

Data Interpretation and Evaluation: 'How Many Songs Do You Have on Your Portable Music Player?'	Stage 4
--	----------------

Scope and sequence summary	Duration: 3 weeks
Substrands: S4 Data Collection and Representation (partial review only) , S4 Single Variable Data Analysis (part)	

Outcomes	Key considerations	Overview
<ul style="list-style-type: none"> > communicates and connects mathematical ideas using appropriate terminology, diagrams and symbols MA4-1WM > applies appropriate mathematical techniques to solve problems MA4-2WM > recognises and explains mathematical relationships using reasoning MA4-3WM > collects, represents and interprets single sets of data, using appropriate statistical displays MA4-19SP > analyses single sets of data using measures of location, and range MA4-20SP 	<p><i>Key inquiry question</i></p> <p>'Is the number of songs that students from a particular Year cohort have on their portable music player representative of other Year cohorts at our school and of the population of the whole school?'</p> <p><i>Key ideas</i></p> <ul style="list-style-type: none"> • Calculate mean, median, mode and range for sets of data • Investigate the effect of outliers on the mean and median • Describe and interpret a variety of data displays using median, mean and range • Calculate and compare summary statistics of different samples drawn from the same population <p><i>Language</i></p> <p>The term 'average', when used in everyday language, generally refers to the mean and describes a 'typical value' within a set of data.</p> <p>Students need to be provided with opportunities to discuss what information can be drawn from the data presented. They need to think about the meaning of the information and to put it into their own words.</p> <p>Language to be developed includes superlatives, comparatives, and expressions such as 'prefer ... over', etc.</p>	<p><i>Students require relevant knowledge, skills and understanding from:</i></p> <ul style="list-style-type: none"> • S4 Data Collection and Representation • S4 Single Variable Data Analysis (part) <p><i>Stimulating interest and discussion</i></p> <p>A media article on music downloads could be presented in the first lesson to stimulate interest and discussion, eg 'Music consumption statistics', 'Listen in gently down the stream', 'Digital music clicking as buyers ditch illegal downloads'.</p> <p><i>Class organisation and preparation</i></p> <p>Divide the class into teams of 3–5 students. Assign one Year cohort of the school to each team, and one subgroup of each Year cohort to each student in the team (eg assign one class to each member of the team).</p> <p>Each student asks every person in their assigned subgroup 'How many songs do you have on your portable music player?' and records the results on a data collection form. Teachers may need to coordinate the data collection process with other teachers and/or inform students in all Year cohorts ahead of time about the information required. An electronic survey application, eg Survey Monkey, could be used as an alternative method for collecting the data.</p> <p>Zero should be recorded for respondents who do not own a portable music device, or who own such a device but do not have any songs. Students should be reminded that a response of zero is a valid data value and needs to be recorded.</p> <p>On most portable music players, a count of the total number of songs can be found at the bottom of the list of all songs on the device.</p> <p><i>Adjustments</i></p> <ul style="list-style-type: none"> • Assign groups carefully to accommodate students' capabilities and social interaction skills. • Some students may require a teacher's assistant or buddy. • Some students may require a script to assist them in collecting data from a class.


