



Varying File Compression Explored

- WinZip

This article adds to the information provided in [Why don't some files compress very much?](#)

Along with the inherent differences of one file type compared to another file type, files of the same type (such as two text files) will often compress by different amounts. When a file is being zipped, the type of information in that file and how that data is formatted will make it easier or harder to compress, regardless of the compression method you choose. Also, the size of a file will affect how much it can be compressed. For example, small files often contain little data with the result that there is not much available to compress.

The following is a table exploring some of these differences. The table displays the results of a test involving fairly large sets of files of certain types. The files also vary widely in their content. Twelve Zip files were created from six different file types. These sets of files were each zipped once using Legacy (Deflate) compression (.zip file) and once using Best method compression (.zipx file). The last row in the table displays information regarding zipping all six sets of files into one .zip file and one .zipx file.

Compression Test 1	Deflate (.zip)		Best m. (.zipx)	
File type	% range	avg	% range	avg
Text file (.txt)	0% to 99%	73%	0% to 99%	81%
Microsoft Word (.doc)	2% to 98%	70%	2% to 99%	81%
Microsoft Excel (.xls)	22% to 92%	84%	23% to 97%	93%
Executable file(.exe)	0% to 92%	47%	0% to 95%	52%
Picture file (.jpg)	0% to 62%	1%	0% to 64%	21%
Dynamic Link Libraries (.dll)	34% to 95%	67%	47% to 97%	75%
Mixed set of files	0% to 99%	42%	0% to 99%	55%

Note: These tests are NOT meant to represent typical results. They are examples based on available large file sets from this office environment. All of the files of each type found in a network drive were zipped for the test Zip files above. Additionally, the following factors should be considered:

1. Microsoft Office 2013/2010/2007 files actually are Zip files with custom file extensions and will not compress well. For this test, only the earlier type Office files (.doc and .xls rather than .docx and .xlsx) were included.
2. Microsoft Word files and many Microsoft Excel files will often include embedded pictures and/or other images, which decrease how well those files will compress. Very few of the Office files on this computer included pictures.
3. Since all of the picture files (.jpg) on this computer were included in the test, there were quite a few very small files that will not compress well even when Best method is used. Additionally, there were a number of .jpg files for use as backgrounds or templates. These latter files, due to their simplicity (often just a solid color) compress much better than the average .jpg file.

This second table represents one other test involving two text files. Both text files being tested are 21.6 MB in size. However, the 21M_a.txt file has fairly typical information, with sentences of various lengths and little repetition, while the 21M_b.txt file has a short section copied over and over again. This shows how files of the same type and size can give different results when they are zipped.

Compression Test 2	Deflate (.zip)	Best m. (.zipx)
File	zipped size	zipped size
21M_a.txt	7.86 MB	4.77 MB
21M_b.txt	102.34 KB	37.15 KB

Compression Test 2 demonstrates that two files of the same type and that are the same size can compress by differing amounts. The simpler, more repetitive text file compresses significantly more than the other.

If you have any questions about this information, please submit a [Technical Support](#) ticket.

Related Articles:

[Default compression - What it is, how to change it](#)

[Additional Compression Methods Specification](#)

[File Compression: Everything You Need to Know](#)

[Settings Tab](#)

[What is the difference between "Legacy" and "Best method" compression?](#)

[What is a Zipx File?](#)

[Using Jpeg compression in WinZip](#)

[Why don't some files compress very much?](#)
