# Wisdom of the crowd

Students explore the effect on measures of centre and spread when data is added or removed.

## Visible learning

### Learning intention

* To understand how data affects the mean, median, range and mode.

### Success criteria

* I can explain what happens to the mean, median, range and mode when an outlier is added.
* I can explain what happens to the mean, median, range and mode when a cluster of numbers is added or removed from a dataset.
* I can add or remove data to increase or decrease the mean, median and mode.

### 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**
* analyses simple datasets using measures of centre, range and shape of the data   
  **MA4-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.

## Activity structure

Please use the associated PowerPoint *Wisdom of the crowd* to display images in this lesson.

### Launch

1. Display the following picture. This can also be found on slide 2 of the PowerPoint *Wisdom of the crowd.*

Figure 1: jar of jellybeans



Image licensed under [Pixabay Content License](https://pixabay.com/service/license-summary/).

Alternatively, you can make your own jar of jellybeans.

1. Ask students to write down on a piece of paper how many jellybeans they think are in the jar.
2. Collect the pieces of paper from the students and collate the data on the board in a stem-and-leaf plot.

To create a stem-and-leaf plot you can use ‘Stem and Leaf Plot Generator’ ([bit.ly/stemandleafgenerator](https://bit.ly/stemandleafgenerator)).

Alternatively, this could be an opportunity for students to further develop their understanding of the concept of constructing stem and leaf plots.

It is recommended that the stem-and-leaf plot be left on the board or in the generator as it will be used later in the lesson.

1. In visibly random groups of 3 ([bit.ly/visiblegroups](https://bit.ly/visiblegroups)) on vertical non-permanent surfaces ([bit.ly/VNPSstrategy](https://bit.ly/VNPSstrategy)) ask students to use the data in the stem and leaf plot to revise their estimate of the number of jellybeans in the jar.

Prompt students to use measures of centre such as mean, median and mode to help calculate their new estimate.

1. Ask students to do a gallery walk ([bit.ly/DLSgallerywalk](https://bit.ly/DLSgallerywalk)) of each group’s solution.
2. Ask random students to explain how their group calculated their estimate of the number of jellybeans.

### Explore

1. Explain to students that the experiment they were conducting was called the ‘wisdom of the crowd’.

Wisdom of the crowd is the idea that if you have a large and diverse enough group of people, then their errors will cancel out and in return you get an answer that is quite accurate.

1. Students will complete the desmos classroom activity ‘Strength in numbers’ (<https://bit.ly/Strength_in_numbers>) to further investigate the concept of ‘Wisdom of the crowd’ and to explore the location of the mean and median in different datasets. For more information on how to set up and assign a desmos classroom activity, visit [bit.ly/desmosclassroomstrategy](https://bit.ly/desmosclassroomstrategy).
2. Students will be asked to consider the impact of a change in the data based on different scenarios.
3. Display slide 3 from the PowerPoint *Wisdom of the crowd*.
4. Read the scenarios outlined on the slide.
5. Returning to their groups of 3 on vertical non-permanent surfaces students are to explore and explain how each scenario would change the mean, median, mode and range of the jellybean guesses.

If the stem-and-leaf plot is still available on the board, values can be added or subtracted from it.

1. Use the Pose-Pause-Pounce-Bounce question strategy (PDF 557 KB) ([bit.ly/posepausepouncebounce](https://bit.ly/posepausepouncebounce)) for students to share their thoughts.
2. Ask students to determine which measure of centre would be the most accurate in each scenario.
3. Use the following assessing and advancing questions table to progress student thinking.

Table 1: assessing and advancing questions

|  |  |
| --- | --- |
| Assessing questions | Advancing questions |
| How has the change in data affected your measures of centre? | If a different cluster of values were lost, which cluster would have the most interesting impact on the mean and the median. |
| What number, if it was called out 10 times, could most have an impact on the mean and median? | If a different number was called out, how would that change your measures of centre and spread? |
| What number, if it was called out 10 times, could have the least impact on the mean and median? | How would you describe what has happened to the mean/median/mode/range, so that you could make a better prediction of their new value in a similar situation? |

### Summarise

1. Display slide 4 from the PowerPoint *Wisdom of the crowd.*
2. Use a Think-Pair-Share for students to match each scenario with the following:

* adding a cluster of data
* removing a cluster of data
* adding an outlier to the data.

1. Distribute Appendix A ‘What happens when’ for each student to complete. This appendix provides sentence scaffolds for students to explain how the mean, median, mode and range are affected by outliers and clusters.
2. Students are to share their sentences with a partner and give feedback using Two stars and a wish ([bit.ly/2starwish](https://bit.ly/2starwish)).

