

Informational Document for Deciphering Coded Messages

How to solve a puzzle written in a cipher? A coded message where all the letters are replaced with other letters, hieroglyphs, or mathematical symbols is also called a substitution cipher (because each letter is substituted with an unknown symbol). One way to decipher such a message is to count how many times each new symbol appears in the text. This way, you can find the frequency of occurrences for each symbol in the message. If you are aware of the language in which the original text is written, you can try comparing the coded message with the most common letters in the language and swap the most frequent symbols with corresponding letters.

When comparing various randomly chosen books, articles, or other pieces of written material in English with sufficient length, similar patterns usually emerge: the most common symbol is space and the most common letters are E, T, A, and O, while the least common letters are J, X, Q, and Z. Depending on the text, the frequencies of letter appearances may vary slightly.

If you don't have a frequency table corresponding to the English language at hand, you can try creating one based on any written text. For example, you can create a frequency table based on the same text you are currently looking at. If you count all the letters, symbols and spaces in the text, there are 2947 of them. If you count all of the times the letter E appears in the text, there are 331 of them. This means that $(331/2947 * 100\% =) 11,23\%$ of the letters in the text are E's. Similarly, you can find the frequency of any other symbol. For example, L appears 117 times, or 3.97%, and Y appears on 66 occasions, or 2.24%. An empty space between adjacent words appears 513 times in this text.

To break the cipher, you need to create a frequency table for the coded message as well. Most likely, the most popular symbol in the cipher corresponds to the most popular symbol in the analyzed text. Therefore, you can try replacing the next popular letters with the correct ones as well. This does not work for all letters. The less frequently a certain symbol appears in the language, the more its frequency may differ in different written sources — especially in shorter texts.

Once the most popular letters have found their place, you need to start looking for patterns in familiar words. For example, if two symbols have been replaced with the correct letters in a three-letter word, you should consider which letter may fit in place of the third symbol. If you replace the third symbol with a suitable letter, you can read out a three-letter word. You can try replacing the substituted symbol with the same letter in other words as well. If logical words form elsewhere, the substitution has been correct. This way, you can gather new clues about the text and suitable letters.

When the entire text is deciphered, you have taken a step closer to solving the puzzle.

Symbol	Count	Percentage	Symbol	Count	Percentage	Symbol	Count	Percentage
/space/	507	17.20%	L	117	3.97%	W	43	1.46%
E	331	11.23%	H	114	3.87%	G	35	1.19%
T	245	8.31%	C	88	2.99%	B	33	1.12%
A	169	5.73%	M	68	2.31%	X	18	0.61%
O	164	5.56%	U	67	2.27%	Q	10	0.34%
R	155	5.26%	Y	66	2.24%	K	8	0.27%
S	150	5.09%	P	58	1.97%	V	8	0.27%
N	134	4.55%	D	56	1.90%	Z	6	0.20%
I	123	4.17%	F	48	1.63%	J	2	0.07%

DUMKTOALGMTAQFGILKOKMALKMGIKKEMFJMGFZUEMDGMTXFZAMALOKKM

LZUEOKEMQKAKOGMIKOMGKCFUEMTMGITCKCOTJAMCDOCZHTADUVMALKM

IHTUKAMKTOALMFUMTMVKFGATADFUTOSMFOXDAMUKKEGMAFMLTPKMTM

PKHFCDASMFJMTIIOFBDQTAKHSMALOKKMRDHFQKAKOGMIKOMGKCFUEMLFNM

QTUSMADQKGMJTGAKOMEFKGMTMGITCKCOTJAMQFPKMCQITOKEMAFMALKM

GFZUEMDUMKTOALGMTAQFGILKOK

Empty table for writing stuff down

Symbol				Symbol				Symbol		
/space/				L				W		
E				H				G		
T				C				B		
A				M				X		
O				U				Q		
R				Y				K		
S				P				V		
N				D				Z		
I				F				J		