**Letter frequency**

From Wikipedia, the free encyclopedia

Jump to: [navigation](https://en.wikipedia.org/wiki/Letter_frequency#mw-head), [search](https://en.wikipedia.org/wiki/Letter_frequency#p-search)

The **frequency of letters** in text has been studied for use in [cryptanalysis](https://en.wikipedia.org/wiki/Cryptanalysis), and [frequency analysis](https://en.wikipedia.org/wiki/Frequency_analysis) in particular, dating back to the Iraqi mathematician [Al-Kindi](https://en.wikipedia.org/wiki/Al-Kindi) (c. 801–873 AD), who formally developed the method (the ciphers breakable by this technique go back at least to the [Caesar cipher](https://en.wikipedia.org/wiki/Caesar_cipher) invented by [Julius Caesar](https://en.wikipedia.org/wiki/Julius_Caesar), so this method could have been explored in classical times).

[](https://en.wikipedia.org/wiki/File:CaliforniaJobCaseLayout.png)

[California Job Case](https://en.wikipedia.org/wiki/California_Job_Case)

Letter frequency analysis gained additional importance in Europe with the development of [movable type](https://en.wikipedia.org/wiki/Movable_type) in 1450 AD, where one must estimate the amount of type required for each letterform, as evidenced by the variations in letter compartment size in typographer's type cases.

Linguists use letter frequency analysis as a rudimentary technique for [language identification](https://en.wikipedia.org/wiki/Language_identification), where it's particularly effective as an indication of whether an unknown writing system is alphabetic, syllablic, or ideographic. For example, the Japanese [Hiragana](https://en.wikipedia.org/wiki/Hiragana) syllabary contains 46 distinct characters, which is more than most phonetic alphabets, e.g. the [Hawaiian alphabet](https://en.wikipedia.org/wiki/Hawaiian_alphabet) which has a mere 13 letters, or [English](https://en.wikipedia.org/wiki/English_language) which has 26.

No exact letter frequency distribution underlies a given language, since all writers write slightly differently. However, most languages have a characteristic distribution which is strongly apparent in longer texts. Even language change as extreme as from old English to modern English (regarded as mutually unintelligible) show strong trends in related letter frequencies: over a small sample of Biblical passages, from most frequent to least frequent, **enaid sorhm tgþlwu (æ)cfy ðbpxz** of old English compares to **eotha sinrd luymw fgcbp kvjqxz** of modern English, with the most extreme differences concerning letterforms not shared.[[1]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-1)

[Linotype machines](https://en.wikipedia.org/wiki/Linotype_machine) for the English language assumed the letter order, from most to least common, to be [**etaoin shrdlu**](https://en.wikipedia.org/wiki/Etaoin_shrdlu) **cmfwyp vbgkjq xz** based on the experience and custom of manual compositors. The equivalent for the French language was **elaoin sdrétu cmfhyp vbgwqj xz**.

[Modern International Morse code](https://en.wikipedia.org/wiki/Morse_code) (generally believed to have been developed by [Alfred Vail](https://en.wikipedia.org/wiki/Alfred_Vail) based on English-language letter frequencies of the 1830s) encodes the most frequent letters with the shortest symbols; arranging the Morse alphabet into groups of letters that require equal amounts of time to transmit, and then sorting these groups in increasing order, yields **e it san hurdm wgvlfbk opxcz jyq**. Similar ideas are used in modern [data-compression](https://en.wikipedia.org/wiki/Data_compression) techniques such as [Huffman coding](https://en.wikipedia.org/wiki/Huffman_coding).

Letter frequency was used by other telegraph systems, such as the [Murray Code](https://en.wikipedia.org/wiki/Murray_Code).

**Contents**

 [hide]

* [1 Introduction](https://en.wikipedia.org/wiki/Letter_frequency#Introduction)
* [2 Relative frequencies of letters in the English language](https://en.wikipedia.org/wiki/Letter_frequency#Relative_frequencies_of_letters_in_the_English_language)
* [3 Relative frequencies of the first letters of a word in the English language](https://en.wikipedia.org/wiki/Letter_frequency#Relative_frequencies_of_the_first_letters_of_a_word_in_the_English_language)
* [4 Relative frequencies of letters in other languages](https://en.wikipedia.org/wiki/Letter_frequency#Relative_frequencies_of_letters_in_other_languages)
* [5 See also](https://en.wikipedia.org/wiki/Letter_frequency#See_also)
* [6 References](https://en.wikipedia.org/wiki/Letter_frequency#References)
* [7 External links](https://en.wikipedia.org/wiki/Letter_frequency#External_links)

**Introduction[**[**edit**](https://en.wikipedia.org/w/index.php?title=Letter_frequency&action=edit&section=1)**]**

Letter frequencies, like [word frequencies](https://en.wikipedia.org/wiki/Word_frequencies), tend to vary, both by writer and by subject. One cannot write an essay about x-rays without using frequent Xs, and the essay will have an idiosyncratic letter frequency if the essay is about the frequent use of x-rays to treat zebras in Qatar. Different authors have habits which can be reflected in their use of letters. [Hemingway](https://en.wikipedia.org/wiki/Ernest_Hemingway)'s writing style, for example, is visibly different from [Faulkner](https://en.wikipedia.org/wiki/William_Faulkner)'s. Letter, [bigram](https://en.wikipedia.org/wiki/Bigram), [trigram](https://en.wikipedia.org/wiki/Trigram), word frequencies, word length, and sentence length can be calculated for specific authors, and used to prove or disprove authorship of texts, even for authors whose styles are not so divergent.

Accurate average letter frequencies can only be gleaned by analyzing a large amount of representative text. With the availability of modern computing and collections of large [text corpora](https://en.wikipedia.org/wiki/Corpus_linguistics), such calculations are easily made. Examples can be drawn from a variety of sources (press reporting, religious texts, scientific texts and general fiction) and there are differences especially for general fiction with the position of 'h' and 'i', with H becoming more common.

[Herbert S. Zim](https://en.wikipedia.org/wiki/Herbert_Zim), in his classic introductory cryptography text "Codes and Secret Writing", gives the English letter frequency sequence as "ETAON RISHD LFCMU GYPWB VKJXQ Z", the most common letter pairs as "TH HE AN RE ER IN ON AT ND ST ES EN OF TE ED OR TI HI AS TO", and the most common doubled letters as "LL EE SS OO TT FF RR NN PP CC".[[2]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-2)

Also, to note that different dialects of a language will also affect a letter's frequency. For example, an author in the United States would produce something where the letter 'z' is more common as words like "analyze", "apologize" or "recognize" contain the letter, whereas in British English they do not. This would highly affect the frequency of the letter 'z' as it is a rarely used letter elsewhere in the English language.[[3]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-3)

