

The following text has been taken from Wikipedia article on Comic Sans font (https://en.wikipedia.org/wiki/Comic_Sans).

Comic Sans

Comic Sans MS, commonly referred to as Comic Sans, is a sans-serif casual script typeface designed by Vincent Connare and released in 1994 by Microsoft Corporation. It is a casual, non-connecting script inspired by comic book lettering, intended for use in informal documents and children's materials.

The typeface has been supplied with Microsoft Windows since the introduction of Windows 95, initially as a supplemental font in the Windows Plus Pack and later in Microsoft Comic Chat. Describing it, Microsoft has explained that "this casual but legible face has proved very popular with a wide variety of people."^[1]

The typeface's widespread use, often in situations for which it was not intended, has been criticized.

History

Microsoft designer Vincent Connare began work on Comic Sans in October 1994. Connare had already created child-oriented fonts for various applications, so when he saw a beta version of Microsoft Bob that used Times New Roman in the word balloons of cartoon characters, he felt that the result was a formal look inappropriate for a program intended to introduce younger users to computers. His decision was to create a new face based on the lettering style of comic books he had in his office, specifically *The Dark Knight Returns* (lettered by John Costanza) and *Watchmen* (lettered by Dave Gibbons).^[3]

He completed the face too late for inclusion in MS Bob, but the programmers of Microsoft 3D Movie Maker, which also used cartoon guides and speech bubbles, began to use it. The speech bubbles eventually were phased out and replaced by actual sound, but Comic Sans stayed for the program's pop-up windows and help sections. The typeface later shipped with the Windows 95 Plus! Pack. It then became a standard font for the OEM version of Windows 95. Finally, the font became one of the default fonts for Microsoft Publisher and Microsoft Internet Explorer. The font is also used in Microsoft Comic Chat, which was released in 1996 with Internet Explorer 3.0.

Optical character recognition

Again, this text is from Wikipedia, this time block-aligned.

Optical character recognition (optical character reader, OCR) is the mechanical or electronic conversion of images of typed, handwritten or printed text into machine-encoded text, whether from a scanned document, a photo of a document, a scene-photo (for example the text on signs and billboards in a landscape photo) or from subtitle text superimposed on an image (for example from a television broadcast).[1] It is widely used as a form of information entry from printed paper data records, whether passport documents, invoices, bank statements, computerised receipts, business cards, mail, printouts of static-data, or any suitable documentation. It is a common method of digitising printed texts so that they can be electronically edited, searched, stored more compactly, displayed on-line, and used in machine processes such as cognitive computing, machine translation, (extracted) text-to-speech, key data and text mining. OCR is a field of research in pattern recognition, artificial intelligence and computer vision.

And some center aligned poetry:

Once upon a midnight dreary, while I pondered, weak and weary,
Over many a quaint and curious volume of forgotten lore—
While I nodded, nearly napping, suddenly there came a tapping,
As of some one gently rapping, rapping at my chamber door.
"Tis some visitor," I muttered, "tapping at my chamber door—
Only this and nothing more."

Ah, distinctly I remember it was in the bleak December;
And each separate dying ember wrought its ghost upon the floor.
Eagerly I wished the morrow;—vainly I had sought to borrow
From my books surcease of sorrow—sorrow for the lost Lenore—
For the rare and radiant maiden whom the angels name Lenore—
Nameless here for evermore.

This text serves as a basis for OCR Comic Sans dataset for FIT BUT School project, font size 16. It will be printed and scanned again. Here's a funny story from reddit.com:

So I'm hoping a load of people are going to come out in support of me here but I've got that sinking feeling I may be alone in this.

Our toilet broke so I was in shopping for new ones and the sales person joked (no doubt for the millionth time) that I'll want one that automatically puts the seat down after I'm finished with it. I 'joked' back and said if I didn't have a wife I could save money and not buy one with a seat and I'd never have to hear women complaining about putting it down again. To which he gave me a strange look and said "but what about when you need to poop?". I naturally pointed out that I'm a guy and therefore don't put the seat down, I sit on the rim of the bowl. Several embarrassing moments later, I realize that I've misunderstood my entire life and that guys do indeed use the toilet seat. I left empty handed and red faced.

Thinking about it now, it makes sense. Especially how men's restrooms have seats. But I just assumed it was a unisex/cost saving/oversight deal.