### Apply

1. Survey another class of students to create a second dataset of estimates about the number of jellybeans in Figure 1.

Alternatively, if other classes have completed the Launch activity you can use that data to complete this section.

1. Students are to create a second stem-and-leaf plot for the new dataset and calculate its mean, median, mode and range.
2. In new visibly random groups of 3 at vertical non-permanent surfaces, students are to compare the 2 datasets and identify if combining the datasets would create any of the following scenarios:

* add an outlier
* add a cluster
* remove a cluster.

1. In their groups of 3 students should describe and estimate what will happen to the measures of centre and spread when the 2 datasets are combined.
2. Students are to then create a stem and leaf plot for the combined datasets and calculate the mean, median, range and mode.
3. As a group, students are to reflect if the combined results make a better estimate of the number of jellybeans in the jar, than each class on their own.

## Assessment and differentiation

### Suggested opportunities for differentiation

**Launch**

* We do not know how many jellybeans are in the jar so all students should be encouraged to participate and share their thoughts. Some students will guess, and some will make an approximation based on calculations and should be encouraged to do so. However, this is a data-gathering exercise so all guesses or approximations would be acceptable.
* Students can use the stem and leaf generator ([bit.ly/stemandleafgenerator](https://bit.ly/stemandleafgenerator)) to support them in creating the graph.
* Students should be encouraged to write out the raw data if it helps them to calculate the mean, mode, median and range.

**Explore**

* Challenge students to decide which data points to remove from a dataset, to create certain changes in the mean, median, range and mode. For example, a lowering of the median.
* Use the assessing and advancing questions to assist students to access the discussion if finding difficulties.

**Summarise**

* Working in pairs will assist students who need support with creating their sentence scaffolds.

**Apply**

* Students can use the stem and leaf generator ([bit.ly/stemandleafgenerator](https://bit.ly/stemandleafgenerator)) to support them in creating the graph.
* Students may prefer to work with raw data to calculate the mean and median.
* To increase student interest, you can change the context of the original problem. For example, how many people would fit in a bus or how many litres are there in a backyard swimming pool.

### Suggested opportunities for assessment

**Launch**

* Students working from their vertical non-permanent surface can be used as formative assessment for calculating mean, median, range and mode.

**Explore**

* When placed in groups of 3, students provide and receive peer feedback on their understanding.
* Students will demonstrate their working mathematically skills in discussions and justifications.

**Summarise**

* Collect Appendix A as evidence of learning.

**Apply**

* Students could choose a topic of interest to investigate and submit as summative assessment.

## Appendix A

### What happens when…?

Complete the table below using the sentence scaffolds.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Add an outlier | Add a cluster | Remove a cluster |
| Mean | When we add an outlier the mean… | When we add a cluster the mean… | When we remove a cluster the mean… |
| Median |  |  |  |
| Mode |  |  |  |
| Range |  |  |  |

## Sample solutions

### Scenarios

The following sample solutions attempt to explain how the scenarios would affect the data in general terms without referencing values. When discussing with your class, these statements could be made more specific to your estimates.

Given the following scenarios explore how the new data changes the mean, median, range and mode of the class data.

* A new student joined the class and guessed there were 30 jellybeans.

The mean decreased, the median moved by 1 data point or less, the range increased, and the mode remained the same.

* A group of 10 people walking by were asked to guess. They all heard a number called out, so they all guessed the same number.

The mean remained similar, the median shifted towards the cluster, the range remained the same, and the mode may be in the cluster.

* Your group lost all the estimates between 280 and 300.

The mean shifted away from the removed data, the median shifted away from the removed cluster, the range remained the same (if the maximum value was larger than 300), and the mode may have changed.

### Appendix A – What happens when…?

Complete the table below using the sentence scaffolds:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Add an outlier | Add a cluster | Remove a cluster |
| Mean | When we add an outlier, the mean moves towards the value of the outlier. | When we add a cluster, the mean may move closer to the added data. | When we remove a cluster, the mean may move slightly away from the removed data. |
| Median | When we add an outlier, the median increases or decreases by at most one data value. | When we add a cluster, the median will move closer to the added data. | When we remove a cluster, the median may move slightly away from the removed data. |
| Mode | When we add an outlier, the mode is not affected. | When we add a cluster, the mode may stay the same, or shift to be within the added cluster. | When we remove a cluster, the mode will remain the same unless contained in the cluster. |
| Range | When we add an outlier, the range increases dramatically as the outlier would be the new maximum or minimum value. | When we add a cluster, the range could remain the same if added within the minimum and maximum values or it could increase if the cluster includes the new maximum value, or minimum value. | When we remove a cluster, the range may remain the same if the cluster was within the maximum and minimum values. The range may decrease if the cluster included the maximum or minimum value. |

## 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>.

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/home](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.

**© 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.