**The Statistics and Probability Strand: Outcome SP7.1**

[Learning Space](#ls) [What Students Should…](#wss) [Key Questions](#kq) [Assessment](#assessment) [Instruction](#instruction)

|  |  |
| --- | --- |
| **Outcome** | **Indicators** |
| SP7.1 Demonstrate an understanding of the measures of central tendency and range for sets of data.[C, CN, PS, R, T]*In support of the K-12 Mathematics goals of Spatial Sense, Number Sense, Logical Thinking and Mathematical Attitude.* | 1. Concretely represent mean, median, and mode and explain the similarities and differences among them.
2. Determine mean, median, and mode for a set of data, and explain why these values may be the same or different.
3. Determine the range of a set of data.
4. Provide a context in which the mean, median, or mode is the most appropriate measure of central tendency to use when reporting findings.
5. Solve a problem involving the measures of central tendency.
6. Analyze a set of data to identify any outliers.
7. Explain the effect of outliers on the measures of central tendency for a data set.
8. Identify outliers in a set of data and justify whether or not they should be included in the reporting of the measures of central tendency.
9. Provide examples of situations in which outliers would and would not be used in reporting the measures of central tendency.
10. Explain why qualitative data, such as colour or favourite activity, cannot be analyzed for all three measures of central tendency.
 |
| **Learning Space** [**Top**](#top) |
| Up until grade 7, much of the students’ exploration of statistics has involved the collection of data, the creation of graphic displays for the data and analysis of the data based upon the graphic displays. In grade 7, the students begin their study of statistical analysis of the data through computation. Of the three measures of central tendency, the mean is likely the most mathematically difficult, but because students have likely had many encounters with “finding their average” or comparing their mark to the average, students often struggle with the other concepts of mode and median which are visually and numerically more obvious. It is important then that the students be invited to explore data sets and make comments about what they see in the data (“Wow – a lot of people travel between 10 and 20 km to work” or “Hey – your height is right in the middle of everyone’s heights”) before the formalization of terminology and the introduction of definitions. As much as possible, these explorations should consider data sets that can be represented concretely so that the students can visualize as well as numerically analyze.Frequently, students do not realize that all three measures of central tendency are called averages, all three of them indicating a type of centre to the data. It is important that this terminology is made clear because it is different from our typical daily usage.It is also important that students realize that outliers are not necessarily faulty pieces of data (an incorrect measurement or exaggerated report). Instead, they are most often naturally occurring anomalies. The reason they are removed from data sets is because they can so strongly affect the measures of central tendency, thus not accurately portraying what could be seen to be the norm.Understanding the affects of an outlier on a data set is an important learning for the students in grade seven.The students should also be observing the spread or dispersion of the data. Thinking about this characteristic of the data is partially addressed by the determination of the range, but students should also be starting to look for and talk about clustering and gaps within the data.Both Health Education and Physical Education can serve as sources to generate data for the students to analyze for measures of central tendency. In addition, data that is collected or researched in Science and Social Studies may prove interesting and engaging to the students. It is important that when carrying out data analysis that the data used is relevant and interesting to the students, otherwise their understanding of the data analysis they do will be limited to procedural activities without reasoning. |
| **What Students Should…** [**Top**](#top) |
| **Know*** The terms “mean”, “median”, “mode”, “range”, and “outliers”.
* Mean, median, and mode are all forms of averages.
 | **Understand*** How mean, median and mode represent different types of central tendencies in the data.
* How mean, median and mode can be used to convey alternate interpretations for a data set.
* Outliers are anomalies within a data set and not mistakes.
* The exclusion of outliers in calculations of measures of central tendency and range must be justified.
* How mean, median, and mode can be affected by the addition or deletion of pieces of data.
* How range can provide insight into the spread or dispersion of the overall data set, but not into trends within the data set.
* The limitations of the three measures of central tendency.
 | **Be Able to Do*** Determine mean, median and mode for a data set.
* Create a data set with a particular mean, median or mode.
* Determine the change in mean, median or mode when outliers are removed or new data is added.
* Determine the range of a data set.
* Present an argument for or against some idea or context based upon the mean, median or mode for a data set.
* Solve problems involving the mean, median, mode or range.
 |

|  |
| --- |
| **Key Questions** [**Top**](#top) |
| * Where do outliers come from?
* Why are mean, median and mode all called averages?
* What does central tendency mean?
* Why might you exclude an outlier when determining a measure of central tendency?
* What would the possible impacts of adding another piece of data be on the measures of central tendency and why?
* What are the strengths and weaknesses of data being represented by each of the three measures of central tendency?
 |
| **Suggestions for Assessment:** [**Top**](#top) |
| **Big Idea:**Measures of central tendency*.***Suggestions for assessment tasks:**1. A project can be used to assess this outcome. The students can be asked to select a context or situation that they are interested in and in which they will be able to collect numerical (quantitative data) for which they are to determine the mean, median and mode and create a presentation of their findings. Within their analysis, the students will need to determine if there are any outliers and then justify the inclusion or exclusion of those outliers in their calculations. Each student is to tell the teacher their topic, and the teacher will generate a single question for the student to answer about their data and a single piece of data that the student is to determine the affects of on the three measures of central tendency. Both the piece of data and the single question can be put in a sealed envelop that the student will open and respond to during their presentation. Along with the *Measures of Central Tendency Rubric*, the class may also want to develop an assessment rubric for the presentations and develop specific expectations for the project.

