ANALYSING AND PRESENTING DATA
PART 1 – ANALYSING DATA

Professional Development Course Book

Presented by Mark Priadko

INSTITUTE OF
PUBLIC ADMINISTRATION
AUSTRALIA

IPAA
SOUTH AUSTRALIA

IPAA Personal Membership

Become part of the IPAA SA Personal Member community and access an extensive range of benefits, opportunities and support for your career.

Personal Membership Will…

Get you future ready

We will improve your career as we educate and inform you of what's to come in the future of the profession and build your skill set to ensure you are ready for it.

Build your capability

We will build your current capability to ensure you perform at the best of your ability and get you ready to face every new challenge throughout your career.

Champion you

We will champion your successes and voice them to the sector.

Build your networks

We will build your networks and connect you to like-minded people, people facing the same challenges as you and people who can help you to gain influence and penetrate the vast public sector.

Connect you

We will connect you with ideas from across the country through our National body and State jurisdictions.

Support you

We will provide you with tangible support throughout your career.

More benefits

- Extensive discounts on IPAA SA training, courses, forums and events.
- Access to exclusive member opportunities, such as our mentorship program and networking sessions.
- Receive resources and learning opportunities relevant to your career level.
- Receive a membership certificate for your CV and use of the MIPAA post nominal.
- Unlimited digital access to the Australian Journal of Public Administration.
- Receive a training voucher to the value of your membership.
- Have your say in IPAA SA governance.

How to join?

Joining is as easy as visiting our website: www.sa.ipaa.org.au/membership. Navigate to the Membership section and select the membership type most applicable to you and fill in your details online. Membership fees are generally tax-deductible.

If you have any questions about membership, contact us on 8212 7555 or email membership@sa.ipaa.org.au

Learning objectives:

- Understand where analysis fits and what purpose it serves
- Introduction to different types of analysis
- Practical application of analysis techniques

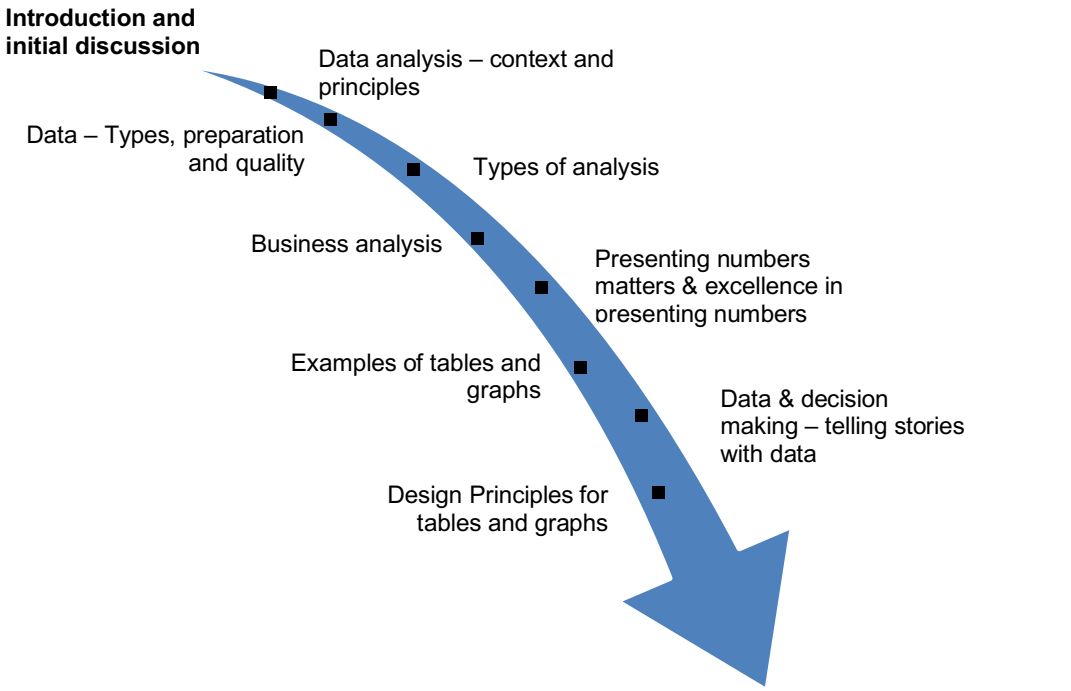
Self-evaluation

My proficiency in analysing and presenting data is:

List of what I need to know to move up one point on this scale

- .....
- .....
- .....
- .....

These course notes are designed to support the presentation of information in the module. They are based on the knowledge and experience of Mark Priadko.



I better understand a range of analytical techniques and options for presenting data that help me use data in the development of evidence-based policy and in reporting.

Table of Contents

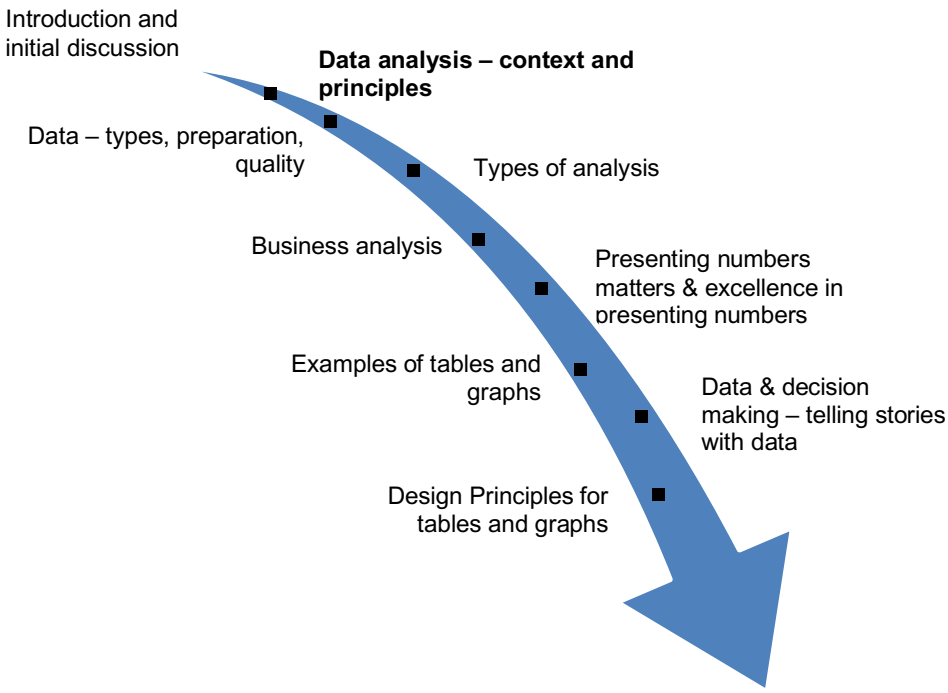
Quick questions

Are you able to guess which countries have seen the largest growth in net migration into Australia over the last ten years?

Are you able to guess which countries may impact more on migration in South Australia?

Discussion notes

What purpose does our data analysis serve? Why is it important?



I better understand a range of analytical techniques and options for presenting data that help me use data in the development of evidence-based policy and in reporting.

Data Analysis – Context and Principles

Data - its benefits and its limitations

Data is an abstraction of a real-world entity (person, object, or event).

Data represents an effort to measure attributes of our world so we can better understand and communicate them.

As an example - for our employees, we use data to understand things about them and their relationship with our business. We collect data on:

- Age
- Gender
- Address
- Classification
- Qualifications.

We can never understand everything about our employees and, for privacy reasons, we never want to record or understand many things about them.

Data can never fully represent the reality of our world - there is no such thing as perfect data. Data can never fully capture the complex details of a real-world entity like an employee.

We use data to move up a knowledge continuum understanding that we can't reach the end at the right-hand side.

What we know about an aspect of our world	Nothing	Everything

We use data to move up the continuum →

We use data to support decision-making. All decision-making occurs in an environment of uncertainty and with the risk of biases. We rely on data to reduce uncertainty and reduce biases in decision-making. Just as data cannot perfectly represent our world, we cannot eliminate uncertainty. Data is used to help us make sense of our world and the decisions we make by reducing uncertainty.

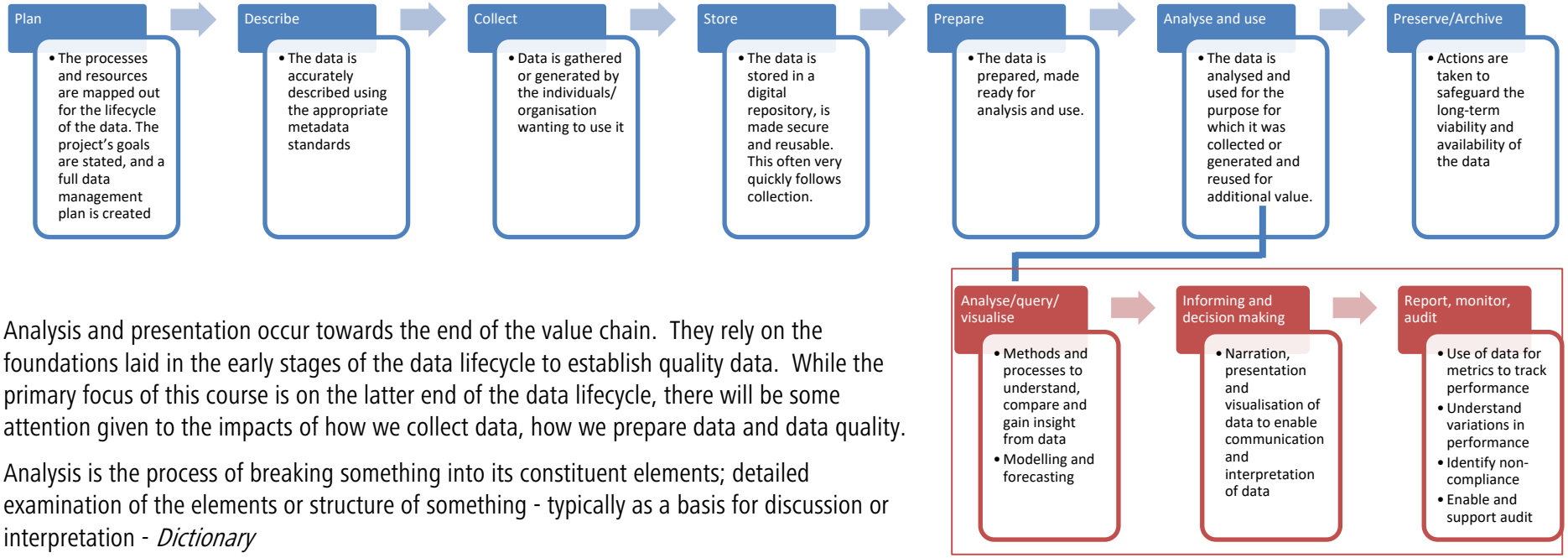
What we know about an aspect of our world	Nothing	Everything
The impact on our decisions	Uncertainty	Certainty

We use data to move up the continuums →

When data is deemed 'fit for purpose', it is fit to assist in decision-making to which the subject is related. It is fit for purpose if it moves us to the right of the continuums. It is not fit for purpose if it risks moving us to the left of the continuums.

There will always be limitations on the ability of data in decision-making. We need to be humble with data.

Data is an organisational asset. The lifecycle over which data is managed is characterised in the diagram below.



Analysis and presentation occur towards the end of the value chain. They rely on the foundations laid in the early stages of the data lifecycle to establish quality data. While the primary focus of this course is on the latter end of the data lifecycle, there will be some attention given to the impacts of how we collect data, how we prepare data and data quality.

Analysis is the process of breaking something into its constituent elements; detailed examination of the elements or structure of something - typically as a basis for discussion or interpretation - *Dictionary*

Analysis is the process of breaking a complex topic or substance into smaller parts to gain a better understanding of it. – *Wikipedia*

Discussion

What are some of the types of data we are working with?

In the public sector, we work with data:

- To monitor and evaluate services
- To understand the community – demographics, use of services, satisfaction with services, attitudes, behaviour, movement
- To understand the economy – labour markets, exports, imports, investment
- To understand the environment – climate, rainfall, temperatures, water flows, waste management, flora and fauna
- To determine funding – hospital and school funding is based on activity data.

We present data for

- Public record
- To inform the public, leaders and managers
- To help us formulate strategies and options
- To help us make decisions.

Principles of analytical design

In his book *Beautiful Evidence*, Edward Tufte outlines some key principles for analytical design. What his principles outline is what we should aspire for when we undertake analysis. These principles are shown below.

Principles of analytical design - from Edward Tufte

Principle 1 - Comparisons - show comparisons contrasts and differences

Principle 2 - Causality, mechanism, structure, explanation

Principle 3 - Multivariate analysis - show more than 1 or 2 variables

Principle 4 - Integration of evidence - integrate words, numbers, images and diagrams

Principle 5 - Documentation – detailing the sources of information, good presentation of tables and charts.

Principle 6 - Content counts most of all - analytical presentations ultimately stand or fall depending on the quality, relevance and integrity of their content.

In real estate, you hear the term 'location, location, location'.

For data analysis, think 'comparison, comparison, comparison'.

Steven Few states: "Comparison is the beating heart of data analysis. What we do when we compare data really encompasses both looking for similarities and looking for differences."

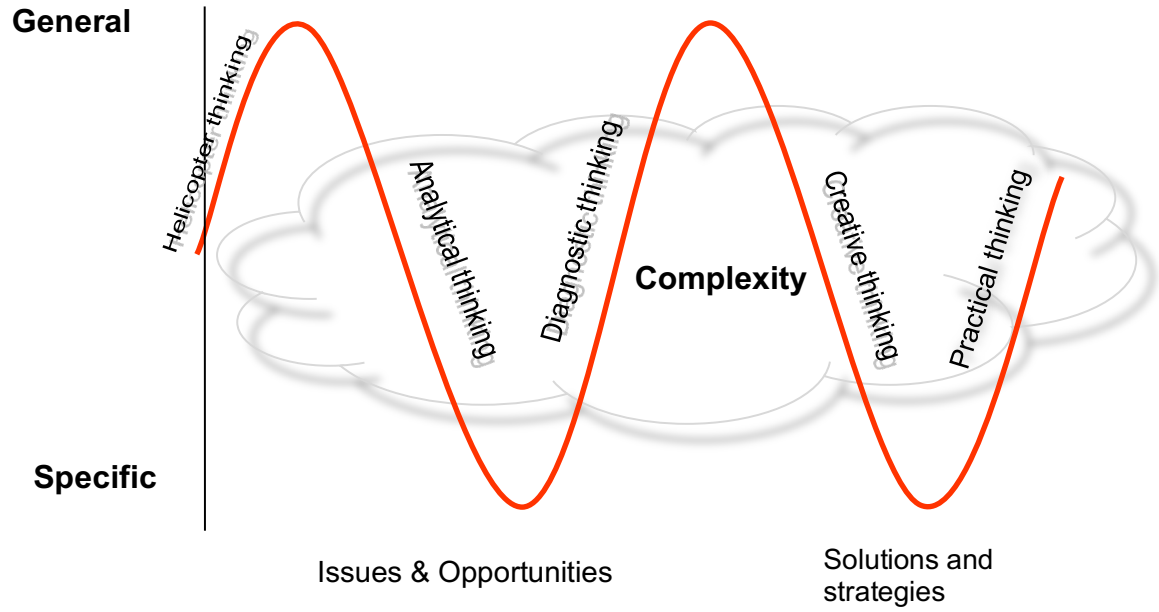
Analysis – one part of thinking things through

Analytical thinking is one of a number of types of thinking necessary to think through an issue.

To think through an issue and its solution involves five different thinking modes.

1. Helicopter thinking
2. Analytical thinking
3. Diagnostic thinking
4. Creative thinking
5. Practical thinking

These five modes of thinking can be included in the model as shown below.



Thinking through an issue involves being able to oscillate between the general (context) of an issue and the specific (details associated with the complexity inherent in the issue) by applying the five modes of thinking. Each mode has its own style and attributes. A summary of the five modes follows.

Helicopter thinking considers our issues as part of a larger view or picture, as part of a system or landscape. This thinking is important to understanding the importance of the matter we are dealing with, defining it and understanding the scope of our work. It requires us to think broadly and holistically to understand the context surrounding our matter.

Analytical thinking is necessary to deliberately work through the details and complexities associated with problems/issues/opportunities to understand them. It involves research and requires that we dive into detail to gather facts and evidence. By definition, analytical thinking requires us to decompose our subject into smaller parts.

Diagnostic thinking is slower and more deliberate and brings these parts back together to draw conclusions, make judgements, identify root causes, reveal insights or reach a diagnosis. Diagnostic thinking is the centrepiece of the thinking approach and the gateway to future solutions or strategies.

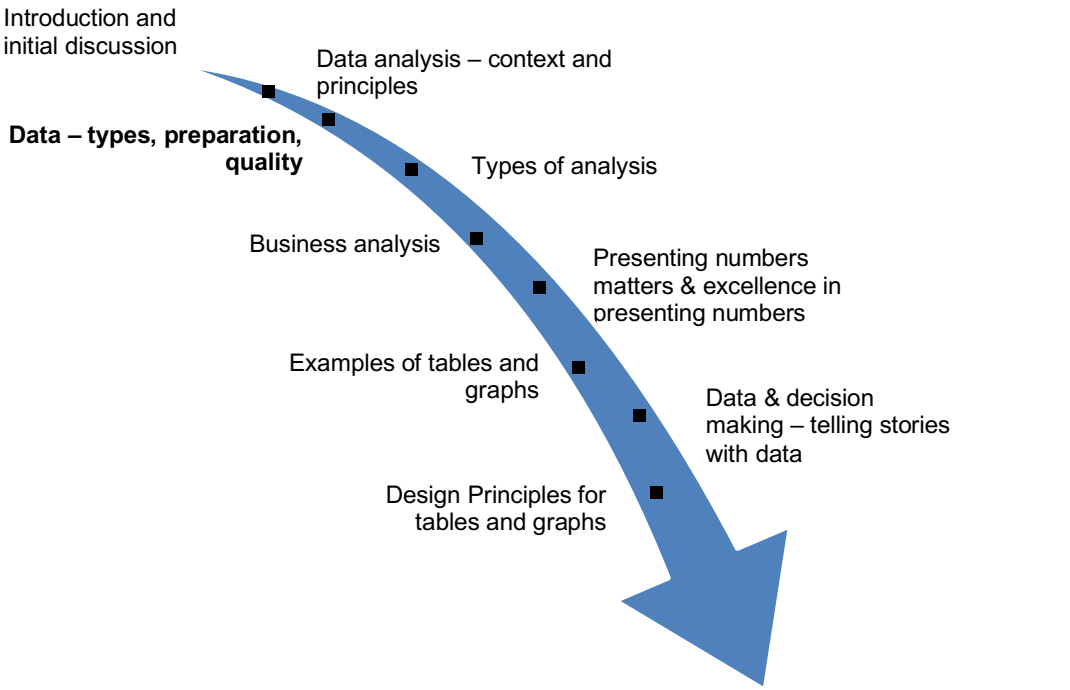
Creative thinking is necessary to develop ways of responding to the revelations of the diagnostic process. Nearly every circumstance I have dealt with has its own unique attributes that require a specific

response. The need for a unique, tailored response demands creating new ways of doing things for that situation.

