1.
Data and information: The basics

Te Pou o te Whakaaro Nui

Data and information: The basics

Data and information can be used by anyone to make a difference in mental health and addiction services. Gathering data and turning it into information creates knowledge, which helps people to understand and inform services.

Data - *the unprocessed facts and figures you gather. This can be numbers, words or observations.*

Information - *data that has been interpreted to make it meaningful to the user.*

What's the difference?

Data - is the unprocessed facts and figures you gather such as numbers, statistics and measurements. It can also be words, observations or other inputs.

For example:

- the number of people who attend a service each month
- the comments left in the waiting room suggestion box
- an outcome measurement score.

There are different uses of data at different levels:

- Individual data: Unit record data related to one thing, event or person.
- Aggregated data: Is data from more than one individual. The most important thing to remember is that when data is aggregated or summarised it can't go back to show individual level of data.

Defining the characteristics of data

There are several characteristics of data that are important to understand.

Each of these characteristics will impact on how data can be presented, analysed or used as information. It will also impact on the kinds of conclusions we are able to draw.

Information - data that has been interpreted to make it meaningful to the user.

For example:

- a graph showing the trend in attendance at a service over time

- A summary of the themes from the comments left in the suggestion box.

Source	Who or what provided the information?	**Self-rated:** Information provided by the tangata whai ora. **Clinician-rated:** Information provided by a staff member. **Other:** For example, whānau rated, eg Hua Oranga.
Type of data	How is the data presented and analysed?	**Qualitative:** Descriptive data that is observed and described but not presented as numbers. **Quantitative:** Data that is presented as numbers.
Collection conditions	When is the data collected?	**Routine:** Collected at set points of time or an activity (eg weekly) **Sporadic:** No set pattern of collection. May be collected at any stage, or never collected for an individual.
Psychometric properties	Do the tools used to collect the data do so effectively?	**Reliability:** Does the tool give the same result every time under the same conditions? **Validity:** Does the tool measure what it is purported to measure?

Commonly used terms

Average

In general when a person refers to an 'average' they are talking about the 'mean'. There are three definitions of an 'average':

1. **Mean:** the sum of the values divided by the number of observations.

2. **Median:** the middle value in an ordered set of values. This is used when there is outlying data that affects the mean or when you are interested in the middle value.

3. **Mode:** the most frequent value in a distribution. This shows what value appears most.

Percentage or proportion

A **percentage** is a calculation of a number with 100 as the denominator. This gives a percentage of the whole.

Example calculation: Investigating the people with a score of less than three

In the equation below the total number of people scoring 3 or over = 4 and the total number of people in the data set (n) = 9

Percentage	4 divided by 9 multiplied by100 equals 44 per cent
Proportion	4 divided by 9 equals 0.44

Number

If you have a small data set it's usually best to describe the data as numbers.

For example, there are four people with a score of three or less.

Variables

The components that are being compared and measured are called **variables**. It is often useful to describe variables as either dependent or independent. The **dependent** variables are what can be seen to change in relation to the particular levels of the **independent** variable. For example when prescribing medication:

- The independent variable is the dose of the medication

- The dependent variable is the frequency or intensity of symptoms.

Presenting information

Tables and graphs can be useful tools for helping you make decisions. However, they only provide part of a story. As well as being able to identify clearly what the graph or table is telling us, **it's also important to identify what parts of the story are missing.**

Tables and graphs present information differently. It's important to consider your audience when choosing how to present the information.

Tables

A table is a way to organise and display information. It shows the interaction between the variables.

Tables show you patterns and relationships between variables. Below is an example of a table you might see in a report.

NB The content in the table and graph examples below are used for demonstration purposes only.

| Team type	Number of collections with no items in clinical range	Total number of collections	Percentage with no items in clinical range
Inpatient services			
Alcohol and drug team	0	294	0%
Eating disorders team	0	43	0%
Forensic team	6	117	5%
Inpatient team	458	10,332	4%
Maternal mental health team	6	102	6%
Older people team	3	27	11%
Total	**473**	**10,919**	**4%**

Graphs

A graph is a diagram representing a system of connections or interrelations among two or more things

Graphs show an overall shape of a variable or the relationships between variables. Types of graphs you may see include:

