



AS91264 (version 3)

Mathematics and Statistics



**Use statistical methods
to make an inference**



Te whai i ngā tikanga o te tūhuratanga
tauanga hei whakaputa hīkaro



Te whai hua - kia ora!

sorted
in Schools

Student name:

Class:

LEVEL

2

CREDITS

4

SORTED THEMES

KiwiSaver
Retirement
Managing my Money

Module Enquiry Context:

Module Enquiry Context Activity

1) a. What is KiwiSaver?

b. When can you access your KiwiSaver funds?

c. Give two reasons why females often have less money than males when they reach retirement age.



Watch the **Sorted in Schools Inequality video**.

- 2) a.** How does the amount of unpaid labour done by females in Aotearoa New Zealand compare with that done by males?

.....

.....

.....

.....

.....

- b.** What are three factors, other than gender, that contribute to income disparities in Aotearoa New Zealand?

.....

.....

.....

.....

.....



Topic One:

Establishing a purpose and an investigative question

Topic 1, Activity 1



Watch the **Sorted in Schools Inequality video**.



Discuss how the gender pay gap and financial wellbeing in retirement might be connected.

Topic 1, Activity 2

Identify whether the following variables are categorical variables **or** numerical variables:

Variable	Categorical or numerical
Gender	
KiwiSaver balance	
Personal income	
Employment status	
Age group	

Topic 1, Activity 3



In the **Sorted in Schools Inequality video**, Sapphie and Simon mention a range of factors that contribute to income inequality.

Identify at least two other factors apart from gender.

Topic 1, Activity 4

Complete this cloze activity, filling in the gaps in the sentences using the words in the box below. Note that you may need to use a word more than once.

parameter	statistic	sample	population	unknown
-----------	-----------	--------	------------	---------

A is any large collection of objects or individuals that we want to investigate.

A is any summary number, such as an average or percentage, that describes the entire population. This value is because the size is too large to find out what it is.

A is a group drawn from the population.

A is any summary number, such as an average or a percentage, that describes the sample.

Because samples are manageable in size, we can work out the value of any

We can then use this to make an informed guess about the value of a population

Topic 1, Activity 5

The table below shows the KiwiSaver investment balances of 15 randomly selected people. The values have been placed in numerical order (smallest to largest). The data has also been plotted on a graph to show you how it is spread out.

ID	KiwiSaver balance	ID	KiwiSaver balance	ID	KiwiSaver balance
1	\$107	6	\$9,160	11	\$34,484
2	\$2,309	7	\$14,191	12	\$40,056
3	\$3,787	8	\$17,727	13	\$63,629
4	\$4,217	9	\$19,015	14	\$70,460
5	\$6,682	10	\$27,929	15	\$143,790

Topic 1, Activity 6

Natasha has developed the following investigative question for her statistical enquiry:

In 2018, was the median KiwiSaver balance of full-time workers in Aotearoa New Zealand (measured in NZ dollars) higher than the median KiwiSaver balance of part-time workers in Aotearoa New Zealand (measured in NZ dollars)?

- a. Fill in the table below using components of Natasha's investigative question to show that she has met all of the requirements of an investigative question.

The variable of interest (including units)	
The two groups being compared	
The parameter being compared	
The direction of the comparison	
The population	

- b. The CFFC barometer data set has information on the following variables:

Age	Year
Age group	18 to 34, 35 to 54, 55 to 64, 65+
Gender	Female, Male, Other
Ethnicity	Note that in the original data set, participants could select more than one ethnicity
Region	Region where the person lives, for example, Auckland, Christchurch, Porirua/Hutt Valley

KiwiSaver balance	How much money the person has in their KiwiSaver account Note that the data set does not include anyone who does not have a KiwiSaver account
Work status	Full-time, Part-time, Unemployed
Personal income	Before tax income measured in New Zealand dollars

Using these variables, write two investigative questions that you could use as the basis of a statistical enquiry.

For each question, make sure that you identify:

- which variable is being measured
- which two groups you are comparing
- which parameter you are comparing
- the direction of the comparison
- the population.

1)

2)

Topic Two:

Selecting a sample

Topic 2, Activity 1

- 1) A survey is emailed to 100 students randomly selected from your school asking how much money they had spent at the school cafeteria the week before.

a. What is the population?

b. What is the variable of interest?

c. Why might the sample not be representative of the population?

- 2) What is one advantage of a stratified sample compared to a simple random sample?

- 3) If you are taking a stratified sample from the New Zealand population, what are some “strata” (groups) that you would want to be represented in your sample?

