

pandas documentation

Date: Apr 10, 2024 Version: 2.2.2

Download documentation: [Zipped HTML](#)

Previous versions: Documentation of previous pandas versions is available at [pandas.pydata.org](#).

Useful links: [Binary Installers](#) | [Source Repository](#) | [Issues & Ideas](#) | [Q&A Support](#) | [Mailing List](#)

pandas is an open source, BSD-licensed library providing high-performance, easy-to-use data structures and data analysis tools for the [Python](#) programming language.



Getting started

New to *pandas*? Check out the getting started guides. They contain an introduction to *pandas*' main concepts and links to additional tutorials.

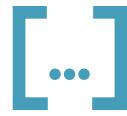
To the getting started guides



User guide

The user guide provides in-depth information on the key concepts of *pandas* with useful background information and explanation.

To the user guide



API reference



Developer guide

[Skip to main content](#)

The reference guide contains a detailed description of the pandas API. The reference describes how the methods work and which parameters can be used. It assumes that you have an understanding of the key concepts.

[To the reference guide](#)

Saw a typo in the documentation? Want to improve existing functionalities? The contributing guidelines will guide you through the process of improving pandas.

[To the development guide](#)

Next

[Getting started](#)



© 2024, pandas via [NumFOCUS, Inc.](#) Hosted by [OVHcloud](#).

Created using [Sphinx](#) 7.2.6.

Built with the [PyData Sphinx Theme](#) 0.14.4.

Date: Apr 10, 2024 Version: 2.2.2

Download documentation: [Zipped HTML](#)

Previous versions: Documentation of previous pandas versions is available at [pandas.pydata.org](#).

Useful links: [Binary Installers](#) | [Source Repository](#) | [Issues & Ideas](#) | [Q&A Support](#) | [Mailing List](#)

pandas is an open source, BSD-licensed library providing high-performance, easy-to-use data structures and data analysis tools for the [Python](#) programming language.



Getting started

New to *pandas*? Check out the getting started guides. They contain an introduction to *pandas*' main concepts and links to additional tutorials.

To the getting started guides



User guide

The user guide provides in-depth information on the key concepts of *pandas* with useful background information and explanation.

To the user guide



API reference



Developer guide

The reference guide contains a detailed description of the pandas API. The reference describes how the methods work and which parameters can be used. It assumes that you have an understanding of the key concepts.

[To the reference guide](#)

Saw a typo in the documentation? Want to improve existing functionalities? The contributing guidelines will guide you through the process of improving pandas.

[To the development guide](#)

Next

[Getting started](#)



© 2024, pandas via [NumFOCUS, Inc.](#) Hosted by [OVHcloud](#).

Created using [Sphinx](#) 7.2.6.

Built with the [PyData Sphinx Theme](#) 0.14.4.

pandas documentation

Date: Apr 10, 2024 Version: 2.2.2

Download documentation: [Zipped HTML](#)

Previous versions: Documentation of previous pandas versions is available at [pandas.pydata.org](#).

Useful links: [Binary Installers](#) | [Source Repository](#) | [Issues & Ideas](#) | [Q&A Support](#) | [Mailing List](#)

pandas is an open source, BSD-licensed library providing high-performance, easy-to-use data structures and data analysis tools for the [Python](#) programming language.



Getting started

New to *pandas*? Check out the getting started guides. They contain an introduction to *pandas*' main concepts and links to additional tutorials.

To the getting started guides



User guide

The user guide provides in-depth information on the key concepts of *pandas* with useful background information and explanation.

To the user guide



API reference



Developer guide

[Skip to main content](#)

The reference guide contains a detailed description of the pandas API. The reference describes how the methods work and which parameters can be used. It assumes that you have an understanding of the key concepts.

[To the reference guide](#)

Saw a typo in the documentation? Want to improve existing functionalities? The contributing guidelines will guide you through the process of improving pandas.

[To the development guide](#)

Next

[Getting started](#)



© 2024, pandas via [NumFOCUS, Inc.](#) Hosted by [OVHcloud](#).

Created using [Sphinx](#) 7.2.6.

Built with the [PyData Sphinx Theme](#) 0.14.4.

pandas documentation

Date: Apr 10, 2024 Version: 2.2.2

Download documentation: [Zipped HTML](#)

Previous versions: Documentation of previous pandas versions is available at pandas.pydata.org.

Useful links: [Binary Installers](#) | [Source Repository](#) | [Issues & Ideas](#) | [Q&A Support](#) | [Mailing List](#)

pandas is an open source, BSD-licensed library providing high-performance, easy-to-use data structures and data analysis tools for the [Python](#) programming language.



Getting started

New to *pandas*? Check out the getting started guides. They contain an introduction to *pandas*' main concepts and links to additional tutorials.

To the getting started guides



User guide

The user guide provides in-depth information on the key concepts of *pandas* with useful background information and explanation.

To the user guide



API reference



Developer guide

[Skip to main content](#)

The reference guide contains a detailed description of the pandas API. The reference describes how the methods work and which parameters can be used. It assumes that you have an understanding of the key concepts.

[To the reference guide](#)

Saw a typo in the documentation? Want to improve existing functionalities? The contributing guidelines will guide you through the process of improving pandas.

[To the development guide](#)

Next

[Getting started](#)



© 2024, pandas via [NumFOCUS, Inc.](#) Hosted by [OVHcloud](#).

Created using [Sphinx](#) 7.2.6.

Built with the [PyData Sphinx Theme](#) 0.14.4.

pandas documentation

Date: Apr 10, 2024 Version: 2.2.2

Download documentation: [Zipped HTML](#)

Previous versions: Documentation of previous pandas versions is available at pandas.pydata.org.

Useful links: [Binary Installers](#) | [Source Repository](#) | [Issues & Ideas](#) | [Q&A Support](#) | [Mailing List](#)

pandas is an open source, BSD-licensed library providing high-performance, easy-to-use data structures and data analysis tools for the [Python](#) programming language.



Getting started

New to *pandas*? Check out the getting started guides. They contain an introduction to *pandas*' main concepts and links to additional tutorials.

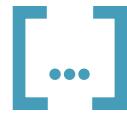
To the getting started guides



User guide

The user guide provides in-depth information on the key concepts of *pandas* with useful background information and explanation.

To the user guide



API reference



Developer guide

[Skip to main content](#)

The reference guide contains a detailed description of the pandas API. The reference describes how the methods work and which parameters can be used. It assumes that you have an understanding of the key concepts.

[To the reference guide](#)

Saw a typo in the documentation? Want to improve existing functionalities? The contributing guidelines will guide you through the process of improving pandas.

[To the development guide](#)

Next

[Getting started](#)



© 2024, pandas via [NumFOCUS, Inc.](#) Hosted by [OVHcloud](#).

Created using [Sphinx](#) 7.2.6.

Built with the [PyData Sphinx Theme](#) 0.14.4.

 > Development

Development

[Contributing to pandas](#)

[Bug reports and enhancement requests](#)

[Finding an issue to contribute to](#)

[Submitting a pull request](#)

[Tips for a successful pull request](#)

[Creating a development environment](#)

[Step 1: install a C compiler](#)

[Step 2: create an isolated environment](#)

[Step 3: build and install pandas](#)

[Contributing to the documentation](#)

[About the pandas documentation](#)

[Updating a pandas docstring](#)

[How to build the pandas documentation](#)

[Previewing changes](#)

[Contributing to the code base](#)

[Code standards](#)

[Pre-commit](#)

[Optional dependencies](#)

[Backwards compatibility](#)

[Type hints](#)

[Testing with continuous integration](#)

[Test-driven development](#)

[Running the test suite](#)

[Running the performance test suite](#)

[Documenting your code](#)

[pandas maintenance](#)

[Skip to main content](#)

[Tasks](#)

[Issue triage](#)

[Investigating regressions](#)

[Closing issues](#)

[Reviewing pull requests](#)

[Backporting](#)

[Cleaning up old issues](#)

[Cleaning up old pull requests](#)

[Becoming a pandas maintainer](#)

[Merging pull requests](#)

[Benchmark machine](#)

[Release process](#)

[Internals](#)

[Indexing](#)

[Subclassing pandas data structures](#)

[Copy on write](#)

[Reference tracking](#)

[Debugging C extensions](#)

[Debugging locally](#)

[Using Docker](#)

[Editor support](#)

[Extending pandas](#)

[Registering custom accessors](#)

[Extension types](#)

[Subclassing pandas data structures](#)

[Plotting backends](#)

[Arithmetic with 3rd party types](#)

[Developer](#)

[Storing pandas DataFrame objects in Apache Parquet format](#)

[Policies](#)

[Version policy](#)

[Python support](#)

[Skip to main content](#)

[Contributor community](#)

[Community meeting](#)

[New contributor meeting](#)

[Calendar](#)

[GitHub issue tracker](#)

[The developer mailing list](#)

[Community slack](#)

< Previous

[pandas.NaT](#)

Next >

[Contributing to pandas](#)

© 2024, pandas via [NumFOCUS, Inc.](#) Hosted by [OVHcloud](#).

