# Introduction to scatter plots

Students collect data and explore how a scatter plot can be used to compare bivariate data.

Students will need at least one digital device per pair to interact with Microsoft Excel and Desmos during this lesson.

Students will need access to a copy of the associated Microsoft Excel file *Introduction to scatter plots* during this lesson.

## Visible learning

### Learning intention

* To be able to use a scatter plot to determine if there is an association between 2 variables.

### Success criteria

* I can explain the difference between 1-variable and 2-variable data.
* I can explain when it is appropriate to represent data using a scatter plot.
* I can represent bivariate data in a scatter plot.

### Syllabus outcomes

A student:

* develops understanding and fluency in mathematics through exploring and connecting mathematical concepts, choosing and applying mathematical techniques to solve problems, and communicating their thinking and reasoning coherently and clearly **MAO-WM-01**
* displays and interprets datasets involving bivariate data **MA5-DAT-C-02**

[Mathematics K–10 Syllabus](https://curriculum.nsw.edu.au/learning-areas/mathematics/mathematics-k-10-2022/overview) © NSW Education Standards Authority (NESA) for and on behalf of the Crown in right of the State of New South Wales, 2022.

Table 1: lesson summary

|  |  |  |  |
| --- | --- | --- | --- |
| Section | Summary of activity | Teaching strategies | Teaching points |
| Launch | Students throw a paper ball into a bin and graph the results. |  | Recognise the limitations of the data and the graphs to represent it. |
| Explore | Using [Appendix A](#_Appendix_A), students collect data from tossing paper into the bin, counting the number of successful shots and the distance they were standing from the bin. They then graph the data using Microsoft Excel *Introduction to scatter plots* and answer questions about the graphs using the prompts in [Appendix B](#_Appendix_B). | Think-Pair-Share  Gallery walk | Common graphs, such as a pie graph (sector graph) and a column graph, are not able to effectively graph bivariate data or represent an association between 2 variables. |
| Summarise | Students consider why some graphs are better for representing an association between 2 variables, define the term ‘bivariate’, and brainstorm ideas of data that might have an association. Students complete a Desmos activity ([bit.ly/RobotScatter plot](https://bit.ly/RobotScatterplot)) to explore how to draw and interpret scatter plots. They write notes using [Appendix C](#_Appendix_C). Students complete [Appendix D](#_Appendix_D) to identify when a scatter plot is the most appropriate graph to use to represent data. | Pose-Pause-Pounce-Bounce  Frayer diagram  Two stars and a wish  Brainstorm | Understand that a scatter plot is a suitable graph to represent bivariate data and that a scatter plot can show when there is an association between 2 variables. |
| Apply | Students use a Desmos activity ([bit.ly/dappercats](https://bit.ly/dappercats)) to make sense of how scatter plots can show an association between variables. |  | A scatter plot can effectively show an association between variables. |

## Activity structure

Please use the associated Microsoft Excel document *Introduction to scatter plots* in this lesson.

### Launch

1. Scrunch up a piece of paper. While standing close to a classroom bin, toss the paper into the bin. Then, ask students if they could land the paper in the bin too.
2. Retrieve the piece of paper or scrunch up another piece of paper and give it to a student, asking them to have a turn at throwing the paper in the bin.
3. Pose the following question to the class: How are we going to record the data before anyone else has a shot?

Suggested ways to record the data could include a tally of successes only, a tally of successes and misses, or a series of ticks and crosses.

1. Continue to ask all students to have a shot, recording the results on the board.

Students may try to move so that they are in a better position to shoot, while others may shoot from wherever they are in the classroom. The aim is to focus student attention on the success rate of the shots, rather than their proximity to the bin.

1. Once all students have had a shot and the data has been recorded, ask students to work in pairs to represent the data graphically.

The graphs drawn will depend on the type of data the class chose to collect. If only the number of successful and unsuccessful shots was collected, graphs may include a sector graph, a column graph and a bar graph.

Students might use Microsoft Excel to record their data and create a graph.

1. Generate a class discussion about the graphs, which types of graphs were drawn, why those graphs were drawn, and what information the graphs show.

