XLR8 – Accelerating Mathematics Learning

YuMiDeadly
Growing community through education

XLR8 Unit 10

Summarising data with statistics

10

2016

Prepared by YuMi Deadly Centre
Faculty of Education, QUT, Kelvin Grove

ACKNOWLEDGEMENTS

We acknowledge the traditional owners and custodians of the lands in which the mathematics ideas for this resource were developed, refined and presented in professional development sessions.

YUMI DEADLY CENTRE

The YuMi Deadly Centre is a Research Centre within the Faculty of Education at Queensland University of Technology which aims to improve the mathematics learning, employment and life chances of Aboriginal and Torres Strait Islander and low socio-economic status students at early childhood, primary and secondary levels, in vocational education and training courses, and through a focus on community within schools and neighbourhoods.

"YuMi" is a Torres Strait Islander Creole word meaning "you and me" but is used here with permission from the Torres Strait Islanders' Regional Education Council to mean working together as a community for the betterment of education for all. "Deadly" is an Aboriginal word used widely across Australia to mean smart in terms of being the best one can be in learning and life.

YuMi Deadly Centre's motif was developed by Blacklines to depict learning, empowerment, and growth within country/community. The three key elements are the individual (represented by the inner seed), the community (represented by the leaf), and the journey/pathway of learning (represented by the curved line which winds around and up through the leaf). As such, the motif illustrates the YuMi Deadly Centre's vision: *Growing community through education*.

More information about the YuMi Deadly Centre can be found at http://ydc.qut.edu.au and staff can be contacted at ydc@qut.edu.au.

Contents

List of Figures

XLR8 Program: Scope and Sequence

	2 year program	3 year program
Unit 01: Comparing, counting and representing quantity Students study countable attributes of their immediate environment, including attributes of the group of students (e.g., more boys than girls, less students with blue eyes than brown eyes) in the classroom and attributes of the school (e.g., quantity of windows in a classroom, quantity of ceiling tiles, length of classroom in steps). This context is limited to those attributes which can be described and represented using whole numbers.	1	1
Unit 02: Additive change of quantities Students extend their investigations of numbers from features of their immediate environment, to features of larger populations in their state, country or world. This context is limited to those features which can be counted using whole numbers and which can be used in additive number stories (for which the total or one of the parts is unknown).	1	1
Unit 03: Multiplicative change of quantities Students explore multiplicative relationships and changes using real-world situations that involve discrete items. This context is limited to those features which can be counted using whole numbers, can be used in multiplicative number stories (for which the product or one of the factors is unknown), and for which divisions also result in whole numbers.	1	1
Unit 04: Investigating, measuring and changing shapes Students explore 3D objects, their 2D surfaces and the 1D attributes of point, line and angle. This context includes measurement of the attribute of turn (angle) and mathematical transformations of 2D shapes and 3D objects including reflection, rotation and translation of shapes and how these may be combined with tessellation to generate and describe designs.	1	1
Unit 05: Dealing with remainders Students extend their investigations of partitioning and quotitioning features of their immediate environment and features of larger populations in their state, country or world to include situations that result in a remainder. Students will explore partitioning whole items into fractions, quotitioning into smaller units, and sharing of remainders of collections. Continuous measures such as length provide useful contexts for partitioning and quotitioning.	1	1
Unit 06: Operations with fractions and decimals Students connect the common fraction representations of tenths, hundredths and thousandths to their decimal fraction representations in contexts common to students' immediate environments including money, measurement and parts of discrete wholes. Students will develop strategies to calculate additive and multiplicative changes involving fractional amounts represented as both common fractions and decimals.	1	2
Unit 07: Percentages Students extend their representations of fractions to include percentage. Percentage is used to compare values multiplicatively and to describe quantity comparisons, recommended daily intake of nutrients, discounts, markups, tax and simple interest. Students will be encouraged to work flexibly between common fractions, decimal fractions and percentages.	1	2

	2 year program	3 year program
Unit 08: Calculating coverage Students extend their investigations of attribute measurement from one-dimensional length measures to two-dimensional measures of coverage or area. This idea starts with coverage which can be counted using whole numbers before extending to fractional measures. Area measurement and calculation provides an opportunity for consolidation of multiplication and division with larger numbers, and consolidation of multiplication and division of fractional quantities expressed as common fractions, mixed numbers or decimal numbers.	2	2
Unit 09: Measuring and maintaining ratios of quantities Students develop their ability to measure duration, convert between units of measure and describe proportional relationships between quantities of discrete items or measurements using ratio notation. Students will also explore changing overall quantities while maintaining consistent proportions between the parts.	2	2
Unit 10: Summarising data with statistics Students develop their ability to gather, organise and represent data from primary and secondary sources. Ideas of sample, population and inference will be used to inform decision making from the gathered data. Students will also develop their ability to analyse measures of central tendency and variation within data sets and learn to represent and interpret these aspects on graphical representations (stem and leaf plots and box and whiskers graphs). Further analysis of the misrepresentation of data will conclude this module's development of ideas surrounding critical analysis and interpretation of data and statistics.	2	2
Unit 11: Describing location and movement Students develop their ability to describe location and movement along a 1D line and in 2D space with respect to an origin and extending from internal to external frames of reference. Generating 2D representations of location and movement on scale maps and grids using alphanumeric coordinates and compass bearings and distance will be extended to include geometric location of points and collections of points on the Cartesian plane. Students will explore Pythagoras' theorem to find diagonal distances travelled.	2	3
Unit 12: Enlarging maps and plans Students develop their ability to describe proportional relationships between quantities of measurements using ratio notation. Ratio will also be used to describe enlargement and reduction transformations to create similar shapes, scale maps and grids, representations of shapes and paths on the Cartesian plane, and plan drawings. Explorations can be extended to trigonometric ratios between similar figures and the application of scale factor to area of similar figures.	2	3
Unit 13: Modelling with linear relationships Students explore parallels between ratio and rate in the context of relationships between measured attributes. These understandings will be extended to algebraic equations which can also be represented on the Cartesian plane to assist with visualisation of relationships and use of equations and algebraic calculations for finding gradient and distances between points on a line.	2	3
Unit 14: Volume of 3D objects Students explore relationships between measurements of solid objects that lead to calculations of formulae, relationships between solid volume and surface area and investigations of contexts that require calculation of solid volume of composite objects.	2	3
Unit 15: Extended probability Students extend upon their ability to determine theoretical probability and make inferences based upon likelihood of an event. Students will explore and compare theoretical and experimental probabilities, recognise when events are mutually inclusive, mutually exclusive or complementary and determine the probability of single-step and multi-step events.	2	3

Overview

Context

In this unit, students will develop their ability to gather, organise and represent data from primary and secondary sources. Ideas of sample, population and inference will be used to inform decision making from the gathered data. Students will also develop their ability to analyse measures of central tendency and variation within data sets and learn to represent and interpret these aspects on graphical representations (stem and leaf plots and box and whiskers graphs). Further analysis of the misrepresentation of data will conclude this unit's development of ideas surrounding critical analysis and interpretation of data and statistics.

Scope

This unit extends students' range of **graphical representations** of **discrete and continuous data sets** from **simple picture**, **bar** and **circle graphs** to include **dot plots**, **stem and leaf plots** and **box plot and whisker graphs**. Students will **gather data** from both **primary sources** using **simple surveys** and from **secondary data** to generate their own **data sets** for **representation** and **analysis**.

Ideas of **sample** and **population** will be explored to enable students to make informed decisions and **inferences** about the data gathered and represented.

Measures of central tendency are calculated to further inform **inferences** and decisions from data as the more general term of average is refined to calculate **mean**, **median**, **mode** and **range**, and interpreted in terms of **spread** and **variation**.

Graphical representation of data and **statistics to describe data** may be used indiscriminately to persuade or influence opinion. Exploration of the various means by which **data may be misrepresented** is important to develop **critical literacy** for clear and effective interpretation of **graphical representations used to support assertions**.

The organisation of these and other related concepts is shown in Figure 1, in which the scope of concepts that is to be developed in this unit is highlighted in blue, concepts that may be connected to and reinforced are highlighted in green and number and algebra concepts and processes that are applied within this area are highlighted in black.

Assessment

This unit provides a variety of items that may be used as evidence of students' demonstration of learning outcomes including:

- *Diagnostic Worksheets:* The diagnostic worksheet should be completed before starting to teach each RAMR cycle. This may show what students already understand. Not all objectives are represented on diagnostic worksheets.
- *Anecdotal Evidence:* Some evidence of student understanding is best gathered through observation or questions. A checklist may be used to record these instances.
- *Summative Worksheet:* The summative worksheet should be completed at the end of teaching the unit. This may be compared with student achievement on the diagnostic worksheets to determine student improvement in understanding.
- *Portfolio task*: The portfolio task *P10: Are we really healthy?* accompanying Unit 10 engages students with exploring the collection and use of data to evaluate complex questions. Students should be encouraged to select their questions, graph and calculate statistics on their data, and then compare with a random sample of 20 similar age students (data drawn from ABS website).

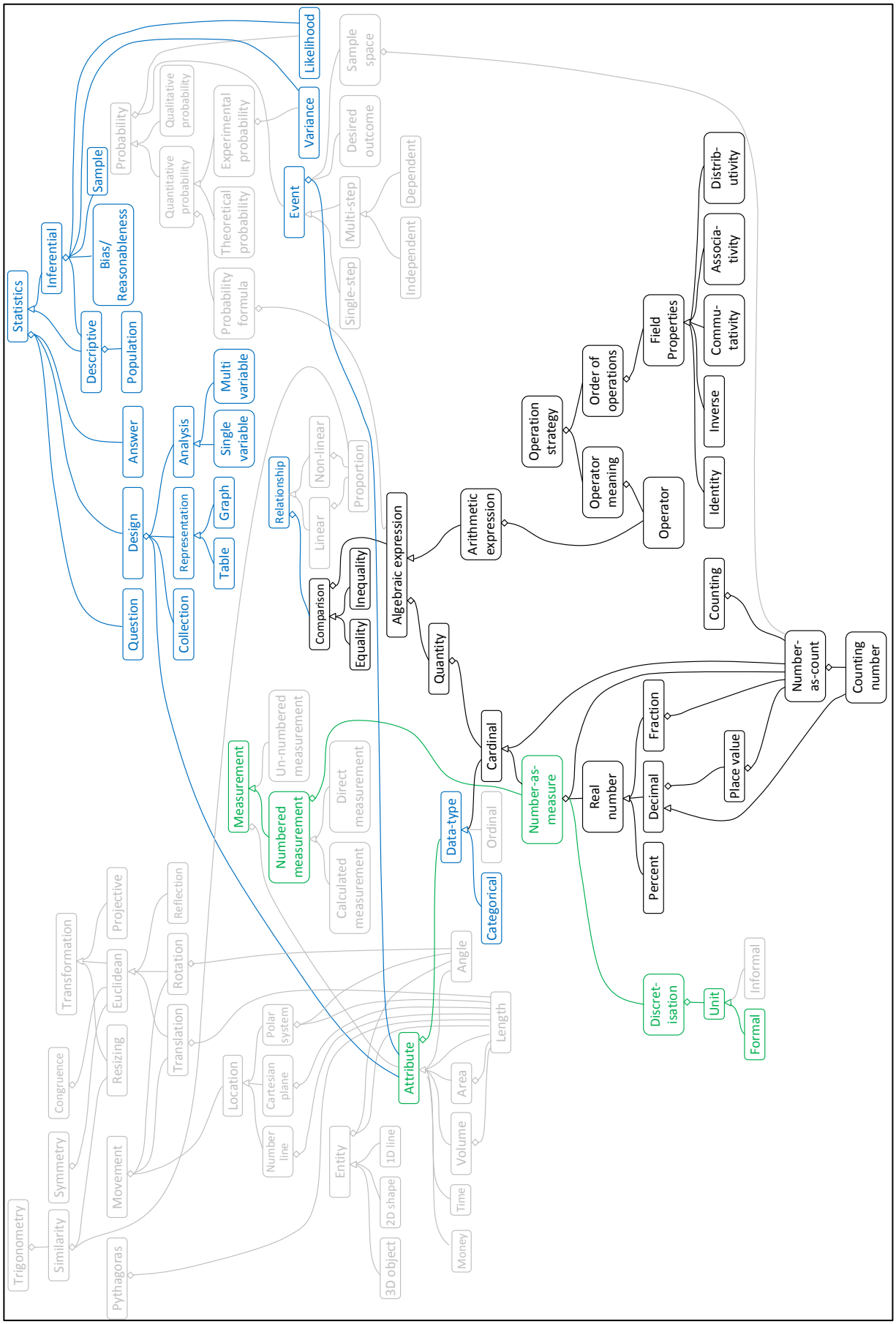


Figure 1. Scope of Unit 10

Cycle Sequence

In this unit, concepts identified in the preceding section are developed in the following sequence:

Cycle 1: Discrete data – Shape and measure

In Unit 01 and Unit 03 students explored pictographs and column graphs of discrete data to develop their understanding of graphical representations of data. This unit extends graphical representations of discrete data to include dot plots and circle graphs, while engaging students with analysis and description of the visual shape of the data with attention to statistical measures of range, spread, mode, median and mean.

Cycle 2: Continuous data – Shape and measure

Within this cycle students will explore the graphical representations and statistical measures of continuous data including histograms (continuous data sorted into categories for representation), stem and leaf graphs and box plot and whiskers. Students will also be engaged with connecting the analysis and description of the data from the graphical representations to the calculation of statistical measures of range, spread, mode, median, quartiles and mean.

Cycle 3: Comparison of data sets

In Cycle 1 and Cycle 2 students explored single data sets. Within this cycle, students will explore the representation of multiple data sets on the one graph for comparison and representation of single data sets that may also be split into smaller categories (e.g., students' heights have been graphed as a single dataset but could be separated by gender into two subsets of the data. Graph types will include side-by-side column graphs, stacked bar graphs, histograms, back-to-back stem and leaf graphs and stacked box plots.

Cycle 4: Sample, population and inference

This cycle engages students with consideration of datasets as possible samples of data that may be applicable to larger populations as well as ideas of variation and uncertainty. Students need to be brought to an awareness of when it is possible to gather data for a complete population and when it is appropriate to gather data from a sample of the population that may be used to infer preferences for the population.

Cycle 5: Misrepresentation of data

This cycle is different from the others in that it looks at how statistics can be used to misrepresent or "lie". This is a powerful way of reinforcing statistical inferential thinking, and also a powerful way to understand how persuasion may be given the credibility of mathematical correctness.

Cycle 6: Correlation

This cycle explores the use of graphical representations to assist with determining when a relationship exists between two continuous variables. For example, time vs. distance travelled, height vs. arm length/leg length/foot length, height vs. age. Students will create scattergrams or scatter plots of data (students' heights vs. arm length) and will consider the idea of a line of best fit in their analysis of the relationship between variables.

Notes on Cycle Sequence:

The proposed cycle sequence outlined may be completed sequentially as it stands or it may be beneficial to complete *Cycle 6: Correlation* before *Cycle 5: Misrepresentation of data*.

Literacy Development

Core to the development of number and operation concepts and their expression at varying levels of representational abstraction (from concrete-enactive through to symbolic) is the use of language that is consistent with the organisation of the mathematical concepts. In this unit the following key language should be explicitly developed with students ensuring that students understand both the everyday and mathematical uses of each term and, where applicable, the differences and similarities between these.

Cycle 1: Discrete data – Shape and measure

Discrete data, tables, graphs, tallies, dot plots, stacked bar graph, stacked column graph, measures of centre, average, mean, median, quartiles, mode, bimodal, range, maximum, minimum, spread, outliers

Cycle 2: Continuous data – Shape and measure

Continuous data, histograms, stem and leaf graph, box plot and whiskers graph

Cycle 3: Comparison of data sets

Comparison, data sets, back-to-back stem and leaf graph, multiple box plot and whiskers graph, side-by-side column/bar graphs, stacked column/bar graphs

Cycle 4: Sample, population and inference

Population, sample, infer, inference, predict, representative group, variation, uncertainty, likelihood, pose, predict, sample, random, population, infer, conclude, decide, certainty

Cycle 5: Misrepresentation of data

Misrepresentation, misleading, interpretation, critical literacy, sample bias, irrelevant statistics, wrong average

Cycle 6: Correlation

Relationship, dependent variable, independent variable, scattergram, scatter plot, line of best fit

Unit 10 Investigation: Gathering Data

Revise the skills for gathering and organising data (explored in Unit 01 and Unit 03).

Students may need to revise the following Unit 01 Objectives:

1.5.1 Pose questions and collect data by observation. [5SP118]

1.5.2 Pose questions and collect data by survey. [5SP118]

1.5.3 Organise data into categories using lists or tables. [3SP069]

1.5.5 Create displays using column graphs. [5SP119]

1. As a class, devise a set of survey questions to gather data about students including the following:
    - Pets
    - Number of People at Home
    - Height
    - Arm Length
    - Foot Length
    - Head Size
    - Shoe Sizes
    - Favourite Takeaway Food
    - Breakfast Choices
    - Favourite Sport
2. Have each student in the class complete the survey.
3. Collate the class survey data into tables or lists. It will be useful to also create these in Excel and save for use throughout Unit 10.
4. Discuss and revise any ambiguous questions to make the survey more streamlined. Conduct the survey with another class to generate a second set of data for use later in Unit 10.

