 Year 12 Mathematics Standard 1

| MS-S3 Further Statistical Analysis | Unit duration |
| --- | --- |
| An understanding of the statistical process will enable students to be global citizens. Students who understand statistics will be able to use statistics and mathematics at home, in the community and in the workplace. | 4 weeks |

| Subtopic focus | Outcomes |
| --- | --- |
| The principal focus of this subtopic is the development of students’ understanding of the purpose and process of statistical investigation, taking into account appropriate basic design principles.Students develop understanding of the complex nature of questionnaire design and potential misconceptions in statistical representations and reasoning.Within this subtopic, schools have the opportunity to identify areas of Stage 5 content which may need to be reviewed to meet the needs of students. | A student:* analyses representations of data in order to make predictions and draw conclusions MS1-12-2
* solves problems requiring statistical processes MS1-12-7
* chooses and uses appropriate technology effectively and recognises appropriate times for such use MS1-12-9
* uses mathematics argument and reasoning to evaluate conclusions, communicating a position clearly to others MS1-12-10

Related Life Skills outcomes: MALS6-2, MALS6-7, MALS6-8, MALS6-13, MALS6-14 |

| Prerequisite knowledge | Assessment strategies |
| --- | --- |
| This unit links to the Stage 5 units MA5.2-16SP Bivariate Data Analysis and MA5.2-9NA Linear Relationships. In addition students should have studied the Stage 6 topic MS-A2 Linear Relationships. |  |

All outcomes referred to in this unit come from [Mathematics Standard Stage 6](https://syllabus.nesa.nsw.edu.au/mathematics-standard-stage6/) Syllabus
© NSW Education Standards Authority (NESA) for and on behalf of the Crown in right of the State of New South Wales, 2017

Glossary of terms

| Term | Description |
| --- | --- |
| Association | An association is a relationship/interconnection between two variables, their values change according to a pattern |
| Bias | Bias generally refers to a systematic favouring of certain outcomes more than others, due to unfair influence (knowingly or otherwise). |
| Bivariate data | Bivariate data is data relating to two variables that have both been measured on the same set of items or individuals. For example, the arm spans and heights of 16-year-olds, the sex of primary school students and their attitude to playing sport. |
| Correlation | measures the strength of the linear relationship between a pair of variables or datasets |
| Dependant variable | A dependent variable within a statistical model is one whose value depends upon that of another. It is represented on the vertical axis of a scatterplot. The dependent variable is also known as the outcome variable or the output of a function. |
| Extrapolation | Extrapolation occurs when the fitted model is used to make predictions using values that are outside the range of the original data upon which the fitted model was based. Extrapolation far beyond the range of the original data is a dangerous process as it can sometimes lead to quite erroneous predictions. |
| Independent variable | An independent variable within a statistical model is one whose outcomes are not due to those of another variable and is represented on the horizontal axis of a scatterplot. The independent variable is also referred to as the input of a function. |
| Interpolation | Interpolation occurs when a fitted model is used to make predictions using values that lie within the range of the original data. |
| Line of best fit | A line of best fit is a line drawn through a scatterplot of data points that most closely represents the relationship between two variables. |

| Lesson sequence | Content | Suggested teaching strategies and resources  | Date and initial | Comments, feedback, additional resources used |
| --- | --- | --- | --- | --- |
| The statistical investigation process(1 lesson) | S3.1 The statistical investigation process for a survey* Understand and use the statistical investigation process: identifying a problem and posing a statistical question, collecting or obtaining data, representing and analysing that data, then communicating and interpreting findings
* Identify the target population to be represented (ACMEM132)
 | **Introducing the statistical investigation process*** The teacher to introduce the statistical investigation process.
* Collaborative ordering activity: Students can discuss what they believe is involved in each phase and where it fits in the process of a statistical investigation.
* Identify a problem
* Pose a statistical question
* Collect or obtain data
* Representing and analysing data
* Communicating and interpreting findings
* The teacher summarises / redirects students to define each phase of a statistical investigation.