Focus the discussion on the limitations of the collected data, the challenges in creating graphs which represented the data in a useful way, and the restricted amount of information available in both the data and the graphs.

### Explore

1. Pose the question ‘If I am throwing a paper ball into the bin, does it matter how far away I am from the bin?’ and encourage students to share their thoughts.
2. Tell students that they are going to work in pairs to conduct an experiment to find out if the distance from the bin affects the success of getting the paper ball in the bin.
3. In a Think-Pair-Share ([bit.ly/thinkpairsharestrategy](https://bit.ly/thinkpairsharestrategy)), ask students to predict what they think the results of the experiment will show.
4. Provide pairs of students with the following equipment:

* a piece of A4 paper to scrunch into a ball
* a ruler or measuring tape
* a receptacle to throw a paper ball into
* a copy of Appendix A ‘Free throws’ to record the data on.

Students can use various items as their receptacles such as a bin, buckets, containers and lids.

It is important for each group to take the same number of shots for each measured distance.

1. Tell the students that they will need to make 5 shots each from each measured distance.
2. Provide students with time to set up and complete their experiment.
3. With one device per pair, ask students to record their data in the Excel spreadsheet *Introduction to scatter plots*.

Students need access to a copy of the Excel spreadsheet in the associated Microsoft Excel file *Introduction to scatter plots*.

1. Using Excel, ask students to draw a graph to represent their data. There are instructions in the Microsoft Excel spreadsheet.
2. Distribute a copy of Appendix B ‘What do the graphs say?’ which asks students to answer questions about 2 graphs.
3. Ask students, in pairs, to do a gallery walk ([bit.ly/DLSgallerywalk](https://bit.ly/DLSgallerywalk)) and select 2 different types of graphs to compare.

Students may choose to compare 2 different graphs from 2 different pairs of students or compare their graph with a different type of graph from another pair of students. Two column graphs with different data are not different types of graphs.

This activity is designed to show students that the types of graphs they are used to drawing are not very suitable for representing bivariate data. It is ok at this point for them to draw graphs that do not accurately represent the data.

### Summarise

1. Use a Pose-Pause-Pounce-Bounce questioning strategy [PDF 557 KB] ([bit.ly/posepausepouncebounce](https://bit.ly/posepausepouncebounce)) to discuss which graphs were more effective at presenting the data from the experiment. Useful question prompts may include:

* What were the graphs showing?
* Why were some graphs better at representing the data? Which graph types were these?
* Why were some graphs not useful for representing the data? Which graph types were these?

Students should recognise:

* that very few of the graphs showed the relationship between the number of successful shots and the distance from the bin.
* that the effective graphs compared the distances from the bin with the number of successful shots.
* that a scatter plot is the most effective graph for this, if it was one of the graphs they compared.
* that the least effective graphs, such as a pie graph (sector graph), did not represent the distance from the bin as a sliding scale, and did not show a comparison between distance and number of successful shots.

If no scatter plots were drawn, the teacher may find it useful to use the data from one pair of students to create and display a scatter plot.

1. State to students that the aim of the graphs was to see if there was an association between 2 variables: the number of successful shots and the distance from the bin and that a scatter plot is particularly effective for displaying 2 variables, called ‘bivariate data’.
2. Use the Pose-Pause-Pounce-Bounce questioning strategy to discuss the meaning of the term ‘bivariate data’.

The word ‘bivariate’ means ‘of, relating to, or involving 2 variables. ‘Bi’ means 2 and ‘variate’ is derived from ‘variable’.

1. Suggest to students that another interesting example of bivariate data might be the amount of sugar versus the number of calories in a food item.
2. As a class, brainstorm ideas to answer the question ‘Can you think of other examples of bivariate data which might have an association which could be explored?’

The NSW Department of Education’s Digital Learning Selector ([bit.ly/DLSbrainstorming](https://bit.ly/DLSbrainstorming)) provides information about the benefits of brainstorming and templates which may be used for brainstorming.