Topic 2, Activity 2

Using the **CFFC Barometer data**, practise taking a sample from the population using either NZGrapher or another statistical tool. Play around with the sampling options so that you become familiar with the various functions of the tool.

1) Take a sample that meets the following requirements:

- a. 50 people aged 35–54
- b. 40 full-time workers and 40 part-time workers
- c. 45 females who identify as Māori.

Hint: to take a sample that meets this definition, you will need to take a sample from a sample.

2) a. If you are comparing two groups based on age, why does it make more sense to select a sample using the “age group” variable rather than simply “age”?



Topic Three:

Displaying data and calculating statistics

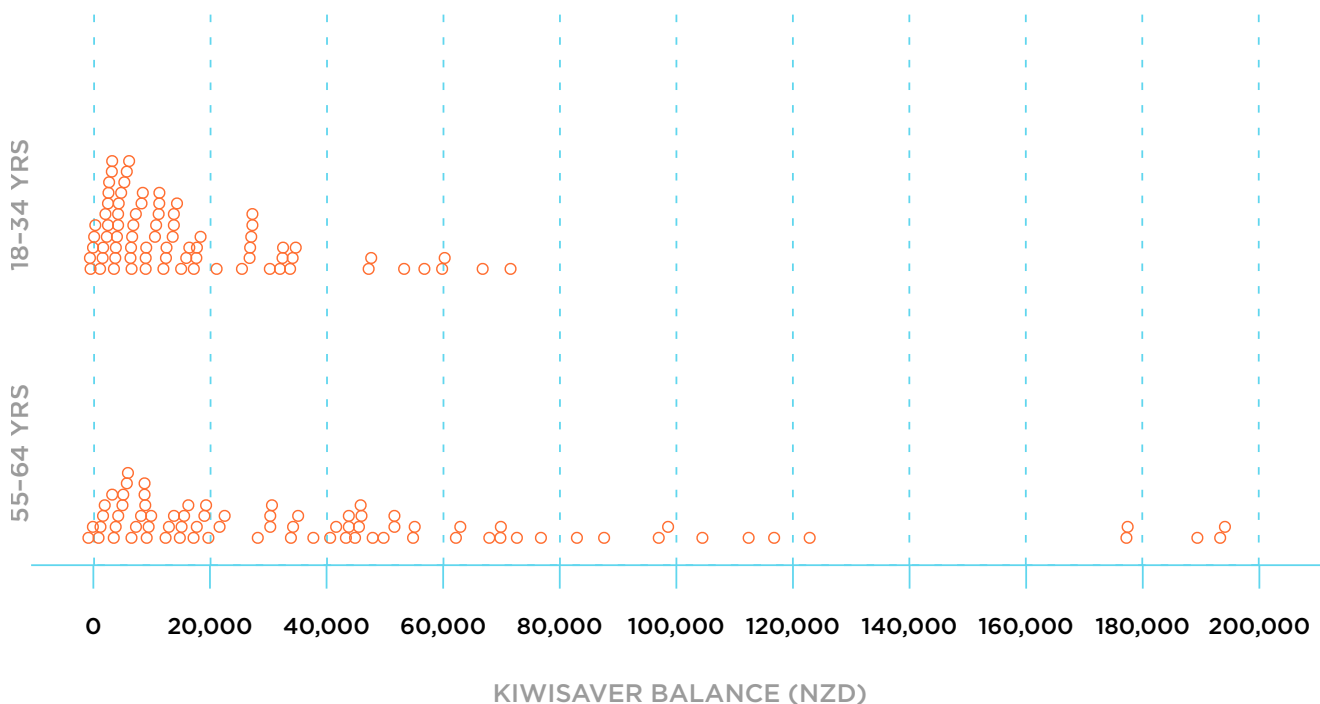
Topic 3, Activity 1

Marcus is investigating whether older KiwiSaver investors tend to have higher KiwiSaver balances (on average) than younger KiwiSaver investors. He takes a sample of:

- the KiwiSaver balances of 80 people aged 18 to 34
- the KiwiSaver balances of 80 people aged 55 to 64.

He creates a dot plot of each sample in order to compare their distributions.

