ECON1310 Introductory Statistics for Social Sciences

Tutorial 1: DESCRIPTIVE STATISTICS I

Tutor: Francisco Tavares Garcia



Who's your Tutor?

Born in 1986 in Ourinhos / São Paulo state, Brazil

2004 - 2008

Bachelor of Computer Science

2008 - 2012

Supervisor at Procter & Gamble

2009 - 2011

MBA - FGV

2012 - 2018

Built and ran a Hostel

2018

Moved to Australia

2021 - current

Bachelor of Economics - UQ

Statistics/Econometrics:

ECON1310 - Introductory Statistics for Social Sciences

ECON2300 - Introductory

Econometrics

ECON2105 - Statistical Theory for

Economists

ECON3350 - Applied Econometrics for

Macroeconomics and Finance

STAT2003 - Mathematical Probability

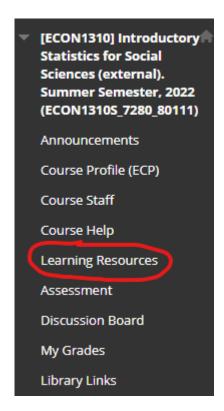
STAT2004 - Statistical Modelling &

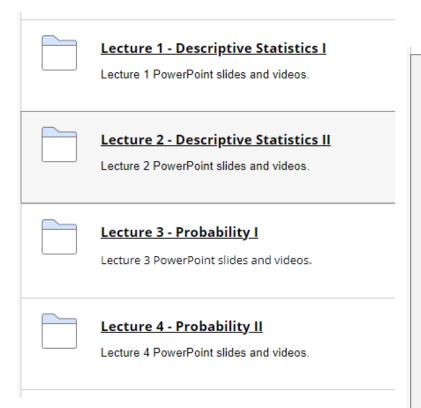
Analysis

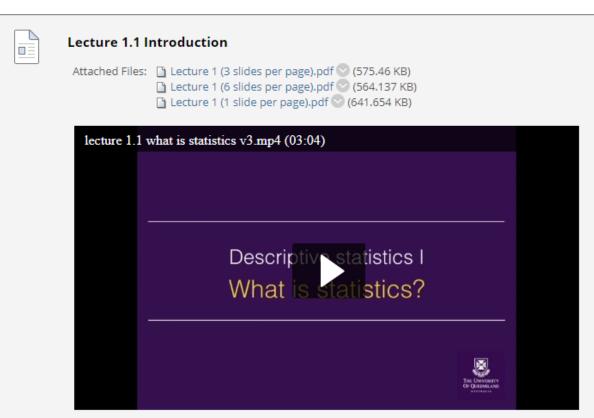




How to watch lectures? Can I attend a lecture?





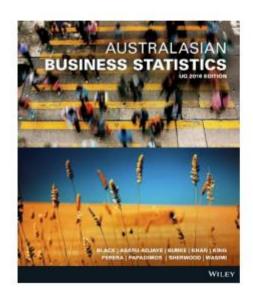


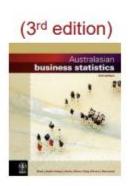


Books we're following

Australasian Business Statistics (Wiley, 4th edition)

by Black, Asafu-Adjaye, Burke, Khan, King, Perera, Papadimos, Sherwood & Wasimi.











Add to Bookshelf

Share Link to Book

Cite Book

Business Analytics and Statistics

Ken Black

Availability

Your institution has access to 6 copies of this book.

Read Online

♣ Not Available for Full Download

Download PDF Chapter

Get up to 136 pages, use any PDF software, does not expire.

Description

Currently not available for this book.

Table of Contents



How to complete this course?

