Year 12 Mathematics

IAS 2.10

Statistical Experiments

Robert Lakeland & Carl Nugent

Contents

nuLake Ltd
Innovative Publisher of Mathematics Texts

NCEA 2 Internal Achievement Standard 2.10 – Statistical Experiments

This achievement standard involves conducting an experiment to investigate a situation using statistical methods.

Achievement	Achievement with Merit	Achievement with Excellence
• Conduct an experiment to investigate a situation using statistical methods.	• Conduct an experiment to investigate a situation using statistical methods, with justification.	• Conduct an experiment to investigate a situation using statistical methods, with statistical insight.

◆ This Achievement Standard is derived from Level 7 of The New Zealand Curriculum and is related to the achievement objective:

❖ carry out investigations of phenomena, using the statistical enquiry cycle

- conducting experiments
- evaluating the choice of measures for variables and data collection methods used
- using relevant contextual knowledge, exploratory data analysis, and statistical inference

in the Statistics strand of the Mathematics and Statistics Learning Area.

◆ Conduct an experiment to investigate a situation using statistical methods involves showing evidence of using each component of the investigation process.

◆ Conduct an experiment to investigate a situation using statistical methods, with justification, involves linking components of the process of investigating a situation by experiment to the context, explaining relevant considerations in the investigation process, and supporting findings with statements which refer to evidence gained from the experiment.

◆ Conduct an experiment to investigate a situation using statistical methods, with statistical insight, involves integrating statistical and contextual knowledge throughout the investigation process which may involve reflecting on the process, or considering other relevant variables.

◆ The process of investigating a situation by experiment involves:

❖ posing an investigative question about a given experimental situation

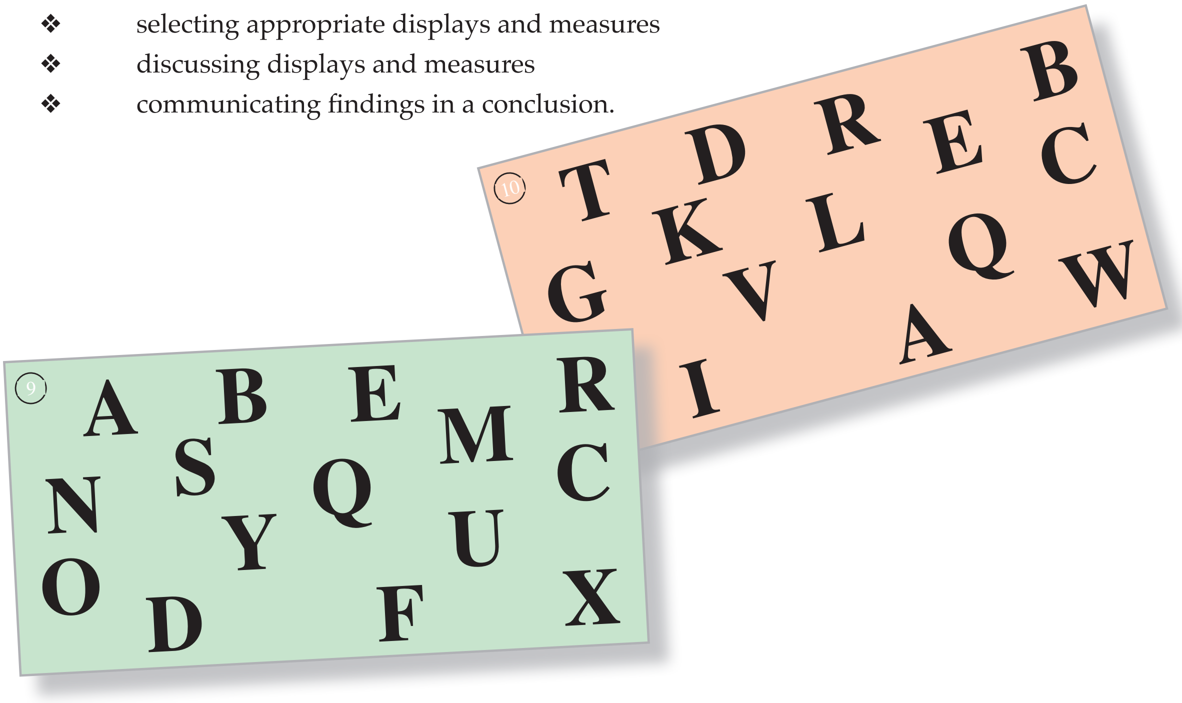
❖ planning the experiment by
- determining appropriate variables and measures
- determining data collection and recording methods

❖ conducting the experiment and collecting data

❖ selecting appropriate displays and measures

❖ discussing displays and measures

❖ communicating findings in a conclusion.



