Mathematics

8.12 The student uses statistical procedures to describe data. The student is expected to:

(A) select the appropriate measure of central tendency or range to describe a set of data and justify the choice for a particular situation.

(C) select and use an appropriate representation for presenting and displaying relationships among collected data, including line plots, line graphs, stem and leaf plots, circle graphs, bar graphs, box and whisker plots, histograms, and Venn diagrams, with and without the use of technology.

Technology Applications

The student demonstrates knowledge and appropriate use of hardware components, software programs, and their connections. The student is expected to:

(1)(a) demonstrate knowledge and appropriate use of operating systems, software applications, and communicate and networking components.

(1)(c) demonstrate the ability to select and use software for a defined task according to quality, appropriateness, effectiveness, and efficiency.

(1)(f) perform basic software application function including, but not limited to, opening an application program and creating, modifying, printing, and saving documents.

The student uses appropriate computer-based productivity tools to create and modify solutions to problems. The student is expected to:

(7)(a) plan, create, and edit documents created with a word processor using readable fonts, alignment, page setup, tabs, and ruler settings.

(7)(b) plan, create, and edit spreadsheet documents using all data types, formulas and functions, and chart information.

(7)(e) create a document using desktop publishing techniques including, but not limited to, the creation of multi-column or multi-section documents with a variety of text-wrapped frame formats.

(7)(g) integrate two or more productivity tools into a document including, but not limited to, tables, charts, and graphs, graphics from paint or draw programs, and mail merge.

The student formats digital information for appropriate and effective communication. The student is expected to:

10)(a) use productivity tools to create effectiveness document files for defined audiences such as slide shows, poster, multimedia presentations, newsletters, brochures, or reports.

(11)(a) publish information in a variety of ways including, but not limited to, printed copy, monitor display, Internet documents, and video.

Materials

Advanced Preparation:

- Pre-cut **TEAM CARDS**, enough for one card per group of 2-3 students
- Copies of **Paper Hockey Puck** Directions, **Go Team!**, (optional) **Purple or Orange**, and **Pure Gold** worksheets
- Access to spreadsheet and large monitor or projector/screen for demonstrations
- Access to **GoTeam** and **PureGold** spreadsheet file for each student or pair of students if there is not enough technology available
- Access to a word processor and/or presentation software for each pair of students and **Purple or Orange** file

For each student:

- **Go Team!** and (optional) **Purple or Orange** worksheets
- **PureGold** worksheet

For each student group of 2 -3 students:

- Blank paper or large index cards, one per student
- One **Paper Hockey Puck Directions**, scissors, and one tape measure
- One **TEAM CARD** and one sheet of chart paper, markers

ENGAGE

The Engage portion of the lesson is designed to create student interest in the concepts addressed. This part of the lesson is designed for groups of 2-3 students and then whole group instruction.

1. In small groups, direct the students to take turns flicking a paper hockey puck so that it glides along the surface of the table, measuring the distance traveled (to the nearest inch), and writing that amount with a marker on a blank sheet of paper (or large index card) until each student in the group has a measurement. (Students may design their own paper hockey pucks or use the **Paper Hockey Puck Directions** handout.)

Facilitation Questions

- (Before conducting the activity of flicking the hockey puck) What factors involved in the actual flicking of the hockey puck and measurement of the distance should we standardize for consistency?
  Answers may vary. Factors may include how you made the hockey puck, how you hold the hockey puck, starting points, surface on which you are flicking the hockey puck, method of flicking the hockey puck, measuring tools, etc.

Facilitation Questions

- Why should we be concerned with consistency in the way we collect our data? In other words, how might inconsistency affect the validity of our results?
  We need to limit the number of variables (factors that may differ) for each event in order to compare distances made when all other factors were the same/controlled. For example, results might be skewed if one group flicked their hockey puck across carpet, and others flick theirs across a slick table.

2. Once each student in the group has recorded his/her measurement, have students line up in numeric order around the room based on the distance traveled by their hockey puck.

Facilitation Questions

- How did you (or how might you) represent the fact that two or more students may have had the same measurement?
  Stand behind each other in one spot
- How did you (or how might you) represent the proportional distance between the values of your distances?
  Consider themselves as a human number line…proportionally spacing the values.

3. Prompt students to determine the median, mode, and range of their data without technology. Record these statistics on a sheet of chart paper or overhead projector for all to see. Have all students return to their seats.

Facilitation Questions

- How did you (or how might you) determine the median of your data without technology? What impact does the spread of the data have on the value of the median? What is the significance of this value?
  Answers may vary. Students could count off from each end of the line to find the middle of their line (median), averaging the values should there be two students in the middle. You might have students raise their hands if their value is above the median, and then do the same for those below the median…to demonstrate that this is a "middle" value. This middle value is not affected by extreme values (outliers) on either end of the data. In this example, it would not be affected by distances that were significantly lower or higher than the rest of the group.

Facilitation Questions

- How did you (or how might you) determine the mode of your data without technology? What is the significance of this value?
  Answers may vary. Students could look for where they have students with the same values lined up behind each other (if any). If a set of data has a mode, it indicates that there is a value that occurred multiple times. In this example, this measure of central tendency may not be the best representative of the data unless there is a measurement that occurs many more times than the others.
- How did you (or how might you) determine the range of your data without technology? What is the significance of this value?
  Answers may vary. Students could have the students on each end of the line (highest and lowest) find the difference in their measurements. The range helps to describe the spread of the data. In this example, it would tell us if the distances achieved were about the same, or if some students were able to flick for distances that were much longer than others.

4. Input the data into a blank spreadsheet file (using a large monitor or projector for viewing) and demonstrate using formulas to calculate the median, mode, and range.
   - To find median: =median (highlight range of cells with data)
   - To find mode (*If there is more than one mode, Excel returns the smallest mode*): =mode (highlight range of cells with data)
   - To find range: =max (highlight range of cells with data) – min (highlight range of cells with data)
5. Prompt the students to estimate the mean. Verify the estimate using the spreadsheet. To find mean: =average (highlight range of cells with data)
6. Add this statistic to the sheet of chart paper or overhead for all to see.
7. Have students return to their original number line position.
8. Direct the students to form a human histogram.

