



MONASH University

Information Technology

FIT1006

Business Information Analysis

Lecture 2

Sampling and Surveying

Topics covered:

- Populations and samples
- Collecting data
- Sources of data
- Designing surveys
- Survey errors
- Choosing a sample

Motivating Problem

- A major bank currently offers Private Banking services to its clients in the CBD. The bank is investigating the possibility of extending its Private Banking business to include several major country towns: Geelong, Ballarat, Hamilton and Bendigo.
- The bank wants to survey potential clients in these districts to determine whether likely demand would make this business viable.
- Suggest a sampling/survey design for the bank to use, outlining the issues you need to consider and any problems you anticipate.

10 Survey questions

These questions could all be addressed using survey data:

1. Who is settling into university better, country or city students?
2. Are students happy with the quality of food in the Campus Centre?
3. What is the average savings of first year students?
4. Do savings differ between students of different faculties?
5. What is the average age of glider pilots?
6. Are OZ Lotto Division 1 winners still happy one year after their win?
7. Where do the majority of visitors to my website click through from – Google? Yahoo? Or?
8. Is Coles cheaper than Woolworths?
9. What proportion of Australians support immigration?
10. Would the Liberal Party win an election held today?

Yesterday, I asked this question ...

Question 1

Are you currently in Australia?

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

The sampling method in Question 1 is an example of:

- A. a random sample.
- B. a convenience sample.
- C. a census.
- D. a cluster sample.
- E. a judgemental sample.

The sampling method in Question 1 is an example of:

A a random sample.



B a convenience sample.



C a census.



D a cluster sample.



E a judgemental sample.



Question 2

- Discussion in groups.

Results improved
after discussion!!

The sampling method in Question 1 is an example of:

A a random sample.

(0) 0%

B a convenience sample.



C a census.



D a cluster sample.



E a judgemental sample.



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

Question 3

The sampling method in Question 1 is an example of:

- A. a random sample.
- B. a convenience sample.
- C. a census.
- D. a cluster sample.
- E. a judgemental sample.

