3.10 Statistical Inferences for Weaker Students

Vikki Clarke - vclarke@westlake.school.nz

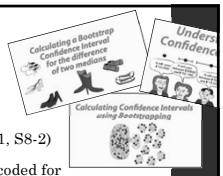
Jake Wills - jwills@westlake.school.nz

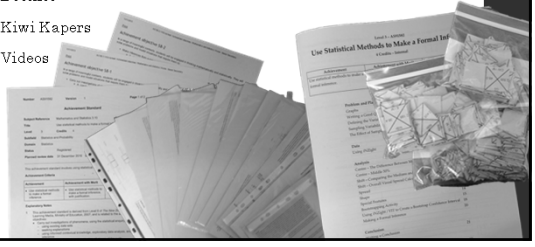
Where do we start?

- Where are the students coming from?
- Where do we need to get to?
- Our cohort
- Our assessment methods

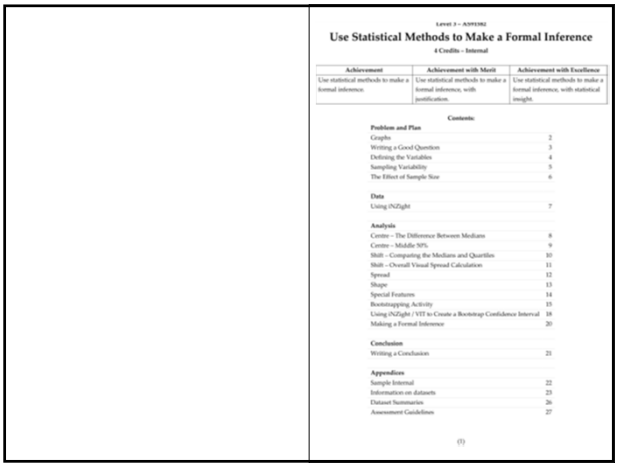
Resources

- Standard and Achievement Objectives (S8-1, S8-2)
- Exemplars (incl. High Excellence) – colour coded for easy reference
- Booklet
- Kiwi Kapers
- Videos





Week	Monday	Tuesday	Wednesday	Thursday	Friday
1	Topic Overview	Recap Level 1 and Level 2 Inference	Recap Level 1 and Level 2 Inference	Looking at Exemplars	Looking at Exemplars
2	Problem and Plan	Writing a Good Question	Defining the Variables	Sampling Variability	The Effect of Sample Size
3	Data – Using iNZight	Centre – The Difference Between Medians, Middle 50%	Shift – Comparing the Medians and Quartiles, Overall Visual Spread Calculation	Shift, Spread, Shape and Special Features	Bootstrapping Activity **INTERNAL ISSUED**
4	Using iNZight / VIT to Create a Bootstrap Confidence Interval	Making a Formal Inference and Writing a Conclusion	Working on Internal, Teacher Checking	Working on Internal, Teacher Checking	Working on Internal, Teacher Checking
5	**INTERNAL DUE IN**				



Level 3 – AS91582

Use Statistical Methods to Make a Formal Inference

4 Credits – Internal

Achievement	Achievement with Merit	Achievement with Excellence
Use statistical methods to make a formal inference.	Use statistical methods to make a formal inference, with justification.	Use statistical methods to make a formal inference, with statistical insight.

Contents:

Problem and Plan
Graphs 2
Writing a Good Question 3
Defining the Variables 4
Sampling Variability 5
The Effect of Sample Size 6

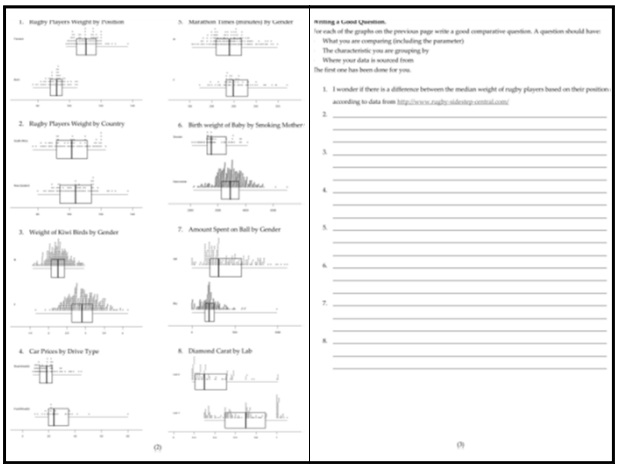
Data
Using iNZight 7

Analysis
Centre – The Difference Between Medians 8
Centre – Middle 50% 9
Shift - Comparing the Medians and Quartiles 10
Shift – Overall Visual Spread Calculation 11
Spread 12
Shape 13
Special Features 14
Bootstrapping Activity 15
Using iNZight / VIT to Create a Bootstrap Confidence Interval 18
Making a Formal Inference 20