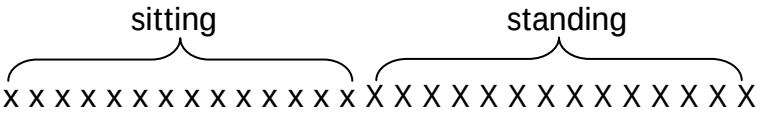
Facilitation Questions

- How might you organize yourselves to create a histogram?
  Students in each group should stand behind each other, forming a line ("bar") for each group/range of data.
- What ranges would be appropriate for the bars in our histogram?
  Answers may vary based on the data.
- What conjectures can we make based on the spread of the data within this human histogram? In other words, what do we notice about the "shape" of the data?
  Answers may vary. Based on the actual spread of the data, students should notice clusters and/or gaps in the spread of the data.

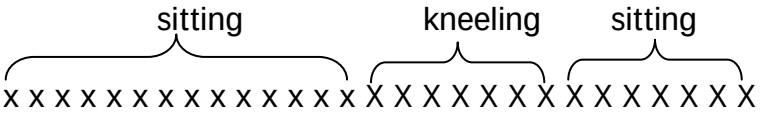
Facilitation Questions

- What kind of information was "lost" when we grouped students together to form the human histogram?
  While we know the number of pieces of data in each group, we do not know where in the range for each group the data lies. For example, if there are 5 values within a range of 20 to 30 inches, they could all be closer to 20 inches, 30 inches, or spread throughout the range.

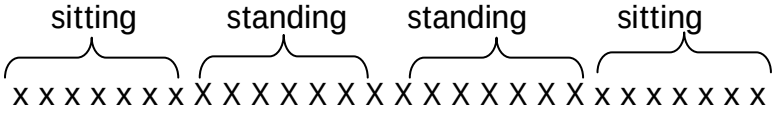
9. Direct the students to form a human box and whisker plot by asking the following questions.
   a. What was the value of our median? (Identify that person or point between two persons if there is an even number of data points.)
   b. Raise your hand if your value is above the median or, in other words, if you are in the upper half of the data? (Once you agree that you have the upper half with their hands raised, ask the lower half to sit on the floor and the upper half can put their hand down.)



   c. Raise your hand if you are above the median of those standing, in other words, if you are in the upper half of the upper half of the data? (Once you agree that you have the upper half with their hands raised, ask the upper-upper half to sit on the floor and the lower-upper half to kneel.)



   d. Repeat the procedure to find the upper half of the lower half. Ask the lower - lower half to sit and the upper - lower half to remain standing. Ask the lower - upper half to now stand.



10. Inform the students that they will create histograms and box and whisker plots using technology in the next activity. They will use this "human" version to help understand what is happening "behind the scenes" as the technology creates the graphs.

Facilitation Questions

- How would you describe the data for the group of students still standing? What is the significance of this group?
  Answers may vary. Students should note that those standing represent the middle half of the data. Essentially this is the middle group that ignores (does not include) the highs and the lows (outliers).

11. Prompt students to return to their seats.

EXPLORE

The Explore portion of the lesson provides the student with an opportunity to be actively involved in the exploration of the mathematical concepts addressed. This part of the lesson is designed for groups of 2-3 students initially and moving to pairs or individual investigation.

1. Divide the class into groups (teams) of 2-3 students each.
2. Give each group a **TEAM CARD** with a list of numbers that represents the number of hits last season by each person on a particular baseball team. (Duplicate sets of data will be distributed to promote comparing/contrasting comments later in the activity.)
3. Ask each group to calculate the mean (average) of the number of hits their team had last season.
4. Prompt each group to write their team name, list of hits, and the mean on a piece of chart paper.

Facilitation Questions

- What do you notice about each team?
  The mean is the same for each team.
- If all three teams had the same mean (average) number of hits last season, what other statistical measures might you examine in order to distinguish between the teams?
  Answers may vary. Lead students to classify teams by measurements such as median, mode, range or the spread of the data.

5. Prompt students (or pairs of students if there is not enough technology available) to open the spreadsheet **GoTeam**. Point out that there are four "worksheets" within the file (Blue Team, Green Team, Red Team, Graphs). Have students click on the various tabs toward the bottom of the screen to become familiar with moving between the pages.
6. Distribute the **Go Team** worksheet. (Include the optional page where students may record a summary of their spreadsheet work to each student, if printing capabilities are not available.) If printing student spreadsheets is an option, wait to distribute this worksheet until step 11.
7. In the **Blue Team** worksheet (spreadsheet), ask the students to fill in the number of hits (using the chart paper/Team Card) for each of the 16 players (in any order) in the light yellow cells (B7 – B22).

Facilitation Questions

- How will the order in which you input the data affect the calculations of central tendencies by the technology?
  Answers may vary. One benefit of using technology to calculate central tendencies is that order of input does not matter. This would be especially helpful if there were a large number of data points.

8. Once the data has been entered in cells B7-B22, students must sort the data in order for the box and whisker plot to graph correctly. (*Excel uses a different formula to find the lower and upper quartiles than the state will use to assess students. To make the box and whisker plot match the state expectations, certain formulas were entered into the spreadsheet. These formulas will only work if the data is sorted in ascending order.*)
9. Guide the students through inserting the statistical formulas in the light yellow cells in column F. (NOTE: Formula hints will be visible when the cursor is over that cell.)
10. Prompt the students to input the data and formulas for the Green Team and the Red Team, in order for the technology to generate the remaining histograms and box and whisker plots.
11. Distribute the **GoTeam** worksheet with questions 1 – 7 to each student.
12. Prompt the students to use the Graphs worksheet to help them answer the 7 questions on the **Go Team** worksheet.
13. Question #7 will be addressed further in the "Explain" phase.

EXPLAIN

The Explain portion of the lesson is directed by the teacher to allow the students to formalize their understanding of the TEKS addressed in the lesson.

1. Debrief the concept of median using the following questions.

Facilitation Questions

- The median is also referred to as the 50^{th} percentile. Why is this?
  50^{th} percentile refers to the value such that, if the data points are sorted from least to greatest, 50% of the data points are less than this value and 50% of the data points are greater than this value.
- Where is the median in your data list? How many values are less than the median and how many are greater than the median?
  The median is 50 hits. There are 8 players with fewer than 50 hits, and 8 players with more than 50 hits.

