

Lesson Title: Heat Wave

Essential Questions:

What tools can I use to describe the shape of a set of data beyond finding its center?
How can I compare sets of quantitative data to each other?

Lesson Rationale:

This is the culminating lesson of the second part of a unit on data. In the first part, students explored methods of displaying qualitative data. In the second part, they work with quantitative data. This lesson follows several lessons on measures of center as well as some informal discussion of how data sets with the same center can look very different. The main focus of this lesson is on box-and-whisker plots as a way of displaying and comparing the centers and spreads of sets of quantitative data.

The context for this lesson is climate, specifically the number of heat wave days in Massachusetts in the years between 1981 and 2010 as recorded by the Centers for Disease Control. This connects back to a previous lesson on weather and climate in which students explored graphs that show two kinds of data at the same time. As students are working toward careers in health care, they will likely encounter box-and-whisker plots in future classes and possibly need to read them in their careers as well.

In a broader sense, the enduring understandings that are the goals of this lesson are:

- 1) As consumers of data, students should look beyond reported measures of center and ask about the shape of the data distribution. This information is often not available in the real world, but students should be aware of what they are missing and what kinds of variability can exist in data sets that have the same or similar centers.
- 2) It is possible to use measures of center and spread to compare multiple sets of quantitative data.

CCRS Standards:

[6.SP.4]: Display numerical data in plots on a number line, and box plots.

[6.SP.5]: Summarize numerical data sets in relation to their context, such as by:

- a. Reporting the number of observations.
- b. Describing the nature of the attribute under investigation, including how it was measured and its units of measurement.
- c. Giving quantitative measures of center (median and/or mean) and variability (interquartile range), as well as describing any overall pattern and any striking deviations from the overall pattern with reference to the context in which the data were gathered.

d. Relating the choice of measures of center and variability to the shape of the data distribution and the context in which the data were gathered.

Academic Skills: <i>(Use DOK verbs)</i>	Sector-Focused Content: <i>(What will students know?)</i>
<p>Organize and synthesize data</p> <p>Reason abstractly and quantitatively (MP.2)</p> <p>Construct viable arguments and critique the reasoning of others (MP.3)</p>	<p>Students will learn about trends in climate change, specifically the increasing number of heat wave days in Massachusetts counties in the years 1981-2010.</p>
Professional/Affective Skills: <i>(Students will be able to)</i>	Academic Content: <i>(not directly career related)</i>
<p>Collaboration, attention to detail, patience, perseverance</p>	<ul style="list-style-type: none"> • Define and find quartiles • Create box-and-whisker plots • Describe the center and spread of data sets • Compare data sets by comparing box-and-whisker plots
Evidence of Learning:	Assessment:
<p><i>What will students produce or do that shows their learning?</i></p> <p>Students will produce at least two box and whisker charts that display data they accessed at wonder.cdc.gov.</p> <p>Students will write a paragraph using what they have learned to convince a climate change skeptic.</p> <p>Students will complete a multiple-choice test-style worksheet.</p> <p>Students will reflect on new learning for homework.</p>	<p><i>What tools will you use to assess student learning?</i></p> <p>See: Heat Wave – Looking Closely for a list of specific things to be looking for to assess student learning.</p>

Materials & Resources:

Computers with internet access, masking tape, 1993 data on cards, [What A Year](#), [Climate Change: Making an Argument](#), [Test Practice](#), [Reflection Questions](#)

Background information: *(What previous learning does this connect to?)*

This lesson connects to the previous lesson in which students learned how to find the median of a data set. It extends the idea of median to quartiles. This lesson also continues the themes of graphical displays of quantitative data and of climate data began in the lesson on Mystery Cities (EMPower).

STEPS**Opening Discussion:**

Ask: Have you heard of the CDC? Do you know what it is?

Students use chromebooks to access wonder.cdc.gov. What do you notice? What do you wonder? Students explore some of the different kinds of data available. Some of them are very complicated. Some we can figure out how to get the information. Find something interesting and share it with the class. [Purpose is to expose students to a data and digital resource and to understand how to submit a form on a website.] (20 min)

We talked about the difference between climate and weather and how climate shows trends over a much greater period of time. I tried to find some data about the climate in Massachusetts. Where do you think I might have looked? [I looked at Heat-Wave Days May-September.]