Note: To reinforce the phases of the statistical investigation process, they can be referred to throughout the topic. **Identifying the target population*** The teacher defines the target population. i.e. the group of individuals or objects you want to research or analyse. Consider, who we are drawing conclusions about. Example:
* We want to know how many parent‘s read the school newsletter to students. The target population would be the parents and students from that school.
* If researching feedback on an advertisement or product, you would target the intended recipient of it.
* Student activity: The teacher to provide or students to brainstorm a number of investigation ideas.
* Students to identify the target population.
* Students to choose an investigation topic to use throughout S3.1. Sample ideas for investigations include:
* The local council has donated money to build a facility for use by teenagers in the local region. The council needs to know to spend the money effectively.
* [AMSI Data Investigation and Interpretation](http://amsi.org.au/teacher_modules/Data_Investigation_and_interpretation5.html#Initial_questions_that_can_motivate_an_investigation)
 |  |  |
| Designing a survey(2 lessons) | * Understand and use the statistical investigation process: identifying a problem and posing a statistical question, collecting or obtaining data, representing and analysing that data, then communicating and interpreting findings
* Investigate questionnaire design principles, such as simple language, unambiguous questions, consideration of number of choices, how data may be analysed to address the original question, issues of privacy and bias, ethics, and responsiveness to diverse groups and cultures **AAM** Aboriginal and Torres Strait Islander histories and cultures icon Asia and Australia’s engagement with Asia icon Ethical understanding icon Difference and diversity icon Work and enterprise icon
 | **Designing surveys*** The teacher introduces the concept of survey design by viewing [Leading Questions – Yes Prime Minister](https://www.youtube.com/watch?v=G0ZZJXw4MTA)

The teacher leads a discussion on good and bad questionnaires and issues that need to be considered when designing them by providing examples of good and bad questions.Resources: [5 common survey question mistakes that’ll ruin your data](https://www.surveymonkey.com/mp/5-common-survey-mistakes-ruin-your-data/), [10 examples of biased survey questions](https://surveytown.com/10-examples-of-biased-survey-questions/)* The teacher defines the characteristics of a well-designed questionnaire:
* Simple Language
* Unambiguous questions
* Consideration of number of choices
* They address the original question
* Issue of privacy and bias
* Ethics
* Responsiveness to diverse groups
* The teacher introduces the concept of piloting a survey to address the issues of non-response or unexpected response.
* Student activity: Designing survey questions
* Students decide on a topic to investigate. Topic ideas: [AMSI Data Investigation and Interpretation](http://amsi.org.au/teacher_modules/Data_Investigation_and_interpretation5.html#Initial_questions_that_can_motivate_an_investigation)
* Students develop a well-designed questionnaire for the topic.
* Students pilot the questionnaire and re-evaluate their quality including identifying possible different interpretations of the question.
* Students develop and pilot a poorly-designed question for the topic.
* Students examine the impact a poorly designed questionnaire has on the results.

Note: The survey could be developed and piloted using technology such as a google form. |  |  |
| Representing and analysing data(2 lessons) | * Understand and use the statistical investigation process: identifying a problem and posing a statistical question, collecting or obtaining data, representing and analysing that data, then communicating and interpreting findings
* Implement the statistical investigation process to answer questions that involve comparing the data across two or more groups Critical and creative thinking icon  Information and communication technology capability icon
 | **Representing and analysing the data*** The teacher leads a review of representing and analysing data:
* Mean
* Median
* Mode
* Range
* Standard deviation
* Inter quartile range
* Outliers
* Graphical representations of data
* Students represent and analyse the data they have collected from previous lessons.

Note: This could be completed using a spreadsheet, such as MS Excel.Resource: Collating-data-using-MS-Excel.DOCX, sample-collating-data.XLSX**Communicating and interpreting the findings*** The teacher models interpreting the data analyses and communicating a response to the original questions/topic.
* Students interpret and communicate their findings.

**Implementing the statistical investigation process*** Students have now considered all phases of the statistical investigation process.
* Identify a problem
* Pose a statistical question
* Collect or obtain data
* Representing and analysing data
* Communicating and interpreting findings
* Students implement the statistical investigation process to answer questions that involve comparing the data across two or more groups. For example:
* A local council wanted to decide where to build a new skate park and was choosing out of two suburbs. Where should it be built?
 |  |  |
|  | S3.2 Exploring and describing data arising from two quantitative variables | **Sources of Data*** Staff may like to use these sources of data throughout the S3.2 unit to contextualise the concepts being taught:
* [Australian Bureau of Statistics](http://www.abs.gov.au/) (ABS)
* [Australian Bureau of Meteorology](http://www.bom.gov.au/) (BOM)
* [Australian Sports Commission](http://www.ausport.gov.au/)
* [Australian Institute of Health and Welfare](http://www.aihw.gov.au/) (AIHW)
* [World bank](https://data.worldbank.org/)
* [Google trends](https://trends.google.com/trends/?geo=US)
* [Statista](https://www.statista.com/)
* [Google Public Data](https://www.google.com/publicdata/directory)
 |  |  |
| Introducing bivariate data(1 lesson) | * Construct a bivariate scatterplot to identify patterns in the data that suggest the presence of an association (ACMGM052) **AAM** **Paperclip icon**  Information and communication technology capability icon
 | **Introducing associations of bivariate data sets*** Teacher to:
* define bivariate data sets (refer to Glossary)
* introduce students to the concept of an association in the context of statistics.
* define linear and non-linear associations
* Student activity: Examination of Scatterplots to:
* identify and describe the type of relationships between the two variables. (Linear or Non-Linear)
* identify dependent and independent variables and that the independent variable is shown on the horizontal axis, the dependant on the vertical.
* Identify that bivariate data sets can contain time as a variable and that time is always an independent variable, shown on the horizontal axis. Refer to the sources of data.