Content	Teaching, learning and assessment	Resources
<p>Construct and compare a range of data displays, including stem-and-leaf plots and dot plots (ACMSP170)</p> <ul style="list-style-type: none"> • use a tally to organise data into a frequency distribution table • construct and interpret frequency histograms and polygons <ul style="list-style-type: none"> ▶ select and use appropriate scales and labels on horizontal and vertical axes (Communicating, Problem Solving, Reasoning) ✨ • construct dot plots <ul style="list-style-type: none"> ▶ explain the importance of aligning data points when constructing dot plots (Communicating, Reasoning) ✨ • construct ordered stem-and-leaf plots, including stem-and-leaf plots with two-digit stems <ul style="list-style-type: none"> ▶ explain the importance of ordering and aligning data values when constructing stem-and-leaf plots (Communicating, Reasoning) ✨ 	<p>The content for this part of the unit was previously studied in the Stage 4 unit Data Collection, Representation and Simple Analysis and should not need to be re-taught explicitly. Students are expected to apply the knowledge, skills and understanding they have gained through the study of this content to represent the data collected for the songs scenario.</p> <p><i>Adjustment</i></p> <ul style="list-style-type: none"> • Pre-assess students' recall of the relevant content from the Stage 4 unit Data Collection, Representation and Simple Analysis, and before proceeding review concepts that are not recalled or understood. <p><i>For the data of their assigned subgroup, and using the same scale or groupings as other students to facilitate comparison, each student constructs each of the following:</i></p> <ul style="list-style-type: none"> • frequency distribution table • frequency histogram • stem-and-leaf plot • dot plot. <p>It may be necessary to use groups or classes to create manageable data displays. Teachers should guide students to determine suitable groupings or class intervals, such as 0–49, 50–99, 100–149, ..., or 0–99, 100–199, 200–299, ... These data displays could be placed on the walls of the classroom for reference by the whole class in relation to subsequent activities. Note: Although the displays may use groups or classes, the raw data is to be used for the analysis using measures of location and spread.</p> <p><i>Adjustment</i></p> <ul style="list-style-type: none"> • Some students may require a scaffold with predetermined groups or classes to assist them in constructing each of the data displays. <p><i>For the data of their assigned Year cohort, and using the same groupings as used for their subgroups, each team enters its raw data into a spreadsheet and uses the spreadsheet to construct each of the following:</i></p> <ul style="list-style-type: none"> • frequency distribution table • frequency histogram • dot plot. <p>Teachers can supply students with a spreadsheet in which the frequency table and dot</p>	<p><i>Creating data displays with ICT</i></p> <ul style="list-style-type: none"> • Applications, such as Excel, Geogebra (freeware), etc • Instructions on how to create a dot plot in MS Excel can be found at www.vertex42.com/ExcelArticles/dot-plot.html

Content	Teaching, learning and assessment	Resources
	<p>plot for each Year cohort is automatically generated. Alternatively, students can investigate the use of the array formula for 'frequency' and the formula for 'rept' (ie repeat) to create frequency tables and dot plots respectively. Once these displays have been completed, the raw data for each Year cohort should be shared with other teams so that each team has the raw data for the whole school in its spreadsheet.</p> <p><i>Adjustment</i></p> <ul style="list-style-type: none"> Some students may enter the raw data they have collected into the spreadsheet that automatically generates the frequency distribution tables and dot plots, but omit using a spreadsheet to construct the frequency histogram. This could be completed by the teacher or by another student. 	
<p>Describe and interpret data displays using mean, median and range (ACMSP172)</p> <ul style="list-style-type: none"> calculate measures of location (mean, median and mode) and the range for data represented in a variety of statistical displays, including frequency distribution tables, frequency histograms, stem-and-leaf plots and dot plots ** 	<p><i>Students analyse:</i></p> <ul style="list-style-type: none"> various statistical displays that include integer, fractional or decimal data values data collected from websites, eg ABS, CensusAtSchool, data.gov.au, sports websites various displays produced by students for each songs subgroup data contained in the songs spreadsheet for each subgroup, Year cohort and the whole school. <p><i>Adjustment</i></p> <ul style="list-style-type: none"> Some students may only analyse sets of data: <ul style="list-style-type: none"> that include integer data values and/or a small number of data values from a selection of the websites based on readability and ease of accessibility using web data that is adapted into simple tables by the teacher. <p><i>Students use calculators to:</i></p> <ul style="list-style-type: none"> calculate the mean without the use of the statistical functions of the calculator calculate the mean using the statistical function of the calculator for mean (\bar{x}) check that the number of data values (n) entered into the calculator matches the number of data values in the set of data. <p><i>Students use a spreadsheet program to:</i></p> <ul style="list-style-type: none"> explore the operation of spreadsheet statistical functions, eg: <ul style="list-style-type: none"> the difference between the functions 'mode', 'mode.sngl' and 'mode.mult' which value(s) is returned for the various 'mode' functions when the data is bi-modal 	<p><i>Scottle resources</i></p> <ul style="list-style-type: none"> M008600 Dataset: Solar Hot Water M007943 Dataset: First Fleet M007103 Dataset: Crime Victims M009079 Information on spreadsheet functions for mean, median and mode

