Tiny modification made on p. 8, 28/3/2013

Statistics Levels 5-7 *(updated 31 July 2009)*
A teacher's guide to informal comparative reasoning

Maxine Pfannkuch (m.pfannkuch@auckland.ac.nz), Chris Wild (c.wild@auckland.ac.nz), Nicholas Horton (nhorton@email.smith.edu), Matt Regan (m.regan@auckland.ac.nz)

This document is a **guide for teachers**. It is **not for students**. It is **not a teaching activity**; other complementary documents will provide ideas for teaching activities.

Future documents will build on this guide for the higher Levels 6, 7, and 8.

This guide details **the thinking processes** that students might activate when using box plots, dot plots and summary statistics to compare two samples.

Students will build up these thinking processes over a number of years. Starting at Level 3 in the curriculum, students describe features of dot plots, and at Level 4 and the first year of Level 5 students make descriptive statements about two samples through visually comparing dot plots. Therefore, the reasoning elements described in the Analysis section of this document (Overall visual comparisons; Shift and overlap; Summary; Spread; Shape, Individual, and Gaps/Clusters) should already be embedded into students' thinking to some extent.

With the introduction of box plots in the second year of Level 5, specific summary statistics and measures of spread, as well as the sampling reasoning element will be new. The concept of making inferences about populations from samples will be introduced. Hence the difference between ***describing sample features*** and ***having inferential thoughts*** will need to be highlighted.

We need to make sure that students are clear as to whether they are *writing descriptively* about the samples or whether they are *wondering what is happening back in the populations*. Careful language usage is needed to highlight this distinction, e.g., when we write "***I notice*** ..." we are **describing what we see in the samples** and when we write "***I wonder*** ..." we are having thoughts about what is happening back in the population(s), i.e., **having inferential thoughts.**

You will note that the language at times may appear awkward, e.g., the repeated use of "*back in the two populations*". The reasons for this usage are: to *stress repetitively* that we are reasoning about two populations, i.e., reasoning inferentially, rather than describing what is happening in the samples; and, to cement in the *concepts* and *imagery* of comparing two population distributions. The 'awkwardness' should attract students' attention and make them remember the phrase which in turn will repeatedly trigger these ideas, concepts and imagery.

The goals of this teacher guide:

1. To give guidance as to what to look for and what to comment on when comparing two samples
2. To emphasise the importance of differentiating sample statements from population statements
3. To give a clear message that *point estimates alone* are *not* sufficient for making calls about what is happening back in the populations
4. To illustrate thinking processes involved and to provide teaching tips when comparing two samples

This guide uses a specific example, comparison of boys' and girls' foot lengths, and gives an exemplar student report together with a commentary. **Please note** that the student report is to a level of sophistication ***far higher*** than any student would be able to emulate but it presents an ideal towards which teachers should aspire to raise their students.