Practical thinking is thinking in a way that focuses on action to ensure that whatever strategies/solutions are developed can be implemented and embedded into practice. It will typically involve answering the multi-layered question – ***who will do what by when and with how much ($)*** to implement the response created.

I present data to help others through some or all of the processes below.

More on this in the section on Presenting Data.



I better understand a range of analytical techniques and options for presenting data that help me use data in the development of evidence-based policy and in reporting.

Data – Types, Preparation and Quality

Types of Data

Primary (direct) and Secondary (indirect) Data

Primary data is data created as original by a researcher through direct efforts and experience, specifically for the purpose of addressing their research problem. Primary data is also known as firsthand or raw data.

Secondary data is data that has already been collected and recorded by another person(s) for a different reason or purpose than that of the current researcher. It is the readily available form of data collected from various sources like government publications, internal records of the organisation or publications for related research.

Primary data is typically more expensive and takes longer to collect but has the benefit of being specific to the question or problem being considered. The collection of primary data is necessary to explore the details behind types of behaviour, opinions or views that cannot be identified from existing data. Collection of this data will often involve the crafting of surveys and questionnaires. These methods can of themselves create distortions and biases in the data collected. As a result, the collection of primary data is a specialist field.

Many analysts underestimate the volumes and value of secondary data available within their organisation. Common sources of existing data include financial information, activity data, transaction volumes, data on suppliers and customers from subsidiary ledgers and data on staff.

Methods of primary (direct) data collection include:

- Surveys administered by an interviewer
- Surveys which are self-enumerated (the information written or entered directly by the respondent)
- In-depth interviews or focus groups that provide the opportunity for discussion and elaboration for collecting more detailed information about a particular issue or issues
- Observational studies in which data are gathered through the direct observation of the population or sample
- Experiments and clinical trials that involve controlled studies where researchers collect data from subset groups taken from the population of interest.

One form of secondary data is administrative data. Administrative data are collected as part of the day-to-day processes and record-keeping of organisations. Administrative data, such as historical data or public records, include: School enrolments; hospital admissions; and records of births, deaths, and marriages. The data are not collected initially for statistical purposes but can be organised to produce statistics.

Census and Sample data

A population may be studied using one of two approaches: taking a census or selecting a sample. Both provide information that can be used to draw conclusions about the whole population.

A census is a study of every unit, everyone or everything, in a population. It is known as a complete count or a complete enumeration.

A sample is a subset of units in a population, selected to represent a population of interest. It is a partial count or partial enumeration. Information from the sampled units is used to estimate the characteristics of the entire population of interest.

Advantages of a census:

- Provides a true measure of the population (no sampling error)
- benchmark data may be obtained for future studies
- detailed information about small sub-groups within the population is more likely to be available.

Disadvantages of a census:

- May be difficult to enumerate all units of the population within the available time
- higher costs, both in staff and monetary terms, than for a sample
- generally takes longer to collect, process, and release data than from a sample.

Advantages of a sample:

- Costs would generally be lower than for a census
- results may be available in less time
- if good sampling techniques are used, the results can be very representative of the actual population.

Disadvantages of a sample:

- Data may not be representative of the total population, particularly where the sample size is small
- often not suitable for producing benchmark data
- as data are collected from a subset of units and inferences made about the whole population, the data are subject to 'sampling' error
- decreased number of units will reduce the detailed information available about sub-groups within a population.

Source: ABS Statistical Language Census & Sample

Quantitative data vs Qualitative data

Quantitative data assigns numbers (quantities) to observations about a subject or population by counting or measuring aspects of the subject or population.

Qualitative data assigns labels to observations about a subject or population or assigns the subject into categories (quality attributes).

Quantitative data about a person	Qualitative data about a person
Height (Mark is 187 centimetres)	Gender (Mark is a male)
Weight (Mark is 90 kilograms)	Marital status (Mark is married)
Age (Mark is 55 years old)	Where they live (Mark lives in Adelaide)

Qualitative data can be added across a population to determine how many males there are or how many married or unmarried people there are.

Word clouds are a means to present qualitative analysis.

A 'word cloud' is a visual representation of word frequency. The more commonly the term appears within the text being analysed, the larger the word appears in the image generated.

Word clouds can be created from add-ins to Word or from online software.

The word cloud below has been created from an assortment of sections from this course using a word add-in.

Data about concrete nouns vs data about abstract nouns

Concrete nouns – people, places and things that we can see. Data counts and measures these nouns more exactly – number of people, money, size, weight etc. This data is more objective. It is about objects we can see. This data is easier to independently validate.

Abstract nouns – ideas, feelings, qualities, states. Data is sought from people about their feelings, ideas, agreements and extent of satisfaction. This data is more subjective. The scores, ratings, evaluations etc are based on what the subject thinks or feels. This data is harder to independently validate.

Preparing for data analysis

The preparation required for good data analysis can be compared to the preparation required to do a good job of painting a house.

To paint well is not just about slapping paint onto a brush or roller and covering a wall.

Good painting at home requires preparation – filling holes, sanding off peeled and rough old paint, cleaning the wall of stains and removing dust and grime. We have to prepare the wall so it is ready to be painted.

So to with data analysis. Before we apply different analysis techniques, we need to spend some time preparing the data so it is ready to be analysed.

Preparing data can include, transforming and flattening the data, cleansing data and reconciling data.

Creating a flat data set or database

Just as we need a good base for painting, we need a good base for data – we need a flat data set or database.

We want all our data properly arranged into rows and columns such that:

- Every column is a field. A field is a single piece of information from a record.
- Each row is a record where each record has a valid entry against each field.

The main benefit of creating a database is that it enables a range of analysis techniques to be performed automatically through tools like data filters, subtotalling and pivot tables.

I have created a database from ABS population data. The data can be analysed in the form presented by the ABS but converting it into a database opens up a range of opportunities for data analysis.

For net migration data, the screen grab below details how the data is provided by the ABS.

Table 4.5 Overseas migrant arrivals and departures, by visa and citizenship groups, South Australia, 2004-05 to 2022-23(a)(b)

Direction	Visa and citizenship groups(c)		2004-05	2005-06	2006-07	2007-08	2008-09
Overseas migrant arrivals(d)	Permanent visas	Family	1,230	1,360	1,590	1,560	1,660
		Skilled (permanent)	2,550	4,530	5,220	5,320	4,710
		Special eligibility & humanitarian	1,220	1,100	1,220	840	1,050
		Other (permanent)	140	130	170	150	170
	Total permanent visas		**5,130**	**7,120**	**8,190**	**7,870**	**7,590**
	Temporary visas	Student - vocational education and training	240	240	360	860	2,650
		Student - higher education	2,380	3,050	3,860	4,570	5,120
		Student - other	1,540	1,390	1,870	1,740	1,620
		Skilled (temporary)	740	1,520	1,700	1,910	2,070
		Working holiday	250	320	470	570	690
		Visitors	1,200	1,300	1,280	1,710	1,480
		Other (temporary)	590	570	530	510	340
	Total temporary visas		**6,930**	**8,380**	**10,060**	**11,860**	**13,960**

Flattening data requires allocated categories in the first three columns across all rows:

This can be done manually by dragging the categories across all rows.

We will also need to change migrant departures to have negative signs rather than positive signs. This can be done by formula rather than by manually changing data.

For our datasets we can also combine jurisdictions by adding another field to each state showing the jurisdiction. We can then 'stack' the datasets together to arrive at a consolidated dataset for net migration by groups across all states and with arrivals with positive data and departures as negative data.

After these changes have occurred, the database appears as follows:

Jurisdiction	Direction	Visa and citizenship groups(d)	Category	Category 1	Category 2	2004-05	2005-06	2006-07	2007-08	2008-09
Australia	**Overseas migrant arrivals(e)**	Permanent visas	Family	Permanent	Family	28,430	30,400	32,480	33,100	35,060
Australia	**Overseas migrant arrivals(e)**	Permanent visas	Skilled (permanent)	Permanent	Skilled	35,540	42,750	47,540	51,600	48,360
Australia	**Overseas migrant arrivals(e)**	Permanent visas	Special eligibility & humanitarian	Permanent	Special eligibility	13,580	12,320	12,400	9,470	11,630
Australia	**Overseas migrant arrivals(e)**	Permanent visas	Other (permanent)	Permanent	Other	3,650	3,640	3,770	3,940	4,000
Australia	**Overseas migrant arrivals(e)**	**Total permanent visas**	**sub total**	**Sub total**	**Sub total**	**81,210**	**89,110**	**96,190**	**98,110**	**99,040**
Australia	**Overseas migrant arrivals(e)**	Temporary visas	Student - vocational education and training	Temporary	Student	7,990	10,500	19,970	31,420	53,570

After flattening the data, it is critical to reconcile the flattened data with the original data set to ensure the integrity of the flat data set. Pivot tables can be used to calculate totals from the flat data set.

The benefit of this setup is that:

- other data can be combined for earlier years or other jurisdictions
- that columns with formulas can be added for analysis purposes
- the database can be filtered and queried to provide summarised data.

The last point is most helpful as it enables the data to be queried and viewed in different ways, in particular using pivot tables so we can see the data from different perspectives.

Data analysis technologies

Improved technologies can make data analysis and data presentation more straightforward.

We are familiar with traditional spreadsheet tools like Excel that enable straightforward data analysis with small and medium-sized data sets.

More advanced tools are now available in the form of Microsoft Power BI and subscription packages like Tableau.

To use the analogy of painting, excel is like a good brush while these other technologies are like rollers and spray guns – you can cover more data more quickly with them.

However, beware, just as painting technologies like rollers and spray guns cannot overcome deficiencies in the surface for cover for lack of preparation, the same is true for data analysis technologies not overcoming deficiencies in the quality of data or the lack of preparation of the data.

It is a criticism of big data that the size of the data set alone is seen as enough to produce better analytical results. The bigger they are, the harder they fall. Bigger data sets can exaggerate biases and deficiencies rather than overcome them. Big data sets can have underlying biases within them. Data captured from certain applications on mobile devices will tell you information about the group of people who use those applications and devices. But that may be a very specific population subset and may not be representative of society. The data will be useful if that subset is your target audience but not so useful if you are using the data to develop policy solutions for a wider population.

Big data has also had to rely on algorithms applied to data to draw conclusions. The development of these algorithms can be dependent on the source data. Example from Tim Harford's book Data Detective (page 166).

> *"In 2014, Amazon, one of the most valuable companies in the world, started using a data driven algorithm to sift resumes, hoping that the computer would find patterns and pick out the very best people based on their similarity to previous successful applicants. Alas, previous successful applicants were disproportionately men. The algorithm then did what algorithms do: it spotted the pattern and ran with it. Observing that men had in the past been preferred, it concluded that men were preferable."*

Think about and question the data

Thoughts from Tim Harford in Data Detective:

- Step back and see the context
- Get the backstory
- Ask who and what is missing
- Demand transparency when algorithms are applied and used
- Don't take strong statistical institutions for granted
- Keep an open mind when working with data.

Questions to ask of data from Stephen Few in the Data Loom:

Required data – What data is required for the task at hand?

Semantics of the data – What do the various fields of data mean?

Relevance of the data – Are all of the data fields relevant to the task at hand?

Source of the data – What is the source of the data and is it credible?

Accuracy of the data – Is the data accurate?

Completeness of the data – Does the data set include all that's needed?

Context of the data – Have I taken all of the relevant context into account?

Representativeness of the data – Is what's revealed in the data typical?

Causes of the behaviours recorded in the data – What is causing this to happen?

Aggregation of the data - Is the level at which the data has been aggregated and the statistical method that was used to produce that aggregation appropriate for the task at hand?

Data Quality

ABS Data quality framework

Data quality is not just about accuracy. Data is generally considered high quality if it is fit for its intended uses and if it correctly represents the real-world construct to which it refers. Data is fit for its intended use if users can do what they need to do with it (analyse, report, make decisions) with reasonable assurance.

There are a number of different dimensions to data quality. We do not dismiss the use of our data because it has some quality shortcomings. We can still assist and improve decision-making with imperfect data but our decision makers need to understand its strengths and limitations.

The Australian Bureau of Statistics provides information on a data quality framework on its website.

The ABS identifies seven different dimensions of data quality.

1. Institutional environment;
2. Relevance;
3. Timeliness;
4. Accuracy;
5. Coherence;
6. Interpretability; and
7. Accessibility.

Institutional environment refers to the institutional and organisational factors that may have a significant influence on the effectiveness and credibility of the agency producing the statistics.

Relevance refers to how well the statistical product or release meets the needs of users in terms of the concept(s) measured, and the population(s) represented.

Timeliness refers to the delay between the reference period (to which the data pertain) and the date at which the data become available; and the delay between the advertised date and the date at which the data become available (i.e. the actual release date).

Accuracy refers to the degree to which the data correctly describe the phenomenon they were designed to measure. This is an important component of quality as it relates to how well the data portrays reality.

Coherence refers to the internal consistency of a statistical collection, product or release, as well as its comparability with other sources of information.

Interpretability refers to the availability of information to help provide insight into the data. Information available that could assist interpretation may include the variables used, and the availability of metadata, including concepts, classifications, and measures of accuracy.

Accessibility refers to the ease of access to data by users, including the ease with which the existence of information can be ascertained, as well as the suitability of the form or medium through which information can be accessed.

Details of the framework can be found on the ABS website by referencing:
https://www.abs.gov.au/statistics/detailed-methodology-information/concepts-sources-methods/australian-system-government-finance-statistics-concepts-sources-and-methods/2015/16-data-quality/part-b-abs-data-quality-framework

Data Cleansing

"Data cleansing or data cleaning is the process of detecting and correcting (or removing) corrupt or inaccurate records from a record set, table, or database and refers to identifying incomplete, incorrect, inaccurate or irrelevant parts of the data and then replacing, modifying, or deleting the dirty or coarse data." Source: Wu, S. (2013), "A review on coarse warranty data and analysis", *Reliability Engineering and System* via Wikipedia

'Dirtiness' includes:

- Duplicate records – the same item inadvertently recorded twice in a data set
- Spelling errors – items where categories have been misspelt or improperly coded (e.g. a 'Camry' recorded as a 'Carmy' in a vehicle register)
- Blank cells – records that have not had particular elements recorded
- Numbers as text – where numbers
- False unique identifiers – dual records with the same unique identifier
- Outliers (false or error records) – unusually large or small values attached to records
- Changes in categories and definitions over time – may require that we rely more on aggregated data or find ways to combine categories.

Basic data cleansing involves:

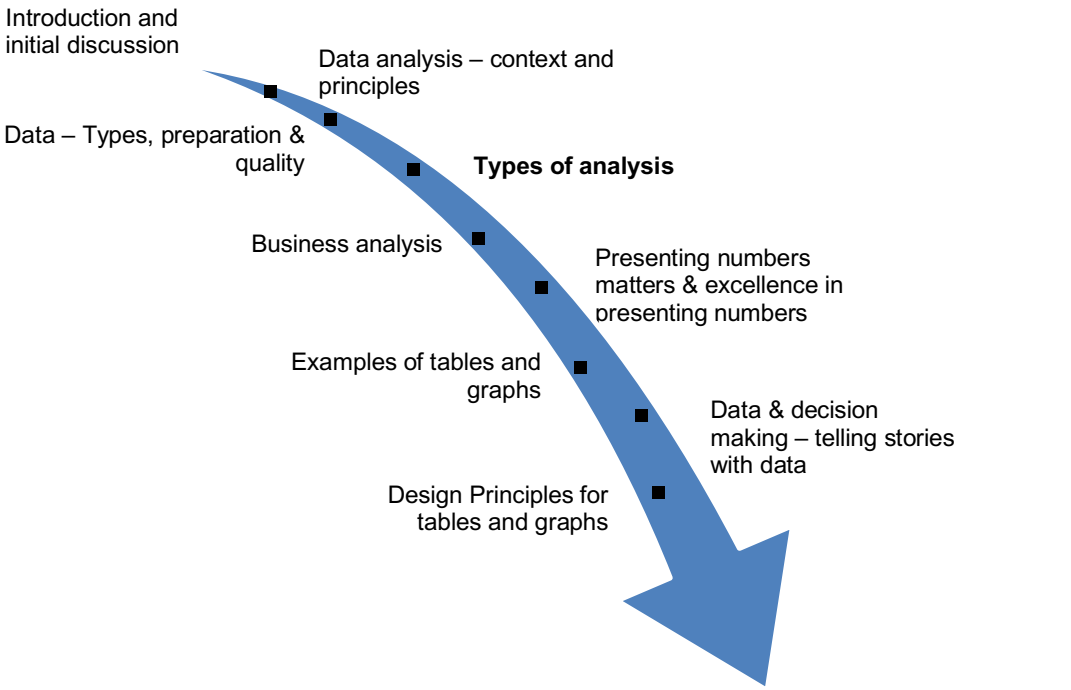
- Sorting data to see duplicates and multiple uses of unique identifiers. Once found, judgements are made about the integrity of the data and, if the data is OK, duplicates are removed and unique identifiers corrected.
- Conditional formatting can help identify unusual records or numbers
- Pivot tables to identify unusual categories and blanks
- Recategorising data using mapping and Vlookup can be used to further assess the quality and categorisation of records.
- Reconciling the parts to totals – checking the parts are consistent with the whole and that they add to the whole to find double counting or gaps.

There are more sophisticated data cleansing methods for large data sets that are beyond the scope of this course and its presenter.

Conclusion

Despite these deficiencies and difficulties with data, I seek to work with the best data available and conclude that working with the best data available is usually better than working with no data at all.

"An approximate answer to the right problem is worth a good deal more than an exact answer to the wrong problem." John Tukey



I better understand a range of analytical techniques and options for presenting data that help me use data in the development of evidence-based policy and in reporting.

Activity – Scenario

The Minister has seen and heard reports of significant growth in international migration and has asked us to do some analysis of data and present our findings to him/her. They have seen these headlines from the ABS

- Overseas migration 2022-23 – a net annual gain of 518,000 people
- Migrant arrivals increased 73% to 737,000 from 427,000 arrivals a year ago
- The largest group of migrant arrivals was temporary visa holders with 554,000 people

What types of data analysis could we consider in developing our advice for the Minister?

Interstate data is available. Would we expect migration numbers to be different across states and why?

Introducing types of information & analysis

We will examine the following types of analysis. They are:

1. Compositional analysis –understanding the components of our data set.
2. Temporal analysis – understanding changes over time (trend and growth analysis)
3. Variance analysis – understanding variances from budget or previous periods
4. Ratio analysis – (primarily used in financial analysis) analysing ratios to assess financial viability and financial health
5. Benchmarking – comparing with external figures
6. Multivariate analysis - combining different data sets (e.g. financial and non-financial data).