Assessment Task	Due Date	Weighting	Learning Objectives
CML Quiz 1 Descriptive Statistics	07 Dec 22 9:00 - 12 Dec 22 16:00 2nd attempt: 14-16 Dec 2022, 09:00-16:00	8%	1, 2, 3, 4, 5
CML Quiz 2 Probability	14 Dec 22 9:00 - 19 Dec 22 16:00 2nd attempt: 21-23 Dec 2022, 09:00-16:00	8%	1, 2, 3, 4, 5
Online Quiz LBRT #1	03 Jan 23 9:00 - 04 Jan 23 16:00 2nd attempt: 5-6 Jan 2023, 09:00-16:00	20%	1, 2, 3, 4, 5
CML Quiz 3 Normal and Sampling Distributions	04 Jan 23 9:00 - 09 Jan 23 16:00 2nd attempt: 11-13 Jan 2023, 09:00-16:00	8%	1, 2, 3, 4, 5
CML Quiz 4 Confidence Intervals	11 Jan 23 9:00 - 16 Jan 23 16:00 2nd attempt: 18-20 Jan 2023, 09:00-16:00	8%	1, 2, 3, 4, 5
Online Quiz LBRT #2	17 Jan 23 9:00 - 18 Jan 23 16:00 2nd attempt: 19-20 Jan 2023, 09:00-16:00	20%	1, 2, 3, 4, 5
CML Quiz 5 Hypothesis Testing	18 Jan 23 9:00 - 23 Jan 23 16:00 2nd attempt: 25-27 Jan 2023, 09:00-16:00	8%	1, 2, 3, 4, 5
CML Quiz 6 Simpler Linear Regression	25 Jan 23 9:00 - 06 Feb 23 16:00 NO SECOND ATTEMPT	8% (Best 5 of 6 CML quizzes. See Section 5.4)	1, 2, 3, 4, 5
Online Quiz LBRT #3	07 Feb 23 9:00 - 08 Feb 23 16:00 2nd attempt: 9-10 Feb 2023, 09:00-16:00	20%	1, 2, 3, 4, 5

But I heard there is an IVA... (Identity Verified Assessment)



What is the Identity Verified Assessment (IVA) process?

Overview of student Identity Verified Assessment (IVA) process.

Purpose: Regular attendance and active participation during tutorials helps support your learning. It also allows tutors to get to know you, which in turn helps them to monitor and compare your *observed* tutorial performances with your *actual* assessment results (CMLs, LBRTs). Importantly, it also allows tutors to identify struggling students so that additional support can be provided. Consequently, tutors are well placed at the end of the semester to help verify students' anticipated grades, providing a valuable means of maintaining academic standards.

Summary of IVA Process: At the end of the semester, you may be invited to attend a 15 minute **mandatory oral interview.** Typically, tutors will invite students who did not attend any (or many) tutorials and/or consultation sessions during the semester, thereby making it difficult to verify a student's *actual* assessment performance.

What to do if you are invited to an interview: The interview will be held with your tutor. It will include a series of questions requiring an oral explanation, or a brief written explanation, involving numerical calculations or theoretical concepts related to any lecture. NOTE: if you have been actively engaging with your tutor during the semester, it will reduce the need to attend an interview.



How to complete this course?

To successfully pass ECON1310, a student needs to achieve all three of the following requirements:

- 1. Receive a **combined total assessment mark of at least 50 marks out of 100 marks** from their CMLs Quizzes (40 marks available) and LBRTs (60 marks available).
- Receive at least 16 marks out of 40 marks from their best 5 out of 6 CML Quizzes (each being worth 8 marks).
- 3. Receive at least 24 marks out of 60 marks from their LBRTs (each being worth 20 marks).

In determining a student's final overall grade for the course, the following will apply:

- If the CML Quizzes aggregate final mark is less than 16/40 and the LBRTs aggregate final mark is less than 24/60, the maximum possible grade will be 2.
- If the CML Quizzes aggregate final mark is less than 16/40 or the LBRTs aggregate final mark is less than 24/60, the maximum possible grade will be 3.

Details regarding the two assessable items (CML Quizzes and LBRTs) are outlined below.



I need HELP!!!