Resource: associations-of-bivariate-data-sets.DOCX* Student activity: Students collect data and construct a scatterplot to identify an association. e.g. height versus arm span.
 |  |  |
| Describing bivariate associations(2 lessons) | * Use bivariate scatterplots (constructing them when needed) to describe the patterns, features and associations of bivariate datasets, justifying any conclusions **AAM** Paperclip icon
* Describe bivariate datasets in terms of form (linear or non-linear) and, in the case of linear, the direction (positive or negative) and strength of any association (strong, moderate or weak)
* Identify the dependent and independent variables within bivariate datasets where appropriate
* Describe and interpret a variety of bivariate datasets involving two numerical variables using real world examples from the media, or freely available from government and business datasets  Information and communication technology capability icon Civics and citizenship icon
 | **Examining the direction and strength of linear associations*** Teacher to lead discussions on the direction (positive or negative) of the trend and the strength of the association (strong/moderate/weak). Approximating a line of best fit may assist students to identify the direction of the association.
* Student activity: Examination of scatterplots to determine the direction of a trend and strength of the association.

Resource: direction-and-strength-of-a-linear-association.docx* Student activity: Students complete [Scatter plot capture](https://teacher.desmos.com/activitybuilder/custom/58cc26d4c722f106146a8310) and use observations to make predictions about future points in the plot. Students focus on linear vs nonlinear association, strong vs weak association, and positive vs negative plots
* Student activity: Students use the [GapMinder](https://www.gapminder.org/tools/#$state$marker$color$data=data_fasttrack&which=income_groups&spaceRef=entities;;;&chart-type=bubbles) website to examine real life trends and correlations between variables such as life expectancy, education and income of countries.
* Teacher to lead discussions on causality.
* Students view [The danger of messing up causality and correlation](https://www.youtube.com/watch?v=8B271L3NtAw)
* Student activity: Examine [Spurious Correlation](http://www.tylervigen.com/) website and discuss whether there is causality.
* Students to conclude that correlation does not equal causation.
* Teachers need to be mindful of any data that may cause distress to students in their class.
* Students use the sources of the data to describe and interpret a variety of bivariate datasets.

Resources: how-to-guide-sourcing-bivariate-data-using-google-trends.DOCX |  |  |
| Developing linear models(2 lessons) | * Model a linear relationship to the data by fitting a line of best fit by eye and by using technology (ACMEM141, ACMEM142) **AAM** **Paperclip icon**  Information and communication technology capability icon
 | **Introducing lines of best fit*** Teacher to lead discussions on what a line of best fit is and why it is important:
* It enables predictions to be made using the model.
* Example: If we have sales history we can then predict how much we will sell in the future.
* Student activity: Fit a line of best fit by eye and consider the concept of residuals using the [Desmos - Line of Best Fit](https://teacher.desmos.com/activitybuilder/custom/56fab6bc1ab86b1f0600369d) activity.

**Developing a line of best fit by eye*** Teacher to model the fitting of a line of best fit by eye using a digital scatterplot and finding the equation of the line of best fit.

Resource: how-to-guide-Desmos-lines-of-best-fit-by-eye.DOCX* Student activity: Fitting a line of best fit by eye

Part 1: Students use technology to produce lines of best by eye.After completion of Part 1, teacher to lead a discussion on * Which graphs they found more difficult to fit by eye? (compare this to the correlation coefficient/or description of the strength of association for each data set)
* Why do students obtain different lines?

Student activity: Part 2: Students develop lines of best fit by eye for practical data and interpret the intercept and gradient of the line with respect to the variables graphed. Note: Interpreting the intercept and gradient is content from MS-S4 (Mathematics Standard 2)Resources: fitting-a-line-of-best-fit-by-eye-activity.DOCX, data-file-1.XLSX, data-file-2.XLSX, how-to-guide-Desmos-lines-of-best-fit-by-eye.DOCX |  |  |
| Making predictions(1 lesson) | * Use the line of best fit to make predictions by either interpolation or extrapolation (ACMEM145) **AAM**  **Paperclip icon**  Information and communication technology capability icon
* Recognise the limitations of interpolation and extrapolation (ACMEM146)
 | **Introducing interpolation and extrapolation*** Teacher to ask students to explain how a model could be used to make predictions.
* Teacher to define:
* Interpolation: A fitted model is used to make predictions using values that lie within the range of the original data.
* Extrapolation: A fitted model is used to make predictions using values that are outside the range of the original data.
* Note: Interpolation and extrapolation can be related to the application of Hawk-eye to tennis and cricket. Refer to The Guardian article [Hawk-eye at Wimbledon](https://www.theguardian.com/science/sifting-the-evidence/2013/jul/08/hawk-eye-wimbledon).