These methods will be applied to migration data. Before proceeding with these, we should not underestimate the value of ranking data. The table below does this.

In 2022-23 net migration figures can be ranked by country to see the largest sources of migration.

Net migration by country ranked – Australia & South Australia (2022-23)

Australia net migration

Country	Net migration 2022-23
India	92,940
China	64,310
Philippines	40,890
Nepal	27,460
Colombia	22,320
UK, CIs & IOM	19,230
Vietnam	17,770
Pakistan	17,290
New Zealand	15,340
Thailand	14,060
Bhutan	13,160
Sri Lanka	13,110
Indonesia	12,380
Brazil	11,900
Ireland	8,850
South Africa	8,780
Taiwan	8,200
Fiji	7,680
Other countries	102,010
Total	**517,680**

South Australia net migration

Country	Net migration 2022-23
India	7,810
China	2,970
Philippines	1,930
Nepal	1,400
Colombia	1,310
Vietnam	1,220
Sri Lanka	1,000
UK, CIs & IOM	760
Afghanistan	730
Pakistan	720
Brazil	690
South Africa	610
Hong Kong	590
Kenya	450
Indonesia	440
Thailand	420
Iran	420
Fiji	360
Other countries	3,950
Total	**27,780**

Source: ABS overseas net migration 2022-23.

Compositional Analysis

This type of analysis is required to understand the makeup or structure of a dataset or the finances of a business or organisation. Compositional analysis is used when you want to understand or get to know a data set or business. This analysis is a starting point for the other forms of analysis.

The table below summarises overseas net migration to show the relative size of different components of the dataset.

Overseas net migration by country compositions - Australia and South Australia (2022-23)

Australia net migration

Country	Net migration 2022-23	Share of total	Cumulative share
India	92,940	18.0%	18.0%
China	64,310	12.4%	30.4%
Philippines	40,890	7.9%	38.3%
Nepal	27,460	5.3%	43.6%
Colombia	22,320	4.3%	47.9%
UK, CIs & IOM	19,230	3.7%	51.6%
Vietnam	17,770	3.4%	55.0%
Pakistan	17,290	3.3%	58.4%
New Zealand	15,340	3.0%	61.3%
Thailand	14,060	2.7%	64.1%
Bhutan	13,160	2.5%	66.6%
Sri Lanka	13,110	2.5%	69.1%
Indonesia	12,380	2.4%	71.5%
Brazil	11,900	2.3%	73.8%
Ireland	8,850	1.7%	75.5%
South Africa	8,780	1.7%	77.2%
Taiwan	8,200	1.6%	78.8%
Fiji	7,680	1.5%	80.3%
Other Countries	102,010	19.7%	100.0%
Total	**517,680**	**100.0%**	

South Australia net migration

Country	Net migration 2022-23	Share of total	Cumulative share
India	7,810	28.1%	28.1%
China	2,970	10.7%	38.8%
Philippines	1,930	6.9%	45.8%
Nepal	1,400	5.0%	50.8%
Colombia	1,310	4.7%	55.5%
Vietnam	1,220	4.4%	59.9%
Sri Lanka	1,000	3.6%	63.5%
UK, CIs & IOM	760	2.7%	66.2%
Afghanistan	730	2.6%	68.9%
Pakistan	720	2.6%	71.5%
Brazil	690	2.5%	73.9%
South Africa	610	2.2%	76.1%
Hong Kong	590	2.1%	78.3%
Kenya	450	1.6%	79.9%
Indonesia	440	1.6%	81.5%
Thailand	420	1.5%	83.0%
Iran	420	1.5%	84.5%
Fiji	360	1.3%	85.8%
Other countries	3,950	14.2%	100.0%
Total	**27,780**	**100.0%**	

Source: ABS overseas net migration 2022-23.

For Australia, around 80% of migration numbers are driven by 18 countries.

For South Australia, over 85% of migration numbers are driven by 18 countries.

Ratio analysis

Ratios can be used to provide indications of numbers per….

Ratio analysis is used in financial analysis to assess liquidity (ratio of liquid assets to current liabilities, profitability (return on assets or return on equity) and levels of debt (debt/equity ratio).

For our example, we can use the data to calculate the average number of net migrations per week and per day. These are shown in the table below.

Overseas net migration ranked and ratios- Australia & South Australia (2022-23)

Australia net migration

Country	Net migration 2022-23	Per week	Per day
India	92,940	1,787	255
China	64,310	1,237	176
Philippines	40,890	786	112
Nepal	27,460	528	75
Colombia	22,320	429	61
UK, CIs & IOM	19,230	370	53
Vietnam	17,770	342	49
Pakistan	17,290	333	47
New Zealand	15,340	295	42
Thailand	14,060	270	39
Bhutan	13,160	253	36
Sri Lanka	13,110	252	36
Indonesia	12,380	238	34
Brazil	11,900	229	33
Ireland	8,850	170	24
South Africa	8,780	169	24
Taiwan	8,200	158	22
Fiji	7,680	148	21
Other Countries	102,010	1,962	279
Total	**517,680**	**9,955**	**1,418**

South Australia net migration

Country	Net migration 2022-23	Per week	Per day
India	7,810	150	21
China	2,970	57	8
Philippines	1,930	37	5
Nepal	1,400	27	4
Colombia	1,310	25	4
Vietnam	1,220	23	3
Sri Lanka	1,000	19	3
UK, CIs & IOM	760	15	2
Afghanistan	730	14	2
Pakistan	720	14	2
Brazil	690	13	2
South Africa	610	12	2
Hong Kong	590	11	2
Kenya	450	9	1
Indonesia	440	8	1
Thailand	420	8	1
Iran	420	8	1
Fiji	360	7	1
Other countries	3,950	76	11
Total	**27,780**	**534**	**76**

Source: ABS overseas net migration 2022-23.

Temporal analysis

To understand current datasets or financial information, insight can come from understanding historical and future trends. This gives us a sense of how the current data compares with the past and how they have and are projected to change. Temporal analysis involves analysing changes over time (this includes trend and growth analysis).

Overseas net migrations - Australia (2022-23 vs 2012-13)

Australia net migration

Country	2012-13	Share of total	Net migration 2022-23	Share of total	Total growth	Annual growth	change in share
India	23,670	10.3%	92,940	18.0%	292.6%	14.7%	7.7%
China	22,590	9.8%	64,310	12.4%	184.7%	11.0%	2.6%
Philippines	14,210	6.2%	40,890	7.9%	187.8%	11.1%	1.7%
Nepal	3,180	1.4%	27,460	5.3%	763.5%	24.1%	3.9%
Colombia	1,560	0.7%	22,320	4.3%	1330.8%	30.5%	3.6%
UK, CIs & IOM	25,030	10.9%	19,230	3.7%	-23.2%	-2.6%	-7.1%
Vietnam	8,000	3.5%	17,770	3.4%	122.1%	8.3%	0.0%
Pakistan	6,250	2.7%	17,290	3.3%	176.6%	10.7%	0.6%
New Zealand	24,740	10.7%	15,340	3.0%	-38.0%	-4.7%	-7.8%
Thailand	3,690	1.6%	14,060	2.7%	281.0%	14.3%	1.1%
Bhutan	590	0.3%	13,160	2.5%	2130.5%	36.4%	2.3%
Sri Lanka	5,140	2.2%	13,110	2.5%	155.1%	9.8%	0.3%
Indonesia	1,560	0.7%	12,380	2.4%	693.6%	23.0%	1.7%
Brazil	1,450	0.6%	11,900	2.3%	720.7%	23.4%	1.7%
Ireland	5,670	2.5%	8,850	1.7%	56.1%	4.6%	-0.7%
South Africa	5,510	2.4%	8,780	1.7%	59.3%	4.8%	-0.7%
Taiwan	7,840	3.4%	8,200	1.6%	4.6%	0.4%	-1.8%
Fiji	1,850	0.8%	7,680	1.5%	315.1%	15.3%	0.7%
Other Countries	68,090	29.5%	102,010	19.7%	49.8%	4.1%	-9.8%
Total	**230,620**	**100.0%**	**517,680**	**100.0%**	**124.5%**	**9.4%**	

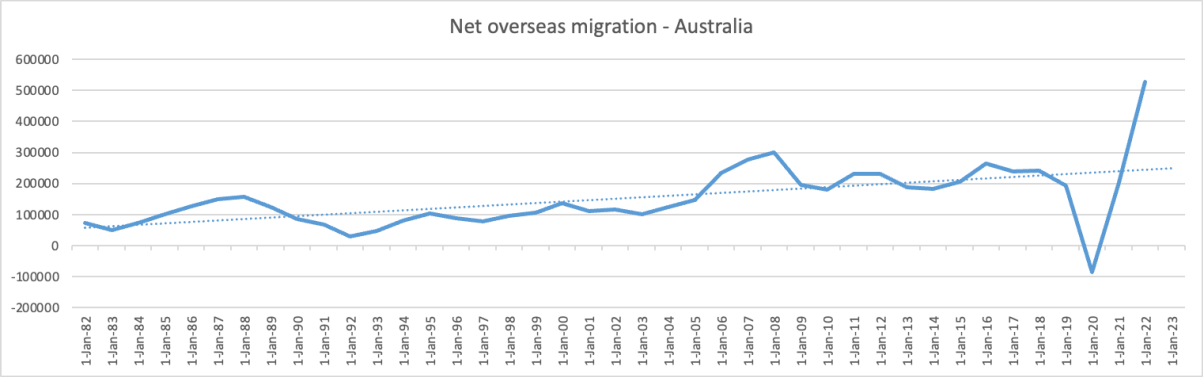
Source: ABS overseas net migration 2022-23.

Annual growth is a compound annual growth rate

Some possible traps associated with temporal analysis:

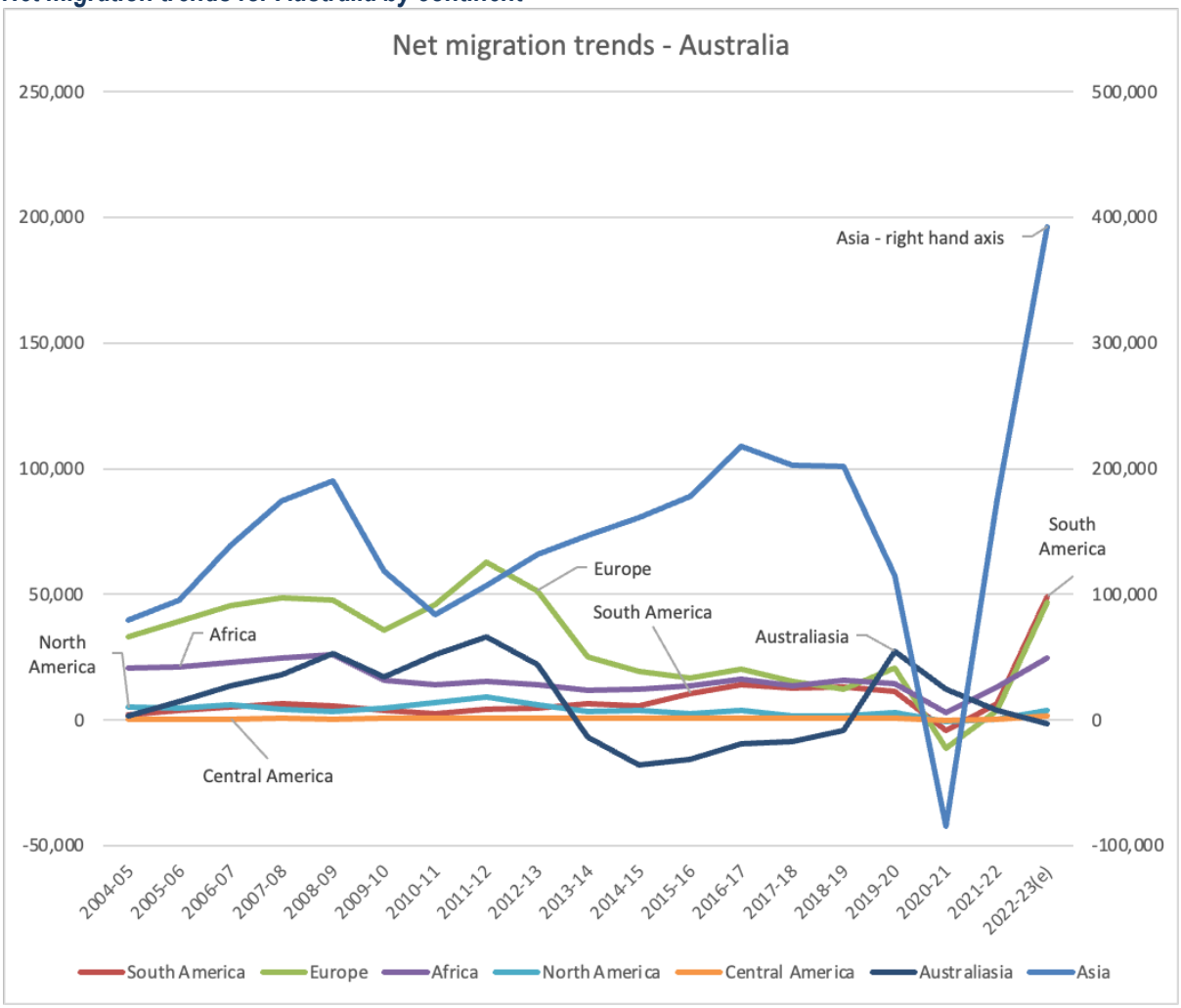
- Selecting beginning and end points to find the 'right' story. This can occur by selecting a low beginning point and high-end point to exaggerate growth or vice-versa (a high beginning point and low-end point to exaggerate declines).
- Growth rates are sensitive to the relative size of the initial point. When the starting point is low, growth rates can be high.
- Quoting growth rates on absolute numbers or nominal numbers when adjustments should be made for underlying changes like population or inflation.

Temporal analysis is aided by trend data presented in a graph over the page.



The long-term trend shows an increase in net migration in the early 2000s.

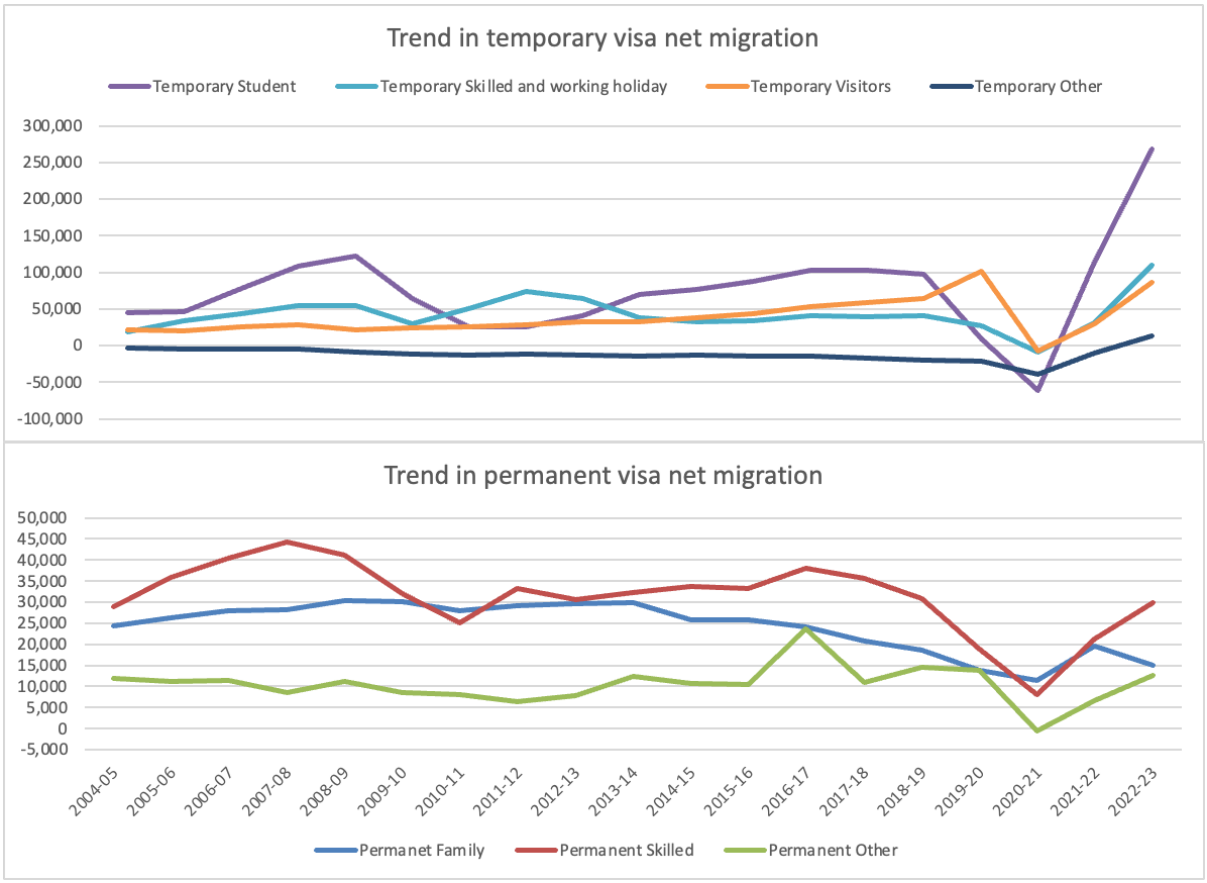
Net migration trends for Australia by continent



We can also look at migration by groups to see another perspective

Australia net migration by group

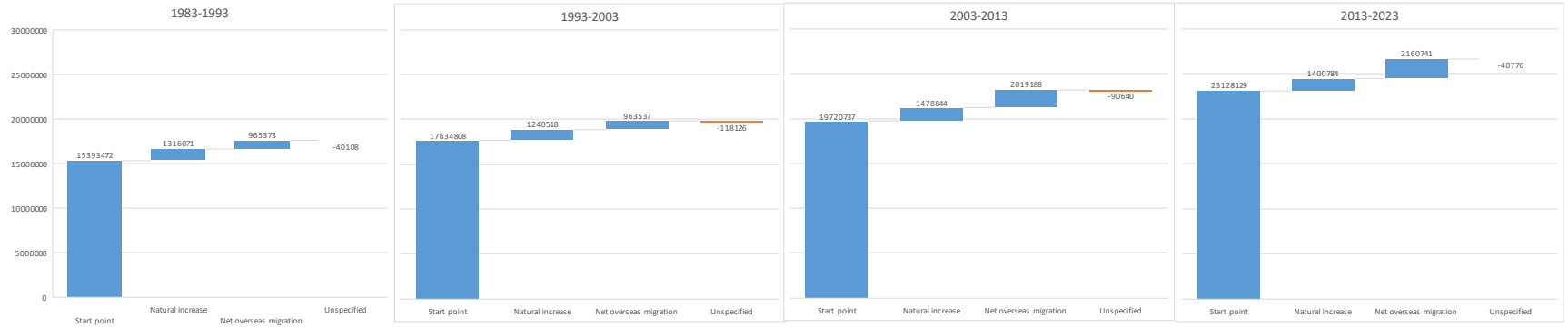
	2012-13	share %	2022-23	share %	Growth	Annual growth
Permanent visas	**68,080**	**31%**	**57,560**	**11%**	**-15.5%**	**-1.7%**
Family	29,670	14%	15,020	3%	-49.4%	-6.6%
Skilled (permanent)	30,540	14%	29,850	6%	-2.3%	-0.2%
Special eligibility & humanitarian	8,400	4%	12,590	2%	49.9%	4.1%
Other (permanent)	-530	0%	100	0%	-118.9%	na
Temporary Visas	**125,030**	**57%**	**477,150**	**91%**	**281.6%**	**14.3%**
Student - higher education	28,130	13%	175,660	34%	524.5%	20.1%
Student - vocational education and training	520	0%	32,100	6%	6073.1%	51.0%
Student - other	12,610	6%	59,960	11%	375.5%	16.9%
Skilled (temporary)	25,220	12%	43,130	8%	71.0%	5.5%
Visitors	32,890	15%	86,660	17%	163.5%	10.2%
Working holiday	38,400	18%	66,560	13%	73.3%	5.7%
Other (temporary)	-12,740	-6%	13,080	2%	-202.7%	na
Australia and New Zealand persons	**25,470**	**12%**	**-10,830**	**-2%**	**-142.5%**	**na**
Total	**218,580**		**523,880**		**139.7%**	**9.1%**



Growth has been in temporary visa net migration – most notably students and temporary skilled and working holiday visas.

