



MONASH University

Information Technology

# Welcome to FIT 1006

## Business Information Analysis

Semester 1, 2021

# FIT 1006

## Business Information Analysis

Teaching Team:

- Lecturer: Mary Poh Lim ( [poh.lim@monash.edu](mailto:poh.lim@monash.edu) )
- Tutors:
  - ❖ Anthony Wong ( [anthony.wong@monash.edu](mailto:anthony.wong@monash.edu) )
  - ❖ Flora Jin ( [flora.jin@monash.edu](mailto:flora.jin@monash.edu) )
  - ❖ Heshan Kumarage ( [heshan.kumarage@monash.edu](mailto:heshan.kumarage@monash.edu) )
  - ❖ Mary Poh Lim ( [poh.lim@monash.edu](mailto:poh.lim@monash.edu) )
  - ❖ Sanaz Nikfalazar ( [sanaz.nikfalazar1@monash.edu](mailto:sanaz.nikfalazar1@monash.edu) )

Please use your Monash account for all emails to lecturer and tutors

# Lectures

- Online lectures: Two 1-hour lecture per week
  - Wednesday 3pm – 4pm
  - Thursday 10am – 11am
  - Zoom link:  
<https://monash.zoom.us/j/89067355304?pwd=cWg0MlJrSVhsMUdwSys3ZDZFQkgwdz09>
- Most lectures will require pre-reading, which will be posted on Moodle
- Lecture slides will be posted after the lecture on Moodle as PDF, one slide/page with annotations.
- Lectures recordings available on Moodle under Class Streaming.
- Some topics will run across multiple lectures.

# Tutorials

- Enrol for tutorials on Allocate+ :
  - <https://my-timetable.monash.edu/odd/student>
- If you need help with your timetabling:
  - <https://connect.apps.monash.edu/students/timetables/allocate/help/>
- Tutorial starts in week 2 (2-hour session per week)
- Download the tutorial sheet each week
- Arrive prepared for each tutorial. This means: revising lectures + reading + practice questions.
- Tutorials will comprise: review of lecture, worked examples, practice questions, manual & computer calculations.
- + 8 hours study/practice/research each week!

# Consultation and Help

- Consultations:
  - Starts from Week 3. Consultation times will be posted on Moodle
- For urgent help or to make an appointment (outside the published consultation times), please email:  
[poh.lim@monash.edu.au](mailto:poh.lim@monash.edu.au)
- Forum: [Ed Discussion Forum](#)

# Resources

- Subject Resources
  - Download all lecture notes, tutorial sheets, past tests and exams from:
  - [FIT1006 Business Information Analysis S1 2021](#)
- [Lecture Recording](#) (under Moodle – Class Streaming)
- Library

Reading List:

<https://rl.talis.com/3/monash/lists/1E595FA4-F8CB-8F4D-FF74-A3B034BCFB45.html?lang=en-GB&login=1>

# Textbook and References

- Prescribed text:

*Australian Business Statistics*, Abridged 7th Ed. (or 8<sup>th</sup> edition)  
Selvanathan et al., Cengage, South Melbourne, 2017.  
(5th and 6th editions are very similar in content)

- Additional Reading

- *Statistics Without Tears*, Derek Rowntree, Penguin, Harmondsworth, 1981.
- *Statistics Explained: An introductory guide for life scientists*, Steve McKillup, Cambridge U.P., 2006.

- Wikipedia – a handy reference for most topics.

- Software:

- Excel and SYSTAT

# Calculators

- A Scientific or Graphing Calculator is ideal.
- Your calculator should be able to calculate:  
 $e^x$ ,  $\pi$  and ! as well as having a statistical mode where you can enter data and perform least squares regression ( $m$ ,  $c$  and  $r$  are the usual outputs).
- If you had a graphing/symbolic calculator in secondary school then use this. You may be able to download a package of statistical functions for probabilities etc.
- You cannot use a smartphone or laptop in exams so make sure you have a calculator you know how to use...



# Mathematics!

- This unit has a mathematics prerequisite of:
- A study score of 25 in VCE Mathematics Methods or Specialist Maths units 3 & 4 or 30 in Further Maths units 3 & 4 or MTH1010 or equivalent.
- If you have come to university via a non-VCE pathway, or are just curious, you can see some sample questions on Moodle.

# Assessment

To pass the unit you must obtain:

- At least 45% in the unit's examination,
- At least 45% in the unit's total non-examination assessment
- And an overall unit mark of 50% or more.