Here are some characters for debugging:

A B C 3 2 1 x c f h 7 A O O 9 6 B 8 W M m I l i L w Q . ? ! / _
ABC321xcfh7AOO96B8WMmIliLwQ.?!/_

EVERYTHING'S BIG

Lorem Ipsum is simply dummy text of the printing and typesetting industry. Lorem Ipsum has been the industry's standard dummy text ever since the 1500s, when an unknown printer took a galley of type and scrambled it to make a type specimen book. It has survived not only five centuries, but also the leap into electronic typesetting, remaining essentially unchanged. It was popularised in the 1960s with the release of Letraset sheets containing Lorem Ipsum passages, and more recently with desktop publishing software like Aldus PageMaker including versions of Lorem Ipsum.

Contrary to popular belief, Lorem Ipsum is not simply random text. It has roots in a piece of classical Latin literature from 45 BC, making it over 2000 years old. Richard McClintock, a Latin professor at Hampden-Sydney College in Virginia, looked up one of the more obscure Latin words, *consectetur*, from a Lorem Ipsum passage, and going through the cites of the word in classical literature, discovered the undoubtable source. Lorem Ipsum comes from sections 1.10.32 and 1.10.33 of "*de Finibus Bonorum et Malorum*" (The Extremes of Good and Evil) by Cicero, written in 45 BC. This book is a treatise on the theory of ethics, very popular during the Renaissance. The first line of Lorem Ipsum, "Lorem ipsum dolor sit amet..", comes from a line in section 1.10.32.

It is a long established fact that a reader will be distracted by the readable content of a page when looking at its layout. The point of using Lorem Ipsum is that it has a more-or-less normal distribution of letters, as opposed to using 'Content here, content here', making it look like readable English. Many desktop publishing packages and web page editors now use Lorem Ipsum as their default model text, and a search for 'lorem ipsum' will uncover many web sites still in their infancy. Various versions have evolved over the years, sometimes by accident, sometimes on purpose (injected humour and the like).

sizes 12, 13, 15 and different line spacing

Here's some text in Esperanto (without accents):

La Voko de Htulo

"El tiaj potencaĵoj aŭ estaĵoj eble estas restanto... restanto de
pratempo kiam... la konscio

manifestigis, saĝe, en formoj kaj bildoj longe malaperitaj antaŭ la
tjado marsanta de la

homaro... bildoj ke nur poeto kaj legendo kaptis iomete kaj nomis
diojn, monstrojn kaj mitajn

estaĵojn ĉiuspecajn..."

-Algernon Blackwood

And some text with big spaces. Each space is actually
three (3) spaces. This is a small text (10) from Wiki:

Computer vision is an interdisciplinary field that deals with how computers can be made to gain high-level understanding from digital images or videos. From the perspective of engineering, it seeks to automate tasks that the human visual system can do.

Computer vision tasks include methods for acquiring, processing, analyzing and understanding digital images, and in general, deal with the extraction of high-dimensional data from the real world in order to produce numerical or symbolic information, e.g., in the forms of decisions.[4][5][6][7] Understanding in this context means the transformation of visual images (the input of the retina) into descriptions of the world that can interface with other thought processes and elicit appropriate action. This image understanding can be seen as the disentangling of symbolic information from image data using models constructed with the aid of geometry, physics, statistics, and learning theory.[8]

Far far away, behind the word mountains, far from the countries Vokalia and Consonantia, there live the blind texts. Separated they live in Bookmarksgrove right at the coast of the Semantics, a large language ocean. A small river named Duden flows by their place and supplies it with the necessary regalia. It is a paradisematic country, in which roasted parts of sentences fly into your mouth. Even the all-powerful Pointing has no control about the blind texts it is an almost unorthographic life One day however a small line of blind text by the name of Lorem Ipsum decided to leave for the far World of Grammar. The Big Oxmox advised her not to do so, because there were thousands of bad Commas, wild Question Marks and devious Semikoli, but the Little Blind Text didn't listen. She packed her seven versalia, put her initial into the belt and made herself on the way. When she reached the first hills of the Italic Mountains, she had a last view back on the skyline of her hometown Bookmarksgrove, the headline of Alphabet Village and the subline of her own road, the Line Lane. Pityful a rethoric question ran over her cheek, then.