Feedback welcomed: m.pfannkuch@auckland.ac.nz

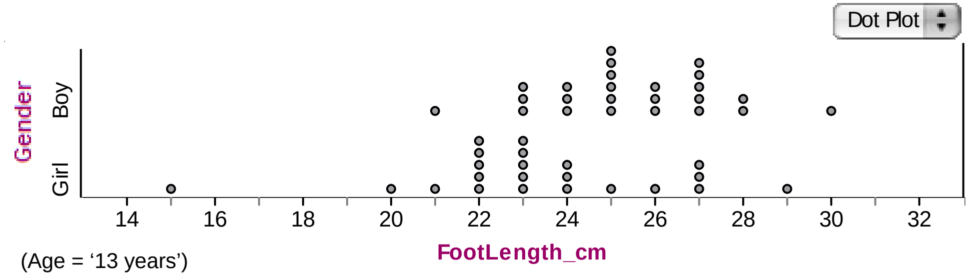
Descriptions and inferences: comparing two samples (Level 5, Year 10)	Teacher Notes
PROBLEM **Background** There is a general belief that 13 year-old boys are bigger than 13 year-old girls. I wish to check out the claim that, for 13 year old students in New Zealand, the boys are bigger than the girls.	
Question Do 13 year-old New Zealand boys have bigger feet than 13 year-old New Zealand girls?	**Teaching Tip:** Discuss the different meanings of 'Population' with students. For example: New Zealand's population is over 4 million; the population of 13 year-old boys; the population distribution for foot lengths of 13 year-old boys; etc. Our question is about **comparing** the **two distributions from two populations** (which are sub-populations of the 13-year-old New Zealand population). Using '**two populations**' rather than 'the population' helps to portray and preserve the image of comparing two population distributions. A difficulty with this question as stated is that it could be interpreted literally to mean: "Do *all* boys have bigger feet than *all* girls?"
Investigative Question Do right foot lengths for 13 year-old NZ boys tend to be bigger than right foot lengths for 13 year-old NZ girls?	The initial question has now been refined and is now referred to as our **investigative question**. The question is structured with the key element, right foot lengths, at the very front of the question. This sentence structure helps to readily invoke an image of two distributions of right foot length measures, one for 13 year-old NZ boys and one for 13 year-old NZ girls, both drawn against the same scale which presumably would be in centimetres. **Technical Note:** Here, the word '*tend*' means that in a plot of the two population distributions the boys' right foot lengths are ***shifted*** further up the scale than the girls'. • This does **not** necessarily mean complete separation of the two population distributions, i.e., it does not imply that, back in the two populations, all the boys' foot lengths are greater than all the girls'. • It does imply that, back in the two populations, the mean/median foot length for the boys will be greater than the girls'. '*tend*' can also be used to describe or compare **sample** features, e.g., "In the samples, the boys tend to have higher values than the girls." meaning that the boys' **sample** distribution is shifted further up the scale than the girls' **sample** distribution.

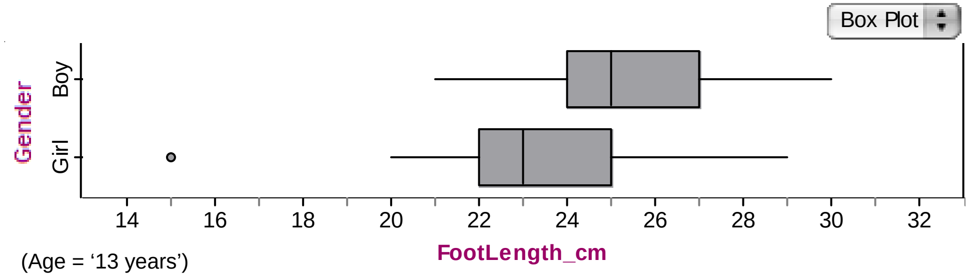
<table>
<tr><th>Descriptions and inferences:
comparing two samples (Level 5, Year 10)</th><th>Teacher Notes</th></tr>
<tr><td></td><td>Teaching Tip:
Students should be asked to sketch the distributional shape and relative location of foot lengths for the populations of 13 year-old boys and girls.
The purpose of getting students to do this sketch before they see any plots is to assist them in:
understanding and clarifying the investigative question
gaining a sense and image of distribution
predicting an appropriate range of foot length values
detecting whether the observed data patterns depart from what is expected – a departure from what is expected leads to further exploration of the data to find out possible reasons
From a teaching perspective, student misconceptions can be revealed. For example, a student draws two distributions with the same shape and centre, but with the boys' hump higher; she says that the higher hump indicates that boys tend to have a greater foot length than girls. Thus confusion between the height of a distribution and the location of a distribution can be ascertained.</td></tr>
<tr><td>PLAN
I will get our two random samples using the CensusAtSchool random sampler.
Take a random sample of 25 boys from the population of 13 year-old NZ boys in the CensusAtSchool database.
Take a random sample of 25 girls from the population of 13 year-old NZ girls in the CensusAtSchool database.
Use the responses on foot length.</td><td>CensusAtSchool website: http://www.censusatschool.org.nz
We assume that random samples taken from the CensusAtSchool data base are two random samples taken from the population of 13-year-old New Zealand boys and the population of 13-year-old New Zealand girls.
Teaching Tip:
Discuss with students the 'reasonableness' of this assumption
Take care when using the phrase a representative sample
Technical Note:
The goal in a sampling process is to obtain a sample to represent the population of interest. In common language usage, a sample is representative of the population if characteristics in the sample are a reflection of those in the parent population. Under this meaning, a truly representative sample almost never exists.
In statistical jargon a representative sample means that the sampling process produces samples in which there is no tendency for certain characteristics to differ from those in the population in some systematic manner, e.g., all random samples could be viewed as representative samples.
Sometimes representative sample is used as jargon simply to signal that some form of stratification has been used in the sampling process.</td></tr>
</table>