KiwiSaver balances by age group



Compare the two sample distributions by identifying what you can see and explaining what it tells you. Try to mention at least two features, for example:

- centre
- symmetry
- overlap
- outliers

Topic 3, Activity 2

- 1) Using the **CFFC Barometer data**, practise taking a sample from the population using NZGrapher or another statistical tool. Make sure that you know how to:
 - select a sample
 - draw a dot plot
 - add a box plot
 - show summary statistics (minimum, lower quartile, median, upper quartile, maximum and mean).
- 2) Use NZGrapher or another statistical tool to create a box plot showing the distribution of KiwiSaver balances for a sample of 60 people. Remember to add a title to the graph.
 - a. Add the summary statistics on the graph you have created.
 - b. Compare the median and the mean of the sample. Explain why the median and the mean are different.

.....

.....

.....

.....

.....

.....

.....



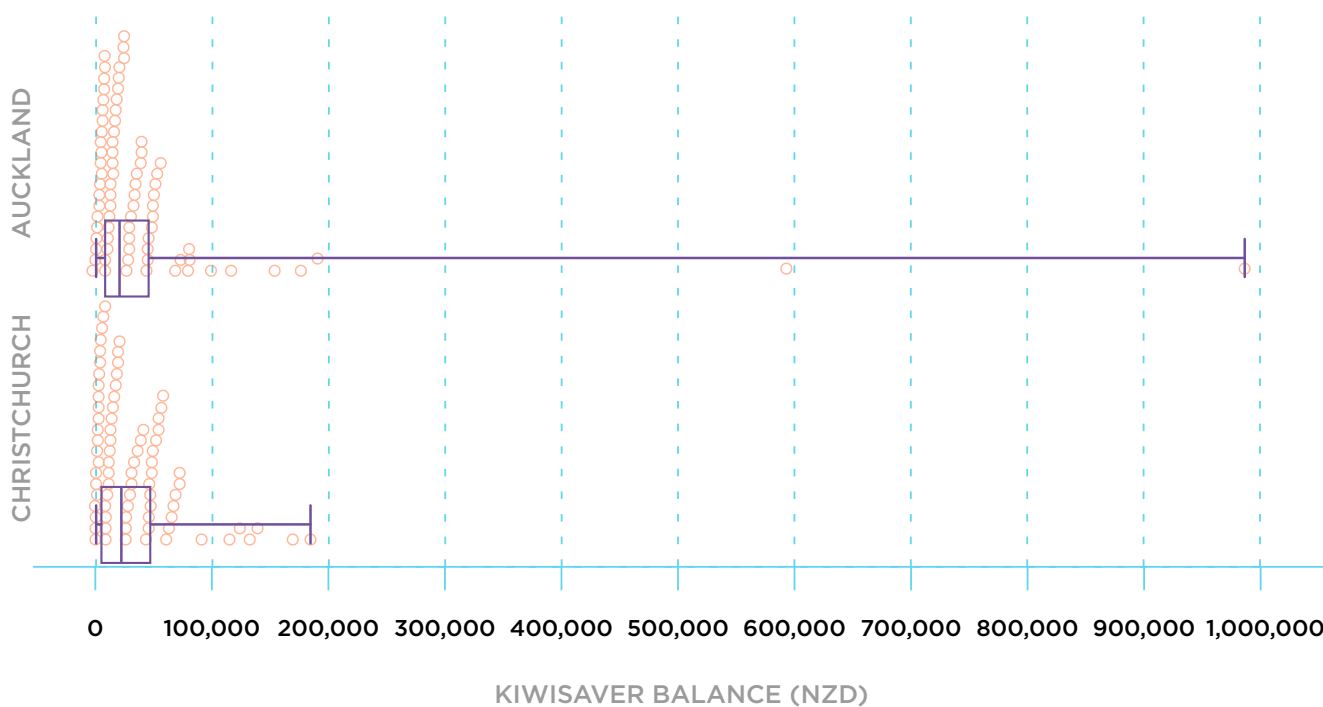
Topic Four:

Comparing box plots and summary statistics

Topic 4, Activity 1

- 1) The graph below shows the KiwiSaver balances of a sample of 80 people from Auckland and 80 people from Christchurch.

KiwiSaver balances of people from Auckland and Christchurch



Summary Statistics		
	Auckland sample	Christchurch sample
Minimum	\$157	\$528
Lower Quartile	\$7,154	\$4,892
Median	\$18,780	\$19,384
Mean	\$50,748	\$34,301
Upper quartile	\$44,726	\$47,998
Maximum	\$989,010	\$184,149

Use the PEEL framework below to make comparisons between the two box plots.

P	Point	Make a statement about what you can see in the dot plot or box plot.
E	Explanation	Explain why you think this.
E	Evidence	Use numbers or facts to back up what you are saying.
L	Link	Link your statements to the population, the context, your research or to a prediction you made at the start of your investigation.