Name:_____ Date:_____

Can you do this? #1

1. Here is data gathered by twelve girls for shoe size.

Girls	01	02	03	04	05	06	07	08	09	10	11	12
Shoe size	7	3	5	4	8	7	6	4	7	4	6	5

(a) Draw a dot plot of the girls' shoe sizes.

Obj. 10.1.1
a)i□
a)ii□
a)iii□
a)iv□
a)v□
a)vi□

(b) Write down the range of shoe sizes in this group. _____ Obj. 10.1.2 b) □

(c) Write down the mode of shoe sizes in this group. _____ Obj. 10.1.3 c) □

(d) Write down the median of shoe sizes in this group. _____ d) □

(e) Calculate the mean of shoe sizes in this group. Obj. 10.1.4 e)i □ e)ii □

(f) If you ran a shoe shop, which size shoes should you stock more of? Obj. 10.1.5 f) □

(g) Why? _____ g) □

Two girls with size 10 shoes join the group.

(h) Will the mode shoe size change? _____ Obj. 10.1.7 h) □

(i) Will the median shoe size change? _____ i) □

(j) If yes, write down the new median shoe size. _____ k) □

(k) Will the mean of the shoe sizes change? _____ Obj. 10.1.3 j) □

(l) If yes, calculate the new mean of the shoe sizes. Obj. 10.1.4 l)i □ l)ii □

Cycle 1: Discrete Data – Shape and Measure

Overview

Big Idea

In Unit 01 and Unit 03 students explored pictographs and column graphs of discrete data to develop their understanding of graphical representations of data. This unit extends graphical representations of discrete data to include dot plots and circle graphs, while engaging students with analysis and description of the visual shape of the data with attention to statistical measures of range, spread, mode, median and mean. Students' *pets* or *number of people at home* data from the Investigation activity may be incorporated within this cycle.

Objectives

By the end of this cycle, students should be able to:

10.1.1 Construct dot plots from discrete data. [7SP170]

10.1.2 Identify range and spread of discrete data sets from graphical representation. [7SP172]

10.1.3 Identify mode and median of discrete data sets from graphical representation. [7SP172]

10.1.4 Calculate mean of discrete data sets. [7SP171]

10.1.5 Describe and interpret graphical representations of discrete data sets using statistical measures in the context of the data. [7SP172]

10.1.6 Interpret and compare a range of data displays of discrete data sets. [6SP147]

10.1.7 Investigate the effect of individual data values, including outliers, on the mode, mean and median of discrete data. [8SP207]

10.1.8 Identify quartiles and interquartile range of discrete data sets from graphical representations. [10SP248]

Conceptual Links

Classifying and sorting skills underlie the gathering and organisation of data. Students will also reinforce previously developed counting and graphical representation skills throughout this cycle.

Analysing and interpreting discrete data using graphical representations and statistical measures in answer to questions is a critical beginning step to data analysis and communication skills addressed in later cycles. The experience of gathering, organising, representing and analysing data also provides students with valuable insights into the processes behind communication of data and how these may be misrepresented in the media. These skills are necessary to develop students' skills as critical and informed citizens.

Materials

For Cycle 1 you may need:

- Counters
- Graph paper
- Decks of cards
- Maths mat or large grid on the floor
- Coloured pens or pencils
- String or wool

Key Language

Discrete data, tables, graphs, tallies, dot plots, stacked bar graph, stacked column graph, measures of centre, average, mean, median, quartiles, mode, bimodal, range, maximum, minimum, spread, outliers

Definitions

Discrete data: data organised in categories where items in each category may be individually counted.

Dot plots: similar to pictograph or column graph where each item of count is represented by a dot.

Mean: calculated measure of centre. Add all values and divide by the number of values

Median: counted measure of centre. Involves identifying the middle value of a set of data

Mode: measure of centre that shows the most common item or value (may also be bi-modal)

Outliers: data values that are separated from the majority of the data set. Outliers increase the distance between the maximum/minimum and the mean or median

Quartiles: measures of centre that show the halfway points between the middle value of the data set and the maximum/minimum

Range: difference between the maximum and minimum values in a data set

Stacked bar/column graph: graph which includes internal detail about the category. For example, further separation of a category by gender. In the context of Cycle 1, a stacked bar/column graph is used to place all collected data into a single strip to facilitate comparison between the categories and the whole sample and to facilitate the creation of a circle graph.

Statistical measures: measures to describe the central tendency and spread of the data. For example, mean, median, mode, quartiles, range, spread, outliers

Assessment

Anecdotal Evidence

Some possible prompting questions:

- What sort of data are you graphing? Are the items separate from each other and countable?
- What can you see when you look at the graph?
- What is the highest point (maximum value)? What is the lowest point (minimum value)? What is the difference between these values (range)?
- What is the most common value (mode)?
- Where is the middle of the data? If you sort the categories from least to most, and count in from each end, where is the middle? Is this nearer the maximum, minimum or halfway between?
- Is there any data that is a long way from other values?

Portfolio Task

This task involves students with selecting survey items to generate data to answer an overarching question as well as gathering data and representing the data graphically.

RAMR Cycle

Reality

Engage students in a discussion about previous data collection and graphing activities. Determine from students their initial understanding of the term *average*. Discuss populations (Australian Bureau of Statistics website has much useful statistical data) as a basis of statistical information or find examples of statistical information (perhaps football scores, cricket run rates, bowling statistics, medal tallies if near Olympic or Commonwealth Games). Discuss how statistics are used and what information they convey. This leads into need for calculations and interpretation of measures. Identify data gathered within the Investigation activity that entails discrete, countable items (e.g., number of pets, number of people living at home).

Abstraction

The abstraction sequence for this cycle uses discrete data sets that students gathered within the Investigation activity (e.g., number of pets). This abstraction phase will explore mean as it is likely the most common measure of centre that students have experienced. Other measures will be explored in the mathematics phase. A suggested sequence of activities is as follows:

1. *Kinaesthetic activity*. Use a maths mat or tape on the floor to create a large axis for students to stand on (generates a human dot plot). Discuss what labels would be needed for a graph of students' pets. Give each student a circle of paper to use as a marker. Have students stand in an appropriate square on the graph for their pet. Photograph the human graph (from as high as possible).
2. *Model with materials*. Have students place their marker in their square and step away from the graph so they can see it. Discuss the features of the graph that students have created (scale, axis labels, column labels, title) and reinforce the reason for these (Linking to Unit 01 and Unit 03). Have each student create a drawn version of the dot plot for later use.
3. *Introduce language and symbols.* Engage students with recreating the large human dot plot using counters on an A4 page axis. Discuss how many pets each student would have if the number of pets were distributed evenly between the people in the class. Ensure that students explain their thinking and how they worked it out.
4. *Model with materials.* Use the counter dot plots to share the taller columns with the shorter columns to make each column the same height. See how many are in each column when all are (as near as possible to the same height). Discuss other ways of finding how many in each column (e.g., could we count all the counters we have and divide it by the number of columns to see how many in each column?).
5. *Reinforce language and symbols.* Remind students of the action in counting all the counters altogether then breaking into the same number of parts as columns. See if students can generalise this to the formula for mean (adding all frequencies and dividing by the number of columns). Fractional answers may result. Ensure that students understand when a fractional answer for discrete data may be relevant or appropriate and when it may not (discrete items in general should not be partitioned although some can be e.g., food items, not living things).
6. *Reasoning.* Discuss when the mean is a useful measure of discrete data sets to know (e.g., when ranking teams according to score). Discuss also when the mean may be a misleading measure of centre (e.g., when there is a great range between highest and lowest value with most values clustered at one end or the other).

Mathematics

Language/symbols and practice

Identify other sources of discrete data of interest to students. These may include eye colour graphs, number of televisions at home or number of people living at home. A range of large datasets are available on the Australian Bureau of Statistics site for school use where students can gather already collated data. Discuss the various sources of discrete data and information available while also considering how the data may have been collected and the reliability of possible sources.

Explore other statistical measures – Range and mode

Return to the maths mat dot plot to explore range and mode. Sort the columns of data from least frequent pet to most frequent pet (and move column label with the column). Use a length of wool or string to highlight the **minimum** frequency across the graph and another length to highlight the **maximum** frequency across the graph. Discuss this as the **range** of the frequencies for this set of data. Discuss the visual shape of the graph. Consider the spread of the data within the range and identify if there is an even spread between the maximum and minimum frequencies or if the data is more concentrated in some areas than other.

Discuss which pet is most common (**mode**). If there is more than one type of pet with the same frequency, discuss bi-modality as two values that are most common. If there is only one mode, discuss what the mode might be if there were an increase in one of the other frequencies to make it the same. Also discuss who might use this information (e.g., if a pet store is catering to this class, what items do they need more or less of?). Compare the mode of the data with the mean. Discuss which of the measures is more useful when considering the *number of pets* data set.

Explore other statistical measures – Median and quartiles

Have students consider what the middle (**median**) frequency is for their data. (Find median by counting in evenly from each end of the sorted graph.) Clearly label the middle value as **median**. (If there are even numbers of students, the median will be halfway between the two central students' frequencies.) Find the middle frequencies between the median and each end. Clearly label these as **quartiles**. Discuss with students what these values might indicate on analysis.

Compare the median with the mode and discuss the differences between these two measures. Discuss when it may be more useful to know the exact middle of a set of data than the most common frequency.

Compare the median with the mean. Discuss with students why these values may be different for the given set of data. Consider in relation to the spread of the data and ask students to suggest why unevenly spread data may result in differences between the mean and the median.

Explore the effects of outliers

Considering the *number of pets* data set, reconsider the notion of spread. Ask students to identify where a frequency is a long way higher or lower than the other frequencies. Explore what happens to the measures of centre if outliers are removed from or added to a set of data.

Explore relationships between mean and median. Investigate what happens to the mean and median if further values are added. This can be done by working through further abstraction cycles using foot lengths or arm lengths and calculating the mean and median as lengths are added to the list (i.e., start with two values and calculate the mean and median [will be the same], add a value and recalculate, record changes, …). Discuss the effects on the mean and median of increasing the range of values in the data sample. Explore what happens when more values are added within the range and above or below the median, outside the initial range.

Compare different graph types

Discuss previous graphing experiences with students (bar or column graphs with categorical data). Discuss with students what statistical measures are apparent when bar or column or dot plots are used. Generate circle graphs of the *number of pets* data (*Resource 10.1.1: Construct a circle graph – protractor, Resource 10.1.2: Construct a bead string circle graph* and *Resource 10.1.3: Interpreting circle graphs* may be useful if students are unfamiliar with circle graph construction). Ensure that students can recognise that circle graphs show relative size of a category compared to other categories and the whole set, and may be indicative of mode, while bar or column graphs and dot plots provide more detailed information that allows for analysis of measures of centre and spread.

Resource

Resource 10.1.1 Construct a circle graph - protractor
Resource 10.1.2 Construct a bead string circle graph
Resource 10.1.3 Interpreting circle graphs

Further activities to consolidate or check student understanding of mean, median and mode include *Resource 10.1.5: Finding what is typical median activity, Resource 10.1.6: Hat sizes median activity* and *Resource 10.1.7: Mean, mode and median card games*. *Resource 10.1.4: Grid papers* may be useful printable resources for students to use for graphing.

Resource

Resource 10.1.4 Grid papers
Resource 10.1.5 Finding what is typical (median) activity
Resource 10.1.6 Hat sizes (median) activity
Resource 10.1.7 Mean, mode and median card games

Reflection

Check the idea

Smarties graph

Multi-packs of Smarties or Skittles may be used to explore graphing and statistical measures. Give each student their own smaller packet of Smarties or Skittles. Students should sort their Smarties into colours and create a dot plot from least of a colour to most and determine the range, mode, median and mean for their packet.

On the whiteboard, collate a class set of Smarties or Skittles colours (this data set should be saved for further comparison in Cycle 3). As each set of values is added to the class set, have students recalculate the mean, median, mode and range. An excel spreadsheet can be useful for generating a quick graphic.

Apply the idea

Choose a topic to gather data on, for example, class shoe sizes. Collate the data on the whiteboard in a table. Each student should generate their own dot plot of the data (disregard the difference in actual sizing between boys and girls shoes). For example:

					0				
					0				
				0	0				
				0	0	0			
			0	0	0	0			
			0	0	0	0			
			0	0	0	0		0	
	0		0	0	0	0	0	0	0
Shoe size:	3	4	5	6	7	8	9	10	11

Calculate mean, median and mode, range and quartiles for the class's shoe sizes. Write a short paragraph to state which measure would be most useful for the following audiences:

- Shoe retailer
- Designer of a shoe rack for the class

Extend the idea

Extend from the application activity to explore the generation of a set of data with the same measures of centre as the class but different actual shoe sizes. Suppose a class with the same number of students had size 3, 4, 5, 9, 10, 11 shoes only with no in-between sizes. Construct a data set of shoe sizes for this class which has the same mean as your class. Answer the following:

- What happens to mean, median and mode compared to the class data set?
- Draw a dot plot. How is it different from the class dot plot?
- Is it possible to adjust the data set to have the same median as the class data set? How?
- If two giants joined the class with shoe sizes of 25 and 27, what will happen to the mean, median and the mode?

What would the mean, median and mode be if some students wore size 2 shoes? Adjust the data set by adding five students with size 2 shoes. Recalculate the mean, median and mode.

Teacher Reflective Notes

This page is provided for you to record any notes with respect to resources you found useful, additional resources, activities and/or models that worked well/not so well.

Name:_____ Date:_____

Can you do this? #2

1. Here is data gathered by twelve girls for arm length.

Girls	01	02	03	04	05	06	07	08	09	10	11	12
Arm length cm	65	34	38	38	63	38	54	34	57	40	50	42

(a) Write down the range of the girls' arm lengths. _____

(b) Use the table to work out the intervals and to tally the frequency of arm lengths in each interval.

Tally									
Interval									
Centre									

(c) Draw a histogram of the girls' arm lengths.

_____

(d) Write down the mode of the girls' arm lengths. _____

(e) Calculate the mean of the girls' arm lengths.

_____

_____

Two girls with arm lengths of 50cm join the group.

(f) Will the mode of the girls' arm lengths change? _____

(g) If yes, write down the new mode/s for the girls' arm lengths.

_____

(h) Will the mean of the girls' arm lengths change? _____

(i) Calculate the changed mean for the girls' arm lengths.

_____

Obj. 10.2.4
a)i □
a)ii □

Obj. 10.2.1
b)i□
b)ii□
b)iii□
c)i□
c)ii□
c)iii□
c)iv□
c)v□
c)vi□

Obj. 10.2.5
d)i □
d)ii □

Obj. 10.2.6
e)i □
e)ii □
e)iii □

Obj. 10.2.9
f) □
g)i □
g)ii □
g)iii □
g)iv □
h)i □
i)i □
i)ii □
i)iii □

2. The following data demonstrates students' heights. Students are all in the same age group.

Person	John	Mary	Frank	Bruce	Sue	Tom	Ben	Amy
Height (cm)	135	121	142	149	118	139	144	135

(a) Construct a stem and leaf graph of the students' heights.

Obj. 10.2.2
a)i□
a)ii□

(b) Write down the median of the students' heights. _____

Obj. 10.2.5
b)i □
b)ii □

(c) Write down the quartiles of the students' heights.

First quartile:_____Third quartile: _____

Obj. 10.2.10
c)i□ii□
iii□iv□

(d) Construct a box plot and whiskers graph of the students' heights.

Obj. 10.2.3
d)i □
d)ii □
d)iii □
d)iv □
d)v □
d)vi □
d)vii □

(e) Calculate the mean of the students' heights.

Obj. 10.2.6
e)i □
e)ii □
e)iii □

(f) Comparing the mean, median and the mode of this data, what conclusion could you draw about this group of students?

Obj. 10.2.7
f) □

(g) You want to represent students' heights on a graph so that the audience can identify individual data values. Circle the graph type you would use.

Obj. 10.2.8
g)i□
g)ii□
g)iii□

Name:_____ Date:_____

(i) histogram (ii) stem and leaf (iii) box plot and whiskers

Cycle 2: Continuous Data – Shape and Measure

Overview

Big Idea

Within this cycle students will explore the graphical representations and statistical measures of continuous data including histograms (continuous data sorted into categories for representation), stem and leaf graphs and box plot and whiskers. Students will also be engaged with connecting the analysis and description of the data from the graphical representations to the calculation of statistical measures of range, spread, mode, median, quartiles and mean.