Descriptions and inferences: comparing two samples (Level 5, Year 10)	Teacher Notes
DATA Managed through CensusAtSchool survey team. The data came from the student responses to the following 2005 CensusAtSchool survey questions: **1.** *Are you:* ◉ *male* ○ *female* **2.** *How old are you?* ☐ *years* **3.** *What is the length of your right foot?* ☐ *cm* In the random sample of 25 boys, one boy did not record a foot length leaving 24 recorded foot lengths, i.e., there was one missing value. In the random sample of 25 girls, there were three missing values leaving 22 recorded foot lengths. ***I worry:*** • about the quality of the foot length data since students measured and recorded their own foot lengths o Were measurements made with shoes on or shoes off? o Would all students have seen 'cm' to the right of the entry box? o To what level of precision did the students make their measurement? o Why were there missing values?	**Teaching Tip:** Get students to answer the relevant survey question(s) for themselves, in particular measure their own right foot length. Discuss any problems or issues with the questions. Answering the questions themselves will prompt any worries or difficulties that they may foresee with the questions. In the 2007 survey, the right foot length question was modified to: *What is the length of your right foot, without a shoe? Answer to the nearest centimetre.* ☐ *cm* **Technical Note:** It is not necessary to have equal sample sizes.

Descriptions and inferences: comparing two samples (Level 5, Year 10)	Teacher Notes

ANALYSIS





FootLength_cm	Gender	
	Boy	Girl
S1 = Min	21	15
S2 = LQ	24	22
S3 = Median	25	23
S4 = UQ	27	25
S5 = Max	30	29
S6 = Count	24	22
S7 = Mean	25.5	23.4

Figure 1: Fathom output,13 year-old boys’ & girls’ foot lengths

Overall Analysis Strategy:

We need to ***describe*** the features we see in the data.

Starting point:

- Overall visual non-numerical comparisons
  - Overlap
  - Shift
  - Unusual features

After the initial overall visual non-numerical comparisons:

- Make more detailed comparative descriptions of the features including use of summary statistics and specific observation values where appropriate
- Reflect and perhaps comment on some of the features using “*I wonder . . .*” and “*I expect . . .*” type statements, i.e., comment on any inferential thoughts

Comparisons are made:

- ***Between*** the groups (e.g., overlap, shift, spread and shape statements)
- ***Within*** each group (e.g., unusual observations)

Teaching Tips:

- Invite each student to add their own right foot length to their dot plot
  This helps to promote the idea that the group is made up of individuals.
- Keep the emphasis on the ***visual*** rather than reading off values

Key principle for plots (to facilitate comparisons):

- Each group must be plotted on the ***same*** scale.