Facilitation Questions

- If the median of the entire group is the 50th percentile, what is the significance of the 25th percentile? How many values are less than the 25th percentile? How many values are greater than the 25th percentile?
  This is the median of the lower half of the data. 25% of the data points are less than this value, leaving 75% of the data points to be greater than this value. In this case there are 4 values below the 25th percentile and 12 values above the 25th percentile.
- If the median of the entire group is the 50th percentile, what is the significance of the 75th percentile? How many values are less than the 75th percentile? How many values are greater than the 75th percentile?
  This is the median of the upper half of the data. 75% of the data points are less than this value, leaving 25% of the data points to be greater than this value. In this case there are 12 values below the 75th percentile and 4 values above the 75th percentile.
- What is the statistical significance of the median of any group of data? In other words, why does knowing the median of a group of data give us more information than just knowing the average (mean)?
  The median represents the "middle" of the data, once the data has been ordered. Since the value of the median is a result of position alone, it is not affected by outliers, whereas even one or two outliers (data points that are significantly higher or lower than the rest of the data) might skew the mean.
- How is this discussion of median versus mean related to the data we have on the number of hits for each player on a team?
  The median number of hits is the value that represents the number of hits that is in the "middle" once the values have been ordered. Since this value is based on position alone, it is not as likely to be affected should certain players suddenly have more or less hits.

2. Using a large monitor or projector/screen for viewing, demonstrate how the median is not affected by outliers.

Facilitation Questions

- Since the median and mean numbers of hits for the Blue team are close to the same values (50 and 50.7 respectively), which would change the most if your best hitter was traded for someone with 500 hits? Justify your prediction.
  Answers may vary. The median would not change because the data points would not change in position if the largest value is exchanged for an even larger value. The mean would change more because there would be a larger total number of hits, therefore a larger average when those hits were divided out among the players. Type "500" in place of the "120" in the computer and note the lack of change in the median and the significant change in the mean.

Facilitation Questions

- With this exchange of players, does the median (50) or the mean (74.4) better describe the data for the team? Justify your selection.
  The median is a better descriptor of the data. Justifications may vary, but students should note that there are very few players on the team that are hitting at or above the mean, while there are still one-half of the players hitting at or above the median (and the other one-half hitting at or below the median).
- How is having the data in the computer helping us justify our thoughts about referring to medians versus means?
  Answers may vary. As we are changing values, the computer instantaneously recalculates the central tendencies so that we can focus our discussion on the meaning and significance of each value, without having to get "bogged down" in the recalculations.
  NOTE: Ensure that students return the "500" value to "120" before proceeding.

3. Prompt the students to explain how the median is reflected in each graphical representation on the **Graphs** worksheet.

Facilitation Questions

- How is the median represented in each of the histograms on the **Graphs** worksheet?
  The median value is contained within the bar representing the group of data on the "41 – 60" bar on each histogram.
- If you did not know the value of the median from the list of data, how could you determine it from the histogram?
  By counting frequencies for each bar, you would be able to narrow it down to the appropriate bar, but you would not be able to identify the exact median from the histograms alone.
- How is the median represented in each of the box and whisker plots on the **Graphs** worksheet?
  The median value is represented by the line within the box on each box and whisker plots.
- If you did not know the value of the median from the list of data, how could you determine it from the box and whisker plot?
  By looking at the position of the line within the box and the number line associated with the plot, you could find the value of the median.

4. Prompt the students to make conjectures about the mean, given the median and the shape of the data in the graphical representations.

Facilitation Questions

- If you had to make an estimate of the mean (average) number of hits, would you rather base your estimate on the histogram or the box and whisker plot? Justify your answer.

  Answers may vary. In box and whisker plots, it is easy to estimate the median and then make conjectures about the value of the mean based on the shape of the data around the median. Because of grouping data in ranges on the histogram, it may make the median harder to pinpoint, therefore making it more difficult to determine the relationship between the mean and median.
- Knowing that the mean number of hits for each team is about 50.7, explain why this value "makes sense" based on what you know about the median and the shape of the data in the box and whisker plot.

  Possible answers:

  Blue team – One might estimate that the value of the mean would be fairly close to the value of the median since the shape of the data in the box and whisker plot indicates that the data on either side of the median is somewhat evenly spread…possibly skewing just a little above the median since there is a little bit larger spread in the data above the median than in the data below the median. You can see this by noting that the right whisker on the box and whisker plot is longer.

  Green team- Similar to the discussion about the Blue team, but this time the data below the median is more spread out, leading you to estimate that the mean is a little less than the median.

  Red team – While the right whisker is longer than the left, the size of the box to the right of the median indicates that those values are very close to the median. Looking at the spread of the data to the left of the median, one could determine that since these values tend to be farther from the median, the mean would be less than the median.

5. Debrief the concept of range by asking the following Facilitation Questions.

Facilitation Questions

- What is the statistical significance of the range of any group of data? In other words, why does knowing the range of a group of data give us more information than just knowing the average (mean)?
  The range is the difference between the maximum and minimum values in a set of data. The range gives you an idea of the spread of the data. The smaller the range, the closer the values of the data points are to each other. As the range increases, so does the spread between the values of the data points.
- How is this discussion of range related to the data we have on the number of hits for each player on a team?
  The range in numbers of hits for a team will help us distinguish between teams where the number of hits for the players are more consistent (closer together) and teams where they are not as consistent (some players have significantly more hits than other players on the same team).

6. Prompt the students to make conjectures about the range, given the shape of the data in the graphical representations.

Facilitation Questions

- Would you feel more confident estimating the range of a set of data from a histogram or a box and whisker plot? Justify your answer.
  Because some values might get "lost" in the bars at the extremes of the histogram, it is sometimes difficult to closely estimate the range given only this graphical representation…but you can determine a "ballpark" range.
- How is the range reflected in each of the histograms on the **Graphs** worksheet ?
  Blue team - While looking at the histogram alone you could only estimate the range to be between 81 and 120, an exact range can be seen on the box and whisker plot by comparing the ends of the whiskers.
  Green team - While looking at the histogram alone you could only estimate the range to be between 41 and 80, an exact range can be seen on the box and whisker plot by comparing the ends of the whiskers.
  Red team - While looking at the histogram alone you could only estimate the range to be between 81 and 120, an exact range can be seen on the box and whisker plot by comparing the ends of the whiskers.