Objectives

By the end of this cycle, students should be able to:

10.2.1 Construct histograms from continuous data. [7SP170]
10.2.2 Construct stem and leaf plots from continuous data. [7SP170]
10.2.3 Construct box plot and whiskers graphs from continuous data. [10SP249]
10.2.4 Identify range and spread of continuous data from graphical representations. [7SP171]
10.2.5 Identify mode and median of continuous data from graphical representations. [7SP171]
10.2.6 Calculate mean of continuous data sets. [7SP171]
10.2.7 Describe and interpret graphical representations of continuous data sets using terms including 'skewed', 'symmetric' and 'bi modal'. [9SP282]
10.2.8 Compare a range of data displays of continuous data sets. [9SP283]
10.2.9 Investigate the effect of individual data values, including outliers, on the mean and median of continuous data sets. [8SP207]
10.2.10 Identify quartiles and interquartile range of continuous data from graphical representations. [10SP248]

Conceptual Links

Classifying and sorting skills underlie the gathering and organisation of data. Students will also reinforce previously developed counting and graphical representation skills throughout this cycle.

Analysing and interpreting continuous data using graphical representations and statistical measures in answer to questions is an extension on the ideas developed within Cycle 1 with further graphical representations including histograms, stem and leaf plots and box and whisker plots. Representation and analysis of continuous data is extended on in later cycles with comparison of data sets, misrepresentation of data and correlation.

Materials

For Cycle 2 you may need:

- Calculators
- Streamers
- Counters
- Real-world data sets
- Unifix cubes
- Sticky tape

Key Language

Continuous data, histograms, stem and leaf graph, box plot and whiskers graph, skewed, symmetric, bi modal

Definitions

Box plot and whiskers: graphical representation which indicates the maximum, minimum, median, quartiles and, in some cases the mean. The section between the first and third quartile is boxed to represent the middle half of the dataset. Whiskers represent the upper and lower quarter of the data. Lengths of whiskers give an indication of the overall spread of the data (particularly the existence of outliers). Analysis of the position of the mean relative to the median may also be considered with this representation

Continuous data: data usually comprised of measures which are not always whole numbers (as for discrete data) but may be approximated to single values for comparison and representation

Histograms: continuous data made discrete by approximation represented on a bar/column graph with no spaces between the bars. Boundaries of bars represent the boundaries of the class or category of data. The height of the bars represent the frequency of values within the class. Individual data values are not apparent. For example, heights may be represented as ranges of 125cm – 130cm, 130cm – 135cm, 135cm – 140cm

Stem and leaf graph: graphical representation which shows a central spine of common and evenly spread values with the leaves comprised of the individual data pieces that fit in each class. This representation maintains greater perspective on the individual data than the histogram

Assessment

Anecdotal Evidence

Some possible prompting questions:

- What sort of data are you graphing?
- Are the items separate from each other and countable or are they measures? Does it make sense to have a part of a whole?
- What can you see when you look at the graph?
- What is the highest point (maximum value)? What is the lowest point (minimum value)? What is the difference between these values (range)?
- What is the most common value (mode)?
- Where is the middle of the data? If you sort the measures from least to most, and count in from each end, where is the middle? Is this nearer the maximum, minimum or halfway between?
- Is there any data that is a long way from other values?

Portfolio Task

This task involves students with calculating measures of centre from gathered data and determining which measure provides the most appropriate indication of the trend within the sample.

RAMR Cycle

Reality

Extend students' understanding of graphical representations and statistical measures to include suitable graphs for continuous data. Discuss sources of continuous data that may be graphed (Australian Bureau of Statistics website has much useful statistical data) such as heights, lengths, masses, rainfall, temperature, prices or costs of items, swimming or running times if near Olympic or Commonwealth Games. Discuss the uses of these statistics and what information they convey. Gather own personal measures, for example, height, foot length, arm length, head size (gathered during the Investigation activity) and make three paper streamer replicas of these measures (not crepe streamers as these will stretch). Ensure students label their streamer with attribute measured, length measured and their name (e.g., Ann Height 154cm).

Abstraction

The abstraction sequence for this cycle uses continuous data sets that students gathered within the Investigation activity (e.g., student heights). This abstraction phase will use a simple chart of students' heights to reinforce measures of centre. Other graphical representations including histograms, stem and leaf graphs and box plot and whisker graphs will be explored in the mathematics phase. A suggested sequence of activities is as follows:

1. *Kinaesthetic activity.* Students line up in order of height (against the wall or a whiteboard). Revise median with students (students should count in from each end to find the middle student).
2. *Model with materials.* Mark each student's height with a line on the whiteboard or wall (alternatively, students may pin one of their paper streamer heights to a pin board with the tail level with the floor). Clearly label the middle value as median. (If there are even numbers of students the median will be halfway between the two central students' heights.)

   Find the middle values between the median and each end. Clearly label these as **quartiles**.

   Identify the tallest value and run a line of string across the chart from left to right across all streamers. Repeat this process for the shortest value. Discuss the difference between the two measures as the **range**.

   Looking at the chart, have students identify any sections where students are the same height. Find the height that occurs most often. Discuss this measure as the **mode**. Label clearly.
3. *Connect to language and symbols.* List students' heights on the whiteboard. Link heights to scale on graph. Students make own graphs to represent class heights and mark in median, quartiles, range and mode as discovered. Record median and mode as numeric values.
4. *Kinaesthetic activity*. Return to original streamer graph of height. Stand the tallest and shortest student together side by side and ask how their height might be made the same or averaged. For example, shorter student might stand on a step while the taller student bends their knees to be shorter. Ask students what could be done to make everyone the same height without gaining or losing overall height. Some suggestions may be to take height off all taller people or add height to all shorter people.
5. *Model with materials*. Place students in groups of four to explore how they might make their spare height streamers the same length. This is time consuming but effective to build an intuitive understanding of the process behind calculating the **mean.** Students may shorten longer streamers to add to shorter streamers to find the average height. Discuss the actions taken in determining the mean. Explore ways of working with the values to make them all the same.

Discuss other possible strategies to make the streamers all the same length, for example, connecting all lengths together with sticky tape and then folding the long length into the required number of equal lengths to cut. This is easily tested in groups of four. Students should be made aware that they are finding the mean of their small group, and each student in the group needs to end up with streamers of the same length.

6. *Generalize formula:* Remind students of the action in putting all streamers together then breaking into the same number of parts as people. Link this to the formula for mean (add all measures and divide by the number in the sample). Connect to the experience of making counter columns all the same height in Cycle 1. Calculate the mean student height for the class (using a calculator). Discuss with students the possibility that they may have a fraction within their answer. Ensure that students realise that this is relevant and appropriate because this is continuous data and it is possible to have a partial measure (unlike discrete data which may or may not be able to be partitioned). Clearly mark the mean student height on the chart along with the mode, median, quartiles and range.
7. *Explore concept of best measure of centre:* Using the student height data, compare the values obtained for mode, median and mean with the shape of the visual display. Discuss whether the graph is symmetric or skewed one way or the other. Discuss what conclusions may be reached about the data from these measures. For example, the mean is greater than the median because there are more tall people in the class; the mode is <height> because there is a group of people the same height; there is more than one mode (bi modal if two). Consider contexts where some statistical measures are more relevant. For example, the needs of designers for clothing, heights of doorways and living spaces, cabin sizes in planes. Discuss with students when mean might be more useful than median or mode in different cases and when it is most relevant.

Mathematics

For this Mathematics phase, the students' height data continues to be used to generate a range of graphs of the same data. This will facilitate student comparison between graph types and assist with understanding of which graph type may be more appropriate for a range of contexts.

Language/symbols and practice

Explore effects of outliers or change to the data set

Explore relationships between mean and median. Investigate what happens to the mean and median if further values are added. This can be done by adding students who were absent when the graph was created or including hypothetical newcomers to the class or possible transfers out of the class. In each case, recalculate the mean, median and mode as lengths are added to/removed from the list (i.e., start with two values and calculate the mean and median [will be the same], add a value and recalculate, record changes, …). Discuss the effects on the mean and median of increasing the range of values in the sample. Explore what happens when more values are added within the range and above or below the median, or outside the initial range at the upper and/or lower ends of the data.

Histograms

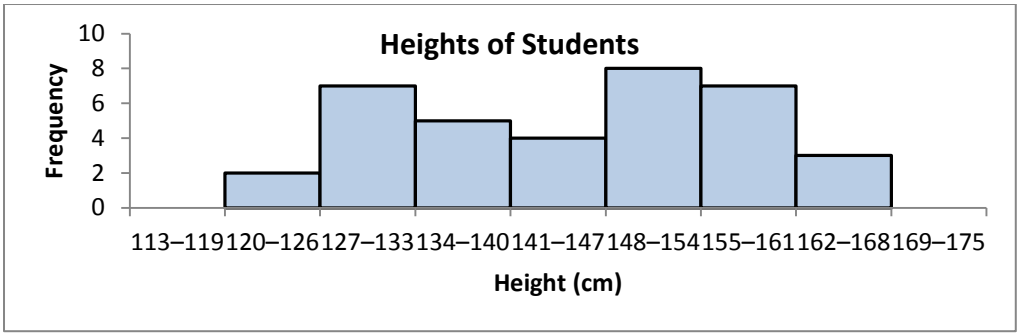
Use the previously generated students' heights chart. List the highest and lowest height. Subtract these two values to find the **range** of the heights. Using this range, determine an interval size that separates the data into sections or intervals (approximately seven sections or intervals is common). For example, a list of heights from 121 cm to 165 cm (a range of 44 cm) could be divided into seven intervals of 7 cm (it is important to have the interval length an **odd number** so that there is a central number to define each bar), this example would result in the following intervals:

120 – 126; 127 – 133; 134 – 140; 141 – 147; 148 – 154; 155 – 161; 162 – 168

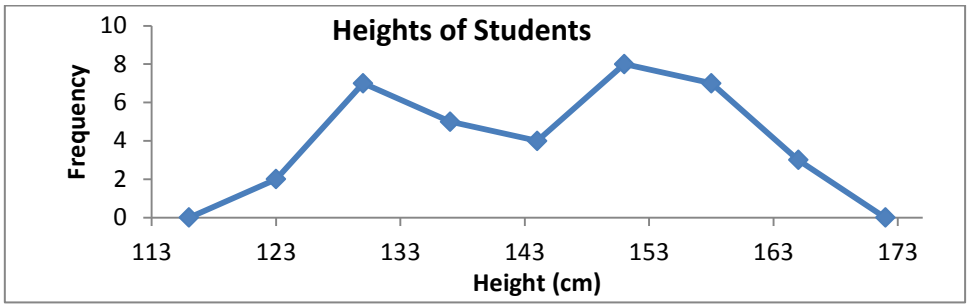
Include the previous and following intervals to this list (so that the ends of the range of data can be shown by the zero values). For example 113 – 119 and 169 – 175 would be added to the above list. Create a table as below and tally into each section the heights that are present in that interval. For example, see the possible tally below:

Interval	113–119	120–126	127–133	134–140	141–147	148–154	155–161	162–168	169–175
Tally		II	𝍸 II	𝍸	IIII	𝍸 III	𝍸 II	III	
Frequency	0	2	7	5	4	8	7	3	0

Use this data to draw a bar graph where bars are connected, as the example below shows.



Convert the bar graph to a line graph (remember the zeros at the end) as the example shows.



Engage students in reflective thinking about the relevance and importance of histograms. For example, How are these different from bar graphs? Do we need this special name? Are they just bar graphs with the bars touching? What information can we still read? What information has been lost between the chart with everybody's height on it and the histogram?

Resource Resource 10.2.1 Constructing histograms
Resource 10.2.2 Grid papers

Stem and leaf graphs

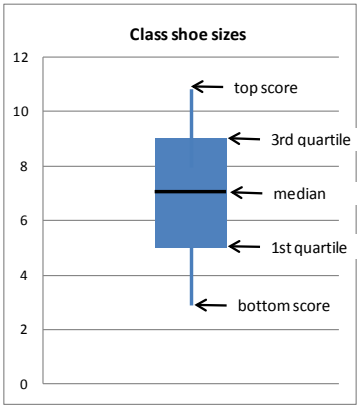
Stem and leaf graphs are similar to histograms in that they separate data into intervals (become the stem) but also show distribution within intervals (leaves). It is also possible to have students arrange themselves into a stem and leaf height graph physically as follows:

1. Use a strip of masking tape for the "stem" (normally vertical), mark sections on the tape for the stem (e.g., 11, 12, 13 and so on). Students stand to the side of the section in which their height belongs, going out from the stem – in order from shortest to tallest with the shortest beside the stem. This is a one-sided, made-with-bodies stem and leaf graph.
2. Get students to place a label with the "end part" of their height on it where they are standing, step back and view what they have made, and then make a pen-and-paper copy of the stem and leaf graph. Discuss the language used when constructing stem and leaf graphs.

Resource Resource 10.2.3 Constructing one-sided stem and leaf graphs

Box plot and whiskers

Use students' height data to construct a box plot and whiskers graph. Find the minimum value, maximum value, median and quartiles. Mark these on the graph as on the right:



Discuss what information about the population or sample is shown on the graph and what information is missing or not apparent.

Resource 10.2.4 Constructing box plot and whisker graphs has further notes for teachers.

Resource Resource 10.2.4 Constructing box plot and whisker graphs

Connections

Compare different graph types

Students should now have generated four different types of graph using their measured heights. Compare and contrast the graphs in terms of clarity of data representation and the information that can be recognised visually from each representation.

Language/symbols and practice

Further activities to consolidate or check student understanding of mean, median and mode may be completed using students' foot lengths or arm lengths (data collected during the Unit 10 Investigation activity or by completing a range of data gathering, representing and analysing activities such as *Resource 10.2.5: Drawing instruments mean activity*. *Resource 10.2.2: Grid papers* has been included here in case printable graph paper is required.

Resource Resource 10.2.2 Grid papers
Resource 10.2.5 Drawing instruments (mean) activity

Reflection

Check the idea

Engage students with *Resource 10.2.6: Measuring smiles (mode) activity* and/or *Resource 10.2.7: Head sizes (median) activity*. Ensure that students are able to calculate statistical measures, identify which measure will be most appropriate and select a suitable graph type to represent the data.

Resource Resource 10.2.6 Measuring smiles (mode) activity
Resource 10.2.7 Head sizes (median) activity

Apply the idea

Choose an investigation like one of the examples below for students to explore. Let the students work out their own way to tackle the question – discuss and reflect upon the different strategies.

- How many metres in 10,000 steps?
- Does Barbie have human dimensions?
- How far does an origami frog jump?
- What is reaction time?
- What is the best design for a paper aeroplane?

Extend the idea

Generalise the effects of changes in the data set on the mean and median.

Explore ways of making the mean and median of a set of figures the same and different. For example, how might the mean and median measures in the students' height data be made the same, or shifted further apart? What height students might need to be included or excluded from the class?

Explore the effect of skewed data on mean values. Reverse from a given set of values to indicate what the data set might look like. For example:

The temperatures for a week were:

S	M	Tu	W	Th	F	S
23	24	23	22	23	24	22

1. What are the mean, median and mode? Did you notice that all three are the same?
2. Can you make the mean different from the median and mode by changing one or more temperatures? Which ones and by how much? Can it be done with a single change?
3. Modify the data set to make a distribution with a low mean.
4. Modify the data set to make a distribution with a high mean.
5. Draw the frequency graphs for (2), (3) and (4). How are they different?

Teacher Reflective Notes

This page is provided for you to record any notes with respect to resources you found useful, additional resources, activities and/or models that worked well/not so well.

Name:_____ Date:_____

Can you do this? #3

Here is data gathered by six girls about their favourite colour.
(R-Red, B-Blue, Y-Yellow, G-Green)

Girls	G1	G2	G3	G4	G5	G6
Favourite colour	R	Y	Y	Y	B	R

Here is data gathered by six boys about their favourite colour.
(R-Red, B-Blue, Y-Yellow, G-Green)

Boys	B1	B2	B3	B4	B5	B6
Favourite colour	G	G	B	Y	R	G

1. (a) Construct a stacked bar graph to show the girls and boys favourite colours.

(b) Write down the mode of the whole group's favourite colour.

(c) Write down the mode of the girls' favourite colours.

(d) Write down the mode of the boys' favourite colours.

(e) What does this tell you about this group's favourite colours?

Obj. 10.3.2
a)i. ☐
a)ii. ☐
a)iii. ☐
a)iv. ☐
a)v. ☐
a)vi. ☐

Obj. 10.3.6
b) ☐
c) ☐
d) ☐
e) ☐

Here is data gathered by six girls about their arm lengths in centimetres.

Girls	G1	G2	G3	G4	G5	G6
Arm length cm	65	34	38	38	63	38

Here is data gathered by six boys about their arm lengths in centimetres.

Boys	B1	B2	B3	B4	B5	B6
Arm length cm	64	54	77	60	70	62

2. (a) Construct a back-to-back stem and leaf graph to show boys' and girls' arm lengths.

Obj. 10.3.4
a)i. □
a)ii. □
a)iii. □

Obj. 10.3.7
b) □
c) □
d) □
e) □
f) □

(b) Calculate the mean arm length of the whole group.

_____

(c) Calculate the mean arm length of the boys.

_____

(d) Calculate the mean arm length of the girls.

_____

Obj. 10.3.6
g)i □
g)ii □
g)iii □
h) □

(e) Are the means the same or different? _____

_____

(f) What does this tell you about the arm lengths of these students?