- <u>d.byrne@uq.edu.au</u> for academic or administrative questions.
- <u>cml.1310@uq.edu.au</u> for CML quiz administration.
- Consultation every weekday!! (Mostly afternoon)

TIME	MON	TUE	WED	THU	FRI
10:00-10:30					
10:30-11:00					
11:00-11:30					
11:30-12:00					
12:00-12:30					DOMINIC (12pm – 1pm) https://uqz.zoom.us/j/5207526654
12:30-13:00					
13:00-13:30		BEN (1pm – 2pm)		BEN (1pm – 2pm)	
13:30-14:00		https://uqz.zoom.us/j/7884658078	PETER (1pm – 3pm)	https://uqz.zoom.us/j/7884658078	
14:00-14:30			https://uqz.zoom.us/j/84419335972		
14:30-15:00					
15:00-15:30					
15:30-16:00					
16:00-16:30	FRANCISCO (4pm - 5pm)		FRANCISCO (4pm - 5pm)		
16:30-17:00	https://uqz.zoom.us/j/3181814065		https://uqz.zoom.us/j/3181814065		



[ECON1310] Introductory Statistics for Social Sciences (external).
Summer Semester, 2022 (ECON1310S_7280_80111)

Announcements

Course Profile (ECP)

Course Staff

Course Help

Learning Resources

Assessment

Discussion Board

My Grades

Library Links



ECON1310 Tutorial 1 – Week 2

DESCRIPTIVE STATISTICS I

At the end of this tutorial you should be able to

- Define basic statistical concepts.
- Discuss the different categories of data.
- · Outline the difference between descriptive statistics and inferential statistics
- Discuss several sampling methods and associated errors.



- Q1. a) What is statistics? Explain the difference between descriptive and inferential statistics. Explain why inferential statistics is not required if we have a census.
 - b) Define: population, sample, parameter, and statistic.



Q1. a) What is statistics? Explain the difference between descriptive and inferential statistics. Explain why inferential statistics is not required if we have a census.



Q1. a) What is statistics? Explain the difference between descriptive and inferential statistics. Explain why inferential statistics is not required if we have a census.

Defining the subject of statistics?

 a branch of mathematics dealing with the analysis and interpretation of data.

Descriptive Statistics

- used to organise, explore, summarise or describe collected data.
- used to present and make sense of data (visually using histograms, pie charts etc.)
- typically needs data to be collected (eg: survey), presented (eg: graph), and characterised (eg:using terms like average).

Example of unorganised data where descriptive statistics could help bring meaning to the data:

24,26,25,33,31,39,28,36,25,34,25,29,41,36,35,27

Inferential statistics

- = the main focus of ECON1310
- drawing conclusions about a much larger population by examining a smaller, representative sample, that is taken from the population.
- requires a systematic set of rules be followed.

Lecture 5 ECON1310

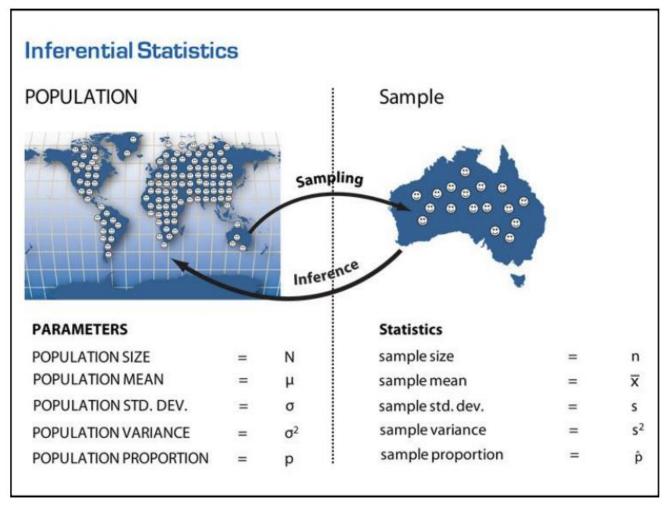
11



- Q1. a) What is statistics? Explain the difference between descriptive and inferential statistics. Explain why inferential statistics is not required if we have a census.
 - b) Define: population, sample, parameter, and statistic.