7. Debrief the concept of mode using the following questions.

Facilitation Questions

- Why does knowing the mode of a group of data give us more information than just knowing the average (mean)?
  The mode is the data point with the greatest frequency. If a set of data has a mode, it tells you which value was most common. Depending on the frequency, the value of the mode may or may not impact the mean.
- How is this discussion of mode related to the data we have on the number of hits for each player on a team?
  The mode in numbers of hits for a team only tells us if a particular number of hits occurs more often than others. In this case, it would not be uncommon for a set of data to have no mode because there are so many possibilities for numbers of hits.
- If you did not know the value of the mode from the list of data, how could it be determined from the histogram or box and whisker plot alone?
  The specific mode is not evident on either representation.

8. Prompt the students to make conjectures about any outliers or clusters, given the shape of the data in the graphical representations.

Facilitation Questions

- For each team, are there any outliers or clusters of data? If so, how are they represented in the histogram and box and whisker plot?
  Possible answers:
  Blue Team – Two players have a significantly higher number of hits (110 and 120) as seen by the gap in the histogram and the long right whisker on the box and whisker plot. According to the histogram, only one player falls within the range including both the mean and median, leading you to believe the data is spread out.
  Green Team – Both the histogram and box and whisker plot show that the vast majority of the players have numbers of hits close to both the mean and median. According to the box and whisker plot, the "middle half" of the data is compressed between the upper 40s and 60, leading you to believe that the numbers of hits for those players are very close together.
  Red Team – While the numbers of hits for players are not as tight around the mean and median as the green team, there are still many players clustered within the upper 30s and 60.

9. To answer question #7 on their **Go Team** sheets, prompt the students to use what they know about the central tendencies and range, as well as the graphical representations to explain which team you would like to join.

Facilitation Questions

- Now that we have explored these graphical representations further, think about your response to #7 on your **Go Team** worksheet. If you had the opportunity to join any of these teams for next season, which would it be? Explain.
  Answers may vary. Look for explanations that use the data as their justification such as--
  The Blue and Red Teams both have players having more than 100 hits.
  The Red Team has the highest median.
  The Green Team has more consistency between players when it comes to number of hits.
  The Blue Team has a number of players hitting well above the median.

ELABORATE

The Elaborate portion of the lesson provides an opportunity for the student to apply the concepts of the TEKS within a new situation. This part of the lesson is designed for pairs of students.

1. Pair up the students, to foster student conversation and extend their thoughts as they interpret graphical representations.
2. Direct each pair of students to open **Purple or Orange** with their word processor. NOTE: They need the file copy (not just a hard copy) in order to have the ability to cut and paste the graphical representations in their newsletter or slide show.
   *(Optional) Distribute a hard copy of the **Purple or Orange** file to each student for reference.*
3. Prompt students to summarize the directions for the task, giving them additional directions on how to name and save their newsletter or slide show.
4. When half of the work time remains, have the student pairs go on a quick (5 - 10 minutes) "spy mission." (For each pair, one student will stay with the work/computer to share their work/thoughts with others, and the other will visit with other students to "compare notes" and possibly hear a different point of view or get a technology tip. The pairs will reunite and continue their work on their newsletter or slide show.)
5. After the students have completed the "spy mission," prompt students to defend their answer to #5 to the whole group.

Facilitation Questions

- What evidence is there in the graphical representation(s) to defend your position?
  Answers may vary.
- If the data point of 100 hits was added, what change (if any) would you notice in the graphical representation(s)?
  Answers may vary.
- If you could call the front office of either team, what question(s) might you ask to help you get a better understanding of the data? In other words, what do you wish you knew that you either don't know or are not sure of based on the graphical representation(s) alone?
  Answers may vary.

6. Allow each pair of students a short amount of time to make adjustments to their newsletter or slide show based on information they gained from the whole group discussion.
7. Upon completion of the **Purple or Orange** activity, the teacher should use a rubric to assess student understanding of the concepts addressed in this lesson.

EVALUATE

The Evaluate portion of the lesson provides the student with an opportunity to demonstrate his or her understanding of the TEKS addressed in the lesson.

1. Distribute the **Pure Gold** activity sheet to each student.
2. Clarify the location of and saving procedure for the **PureGold** spreadsheet.
3. Upon completion of the **Pure Gold** activity sheet, the teacher should use a rubric to assess student understanding of the concepts addressed in this lesson.

Answers and Error Analysis for selected response questions

Question Number	TEKS	Correct Answer	Conceptual Error	Conceptual Error	Procedural Error	Procedural Error	Guess
1	8.12(C)	B	A	D	C		
2	8.12(C)	C	A	B			D
3	8.12(A)	D	A	B	C		
4	8.12(C)	A	C	D	B		

Team Stats

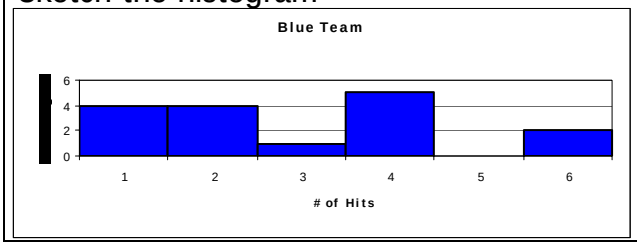
(Give this page to students if printing their spreadsheet is not an option)

See GoTeam-Key spreadsheet.