Experimental versus Observational Studies

Types of Statistical Studies

Observational Studies

An observational study is when we sample different subgroups from a population and attempt to infer, from our sample, differences that may exist in the population. By definition, an observational study requires no intervention (the researcher does nothing to change the result) but just observes differences in the samples. An example of an observational study is a survey of the amount of homework completed by male and female students in a school.

It is impossible to conclude a cause and effect from an observational study as differences may be caused by some variable other than the variable identified by the researcher. The two groups the researcher is observing may be different in many respects other than the variable the researcher has decided to study. Any difference in homework between male and female students may be due to rates of participation in sport, for example.

All observational statistics show smokers are more likely to get lung cancer but they do not show smoking causes lung cancer. It could have been that the people who choose to smoke were genetically programmed to get lung cancer and to smoke.

Just because an observational study shows there is a difference does not mean that the variable being researched caused the difference.

Experimental Studies

In an experimental study we start with two groups that are as similar as possible. We then intervene with one group by changing something and see if the results of the two groups vary.

In many experiments, the participants are randomly allocated to the two groups so natural differences between the two groups are evenly distributed. One group will be left unchanged (the control group)

Unfortunately there are many situations where it would be unethical or immoral to do an experimental study.

Since the 1920s researchers have suggested that there was a link between smoking cigarettes and getting lung cancer or heart disease. By observational studies they could show there was a difference in cancer and heart disease rates between smokers and nonsmokers but this did not demonstrate that smoking caused cancer or heart disease. It could be that people who choose to smoke may have been going to get sick whether they smoked or not because of their living conditions or a genetic weakness.

To establish a casual relationship they would have to conduct an experiment for example, by using two random groups of fifteen year old students.

The control would not be allowed to smoke and the experimental group made to smoke for the next 30 years. If they got differences in cancer rates between the two groups then they could conclude that smoking causes cancer.

It would be unethical to do this study due to the risk to the participant's health.

It is considered unethical to conduct an experiment that results in a risk to the health of participants.

Example

A researcher was investigating whether school assessments affect class attendance. She randomly allocated 56 students to a control and a treatment group. She told the students in the treatment group that they would have an important assessment on Friday while she said nothing to the control group.

She recorded the number of periods missed by the students in the day before the assessment.

Find the mean and median of each group.

Periods missed	Control group	Treatment group
0	15	17
1	5	2
2	1	0
3	2	1
4	1	1
5	4	7

We independently calculate the mean and median for each of the two groups.

On the TI-84 Plus press STAT 1 and delete any existing columns.

Move the cursor to column one and enter the data values 0, 1, 2, 3, 4 and 5, then move to column two and enter the corresponding frequencies 15, 5, 1, 2, 1 and 4.

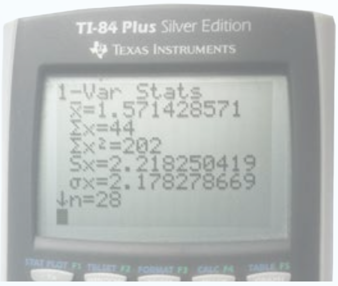
Press STAT and then select CALC by using the arrow keys and select 1 then enter

2ND 1 ▼ 2ND 2 ENTER ENTER

Scroll down to find the mean ($\bar{x}$) and median (Med).

Repeat the procedure for the Treatment group. The results are as follows:

Averages	Control	Treatment
Mean ($\bar{x}$)	1.321	1.571
Median	0	0



Example

A researcher was investigating whether threats of being held in after school could improve a Year 9 class's commitment to working at home. For her experiment the researcher used a test – retest approach. At the start she recorded the 30 student's homework hours (Control group). The teacher then explained that all students who did not pass assessments would be kept in after school for tutorials.

Two weeks later the homework hours were again recorded (Treatment group).

Find the mean and median of each group.