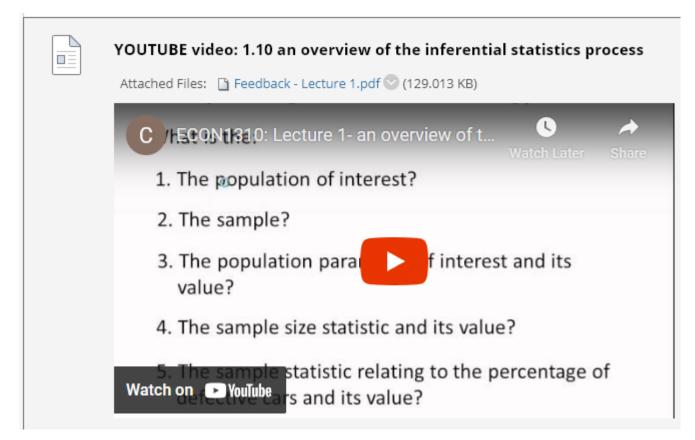


- Q1. a) What is statistics? Explain the difference between descriptive and inferential statistics. Explain why inferential statistics is not required if we have a census.
 - b) Define: population, sample, parameter, and statistic.





- Q1. a) What is statistics? Explain the difference between descriptive and inferential statistics. Explain why inferential statistics is not required if we have a census.
 - b) Define: population, sample, parameter, and statistic.



https://youtu.be/iVn8lbQVZUg



- Q2. a) A manufacturer of mint lollies claims that less than 10% of his product is defective. When 1,000 lollies were drawn from a large production run, 7.5% were found to be defective. What is: the population of interest? the sample? the value of the parameter? the value of the statistic?
 - b) A large company wanted to determine client satisfaction with their products and services. The Director of Marketing developed a questionnaire that yields a satisfaction score between 10 and 50 for participant responses. A random sample of the company's 900 clients is asked to complete a satisfaction survey. The satisfaction scores for the 35 participants are averaged to produce a mean satisfaction score. What is the population for this study? the sample? the statistic? What would be a parameter for this study?



Q2. a) A manufacturer of mint lollies claims that less than 10% of his product is defective. When 1,000 lollies were drawn from a large production run, 7.5% were found to be defective. What is: the population of interest? the sample? the value of the parameter? the value of the statistic?

1. What is: the population of interest? the sample? the value of the parameters of the statistic? (Single Choice)	eter? the	e
1000 Iollies; large production run; 10%; 7.5%		
large production run; 1000 Iollies; 7.5%; 10%		
10%; 7.5%; large production run; 1000 Iollies		
large production run; 1000 Iollies; 10%; 7.5%		
1000 Iollies; large production run; 7.5%; 10%		
	<i>.</i>	••



Q2. a) A manufacturer of mint lollies claims that less than 10% of his product is defective. When 1,000 lollies were drawn from a large production run, 7.5% were found to be defective. What is: the population of interest? the sample? the value of the parameter? the value of the statistic?

1. What is: the population of interest? the sample? the value of the parameter? the value of the statistic? (Single Choice)		he
1000 Iollies; large production run; 10%; 7.5%		
large production run; 1000 Iollies; 7.5%; 10%		
10%; 7.5%; large production run; 1000 Iollies		
large production run; 1000 Iollies; 10%; 7.5%		
1000 Iollies; large production run; 7.5%; 10%		
	O 1	•••

i) large production run ii) 1000 lollies iii) 10% iv) 7.5%



b) A large company wanted to determine client satisfaction with their products and services. The Director of Marketing developed a questionnaire that yields a satisfaction score between 10 and 50 for participant responses. A random sample of the company's 900 clients is asked to complete a satisfaction survey. The satisfaction scores for the 35 participants are averaged to produce a mean satisfaction score. What is the population for this study? the sample? the statistic? What would be a parameter for this study?