Content	Teaching, learning and assessment	Resources
	<ul style="list-style-type: none"> – which values or cells are counted when using the ‘count’ function • calculate the following measures for each subgroup, each Year cohort and the whole school in their songs spreadsheet: <ul style="list-style-type: none"> – count – minimum value – maximum value – mean – median – mode – range • explore methods of repeating formulas for different ranges of cells, eg using the ‘fill’ feature (or equivalent) • write simple formulas that result in the same measure as the equivalent spreadsheet statistical function, eg =sum(C4:C33)/count(C4:C33) returns the same result as =mean(C4:C33). <p><i>Adjustment</i></p> <ul style="list-style-type: none"> • Some students may require assistance to enter formulas into the spreadsheet. 	
<ul style="list-style-type: none"> • draw conclusions based on the analysis of data displays using the mean, median and/or mode, and range 🛠️🎓 	<p><i>Students use measures of location (centre) and spread to draw conclusions based on summary statistics calculated for various data displays, including those:</i></p> <ul style="list-style-type: none"> • that include integer, fractional or decimal values • sourced from the media and/or websites • produced by students for each songs subgroup and Year cohort. <p><i>Adjustments</i></p> <ul style="list-style-type: none"> • Some students may only draw conclusions: <ul style="list-style-type: none"> – from sets of data that include integer data values and/or a small number of data values – from a selection of the websites based on readability and ease of accessibility – from web data that is adapted into simple tables by the teacher – using cloze passages with word banks – using a list of prepared ‘if ... then ...’ features of sets of data to assist them, eg ‘If the range is a large number, <i>then</i> the scores are spread out’, ‘If the range is a small number, <i>then</i> the scores are close together’, etc. 	

Content	Teaching, learning and assessment	Resources
<p>Investigate the effect of individual data values, including outliers, on the mean and median (ACMSP207)</p> <ul style="list-style-type: none"> identify any clusters, gaps and outliers in sets of data 📊 	<p><i>Teachers explain, informally, with appropriate visual aids:</i></p> <ul style="list-style-type: none"> the meaning of the terms 'cluster', 'gap' and 'outlier' (statistical definitions for clusters, gaps and outliers are not required) how to identify clusters, gaps and outliers in various data displays. <p><i>Students identify clusters, gaps and outliers in:</i></p> <ul style="list-style-type: none"> various frequency histograms, stem-and-leaf plots and dot plots the subgroups, the Year cohorts and the whole school in the songs scenario (there should be an outlier in at least one subgroup). 	
<ul style="list-style-type: none"> investigate the effect of outliers on the mean, median, mode and range by considering a small set of data and calculating each measure, with and without the inclusion of an outlier ⚙️ <ul style="list-style-type: none"> explain why it is more appropriate to use the median than the mean when the data contains one or more outliers (Communicating, Reasoning) ⚙️⚙️ determine situations when it is more appropriate to use the median or mode, rather than the mean, when analysing data, eg median for property prices, mode for shoe sizes (Reasoning) ⚙️ 	<p><i>Students predict and investigate the effect on summary statistics of adding one or more particular data values to a particular set of data:</i></p> <ul style="list-style-type: none"> by adding a particular data value to a small set of data and recalculating the summary statistics by using a visual representation of the data values, eg using an online application that shows scores on a number line. <p><i>Students investigate the effect of outliers on summary statistics by calculating them with and without the inclusion of outliers for:</i></p> <ul style="list-style-type: none"> a variety of small sets of data data collected from websites, eg house prices in a particular locality that contains predominantly apartments and a small number of waterfront mansions the subgroups, the Year cohorts and the whole school in the songs scenario. <p><i>Adjustments</i></p> <ul style="list-style-type: none"> Some students may only: <ul style="list-style-type: none"> work with small data sets and omit data from websites focus on identifying outliers and their effects and omit investigation and/or explanation. <p><i>Students explain:</i></p> <ul style="list-style-type: none"> why outlier values may occur in certain situations, eg the locality may contain predominantly apartments and a small number of waterfront mansions which measure of location is the most appropriate for a variety of sets of data, considering the presence/absence of outliers, the context of the data, and the way 	<p><i>Scottle resources</i></p> <ul style="list-style-type: none"> L6564 measures of central tendency M012006 Ploplt applet gives visual demonstration of the effect of outliers on the mean, median and mode M009074 Effect of outliers

