

Level 7 Statistical Investigations

Jeanette Saunders
St Cuthbert's College

NZC Objective S7-1

In a range of meaningful contexts, students will be engaged in thinking mathematically and statistically. They will solve problems and model situations that require them to:

Carry out investigations of phenomena, using the statistical enquiry cycle:

A conducting surveys that **require random sampling techniques**, conducting experiments, **and using existing data sets**;

B evaluating the choice of measures for variables and the sampling and data collection methods used;

C **using relevant contextual knowledge, exploratory data analysis, and statistical inference.**

NZC Objective S7-2

In a range of meaningful contexts, students will be engaged in thinking mathematically and statistically. They will solve problems and model situations that require them to:

Make inferences from surveys and experiments:

- A. making informal predictions, interpolations, and extrapolations
- **B. using sample statistics to make point estimates of population parameters**
- **C. recognising the effect of sample size on the variability of an estimate.**

Indicators for AO 7.2 C

Recognising the effect of sample size on the variability of an estimate:

Within the context of an investigation and statistical plots of observed data:

- Finds **informal confidence intervals for population medians.**
- Plots sample data showing informal confidence intervals ($\text{median} \pm 1.5 \text{ IQR} / \sqrt{n}$) on boxplots.
- **Uses an informal confidence interval** to make an inference about the population median from sample data plot.

Indicators for AO 7.2 C continued

- Makes a claim about whether one group has larger values than another group using informal confidence intervals for the population medians.
- Explains the connections among sample, population, sampling variability, sample size effect, informal confidence interval, and degree of confidence.

What is a point estimate?

A number calculated from a *random sample* that is used as an approximate value for a *population parameter*.

Example

- A *sample proportion*, calculated from a *random sample* taken from a *population*, is a point estimate of the *population proportion*.

What is an interval estimate?

A range of numbers, calculated from a *random sample* taken from the *population*, of which any number in the range is a possible value for a *population parameter*.

Level 7 investigations

- Involve making and using interval estimates for a population parameter, such as the population median.
- Same thinking can be applied to estimating other parameters, such as quartiles, means . . .
- For teaching and learning, medians are used.

Welcome to Teacher Land

- Teacher land is not the real world;
- Students understand Teacher Land is a special place for learning;
- Teacher Land is where students learn how to make inferences from a known population;
- Real world inference is based on one sample only;
- Real world inferences can be relied on because of lessons learned in Teacher Land.

Kiwi Kapers 1, 2 and 3

- Are teacher guides for Level 7 inferential reasoning;
- KK1 and KK2 can be downloaded from <http://seniorsecondary.tki.org.nz/>
- KK3 is still in development. Current version is available at <http://nzstatsedn.wikispaces.com/Gisborne+2012>

KK 1, 2 and 3

- Are lesson plans, with comments from trialling, and a possible script for you to use.
- They present learning sequences to develop the big ideas in Level 7 inference.
- You will adapt them for your students in your school.











Kiwi Population

These lessons are based on a created population of 700 kiwis:

- males and females,
- weight in kg
- height in cm
- 3 species: great spotted, North Island brown and tokoeka.

Senior Secondary guides on TKI have all the information and data.

Files for KK1 on Senior Secondary Guides - everything you will need

- Teaching notes
 -  [Kiwi Kapers 1](#) (Word, 283 KB)
- Data cards
 -  [Great spotted data cards](#) (PDF, 35 KB)
 -  [North Island brown data cards 1](#) (PDF, 39 KB)
 -  [North Island brown data cards 2](#) (PDF, 22 KB)
 -  [Tokoeaka data cards](#) (PDF, 43 KB)
 -  [Kiwi data cards spreadsheet 1](#) (Excel, 868 KB)
- Information about the variables
 -  [Kiwi population information](#) (Word, 55 KB)
- Population graphs and statistics
 -  [Kiwi population graphs and statistics](#) (Word, 44 KB)
- Blank dot plot sheet for collecting weights
 -  [Weight samples](#) (Word, 47 KB)
- Kiwi data in Excel
 -  [Kiwi populations](#) (Excel, 99 KB)
- Kiwi data in Fathom
 - [Kiwi populations](#) (Unknown Binary File, 80 KB)

5 Species of Kiwi

The five formally described kiwi species are:

