Data representation

All data is represented as binary digits, whether it is numbers, text, images or sound. Calculations are also done in binary.

Units

In a computer, all data is stored in binary form. A binary digit has two possible states, 1 and 0.

A binary digit is known as a bit. A bit is the smallest unit of data a computer can use. The binary unit system is used to describe bigger numbers too.

Eight bits are known as a byte.

The binary unit system is as follows:

Size	Unit
8 bits	1 byte (B)
1,000 bytes (1,000 B)	1 kilobyte (KB)
1,000 kilobytes (1,000 KB)	1 megabyte (MB)
1,000 megabytes (1,000 MB)	1 gigabyte (GB)
1,000 gigabytes (1,000 GB)	1 terabyte (TB)
1,000 terabytes (1,000 TB)	1 petabyte (PB)

Four bits or half a byte is known as a nibble.

Analogue data and digital data

Analogue data is a real-life signal that can vary greatly in value. Examples include:

- sound waves
- pressure
- temperature

Digital data is binary data which represents analogue data. Computers work with digital data. Analogue data must be converted to digital before a computer can use it. A device known as an analogue-to-digital convertor (ADC) is used to generate digital data from analogue signals. In the same way, a digital signal can be converted back to an analogue signal using a digital-to-analogue convertor (DAC).

Binary and denary

Humans tend to use the denary number system. However, computers work in binary. Denary numbers must be converted into their binary equivalent before a computer can use them.

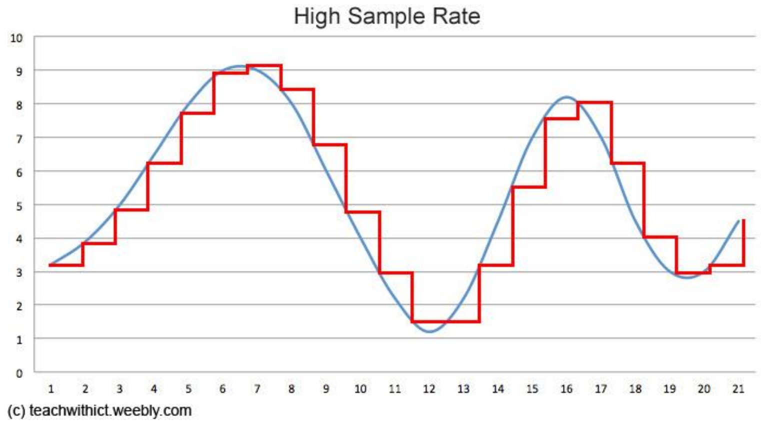
The denary system has ten digits (0, 1, 2, 3, 4, 5, 6, 7, 8 and 9). Each denary place value is calculated by multiplying the previous place value by ten.

Sound Representation

How can sound be sampled and stored in digital form?
A microphone converts sound waves into voltage changes. If a microphone is plugged into a sound card then the voltage can be sampled at regular intervals (the sample rate) and each value converted into a binary number. This digitising of the sound is carried out by the Analogue to Digital Convertor (ADC) on the sound card and the series of binary numbers can be stored as a sound file.

Sampling is therefore the process of measuring the sound level (as a voltage from a microphone) **at set intervals of time** (the sample interval) and storing the values as binary numbers.

The sound card can recreate the stored sound using a Digital to Analogue Convertor (DAC). This converts the series of binary numbers back into a changing voltage which can make a speaker (in a set of headphones or external speakers) vibrate to reproduce the sound.



What affects the size of a sound file and the playback quality?
Sampling rate/sample interval

The **sample rate** is the number of samples of the analogue sound wave taken per second. This frequency is measured in Hertz (Hz).
The **sample interval** is the time period between each sample. It is therefore **the reciprocal of the sample rate** – as the sample interval decreases, the sample rate increases.
For an audio compact disk (CD) the sampling rate is 44.1KHz or 44100 samples per second so the sample interval would be the reciprocal of this so about 0.024 Milliseconds. At this sample rate, 1 minute of audio would use 10Mb of memory (using 16 bit rate sampling).
The smaller the sample interval then the more often the sound is sampled and the closer the match between the original analogue sound wave and the digitised version.

However, higher sample rates need greater space on storage devices, need faster processors to manipulate the data and files will be slower to transfer over networks and the Internet.

A higher sample rate (lower sample period) gives a better quality recording but a larger file size.

Bit rate

The bit rate describes **the number of memory bits that are used to store each sound sample.**
If 1 bit is used then only 2 different levels of sound can be recorded and the sound would be unrecognisable when played back. If 3 bits (as illustrated in the animations above) are used then 8 different levels to be stored but the result would still be an extremely poor quality of digitised sound.
As the bit rate becomes higher, the number of **different** levels it is possible to record becomes greater and the closer the value stored in binary will be to the actual value that was sampled so the quality of the recording improves.
An audio compact disk (CD) uses **16 bit rate** sampling which in theory gives **65,536** different levels of sound, enough for the playback quality to be difficult or impossible to distinguish from the original analogue source.

A high bit rate gives a better quality recording but a larger file size.

File Compression

Both lossy and lossless file compression can be used with sound files although the former is far more effective in reducing file sizes.

The **greater the file compression** the **smaller the file size**. If the compression is lossy then the increased file compression will also result in a **lower sound quality**.