Python for Data Analysis, 3rd edition

Data Wrangling with pandas, NumPy, and Jupyter

Stephen CUI[1]

January 4, 2022

[1]cuixuanStephen@gmail.com

Contents

Chapter 1

Data Loading, Storage, and File Formats

1.1   Reading and Writing Data in Text Format

I'll give an overview of the mechanics of these functions, which are meant to convert text data into a DataFrame. The optional arguments for these functions may fall into a few categories:

Indexing Can treat one or more columns as the returned DataFrame, and whether to get column names from the file, arguments you provide, or not at all.

Type inference and data conversion Includes the user-defined value conversions and custom list of missing value markers.

Date and time parsing Includes a combining capability, including combining date and time information spread over multiple columns into a single column in the result.

Iterating Support for iterating over chunks of very large files.

Unclean data issues Includes skipping rows or a footer, comments, or other minor things like numeric data with thousands separated by commas.

Chapter 2

Data Cleaning and Preparation

Chapter 3

A Brief matplotlib API Primer

Chapter 4

Appendix A