

Contents and Information

This document includes:

1. Lesson Plan Description (4 pages)
2. Worksheet 1 for Data Set 1 (3 pages)
3. Data Sets 2-6 (5 pages total)
4. Worksheet 2 (3 pages)
5. Worksheet 3 (1 page)
6. Assignment (1 page)

NOTE: Worksheet 1 includes calculated values for the data set students are working with, so it is tailored to each data set. Worksheet 1 for Data Sets 2-6 are not included in the document.

CONTACT INFO

If you would like copies for use or have any feedback or questions, please contact me at katri@live.unc.edu.

Demystifying Data

An Introduction to Data and Statistics

Grade Level: 6th Grade

Class Size: 18-30 students

Setting: Classroom

Materials: Pencil, paper, and calculator

Length: four 20 minute sections

Prerequisites:

Students are expected to be able to

- List numbers in ascending order
- Manipulate numbers using

Some comfort reading a table and interpreting a dot plot is beneficial, but the data is presented in multiple ways, so familiarity with dot plots is not essential

Standards Addressed:

Common Core Math 6th Grade Statistics and Probability

- **6.SP.1** Identify statistical questions
- **6.SP.2** Understand data sets have distributions that can be described by the center, spread, and overall shape
- **6.SP.3** Recognize a measure of center summarizes a data set with a single number
- **6.SP.4** Display numerical data in dot plots
- **6.SP.5** Relate the choice of measures of center to the shape of the data and the context in which they were gathered

Learning Objectives:

By the end of the lesson, students will be able to:

- Define “statistical question”
- Name and define three ways of finding the center of a data set
- List center and spread as two important characteristics of a data set
- Explain why there is no one best definition of center
- Give original examples of statistical questions
- Explain why both center and spread are necessary to meaningfully describe a data set
- Collect and record data
- Calculate mean, median, and mode
- Justify their choice of measure of center based on the spread and context of a data set

ENGAGE

Begin the lesson with a group discussion. After the discussion, students are expected to

- Know the definition of statistical question and be able to give examples
- Be able to provide examples of situations where a statistical question would be asked
- Define data collection as the process of asking a statistical question to a group of people and recording the responses

Most of the discussion should be led by the students, so they can reason on their own why statistical questions might be useful and how to go about answering them.

1. Assess background knowledge and current assumptions. Students likely have heard terms like statistics, data, and average before and may have some intuitive idea of what those things mean. Example questions to ask the class:
“What (if anything) do the words ‘statistics’ and ‘data’ mean to you?”
“How many people have heard of Statistics? What are people talking about when they use that word?”
2. Define a statistical question as one where a variety of answers are expected. Give a couple of examples of non-statistical questions and statistical questions. Ask the class for more examples, and create a list on the board.
3. Think/Pair/Share: Choose 2 or 3 of the statistical questions on the board. Give an example of why someone would want to ask that question. How would you go about getting the information needed in each of the examples?
Note: Pick some examples the class provides to use later. You can either record some on the board, or simply make notes for yourself of examples you’d like to return to.
4. Activity: Model the data collection process. You or a student will likely have suggested a statistical question about age. Go through collecting the data of the ages of students in your class. Record it in a chart on the blackboard then create a dot plot of the answers.

EXPLORE

Divide students into 6 approximately equal groups, preferably of 3-5 people. Give students Part I of the attached worksheet. All members of one group are given the same data set, but each group will receive different data sets.

These numbers aren’t set in stone, but decreasing the number of groups will affect the next part of the lesson plan. The data sets are designed to have different amounts of spread, and students will be asked to think about how spread seems to affect measures of center.

Worksheet 1 gets students to use their own intuition to think about how they would choose a “center” of a data set. Although it may not be too hard to come up with a choice of center that

feels reasonable, there's not one clear right answer. This will help to motivate why we talk about mean, median, and mode rather than choosing one method of measuring central tendency.

After students complete Worksheet 1, review the definitions of mean, median, and mode that were given in the worksheet with the class.

EXPLAIN

Students form new groups. Each of the new group should have at least 1 member from each group formed during the Explore activity. Give students Worksheet 2. All students receive the same Worksheet 2.

