**Lesson 10 - Data Challenges**

Choose 3-5 questions from each section to discuss and complete. You may work alone or with a partner. Please record your work somehow, as you may be asked to show and explain it to the teacher.

**Open Questions**

**Collecting/Organizing Data**

1. Think of a situation you are interested in where someone might use information from a survey to make a decision that affects many more people than were surveyed. For example, a commercial gym might ask just some clients about what hours they would prefer and assume other clients would agree. Once you have decided on your situation, discuss how many people and what people you would survey to have a **representative** **sample**.
2. Suppose you want to suggest what new intramural sport your school should offer. Who would you ask, how many people would you ask, and where and when would you collect the data in order to get **unbiased**, reliable data?
3. How **reliable** are your results if you only **sample** girls to determine the favourite type of book read by students in your grade?
4. How would you collect information about whether the amount of housework done by Canadian men has increased in the last 10 years?
5. Look up a “fact” on different internet sites to see if you get the same data from each site (e.g., the population of a country). *How can you decide which data sources to trust*?

**Measures of Central Tendency**

1. You gather data about how long it takes various people of your age to walk 1 mile. How spread out might the data be? What might cause some of the data to be very unusual? How should that affect your **sample**? Discuss **outliers**.
2. Create a set of data in which the **mean** is greater than the **median**.
3. Jane and Fatima are in two different classes. Each had a history test. Describe the possible data showing the marks in each class if Jane’s mark of 83% might actually be viewed as better than Fatima’s mark of 90%. Explain what about your data makes this a reasonable claim.
4. Create two data sets with the same **mean**, one with all values near the mean and one with all values far from the mean. (W*hich set does the mean describe better? Why?*)
5. Use a set of data whose **distribution** across its **range** looks symmetrical, and change some of the values so that the distribution no longer looks symmetrical. Does the change affect the **median** more than the **mean**? Explain your thinking.
6. Create two sets of data with the same mode, one where the **mode** is an appropriate representation and one where the mode is not appropriate. (*What were you thinking when you made each set of data?)*
7. Create two sets of data that both have a **range** of 4. One set should have a **mean** that is much greater than the other. (*Why is the range alone not a good description of a set of data?)*
8. Use data to make a **convincing argument** that the environment is becoming increasingly polluted.

**Graphing (constructing, interpreting, drawing conclusions)**

1. Tomken students ordered the following lunch items: 90 students chose 3 slices of pizza, and 60 students chose 1 slice of pizza. (*How would you graph the data to make it seem like a lot more students chose 2 slices of pizza than 1 slice?*
2. Create a **scatter plot** to compare the lengths of the bases of several similar triangles with their areas.
3. A **scatter plot** has a couple of **outliers**, but the two **variables** that are graphed seem to have an almost **linear relationship**. Draw two possible scatter plots.
4. How is a **histogram** similar to and different from a bar graph? Use examples to support your answer.
5. Investigate websites for graphs that use equal and unequal numerical **intervals** (e.g., many Statistics Canada graphs use unequal age categories). Ask: *Do you think it is appropriate to use unequal intervals*?
6. There were people of these ages at a party: 2, 2, 4, 5, 8, 9, 10, 12, 17, 25, 30, 36, 37, 38, 40. *How might you create intervals for a* ***histogram*** *to show how many people were at the party to fairly reflect the range of ages? How would you create intervals to give a misleading impression of the age range*?
7. How do you know that the three sectors of a **circle graph** cannot represent 1/3, 1/4, and 3/10 of a set of data?
8. The graph below shows the number of paper clips students could clip together in 2 minutes. Collect your own data similar to these data. Graph your data using different intervals. How do the sizes of the **intervals** affect the impression the graph gives?

Paper Clips in 2 Minutes

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 10 - 19 |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 20 – 29 |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 30 – 39 |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 40 - 49 |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |

Number of paper clips used

**Parallel Tasks**

1. **Option #1**: A set of nine pieces of data, all of which are different, has a mean of 30 and a median of 10. What could the data values be?

**Option #2**: A set of nine pieces of data, all of which are different, has a median of 10. What could the data values be?

**Debrief Questions**:

* How do you know your median is 10?
* What is the least value you had? Did it have to be less than 10? Why?
* What was the greatest value you had? Did it have to be greater than 10?
* How did you determine your set of data values?

1. **Option #1**: Create a set of five pieces of data with a mean of 6. No more than one of the values can be 6.

**Option #2**: Create a set of five pieces of data including the values 4, 2, and 2, where the mean is 20.

**Debrief Questions**:

* What is a mean?
* How do you know that your mean is correct?
* How did you figure out the other numbers?

1. **Option #1**: Create a set of data that can be appropriately described by using a histogram. Create that histogram.

**Option #2**: Create a set of data that can be appropriately described by using a stem and leaf plot. Describe that plot.

**Debrief Questions**:

* What are the groupings for the data that you used?
* Why did it make sense to use that type of graph for these data?
* Looking at the graph, what picture of the data does it give the viewer?

1. The line y = 3x + 2 is a good, but not perfect, model to describe how two sets of data are related. What could the data values in each set be?

**Option #1:** Determine the data values by using a graph.

**Option #2:** Determine the data values without using a graph.