

T102 SECTION 8.2 MEASURES OF CENTRAL TENDENCY AND VARIATION

MEASURES OF CENTRAL TENDENCY

* Be sure to read all examples and do the "**NOW TRY THIS**" examples in the book.

In this section we will be doing some "number crunching" with data sets to help us to get "a feel" for the data:

The CENTER of the data. Where is the center of the data set? Do the numbers tend to converge about a certain value? Is there a single value that describes the data set? We use 3 main types of Measures of Central Tendency to answer the above:

- 1.
- 2.
- 3.

MEAN: The arithmetic mean of the numbers $x_1, x_2, x_3, \dots, x_n$ is denoted _____ and is calculated:
(This is commonly referred to as "the average")

We can think of the mean as the **balance point** of our data as they are placed on a line plot. (See page 526-527). The total distances from the mean on one side is equal to the total distances from the mean on the other side.

MEDIAN The value exactly in the _____ of the data when the data is arranged from least to greatest. In other words, half of the data will be less than the median and half will be greater. (Note: If there is an even number of data values, then the median is the average of the middle 2 data values.)

MODE The data value that appears _____, if there is one. Be aware that data sets can have more than one (or two) mode(s) or none at all.

EXAMPLES:

Find the mean, median, and mode of the following data set: 16, 11, 12, 10, 14

Find the mean, median, and mode of the following data set: 2, 5, 8, 4, 11, 8

Find the mean, median, and mode of the following data set: 60, 70, 95, 100, 60, 95

OBSERVATIONS: Which measures of central tendency were affected by extreme values?
Which were always members of the original set of data?
Which were sometimes/sometimes not?

MORE EXAMPLES INVOLVING CENTRAL TENDENCY:

Suppose a company employs 20 people. The president of the company earns \$200,000, the VP earns \$75,000, and the 18 employees each earn \$10,000. What is the mean? Is this the best number to represent the "average" salary of the company? What is the median and mode?

Suppose you own a hat shop and decide to order hats in only one size for the coming season. To decide which size to order, you look at last year's sales figures, which are itemized according to size. Should you find the mean, median or mode for the data? Why?

Students of Dr. Van Horn were asked to keep track of their own grades. One day, Dr. Van Horn asked the students to report their grades. One student had lost the papers but claims to remember the grades on 4 of the 6 assignments: 100, 82, 74, and 60. In addition, the student remembered that the mean of all 6 papers was 69 and the other 2 papers had identical grades. What were the grades on the other 2 papers?

Faith has an average of 76% of her 3 chemistry tests. What grade would she have to make on the 4th test to have an average of 80%?

A class of 23 students had a mean of 78 on a chemistry exam. The 10 boys in the class had a mean of 76.2. What was the mean of the girls' test scores? Round to 2 decimal places

GRADE POINT AVERAGE

Find the grade point average for the following report:

Assume that the grade point values are 4.0 for an A, 3.0 for a B, etc.

Course	Credits	Grade
T102	3	B
FYS	1	A
M333	2	C
P170	3	C
E241	3	A

What if this student put more effort into the P170 course and less in the FYS?

Course	Credits	Grade
T102	3	B
FYS	1	C
M333	2	C
P170	3	A
E241	3	A

BOX-AND-WHISKER PLOTS

We have been discussing measures of central tendency and measures of dispersion. We are going to elaborate on the concept of range and then use some of these new terms to construct another type of graph called **BOX-AND-WHISKER PLOTS**.

We know that the RANGE is: _____

A box-and-whisker plot is a way to display data visually and draw informal conclusions from the data. They are typically used instead of line plots when there is a large amount of data. They are very useful when comparing 2 sets of data.

These plots are visual representations of the five-number summary. The five numbers we need to construct a box-and-whisker plot are:

- 1.
- 2.
- 3.
- 4.
- 5.

From the construction of the plot, the center, the spread of the data, and the overall range of the data are immediately evident upon observation.

