

Mathematics Standard 1 Year 12

Statistical Analysis Topic Guidance

Mathematics Standard 1 Year 12 Statistical Analysis Topic Guidance

[Topic focus 3](#_Toc496185267)

[Terminology 3](#_Toc496185268)

[Use of technology 3](#_Toc496185269)

[Background information 3](#_Toc496185270)

[General comments 4](#_Toc496185271)

[Future study 4](#_Toc496185272)

[Subtopics 4](#_Toc496185273)

[MS-S3: Further Statistical Analysis  5](#_Toc496185274)

[Subtopic focus 5](#_Toc496185275)

[S3.1 The statistical investigation process for a survey 5](#_Toc496185276)

[Considerations and teaching strategies 5](#_Toc496185277)

[Suggested applications and exemplar questions 5](#_Toc496185278)

[S3.2 Exploring and describing data arising from two quantitative variables 5](#_Toc496185279)

[Considerations and teaching strategies 5](#_Toc496185280)

[Suggested applications and exemplar questions 6](#_Toc496185281)

# Topic focus

*Statistical Analysis* involves the collection, display, analysis and interpretation of data to identify and communicate key information.

Knowledge of statistical analysis enables the careful interpretation of situations and raises awareness of contributing factors when presented with information by third parties, including the possible misrepresentation of information.

The study of statistical analysis is important in developing students’ understanding of how conclusions drawn from data can be used to inform decisions made by groups such as scientific investigators, business people and policy-makers.

# Terminology

|  |  |  |
| --- | --- | --- |
| associationbiasbivariatedatasetdependentextrapolationhousehold finance | independent interpolationline of best fit ****linearlinear relationshipnon-linearquestionnaire design | scatterplot statistical investigation processstatistical questionsurveysustainabilitytarget populationvariable |

# Use of technology

Statistical software and spreadsheets should be used in the teaching and learning of the topic area. Spreadsheets are widely used in the workplace and are a suitable tool for tabulating, sorting and graphing data.

Graphing software can be used to fit a line of best fit to data and make predictions by interpolation or extrapolation.

Real data that is relevant to students’ experience and interest areas can be sourced online.

Students could use online survey tools to collect data.

# Background information

The history of gathering and recording official statistics begins back in Babylonian times, just after the invention of writing, when agricultural data was collected. There is also evidence that ancient cultures collected official statistics in order to assess military strength, tax obligations and political status.

In the Inca Empire (1000–1500 AD) each tribe had a designated statistician who kept records of information such as the number of people, llamas, marriages and possible army recruits in the tribe. Data was stored using a system of knots in coloured ropes.

The modern form of a periodic census is closely associated with the rise of democracy. This is because a count of the population is essential for a truly representative government.

In the 1880s Sir Francis Galton, Charles Darwin’s cousin produced the first bivariate scatterplot, which showed a correlation between children’s height and their parent’s height.

# General comments

Materials used for teaching, learning and assessment should include or use current information from a range of sources including, but not limited to, surveys, newspapers, journals, magazines and the internet.

Real data should be used in the teaching and learning of this topic. Online data sources include the [Australian Bureau of Statistics](http://www.abs.gov.au/) (ABS), the [Australian Bureau of Meteorology](http://www.bom.gov.au/) (BOM), the [Australian Sports Commission](http://www.ausport.gov.au/) and the [Australian Institute of Health and Welfare](http://www.aihw.gov.au/) (AIHW) websites.

In their study of statistics, students should develop an understanding of the importance of analysing data in planning and decision-making by governments and businesses.

# Future study

Students may be asked to analyse data and produce a report in subjects that they are studying for the HSC or in post-school contexts and training areas. This topic will set a firm baseline for knowledge, understanding and skills in statistical analysis.

The ability to analyse and critically evaluate statistical information will provide students with the confidence and skills that help them become discerning citizens.

# Subtopics

* MS-S3: Further Statistical Analysis 

## MS-S3: Further Statistical Analysis Paperclip icon

### Subtopic focus

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 awareness 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.

## S3.1 The statistical investigation process for a survey

### Considerations and teaching strategies

* Students look at actual data to investigate examples of datasets where the statistical investigation process could apply, for example data regarding average income per day for the world population in different years.
* Effective questionnaire design includes considerations such as:
	+ simple language
	+ unambiguous questions
	+ requirements for privacy
	+ freedom from bias
	+ the number of choices of answers for questions, for example if only two choices are given, this may force an opinion from the respondents in relation to a particular question, while for other questions it may be appropriate to allow a neutral choice.

### Suggested applications and exemplar questions

* Prepare questionnaires and discuss consistency of presentation and possible different interpretations of the questions.
* Teachers could briefly address the issues of non-response or unexpected response and discuss why questionnaires are often ‘piloted’ before being finalised.

## S3.2 Exploring and describing data arising from two quantitative variables

### Considerations and teaching strategies

* Suitable datasets for statistical analysis could include, but are not limited to, home versus away sports scores, male versus female data (for example height), young people versus older people data (for example blood pressure), population pyramids of countries over time, customer waiting times at fast-food outlets at different times of the day, and monthly rainfall for different cities or regions.
* Biometric data could be obtained to construct lines of fit and to compare dependent and independent variables by using a spreadsheet or other appropriate digital technology. Scatterplots drawn could include height versus arm-span, height versus weight, hip-height versus stride length, hand-span versus height, or foot-length versus height. Gender could be added, where appropriate, as an additional variable.

### Suggested applications and exemplar questions

* Students compare the cost of renting similar properties in different locations.
* Predictions could be made using the line of fit, for example predicting a person’s height based on their arm-span or arm-length. Students should assess the accuracy of the predictions by measurement and calculation in relation to an individual not in the original data set.

For example:

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.



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?