to Python for Data Science and/or Machine Learning
TTPS4874	Applied Python for Data Science and Engineering

Agenda

Please note that this list of topics is based on our standard course offering, evolved from typical industry uses and trends. We will work with you to tune this course and level of coverage to target the skills you need most. Course agenda, topics and labs are subject to adjust during live delivery in response to student skill level, interests and participation.

Introduction to Data Analysis

- Fundamentals of data analysis
- Statistical foundations
- Setting up a virtual environment

Working with Pandas DataFrames

- Pandas data structures
- Bringing data into a pandas DataFrame
- Inspecting a DataFrame object
- Grabbing subsets of the data
- Adding and removing data

Data Wrangling with Pandas

- What is data wrangling?
- Collecting temperature data
- Cleaning up the data
- Restructuring the data
- Handling duplicate, missing, or invalid data

Aggregating Pandas DataFrames

- Database-style operations on DataFrames
- DataFrame operations
- Aggregations with pandas and numpy
- Time series

Visualizing Data with Pandas and Matplotlib

- An introduction to matplotlib
- Plotting with pandas
- The pandas.plotting subpackage

Plotting with Seaborn and Customization Techniques

- Utilizing seaborn for advanced plotting
- Formatting
- Customizing visualizations

Financial Analysis - Bitcoin and the Stock Market

- Building a Python package
- Data extraction with pandas
- Exploratory data analysis
- Technical analysis of financial instruments
- Modeling performance

Rule-Based Anomaly Detection

- Simulating login attempts
- Exploratory data analysis
- Rule-based anomaly detection

Getting Started with Machine Learning in Python

- Learning the lingo
- Exploratory data analysis
- Preprocessing data
- Clustering
- Regression
- Classification

Making Better Predictions - Optimizing Models

- Hyperparameter tuning with grid search
- Feature engineering
- Ensemble methods
- Inspecting classification prediction confidence
- Addressing class imbalance
- Regularization

Machine Learning Anomaly Detection

- Exploring the data
- Unsupervised methods
- Supervised methods
- Online learning

Optional: AI-Driven Enhancements to Data Analytics

Incorporating AI in Data Analysis Enhancements

- Exploring AI-driven analytics platforms
- Integrating predictive analytics into data projects
- Enhancing decision-making with prescriptive analytics
- Building a simple AI model to derive insights from data

Advancing Data Wrangling with Machine Learning

- Implementing ML algorithms for data cleaning and preprocessing
- Utilizing natural language processing (NLP) for data categorization and tagging
- Case study: Automating data wrangling in industry scenarios
- Optional Lab: Using ML tools for handling complex data transformations

Deep Learning for Predictive Data Visualizations

- Fundamentals of neural networks in visualization
- Creating advanced visuals with deep learning models
- Interactive data stories using AI-powered tools
- Optional Lab: Designing a predictive dashboard using AI visual tools

The Road Ahead

- Data resources
- Practicing working with data
- Python practice

Follow On Courses

TTAI2360	Applied AI: Building Recommendation Systems with Python
TTAI2361	Applied AI : Quick Start to Building AI-Driven, Intelligent Web Applications
TTML5506-P	Machine Learning Essentials with Python
TTML5511	Machine Learning Boot Camp / Deep Dive Skills Workshop

Related Courses

TTPS4879	Hands-On Predictive Analytics with Python
----------	---

All applicable course software, digital courseware files or course notes, labs, data sets and solutions, live coaching support channels, CodeCoach.AI anytime tutor access, and rich extended learning and post training resources are provided for you in our “easy access, single source, no install required” online **Learning Experience Platform (LXP)**, remote lab and content environment. Access periods vary by course. We’ll collaborate with you to ensure your team is set up and ready to go well in advance of the class. Please inquire about set up details and options for your specific course of interest.

For More Information

Please [contact us](#) or call 844-475-4559 toll free for more information about our training services (instructor-led, self-paced or blended), coaching and mentoring services, public

course enrollment or questions, partner programs, courseware licensing options and more.