2025

Data & Encoding

Dan Collingbourne

ORB Education

Do This	Section
☐	1. Wordlist
☐	2. Data in our Lives
☐	3. Basic Data Tasks
☐	4. Searching the Web
☐	5. Data Authenticity
☐	6. Data Processing
☐	7. Surveys
☐	8. Binary Systems
☐	9. Binary Counting
☐	10. Bits and Bytes
☐	11. Text Characters
☐	12. Colours
☐	13. Bitmap Images
☐	14. Vector Graphics
☐	15. Audio Encoding
☐	16. Image Encoding

Computer data includes numbers, text, images, documents, sounds and code. Data is stored on hard drives, discs and other storage devices. It is transferred around the world via the cables and airways that form the internet. There is now an incredible amount of digital data in existence.

***Aim:** To learn about data and how it affects us.*

Task 1 – What Data do you Use?

Write down answers to the questions below. They do not need to be too accurate; approximations are fine.

a. How many photos do you have stored on your phone, computers and social media sites? ____________

b. Do you store music on your devices? If so, how many songs do you have? ____________

c. How many emails are sitting in all your email accounts? ____________

d. Roughly how many minutes of video do you watch through the internet each week? ____________

e. How many files containing schoolwork do you store, both at school and at home? ____________

f. What is the total storage capacity of all your devices and computer drives, in gigabytes? ____________

Task 2 – Big Data

There is so much data in the world now that many companies have outgrown the databases and other traditional methods used to organise it. This is *Big Data*.

Try and match the big numbers with the descriptions given. Look on the internet for help if needed.

	Number
1	500 million
2	1 billion
3	3 billion
4	5.6 billion
5	14 billion
6	512 billion
7	400 billion billion
8	149,000 billion billion

	Description
a	Number of people with access to the internet.
b	Images shared each day through social media.
c	Tweets per day.
d	Hours of content watched on YouTube every day.
e	Total bytes of data stored in the world.
f	Monthly active users on Facebook.
g	Bytes of data stored on a top end iPhone 16.
h	New bytes of data stored each day in the world.

***Note:** Numbers quoted from various sources as of 2024. Some are still growing quickly.*

Searching the Web

Aim: *To compare search engines and practice basic web searches.*

These days, you can generally type a few words into Google and it will turn up a good set of results for your search subject. This didn't happen accidentally. To provide this service, Google does the following:

- *It indexes hundreds of billions of webpages, collecting information about the text, images and other content. It also looks at how often the pages are updated.*
- *It measures the quality of each page by analysing its information. It also looks at the visitors, the technical structure and the other websites that link to it.*
- *It works out what you are actually searching for, predicting text as well as meaning. It achieves this by analysing huge numbers of searches and the behaviour of the people making them.*

Google generally returns results from its huge data banks in about 1/8th of a second.

Task 1 – A Search Race

Google isn't the only search engine. Other examples include *Bing* and *Duck Duck Go*. Your job as a class is to try and find out which of these three search engines is the fastest when it comes to locating answers to some general knowledge questions. Here is the plan:

vs vs

a. Divide the class into 3 groups. Everyone in one group should use *Google*, the second group all use *Bing* and the last group *Duck Duck Go*. The web addresses are **google.com**, **bing.com** and **duckduckgo.com**.

b. Time and record how long it takes each student to note answers to the 10 questions *on the last page of this resource* (and no peeking until everyone is ready to start).

c. Check the answers and add a minute's penalty for each incorrect one.

The questions are on the last page, but don't look until everyone is ready.

Analysis

d. Collect the times for each student. If you know how, enter these into three columns in a spreadsheet.

e. Delete the fastest and slowest time from each of the three sets of results. These are the outliers. One very slow time, for example, can greatly affect the average. The person may have been distracted.

f. Calculate the average time taken on each of the three search engines using the remaining data.

g. Create a column chart displaying the three averages. Format your spreadsheet and chart, then save your workbook as '**04. Search Race**'.

Data processing means to take raw facts and figures and change them into useful information that we can all understand. Examples of data processing include:

Aim: *To learn about the stages in data processing.*

- *A weather bureau taking temperature measurements at various times throughout the day and presenting them as a chart in a newspaper.*
- *A company analysing sales figures from the year and presenting a report to the board.*
- *A self-driving car analysing the mass of data from its sensors and deciding whether it needs to urgently stop before someone gets hurt.*

Task 1 – Data Processing Steps

The 6 steps in data processing are listed below. Try and match each step to the correct definition. You may be able to work out the matches yourself, otherwise look online for help.