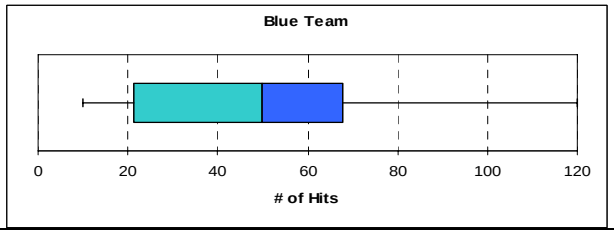
BLUE TEAM

Minimum:	*10*	Mean:	*50.7*	25th %-tile:	*19.5*
Maximum:	*120*	Mode:	*10*	Median:	*50*
Range:	*110*			75th %-tile:	*69.5*

Sketch the histogram



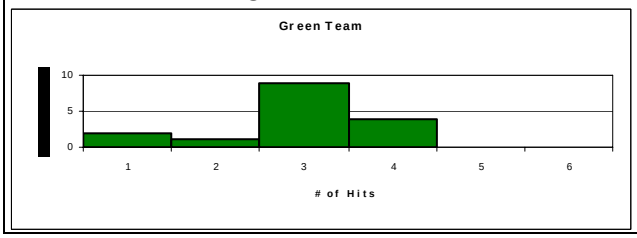
Sketch the box and whisker plot



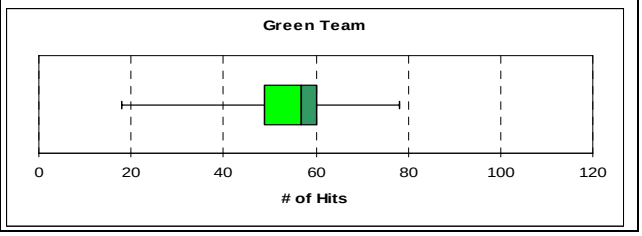
GREEN TEAM

Minimum:	*18*	Mean:	*50.7*	25th %-tile:	*48.5*
Maximum:	*78*	Mode:	*57, 61*	Median:	*57*
Range:	*60*			75th %-tile:	*60.5*

Sketch the histogram



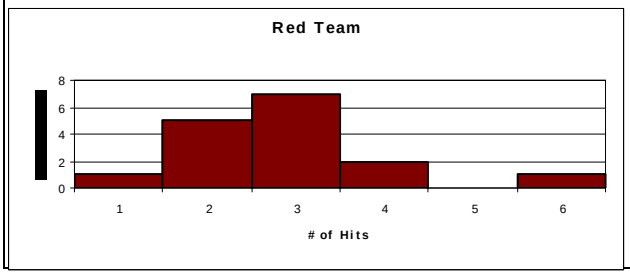
Sketch the box and whisker plot



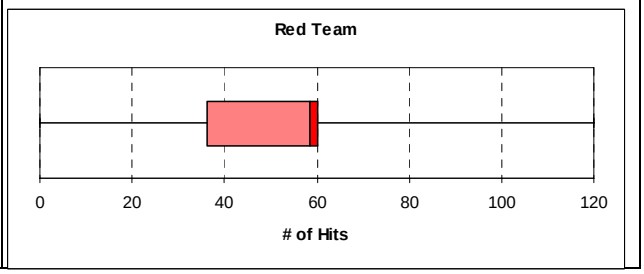
RED TEAM

Minimum:	*0*	Mean:	*50.7*	25th %-tile:	*34.5*
Maximum:	*120*	Mode:	*60*	Median:	*58.5*
Range:	*120*			75th %-tile:	*60*

Sketch the histogram



Sketch the box and whisker plot



Go Team!

Use the terms in the word bank below to complete the statements about the statistics and graphic representations in your spreadsheet. Each term can be used only once.

median	Red	mode	Blue
Green	range	outlier	mean

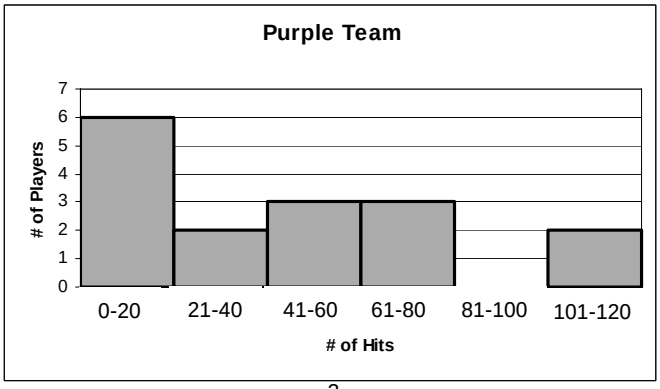
1. The box and whisker plot of the ***Red*** Team has the longest whisker. This is usually an indication that the set of data contains at least one ***outlier*** .
2. The ***mode*** of the data is the central tendency for which the graphic representations give us the least information.
3. The graphic representation with the smallest box (on the box and whisker plot) or with the middle bars significantly taller than the outer bars (on the histogram) for the ***Green*** Team reflects the fact that the number of hits for many of the players on that team is close to the ***median***.
4. While the data for each of the three teams is very different, the ***mean*** number of hits is the same for all.
5. The ***Blue*** and Red Teams both have players with more than 100 hits.
6. The ***range*** of the number of hits was the smallest for the Green Team.

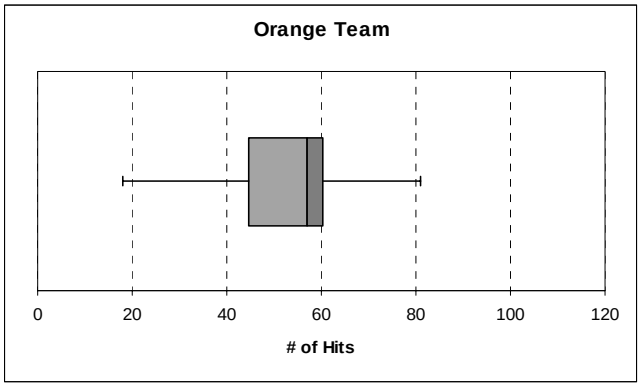
---

7. If you had the opportunity to join any of these teams for next season, which would it be? Explain using statistics and/or the graphical representation(s) to justify your selection.
   Answers may vary.

Purple or Orange? (hard copy of Microsoft Word file)

Below are graphical representations of the number of hits last season by members of the Purple and Orange teams.





Hard Hitting Harold (H^3 for short) has offers to join either the Purple team or the Orange team. H^3 had 100 hits last season.

As a local sports reporter, you have received the task of analyzing the impact for each team, should H^3 join either the Purple or the Orange team. You must base your analysis on what you can gather from the graphical representations you have received.

Use either a word processor to create a newsletter or presentation software to create a slide show that will communicate your interpretations. Copy and paste the graphical representations into your newsletter or slide show and use the drawing tools to help make your points.
Your newsletter or slide show should answer the following questions.