Students share with their groups how they dealt with different data sets and record the results of different groups. The last question asks them to identify patterns relating measures of center to the spread of a data set.

Follow up worksheet 2 by having students share their thoughts with the class. This will help solidify the concepts in Part II by having students explain their ideas and making sure everyone sees that center is easier to define for distributions without a lot of spread.

Patterns they should notice in worksheet 2:

- All data sets had a mean of 6
- The data sets that had less spread had mean, median, and mode closer together
- Group members came to a consensus more easily for data sets with less spread
- The mean was generally a reasonable choice
- The mode was most frequently not a good choice

EXPAND

Students apply their knowledge by collecting and analyzing their own data set. Have students complete Worksheet 3. To save time you may want to have students collect data in class and finish the worksheet at home.

Guide students through discussion questions to get them thinking more deeply about data collection and analysis.

- Did anyone think they're data set could be useful? Why or why not?
- What if the same question is asked to a different group of people? Could that data set be useful to someone?
- Revisit some examples given during the Engage activity. Which, if any, of the mean, median, or mode would they be interested in? What other information would be useful to the people asking the statistical question? How easy is that information to calculate?

Lesson Plan Description: 4/4

- Establishing some sort of center is often useful, but is it enough? What questions might you have if all you knew about a data set is its mean?
- The worksheets touched on the idea of spread. Any thoughts on how you might try to measure it? (I'd guess people would come up with range, but that may be the only suggestion. Which is fine – let them know they'll be learning more about it soon!)

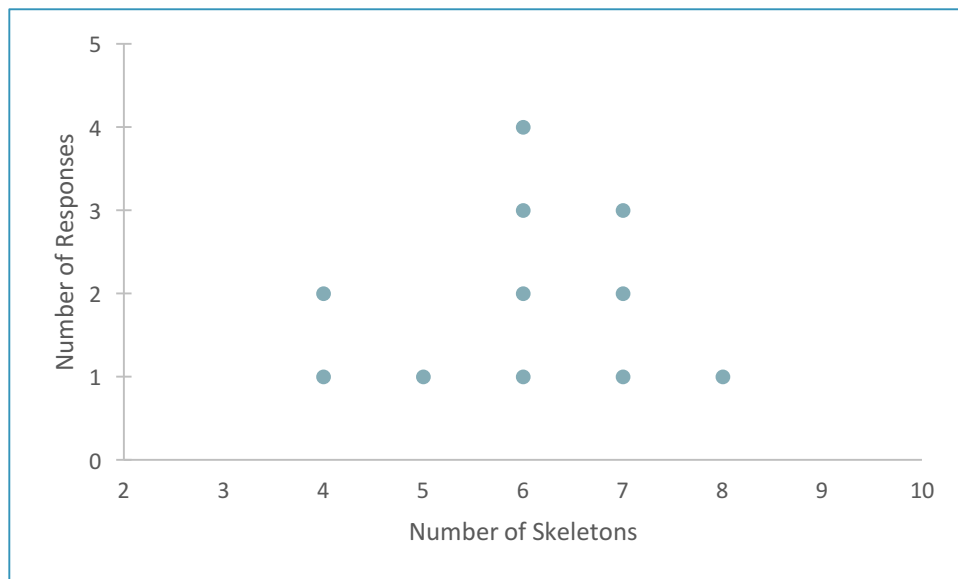
EVALUATE

Collect the completed worksheets, leaving worksheet 3 for homework if you wish. There is also an assignment in the lesson plan that asks students to identify real world statistical questions and think about the types of information about the resulting data sets that would be useful.