Hours of Hmwk	Control freq.	Treatment freq.
0 –	2	5
0.5 –	6	2
1 –	7	3
1.5 –	11	14
2 –	3	2
2.5 –	1	4

We assume that all the data for each class is at the midpoint of the class and all classes are the same size. For example the 2.5 – class is assumed to be from 2.5 to 3.0 hours with a midpoint of 2.75 hours.

Hours of Hmwk	Control freq.	Treatment freq.
0.25	2	5
0.75	6	2
1.25	7	3
1.75	11	14
2.25	3	2
2.75	1	4

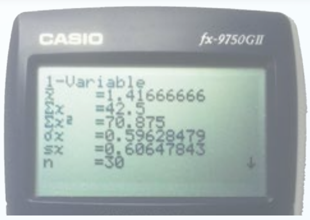
On the Casio 9750GII select 2 from the menu to get to STAT. Then delete any existing data. Move the cursor to column one and enter the data 0.25, 0.75, 1.25, 1.75, 2.25 and 2.75, then move to column two and enter frequencies 2, 6, 7, 11, 3 and 1.

Select F2 (Calc) F6 (Set) and check that 1Var XList: List1 and 1Var Freq: List 2, then press EXE F1.

Scroll down to find the mean ($\bar{x}$) and median.

Repeat for the procedure for the Treatment group The results are as follows:

Averages	Control	Treatment
Mean ($\bar{x}$)	1.417	1.55
Median	1.5	1.75



Achievement – Answer the following questions, using an appropriate statistical graph.

31. Draw a suitable graph to compare the results of this experiment. Comment on what your graph shows you about the experiment.

Control	38, 43, 45, 45, 45, 48, 49, 53, 54, 55, 56, 56, 56, 58, 61, 63, 64, 65, 65, 67, 75, 78
Treat.	45, 45, 46, 56, 56, 62, 63, 65, 67, 67, 68, 68, 72, 75, 75, 76, 78, 78, 78, 81, 84, 85

32. Draw a scatter graph of the time to run 400 metres by hockey players before and after a week with no training. Compare the running times before and after the break.
Data (before, after) in seconds.

(62, 57), (65, 62), (63, 65), (57, 52), (55, 58), (54, 50), (69, 70), (57, 53), (61, 55), (61, 58), (55, 51), (65, 58)

33. Draw a suitable graph to compare the classes missed by the same group of students before and after the introduction of a motivation programme. Compare and comment on the classes missed before and after the motivation programme. Data is in the form (before, after) in missed classes per week.

(0, 5), (3, 4), (11, 8), (0, 0), (18, 12), (24, 11), (5, 0), (17, 9), (6, 2), (7, 2), (11, 4), (15, 9), (3, 8), (7, 8), (13, 6), (0, 2), (21, 7), (14, 5), (7, 1), (14, 11), (7, 2), (9, 0), (11, 4), (13, 9)

37. In an experiment to see if a person's ability to estimate the mass of an object is affected by the density of the object, ten objects were assembled. The five dense objects were stones with masses of 105 g, 175 g, 220 g, 220 g and 395 g. The five less dense objects were off cuts of a light wood, cut so their masses were the same at 105 g, 175 g, 220 g, 220 g and 395 g. The objects were randomly labelled with a letter of the alphabet so estimates could be recorded accurately.

Each of six volunteers were told it was an experiment in correctly estimating the mass of ten different objects. They were not told there were pairs the same mass. They were told no object has a mass less than 80 g or greater than 450 g and were given a 150 g object as a benchmark at the start.

Each subject estimated the mass of the ten objects in a random order. The results are in ordered pairs (stone, wood) for the same mass. E.g. (260, 180) represents estimates of stone and wood each of 395 g mass.

The results are;

(260, 180), (180, 150), (100, 80), (200, 130), (300, 270), (220, 170), (160, 150), (90, 80), (100, 160), (180, 180), (220, 210), (110, 165), (80, 55), (180, 170), (235, 120), (400, 220), (170, 180), (80, 140), (170, 180), (325, 200), (350, 310), (180, 155), (100, 80), (230, 190), (260, 220), (375, 310), (200, 175), (125, 90), (210, 180), (290, 230).

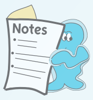
a) Graph the data on a scatter graph with the same scale on the horizontal (stone mass) and vertical (wood mass) scales. Draw in the line y = x on the graph.

b) Why was a scatter graph the best graph for this particular experiment.

c) What conclusions can you draw from this experiment? You need to justify any conclusion by referring to your graph or to statistics that you have calculated.

d) Draw side-by-side box and whisker graphs of the stone and wood results. What additional information does this graph add?