_____

_____

(g) Looking at the shape of the graph for each group of students, is the data skewed or symmetric? Boys _____

Girls _____ Whole group _____

(h) How do the girls' arm lengths affect the mean of the whole group of students? _____

_____

Cycle 3: Comparison of Data Sets

Overview

Big Idea

In Cycle 1 and Cycle 2 students explored single data sets. Within this cycle, students will explore the representation of multiple data sets on the one graph for comparison and representation of single data sets that may also be split into smaller categories (e.g., students' heights have been graphed as a single dataset but could be separated by gender into two datasets. Graph types will include side-by-side column graphs, stacked bar/column graphs, multiple histograms, back-to-back stem and leaf graphs and stacked box plot and whiskers graphs.

Objectives

By the end of this cycle, students should be able to:

10.3.1 Construct side-by-side column graphs. [6SP147]
10.3.2 Construct stacked bar/column graphs. [6SP147]
10.3.3 Construct multiple line histograms. [9SP282]
10.3.4 Construct back-to-back stem and leaf graphs. [9SP282]
10.3.5 Construct stacked box plot and whiskers graphs. [10SP249]
10.3.6 Interpret and describe data sets, using terms including 'skewed', 'symmetric' and 'bi modal'. [9SP282]
10.3.7 Compare multiple data sets represented on a single graph using statistical measures. [9SP283]
10.3.8 Interpret and compare data sets represented as stacked box plot and whiskers graphs. [10SP249]
10.3.9 Interpret and compare a range of displays of multiple data sets. [10SP250]

Conceptual Links

Through this cycle students should be encouraged to link back to previous graphing experiences including gathering, organising and representing data as well as ideas surrounding the generation of suitable scales on graphs. Counting and operation skills are also able to be practised and reinforced.

Comparison of data sets using visual representations and statistical measures are skills needed to extend to decision making and inference in Cycle 4.

Materials

For Cycle 3 you may need:

- Data set activity
- Measuring tools
- Graphs generated in Cycle 1

Key Language

Comparison, data sets, back-to-back stem and leaf graph, stacked box plot and whiskers graph, side-by-side bar/column graphs, stacked bar/column graphs

Definitions

Back-to-back stem and leaf: graphical representation which shows a central spine of common and evenly spread values with the leaves comprised of the individual data pieces that fit in each class. This representation maintains greater perspective on the individual data than the histogram and is able to show two sets of data with a related theme. For example, heights of a group separated by gender

Side-by-side bar/column graphs or stacked bar/column graphs: bar/column graphs for subsets of data within a larger population

Stacked box plot and whiskers: graphical representation which indicates the maximum, minimum, median, quartiles and, in some cases the mean. The section between the first and third quartile is boxed to represent the middle half of the dataset. Whiskers represent the upper and lower quarter of the data. Stacks of box plots can be used to compare similar datasets from within a population

Assessment

Anecdotal Evidence

Some possible prompting questions:

- What do you notice about the set of data?
- What is the same/different about the two parts of the data set?
- Compare the mean/median/mode of each part of the data set. How are the same/different?
- What conclusions can you draw about the two parts of the data set?

Portfolio Task

This task provides students with the opportunity to compare their class data with a random sample of twenty other students (data from ABS website).

RAMR Cycle

Reality

Cycle 1 and Cycle 2 activities provide an ideal starting point for this cycle and can be explored further in practice activities. Remind students of their data gathering activities, the data sets explored and data gathered. Focus on the students' height data. Discuss ways of representing the data to show more detail. For example, visual representations explored so far only show frequency of students' heights. Consider other classifications that could be applied within the data set (e.g., gender or age).

Abstraction

The abstraction sequence for this cycle engages students with more detailed data representations that show subsets within a larger data set for comparison. For this cycle, side-by-side bar/column graphs will be explored. Other graphical representations will be explored in the Mathematics phase. A suggested sequence of activities is as follows:

1. *Kinaesthetic activity.* Using students' height data, generate suitable ranges for representing the data set as a bar/column graph. Give each student a sticky note to write their height onto (different colours for boys and girls). Have students' stand on a maths mat or large floor grid in the appropriate bar/column for their height. Separate each bar/column into two to create a bar/column of boys and girls for each height range. Students should place their sticky note down on the graph at their feet before they step away.
2. *Represent with materials.* Use sticky notes to generate a side-by-side bar/column graph of students heights on the whiteboard. Unifix cubes may also be used to generate a physical representation of the side-by-side bar/column graph.
3. *Represent on paper.* Recreate sticky note/unifix cubes graphs on paper.
4. *Connect to language.* Discuss the shape of the graph now that it has been separated by gender. Reinforce statistical measures explored (mean, median, mode, range) by locating these for boys, girls and the whole group. Ensure that students recognise the need to add the frequency of girls and boys together to generate the frequency of the height for the whole group.
5. Discuss the differences in the statistical measures between boys, girls and the whole group. Consider the range and spread of each data set. Discuss the effect of adding the boys' heights to the graph of girls' heights on the location of mean, median, mode for the whole group.

Mathematics

Language/symbols and practice

Once students have developed skills and language to compare more than one data set on a single visual representation, explore other graphical representations that may be used in this way including stacked bar/column graphs, double-sided stem and leaf plots, stacked box plot and whiskers graphs and multiple histograms.

Stacked bar/column graphs

Stacked bar/column graphs are very similar in construction to side-by-side bar/column graphs but the two data sets are stacked vertically instead of side by side. Ensure that students recognise that the frequency of the whole group and the lower part of the stack may be read directly from the scale on the graph. However, the upper part of the stack will need to be determined by finding the difference between the lower stack and the whole group.

Double-sided stem and leaf plots

Double-sided stem and leaf graphs are similar to one-sided stem and leaf graphs but in that they separate data into intervals (become the stem), show distribution within intervals (leaves), but also show separate groups within intervals on either side of the stem. It is also possible to have students arrange themselves into a double-sided stem and leaf height graph physically as they did for a one-sided stem and leaf graph.

1. Use a strip of masking tape for the "stem" (normally vertical), mark sections on the tape for the stem (e.g., 11, 12, 13 and so on). Students stand to the side of the section in which their height belongs, going out from the stem, with boys on one side and girls on the other – in order from shortest to tallest with the shortest beside the stem. This is a double-sided, made-with-bodies stem and leaf graph.
2. Get students to place a label with the "end part" of their height on it where they are standing, step back and view what they have made, and then make a pen-and-paper copy of the stem and leaf graph. Discuss the language used when constructing stem and leaf graphs.
3. Discuss the location/calculation of statistical measure on the double-sided stem and leaf graph.

Resource Resource 10.3.1 Constructing double-sided stem and leaf graphs

Stacked box plot and whiskers graphs

Use students' height data to construct a box plot and whiskers graph. Find the minimum value, maximum value, median and quartiles.

Separate the data into boys' heights and girls' heights. Generate individual box plots and whiskers for each subset of the data. Place on the same axis as for the whole group. Discuss what information about each subset of the sample is shown on the graph and what comparisons may be made.

Resource Resource 10.3.2 Constructing stacked box plot and whisker graphs

Multiple histograms

Use the side-by-side column graph constructed in the abstraction phase as a base for generating multiple histograms. Use different colours for the histogram line for each dataset.

Resource Resource 10.3.3 Constructing multiple histograms
Resource 10.3.4 Grid papers

Additional practice

Use the additional data sets gathered during the Investigation activity to practise generating multiple graph sets from each collection of data and analyse the relationships between the parts of the data set and the relationships between the parts of the data and the whole set.

Reflection

Check the idea

Use *Resource 10.3.5: Data set activity* as a quick set of data similar to what students have engaged with during the Abstraction and Mathematics phases of the cycle. Students will need to decide which data they will use, how to best display their data and analyse the relationships between the parts of the group and the whole group.

Collate a class data set to further compare with the sample data set. Keep these graphs and statistical measures for use in Cycle 4.

Resource Resource 10.3.5 Data Set Activity

Apply the idea

Choose an investigation like one of those in the examples below. Try to make it relevant and motivating for your students. Let the students work out their own way to tackle the question – discuss and reflect. Use every opportunity to direct attention to and reinforce the movement from statistical literacy to statistical reasoning, and take every opportunity to meet the four challenges and use the three types of questions.

- How much do we spend at the fete?
- What is a typical hand span?
- How much time do we spend on social media?
- How many commercials do we watch?

Alternatively, choose a topic that has rich online data, for example, crime statistics, road safety statistics, weather statistics, sport statistics, TV ratings. Engage students in a discussion to see what problems are posed.

Extend the idea

Compare and contrast graphical representations

Compare and contrast the multiple graphs constructed. Discuss which graphs show the data most clearly and which statistical measures are most easily compared with each graphical representation. Consider what contexts may be most applicable for each graph type and where these graphs are usually seen.

Teacher Reflective Notes

This page is provided for you to record any notes with respect to resources you found useful, additional resources, activities and/or models that worked well/not so well.

Name:_____ Date:_____

Can you do this? #4

Here is data gathered by six girls about their favourite colour.
(R-Red, B-Blue, Y-Yellow, G-Green)

Girls (Sample G)	G1	G2	G3	G4	G5	G6
Favourite colour	R	Y	Y	Y	B	R

A different group of six girls gathered data about their favourite colour.
(R-Red, B-Blue, Y-Yellow, G-Green)

Girls (Sample F)	F1	F2	F4	F4	F5	F6
Favourite colour	G	G	B	Y	R	G

1. Draw a side-by-side bar graph of Sample F and Sample G favourite colours.

Obj. 10.3.1
1. i. ☐
ii. ☐
iii. ☐
iv. ☐
v. ☐
vi. ☐

2. Compare the two samples.
   (a) What is the mode of Sample G girls' favourite colours? _____
   (b) What is the mode of Sample F girls' favourite colours? _____
   (c) How are the two samples the same or different? _____

Obj. 10.4.1
2. a) ☐
b) ☐
c) ☐

   (d) Is it possible to predict colour preference of girls from these two samples? _____ Why/why not? _____
   (e) What would make the sample more effective? _____

Obj. 10.4.3
d)i. ☐
d)ii. ☐
e) ☐

Here is data gathered by six girls about their shoe sizes.

Girls (Sample G)	G1	G2	G3	G4	G5	G6
Shoe size	7	3	5	4	8	7

A different group of six girls gathered data about their shoe sizes.

Girls (Sample F)	F1	F2	F3	F4	F5	F6
Shoe size	6	4	7	4	6	5

The girls are all the same age.

3. (a) Calculate the mean shoe size for Sample G girls.

_____

(b) Calculate the mean shoe size for Sample F girls.

_____

(c) What is the variation between these two means? _____

(d) What conclusion can you draw from this comparison?

_____

(e) Are these effective samples for predicting the shoe sizes of girls in this age group? _____

(f) Calculate the mean shoe size for the two samples together as one group. _____

_____

(g) Select a sample of six shoe sizes from the whole group that have the same mean as the whole group. _____

_____

(h) What information other than age and shoe size would you need to more effectively predict shoe sizes for girls? _____

_____

_____

(i) How would you gather this information? _____

(j) What questions could you ask to find out all the information you need? _____

_____

_____

Obj. 10.4.1
3. a)i ☐
a)ii ☐
b)i ☐
b)ii ☐
f)i ☐
f)ii ☐

Obj. 10.4.7
3. c) ☐
d) ☐
g)i ☐
g)ii ☐
g)iii ☐

Obj. 10.4.3
e) ☐

Obj. 10.4.6
h) ☐
j)i ☐
j)ii ☐
j)iii ☐

Obj. 10.4.2
i) ☐

Cycle 4: Sample, Population and Inference

Overview

Big Idea

This cycle engages students with consideration of datasets as possible samples of data that may be applicable to larger populations as well as ideas of variation and uncertainty. Students need to be brought to an awareness of when it is possible to gather data for a complete population and when it is appropriate to gather data from a sample of the population that may be used to infer preferences for the population.

Objectives

By the end of this cycle, students should be able to:

10.4.1 Describe and interpret different sample sets for the same data in context. [5SP120]

10.4.2 Investigate census as a data collection technique. [8SP284]

10.4.3 Investigate sampling as a data collection technique. [8SP284]

10.4.4 Collect data using sampling. [8SP206]

10.4.5 Collect data from secondary sources. [9SP228]

10.4.6 Identify everyday questions and issues involving at least one numerical and at least one categorical variable. [9SP228]

10.4.7 Explore the variation of means and proportions of random samples drawn from the same population. [8SP293]

Conceptual Links

Through this cycle students should be encouraged to link back to previous graphing experiences including gathering, organising and representing data as well as ideas surrounding the generation of suitable scales on graphs. Counting and operation skills are also able to be practised and reinforced.

Notions of sample, population and inference can be further connected with notions of likelihood to better interpret and make sound decisions from data represented graphically.

Materials

For Cycle 4 you may need:

- Data set activity
- Measuring tools
- Graphs generated in Cycle 1

Key Language

Population, sample, infer, inference, predict, representative group, variation, uncertainty, likelihood, pose, predict, sample, random, population, infer, conclude, decide, certainty

Definitions

Infer/inference: valid conclusion drawn from known data

Population: the entirety of a group with a given or stated attribute

Sample: a portion of a group with a given or stated attribute. When used in statistics, a random sample may be chosen to represent a population. Judgements then need to be made about the effectiveness or representativeness of the sample from which inferences about the population may be made

Variation: differences between statistical measures between samples of a population

Assessment

Anecdotal Evidence

Some possible prompting questions:

- What information do you know?
- What does the data tell you?
- What statistical measures are there?
- Are these samples similar or different?
- How are they different?
- How much are they different by?
- If these samples are similar, does this make them more or less likely to be representative of the population?
- If these samples are too different, how can you select a better sample of the population?

Portfolio Task

This task provides students with the opportunity to compare their class data with a random sample of twenty other students (data from ABS website).

RAMR Cycle

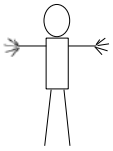
Reality

Cycle 1 and Cycle 2 activities provide an ideal starting point for this cycle and can be explored further in practice activities. Remind students of their data gathering activities, the data sets explored and data gathered. Check that students can identify similarities and differences between the results of data collection. Discuss what variations in the data might occur if the students in the next classroom were surveyed. Consider what variations in the data might occur if compared to classes across South East Queensland. A simple context for exploring variation and uncertainty is simple length measurement.

Abstraction

The abstraction sequence for this cycle engages students with gathering measurement data about a student. Variation in the measurements is analysed and considered as a basis for beginning inferential reasoning. A suggested sequence of activities is as follows:

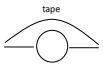
1. *Kinaesthetic activity.* Choose a student to stand with arms outstretched and organise all students to measure the student's arm span in centimetres (millimetres as decimals) from fingertip to fingertip. Ensure that students understand that it is acceptable for them to have differing results for this measure to prevent them from changing their responses.



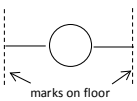
2. Record all the students' data on the board to nearest cm. Organise the students to graph this data in terms of individual lines or frequencies.
3. *Connect to language.* Use the data and the graphs to reinforce the following mathematical ideas: range, median, outliers, certainty, and mode (typical or most common).
4. Ask the students to make decisions from the data to identify: (a) What is the student's arm span?; (b) What is the variation and why is it occurring?; (c) What is certain/uncertain?.

In particular, ask what the students would do if all they had was the data and no way to re-measure the student's arm span. What do they think is the "best answer" and why? This activity can lead to good discussion on data such as: What is the middle?; What is the average?; and, Are there outliers?; What are these?; and, Why would they occur?. Finally, asking the students to come to a conclusion or consensus just from the data can lead to great discussion justifying different outcomes. This then is an opportunity to discuss the idea of error but, even more importantly, the idea of uncertainty.

Points for discussion: In most examples of this activity, there is a wide variation in measures. This gives an opportunity to discuss errors in measurement and the way in which different ways of measuring may lead to different measures. For example, the following can lead to error:

(a) 

Measure too long because of bend in measuring tape.

(b) 

Discuss possible accuracy of different ways of measuring (e.g., laying the student on the ground, marking lines on the ground and measuring between these lines).

(c) Students not accurate when measuring with a tape – particularly if they have a 1 m tape and it has to be moved for the overall measure.

(d) The student moves their arms between different students measuring – the arms could move backwards and forwards, after a time tiredness may make the arms sag so they are not at right angles to the body, the fingers may curl, and so on.

Mathematics

Language/symbols and practice

Foci of practice activities

There are four foci for consolidating informal ideas surrounding error, uncertainty and decision making. Teachers need to choose which of these is/are appropriate for their students:

Ensuring basics. It is important to run activities that build the abilities that underlie inference. The first of these includes gathering data, recording data, graphing data and describing data. For example, it is possible to gather all students' shoe sizes by using shoes to make the graph. This allows for recording (tallies and tables), graphing (bar graphs and frequency bar graphs) and informally discussing what "most students' size is" (e.g., centre, average or typical) and what are unusual sizes (e.g., range, outliers). The second of these is to return to measuring. This leads to discussion of errors, how they can be made and how to take account of them.

