**Barron’s Let’s Review Regents – Algebra I**

# Chapter 13: Statistics

## 13.1 Measures of Central Tendency

A measure of central tendency is a number that summarizes a list of numbers. Since one number cannot describe everything about a set of data, there are many different measures of central tendency.

**(Sample) Mean**

The *mean* of a data set is also known as the arithmetic mean or the average. To calculate the mean, add up all the numbers in the data set and divide that result by the amount of numbers in the set. The symbol for the mean is .

**Mode**

The *mode* of a data set is the most “popular” number.

**Median**

The *median* is the “middle” number after the numbers have been arranged from least to greatest.

If the list has an even amount of numbers, add the two middle numbers and divide by 2.

**Quartiles**

The *median* is also called the *second quartile* since it is the number that is greater than 50% of the numbers in the set. The *first quartile* is the number that is greater than 25% of the numbers in the list The *third quartile* is the number that is greater than 75% of the numbers in the list.

**Interquartile Range**

The difference between the third quartile and the first quartile is called the *interquartile range*.

**Using the Graphing Calculator to determine the Maximum, Minimum, Median, First Quartile, and Third Quartile**

To calculate the maximum, minimum, median, first quartile (Q1), and third quartile (Q3) in Excel, use these formulas: =MAX(data\_range), =MIN(data\_range), =MEDIAN(data\_range), =QUARTILE.INC(data\_range,1), and =QUARTILE.INC(data\_range,3) respectively.

To calculate the population and sample mean in Excel, use the AVERAGE function for both, but be sure to use the correct data range: AVERAGE(data\_range) for sample mean and AVERAGE(population\_data\_range) for population mean.

1. Understanding the Concepts:

**Population Mean:**

The average of all values within the entire group or dataset you are analyzing.

**Sample Mean:**

The average of values within a subset or sample taken from the larger population.

**Calculating the Standard Deviation**

When the numbers in a data set are very close together, the data set is said to have a low standard deviation.

The is for the sample standard deviation, and the (symbol is sigma) is for the population standard deviation.

**Population Mean**

The population mean is denoted as (mu).

|  |  |
| --- | --- |
| N | = the size of the population |
|  | = each value from the population |
|  | = the population mean |

**Population Standard Deviation**

|  |  |
| --- | --- |
|  | = population standard deviation |
| N | = the size of the population |
|  | = each value from the population |
|  | = the population mean |

**Two-Way Frequency Tables**

A *two-way frequency table* is a good way to organize collected data from a survey. Instead of just displaying the results, the two-way frequency table is more specific, displaying how certain categories of people responded to the survey.

Interpreting a two-way frequency table based on a survey given to 40 menu and 60 women about whether or not math is their favorite subject.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Yes** | **No** | **Total** |
| Men | 32 | 8 | 40 |
| Women | 54 | 6 | 60 |
| Total | 86 | 14 | 100 |

**Creating a Two-Way Frequency Table**

**Creating a Relative Two-Way Frequency Table**

A relative two-way frequency table is much like a regular two-way frequency table except that all the numbers are recorded as percentages of the total.

To convert the table from the last section into a relative two-way frequency table, divide each number by the total, which is 12.

### Check Your Understanding of Section 13.1

1. Multiple-Choice
2. Find the mean of this set of numbers  
   { 4, 5, 8, 8, 8, 10, 10, 13, 15, 17, 23 }.  
   **(4) 11**
3. Find the median of this set of numbers  
   { 4, 5, 8, 8, 8, 10, 10, 13, 15, 17, 23 }.  
   **(2) 10**
4. Find the mode of this set of numbers  
   { 4, 5, 8, 8, 8, 10, 10, 13, 15, 17, 23 }.  
   **(4) 8**
5. Find the first quartile of this set of numbers  
   { 4, 5, 8, 8, 8, 10, 10, 13, 15, 17, 23 }.  
   **(1) 8**
6. Find the third quartile of this set of numbers  
   { 4, 5, 8, 8, 8, 10, 10, 13, 15, 17, 23 }.  
   **(2) 15**
7. Find the interquartile range of this set of numbers: { 4, 5, 8, 8, 8, 10, 10, 13, 15, 17, 23 }.  
   15 – 8 = 7  
   **(1) 7**
8. For the first fourd days of a five-day vacation, the mean temperature was 80 degrees. What must be the terperature be on the fifth day in order for the man temperature to be 82 degrees?  
   (4\*80)+x)/5 = 82  
   320+x = 410  
   x = 90  
   **(3) 90**
9. What is the median of the set of numbers   
   { 12, 4, 8, 3, 1, 4, 9, 5 }?  
   { 1, 3, 4, 4, 5, 8, 9, 12 }  
   (4 + 5)/2 = 4.5  
   **(4) 4.5**
10. In a set of seven numbers, the largest number is increased by 10. Which measure of central tendency must increase because of this?  
    **(1) Mean**
11. What is the sample standard deviation, rounded to the nearest hundredth of this data set: 20, 25, 28, 30, 32, 40?  
    Microsoft Excel: =STDEV.S(A1:A6)  
    **(2) 6.77**
12. *Show how you arrived at your answers*.
13. Seven numbers out of a set of nine numbers are 16, 17, 19, 19, 21, 21, and 25. If the mode of the eight numbers is 19 and the mean is 20, what are the other two numbers?  
    **One number will be 19** since it is the mode, and there are two instances of the number 21, so there must be three instances of 19 to be the mode.  
    16, 17, 19, 19, 19, 21, 21, and 25 and x  
    x + 157 = 9(20) = 180  
    x = 180-157 = **23**
14. Two basketball teams each have a mean height of 6 feet and a median height of 6 feet, but the interquartile range of the first team is much greater than the interquartile range of the second team. How can this be? Explain.  
      