Built with the [PyData Sphinx Theme](#) 0.14.4.

Created using [Sphinx](#) 7.2.6.

 > Development

Development

Contributing to pandas

Bug reports and enhancement requests

Finding an issue to contribute to

Submitting a pull request

Tips for a successful pull request

Creating a development environment

Step 1: install a C compiler

Step 2: create an isolated environment

Step 3: build and install pandas

Contributing to the documentation

About the pandas documentation

Updating a pandas docstring

How to build the pandas documentation

Previewing changes

Contributing to the code base

Code standards

Pre-commit

Optional dependencies

Backwards compatibility

Type hints

Testing with continuous integration

Test-driven development

Running the test suite

Running the performance test suite

Documenting your code

pandas maintenance

[Tasks](#)

[Issue triage](#)

[Investigating regressions](#)

[Closing issues](#)

[Reviewing pull requests](#)

[Backporting](#)

[Cleaning up old issues](#)

[Cleaning up old pull requests](#)

[Becoming a pandas maintainer](#)

[Merging pull requests](#)

[Benchmark machine](#)

[Release process](#)

[Internals](#)

[Indexing](#)

[Subclassing pandas data structures](#)

[Copy on write](#)

[Reference tracking](#)

[Debugging C extensions](#)

[Debugging locally](#)

[Using Docker](#)

[Editor support](#)

[Extending pandas](#)

[Registering custom accessors](#)

[Extension types](#)

[Subclassing pandas data structures](#)

[Plotting backends](#)

[Arithmetic with 3rd party types](#)

[Developer](#)

[Storing pandas DataFrame objects in Apache Parquet format](#)

[Policies](#)

[Version policy](#)

[Python support](#)

[Skip to main content](#)

[Contributor community](#)

[Community meeting](#)

[New contributor meeting](#)

[Calendar](#)

[GitHub issue tracker](#)

[The developer mailing list](#)

[Community slack](#)

< Previous

[pandas.NaT](#)

Next >

[Contributing to pandas](#)

© 2024, pandas via [NumFOCUS, Inc.](#) Hosted by [OVHcloud](#).

Built with the [PyData Sphinx Theme](#) 0.14.4.

Created using [Sphinx](#) 7.2.6.

```
:notoc:

.. pandas documentation master file, created by

.. module:: pandas

*****
pandas documentation
*****


**Date**: |today| **Version**: |version|


**Download documentation**: `Zipped HTML <pandas.zip>`__


**Previous versions**: Documentation of previous pandas versions is available at
`pandas.pydata.org <https://pandas.pydata.org/>`__.


**Useful links**:
`Binary Installers <https://pypi.org/project/pandas>`__ |
`Source Repository <https://github.com/pandas-dev/pandas>`__ |
`Issues & Ideas <https://github.com/pandas-dev/pandas/issues>`__ |
`Q&A Support <https://stackoverflow.com/questions/tagged/pandas>`__ |
`Mailing List <https://groups.google.com/g/pydata>`__


:mod:`pandas` is an open source, BSD-licensed library providing high-performance,
easy-to-use data structures and data analysis tools for the `Python <https://www.python.org/>`__
programming language.

.. grid:: 1 2 2 2
   :gutter: 4
   :padding: 2 2 0 0
   :class-container: sd-text-center

.. grid-item-card:: Getting started
   :img-top: _static/index_getting_started.svg
   :class-card: intro-card
   :shadow: md

New to *pandas*? Check out the getting started guides. They contain an
introduction to *pandas'* main concepts and links to additional tutorials.

+++
.. button-ref:: getting_started
   :ref-type: ref
   :click-parent:
   :color: secondary
   :expand:

   To the getting started guides

.. grid-item-card:: User guide
   :img-top: _static/index_user_guide.svg
   :class-card: intro-card
   :shadow: md

The user guide provides in-depth information on the
key concepts of pandas with useful background information and explanation.

+++
.. button-ref:: user_guide
   :ref-type: ref
   :click-parent:
   :color: secondary
   :expand:

   To the user guide
```

```
.. grid-item-card:: API reference
  :img-top: _static/index_api.svg
  :class-card: intro-card
  :shadow: md
```

The reference guide contains a detailed description of the pandas API. The reference describes how the methods work and which parameters can be used. It assumes that you have an understanding of the key concepts.

+++

```
.. button-ref:: api
  :ref-type: ref
  :click-parent:
  :color: secondary
  :expand:
```

To the reference guide

```
.. grid-item-card:: Developer guide
  :img-top: _static/index_contribute.svg
  :class-card: intro-card
  :shadow: md
```

Saw a typo in the documentation? Want to improve existing functionalities? The contributing guidelines will guide you through the process of improving pandas.

+++

```
.. button-ref:: development
  :ref-type: ref
  :click-parent:
  :color: secondary
  :expand:
```

To the development guide

```
.. toctree::
  :maxdepth: 3
  :hidden:
  :titlesonly:
```

getting_started/index
user_guide/index
reference/index
development/index
whatsnew/index

User Guide

The User Guide covers all of pandas by topic area. Each of the subsections introduces a topic (such as “working with missing data”), and discusses how pandas approaches the problem, with many examples throughout.

Users brand-new to pandas should start with [10 minutes to pandas](#).

For a high level summary of the pandas fundamentals, see [Intro to data structures](#) and [Essential basic functionality](#).

Further information on any specific method can be obtained in the [API reference](#).

How to read these guides

In these guides you will see input code inside code blocks such as:

```
import pandas as pd
pd.DataFrame({'A': [1, 2, 3]})
```

or:

```
In [1]: import pandas as pd
In [2]: pd.DataFrame({'A': [1, 2, 3]})
Out[2]:
   A
0  1
1  2
2  3
```

The first block is a standard python input, while in the second the `In [1]:` indicates the input is inside a [notebook](#). In Jupyter Notebooks the last line is printed and plots are shown inline.

For example:

In [4]: a
Out[4]: 1

is equivalent to:

```
a = 1  
print(a)
```

Guides

[10 minutes to pandas](#)

[Basic data structures in pandas](#)

[Object creation](#)

[Viewing data](#)

[Selection](#)

[Missing data](#)

[Operations](#)

[Merge](#)

[Grouping](#)

[Reshaping](#)

[Time series](#)

[Categoricals](#)

[Plotting](#)

[Importing and exporting data](#)

[Gotchas](#)

[Intro to data structures](#)

[Series](#)

[DataFrame](#)

[Essential basic functionality](#)

[Head and tail](#)

[Attributes and underlying data](#)

[Accelerated operations](#)

[Flexible binary operations](#)

[Function application](#)

[Reindexing and altering labels](#)

[Iteration](#)

[.dt accessor](#)

[Vectorized string methods](#)

[Sorting](#)

[Copying](#)

[dtypes](#)

[Selecting columns based on `dtype`](#)

[IO tools \(text, CSV, HDF5, ...\)](#)

[CSV & text files](#)

[JSON](#)

[HTML](#)

[LaTeX](#)

[XML](#)

[Excel files](#)

[OpenDocument Spreadsheets](#)

[Binary Excel \(.xlsm\) files](#)

[Calamine \(Excel and ODS files\)](#)

[Clipboard](#)

[Pickling](#)

[msgpack](#)

[HDF5 \(PyTables\)](#)

[Feather](#)

[Parquet](#)

[ORC](#)

[SQL queries](#)

[Google BigQuery](#)

[Stata format](#)

[SAS formats](#)

[SPSS formats](#)

[Other file formats](#)