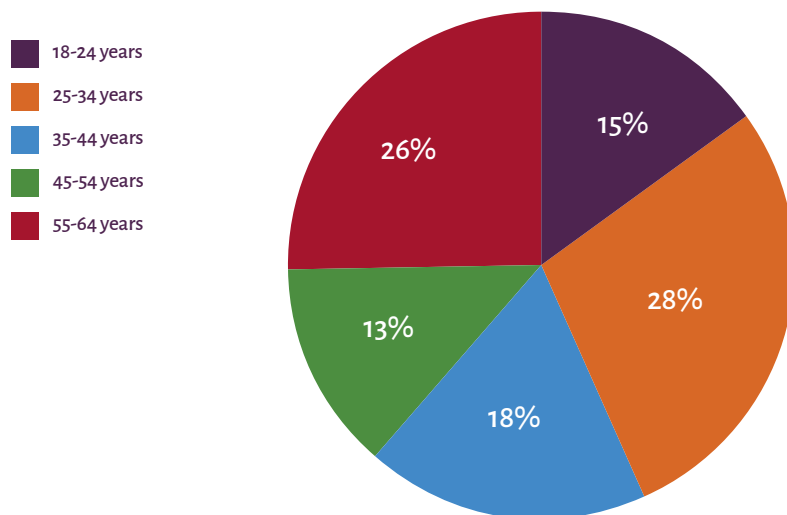**Bar graph** - used to compare different groups or to track changes over time.
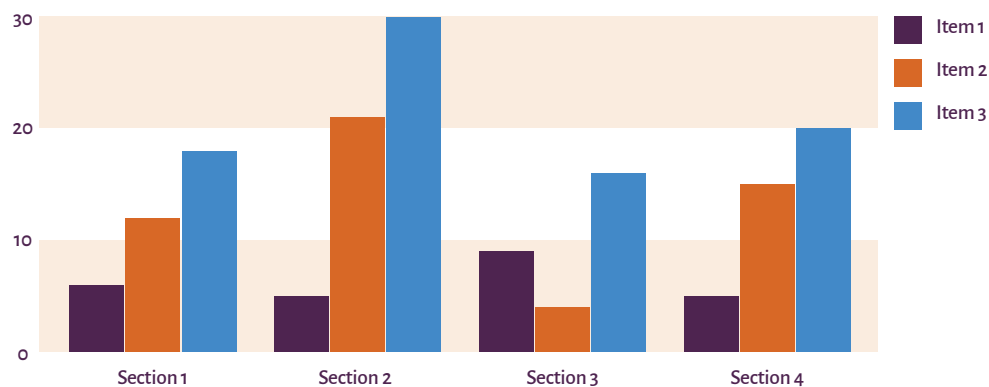
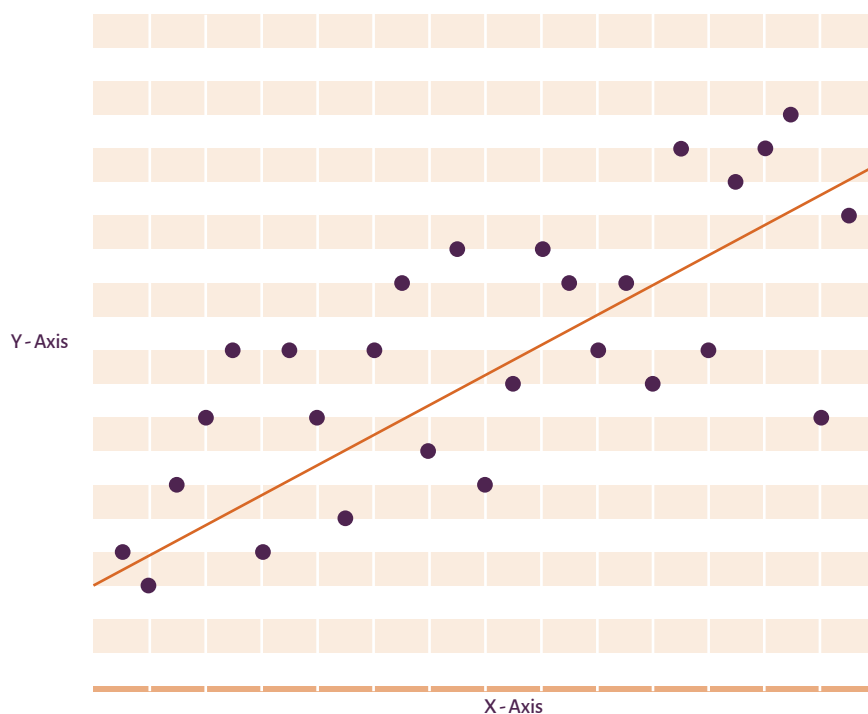**Line graph** - used to track changes over time.
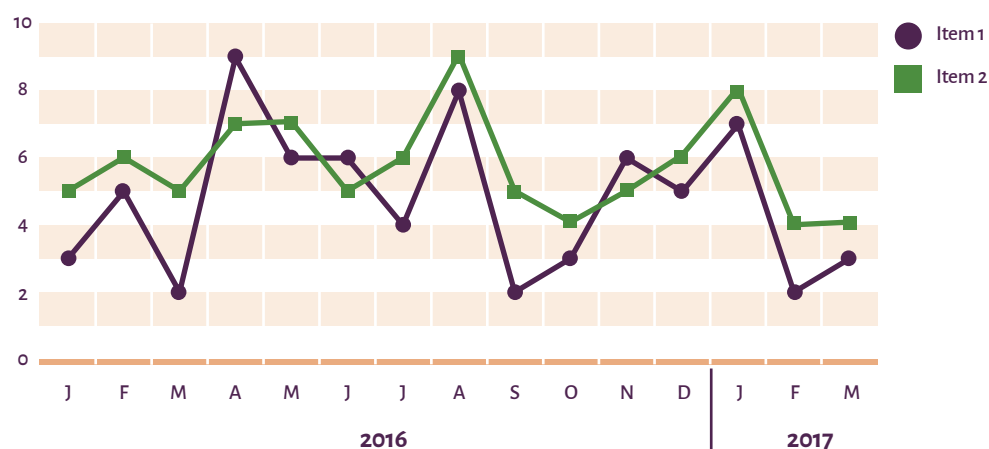
Pie chart - used to show the parts that make up a whole.

- 18-24 years
- 25-34 years
- 35-44 years
- 45-54 years
- 55-64 years

15%
28%
18%
13%
26%

Scatter graph - used to display the relationship between two different variables.

Y - Axis

X - Axis

Te Pou o te Whakaaro Nui