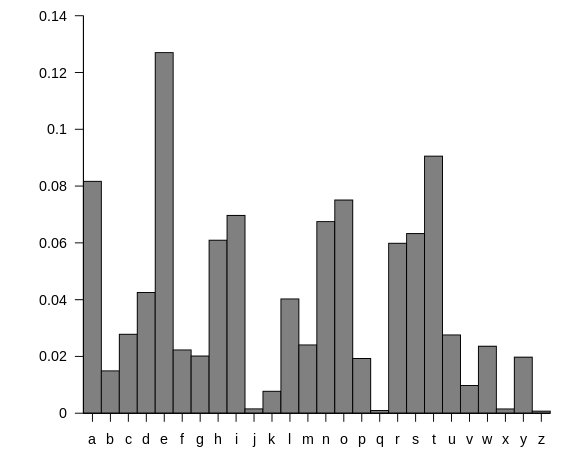
The "top twelve" letters constitute about 80% of the total usage. The "top eight" letters constitute about 65% of the total usage. Letter frequency as a function of rank can be fitted well by several rank functions, with the two-parameter Cocho/Beta rank function being the best.[[4]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-4) Another rank function with no adjustable free parameter also fits the letter frequency distribution reasonably well[[5]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-5) (the same function has been used to fit the amino acid frequency in protein sequences.[[6]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-6)) A spy using the [VIC cipher](https://en.wikipedia.org/wiki/VIC_cipher) or some other cipher based on a straddling checkerboard typically uses a mnemonic such as "a sin to err" (dropping the second "r")[[7]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-7)[[8]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-8) or "at one sir"[[9]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-9) to remember the top eight characters.

The use of letter frequencies and [frequency analysis](https://en.wikipedia.org/wiki/Frequency_analysis) plays a fundamental role in [cryptograms](https://en.wikipedia.org/wiki/Cryptograms) and several word puzzle games, including [Hangman](https://en.wikipedia.org/wiki/Hangman_(game)), [Scrabble](https://en.wikipedia.org/wiki/Scrabble) and the television game show [*Wheel of Fortune*](https://en.wikipedia.org/wiki/Wheel_of_Fortune_(U.S._game_show)). One of the earliest descriptions in classical literature of applying the knowledge of English letter frequency to solving a cryptogram is found in [E.A. Poe](https://en.wikipedia.org/wiki/Poe)'s famous story [*The Gold-Bug*](https://en.wikipedia.org/wiki/The_Gold-Bug), where the method is successfully applied to decipher a message instructing on the whereabouts of a treasure hidden by [Captain Kidd](https://en.wikipedia.org/wiki/William_Kidd).[[10]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-10)

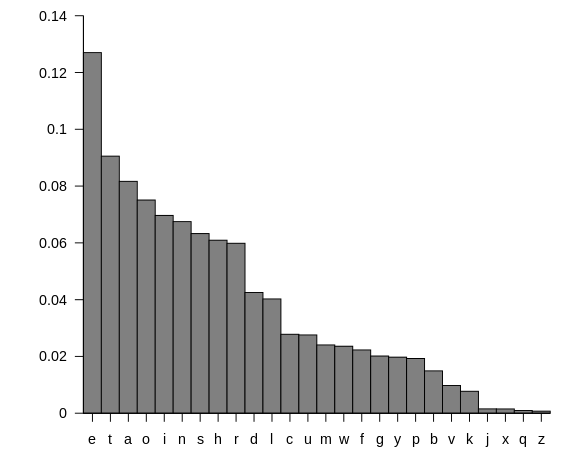
Letter frequencies had a strong effect on the design of some [keyboard layouts](https://en.wikipedia.org/wiki/Keyboard_layout). The most frequent letters are on the bottom row of the [Blickensderfer typewriter](https://en.wikipedia.org/wiki/Blickensderfer_typewriter), and the [home row](https://en.wikipedia.org/wiki/Home_row) of the [Dvorak Simplified Keyboard](https://en.wikipedia.org/wiki/Dvorak_Simplified_Keyboard).

**Relative frequencies of letters in the English language[**[**edit**](https://en.wikipedia.org/w/index.php?title=Letter_frequency&action=edit&section=2)**]**

There are three primary ways to count letter frequency that result in very different charts for common letters. The first method, used in the chart below, is to count letter frequency in root words of a dictionary. The second is to include all word variants when counting, such as "abstracts", "abstracted" and "abstracting" and not just the root word of "abstract", which results in letters like "s" appearing much more frequently. A final variant is to count letters based on their frequency of use in actual texts, resulting in certain letter combinations like "th" becoming more common due to the frequent use of common words like "the".

[](https://en.wikipedia.org/wiki/File:English_letter_frequency_(alphabetic).svg)

Relative frequencies of letters in text.

[](https://en.wikipedia.org/wiki/File:English_letter_frequency_(frequency).svg)

Relative frequencies ordered by frequency.

An analysis of entries in the Concise Oxford dictionary, ignoring frequency of word use, gives an order of "EARIOTNSLCUDPMHGBFYWKVXZJQ".[[11]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-11)

The letter-frequency table below is taken from Pavel Mička's website, which cites Robert Lewand's *Cryptological Mathematics*.[[12]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-12)

| **Letter** | **Relative frequency in the English language** | |
| --- | --- | --- |
| **a** | 8.167% | 8.167 |
| **b** | 1.492% | 1.492 |
| **c** | 2.782% | 2.782 |
| **d** | 4.253% | 4.253 |
| **e** | 12.702% | 12.702 |
| **f** | 2.228% | 2.228 |
| **g** | 2.015% | 2.015 |
| **h** | 6.094% | 6.094 |
| **i** | 6.966% | 6.966 |
| **j** | 0.153% | 0.153 |
| **k** | 0.772% | 0.772 |
| **l** | 4.025% | 4.025 |
| **m** | 2.406% | 2.406 |
| **n** | 6.749% | 6.749 |
| **o** | 7.507% | 7.507 |
| **p** | 1.929% | 1.929 |
| **q** | 0.095% | 0.095 |
| **r** | 5.987% | 5.987 |
| **s** | 6.327% | 6.327 |
| **t** | 9.056% | 9.056 |
| **u** | 2.758% | 2.758 |
| **v** | 0.978% | 0.978 |
| **w** | 2.360% | 2.36 |
| **x** | 0.150% | 0.15 |
| **y** | 1.974% | 1.974 |
| **z** | 0.074% | 0.074 |

According to Lewand, arranged from most to least common in appearance, the letters are: **etaoinshrdlcumwfgypbvkjxqz** Lewand's ordering differs slightly from others, such as Cornell University Math Explorer's Project, which produced a table after measuring 40,000 words.[[13]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-13)

In English, the space is slightly more frequent than the top letter (e) [[14]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-14) and the non-alphabetic characters (digits, punctuation, etc.) collectively occupy the fourth position (having already included the space) between *t* and *a*.[[15]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-15)

**Relative frequencies of the first letters of a word in the English language[**[**edit**](https://en.wikipedia.org/w/index.php?title=Letter_frequency&action=edit&section=3)**]**

The frequency of the first letters of words or names is helpful in pre-assigning space in physical files and indexes.[[16]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-ohlman-16) Given 26 [filing cabinet](https://en.wikipedia.org/wiki/Filing_cabinet) drawers, rather than a 1:1 assignment of one drawer to one letter of the alphabet, it is often useful to use a more equal-frequency-letter code by assigning several low-frequency letters to the same drawer (often one drawer is labeled VWXYZ), and to split up the most-frequent initial letters — S, A, and C - into several drawers (often 6 drawers Aa-An, Ao-Az, Ca-Cj, Ck-Cz, Sa-Si, Sj-Sz). The same system is used in some multi-volume works such as some [encyclopedias](https://en.wikipedia.org/wiki/Encyclopedia). [Cutter numbers](https://en.wikipedia.org/wiki/Cutter_number), another mapping of names to a more equal-frequency code, are used in some libraries.