Performance considerations

PyArrow Functionality

Data Structure Integration

Operations

I/O Reading

Indexing and selecting data

Different choices for indexing

Basics

Attribute access

Slicing ranges

Selection by label

Selection by position

Selection by callable

Combining positional and label-based indexing

Selecting random samples

Setting with enlargement

Fast scalar value getting and setting

Boolean indexing

Indexing with isin

The `where()` Method and Masking

Setting with enlargement conditionally using `numpy()`

The `query()` Method

Duplicate data

Dictionary-like `get()` method

Looking up values by index/column labels

Index objects

Set / reset index

Returning a view versus a copy

MultIndex / advanced indexing

Hierarchical indexing (MultIndex)

Advanced indexing with hierarchical index

Sorting a `MultiIndex`

[Index types](#)

[Miscellaneous indexing FAQ](#)

[Copy-on-Write \(CoW\)](#)

[Previous behavior](#)

[Migrating to Copy-on-Write](#)

[Description](#)

[Chained Assignment](#)

[Read-only NumPy arrays](#)

[Patterns to avoid](#)

[Copy-on-Write optimizations](#)

[How to enable CoW](#)

[Merge, join, concatenate and compare](#)

[concat\(\)](#)

[merge\(\)](#)

[DataFrame.join\(\)](#)

[merge_ordered\(\)](#)

[merge_asof\(\)](#)

[compare\(\)](#)

[Reshaping and pivot tables](#)

[pivot\(\) and pivot_table\(\)](#)

[stack\(\) and unstack\(\)](#)

[melt\(\) and wide_to_long\(\)](#)

[get_dummies\(\) and from_dummies\(\)](#)

[explode\(\)](#)

[crosstab\(\)](#)

[cut\(\)](#)

[factorize\(\)](#)

[Working with text data](#)

[Text data types](#)

[String methods](#)

[Splitting and replacing strings](#)

Concatenation

[Skip to main content](#)

[Extracting substrings](#)

[Testing for strings that match or contain a pattern](#)

[Creating indicator variables](#)

[Method summary](#)

[Working with missing data](#)

[Values considered “missing”](#)

[NA semantics](#)

[Inserting missing data](#)

[Calculations with missing data](#)

[Dropping missing data](#)

[Filling missing data](#)

[Duplicate Labels](#)

[Consequences of Duplicate Labels](#)

[Duplicate Label Detection](#)

[Disallowing Duplicate Labels](#)

[Categorical data](#)

[Object creation](#)

[CategoricalDtype](#)

[Description](#)

[Working with categories](#)

[Sorting and order](#)

[Comparisons](#)

[Operations](#)

[Data munging](#)

[Getting data in/out](#)

[Missing data](#)

[Differences to R's `factor`](#)

[Gotchas](#)

[Nullable integer data type](#)

[Construction](#)

[Operations](#)

[Scalar NA Value](#)

[Indexing with NA values](#)

[Kleene logical operations](#)

[Chart visualization](#)

[Basic plotting: `plot`](#)

[Other plots](#)

[Plotting with missing data](#)

[Plotting tools](#)

[Plot formatting](#)

[Plotting directly with Matplotlib](#)

[Plotting backends](#)

[Table Visualization](#)

[Styler Object and Customising the Display](#)

[Formatting the Display](#)

[Styler Object and HTML](#)

[Methods to Add Styles](#)

[Table Styles](#)

[Setting Classes and Linking to External CSS](#)

[Styler Functions](#)

[Tooltips and Captions](#)

[Finer Control with Slicing](#)

[Optimization](#)

[Builtin Styles](#)

[Sharing styles](#)

[Limitations](#)

[Other Fun and Useful Stuff](#)

[Export to Excel](#)

[Export to LaTeX](#)

[More About CSS and HTML](#)

[Extensibility](#)

[Group by: split-apply-combine](#)

[Splitting an object into groups](#)

[Iterating through groups](#)

[Aggregation](#)[Transformation](#)[Filtration](#)[Flexible `apply`](#)[Numba Accelerated Routines](#)[Other useful features](#)[Examples](#)[Windowing operations](#)[Overview](#)[Rolling window](#)[Weighted window](#)[Expanding window](#)[Exponentially weighted window](#)[Time series / date functionality](#)[Overview](#)[Timestamps vs. time spans](#)[Converting to timestamps](#)[Generating ranges of timestamps](#)[Timestamp limitations](#)[Indexing](#)[Time/date components](#)[DateOffset objects](#)[Time Series-related instance methods](#)[Resampling](#)[Time span representation](#)[Converting between representations](#)[Representing out-of-bounds spans](#)[Time zone handling](#)[Time deltas](#)[Parsing](#)[Operations](#)[Reductions](#)[Skip to main content](#)

[Attributes](#)

[TimedeltaIndex](#)

[Resampling](#)

[Options and settings](#)

[Overview](#)

[Available options](#)

[Getting and setting options](#)

[Setting startup options in Python/IPython environment](#)

[Frequently used options](#)

[Number formatting](#)

[Unicode formatting](#)

[Table schema display](#)

[Enhancing performance](#)

[Cython \(writing C extensions for pandas\)](#)

[Numba \(JIT compilation\)](#)

[Expression evaluation via `eval\(\)`](#)

[Scaling to large datasets](#)

[Load less data](#)

[Use efficient datatypes](#)

[Use chunking](#)

[Use Other Libraries](#)

[Sparse data structures](#)

[SparseArray](#)

[SparseDtype](#)

[Sparse accessor](#)

[Sparse calculation](#)

[Interaction with `scipy.sparse`](#)

[Frequently Asked Questions \(FAQ\)](#)

[DataFrame memory usage](#)

[Using if/truth statements with pandas](#)

[Mutating with User Defined Function \(UDF\) methods](#)

[Missing value representation for NumPy types](#)

[Skip to main content](#)

[Thread-safety](#)

[Byte-ordering issues](#)

[Cookbook](#)

[Idioms](#)

[Selection](#)

[Multiindexing](#)

[Missing data](#)

[Grouping](#)

[Timeseries](#)

[Merge](#)

[Plotting](#)

[Data in/out](#)

[Computation](#)

[Timedeltas](#)

[Creating example data](#)

[Constant series](#)

[Previous](#)

[Next](#)

© 2024, pandas via [NumFOCUS, Inc.](#) Hosted by [OVHcloud](#).

Built with the [PyData Sphinx Theme](#) 0.14.4.

Created using [Sphinx](#) 7.2.6.

User Guide

The User Guide covers all of pandas by topic area. Each of the subsections introduces a topic (such as “working with missing data”), and discusses how pandas approaches the problem, with many examples throughout.

Users brand-new to pandas should start with [10 minutes to pandas](#).

For a high level summary of the pandas fundamentals, see [Intro to data structures](#) and [Essential basic functionality](#).

Further information on any specific method can be obtained in the [API reference](#).

How to read these guides

In these guides you will see input code inside code blocks such as:

```
import pandas as pd
pd.DataFrame({'A': [1, 2, 3]})
```

or:

```
In [1]: import pandas as pd
In [2]: pd.DataFrame({'A': [1, 2, 3]})
Out[2]:
A
0 1
1 2
2 3
```

The first block is a standard python input, while in the second the `In [1]:` indicates the input is inside a [notebook](#). In Jupyter Notebooks the last line is printed and plots are shown inline.

For example:

In [4]: a
Out[4]: 1

is equivalent to:

```
a = 1  
print(a)
```

Guides

[10 minutes to pandas](#)

[Basic data structures in pandas](#)

[Object creation](#)

[Viewing data](#)

[Selection](#)

[Missing data](#)

[Operations](#)

[Merge](#)

[Grouping](#)

[Reshaping](#)

[Time series](#)

[Categoricals](#)

[Plotting](#)

[Importing and exporting data](#)

[Gotchas](#)

[Intro to data structures](#)

[Series](#)

[DataFrame](#)

[Essential basic functionality](#)

[Head and tail](#)

[Attributes and underlying data](#)

[Accelerated operations](#)

[Flexible binary operations](#)

[Function application](#)

[Reindexing and altering labels](#)

[Iteration](#)

[.dt accessor](#)

[Vectorized string methods](#)

[Sorting](#)

[Copying](#)

[dtypes](#)

[Selecting columns based on `dtype`](#)

[IO tools \(text, CSV, HDF5, ...\)](#)

[CSV & text files](#)

[JSON](#)

[HTML](#)

[LaTeX](#)

[XML](#)

[Excel files](#)

[OpenDocument Spreadsheets](#)

[Binary Excel \(.xlsm\) files](#)

[Calamine \(Excel and ODS files\)](#)

[Clipboard](#)

[Pickling](#)

[msgpack](#)