Unit Assessment		
Assessment	Due Date	% of Mark
Assignment	Thu 1 <sup>st</sup> April 11pm	20
Test	Wed 21 <sup>st</sup> April 4 – 5pm	20
eExam (2 hours)	TBA	60

# Brief outline

- Collecting and presenting data;
- Calculating and interpreting descriptive statistics;
- Calculations by hand using your calculator, and with EXCEL and SYSTAT;
- Modelling relationships between variables;
- Probability theory, probability distributions;
- Inferential statistics and hypothesis testing;
- Time series analysis;
- **See: FIT1006 Unit Outline 2021.pdf (on Moodle or in the next 2 slides)**

# Unit outline

Lect	Date	Theme	Topics
1	3-Mar	Introduction	Introduction to the unit. The broad context of statistical problems. The process/practice of statistics. Practical matters.
2	4-Mar	Surveys and Data Collection	Populations and samples. Reasons for sampling. Methods of choosing samples. Survey methods. Sampling errors.
		No Tutorial during Week 1	
3	10-Mar	Graphical Presentation of Data	Types of data, Tally, Frequency Table, Stem and Leaf Plot, Histogram, Visual interpretation.
4	11-Mar	Descriptive Statistics, Measures of Centre and Dispersion	Mean, Median, Trimmed Mean, Robust statistics. Variance and Standard Deviation, Quartiles, Interquartile Range, Boxplots.
		Tutorial 1	Introductions, Surveys
5	17-Mar	Introduction to statistical software ( <i>assignment discussion</i> )	Calculating descriptive statistics with EXCEL and SYSTAT. Comparing groups.
6	18-Mar	Analysing Data, Writing a Statistical Report	Visualising data, Using appropriate statistics, Describing data, The structure of a report, Figures and tables, Citing references.
		Tutorial 2	Disp Data, Desc Stats
7	24-Mar	Correlation and Regression	Introduction to correlation and regression, The general linear model, q-Correlation, Pearson's $r$ .
8	25-Mar	Correlation and Regression	Least squares regression. SYSTAT and Excel calculation. Interpreting diagnostic output, Recognising outliers and extreme values.
		Tutorial 3	Stat Software Report
9	31-Mar	Introduction to Probability	Definition of probability, Introduction to set theory, Probability distributions, Expectation.
10	1-Apr	Probability <i>Assignment due 1st Apr 11pm</i>	Independence, Conditional probability, Bayes' theorem. Background mathematics for probability distributions.
		Tutorial 4	Correlation, Regression
		Mid Semester Break (5th April - 9th April 2021)	

Lect	Date	Theme	Topics
11	14-Apr	Binomial and Poisson Distributions	Characteristics of Binomial and Poisson Distributions including mean and variance, Applications of these distributions to typical situations.
12	15-Apr	The Normal Distribution	Characteristics, The Standard Normal, Standardising variables, Mean and variance, Calculating Normal probabilities using tables and Excel.
		Tutorial 5	Probability
13	21-Apr	<i>Mid-semester test</i>	<i>4pm - 5pm</i>
14	22-Apr	Index Numbers – Applications. The Consumer Price Index	Calculating indices, Using indices in practice, The Australian Consumer Price Index (CPI), Interpreting the quarterly CPI Report. + answer some questions using the CPI (inflate/deflate/chain/compare sectors)
		Tutorial 6	Probability Distributions
15	2-May	Theoretical Sampling Distributions	Central Limit Theorem, Sampling distribution of mean and proportion, Standard error of sample mean.
16	5-May	Test review and feedback	Feedback on mid-semester test.
		Tutorial 7	Index Numbers
17	9-May	Estimation	Estimators, Confidence Intervals (C.I.), C.I. for the mean and proportion, difference of means and proportions.
18	12-May	Estimation	t-Distribution, C.I. for the difference of means and proportions. Polling.
		Tutorial 8	Sampling Distributions
19	16-May	Hypothesis Testing	The Null and Alternative Hypothesis, The test procedure, The hypothesis test for a population mean and proportion.
20	19-May	Hypothesis Testing	The hypothesis test for the difference of means, Type I and II errors, the significance and power of a test.
		Tutorial 9	Estimation
21	16-May	Time Series Analysis	Time series data, Components of a time series, Smoothing with moving averages and medians, Exponential smoothing.
22	19-May	Time Series Analysis	Seasonal Indices, Calculating multiplicative seasonal indices, Regression based forecasting, The accuracy of forecast.
		Tutorial 10	Hypothesis Testing
23	23-May	Review of course/ Exam Prep	Applied questions from hypothesis testing, probability etc.
24	26-May		Motivating questions from past lectures, Review of the course.
		Tutorial 11	Time Series Analysis