Making decisions. It is important, after or during basics, that activities like measuring arm span are undertaken to (a) extend discussion to **error and uncertainty**; (b) build ability to **make decisions** from the data; and (c) **defend decisions** from the data in relation to the specifics of the measuring.

Posing problems and devising data gathering. As well as making decisions, it is important to build the students' ability to work from the problem only. So we need to build the ability to pose questions and devise ways to gather data for their answers. We need to reduce support for the students – just ask a question like "How far do we jump?" and allow the students to work out ways to gather data for this. Then, justification for the inferences made are not just in terms of data but also in terms of relevance of the data. This is an idea that is extended upon in *Cycle 5: Misrepresenting data.*

Building complexity. The introductory activity on measuring arm span is specific, yet filled with uncertainty. The next step that could be undertaken is to begin to add in extra uncertainty by asking not for a specific arm span but for a "typical" arm span.

Investigations

Choose an investigation like one of those in the examples below. Try to make it relevant and motivating for your students. Let the students work out their own way to tackle the question – discuss and reflect. Use every opportunity to direct attention to and reinforce the outcomes for this cycle and the four foci in the activities as appropriate to the students.

- How tall is John?
- How far do we jump?
- What is the best recipe for play dough?
- What kind of books do we like?
- What is the best design for an obstacle course?
- What makes a toy car go further – a steep or a low ramp?
- How long does it take to tie a shoelace?
- Do most kids in class have brown eyes?
- Are we getting better at skipping?

Consider investigations that have more than one way of arriving at a solution. For example:

- How many advertisements do they have on TV? Does it change for different programs?
- What is the most popular car colour?
- Who is the best player on a football team? [Class chooses team]
- Which is the best class from the maths test?

Development of Inferential Reasoning

Once students have developed statistical literacy, it is possible to extend their thinking to statistical reasoning. Statistical reasoning focuses on reasoning and making sense of data in context – it utilises data, graphs and statistics information to understand problems and situations. It is important to ensure that students are engaged with converting frequency data to percentages, using the language of inference (pose, predict, sample, random, population, infer, conclude, decide, certainty) as well as understanding the use of random samples to make predictions about a larger population. It is important to consider whether a sample is representative of the larger population to determine how appropriate it is to use the sample data to make inferences or predictions about the population.

The practicalities for this include shifting focus from data points (e.g., "Kym watches 10 hours of TV") to holism or characterising groups (e.g., "most of my class watch between 10 and 15 hours of TV a week"). This has been characterised as having the ability to "distinguish signal from noise" (like tuning a small radio). It is best undertaken through investigating questions by allowing students to draw inferences from data they have gathered.

It is important to go beyond routine questions and engage students with interpretation of information from data sets and graphs. Inference includes taking account of variation, predicting, hypothesising, and criticising. It has three components: generalising beyond the data, using data as evidence, and acknowledging uncertainty in the conclusion. There are three types of questions:

a) from the data (answers can be read directly from the graph);
b) between the data (involve comparing categories on the graph); and
c) beyond the data (students infer reasons why or predict from the graph).

As a reversing activity students can be asked to suggest what questions may have been asked to generate the data in the graph. Sharing questions and responses can lead to significant engagement with the data and may suggest further avenues that students can explore in generating their own data collections.

Reflection

Check the idea

Softball throwing

A class had to choose a representative for a softball throwing contest. Three children volunteered. Each volunteer was asked to make five throws which were measured with a trundle wheel to the nearest metre. The results were as shown on the table below.

Volunteers	**Their 5 throws (to the nearest metre)**
Rachel	28, 23, 22, 24, 27
Betty	24, 23, 27, 24, 27
Tony	23, 27, 29, 18, 26

(a) Who would be the best representative? Who is the most consistent? Who has the longest throw?
(b) What should our criteria be for selecting the best representative? Who has the best typical throw? How do we define typical? Is consistency important? Should we have measured more or less than five throws? Should bad throws be excluded? Is anything important lost in rounding to the nearest metre?
(c) Develop an argument for your choice.

Note 1: Students can be allowed to put information on tables, or to tally throws into sections, say 15–19, 20–24, 25–29, etc. if this helps them. Students can also graph the results and work out averages if this also helps. Encourage students to take into account the context (one big throw wins) and their analysis of data in their arguments for their representative.

Note 2: If students refine the way to represent the data in arriving at their inference (e.g., from a simple plot of points to a frequency bar graph), this is said to be an example of "transnumeration" (Wild & Pfannkuch, 1999) – "changing representations to engender understanding" (p. 227). This shows strong knowledge growth and increased understanding. Observe students to see if this happens.

Apply the idea

Canned food

There are 23 brands of baked sausage in plum sauce. Sixteen use microwaves to cook the sausages and sauce, and 7 use steam. The cans always display how they are cooked but stores always seem to have different brands or unbranded cans. To enable consumers to determine which type of can to buy, a consumer group tests all brands and marks them out of 20 on quality. The marks for each of the 23 brands are given in the table below.

Type of can	**Marks for all brands of that type**
Microwaved	7 4 18 5 4 12 9 3 4 3 16 9 5 13 7 18
Steamed	10 14 5 11 15 12 9

(a) Which type, microwaved or steamed, is best to buy when there is no brand? Which type has the best mark? Which type has the most consistent mark?

(b) Determine criteria for making a judgement. Develop an argument for your choice.

Note: Again let students use any form of table, graph or determination of average that they think will help them. Check for the misconception that the type with the highest score is best even if other scores are low – in this context, it could be argued that consistency is needed as exactly what is being bought is unknown.

Extend the idea

Choose an investigation like one of those in the examples below. Try to make it relevant and motivating for your students. Let the students work out their own way to tackle the question – discuss and reflect. Use every opportunity to direct attention to and reinforce the movement from statistical literacy to statistical reasoning, and take every opportunity to meet the four challenges and use the three types of questions.

- What year level has the healthiest lunch?
- What is the shortest shadow?
- How much do we spend at the fete?
- What is a typical hand span?
- Is rolling your tongue hereditary?
- How many commercials do we watch?

Alternatively, choose a topic that has rich online data, for example, something from statistics that can be found on the Internet such as crime statistics, road safety statistics, weather statistics, sport statistics, TV ratings. Engage students in a discussion to see what problems are posed. Investigate the question, infer and defend a conclusion.

Teacher Reflective Notes

This page is provided for you to record any notes with respect to resources you found useful, additional resources, activities and/or models that worked well/not so well.

Name:_____ Date:_____

Can you do this? #5

1. Describe the following ways of misrepresenting statistics.

   (a) Sample bias: _____

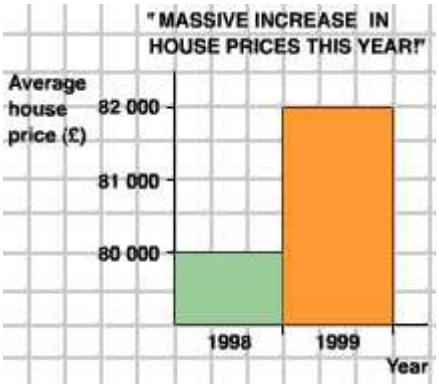
   _____

   (b) Missing information: _____

   _____

   (c) Irrelevant statistics: _____

   _____



2. Here is a graph showing average house price for 1998 and 1999.

   (a) Is this an appropriate representation of this data? _____

   (b) Why/why not? _____

   _____

   (c) Draw a graph that would fit the title:

   "House prices remain stable"

Obj. 10.5.1
a) ☐
b) ☐
c) ☐

Obj. 10.5.2
a) ☐
b)i ☐
b)ii ☐
c)i ☐
c)ii ☐
c)iii ☐
c)iv ☐

Cycle 5: Misrepresentation of Data

Overview

Big Idea

This cycle is different from the others in that it looks at how statistics can be used to misrepresent or "lie". This is a powerful way of reinforcing statistical inferential thinking, and also a powerful way to understand how persuasion may be given the credibility of mathematical correctness.

Objectives

By the end of this cycle, students should be able to:

10.5.1 Recognise ways that data may be misrepresented. [6SP148]

10.5.2 Identify appropriate data displays to represent data. [7SP170]

Conceptual Links

This cycle links graphical representations of data and statistical measures from previous experiences to further consideration of the appropriateness of representations and recognition of when graphical representations and statistical measures are being used to misrepresent, misinform or mislead the reader.

Materials

For Cycle 5 you may need:

- Data sets
- Poster paper
- Markers
- Rulers
- Examples of misrepresented or misleading graphs
- Examples of misrepresented or misleading use of statistics

Key Language

Misrepresentation, misleading, interpretation, critical literacy, sample bias, irrelevant statistics, wrong average

Definitions

Irrelevant statistics: statistics are quoted that are not related to the data or are not appropriate for the context

Misrepresentation: use of visual or statistical means to distort the readers' view of the data

Sample bias: use of a sample that is not representative of the population to make inferences

Assessment

Anecdotal Evidence

Some possible prompting questions:

- What is the title of this graph? Has it been worded to sway your opinion?
- Is the graph type appropriate for the data?
- Is the scale on the axes appropriate?
- Is the sample representative of the population or is it biased?
- How do you know?
- What would you do to fix the graph/representation?
- Are the statistics quoted appropriate for the question?
- Has an appropriate statistical measure been used?
- Is there other language in fine print that changes the measure used? (For example, prices quoted as “from <minimum value>”, “below <maximum value>”.)

Portfolio Task

Portfolio Task 10: Are we really healthy? does not directly include misrepresentation of data although it is possible to challenge students to further present their data to inform or misinform an audience about the class’s sleep, diet and physical activity statistics (perhaps as a newsletter item for parents).

RAMR Cycle

This cycle is different from the others in that it looks at how statistics can be used to misrepresent or "lie". This is a powerful way of reinforcing statistical inferential thinking, and also a powerful way to understand how persuasion may be given the credibility of mathematical correctness. The cycle provides information for teachers of methods for misrepresentation and then at activities and investigations that may be carried out with students.

Reality

Start with examples of graphs that are presented within media that include Headllines, Titles or Text that are designed to sway the readers' interpretation of the data. Start with simple and obvious examples and encourage students to look more deeply at the graphical representation of the data.

Resource Resource 10.5.1 Critical thinking scenarios

Abstraction

It is more effective to first teach students some of the ways that data may be misrepresented and then engage them with generating graphs that are designed to mislead or sway a reader's perception of the data. Darrell Huff in his book *How to lie with statistics* (Gollancz, London, 1969 – now out in Penguin) describes a series of ways in which statistics can misrepresent as follows:

Sample bias

Statistics showing central tendency for a large population are nearly always based on a small sample. Such samples may not represent the population as a whole and so bias the statistic. For example, Oxendorf University may trumpet that its graduates of 10 years standing earn on average $145 165 per year. Is this really correct? A closer look at how this statistic was arrived at may provide insight.

It is likely that the information on salaries was collected by replies to an emailed survey. In this case, only those graduates whose email addresses were known and who bothered to reply were included in the result. Furthermore, the statistic was calculated not on their actual salary but on what they said their salary was, with all the problems of lying and misunderstanding that this involves. The result is likely to be a bias towards more successful graduates, whose addresses are known and who have support staff to reply to the email (and who may inflate their salaries).

The statistics of average salary being $145 165 may also conceal large differences. What about deviation? The statement seems to imply that such a salary is what **every** graduate can expect!

We should always realise that all samples have a bias – towards people with more money, more education, more information and alertness, better clothing and more conventional and settled appearance – because these are the people who most interviewers feel more at ease with. So, when faced with a survey result, say "how was the information collected?".

Wrong average

In a factory or enterprise, there may be, for example:

- an owner earning $900 000/year
- a partner earning $300 000/year
- 2 assistant managers earning $200 000/year
- a sales manager earning $114 000/year
- 3 sales people earning $100 000/year
- 4 IT staff earning $74 000/year
- a foreman earning $60 000/year
- 12 workers earning $40 000/year

In this case, the **mode** (the wage/salary occurring most frequently) is $40 000 (the workers at the bottom of the range). The **median** (the wage/salary in the middle – 12 people earn more, 12 people earn less) is $60 000 per year (the foreman), while the **mean** (the average) is $114 000 per year, but only 4 of the 25 people earn more than this. Depending on the data, mean, median and mode may be the same or differ widely. Where there is a large range of values which contains a few very large values and many close together low values (which is typical of income statistics), the mean is high and the mode low. The median is the best measure of centre. So, when faced with an average, say "which average?".

Missing information

Statistically inadequate samples (small ones) can produce just about any result. Therefore if we ignore unfavourable samples, we can end up with an "independent laboratory test" certified by a "public accountant" proving just about anything. Four out of five people liking "Exo teeth licorice" can be just that – groups of five people were asked if they liked "Exo" until one group was found where four out of five did. The 30 previous groups are not included in the data sample.

The average alone can be misleading. A town with cold nights and hot days can end up with a delightful average temperature. We need information on range and deviation as well as average.

Words may be used ambiguously. What do statisticians mean when they say that "Tuffo" cleans twice as bright? Is this twice as bright as other cleaners or twice as bright as before cleaning?

Irrelevant statistics

Darryl Huff gives the old adage: "if you cannot prove what you want, demonstrate something else and pretend it is the same". Statistics about related matters are often used to support arguments for which there is no direct support. The statistic may be true, but not for the situation it is applied to.

For example, "laboratory controlled tests" may show that "Basho" destroys 9 out of 10 germs when used in high concentrations in a test tube, but will it do anything in your mouth in dilute concentrations? Young people from 16 to 21 may have more car accidents than the 50 to 55 age range, but this may be due to driving more. Accidents per person **per kilometres driven** may show that it is safer, for a 100 kilometre drive, to be with the young person!

Direct misrepresentation

Statistical data can be directly misrepresented. For example, juvenile delinquency figures can take a large jump when the courts change their recording procedures to count charges for group activities to each individual. Five youths stealing from a house can change from one offence to five break and enter offences, five being unlawfully on premises and five stealing offences (15 offences in all).

Percentages can make increases smaller or larger, depending on what you want, by choosing the appropriate base. Percentage increases can also look different to absolute increases. For example, someone taking a 50% pay cut from $800 to $400 per week is not going to be happy when that 50% is returned, but on the $400, i.e. to $600. We would not think it right if our 50% rabbit burger was made by mixing one rabbit with one bullock.

Resource Resource 10.5.2 Misrepresenting data

Mathematics

In the mathematics phase, students should practice asking questions of data and statistics to appropriately interpret when these are being misrepresented.

Language/symbols and practice

Talking back to statistics

The following five questions should be asked of statistics. Make these the basis of teaching.

Who says so? Look for bias, missing information, wrong measures, ambiguous statements and value-laden names (the prestigious university). What are the interests of the claimants?

How do you know? Look at how the statistic was calculated (the gathering of the data, sample size, type of data, calculations, and so on).

Is there anything missing? Is there range and deviation as well as average? What average? Try to look past percentages to raw scores. Look to see if key words are properly defined. Be wary of value-laden labels. Look at the groupings and categories. How were these selected?

Did someone change the subject? Watch for the switch from data to conclusion. More reports of cases do not mean more cases! What people say they do may not be what they do. Watch comparisons. Are they between different things?

Does it make sense? Many a statistic is false on its face. The magic of numbers suspends belief. Be wary of the decimal for the poverty line for a family of four. This is near impossible to directly measure. It has to be calculated from estimates.

Reflection

Check the idea

"Against the assertion" poster.

The world is full of assertions and sayings. For example, "there is more crime now than in the past"; "a stitch in time saves nine"; "the unemployed are dole bludgers"; "many hands make like work"; "people are less friendly than in the past".

Choose an assertion that most people agree with. Gather data regarding the assertion, represent the data and draw inferences from the data, to make a poster that disagrees with the assertion.

(a) Look at Huff's five ways of misrepresenting. For example – gather data from a sample which is biased by choice of sample or choice of interviewer or statements in the interview; reinterpret data so that there is a different result (e.g. look at accidents per km driven not total accidents, or look at crimes per person not crimes on their own); gather data on something that looks like it is the same but is not, but can act as if it is; or use percentages of centres differently.

(b) Construct the poster. Try to be convincing in data and arguments. Make it really attractive and appealing.

(c) Did you find that some data cannot be misrepresented easily? Why?

Apply the idea

"Both sides" posters.

Prepare a double poster display for an assertion as below.

Choose an assertion (e.g. Young drivers are more dangerous than old drivers) that is suitable for (b) below.

a) Gather data and prepare two side-by-side posters, one presenting data and drawing inferences in a way that supports the assertion, and the other presenting data and drawing inferences in a way that rejects the assertion (make the posters attractive and appealing on both sides, with graphs and headings and so on).
b) Think of opposing ideas and how you could be able to find opposing data – remember that there may need to be a different way of looking at the data. Think of different ways your data could be biased and your graphs look better in their support of bias.
c) Display your posters side by side.

Extend the idea

"Paradox" poster.