The first letter of an English word, from most to least common, **t o a w b c d s f m r h i y e g l n p u v j k q z x**[[17]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-17)

Both the overall letter distribution and the word-initial letter distribution approximately match the [Zipf distribution](https://en.wikipedia.org/wiki/Zipf_distribution) and even more closely match the [Yule distribution](https://en.wikipedia.org/wiki/Yule_distribution).[[18]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-pande-18)

Often the frequency distribution of the first digit in each datum is significantly different from the overall frequency of all the digits in a set of numeric data—see [Benford's law](https://en.wikipedia.org/wiki/Benford%27s_law) for details.

|  |  |
| --- | --- |
| https://upload.wikimedia.org/wikipedia/commons/thumb/b/b4/Ambox_important.svg/60px-Ambox_important.svg.png | This section **possibly contains** [**original research**](https://en.wikipedia.org/wiki/Wikipedia:No_original_research). Please [improve it](https://en.wikipedia.org/w/index.php?title=Letter_frequency&action=edit) by [verifying](https://en.wikipedia.org/wiki/Wikipedia:Verifiability) the claims made and adding [inline citations](https://en.wikipedia.org/wiki/Wikipedia:Citing_sources#Inline_citations). Statements consisting only of original research should be removed. *(February 2014)* *(*[*Learn how and when to remove this template message*](https://en.wikipedia.org/wiki/Help:Maintenance_template_removal)*)* |

Analysis of a subset of [Project Gutenberg](https://en.wikipedia.org/wiki/Project_Gutenberg) text which shows the frequencies which various letters are found at the beginnings of words:[[19]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-19)

| **Letter** | **Relative frequency as the first letter of an English word** | |
| --- | --- | --- |
| **a** | 11.602% | 11.602 |
| **b** | 4.702% | 4.702 |
| **c** | 3.511% | 3.511 |
| **d** | 2.670% | 2.67 |
| **e** | 2.007% | 2.007 |
| **f** | 3.779% | 3.779 |
| **g** | 1.950% | 1.95 |
| **h** | 7.232% | 7.232 |
| **i** | 6.286% | 6.286 |
| **j** | 0.597% | 0.597 |
| **k** | 0.590% | 0.59 |
| **l** | 2.705% | 2.705 |
| **m** | 4.383% | 4.383 |
| **n** | 2.365% | 2.365 |
| **o** | 6.264% | 6.264 |
| **p** | 2.545% | 2.545 |
| **q** | 0.173% | 0.173 |
| **r** | 1.653% | 1.653 |
| **s** | 7.755% | 7.755 |
| **t** | 16.671% | 16.671 |
| **u** | 1.487% | 1.487 |
| **v** | 0.649% | 0.649 |
| **w** | 6.753% | 6.753 |
| **x** | 0.017% | 0.017 |
| **y** | 1.620% | 1.62 |
| **z** | 0.034% | 0.034 |

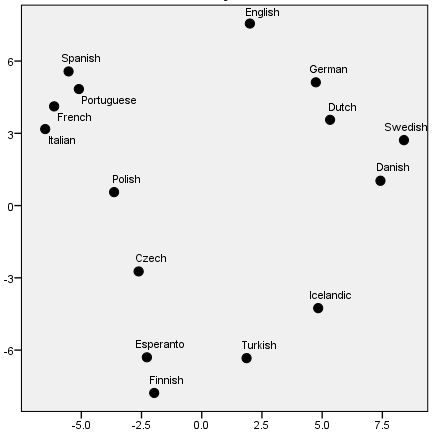
**Relative frequencies of letters in other languages[**[**edit**](https://en.wikipedia.org/w/index.php?title=Letter_frequency&action=edit&section=4)**]**

|  |  |
| --- | --- |
| [Text document with red question mark.svg](https://en.wikipedia.org/wiki/File:Text_document_with_red_question_mark.svg) | This article **possibly contains inappropriate or misinterpreted** [**citations**](https://en.wikipedia.org/wiki/Wikipedia:CITE) **that do not** [**verify**](https://en.wikipedia.org/wiki/Wikipedia:V) **the text**. Please help [improve this article](https://en.wikipedia.org/w/index.php?title=Letter_frequency&action=edit) by checking for citation inaccuracies. *(July 2014)* *(*[*Learn how and when to remove this template message*](https://en.wikipedia.org/wiki/Help:Maintenance_template_removal)*)* |