# Peer-assisted learning

We're using peer-assisted learning for you to participate more fully during lectures. This means:

- Come prepared to each lecture having done the pre-reading and other activities. (See Moodle)
- Bring your mobile or laptop or any device that will allow you to connect to the internet.
- Feel free to private chat with your friends to discuss the questions/answers

# Instructions to participate in the poll:

- Audience Response System – FLUX
- Use your smartphone, iPad or computer etc.
- Search for <https://flux.qa>
- Login via your Monash account
- Join an audience: use the + button
- Type 6 digit code: **SJ6KGV**
- Select today's lecture
- Put in your response while the Poll is open

# Flux.qa for lecture participation

You should see something like this on your device.



To participate, go to

**flux.qa/SJ6KGV**

GOT IT



<https://flux.qa> (Feed code: SJ6KGV)

## Question 1

Are you currently in Australia?

---

- A. Yes, I'm in Australia (but not in Melbourne).
- B. Yes, I'm in Melbourne, Australia.
- C. No, I'm not in Australia at the moment.

<https://flux.qa> (Feed code: SJ6KGV)

## Question 2

What degree are you studying?

---

- A. B. Information Technology (include doubles)
- B. B. Computer Science
- C. B. Software Engineering
- D. Humanities (any)
- E. Sciences (any)
- F. Something else

# About FIT1006...

## Quantitative analysis

- Some real-world issues where quantitative analysis may lead to improved outcomes:
  - Health: patterns of disease, effectiveness of treatments, the inequality of nations;
  - Civilization: the science of food production, the risk of catastrophe – the GFC, climate change, global warming, social disadvantage;
  - Commerce: the volatility of investments, the risk of borrowers defaulting on loans...;
  - Sport: player statistics and rankings.

# Quantitative analysis

- Quantitative analysis shows us the big picture.
- Using data (usually numerical) to:
  - Generalise: what has happened now and in the past (and may happen in the future), what is typical behaviour?
  - Normalise: what is normal? What is exceptional?
  - Contextualise: how do we compare with others around us? How are things changing over time?

# Quantitative analysis

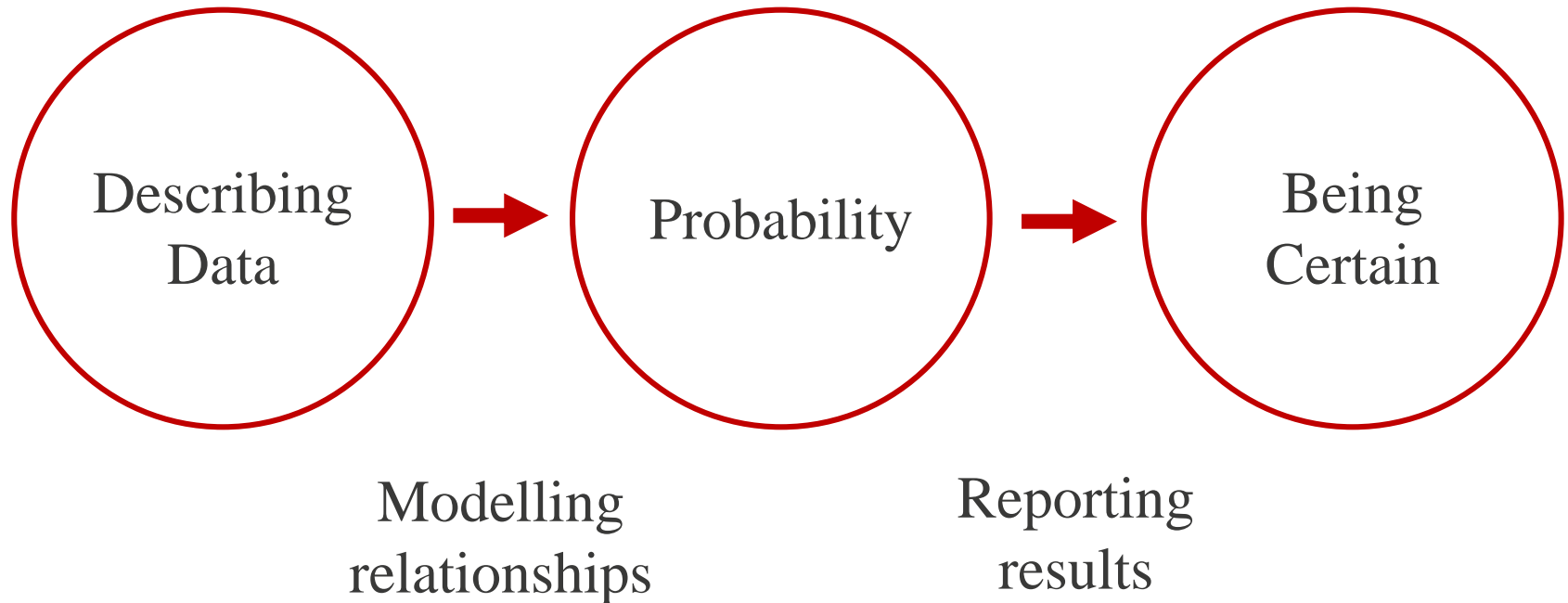
- Quantitative analysis for business:
  - Customer behaviour, randomness, risk;
  - Demand patterns, trends – change over time;
  - Determining public sentiment;
  - Cause and effect relationships;
  - Identifying the systematic from the *ad hoc*...
  - A good example of a company using quantitative analysis to good effect is Google. Their page rank algorithm changed the world. SEO was an industry unknown 15 years ago...