There is a paradox in percent in that two sets of data can show a reduction yet the combined data show an increase. For example, there could be two groups – in the first, most prefer A and in the second most also prefer A; however, when put together, the combined data shows most prefer B.

a) Find this paradox. Describe it.
b) Why is it possible? Represent it on a poster.

Teacher Reflective Notes

This page is provided for you to record any notes with respect to resources you found useful, additional resources, activities and/or models that worked well/not so well.

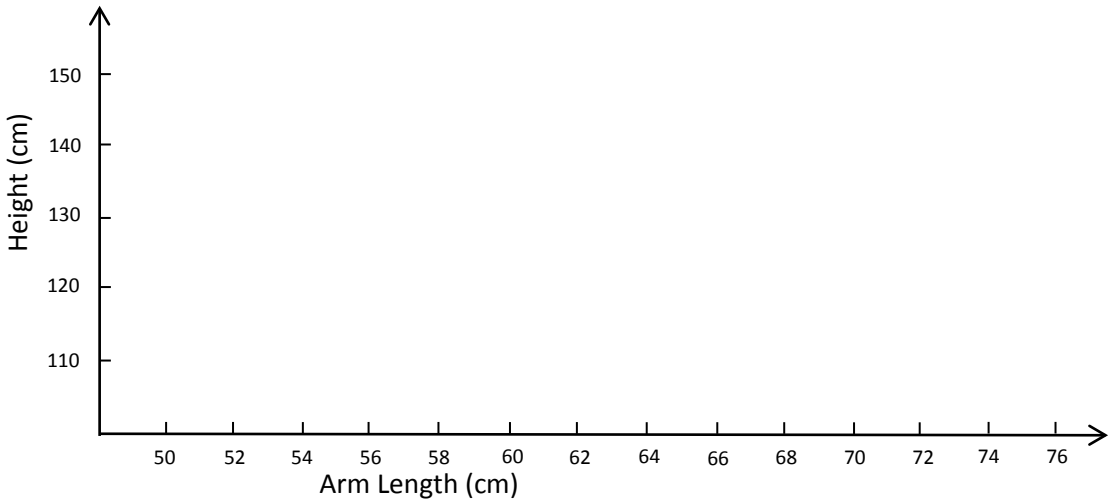
Name:_____ Date:_____

Can you do this? #6

The following table contains students' heights and arm lengths. The students are all in the same age group.

Person	Height	Arm length
John	135 cm	68 cm
Mary	121 cm	55 cm
Frank	142 cm	70 cm
Bruce	149 cm	75 cm
Sue	118 cm	52 cm
Tom	139 cm	70 cm
Ben	144 cm	71 cm

1. (a) Represent the data on a scatter plot.



(d) What does this data say about height and arm length?

_____

(e) What age do you think these students are and why? _____

_____

(f) What could this data suggest about girls in this age group?

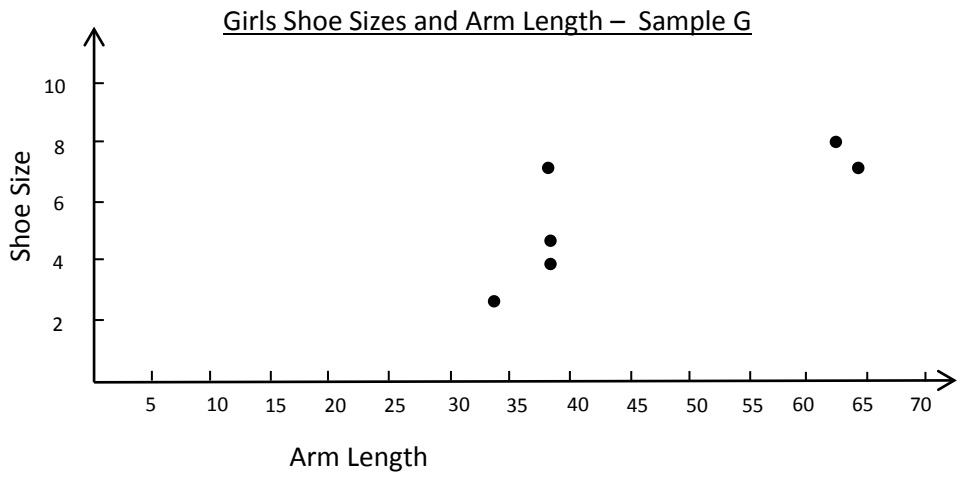
_____

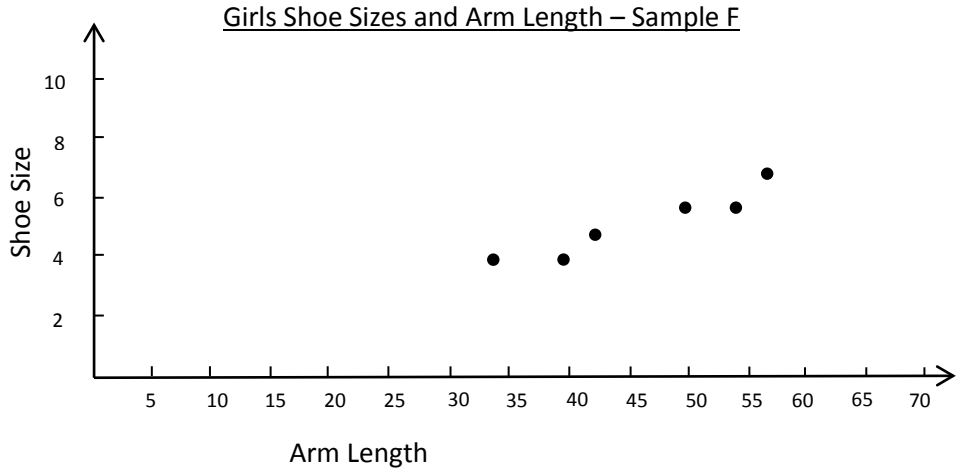
_____

Obj. 10.6.1
a)i. ☐
ii. ☐
iii. ☐
iv. ☐
v. ☐
vi. ☐
vii. ☐

Obj. 10.6.2
b) ☐
c) ☐
d) ☐

Two groups of girls made a scatter plot of their shoe sizes against their arm lengths.





2. Compare the scatter plots from the two groups.

   (a) Do the graphs of shoe size against arm length for each group look similar? _____

   (b) What might this suggest about the relationship between shoe size and arm length for girls? _____

   _____

   (c) Draw a line of best fit on the graph for Sample F.

   (d) Plot the points from Sample G onto the graph for Sample F.

   (e) Does the line of best fit on the graph apply to the larger sample?

   _____

Obj. 10.6.2
a) ☐
b) ☐

Obj. 10.6.1
d)i. ☐
ii. ☐
iii. ☐
iv. ☐
v. ☐
vi. ☐

Obj. 10.6.3
c) ☐
e) ☐

Cycle 6: Correlation

Overview

Big Idea

This cycle explores the use of graphical representations to assist with determining when a relationship exists between two continuous variables. For example, time vs. distance travelled, height vs. arm length/leg length/foot length, height vs. age. Students will create scattergrams or scatter plots of data (students' heights vs. arm length) and will consider the idea of a line of best fit in their analysis of the relationship between variables.

Objectives

By the end of this cycle, students should be able to:

10.6.1 Represent numerical variables on a scatterplot. [10SP251]

10.6.2 Comment on the relationship between two numerical variables. [10SP251]

10.6.3 Identify line of best fit on a scatterplot. [10SP251]

Conceptual Links

This cycle extends from students' previous data gathering and representation activities of categorical and continuous data. Students need to understand and recognise the differences between categorical data, discrete data and continuous data.

In this cycle the visual representation of two continuous values are mapped on a scattergram or scatter plot in order to determine the existence of a relationship between the variables. As such, students need to recognise that each plotted coordinate on the graph represents the relationship between the variables on the axes. If there is a recognisable trend within the data and many values appear to be clustered, a line of best fit may be drawn. In the event that there is no recognisable trend, it is not possible to deduce a relationship between the variables. These ideas will extend into Unit 13: Modelling with linear relationships.

Materials

For Cycle 6 you may need:

- Graph paper
- Marking pens
- Masking tape
- String/wool

Key Language

Direct relationship, indirect relationship, scattergram, scatter plot, line of best fit

Definitions

Direct relationship: a relationship between two variables where an increase or decrease in one variable causes the same increase or decrease in the other variable.

Indirect relationship: a relationship between two variables where an increase in one variable causes a decrease in the other variable or a decrease in one variable causes an increase in the other variable.

Line of best fit: a line on a graph showing the general direction that a group of points seem to be heading. Not all points will be on the line of best fit but most should be quite close.

Scattergram or Scatter plot: a graph of plotted points that show the relationship between two sets of data

Assessment

Anecdotal Evidence

Some possible prompting questions:

- Which variables are on the graph?
- Are they clustered together or very spread apart?
- Is there a thicker part of the cluster that runs in a particular direction?
- Could this be a trend?
- If we drew a line here, could we state that most values are around this line?
- Are both variables going up?
- Is one variable going up while the other variable goes down?
- Is there a relationship that can be inferred?

Portfolio Task

Analysis of correlation between variables has not been included in *Portfolio Task 10: Are we really healthy?* but there is scope within the data set chosen to generate some of these graphs. It may be possible to graph students' sleep against their physical activity, physical activity against their leisure choices or physical activity against dietary choices to see if these variables are related.

RAMR Cycle

Reality

Discuss related data with students. Consider data sets they have explored so far in terms of ideas that may be related. For example, is there a relationship between students' heights and their arm length; students' heights and foot lengths; number of people living at home and the number of pets a household has? Have students hypothesise on whether the data may be related (dependent) or unrelated (independent).

Abstraction

The abstraction sequence for this cycle engages students with exploring possible relationships between variables using previously gathered data of students' heights and arm lengths. A suggested sequence of activities is as follows:

1. *Kinaesthetic activity.* Create a large grid on the floor with tape or use a maths mat. Label the axes with height and arm lengths. Have students locate themselves on the maths mat at the point where these intersect ensuring that they go across the x-axis and then up the y-axis (informally experience coordinates and ways of reading ordered pairs). Ensure that students have a sticky note with their name on it so that if an area of the graph becomes congested, students may be replaced by sticky notes.

   Identify points of congestion on the graph. Has a line developed (may be rough) on the graph? Ask students to suggest whether arm length is getting longer along with height.

2. *Represent with materials.* Have students place their sticky notes down on the large graph and step away. Highlight the general direction of the data with a length of string or wool. Reinforce with students that arm length appears to be increasing as students' height is increasing. Adjust the string or wool until a line of best fit that best represents the trend in the data is reached.

3. *Represent with symbols.* Have students recreate the resulting scatterplot and line graph in their books, using a ruler to draw the line of best fit.

Mathematics

Language/symbols and practice

In the mathematics phase, students should practice graphing continuous variables against other continuous variables and analyse the resulting visual representation for trends in the data or possible relationships that may be described by a line of best fit. Ensure that students also experience situations where there is no relationship between the variables.

Examples of datasets where a relationship exists that students may wish to investigate could include:

- Height against foot length
- Height against shoe size
- Height of a birthday candle against time (mark candle in 5mm segments; record time as the top of the candle reaches each mark)
- Plant growth against time
- Length of shadow against time
- Dimensions of an ice cube against time (or area of the puddle created by a melting ice cube; or area of a puddle against time)

Examples of datasets that may not be related could include:

- Distance from home to school against students' heights
- Number of people at home against students' heights
- Shoe size against head size
- Hand length against head size
- Number of pets against number of people at home
- Distance from home to school against number of pets

Reflection

Check the idea

Engage students with investigating relationships between variables. Some of the suggested investigations for the Mathematics phase are suitable here. Alternatively, students may like to explore the following:

- Index finger length against time taken to tie a shoelace
- Height against length of a standing jump
- Hand length against longest finger length
- Hand length against hand span

Apply the idea

Investigate a student-suggested relationship between variables. Encourage them to consider things they do that are measurable. If no suggestions are forthcoming, consider any investigations that have not yet been explored during this cycle.

Extend the idea

This idea will be extended to linear relationships in Unit 13. For now, ensure that students can generalise the general trend in the direction of the line of best fit where variables are both increasing, both decreasing, or one is increasing while the other decreases.

Teacher Reflective Notes

This page is provided for you to record any notes with respect to resources you found useful, additional resources, activities and/or models that worked well/not so well.

Unit 10 Portfolio Task – Teacher Guide

Are we really healthy?

Content Strand/s: Statistics and Probability

Resources Supplied:

- Task sheet
- Teacher guide

Other Resources Needed:

- Reference: http://www.kidshelp.com.au/teens/get-info/hot-topics/healthy-lifestyles.php

Summary:
Students gather data and make comparisons with ABS data sample to determine the important aspects of a healthy lifestyle.

Variations:

- Students can gather data online, or find their own data sets to investigate.

ACARA Proficiencies Addressed:
Understanding
Fluency
Problem Solving
Reasoning

Content Strands:

Statistics and Probability

10.1.1 Construct dot plots. [7SP170]
10.1.3 Identify mode and median of discrete data sets. [7SP171]
10.1.4 Calculate mean of discrete data sets. [7SP171]
10.1.5 Describe and interpret graphical representations of discrete data sets using statistical measures in the context of the data. [7SP171]
10.1.6 Interpret and compare a range of data displays of discrete data sets.
10.2.1 Construct histograms from continuous data.[7SP170]
10.2.5 Identify mode and median of continuous data sets. [7SP171]
10.2.6 Calculate mean of continuous data sets. [7SP171]
10.2.8 Interpret and compare a range of data displays of continuous data sets. [7SP172]
10.3.1 Construct side-by-side column graphs. [6SP147]
10.4.1 Describe and interpret different sample sets for the same data in context. [5SP120]
10.4.4 Collect data using sampling. [8SP206]
10.4.6 Identify everyday questions and issues involving at least one numerical and at least one categorical variable. [9SP228]

Are we Really Healthy?

Name	
Teacher	
Class	

Your Task:

It is your task to decide if your class and age group are really healthy. To do this you will have to use:

- data
- statistics
- inference

Within Portfolio Task 10, your work has demonstrated the following characteristics:

			A	B	C	D	E
Understanding and Fluency	Mathematical language and symbols	10.1.3 Identify mode and median of discrete data sets. 10.1.4 Calculate mean of discrete data sets. 10.2.5 Identify mode and median of continuous data sets. 10.2.6 Calculate mean of continuous data sets.	Effective and clear use of appropriate mathematical terminology, diagrams, conventions and symbols	Consistent use of appropriate mathematical terminology, diagrams, conventions and symbols	Satisfactory use of appropriate mathematical terminology, diagrams, conventions and symbols	Use of aspects of mathematical terminology, diagrams and symbols	Use of everyday language
Problem Solving and Reasoning	Problem solving approaches	10.4.4 Collect data using sampling. 10.4.6 Identify everyday questions and issues involving at least one numerical and at least one categorical variable.	Systematic application of relevant problem-solving approaches to investigate a range of situations, including some that are complex unfamiliar	Application of relevant problem-solving approaches to investigate complex familiar or simple unfamiliar situations	Application of problem-solving approaches to investigate simple familiar situations	Some selection and application of problem-solving approaches in simple familiar situations.	Partial selection of problem-solving approaches
	Mathematical modelling	10.3.1 Construct side-by-side column graphs. 10.1.1 Construct dot plots. 10.2.1 Construct histograms from continuous data.	Development of mathematical models and representations in a range of situations, including some that are complex unfamiliar	Development of mathematical models and representations in complex familiar or simple unfamiliar situations	Development of mathematical models and representations in simple familiar situations	Statements about simple mathematical models and representations	Isolated statements about given mathematical models and representations
	Reasoning and justification	10.1.5 Describe and interpret discrete data sets using statistics. 10.1.6 Interpret and compare data displays of discrete data sets. 10.2.8 Interpret and compare data displays of continuous data sets. 10.4.1 Describe and interpret different sample sets for the same data in context.	Clear explanation of mathematical thinking and reasoning, including justification of choices made, evaluation of strategies used and conclusions reached	Explanation of mathematical thinking and reasoning, including reasons for choices made, strategies used and conclusions reached	Description of mathematical thinking and reasoning, including discussion of choices made, strategies used and conclusions reached	Statements about choices made, strategies used and conclusions reached	Isolated statements about given strategies or conclusions

Comments:

Do you ever feel like every time you turn on the TV or pick up a magazine, that all that everyone is talking about is "healthy lifestyles"? It seems to be everyone's favourite topic. But what does it really mean to have a healthy lifestyle and is it just about food and exercise?

What is a Healthy Lifestyle?

Having a healthy lifestyle is not just about eating healthy and exercising. There is lots of information out there that talks about the importance of having a healthy lifestyle and this often refers to having a work/life balance. As a teenager, you may have a lot of things that you are trying to balance such as school or study, friendships/relationships, extracurricular activities, sports, part-time job, and responsibilities at home. Sometimes this can start to feel like a huge juggling act and can easily become overwhelming. That's why it's important to find balance in your life which is really what having a healthy lifestyle is all about.

So, what are the important aspects of having a healthy lifestyle and how can you start to develop habits now that will set the foundations for a healthy future?