- Little spotted kiwi (*A. owenii*) on several offshore islands and at Karori Sanctuary in Wellington
- Great spotted/Roroa (*A. haastii*) in the northern South Island
- Brown kiwi (*Apteryx mantelli*) in the North Island
- Rowi (*A. rowi*) at Okarito, on the West Coast of the South Island
- Tokoeka (*A. australis*) in the South Island (Fiordland, the Haast Range and on Stewart and Kapiti Islands)

Web link for information

<http://www.kiwisforkiwi.org/>

Background radio podcast

[This Way Up Radio New Zealand](#)

Search Radio New Zealand podcasts for
This Way Up: Kiwi Tracking

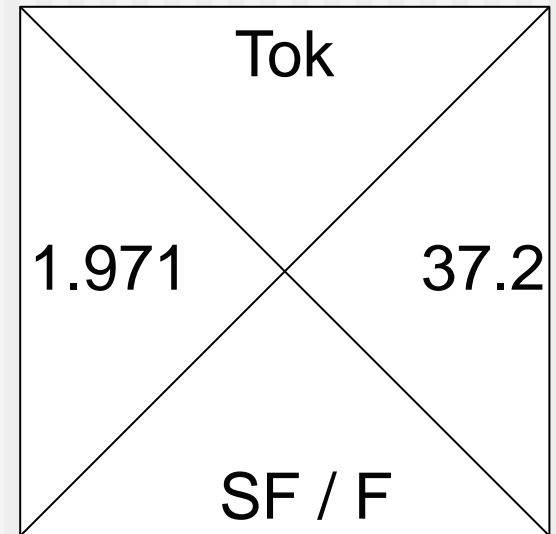
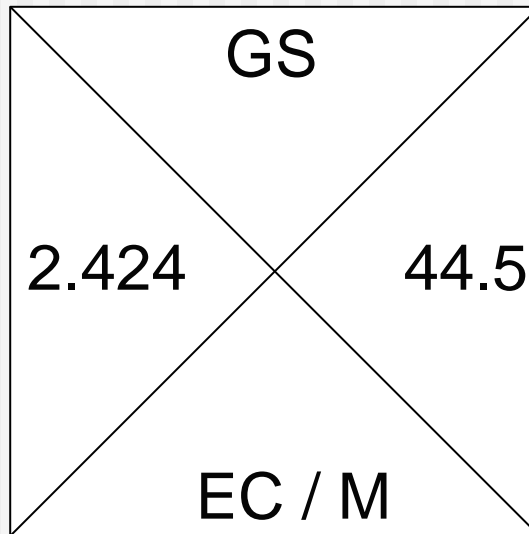
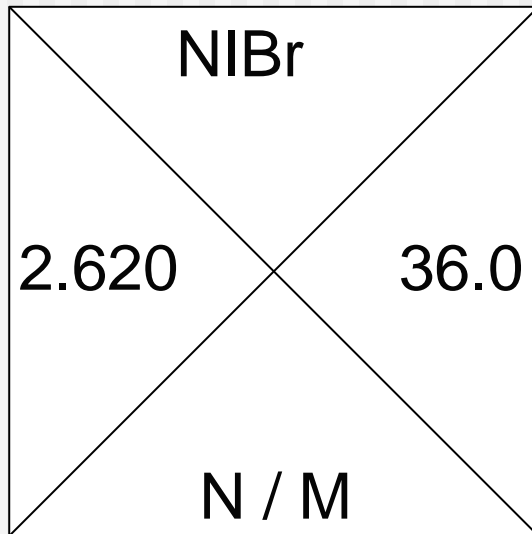
<http://www.radionz.co.nz/national/programmes/thiswayup/audio/2518442/kiwi-tracking>

KK 1 Big ideas

- **Sample medians vary** (at Level 6, we said samples vary) and medians vary more for small samples and less for large samples.
- Possible range for variation in medians can be described eg Evidence from my samples of size 15 suggests the median weight of kiwis is between 2.2 and 3.0 kg.

Kiwi Data Cards

- Population is available as data cards:



-
- Read page 1 of KK1
 - Discuss briefly with your neighbour
 - Reflect on setting the scene for a series of lessons: what will you say, what will you expect students to say/ask?

-
- Read page 2 of KK1
 - Notice definitions of question types and features of good questions
 - Reflect on Learning question 1: What is its purpose, how would you use it in your lesson?