Shape

Point	<div></div> <div></div> <div></div>
Explanation	<div></div> <div></div> <div></div>
Evidence	<div></div> <div></div> <div></div>
Link	<div></div> <div></div> <div></div>



Symmetry

Point	<div></div> <div></div> <div></div>
Explanation	<div></div> <div></div> <div></div>
Evidence	<div></div> <div></div> <div></div>
Link	<div></div> <div></div> <div></div>

Spread

Point	<div></div> <div></div> <div></div>
Explanation	<div></div> <div></div> <div></div>
Evidence	<div></div> <div></div> <div></div>
Link	<div></div> <div></div> <div></div>

Overlap

Point	<div></div> <div></div> <div></div>
Explanation	<div></div> <div></div> <div></div>
Evidence	<div></div> <div></div> <div></div>
Link	<div></div> <div></div> <div></div>

2) Repeat this process using a sample of two groups of your choice from the CFFC Barometer data set.

Shape

Point	<div></div> <div></div> <div></div>
Explanation	<div></div> <div></div> <div></div>
Evidence	<div></div> <div></div> <div></div>
Link	<div></div> <div></div> <div></div>

Symmetry

Point	<div></div> <div></div> <div></div>
Explanation	<div></div> <div></div> <div></div>
Evidence	<div></div> <div></div> <div></div>
Link	<div></div> <div></div> <div></div>

Spread

Point	<div></div> <div></div> <div></div>
Explanation	<div></div> <div></div> <div></div>
Evidence	<div></div> <div></div> <div></div>
Link	<div></div> <div></div> <div></div>

Overlap

Point	<div></div> <div></div> <div></div>
Explanation	<div></div> <div></div> <div></div>
Evidence	<div></div> <div></div> <div></div>
Link	<div></div> <div></div> <div></div>

Topic Five:

Confidence intervals

Topic 5, Activity 1

- 1) Nate and Emily each take a sample from the same population, but when they calculate the mean and median of their samples, they get different results.
- Explain why.

Topic 5, Activity 2

- 1) Rauhina has taken a sample of the KiwiSaver balances of 100 females and 100 males. Here are the summary statistics from her sample:

Summary Statistics		
	Female sample	Male sample
Minimum	\$181	\$299
Lower Quartile (LQ)	\$5,671	\$7,584
Median	\$14,616	\$18,307
Upper quartile (UQ)	\$33,486	\$45,555
Maximum	\$149,210	\$594,080
Interquartile range (IQR)	\$27,815	\$37,971
Sample size (n)	100	100

Use the formulae below to construct informal confidence intervals for the following two samples:

$$\text{lower limit} = \text{median} - 1.5 \times \frac{\text{IQR}}{\sqrt{n}}$$

and

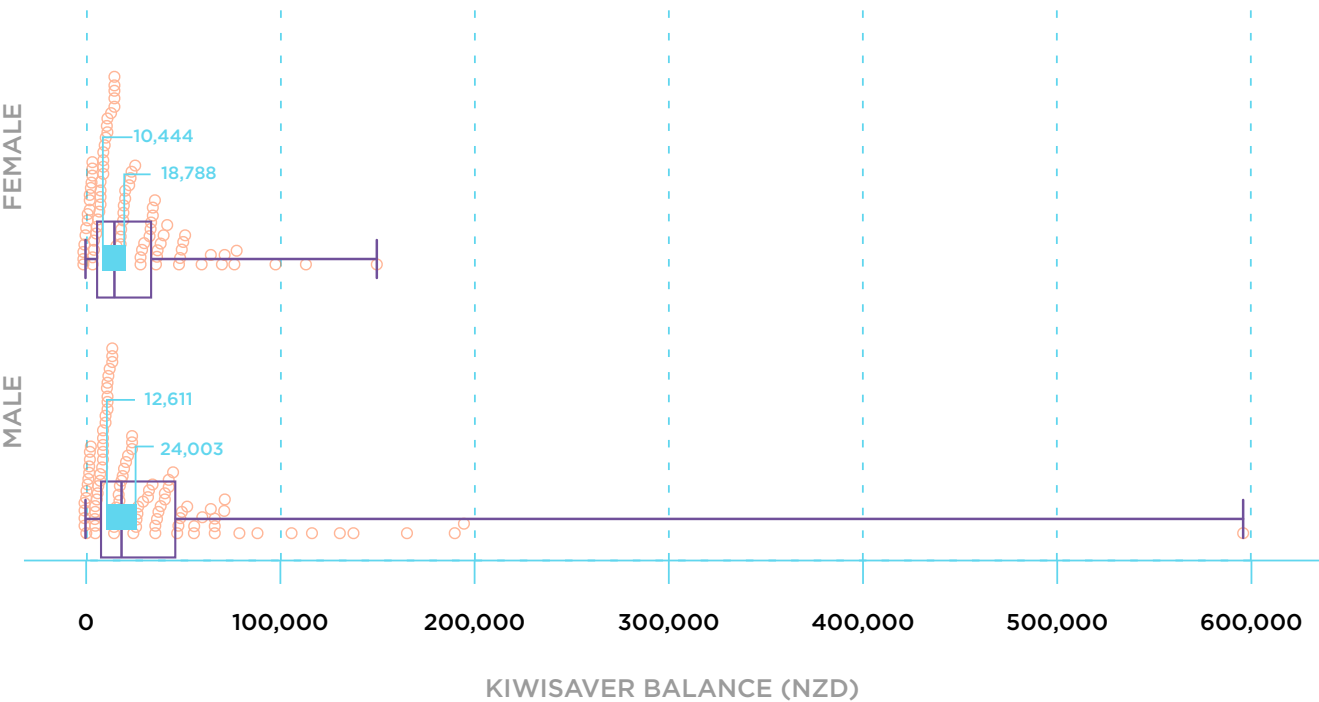
$$\text{upper limit} = \text{median} + 1.5 \times \frac{\text{IQR}}{\sqrt{n}}$$