Briefly review this definition of a “heat wave day.” The purpose of defining it rigorously is to show that we need to have terms clearly defined before talking about data. It is not actually important that students fully understand or remember this complicated definition:

Relatively simply, the North America Land Data Assimilation System (NLDAS) looked at the maximum temperatures in an area in May through September in the 30 years from 1981 to 2010. Any time the temperature went above 95% of those temperatures for two or more days in a row, each of the days was considered a “heat wave day.” [Note: What counts as a heat wave in one place may not in another because it is based on how the temperature compares to temperatures in that particular place.]

Ask students to find the number of heat wave days in Essex County, MA and in St. Clair County, IL in 2006 [Essex: 5; St. Clair: 24]

Activity 1: The Middle of the Halves

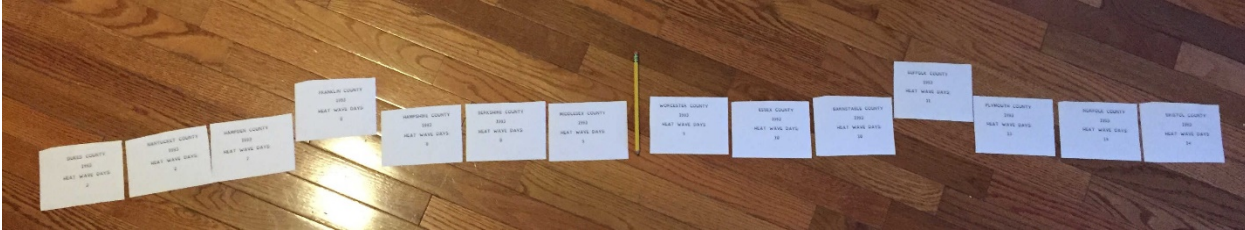
Hand students cards with one data point on each one showing the number of heat wave days in 1993 in the 14 Massachusetts counties – one set of cards for the whole class. Ask students to organize the data from fewest to most in one long line (don’t stack up numbers that are the same).

Have students identify the median of the set. [The median is 9 with 7 counties on either side of it. The median falls between two counties that both had 9 heat wave days.] Since the median falls between two data points, use a pencil or ruler or some other marker to mark where it is.

Now turn your attention to the “bottom half” of the data – those seven counties that are to the left of the median line – and ask the students to find the median of that set. This is called the **first quartile or Q1**. Note that the first quartile is the number that is the median of the bottom half of the

data, not the county that had that number of heat wave days. For this data, the first quartile is 8. Slide the card that is at the first quartile slightly up to mark it. (Three counties had 8 heat wave days, but only one of them is marking the first quartile – the fourth card in.)

Do the same with the top half of the data to locate the **third quartile or Q3**. Slide the card that marks it [Suffolk County – 11 heat wave days] up slightly.



Ask the students to make observations about what they see. The following questions may be helpful:

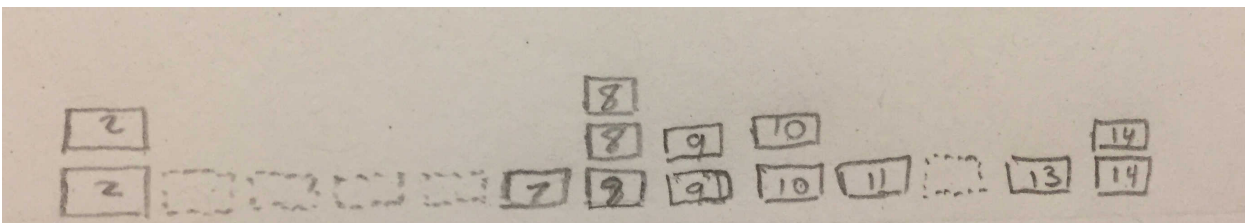
- Why do you think the quartiles are called quartiles? [They divide the data into quarters.]
- Why do you think we didn't identify a second quartile? [The median is the second quartile.]

Make sure the students recognize that quartiles, including the median, divide the data set into four equal parts. The cards that mark the first and third quartiles are not part of the equal parts just as the median is not part of the two equal halves of the data when there are an odd number of data points (established in a prior lesson).

Make note on the board of the **five number summary** of the data:

Minimum: 2
First Quartile: 8
Median: 9
Third Quartile: 11
Maximum: 14

Now suggest that students stack cards that have the same value so that all are visible and space them out as if they were on a number line. (This is like making a frequency graph with quantitative data (also called a dot plot when using dots instead of data cards.... Now the distance between the points means something.) Students may use extra cards to get the spacing right.



Start to build a number line by laying a piece of masking tape from the beginning to the end of the data.

Use masking tape to place tick marks at the minimum and maximum data points (the masking tape should not extend beyond these points).