(Poll)



b) A large company wanted to determine client satisfaction with their products and services. The Director of Marketing developed a questionnaire that yields a satisfaction score between 10 and 50 for participant responses. A random sample of the company's 900 clients is asked to complete a satisfaction survey. The satisfaction scores for the 35 participants are averaged to produce a mean satisfaction score. What is the population for this study? the sample? the statistic? What would be a parameter for this study?

1. What is the population for this study? the sample? the statistic? What would be a
parameter for this study? (Single Choice)
 35 participants, 900 clients, satisfaction (35 part), satisfaction (900 cli)
900 clients, 35 participants, satisfaction (35 part), satisfaction (900 cli)
 35 participants, 900 clients, satisfaction (900 cli), satisfaction (35 part)
900 clients, 35 participants, satisfaction (900 cli), satisfaction (35 part)



b) A large company wanted to determine client satisfaction with their products and services. The Director of Marketing developed a questionnaire that yields a satisfaction score between 10 and 50 for participant responses. A random sample of the company's 900 clients is asked to complete a satisfaction survey. The satisfaction scores for the 35 participants are averaged to produce a mean satisfaction score. What is the population for this study? the sample? the statistic? What would be a parameter for this study?

What is the population for this study? the sample? the statistic? What would be a parameter for this study? (Single Choice)
 35 participants, 900 clients, satisfaction (35 part), satisfaction (900 cli)
 900 clients, 35 participants, satisfaction (35 part), satisfaction (900 cli)
 35 participants, 900 clients, satisfaction (900 cli), satisfaction (35 part)
 900 clients, 35 participants, satisfaction (900 cli), satisfaction (35 part)

b) population is 900 clients of the company sample is 35 participants statistic is the mean satisfaction score from the 35 participants parameter would be the average satisfaction score for all the 900 clients



Q3. Coca-Cola and Pepsi display intense competition in their marketing campaigns featuring TV stars, rock videos, endorsements by well-known athletes etc. Suppose, as part of a Pepsi marketing campaign, 1000 [cola] consumers are given a blind taste test (i.e. a taste test in which the two brand names are not recognizable by the participants). Each consumer is asked to state a preference for Brand A or Brand B. Describe the population, the variable of interest, the sample and the inference.



Q3. Coca-Cola and Pepsi display intense competition in their marketing campaigns featuring TV stars, rock videos, endorsements by well-known athletes etc. Suppose, as part of a Pepsi marketing campaign, 1000 [cola] consumers are given a blind taste test (i.e. a taste test in which the two brand names are not recognizable by the participants). Each consumer is asked to state a preference for Brand A or Brand B. Describe the population, the variable of interest, the sample and the inference.

Population: all cola consumers

Variable of interest: proportions of consumers who prefer A to B

Sample: 1000 consumers given a blind taste test

The inference: To use the proportion we found in the blind test

(sample) to infer something about the population.



Q3. Coca-Cola and Pepsi display intense competition in their marketing campaigns featuring TV stars, rock videos, endorsements by well-known athletes etc. Suppose, as part of a Pepsi marketing campaign, 1000 [cola] consumers are given a blind taste test (i.e. a taste test in which the two brand names are not recognizable by the participants). Each consumer is asked to state a preference for Brand A or Brand B. Describe the population, the variable of interest, the sample and the inference.

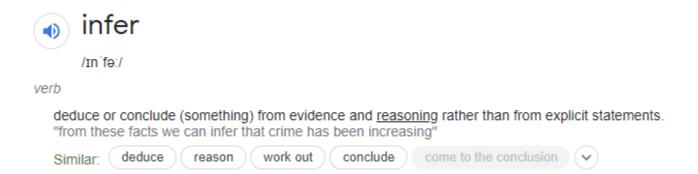
Population: all cola consumers

Variable of interest: proportions of consumers who prefer A to B

Sample: 1000 consumers given a blind taste test

The inference: To use the proportion we found in the blind test

(sample) to infer something about the population.