    **The first team has a greater range of heights that average 6 feet, while the second team has a smaller range of heights that average 6 feet.**
15. Zabra says that on five tests she has a mode of 81 but a mean and median of 90. How is this possible?  
      
    **The mode represents the most “popular” number, so Zabra coud have scored an 81 on two tests, and that would be the most popular number.**  
      
    **To get a mean of 90, the five scores must total 450.  
    81 + 81 + 90 + x + y = 450  
    252 + x + y = 198  
    x + y = 198  
    x = 98, y = 100k (can’t both be 99, because that would change the mode)  
      
    To achieve a mean of 90, she could have had scores of 81, 81, 90, 98 and 100.**
16. What fraction of these 11 numbers is between the first quartile and the third quartile?  
    3, 17, 4, 8, 4, 9, 9, 15, 14, 19, 8  
    3, 4, 4, 8, 8, 9, 9, 14, 15, 17, 19  
      
    The amount of numbers between the first quartile and the third quartile is 5 numbers which is **.**
17. Lila has taken five algebra tests. Her median score is 90, and her mode score is 90, but her mean score is greater than 90. What is the highest her mean score can be?  
    90, 90, 90, 100, 100  
    Highest possible mean score: 94

## 13.2 Graphically Representing Data

When data are listed as just a series of numbers, man aspects of the data set are not clear. With a graphical representation of the data, certain information can be seen, even without doing any calculations. Three ways to represent data graphically are dot plots, histograms, and box plots. Each provides more information than a list of numbers would.

If the data set for the heights, in inches, of 11 professional men’s basketball players are:

76, 67, 76, 77, 72, 77, 80, 70, 84, 77, 80

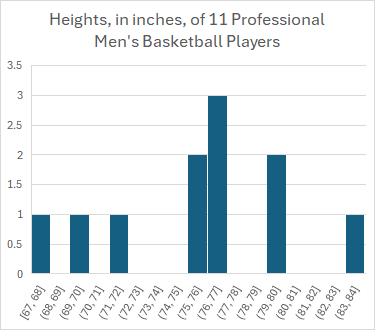
**Dot Plots**

In a *dot plot*, each piece of data is represented by a circle. When two data points represent the same value, they are stacked in a vertical line. With a dot plot, the mode can be easily determined. It is even possible to estimate the interquartile range from a dot plot.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  | ◯ |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  | ◯ | ◯ |  |  | ◯ |  |  |  |  |
| ◯ |  |  | ◯ |  | ◯ |  |  |  | ◯ | ◯ |  |  | ◯ |  |  |  | ◯ |
| 67 | 68 | 69 | 70 | 71 | 72 | 73 | 74 | 75 | 76 | 77 | 78 | 79 | 80 | 81 | 82 | 83 | 84 |

**Histograms**

The same data set can be represented graphically with a histogram. In a histogram, the data are represented by rectangles. Unlike a bar graph, in a histogram there are not spaces between bars representing the two 76s and the bar representing the three 77s.



**Box Plots**

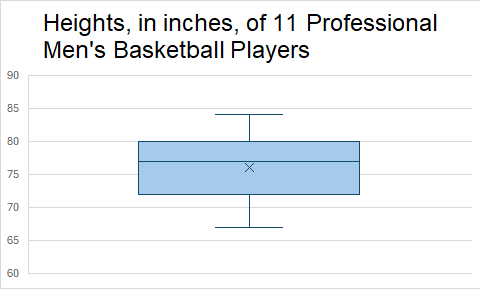
A box plot is a way to represent graphically the minimum value, the maximum value, the first quartile, the median, the third quartile, and the maximum value on the same diagram.