Ask a student to place a masking tape tick mark at the median of the data set [9].

Ask: Is the median of the data in the middle of our number line? [No. The number line goes from 2 to 14, so the middle would be at 8.]

Ask: Can you explain how the median represents a middle value even though it's not in the middle of the number line? [It's in the middle of the data. Half the data is equal to or greater than the median and half is equal to or less than the median.]

Ask a student to place a masking tape tick mark at the first quartile [8].

Revisit the question of why the first quartile divides the bottom half of the data into two equal parts but does not necessarily land in the middle of the interval, just like we saw with the median.

Do the same with the third quartile [11] as with the first.

Connect the 1st and 3rd quartiles with strips of masking tape above and below the center line to create a box. This is a **box and whisker plot**.

Key Points:

- Quartiles are found by cutting the data in half and then in half again, just like how we found *quarters* when we did pie charts at the beginning.
- When making a box and whisker plot, the data should be spaced out as it would be on a number line.

Discussion: What Does It Show?

Show 1993 box plot PowerPoint which shows how all the parts of the box and whisker plot from the activity come together. On the first slide, which just has the plot, ask each student to write down three observations about the graph. Ask each student to share at least one observation. Address the following questions in discussion:

Suppose this graph is what we started with. What would we be able to tell about what happened in 1993?

What does the graph show us? What information is lost? [We can see the center (median) and range of the data. We cannot see how many data points there are.]

What do you think is the significance of the box? What percent of the data is in the box? [50% or half of the data is in the box. The box shows us the middle half of the data.]

What do you think is the significance of the whiskers? What do they show you about the data set? [They give a good picture of the range and show where the maximum and minimum values are.]

How do the quartiles divide up the data? [The quartiles and the median (which is also called the second quartile) divide the data into fourths. Connections can be made here with the fraction equivalence $\frac{1}{2} = \frac{2}{4}$.]

Key Points:

- The box represents the middle 50% of the data. The box and whisker plot gives us more information about the shape of the data set because we can see the center, the range, and where the middle half of the data sit.
- Box and whisker plots are useful for picturing the shape of large data sets. With small sets, we can use frequency graphs or bar graphs, but with large sets, those become cumbersome.

Activity 2: What a Year!

Hand out What a Year! In pairs, students choose a year and look up the data for Massachusetts for that year. They review the process and vocabulary involved in creating a box and whisker plot as they create their own and reflect on what it shows them. They also compare their graph to another pair's and make observations about the comparison. This sheet also introduces students to the term "interquartile range." Students are invited to speculate about what the inter-quartile range tells them about the data. (If students want to look up the data for a different state, they can do that after doing one for Massachusetts because they need to be able to compare their graph to another pair's.) [Note: 1992 had *zero* heat wave days in MA! If students choose this year, allow them to discover this on their own, make note of this interesting data, and ask them to choose another year to make their plot.]

When debriefing the exercise, elicit/make the following points about interquartile range:

- The IQR is called a measure of variability. The larger the IQR, the more spread out the data are. The smaller the IQR, the more consistent the data are. Get at this point by asking students to compare data sets with different IQRs.
- The IQR is a better measure than the range of how spread out the data are because one or two extreme points can have a big effect on the range and make the data seem more spread out. If there is a small IQR and long whiskers, the data points at the end may be considered **outliers**.

- The IQR is related to the median. There is another measure of variability that is tied to the mean. It is called “Standard Deviation.” Standard deviation is beyond the scope of this course, but you may see data reported with a mean and standard deviation and you should know that it is similar to the IQR in that the larger it is, the more spread out the data are.

Activity 3: Is There a Trend?

Students cut out their box and whisker plots from the previous activity and organize them by year. Line up the plots so that the number lines align and so there is an appropriate distance between years (i.e. a greater distance between 1985 and 1995 than between 1985 and 1987). Depending on the years they chose, a trend may or may not be visible.

Ask: Do we see evidence of climate change in this data? What arguments might be made that this data shows evidence of climate change? [If a trend is visible, it appears to.] What arguments might be made that this data does not show evidence of climate change? [Even if a trend does appear, we aren't looking at very many years, so maybe it's a coincidence. It could be due to natural variability in weather.]

Ask: What would we need to do to satisfy ourselves that there is or is not a trend in increasing numbers of heat wave days over the 30 years from 1981 to 2010? [One possibility is that we could draw plots for each one of the years and see if we see a general trend over time. That would take a long time. To get a bigger picture without so much work, we can group the years and look at the combined data.]

