Data compression is making the same amount of data into a smaller size for sending or storing.



**TASK 1**

Write your name down on the strips of paper provided.

To ‘compress’ the data, fold the paper along the lines neatly - if we want to ‘recognise’ and ‘uncompress’ the data then the uncompression needs to be the EXACT REVERSE of the compression.

What can you say about the ‘compressed’ version of your name?

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Why would this ‘compression’ of your data be useful when storing the data?

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Why would this ‘compression’ of your data be useful when sending the data to another computer or another user e.g. by e-mail?

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**TASK 2**

Using TWO A4 sheets of paper – write your name down on BOTH sheets.

Fold or ‘compress’ the data for one sheet in one way – REMEMBER… IT HAS TO BE NEAT!

Fold or ‘compress’ the data for the second sheet in a different way – REMEMBER… IT HAS TO BE NEAT!

You have invented two **different** ‘compression algorithms’!

**TASK 3**

Go to the website:

<http://www.bagill.com/text-converter.php>



Convert your name into binary.

**Run-length encoding** is a simple computer method of LOSSLESS compressing or encoding data.

When there is a pair or more than TWO bits the same (a ‘run length’ of common digits), we can compress those bits by writing how many - length - of those common bits (run) there are.

For example:

11000010 🡪 114010 - this means there is a string of four zeroes in the sequence

|  |  |
| --- | --- |
| **Data** | **Number of bits** |
| 11000010 | 8 bits |
| 214010 | 6 bits (‘compressed’) |

***EXAMPLE***

|  |  |  |
| --- | --- | --- |
| **MY NAME (ASCII characters)** | JOHN SMITH | **HOW MANY BITS?** |
| **MY NAME (in BINARY - uncompressed)** | 01001010011011110110100001101110001000000101001101101101011010010111010001101000 | 80 bits |
| **MY NAME (in BINARY – compressed)** | 0120101202104102101402103130160101202102102101021012010310130210130 | 67 bits= 17% compression |

\* COUNTING BITS IN THE ‘UNCOMPRESSED’ NAME 🡪 EACH LETTER = 1 BYTE = 8 BITS

**TASK 4**

Using the website:

 <http://www.bagill.com/text-converter.php>

- fill in the table below with:

 - your name

 - your name in BINARY

 - your name ‘compressed’ using **RUN-LENGTH ENCODING**

 - how many bits storage space your name takes UNCOMPRESSED

 - how many bits storage space your name takes COMPRESSED

 - the compression %age

|  |  |  |
| --- | --- | --- |
| **MY NAME (ASCII characters)** |  | **HOW MANY BITS?** |
| **MY NAME (in BINARY - uncompressed)** |  | \_\_\_ bits |
| **MY NAME (in BINARY – compressed)** |  | \_\_\_ bits= \_\_\_ % compression |

\* COUNTING BITS IN THE ‘UNCOMPRESSED’ NAME 🡪 EACH LETTER = 1 BYTE = 8 BITS