Data Set 1

You and your friends are planning a Halloween party. You're in charge of the decorations, but you can't decide how many spooky skeletons to have. You want everyone to be happy with your choice, so you ask your friends how many spooky skeletons they want. Their answers are shown below both in a table and in a dot plot:

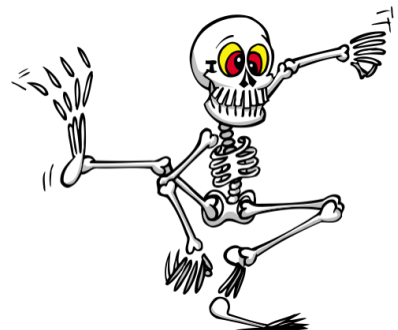
Preference for Number of Skeletons	Number of Responses
3	0
4	2
5	1
6	4
7	3
8	1
9	0
10	0



Demystifying Data

Worksheet 1

1. Your group was given a data set. The data shows the answers to a specific statistical question. What is the statistical question?
2. On your own, look at the responses and decide how many skeletons to get. **Justify your answer based on the data.** Write down your choice and justification. (There is no right or wrong answer. The important thing is to explain your choice based on the responses.)
3. In your group, take turns sharing your answers. Discuss your reasoning. List your group members' choices.
4. As a group decide how many skeletons you will have. Write down your group's final choice and explain your decision.



You're having trouble making a decision. Luckily your older sister who knows statistics offers to help out. She says what you're looking for is a number at the center of the data.

One way to measure the center of a data set is the **mode**. The mode is the most commonly given answer. In the data set above, the mode is 6 skeletons.

5. If you use the mode, how many skeletons will you get? Do you think this is a good choice? Why or why not?

Another way to measure the center of a data set is the **median**. The median is the answer in the middle of the data when it's listed in ascending order. In the data set above, we find the list

4,4,5,6,6,6,6,7,7,7,8

so the median is 6, since it's in the middle of the list.

6. If you use the median, how many skeletons will you get? Do you think this is a good choice? Why or why not?

The third way to measure the center of a data set is the **mean**. The mean is calculated by adding up all the responses and dividing by the number of people who answered. In the data set above, we find the sum of the responses is:

$$4+4+5+6+6+6+6+7+7+7+8 = 66$$

There were 11 people who answered the question, so we find that the mean is $66/11 = 6$

7. If you use the mean, how many skeletons will you get? Do you think this is a good choice? Why or why not?

8. Did your choice match up with the median, mode, or mean for the data set? If so, which one(s) were the same as your choice?

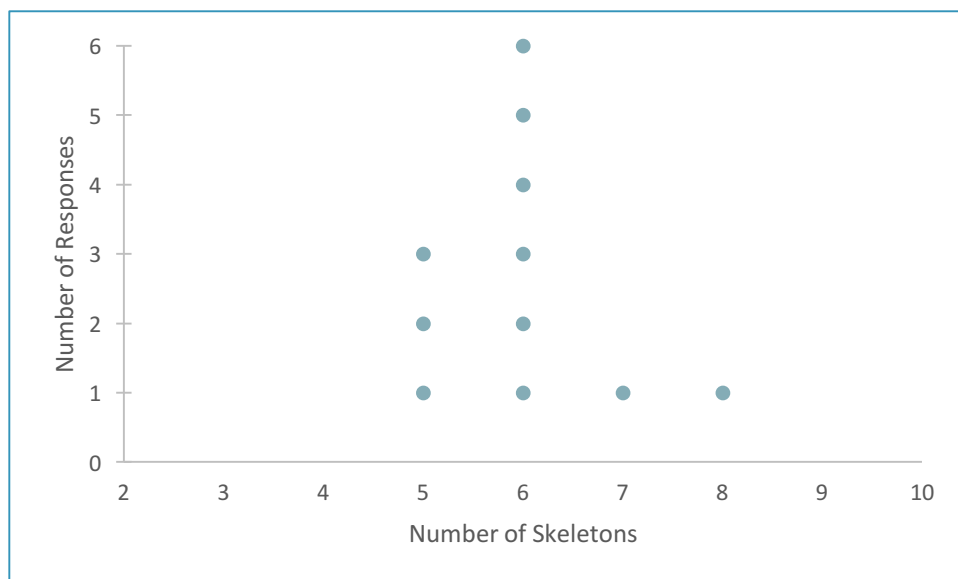
Data Sets 2-6

(corresponding Worksheet 1 omitted contact katri@live.unc.edu for Worksheet 1)