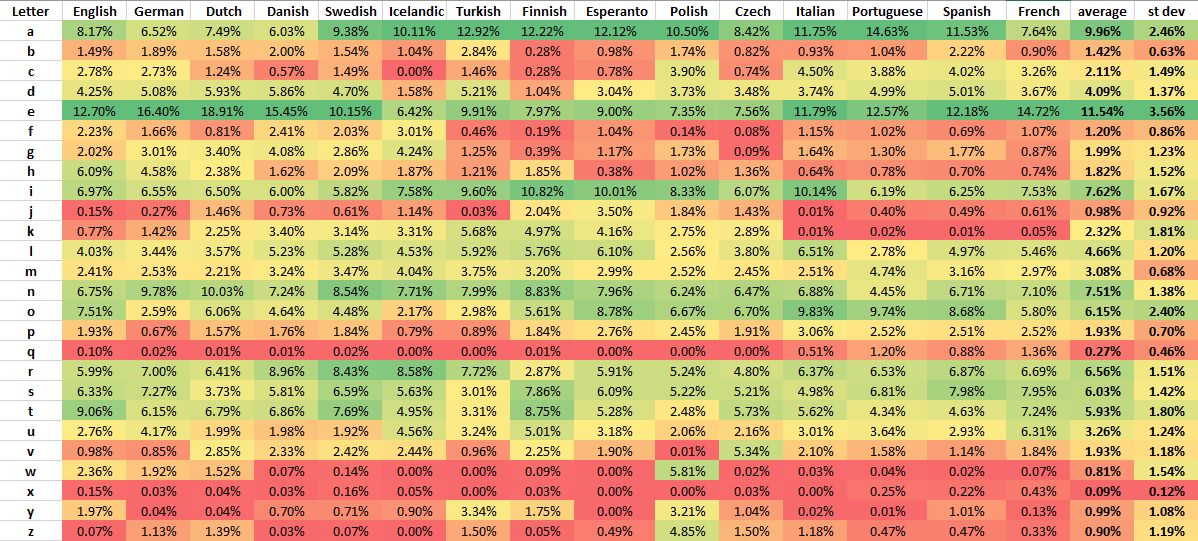
| **Letter** | [**English**](https://en.wikipedia.org/wiki/English_language) | [**French**](https://en.wikipedia.org/wiki/French_language)[[20]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-20) | [**German**](https://en.wikipedia.org/wiki/German_language)[[21]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-21) | [**Spanish**](https://en.wikipedia.org/wiki/Spanish_language)[[22]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-22) | [**Portuguese**](https://en.wikipedia.org/wiki/Portuguese_language)[[23]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-23) | [**Esperanto**](https://en.wikipedia.org/wiki/Esperanto_language)[[24]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-24) | [**Italian**](https://en.wikipedia.org/wiki/Italian_language)[[25]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-25) | [**Turkish**](https://en.wikipedia.org/wiki/Turkish_language)[[26]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-26) | [**Swedish**](https://en.wikipedia.org/wiki/Swedish_language)[[27]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-27) | [**Polish**](https://en.wikipedia.org/wiki/Polish_language)[[28]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-28) | [**Dutch**](https://en.wikipedia.org/wiki/Dutch_language)[[29]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-OnzeTaal-29) | [**Danish**](https://en.wikipedia.org/wiki/Danish_language)[[30]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-30) | [**Icelandic**](https://en.wikipedia.org/wiki/Icelandic_language)[[31]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-31) | [**Finnish**](https://en.wikipedia.org/wiki/Finnish_language)[[32]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-32) | [**Czech**](https://en.wikipedia.org/wiki/Czech_language) |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **a** | 8.167% | 7.636% | 6.516% | 11.525% | 14.634% | 12.117% | 11.745% | 12.920% | 9.383% | 10.503% | 7.486% | 6.025% | 10.110% | 12.217% | 8.421% |
| **b** | 1.492% | 0.901% | 1.886% | 2.215% | 1.043% | 0.980% | 0.927% | 2.844% | 1.535% | 1.740% | 1.584% | 2.000% | 1.043% | 0.281% | 0.822% |
| **c** | 2.782% | 3.260% | 2.732% | 4.019% | 3.882% | 0.776% | 4.501% | 1.463% | 1.486% | 3.895% | 1.242% | 0.565% | 0 | 0.281% | 0.740% |
| **d** | 4.253% | 3.669% | 5.076% | 5.010% | 4.992% | 3.044% | 3.736% | 5.206% | 4.702% | 3.725% | 5.933% | 5.858% | 1.575% | 1.043% | 3.475% |
| **e** | 12.702% | 14.715% | 16.396% | 12.181% | 12.570% | 8.995% | 11.792% | 9.912% | 10.149% | 7.352% | 18.91% | 15.453% | 6.418% | 7.968% | 7.562% |
| **f** | 2.228% | 1.066% | 1.656% | 0.692% | 1.023% | 1.037% | 1.153% | 0.461% | 2.027% | 0.143% | 0.805% | 2.406% | 3.013% | 0.194% | 0.084% |
| **g** | 2.015% | 0.866% | 3.009% | 1.768% | 1.303% | 1.171% | 1.644% | 1.253% | 2.862% | 1.731% | 3.403% | 4.077% | 4.241% | 0.392% | 0.092% |
| **h** | 6.094% | 0.737% | 4.577% | 0.703% | 0.781% | 0.384% | 0.636% | 1.212% | 2.090% | 1.015% | 2.380% | 1.621% | 1.871% | 1.851% | 1.356% |
| **i** | 6.966% | 7.529% | 6.550% | 6.247% | 6.186% | 10.012% | 10.143% | 9.600%\* | 5.817% | 8.328% | 6.499% | 6.000% | 7.578% | 10.817% | 6.073% |
| **j** | 0.153% | 0.613% | 0.268% | 0.493% | 0.397% | 3.501% | 0.011% | 0.034% | 0.614% | 1.836% | 1.46% | 0.730% | 1.144% | 2.042% | 1.433% |
| **k** | 0.772% | 0.049% | 1.417% | 0.011% | 0.015% | 4.163% | 0.009% | 5.683% | 3.140% | 2.753% | 2.248% | 3.395% | 3.314% | 4.973% | 2.894% |
| **l** | 4.025% | 5.456% | 3.437% | 4.967% | 2.779% | 6.104% | 6.510% | 5.922% | 5.275% | 2.564% | 3.568% | 5.229% | 4.532% | 5.761% | 3.802% |
| **m** | 2.406% | 2.968% | 2.534% | 3.157% | 4.738% | 2.994% | 2.512% | 3.752% | 3.471% | 2.515% | 2.213% | 3.237% | 4.041% | 3.202% | 2.446% |
| **n** | 6.749% | 7.095% | 9.776% | 6.712% | 4.446% | 7.955% | 6.883% | 7.987% | 8.542% | 6.237% | 10.032% | 7.240% | 7.711% | 8.826% | 6.468% |
| **o** | 7.507% | 5.796% | 2.594% | 8.683% | 9.735% | 8.779% | 9.832% | 2.976% | 4.482% | 6.667% | 6.063% | 4.636% | 2.166% | 5.614% | 6.695% |
| **p** | 1.929% | 2.521% | 0.670% | 2.510% | 2.523% | 2.755% | 3.056% | 0.886% | 1.839% | 2.445% | 1.57% | 1.756% | 0.789% | 1.842% | 1.906% |
| **q** | 0.095% | 1.362% | 0.018% | 0.877% | 1.204% | 0 | 0.505% | 0 | 0.020% | 0 | 0.009% | 0.007% | 0 | 0.013% | 0.001% |
| **r** | 5.987% | 6.693% | 7.003% | 6.871% | 6.530% | 5.914% | 6.367% | 7.722% | 8.431% | 5.243% | 6.411% | 8.956% | 8.581% | 2.872% | 4.799% |
| **s** | 6.327% | 7.948% | 7.270% | 7.977% | 6.805% | 6.092% | 4.981% | 3.014% | 6.590% | 5.224% | 3.73% | 5.805% | 5.630% | 7.862% | 5.212% |
| **t** | 9.056% | 7.244% | 6.154% | 4.632% | 4.336% | 5.276% | 5.623% | 3.314% | 7.691% | 2.475% | 6.79% | 6.862% | 4.953% | 8.750% | 5.727% |
| **u** | 2.758% | 6.311% | 4.166% | 2.927% | 3.639% | 3.183% | 3.011% | 3.235% | 1.919% | 2.062% | 1.99% | 1.979% | 4.562% | 5.008% | 2.160% |
| **v** | 0.978% | 1.838% | 0.846% | 1.138% | 1.575% | 1.904% | 2.097% | 0.959% | 2.415% | 0.012% | 2.85% | 2.332% | 2.437% | 2.250% | 5.344% |
| **w** | 2.360% | 0.074% | 1.921% | 0.017% | 0.037% | 0 | 0.033% | 0 | 0.142% | 5.813% | 1.52% | 0.069% | 0 | 0.094% | 0.016% |
| **x** | 0.150% | 0.427% | 0.034% | 0.215% | 0.253% | 0 | 0.003% | 0 | 0.159% | 0.004% | 0.036% | 0.028% | 0.046% | 0.031% | 0.027% |
| **y** | 1.974% | 0.128% | 0.039% | 1.008% | 0.006% | 0 | 0.020% | 3.336% | 0.708% | 3.206% | 0.035% | 0.698% | 0.900% | 1.745% | 1.043% |
| **z** | 0.074% | 0.326% | 1.134% | 0.467% | 0.470% | 0.494% | 1.181% | 1.500% | 0.070% | 4.852% | 1.39% | 0.034% | 0 | 0.051% | 1.503% |
| **à** | 0 | 0.486% | 0 | 0 | 0.072% | 0 | 0.635% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **â** | 0 | 0.051% | 0 | 0 | 0.562% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **á** | 0 | 0 | 0 | 0.502% | 0.118% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.799% | 0 | 0.867% |
| **å** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.338% | 0 | 0 | 1.190% | 0 | 0.003% | 0 |
| **ä** | 0 | 0 | 0.578% | 0 | 0 | 0 | 0 | 0 | 1.797% | 0 | 0 | 0 | 0 | 3.577% | 0 |
| **ã** | 0 | 0 | 0 | 0 | 0.733% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ą** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.699% | 0 | 0 | 0 | 0 | 0 |
| **æ** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.872% | 0.867% | 0 | 0 |
| **œ** | 0 | 0.018% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ç** | 0 | 0.085% | 0 | 0 | 0.530% | 0 | 0 | 1.156% | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ĉ** | 0 | 0 | 0 | 0 | 0 | 0.657% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ć** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.743% | 0 | 0 | 0 | 0 | 0 |
| **č** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.462% |
| **ď** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.015% |
| **ð** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4.393% | 0 | 0 |
| **è** | 0 | 0.271% | 0 | 0 | 0 | 0 | 0.263% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **é** | 0 | 1.504% | 0 | 0.433% | 0.337% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.647% | 0 | 0.633% |
| **ê** | 0 | 0.218% | 0 | 0 | 0.450% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ë** | 0 | 0.008% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ę** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.035% | 0 | 0 | 0 | 0 | 0 |
| **ě** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.222% |
| **ĝ** | 0 | 0 | 0 | 0 | 0 | 0.691% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ğ** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.125% | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ĥ** | 0 | 0 | 0 | 0 | 0 | 0.022% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **î** | 0 | 0.045% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ì** | 0 | 0 | 0 | 0 | 0 | 0 | 0.030% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **í** | 0 | 0 | 0 | 0.725% | 0.132% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.570% | 0 | 1.643% |
| **ï** | 0 | 0.005% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ı** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5.114%\* | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ĵ** | 0 | 0 | 0 | 0 | 0 | 0.055% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ł** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2.109% | 0 | 0 | 0 | 0 | 0 |
| **ñ** | 0 | 0 | 0 | 0.311% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ń** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.362% | 0 | 0 | 0 | 0 | 0 |
| **ň** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.007% |
| **ò** | 0 | 0 | 0 | 0 | 0 | 0 | 0.002% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ö** | 0 | 0 | 0.443% | 0 | 0 | 0 | 0 | 0.777% | 1.305% | 0 | 0 | 0 | 0.777% | 0.444% | 0 |
| **ô** | 0 | 0.023% | 0 | 0 | 0.635% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ó** | 0 | 0 | 0 | 0.827% | 0.296% | 0 | 0 | 0 | 0 | 1.141% | 0 | 0 | 0.994% | 0 | 0.024% |
| **ø** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.939% | 0 | 0 | 0 |
| **ř** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.380% |
| **ŝ** | 0 | 0 | 0 | 0 | 0 | 0.385% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ş** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.780% | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ś** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.814% | 0 | 0 | 0 | 0 | 0 |
| **š** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.688% |
| **ß** | 0 | 0 | 0.307% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ť** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.006% |
| **þ** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.455% | 0 | 0 |
| **ù** | 0 | 0.058% | 0 | 0 | 0 | 0 | 0.166% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ú** | 0 | 0 | 0 | 0.168% | 0.207% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.613% | 0 | 0.045% |
| **û** | 0 | 0.060% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ŭ** | 0 | 0 | 0 | 0 | 0 | 0.520% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ü** | 0 | 0 | 0.995% | 0.012% | 0.026% | 0 | 0 | 1.854% | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **ů** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.204% |
| **ý** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.228% | 0 | 0.995% |
| **ź** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.078% | 0 | 0 | 0 | 0 | 0 |
| **ż** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.706% | 0 | 0 | 0 | 0 | 0 |
| **ž** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.721% |