-
- Read page 3 of KK1
 - Sampling begins here, second lesson of plan.
 - How could collecting samples of 15 and 30 be efficiently managed?

weightsamples.doc

KIWI KAPERS

WEIGHT

SAMPLE SIZE _____

1.2 1.4 1.6 1.8 2.0 2.2 2.4 2.6 2.8 3.0 3.2 3.4 3.6 3.8 4.0 4.2

Analysis page 4

- Students draw dot plots and box plots of their samples of size 15 and size 30 and write ‘I notice . . . ‘ and ‘I wonder . . . ‘ statements
- We didn’t copy and reduce the samples, but lined them up full size on noticeboard, marked medians and quartiles with felt pens to show overall picture (page 5).
- What do you expect students to learn from the sets of graphs, $n = 15$ and $n = 30$?
- **Collect sample medians for use in later lessons.**

Analysis page 5

Key ideas about sampling;

- Are the samples fair?
- Did each member of the population have an equal chance of being selected?
- What do you notice about the spread of median weights?
- If you were making a prediction of the population median, what would you be comfortable saying after this lesson?

Answering the question

- Read page 6 conclusion
- The population median weight is likely to be
- Reflect on the conclusion your students might reach and what you would say to guide them to acknowledge that an interval would be useful for describing range of possible values for population median.

KK2 Big Ideas

- Technology can be used to take samples and record medians.
- Samples of size 30 or 50 are large enough for making a reliable inference of the population median.

KK2

- Explores more sample sizes, $n = 50$ and $n = 100$, using technology:
- www.censusatschool.org.nz/2010/data-viewer/
- Decision made about a sample size that is big enough to give a reliable estimate, but not too big, given the time and effort needed to catch and weigh kiwis.

Written for using Fathom

- But is easy to do using Data Viewer from www.censusatschool.org.nz
- The big idea is that now students will be working with distributions of sample medians, which will be less spread out than the population. Quantifying that spread and relating it to the size of the sample is the crucial learning to be guided.

Census at School data viewer

PROBLEM

What is your **investigative question**?

I wonder ...

My variable(s) are ...

I am a year 12 student

PLAN

What is the **population** you want to **sample** from?



What is your total **sample size**?

Screen Shot

DATA

ANALYSIS

Sample: 11-21-59481947

What **variable(s)** do you want to analyse?

Select one or two variables below and click "Do Analysis" to add a graph.