Conclusion
Writing a Conclusion 21

Appendices
Sample Internal 22
Information on datasets 23
Dataset Summaries 26
Assessment Guidelines 27

(1)



1. Rugby Players Weight by Position
2. Rugby Players Weight by Country
3. Weight of Kiwi Birds by Gender
4. Car Prices by Drive Type
5. Marathon Times (minutes) by Gender
6. Birth weight of Baby by Smoking Mother
7. Amount Spent on Ball by Gender
8. Diamond Carat by Lab

(2)

Writing a Good Question.

For each of the graphs on the previous page write a good comparative question. A question should have:
- What you are comparing (including the parameter)
- The characteristic you are grouping by
- Where your data is sourced from

The first one has been done for you.

1. I wonder if there is a difference between the median weight of rugby players based on their position according to data from http://www.rugby-sidestep-central.com/
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____

(3)

Analysis
We now start on the Analysis section of our report. This section can be abbreviated to CSI.
The C stands for Centre, then there are 4 S's, Shift, Spread, Shape and Special Features. I stands for inference.

Centre – The Difference Between Medians
We now need to state what the difference between the medians is. This is calculated by subtracting one median from the other.
Again the first one has been done for you.

1. The forwards' median weight is 18.50 kg higher than the backs' median weight.
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____

(8)

Centre – Middle 50%
The centre is looking at what is happening with the middle 50% of the data, which is between the lower quartile (1^{st} Q_1) and the upper quartile (3^{rd} Q_3).

Discuss the centre for each of the sets of data, the first one has been done for you.

1. The middle 50% of the forward's weights are between 104.8 kg and 117.0 kg whereas the middle 50% of the back's weights are between 88.0 kg and 95.5 kg.
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____

(9)

Shift – Comparing the Medians and Quartiles
With the shift we need to look at what parts of the box and whisker graphs overlap, and which parts are shifted along. You need to consider where the median and upper / lower quartiles are for the two groups of beta.

Compare the medians and quartiles for each of the sets of data, the first one has been done for you.

1. The lower quartile for the forwards weight is higher than the upper quartile of the weight of the backs.
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____

(10)

Shift – Overall Visual Spread Calculation
You also need to consider the difference in the medians (which we calculated earlier) in relation to the overall visual spread (the highest upper quartile minus the lowest lower quartile).

The calculation that you need to do is $\frac{\text{difference between medians}}{\text{overall visual spread}}$ to tell you how significant the difference is. In the example we have been working through this would be $\frac{18.5}{117.0-88.0} = 0.638$. The closer this number is to one the more significant the difference is.

There is a significant difference between the samples if the number is bigger than

Sample Size	Calculation Bigger Than
30	0.33
100	0.20
1000	0.10

Always use the smaller sample size when making the call.

Discuss the shift for each of the sets of data, the first one has been done for you.

1. The difference between the medians is 18.5 kg which is 0.638 of the overall visual spread which is a significant difference.
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____

(11)

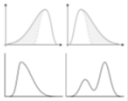
Spread
To calculate the spread we normally look at the inter-quartile range (IQR) for the two data sets. The IQR is calculated by subtracting the lower quartile off the upper quartile. You can also look at the standard deviation for each of the two data sets. You should also comment on what you see visually.

Discuss the shape for each of the sets of data, the first one has been done for you.

1. The inter quartile range for the forwards is 12.2 kg whereas the interquartile range or the backs is 7.5 kg, indicating that the forwards have more variation in their weights than the backs. The standard deviation is also higher for the forwards. Overall visually the forwards seem to be slightly more spread out than the backs.
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____

(12)

Shape
In the shape we need to look at two things… the skew and the modality.
If the distribution has a long tail to the left, it is skewed to the *left* (like left diagram).
If it has a long tail to the right it is skewed to the *right* (like right diagram).



