Identification of Misleading Statistics

Jen Kershaw

Say Thanks to the Authors
Click http://www.ck12.org/saythanks
(No sign in required)
To access a customizable version of this book, as well as other interactive content, visit www.ck12.org

CK-12 Foundation is a non-profit organization with a mission to reduce the cost of textbook materials for the K-12 market both in the U.S. and worldwide. Using an open-source, collaborative, and web-based compilation model, CK-12 pioneers and promotes the creation and distribution of high-quality, adaptive online textbooks that can be mixed, modified and printed (i.e., the FlexBook® textbooks).

Copyright © 2015 CK-12 Foundation, www.ck12.org

The names “CK-12” and “CK12” and associated logos and the terms “FlexBook®” and “FlexBook Platform®” (collectively “CK-12 Marks”) are trademarks and service marks of CK-12 Foundation and are protected by federal, state, and international laws.

Any form of reproduction of this book in any format or medium, in whole or in sections must include the referral attribution link http://www.ck12.org/saythanks (placed in a visible location) in addition to the following terms.

Except as otherwise noted, all CK-12 Content (including CK-12 Curriculum Material) is made available to Users in accordance with the Creative Commons Attribution-Non-Commercial 3.0 Unported (CC BY-NC 3.0) License (http://creativecommons.org/licenses/by-nc/3.0/), as amended and updated by Creative Commons from time to time (the “CC License”), which is incorporated herein by this reference.

Complete terms can be found at http://www.ck12.org/about/terms-of-use.

Printed: December 2, 2015
Chapter 1: Identification of Misleading Statistics

Here you’ll learn to identify and analyze misleading statistics.

Remember the Iditarod? Take a look at this dilemma.

Hayley is investigating the prize money that mushers have earned in the Iditarod. In completing her research, she finds that the prize money for first place hasn’t changed very much in the past few years and hovers around $69,000. The overall prize money has changed significantly though.

She notices that there was a big year for race money in 2008 with mushers earning a total of $935,000. She shows this to her friend Tiffany.

“That is a lot of money,” Tiffany comments.

“Yes, but they all don’t get the same amount. The first place winner receives about the same amount, but then the other mushers get less based on where they finish. There is less money being given out though this year than last.”

Hayley creates this line graph of the data that she has discovered.

When finished, she shows it to Tiffany.

“Here is the graph that I created,” Hayley says to Tiffany.

“This won’t work. It isn’t accurate.”

“What do you mean?”

Hayley is puzzled by Tiffany’s comment. Do you see what is misleading in Hayley’s graph? Creating graphs that accurately display data can be a tricky thing. Hayley thought that her graph was accurate, but Tiffany disagrees.

This Concept is all about misleading graphs and data. By the end of it, you will know why Tiffany made the comment about Hayley’s graph.
Guidance

Graphs provide a visual picture of data. Graphs can be used to present, persuade, or even mislead the viewer. The same set of data can be presented on a graph in different ways. Sometimes, the way that a graph is drawn can present only one side of the statistics.

Why would this happen?

Sometimes, companies try to sell more of a specific product by creating misleading graphs. These graphs can be divided to make it look like a product is more popular than another product.

Companies can also do this with purchases by showing that a percentage of people like their product better than another one. However, if you look at how the graph is divided and created, it will give you an idea whether or not the statistics represented give a good overall picture of the data.

There are several ways that data can be misrepresented. By the end of this lesson, you will know how to identify misleading data.

The first way is to show a break in the vertical axis.

A break in the scale on the vertical axis, will show more detail or emphasize an increase or decrease in data values. A break in the scale means that the vertical axis does not start at zero. If data does not start at zero it is difficult to determine where the data actually begins and where it ends.

Using a different scale or spacing along the horizontal or vertical axis will also change the appearance of the data. You may choose to space the values closely together to depict a greater change in the data. Or, you may choose to spread the values out to depict less of a change in the data. If you can’t detect a change in the data, then it is misleading.

Let’s look at two situations where this is the case.

A survey was taken to determine people’s favorite fruit. Both graphs depict the results of the survey. Which graph provides a more reliable view of the data?

The vertical axis scale on Graph 1 is smaller than the scale on Graph 2. The vertical axis scale on Graph 1 is twenty-five. The vertical axis scale on Graph 2 is hundreds. Because the scale is smaller, detecting changes in the data is easier on Graph 1 than on Graph 2. Therefore, Graph 1 depicts the data in a more reliable manner because it shows more of the data.
Both graphs below depict the average monthly water temperature in Hawaii. Which graph provides the most reliable view of the data?

Graph 2 shows a sharp increase in the change in water temperature. The change in temperature does not vary greatly from month to month and therefore graph 2 is not a reliable display. **Graph 1 reveals the change in water temperature as more gradual and is therefore more reliable.**

Now that you have an idea of a couple of ways that data can be misrepresented, let’s look at identifying why data is misleading.

**Remember, you are looking for a couple of things.**

1. **Does the data show an accurate change from one piece of information to the next?**
2. **Does the data start at zero or is there a break in the vertical axis?**
3. **Is the scale that is being used one that makes sense? Are the sections even or spread too far out?**
Write these three things down in your notebooks under “Data Display Checking.”
Sometimes, you can find something else on a graph too. You can find data that isn’t connected. When you look for the above three things on a graph, also be sure that the things that are being compared are similar. Explain why the data compared on the bar graph is misleading.

While the incidence rates of shark attacks are correct, one must take into consideration several things when analyzing the data. First, each state on the graph has a different size coastline. For example, Florida’s vast amount of coastline may contribute to the fact that it has a far higher incidence of shark attacks than any other U.S. state. As well, California, Florida, and Hawaii are all big beach destination states; therefore the incidence of shark attacks will be greater than states such as Alabama. To make the data more reliable, a graph should be created comparing the incidence of shark attacks among states that are more geographically similar.