Weight



variable 2



Do Analysis



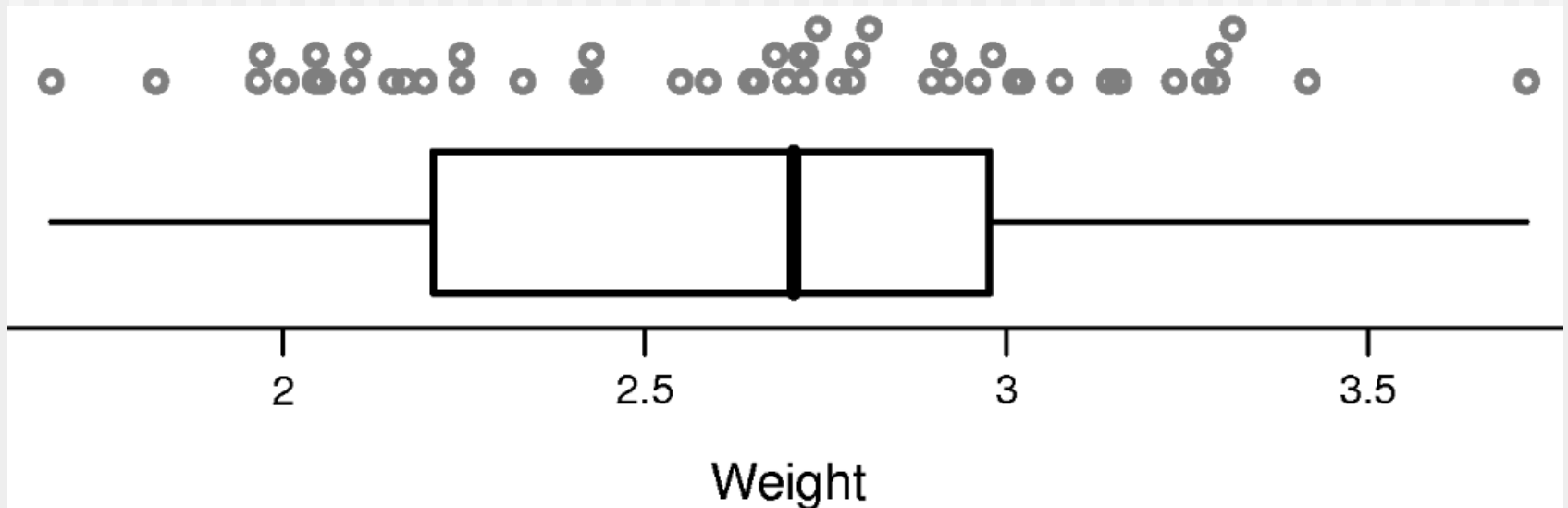
Add summaries

Further exploration:



Do Analysis

Data Viewer Analysis



Summary of Weight

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std.dev	n.obs
1.68	2.21	2.71	2.64	2.98	3.72	0.48	50.00

KK3 Big Ideas

- A Level 7 question asks about a population parameter:
- I wonder what is the **median** weight of kiwis?
- Informal confidence interval recipe:
sample median $\pm k \times \frac{IQR}{\sqrt{n}}$

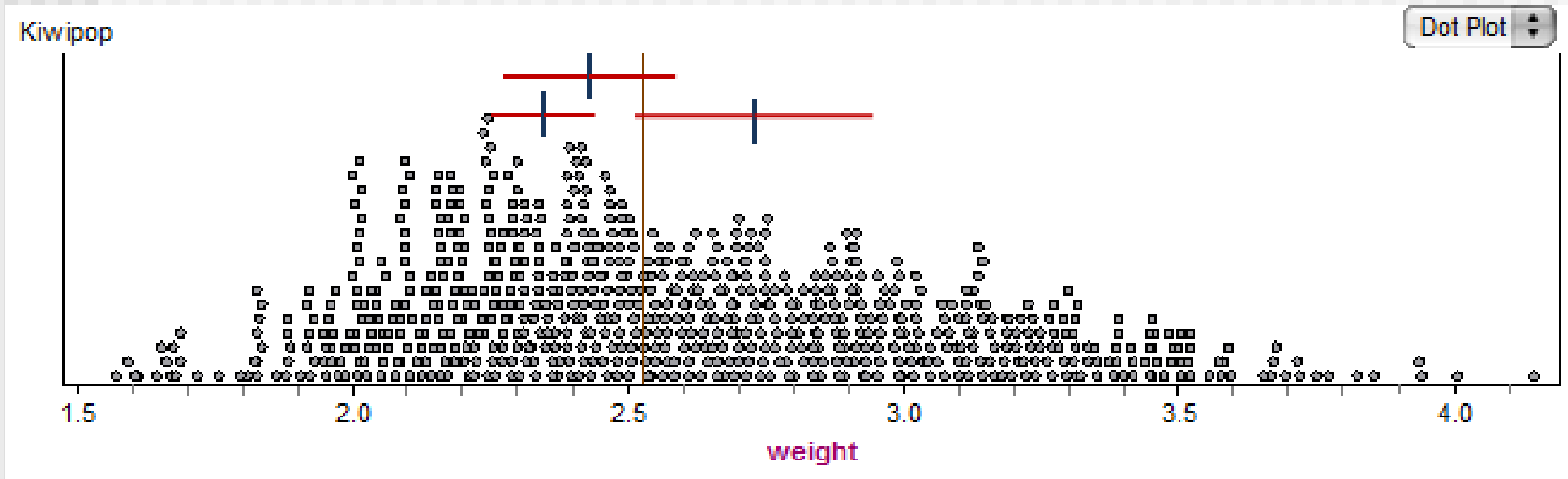
KK3

- Students cannot be expected to discover the informal confidence interval recipe for themselves and should be guided to it.
- Read page 1 of KK3.
- The teacher notes on page 2 have been trialled with teachers and adapted after feedback. You have the latest version.