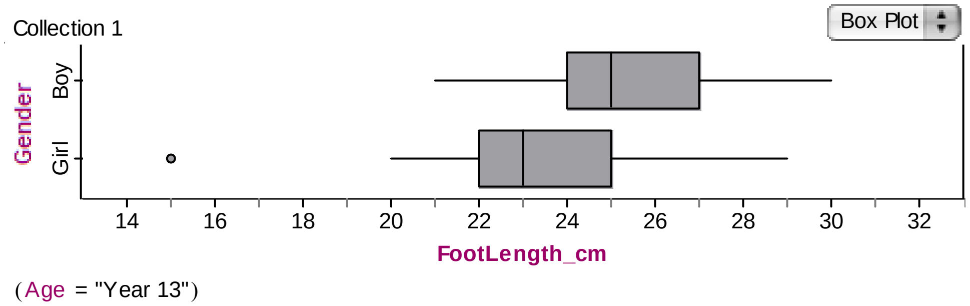
Descriptions and inferences: comparing two samples (Level 5, Year 10)	Teacher Notes
	Technical Note: Dot plots: • preserve the idea that the group is made up of individuals • keep the idea that we are talking about data to the forefront • show the shape of the distribution (e.g., modes, skewness, subgroups) of the data • suggest the shape of the underlying (population) distribution Box plots: • allow for quick visual comparisons • allow for approximate reading of summary statistics (detailed reading of values from plots should not be a major focus) • obscure the individual which may increase the risk of students treating the boxes as pictures or mistakenly interpreting the areas of segments of the box using their fraction knowledge or histogram frequency knowledge • give limited information about the shape of the distribution (symmetry/skewness) of the sample data or population • don't show modality • are not good for summarising small samples ($n \leq 15$ or so)
Overall visual comparisons ***I notice:*** • there is a lot of overlap between the boys' and girls' foot lengths • the boys' foot lengths are shifted further up the scale • one of the girls has a recorded foot length *far* shorter than any other girl	**Teaching Tip:** • Ensure students know what is meant by 'overlap' Perhaps best explained to students visually, e.g., moving dot or box plots around on whiteboard or paper .
Shift and overlap ***I notice:*** • the middle 50% of the boys' foot lengths (the box) is shifted *much* further along the scale than the middle 50% of the girls' • there is *some* overlap for the middle 50% of the boys' right foot lengths and the middle 50% of the girls' • some of the boys have bigger right foot lengths than some of the girls and vice versa	**Technical Note:** • The key idea is to use a central proportion to remove the influence of the relatively few high and low values, i.e., use the central bulk of the data • The '50%' central proportion ties in with the box of the box plot summary.
Summary (looking at summary statistics) ***I notice:*** • the boys' median foot length is bigger than the girls' median by 2cm (boys' median is 25cm, girls' median is 23cm) i.e., there is a difference in the medians of 2cm. Half of the boys have a foot length of at least 25cm	Use the plots **and** the numerical summaries to compare appropriate summary statistics. Interpret statements where appropriate. At higher levels measures of centre (e.g., median and means) are going to be used as point estimates. Students should be encouraged to start working with them now. **Point estimates** (e.g., medians) alone are ***not sufficient*** to

Descriptions and inferences: comparing two samples (Level 5, Year 10)	Teacher Notes
whereas half of the girls have a foot length of at least 23cm. • the boys' median and the girls' upper quartile are the same, i.e., half of the boys have a foot length at least 25cm whereas only a quarter of the girls do	answer the investigative question since sampling variability has not been taken into account. Students will often conclude that, back in the two populations, boys tend to have longer foot lengths than girls *solely* on the basis that the sample median for the boys is bigger than the sample median for the girls. A major aim at this level is to counter this incorrect reasoning by students.
Spread ***I notice:*** • the middle 50% of boys have a right foot measuring between 24cm and 27cm (IQR = 3cm) whereas the middle 50% of the girls are between 22 and 25cm (IQR = 3cm). This means that the foot lengths for these boys vary by about the same amount as these girls' do. • the boys' foot lengths went from a minimum of 21cm to a maximum of 30cm whereas the girls went from 15 to 29cm. ***I wonder:*** • if boys' and girls' foot length distributions back in the two populations have similar variability. I expect so.	Spread is one aspect of the **overall pattern of variability** in a distribution. (Another aspect is shape.) The **range** should **not** be **used** as it is very inclined to be an unstable estimate of the population spread. The range is highly likely to vary greatly from sample to sample for samples of these sizes. The range is also prone to be severely affected by the occasional extreme observation. Students should be encouraged to use other more resistant measures of spread such as the **IQR**. The IQR is not disturbed by the presence of a few very large or very small observations. It is **important to convey** whether we are talking about the **samples** or the **populations**. Sometimes statements are made *without explicitly stating* whether we are referring to the samples or the populations. When we say "*the* boys" then it is understood to mean that we are talking about the boys in the sample, whereas "boys" (without using 'the') means we are talking about boys in the population. (See the questions in the Problem section.) When we write ***"I expect …"*** (as part of an "I wonder …" statement) we are asking students to draw on their common sense knowledge. That is, relate the knowledge they already have about the world they live in to features they observe in the samples or features they wonder about in the populations. When using the data to make inferences about the populations we draw on both statistical knowledge and our own contextual knowledge. For example, our statistical knowledge tells us that range is not appropriate. Our contextual knowledge leads us to believe that boys' and girls' foot lengths may have similar variability.