[HDF5 \(PyTables\)](#)

[Feather](#)

[Parquet](#)

[ORC](#)

[SQL queries](#)

[Google BigQuery](#)

[Stata format](#)

[SAS formats](#)

[SPSS formats](#)

[Other file formats](#)

[Skip to main content](#)

Performance considerations

PyArrow Functionality

Data Structure Integration

Operations

I/O Reading

Indexing and selecting data

Different choices for indexing

Basics

Attribute access

Slicing ranges

Selection by label

Selection by position

Selection by callable

Combining positional and label-based indexing

Selecting random samples

Setting with enlargement

Fast scalar value getting and setting

Boolean indexing

Indexing with isin

The `where()` Method and Masking

Setting with enlargement conditionally using `numpy()`

The `query()` Method

Duplicate data

Dictionary-like `get()` method

Looking up values by index/column labels

Index objects

Set / reset index

Returning a view versus a copy

MultIndex / advanced indexing

Hierarchical indexing (MultIndex)

Advanced indexing with hierarchical index

Sorting a `MultiIndex`

[Index types](#)

[Miscellaneous indexing FAQ](#)

[Copy-on-Write \(CoW\)](#)

[Previous behavior](#)

[Migrating to Copy-on-Write](#)

[Description](#)

[Chained Assignment](#)

[Read-only NumPy arrays](#)

[Patterns to avoid](#)

[Copy-on-Write optimizations](#)

[How to enable CoW](#)

[Merge, join, concatenate and compare](#)

[concat\(\)](#)

[merge\(\)](#)

[DataFrame.join\(\)](#)

[merge_ordered\(\)](#)

[merge_asof\(\)](#)

[compare\(\)](#)

[Reshaping and pivot tables](#)

[pivot\(\) and pivot_table\(\)](#)

[stack\(\) and unstack\(\)](#)

[melt\(\) and wide_to_long\(\)](#)

[get_dummies\(\) and from_dummies\(\)](#)

[explode\(\)](#)

[crosstab\(\)](#)

[cut\(\)](#)

[factorize\(\)](#)

[Working with text data](#)

[Text data types](#)

[String methods](#)

[Splitting and replacing strings](#)

Concatenation

[Skip to main content](#)

[Extracting substrings](#)

[Testing for strings that match or contain a pattern](#)

[Creating indicator variables](#)

[Method summary](#)

[Working with missing data](#)

[Values considered “missing”](#)

[NA semantics](#)

[Inserting missing data](#)

[Calculations with missing data](#)

[Dropping missing data](#)

[Filling missing data](#)

[Duplicate Labels](#)

[Consequences of Duplicate Labels](#)

[Duplicate Label Detection](#)

[Disallowing Duplicate Labels](#)

[Categorical data](#)

[Object creation](#)

[CategoricalDtype](#)

[Description](#)

[Working with categories](#)

[Sorting and order](#)

[Comparisons](#)

[Operations](#)

[Data munging](#)

[Getting data in/out](#)

[Missing data](#)

[Differences to R's `factor`](#)

[Gotchas](#)

[Nullable integer data type](#)

[Construction](#)

[Operations](#)

[Scalar NA Value](#)

[Indexing with NA values](#)

[Kleene logical operations](#)

[Chart visualization](#)

[Basic plotting: `plot`](#)

[Other plots](#)

[Plotting with missing data](#)

[Plotting tools](#)

[Plot formatting](#)

[Plotting directly with Matplotlib](#)

[Plotting backends](#)

[Table Visualization](#)

[Styler Object and Customising the Display](#)

[Formatting the Display](#)

[Styler Object and HTML](#)

[Methods to Add Styles](#)

[Table Styles](#)

[Setting Classes and Linking to External CSS](#)

[Styler Functions](#)

[Tooltips and Captions](#)

[Finer Control with Slicing](#)

[Optimization](#)

[Builtin Styles](#)

[Sharing styles](#)

[Limitations](#)

[Other Fun and Useful Stuff](#)

[Export to Excel](#)

[Export to LaTeX](#)

[More About CSS and HTML](#)

[Extensibility](#)

[Group by: split-apply-combine](#)

[Splitting an object into groups](#)

[Iterating through groups](#)

[Aggregation](#)[Transformation](#)[Filtration](#)[Flexible **apply**](#)[Numba Accelerated Routines](#)[Other useful features](#)[Examples](#)[Windowing operations](#)[Overview](#)[Rolling window](#)[Weighted window](#)[Expanding window](#)[Exponentially weighted window](#)[Time series / date functionality](#)[Overview](#)[Timestamps vs. time spans](#)[Converting to timestamps](#)[Generating ranges of timestamps](#)[Timestamp limitations](#)[Indexing](#)[Time/date components](#)[DateOffset objects](#)[Time Series-related instance methods](#)[Resampling](#)[Time span representation](#)[Converting between representations](#)[Representing out-of-bounds spans](#)[Time zone handling](#)[Time deltas](#)[Parsing](#)[Operations](#)[Reductions](#)[Skip to main content](#)

[Attributes](#)

[TimedeltaIndex](#)

[Resampling](#)

[Options and settings](#)

[Overview](#)

[Available options](#)

[Getting and setting options](#)

[Setting startup options in Python/IPython environment](#)

[Frequently used options](#)

[Number formatting](#)

[Unicode formatting](#)

[Table schema display](#)

[Enhancing performance](#)

[Cython \(writing C extensions for pandas\)](#)

[Numba \(JIT compilation\)](#)

[Expression evaluation via `eval\(\)`](#)

[Scaling to large datasets](#)

[Load less data](#)

[Use efficient datatypes](#)

[Use chunking](#)

[Use Other Libraries](#)

[Sparse data structures](#)

[SparseArray](#)

[SparseDtype](#)

[Sparse accessor](#)

[Sparse calculation](#)

[Interaction with `scipy.sparse`](#)

[Frequently Asked Questions \(FAQ\)](#)

[DataFrame memory usage](#)

[Using if/truth statements with pandas](#)

[Mutating with User Defined Function \(UDF\) methods](#)

[Missing value representation for NumPy types](#)

[Skip to main content](#)

[Thread-safety](#)

[Byte-ordering issues](#)

[Cookbook](#)

[Idioms](#)

[Selection](#)

[Multiindexing](#)

[Missing data](#)

[Grouping](#)

[Timeseries](#)

[Merge](#)

[Plotting](#)

[Data in/out](#)

[Computation](#)

[Timedeltas](#)

[Creating example data](#)

[Constant series](#)

[Previous](#)

[Next](#)

© 2024, pandas via [NumFOCUS, Inc.](#) Hosted by [OVHcloud](#).

Built with the [PyData Sphinx Theme](#) 0.14.4.

Created using [Sphinx](#) 7.2.6.

Release notes

This is the list of changes to pandas between each release. For full details, see the [commit logs](#). For install and upgrade instructions, see [Installation](#).

Version 2.2

[What's new in 2.2.2 \(April 10, 2024\)](#)

[Pandas 2.2.2 is now compatible with numpy 2.0](#)

[Fixed regressions](#)

[Bug fixes](#)

[Other](#)

[Contributors](#)

[What's new in 2.2.1 \(February 22, 2024\)](#)

[Enhancements](#)

[Fixed regressions](#)

[Bug fixes](#)

[Other](#)

[Contributors](#)

[What's new in 2.2.0 \(January 19, 2024\)](#)

[Upcoming changes in pandas 3.0](#)

[Enhancements](#)

[Notable bug fixes](#)

[Deprecations](#)

[Performance improvements](#)

[Bug fixes](#)

[Contributors](#)

What's new in 2.1.4 (December 8, 2023)

[Fixed regressions](#)

[Bug fixes](#)

[Contributors](#)

What's new in 2.1.3 (November 10, 2023)

[Fixed regressions](#)

[Bug fixes](#)

[Contributors](#)

What's new in 2.1.2 (October 26, 2023)

[Deprecations](#)

[Fixed regressions](#)

[Bug fixes](#)

[Other](#)

[Contributors](#)

What's new in 2.1.1 (September 20, 2023)

[Fixed regressions](#)

[Bug fixes](#)

[Other](#)

[Contributors](#)

What's new in 2.1.0 (Aug 30, 2023)

[Enhancements](#)

[Backwards incompatible API changes](#)

[Deprecations](#)

[Performance improvements](#)

[Bug fixes](#)

[Contributors](#)