\*See [Dotted and dotless I](https://en.wikipedia.org/wiki/Dotted_and_dotless_I).

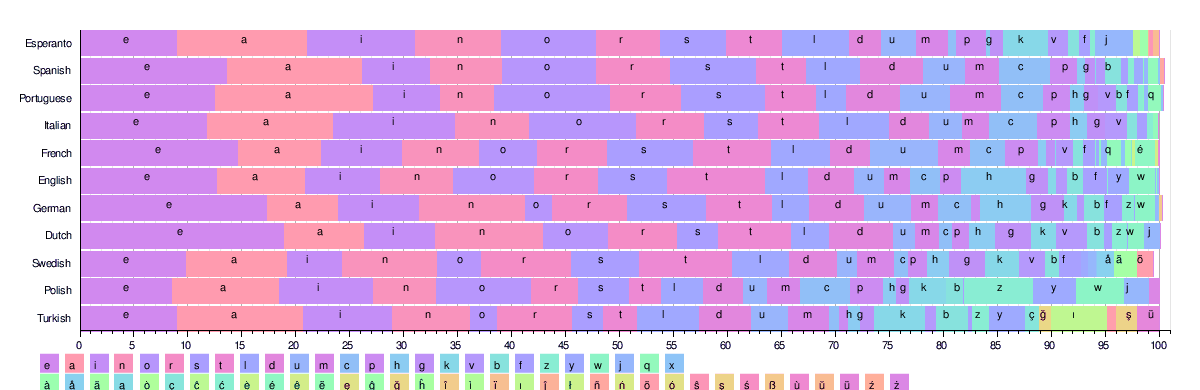
Leah Borovoi from the [InfinityLabs](http://infinitylabs.co.il/) calculated the relative frequencies of letters in the dictionaries of 15 languages. The data was analyzed using the [Mutidimensional Prefscal](https://en.wikipedia.org/wiki/Multidimensional_analysis) Procedure in [SPSS software](https://en.wikipedia.org/wiki/SPSS) to uncover the similarity between languages. The closer languages appear on the map, the more similar they are; the more distant languages appear on the map, the more dissimilar they are in terms of letter frequencies. The figure below illustrates the language map as shown in the Multidimensional Prefscal analysis. It confirms the usual family structure of languages (Romance, Germanic etc.).

[](https://en.wikipedia.org/wiki/File:SSA-language-analysis.jpg)

According to the Multidimensional Prefscal map the following table was developed, where close languages appear in the near columns. Red color represents low frequency, while green color represents high frequency of the letter in the language. It can be seen from the table that the most distinctive ordinary ABC letters between languages are a,e,o, i and l,n,t,d.

[](https://en.wikipedia.org/wiki/File:Letter_languages.jpg)

The figure below illustrates the frequency distributions of the 26 most common Latin letters across some languages. All of these languages use a similar 25+ character alphabet.