Descriptions and inferences: comparing two samples (Level 5, Year 10)	**Teacher Notes**
Shape ***I notice:*** • the sample distribution for the boys' foot lengths is roughly symmetrical with a mound around 23 to 27cm, i.e., unimodal *24 changed to 23 on 28/3/2013 • the sample distribution for the girls' foot lengths shows a large mound around 22 to 24 cm and a hint of a small mound around 27cm, i.e., a hint of bimodality ***I wonder:*** • if boys' and girls' foot length distributions back in the two populations are roughly symmetric and unimodal. I expect so for a body measurement such as foot length for both girls and boys.	When we consider the **shape** of a distribution we are trying to understand another aspect of the overall pattern of variability for foot length. *'Bell-shaped'* is often used to describe a *symmetric mound*. We suspect that the (hint of) bimodality displayed in the girls' foot lengths is just a manifestation of sampling variability and will NOT be present in the population and therefore we would not normally comment on it. We would fleetingly notice it, decide that they were simply manifestations of sampling variability and that would have been the end of it. Inferential thoughts tell us which data features we should comment on and which ones we should ignore. Inferential thoughts help govern what descriptive statements to make. (See similar commentary in the Gaps/Clusters section.) **Teaching Tip:** Discourage • the over-interpretation of modality in the data The sample distributions lead us to believe that the patterns of variability in the population distributions are as expected. If the sample distribution patterns suggest a departure from what we would expect in the population distribution patterns then we would want to explore reasons for this departure. For example, if the observed data pattern led us unexpectedly to believe that the population distribution were bimodal then we would want to seek explanations for this bimodality. **Teaching Tip:** Invite students to look back at their predicted sketches for the foot length population distributions (Problem Phase) and compare with the two sample distributions.

Descriptions and inferences: comparing two samples (Level 5, Year 10)	**Teacher Notes**
Individual ***I notice:*** • one of the girls has a foot length (15cm) far smaller than any other girl ***I worry:*** • that this may be a mistake. It could be a measurement or recording mistake or perhaps this girl is much younger than 13 years. I wouldn't expect a 13 year-old girl to have a foot size this small. I need to check her other measurements such as age, height etc. to further investigate this extreme value.	**Teaching Tips:** • Describe any unusual observations **in context**. • Avoid using the word *outlier* with students at this level Use of the word *outlier* can create problems. Students tend to see *outlier* as a *strictly technical* word (jargon) and don't see its common language meaning, *to lie outside*, within the word itself. They often believe that if they label an observation as an outlier then they have license to discard the observation. It is better to use "really unusual observation " or "oddball". **Technical Note:** • Unusual observations which stand by themselves, i.e., are a long way away from the main body of the data are ***oddballs*** • Oddballs can have a **big effect on conclusions** reached • Avoid the temptation to discard an oddball just because it is odd. We need to be *very careful* about discarding any data • We need to check whether an oddball is a mistake or whether there is something unusual and interesting going on • If possible, go back to the original source of the data to see whether the oddball is a mistake • If the oddball is a mistake then correct it or if the mistake can't be corrected then discard it and report this action • If the oddball is NOT a mistake (or cannot be confirmed as a mistake) then it must NOT be discarded • If the oddball is not a mistake then we should seek to explain what caused it. If the cause is undetermined then simply accept that large unexplained variations occur from time to time
Gaps/Clusters ***I notice:*** • the dots are stacked on whole numbers. This is because the foot lengths are measured to the nearest cm. • there is a gap in the girls' group at 28cm and gaps in the boys' group at 22 and 29cm ***I wonder:*** • if boys' and girls' foot length distributions back in the two populations would have gaps at these same values. I don't expect so because I don't know any reason for this to happen.	Absence of 22 and 29cm measurements for the boys and 28cm for the girls is NOT remarkable, i.e., we suspect (based on the number in the group, context, and common sense) that back in the two populations there would be girls and boys who have these foot lengths. We suspect that these 'gaps' in the data have just happened by chance. (Note this is an inferential thought.) We would not normally comment on such 'gaps' in the data.