Choosing a Sample

- Probability Designs

- Random Sampling
- Systematic Sampling
- Stratified Sampling
- Cluster Sampling

- Non-Probability Designs

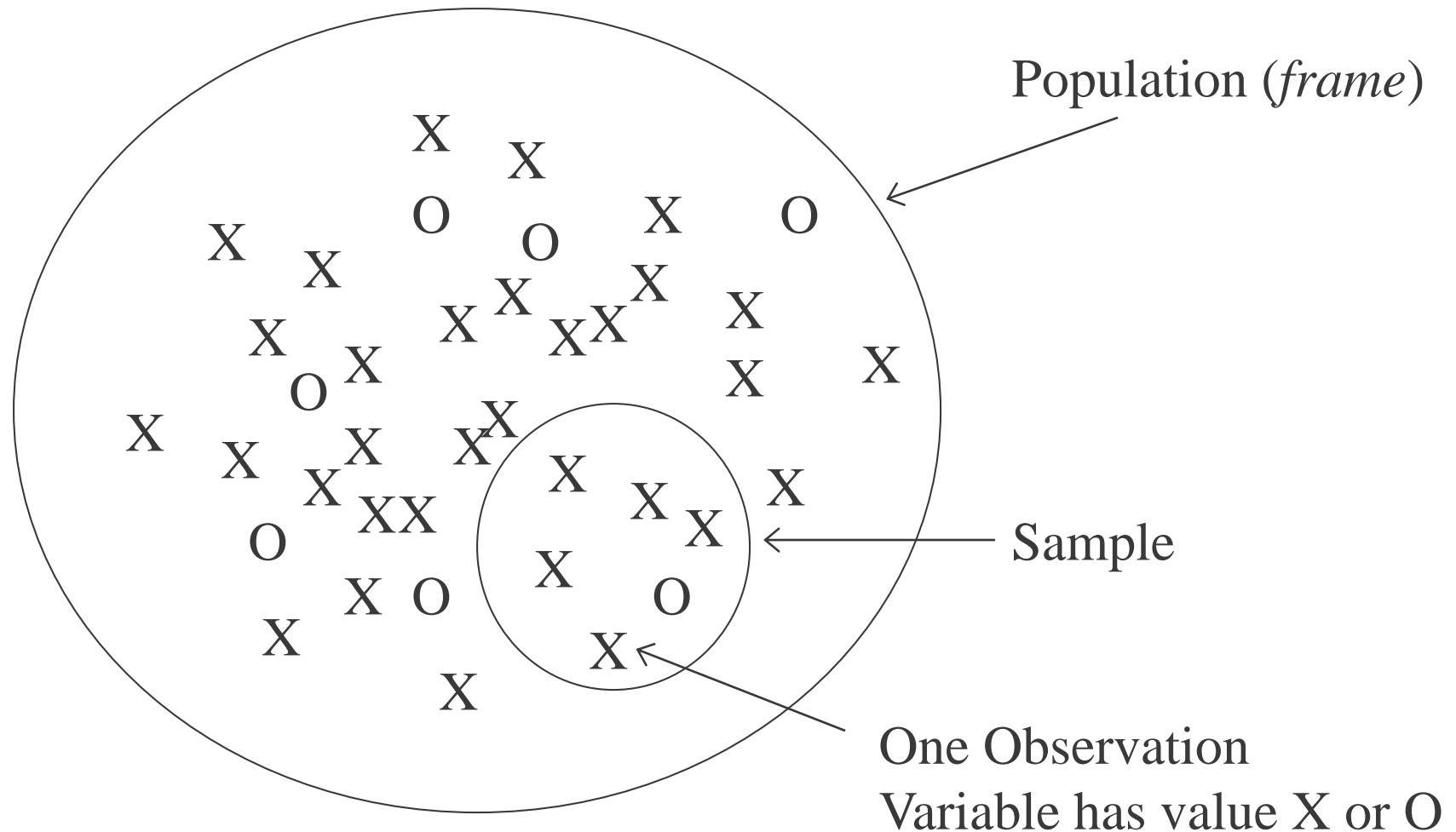
- Convenience Sampling
- Judgemental Sampling
- Quota Sampling
- Snowball Sampling

Non-probability designs are more prone to bias.

Populations and Samples

- Population: whole collection of what we are observing.
- Sample: subset we actually observe.
- *We look at the sample to make an inference about the population.*
- Population Parameter: the thing of interest.
- Sample Statistic: what we actually measure.
- *We estimate the value of the population parameter from the sample statistic.*

The General Picture



Reasons for Using Samples

- High cost of sampling a whole population.
- Length of time required to sample whole population.
- May have too much data.
- May not have access to whole population.
- Sampling can be destructive
 - Testing seat belts for breaking strain.
- A census is a survey of a whole population.
 - The Australian population census is conducted once every five years.

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

Question 4

I want to know whether Monash students support a Rail link from city. Which sample is least biased?

- A. Email all students and record the first 100 responses.
- B. Randomly choose 100 students in the Campus Centre at lunch time.
- C. Randomly choose a subject (say MED1000) and randomly choose 100 students from that class.
- ✓ D. Choose 100 students at random from the University's register and email them.

Examples of surveys

- Television ratings
- Market research
- Pre-election polls
- Product registration information
- Product Testing

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

Question 5

I want to know whether people with early-onset Alzheimer's disease have had trouble managing their personal finances. The best survey method is:

- ✓ A. Face-to-face interview.
- B. Telephone interview.
- C. A postal survey.
- D. An Internet survey.

Types of Surveys

- There are three main types of survey methods:
 - Personal (face to face) Interview
 - Telephone Survey
 - Self-administered survey (postal/Internet)
- Two main issues are at stake:
 - How to encourage people to respond truthfully and accurately?
 - How to conduct the survey with the minimum cost?