Conducting an Experiment – A Case Study



Posing a Question

If we look at the statistical enquiry cycle we see the first part is labelled the problem. What do you want to investigate with an experiment? Choose a question that you do not already know the answer to. Do some research on the internet or look at ideas in this workbook. Start off with 'I wonder if ...' and refine your question so you have stated the explanatory variable and the response variable.

Example: A Case Study

The Mozart effect is based on research that indicates that listening to Mozart's music may induce a short-term improvement in mental retention or 'spatial-temporal reasoning'.

Consider the statement *"I wonder if listening to music by Mozart can improve a person's memory retention?"*

Such a statement could lead to the investigative question *"Does listening to music by Mozart improve a person's ability to retain information?"*

The case study experiment process is a test-retest on the same group of subjects.

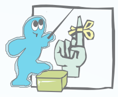
The explanatory variable is whether the person has listened to Mozart music or not for ten minutes prior to a retention test being administered.

The response variable for the experiment will be the retentiveness percentage (the number of correct letters out of the total number of letters in the test) which a person can remember.

The control for this experiment is the results for the number of letters remembered when there is no intervention.

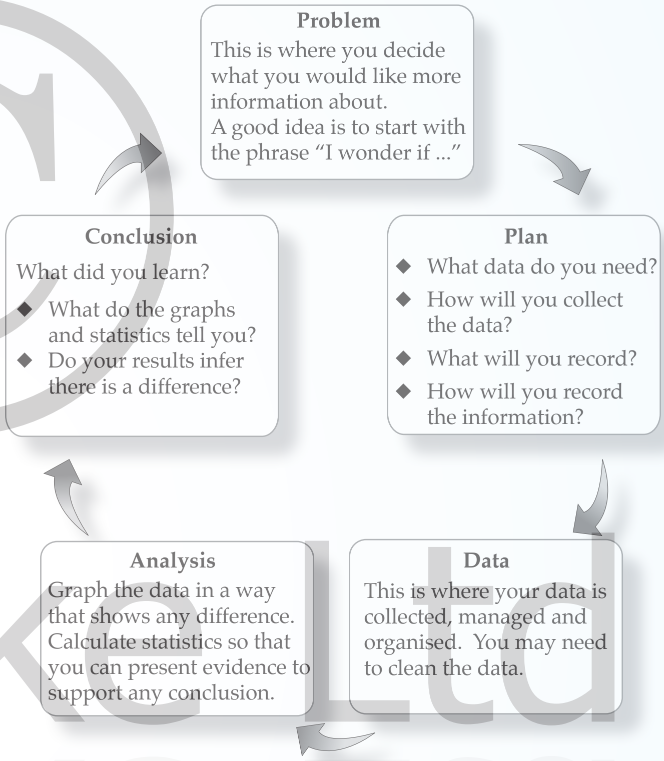
The experimental or treatment results will be the number of letters remembered on the same group of volunteers who sit a similar memory test after listening to Mozart for ten minutes prior to a retention test being administered.

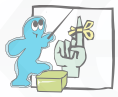
The result of the number of letters will then be compared directly to their result from a similar test at another time where no music is played beforehand.



The Statistical Enquiry Cycle

When statistical researchers want to investigate something they go through five steps called the statistical enquiry cycle. When you have completed one complete cycle often an investigation will identify other problems which means the cycle can start again.





Once you have identified the basis of your statistical investigation and proposed experiment it is a good idea to do some background research.

The internet is a good source. Include some of these background details with applicable references in your final report.

For the Mozart Effect we are investigating we could possibly include the following as background details.