Ideas could include:

* weight versus height
* temperature versus wind speed
* number of orders versus profit
* hours studying versus exam result

1. With one device between each pair of students, direct them to join and complete the Desmos classroom activity ‘Robots: Points in a Scatter Plot’ ([bit.ly/RobotScatter plot](https://bit.ly/RobotScatterplot)).

Before completing this activity, you will need to set up a Desmos classroom ([bit.ly/createdesmosclassroom](https://bit.ly/createdesmosclassroom)).

1. Distribute Appendix C ‘Frayer diagram’ to each student and ask them to complete the Frayer diagram ([bit.ly/frayerdiagram](http://bit.ly/frayerdiagram)) in pairs.
2. Students are to swap Frayer diagrams and complete peer feedback in the form of Two stars and a wish ([bit.ly/DLSpeerfeedback](https://bit.ly/DLSpeerfeedback)).
3. Students then respond to the feedback to improve their Frayer diagrams.
4. Distribute Appendix D ‘Which graph is best?’ to pairs of students. Ask students to discuss which of the graphs is more suitable for the data according to each question, explaining their reasoning.

### Apply

1. Direct students to the Desmos activity ‘Dapper Cats toys’ ([bit.ly/dappercats](https://bit.ly/dappercats)).

Before doing this activity, you will need to set up a Desmos classroom ([bit.ly/createdesmosclassroom](https://bit.ly/createdesmosclassroom)).

## Assessment and differentiation

### Suggested opportunities for differentiation

**Launch**

* Students may need the bin moved closer to improve their ability to participate.
* Students may benefit from being provided with templates of different graphs.
* Students may be prompted to evaluate the use of statistical measures, such as median, mean, and mode to analyse the data.

**Explore**

* Students may be supported by having the distances prescribed when collecting the data.
* Rather than telling students to take 5 shots from each distance, students could record their success rate as a percentage.
* Students may be supported by being provided with pairs of graphs to compare.

**Summarise**

* Students might be provided with a scatter plot and another graph to examine how a scatter plot is different from other graphs.
* Students might be given a different set of data and be asked to select data sets that might be interesting to compare.
* Provide students with statements, such as ‘The higher the temperature, the windier it is’ and ask them to identify the bivariate data which could be collected to assess the relationship.
* Students might benefit from revisiting types of data, such as quantitative and qualitative data, and so on. Students were introduced to types of data in Lesson 4 – What does Google know about me? of Year 7, Unit 2 – making decisions.

**Apply**

* Use the teacher dashboard to pace students through the activity.

### Suggested opportunities for assessment

**Launch**

* Monitor the graphs drawn and the student discussion to determine existing knowledge and understanding of types of graphs.

**Explore**

* Collect a copy of the graphs and the responses to Appendix B to determine student understanding of how to interpret a graph and determine the suitability of a graph.

**Summarise**

* Collect a copy of student’s Frayer diagrams to assess student understanding of scatter plots.
* Monitor responses in the Desmos activity to assess whether students understand how to plot bivariate data.
* Monitor responses to the ‘Which graph is best’ activity to assess student reasoning when justifying which graph is best.

**Apply**

* Monitor responses in the Desmos activity to assess student understanding of how a scatter plot represents an association between bivariate data.

## Appendix A

### Free throws

Record your data in the table below.

|  |  |
| --- | --- |
| Distance from bin | Number of successful shots |
|  |  |
| 100 cm |  |
|  |  |
|  |  |
| 250 cm |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

## Appendix B

### What do the graphs say?

|  |  |  |
| --- | --- | --- |
| Question | Graph 1 | Graph 2 |
| What type of graph is it? |  |  |
| How many successful shots were made from one metre? |  |  |
| How many successful shots were made from 250 cm? |  |  |
| How is the distance from the bin represented in the graph? |  |  |
| Does the graph show the data effectively? Explain why or why not. |  |  |

## Appendix C

### Frayer diagram

Frayer diagram template for scatter plots. 
There are 4 quadrants: Definition, Facts/characteristics, Examples and Non-examples.