Version 2.0

What's new in 2.0.3 (June 28, 2023)

[Fixed regressions](#)

[Bug fixes](#)

[Skip to main content](#)

Contributors

What's new in 2.0.2 (May 29, 2023)

Fixed regressions

Bug fixes

Other

Contributors

What's new in 2.0.1 (April 24, 2023)

Fixed regressions

Bug fixes

Other

Contributors

What's new in 2.0.0 (April 3, 2023)

Enhancements

Notable bug fixes

Backwards incompatible API changes

Deprecations

Removal of prior version deprecations/changes

Performance improvements

Bug fixes

Contributors

Version 1.5

What's new in 1.5.3 (January 18, 2023)

Fixed regressions

Bug fixes

Other

Contributors

What's new in 1.5.2 (November 21, 2022)

Fixed regressions

Bug fixes

Other

What's new in 1.5.1 (October 19, 2022)

Behavior of `groupby` with categorical groupers (GH 48645)

Fixed regressions

Bug fixes

Other

Contributors

What's new in 1.5.0 (September 19, 2022)

Enhancements

Notable bug fixes

Backwards incompatible API changes

Deprecations

Performance improvements

Bug fixes

Contributors

Version 1.4

What's new in 1.4.4 (August 31, 2022)

Fixed regressions

Bug fixes

Other

Contributors

What's new in 1.4.3 (June 23, 2022)

Behavior of `concat` with empty or all-NA DataFrame columns

Fixed regressions

Bug fixes

Other

Contributors

What's new in 1.4.2 (April 2, 2022)

Fixed regressions

Bug fixes

Contributors

[Fixed regressions](#)

[Bug fixes](#)

[Other](#)

[Contributors](#)

[What's new in 1.4.0 \(January 22, 2022\)](#)

[Enhancements](#)

[Notable bug fixes](#)

[Backwards incompatible API changes](#)

[Deprecations](#)

[Performance improvements](#)

[Bug fixes](#)

[Contributors](#)

Version 1.3

[What's new in 1.3.5 \(December 12, 2021\)](#)

[Fixed regressions](#)

[Contributors](#)

[What's new in 1.3.4 \(October 17, 2021\)](#)

[Fixed regressions](#)

[Bug fixes](#)

[Other](#)

[Contributors](#)

[What's new in 1.3.3 \(September 12, 2021\)](#)

[Fixed regressions](#)

[Performance improvements](#)

[Bug fixes](#)

[Contributors](#)

[What's new in 1.3.2 \(August 15, 2021\)](#)

[Fixed regressions](#)

[Bug fixes](#)

[Contributors](#)

[Fixed regressions](#)

[Bug fixes](#)

[Contributors](#)

[What's new in 1.3.0 \(July 2, 2021\)](#)

[Enhancements](#)

[Notable bug fixes](#)

[Backwards incompatible API changes](#)

[Deprecations](#)

[Performance improvements](#)

[Bug fixes](#)

[Contributors](#)

Version 1.2

[What's new in 1.2.5 \(June 22, 2021\)](#)

[Fixed regressions](#)

[Contributors](#)

[What's new in 1.2.4 \(April 12, 2021\)](#)

[Fixed regressions](#)

[Contributors](#)

[What's new in 1.2.3 \(March 02, 2021\)](#)

[Fixed regressions](#)

[Contributors](#)

[What's new in 1.2.2 \(February 09, 2021\)](#)

[Fixed regressions](#)

[Bug fixes](#)

[Contributors](#)

[What's new in 1.2.1 \(January 20, 2021\)](#)

[Fixed regressions](#)

[Calling NumPy ufuncs on non-aligned DataFrames](#)

[Bug fixes](#)

[Other](#)

What's new in 1.2.0 (December 26, 2020)

[Enhancements](#)

[Notable bug fixes](#)

[Deprecations](#)

[Performance improvements](#)

[Bug fixes](#)

[Contributors](#)

Version 1.1

What's new in 1.1.5 (December 07, 2020)

[Fixed regressions](#)

[Bug fixes](#)

[Other](#)

[Contributors](#)

What's new in 1.1.4 (October 30, 2020)

[Fixed regressions](#)

[Bug fixes](#)

[Contributors](#)

What's new in 1.1.3 (October 5, 2020)

[Enhancements](#)

[Fixed regressions](#)

[Bug fixes](#)

[Other](#)

[Contributors](#)

What's new in 1.1.2 (September 8, 2020)

[Fixed regressions](#)

[Bug fixes](#)

[Other](#)

[Contributors](#)

What's new in 1.1.1 (August 20, 2020)

[Fixed regressions](#)

[Contributors](#)

[What's new in 1.1.0 \(July 28, 2020\)](#)

[Enhancements](#)

[Notable bug fixes](#)

[Backwards incompatible API changes](#)

[Deprecations](#)

[Performance improvements](#)

[Bug fixes](#)

[Contributors](#)

Version 1.0

[What's new in 1.0.5 \(June 17, 2020\)](#)

[Fixed regressions](#)

[Bug fixes](#)

[Contributors](#)

[What's new in 1.0.4 \(May 28, 2020\)](#)

[Fixed regressions](#)

[Bug fixes](#)

[Contributors](#)

[What's new in 1.0.3 \(March 17, 2020\)](#)

[Fixed regressions](#)

[Bug fixes](#)

[Contributors](#)

[What's new in 1.0.2 \(March 12, 2020\)](#)

[Fixed regressions](#)

[Indexing with nullable boolean arrays](#)

[Bug fixes](#)

[Contributors](#)

[What's new in 1.0.1 \(February 5, 2020\)](#)

[Fixed regressions](#)

[Deprecations](#)

Contributors

What's new in 1.0.0 (January 29, 2020)

New deprecation policy

Enhancements

Experimental new features

Other enhancements

Backwards incompatible API changes

Deprecations

Removal of prior version deprecations/changes

Performance improvements

Bug fixes

Contributors

Version 0.25

What's new in 0.25.3 (October 31, 2019)

Bug fixes

Contributors

What's new in 0.25.2 (October 15, 2019)

Bug fixes

Contributors

What's new in 0.25.1 (August 21, 2019)

IO and LZMA

Bug fixes

Contributors

What's new in 0.25.0 (July 18, 2019)

Enhancements

Backwards incompatible API changes

Deprecations

Removal of prior version deprecations/changes

Performance improvements

Bug fixes

Version 0.24

[What's new in 0.24.2 \(March 12, 2019\)](#)

[Fixed regressions](#)

[Bug fixes](#)

[Contributors](#)

[What's new in 0.24.1 \(February 3, 2019\)](#)

[API changes](#)

[Fixed regressions](#)

[Bug fixes](#)

[Contributors](#)

[What's new in 0.24.0 \(January 25, 2019\)](#)

[Enhancements](#)

[Backwards incompatible API changes](#)

[Extension type changes](#)

[Deprecations](#)

[Removal of prior version deprecations/changes](#)

[Performance improvements](#)

[Bug fixes](#)

[Contributors](#)

Version 0.23

[What's new in 0.23.4 \(August 3, 2018\)](#)

[Fixed regressions](#)

[Bug fixes](#)

[Contributors](#)

[What's new in 0.23.3 \(July 7, 2018\)](#)

[Contributors](#)

[What's new in 0.23.2 \(July 5, 2018\)](#)

[Logical reductions over entire DataFrame](#)

[Fixed regressions](#)

[Build changes](#)

[Skip to main content](#)

[Bug fixes](#)

[Contributors](#)

[What's new in 0.23.1 \(June 12, 2018\)](#)

[Fixed regressions](#)

[Performance improvements](#)

[Bug fixes](#)

[Contributors](#)

[What's new in 0.23.0 \(May 15, 2018\)](#)

[New features](#)

[Backwards incompatible API changes](#)

[Deprecations](#)

[Removal of prior version deprecations/changes](#)

[Performance improvements](#)

[Documentation changes](#)

[Bug fixes](#)

[Contributors](#)

Version 0.22

[Version 0.22.0 \(December 29, 2017\)](#)

[Backwards incompatible API changes](#)

[Compatibility](#)

[Contributors](#)

Version 0.21

[Version 0.21.1 \(December 12, 2017\)](#)

[Restore Matplotlib datetime converter registration](#)

[New features](#)

[Deprecations](#)

[Performance improvements](#)

[Bug fixes](#)

[Skip to main content](#)

Version 0.21.0 (October 27, 2017)

[New features](#)

[Backwards incompatible API changes](#)