Stages of a Survey

- Exploratory interview with focus group
 - Identify issues, form hypothesis and questions
- Questionnaire designed and tested
- Experimental design and population *frame* is determined
- Samples are selected
- Questionnaire is administered
- Analysis and reporting

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

Question 6

Sampling 8 (one at a time) at random. Which sample could not have come from the population (on the RHS)?

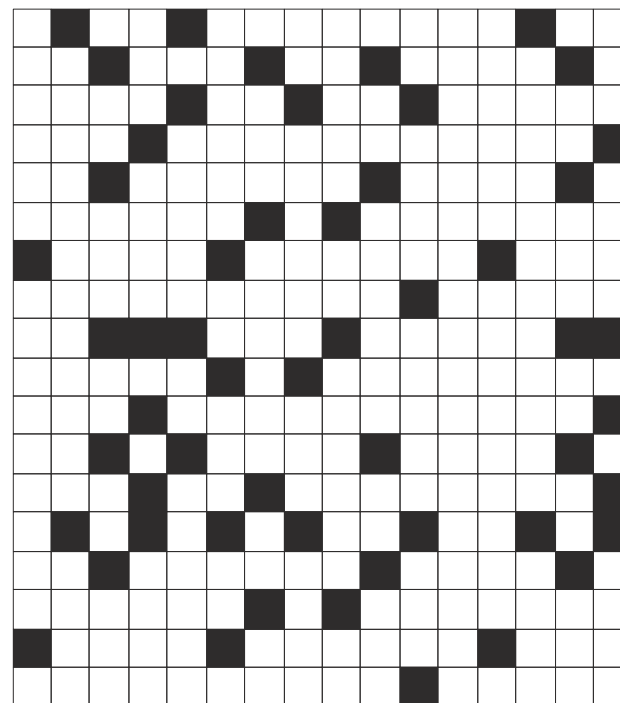
A. 

B. 

C. 

D. 

✓ E. All equally likely.



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

Question 7

Which sample is more likely to have come from the population (on the right-hand side)?

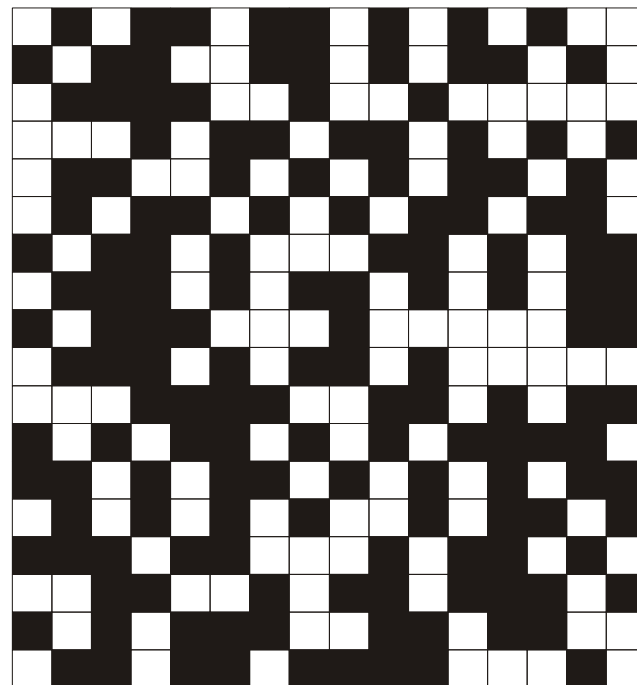
A. 

B. 

C. 

D. 

✓ E. All equally likely.




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

Question 8

Which sample is more likely to have come from the population (on the right-hand side)?