In 1993 Rauscher made the surprising claim that, after listening to a Mozart sonata for ten minutes, normal subjects showed significantly better retention skills. The mean retention rates were higher after listening to the music than not. The enhancing effect did not extend beyond 10 - 15 minutes. These results proved controversial. Some investigators were unable to reproduce the findings.... etc.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1281386/

Card Set Two (after listening to ten minutes of Mozart's music)

Card 1
C B
O J
R L V

Card 2
D F E
J L
W N P

Card 3
B C R
Y L E
O K G

Card 4
I O B A
R W
U E H Z

Card 5
Y A L C
T K F
Z B P R

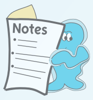
Card 6
N T A O
E M C W
H S R G

Card 7
T S Z N O
C U J P
Q F Y K

Card 8
D O I V M
X T F C
B L A
R Z Q

Card 9
J C F T
B N H M I
W V R
E Z Y

Card 10
Y L U A
P B D K Z
X T O
M J G



The Conclusion

The final phase of the statistical enquiry cycle is the conclusion. The conclusion should relate directly back to the explanatory variable defined in the original question.

It is essential that you give reasons that justify your conclusions using appropriate statistical language.

In addition to the summary statistics you should look at the graphs. You will need to answer the question 'Does the experiment tend to result in a change in the response variable'.

The CensusAtSchool team of Chris Wild, Maxine Pfannkuch, Nicholas Horton and Matt Regan have developed a statistic that gives an estimate of the range of the sample median in which you are likely to find the population median. They call this statistic the 'informal margin of error'. You expect that within one margin of error either side of the sample median you are likely to find the population median.

$$\text{Informal M. of E.} = \pm 1.5\left(\frac{\text{IQR}}{\sqrt{n}}\right)$$

If the IQR = 18.5 and the sample size was 60 then you would expect the median to vary by

$$\text{Informal M. of E.} = \pm 1.5\left(\frac{18.5}{\sqrt{60}}\right)$$
$$= \pm 3.6$$

If you can show that the range for the two medians overlaps then you will not be able to make a call that the experiment resulted in a change in the response variable.

If you have drawn scatter graphs as part of a paired analysis you need to talk about whether the majority of points are now above or below the y = x line (the line of no change). Look at the percentage that have shown an improvement versus the percentage that have shown a reduction.

You may identify possible extremes in the data giving reasons and/or justifying why they are there or whether they should be removed as outliers.

You should discuss the distribution of the data, referring to the range, standard deviation and the quartiles. The data may be skewed or the results all clustered within a certain range.

Students should also highlight any other points or ideas that present themselves as part of their experiment.

It is important to justify all conclusions.
Simply stating that participants tended to do better in the experimental test rather than the control test by 2% without supporting evidence is likely to see you limited to Achievement level. State clearly the evidence that supports and justifies every conclusion.

In the conclusion you should strive to show statistical insight and to link findings to the original question and experimental situation by using supporting statements which relate to key aspects of the experimental process.

One area that you can demonstrate insight is by doing research of experiments similar to the experiment you have completed. If you find information quote the source or website where you found it. You should also list other possible areas of investigation or follow up, that the experiment may lead to.

The quality of your report, including your interpretation, discussion and reasoning about the experimental process as well as the way you tie it into context, will determine your grade.

The difference between an Achievement, Merit or Excellence Assessment

For **Achievement**, you need to conduct an experiment using statistical methods. This involves you:

1. Posing an investigative question
2. Planning the experiment
3. Conducting the experiment and collecting data
4. Selecting appropriate displays and measures
5. Discussing displays and measures
6. Communication of findings in a conclusion which needs to include a 'suggestive inference'.

For **Merit**, you must do all that is required for Achievement with justification.
This could involve the following:

1. Linking components of the experimental design process to the context
2. Explaining relevant considerations in the investigation process such as timing
3. Supporting findings with statements which refer to evidence gained from the experiment.

For **Excellence**, you need to conduct an experiment with statistical insight. This could involve the following:

1. Reflecting on the process and use your contextual knowledge to explain what has been observed in the experiment
2. Considering other relevant or lurking variables and their possible effects
3. Discussing how your findings could be used to improve the experimental design.

