

In [computer programming](#), **pandas** is a [software library](#) written for the [Python programming language](#) for data manipulation and analysis. In particular, it offers data structures and operations for manipulating numerical tables and [time series](#). It is [free software](#) released under the [three-clause BSD license](#).¹ The name is derived from the term "[panel data](#)", an [econometrics](#) term for data sets that include observations over multiple time periods for the same individuals.

Library features

- DataFrame object for data manipulation with integrated indexing.
- Tools for reading and writing data between in-memory data structures and different file formats.
- Data alignment and integrated handling of missing data.
- Reshaping and pivoting of data sets.
- Label-based slicing, fancy indexing, and subsetting of large data sets.
- Data structure column insertion and deletion.
- Group by engine allowing split-apply-combine operations on data sets.
- Data set merging and joining.
- Hierarchical axis indexing to work with high-dimensional data in a lower-dimensional data structure.
- Time series-functionality: Date range generation^[3] and frequency conversion, moving window statistics, moving window linear regressions, date shifting and lagging.

The library is highly optimized for performance, with critical code paths written in [Cython](#) or [C](#).

¹ [License – Package overview – pandas 0.21.1 documentation](#). *pandas*. 12 December 2017. Retrieved 13 December 2017