Based on these tables, the '[etaoin shrdlu](https://en.wikipedia.org/wiki/Etaoin_shrdlu)'-equivalent results for each language is as follows:

* French: 'esait nruol'; (Indo-European: Romance; traditionally, 'esartinulop' is used, in part for its ease of pronunciation[[33]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-33))
* Spanish: 'eaosr nidlt'; (Indo-European: Romance)
* Portuguese: 'aeosr idmnt' (Indo-European: Romance)
* Italian: 'eaion lrtsc'; (Indo-European: Romance)
* Esperanto: 'aieon lsrtk' (artificial language – influenced by Indo-European languages, Romance, Germanic mostly)
* German: 'enisr atdhu'; (Indo-European: Germanic)
* Swedish: 'eanrt sildo'; (Indo-European: Germanic)
* Turkish: 'aeinr lkdım'; (Altaic: Turkic)
* Dutch: 'enati rodsl'; (Indo-European: Germanic)[[29]](https://en.wikipedia.org/wiki/Letter_frequency#cite_note-OnzeTaal-29)
* Polish: 'aieon wrszc'; (Indo-European: Slavic)
* Danish: 'ernta idslo'; (Indo-European: Germanic)
* Icelandic: 'arnie stulð'; (Indo-European: Germanic)
* Finnish: 'ainte slouk'; (Uralic: Finnic)
* Czech: 'aeoni tvsrl'; (Indo-European: Slavic)

**See also[**[**edit**](https://en.wikipedia.org/w/index.php?title=Letter_frequency&action=edit&section=5)**]**

* [Corpus linguistics](https://en.wikipedia.org/wiki/Corpus_linguistics)
* [ETAOIN SHRDLU](https://en.wikipedia.org/wiki/ETAOIN_SHRDLU)
* [RSTLNE (Wheel of Fortune)](https://en.wikipedia.org/wiki/Wheel_of_Fortune_(U.S._game_show)#Bonus_round)
* [Frequency analysis (cryptanalysis)](https://en.wikipedia.org/wiki/Frequency_analysis_(cryptanalysis))
* [Linotype machine](https://en.wikipedia.org/wiki/Linotype_machine)
* [Most common words in English](https://en.wikipedia.org/wiki/Most_common_words_in_English)
* [Scrabble](https://en.wikipedia.org/wiki/Scrabble)
* [Arabic Letter Frequency](https://en.wikipedia.org/wiki/Arabic_Letter_Frequency)

**References[**[**edit**](https://en.wikipedia.org/w/index.php?title=Letter_frequency&action=edit&section=6)**]**

* 1. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-1) *Moreno, Marsha Lynn (Spring 2005).* [*"Frequency Analysis in Light of Language Innovation"*](http://www.math.ucsd.edu/~crypto/Projects/MarshaMoreno/TimeComparisonFrequency.pdf) *(PDF). Math UCSD. Retrieved 19 February 2015.*
  2. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-2) *Zim, Herbert Spencer. (1961). Codes & Secret Writing: Authorized Abridgement. Scholastic Book Services.* [*OCLC*](https://en.wikipedia.org/wiki/OCLC)[*317853773*](https://www.worldcat.org/oclc/317853773)*.*
  3. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-3) <http://www.oxforddictionaries.com/words/british-and-american-spelling>
  4. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-4) *Li, Wentian; Miramontes, Pedro (2011). "Fitting ranked English and Spanish letter frequency distribution in US and Mexican presidential speeches". Journal of Quantitative Linguistics.* ***18*** *(4): 359.* [*doi*](https://en.wikipedia.org/wiki/Digital_object_identifier)*:*[*10.1080/09296174.2011.608606*](https://dx.doi.org/10.1080%2F09296174.2011.608606)*.*
  5. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-5) [*Gusein-Zade, S.M.*](https://en.wikipedia.org/wiki/Sabir_Gusein-Zade) *(1988). "Frequency distribution of letters in the Russian language". Probl. Peredachi Inf.* ***24*** *(4): 102–7.*
  6. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-6) *Gamow, George; Ycas, Martynas (1955).* [*"Statistical correlation of protein and ribonucleic acid composition"*](http://www.pnas.org/content/41/12/1011.full.pdf) *(PDF). Proc. Natl. Acad. Sci.* ***41*** *(12): 1011–19.* [*doi*](https://en.wikipedia.org/wiki/Digital_object_identifier)*:*[*10.1073/pnas.41.12.1011*](https://dx.doi.org/10.1073%2Fpnas.41.12.1011)*.* [*PMC*](https://en.wikipedia.org/wiki/PubMed_Central)[*528190*](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC528190)*Freely accessible.* [*PMID*](https://en.wikipedia.org/wiki/PubMed_Identifier)[*16589789*](https://www.ncbi.nlm.nih.gov/pubmed/16589789)*.*
  7. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-7) Friedrich L. Bauer. ["Decrypted Secrets: Methods and Maxims of Cryptology"](https://books.google.com/books?id=hfWTDr_bvMwC). 2006. p. 57.
  8. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-8) Greg Goebel. ["The Rise Of Field Ciphers: straddling checkerboard ciphers"](http://www.vectorsite.net/ttcode_03.html) 2009.
  9. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-9) Dirk Rijmenants. ["One-time Pad"](http://users.telenet.be/d.rijmenants/en/onetimepad.htm)
  10. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-10) *Poe, Edgar Allan.* [*"The works of Edgar Allan Poe in five volumes"*](http://www.gutenberg.org/catalog/world/readfile?fk_files=1977099)*. Project Gutenberg.*
  11. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-11) [*"What is the frequency of the letters of the alphabet in English?"*](http://oxforddictionaries.com/words/what-is-the-frequency-of-the-letters-of-the-alphabet-in-english)*. Oxford Dictionary. Oxford University Press. Retrieved 29 December 2012.*
  12. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-12) *Mička, Pavel.* [*"Letter frequency (English)"*](http://en.algoritmy.net/article/40379/Letter-frequency-English)*. Algoritmy.net.*
  13. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-13) [*"Frequency Table"*](http://www.math.cornell.edu/~mec/2003-2004/cryptography/subs/frequencies.html)*. cornell.edu.*
  14. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-14) [*"Statistical Distributions of English Text"*](http://www.data-compression.com/english.html)*. data-compression.com.*
  15. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-15) *Lee, E. Stewart.* [*"Essays about Computer Security"*](http://www.cl.cam.ac.uk/~mgk25/lee-essays.pdf) *(PDF). University of Cambridge Computer Laboratory. p. 181.*
  16. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-ohlman_16-0) [Herbert Marvin Ohlman](https://en.wikipedia.org/wiki/Herbert_Marvin_Ohlman). "Subject-Word Letter Frequencies with Applications to [Superimposed Coding](https://en.wikipedia.org/wiki/Superimposed_code)". [[1]](http://books.nap.edu/openbook.php?record_id=10866&page=903) Proceedings of the International Conference on Scientific Information (1959).
  17. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-17) [*Singh, Simon*](https://en.wikipedia.org/wiki/Simon_Singh)*.* [*"The Black Chamber: Hints and Tips"*](http://www.simonsingh.net/The_Black_Chamber/hintsandtips.html)*. Retrieved 7 March 2016.*
  18. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-pande_18-0) Hemlata Pande and H. S. Dhami. ["Mathematical Modelling of Occurrence of Letters and Word’s Initials in Texts of Hindi Language"](http://www.skase.sk/Volumes/JTL16/pdf_doc/02.pdf).
  19. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-19) Calculated from "Project Gutenberg Selections" available from the [NLTK Corpora](http://nltk.googlecode.com/svn/trunk/nltk_data/index.xml)
  20. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-20) [*"CorpusDeThomasTempé"*](http://gpl.insa-lyon.fr/Dvorak-Fr/CorpusDeThomasTemp%C3%A9)*. Retrieved 2007-06-15.*
  21. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-21) *Beutelspacher, Albrecht (2005). Kryptologie (7 ed.). Wiesbaden: Vieweg. p. 10.* [*ISBN*](https://en.wikipedia.org/wiki/International_Standard_Book_Number)[*3-8348-0014-7*](https://en.wikipedia.org/wiki/Special:BookSources/3-8348-0014-7)*.*
  22. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-22) *Pratt, Fletcher (1942). Secret and Urgent: the Story of Codes and Ciphers. Garden City, N.Y.: Blue Ribbon Books. pp. 254–5.* [*OCLC*](https://en.wikipedia.org/wiki/OCLC)[*795065*](https://www.worldcat.org/oclc/795065)*.*
  23. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-23) [*"Frequência da ocorrência de letras no Português"*](http://www.numaboa.com/criptografia/criptoanalise/310-Frequencia-no-Portugues)*. Retrieved 2009-06-16.*
  24. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-24) [*"La Oftecoj de la Esperantaj Literoj"*](http://lingvakritiko.com/2007/09/13/literoftecoj-kaj-tabelvortoftecoj/)*. Retrieved 2007-09-14.*
  25. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-25) *Singh, Simon; Galli, Stefano (1999). Codici e Segreti (in Italian). Milano: Rizzoli.* [*ISBN*](https://en.wikipedia.org/wiki/International_Standard_Book_Number)[*978-8-817-86213-4*](https://en.wikipedia.org/wiki/Special:BookSources/978-8-817-86213-4)*.* [*OCLC*](https://en.wikipedia.org/wiki/OCLC)[*535461359*](https://www.worldcat.org/oclc/535461359)*.*
  26. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-26) Sefik Ilkin Serengil, Murat Akin. "[Attacking Turkish Texts Encrypted by Homophonic Cipher](http://www.wseas.us/e-library/conferences/2011/Cambridge/NEHIPISIC/NEHIPISIC-20.pdf)" Proceedings of the 10th WSEAS International Conference on Electronics, Hardware, Wireless and Optical Communications, pp.123-126, Cambridge, UK, February 20–22, 2011.
  27. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-27) [*"Practical Cryptography"*](http://practicalcryptography.com/cryptanalysis/letter-frequencies-various-languages/swedish-letter-frequencies/)*. Retrieved 2013-10-30.*
  28. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-28) [Wstęp do kryptologii](http://www.korzen.org/wsisiz/wstep%20do%20kryptologii/wyklady/Wykad%202.doc), counting [space] 17.2%, [dot point] 0.9%, [comma] 0.9% and [semicolon] 0.5%
  29. ^ [Jump up to: ***a***](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-OnzeTaal_29-0) [***b***](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-OnzeTaal_29-1) [*"Letterfrequenties"*](http://www.onzetaal.nl/advies/letterfreq.php)*. Genootschap OnzeTaal. Retrieved 2009-05-17.*
  30. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-30) [*"Practical Cryptography"*](http://practicalcryptography.com/cryptanalysis/letter-frequencies-various-languages/danish-letter-frequencies/)*. Retrieved 2013-10-24.*
  31. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-31) [*"Practical Cryptography"*](http://practicalcryptography.com/cryptanalysis/letter-frequencies-various-languages/icelandic-letter-frequencies/)*. Retrieved 2013-10-24.*
  32. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-32) [*"Practical Cryptography"*](http://practicalcryptography.com/cryptanalysis/letter-frequencies-various-languages/finnish-letter-frequencies/)*. Retrieved 2013-10-24.*
  33. [**Jump up ^**](https://en.wikipedia.org/wiki/Letter_frequency#cite_ref-33) Perec, Georges; *Alphabets*; Éditions Galilée, 1976