1. From the given graphical representations, what do you know about the spread of the data (numbers of hits per player) for the Purple team? for the Orange team? (Include a "discussion" of any clusters, gaps, and/or outliers.)
   Answers may vary. Purple…spread between 0 and 120 with half below 41 and a couple of outliers between 101 and 120. Orange…middle half clustered between about 45 and 60…overall spread between about 18 and 81.

2. Should H^3 join the team, how would his number of hits (100) impact the current spread of the data for the Purple team? for the Orange team?
   Answers may vary. Purple…would appear to fill the 81-100 gap in the histogram, but is just one hit away from being included in the 101-120 group.
   Orange…would increase the overall spread of the data since the current maximum is around 81…would likely be an outlier, almost 20 greater than the current maximum.

3. From the given graphical representations, what do you know about the current range, median, and mean number of hits for the Purple team? for the Orange team?
   Answers may vary. Purple…range is somewhere between 80 (20 to 101) and 120 (0 to 120)…median falls between the highest value in the 21-40 bar and the lowest value in the 41-60 bar…mean is likely higher than the median as a result of the outliers over 100. Orange…range is just over 60…median is around 57…mean will likely be less than the median due to the larger spread (not number of data points) of the data below the median

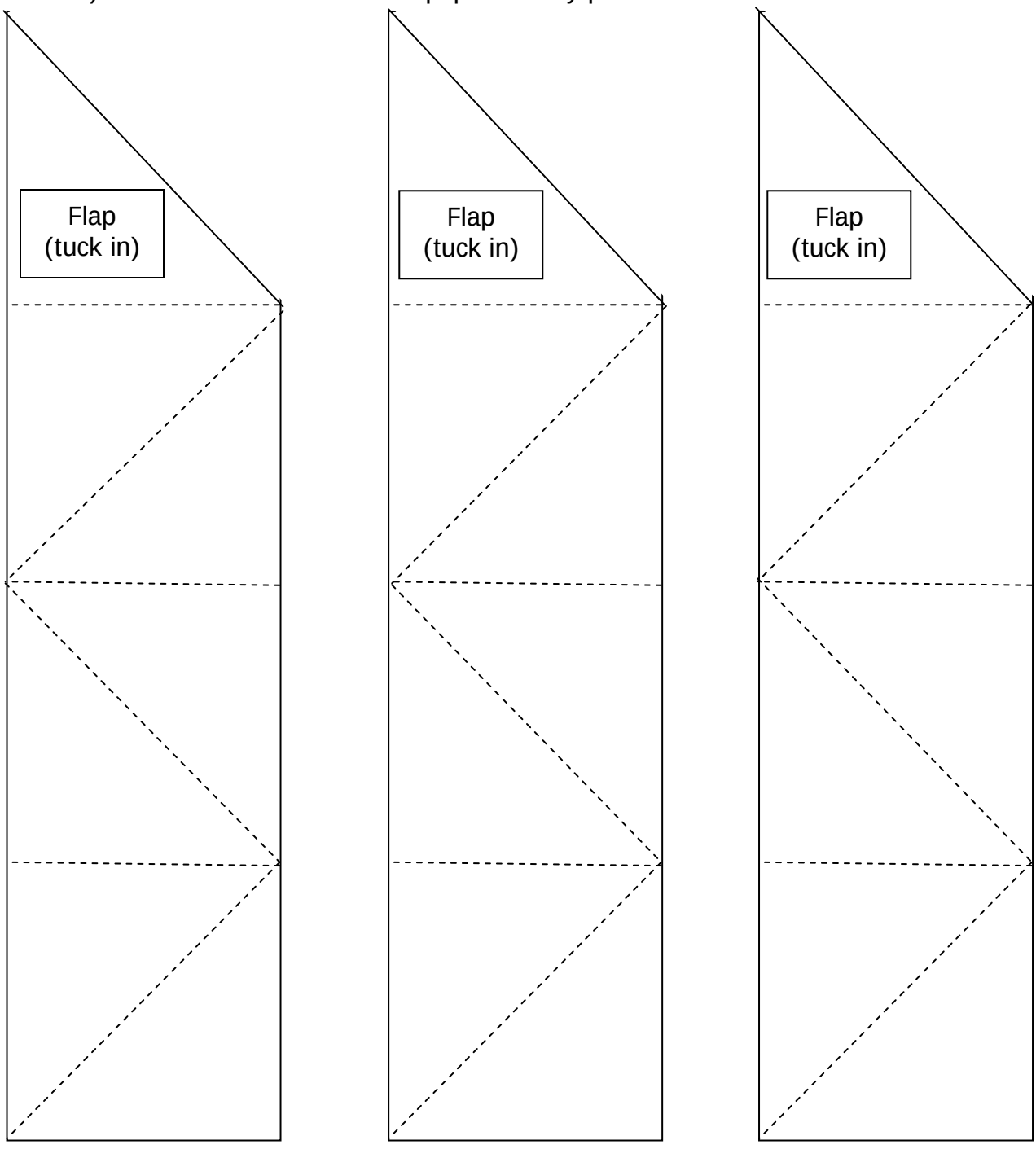
4. Should H^3 join the team, what would be the impact on the range, median, and mean number of hits for the Purple team? for the Orange team? *Answers may vary. Purple…will not change the range…will likely increase the median and mean (although hard to say by how much because of the ranges within the bars) because it is greater than both. Orange…will increase the range by around 20,,,will likely increase the median and mean because it is greater than both (probably the mean more than the median since the 100 is significantly greater than the current maximum)*

5. In your opinion, which team would benefit the most from having H^3 join their team?
   Answers may vary. One could make a case for the Purple team since the 100 would begin to fill in the current gap and/or help to balance against those with very few hits, while the Orange team would gain a player with significantly more hits than the rest of their players, therefore increasing their average number of hits. Accept students' opinions if they can support that opinion with interpretations of the statistics.