**What to look for:*** See: [*Measures of Central Tendency Rubric*](file:///C%3A%5CUsers%5Ckw426%5CAppData%5CLocal%5CTemp%5CMeasures%20of%20Central%20Tendency%20Rubric.doc)*.*
* Develop an assessment rubric to be used for assessing the project. This could include both teacher and student assessment components.
 |

|  |
| --- |
| **Suggestions for Instruction:** [**Top**](#top) |
| **Big Idea:***Measures of central tendency.***Suggestions for instructional activities**1. Have the students go in groups of 6 to 8 students and provide each group six-sided dice and linking cubes. Have each student roll a die and make a tower out of the rolled number of linking cubes. Then, have each group answer the following questions:
	* What is the smallest tower in your group?
	* What is the largest tower in your group?
	* What is the difference in size between the largest and smallest towers in your group?
	* Is there a most common tower size? Is there more than one tower size that is equally common?
	* When you organize your towers in order of smallest to largest, whose tower is in the centre of the towers and how big is it?

Have each group share their answers with the class, comparing the results and the strategies they used to get those answers. Once the students have shared and discussed their results, identify that the first three questions were designed to have them determine the range of the data. Ask the students to define “range”. After discussing these definitions, and double checking the definition against the answers to the questions, have the students write a definition of range in their journal.Next, tell the students what they determined in the fourth question is called the mode. Have the students discuss how they would define mode and then have them record their ideas in their journal.Finally, tell the students that what they have determined in the last question is called the median. Have the students discuss how they would define median and then have them record their ideas in their journal.1. Next, ask the students to go back into their groups with the same towers and ask them to determine how many linking cubes each student should get in order for each student to have the same number of linking cubes. Do not worry if the total number of linking cubes is not divisible by the number of students in the group as they will come up with ways to deal with the issue such as talking about fractions of a linking cube or removing the remainder cubes. The key is that they are getting the idea of this value being related to equal sharing across the total number. Some groups may approach the task by taking cubes from the tallest towers and giving them to the shortest towers until all towers are equal in height. Others might put all the towers together and then sort them into the total number of groups. Either approach is appropriate. Have the students share their strategies and results with the class and have the class discuss how the strategies and results are related. Some time can also be spent considering different ways to deal with the “left-overs”. Have the students generalize a procedure for how to determine how many cubes each student should have for the cubes to be shared equally amongst the students. Tell the students that this number is called the mean. Discuss how sometimes outside of mathematics classes it’s also called the average. Explain that this is not helpful in math because mean, median and mode are all considered averages or measures of the central tendency in mathematical language.
2. Have the students write a journal entry about the mean and how to determine it. Also, have the students reflect on why the mean, median and mode would be referred to as measures of central tendency.
3. In their same groups, have the students measure and record one of their heights, hand span, shoe size... Have them analyze the resulting data for its mean, median, mode and range and create a poster or other type of presentation that shows the data collected and their determination of the mean, median, mode and range.
4. If possible, use one of the group’s data and results to start a discussion of outliers. If there isn’t an appropriate data set (one that has at least one obvious outlier), then use another set of data that you have collected or researched that will be relevant to the students in the class. Have the students consider the data, without determining any of the measures of central tendency or the range and ask them to talk about things that they notice about the data. When the students mention the one or two pieces of data that are very different from the rest, tell them that such pieces of data are called “outliers”. Have the students discuss why they think the outliers occur. It is important that they do not believe that they are mistakes in the data, but rather statistical anomalies. Next, ask the students to individually hypothesize what the affects on the measures of central tendency for the data set would be if the outliers were removed. Once the individual students have a hypothesis accompanied by their reasoning, have the students share their ideas with a partner. Let the partners discuss and explore their reasoning. They may even carry out calculations, but let that happen on its own. Finally, bring the class together to discuss their ideas and to carry on discussing and developing ideas about the impact of outliers on the measures of central tendency and range. Also, have the students discuss the advantages and disadvantages of excluding outliers.
5. Provide the students with a new set of data and ask them to identify what they believe to be outliers and why, explain whether or not they would want to eliminate the outliers before determining the measures of central tendency and the range, and then determine the four values.
6. Provide the students with a new set of data and have them determine the three measures of central tendency and the range. Then, provide the students with a new piece of data and ask them to hypothesize what the impact of adding that piece of data would be on each of the values that they had determined. Have the students determine the new values only if there is disagreement on what would happen or if they are incorrect in their reasoning. Repeat the process using different new values (always returning to the original data set) that impact one or more of the four values.
 |