Notes: Standard Deviation and the Normal Model

Standard deviation is a measure of spread, or _______________. The smaller the standard deviation, the __________ variability is present in the data. The larger the standard deviation, the __________ variability is present in the data.

Standard deviation can be used as a ruler for measuring how an individual compares to a __________.

To measure how far above or below the mean any given data value is, we find its ______________________________, or ______________________________.

$$z =$$

To standardize a value, subtract the __________and divide by the _______________________.

Measure your height in inches. Calculate the standardized value for your height given that the average height for women is 64.5 inches with a standard deviation of 2.5 inches and for men is 69 inches with a standard deviation of 2.5 inches. *Are you tall?*

$$z_{height} =$$

Suppose the average woman's shoe size is 8.25 with a standard deviation 1.15 and the average male shoe size is 10 with a standard deviation of 1.5. *Do you have big feet?*

$$z_{shoe} =$$

Suppose Sharon wears a size 9 shoe and Andrew wears a size 9. *Does Sharon have big feet? Does Andrew?*

$z_{Sharon} =$ $\qquad\qquad\qquad$ $z_{Andrew} =$

In order to compare values that are measured using different scales, you must first _______________the values. The standardized values have no _______________and are called _______________. Z-scores represent how far the value is above the __________ (if _______________) or below the ___________(if _______________).

Example: z = 1 means the value is ________standard deviation ____________the mean
z = -0.5 means the value is ________of a standard deviation __________the mean

The _______________the z-score, the more unusual it is.

Standardized values, because they have no units, are therefore useful when comparing values that are measured on different _______________, with different _______________, or from different _______________.

Adding a constant to all of the values in a set of data adds the same constant to the measures of ________________. It does not, however, affect the ________________.

Example: Add 5 to each value in the given set of data (on the left) to form a new set of data (on the right). Then find the indicated measures of center and spread.

{5, 5, 10, 35, 45}.	**{____, ____, ____, ____, ____}.**
Center:	Center:
$\overline{x}$ =	$\overline{x}$ =
M =	M =
Mode =	Mode =
Spread:	Spread:
Range =	Range =
IQR =	IQR =
SD =	SD =

Multiplying a constant to all of the values in a set of data multiplies the same constant to the measures of ________________and ________________.

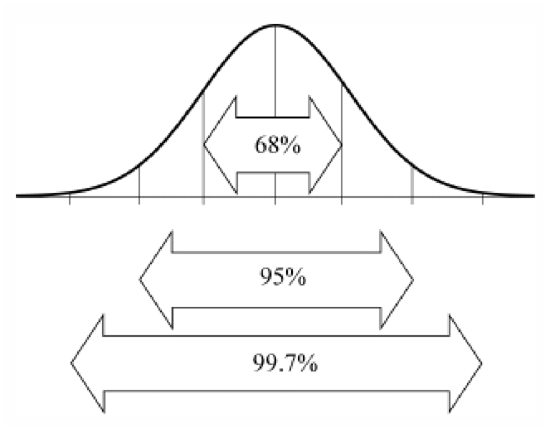
Example: Multiply each value in the given set of data (on the left) by 2 to form a new set of data (on the right). Then find the indicated measures of center and spread.

{5, 5, 10, 35, 45}.	**{____, ____, ____, ____, ____}.**
Center:	Center:
$\overline{x}$ =	$\overline{x}$ =
M =	M =
Mode =	Mode =
Spread:	Spread:
Range =	Range =
IQR =	IQR =
SD =	SD =

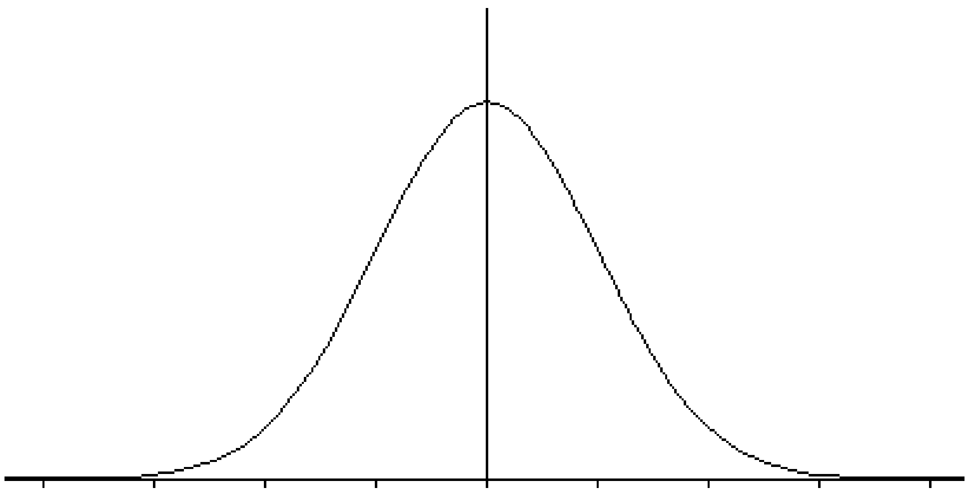
By standardizing values, we shift the distribution so that the mean is _____, and rescale it so that the standard deviation is _____. Standardizing does not change the ________________of the distribution.