Think about misleading data to answer each of the following questions.

Example A
Why would data about ocean swimmers in Michigan be misleading?
Solution: Michigan does not border on an ocean.

Example B
Why would data about temperatures below 32° comparing Georgia and Tennessee with North Dakota be misleading?
Solution: Because the climate in North Dakota is much colder than Georgia and Tennessee.

Example C
Why would a graph whose vertical axis starts at 50 be misleading?
Solution: Because it would look like the data started there and not at zero.

Here is the original problem once again. Reread it and then think about why Hayley’s graph is misleading. Read the comments that Tiffany makes at the end of the problem.
Hayley is investigating the prize money that mushers have earned in the Iditarod. In completing her research, she finds that the prize money for first place hasn’t changed very much in the past few years and hovers around $69,000. The overall prize money has changed significantly though.

She notices that there was a big year for race money in 2008 with mushers earning a total of $935,000. She shows this to her friend Tiffany.

“That is a lot of money,” Tiffany comments.

“Yes, but they all don’t get the same amount. The first place winner receives about the same amount, but then the other mushers get less based on where they finish. There is less money being given out though this year than last.”

Hayley creates this line graph of the data that she has discovered.

When finished, she shows it to Tiffany.

“Here is the graph that I created,” Hayley says to Tiffany.

“This won’t work. It isn’t accurate.”

“What do you mean?”

**Tiffany looks at the graph over Hayley’s shoulder.**

“There are two things wrong with this graph. The first one is that your spacing isn’t even. That means that the money amounts aren’t clearly represented. It makes it look like the jumps are more drastic than they are.”

“The second thing is that the graph doesn’t start at 0. Because it doesn’t start at zero, you can’t get a good idea of how the money has changed over time.”

Hayley looked at her graph again. With Tiffany’s help, she now has an idea how to fix her graph.

Now it is your turn. Look at Hayley’s graph and use Tiffany’s comments to revise the graph. Be sure the graph is accurate and not misleading.

**Guided Practice**

Here is one for you to try on your own.

Both graphs below depict the population of five animals on the endangered species list. Explain why Graph 1 is misleading and Graph 2 is more reliable.
Answer

Graph 1 has a break in the scale on the vertical axis. You can see that the scale begins at 1,000. This provides a distorted and misleading view of the data.

Video Review

This is a video on statistics. It is a supporting video for this Concept.

Explore More

Directions: Answer each question regarding misleading data.
1. Is this a misleading graph?
2. What is one thing that makes it a misleading graph?
3. What is one thing that you could do to fix this graph?

The data table below depicts the amount of time students at different grade levels spend on homework and studying. Ensure that the second graph shows that time spent on homework in twelfth grade is triple that of sixth grade.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>6th</td>
<td>1.75</td>
</tr>
<tr>
<td>7th</td>
<td>2</td>
</tr>
<tr>
<td>8th</td>
<td>2.25</td>
</tr>
<tr>
<td>9th</td>
<td>2.5</td>
</tr>
<tr>
<td>10th</td>
<td>2.75</td>
</tr>
<tr>
<td>11th</td>
<td>3</td>
</tr>
<tr>
<td>12th</td>
<td>3.5</td>
</tr>
</tbody>
</table>

4. Use the data below to create two bar graphs. One that shows the data accurately, that time spent on homework in twelfth grade is double that of sixth grade.

5. Ensure that the second graph shows that time spent on homework in twelfth grade is triple that of sixth grade.

6. If the students doubled the time that they spend on homework in the 7th grade, how many hours would they be spending?

7. If the students in the 11th grade spent half as much time on homework, how many more hours of free time would they gain?

8. True or false. All students spend at least one hour on homework.

The data table below depicts the sales tax rate for several U.S. states.

<table>
<thead>
<tr>
<th>State</th>
<th>Sales Tax Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alaska</td>
<td>0</td>
</tr>
<tr>
<td>Alabama</td>
<td>4.0</td>
</tr>
<tr>
<td>Arizona</td>
<td>5.6</td>
</tr>
<tr>
<td>California</td>
<td>6.25</td>
</tr>
<tr>
<td>New Jersey</td>
<td>7.0</td>
</tr>
</tbody>
</table>
9. Use the information on the data table to create two graphs. One graph should depict the data accurately. On this graph, the sales tax rate for New Jersey is almost double the sales tax rate in Alabama.

10. The second graph should present the data in a misleading manner to suggest that the sales tax rate in New Jersey is more than triple the tax rate in Alabama.

11. Which state has the highest state tax?

12. If Alaska doesn’t have a state tax, does it make sense to put it on the list?

The data below depicts the daily temperature in Juneau, Alaska for ten days.

<table>
<thead>
<tr>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
</tr>
<tr>
<td>60</td>
</tr>
<tr>
<td>57</td>
</tr>
<tr>
<td>55</td>
</tr>
<tr>
<td>49</td>
</tr>
<tr>
<td>57</td>
</tr>
<tr>
<td>58</td>
</tr>
<tr>
<td>60</td>
</tr>
<tr>
<td>59</td>
</tr>
<tr>
<td>56</td>
</tr>
</tbody>
</table>

13. Draw a line graph that depicts a sharp decrease in temperature.

14. Draw another line graph that depicts the decrease accurately.

15. What is the highest temperature on the list?

16. What is the lowest temperature on the list?

17. –20. Look through a newspaper and choose three different graphs. Then write a few sentences about each one explaining how the data represented is or is not misleading and why.