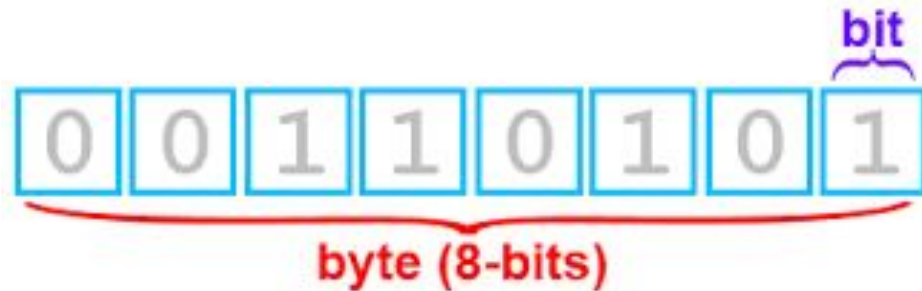# Unit 2: Digital Information

Lesson 1: Bytes and File Sizes

What is a byte?

A byte is a unit of data that is 8 bits long. A byte is the standard "chunk size" for binary information in most modern computers

Why 8? Much of the early computing world relied on representing and exchanging messages encoded in ASCII text, and so an 8-bit chunk, or "byte," became a very common chunk-size for representing information. It quickly became the fundamental unit with which we measure the "size" of data on computers.

Today most computers only let you save data as combinations of whole bytes; even if you only want to store 1 bit of information, you have to use a whole byte to do it.

Why the name byte?

There are various accounts about why it was called a "byte" but most point to early days at IBM where "bite" was used to to refer to groups of 8-bits that a computer was processing, as in it could "bite" off 8 bits at time. The spelling was changed to "byte" to avoid confusion with "bit".

How many bits and bytes?

REVIEW…

If we were to store the word `"hello"` in a computer in plain text ASCII format, how many <u>bits</u> would it take up?

How many <u>bytes </u>would that be?

Make a prediction...

In a prior lesson we learned that in addition to the actual text of a document, it is usually necessary to store the <u>formatting information </u>that allows the text to be displayed correctly. We might wonder just how much extra information, i.e. how many extra bytes, we need to store when we include all of this formatting.

Predict: How many times more bytes will a word processing program like Word require to store the word `"hello"`? Twice as much? Ten times as much? One hundred times as much?
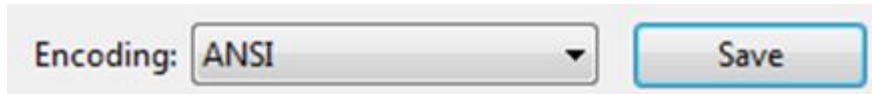
Activity #1: File Sizes

1) **Type the word "hello" in a plain text document. How large is the file?**

   PC/Windows: use **Notepad**

   Encoding: ANSI ▼ Save

   Mac: use **TextEdit** (Note: TextEdit needs to be switched into plain text mode from rich text. Go to Format → Make Plain Text)

   ANSWER:

2) **Now go type the word `hello` in a Microsoft Word. (If you have Word on your computer!) How large is this new, formatted file?**
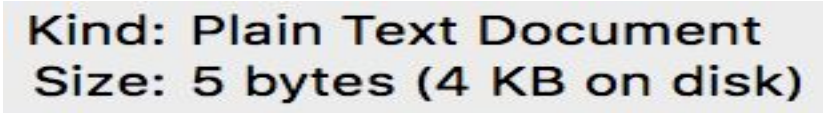
   ANSWER:

Activity #1: File Sizes

Look back at your predictions to see how close they were.

- The big difference in file size between `.txt` and `.docx` is due to the <u>extensive formatting information</u> included along with the actual text in `.docx`.

NOTE: A 5-byte file is so small that some computers won't allocate a chunk of memory that small. For example you might see something like this:

Kind: Plain Text Document
Size: 5 bytes (4 KB on disk)

Which indicates that even though the file is 5 bytes, it's taking up 4 Kilobytes of memory on your computer.

Sizes

Each time you go up a row, you are increasing by a factor of 1,024

Byte	Burger
KiloByte	King
MegaByte	Makes
GigaByte	Giant
TeraByte	Tasty
PetaByte	Potatoes (with)
ExaByte	Extra
ZettaByte	Zesty
YottaByte	Yams

Question	Answer
How many bytes are in 1 kilobyte?	
How many bytes are in 1 gigabyte?	
How many bytes are in 1 petabyte?	

Activity #3: How Big are the Files I Use Every Day?

File type	Size as # of pages, minutes, seconds, or dimensions	Size of file in Bytes, KB, MB, or GB, etc.
page of plain text (.txt)	About 500 words, or 2500 characters	2500 Bytes, 2.5KB
.jpg image		
animated .gif image		
.pdf file		
Audio file as .mp3		
movie file such as .mov or .mp4		

Larger Chunks

On modern computers the amount of information we can create and store has grown so large that we need new units of measurement to describe the size of our data.

Use these resources to fill in the table on the next slide. If you've done any work in the metric system, some of these new words might be familiar.

Use this websites for your research:

Computer Hope - How much is 1 byte, kilobyte, megabyte, gigabyte, etc.?
http://www.computerhope.com/issues/chspace.htm

Activity #2: File Sizes

Unit	Example of File Type or Data Measured in this Unit
Kilobyte (KB)	
Megabyte (MB)	
Gigabyte (GB)	

Unit	Example of File Type or Data Measured in this Unit
Terabyte (TB)	
Petabyte (PB)	
Exabyte (EB)	

SUMMARY

Early computers stored and ran 8-bit instructions. Much of the early computing world relied on representing and exchanging messages encoded in ASCII text, and so an 8-bit chunk, or "byte," became a very common chunk-size or unit of data for representing information. It quickly became the fundamental unit with which we measure the "size" of data on computers, and in fact, today most computers only let you save data as combinations of whole bytes; even if you only want to store 1 bit of information, you have to use a whole byte to do it.

The salesperson in a cell phone store is telling me that the phone I'm considering has 8GB of memory, which means I can save 10,000 photos taken with the phone's camera!

Is the salesperson telling me the truth? Why or why not?

Homework #2

Shakespeare's complete works have approximately 3.5 million characters. Which is bigger in file size: Shakespeare's complete works stored in plain ASCII text or a 4 minute song on mp3? How much bigger?

Homework #3

Alice has 600 MB of data. Bob has 2000 MB of data. Will it all fit on Alice's 4 GB thumb drive? Explain why or why not.

Alice has 100 small images, each of which is 500 KB. How much space do they take up overall in MB?

Homework #5

Your ghost hunting group is recording the sound inside a haunted classroom for 20 hours as MP3 audio files. About how much data will that be, expressed in GB?