Data Set 2

You and your friends are planning a Halloween party. You're in charge of the decorations, but you can't decide how many spooky skeletons to have. You want everyone to be happy with your choice, so you ask your friends how many spooky skeletons they want. Their answers are shown below both in a table and in a dot plot:

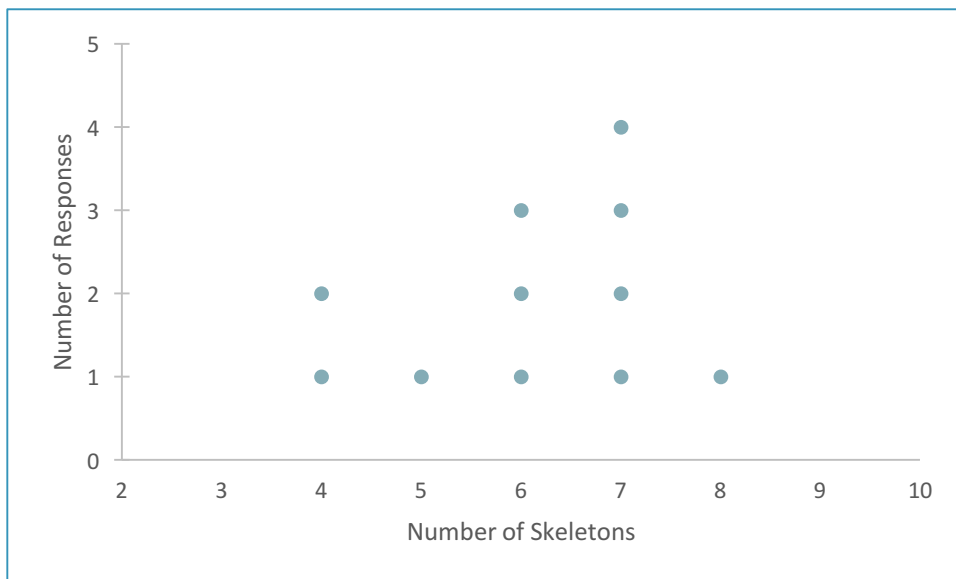
Preference for Number of Skeletons	Number of Responses
3	0
4	0
5	3
6	6
7	1
8	1
9	0
10	0



Data Set 3

You and your friends are planning a Halloween party. You're in charge of the decorations, but you can't decide how many spooky skeletons to have. You want everyone to be happy with your choice, so you ask your friends how many spooky skeletons they want. Their answers are shown below both in a table and in a dot plot:

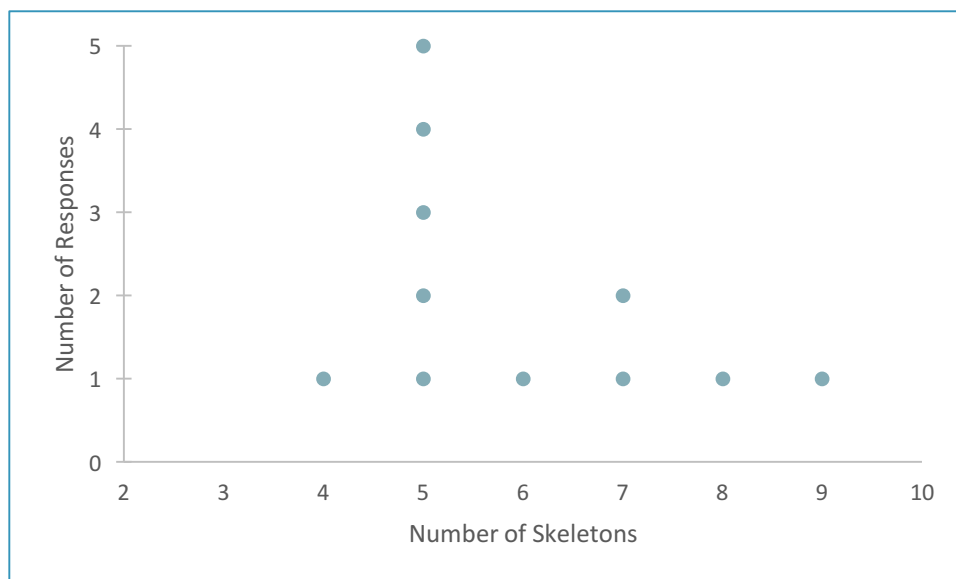
Preference for Number of Skeletons	Number of Responses
3	0
4	2
5	1
6	3
7	4
8	1
9	0
10	0