# Overview

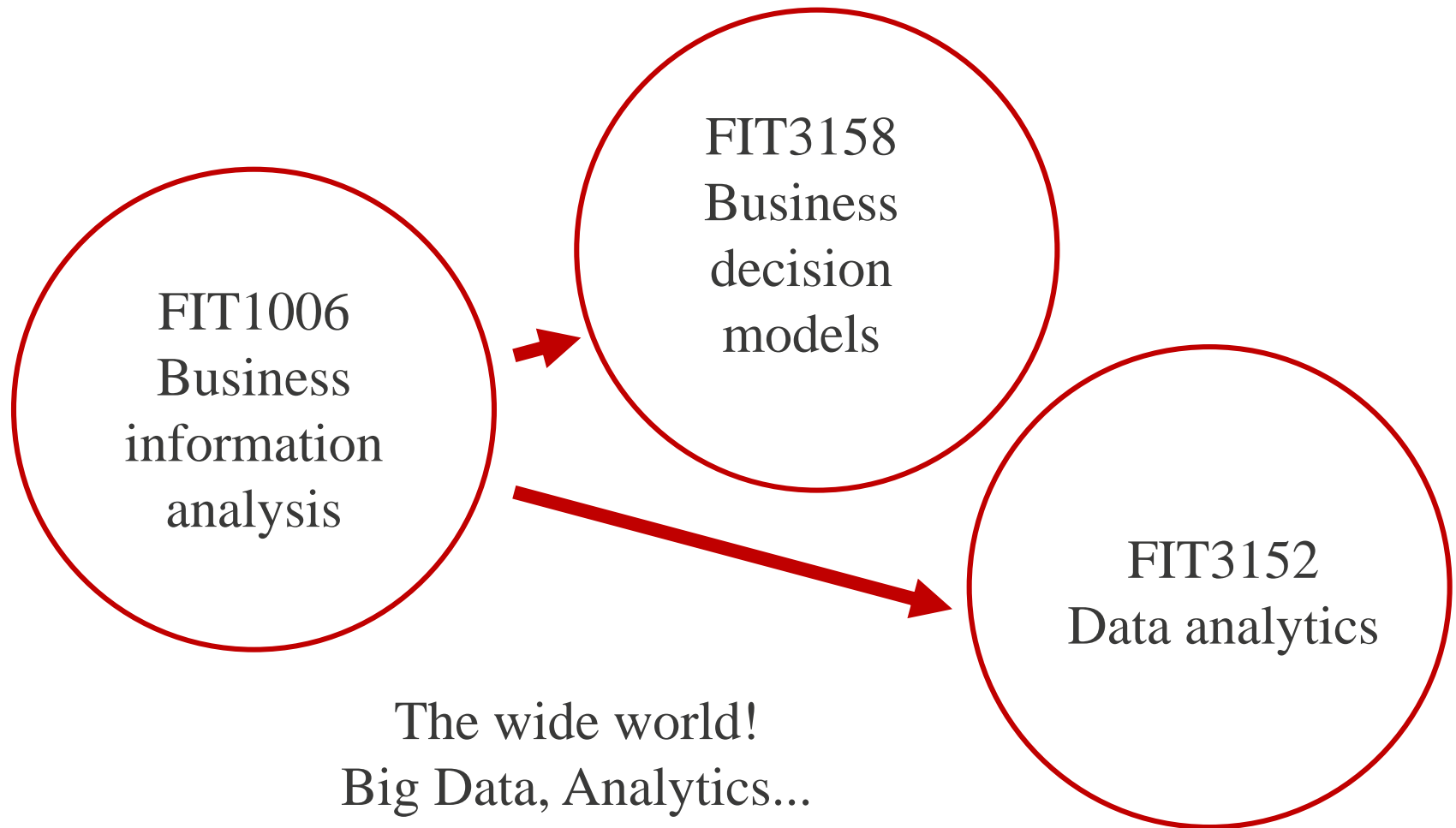
- Quantitative techniques include:
  - Descriptive techniques applicable to a wide range of data – sample, financial, Internet, etc.;
  - Conducting surveys and analysing the results;
  - Modelling relationships and trends;
  - Forecasting based on historical data;
  - Working with large(ish) data sets;
  - Presenting results as summaries and reports;
  - Testing hypotheses.