Answers

Page 7

1. $\bar{x} = 53.5$ (1 dp)
Median = 50.5

2. $\bar{x} = 4.59$ (2 dp)
Median = 4.58

3. $\bar{x} = 46.8$ (1 dp)
Median = 45.5

4. $\bar{x} = 7.2$ (1 dp)
Median = 7.55

5. Control: $\bar{x} = 1.18$ (2 dp)
Median = 0
Treatment: $\bar{x} = 1.33$ (2 dp)
Median = 1

6. Control: $\bar{x} = 2.10$ kg (2 dp)
Median = 2.1 kg
Treatment: $\bar{x} = 2.12$ kg (2 dp)
Median = 2.1 kg

7. Control: $\bar{x} = 5.25$ kg (2 dp)
Median = 4.86 kg
Treatment: $\bar{x} = 5.11$ kg (2 dp)
Median = 4.83 kg

Page 9

8. LQ = 2.36
UQ = 7.25
IQ range = 4.89
Range = 9.24

9. LQ = 27
UQ = 66
IQ range = 39
Range = 98

10. LQ = 4.1
UQ = 8.5
IQ range = 4.4
Range = 11.4

11. LQ = 192
UQ = 784
IQ range = 592
Range = 857

12. a) LQ = 7, UQ = 8
IQ range = 1, Range = 5
b) LQ = 4.5, UQ = 8
IQ range = 3.5, Range = 5
c) Range in both cases is 5, yet the distribution in Question 12 b) has more data further from the median so the IQR is larger.

Page 10

13. a) Mean = 23.9, Median = 24
LQ = 24, UQ = 25
IQ range = 1, Range = 8
b) Mean = 23.9, Median = 25.5
LQ = 24, UQ = 27
IQ range = 3, Range = 28
c) Mean same but median up 1.5. Range is now 28, while the IQR has increased from 1 to 3. The results are spread more but the IQR is not affected by data at the extremes (results of 0 etc.).

Page 12

14. $\bar{x} = 4.93$ (3 sf)
s = 2.77 (3 sf)

15. $\bar{x} = 48.7$ (3 sf)
s = 29.9 (3 sf)

16. $\bar{x} = 501.4$ (4 sf)
s = 283.2 (4 sf)

17. $\bar{x} = 50.7$ (3 sf)
s = 30.5 (3 sf)

18. a) $\bar{x} = 61.5$ (3 sf), s = 13.1 (3 sf)
b) From 48.4% to 74.6% there are 20 (67%) of the results.

19. a) Group B has the greatest standard deviation.
The data is spread over a greater range (10.93 as opposed to 7).
b) Group A and C have exactly the same spread but each data value in group C is exactly 222 greater than the corresponding value in group A.

Page 13

20. $\bar{x} = 1.2$ (1 dp)
s = 1.8 (1 dp)

21. $\bar{x} = 2.5$ (1 dp)
s = 2.1 (1 dp)

22. $\bar{x} = 39.0$ (1 dp)
s = 3.5 (1 dp)

23. $\bar{x} = 23.0$ (1 dp)
s = 2.0 (1 dp)

24. $\bar{x} = 42.5$ (1 dp)
s = 5.2 (1 dp)

25. $\bar{x} = 70.4$ (1 dp)
s = 19.0 (1 dp)

Page 13 cont...

26. Before
$\bar{x} = 1.9$ days (1 dp)
s = 1.9 days (1 dp)
After
$\bar{x} = 1.4$ days (1 dp)
s = 1.5 days (1 dp)

Despite the number of 0 days dropping the system is reporting fewer absences (reduced by 0.5 days) and less spread of results (Std. Dev. down by 0.4 days).

27. School
$\bar{x} = 2.3$ grades (1 dp)
s = 1.1 grades (1 dp)
Homework
$\bar{x} = 2.3$ grades (1 dp)
s = 1.6 grades (1 dp)
The mean grade was the same but the spread increased (Std. Dev. up by 0.5) implying there were higher and lower grades. Possibly some cheated and some did not do their homework.

Page 14

28. Control $\bar{x} = 5.28$ (2 dp)
s = 1.31 (2 dp)
Exp. Grp. $\bar{x} = 5.35$ (2 dp)
s = 0.82 (2 dp)
The mean has increased slightly as a result of the instruction but the standard deviation has dropped considerably showing less variation of results after the instruction.

29. a) $\bar{x} = 3.03$ people (3 sf)
s = 1.68 people (3 sf)
b) $\bar{x} = 2.43$ people (3 sf)
s = 1.55 people (3 sf)
c) There are typically 3 people in a random NZ house ($\bar{x} = 3.03$), but in their Hutt Valley sample the mean drops by half a person per house ($\bar{x} = 2.43$). There is slightly more variation in the NZ sample as it has a higher standard deviation.

30. $\bar{x} = 0.83$ for both groups but the standard deviation shows the spread has increased from s = 1.6 for the control group to s = 1.8 for the experimental group. There is a greater spread of results even though the means are the same.