

Example	Storage Usage in Mbit	Entropy H		Compression System	Processes	Compressed Size	Average Code Word Length		Links to Examples			
Wordnik	20.5	4.25		Huffman	2565462	11	4.29		Wordnik	https://raw.githubusercontent.com/wordnik/wordlist/refs/heads/main/wordlist-20210729.txt		
Shakespeare	44.6	4.84		Shannon	2565500	12	4.7		Shakespeare	https://gutenberg.org/cache/epub/100/pg100.txt		
alice29	1.2	4.57		Fano	2565349	11.4	4.44		Remaining...	https://corpus.canterbury.ac.nz/descriptions/index.html		
asyoulk	1	4.86		Comparison with Wordnik								
cp	0.2	5.27										
fields	0.09	5.05		Compression System	Processes	Compressed Size	Average Code Word Length					
lct10	3.4	4.67		Huffman	5575693	27.2	4.88					
plrabn12	3.8	4.58		Shannon	5576446	29.6	5.32					
e.coli	37.1	2		Fano	5575268	27.4	4.91					
world192	19.8	5		Comparison with Shakespeare								
The number of processes is relative, different processes require a different amount of time and use different components, this number could be interpreted vaguely.				Compression System	Processes	Compressed Size	Average Code Word Length					
Compression size does not include the storage space required for the compression scheme, the larger the data to be compressed the less statistically relevant this size gets, for simplicity we won't add this but for smaller examples you should remember this fact				Huffman	152529	0.7	4.61					
For simplicity the storage usage for the examples are calculated by the number of characters multiplied by 8. This is comparable to the ASCII coded character set but mostly inaccurate due to the fact most systems use UTF8 or UTF32				Shannon	152871	0.8	5.11					
				Fano	152236	0.7	4.67					
				Comparison with alice29								
				Compression System	Processes	Compressed Size	Average Code Word Length					
				Huffman	129705	0.6	4.89					
				Shannon	129947	0.7	5.33					
				Fano	129432	0.6	4.92					
				Comparison with asyoulik								
				Compression System	Processes	Compressed Size	Average Code Word Length					
				Huffman	25735	0.1	5.31					
				Shannon	26066	0.1	5.86					
				Fano	25394	0.1	5.37					
				Comparison with cp								
				Compression System	Processes	Compressed Size	Average Code Word Length					
				Huffman	12123	0.06	5.09					
				Shannon	12461	0.06	5.56					
				Fano	11762	0.06	5.2					
				Comparison with fields								
				Compression System	Processes	Compressed Size	Average Code Word Length					
				Huffman	426969	2	4.7					
				Shannon	427358	2.2	5.21					
				Fano	426636	2.1	4.85					
				Comparison with lct10								
				Compression System	Processes	Compressed Size	Average Code Word Length					
				Huffman	470950	2.2	4.61					
				Shannon	471400	2.4	5.06					
				Fano	470629	2.2	4.73					
				Comparison with plrabn12								
				Compression System	Processes	Compressed Size	Average Code Word Length					
				Huffman	4638710	9.3	2					
				Shannon	4638704	11.6	2.49					
				Fano	4638697	9.3	2					
				Comparison with e.coli								
				Compression System	Processes	Compressed Size	Average Code Word Length					
				Huffman	2473960	12.5	5.04					
				Shannon	2474426	13.6	5.51					
				Fano	2473587	12.6	5.09					
				Comparison with world192								

Org.Storage to Avg Code Word Length	fields	cp	asyoulik	alice29	lct10	plrabn12	world192	Wordnik	e.coli	Shakespeare
Huffman		0.09	0.2	1	1.2	3.4	3.8	19.8	20.5	37.1
		5.09	5.31	4.89	4.61	4.7	4.61	5.04	4.29	4.88
Shannon		5.56	5.86	5.33	5.11	5.21	5.06	5.51	4.7	2.49
Fano		5.2	5.37	4.92	4.67	4.85	4.73	5.09	4.44	2
Entropy to Avg Code Word Length	e.coli	Wordnik	alice29	plrabn12	lct10	Shakespeare	asyoulik	world192	fields	cp
Huffman		2	4.25	4.57	4.58	4.67	4.84	4.86	5	5.05
		2	4.29	4.61	4.61	4.7	4.88	4.89	5.04	5.31
Shannon		2.49	4.7	5.11	5.06	5.21	5.32	5.33	5.51	5.56
Fano		2	4.44	4.67	4.73	4.85	4.91	4.92	5.09	5.37