We also need to say if there is one mode (unimodal, left diagram) or two modes (bimodal, right diagram).

Discuss the shape for each of the sets of data, the first one has been done for you.

1. The forwards weights appear to be skewed to the right whereas the backs weights seem reasonably symmetrical. The backs appear to be unimodal whereas the forwards are potentially bimodal.
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____

(13)

Special Features

We also need to discuss any unusual features that we notice with the data sets. This could be an extreme value (a point with a much higher value than the others) or anything else that you notice. It is good to give a possible explanation for anything you notice. Going back to the original data set to find out more information about the data point is often useful as well.

Discuss the unusual for each of the sets of data, the first one has been done for you.

1. Looking at the graphs I can see that the forwards have one player that weighs more than most of the other forwards. He is a New Zealander weighing 137 kg and is 1.81 m tall. This could be because he is a stockier player that is quite large with more muscles causing him to weigh more.
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____

(14)

Bootstrapping Activity

Below is all of the rugby players from Data Set 1. You will need to cut them all out in order to do the activity on page 17.

Back 82	Back 84	Back 88	Forward 119	Forward 102
Back 84	Back 90	Back 96	Forward 100	Forward 120
Back 93	Back 99	Back 79	Forward 109	Forward 98
Back 93	Back 83	Back 97	Forward 114	Forward 115
Back 105	Back 87	Back 101	Forward 115	Forward 99
Back 82	Back 88	Forward 116	Forward 117	Forward 100
Back 93	Back 85	Forward 120	Forward 105	Forward 103
Back 89	Back 93	Forward 102	Forward 108	Forward 110
Back 90	Back 96	Forward 110	Forward 107	Forward 115
Back 85	Back 105	Forward 137	Forward 111	Forward 103
Back 101	Back 89	Forward 102	Forward 117	Forward 115
Back 89	Back 92	Forward 112	Forward 118	Forward 124
Back 94	Back 93	Forward 103	Forward 102	Forward 110
Back 85	Back 95	Forward 123	Forward 103	Forward 116
Back 87	Back 97	Forward 114	Forward 107	Forward 99
Back 93	Back 92	Forward 115	Forward 117	Forward 101
Back 88	Back 92	Forward 116	Forward 107	Forward 110
Back 89	Back 94	Forward 118	Forward 113	Forward 110
Back 100	Back 77	Forward 125	Forward 106	Forward 106
Back 104	Back 92	Forward 102	Forward 113	Forward 106
Back 92	Back 87	Forward 120	Forward 101	Forward 112
Back 92	Back 96	Forward 101	Forward 108	Forward 114
Back 94	Back 89	Forward 104	Forward 106	Forward 114
Back 95	Back 91	Forward 107	Forward 115	Forward 117
Back 97	Back 94	Forward 109	Forward 104	Forward 120
Back 104	Back 93	Forward 118	Forward 110	Forward 119
Back 80	Back 99	Forward 127	Forward 129	Forward 120

(15)

This page has been deliberately left blank (as you are cutting out the other side!

(16)

Bootstrapping Activity

bootstrapping is sampling from the sample with replacement. It normally involves sampling until you have the same number as in your original sample, but for the sake of this activity when we are doing it manually we are just going to take samples of 30 in total, which means we may end up with different numbers of forwards and backs.

Record the weights of the forwards and backs below (you won't end up filling up the whole table), and then use your calculator to work out the median for the forwards and the backs from the bootstrap, and find the difference between the two.

Bootstrap 1		Bootstrap 2		Bootstrap 3		Bootstrap 4	
Forwards	Backs	Forwards	Backs	Forwards	Backs	Forwards	Backs
Med:	Med:	Med:	Med:	Med:	Med:	Med:	Med:
Difference		Difference		Difference		Difference	

Plot the differences from both your bootstraps, as well as the bootstraps from your class as a dot plot on the axis below.

This gives us a fairly good idea of how accurate our samples are going to be, and if there is going to be a difference between the two groups (in this case the forwards' and the backs' weights). It is a very tedious process though, so we normally will us a computer to speed it up.

(17)

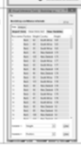
Using iNZight / VIT to Create a Bootstrap Confidence Interval

The next thing we need to do is to create a bootstrap distribution. To do this we need to load the bootstrap confidence interval module of VIT.