	Step
1	Collection
2	Preparation
3	Data input
4	Processing
5	Output
6	Storage

	Definition
a	Working out how you can turn the data into useful information.
b	Keeping the data safe for later use.
c	Gathering data from trustworthy sources such as quality websites.
d	Creating reports, charts, videos or other usable forms of information.
e	Entering the data into an application such as a database.
f	Cleaning and organising the data. Checking for errors.

Task 2 – Your Data Processing Experience

You may be more experienced in data processing than you think (we have already tried some in the previous tasks in this resource). Which of the following things have you done before?

____ *Searched for something on the internet (data collection)?*

____ *Selected the best results from a search and ignored ones that don't seem relevant (preparation)?*

____ *Created a table of information, or entered data into a spreadsheet (data input)?*

____ *Looked at some facts or figures and wondered how to make sense of them (processing)?*

____ *Created a chart or graph from some data in science, maths or geography lessons (output)?*

____ *Backed up your files online or to a USB so that you don't lose them (storage)?*

Task 3 – Data Processing Example Task

We're going to work through a data processing activity to get some experience of the whole process. We'll look at the distribution of first nations people across the different states of Australia in 2021.

A. Data Collection

Gathering data from trustworthy sources such as quality websites.

Look at the different sources of information below and decide which you will trust and which you will ignore.

2021 Census Data

State	First Nations People
NSW	339710
VIC	78696
QLD	273119
ACT	9525

My gran thinks there are 10,000 first nations people in Tasmania

Australian Bureau of Statistics 2021

Qld	
SA	273,119
WA	52,069
Tas.	120,006
NT	33,857
ACT	76,487
	9,525

Google

All Images News Shopping Videos Web Forums

New South Wales had the largest Aboriginal and Torres Strait Islander population (339,700 people), followed by Queensland (273,100 people) and Western Australia (120.000 people).

Wikipedia

Population distribution by state/territory	
New South Wales	265,685 (3.55%)
Queensland	221,276 (4.57%)
Western Australia	100,512 (3.93%)
Northern Territory	74,546 (30.34%)
Victoria	57,767 (0.94%)

NT Government Website

Table 2: Aboriginal population

	ERP June 2021 Number
Northern Territory	76,487
New South Wales	339,710
Victoria	78,696

Search

MeKnowEverything
Monday at 11:00 AM

victoria has a first nations population of 405000

https://healthinfonet.ecu.edu.au/learn/locations/wa/

In the 2021 Census, 3.3% (89,000 people) of the WA population identify as Aboriginal and/or Torres Strait Islander people compared with 3.2% nationally [2]. Of the 89,000 people who identify as Aboriginal and/or Torres Strait Islander in WA, 95.8% identify as Aboriginal, 1.8% as Torres Strait Islander, and 2.3% as both Aboriginal and Torres Strait Islander.

Aim: To learn how to count in binary and convert binary numbers into decimal.

Most of the time, we like to count in base 10. Historically, this probably occurred because we have ten fingers and people used them to add up. When counting in base 10, if any digit goes past 9, we add one to the column on the left and start again. Adding 1 to 9 makes 10; adding 1 to 39 makes 40; adding 1 to 99 makes 100 etc.

Computers don't have fingers, but they do have switches. Switches can only be on or off; 1 or 0. If one of the tiny electrical circuits is carrying electricity, then it is said to be ON. This can be represented by a 1. If the circuit is not carrying electricity, then it is OFF, represented by a 0.

Computers can therefore really only count in base 2, or binary. There isn't a number 2 in binary, so when a digit goes past 1, we add one to the column on the left and start again. Adding 1 to 1 makes 10 and adding 1 to 11 makes 100.

Task 1 – Binary Computing

Complete the passage using the words in the box.

transistors binary ON base billions OFF

Modern computers are built using ________________ of tiny electronic switches called ________________. Each switch can either be ON or OFF. If the switch is turned ________________, it is represented by the digit 1. If it is turned ________________, then it is represented by the digit 0. As data is stored on the computer using these two digits, operations must be carried out using the ________________ system of counting. The binary system involves working with numbers in ________________ two.

Task 2 – Binary Counting

Complete the table below showing how to count in binary. Remember that any time a digit goes past 1, you add one to the column on the left and start again. Find help online if required.

Dec	Binary				
0					0
1					1
2				1	0
3				1	1
4			1	0	0
5					
6					
7					

Dec	Binary				
8		1	0	0	0
9		1	0	0	1
10		1	0	1	0

Dec	Binary				
16	1	0	0	0	0

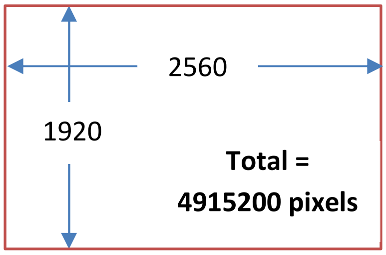
Dec	Binary				

Aim: *To investigate pixels and create a bitmap image using Paint.*

Bitmaps are images made from a grid of tiny squares called pixels. Each pixel can be a different colour. The idea is that the individual pixels are so small that they cannot be seen by the human eye. It's your job when creating bitmaps (either using a camera or a graphics program) to ensure that this is the case. Examples of bitmaps are PNG and JPEG images, GIFs and the BMP files saved by the Windows Paint application.