A screenshot of a graph

AI-generated content may be incorrect.

A screenshot of a graph

AI-generated content may be incorrect.

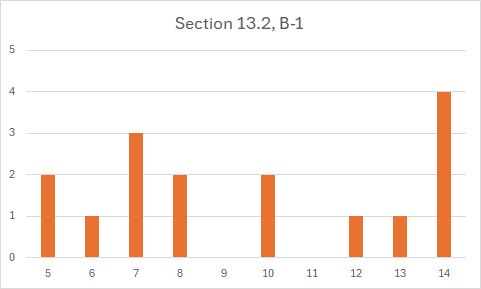


### Check Your Understanding of Section 13.2

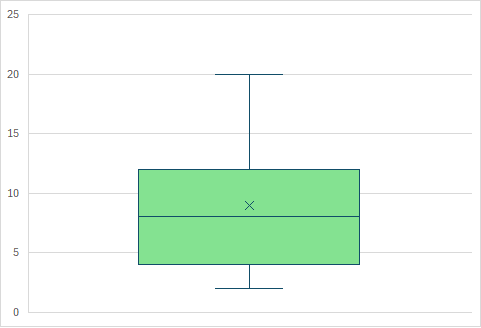
1. Multiple-Choice
2. What is the mode of the data in this dot plot?  
   **(2) 6**
3. What is the median of the data plot in this dot plot?  
   **(3) 5**
4. What is the mean of the data in this dot plot, rounded to the nearest hundredth?  
   **(2) 5.27**
5. What is the mode of the data in this histogram?  
   **(4) 16**
6. What is the median of the data in this histogram?  
   **(2) 35**
7. What is the median in this box plot?  
   **(1) 17**
8. What is the interquartile range of the data in this box plot?  
   **(3) 18**
9. Which box plot is based on data from this histogram?  
   **(4)**
10. What is true about this data set:   
    1, 2, 10, 11, 11?  
    Median: 10, Mean: 7, Mode: 11  
    **(1) The median is greater than the mean**
11. For which data set is the interquartile range equal to 0?  
    **(2)**
12. *Show how you arrived at your answers*.
13. Create a dot plot for the following set:  
    5,5,6,7,7,7,8,8,10,10,12,13,14,14,14,14  
    A screenshot of a table

    AI-generated content may be incorrect.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  | ◯ |
|  |  | ◯ |  |  |  |  |  |  | ◯ |
| ◯ |  | ◯ | ◯ |  | ◯ |  |  |  | ◯ |
| ◯ | ◯ | ◯ | ◯ |  | ◯ |  | ◯ | ◯ | ◯ |
| 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |



1. Create a box plot for the following set.  
   20,12,4,8,4,8,12,17,2,12,9,3,2,6,15  
   Average: 8.9, Median: 8, Mode: 12  
   A screenshot of a table

   AI-generated content may be incorrect.  
     
   
2. Two basketball teams each have a mean height of 72 inches and a median height of 72 inches. The box plots for the two teams are below. How is it that two teams with the same mean and median can have such different box plots?  
     
   The first team has members that are closer to the mean of 72 inches, and the second team has players more widely distributed and farther from the mean of 72 inches.  
     
   The first team has a first quartile of 70 inches and third quartile of 73 inches.  
     
   The second team has a first quartile of 60 inches and third quartile of 80 inches.
3. The box plot below is based on 11 numbers. Find two different sets of data that would produce this plot.  
   Range: 10 to 30, Median: 18  
   10, 12, 14, 16, 17, 18, 22, 24, 26, 28, 30  
   10, 11, 13, 15, 17, 18, 21, 23, 25, 27, 30
4. Make a histogram for this data set using 6 bars, each with a width of 5, using 0 to 5, including 5, but not 0, 5 to 10 including 10 but not 5, 10 to 15 including 15 but not 10, 15 to 20 including 15 but not 20, 20 to 25 including 25 but not 20, and 25 to 30 including 30 but not 25.  
     
   2, 2, 4, 6, 7, 7, 8, 8, 9, 11, 12, 12, 13, 16, 17, 17, 18, 18, 18, 19, 21, 21, 22, 22, 22, 23, 24, 28, 29, 29  
     
   A graph with numbers and a bar

   AI-generated content may be incorrect.  
     
   Using Microsoft Excel Analysis Pak add-in, using Data Analysis on Data tab, using Histogram  
     
   A graph on a sheet of paper

   AI-generated content may be incorrect.  
     
   