Notes

Some useful tables for single letter, digram, trigram, tetragram, and pentagram frequencies based on 20,000 words that take into account word-length and letter-position combinations for words 3 to 7 letters in length. The references are as follows:

1. *Mayzner, M.S.; Tresselt, M.E. (1965). "Tables of single-letter and digram frequency counts for various word-length and letter-position combinations". Psychonomic Monograph Supplements.* ***1*** *(2): 13–32.* [*OCLC*](https://en.wikipedia.org/wiki/OCLC)[*639975358*](https://www.worldcat.org/oclc/639975358)*.*
2. *Mayzner, M.S.; Tresselt, M.E.;Wolin, B.< R.< (1965). "Tables of trigram frequency counts for various word-length and letter-position combinations". Psychonomic Monograph Supplements.* ***1*** *(3): 33–78.*
3. *Mayzner, M.S.; Tresselt, M.E.;Woliin, B.< R,.. (1965). "Tables of tetragram frequency counts for various word-length and letter-position combinations". Psychonomic Monograph Supplements.* ***1*** *(4): 79–143.*
4. *Mayzner, M.S.; Tresselt, M.E.Wolin, B,.< R.> (1965). "Tables of pentagram frequency counts for various word-length and letter-position combinations". Psychonomic Monograph Supplements.* ***1*** *(5): 144–190.*

**External links[**[**edit**](https://en.wikipedia.org/w/index.php?title=Letter_frequency&action=edit&section=7)**]**

* [A site with content of *Cryptographical Mathematics*](http://pages.central.edu/emp/LintonT/) by Robert Edward Lewand
* [Some examples of letter frequency rankings in some common languages](http://www.bckelk.ukfsn.org/words/etaoin.html)
* [JavaScript Heatmap Visualization showing letter frequencies of texts on different keyboard layouts](http://www.patrick-wied.at/projects/heatmap-keyboard/)
* [An updated version of Mayzner's work using Google books Ngrams data set](http://norvig.com/mayzner.html) by Peter Norvig

<img src="//en.wikipedia.org/wiki/Special:CentralAutoLogin/start?type=1x1" alt="" title="" width="1" height="1" style="border: none; position: absolute;" />

Retrieved from "<https://en.wikipedia.org/w/index.php?title=Letter_frequency&oldid=767208605>"

[Categories](https://en.wikipedia.org/wiki/Help:Category):

* [Cryptography](https://en.wikipedia.org/wiki/Category:Cryptography)
* [Linguistics terminology](https://en.wikipedia.org/wiki/Category:Linguistics_terminology)
* [Quantitative linguistics](https://en.wikipedia.org/wiki/Category:Quantitative_linguistics)
* [Phonology](https://en.wikipedia.org/wiki/Category:Phonology)
* [Grammatology](https://en.wikipedia.org/wiki/Category:Grammatology)
* [Frequency distribution](https://en.wikipedia.org/wiki/Category:Frequency_distribution)

Hidden categories:

* [CS1 Italian-language sources (it)](https://en.wikipedia.org/wiki/Category:CS1_Italian-language_sources_(it))
* [Articles that may contain original research from February 2014](https://en.wikipedia.org/wiki/Category:Articles_that_may_contain_original_research_from_February_2014)
* [All articles that may contain original research](https://en.wikipedia.org/wiki/Category:All_articles_that_may_contain_original_research)
* [Articles lacking reliable references from July 2014](https://en.wikipedia.org/wiki/Category:Articles_lacking_reliable_references_from_July_2014)
* [All articles lacking reliable references](https://en.wikipedia.org/wiki/Category:All_articles_lacking_reliable_references)