Descriptions and inferences: comparing two samples (Level 5, Year 10)	Teacher Notes
	Inferential thoughts tell us which data features we should comment on and which ones we should ignore. Inferential thoughts help govern what descriptive statements to make. For this example, we would normally make no comment on these gaps. We would fleetingly notice them, decide that they were simply manifestations of sampling variability and that would have been the end of it. (See similar commentary in the Shape section.) It is not usual to comment on sample features which are absent. **Teaching Tips:** Discourage • the over-interpretation of 'gaps' in the data • commenting on sample features (e.g., skewness, oddballs, clusters etc) which are not present in the data **Technical Note:** If the presence of any clusters in a sample distribution lead us to believe that there are subgroups back in the population (the pattern in the distribution has not just happened by chance), then we should seek to identify what defines these subgroups.
Sampling If a new random sample of 24 13-year-old boys and a new random sample of 22 13-year-old girls were taken I would expect the plots to look different because of sampling variability. With these sample sizes, I would expect each IQR spread to change slightly and that each box would be slightly further down or up the scale. ***I wonder:*** • if I repeated this sampling process many times the boys' foot lengths would, just about always, be shifted further up the scale than the girls' • if boys tend to have a greater foot length than girls back in the two populations • if the median foot length of boys *really* is greater than that of girls back in the two populations	If a data pattern (e.g., the shift pattern between two groups) comes up again and again in repeated sampling, then this data pattern is a reflection of what is happening back in the population, i.e., **a *real* pattern**. That is, the data pattern has **not just happened by chance,** i.e., the pattern is not just as a result of who, *by chance*, we happened to randomly select for our sample. We are asking ourselves whether the data pattern *conforms to a chance explanation*, or whether the data pattern is *implausible under a chance explanation*. **Teaching Tip:** Draw students' attention to the fact that in practice we don't repeatedly sample.

<table>
<tr><th>Descriptions and inferences:
comparing two samples (Level 5, Year 10)</th><th>Teacher Notes</th></tr>
<tr><td>

I notice:
that more than half of the girls have foot lengths shorter than the foot lengths of ¾ of the boys</td><td>At this point the students should be starting to think about what they need to look at in the plot in order to apply the ‘rule’: at least one of the groups’ medians is outside the box of the other group.</td></tr>
</table>

Descriptions and inferences: comparing two samples (Level 5, Year 10)	Teacher Notes
CONCLUSION The shift in these two samples makes me want to claim that right foot lengths of 13 year-old New Zealand boys tend to be longer than right foot lengths of 13 year-old New Zealand girls back in the two populations. I am prepared to make this call because, in my data, the difference between the boys' and the girls' foot lengths is big enough for my two sample sizes. To make this call, with sample sizes between 20 and 40, the rule requires that more than half of the girls' foot lengths must be smaller than ¾ of the boys' (or more than half of the boys' foot lengths must be longer than ¾ of the girls'), i.e., the median foot length for the girls must be outside the box for the boys. By making the call, I am saying that, in my data, the pattern of the boys tending to have longer foot lengths than the girls is not just due to who happened to be randomly selected in the girls' group and who happened to be randomly selected in the boys' group, i.e., has not just happened by chance. I claim that this pattern in the data is *real*, i.e., that this pattern persists back in the two populations.	We use '… *right* foot length …' because the investigative question asks about the right foot length. Using statistics there is always the possibility that the calls (decisions) that we make are wrong, i.e., we are making calls in the face of uncertainty. For example, we want to make a call on who tends to be taller (back in the two populations), 13 year-old boys or 13 year-old girls. We may make the call that it's 13 year-old boys when in fact it's girls who tend to be taller. Or, we may *not* want to make a call even though boys tend to be taller than girls. **Technical Note:** ***Making a call* versus *making a claim* versus *making a conclusion*** At this level we prefer to use the phrase *making a call* because it has a strong connotation of making a decision through a *weighing-up* or an *on-balance* type reasoning and, inherent in that, is the possibility of the 'call' being wrong. The phrase *making a claim* has, albeit to a slightly lesser extent, the same connotation. The phrase *I conclude* has greater connotations of certainty and therefore its use, at this level, runs the risk of the students losing sight of the fact any so called 'conclusions' are made in the face of uncertainty.
Explanatory I expected boys to have bigger feet than girls. This study gives me enough information to be able to make the call that 13 year-old boys tend to have bigger feet than 13 year-old girls. I can't think of any other factor which can explain the difference in foot size other than gender.	In this explanatory element we ask ourselves if our conclusion makes sense with what we know, i.e., whether our contextual knowledge matches our conclusions. We must try to think of other factors which may lead to alternative explanations when measuring foot lengths. These suggestions should also be present in the conclusion. As another example, samples from the NZ Year 5 to Year 10 Census At School database may lead us to claim that students who own a cell phone tend to have less hours of sleep per night than students who don't. However, the explanation for the number of hours slept per night might not be cell phone ownership but rather the age of the students. That is, those students who own cell phones tend to be older students and older students tend to sleep less. Sometimes we will not be able to think of any other factors that should be taken into account and we would simply say that.