- **Q4.** Explain the difference between a census and a sample survey. List 3 examples (for each part) where we rely on sample information rather than a census because:
 - i) a census would be too expensive,
 - ii) the product is destroyed in the process of making the observation.

What other reasons are there for using sample data rather than census data?



Q4. Explain the difference between a census and a sample survey. List 3 examples (for each part) where we rely on sample information rather than a census because:

i) a census would be too expensive,

ii) the product is destroyed in the process of making the observation.

What other reasons are there for using sample data rather than census data?



- Q4. Explain the difference between a census and a sample survey. List 3 examples (for each part) where we rely on sample information rather than a census because:
- A census will survey the whole population
- A sample survey will only survey part of the population (the sample)
- i) a census would be too expensive,
- Market research to launch a new product.
- Election Polls.
- Everything that has a very large target audience.

- ii) the product is destroyed in the process of making the observation.
 - Tasting chips at the end of production.
 - Testing the resistance of a brick.
- Quality assurance in a manufacturing line,

What other reasons are there for using sample data rather than census data?

- Not enough time to census the population.
- Data availability (secondary data).
- Ethical reasons (drug testing).



- Q5. a) For each of the following random variables,
 - (i) determine whether the variable is categorical or numerical (then continuous or discrete);
 - (ii) determine whether the variable is nominal, ordinal, interval, or ratio.
 - a. the amount of money you spent buying clothes last month (to the nearest whole dollar).
 - the number t-shirts you own.
 - c. the job title of different employees in a large department store.
 - d. the time you spent shopping for clothes last month (in hours, to two decimal places).
 - e. when you go shopping for clothes during the week (eg: weekday, weeknight, weekend).
 - f. the type of transport you can use to go shopping.



Q5. a) For each of the following random variables,

categorical or numerical (then continuous or discrete);

- (i) determine whether the variable is categorical or numerical (then continuous or discrete);
- (ii) determine whether the variable is nominal, ordinal, interval, or ratio.

nominal, ordinal, interval, or ratio.

- a. the amount of money you spent buying clothes last month (to the nearest whole dollar).
- the number t-shirts you own.
- c. the job title of different employees in a large department store.
- d. the time you spent shopping for clothes last month (in hours, to two decimal places).
- e. when you go shopping for clothes during the week (eg: weekday, weeknight, weekend).
- f. the type of transport you can use to go shopping.



Q5. a) For each of the following random variables,

categorical or numerical
(then continuous or discrete);

(i) determine whether the variable is categorical or numerical (then continuous or discrete);

(ii) determine whether the variable is nominal, ordinal, interval, or ratio.

nominal, ordinal, interval, or ratio.

Numerical discrete

a. the amount of money you spent buying clothes last month (to the nearest whole dollar). Ratio

Ratio

Numerical discrete

b. the number t-shirts you own.

Nominal

Categorical

c. the job title of different employees in a large department store.

Numerical continuous

d. the time you spent shopping for clothes last month (in hours, to two decimal places).

Ratio

Categorical

e. when you go shopping for clothes during the week (eg: weekday, weeknight, weekend).

Nominal

Categorical

f. the type of transport you can use to go shopping.

Nominal



- Select the type of sample (SRS, Systematic, stratified, cluster) that has been formed in each survey.
 - i) The PPP company has 2000 employees. The CEO wants to investigate job satisfaction. The employees are numbered from 1 to 2000 and then 50 different random numbers are generated between 1 and 2000. A sample is formed by choosing the 50 employees whose numbers were generated.
 - ii) A small town has 10000 people who are investors in one way or another. The proportions are 1:3:6 for government bonds: real estate: stocks. To investigate investors' returns a researcher randomly selects 20 bond holders, 60 real estate investors and 120 stock investors to form a sample of 200.
 - iii) The CEO of PPP company decides to break the 2000 employees into 50 groups, each containing 40 employees. He selects randomly a number between 1 and 40 and selects the employee with this number. He then selects every 40th employee after this.