We can look at the relative scale of these figures by looking at how they compare with changes in Australia's overall population.

Changes in population – Australia



The changes in population show net overseas migration contributing just under 100,000 per annum to the population between the years of 1983 and 2003. Since 2003, net migration has added over 200,000 people per annum to the population.

Variance analysis

To date, we have been looking at the composition of our data and comparing changes over time. Now we need to make comparisons with other measures. For net migration data, we can see if there are variances from projections made by the ABS in 2004 and variances by state for different countries. For financial data, the most common form of variance analysis is to examine whether or not actual results vary from budget.

Net overseas migration – actual for 2023 compared to 2023 projections made in 2004

Age	Projected in 2004	Actual in 2023	Variance	Variance %
0-4	1,328,086	1,513,072	184,986	13.9%
5-9	1,350,176	1,612,127	261,951	19.4%
10-14	1,381,488	1,638,761	257,273	18.6%
15-19	1,414,576	1,533,919	119,343	8.4%
20-24	1,458,427	1,637,698	179,271	12.3%
25-29	1,559,135	1,827,118	267,983	17.2%
30-34	1,634,479	1,915,630	281,151	17.2%
35-39	1,634,873	1,894,505	259,632	15.9%
40-44	1,615,879	1,704,808	88,929	5.5%
45-49	1,550,183	1,622,490	72,307	4.7%
50-54	1,618,336	1,653,689	35,353	2.2%
55-59	1,492,677	1,534,811	42,134	2.8%
60-64	1,507,649	1,492,429	-15,220	-1.0%
65-69	1,332,951	1,302,941	-30,010	-2.3%
70-74	1,192,154	1,144,905	-47,249	-4.0%
75-79	968,246	872,845	-95,401	-9.9%
80-84	619,402	565,473	-53,929	-8.7%
85 and over	622,604	547,178	-75,426	-12.1%
	24,281,321	**26,014,399**	**1,733,078**	**7.1%**

In this context, variance analysis is different from analysis of variance

The first (variance analysis) is used in areas like finance where variances are used to show a difference between a budget (or forecast) and an actual. The variances are presented in dollar terms and of often presented in percentage terms. Different organisations will have views about the materiality of these variances. A materiality threshold could be expressed in dollar terms (e.g. all variances over $10,000 are deemed material and need to be explained). The materiality threshold could also be expressed in percentage terms (e.g. all variances greater than 2% or less than -2% need to be explained along with remedies).

Variances have a different meaning in statistics.

In statistics, variance measures variability from the average or mean.

Variance tells you the degree of spread in your data set. The more spread the data, the larger the variance is in relation to the average.

Analysis of variance is used when testing a hypothesis. It is typically used when comparing two populations (e.g. a test group using a new drug and a control group using a placebo). For each group, calculations will include a sample size, an average result and a variation.

Overseas migration data - SA Variance from Australia (2022-23)

	% share		Variance
Country of birth	**Australia**	**South Australia**	**South Australia**
India	17.9%	28.1%	10.2%
China	12.4%	10.7%	-1.7%
Philippines	7.9%	6.9%	-0.9%
Nepal	5.3%	5.0%	-0.3%
Colombia	4.3%	4.7%	0.4%
Vietnam	3.4%	4.4%	1.0%
UK, CIs & IOM	3.7%	2.7%	-1.0%
Pakistan	3.3%	2.6%	-0.7%
Sri Lanka	2.5%	3.6%	1.1%
Thailand	2.7%	1.5%	-1.2%
New Zealand	3.0%	1.3%	-1.7%
Indonesia	2.4%	1.6%	-0.8%
Brazil	2.3%	2.5%	0.2%
Bhutan	2.5%	0.3%	-2.3%
Ireland	1.7%	0.2%	-1.5%
Taiwan	1.6%	0.8%	-0.8%
South Africa	1.7%	2.2%	0.5%
Hong Kong	1.4%	2.1%	0.7%
Afghanistan	1.2%	2.6%	1.4%
Fiji	1.5%	1.3%	-0.2%
Other countries	17.1%	14.8%	-2.3%
	100.0%	**100.0%**	

Source: ABS overseas net migration 2022-23.

Benchmarking analysis

To date, we have been analysing a data set for one unit (Australia) on its own. An important aspect of data analysis is to compare our unit with external parties using benchmark analysis.

We can compare other aspects of the data with other jurisdictions or between jurisdictions in Australia

A common requirement of benchmarking is that you must compare like with like or apples with apples. This is often taken too far. No two entities are alike. We necessarily need to compare different entities to identify the nature and possible causes of differences.

For our dataset, we can compare overseas migration across states like South Australia, Tasmania and Victoria.

Overseas migration by top 20 countries – Victoria, South Australia & Tasmania (2022-23)

Country of birth	Australia	Victoria	South Australia	Tasmania
India	92,940	35,460	7,810	950
China	64,320	21,270	2,970	190
Philippines	40,890	10,220	1,930	290
Nepal	27,450	4,270	1,400	510
Colombia	22,330	7,470	1,310	60
Vietnam	17,770	7,360	1,220	140
UK, CIs & IOM	19,230	4,650	760	60
Pakistan	17,280	5,190	720	220
Sri Lanka	13,120	6,900	1,000	210
Thailand	14,060	6,220	420	80
New Zealand	15,340	4,600	350	40
Indonesia	12,370	3,340	440	80
Brazil	11,910	1,200	690	30
Bhutan	13,160	380	80	40
Ireland	8,850	2,330	50	10
Taiwan	8,200	2,510	220	70
South Africa	8,770	1,360	610	110
Hong Kong	7,480	2,550	590	70
Afghanistan	6,340	2,880	730	90
Fiji	7,680	1,800	360	50
Other countries	88,560	22,280	4,120	650
	518,050	**154,240**	**27,780**	**3,950**

Source: ABS overseas net migration 2022-23.

What is the main problem with this comparison?

Multivariate analysis

To date, all the analysis we have done has been of specific data in isolation from other data. Data does not exist in isolation from other variables. Therefore understanding our dataset better will involve combining our data with other variables.

In some cases, we can combine data relating to finances with output data or caseload data to determine costs per unit of output or cost per case. This is often important in doing benchmarks as we need to take into account the different sizes of states or organizations to make data more comparable.

In the case of comparing states, we can compare data figures with the population of each state in order to determine indicators per capita.

The tables on the following pages present our data on a per capita basis.

Net overseas migration per 1,000,000 of the population – Victoria, South Australia & Tasmania (2022-23)

Country of birth	Victoria	South Australia	Tasmania
India	5,278	4,254	1,660
China	3,166	1,618	332
Philippines	1,521	1,051	507
Nepal	636	763	891
Colombia	1,112	714	105
Vietnam	1,095	665	245
UK, CIs & IOM	692	414	105
Pakistan	772	392	384
Sri Lanka	1,027	545	367
Thailand	926	229	140
New Zealand	685	191	70
Indonesia	497	240	140
Brazil	179	376	52
Bhutan	57	44	70
Ireland	347	27	17
Taiwan	374	120	122
South Africa	202	332	192
Hong Kong	380	321	122
Afghanistan	429	398	157
Fiji	268	196	87
Other countries	3,316	2,244	1,136
	22,956	**15,132**	**6,903**

Source: ABS overseas net migration 2022-23.

States net migration by group for 2022-23 (per 1 million of population)

	Victoria	South Australia	Tasmania
Permanent visas	**2,834**	**3,148**	**1,678**
Family	771	534	140
Skilled (permanent)	1,360	1,983	979
Special eligibility & humanitarian	643	681	577
Other (permanent)	60	-49	-17
Temporary Visas	**20,347**	**12,800**	**6,151**
Student - higher education	7,641	5,866	2,219
Student - vocational education and training	1,268	790	437
Student - other	2,688	1,738	507
Skilled (temporary)	1,828	828	629
Visitors	3,998	2,146	1,241
Working holiday	2,407	752	542
Other (temporary)	518	681	577
Australia and New Zealand persons	**-63**	**-665**	**-594**
Total	**23,118**	**15,284**	**7,235**

Source: ABS overseas migration 2022-23.

Discussion

To date, the data considered has looked at migrants and students presuming each is the same (i.e. each is equal to one). For the purposes of developing policy, we may need to be more sophisticated than this. What can we do to improve the quality of our analysis?

Standardising data

When using financial data, a dollar is a dollar. Finances compare across organisations and functions because of the standard method of valuation - the dollar. Other forms of data cannot always rely on such standardisation or consistency in value.

Relevant examples of where data is not easily comparable include:

- Visitors to hospitals – counting the raw number of visitors does not take into account the relative complexity of each visitor
- Number of court trials or cases– counting the number of trials or cases masks the relative complexity of each trial.
- Number of transactions in a customer service centre – this data will mask both the value and the complexity of transactions.

As a result, methods of standardising data are needed to provide data sets and summary analysis that are more sophisticated than raw counts. Using the examples above, standardising data will require the application of weights that represent the relative complexity of each transaction. In health activity measures in hospitals are weighted to enable an analysis of standard measures of activity. Sophisticated weightings have been developed for different groups of patient treatments and episodes of care. In education, weights have been developed for different student groups to determine the funding needs of schools.

For customer service transactions, the raw count of transactions can be weighted by the time for each transaction. For example:

- Renewing a motor vehicle registration takes 3 minutes
- Renewing a licence (that includes a photo) takes 6 minutes.
- The registration renewal will be deemed a standard transaction.
- Each licence renewal will therefore represent 2 standard transactions.

A simple table showing transaction volumes compares the raw transaction count compared with the standardised or adjusted transaction count.

Number of transactions	Raw	Standardised
Registration renewals	1,000	1,000
Licence renewals	500	1,000
Total Transactions	**1,500**	**2,000**

Based on raw data, there were 1,500 transactions. However, after taking into account the relative complexity of the licence renewals, the standardised data shows 2,000 transactions.

A **full-time equivalent (FTE)** is a measure of standardised data. We may have four employees working in a team, but if two of those employees are working part-time and two are full-time, they should not each be counted as one employee for budgeting purposes. If the part-time staff are working three-day weeks, they are working 60% of the hours of a full-time employee and therefore would be counted as 0.6 FTE for budgeting purposes. Therefore our team of four employees is counted as 3.2 FTE for budgeting purposes.

An example of standardised data used in education is EFTSL. Equivalent full-time student load (EFTSL) is the measure used to determine a student's study load. EFTSL is required by external organisations or government departments, as it provides a universal measurement of study load and is typically used to determine eligibility for funding, payments and allowances.

Links to other data – International student enrolments and the economy

One of the main drivers of migration figures has been the growth in temporary visas, in particular, for students. This provides the opportunity to do some analysis of student data.

The table below summarises the increases in enrolments in education from international students since 2005.

International enrolments for Australia and South Australia

	Australia enrolments				South Australia enrolments			
	2005	2023	Growth	CAGR	2005	2023	Growth	CAGR
Higher Education	150,968	334,819	121.8%	4.5%	8,454	21,857	158.5%	5.4%
VET	36,980	231,886	527.1%	10.7%	968	12,505	1191.8%	15.3%
ELICOS	39,920	103,182	158.5%	5.4%	1,724	4,051	135.0%	4.9%
Non-award	16,737	18,871	12.8%	0.7%	1,222	1,102	-9.8%	-0.6%
Schools	21,461	11,927	-44.4%	-3.2%	1,655	1,494	-9.7%	-0.6%
Total	**266,066**	**700,685**	**163.4%**	**5.5%**	**14,023**	**41,009**	**192.4%**	**6.1%**

Source: Commonwealth Provider Registration and International Student Management System (PRISMS) database.

We can also analyse this data by country to see the top-ranked countries of origin for Australia and South Australia.

International student enrolments by country for Australia and South Australia

	Australian Enrolments					
	2005	share	2023	share	Growth	CAGR
China	64,340	24.2%	143,641	20.5%	123.3%	4.6%
India	19,759	7.4%	116,543	16.6%	489.8%	10.4%
Nepal	839	0.3%	62,338	8.9%	7330.0%	27.0%
Colombia	1,188	0.4%	37,109	5.3%	3023.7%	21.1%
Thailand	12,189	4.6%	28,609	4.1%	134.7%	4.9%
Philippines	875	0.3%	27,237	3.9%	3012.8%	21.0%
Vietnam	4,103	1.5%	26,454	3.8%	544.7%	10.9%
Brazil	4,203	1.6%	24,351	3.5%	479.4%	10.3%
Pakistan	1,667	0.6%	21,635	3.1%	1197.8%	15.3%
Indonesia	13,315	5.0%	18,754	2.7%	40.8%	1.9%
Malaysia	17,204	6.5%	15,967	2.3%	-7.2%	-0.4%
Sri Lanka	3,000	1.1%	13,462	1.9%	348.7%	8.7%
Hong Kong	15,825	5.9%	13,342	1.9%	-15.7%	-0.9%
Korea, Republic of (South)	19,625	7.4%	12,836	1.8%	-34.6%	-2.3%
Japan	14,746	5.5%	10,015	1.4%	-32.1%	-2.1%
Other countries	73,188	27.5%	128,392	18.3%	75.4%	3.2%
Total	**266,066**	**100.0%**	**700,685**	**100.0%**	**163.4%**	**5.5%**

	South Australian Enrolments					
	2005	share	2023	share	Growth	CAGR
India	1,103	7.9%	11,744	28.6%	964.7%	14.0%
China	3,931	28.0%	7,427	18.1%	88.9%	3.6%
Nepal	13	0.1%	2,924	7.1%	22392.3%	35.1%
Vietnam	234	1.7%	2,194	5.4%	837.6%	13.2%
Colombia	25	0.2%	2,013	4.9%	7952.0%	27.6%
Hong Kong	1,128	8.0%	1,790	4.4%	58.7%	2.6%
Philippines	31	0.2%	1,311	3.2%	4129.0%	23.1%
Brazil	46	0.3%	1,074	2.6%	2234.8%	19.1%
Sri Lanka	74	0.5%	1,010	2.5%	1264.9%	15.6%
Pakistan	96	0.7%	836	2.0%	770.8%	12.8%
Kenya	53	0.4%	809	2.0%	1426.4%	16.3%
Malaysia	1,908	13.6%	795	1.9%	-58.3%	-4.7%
Thailand	336	2.4%	632	1.5%	88.1%	3.6%
Bangladesh	76	0.5%	507	1.2%	567.1%	11.1%
Japan	884	6.3%	484	1.2%	-45.2%	-3.3%
Other countries	4,085	29.1%	5,459	13.3%	33.6%	1.6%
Total	**14,023**	**100.0%**	**41,009**	**100.0%**	**192.4%**	**6.1%**

Source: Commonwealth Provider Registration and International Student Management System (PRISMS) database.

The increase in international student enrolments is also reflected in economic data for service exports. The service export data shows the increasing significance of international students to our exports and our economy.

The tables below show the growth in education services exports relative to total exports and as a share of the total economy (GDP & GSP).

Education service exports for Australia and South Australia

	2004-05	2022-23	growth	CAGR
Australia				
Education services exports ($m)	8,570	36,423	325%	8.37%
Total exports ($m)	167,669	686,043	309%	8.14%
Education services exports % of exports	5.1%	5.3%		
GDP ($m)	1,509,274	2,403,614	59%	2.62%
Education service exports % GDP	0.6%	1.5%		

Source: ABS international trade and national accounts

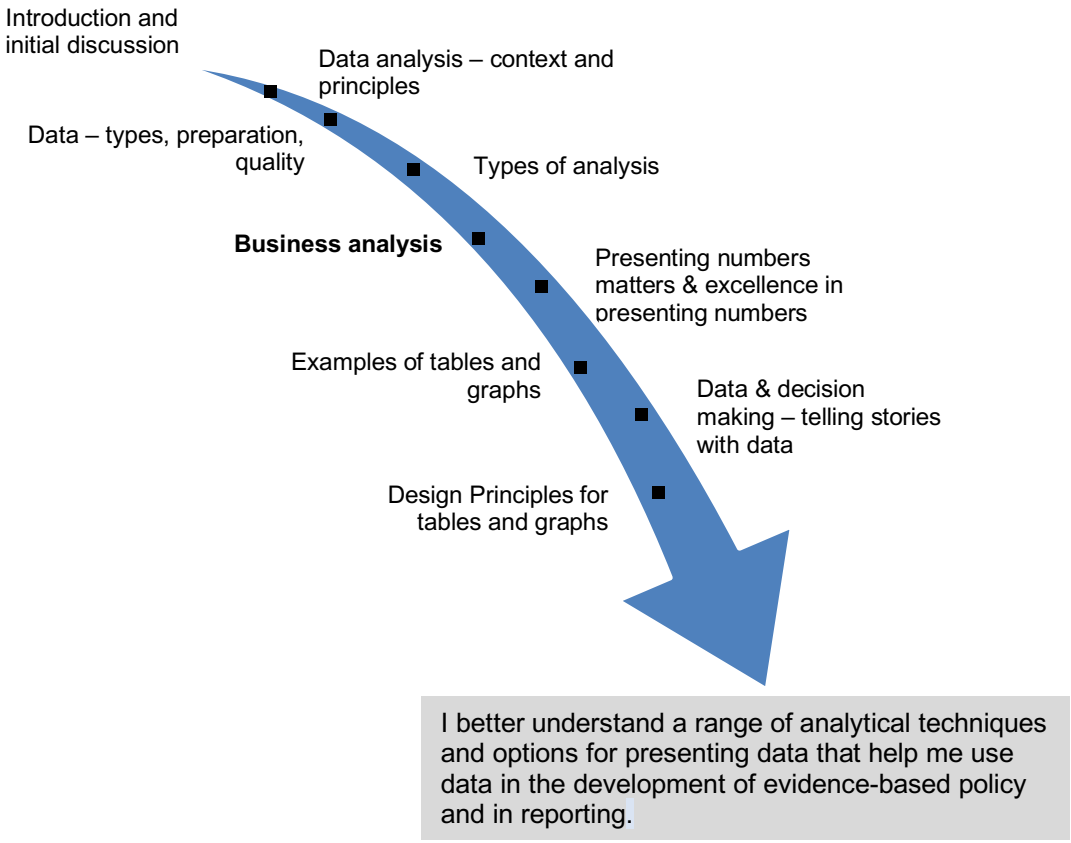
	2004-05	2022-23	growth	CAGR
South Australia				
Education services exports ($m)	459	2,477	440%	9.82%
Total exports ($m)	9,199	21,042	129%	4.70%
Education services exports % of exports	5.0%	11.8%		
GSP ($m)	62,984	142,035	126%	4.62%
Education service exports % GSP	0.7%	1.7%		

Source: ABS international trade and national accounts

The table shows the importance of education services exports to the South Australian economy with it growing to 11.8% of all exports and more than doubling its share of Gross State Product.