[Deprecations](#)

[Removal of prior version deprecations/changes](#)

[Performance improvements](#)

[Documentation changes](#)

[Bug fixes](#)

[Contributors](#)

Version 0.20

Version 0.20.3 (July 7, 2017)

[Bug fixes](#)

[Contributors](#)

Version 0.20.2 (June 4, 2017)

[Enhancements](#)

[Performance improvements](#)

[Bug fixes](#)

[Contributors](#)

Version 0.20.1 (May 5, 2017)

[New features](#)

[Backwards incompatible API changes](#)

[Reorganization of the library: privacy changes](#)

[Deprecations](#)

[Removal of prior version deprecations/changes](#)

[Performance improvements](#)

[Bug fixes](#)

[Contributors](#)

Version 0.19

[Enhancements](#)

[Performance improvements](#)

[Bug fixes](#)

[Contributors](#)

[Version 0.19.1 \(November 3, 2016\)](#)

[Performance improvements](#)

[Bug fixes](#)

[Contributors](#)

[Version 0.19.0 \(October 2, 2016\)](#)

[New features](#)

[API changes](#)

[Deprecations](#)

[Removal of prior version deprecations/changes](#)

[Performance improvements](#)

[Bug fixes](#)

[Contributors](#)

Version 0.18

[Version 0.18.1 \(May 3, 2016\)](#)

[New features](#)

[Sparse changes](#)

[API changes](#)

[Performance improvements](#)

[Bug fixes](#)

[Contributors](#)

[Version 0.18.0 \(March 13, 2016\)](#)

[New features](#)

[Backwards incompatible API changes](#)

[Performance improvements](#)

[Bug fixes](#)

[Contributors](#)

Version 0.17

Version 0.17.1 (November 21, 2015)

[New features](#)

[Enhancements](#)

[API changes](#)

[Performance improvements](#)

[Bug fixes](#)

[Contributors](#)

Version 0.17.0 (October 9, 2015)

[New features](#)

[Backwards incompatible API changes](#)

[Performance improvements](#)

[Bug fixes](#)

[Contributors](#)

Version 0.16

Version 0.16.2 (June 12, 2015)

[New features](#)

[API changes](#)

[Performance improvements](#)

[Bug fixes](#)

[Contributors](#)

Version 0.16.1 (May 11, 2015)

[Enhancements](#)

[API changes](#)

[Index representation](#)

[Performance improvements](#)

[Bug fixes](#)

[Contributors](#)

Version 0.16.0 (March 22, 2015)

... more

[Skip to main content](#)

[Backwards incompatible API changes](#)

[Performance improvements](#)

[Bug fixes](#)

[Contributors](#)

Version 0.15

[Version 0.15.2 \(December 12, 2014\)](#)

[API changes](#)

[Enhancements](#)

[Performance](#)

[Bug fixes](#)

[Contributors](#)

[Version 0.15.1 \(November 9, 2014\)](#)

[API changes](#)

[Enhancements](#)

[Bug fixes](#)

[Contributors](#)

[Version 0.15.0 \(October 18, 2014\)](#)

[New features](#)

[Backwards incompatible API changes](#)

[Enhancements](#)

[Performance](#)

[Bug fixes](#)

[Contributors](#)

Version 0.14

[Version 0.14.1 \(July 11, 2014\)](#)

[API changes](#)

[Enhancements](#)

[Performance](#)

[Bug fixes](#)

[Contributors](#)

[Version 0.14.0 \(May 31, 2014\)](#)

[API changes](#)

[Display changes](#)

[Text parsing API changes](#)

[GroupBy API changes](#)

[SQL](#)

[Multi-indexing using slicers](#)

[Plotting](#)

[Prior version deprecations/changes](#)

[Deprecations](#)

[Known issues](#)

[Enhancements](#)

[Performance](#)

[Experimental](#)

[Bug fixes](#)

[Contributors](#)

Version 0.13

[Version 0.13.1 \(February 3, 2014\)](#)

[Output formatting enhancements](#)

[API changes](#)

[Prior version deprecations/changes](#)

[Deprecations](#)

[Enhancements](#)

[Performance](#)

[Experimental](#)

[Bug fixes](#)

[Contributors](#)

[Version 0.13.0 \(January 3, 2014\)](#)

[Skip to main content](#)

[Prior version deprecations/changes](#)

[Deprecations](#)

[Indexing API changes](#)

[Float64Index API change](#)

[HDFStore API changes](#)

[DataFrame repr changes](#)

[Enhancements](#)

[Experimental](#)

[Internal refactoring](#)

[Bug fixes](#)

[Contributors](#)

Version 0.12

[Version 0.12.0 \(July 24, 2013\)](#)

[API changes](#)

[IO enhancements](#)

[Other enhancements](#)

[Experimental features](#)

[Bug fixes](#)

[Contributors](#)

Version 0.11

[Version 0.11.0 \(April 22, 2013\)](#)

[Selection choices](#)

[Selection deprecations](#)

[Dtypes](#)

[Dtype conversion](#)

[Dtype gotchas](#)

[Datetime conversion](#)

[API changes](#)

[Skip to main content](#)

Version 0.10

[Version 0.10.1 \(January 22, 2013\)](#)

[API changes](#)[New features](#)[HDFStore](#)[Contributors](#)

[Version 0.10.0 \(December 17, 2012\)](#)

[File parsing new features](#)[API changes](#)[New features](#)[Wide DataFrame printing](#)[Updated PyTables support](#)[N dimensional panels \(experimental\)](#)[Contributors](#)

Version 0.9

[Version 0.9.1 \(November 14, 2012\)](#)

[New features](#)[API changes](#)[Contributors](#)

[Version 0.9.0 \(October 7, 2012\)](#)

[New features](#)[API changes](#)[Contributors](#)

Version 0.8

[Version 0.8.1 \(July 22, 2012\)](#)

Performance improvements

Contributors

Version 0.8.0 (June 29, 2012)

Support for non-unique indexes

NumPy datetime64 dtype and 1.6 dependency

Time Series changes and improvements

Other new features

New plotting methods

Other API changes

Potential porting issues for pandas <= 0.7.3 users

Contributors

Version 0.7

Version 0.7.3 (April 12, 2012)

New features

NA boolean comparison API change

Other API changes

Contributors

Version 0.7.2 (March 16, 2012)

New features

Performance improvements

Contributors

Version 0.7.1 (February 29, 2012)

New features

Performance improvements

Contributors

Version 0.7.0 (February 9, 2012)

New features

API changes to integer indexing

API tweaks regarding label-based slicing

Changes to Series [1] operator

[Performance improvements](#)

[Contributors](#)

Version 0.6

[Version 0.6.1 \(December 13, 2011\)](#)

[New features](#)

[Performance improvements](#)

[Contributors](#)

[Version 0.6.0 \(November 25, 2011\)](#)

[New features](#)

[Performance enhancements](#)

[Contributors](#)

Version 0.5

[Version 0.5.0 \(October 24, 2011\)](#)

[New features](#)

[Performance enhancements](#)

[Contributors](#)

Version 0.4

[Versions 0.4.1 through 0.4.3 \(September 25 - October 9, 2011\)](#)

© 2024, pandas via [NumFOCUS, Inc.](#) Hosted by [OVHcloud](#).

Created using [Sphinx 7.2.6](#).

Built with the [PyData Sphinx Theme](#) 0.14.4.

pandas documentation

Date: Jul 25, 2024 Version: 3.0.0.dev0+1252.gd6c9941b73

Download documentation: [Zipped HTML](#)

Previous versions: Documentation of previous pandas versions is available at [pandas.pydata.org](#).

Useful links: [Binary Installers](#) | [Source Repository](#) | [Issues & Ideas](#) | [Q&A Support](#) | [Mailing List](#)

pandas is an open source, BSD-licensed library providing high-performance, easy-to-use data structures and data analysis tools for the [Python](#) programming language.



Getting started

New to *pandas*? Check out the getting started guides. They contain an introduction to *pandas*' main concepts and links to additional tutorials.

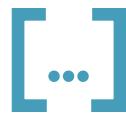
To the getting started guides



User guide

The user guide provides in-depth information on the key concepts of *pandas* with useful background information and explanation.

To the user guide



API reference

The reference guide contains a detailed description of the pandas API. The reference describes how the methods work and which parameters can be used. It assumes that you have an understanding of the key concepts.