## Appendix D

### Which graph is best?

1. What was the volume of storage at Menindee reservoir at the end of July 2023?

**Line graph**

**A line graph showing the storage at Menindee Reservoir, Copi Hollow, from February 2023 to February 2024. The x-axis is labelled 'Month' and lists the months from February 2023 to February 2024. The y-axis is labelled 'Volume (Megalitres) and ranges from zero to 12000.
The line shows that storage levels were just over 11 000 Megalitres in February 2023. They dropped to about 7000 Megalitres in April and stayed at that level until July. They rose over the next few months to 10 000 Megalitres from September to October before dropping to 8000 Megalitres by February 2024.**

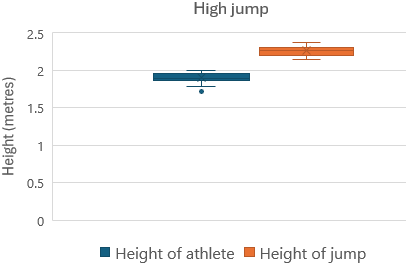
**Scatter plot**

**A scatter plot showing the storage at Menindee Reservoir, Copi Hollow, from February 2023 to February 2024. The x-axis is labelled 'Month' and has numbers ranging from 0 to 14. The y-axis is labelled 'Volume' (Megalitres)' with a range of zero to 12000.
The data shows the volume for each month, starting at 11 000 Megalitres, dropping to about 7000 Megalitres, rising to about 10 000 Megalitres and then dropping to about 8000 Megalitres.**

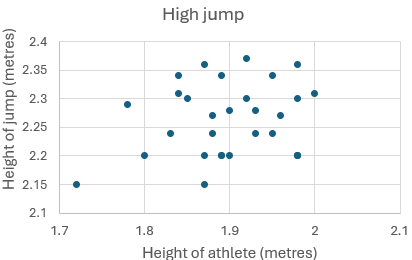
Source: Adapted from Murray-Darling Basin Authority (2024)

1. Which graph is better for answering the question? Explain why.
2. What is the relationship between the height of an athlete and the height of their jump in a high jump competition?

**Box plot**



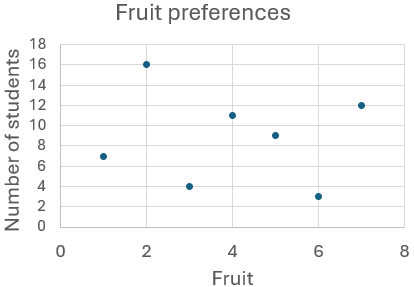
**Scatter plot**



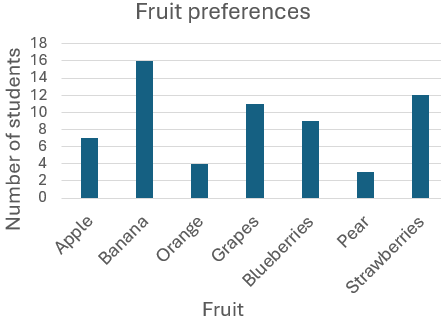
Source: Adapted from International Olympic Committee (2024)

1. Which graph is better for answering the question? Explain why.
2. What are the fruit preferences of students in the class?

**Scatter plot**

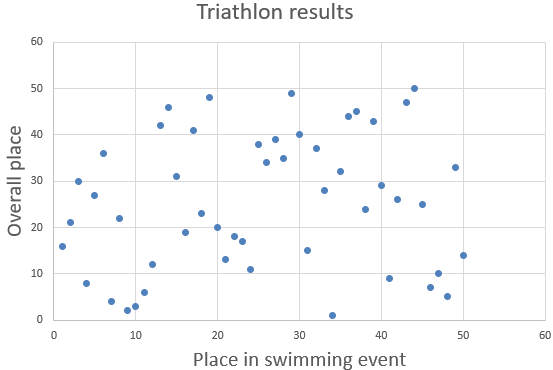


**Column graph**

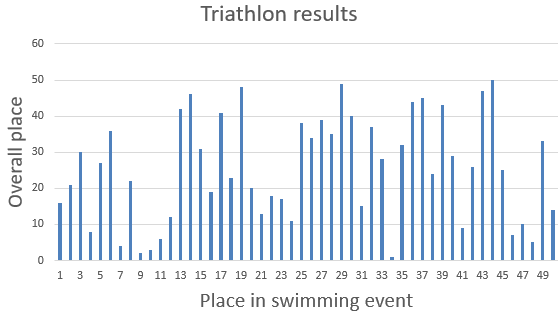


1. Which graph is better for answering the question? Explain why.
2. Is there a relationship between an athlete’s place in the swim event and their overall place in a triathlon?