# FIT1006: The big picture

Application of theory  
to practical problems



# FIT1006: The bigger picture



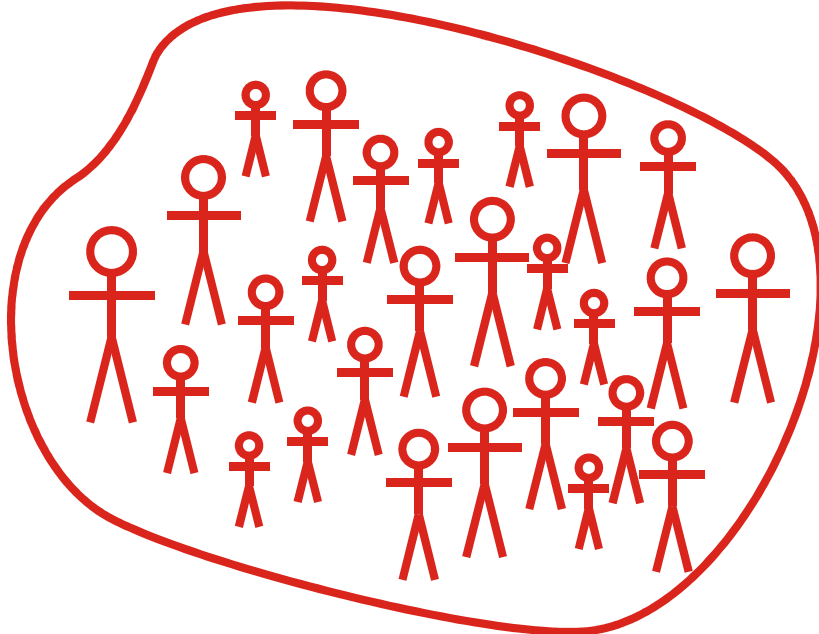


# Descriptive statistics

- A statistic is a summary of data. For example, the *number* of students enrolled in this subject is a statistic. Statistics are a way of summarizing the essential features of a large quantity of data.
- *Statistics* gets its name from the collection of data about the *State*.
- *Data* is a plural noun. *Datum* is the singular form.
- Descriptive statistics describe the features of a data set; for example, the *mean and standard deviation*.

