

Data Compression and Archiving

Although modern computer systems have an ever increasing storage capacity, the growth of data being produced is unrelenting. Lossless compression algorithms make up for some of the shortfall in capacity by trading time spent compressing or decompressing data for the space needed to store it. Python includes interfaces to the most popular compression libraries so it can read and write files interchangeably.

[zlib](#) and [gzip](#) expose the GNU zip library, and [bz2](#) provides access to the more recent bzip2 format. Both formats work on streams of data, without regard to input format, and provide interfaces for reading and writing compressed files transparently. Use these modules for compressing a single file or data source.

The standard library also includes modules to manage *archive* formats, for combining several files into a single file that can be managed as a unit. [tarfile](#) reads and writes the Unix tape archive format, an old standard still widely used today because of its flexibility. [zipfile](#) works with archives based on the format popularized by the PC program PKZIP, originally used under MS-DOS and Windows, but now also used on other platforms because of the simplicity of its API and portability of the format.

- [zlib](#) — GNU zlib Compression
- [gzip](#) — Read and Write GNU zip Files
- [bz2](#) — bzip2 Compression
- [tarfile](#) — Tar Archive Access
- [zipfile](#) — ZIP Archive Access

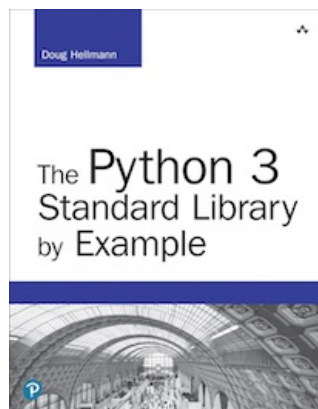
[← csv — Comma-separated Value Files](#)

[zlib — GNU zlib Compression →](#)

This page was last updated 2016-12-03.

Navigation

- [← csv — Comma-separated Value Files](#)
- [zlib — GNU zlib Compression](#)



[Get the book](#)

The output from all the example programs from PyMOTW-3 has been generated with Python 3.7.1, unless otherwise noted. Some of the features described here may not be available in earlier versions of Python.

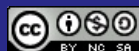
Looking for [examples for Python 2?](#)

This Site

- [Module Index](#)
- [Index](#)



© Copyright 2019, Doug Hellmann



Other Writing

 [Blog](#)

 [The Python Standard Library By Example](#)