The quick, brown fox jumps over a lazy dog. DJs flock by when MTV ax quiz prog. Junk MTV quiz graced by fox whelps. Bawds jog, flick quartz, vex nymphs. Waltz, bad nymph, for quick jigs vex! Fox nymphs grab quick-jived waltz. Brick quiz whangs jumpy veldt fox. Bright vixens jump; dozy fowl quack. Quick wafting zephyrs vex bold Jim. Quick zephyrs blow, vexing daft Jim. Sex-charged fop blew my junk TV quiz. How quickly daft jumping zebras vex. Two driven jocks help fax my big quiz. Quick, Baz, get my woven flax jodhpurs! "Now fax quiz Jack!" my brave ghost pled. Five quacking zephyrs jolt my wax bed. Flummoxed by job, kvetching W. zaps Iraq. Cozy sphinx waves quart jug of bad milk. A very bad quack might jinx zippy fowls. Few quips galvanized the mock jury box. Quick brown dogs jump over the lazy fox. The jay, pig, fox, zebra, and my wolves quack! Blowzy red vixens fight for a quick jump. Joaquin Phoenix was gazed by MTV for luck. A wizard's job is to vex chumps quickly in fog. Watch "Jeopardy!", Alex Trebek's fun TV quiz game. Woven silk pyjamas exchanged for blue quartz. Brawny gods just.

A wonderful serenity has taken possession of my entire soul, like these sweet mornings of spring which I enjoy with my whole heart. I am alone, and feel the charm of existence in this spot, which was created for the bliss of souls like mine. I am so happy, my dear friend, so absorbed in the exquisite sense of mere tranquil existence, that I neglect my talents. I should be incapable of drawing a single stroke at the present moment; and yet I feel that I never was a greater artist than now. When, while the lovely valley teems with vapour around me, and the meridian sun strikes the upper surface of the impenetrable foliage of my trees, and but a few stray gleams steal into the inner sanctuary, I throw myself down among the tall grass by the trickling stream; and, as I lie close to the earth, a thousand unknown plants are noticed by me: when I hear the buzz of the little world among the stalks, and grow familiar with the countless indescribable forms of the insects and flies, then I feel the presence of the Almighty, who formed us in his own image, and the breath.

```

aaaaaaaaaaaaabbbbbbbbbbbbbbbccccccccccccccddddd
ddddddeeeeeeeeeeffffffffffffffffffggggggggggghhh
hhhhhhhhhhhhiiiiiiiiiiiiiiiijjjjjjjjjjjkkkkkkkkkklll
llllllllllllmmmmmmmmmmnnnnnnnnnnnoooooooooooooo
ppppppppppppppqqqqqqqqqqrrrrrrrrrrrrrrrrrrrrssss
ssssstttttttttttuuuuuuuuuuuvvvvvvvvvvvvvvvvvvv
wxxxxxxxxxxxxxxxxxyyyyyyyyyyyyyyyzzzzzzzzzzzzzz00000000
00011111111111222222222233333333333333334444444444555555
555555556666666666777777777788888888889999999999
9?????????????.....-!!!!

```

a a a a a a a a a a b b b b b b b b b b b b c c c c c c c c c c c c
d d d d d d d d d d d d d d d d e e e e e e e e e e f f f f f f f f f f
f f f f f f f f f g g g g g g g g g g h h h h h h h h h h i i i i i i i i
i i i i i i i i j j j j j j j j j k k k k k k k k k l l l l l l l l
l l l l l l m m m m m m m m m n n n n n n n n n o o o o o o o o
o o o o o o p p p p p p p p p q q q q q q q q r r r r r r r r
r r r r r r r r s s s s s s s s s t t t t t t t t u u u u u u u u
u u u u u v v v v v v v v w w w w w w x x x x x x x x
x x x x x y y y y y y z z z z z z z z z z z z z z 0 0 0 0
0 0 0 0 0 1 1 1 1 1 1 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3
4 4 4 4 4 4 4 5 5 5 5 5 5 6 6 6 6 6 6 7 7 7 7 7 7
7 7 7 8 8 8 8 8 8 9 9 9 9 9 9 ? ? ? ? ? ? ? ? ? ?
..... - - - - - ! ! ! !