Sleep

To start with, make sure you get plenty of sleep. Some research suggests that teenagers need more sleep than adults however it's often the case that they get less than they need. With homework and social obligations, before you know it you can be staying up late but still have to get up early to get to classes, work or other activities. To function properly, your body and mind need a minimum of eight hours of sleep each night. Sleep is important as it improves your alertness and concentration and is the only part of the day where your body gets a chance to fully recuperate, recover and heal.

Getting enough sleep will also help increase your energy levels to help you get through the day. Some tips for getting a good night's sleep include:

- have a routine for going to bed, e.g. having a shower, getting into your PJ's, having a warm drink etc.
- close the blinds or curtains before you go to bed and minimize any light sources in the bedroom
- turn off the TV and if you need to have music to help you sleep, make sure you keep the volume down
- turn off your phone. Everyone says it right, and it can be really hard to do, but it does help if you turn your phone off so you don't get woken up by messages or notifications while you're trying to get to sleep.

Diet

Eat a well balanced diet. This sounds much easier than it is. Did you know though that if you don't eat well that you could be put yourself at risk of:

- becoming overweight,
- being more susceptible to illnesses, cardiovascular diseases and acne,
- having erratic moods and experiencing depression and,
- experiencing a lower level of enjoyment of life.

However lots of people are really conscious about what they eat, and try to make healthy choices. Remember, even where it comes to eating habits, it's all about balance. If you think you have too much junk food and want to make some changes to your diet, don't try and eliminate all junk food. If you do, it normally leads to intense junk food cravings which sooner or later, you're likely to give in and binge on. The trick is to enjoy the unhealthy stuff occasionally, and substitute or swap the unhealthy stuff for the healthy stuff regularly! If you're trying to eat healthy, a few simple tips are:

- drink plenty of water
- eat a balance of meat/protein, whole grains, fruit and vegetables daily
- don't skip breakfast. Eat something for breakfast everyday
- avoiding eating just because you're bored
- eat regular meals

Physical Activity

Exercising regularly is part of having a healthy lifestyle. Physical activity helps build a strong body and mind. People, who exercise in conjunction with eating a balanced diet, benefit from a healthy physique, a mind that can think more clearly and are more likely to have less stress. Doing physical activity releases endorphins, which are a chemical your body produces, that gives you a good feeling. Physical activity is also an effective way to manage moods and is a really good lifestyle choice to improve your overall wellbeing.

Sharing the Load

Having friends and family that you can confide in is really important. To be a healthy person, you need to have people around you that you can talk to about what's happening for you, to talk to about your problems rather than dealing with them alone or not dealing with them at all. If you feel stressed or overwhelmed, it helps to speak with someone you trust. Although it might seem obvious to you, those around may not always realise what you are going through so it's sometimes important to let someone know.

Play

Another important part of having a healthy lifestyle is remembering to "play". Just taking time to laugh, have fun and being around people who make you feel good is part of having balance in your life. Take time to learn about what makes you feel good and remember to include some of those things in your day.

It's All About Balance

Getting enough sleep, having good food habits, doing some daily physical activity, having trusted people to confide in and finding time to "play" will mean you are well on your way to having a healthy lifestyle. However, when you're busy trying to fit in everything you have to do in a day, it can be hard to find a healthy balance. It is really easy to slip into the habit of choosing unhealthy snacks and take-away foods or spending your free time watching TV or in front of the computer. In moderation, these can be fine, but, if they are out of balance, they can be dangerous for your health, both now and in the long-term. Habits that you form as a young person will set the foundation for the future. Forming good habits of a healthy lifestyle will be something that you will take with you into your adult years.

At KHL we understand that healthy lifestyle choices are influenced by many things and that information/education is important as well as understanding yourself and the factors that influence your choices. We aim to help our callers to gain awareness about their lifestyle choices, explore the positive and negative impacts of lifestyle choices and if needed, explore healthier options. If you would like to talk to someone about having a healthy life style you can call Kids Helpline on 1800 55 1800, 24 hours a day, 7 days a week or use our email or web counselling services.

References

Teenagers and sleep - Better Health Channel

Published: 4 November 2013

Adapted from Healthy Lifestyles – Kids Help Line Webpage:

http://www.kidshelp.com.au/teens/get-info/hot-topics/healthy-lifestyles.php

1) What are the important aspects of having a healthy lifestyle?

(**Hint**: Look at the headings and the pictures.)

(a) _____ (b) _____

(c) _____ (d) _____

(e) _____

2) Look at the survey questions. Highlight the questions you could ask to work out if your class and age group are really healthy?

(Questionnaire questions from Census at School, Variables List 2014)

1. Are you male or female?
2. When were you born?
6. What year level are you in at school?
7. Are you of Aboriginal or Torres Strait Islander origin?
8. What colour are your eyes?
9. How tall are you without your shoes on?
17. What is your favourite type of take-away food?
18. What did you have for breakfast this morning?
    (a) Did not eat breakfast
    (b) Bread or bread product (e.g. toast, pancakes, dried biscuits)
    (c) Breakfast cereal
    (d) Breakfast bar/snack bar or pack
    (e) Baked beans/spaghetti
    (f) Eggs
    (g) Fruit/fruit juice
    (h) Noodles/rice
    (i) Meat or meat product
    (j) Milk or milk product
    (k) Lollies/potato chips
    (l) Soft drink
    (m) Tea/coffee
    (n) Other foods
19. How many hours of sleep do you usually get on a school night?
20. What is the main method of travel that you usually use to get to school?
21. How long does it usually take you to travel to school? (minutes)
22. How do you access the Internet at your home?
23. In what sport or activity do you most enjoy participating?
24. What is your favourite type of music?

26. Estimate how many hours a week you usually spend doing these activities.
    (a) Hanging out with friends?
    (b) Doing homework?
    (c) Doing things with family?
    (d) Playing sports/outdoor games or activities?
    (e) Playing computer/video games?
    (f) Using the computer/Internet?
    (g) Watching TV?
    (h) Paid work?
    (i) Volunteer community work?
    (j) Housework/jobs at home?

27. How do you usually spend your time on the Internet?
    (a) Social networking sites
    (b) Researching for school work
    (c) Emailing family and friends
    (d) Searching/browsing for information
    (e) Uploading/downloading (music, photos, videos)
    (f) Playing games
    (g) Video sharing sites (YouTube)
    (h) Buying and/or selling things

31. What is your resting pulse rate?

3) Choose the ten most important questions to ask and gather data from your class.

4) Work out the mean, median and mode amounts of sleep in your class. Which of these is the most appropriate measure of centre for how much sleep your class has each school night?

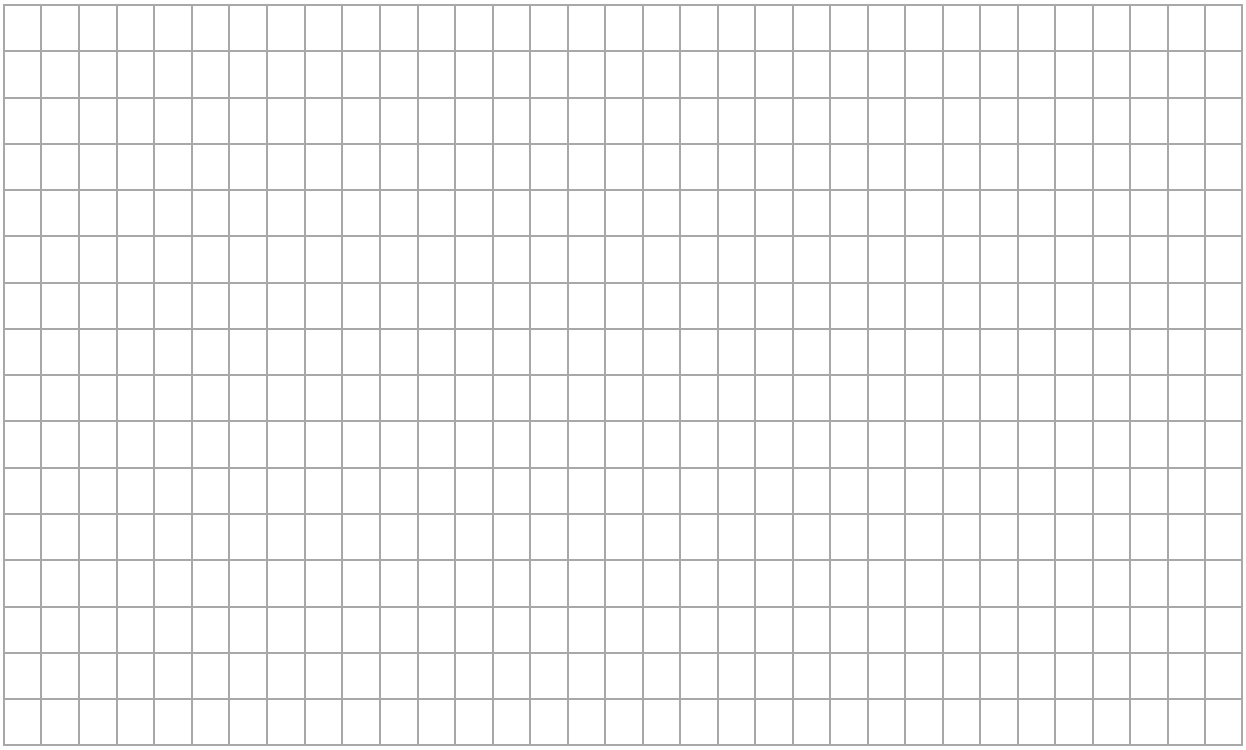
_____

_____

5) How does this match up with the recommended amount of sleep for teenagers?

_____

6) a. Graph the favourite types of takeaway food and breakfast information for your class.



6) b. Identify healthy options and less healthy options. How healthy do you think the food choices for your class are?

_____

_____

_____

On the following pages there are three data displays.

7) Look at the data that shows how much physical activity people do. What is the mean, median and mode amount of physical activity in your class? Which is the most appropriate measure to use in this case?

_____

_____

_____

_____

_____

8) Work out the mean, median and mode amounts of sleep in this sample. How does this compare with your class data? (Is your class similar or different from this sample? Which do you think is healthier?)

_____

_____

9) How does this match up with the recommended amount of sleep for teenagers?

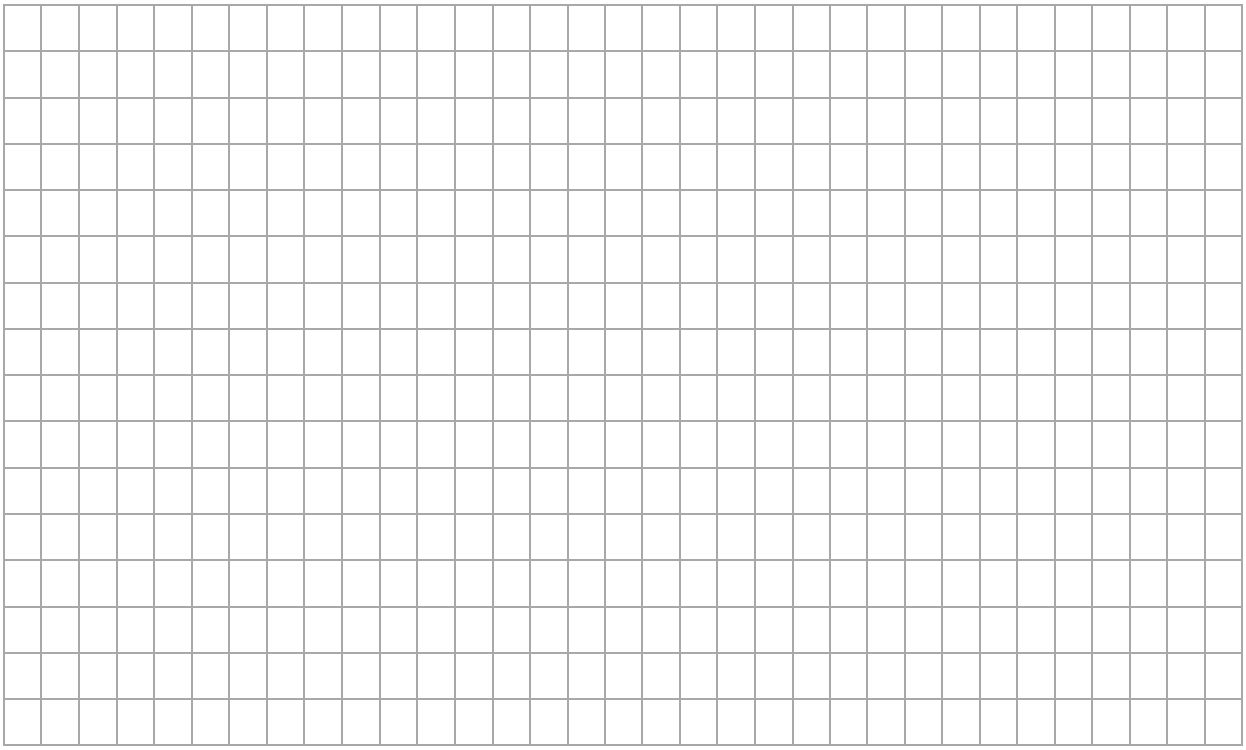
_____

Gender	Activity Estimate (hours per week)									
	Hanging out with friends	Doing homework	Doing things with family	Playing sport / outdoor games	Playing computer/ video games	Using the computer/ internet	Watching TV	Paid work	Volunteer community work	Housework/ jobs at home
Male	30	6	11	17	40	40	19	0	0	6
Female	40	6	5	5	0	25	0	0	0	3
Male	20	4	40	20	10	20	40		0	40
Male	1	0	1	0	40	40	0	0	0	0
Male	40	2	40	20	30	31	10	11	0	7
Female	35	10	11	4	10	20	9			2
Female	10	1	40			10	40	10		20
Male	11	3	3	15	1	11	5	23	3	2
Female	25	5	40	40	1	1	3	15	1	4
Male	4	4	12	8	2	10	5	0		2
Female	8	1	2	3	0	6	3	1		2
Female	10	2	4	7	2	10	3	1	2	1
Female	23	5	3	8		12	5	6		4
Male	40	0	4	4	7	13	4	0	0	2
Male	5	3	2	5	8	8	3	7	3	5
Female	10	3	7	18		40	12		9	2
Male	16	5	10	12	40	40	0	0	0	4
Male	5	3	7	20	1	4	1	0		2
Male	4	10	2	8	12		10			6
Male	10	1	4	8	15	15	6	0		3

Gender	Breakfast choices													
	No Breakfast	Bread	Cereal	Snack bar	Beans/ Spaghetti	Eggs	Fruit/ juice	Noodles/ rice	Meat	Milk	Lollies/ Chips	Soft drink	Tea/ coffee	Other
Male	No	Yes	No	No	No	No	No	No	No	No	No	No	No	No
Female	No	No	Yes	No	No	No	No	No	No	No	No	No	No	No
Male	No	Yes	No	No	No	No	No	No	No	No	No	No	No	No
Male	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No
Male	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No
Female	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No
Female	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No
Male	No	Yes	No	No	No	No	No	No	No	No	No	No	Yes	No
Female	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No
Male	No	Yes	No	No	No	No	No	No	No	No	No	No	No	No
Female	No	Yes	No	No	No	No	No	No	No	No	No	No	No	No
Female	Yes	No	No	No	No	No	Yes	No	No	No	No	No	No	No
Female	No	No	Yes	No	No	No	No	No	No	No	No	No	No	No
Male	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No
Male	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No
Female	No	Yes	No	No	No	No	No	No	No	No	No	No	Yes	No
Male	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Male	No	No	Yes	No	No	No	No	No	No	No	No	No	No	No
Male	No	Yes	Yes	No	No	No	No	No	No	No	No	No	No	No
Male	No	No	Yes	No	No	No	No	No	No	No	No	No	No	No

Gender	Favourite Takeaway	Hours sleep	Travel to school	How long trip to school takes	Resting pulse rate
Male	Chips/Fries	7	Car	20	64
Female	Pizza/Pasta	7	Walk	3	67
Male	Kebabs/Wraps	10	Walk	20	75
Male	Chicken (e.g.BBQ chicken)	10	Car	3	60
Male	Pies/Pasties	8	Walk	1	62
Female	Pizza/Pasta	7	Bus	30	86
Female	Chips/Fries	5	Walk	30	98
Male	Pizza/Pasta	6	Bus	120	76
Female	Chicken (e.g.BBQ chicken)	6	Bus	60	65
Male	Pizza/Pasta	8	Bus	20	44
Female	Chips/Fries	7	Bus	23	80
Female	Chips/Fries	9	Car	4	52
Female	Chicken (e.g.BBQ chicken)	10	Bus	30	99
Male	Kebabs/Wraps	7	Walk	7	71
Male	Chips/Fries	8	Car	15	60
Female	Pizza/Pasta	5	Car	3	64
Male	Chicken (e.g.BBQ chicken)	9	Walk	10	980
Male	Pizza/Pasta	8	Car	12	40
Male	Chicken (e.g.BBQ chicken)	7	Walk	2	76
Male	Chicken (e.g.BBQ chicken)	8	Walk	10	90

10) a. Graph the favourite types of takeaway food and breakfast information for this sample.