[To the reference guide](#)

Developer guide

Saw a typo in the documentation? Want to improve existing functionalities? The contributing guidelines will guide you through the process of improving pandas.

[To the development guide](#)

Next

[Getting started >](#)

© 2024, pandas via [NumFOCUS, Inc.](#) Hosted by [OVHcloud](#).

Created using [Sphinx](#) 7.4.7.

Built with the [PyData Sphinx Theme](#) 0.14.4.

API reference

This page gives an overview of all public pandas objects, functions and methods. All classes and functions exposed in `pandas.*` namespace are public.

The following subpackages are public.

- `pandas.errors`: Custom exception and warnings classes that are raised by pandas.
- `pandas.plotting`: Plotting public API.
- `pandas.testing`: Functions that are useful for writing tests involving pandas objects.
- `pandas.api.extensions`: Functions and classes for extending pandas objects.
- `pandas.api.indexers`: Functions and classes for rolling window indexers.
- `pandas.api.interchange`: DataFrame interchange protocol.
- `pandas.api.types`: Datatype classes and functions.
- `pandas.api.typing`: Classes that may be necessary for type-hinting. These are classes that are encountered as intermediate results but should not be instantiated directly by users. These classes are not to be confused with classes from the [pandas-stubs](#) package which has classes in addition to those that occur in pandas for type-hinting.

In addition, public functions in `pandas.io` and `pandas.tseries` submodules are mentioned in the documentation.

⚠ Warning

The `pandas.core`, `pandas.compat`, and `pandas.util` top-level modules are PRIVATE. Stable functionality in such modules is not guaranteed.

Input/output

[Pickling](#)

[Flat file](#)

[Clipboard](#)

[Excel](#)

[HTML](#)[XML](#)[Latex](#)[HDFStore: PyTables \(HDF5\)](#)[Feather](#)[Parquet](#)[ORC](#)[SAS](#)[SPSS](#)[SQL](#)[Google BigQuery](#)[STATA](#)

[General functions](#)

[Data manipulations](#)[Top-level missing data](#)[Top-level dealing with numeric data](#)[Top-level dealing with datetimelike data](#)[Top-level dealing with Interval data](#)[Top-level evaluation](#)[Datetime formats](#)[Hashing](#)[Importing from other DataFrame libraries](#)

[Series](#)

[Constructor](#)[Attributes](#)[Conversion](#)[Indexing, iteration](#)[Binary operator functions](#)[Function application, GroupBy & window](#)[Computations / descriptive stats](#)[Reindexing / selection / label manipulation](#)[Missing data handling](#)

[Combining / comparing / joining / merging](#)

[Time Series-related](#)

[Accessors](#)

[Plotting](#)

[Serialization / IO / conversion](#)

[DataFrame](#)

[Constructor](#)

[Attributes and underlying data](#)

[Conversion](#)

[Indexing, iteration](#)

[Binary operator functions](#)

[Function application, GroupBy & window](#)

[Computations / descriptive stats](#)

[Reindexing / selection / label manipulation](#)

[Missing data handling](#)

[Reshaping, sorting, transposing](#)

[Combining / comparing / joining / merging](#)

[Time Series-related](#)

[Flags](#)

[Metadata](#)

[Plotting](#)

[Sparse accessor](#)

[Serialization / IO / conversion](#)

[pandas arrays, scalars, and data types](#)

[Objects](#)

[Utilities](#)

[Index objects](#)

[Index](#)

[Numeric Index](#)

[CategoricalIndex](#)

[IntervalIndex](#)

[MultiIndex](#)

[Skip to main content](#)

[TimedeltaIndex](#)

[PeriodIndex](#)

[Date offsets](#)

[DateOffset](#)

[BusinessDay](#)

[BusinessHour](#)

[CustomBusinessDay](#)

[CustomBusinessHour](#)

[MonthEnd](#)

[MonthBegin](#)

[BusinessMonthEnd](#)

[BusinessMonthBegin](#)

[CustomBusinessMonthEnd](#)

[CustomBusinessMonthBegin](#)

[SemiMonthEnd](#)

[SemiMonthBegin](#)

[Week](#)

[WeekOfMonth](#)

[LastWeekOfMonth](#)

[BQuarterEnd](#)

[BQuarterBegin](#)

[QuarterEnd](#)

[QuarterBegin](#)

[BYearEnd](#)

[BYearBegin](#)

[YearEnd](#)

[YearBegin](#)

[FY5253](#)

[FY5253Quarter](#)

[Easter](#)

[Tick](#)

[Day](#)

[Hour](#)[Minute](#)[Second](#)[Milli](#)[Micro](#)[Nano](#)

Frequencies

[pandas.tseries.frequencies.to_offset](#)

Window

[Rolling window functions](#)[Weighted window functions](#)[Expanding window functions](#)[Exponentially-weighted window functions](#)[Window indexer](#)

GroupBy

[Indexing, iteration](#)[Function application helper](#)[Function application](#)[**DataFrameGroupBy** computations / descriptive stats](#)[**SeriesGroupBy** computations / descriptive stats](#)[Plotting and visualization](#)

Resampling

[Indexing, iteration](#)[Function application](#)[Upsampling](#)[Computations / descriptive stats](#)

Style

[Styler constructor](#)[Styler properties](#)[Style application](#)[Builtin styles](#)[Style export and import](#)

[pandas.plotting.andrews_curves](#)
[pandas.plotting.autocorrelation_plot](#)
[pandas.plotting.bootstrap_plot](#)
[pandas.plotting.boxplot](#)
[pandas.plotting.deregister_matplotlib_converters](#)
[pandas.plotting.lag_plot](#)
[pandas.plotting.parallel_coordinates](#)
[pandas.plotting.plot_params](#)
[pandas.plotting.radviz](#)
[pandas.plotting.register_matplotlib_converters](#)
[pandas.plotting.scatter_matrix](#)
[pandas.plotting.table](#)

Options and settings

[Working with options](#)
[Numeric formatting](#)

Extensions

[pandas.api.extensions.register_extension_dtype](#)
[pandas.api.extensions.register_dataframe_accessor](#)
[pandas.api.extensions.register_series_accessor](#)
[pandas.api.extensions.register_index_accessor](#)
[pandas.api.extensions.ExtensionDtype](#)
[pandas.api.extensions.ExtensionArray](#)
[pandas.arrays.NumpyExtensionArray](#)
[pandas.api.indexers.check_array_indexer](#)

Testing

[Assertion functions](#)
[Exceptions and warnings](#)
[Bug report function](#)
[Test suite runner](#)

Missing values

[pandas.NA](#)

< [pandas.NaT](#)

Next >

Input/output

[Skip to main content](#)

API reference

This page gives an overview of all public pandas objects, functions and methods. All classes and functions exposed in `pandas.*` namespace are public.

The following subpackages are public.

- `pandas.errors`: Custom exception and warnings classes that are raised by pandas.
- `pandas.plotting`: Plotting public API.
- `pandas.testing`: Functions that are useful for writing tests involving pandas objects.
- `pandas.api.extensions`: Functions and classes for extending pandas objects.
- `pandas.api.indexers`: Functions and classes for rolling window indexers.
- `pandas.api.interchange`: DataFrame interchange protocol.
- `pandas.api.types`: Datatype classes and functions.
- `pandas.api.typing`: Classes that may be necessary for type-hinting. These are classes that are encountered as intermediate results but should not be instantiated directly by users. These classes are not to be confused with classes from the [pandas-stubs](#) package which has classes in addition to those that occur in pandas for type-hinting.

In addition, public functions in `pandas.io` and `pandas.tseries` submodules are mentioned in the documentation.

⚠ Warning

The `pandas.core`, `pandas.compat`, and `pandas.util` top-level modules are PRIVATE. Stable functionality in such modules is not guaranteed.