Reflection

What has been my most important learning and why?



Business Analysis

Analysis is the process of breaking something into its constituent elements.

Consider an enterprise, business or project.

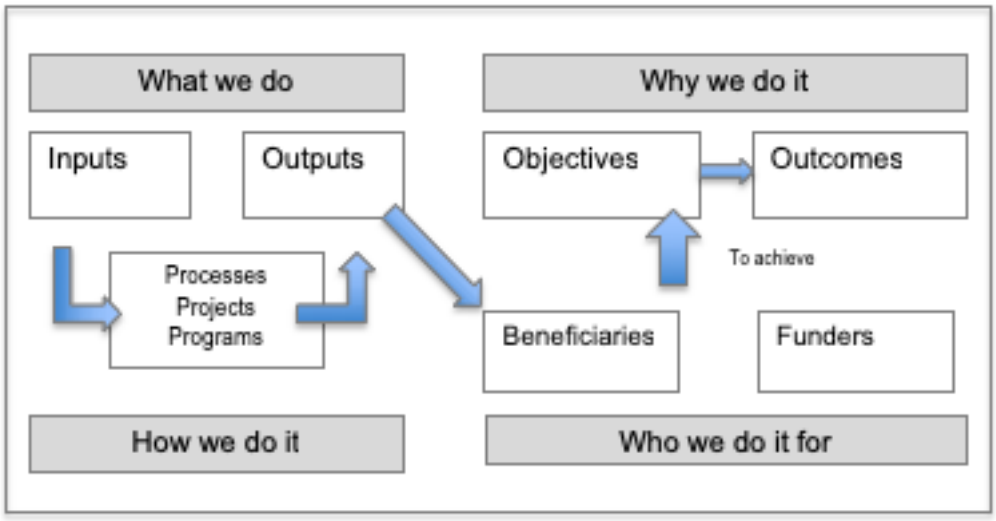
Discuss how can we break these down into their constituent elements to gain a better understanding of them.

Business Model Perspective

"A business model describes the rationale of how an organisation creates, delivers and captures value."

> Osterwalder, A., & Pigneur, Y. (2010). *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers.* New Jersey: John Wiley & Sons.

Consider the business model below that presents the logic of an organisation in terms of what it does, how it does it, who it does it for and why it does it.



Key Components of a Business Model

What the business does:

- Converts inputs (staff, goods and services and technology)
- Into outputs (goods and services)

How the organisation does it:

- Processes
- Projects
- Programs

Who is involved?

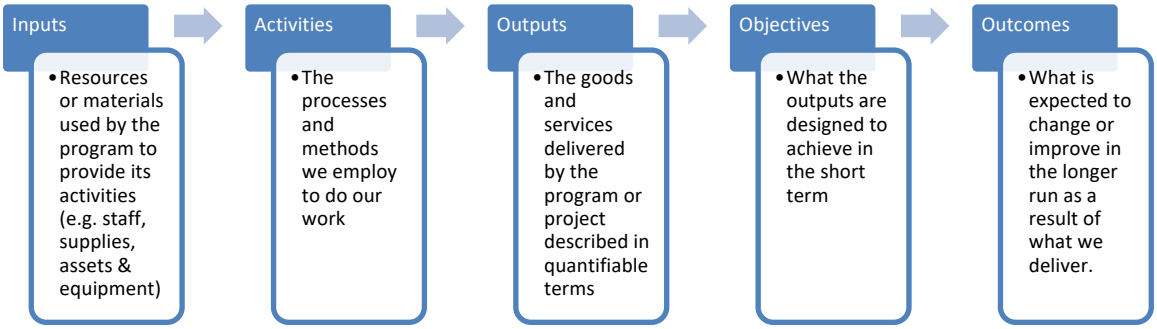
- Suppliers
- Employees
- Beneficiaries
- Funders/Financiers
- Owners

Why?

- Purpose
- Objectives
- Outcomes

The combination of answers to these questions starts to define the business model of an organisation or enterprise.

Embedded within a business model is a basic logic:



This logic can provide a framework for the types of data we might want to analyse to understand a business. It can help ensure we distinguish the following types of data:

- Input data
- Activity data
- Output data
- Objective and
- Outcome data.

In your groups, consider examples of types of data that may belong in each category

Category	**Example**	**Example**
Input data		
Activity data		
Output data		
Objective and		
Outcome data.		

Examples of performance measures using a logic model are presented below.

	Training organisation	Fire service
Input measures	number of staff Facilities (rooms, PCs, desks) annual total budget	number of firefighters number of fire appliances number of fire stations annual total budget
Activity measures	Training sessions scheduled Registrations for sessions Development of materials & resources	average time to dispatch a fire truck percentage of incidents reached by an appliance within X minutes
Output measures	Materials produced Courses delivered Assessments completed	number of incidents attended
Objective measures	Learning goals achieved students qualified attendee satisfaction	percentage of incidences where fire is contained citizen satisfaction
Outcome measures	Career progression of attendees Manager/business satisfaction	number of deaths per thousand fires perceptions of public safety

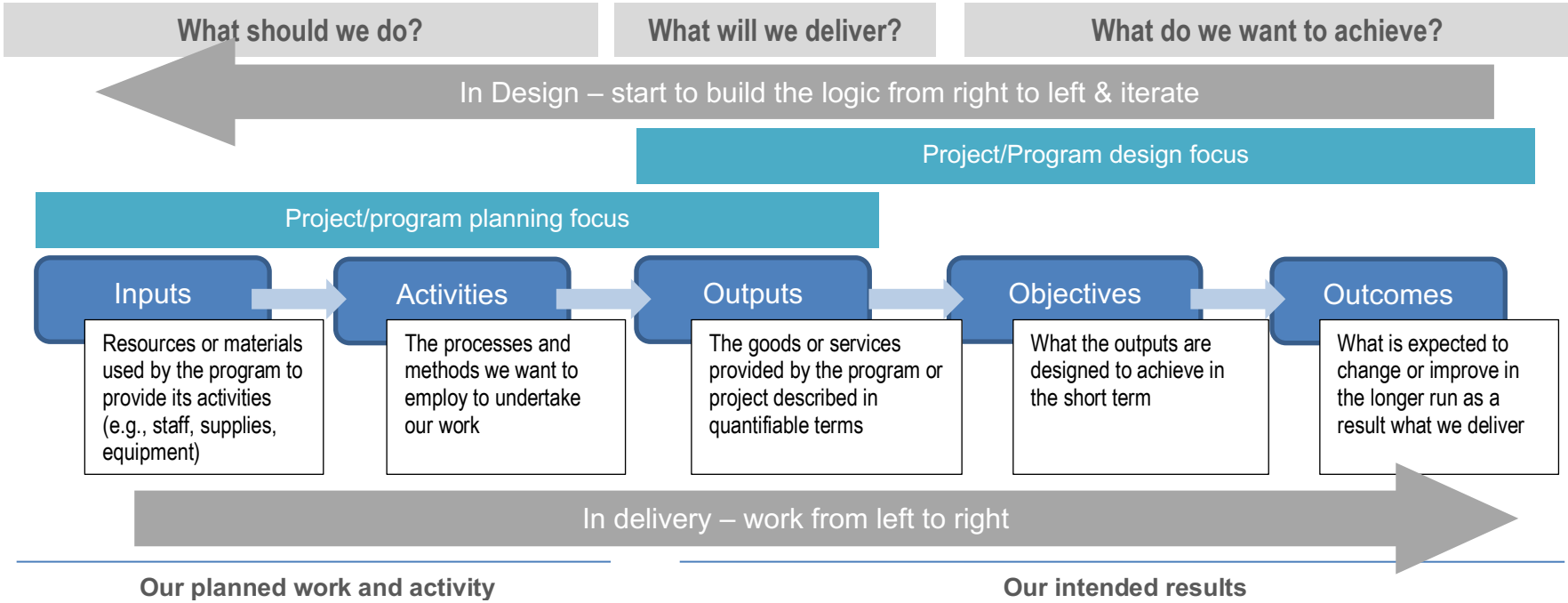
Source (fire service): *Financial Management and Accounting in the Public Sector* - Author Gary Bandy.

It is also possible to take this further, by breaking the data into:

- Quantity
- Quality
- Cost and
- Effectiveness data.

A useful tool for the design and planning of initiatives and the evaluation of initiatives is a logic model.

"A logic model is a systematic and visual way to present and share your understanding of the relationships among the resources you have to operate your program (or project), the activities you plan, and the changes or results you hope to achieve." Kellogg's Foundation.



A logic model can help create a shared understanding of, and focus on, the goals and methodology of an initiative and in relating inputs and activities to projected outcomes.

The design phase of an initiative or program is primarily focussed on the relationship between outputs, objectives and outcomes (the three elements to the right of the model) – will the work deliver what is needed to achieve goals (objectives as more immediate goals and outcomes as longer-term goals).

The planning phase examines in more detail the relationships between inputs, activities and outputs (the three elements to the left of the model)

Logic models are used as tools to support funding bids or grant applications with applicants required to demonstrate links between inputs and funding required and outcomes sought.

The logic model is based on a series of 'if–then' statements. If we have these inputs, we can undertake these activities and produce these outputs. If we deliver these outputs, we will achieve certain objectives. If we achieve certain objectives, then we will achieve certain outcomes.

When designing a strategy or initiative data is often sought and analysed to establish these 'if-then' relationships. For example, when designing a road safety strategy, data will be collected and analysed to understand the root causes of accidents, casualties and fatalities.

The model is used by the Victorian Government as the basis for funding bids and departmental funding models. Below is an extract from the Victorian Government Resource Management Framework

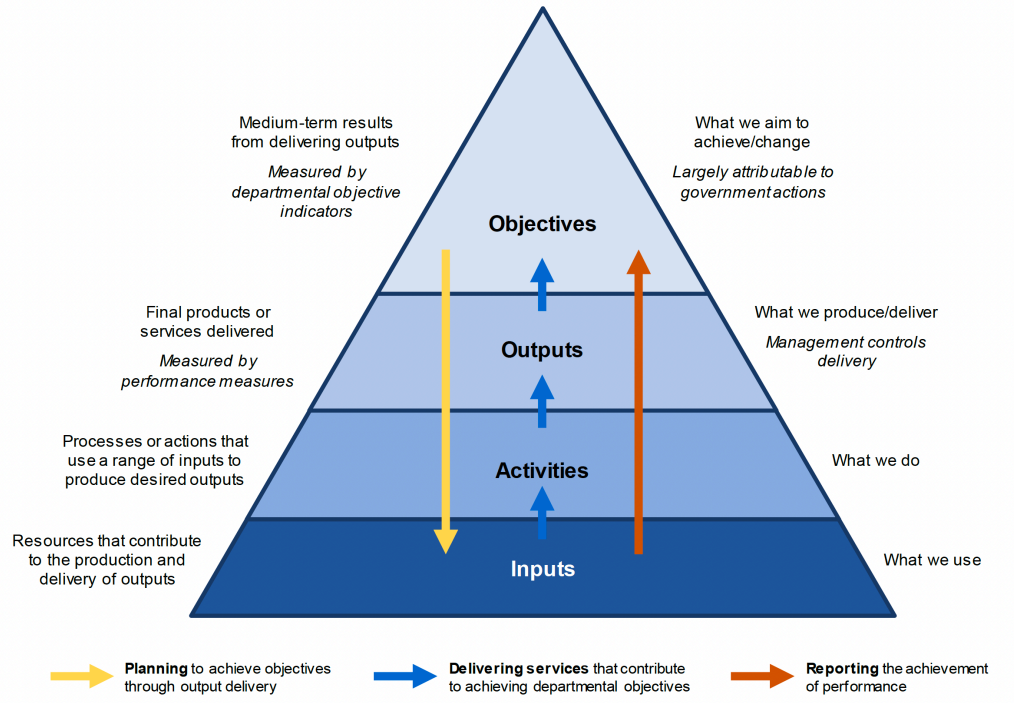
The Departmental Funding Model

Victoria has a system that devolves responsibility to portfolio ministers and departments to manage a global budget in delivering certain agreed outputs, which in turn are aligned with departmental objectives. This is known as the Departmental Funding Model (DFM).

Under the DFM, the Government contracts the delivery of goods and services to departments. Departments have flexibility in the way services are delivered but must account for the delivery of these services. This is an incentive for departments to perform and improve service delivery over time.

The DFM contains the concepts of outputs, objectives, and their associated indicators and performance measures. These are explained further in Figure 6, which illustrates the relationship between objectives and outputs, and the activities and inputs that contribute to outputs.

Figure 6: Performance management concepts



Reflection

What has been a key learning for me and why?

References

Australian Bureau of Statistics

Catalogue 1500.0 - A guide for using statistics for evidence-based policy, 2010

Understanding statistical language - https://www.abs.gov.au/websitedbs/D3310114.nsf/Home/Statistical+Language?OpenDocument

ABS Data Quality Framework

(http://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/1520.0Main%20Features10May%202009?opendocument&tabname=Summary&prodno=1520.0&issue=May%202009&num=&view=) – to help you assess the quality of a dataset and design data collections which are fit for purpose.

Moment of Clarity (2014), Christian Madsbjerg, Mikkel Vedby Rasmussen

Naked Statistics (2013), Charles Wheelan

Everybody Lies (2017), Seth Stephens-Davidowitz

I Think You'll Find It's a Bit More Complicated Than That (2015), Ben Goldacre

DataStory (2019), Nancy Duarte

Storytelling with Data (2015), Cole Nussbaumer Knaflic

The Data Detective (2021), Tim Harford

Style Manual, Commonwealth of Australia

Edward Tufte:

Website: www.edwardtufte.com/tufte/index

Books:

The Visual Display of Quantitative Information

Visual Explanations

Envisioning Information

Beautiful Evidence

Stephen Few

Website: www.perceptualedge.com

Books:

Show Me the Numbers

Signal

Information Dashboard Design

The Data Loom

Sally Bigwood & Melissa Spore

Website: www.plainfigures.com

Book: Presenting Numbers, Tables and Charts

Appendix 1: Assorted terminology and principles

Ackoff's Path

Analysis helps me navigate along a path described by Russell Ackoff (from Data to Wisdom)

Data – Information – Knowledge – Understanding - Wisdom

I do analysis to organise, discover and learn from data. I present data and analysis to help others discover and learn from data.

Data – a set of symbols or signals. In its raw form, it is unorganised and unprocessed. Census takers collect data.

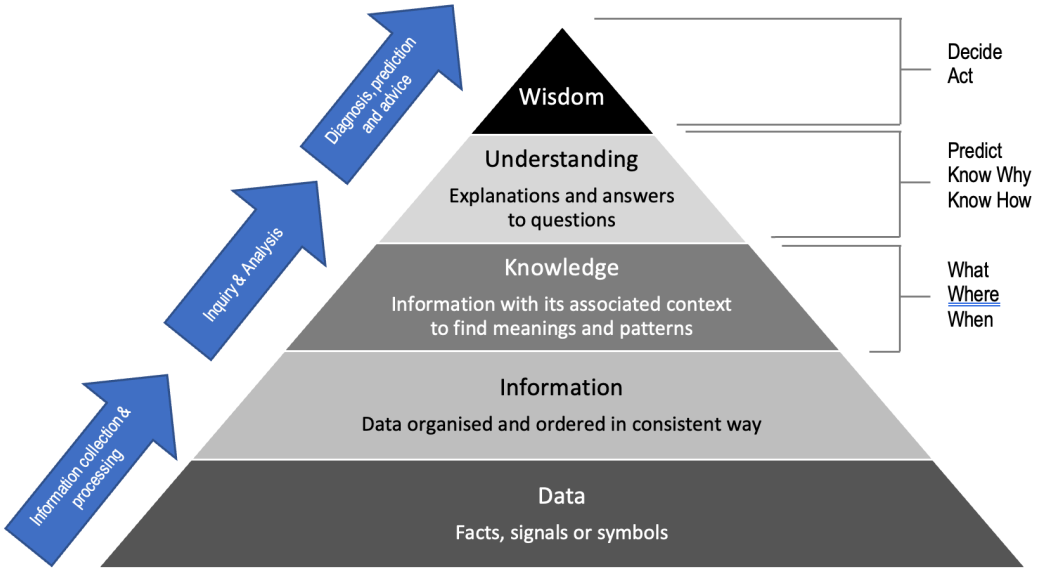
Information – consists of processed data. We apply systems to classify, organise and arrange data. Information is data with meaning. Data from the census is converted into tables that present the results from the census takers.

Knowledge – While information can help us understand relationships, knowledge is about finding meaning and patterns. Knowledge is conveyed by instructions – answers to how-to questions. From the census, we know what the average age of Australians is or how many Australians are living in different locations. We need to do some analysis to determine this.

Understanding – understanding is conveyed by explanations – answers to why questions. We need to do further analysis to answer why questions. The difference between knowledge and understanding is the difference between memorising (knowledge) and learning (understanding). Our analysis helps us understand why the average age has changed or why Australians have changed where they live.

Wisdom – apply our knowledge for future prediction and decision-making. Wisdom involves the exercising of judgement – to adapt our understanding to different circumstances requiring discernment. We can make policy decisions that improve health outcomes or improve quality of life.

Analysis is most heavily used to get from Information to Knowledge and from Knowledge to Understanding.



The MECE Principle (McKinsey Way)

"The **MECE principle**, pronounced 'me see', is a grouping principle for separating a set of items into subsets that are mutually exclusive and collectively exhaustive." Wikipedia.

The term has been popularized by McKinsey as is deemed part of the McKinsey Way.

These terms are traditionally associated with probability theory but combine well as a principle for analysing data and setting up data sets.