10) b. Identify healthy options and less healthy options. How healthy do you think the food choices for this sample are? How does this compare with your class data? (Is your class similar or different from this sample? Which do you think is healthier?)

_____

_____

_____

11) Look at the data that shows how much physical activity people do. What is the mean, median and mode amount of physical activity in this sample? How does this compare with your class data? (Is your class similar or different from this sample? Which do you think is healthier?)

_____

_____

_____

_____

Name:_____ Date:_____

Can you do this now? Unit 10

1. Here is data gathered by twelve girls for shoe size.

Girls	01	02	03	04	05	06	07	08	09	10	11	12
Shoe size	7	3	5	4	8	7	6	4	7	4	6	5

(a) Draw a dot plot of the girls' shoe sizes.

Obj. 10.1.1
1.a)i□
a)ii□
a)iii□
a)iv□
a)v□
a)vi□

(b) Write down the range of shoe sizes in this group. _____ Obj. 10.1.2 b) □

(c) Write down the mode of shoe sizes in this group. _____ Obj. 10.1.3 c) □

(d) Write down the median of shoe sizes in this group. _____ d) □

(e) Calculate the mean of shoe sizes in this group. Obj. 10.1.4 e)i □

_____ e)ii □

(f) If you ran a shoe shop, which size shoes should you stock more of? Obj. 10.1.5 f) □

_____ g) □

(g) Why? _____

Two girls with size 9 shoes join the group.

(h) Will the mode shoe size change? _____ Obj. 10.1.7 h) □

(i) Will the median shoe size change? _____ i) □

(j) If yes, write down the new median shoe size. _____ k) □

(k) Will the mean of the shoe sizes change? _____ Obj. 10.1.3 j) □

(l) If yes, calculate the new mean of the shoe sizes. Obj. 10.1.4 l)i □

_____ l)ii □

2. Here is data gathered by twelve girls for arm length.

Girls	01	02	03	04	05	06	07	08	09	10	11	12
Arm length cm	65	34	38	38	63	38	54	34	57	40	50	42

(a) Write down the range of the girls' arm lengths. _____

(b) Use the table to work out the intervals and to tally the frequency of arm lengths in each interval.

Tally									
Interval									
Centre									

(c) Draw a histogram of the girls' arm lengths.

_____

_____

(d) Write down the mode of the girls' arm lengths. _____

(e) Calculate the mean of the girls' arm lengths.

_____

_____

Two girls with arm lengths of 60cm join the group.

(f) Will the mode of the girls' arm lengths change? _____

(g) If yes, write down the new mode/s for the girls' arm lengths.

_____

(h) Will the mean of the girls' arm lengths change? _____

(i) Calculate the changed mean for the girls' arm lengths.

_____

Obj. 10.2.4
a)i ☐
a)ii ☐

Obj. 10.2.1
b)i☐
b)ii☐
b)iii☐
c)i☐
c)ii☐
c)iii☐
c)iv☐
c)v☐
c)vi☐

Obj. 10.2.5
d)i ☐
d)ii ☐

Obj. 10.2.6
e)i ☐
e)ii ☐
e)iii ☐

Obj. 10.2.9
f) ☐
g)i ☐
g)ii ☐
g)iii ☐
g)iv ☐
h) ☐
i)i ☐
i)ii ☐
i)iii ☐

3. The following data demonstrates students' heights. Students are all in the same age group.

Person	John	Mary	Frank	Bruce	Sue	Tom	Ben	Amy
Height (cm)	135	121	142	149	118	139	144	135

(a) Construct a stem and leaf graph of the students' heights.

Obj. 10.2.2 a)i□ a)ii□

(b) Write down the median of the students' heights. _____

Obj. 10.2.5 b)i □ b)ii □

(c) Write down the quartiles of the students' heights.

First quartile:_____Third quartile: _____

Obj. 10.2.10 c)i□ii□ iii□iv□

(d) Construct a box plot and whiskers graph of the students' heights.

Obj. 10.2.3 d)i □ d)ii □ d)iii □ d)iv □ d)v □ d)vi □ d)vii □

(e) Calculate the mean of the students' heights.

Obj. 10.2.6 e)i □ e)ii □ e)iii □

(f) Comparing the mean, median and the mode of this data, what conclusion could you draw about this group of students?

Obj. 10.2.7 f) □

(g) You want to represent students' heights on a graph so that the audience can identify individual data values. Circle the graph type you would use.

Obj. 10.2.8 g)i□ g)ii□ g)iii□

(i) histogram (ii) stem and leaf (iii) box plot and whiskers

4. Here is data gathered by six girls about their favourite colour.
(R-Red, B-Blue, Y-Yellow, G-Green)

Girls	G1	G2	G3	G4	G5	G6
Favourite colour	R	Y	Y	Y	B	R

Here is data gathered by six boys about their favourite colour.
(R-Red, B-Blue, Y-Yellow, G-Green)

Boys	B1	B2	B3	B4	B5	B6
Favourite colour	G	G	B	R	R	G

(a) Construct a stacked bar graph to show the girls and boys favourite colours.

Obj. 10.3.2
4.a)i. ☐
a)ii. ☐
a)iii. ☐
a)iv. ☐
a)v. ☐
a)vi. ☐

Obj. 10.3.6
b) ☐
c) ☐
d) ☐
e) ☐

(b) Write down the mode of the whole group's favourite colour.

(c) Write down the mode of the girls' favourite colours.

(d) Write down the mode of the boys' favourite colours.

(e) What does this tell you about this group's favourite colours?

5. Here is data gathered by six girls about their arm lengths in centimetres.

Girls	G1	G2	G3	G4	G5	G6
Arm length cm	65	34	38	38	63	38

Here is data gathered by six boys about their arm lengths in centimetres.

Boys	B1	B2	B3	B4	B5	B6
Arm length cm	64	54	77	60	70	62

(a) Construct a back-to-back stem and leaf graph to show boys' and girls' arm lengths.

Obj. 10.3.4
5.a)i. □
a)ii. □
a)iii. □

(b) Calculate the mean arm length of the whole group.

Obj. 10.3.7
b) □
c) □
d) □
e) □
f) □

_____

(c) Calculate the mean arm length of the boys.

_____

(d) Calculate the mean arm length of the girls.

Obj. 10.3.6
g)i □
g)ii □
g)iii □
h) □

_____

(e) Are the means the same or different? _____

_____

(f) What does this tell you about the arm lengths of these students?

_____

_____

(g) Looking at the shape of the graph for each group of students, is the data skewed or symmetric? Boys _____

Girls _____ Whole group _____

(h) How do the girls' arm lengths affect the mean of the whole group of students? _____

_____

Here is data gathered by six girls about their favourite colour.
(R-Red, B-Blue, Y-Yellow, G-Green)

Girls (Sample G)	G1	G2	G3	G4	G5	G6
Favourite colour	R	Y	Y	Y	B	R

A different group of six girls gathered data about their favourite colour.
(R-Red, B-Blue, Y-Yellow, G-Green)

Girls (Sample F)	F1	F2	F4	F4	F5	F6
Favourite colour	G	G	B	Y	R	G

6. Draw a side-by-side bar graph of Sample F and Sample G favourite colours.

Obj. 10.3.1
6. i. ☐
ii. ☐
iii. ☐
iv. ☐
v. ☐
vi. ☐

7. Compare the two samples.

(a) What is the mode of Sample G girls' favourite colours? _____

(b) What is the mode of Sample F girls' favourite colours? _____

(c) How are the two samples the same or different? _____

_____

(d) Is it possible to predict colour preference of girls from these two samples? _____ Why/why not? _____

(e) What would make the sample more effective? _____

_____

Obj. 10.4.1
7. a) ☐
b) ☐
c) ☐

Obj. 10.4.3
d)i. ☐
d)ii. ☐
e) ☐

Here is data gathered by six girls about their shoe sizes.

Girls (Sample G)	G1	G2	G3	G4	G5	G6
Shoe size	7	3	5	4	8	7

A different group of six girls gathered data about their shoe sizes.

Girls (Sample F)	F1	F2	F3	F4	F5	F6
Shoe size	6	4	7	4	6	5

The girls are all the same age.

8. (a) Calculate the mean shoe size for Sample G girls.

_____

(b) Calculate the mean shoe size for Sample F girls.

_____

(c) What is the variation between these two means? _____

(d) What conclusion can you draw from this comparison?

_____

(e) Are these effective samples for predicting the shoe sizes of girls in this age group? _____

(f) Calculate the mean shoe size for the two samples together as one group. _____

_____

(g) Select a sample of six shoe sizes from the whole group that have the same mean as the whole group. _____

_____

(h) What information other than age and shoe size would you need to more effectively predict shoe sizes for girls? _____

_____

_____

(i) How would you gather this information? _____

(j) What questions could you ask to find out all the information you need? _____

_____

_____

Obj. 10.4.1
8. a)i ☐
a)ii ☐
b)i ☐
b)ii ☐
f)i ☐
f)ii ☐

Obj. 10.4.7
c) ☐
d) ☐
g)i ☐
g)ii ☐
g)iii ☐

Obj. 10.4.3
e) ☐

Obj. 10.4.6
h) ☐
j)i ☐
j)ii ☐
j)iii ☐

Obj. 10.4.2
i) ☐

9. Describe the following ways of misrepresenting statistics.

(a) Sample bias: _____

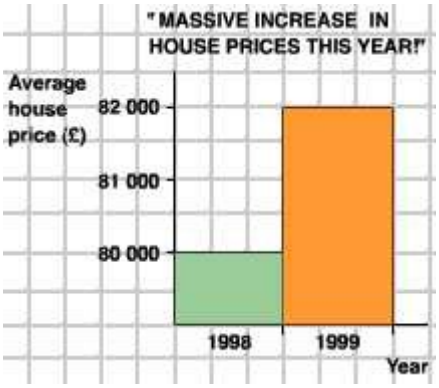
_____

(b) Missing information: _____

_____

(c) Irrelevant statistics: _____

_____



10. Here is a graph showing average house price for 1998 and 1999.

(a) Is this an appropriate representation of this data? _____

(b) Why/why not? _____

_____

(c) Draw a graph that would fit the title:

"House prices remain stable"

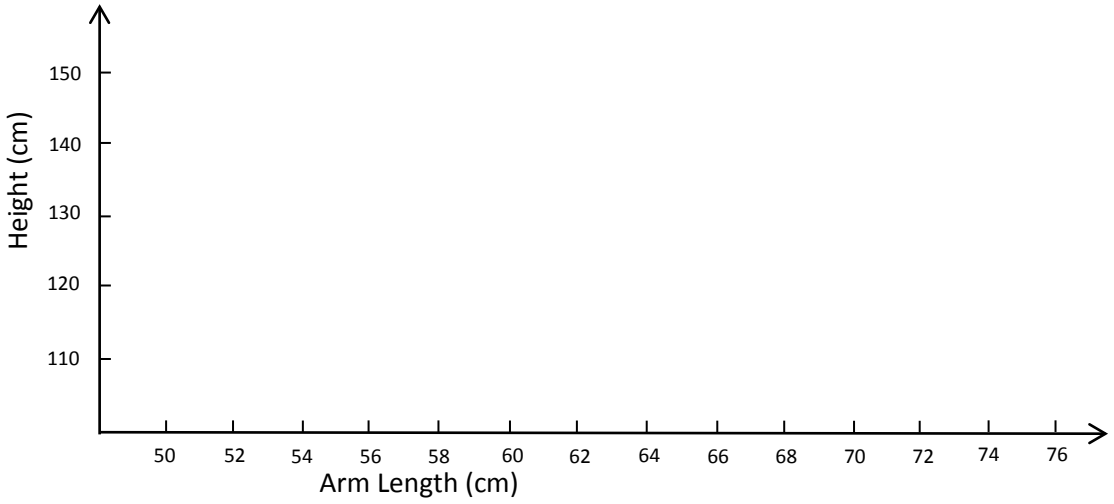
Obj. 10.5.1
9.a) ☐
b) ☐
c) ☐

Obj. 10.5.2
10.a) ☐
b)i ☐
b)ii ☐
c)i ☐
c)ii ☐
c)iii ☐
c)iv ☐

11. The following table contains students' heights and arm lengths. The students are all in the same age group.

Person	Height	Arm length
John	135 cm	68 cm
Mary	121 cm	55 cm
Frank	142 cm	70 cm
Bruce	149 cm	75 cm
Sue	118 cm	52 cm
Tom	139 cm	70 cm
Ben	144 cm	71 cm

(a) Represent the data on a scatter plot.



(b) What does this data say about height and arm length?

_____

(c) What age do you think these students are and why? _____

_____

(d) What could this data suggest about girls in this age group?

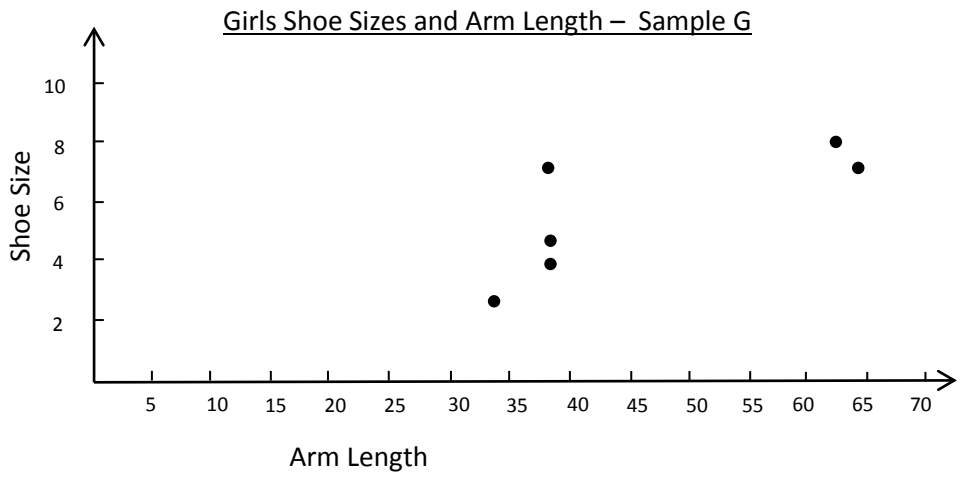
_____

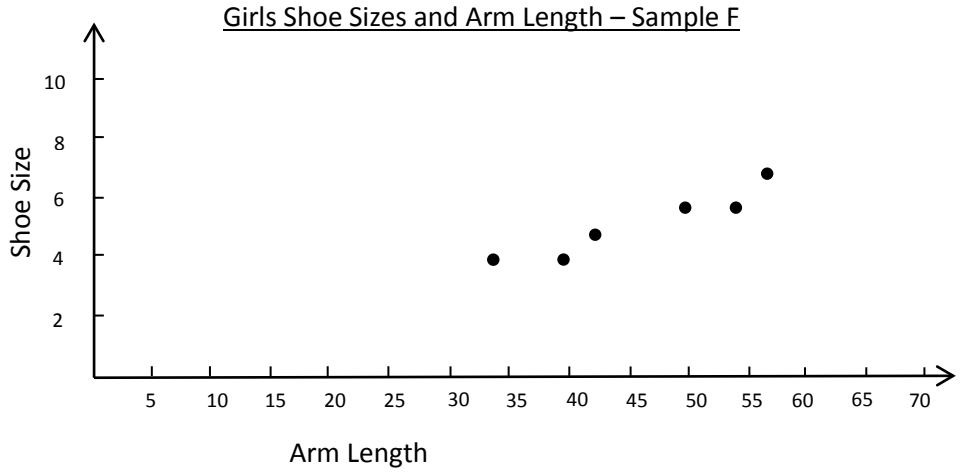
_____

Obj. 10.6.1
11a)i ☐
ii ☐
iii ☐
iv ☐
v ☐
vi ☐
vii ☐

Obj. 10.6.2
b) ☐
c) ☐
d) ☐

12. Two groups of girls made a scatter plot of their shoe sizes against their arm lengths.





Compare the scatter plots from the two groups.

(a) Do the graphs of shoe size against arm length for each group look similar? _____

(b) What might this suggest about the relationship between shoe size and arm length for girls? _____

_____

(c) Draw a line of best fit on the graph for Sample F.

(d) Plot the points from Sample G onto the graph for Sample F.

(e) Does the line of best fit on the graph apply to the larger sample?

_____

Obj. 10.6.2
12.a) ☐
b) ☐

Obj. 10.6.1
d)i. ☐
ii. ☐
iii. ☐
iv. ☐
v. ☐
vi. ☐

Obj. 10.6.3
c) ☐
e) ☐

XLR8 – Accelerating Mathematics Learning

YuMiDeadly
Growing community through education

Faculty of Education
School of Curriculum
S Block, Room 404
Victoria Park Road
KELVIN GROVE QLD 4059
CRICOS No. 00213J

Phone: +61 7 3138 0035
Fax: +61 7 3138 3985
E-mail: ydc@qut.edu.au
Website: ydc.qut.edu.au