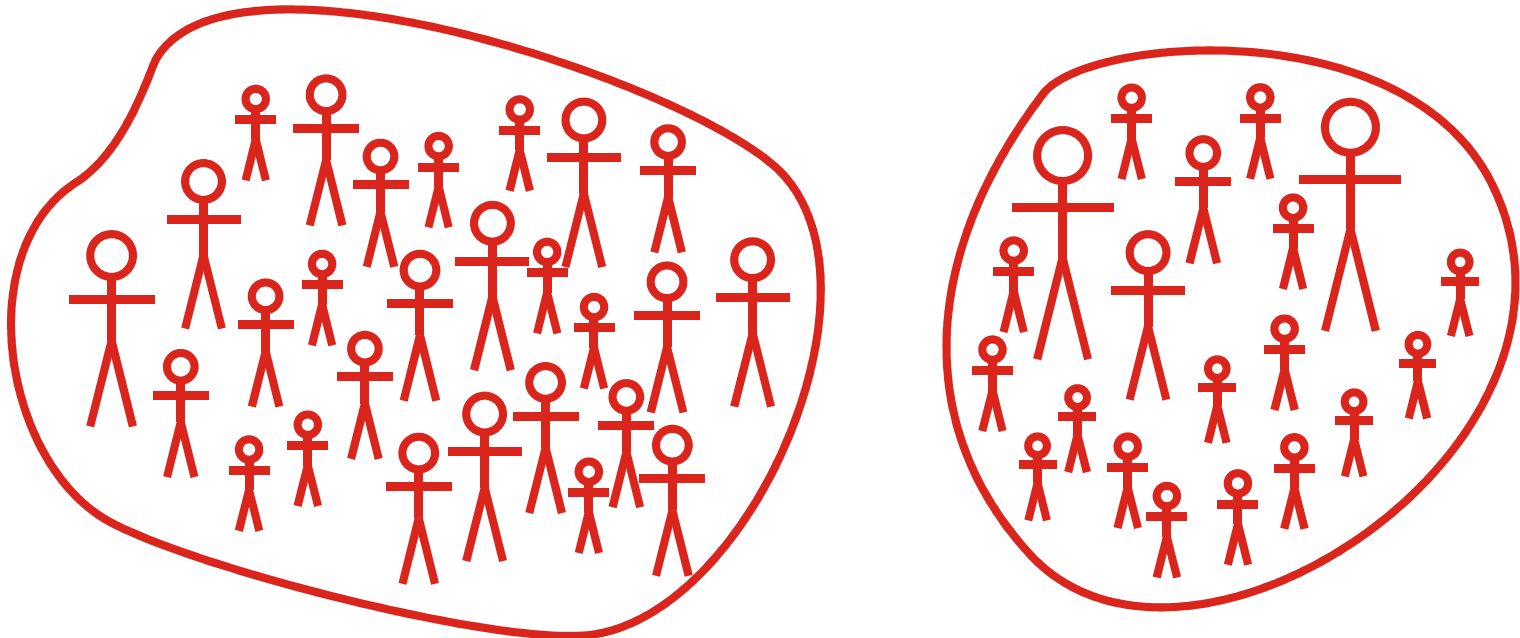
# Describing data (a)

- How was the data sample obtained; could this introduce doubt into our conclusions?
- Representing the data graphically; calculating summary statistics: average height for example.



## Describing data (b)

- What are the major differences between the two groups as suggested by the summary statistics?
- Can we represent this graphically?

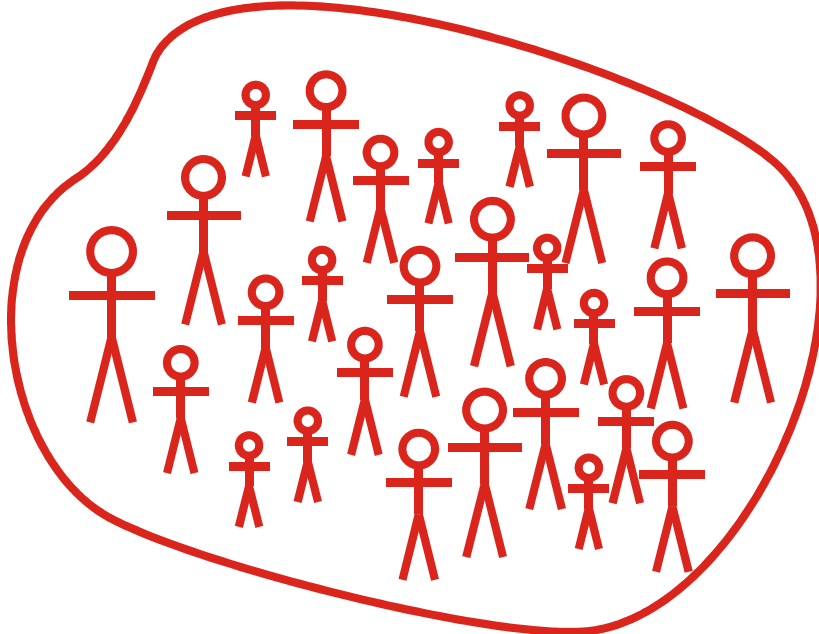


# Statistical reports

- Having responsibly obtained our data and accurately calculated statistics, the next most important skill is the *interpretation* and *communication* of the results.
- Reports should discuss the *significance* and *consequences* of findings as well as any *assumptions* made in analysis.
- Statistical summaries should be presented in an *easy to read* form - such as a table.
- Your report should be understandable to a person without a detailed statistical knowledge.