(Which employees (numbered from 1 to 2000) would be chosen in (iii) if the randomly selected number was 35?)



(SRS, Systematic, stratified, cluster)

- Select the type of sample (SRS, Systematic, stratified, cluster) that has been formed in each survey.
- i) The PPP company has 2000 employees. The CEO wants to investigate job satisfaction. The employees are numbered from 1 to 2000 and then 50 different random numbers are generated between 1 and 2000. A sample is formed by choosing the 50 employees whose numbers were generated.
- ii) A small town has 10000 people who are investors in one way or another. The proportions are 1:3:6 for government bonds: real estate: stocks. To investigate investors' returns a researcher randomly selects 20 bond holders, 60 real estate investors and 120 stock investors to form a sample of 200.
- iii) The CEO of PPP company decides to break the 2000 employees into 50 groups, each containing 40 employees. He selects randomly a number between 1 and 40 and selects the employee with this number. He then selects every 40th employee after this.

(Which employees (numbered from 1 to 2000) would be chosen in (iii) if the randomly selected number was 35?)



(SRS, Systematic, stratified, cluster)

Select the type of sample (SRS, Systematic, stratified, cluster) that has been formed in each survey.

SRS Simple Random Sampling

i) The PPP company has 2000 employees. The CEO wants to investigate job satisfaction. The employees are numbered from 1 to 2000 and then 50 different random numbers are generated between 1 and 2000. A sample is formed by choosing the 50 employees whose numbers were generated.

Stratified

ii) A small town has 10000 people who are investors in one way or another. The proportions are 1:3:6 for government bonds: real estate: stocks. To investigate investors' returns a researcher randomly selects 20 bond holders, 60 real estate investors and 120 stock investors to form a sample of 200.

Systematic

iii) The CEO of PPP company decides to break the 2000 employees into 50 groups, each containing 40 employees. He selects randomly a number between 1 and 40 and selects the employee with this number. He then selects every 40th employee after this.

35th, 75th, 115th... (Which employees (numbered from 1 to 2000) would be chosen in (iii) if the randomly selected number was 35?)



Q6. A company employs 3500 people and a random sample of 175 of these employees has been taken by systematic sampling. The researcher would start the sample selection between what two values? How would the researcher choose subsequent members of the sample? Where might the researcher obtain a frame for this study?



Q6. A company employs 3500 people and a random sample of 175 of these employees has been taken by systematic sampling. The researcher would start the sample selection between what two values? How would the researcher choose subsequent members of the sample? Where might the researcher obtain a frame for this study?

Decide on sample size (n) and calculate k:

k = N/n (round to nearest whole number if required)



Q6. A company employs 3500 people and a random sample of 175 of these employees has been taken by systematic sampling. The researcher would start the sample selection between what two values? How would the researcher choose subsequent members of the sample? Where might the researcher obtain a frame for this study?

Decide on sample size (n) and calculate k:

k = N/n (round to nearest whole number if required)

3500 / 175 = 20

The researcher would start the sample selection between 1 and 20.

After selecting the first member, the researcher would select the member who is 20 positions after the first one. This process would continue until 175 members are selected. (16th, 36th, 56th, etc.)

The researcher could have a list of employees by age, first name, or last name, sector, etc.

Decide on sample size (n) and calculate k:
 k = N/n (round to nearest whole number if required)

Randomly select a starting point between 1 and k.

Select every kth individual after the starting point.



ECON1310 Tutorial 1 – Week 2

DESCRIPTIVE STATISTICS I

At the end of this tutorial you should be able to

- Define basic statistical concepts.
- Discuss the different categories of data.
- Outline the difference between descriptive statistics and inferential statistics
- Discuss several sampling methods and associated errors.



Thank you

Francisco Tavares Garcia

Academic Tutor | School of Economics

tavaresgarcia.github.io

Reference

Black et al. (2016), Australasian Business Statistics, 4th Edition, Wiley Australia.

CRICOS code 00025B