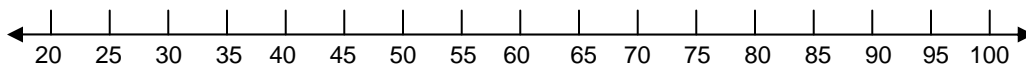
EXAMPLE:

Copy the information from the above example.

20 25 40 50 50 60 70 75 80 80 90 100 100

Let's construct a box-and-whisker plot for the above data.

- 1st: Draw a number line that extends from the least to the greatest value.
- 2nd: Draw a vertical line at Q_1 , Q_2 (the median), and Q_3 .
- 3rd: Connect these vertical lines with 2 horizontal lines to form a box.
- 4th: Draw segments from each end of the box to the extreme values to form the whiskers.



Sometimes there are **OUTLIERS** or values that are widely separated from the rest. Technically defined as any number that is more than 1.5 times the IQR above Q_3 or below Q_1 .

Identify any outliers. Then create a **box-and-whisker** plot for this data set.

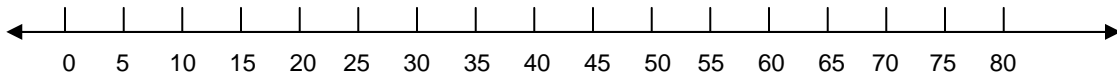
3 15 17 18 21 21 22 25 27 30 38 49 68

IQR = _____

$1.5 \cdot \text{IQR} =$ _____

$Q_1 - (1.5 \cdot \text{IQR}) =$ _____

$Q_3 + (1.5 \cdot \text{IQR}) =$ _____



Identify any outliers, then create a **box-and-whisker** plot for this data set.

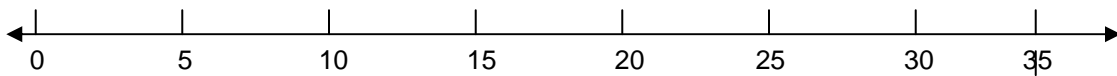
5 8 11 12 12 15 16 16 17 19 20 33

IQR = _____

$1.5 \cdot \text{IQR} =$ _____

$Q_1 - (1.5 \cdot \text{IQR}) =$ _____

$Q_3 + (1.5 \cdot \text{IQR}) =$ _____



The next two measures are very closely related and they are based on how far the data values lie from the mean.

VARIANCE The sum of the squares of the differences between each data value and the mean divided by the number of values in the data set, n .

V =

STANDARD DEVIATION

Denoted by the letter s , the standard deviation is the square root of the *variance*.

S =

EXAMPLE: Dr. Abel and Dr. Babel each taught a section of mathematics with 6 students. Both gave a final exam with the following results. Find the mean, median, and mode of each class. Construct back-to-back stem and leaf plots for the 2 data sets.

Dr. Abel's Scores	Dr. Babel's Scores
100	70
80	70
70	60
50	60
50	60
10	40
$\bar{x} =$	$\bar{x} =$
median	median
mode	mode

Now let's find the variance and standard deviation using the above formulas.

Dr. Abel's Scores			Dr. Babel's Scores		
x	$(x - \bar{x})$	$(x - \bar{x})^2$	x	$(x - \bar{x})$	$(x - \bar{x})^2$
100			70		
80			70		
70			60		
50			60		
50			60		
10			40		
$\Sigma=$	$\Sigma=$	$\Sigma=$	$\Sigma=$	$\Sigma=$	$\Sigma=$
mean			mean		
Variance			Variance		
Standard Deviation			Standard Deviation		

The standard deviation is a larger number when the values from a set of data are widely spread and a smaller number (close to 0) when the data values are close together.

EXAMPLE:

Given the following data: 85 70 68 94 90

Find the standard deviation.

EXAMPLE: Dr. Abel gave two group exams.