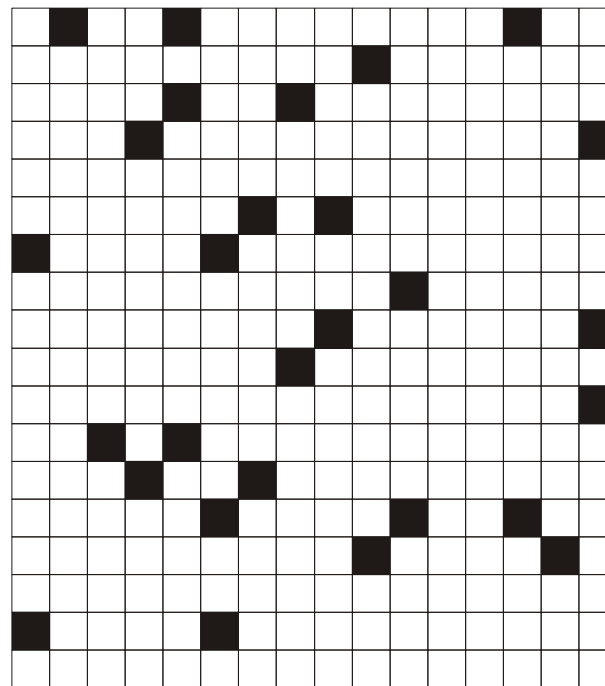
A. 

✓ B. 

C. 

D. 

E. All equally likely.



Survey Errors

- Sampling Errors
 - The characteristics of the sample do not match those of the population.
 - This is usually addressed by taking a larger sample.
- Non-sampling Errors
 - Errors in response or in recording data
 - Misclassification or inaccurate response
 - Bias in the selection of the sample
 - Non-response bias, Self-selection of respondents.

Motivating Problem

- A major bank currently offers Private Banking services to its clients in the CBD. The bank is investigating the possibility of extending its Private Banking business to include several major country towns: Geelong, Ballarat, Hamilton and Bendigo.
- The bank wants to survey potential clients in these districts to determine whether likely demand would make this business viable.
- Suggest a sampling/survey design for the bank to use, outlining the issues you need to consider and any problems you anticipate.

Key Ideas

- **You should be familiar with the following:**
 - Parameter, Sample Statistic, Variable, Observation;
 - The reasons for sampling;
 - The 3 main methods of surveying (advantages / disadvantages of each);
 - Causes of error in surveys;
 - Methods of choosing samples (4 random / 4 non-random methods).

Key Ideas cont...

■ Main points:

- Think of sampling as the first stage in the process of statistical analysis so errors here undermine any conclusions we can draw from subsequent analysis.
- Relevant issues are: sampling method, survey design, potential bias in sample selection.
- In later lectures we look at how sample size affects the validity of conclusions we draw.

Key Ideas cont...