**Scatter plot**



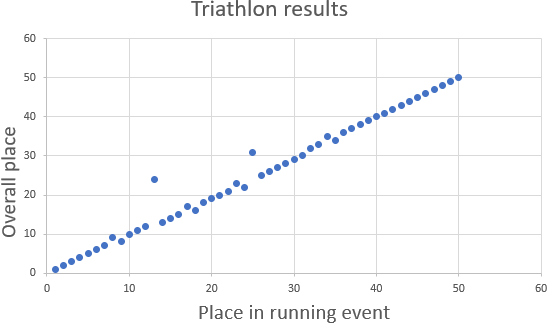
**Column graph**



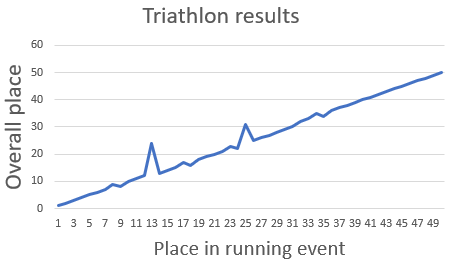
Source: Adapted from International Olympic Committee (2008)

1. Which graph is better for answering the question? Explain why.
2. Is there a relationship between an athlete’s place in the run event and their overall place in a triathlon?

**Scatter plot**



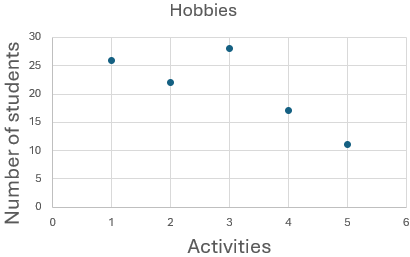
**Column graph**



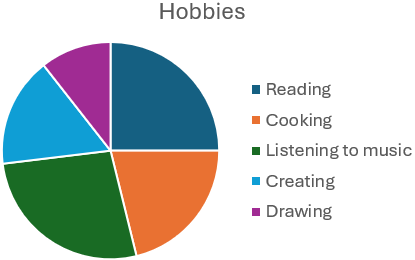
Source: Adapted from International Olympic Committee (2008)

1. Which graph is better for answering the question? Explain why.
2. What do students like to do in their spare time?

**Scatter plot**



**Sector graph**



1. Which graph is better for answering the question? Explain why.

## Sample solutions

### Appendix C – Frayer diagram

An example of a completed Frayer diagram for scatter plots. The definition for a scatter plot is 'A graph showing 2 data sets (bivariate data) where the position of the dots on a Cartesian plane shows the relationship between the data sets.' The facts/characteristics are: 'The points are placed higher or lower, left or right, depending on the variables' and 'Scatter plots are used to look for relationships between variables.'
There is an example of a scatter plot and there is a sector graph and a line graph shown as non-examples.

### Appendix D – Which graph is best?

1. The line graph is best. It shows the volume of water in storage and the year as continuous data. So, it’s possible to find, for example, the volume at any time, from February 2023 to February 2024 and it’s easier to see the changes over time. The scatter plot represents the data as discrete data. It’s possible to see the changes over time, but it’s not possible to use the scatter plot to predict the volume between the data points.
2. The scatter plot is best. It shows how the height of the jump is related to the person’s height, with each dot representing a person. The box plot has separated the data. It does show the range and the IQR for the height of the jump and the height of the athletes, but it doesn’t show the association effectively.
3. The column graph is best. It shows the different fruits and how many people preferred each fruit. The scatter plot has numbers instead of data labels for the fruit, which implies that the fruit has an order. Scatterplots aren’t used for categorical nominal data for this reason.
4. The scatter plot is best. It shows that the result in the swim event doesn’t have an association with the final place. The column graph has too many columns.
5. The scatter plot is best. It shows an association between the run event and the final place. The line graph makes the data continuous, so that someone could get a place of 5.5, for example, and it shows the 2 athletes that don’t fit in the line as dips in the data.
6. The sector graph is best. It clearly shows the proportion of people for each activity. The scatter plot doesn’t represent the data clearly. It doesn’t show the activities on the horizontal axis and suggests there is an order to the activities.