Divide up the time into six 5-year spans: 1981-1985, 1986-1990, 1991-1995, 1996-2000, 2001-2005, 2006-2010. Assign a 5-year span to each group of students. If there aren't enough students, each group can do two or groups that finish faster can pick up ones that haven't been done yet. Each group gets the data for their 5-year span and creates a box and whisker chart.

Line up the charts to see that there is a trend of increasing numbers of heat wave days. [The numbers actually go down for 1996-2000, but there is still a general upward trend.] If students have not used the same scale to make their charts, they will need to agree on a scale and redraw all the charts to line up with each other.

Activity 4: Convince a Skeptic

Hand out [Climate Change: Making an Argument](#). Students write a paragraph in which they use data and box plots to counter the argument of a climate change skeptic. Students may discuss the question with each other, but should each write their own paragraph. (This activity could also be a homework assignment.)

Extensions & Content Connections

Possible extensions:

- Exploring outliers with the IQR (using $1.5 \times \text{IQR}$ as the limit beyond which a point is an outlier)
- Showing trends over time with box plots in a graph

Possible connections to health care careers:

- What are the health effects of a heat wave?
- Who is most at risk during a heat wave?
- What should hospital staff be prepared to deal with during a heat wave?

Notes

If teaching this lesson over several class periods, reflection questions can be broken up to assign the relevant ones for homework each night.

This lesson plan template was adapted from *Understanding by Design*.

Wiggins, G. P., McTighe, J., Kiernan, L. J., Frost, F., & Association for Supervision and Curriculum Development. (1998). *Understanding by design*. Alexandria, Va: Association for Supervision and Curriculum Development.

Looking Closely

Observe whether students are able to...

Line up their box plots with a number line:

- Tick marks on the number line are evenly spaced
- No numbers are left out
- A scale may be used

Identify the parts of a box-and-whisker plot

- Identify the five-number-summary from a box plot
- Know that the median and the second quartile are the same
- Recognize that the box represents the “middle half” of the data

Reason about the spread of the data from a box-and-whisker plot

- Calculate both the inter-quartile range and the range from a box plot
- Informally reason about possible outliers by comparing the length of the whiskers to the IQR (i.e. long whiskers and a small IQR indicates possible outliers.)

Compare data sets by comparing their box-and-whisker plots

- Compare medians, minimums, maximums, and IQR's.
- Argue for or against the presence of a trend in looking at box plots for data changing over time.

DUKES COUNTY

1993

HEAT WAVE DAYS:

2

HAMPDEN COUNTY

1993

HEAT WAVE DAYS:

7

NANTUCKET COUNTY

1993

HEAT WAVE DAYS:

2

BERKSHIRE COUNTY

1993

HEAT WAVE DAYS:

8

FRANKLIN COUNTY

1993

HEAT WAVE DAYS:

8

MIDDLESEX COUNTY

1993

HEAT WAVE DAYS:

9

HAMPSHIRE COUNTY

1993

HEAT WAVE DAYS:

8

WORCESTER COUNTY

1993

HEAT WAVE DAYS

9

BARNSTABLE COUNTY

1993

HEAT WAVE DAYS:

10

SUFFOLK COUNTY

1993

HEAT WAVE DAYS:

11

ESSEX COUNTY

1993

HEAT WAVE DAYS:

10

PLYMOUTH COUNTY

1993

HEAT WAVE DAYS:

13

BRISTOL COUNTY

1993

HEAT WAVE DAYS:

14

NORFOLK COUNTY

1993

HEAT WAVE DAYS:

14

What a Year!

Working with a partner, choose a year between 1981 and 2010 and look up the data for heat wave days in Massachusetts at wonder.cdc.gov. Fill in the data below.

Heat Wave Days in Massachusetts Counties in _____(year):

Barnstable County, MA	
Berkshire County, MA	
Bristol County, MA	
Dukes County, MA	
Essex County, MA	
Franklin County, MA	
Hampden County, MA	
Hampshire County, MA	
Middlesex County, MA	
Nantucket County, MA	
Norfolk County, MA	
Plymouth County, MA	
Suffolk County, MA	
Worcester County, MA	

What is the smallest (minimum) number of heat wave days? _____

What is the largest (maximum) number of heat wave days? _____

What was the median number of heat wave days? _____

List in order the “bottom half” of the data: _____

The first quartile (Q1) is the median of the bottom half of the data. What is the first quartile? _____

