Python and Pandas for Data Manipulation

This Python and Pandas for Data Manipulation training teaches how to efficiently manipulate and analyze data from multiple sources. Learn to clean, filter, sort, and combine data to extract insights. Perfect for onboarding data professionals or as a reference for IT teams. Topics include data cleaning, statistics, and visualization for smarter decision-making.

CBT Nuggets course material →

STUDY PLAN

WEEK 1

Installation and Setup — 152 min.

What is pandas?	8
What is Jupyter Notebook?	5
Anaconda Installation	10
Conda Environments	7
Challenge	1
Challenge Question Answers (optional)	14
Supplemental Files	1

Jupyter Notebook

Introduction	2
Brief History	6
User Interface	15
Data Types Review	6
Cell Types	7
Shortcuts	4
Code Challenge	1

Series Introduction

Introduction	2
Create a Series from a list	8
Create a Series from a dictionary	4
Read CSV files	12
Read Excel files	7
Head and tail functions	5
Series attributes	4
Series methods	4
Supplemental Files	1

Series Attributes and Methods

Introduction	1
Parameter and arguments	9
Sorting values	8

WEEK 2

153 min.

Series attributes	8
Series Methods	8
Inplace Mutation	7
Sorting Series Indices	7
Challenge	1
Supplemental Files	1

Series Basics

Introduction	1
The in keyword	7
Extract by position	7
Extract by label	16
The get() method	9
Math methods	7
The idxmin() and idxmax() methods	4
Unique values	5
The apply() method	6
Challenge	1
Supplemental Files	1

DataFrame Introduction

DataFrame Introduction	11
Series shared attributes	7
Shared methods	8
Extracting columns	6
Extracting two or more columns	5
Adding columns	7
Broadcasting Operations	7
DataFrames value_counts( )	6

WEEK 3

158 min.

Challenge	1

DataFrame Cleaning

Introduction	1
Handling null values	7
Drop null values	14
Impute missing values	7
Value counts for DataFrames	5
Detect null and not null values	6
Challenge	1

DataFrame Sorting

Introduction	1
Changing data types	16
Sorting values	17
Sort by indices	5
Ranking a Series	14
Challenge	1

Filtering Data

Introduction	2
Optimization	14
Conditional Filtering	20
Filtering with AND and OR	18
Inclusion method	8

WEEK 4

164 min.

Challenge	1

Filtering Duplicates

Introduction	1
Checking for duplicates	10
Drop duplicates	9
Unique values	6
Inclusion with between()	11
Challenge	1
Solution video	7

Extracting Values

Introduction	1
Setting and resetting indices	14
Extraction with loc	14
Extraction with iloc	13
Setting new values	4
Set multiple values	7
Challenge	1

Extraction Methods

Introduction	1
The drop method	9
Returning smallest and largest values	10
The where method	8
The query method	10
The copy method	10
Challenge	1

Text Data Basics

Introduction	1
Manipulating text data	14

WEEK 5

165 min.

String methods	14
The replace string method	15
Filtering string methods	11
Challenge	1

Splitting and Stripping Text Data

Introduction	2
Strip strings	21
Column and index methods	7
Splitting strings	8
More splitting	8
Challenge	1

Grouping Methods

Introduction	1

Grouping	10
group_by operations	11
get_group method	9
The group_by methods	13
Challenge	1

Combining DataFrames

Introduction	3
Combining DataFrames	8
Concatenation	21

WEEK 6

161 min.

Inner joins	14
Outer joins	11
Challenge	1

Time Series Data

Introduction	1
Python Datetime	9
Pandas Timestamp	8
DatetimeIndex	6
The to_datetime method	11
Date Ranges Introduction	8
Challenge	1

Date Ranges

Introduction	1
Date ranges part 1	28
Date ranges part 2	15
Date ranges part 3	8
The dt accessor	12
Challenge	1
Supplemental Files	1

DataReader

Introduction & setup	9
Reading cryptocurrency data	16

WEEK 7

91 min.

Selecting Datetime rows	12
Timestamp attributes & methods	13
Challenge	1
Supplemental Files	1

Visualization

Introduction	2
Matplotlib & PyPlot	9
Visualizing cryptocurrencies	23
Customizing visualizations	15
Creating charts	13
Challenge	1
Supplemental Files	1

Last updated 12.3.2024