## References

This resource contains NSW Curriculum and syllabus content. The NSW Curriculum is developed by the NSW Education Standards Authority. This content is prepared by NESA for and on behalf of the Crown in right of the State of New South Wales. The material is protected by Crown copyright.

Please refer to the NESA Copyright Disclaimer for more information [https://educationstandards.nsw.edu.au/wps/portal/nesa/mini-footer/copyright](https://aus01.safelinks.protection.outlook.com/?url=https%3A%2F%2Feducationstandards.nsw.edu.au%2Fwps%2Fportal%2Fnesa%2Fmini-footer%2Fcopyright&data=05%7C01%7CCaitlin.Pace1%40det.nsw.edu.au%7C9c2c1a9f59c94d2df30708dafa7edb23%7C05a0e69a418a47c19c259387261bf991%7C0%7C0%7C638097720042599463%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=BzQh0UsffVZE3eO22b2Xba3p0VMOBZSHfS21FGHXtZM%3D&reserved=0).

NESA holds the only official and up-to-date versions of the NSW Curriculum and syllabus documents. Please visit the NSW Education Standards Authority (NESA) website <https://educationstandards.nsw.edu.au/> and the NSW Curriculum website <https://curriculum.nsw.edu.au/>.

[Mathematics K–10 Syllabus](https://curriculum.nsw.edu.au/learning-areas/mathematics/mathematics-k-10-2022/overview) © NSW Education Standards Authority (NESA) for and on behalf of the Crown in right of the State of New South Wales, 2022.

International Olympic Committee (2024) ‘[Entries Athletics](https://olympics.com/en/paris-2024/entries/athletics/all-npc/men-s-high-jump)’, Paris 2024, Olympics website, accessed 11 October 2024

International Olympic Committee (2008) ‘[Beijing 2008 Individual Men Results: Olympic triathlon](https://olympics.com/en/olympic-games/beijing-2008/results/triathlon/individual-men)’, Beijing 2008, Olympics website, accessed 11 October 2024

Murray-Darling Basin Authority (2024) ‘[Menindee Lakes Combined Storage (Darling)](https://riverdata.mdba.gov.au/list-view/menindee-lakes-combined-storage)’, River Murray Data, accessed 11 October 2024

**© State of New South Wales (Department of Education), 2024**

The copyright material published in this resource is subject to the *Copyright Act 1968* (Cth) and is owned by the NSW Department of Education or, where indicated, by a party other than the NSW Department of Education (third-party material).

Copyright material available in this resource and owned by the NSW Department of Education is licensed under a [Creative Commons Attribution 4.0 International (CC BY 4.0) license](https://creativecommons.org/licenses/by/4.0/).

[](https://creativecommons.org/licenses/by/4.0/)

This license allows you to share and adapt the material for any purpose, even commercially.

Attribution should be given to © State of New South Wales (Department of Education), 2024.

Material in this resource not available under a Creative Commons license:

* the NSW Department of Education logo, other logos and trademark-protected material
* material owned by a third party that has been reproduced with permission. You will need to obtain permission from the third party to reuse its material.

**Links to third-party material and websites**

Please note that the provided (reading/viewing material/list/links/texts) are a suggestion only and implies no endorsement, by the New South Wales Department of Education, of any author, publisher, or book title. School principals and teachers are best placed to assess the suitability of resources that would complement the curriculum and reflect the needs and interests of their students.

If you use the links provided in this document to access a third-party's website, you acknowledge that the terms of use, including licence terms set out on the third-party's website apply to the use which may be made of the materials on that third-party website or where permitted by the *Copyright Act 1968* (Cth). The department accepts no responsibility for content on third-party websites.