Input/output

[Pickling](#)

[Flat file](#)

[Clipboard](#)

[Excel](#)

[HTML](#)[XML](#)[Latex](#)[HDFStore: PyTables \(HDF5\)](#)[Feather](#)[Parquet](#)[ORC](#)[SAS](#)[SPSS](#)[SQL](#)[Google BigQuery](#)[STATA](#)

[General functions](#)

[Data manipulations](#)[Top-level missing data](#)[Top-level dealing with numeric data](#)[Top-level dealing with datetimelike data](#)[Top-level dealing with Interval data](#)[Top-level evaluation](#)[Datetime formats](#)[Hashing](#)[Importing from other DataFrame libraries](#)

[Series](#)

[Constructor](#)[Attributes](#)[Conversion](#)[Indexing, iteration](#)[Binary operator functions](#)[Function application, GroupBy & window](#)[Computations / descriptive stats](#)[Reindexing / selection / label manipulation](#)[Missing data handling](#)

[Combining / comparing / joining / merging](#)

[Time Series-related](#)

[Accessors](#)

[Plotting](#)

[Serialization / IO / conversion](#)

[DataFrame](#)

[Constructor](#)

[Attributes and underlying data](#)

[Conversion](#)

[Indexing, iteration](#)

[Binary operator functions](#)

[Function application, GroupBy & window](#)

[Computations / descriptive stats](#)

[Reindexing / selection / label manipulation](#)

[Missing data handling](#)

[Reshaping, sorting, transposing](#)

[Combining / comparing / joining / merging](#)

[Time Series-related](#)

[Flags](#)

[Metadata](#)

[Plotting](#)

[Sparse accessor](#)

[Serialization / IO / conversion](#)

[pandas arrays, scalars, and data types](#)

[Objects](#)

[Utilities](#)

[Index objects](#)

[Index](#)

[Numeric Index](#)

[CategoricalIndex](#)

[IntervalIndex](#)

[MultiIndex](#)

[Skip to main content](#)

[TimedeltaIndex](#)

[PeriodIndex](#)

[Date offsets](#)

[DateOffset](#)

[BusinessDay](#)

[BusinessHour](#)

[CustomBusinessDay](#)

[CustomBusinessHour](#)

[MonthEnd](#)

[MonthBegin](#)

[BusinessMonthEnd](#)

[BusinessMonthBegin](#)

[CustomBusinessMonthEnd](#)

[CustomBusinessMonthBegin](#)

[SemiMonthEnd](#)

[SemiMonthBegin](#)

[Week](#)

[WeekOfMonth](#)

[LastWeekOfMonth](#)

[BQuarterEnd](#)

[BQuarterBegin](#)

[QuarterEnd](#)

[QuarterBegin](#)

[BYearEnd](#)

[BYearBegin](#)

[YearEnd](#)

[YearBegin](#)

[FY5253](#)

[FY5253Quarter](#)

[Easter](#)

[Tick](#)

[Day](#)

[Hour](#)[Minute](#)[Second](#)[Milli](#)[Micro](#)[Nano](#)

Frequencies

[pandas.tseries.frequencies.to_offset](#)

Window

[Rolling window functions](#)[Weighted window functions](#)[Expanding window functions](#)[Exponentially-weighted window functions](#)[Window indexer](#)

GroupBy

[Indexing, iteration](#)[Function application helper](#)[Function application](#)[**DataFrameGroupBy** computations / descriptive stats](#)[**SeriesGroupBy** computations / descriptive stats](#)[Plotting and visualization](#)

Resampling

[Indexing, iteration](#)[Function application](#)[Upsampling](#)[Computations / descriptive stats](#)

Style

[Styler constructor](#)[Styler properties](#)[Style application](#)[Builtin styles](#)[Style export and import](#)

[pandas.plotting.andrews_curves](#)
[pandas.plotting.autocorrelation_plot](#)
[pandas.plotting.bootstrap_plot](#)
[pandas.plotting.boxplot](#)
[pandas.plotting.deregister_matplotlib_converters](#)
[pandas.plotting.lag_plot](#)
[pandas.plotting.parallel_coordinates](#)
[pandas.plotting.plot_params](#)
[pandas.plotting.radviz](#)
[pandas.plotting.register_matplotlib_converters](#)
[pandas.plotting.scatter_matrix](#)
[pandas.plotting.table](#)

Options and settings

[Working with options](#)
[Numeric formatting](#)

Extensions

[pandas.api.extensions.register_extension_dtype](#)
[pandas.api.extensions.register_dataframe_accessor](#)
[pandas.api.extensions.register_series_accessor](#)
[pandas.api.extensions.register_index_accessor](#)
[pandas.api.extensions.ExtensionDtype](#)
[pandas.api.extensions.ExtensionArray](#)
[pandas.arrays.NumpyExtensionArray](#)
[pandas.api.indexers.check_array_indexer](#)

Testing

[Assertion functions](#)
[Exceptions and warnings](#)
[Bug report function](#)
[Test suite runner](#)

Missing values

[pandas.NA](#)

< [pandas.NaT](#)

Next >

Input/output

[Skip to main content](#)

Getting started

Installation

Working with conda?

pandas is part of the [Anaconda](#) distribution and can be installed with Anaconda or Miniconda:

```
conda install -c conda-forge pandas
```

Prefer pip?

pandas can be installed via pip from [PyPI](#).

```
pip install pandas
```

In-depth instructions?

Installing a specific version? Installing from source? Check the advanced installation page.

[Learn more](#)

Intro to pandas



What kind of data does pandas handle?

[Straight to tutorial...](#)

How do I read and write tabular data?

[Straight to tutorial...](#)

How do I select a subset of a table?

[Straight to tutorial...](#)[Skip to main content](#)



How to create plots in pandas?

[Straight to tutorial...](#)



How to create new columns derived from existing columns?

[Straight to tutorial...](#)



How to calculate summary statistics?

[Straight to tutorial...](#)



How to reshape the layout of tables?

[Straight to tutorial...](#)



How to combine data from multiple tables?

[Straight to tutorial...](#)



How to handle time series data?

[Straight to tutorial...](#)



How to manipulate textual data?

[Straight to tutorial...](#)

Coming from...

Are you familiar with other software for manipulating tabular data? Learn the pandas-equivalent operations compared to software you already know:



The [R programming language](#) provides the `data.frame` data structure and multiple packages, such as [tidyverse](#) use and extend `data.frame` for convenient data handling functionalities similar to pandas.

[Learn more](#)



Already familiar to `SELECT`, `GROUP BY`, `JOIN`, etc.? Most of these SQL manipulations do have equivalents in pandas.

[Learn more](#)



The `data set` included in the [STATA](#) statistical software suite corresponds to the pandas `DataFrame`. Many of the operations known from STATA have an equivalent in pandas.

[Learn more](#)



Users of [Excel](#) or other spreadsheet programs will find that many of the concepts are transferrable to pandas.

[Learn more](#)



The [SAS](#) statistical software suite also provides the `data set` corresponding to the pandas `DataFrame`. Also SAS vectorized operations, filtering, string

[Skip to main content](#)

 > Getting started

Getting started

Installation

Working with conda?

pandas is part of the [Anaconda](#) distribution and can be installed with Anaconda or Miniconda:

```
conda install -c conda-forge pandas
```

Prefer pip?

pandas can be installed via pip from [PyPI](#).

```
pip install pandas
```

In-depth instructions?

Installing a specific version? Installing from source? Check the advanced installation page.

[Learn more](#)

Intro to pandas



What kind of data does pandas handle?

[Straight to tutorial...](#)



How do I read and write tabular data?

[Straight to tutorial...](#)



How do I select a subset of a table?

[Straight to tutorial...](#)

[Skip to main content](#)



How to create plots in pandas?

[Straight to tutorial...](#)



How to create new columns derived from existing columns?

[Straight to tutorial...](#)



How to calculate summary statistics?

[Straight to tutorial...](#)



How to reshape the layout of tables?

[Straight to tutorial...](#)



How to combine data from multiple tables?

[Straight to tutorial...](#)



How to handle time series data?

[Straight to tutorial...](#)



How to manipulate textual data?

[Straight to tutorial...](#)

Coming from...

Are you familiar with other software for manipulating tabular data? Learn the pandas-equivalent operations compared to software you already know:



The [R programming language](#) provides the `data.frame` data structure and multiple packages, such as [tidyverse](#) use and extend `data.frame` for convenient data handling functionalities similar to pandas.

[Learn more](#)



Already familiar to `SELECT`, `GROUP BY`, `JOIN`, etc.? Most of these SQL manipulations do have equivalents in pandas.

[Learn more](#)



The `data set` included in the [STATA](#) statistical software suite corresponds to the pandas `DataFrame`. Many of the operations known from STATA have an equivalent in pandas.

[Learn more](#)



Users of [Excel](#) or other spreadsheet programs will find that many of the concepts are transferrable to pandas.

[Learn more](#)



The [SAS](#) statistical software suite also provides the `data set` corresponding to the pandas `DataFrame`. Also SAS vectorized operations, filtering, string

[Skip to main content](#)