LEVEL 6 & 7

The major differences at Levels 6 and 7 are the use of different informal decision making rules. For Levels 6 and 7 the reasoning elements of the Analysis phase would be similar to those shown in Level 5 except for the Sampling element, which will give rise in the Conclusion phase to explicit statements about the *relative* distance between groups' medians.
We will continue using the same context as above, but at Level 7 students will be using more substantive and richer contexts.

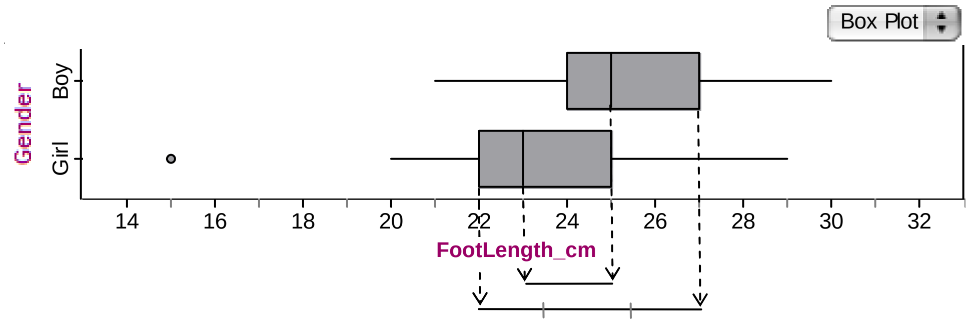
Descriptions and inferences: comparing two samples (Level 6, Year 11)

Sampling

If a new random sample of 24 13-year-old boys and a new random sample of 22 13-year-old girls were taken I would expect the plots to look different because of sampling variability. With these sample sizes, I would expect each IQR spread to change slightly and that each box would be slightly further down or up the scale.

I wonder:

- if I repeated this sampling process many times the boys' foot lengths would, just about always, be shifted further up the scale than the girls'
- if boys tend to have a greater foot length than girls back in the two populations
- if the median foot length of boys *really* is greater than that of girls back in the populations



"overall visible spread"

I notice:

- that the distance between the medians is greater than 1/3 of the "overall visible spread"

Teacher Notes

See **Level 5** Sampling commentary.

Could the difference or *distance* between the two sample medians just be due to sampling variability?

- If the distance between the two sample medians is *small*, then we could just write it off as being due to sampling variability and not necessarily a reflection of what is happening back in the populations.
- If the distance is *big*, then we can't say that it is just due sampling variability alone, it must be, at least partially, due to what is happening back in the populations.

Major Principles:

1. *Bigness* and *smallness* are relative ideas, e.g., the distance between myself and a person sitting next to me is a **BIG** distance *relative* to the width of a sewing needle whereas it is a **small** distance *relative* to the length of a football field.
2. Consider the distance between the two medians *relative* to some measure of variability (spread) of the two samples.