Mutually Exclusive

Related in such a way that each thing makes the other thing impossible: not able to be true at the same time or to exist together

Pertains to the categorization of data – categories should be mutually exclusive to ensure no double counting and to remove ambiguity on where items belong.

Completely Exhaustive

Does the data set capture the totality of what it is you are analysing? If so, the data set is completely exhaustive. It is a common failure of analysis to focus too narrowly on a subset of data in isolation from the total data set. This causes valuable information to be lost and can cause a change in conclusions from changes in scope.

Correlation vs Causation

The volume of data does not adequately address what is causing what. We need to look at causation vs correlation.

Correlation - Correlation is a statistical measure that indicates the extent to which two or more variables fluctuate together. A positive correlation indicates the extent to which those variables increase or decrease in parallel; a negative correlation indicates the extent to which one variable increases as the other decreases

Causation - Connection between two events or states such that one produces or brings about the other; where one is the cause and the other its effect. Also called causality

Correlation can often be confused with causation. Two variables can move together without it being clear which is causing which or with the possibility that a third variable could be triggering both.

The Importance of Metadata

Imagine you have just joined a new team and you have been asked to analyse data for a monthly report.

You look at the data file and all the column headers are abbreviations, and you aren't sure what the data is in each of the columns.

You ask your colleague for written instructions on how the data is compiled and analysed so you can repeat the process, but nothing is written down and you must be shown.

You look at the last report to get further information and can see the data refers to licenses, service providers and consumers. Still, there is no further information about who service providers and consumers are, and what sort of licenses are included in the report.

Your team members can verbally provide you with some information about the data, but in some cases, no one knows because your predecessor "managed all that".

Your job would be so much easier if information about the data had been written down…

Data is an asset, and its users need information about the data to help them make better use of it. The descriptive information about the data we use is referred to as **metadata**.

Definition

Metadata is the information that defines and describes data.

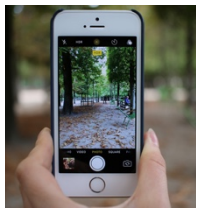
Metadata is the information you need to have to have an accurate understanding of the (associated) dataset.

Metadata helps people to correctly interpret the data. It also helps future producers of the same or similar collections of data to understand the collection or issues that might arise in a new but similar collection.

Metadata includes details of what the data is measuring, that is, exactly how the data items are defined. Any standards of classifications used to categorise data items are also part of the metadata, as is the sort of information contained in many explanatory notes, such as giving details of the scope of the collection. Source: ABS

For an everyday example consider what happens when you take a photo. You are creating data in the form of a saved digital image.

While what you see is the photo, your smartphone also creates metadata accompanying that image such as:



- Image size
- The time when the photo was taken
- Location of where the photo was taken
- A thumbnail of the image

Your phone manages this metadata, allowing the image to be easily searched and identified.

These descriptions will not only help those using the data, but they will also help those who are designing and building datasets as the descriptions provide rules for how data should be collected and structured. They will also help people providing ACMA data as the descriptions help them to understand exactly what data we are asking for.

Big data

"Extremely large data sets that may be analysed computationally to reveal patterns, trends, and associations, especially relating to human behaviour and interactions." Source Dictionary.com

"Big data is a term applied to data sets whose size of type is beyond the ability of traditional relational databases to capture, manage, and process the data with low latency. The data has one or more of the following characteristics – high volume, high velocity or high variety. Big data comes from sensors, devices, video/audio, networks, log files, transactional applications, the web and social media. Much of this data is generated in real-time and on a very large scale." Source IBM Analytics

Examples of how Government Agencies could use big data

1. To Assist in Fraud Detection and Financial Market Analysis – Centrelink and Social Security
2. To Assist in Health-Related Research – Food and Drug Administration for foodborne illnesses
3. To Assist in Government Oversight and Education – tracking legislative changes or school results
4. To Assist in Fighting Crime – Home Affairs and Homeland Security
5. To Assist in Environmental Protection and Energy Exploration.

See: http://www.businessofgovernment.org/BigData3Blog.html for more details

Big data will include data sets that are too large to view. Information about the data is sought through graphs or data summarisation. Working with big data will also require analytics to be done to test the quality of the data for things like missing values, outliers etc.

With big data, visualisation techniques like different types of graphs are used to understand the data. With such large data sets the size of the data set prohibits being able to draw conclusions from within the data.

The analysis of big data sets usually requires some programming skills which is beyond the scope of this course. The skills are required to cleanse the data and to run hypothesis testing to identify relationships between different variables. Skills are also required to convert information that is not quantified data into some numbers. For data sets of over 1 million rows, this work cannot be done in traditional software like Microsoft Excel and requires more specialist software like Microsoft R, and Hadoop (from SAS).

Another software package called Microsoft Power BI, is a data visualisation package that enables summary tables like those from pivot tables in Microsoft Excel and it enables quick graphics to be produced to understand relationships within the data.

With big datasets, Data visualisation is part of the analytical process to establish a view into the underlying data, its quality and to test possible relationships within it.

Other terms

Predictive analytics is the use of data, statistical algorithms and machine learning techniques to identify the likelihood of future outcomes based on historical data.

The core of predictive analytics relies on capturing relationships between explanatory variables and the predicted variables from past occurrences and exploiting them to predict the unknown outcome.

It uses historical data to find patterns that can be extrapolated into the future. This has been done for years in financial markets, sports and gaming. Predictive analytics aims to use correlations and causations to predict future directions of data.

Data cleansing or **data cleaning** is the process of detecting and correcting (or removing) corrupt or inaccurate records from a record set, table, or database and refers to identifying incomplete, incorrect, inaccurate or irrelevant parts of the data and then replacing, modifying, or deleting the dirty or coarse data

Missing values: In statistics, **missing data**, or **missing values**, occur when no data value is stored for the variable in an observation. Missing data is a common occurrence and can have a significant effect on the conclusions that can be drawn from the data.

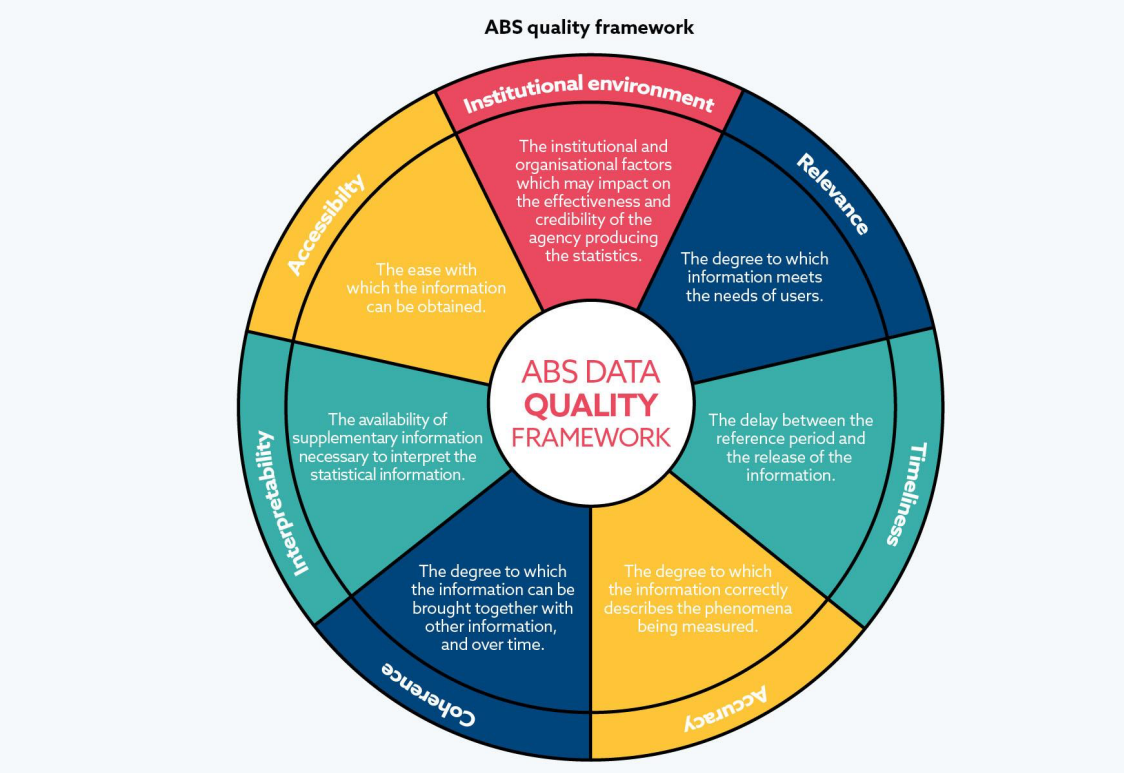
The presence of missing values means we may have to exclude the data or in some cases impute values.

Appendix 2: ABS data quality framework

The Australian Bureau of Statistics provides information on a data quality framework on its website. Details of the framework can be found on the ABS website by referencing:

1520.0 - ABS Data Quality Framework, May 2009

A graphic overview of the framework is below.



The framework refers to seven elements for assessing data quality.

1. Institutional Environment - Who produced or collected the data? Under what authority or legislation were the data collected? To what extent are quality guidelines documented by the agency? Agency capability to meet production or collection of data
2. Relevance - Details on exactly what data was collected (e.g. Key data items and definitions, Geographies, Populations from which the data came, Relevant time periods)
3. Timeliness – how old is the data and what is the reference period (the gap between the date of reference and use), when data becomes available (publication dates) and the frequency of data collection
4. Accuracy – considers aspects like: Sampling error, non-response error, coverage error and any revisions that have been applied to the data
5. Coherence – ability to link or compare with other data and has regard to whether the data has been confronted with other data sources and are the messages consistent from all data sources.
6. Interpretability - is there supporting information (e.g. concepts, sources and methods) to enable interpretation of the data
7. Accessibility – ease or difficulty of access and use (e.g. what format does the data come in or is it confined to particular types of software?).

Appendix 3: Establishing data sets

Everything is being converted to data

- Our names
- Our addresses
- Our movements are tracked by GPS
- The type of car we drive
- Our age
- Our marital status
- The medications we take
- Our taxation history.

Data can take many forms. I most commonly work with financial data or with transactional data.

The first starting point is to ensure our data is arranged into different fields that contain categories or values. In Excel, datasets will be presented in fields, each represented as a column.

The data I work with relies on being able to establish the following types of fields:

- Categories – e.g. teams, products/services, accounts, regions
- Time – ensuring different time periods are separated
- Quantified values – dollar values, transaction volumes, number of people
- Weights – values assigned to recognise the complexity, difficulty or time associated with a category.
- Weighted values – where original values are converted into weighted values.

This often requires that modifications be done to the original data to get it into a form possible to undertake and present analysis.

This is best demonstrated by way of example.

Working with vehicle registration data

The largest data set I have ever worked with is the register of vehicles in South Australia.

Although large (about 243,000 lines) I could work with the data in Excel. That means I could see the data and scroll through it, albeit glancing only.

The data set included:

1. The make of vehicles
2. The model of the vehicle
3. The year of manufacture
4. The number of cylinders
5. The number of vehicles in these categories

The data included no names or addresses. The only locational data was the fact the register was of vehicles in South Australia.

The data was provided in a table format as per below:

Source data	A	B	C	D	E
1	Make	Model	Year of Manufacture	Number of Cylinders	Number of vehicles
2	Toyota	Camry	1997	6	450
3	Honda	Jazz	2002	4	321
4	Ford	Mondeo	2004	6	150
5	Holden	Jackeroo	2000	6	240

A data set like might be used to explore the impact of different options for registering or treating vehicles. Options could include:

- Register by age
- Register by mass
- Register by length/size.

However, the original data does not come with information about the mass or length of vehicles. This is quite common in working with data – recategorising the data. In this case, cars would have to be recategorised by mass.

To recategorise data, I create data maps. Data maps will seek to assign existing categories to a new category. For example, a data map for vehicle models to a new category is:

MAP	A	B
1	**Current category**	**New Category – size**
2	Camry	Medium light vehicle
3	Jackeroo	Large light vehicle
4	Mondeo	Medium light vehicle
5	Jazz	Small light vehicle

In Excel, we can use a function called VLOOKUP to assign a new value to an existing category.

To do this, we can save a new version of our data set and add another column for "Size". I usually insert the new category before the values or numbers columns. We can then use the VLOOKUP function to assign the sizes in our map to the vehicles in our dataset.

For the Toyota Camry, the formula would be:

=VLOOKUP(Source Data B2,MAP $A$1:$B$5,2,=FALSE)

This would result in the following table:

Make	Model	Year of Manufacture	Number of Cylinders	Size (using the VLOOKUP formula)	Number of vehicles
Toyota	Camry	1997	6	Medium light vehicle	450
Honda	Jazz	2002	4	Small light vehicle	321
Ford	Mondeo	2004	6	Medium light vehicle	150
Holden	Jackeroo	2000	6	Large light vehicle	240

This provides a new way of categorising and therefore, analysing the vehicles.

The process can be repeated for other new categories.

Once a good dataset is established, Excel tools like pivot tables and a range of formulas can be used to summarise the data.

A ***pivot table*** is a table that summarizes data in another table and is made by applying an operation such as sorting, averaging, or summing to data in the first table, typically including grouping of the data. Source: Wikipedia

A pivot table doesn't change the original spreadsheet or database itself.

For motor vehicles, a pivot table could generate a summarised table like the one below

	Number of cylinders				
Size	**4**	**6**	**8**	**10**	**12**
Small	#	#	#	#	#
Medium	#	#	#	#	#
Large	#	#	#	#	#

Appendix 4: Insights about analysis

The analytical process

For this purpose, I will use the analogy of mining for a diamond as our goal to outline the analytical process.

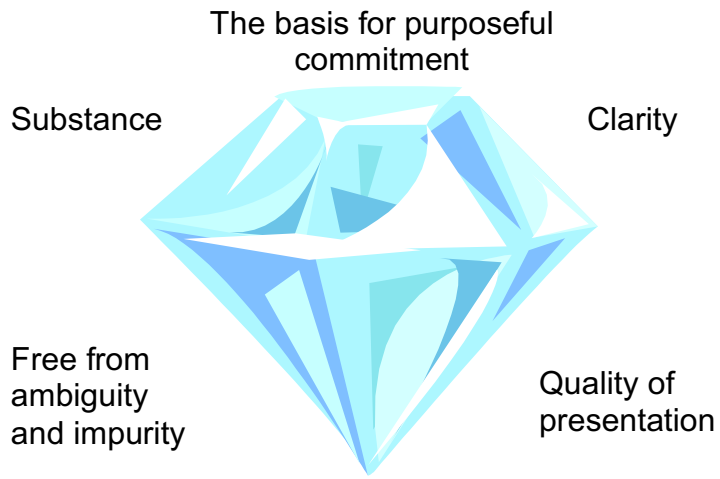
Analysis is a key step to revealing insight. Using this analogy, insight is our diamond. It is the value that we are seeking, to help others make a decision or to make a change. A diamond is also a symbol of engagement. Engagement is a critical element in getting decisions made with our analysis.

Engagement – when we want others to decide

A diamond is a symbol of engagement. A diamond also represents high quality.

When we want the engagement of our customers, our managers and our colleagues, we need to come to them with a diamond. Anything less will not cut it.

Many leaders seek the engagement of their staff and the engagement of their customers. Engagement is connecting to achieve commitment. In the workplace, it is to get a commitment to action and to change. The process of engagement is a process of seeking a decision from others to move forward with us.



When we are seeking to engage, we want the highest possible quality decisions from our work. We want:

- True value based on substance
- Value that is obvious to others
- Clarity and elimination of ambiguity.
- Incisive action and commitment

These are all the attributes of a high-quality diamond: High value, high clarity, elimination of ambiguity and a true substance that will cut through anything.

The analytical process is therefore akin to the exploration and digging processes used in mining for a diamond. In mining information, we are looking to pull it apart and find out what's inside. We are looking to identify areas that are of important value to us in revealing information that may sway decision-makers. We are looking to reveal information that may help us decide how to change the course of our project or team, to keep it viable, within budget and achieve our goals.

Following are the key steps involved in undertaking data analysis.

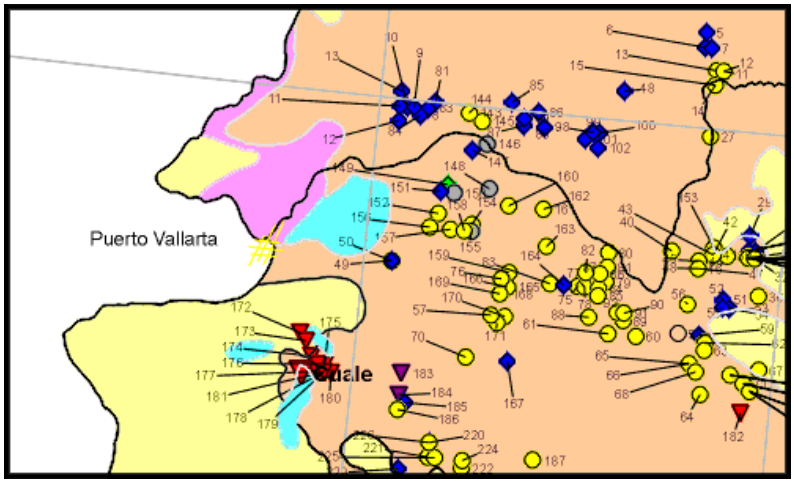
1. Be clear on the purpose of analysis – which decision(s) is it informing
2. Pre-drilling – initial exploration of data
3. Drilling into detail
4. Identify key findings
5. Start preparing for presentation – remove ambiguities
6. Polish to enhance clarity

Purpose - focus on the decision

The first step is to understand the purpose of the analysis and the key decisions that you are looking to inform or influence as a result of doing this analysis. It is important that we have clarity at this point, for our analysis to be as focused as possible.

Pre-drilling – explore overall data and identify priority areas

The second step is to source our basic data starting from the top of the relevant reporting unit and working down. It is important with data analysis that we ensure that the story we are telling is compiled in the context of the whole organisation within which we are working. Rather than being selective, and in some cases taking a convenient snapshot, it is important that the data is seen as part of a holistic picture, if possible, to ensure that we cannot be accused of being selective or missing key elements because we are too focused on our slice of data. Initial exploration of our basic data is like exploring to identify the plot or claim that we want to set aside for further drilling and ultimately mining.



This step involves working through some basic analytical techniques to establish priority areas of focus. Initially, I will use compositional analysis to identify major elements and then undertake temporal analysis to understand significant growth or changes over time. These will start to identify where we think we need to drill further. As with mining, we are keen to ensure that our drilling efforts are as focused as possible because it can be, like mining, time-consuming and expensive.