Select as circled and click on the 'Run selected VIT module' button at the bottom of the window.

You will need to import the data again, and once imported choose the variables. Variable one should be weight and variable 2 should be the position.

This should give you a window that looks like the one on the right.



The next step is to click on the 'Analyse' tab.

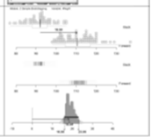
(18)

You need to change the Quantity to 'median' and then click record my choices.

Then click in the bottom section on 1000 repetitions and then click go, as shown to the right.

Once done you need to click on 'Show CI' to get the confidence interval shown on the graph.

This gives the output shown to the right, which tells us the difference between the medians is 18.5Kg, but that we can be reasonably confident that forwards will be between 16kg and 23kgs on average heavier than the backs.



Now it is your turn. For each dataset you need to produce the bootstrap confidence interval… don't forget to press the show CI button and write down the confidence intervals so you can refer back to them later.

1. 16 kg to 23 kg
2. _____ to _____
3. _____ to _____
4. _____ to _____
5. _____ to _____
6. _____ to _____
7. _____ to _____
8. _____ to _____

(19)

Making a Formal Inference

We now come to the most important part of the internal, where we have been leading up to the whole time, making a formal inference. This is about linking it back to the population that we care about. To get the interval we look at the bootstrap distributions that we produced earlier.

Make a formal inference for each of the sets of data, the first one has been done for you.

1. From the bootstrapping confidence interval I can be reasonably confident that forwards will weigh between 16.0 kg and 23.0 kg more than backs on average.
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____

Writing a Conclusion

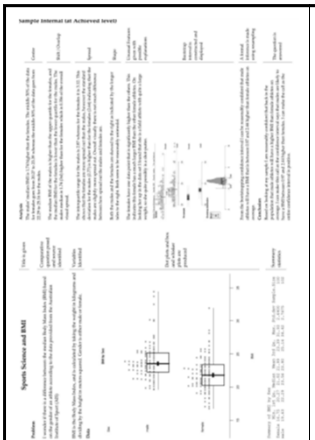
We also need to make a conclusion that summarises what we have found so far. We need to say what the call is that we are making and why we can make the call (or if we can't make the call). We can only make he call if the entire interval is positive or the entire interval is negative, as if zero is in the interval then here might be a difference of zero or the difference might be the other way round.

Make a conclusion for each of the sets of data, the first one has been done for you.

1. Based on looking at my sample I am reasonably confident that back in the population that forwards will weigh more than backs on average. I can make this call as the confidence interval says that forwards are likely to weigh between 16.0 kg and 23.0 kg more than backs. I can make the call as the entire confidence interval is positive.
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____

Congratulations, you now have written up a report for 7 different sets of data, so you now should be able o write up your own internal. Don't forget to give your report a title.

Sample Internal (at Achieved level)



Sports Science and BMI

Data Set Information:

babies

Various health measures on new born babies and their mothers can give an indication of the future health # the infant. In particular, low birth weight is known to be associated with increased morbidity and poor wealth outcomes.

Data is routinely collected by all birthing centres in New Zealand concerning various health measures of nothers and their new born babies. A random sample of 550 records was selected in 2011 by a team of nedical researchers from a birthing centre in a large teaching hospital.

You have been supplied with the dataset containing some of the variables for the random sample collected n 2011.

Variable	Description
bloodsugar	GDM = mother has gestational diabetes Normal = mother has normal blood sugar levels
smoking	Smoker = mother smoked during pregnancy NonSmoker = mother was a non-smoker during pregnancy
neonatalsexgroup	Male = new born infant is male Female = new born infant is female
birthweight	Weight of infant at birth (in grams)
gestationalage	Length of pregnancy (in weeks)
fastingBloodglucose	Results from a routine blood test during pregnancy (mmol/L)

SuitWear

Data was recorded of students going to the school ball in 2012 as to how much they spent on their clothing and accessories.

Variable	Description
Gender	Boy = new student is male Girl = new student is female
Amount.spent	The amount spent on clothing and accessories in New Zealand Dollars.

Cars

With rising costs of owning and running a car, and environmental awareness, buyers are becoming more conscious of the features when purchasing new cars. The data supplied is for new vehicles sold in America in 1993.