Content	Teaching, learning and assessment	Resources
	<p>in which the measure will be used, eg when determining what stock to buy for a shoe shop, the modal shoe size is likely to be more useful</p> <ul style="list-style-type: none"> situations where the median or mode is a more useful measure of location than the mean, eg median for house prices in the newspaper. <p><i>Students generalise their observations about the effect of outliers on summary statistics to determine and explain:</i></p> <ul style="list-style-type: none"> the most appropriate measure of location (centre) when the data contains an outlier whether the range is an appropriate representation of the spread of the data when the data contains an outlier. 	
<ul style="list-style-type: none"> analyse collected data to identify any obvious errors and justify the inclusion of any individual data values that differ markedly from the rest of the data collected 🌀 ⚖️ 	<p><i>Teachers explain, with appropriate examples, possible sources of error in data, including:</i></p> <ul style="list-style-type: none"> the process used to collect the data, eg a misinterpretation of the question by the respondent, a transcription error in a survey response the process used to record the data, such as a typographical error in the data-entry process, eg 5 becoming 55 the practicalities of the issue being investigated, eg the impact of the storage capacity of the portable music player on the data. <p><i>Students examine data for errors and justify the inclusion of data values suspected of being errors in:</i></p> <ul style="list-style-type: none"> a variety of sets of data, eg CensusAtSchool data the subgroups, the Year cohorts and the whole school in the songs scenario. <p><i>Adjustment</i></p> <ul style="list-style-type: none"> Some students may focus only on identifying sources of error and omit explanations. 	<p><i>Scottle resources</i></p> <ul style="list-style-type: none"> M008600 Dataset: Solar Hot Water M007943 Dataset: First Fleet M007103 Dataset: Crime Victims
<p>Explore the variation of means and proportions of random samples drawn from the same population (ACMSP293)</p> <ul style="list-style-type: none"> investigate ways in which different random samples may be drawn from the same population, eg random samples from a census may be chosen by gender, postcode, state, etc 	<p>The concepts of ‘census’, ‘sample’ and ‘random sample’ were introduced in the Stage 4 unit Data Collection, Representation and Simple Analysis and should not need to be taught explicitly.</p> <p><i>Adjustment</i></p> <ul style="list-style-type: none"> Pre-assess students’ recall of the relevant content from the Stage 4 unit Data Collection, Representation and Simple Analysis, and review concepts that are not recalled or understood before proceeding. 	<p>Random samples of data for a variety of issues can be obtained from CensusAtSchool using its ‘random sampler’ function.</p>

Content	Teaching, learning and assessment	Resources
	<p><i>Using existing knowledge, skills and understanding, students explain:</i></p> <ul style="list-style-type: none"> • the difference between a census and a sample of the population • issues that may result in bias when selecting a sample • processes by which random samples may be selected to obtain representative data in a range of situations, eg <ul style="list-style-type: none"> – to design a new uniform for a particular school – to determine what facilities need to be built in the local council area – to rate a particular movie – to determine the mean height of 13-year-olds in Australia – to determine typical household water usage in Australia • the randomness or bias of particular samples, given the method of selection and the issue to be investigated • the subgroups, the Year cohorts and whole school in the songs scenario using the terms 'census' and 'sample' • whether or not the subgroups and Year cohorts in the songs scenario are 'random samples' that will provide data representative of the whole school or data that is biased in some way. <p><i>Adjustment</i></p> <ul style="list-style-type: none"> • Some students may have difficulty explaining in their own words why a sample is biased or random. These students could select an appropriate reason from a set of reasons provided by the teacher. 	
<ul style="list-style-type: none"> • calculate and compare summary statistics (mean, median, mode and range) of at least three different random samples drawn from the same population <ul style="list-style-type: none"> ▶ use a spreadsheet to calculate and compare summary statistics of different random samples drawn from the same population (Communicating, Problem Solving)  ▶ recognise that summary statistics may vary from sample to sample (Communicating) 	<p><i>Students use the summary statistics of the subgroups, the Year cohorts and the whole school in the songs scenario to compare and contrast:</i></p> <ul style="list-style-type: none"> • each subgroup with its Year cohort • each subgroup with another subgroup from the same Year cohort • each subgroup with a subgroup from another Year cohort • each Year cohort with another Year cohort • each Year cohort with the whole school. <p><i>Adjustment</i></p> <ul style="list-style-type: none"> • Some students may complete a guided response sheet and/or a comparison table. 	<p><i>Scottle resources</i></p> <ul style="list-style-type: none"> • M009018 What is 'typical'?

