PROGRESS OUTCOME 6

# Why compression matters

## Context

Sarah has been investigating the concept of compression coding, by looking at the different ways images can be compressed and the trade-offs of using different methods of compression in relation to the size and quality of images. She researches the topic by doing several web searches and produces a report with her findings.

## Insight 1: Reasons for compressing files

I realised that data compression is extremely useful as it reduces the amount of space needed to store files. Computers and storage devices have limited space, so using smaller files allows you to store more files. Compressed files also download more quickly, which saves time – and money, because you pay less for electricity and bandwidth. Compression is also important in terms of HCI (human-computer interaction), because if images or videos are not compressed they take longer to download and, rather than waiting, many users will move on to another website.

## Insight 2: Representing images in an uncompressed form

Most people don't realise that a computer image is made up of tiny dots of colour, called pixels (short for "picture elements"). Each pixel has components of red, green and blue light and is represented by a binary number. The more bits used to represent each component, the greater the depth of colour in your image. For example, if red is represented by 8 bits, green by 8 bits and blue by 8 bits, 16,777,216 different colours can be represented, as shown below:

8 bits corresponds to 256 possible numbers in binary

red x green x blue = 256 x 256 x 256 = 16,777,216 possible colour combinations

## Insight 3: Lossy compression and human perception

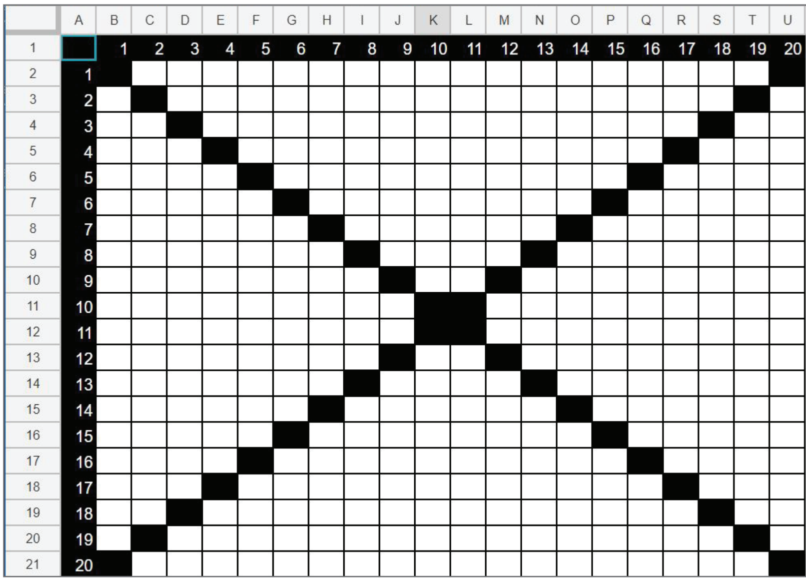
There are many different formats for file compression such as JPEG (image compression), MP3 (audio compression) and MPEG (audio and video compression). These formats use lossy compression, in which some of the original data is removed when the file is compressed. The benefit is a smaller file that downloads more quickly, but it does mean that you cannot convert the file back to its original full quality.

We use lossy compression because most image and sound files have more data than the human eye or ear can perceive and if data is removed it isn't usually noticeable. If we compress a file too much, people will start to notice – for example, an image may look pixelated or blurry. Sometimes it's not appropriate to compress an image because the quality is extremely important, such as in a medical scan.

## Insight 4: Lossless compression

Lossless compression compresses a file without removing any data. Instead a computer uses an algorithm to map the repeating patterns or information in the file (such as colours for an image or letters in a document). Most people use zip compression for general files. For images, lossless compression methods include PNG (Portable Network Graphics) and RLE (Run Length Encoding). Because lossless compression works by storing a map of repeating patterns of data, it is most effective when there is a lot of repeated data.

Below is an example of how lossless compression could work for a simple black and white image using RLE. The image on the left could be exactly replicated by a computer using the encoding map on the right. The computer has all the information it needs to recreate the image pixel by pixel (how many bits are black and how many bits are white).



| W | X | Y | Z | AA |
|---|---|---|---|---|
| Characters for encoding | | | | |
| 0 | 1 | 18 | 1 | |
| 1 | 1 | 16 | 1 | 1 |
| 2 | 1 | 14 | 1 | 2 |
| 3 | 1 | 12 | 1 | 3 |
| 4 | 1 | 10 | 1 | 4 |
| 5 | 1 | 8 | 1 | 5 |
| 6 | 1 | 6 | 1 | 6 |
| 7 | 1 | 4 | 1 | 7 |
| 8 | 1 | 2 | 1 | 8 |
| 9 | 2 | 0 | | |
| 9 | 2 | 0 | | |
| 8 | 1 | 2 | 1 | 8 |
| 7 | 1 | 4 | 1 | 7 |
| 6 | 1 | 6 | 1 | 6 |
| 5 | 1 | 8 | 1 | 5 |
| 4 | 1 | 10 | 1 | 4 |
| 3 | 1 | 12 | 1 | 3 |
| 2 | 1 | 14 | 1 | 2 |
| 1 | 1 | 16 | 1 | 1 |
| 0 | 1 | 18 | 1 | |


ISBN: 978-1-77669-237-8

MINISTRY OF EDUCATION
TE TĀHUHU O TE MĀTAURANGA