List in order the “top half” of the data: _____

The third quartile (Q3) is the median of the top half of the data. What is the third quartile? _____

The minimum, Q1, median, Q3, and maximum values form a “Five-Number Summary” of the data. List those values here:

Minimum: _____

First Quartile: _____

Median: _____

Third Quartile: _____

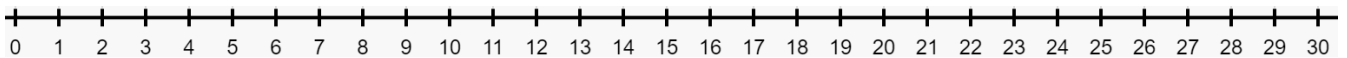
Maximum: _____

What does the five-number summary tell you about the whole data set?

The interquartile range (IQR) is the distance from the first quartile to the third quartile. What is the interquartile range for your data? _____

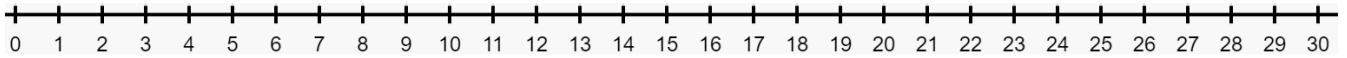
What does the IQR tell you about your data?

Draw a box and whisker plot of your data using the number line below:



Compare your plot to another group’s plot. What do you notice? What do you wonder?

Draw a copy of your box and whisker plot here for the next activity:



Name: _____

Climate Change: Making an Argument

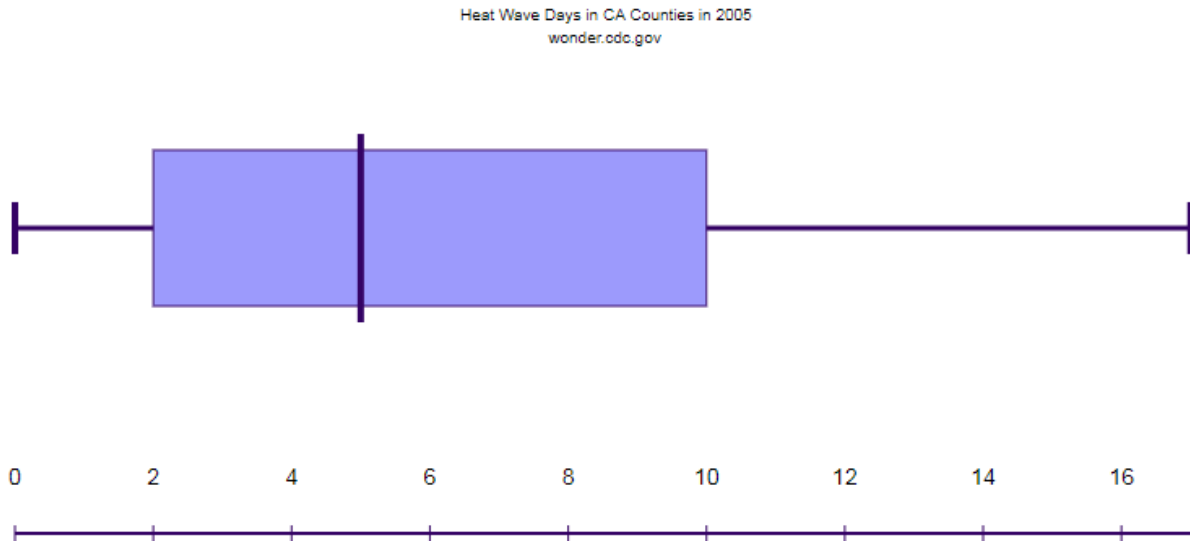
Your friend Livia has been looking at the data from wonder.cdc.gov and presents you with the following argument:

People claim that the planet is warming up, but I've looked at the data and I don't agree. In 1992 there was not a single heat wave day in Massachusetts but in 1991 there were 9 heat wave days just in Bristol County. How can you say the planet is warming up when it was clearly hotter in 1991 than it was in 1992?

Write a paragraph in which you address Livia's points and share your conclusions based on the data. You may want to include:

- data
- a box and whisker plot (or several)
- discussion of median and inter-quartile range.

Name: _____

Test PracticeUse the chart below for questions 1-4:

1) What was the median number of heat wave days in California counties in 2005?

- (1) 0
- (2) 2
- (3) 5
- (4) 10
- (5) 17

2) According to the chart, which of the following statements is not true?