Content	Teaching, learning and assessment	Resources
<ul style="list-style-type: none"> ▶ recognise that the way in which random samples are chosen may result in significant differences between their respective summary statistics, eg a random sample of girls compared to a random sample of boys from the population, random samples chosen by postcode (Communicating, Reasoning) ** ▶ suggest reasons why different random samples drawn from the same population may have different summary statistics (Communicating, Reasoning) ** 	<p><i>Students suggest reasons why:</i></p> <ul style="list-style-type: none"> • the summary statistics in the songs scenario vary from subgroup to subgroup • the summary statistics for some subgroups vary more than others from the summary statistics for the whole school. <p><i>With reference to statistics, students justify:</i></p> <ul style="list-style-type: none"> • which subgroup of their assigned Year cohort is most representative of the whole school • which Year cohort is the most representative of the whole school population. <p><i>Adjustment</i></p> <ul style="list-style-type: none"> • Some students could be informed which subgroup/Year cohort is the most representative of the school and select appropriate reasons from a set of reasons provided by the teacher. 	

Assessment overview

Each team submits a report or presents an oral presentation titled 'Analysis of how many songs the students in our school have on their portable music player'. Components are clearly attributed to either an individual or the team to allow assessment individually (I) and as a team (T) for each of the following:

- data displays for each subgroup, including a frequency distribution table, frequency histogram, stem-and-leaf plot, dot plot (I: each student creates displays for his or her assigned subgroup)
- a spreadsheet of raw data and summary statistics for each subgroup, the Year cohort and the whole school calculated using the statistical functions of the spreadsheet application (T)
- data displays constructed for their assigned Year cohort, including a frequency distribution table, frequency histogram and dot plot (T)
- identification and comparison of similarities and differences in the summary statistics, clusters, gaps and outliers of each assigned subgroup with its Year cohort, including consideration and discussion of the effect of outliers (where they exist) on the summary statistics and determination of the most appropriate measure of location for the context (I: each student compares his or her assigned subgroup with the Year cohort)
- discussion of why or why not each subgroup represents a random sample of its Year cohort with appropriate references to statistical measures (I: each student provides an explanation for his or her assigned subgroup)
- identification and comparison of similarities and differences in the summary statistics, clusters, gaps and outliers of the assigned Year cohort with the whole school population, including consideration and discussion of the effect of outliers (where they exist) on the summary statistics and determination of the most appropriate measure of location for the context (T)
- a discussion of why or why not the assigned Year cohort represents a random sample of the whole school population with appropriate references to statistical measures (T)
- identification and justification of which Year cohort in the school is believed to be the most representative of the whole school population (T).

Digital technologies appropriate for use in the preparation and delivery of the report/presentation

- An online file-sharing facility (such as the school portal, Google Docs, SkyDrive, etc) to assist students in working collaboratively in one file
- A spreadsheet application (such as Excel, Numbers, etc)
- Applications that assist students in creating data displays (such as Excel, Calc, Math, Geogebra, Microsoft Mathematics)
- A presentation application (such as PowerPoint, Impress, Notebook, Keynote, Prezi, etc) for an oral presentation
- A word processing/publishing application (such as Word, Writer, Publisher, Pages, etc) for a written report