Dockerfile	fix: add pytest-qt deps to dockerfi…	3 weeks ago
LICENSE	Update copyright year (#56960)	last week
MANIFEST.in	CI: Build wheel from sdist (#53087)	8 months ago
README.md	DOC: Update build instructions in …	5 months ago
codecov.yml	CI/DOC: replace master -> main b…	3 years ago
environment.yml	TYP: some return types from ruff (…	last month
generate_pxi.py	TYP: simple return types from ruff …	last month
generate_version.py	TYP: simple return types from ruff …	last month
meson.build	Compiled pandas with -Wextra (#…	last month
pyproject.toml	STY: Enable B904 (#56941)	last week
pyright_reportGeneralTypeI…	TYP: Update pyright (#56892)	2 weeks ago
requirements-dev.txt	TYP: some return types from ruff (…	last month
setup.py	TYP: simple return types from ruff …	last month

pandas: powerful Python data analysis toolkit

Testing	Unit Tests passing codecov 83%
Package	pypi v2.2.0 PyPI downloads 152M/month Anaconda.org 2.2.0 Conda downloads 47M
Meta	powered by NumFOCUS DOI 10.5281/zenodo.3509134 license BSD join Slack information

What is it?

pandas is a Python package that provides fast, flexible, and expressive data structures designed to make working with "relational" or "labeled" data both easy and intuitive. It aims to be the fundamental high-level building block for doing practical, **real world** data analysis in Python. Additionally, it has the broader goal of becoming **the most powerful and flexible open source data analysis / manipulation tool available in any language**. It is already well on its way towards this goal.

Table of Contents

Main Features

Here are just a few of the things that pandas does well:

- Easy handling of **missing data** (represented as `NaN`, `NA`, or `NaT`) in floating point as well as non-floating point data
- Size mutability: columns can be **inserted and deleted** from DataFrame and higher dimensional objects
- Automatic and explicit **data alignment**: objects can be explicitly aligned to a set of labels, or the user can simply ignore the labels and let `Series`, `DataFrame`, etc. automatically align the data for you in computations
- Powerful, flexible **group by** functionality to perform split-apply-combine operations on data sets, for both aggregating and transforming data
- Make it **easy to convert** ragged, differently-indexed data in other Python and NumPy data structures into DataFrame objects
- Intelligent label-based **slicing**, **fancy indexing**, and **subsetting** of large data sets
- Intuitive **merging** and **joining** data sets
- Flexible **reshaping** and **pivoting** of data sets
- **Hierarchical** labeling of axes (possible to have multiple labels per tick)
- Robust IO tools for loading data from **flat files** (CSV and delimited), **Excel files**, **databases**, and saving/loading data from the ultrafast **HDF5 format**
- **Time series**-specific functionality: date range generation and frequency conversion, moving window statistics, date shifting and lagging