# Probability

- What is the chance that a certain person selected at random will exhibit a certain property?
- Does that chance change if the person belongs to a subgroup of the population?

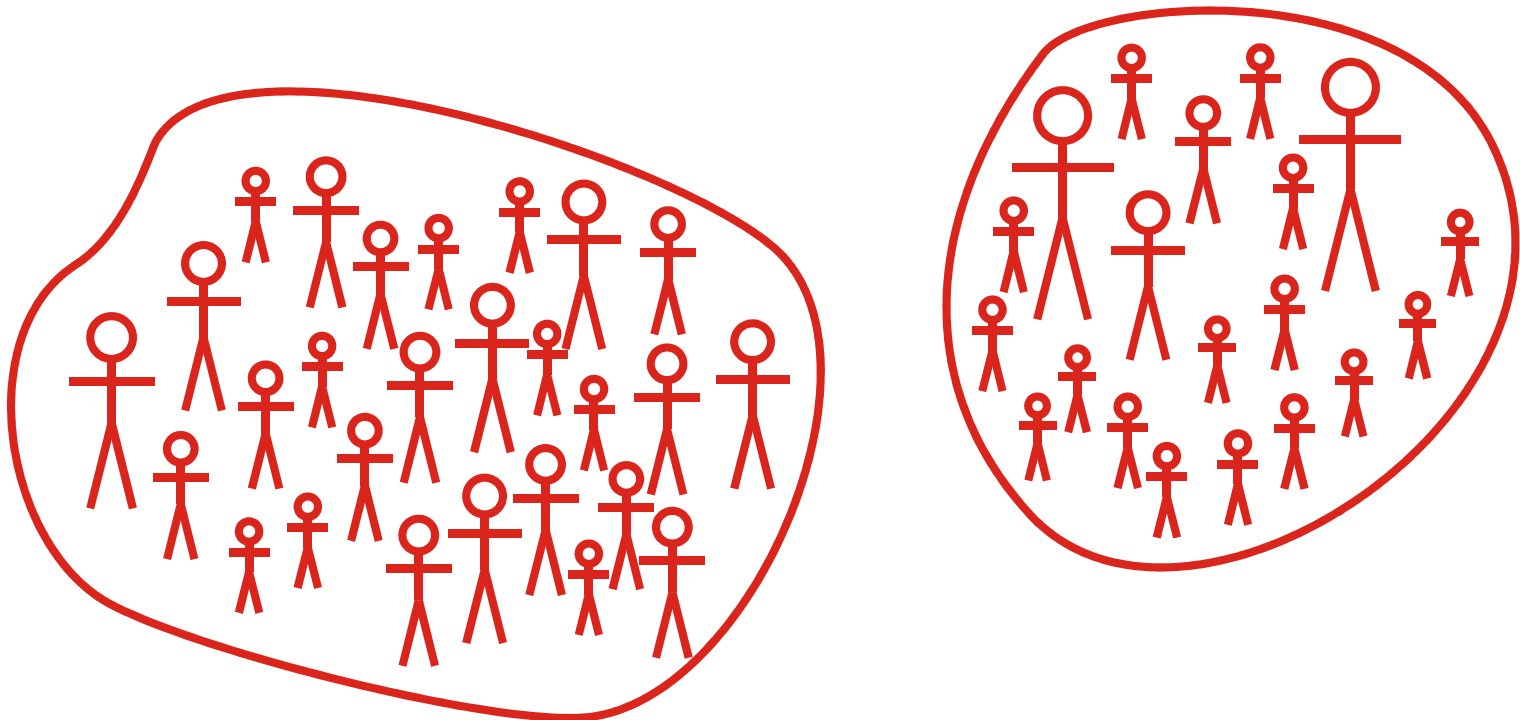


# Inferential statistics

- Having observed some differences between groups of data we may wish to know whether these differences are due to some *systematic* cause or just due to *random* fluctuations.
- This is an important part of product testing, medical treatments etc.
- We generally use descriptive statistics to observe differences between groups and then use inferential statistics to test whether the difference is *significant*.