6. As an added note or disclaimer, compare and contrast the amount and type of information you were able to get from the histogram versus the box and whisker plot when you addressed questions #1 and 3. What information might you get from a histogram that you would not get from a box and whisker plot? What information might you get from a box and whisker plot that you would not get from a histogram?
   Answers may vary. Both help us to see the spread in the data. Since histograms often have more than 4 groupings (6 in this case), gaps and outliers may be more evident. Box and whisker plots, on the other hand, often allow you to estimate the median and range better.

PAPER HOCKEY PUCK DIRECTIONS

Cut out, fold on the dotted lines, and tuck in the flap. (It will form a right triangle when folded.) Each student will need one paper hockey puck.



Use your thumb and pointer of one hand to hold the triangle vertically by the vertices on either side of the hypotenuse. Use your other hand to flick the triangle (hockey puck).

TEAM CARDS

Blue Team 10, 120, 15, 16, 23, 27, 66, 39, 40, 60, 10, 61, 64, 73, 77, 110	**Blue Team** 10, 120, 15, 16, 23, 27, 66, 39, 40, 60, 10, 61, 64, 73, 77, 110
Green Team 20, 51, 21, 78, 48, 49, 50, 57, 18, 57, 58, 59, 60, 61, 63, 61	**Green Team** 20, 51, 21, 78, 48, 49, 50, 57, 18, 57, 58, 59, 60, 61, 63, 61
Red Team 29, 60, 30, 38, 60, 40, 42, 120, 57, 60, 0, 60, 61, 63, 31, 60	**Red Team** 29, 60, 30, 38, 60, 40, 42, 120, 57, 60, 0, 60, 61, 63, 31, 60

Team Stats

BLUE TEAM

Minimum:	_____	Mean:	_____	25th %-tile:	_____
Maximum:	_____	Mode:	_____	Median:	_____
Range:	_____			75th %-tile:	_____

Sketch the histogram	Sketch the box and whisker plot

GREEN TEAM

Minimum:	_____	Mean:	_____	25th %-tile:	_____
Maximum:	_____	Mode:	_____	Median:	_____
Range:	_____			75th %-tile:	_____

Sketch the histogram	Sketch the box and whisker plot

RED TEAM

Minimum:	_____	Mean:	_____	25th %-tile:	_____
Maximum:	_____	Mode:	_____	Median:	_____
Range:	_____			75th %-tile:	_____

Sketch the histogram	Sketch the box and whisker plot

Go Team!

Use the terms in the word bank below to complete the statements about the statistics and graphic representations in your spreadsheet. Each term can be used only once.

median	Red	mode	Blue
Green	range	outlier	mean

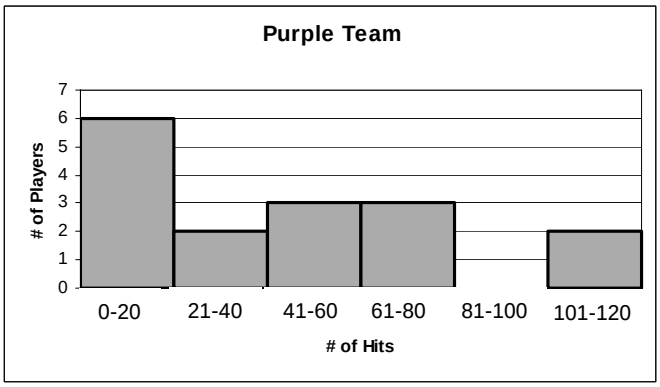
1. The box and whisker plot of the _____ Team has the longest whisker. This is usually an indication that the set of data contains at least one _____ .
2. The _____ of the data is the central tendency for which the graphic representations give us the least information.
3. The graphic representation with the smallest box (on the box and whisker plot) or with the middle bars significantly taller than the outer bars (on the histogram) for the _____ Team reflects the fact that the number of hits for many of the players on that team is close to the _____.
4. While the data for each of the three teams is very different, the _____ number of hits is the same for all.
5. The _____ and Red Teams both have players with more than 100 hits.
6. The _____ of the number of hits was the smallest for the Green Team.

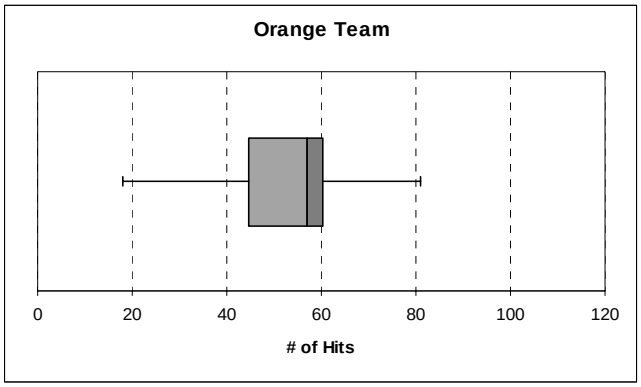
---

7. If you had the opportunity to join any of these teams for next season, which would it be? Explain using statistics and/or the graphical representation(s) to justify your selection.

Purple or Orange? (hard copy of Microsoft Word file)

Below are graphical representations of the number of hits last season by members of the Purple and Orange teams.





Hard Hitting Harold (H^3 for short) has offers to join both the Purple team and the Orange team. H^3 had 100 hits last season.

As a local sports reporter, you have received the task of analyzing the impact for each team, should H^3 join either the Purple or the Orange team. You must base your analysis on what you can gather from the graphical representations you have received.

Use either a word processor to create a newsletter or presentation software to create a slide show that will communicate your interpretations. Copy and paste the graphical representations into your newsletter or slide show and use the drawing tools to help make your points. Your newsletter or slide show should answer the following questions.

1. From the given graphical representations, what do you know about the spread of the data (numbers of hits per player) for the Purple team? for the Orange team? (Include a "discussion" of any clusters, gaps, and/or outliers.)

2. Should H^3 join the team, how would his number of hits (100) impact the current spread of the data for the Purple team? for the Orange team?

3. From the given graphical representations, what do you know about the current range, median, and mean number of hits for the Purple team? for the Orange team?

