Nau mai haere mai ki
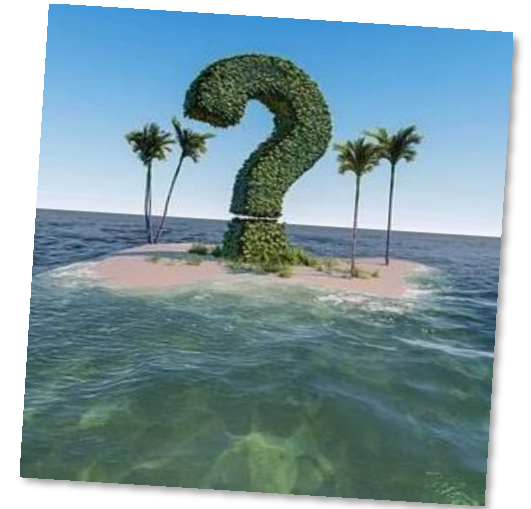
Introduction to Data

data.govt.nz

▶ Next

Am I in the right place?

The intended audience for this e-learning is:

- NZ government departments
- NZ not-for-profit organisations
- Iwi and Māori organisations

The content is for those who are
new to data and its management.

There is no assessment at the end of this e-learning.

What can I expect from this e-learning?

When you complete this e-learning you will:

- Understand what data is.

- Learn the meaning of terms related to data.

- Understand why data is important.

- Recognise the three different forms of data.

- Know the FAIR principles for data usage.

- Understand the value of a good data management plan.

- Know where to go next for more information or help.

What is data?

Data is a type of information (especially facts or numbers) that is collected to be categorised, analysed, and/or used to help decision-making.

Adapted from the Cambridge Dictionary [definition](#)

Important data-related terms

Like many topics, data science has its own language.
Here are some of the terms it is useful to know:

- Dataset: a particular *collection* of data, gathered for a purpose.

- Re-use: using data for a purpose *different* to the original one.

- Metadata: data that *describes* and gives the *context* for the data (allowing discovery and re-use).

- Discovery: through good metadata, being able to *find* the data you are looking for.

- Statistics: a type of *result* from analysing and interpreting raw data.

Why is data important?

Data:

- Supports good decision-making and problem-solving.

- Informs research and policy.

- Enables an organisation to measure performance and success.

- Results in products/services more aligned with customer needs.

- Supports better policies/strategies.

- Provides a record of business activity.

What are the three different forms of data?

Open data is data that anyone can access, use and share, with full permission to use any way they like.

Shared data is data that can be shared with a specific group of people for a specific purpose.

Closed data is data that can only be accessed by those who collected it or are accountable for it.

Source: adapted from The Open Data Institute's "Open/Shared/Closed: The World of Data"

What are the principles of responsible data usage?

- Data is a valuable resource. Unfortunately it can be used inappropriately, on purpose or by accident.

- To help avoid this, a number of different principles exist to ensure that data be as accessible, usable and ethically governed as possible.

- Examples include the NZ Data & Information Management Principles, the NZ Privacy, Human Rights & Ethics Framework and the CARE principles for indigenous data governance.

- A good international and well-recognised set of data principles are the FAIR principles: **F**indable, **A**ccessible, **I**nteroperable and **R**e-usable.

Continued on the next slide...

The FAIR principles

🔍 **F**indable: data and metadata should be easy
to find for both humans and computers.

🔒 **A**ccessible: once you have found the data, it should be easy to
access, and authorisation processes should be clear.

🔗 **I**nteroperable: the data should be easily combined with other data,
and easily work within standard applications.

♻ **R**e-usable: data and metadata should be well-described
so that they can be re-purposed.

Source: *www.go-fair.org/fair-principles/*

How is data managed?

The best way to manage data is by creating and using a "data management plan". A good plan outlines how you are going to:

- Collect data.
- Check its accuracy and quality.
- Store it.
- Use it securely and efficiently.

Having a plan means others can understand a lot about your data without having to ask you, saving time and effort.

A plan can be simple, or complex, depending on the amount and variety of data you may have.
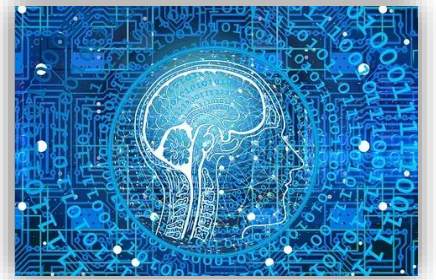
How can I learn more?

The next introductory module on this site is: Introduction to Data Management Part One. Take a look!

Explore the data.govt.nz site.

Contact: datalead@stats.govt.nz

Learning....