KK3 Worksheet 1

- Students would add their own box from their sample medians of size 30 (or 50) and then use a ruler to measure widths of boxes.
- If you have a ruler, you may like to measure the boxes and fill in the fractions.
- Results will be surprisingly close to $\frac{1}{\sqrt{n}}$

KK3 Part 2 page 3



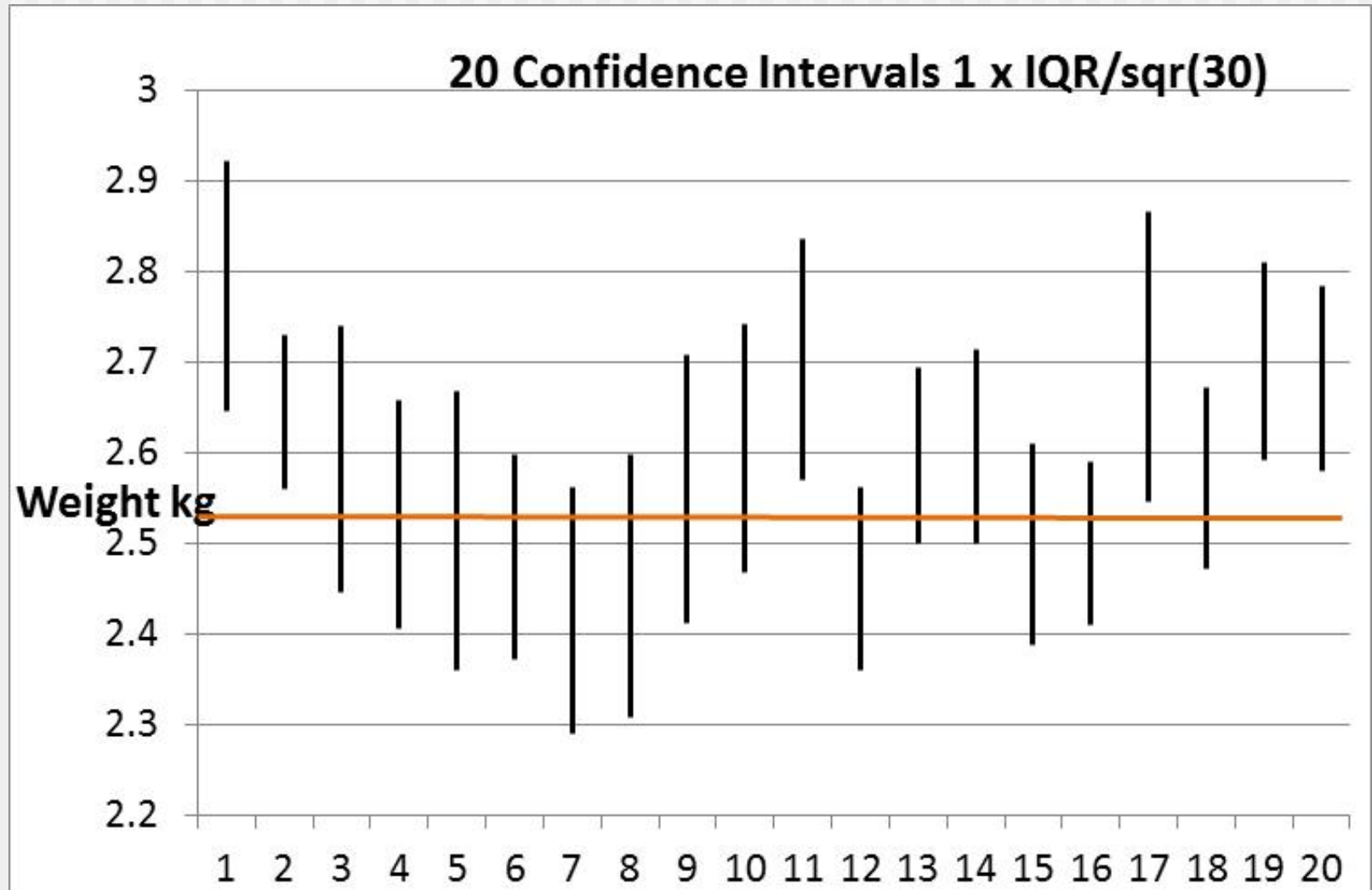
Informal confidence intervals are calculated using:

$$\text{sample median} \pm k \times \frac{IQR}{\sqrt{n}}$$

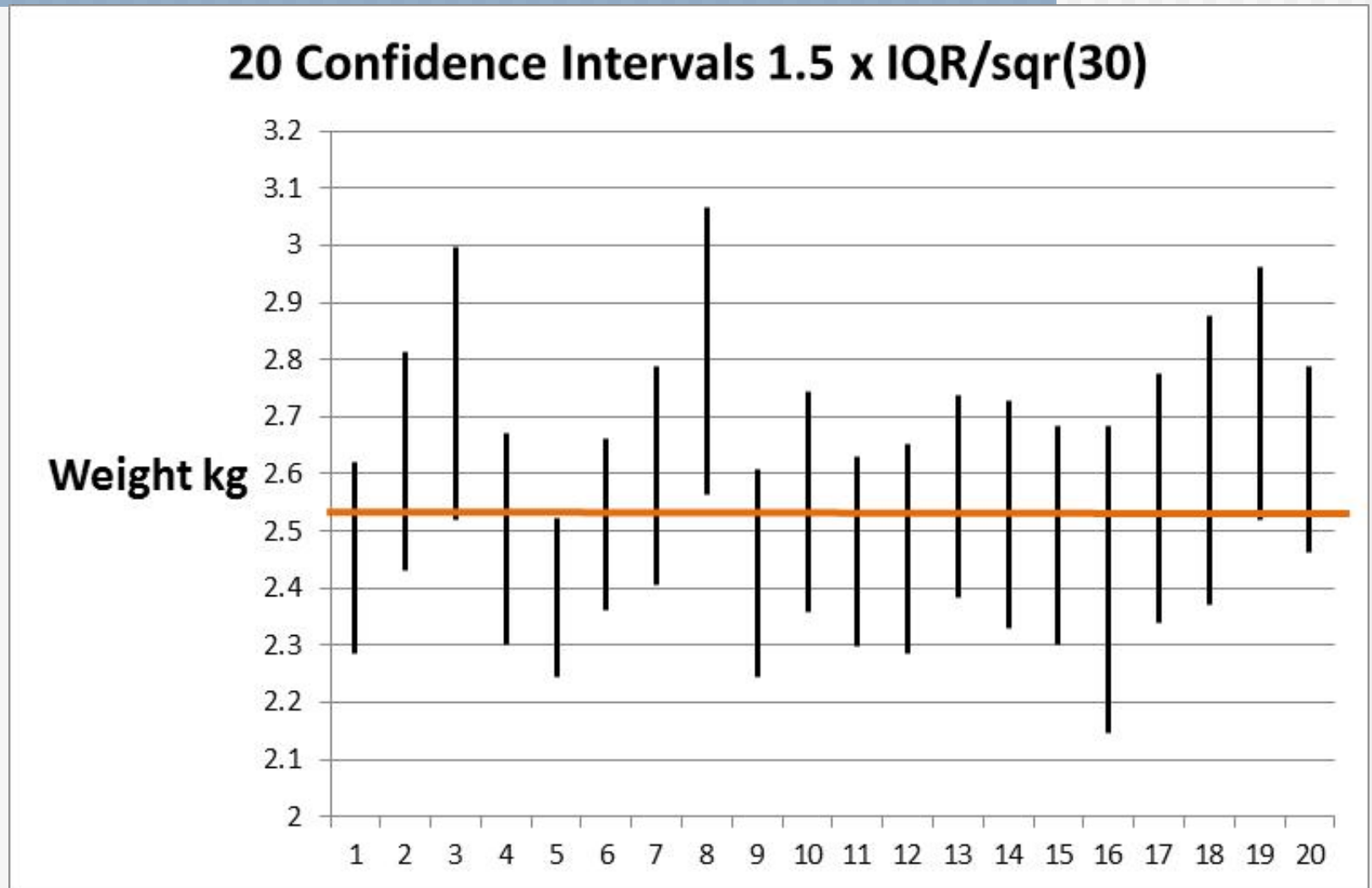
KK3 pages 4 - 8

- Are an alternative teaching sequence for deciding how wide to make the informal confidence interval, using $\frac{IQR}{\sqrt{n}}$ as the yardstick or unit.
- I am suggesting a more visual method, developed by a St C's teacher, Vicky Binedell.

