

- [10 minutes to pandas](#)
- [Intro to data structures](#)
- [Essential basic functionality](#)
- [IO tools \(text, CSV, HDF5, ...\)](#)
- [Indexing and selecting data](#)
- [MultiIndex / advanced indexing](#)
- [Merge, join, concatenate and compare](#)
- [Reshaping and pivot tables](#)
- [Working with text data](#)
- [Working with missing data](#)
- [Duplicate Labels](#)
- [Categorical data](#)
- [Nullable integer data type](#)
- [Nullable Boolean data type](#)
- [Chart Visualization](#)
- [Table Visualization](#)
- [Computational tools](#)
- [Group by: split-apply-combine](#)
- [Windowing Operations](#)
- [Time series / date functionality](#)
- [Time deltas](#)
- [Options and settings](#)
- [Enhancing performance](#)
- [Scaling to large datasets](#)
- [Sparse data structures](#)
- [Frequently Asked Questions \(FAQ\)](#)
- [Cookbook](#)

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](#).

- [10 minutes to pandas](#)
 - [Object creation](#)
 - [Viewing data](#)
 - [Selection](#)
 - [Missing data](#)
 - [Operations](#)
 - [Merge](#)
 - [Grouping](#)
 - [Reshaping](#)
 - [Time series](#)
 - [Categoricals](#)
 - [Plotting](#)
 - [Getting data in/out](#)
 - [Gotchas](#)
- [Intro to data structures](#)
 - [Series](#)
 - [DataFrame](#)
- [Essential basic functionality](#)
 - [Head and tail](#)
 - [Attributes and underlying data](#)
 - [Accelerated operations](#)
 - [Flexible binary operations](#)
 - [Descriptive statistics](#)
 - [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 \(.xlsb\) files](#)
 - [Clipboard](#)
 - [Pickling](#)
 - [msgpack](#)
 - [HDF5 \(PyTables\)](#)
 - [Feather](#)
 - [Parquet](#)
 - [ORC](#)
 - [SQL queries](#)
 - [Google BigQuery](#)

- [Stata format](#)
- [SAS formats](#)
- [SPSS formats](#)
- [Other file formats](#)
- [Performance considerations](#)
- [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](#)
 - [Indexing with list with missing labels is deprecated](#)
 - [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](#)
- [MultiIndex / advanced indexing](#)
 - [Hierarchical indexing \(MultiIndex\)](#)
 - [Advanced indexing with hierarchical index](#)
 - [Sorting a MultiIndex](#)
 - [Take methods](#)
 - [Index types](#)
 - [Miscellaneous indexing FAQ](#)
- [Merge, join, concatenate and compare](#)
 - [Concatenating objects](#)
 - [Database-style DataFrame or named Series joining/merging](#)
 - [Timeseries friendly merging](#)
 - [Comparing objects](#)
- [Reshaping and pivot tables](#)
 - [Reshaping by pivoting DataFrame objects](#)
 - [Reshaping by stacking and unstacking](#)
 - [Reshaping by melt](#)
 - [Combining with stats and GroupBy](#)
 - [Pivot tables](#)
 - [Cross tabulations](#)
 - [Tiling](#)
 - [Computing indicator / dummy variables](#)
 - [Factorizing values](#)
 - [Examples](#)
 - [Exploding a list-like column](#)
- [Working with text data](#)
 - [Text data types](#)
 - [String methods](#)
 - [Splitting and replacing strings](#)
 - [Concatenation](#)
 - [Indexing with `.str`](#)
 - [Extracting substrings](#)
 - [Testing for strings that match or contain a pattern](#)
 - [Creating indicator variables](#)
 - [Method summary](#)
- [Working with missing data](#)
 - [Values considered "missing"](#)
 - [Inserting missing data](#)

- [Calculations with missing data](#)
- [Sum/prod of empties/nans](#)
- [NA values in GroupBy](#)
- [Filling missing values: fillna](#)
- [Filling with a PandasObject](#)
- [Dropping axis labels with missing data: dropna](#)
- [Interpolation](#)
- [Replacing generic values](#)
- [String/regular expression replacement](#)
- [Numeric replacement](#)
- [Experimental **NA** scalar to denote missing values](#)
- [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](#)
- [Nullable Boolean data type](#)
 - [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 HTML](#)
 - [Formatting the Display](#)
 - [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](#)
- [Computational tools](#)
 - [Statistical functions](#)
- [Group by: split-apply-combine](#)
 - [Splitting an object into groups](#)
 - [Iterating through groups](#)

- [Selecting a group](#)
- [Aggregation](#)
- [Transformation](#)
- [Filtration](#)
- [Dispatching to instance methods](#)
- [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](#)
 - [Frequency conversion](#)
 - [Attributes](#)
 - [TimedeltaIndex](#)
 - [Resampling](#)
- [Options and settings](#)
 - [Overview](#)
 - [Getting and setting options](#)
 - [Setting startup options in Python/IPython environment](#)
 - [Frequently used options](#)
 - [Available 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](#)
 - [Migrating](#)
 - [Interaction with `scipy.sparse`](#)
- [Frequently Asked Questions \(FAQ\)](#)
 - [DataFrame memory usage](#)
 - [Using if/truth statements with pandas](#)

- [Mutating with User Defined Function \(UDF\) methods](#)
- [NaN, Integer NA values and NA type promotions](#)
- [Differences with NumPy](#)
- [Thread-safety](#)
- [Byte-ordering issues](#)
- [Cookbook](#)
 - [Idioms](#)
 - [Selection](#)
 - [Multiindexing](#)
 - [Missing data](#)
 - [Grouping](#)
 - [Timeseries](#)
 - [Merge](#)
 - [Plotting](#)
 - [Data in/out](#)
 - [Computation](#)
 - [Timedeltas](#)
 - [Creating example data](#)

[<< Community tutorials](#)

[10 minutes to pandas >>](#)