**Predicting by Interpolating and Extrapolating*** The teacher models methods of making predictions by interpolating and extrapolating. Teachers may like to use this [Geogebra app](https://ggbm.at/XwC4p4Qc) linking Fuel Use and Engine Size to illustrate the concept:
* The graphical method: use the graph of a line of best fit to read predicted values.

Students may need to extend their line of best fit to allow extrapolation.* The algebraic method: substitute into the equation of a line of best fit and then evaluate the resulting expression or solve the resulting equation to make a prediction.
* Students compare the predictions obtained using the two methods and consider the benefits of each.
* Student activity: Students interpolate and extrapolate to make predictions, including in context, and examines variations in the predictions obtained using the graphical and algebraic methods.
* Resources: interpolating-and-extrapolating-activity.DOCX, data-file-1.XLSX, data-file-2.XLSX, how-to-guide-Desmos-lines-of-best-fit-by-eye.DOCX

**Examining the limitations of interpolation and extrapolation*** Student activity: Students use the [Alligator Investigation](https://teacher.desmos.com/activitybuilder/custom/5670aca10255cd610798be54) to look at the limits of linear models. “An enormous alligator lurks in the swamp. Can scatterplots and least-squares regression tell you if you have enough animal tranquilizer to stay safe?”
* Teacher to lead students to identifying the limitations of making predictions using a model. For example:
* If a model’s correlation is weak, the accuracy of predictions will diminish.
* How confident are we that the model will behave in the same pattern? When extrapolating, we need to determine if it is reasonable to assume the model is valid outside of our data range.
* NESA exemplar question

Ahmed collected data on the age (a) and height (h) of males aged 11 to 16 years. He created a scatterplot of the data and constructed a line of best fit to model the relationship between the age and height of males.A graph showing the age and height of males and a line of best fit for the data.* 1. Determine the gradient of the line of best fit shown on the graph.
	2. Explain the meaning of the gradient in the context of the data.
	3. Determine the equation of the line of best fit shown on the graph.
	4. Use the line of best fit to predict the height of a typical 17-year-old male.
	5. Why would this model not be useful for predicting the height of a typical 45-year-old male?

Resource: ms-s3-nesa-exemplar-question-solutions.DOCX |  |  |
| Collecting, modelling and interpreting data(2 lessons) | * Collect data, interpret and construct graphs using contexts, for example, sustainability, household finance and the human body **AAM** Sustainability icon Personal and social capability icon
 | Collecting, modelling and interpreting data* Students are to collect bivariate data, graph the data using a scatterplot, develop a line of best fit by eye and interpret results using contexts. Sample ideas for some contexts are listed below:
* Mobile phone battery life. Resource: mobile-phone-battery-life-activity.docx
* The human body: Students test the ratios identified by da Vinci in his [Vitruvian Man](https://www.academickids.com/encyclopedia/index.php/Vitruvian_Man) by comparing them with a bivariate model developed from measurements taken from within the class:
* Sustainability: [Global climate change data](https://data.worldbank.org/topic/climate-change?view=chart) can be sourced from the World Bank and downloaded in MS Excel: Ideas to model includes:
* [Renewable energy consumption](https://data.worldbank.org/indicator/EG.FEC.RNEW.ZS?view=chart) over time
* [Greenhouse gas emissions](https://data.worldbank.org/indicator/EN.ATM.GHGT.KT.CE?view=chart) over time
* [Forest abundance](https://data.worldbank.org/indicator/AG.LND.FRST.ZS?view=chart) over time
* [Forest abundance](https://data.worldbank.org/indicator/AG.LND.FRST.ZS?view=chart) to temperature
* [Urbanisation](https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS) over time
* Distance to coast and temperature
* Living conditions and impact on children
* Household finance:
* [ABS](https://www.abs.gov.au/AUSSTATS/abs%40.nsf/DetailsPage/6523.02017-18?OpenDocument) income, wealth and debt data
* [Inflation](https://data.worldbank.org/indicator/FP.CPI.TOTL.ZG?view=chart) over time
* [Lending interest rates](https://data.worldbank.org/indicator/FR.INR.LEND?view=chart) over time
* [Deposit interest rates](https://data.worldbank.org/indicator/FR.INR.DPST?view=chart) over time
* Level of education to level of income
* Compare the cost of renting similar properties in different locations over time
* Compare the median price of properties in different locations over time
 |  |  |

Reflection and evaluation

Please include feedback about the engagement of the students and the difficulty of the content included in this section. You may also refer to the sequencing of the lessons and the placement of the topic within the scope and sequence. All ICT, literacy, numeracy and group activities should be recorded in the ‘Comments, feedback, additional resources used’ sections.