At Level 6, we consider the distance between the two medians **relative** to the *"overall visual spread"*.

Teaching Tips

- Stress ***visual*** measuring **not** measuring using a ruler **nor** reading off the scale (even although in this exercise the actual distances are very easy to read off).
- **By eye**, divide the 'overall visible spread' into thirds.

Descriptions and inferences: comparing two samples (Level 6, Year 11)	Teacher Notes
CONCLUSION I am going to claim that the right foot lengths of 13 year-old New Zealand boys tend to be longer than the right foot lengths of 13 year-old New Zealand girls back in the two populations. I am prepared to make this call because, in my data, the distance between the boys' and the girls' median foot lengths is big relative to the overall visible spread. To make this call, with sample sizes of around 30, the difference between the two foot length medians needs to be more than about 1/3 of the overall visible spread. This is true for my data. I don't believe that the pattern in my data of the boys tending to have longer foot lengths than the girls is just due to who happened to be randomly selected in the girls' group and who happened to be randomly selected in the boys' group, i.e., I don't believe this data pattern has just happened *by chance*. I am prepared to claim that this pattern in the data is *real*, i.e., that this pattern persists back in the two populations.	See **Level 5** Conclusion commentary
Explanatory I expected that boys tend to have bigger feet than girls back in the populations and the information I collected (my data) supports this belief. I can't think of any other factor which can explain the difference in foot size other than gender.	See **Level 5** Explanatory commentary

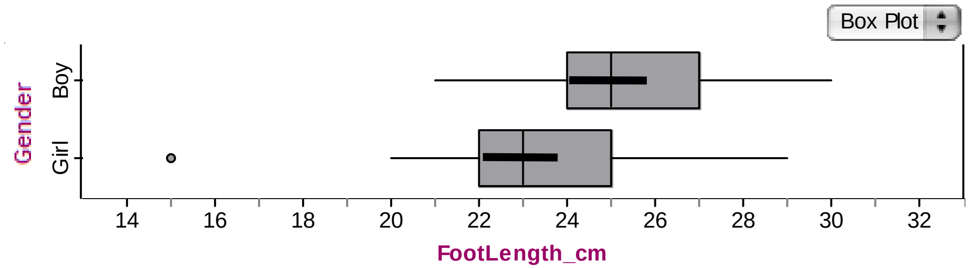
Descriptions and inferences: comparing two samples (Level 7, Year 12)	**Teacher Notes**
Sampling	See **Level 6** sampling commentary

Sampling

If a new random sample of 24 13-year-old boys and a new random sample of 22 13-year-old girls were taken I would expect the plots to look different because of sampling variability. With these sample sizes, I would expect each IQR spread to change slightly and that each box would be slightly further down or up the scale.

I wonder:

- if I repeated this sampling process many times the boys' foot lengths would, just about always, be shifted further up the scale than the girls'
- if boys tend to have a greater foot length than girls back in the two populations
- if the median foot length of boys *really* is greater than that of girls back in the populations



I notice:

- that the informal confidence intervals for the population medians do not overlap

Descriptions and inferences: comparing two samples (Level 7, Year 12)	Teacher Notes
CONCLUSION I am going to claim that, on average, the right foot length of 13 year-old New Zealand boys is longer than the right foot length of 13 year-old New Zealand girls back in the two populations. I am prepared to make this call because, from my data, we are reasonably sure that the possible values for the boys' and girls population medians are somewhere within their respective informal confidence intervals. To make this call, with sample sizes of around 30, these informal confidence intervals for the population medians must not overlap. This is true for my data. I don't believe that the pattern in my data of the boys' median foot length being greater than the girls' has just happened *by chance*. I am prepared to claim that this pattern in the data is *real*, i.e., that population median foot length is greater than the population median foot length for the girls.	At this level, there is a greater emphasis on using measures of centres and the investigative question is more likely to be "Is the average right foot length for 13 year-old New Zealand boys bigger than the average right foot length for 13 year-old New Zealand girls?".
Explanatory I expected that, on average, boys have bigger feet than girls back in the populations and the information I collected (my data) supports this belief. I can't think of any other factor which can explain the difference in foot size other than gender.	