**Navigation menu**

**Personal tools**

* Not logged in
* [Talk](https://en.wikipedia.org/wiki/Special:MyTalk)
* [Contributions](https://en.wikipedia.org/wiki/Special:MyContributions)
* [Create account](https://en.wikipedia.org/w/index.php?title=Special:CreateAccount&returnto=Letter+frequency)
* [Log in](https://en.wikipedia.org/w/index.php?title=Special:UserLogin&returnto=Letter+frequency)

**Namespaces**

* [Article](https://en.wikipedia.org/wiki/Letter_frequency)
* [Talk](https://en.wikipedia.org/wiki/Talk:Letter_frequency)

**Variants**

**Views**

* [Read](https://en.wikipedia.org/wiki/Letter_frequency)
* [Edit](https://en.wikipedia.org/w/index.php?title=Letter_frequency&action=edit)
* [View history](https://en.wikipedia.org/w/index.php?title=Letter_frequency&action=history)

**More**

**Search**

Top of Form



Bottom of Form

**Navigation**

* [Main page](https://en.wikipedia.org/wiki/Main_Page)
* [Contents](https://en.wikipedia.org/wiki/Portal:Contents)
* [Featured content](https://en.wikipedia.org/wiki/Portal:Featured_content)
* [Current events](https://en.wikipedia.org/wiki/Portal:Current_events)
* [Random article](https://en.wikipedia.org/wiki/Special:Random)
* [Donate to Wikipedia](https://donate.wikimedia.org/wiki/Special:FundraiserRedirector?utm_source=donate&utm_medium=sidebar&utm_campaign=C13_en.wikipedia.org&uselang=en)
* [Wikipedia store](https://shop.wikimedia.org/)

**Interaction**

* [Help](https://en.wikipedia.org/wiki/Help:Contents)
* [About Wikipedia](https://en.wikipedia.org/wiki/Wikipedia:About)
* [Community portal](https://en.wikipedia.org/wiki/Wikipedia:Community_portal)
* [Recent changes](https://en.wikipedia.org/wiki/Special:RecentChanges)
* [Contact page](https://en.wikipedia.org/wiki/Wikipedia:Contact_us)

**Tools**

* [What links here](https://en.wikipedia.org/wiki/Special:WhatLinksHere/Letter_frequency)
* [Related changes](https://en.wikipedia.org/wiki/Special:RecentChangesLinked/Letter_frequency)
* [Upload file](https://en.wikipedia.org/wiki/Wikipedia:File_Upload_Wizard)
* [Special pages](https://en.wikipedia.org/wiki/Special:SpecialPages)
* [Permanent link](https://en.wikipedia.org/w/index.php?title=Letter_frequency&oldid=767208605)
* [Page information](https://en.wikipedia.org/w/index.php?title=Letter_frequency&action=info)
* [Wikidata item](https://www.wikidata.org/wiki/Q520562)
* [Cite this page](https://en.wikipedia.org/w/index.php?title=Special:CiteThisPage&page=Letter_frequency&id=767208605)

**Print/export**

* [Create a book](https://en.wikipedia.org/w/index.php?title=Special:Book&bookcmd=book_creator&referer=Letter+frequency)
* [Download as PDF](https://en.wikipedia.org/w/index.php?title=Special:Book&bookcmd=render_article&arttitle=Letter+frequency&returnto=Letter+frequency&oldid=767208605&writer=rdf2latex)
* [Printable version](https://en.wikipedia.org/w/index.php?title=Letter_frequency&printable=yes)

**Languages**

* [Català](https://ca.wikipedia.org/wiki/Freq%C3%BC%C3%A8ncia_de_les_lletres)
* [Deutsch](https://de.wikipedia.org/wiki/Buchstabenh%C3%A4ufigkeit)
* [Español](https://es.wikipedia.org/wiki/Frecuencia_de_aparici%C3%B3n_de_letras)
* [Esperanto](https://eo.wikipedia.org/wiki/Literofteco)
* [Français](https://fr.wikipedia.org/wiki/Fr%C3%A9quence_d%27apparition_des_lettres_en_fran%C3%A7ais)
* [हिन्दी](https://hi.wikipedia.org/wiki/%E0%A4%B5%E0%A4%B0%E0%A5%8D%E0%A4%A3_%E0%A4%86%E0%A4%B5%E0%A5%83%E0%A4%A4%E0%A5%8D%E0%A4%A4%E0%A4%BF)
* [עברית](https://he.wikipedia.org/wiki/%D7%A9%D7%9B%D7%99%D7%97%D7%95%D7%AA_%D7%90%D7%95%D7%AA%D7%99%D7%95%D7%AA_%D7%91%D7%A2%D7%91%D7%A8%D7%99%D7%AA)
* [Português](https://pt.wikipedia.org/wiki/Frequ%C3%AAncia_de_letras)
* [Русский](https://ru.wikipedia.org/wiki/%D0%A7%D0%B0%D1%81%D1%82%D0%BE%D1%82%D0%BD%D0%BE%D1%81%D1%82%D1%8C)
* [Slovenščina](https://sl.wikipedia.org/wiki/Frekvence_%C4%8Drk)
* [Svenska](https://sv.wikipedia.org/wiki/Bokstavsfrekvens)
* [中文](https://zh.wikipedia.org/wiki/%E5%AD%97%E6%AF%8D%E9%A2%91%E7%8E%87)

[Edit links](https://www.wikidata.org/wiki/Q520562#sitelinks-wikipedia)

* This page was last modified on 24 February 2017, at 15:06.
* Text is available under the [Creative Commons Attribution-ShareAlike License](https://en.wikipedia.org/wiki/Wikipedia:Text_of_Creative_Commons_Attribution-ShareAlike_3.0_Unported_License); additional terms may apply. By using this site, you agree to the [Terms of Use](https://wikimediafoundation.org/wiki/Terms_of_Use) and [Privacy Policy](https://wikimediafoundation.org/wiki/Privacy_policy). Wikipedia® is a registered trademark of the [Wikimedia Foundation, Inc.](https://www.wikimediafoundation.org/), a non-profit organization.
* [Privacy policy](https://wikimediafoundation.org/wiki/Privacy_policy)
* [About Wikipedia](https://en.wikipedia.org/wiki/Wikipedia:About)
* [Disclaimers](https://en.wikipedia.org/wiki/Wikipedia:General_disclaimer)
* [Contact Wikipedia](https://en.wikipedia.org/wiki/Wikipedia:Contact_us)
* [Developers](https://www.mediawiki.org/wiki/Special:MyLanguage/How_to_contribute)
* [Cookie statement](https://wikimediafoundation.org/wiki/Cookie_statement)
* [Mobile view](https://en.m.wikipedia.org/w/index.php?title=Letter_frequency&mobileaction=toggle_view_mobile)
* [Wikimedia Foundation](https://wikimediafoundation.org/)
* [Powered by MediaWiki](https://www.mediawiki.org/)