# Being certain

- Are the two groups different?
- What measure are we comparing?
- What level of confidence? Are you 99% sure – or only 50% sure?



# Why is this unit important?

- We are constantly being told certain things are true. This subject gives you the tools to track through the evidence, apply the theory, and verify for yourself.
- For example, from the past and recent press:
  - New Johnson & Johnson Shot Prevents Severe COVID As Well As Existing Vaccines Do, Experts Say. (Scientific American, Mar 2021)
  - Belly fat may be resistant to weight loss when intermittent fasting (The Age, Mar 2021)
  - Climate change sceptics are more likely to be conservative white males. (Scientific American Feb 2016)
  - Women are better than men detecting infidelity from facial expressions. (The Age Dec 2012)



## Drinking soft drink linked to cancer risk

By Sarah Wiedersehn

Updated 22 February 2018 – 10:54am, first published at 10:42am



People who regularly consume at least one soft drink a day, no matter the size of their waist, could be at increased risk of cancer, according to a new study.

Researchers at Cancer Council Victoria and University of Melbourne analysed more than 3000 cases of 11 obesity-related cancers including breast, liver and prostate reported between 1990-1994 and 2003-2007 through The Melbourne Collaborative Cohort Study.



Published on Thursday, the study found a positive association between soft drink consumption and cancer risk independent of obesity.

Photo: Jeff Chiu

Published on Thursday, the study found a positive association between soft drink consumption and cancer risk independent of obesity after statistically adjusting for waist circumference.

<https://www.theage.com.au>

# Soft Drink

- ... the study found a positive association between soft drink consumption and cancer risk independent of obesity after statistically adjusting for waist circumference. ...
- ... "Initially our hypothesis was that drinking soft drinks would cause obesity which would then cause an association with obesity-related cancers but we found that there was more beyond the affect of obesity," ...
- ... "According to the research, the more sugary soft drinks participants drank the higher their risk of cancer. However, this was not the case with those who drank diet soft drinks, suggesting sugar could be the key, says Professor Hodge. ...
- <https://www.theage.com.au>

# Quantitative claims using data

For any quantitative argument based on data, you should get in the habit of thinking about the following three considerations:

1. Data: how was it collected, how much have you got?
2. Model: are you using the correct model for the argument? (average, linear equation, proportion, time series or forecast...?)
3. Randomness: how does variability in the data affect the reliability of the conclusion?

<https://flux.qa> (Feed code: SJ6KGV)

## Question 3

Do you believe the Manufacturer's claim?

---

- A. Yes
- B. No



<https://flux.qa> (Feed code: SJ6KGV)

## Question 4

What extra evidence would most convince you?

---

- A. Number of situations tested?
- B. Types of germs studied?
- C. How germs are counted?
- D. Number of trials conducted?
- E. How experiments are designed?



# What to do this week (today)

- Get Flux.qa set up if you have not done so.
- **Do the pre-reading for Lecture 2.**
- Log in to Moodle and check that you have access to the FIT1006 page.
- Check your tutorial time and lab using Allocate<sup>+</sup>.
- Obtain a textbook and calculator if required.

## Seek assistance as a preventative measure



Take the following relevant preventative measures if you are falling behind in your studies:

- Study difficulties: Discuss any difficulties you are experiencing with your course leader, unit coordinator, lecturer or tutor.
  - These staff members can assist you in identifying your problem areas and explore the options available to you in your course.
- Language and learning online can help you with study methods, language skills and work presentation  
<http://www.monash.edu.au/lis/llonline/>
- Student life and support services can be found at:  
<http://monash.edu/students/support/>

and include: Health services, support and services, clubs and sports etc

# Disability Support Services

**Do you have a disability, medical or mental health condition that may impact on your study?**

Disability Support Services provides a range of services for registered students including:

- Note takers and Auslan interpreters
- Readings in alternative formats
- Adaptive equipment and software
- Alternative arrangements for exam and class tests

**Disability Support Services also support students who are carers of a person with a disability, medical or mental health condition, or who is aged and frail.**

For further information and details about how to register:

T: 03 9905 5704

E: [disabilitysupportservices@monash.edu](mailto:disabilitysupportservices@monash.edu)