Data Set 4

You and your friends are planning a Halloween party. You're in charge of the decorations, but you can't decide how many spooky skeletons to have. You want everyone to be happy with your choice, so you ask your friends how many spooky skeletons they want. Their answers are shown below both in a table and in a dot plot:

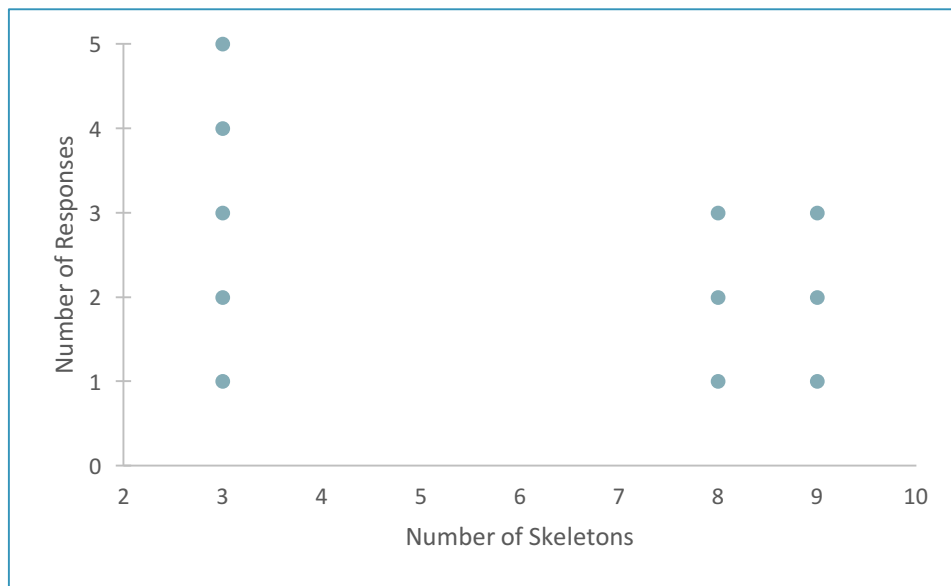
Preference for Number of Skeletons	Number of Responses
3	0
4	1
5	5
6	1
7	2
8	1
9	1
10	0



Data Set 5

You and your friends are planning a Halloween party. You're in charge of the decorations, but you can't decide how many spooky skeletons to have. You want everyone to be happy with your choice, so you ask your friends how many spooky skeletons they want. Their answers are shown below both in a table and in a dot plot:

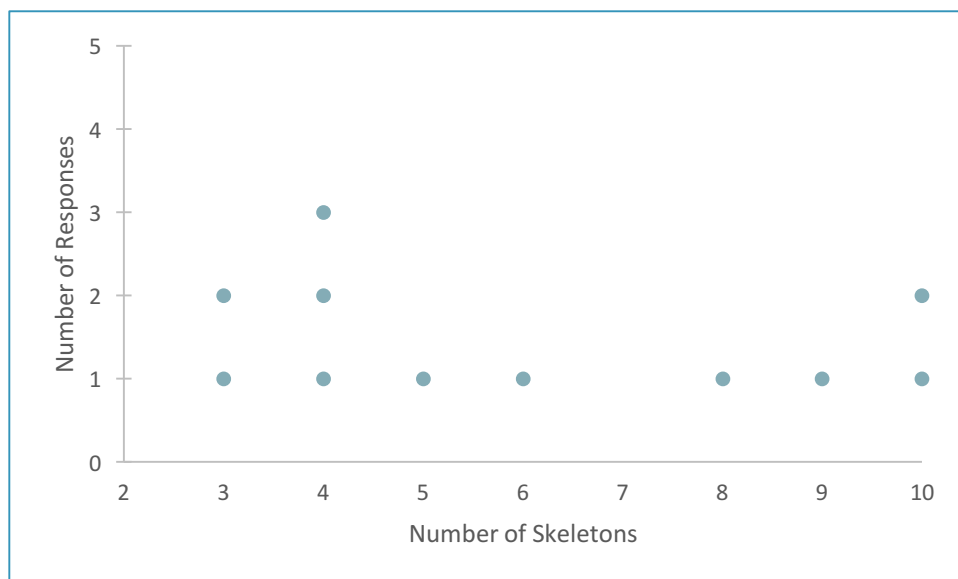
Preference for Number of Skeletons	Number of Responses
3	5
4	0
5	0
6	0
7	0
8	3
9	3
10	0