Look at the photograph below left. It appears to be a picture of a crane made from nice smooth bars of metal. However, when we look at it close up (as below right) we can see that it is in fact an array of tiny, coloured pixels. The pixels are small enough that our eyes don't normally notice them.

The original image was made from a grid 2560 pixels across and 1920 pixels down. This is a total of 4915299 pixels (or about 5 million). The camera was set to take a 5-megapixel photograph.



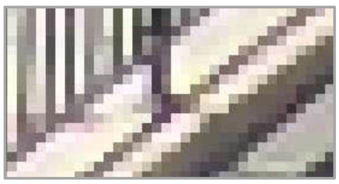
Note: *The activities in this resource will use Paint – a simple graphics program included in all versions of Windows. Mac users can complete the tasks in Paintbrush, which is very similar.*

Task 1 – Viewing Pixels

a. Select a photograph, either saved from the web or from your own collection.

b. Double-click on the image file to open it in your default image viewer. Zoom in and see whether you notice the pixels. This depends on the quality of the photograph and how much the application allows you to zoom. You will probably find some edges that look a little jagged.

c. Now open the same picture in MS Paint (or Paintbrush, on a Mac). In Windows, Paint can be found in the *Accessories* collection, or by searching for Paint. You can either open Paint first and then open your photograph from within, or right click on the photograph file in Windows and select 'Open with / Paint'.



d. Zoom in again on your photograph. Use either the zoom controls under the *View* tab in the main menu or hold down the *Ctrl* key and scroll with the mouse.

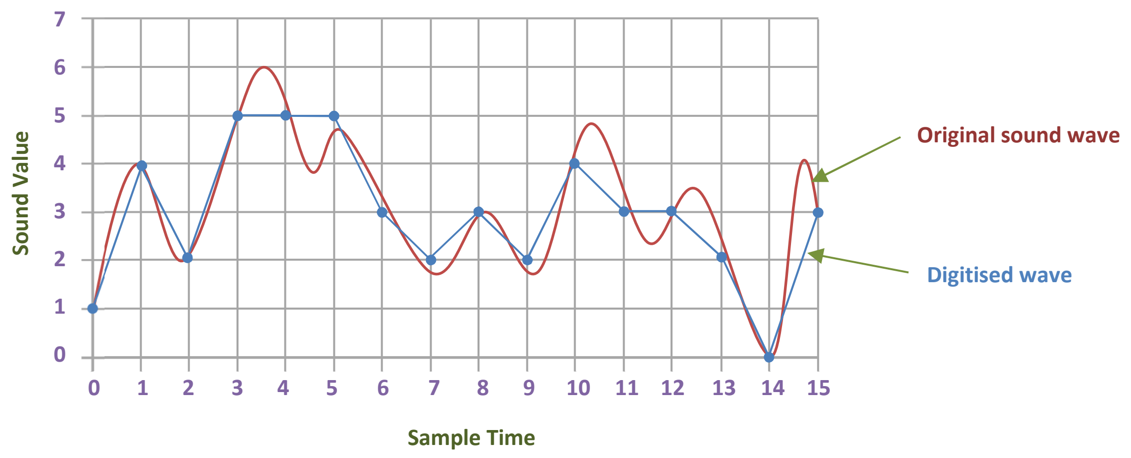
e. View a variety of pictures of different quality and size.

As you may have learned in earlier tasks, data must be encoded as binary so that it can be processed, stored and transferred across networks. To store and listen to music using your digital devices, the sound must be encoded in this way. To encode sound, a sample is taken at frequent intervals.

Aim: *To investigate how sound is encoded as binary and the effect of sample rates.*

Task 1 – Sampling Sound

The diagram below shows how a sound wave can be changed into a digital signal represented by numbers. The original sound is the very wobbly (analogue) red line. The digital sound is the blue line. The blue dots show the sample points. They are rounded off to the nearest whole value.



a. The blue dots show the samples that have been taken. There are sixteen altogether, taken at regular intervals (shown on the horizontal axis). The values for each sample are shown on the vertical axis. Complete the table showing the sound values for each sample.

Sample Time	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Sound Value	1	4	2													

b. What is the highest sound value for the digital signal? What appears to be the highest value of the original analogue signal?

c. The two waves are shown separately on the right. Describe some of the other differences between them.