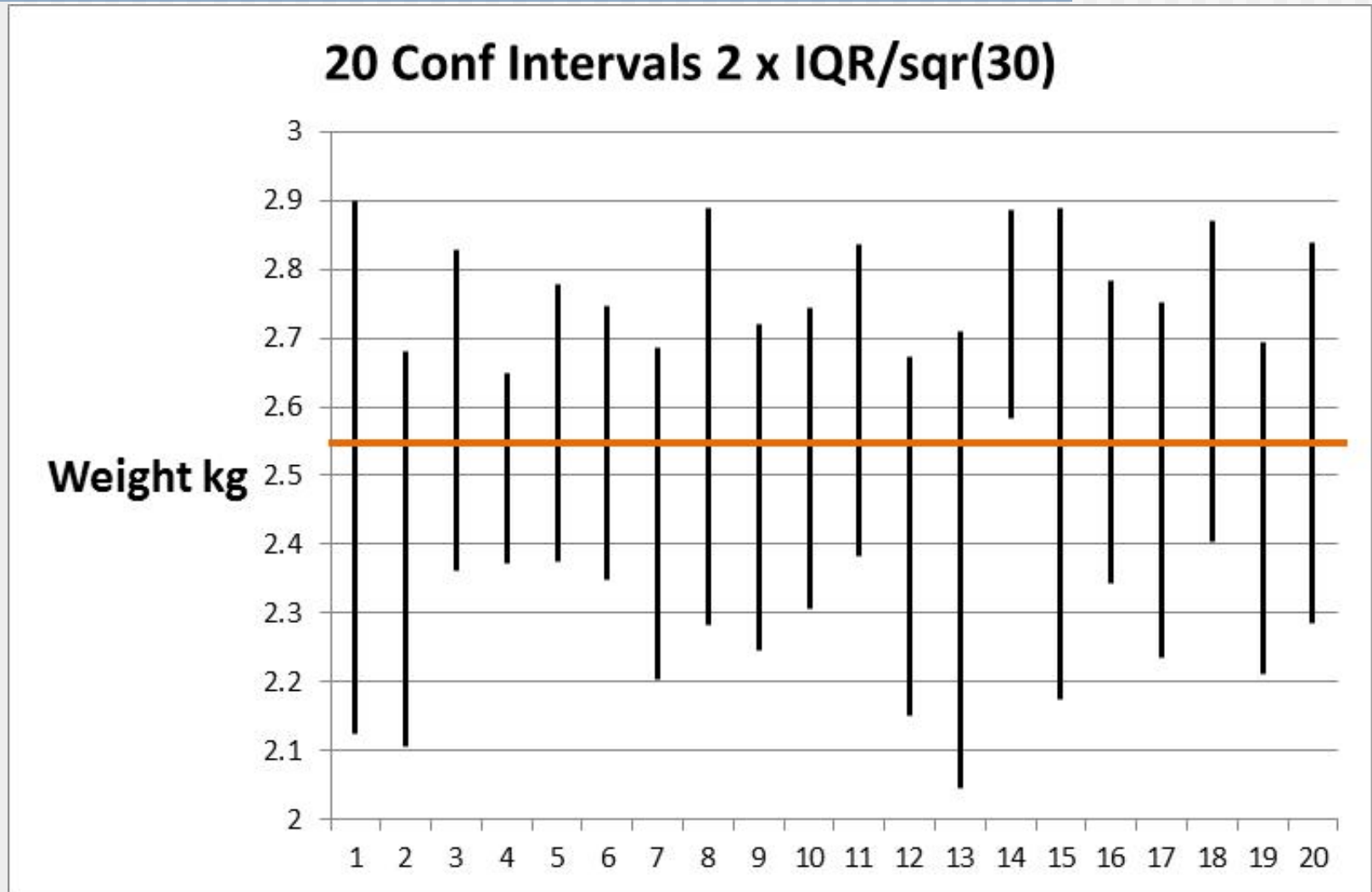
Set of informal confidence intervals, 1



Set of intervals, 1.5



Set of intervals, 2



What is a sensible value for k ?

- What do you notice about the sets of intervals?
- What is the same and what is different about them?
- Discuss with your neighbour.