	Lower limit of confidence interval	Upper limit of confidence interval
Female sample		
Male sample		

Topic 5, Activity 3

Rauhina uses NZGrapher to add informal confidence intervals to her samples of female and male KiwiSaver balances.

KiwiSaver balances by gender



	Lower limit of confidence interval	Upper limit of confidence interval
Female sample	\$10,444	\$18,788
Male sample	\$12,611	\$24,003

a. Interpret the two informal confidence intervals.

b. Use the informal confidence intervals to draw a conclusion based on this sample about whether the median KiwiSaver balance of females in Aotearoa New Zealand is lower than the median KiwiSaver balance of males in Aotearoa New Zealand.



Topic Six:

Writing your report

Topic 6, Activity 1

This extended activity provides an opportunity for you to work your way through the process of using statistical methods to make an inference.

The data set you will be using is the **ANZ Financial Wellbeing Survey 2018 Data set**. Although this data set is itself a sample, for the purpose of this task you can treat it as the population of Aotearoa New Zealand.

This data set contains the following variables:

Variable	Description
Location	Region of Aotearoa that the person lives in
Age	Age in years

Ethnic group	The ethnic group that the participant identifies as. Note that in the original data set, respondents could select more than option.
Gender	Male, female
Savings	The amount of money the person has the in bank (NZ dollars)
KiwiSaver balance	The amount of money the person has in their KiwiSaver fund
Debt	The amount of money the person owes.
Financial wellbeing rating	The rating the person selected to describes their financial wellbeing. The four options are: <ul style="list-style-type: none"> • Struggling • Getting by • Doing okay • No worries.

Use the statistical enquiry cycle outlined below to make an inference about two groups in this population.

Problem

- Familiarise yourself with the ANZ Financial Wellbeing Survey 2018 Data set data set by uploading it to NZGrapher or another statistical tool.
- Choose an area of investigation.
- Explain the purpose of your investigation, for example, why it is useful or interesting.
- Pose a comparative investigative question that can be answered using data provided in the ANZ financial wellbeing data set.

Make sure that your investigative question includes:

- The variable you are going to focus on, including units
- The two groups you are comparing
- The direction of your comparison, for example, investigating whether the median of one group is bigger (or smaller) than the median of another group

- Which parameter (median or mean) you are making an inference about.

Plan and data

Select random samples, and explain your sampling method and sample size.

Analysis

- Use graphs to display your data and provide summary statistics
- Discuss the sample distributions by comparing at least three of their key features, for example, their:
 - shape
 - overlap
 - shift
 - spread
 - middle 50%
 - any unusual or interesting features.

You may like to use the PEEL framework to craft your responses:

P	Point	Make a statement about what you can see in the dot plot or box plot.
E	Explanation	Explain why you think this.
E	Evidence	Use numbers or facts to back up what you are saying.
L	Link	Link your statements to the population, the context, your research or to a prediction you made at the start of your investigation.

Conclusion

Answer your investigative question and explain how it relates to the purpose of your investigation. Include an evaluation of the process you have used to make your inference.