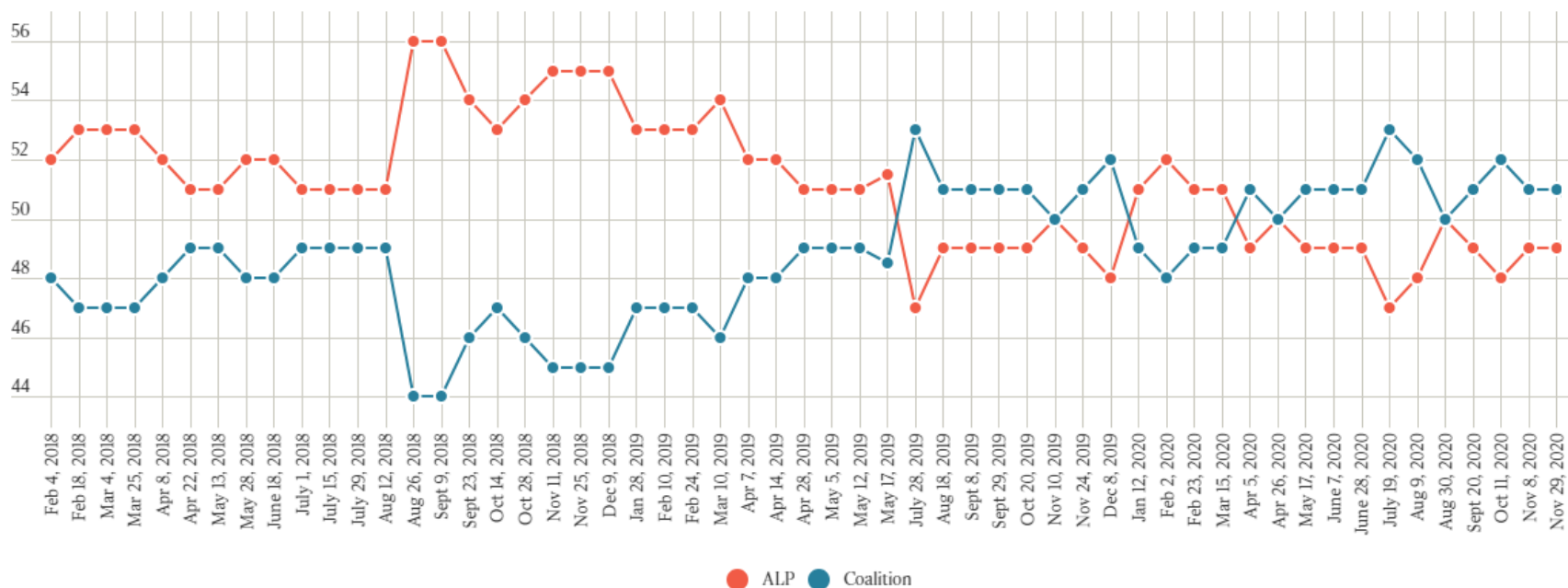
■ Think about:

- How we sample,
- Why we sample,
- How we conduct a survey and collect the results,
- Problems or considerations that emerge depending on our choice of sample/survey method.

Australian – News poll (Feb 2021)

Two-party preferred

Preference flows based on recent federal and state elections



<http://www.theaustralian.com.au/national-affairs/newspoll/vic>

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

Question 9

Based on a poll result of 49% ALP and 51% Coalition with a sample size of 1,205 we can conclude that:

- A. 95% sure ALP would have won an election held over the sampling period.
- B. 95% sure Coalition would have won the election.
- C. Results are too close to call.
- D. Sample size is too small.
- E. Don't know.

We will revisit this question later in our course

Expert attacks 'healthy' food tick

Excerpt:

- “Dr Rosemary Stanton, in an article published online in *The Medical Journal of Australia*, said food manufacturers paid to use the tick for products that might be less healthy than alternatives.
- Manufacturers pay a licence fee, based on a sliding scale of annual sales, to use the tick logo ...
- But Dr Stanton said comparable products that did not carry the tick could be a better choice, and were usually cheaper because their prices were not inflated by licence fees...”
- Design a survey to test whether ticked items are more expensive than their non-ticked counterparts.
- <http://www.theage.com.au/national/expert-attacks-healthy-food-tick-20110228-1bbqy.html>

Reading

- General

- Selvanathan Chapter 1.

- Sampling

- Selvanathan Chapter 2.
- McClennan, W., *An Introduction to Sample Surveys*, Australian Bureau of Statistics, 1999.Cat 1299.0

<http://abs.gov.au/ausstats/abs@.nsf/mf/1299.0>

- Mission Australia Youth Survey 2015

<https://www.missionaustralia.com.au/what-we-do/research-evaluation/youth-survey>

Questions – for Tutorial 1

- General

Selvanathan: Questions 1.1 – 1.7

(For thinking about...)

- Sampling

Selvanathan:

- 2.18, 2.21, 2.22, 2.24, 2.25, 2.26.

Tutorial 1 Questions.