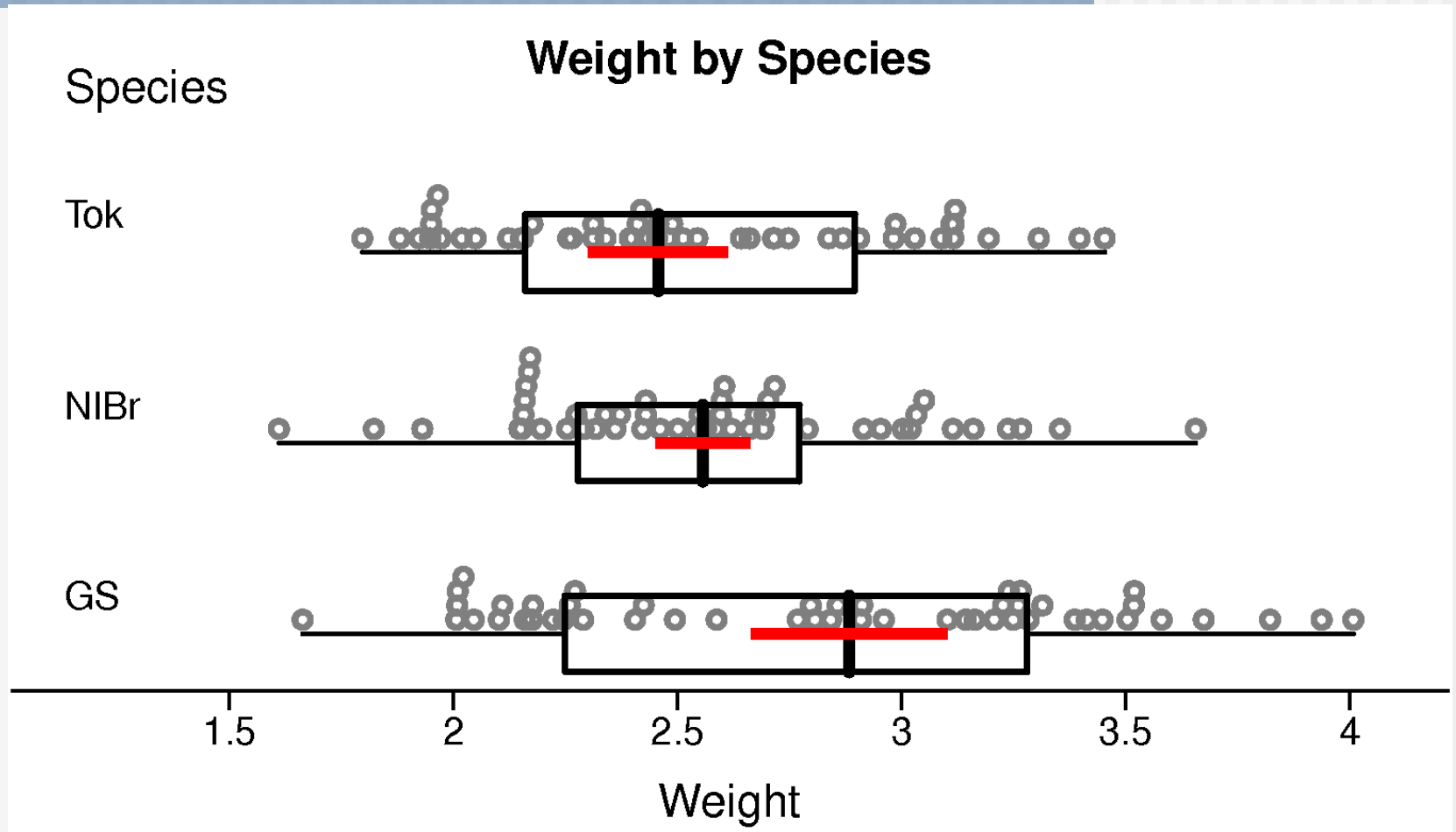
What is a sensible value for k ?

- We should choose a value for k so that about 90% of intervals cover the population median.
- Working from one sample, there is a small, 10%, chance the informal confidence interval will miss the population median, but our method will work for 90% of samples.
- It seems that 1.5 is a sensible value.

iNZight

- Can be downloaded from Census at School, and the latest version has automatic updates.
- The videos show you how to use iNZight, which is a powerful visual way of showing confidence intervals.

Comparing population medians



Word of warning

- Comparing kiwi sample data for weight or height by gender is not suitable for Level 7 as the differences between genders are so great the Level 5 rule would work.
- Find data sets where differences between groups is small so that confidence intervals are needed to make a decision about which group has the bigger parameter (median).