- (1) 50% of California counties had between 2 and 10 heat wave days in 2005.
- (2) 50% of California counties had 5 or more heat wave days in 2005.
- (3) At least one county in California had no heat wave days in 2005.
- (4) The greatest number of heat wave days in a California county in 2005 was 16.
- (5) 25% of California counties had between 2 and 5 heat wave days in 2005.

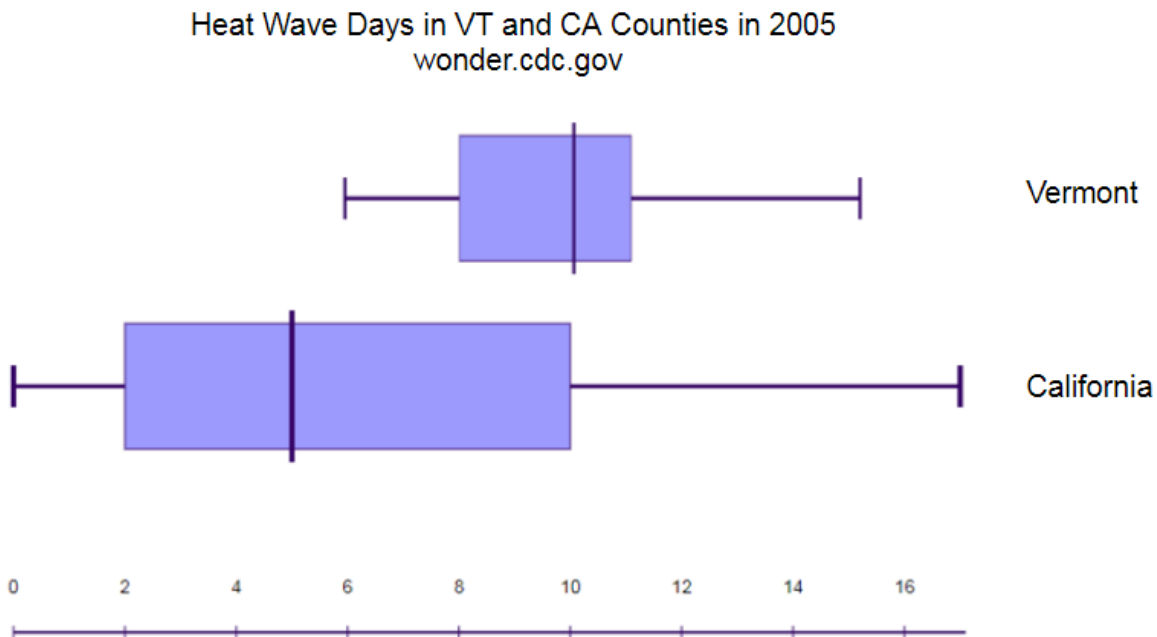
3) There are 58 counties in California. With that information, which of the following statements can be concluded from the chart?

- (1) Exactly 29 counties had between 2 and 10 heat wave days in 2005.
- (2) At least 29 counties had between 2 and 10 heat wave days in 2005.
- (3) Exactly 29 counties had more than 10 heat wave days in 2005.
- (4) Exactly 15 counties had 10 or more heat wave days in 2005.
- (5) All 58 counties had fewer than 17 heat wave days in 2005.

4) What is the interquartile-range for the California heat wave data shown in the chart?

- (1) 2
- (2) 8
- (3) 10
- (4) 16
- (5) 17

Use the chart below for questions 5-6:



5) Which of the following statements cannot be concluded from the chart above?

- (1) There is more variability in the number of heat wave days in California than in Vermont.
- (2) There are more counties in Vermont that had 6 or more heat wave days than in California.
- (3) Every county in Vermont had at least 6 heat wave days in 2005.

(4) The median number of heat wave days was lower in California than it was in Vermont.

(5) At least half of the counties in Vermont and three fourths of the counties in California had 10 or fewer heat wave days in 2005.

6) There are 14 counties in Vermont. Create a data set with 14 numbers that could be the Vermont data from 2005. (In other words, create a data set that fits the graph.)

7) Bonus! What do you think accounts for the differences between the distributions of heat wave days in Vermont and California?

Heat Wave – Reflection Questions

What is something you learned about quartiles?

What is something you learned about drawing a box and whisker chart?

What is something you learned about the box in a box and whisker chart?

What is something you learned about the whiskers in a box and whisker chart?

What is something you learned about comparing sets of data using box and whisker charts?

What is something you learned about trends in data?

What is something you learned about heat wave days in Massachusetts?