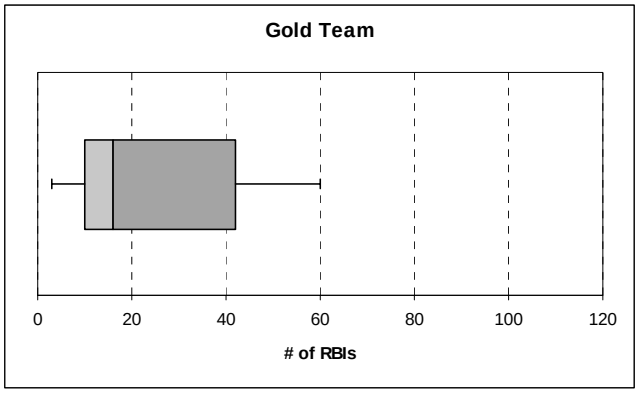
4. Should H^3 join the team, what would be the impact on the range, median, and mean number of hits for the Purple team? for the Orange team?

5. In your opinion, which team would benefit the most from having H^3 join their team?

6. As an added note or disclaimer, compare and contrast the amount and type of information you were able to get from the histogram versus the box and whisker plot when you addressed questions #1 and 3. What information might you get from a histogram that you would not get from a box and whisker plot? What information might you get from a box and whisker plot that you would not get from a histogram?

Pure Gold

You have just been hired as the manager of the Gold Team. A plot of the number of RBIs (runs batted in) of your team is shown below.

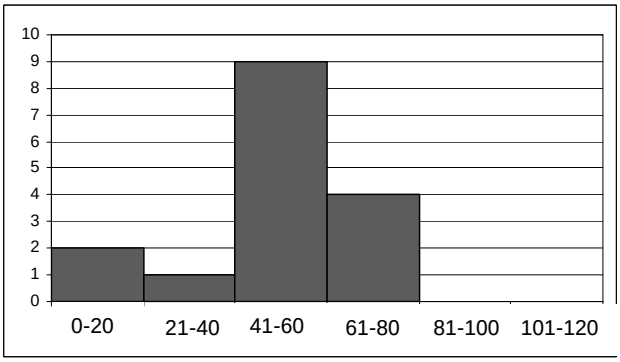


Your first job as team manager is to add 3 players (to replace 3 that retired) to the team. You must meet these goals.

- Do not increase the current range of RBIs.
- Keep the various numbers of RBIs as clustered around the median as possible.

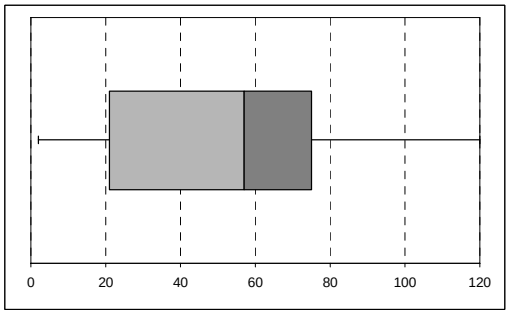
a. Open the **PureGold** spreadsheet to see the RBI statistics on your current players and the players that are available to join your team.
b. Add 3 players to get the desired results.
c. Prepare a statement for the press that lists the RBIs of the players you added and describes the impact of these additions on each of the following statistical measures for your team.
d. Justify your statement by including the amount of change (if any) from the original statistics and original box and whisker plot, along with how these additions might benefit the team.
   - Range
   - Median
   - Mean

1.

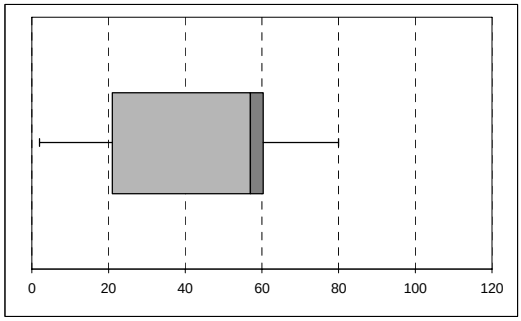


Which of the following box and whisker plots would contain data similar to the histogram above?

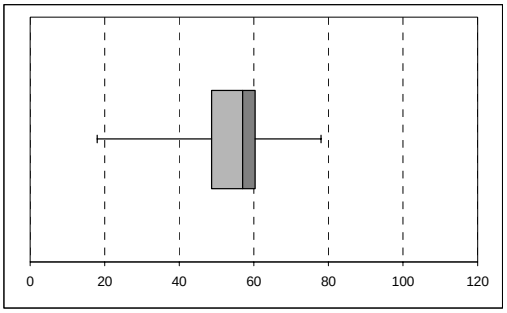
A.



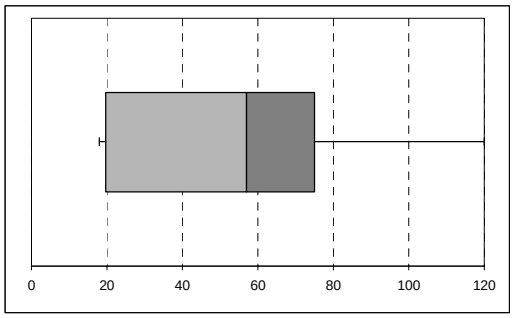
C.



B.



D.

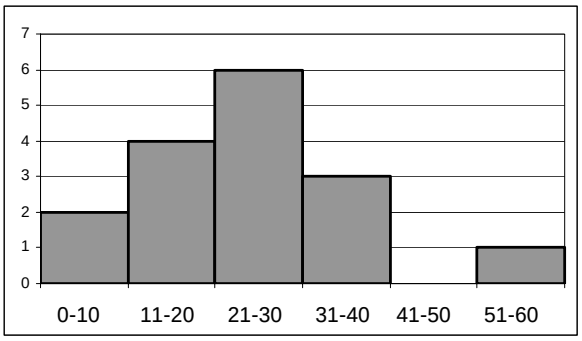


2. Which statistical measure is NOT evident on a box and whisker plot?

A. range
B. median
C. mode
D. all are evident

Use the information below to answer questions 3 and 4.

A police officer sat on the side of the road and monitored the speed of the traffic with a radar gun. The histogram below represents the speeds of the first sixteen cars to go by.



3. What was the range in speed of the cars?

A. 60 mph
B. 6 mph
C. 50 mph
D. cannot be determined from the graph

4. If the road the officer was monitoring was a school zone (speed limit of 20 mph), how many of those cars were speeding?

A. 10
B. 40
C. 6
D. not enough information