Hospital details have not been provided to ensure privacy of all research participants.

Variable	Description
Vehicle Name	
Origin	Country of manufacture • America • Foreign
Price	US $1000
Type	Small, midsize, large, compact, sporty, van
City	Fuel efficiency in litres per 100km in cities and on motorways
Open	Fuel efficiency in litres per 100km on country open road
Drive Train	Front Wheel Drive Rear Wheel Drive
Engine Size	Size in litres
Manual Transmission	Yes No
Weight	Weight of car in Kg

Diamonds

Every diamond is unique, and there are a variety of factors which affect the price of a diamond. Insurance companies in particular are concerned that stones are valued correctly.

Data on 308 round diamond stones was collected from a Singapore based retailer of diamond jewellery, who had the stones valued.

Variable	Description
Carat	Weight of diamond stones in carat units 1 carat =0.2 grams
Colour	Numerical value given for quality of colour ranging from 1=colourless to 6=near colourless
Clarity	Average = score 1, 2 or 3 Above average = score 4, 5 or 6
Lab	Laboratory that tested & valued the diamond 1 = laboratory 1 2 = laboratory 2
Price	Price in US dollars

Kiwi

Data on kiwi birds around New Zealand was collected in order to help with conservation efforts.

Variable	Description
Species	GS-Great Spotted NBS-Northland Brown Tok-Southern Tokoeka
Gender	M-Male F-female
Weight(kg)	The weight of the kiwi bird in kg
Height(cm)	The height of the kiwi bird in cm
Location	NWN-North West Nelson CW-Central Westland EC-Eastern Canterbury SS-Stewart Island NF-North Fiordland SF-South Fiordland N-Northland E-East North Island W-West North Island

Marathon

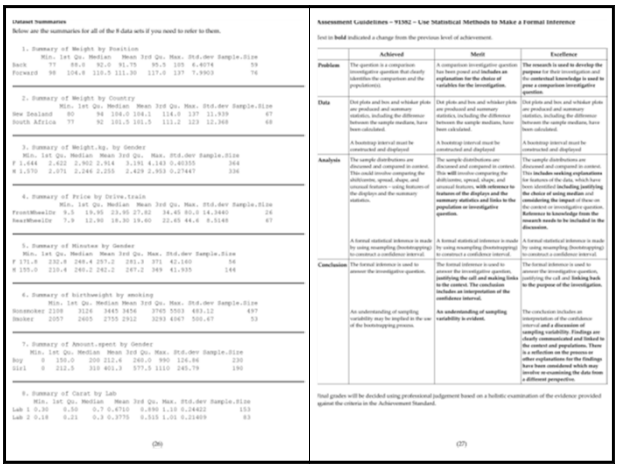
The data is a sample taken from marathons in NZ.
It is a simple random sample of 200 athletes.

Variable	Description
Minutes	How many minutes they completed the marathon in
Gender	Male (M) or Female (F)
AgeGroup	Younger (under 40) or older (over 40)
StrideLengthCM	The persons average stride length over the marathon in cm.

Rugby

The data is real data and comes from http://www.rugby-sidestep-central.com/

Variable	Description
Country	New Zealand or South Africa
Position	Forward or Back
Weight	The weight of the player in kilograms (kg)
Height	The height of the player in metres (m)



Dataset Summaries

Below are the summaries for all of the 8 data sets if you need to refer to them.

1. Summary of Weight by Position

```
     Min. 1st Qu. Median   Mean 3rd Qu. Max. Std.dev Sample.Size
Back      77   88.0   92.0  91.75   95.5  105  6.4074          59
Forward   98  104.8  110.5 111.30  117.0  137  7.9903          76
```

2. Summary of Weight by Country

```
              Min. 1st Qu. Median  Mean 3rd Qu. Max. Std.dev Sample.Size
New Zealand    80      94  104.0 104.1   114.0  137  11.939          67
South Africa   77      92  101.5 101.5   111.2  123  12.368          68
```

3. Summary of Weight.kg. by Gender

```
  Min. 1st Qu. Median  Mean 3rd Qu.  Max. Std.dev Sample.Size
F 1.444   2.422  2.902 2.914   3.191 4.143 0.40355         364
M 1.570   2.071  2.246 2.255   2.429 2.953 0.27447         336
```