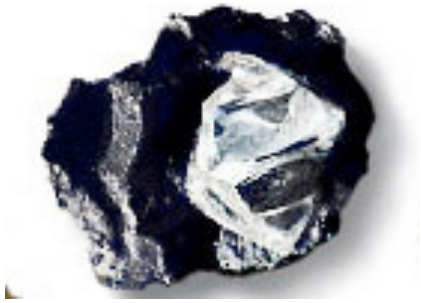
Start drilling (cautiously at first), some construction may be required

Having established some key areas for drilling, it is important to assess the quality of data, into which we are drilling. In some cases, data will be well formulated and consistent across time, which will enable us to quickly move to using more sophisticated analytical techniques. However, there will be situations where the quality of data is poor and effort will need to go into constructing a solid data set that is comparable over time, or across units. Using the mining analogy, there may be cases where we need to reinforce areas of ground before entering them, to make sure that mining is safe. In these cases, we need to construct some mechanisms to support our analysis.

This will be particularly the case where we are comparing specific transactional information over time, where we are developing ratios, and comparisons over time or where we are constructing benchmarks to compare different entities. In these cases some basic methods of reinforcement to the data need to occur, to ensure it is robust.

Stop to identify and assess findings

By drilling into details, we can pick apart detailed information to draw some conclusions. It is at this point in the process there we are seeking to spot diamonds. It is at this stage of the process that we often need to slow down to be more careful about what we are observing.



Diamonds will not just jump out at us. We need to scratch a little, we need to manoeuvre the data in different ways, compare it in different ways and try different forms of analysis before the insight reveals itself. As with the mining process, often our diamonds are buried amongst the rubble and are not always that easy to find. This demands some reflection and careful scanning to reveal our findings.

Start preparing for presentation – remove ambiguities

When we start to draw our conclusions, our data will often be buried amongst rubble. At this point, it is not fit for presentation. Once we have revealed key causes from within our data, we need to move to present our information to others. It is often the case that we can use many different instances of data to come up with some very simple and basic conclusions. In these cases, it is important that we leave a trail behind on how we come up with the data we use so that we can present to others, our methodology.

Having gotten our hands dirty to reveal something in the data, the next step is to start to work on the presentation of our information. We do not want to present to our audience a pile of rubble and let them know that amongst it is a diamond. Rather, we would prefer to present to them the diamond in all its glory and give them some background on how we arrived upon it to demonstrate substance and credibility.

Final polish to enhance clarity

A key quality indicator of our data analysis is transparency. Transparency and clarity are key elements of the quality of the diamond and are also key elements in a quality decision that others make with our data analysis and information.

In this way, we see a flipside to analysis. Analysis is the deconstruction of information into its constituent elements. In contrast, the presentation of our analysis requires the reconstruction of data, based on key insights. Therefore an important aspect of analysing data is the ability to present the findings of our analysis to others in a polished and clear way.

To do anything less will compromise the extent of engagement and buy-in we get from the decision-makers receiving our analysis.

Conclusion

Analysis is the process of breaking something into its constituent elements; a detailed examination of the elements or structure of something - typically as a basis for discussion or interpretation.

We are trying to get to the bottom of an issue within an organisation to understand the key drivers behind it. By starting with basic analysis and then drilling we are seeking to understand the cause and effect behind data and changes in the data.

We can go through a lot of information to reach a small number of conclusions. In fact, the more thorough our analysis, we can have a situation where there is an inverse relationship between the amount of information we analyse and the number of conclusions we reach or insights we find. This is a key indicator for me of the quality of analysis.

NOTES

Appendix 5: Financial Analysis

Introducing Financial statements

Financial information for organisations starts with three basic statements:

1. The operating statement (Statement of Financial performance)
2. The balance sheet (Statement of Financial position)
3. Cash flow statement

Analysis of the finances of any organisation or business unit will start with an analysis of these financial statements. This demands that we have some understanding of what each statement is presenting to us.

The **operating statement** presents flows of revenues and expenses to show the profitability or cost of service associated with an organisation

The **balance sheet** presents the financial position of a business comparing the value of its assets with the value of its liabilities to determine the net worth or value of the business. It is a measure of stocks at a point in time - assets and liabilities.

The **cash flow** statement presents the inflows and outflows of cash for a business demonstrating how the amount of cash at the end of the reporting period has been arrived at given the amount of cash at the beginning of the year.

We will start by examining each of these financial statements for the Victorian General Government sector

Financial performance - The operating statement or income statement

While the balance sheet states a company's assets and liabilities, the statement of financial performance, also referred to as a profit and loss statement (P&L), income statement (IS) or operating statement, shows a company's revenues and expenses for a given period of time. The income statement divulges a company's profitability, which reflects the company's performance and how much income can be reinvested into the company or passed onto investors in the form of dividends.

Profit = revenue (or income) less expenses

Profit is frequently used as a measure of performance or as the basis for other measures, such as return on investment or earnings per share. The elements directly related to the measurement of profit are income and expenses.

The elements of income and expenses are defined as follows.

Income is increases in economic benefits during the accounting period in the form of inflows or enhancements of assets or decreases of liabilities that result in increases in equity, other than those relating to contributions from equity participants.

Expenses are decreases in economic benefits during the accounting period in the form of outflows or depletions of assets or incurrence of liabilities that result in decreases in equity, other than those relating to distributions to equity participants.

SOURCE: **http://www.aasb.com.au/**

A copy of the operating statement for the Victorian General Government sector for the 2015-16 result is shown below.

Table 1 Victoria State General Government Operating Statement

	2015-16 $m
GFS Revenue	
Taxation revenue	20,027
Current grants and subsidies	24,737
Sales of goods and services	6,423
Interest income	786
Other revenue	4,869
Total GFS revenue	**56,842**
less	
GFS Expenses	
Gross operating expenses	
Depreciation	2,504
Employee expenses	22,125
Other operating expenses	17,443
Total gross operating expenses	***42,072***
Nominal superannuation interest expenses	878
Other interest expenses	2,076
Current transfers	
Grant expenses	5,095
Subsidy expenses	2,545
Other current transfers	36
Capital transfers	
Grants to local governments	0
Other capital transfers	1,215
Total GFS expenses	**53,916**
equals	
GFS Net operating balance	**2,924**
less	
Net acquisition of non-financial assets	
Gross fixed capital formation	5,186
less Depreciation	2,504
plus Change in inventories	43
plus Other transactions in non-financial assets	115
Total net acquisition of non-financial assets	**2,841**
equals	
GFS NET LENDING(+)/BORROWING(-)	**84**

Source: ABS catalogue 5512.0

NOTES

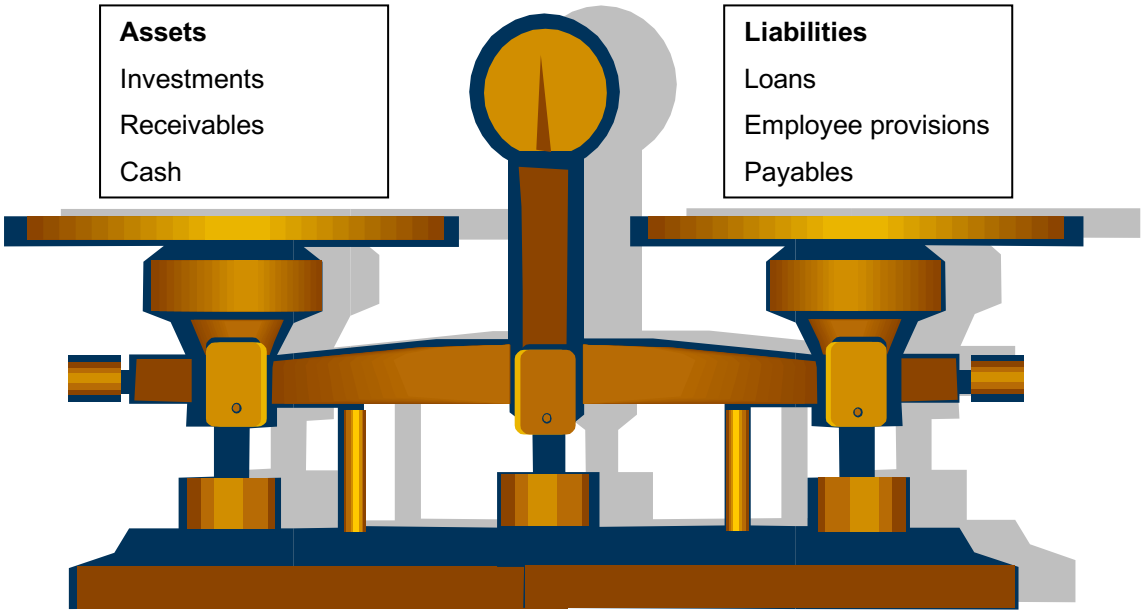
Financial Position – The Balance Sheet

A financial statement that summarizes a company's assets, liabilities and shareholders' equity at a specific point in time. These three balance sheet segments give investors an idea as to what the company owns and owes, as well as the amount invested by the shareholders

The balance sheet must follow the following formula:

Assets = Liabilities + Shareholders' Equity

This is known as the accounting equation



Each of the three segments of the balance sheet will have many accounts within it that document the value of each. Accounts such as cash, inventory and property are on the asset side of the balance sheet, while on the liability side, there are accounts such as accounts payable or long-term debt. The exact accounts on a balance sheet will differ by company and by industry, as there is no one set template that accurately accommodates the differences between different types of businesses.

It's called a balance sheet because the two sides balance out. This makes sense: a company has to pay for all the things it has (assets) by either borrowing money (liabilities) or getting it from shareholders (shareholders' equity).

The balance sheet is one of the most important pieces of financial information issued by a company. It is a snapshot of what a company owns and owes at that point in time. The income statement, on the other hand, shows how much revenue and profit a company has generated over a certain period. Neither statement is better than the other - rather, the financial statements are built to be used together to present a complete picture of a company's finances.

Assets

An asset is a resource controlled by the entity as a result of past events and from which future economic benefits are expected to flow to the entity.

Table 3 Victoria State General Government Balance Sheet

	2015-16 $m
Assets	
Financial assets	
Cash and deposits	4,772
Advances paid	4,524
Investments, loans and placements	2,789
Other non-equity assets	6,021
Equity	91,823
Total financial assets	**109,928**
Non-financial assets	
Land and fixed assets	115,364
Other non-financial assets	295
Total non-financial assets	**115,660**
Total assets	**225,588**
less	
Liabilities	
Deposits held	582
Advances received	0
Borrowing	33,808
Unfunded superannuation liability & other employee entitlements	35,428
Other provisions	764
Other non-equity liabilities	5,759
Total liabilities	**76,342**
less	
GFS NET WORTH	**149,246**
Net debt	22,306
Net financial worth	33,587

Source: ABS catalogue 5512.0

Liabilities

A liability is a present obligation of the entity arising from past events, the settlement of which is expected to result in an outflow from the entity of resources embodying economic benefits.

Equity (or Capital)

Equity is the residual interest in the assets of the entity after deducting all its liabilities.

Double Entry bookkeeping – maintaining the balance

All transactions will be such that the accounting equation remains in balance. Everything that is purchased will be paid for and everything that is sold will result in additional cash. Each transaction has a minimum of two components.

Cash flow statement

This document provides aggregate data regarding all cash inflows a company receives from both its ongoing operations and external investment sources, as well as all cash outflows that pay for business activities and investments during a given quarter.

Because companies tend to use accrual accounting, the income statements they release may not necessarily reflect changes in their cash positions. For example, if a company lands a major contract, this contract would be recognized as revenue (and therefore income), but the company may not yet actually receive the cash from the contract until a later date. While the company may be earning a profit in the eyes of accountants (and paying income taxes on it), the company may, during the quarter, end up with less cash than when it started the quarter. Even profitable companies can fail to adequately manage their cash flow, which is why the cash flow statement is important: it helps investors see if a company is having trouble with cash.

A copy of the cash flow statement for the General Government budget sector for the 2015-16 estimated result is shown below.

Table 2 Victoria State General Government Cash Flow Statement

	2015-16 $m
Net cash flows from operating activities	
Cash receipts from operating activities	
Taxes received	19,775
Receipts from sales of goods and services	7,254
Grants and subsidies received	25,271
Other receipts	4,213
Total cash receipts from operating activities	**56,512**
Cash payments for operating activities	
Payments for goods and services	-39,873
Grants and subsidies paid	-8,481
Interest paid	-2,040
Other payments	-698
Total cash payments for operating activities	**-51,093**
Total net cash flows from operating activities	**5,419**
Net cash flows from investments in non-financial assets	
Sales of non-financial assets	190
Purchases of new non-financial assets	-4,381
Total net cash flows from investments in non-financial assets	**-4,191**
Total net cash flows from investments in financial assets for policy purposes	**109**
Total net cash flows from investments in financial assets for liquidity purposes	**539**
Net cash flows from financing activities	
Advances received (net)	0
Borrowing (net)	-1,335
Deposits received (net)	65
Other financing (net)	-116
Total net cash flows from financing activities	**-1,386**
NET INCREASE(+)/DECREASE(-) IN CASH HELD	**490**

Source: ABS catalogue 5512.0

Supporting financial information

These core financial statements are often supported by more detailed information.

There is limit as to how much information you can get from financial statements. Understanding the finances of an organisation demands that we drill into more detail regarding the main revenue and expense items to understand the key drivers of this business. Further analysis demands that we have access to a wider set of data. Typically financial statements are supported by sets of notes. In the case of general government finances, the financial statements included in the budget papers are supported by other chapters that contain more details on each of the elements of the financial statements. For example, the budget statement includes the following chapters:

Chapter 1: Fiscal Strategy and Budget Priorities

Chapter 2: Expenditure

Chapter 3: Revenue

Chapter 4: Intergovernmental Finances

Chapter 5: Managing the state's assets and liabilities

Each of these provides more detail on these elements of the financial statements.

In the case of reports within an organisation to support team leaders, cost centre managers and project managers, transaction reports are usually available to support analysis.

Financial analysis

Compositional Analysis

This type of financial analysis is required to understand the makeup or structure of the finances of a business or organization. Compositional analysis is used when you want to understand or get to know a business. You may need to do this analysis if you are asked to define a financial strategy or to assess financial risks. This analysis is a basic starting point for other forms of financial analysis. This provides us with a basic understanding of the organisation we are dealing with.

Following are the operating statement and the balance sheet for the Victorian Government. The tables are like those shown previously, however this time, we have some analysis of the percentage composition of each revenue and expense item. Further, following these statements, we have some additional tables sourced from within the budget papers that give us more detail on the composition of the budget.

Table 1 Victoria State General Government Operating Statement

	2015-16 $m	% share of revenue or expenses
GFS Revenue		
Taxation revenue	20,027	35.2%
Current grants and subsidies	24,737	43.5%
Sales of goods and services	6,423	11.3%
Interest income	786	1.4%
Other revenue	4,869	8.6%
Total GFS revenue	**56,842**	**100.0%**
less		
GFS Expenses		
Gross operating expenses		
Depreciation	2,504	4.6%
Employee expenses	22,125	41.0%
Other operating expenses	17,443	32.4%
Total gross operating expenses	***42,072***	***78.0%***
Nominal superannuation interest expenses	878	1.6%
Other interest expenses	2,076	3.9%
Current transfers		
Grant expenses	5,095	9.4%
Subsidy expenses	2,545	4.7%
Other current transfers	36	0.1%
Capital transfers		
Grants to local governments	0	0.0%
Other capital transfers	1,215	2.3%
Total GFS expenses	**53,916**	**100.0%**

Examples of some further information to support the financial statements include a breakdown of taxes and a breakdown of grants and details on how expenses are split across the various functions of government.

Table: Taxation Revenue ($million)

	2017-18 Budget	% of total taxes
Taxes on employers' payroll and labour force	5,898	27.0%
Taxes on immovable property		
Land taxes	2,366	10.8%
Fire Services Property tax	674	3.1%
Congestion Levy	118	0.5%
Metropolitan improvement levy	162	0.7%
Financial and capital transactions		
Land transfer duty	6,164	28.2%
Metropolitan planning levy	27	0.1%
Financial accommodation levy	154	0.7%
Growth areas infrastructure contributions	175	0.8%
Total taxes on property	9,840	45.1%
Taxes on the provision of goods and services		
Gambling taxes		
Public lotteries	438	2.0%
Electronic Gaming Machines	1,126	5.2%
Casino	236	1.1%
Racing	74	0.3%
Other	29	0.1%
Levies on statutory corporations	112	0.5%
Taxes on insurance	1,289	5.9%
Total taxes on the provision of goods and services	3,304	15.1%
Taxes on the use of goods and performance of activities		
Vehicle Registation fees	1,594	7.3%
Duty on Vehicle registrations and transfers	925	4.2%
Liquor licence fees	23	0.1%
Other	243	1.1%
Total taxes on the use of goods and performance of activities	2,785	12.8%
Total taxation revenue	21,827	100.0%

Source: 2016-17 Final Budget Outcome

As well as getting a more detailed breakdown, another perspective on revenues and expenses can help us understand the data. For example, the budget papers provide a breakdown of General government expenses by function. This table is shown below. From this, we can understand the composition of government spending by function.

Table 4 Victoria State General Government Expenses by Purpose

	2015-16 $m	share of expenses %
General public services	1,362	2.5%
Defence		
Public order and safety	5,999	11.1%
Education		
Primary and secondary	9,905	18.4%
University	103	0.2%
Technical and further education	2,110	3.9%
Other education	1,377	2.6%
Total education	*13,494*	*25.0%*
Health		
Acute care institutions	11,316	21.0%
Other health institutions	90	0.2%
Community health services	2,353	4.4%
Pharmaceutical	407	0.8%
Other health	999	1.9%
Total health	*15,165*	*28.1%*
Social security and welfare		
Social security	513	1.0%
Welfare services	3,993	7.4%
Total social security and welfare	*4,506*	*8.4%*
Housing and community amenities		
Housing and community development	1,584	2.9%
Water supply	121	0.2%
Sanitation and protection of the environment	481	0.9%
Other community amenities	9	0.0%
Total housing and community amenities	*2,195*	*4.1%*
Recreation and culture		
Recreational facilities and services	386	0.7%
Cultural facilities and services	409	0.8%
Total recreation and culture	*794*	*1.5%*
Fuel and energy	126	0.2%
Agriculture, forestry and fishing	377	0.7%
Transport and communications		
Road transport	2,315	4.3%
Water transport	0	0.0%
Rail transport	3,671	6.8%
Communications and other transport	112	0.2%
Total transport and communications	*6,098*	*11.3%*
Other economic affairs	804	1.5%
Nominal interest on superannuation	878	1.6%
Public debt transactions	2,076	3.9%
Other purposes	45	0.1%
Total Expenses	**53,916**	**100.0%**

Source: ABS Catalogue 5512.0

Temporal analysis

To understand current financial information, it often pays to understand historical and future trends. This gives us a sense of how the current finances compare with the past and how they have and are projected to change.

Temporal analysis involves analysing changes over time (this includes trend and growth analysis). To understand the finances we have observed to date, we need to add a few more years of data and compare changes across years.

The tables over the page show figures across years for the Victorian general government sector and the South Australian general government sector.

The first presents historical data for Victoria – showing how 2006-07 compares with 2015-16.

Analysis tools for forward estimates data include –

- calculating annual growth for different items
- calculating average growth across the forward estimates for different items
- Graphing the data to compare growth.