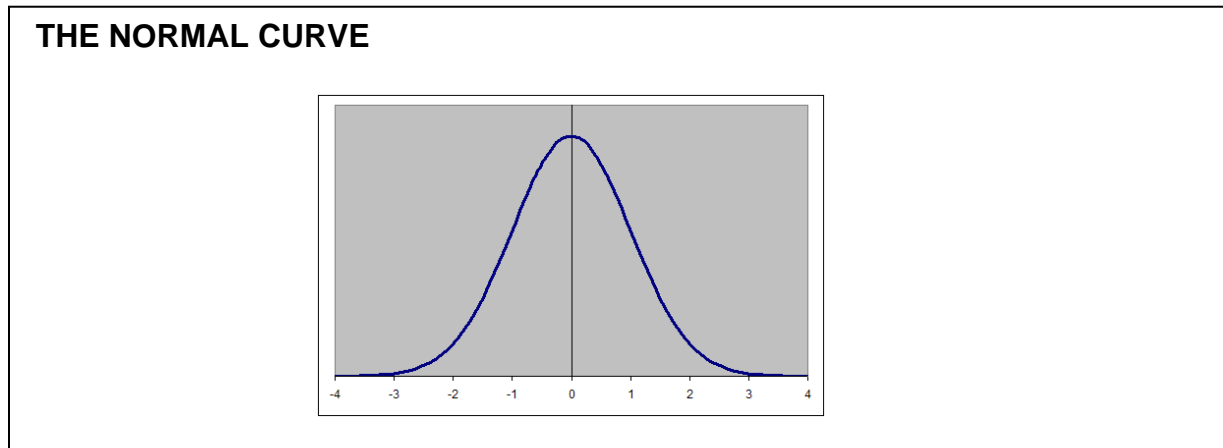
Exam A had grades of 0, 0, 0, 100, 100, 100
Exam B had grades of 50, 50, 50, 50, 50, 50

Find the mean, median, mode, range and standard deviation for both exams.

NORMAL DISTRIBUTION CURVE

To better understand how standard deviations are used as measures of dispersion, we next consider normal distributions. The graphs of normal distributions are the _____ curves that we call _____ curves. These curves are used when describing IQ scores or standardized test scores.

A normal curve is a smooth bell-shaped curve that depicts the frequency of the data values symmetrically about the mean. (NOTE: When using a normal curve, the mean, median and mode all have the same value.)



When discussing the normal curve, we talk about “how many standard deviations a particular score is from the mean” .

Within _____ standard deviation of the mean lie 68% of the data values.
This means 34% to the left and 34% to the right of the mean.

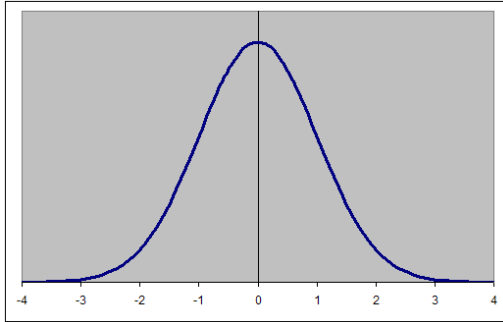
Within _____ standard deviations of the mean lie 95% of the data values.
This means _____% to the left and _____% to the right of the mean.

Within _____ standard deviations of the mean lie 99.8% (or 99.7% in some texts) of the data values.
This means _____% to the left and _____% to the right of the mean.

These percentages represent approximations of the total percent of area under the curve.

EXAMPLE

Psychologists claim that the mean IQ is 100 and the standard deviation is 15. They also claim that an IQ score of over 130 represents a superior score. What percent of the population of the US fall into this "superior" category?

**PERCENTILES**

We have all taken standardized examination such as the SAT or the ACT. When given your score, you are also given a percentile. The percentile shows your score in relation to all the other scores. If you are in the 92nd percentile this means that approximately _____% of those taking the exam scored lower than you and _____% score higher than you.

My friend Karen called the other day. She was upset because her child scored in the 60th percentile on her SAT. She said, "When I was in school, 60% was a D or an F. She needs to do MUCH BETTER THAN THAT!"

a. Is her child's score:

BETTER THAN AVERAGE OR AVERAGE OR BELOW AVERAGE?

b. What would you tell her about being in the 60th percentile?