4. Summary of Price by Drive.train

```
             Min. 1st Qu. Median  Mean 3rd Qu. Max. Std.dev Sample.Size
FrontWheelDr  9.5   10.95  23.95 27.82  34.45 60.0 14.3440          26
RearWheelDr   7.9   12.90  18.30 19.80  22.65 44.6  8.3149          67
```

5. Summary of Minutes by Gender

```
  Min. 1st Qu. Median  Mean 3rd Qu. Max. Std.dev Sample.Size
F 171.8   232.8  249.4 257.2   281.3  371  42.160          56
M 155.0   210.4  240.2 242.2   267.2  369  41.935         144
```

6. Summary of birthweight by smoking

```
           Min. 1st Qu. Median Mean 3rd Qu. Max. Std.dev Sample.Size
Nonsmoker 2108    3126   3445 3456    3765 5553  483.12         497
Smoker    2057    2605   2755 2912    3293 4067  500.47          53
```

7. Summary of Amount.spent by Gender

```
     Min. 1st Qu. Median  Mean 3rd Qu. Max. Std.dev Sample.Size
Boy     0   150.0  200.0 212.6   240.0  990  126.86         230
Girl    0   212.5  310.0 401.3   577.5 1110  245.79         190
```

8. Summary of Carat by Lab

```
      Min. 1st Qu. Median   Mean 3rd Qu. Max. Std.dev Sample.Size
Lab 1 0.30    0.50   0.7 0.6711   0.890 1.10 0.24422         153
Lab 2 0.18    0.21   0.3 0.3775   0.515 1.01 0.21409          83
```

(36)

Assessment Guidelines – 91592 – Use Statistical Methods to Make a Formal Inference

Text in **bold** indicated a change from the previous level of achievement.

	Achieved	Merit	Excellence
Problem	The question is a comparison investigative question that clearly identifies the comparison and the population(s).	A comparison investigative question has been posed and **includes an explanation for the choice of variables for the investigation.**	**The research is used to develop the purpose** for their investigation and the **contextual knowledge is used to pose a comparison investigative question.**
Data	Dot plots and box and whisker plots are produced and summary statistics, including the difference between the sample medians, have been calculated. A bootstrap interval must be constructed and displayed	Dot plots and box and whisker plots are produced and summary statistics, including the difference between the sample medians, have been calculated. A bootstrap interval must be constructed and displayed	Dot plots and box and whisker plots are produced and summary statistics, including the difference between the sample medians, have been calculated. A bootstrap interval must be constructed and displayed
Analysis	The sample distributions are discussed and compared in context. This could involve comparing the shift/centre, spread, shape, and unusual features – using features of the displays and the summary statistics. A formal statistical inference is made by using resampling (bootstrapping) to construct a confidence interval.	The sample distributions are discussed and compared in context. This **will** involve comparing the shift/centre, spread, shape, and unusual features, **with reference to features of the displays and the summary statistics and links to the population or investigative question.** A formal statistical inference is made by using resampling (bootstrapping) to construct a confidence interval.	The sample distributions are discussed and compared in context. This **includes seeking explanations** for features of the data, which have been identified **including justifying the choice of using median** and **considering the impact** of these on the context or investigative question. **Reference to knowledge from the research needs to be included in the discussion.** A formal statistical inference is made by using resampling (bootstrapping) to construct a confidence interval.
Conclusion	The formal inference is used to answer the investigative question. An understanding of sampling variability may be implied in the use of the bootstrapping process.	The formal inference is used to answer the investigative question, **justifying the call and making links to the context. The conclusion includes an interpretation of the confidence interval.** **An understanding of sampling variability is evident.**	The formal inference is used to answer the investigative question, justifying the call and **linking back to the purpose of the investigation.** The conclusion includes an interpretation of the confidence interval **and a discussion of sampling variability. Findings are clearly communicated and linked to the context and populations. There is a reflection on the process or other explanations for the findings have been considered which may involve re-examining the data from a different perspective.**

Final grades will be decided using professional judgment based on a holistic examination of the evidence provided against the criteria in the Achievement Standard.

(37)

Want Resources? Email us

We believe in sharing… we're all in this to help the kids do the best that they can.

jwills@westlake.school.nz