These techniques are presented in subsequent tables.

Table 4 Victoria State General Government Expenses by Purpose

	2006-07 $m	2015-16 $m	Growth since 2006-07	Average per annum growth
General public services	257	1,362	430.0%	20.4%
Defence				
Public order and safety	3,718	5,999	61.4%	5.5%
Education				
Primary and secondary	6,720	9,905	47.4%	4.4%
University	99	103	4.0%	0.4%
Technical and further education	1,350	2,110	56.3%	5.1%
Other education	765	1,377	80.0%	6.7%
Total education	*8,934*	*13,494*	*51.0%*	*4.7%*
Health				
Acute care institutions	6,958	11,316	62.6%	5.6%
Other health institutions	40	90	125.0%	9.4%
Community health services	1,311	2,353	79.5%	6.7%
Pharmaceutical	102	407	299.0%	16.6%
Other health	587	999	70.2%	6.1%
Total health	*8,998*	*15,165*	*68.5%*	*6.0%*
Social security and welfare				
Social security	121	513	324.0%	17.4%
Welfare services	2,533	3,993	57.6%	5.2%
Total social security and welfare	*2,655*	*4,506*	*69.7%*	*6.1%*
Housing and community amenities				
Housing and community development	1,330	1,584	19.1%	2.0%
Water supply	228	121	-46.9%	-6.8%
Sanitation and protection of the environmen	290	481	65.9%	5.8%
Other community amenities	102	9	-91.2%	-23.6%
Total housing and community amenities	*1,950*	*2,195*	*12.6%*	*1.3%*
Recreation and culture				
Recreational facilities and services	451	386	-14.4%	-1.7%
Cultural facilities and services	294	409	39.1%	3.7%
Total recreation and culture	*745*	*794*	*6.6%*	*0.7%*
Fuel and energy	85	126	48.2%	4.5%
Agriculture, forestry and fishing	424	377	-11.1%	-1.3%
Transport and communications				
Road transport	1,472	2,315	57.3%	5.2%
Water transport	14	0	-100.0%	na
Rail transport	1,929	3,671	90.3%	7.4%
Communications and other transport	167	112	-32.9%	-4.3%
Total transport and communications	*3,583*	*6,098*	*70.2%*	*6.1%*
Other economic affairs	379	804	112.1%	8.7%
Nominal interest on superannuation	419	878	109.5%	8.6%
Public debt transactions	459	2,076	352.3%	18.3%
Other purposes	918	45	-95.1%	-28.5%
Total Expenses	**33,521**	**53,916**	**60.8%**	**5.4%**

Variance analysis

To date, we have been looking at the composition of our statements and have been comparing changes over time. Now we need to make comparisons with other measures, including the budget

Variance analysis occurs when you compare a budget with actuals or a previous period with actuals to understand the differences between the two. When we are reading, financial statements for variance, the bottom line we are looking at is the bottom-line variance, and the big numbers that we will need to focus on will be the large variances.

Below are the 2016-17 actuals for the budget sector compared with the original budget.

Table A.1 South Australian General government sector operating statement ($ million)

		2016-17 Budget	2016-17 Actual	2016-17 Variance
Revenue				
Taxation revenue		4,517	4,431	-86
Grants		10,246	9,958	-288
Sales of goods and services		2,487	2,568	81
Interest income		24	25	1
Dividend and income tax equivalent income		326	645	319
Other		664	852	188
Total revenue		18,263	18,480	213
less				
Expenses				
Employee expenses		7,899	8,010	-111
Superannuation expenses				
	Superannuation interest cost	392	345	47
	Other superannuation expenses	859	767	92
Depreciation and Amortisation		963	928	35
Interest expenses		346	192	154
Other property expenses		-	-	0
Other operating expenses		4,613	4,771	-158
Grants		2,937	3,025	-88
Total expenses		18,009	18,037	-28
equals				
GFS net operating balance		254	443	185
less				
Net acquisition of non-financial assets				
Purchases of non-financial assets		4,630	4,460	170
less	Sales of non-financial assets	976	722	254
less	Depreciation	963	928	35
plus	Change in inventories	0	3	-3
plus	Other movements in non-financial assets			
equals	Total net acquistion of non-financial assets	2,690	2,814	-122
equals				
GFS net lending/borrowing		-2,436	-2,371	65

SOURCE: 2016-17 Budget papers and 2016-17 Final Budget Outcome

Variances in general government expenses by function between the 2016-17 budget and the 2016-17 actual result are shown below.

South Australia General government expenses by function ($ million)

	2016-17 Budget	2016-17 Actual	2016-17 Variance	%
General public services	408	344	64	15.7%
Public order and safety	1,752	1,807	-55	-3.1%
Education	4,364	4,505	-141	-3.2%
Health	5,618	5,623	-5	-0.1%
Social Security and Welfare	1,437	1,711	-274	-19.1%
Housing and community amenities	1,275	952	323	25.3%
Recreation and culture	421	432	-11	-2.6%
Fuel and energy	74	76	-2	-2.7%
Agriculture forestry, fishing etc	213	169	44	20.7%
Mining and mineral resources	73	72	1	1.4%
Transport and communications	1,219	1,295	-76	-6.2%
Other economic affairs	390	311	79	20.3%
Other purposes	766	739	27	3.5%
Total expenses	**18,009**	**18,037**	**-28**	**-0.2%**

Ratio analysis

Ratio analysis is a more sophisticated form of financial analysis that is most often used in assessing the financial health and viability of an organization. It is used by rating agencies in rating organisations including in rating governments.

We can use ratio analysis to try to tell us whether an organisation

1. is profitable or viable
2. has enough money to pay its bills (liquidity)
3. is using its assets efficiently
4. has a gearing problem
5. is a candidate for being bought by another company or investor?

Profitability analysis ratios

Financial ratios used to assess the profitability of a business include:

Return on assets = Net income/Average total assets
Average total assets = (beginning total assets + ending total assets)/2

Return on equity = Net income/Average stockholders' equity
Average stockholders' equity = (beginning stockholders' equity + ending stockholders' equity)/2

Profit Margin = Net income/sales

Liquidity analysis ratios

Financial ratios used to assess the liquidity of a business include:

Current ratio = Current assets/current liabilities

Quick ratio = Quick assets/current liabilities
Quick assets = current assets – inventories

Net working capital ratio = net working capital/total assets
Net working capital = current assets – current liabilities

Activity analysis ratios

Asset turnover ratio = Sales/Average total assets
Average total assets = (beginning total assets + ending total assets)/2

Accounts receivable turnover ratio = sales/Average accounts receivable
Average account receivable = (beginning accounts receivable + ending accounts receivable)/2

Inventory turnover ratio = Cost of Goods sold/Average inventories
Average inventories = (beginning inventories + ending inventories)/2

Gearing or debt ratios

Debt to equity ratio = Total liabilities/Total stockholders' equity

Interest coverage ratio = Income before interest and tax/Interest expense

Debt to revenue ratio = Net debt/Total revenue

Net financial liabilities to revenue = net financial liabilities/total revenue
Net financial liabilities = total financial liabilities – total financial assets

Other ratios used regularly by potential investors include returns on assets, price-earnings (PE) ratio, dividend yield, and dividend payout (ratio).

More information on these ratios can be found at www.investopedia.com

Who is interested in what?

The use of ratios depends on the type of organization you are (public, private, household) and your relationship with a business (are you an investor, a supplier, a customer, an owner, a lender, providing grants).

Different financial targets and ratios will be relevant to different audiences.

Shareholder

Shareholders are interested in returns on the funds they have invested. Their primary target will be the return on capital or the returns on assets. In making their investing decisions, shareholders will consider a range of indicators, but the most important among them will be return on capital and trends in total profit, revenue and debt-equity ratios. These provide indications of sustainability and risk.

CEO

While also interested in returns on equity and the range of balance sheet measures, the CEO's interest will work further down the organisation. CEO will consider indicators that try to distinguish different impacts on the business. For instance differentiating the impact of financing decisions, from asset management decisions and market-based actions.

This is where consideration of different profit measures is of importance to the CEO.

- EBIT measures profit before the impact of financing decisions and tax implications.
- EBITDA measures the impact of profit before the impact of asset management decisions.

The CEO will be keen to assess the contributions to profit from different divisions and different parts of the business.

Division/project leader

Division leaders are less interested in measures that relate to how the entire company is financed and more interested in how their part of the organisation is contributing to overall results.

In many organisations, aspects like assets, investments and liabilities are centrally managed and are out of the control of divisions.

Division leaders are therefore mainly interested in income and profit measures like EBIT or EBITDA as they relate to their division. It is important for division leaders to identify those components of these measures that they control directly (their own staff, sales levels etc.) from those components of the measures that they do not control (e.g. corporate overheads).

In businesses where revenue is centralised, the focus of division and project leaders is likely to be on cost control and expense management.

There are likely to be some measures specific to each division in an organisation that are unique to that division and therefore will be monitored specifically (e.g. sales for the sales division, taxation expenses for the accounting division, legal expenses for legal etc.)

Lenders and Suppliers

Lenders will look at the overall health of the balance sheet, with particular interest in the debt/equity ratios to assess the credit risk of a business.

Suppliers are keen to ensure that the company can trade in the short term to ensure they can meet their bills as and when they fall due. As a result, suppliers are interested in working capital, working capital turnover and liquidity ratios that are indicators of an organisation's ability to pay and creditworthiness.

Many suppliers will rely on external agencies to rate companies and their creditworthiness. To this end, organisations like Dun & Bradstreet, Moody's and Standard and Poor's provide rating services to prospective suppliers and lenders.

NOTES

Not all of these ratios are relevant for use in the public sector.

Only a subset of ratios will be relevant to government agencies. These include liquidity and possible ratios regarding asset usage.

There are circumstances where the government deals with external businesses and needs to understand the health of their finances as part of doing business. This will be the case in tenders for major contracts, in the provision of grants to organisations to run programs on behalf of the government or for situations where the government may lend money to assist firms or industries to grow. In these cases, the government agency will need to undertake some financial analysis to be assured of the financial health of the business it is dealing with.

Financial ratios in government are relevant for an assessment of the financial health of the government and the budget as an organization. This is done primarily by the rating agencies Standard and Poors and Moodys. These agencies are heavy users of financial information to assess financial health and for rating an organisation or a government.

Key ratios for the state government

Included in the presentation of the state budget are some key ratios that are used as indicators of the financial health of the budget.

The ratios most relevant to the state government relate to indicators of the burdens of debt and financial liabilities.

Ratios for South Australia's General Government Sector

As at 30 June	2016 Actual	2017 Estimated result	2018 Budget	2019 Estimate	2020 Estimate	2021 Estimate
Net Debt						
$ m	4,393	6,297	6,072	6,733	6,808	6,687
% of total revenue	25.3	34.5	31.7	34.9	34.8	33.1
% of GSP	4.4	6.0	5.6	6.0	5.8	5.4
Net financial liabilities						
$ m	21,372	21,850	21,483	21,948	21,882	21,546
% of total revenue	123.1	119.6	112.2	113.7	111.9	106.7
Net financial worth						
$ m	-2,049	-1,926	-1,376	-1,453	-968	-219
% of total revenue	-11.8	-10.5	-7.2	-7.5	-5.0	-1.1
Net worth						
$ m	37,741	41,523	41,943	42,714	43,640	44,824
% of total revenue	217.4	227.3	219.0	221.2	223.1	221.9

SOURCE: 2017-18 Budget Statement

Victorian Budget Sector Aggregates

Table 4.1: General government fiscal aggregates

	Unit of measure	*2015-16 actual*	*2016-17 revised*	*2017-18 budget*	*2018-19 estimate*	*2019-20 estimate*	*2020-21 estimate*
Net result from transactions	$ billion	2.7	1.3	1.2	2.0	2.4	2.7
Government infrastructure investment [(a)(b)]	$ billion	4.8	9.3	10.1	9.8	10.1	8.4
Net debt	$ billion	22.3	18.1	23.8	25.1	27.4	28.9
Net debt to GSP	per cent	6.0	4.6	5.8	5.8	6.0	6.0

Source: Department of Treasury and Finance

Notes:

(a) Includes general government net infrastructure investment and estimated cash flows for Partnerships Victoria projects and the State contribution to the West Gate Tunnel Project.

(b) Excludes the impact of the medium-term lease over the operations of the Port of Melbourne.

Benchmarking analysis

To date, we have been analysing the financial information for an entity or a unit on its own. An important aspect of financial analysis is to compare our entity or a unit with external parties using benchmark analysis.

To some extent, ratio analysis, where there are target ratios, is a form of benchmark analysis. However, we can also compare other aspects of financial information with other jurisdictions.

A common requirement of benchmarking is that you must compare like with like or apples with apples. This is often taken too far. No two entities are alike. We necessarily need to compare different entities to identify the nature and possible causes of differences.

Below is a comparison of the composition of general government revenues and expenses across three states.

Table A.1 General government sector operating statement ($ million)
2015-16

		Vic	%	WA	%	SA	%
Revenue							
Taxation revenue		20,027	35%	8,997	34%	4,426	25%
Grants		24,737	44%	7,407	28%	8,671	50%
Sales of goods and services		6,423	11%	2,155	8%	2,455	14%
Interest income		786	1%	192	1%	36	0%
Other		4,869	9%	7,646	29%	1,790	10%
Total revenue		56,842	100%	26,397	100%	17,378	100%
less							
Expenses							
Employee expenses		22,125	41%	12,910	45%	8,465	50%
Superannuation interest cost		878	2%	218	1%	402	2%
Depreciation and Amortisation		2,504	5%	1,359	5%	890	5%
Interest expenses		2,076	4%	664	2%	220	1%
Other operating expenses		17,443	32%	9,482	33%	4,435	26%
Grants		8,891	16%	3,873	14%	2,679	16%
Total expenses		53,917	100%	28,506	100%	17,091	100%
equals							
GFS net operating balance		2,925		-2,109		287	
less							
Net acquisition of non-financial assets							
Purchases of non-financial assets		5,186		2,296		1,098	
less	Sales of non-financial assets						
less	Depreciation	2,504		1,359		890	
plus	Change in inventories	43		15		-3	
plus	Other movements in non-financial assets	115		92		-15	
equals	Total net acquistion of non-financial assets	2,840		1,043		190	
equals							
GFS net lending/borrowing		85		-3,152		97	

Source: ABS catalogue number 5512.0

Comparison by function

The table below compares the expenses by function across each of the three states. These comparisons are made both on a dollar basis and on a percentage basis.

General government expenses by function (2015-16)

	Vic $ m	Vic %	WA $ m	WA %	SA $ m	SA %
Health	15,165	28.1%	8,306	29.1%	5,522	32.3%
Education	13,494	25.0%	7,017	24.6%	4,347	25.4%
Public order and safety	5,999	11.1%	3,372	11.8%	1,750	10.2%
General public services	1,362	2.5%	177	0.6%	289	1.7%
Social Security and Welfare	4,506	8.4%	2,114	7.4%	1,533	9.0%
Transport and communications	6,098	11.3%	2,091	7.3%	1,030	6.0%
Housing and community amenities	2,195	4.1%	2,043	7.2%	970	5.7%
Recreation and culture	794	1.5%	746	2.6%	393	2.3%
Other economic affairs	804	1.5%	620	2.2%	279	1.6%
Agriculture forestry, fishing etc	377	0.7%	386	1.4%	169	1.0%
Mining and mineral resources	na	na	200	0.7%	86	0.5%
Fuel and energy	126	0.2%	464	1.6%	32	0.2%
Other purposes	2,999	5.6%	968	3.4%	691	4.0%
Total expenses	**53,919**	**100%**	**28,504**	**100%**	**17,091**	**100%**

Source: ABS catalogue number 5512.0

Multivariate analysis - combining financial and non-financial data

To date, all the analysis we have done has been of financial data in isolation from any other data. Our finances do not exist in isolation from other variables. Therefore understanding our finances better will involve combining financial data with some non-financial data.

In some cases, we can combine data relating to finances with output data or caseload data to determine costs per unit of output or cost per case. This is often important in doing benchmarks as we need to consider the different sizes of states or organizations by making data more comparable.

In the case of comparing states, we can compare expense and revenue figures with the size of the economies or with the population of each state to determine key financial indicators per capita or as a % of GDP.

Important indicators for state governments will include operating expenses per capita, operating revenue per capita, capital expenditure per capita and each of these as a % of GDP.

The tables on the following pages present the finances of three states on a per capita basis.

Table A.1 General government sector operating statement ($ per capita)
2015-16 outcome

	Vic	WA	SA
Revenue			
Taxation revenue	3,284	3,527	2,593
Grants	4,057	2,904	5,081
Sales of goods and services	1,053	845	1,438
Interest income	129	75	21
Other	799	2,997	1,049
Total revenue	9,322	10,348	10,183
less			
Expenses			
Employee expenses	3,628	5,061	4,960
Superannuation interest cost	144	85	236
Depreciation and Amortisation	411	533	521
Interest expenses	340	260	129
Other operating expenses	2,861	3,717	2,599
Grants	1,458	1,518	1,570
Total expenses	8,842	11,175	10,014
equals			
GFS net operating balance	480	-827	168
less			
Net acquisition of non-financial assets			
Purchases of non-financial assets	850	900	643
less Sales of non-financial assets	0	0	0
less Depreciation	411	533	521
plus Change in inventories	7	6	0
plus Other movements in non-financial assets	19	36	0
equals Total net acquistion of non-financial assets	466	409	122
equals			
GFS net lending/borrowing	13	-1,237	46

Source: ABS catalogue number 5512.0

Table B.6: General government expenses by function (2015-16)

	Vic $ per capita	WA $ per capita	SA $ per capita
Health	2,487	3,256	3,236
Education	2,213	2,751	2,547
Public order and safety	984	1,322	1,025
General public services	223	69	169
Social Security and Welfare	739	829	898
Transport and communications	1,000	820	604
Housing and community amenities	360	801	568
Recreation and culture	130	292	230
Other economic affairs	132	243	163
Agriculture forestry, fishing etc	62	151	99
Mining and mineral resources	na	78	50
Fuel and energy	21	182	19
Other purposes	492	379	405
Total expenses	**8,843**	**11,174**	**10,014**

Source: ABS catalogue number 5512.0

Comparing Queensland with medium-sized states

Table B.6: General government expenses by function (2015-16)

	Qld $ per capita	WA $ per capita	SA $ per capita
Health	3,163	3,256	3,236
Education	2,453	2,751	2,547
General public services	311	69	169
Public order and safety	924	1,322	1,025
Social Security and Welfare	585	829	898
Housing and community amenities	302	801	568
Recreation and culture	230	292	230
Fuel and energy	118	182	19
Agriculture forestry, fishing etc	144	151	99
Mining and mineral resources	59	78	50
Transport and communications	1,102	820	604
Other economic affairs	182	243	163
Other purposes	753	379	405
Total expenses	**10,327**	**11,174**	**10,014**

Source: ABS catalogue number 5512.0

Reflection

What has been a key learning for me and why?