Data Set 6

You and your friends are planning a Halloween party. You're in charge of the decorations, but you can't decide how many spooky skeletons to have. You want everyone to be happy with your choice, so you ask your friends how many spooky skeletons they want. Their answers are shown below both in a table and in a dot plot:

Preference for Number of Skeletons	Number of Responses
3	2
4	3
5	1
6	1
7	0
8	1
9	1
10	2



Demystifying Data

Worksheet 2

1. Look at the data sets the other groups were given. For each set think about whether people seem to have similar opinions or if there's a lot of variety in the preferences.
 - a. Which three data sets have the most variety in the responses?
 - b. Which three data sets have the least variety in the responses?
2. Take turns sharing your group's decision on Worksheet 1 and explaining your reasoning. Record the requested information about each of the groups' data sets.

DATA SET 1

Group choice:	
Mode:	
Median:	
Mean:	

1. How many different decisions did the group members come up with on their own?
2. Which of the mode, median, and mean gave a good choice?
3. Did you say this set had low, medium, or high variety?

DATA SET 2

Group choice:	
Mode:	
Median:	
Mean:	

1. How many different decisions did the group members come up with on their own?
2. Which of the mode, median, and mean gave a good choice?
3. Did you say this set had low, medium, or high variety?

DATA SET 3

Group choice:	
Mode:	
Median:	
Mean:	

1. How many different decisions did the group members come up with on their own?
2. Which of the mode, median, and mean gave a good choice?
3. Did you say this set had low, medium, or high variety?

DATA SET 4

Group choice:	
Mode:	
Median:	
Mean:	

1. How many different decisions did the group members come up with on their own?
2. Which of the mode, median, and mean gave a good choice?
3. Did you say this set had low, medium, or high variety?

DATA SET 5

Group choice:	
Mode:	
Median:	
Mean:	

1. How many different decisions did the group members come up with on their own?
2. Which of the mode, median, and mean gave a good choice?
3. Did you say this set had low, medium, or high variety?

DATA SET 6

Group choice:	
Mode:	
Median:	
Mean:	

1. How many different decisions did the group members come up with on their own?
2. Which of the mode, median, and mean gave a good choice?
3. Did you say this set had low, medium, or high variety?

3. What patterns do you see in the information above? Does the amount of variety in responses affect measures of center?

Demystifying Data

Worksheet 3

Now that you know the basics of dealing with data, collect your own!

1. Think about what you'd like to know about your class. Come up with a statistical question to ask some of your classmates. Be sure to choose a numerical question (one where the expected answers are numbers). What question will you use?
2. Ask your question to 10 students in your class. Record their responses below:
3. What is the mode of your data?
4. What is the median of your data?
5. What is the mean of your data?
6. Describe a scenario where someone might find your statistical question useful. (Either using the data you collected, or collecting their own data by asking your statistical question to a different group of people).
7. In the scenario you described, are the mean, median, or mode be helpful pieces of information? If so, which one do you think would be the most useful?

Demystifying Data

Assignment

Find a real world example of a statistical question being used to collect and analyze data. The news is a good source for examples, but see if you can find other places statistics may come up.

1. Briefly describe the example you found.
2. What is the statistical question being asked?
3. Who was asked the question?
4. What was the purpose of collecting the data? How was the data used?
5. What information about the data set do you think would be most useful for the person or group asking the question?
6. Would any of the measures of center you learned in class be useful? If so, which one do you think would be most useful and why? If not, explain why not.