The Normal model:

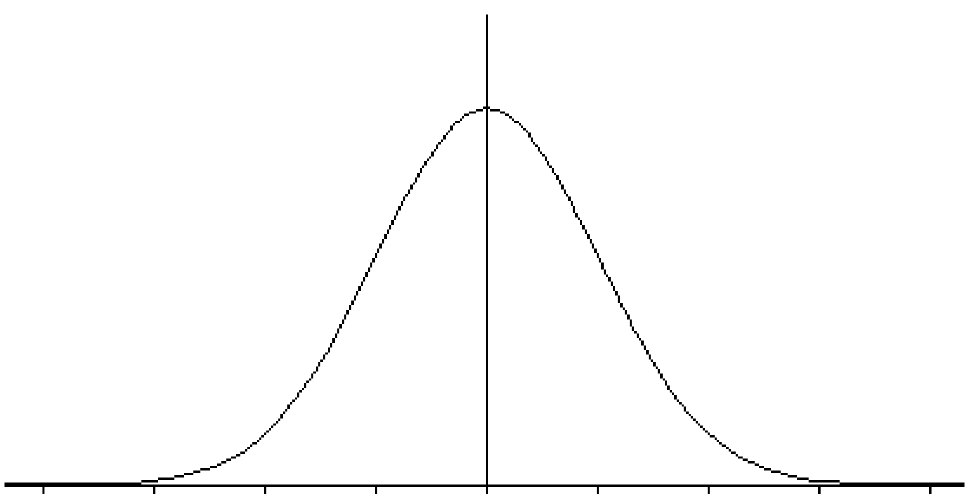
- is ________________and ________________.
- follows the ________________________________
  - About _____ of the values fall within ___________ standard deviation of the mean.
  - About _____ of the values fall within ___________standard deviations of the mean.
  - About _____ (almost all) of the values fall within ___________standard deviations of the mean.



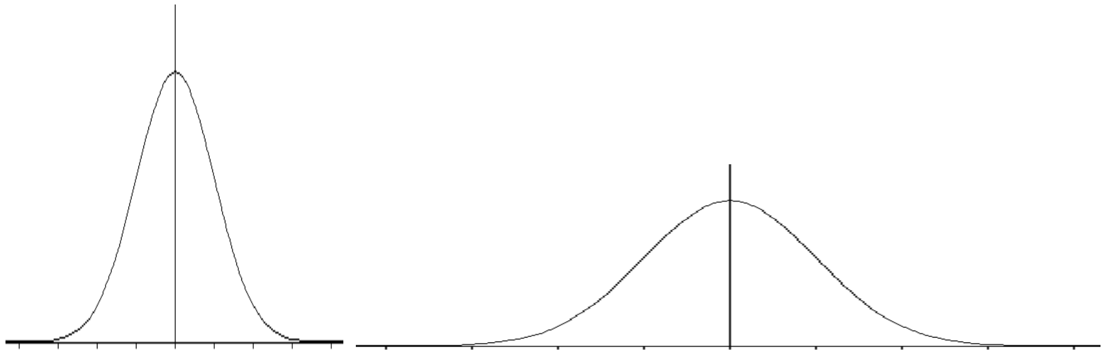
The standard Normal model has mean _____ and standard deviation _____.



The Normal model is determined by __________and __________. We use the Greek letters sigma and mu because this is a __________; it does not come from actual __________. Sigma and mu are the _______________that specify the model.



The larger sigma, the ___________spread out the normal model appears. The inflection points occur a distance of ___________on either side of ___________.



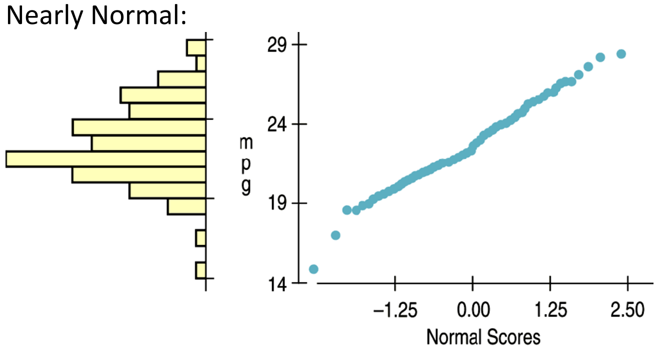
To standardize Normal data, subtract the __________ (_____) and divide by the ______________________ (___________).

$$z =$$

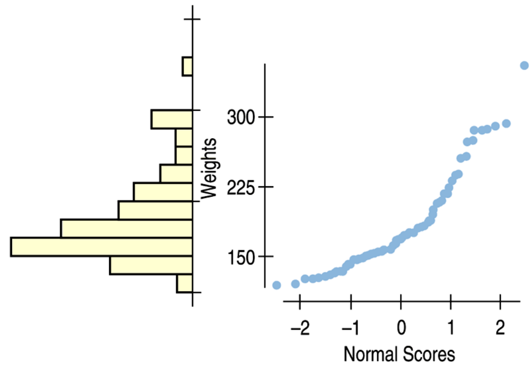
To assess normality:

- Examine the ___________of the histogram or stem-and-leaf plot. A normal model is ___________about the mean and ________________.
- Compare the mean and median. In a Normal model, the mean and median are ___________.
- Verify that the ______________________holds.
- Construct a ____. If the graph is linear, the model is Normal.

Nearly Normal:



Skewed distribution:



Notes: Standard Deviation and the Normal Model

Standard deviation is a measure of spread, or **variability**. The smaller the standard deviation, the **less** variability is present in the data. The larger the standard deviation, the **more** variability is present in the data.

Standard deviation can be used as a ruler for measuring how an individual compares to a **group**.

To measure how far above or below the mean any given data value is, we find its **standardized value**, or **z-score**.

$$z = \frac{y - \overline{y}}{s}$$

To standardize a value, subtract the **mean** and divide by the **standard deviation**.

Measure your height in inches. Calculate the standardized value for your height given that the average height for women is 64.5 inches with a standard deviation of 2.5 inches and for men is 69 inches with a standard deviation of 2.5 inches. *Are you tall?*

$$z_{height} =$$

Suppose the average woman's shoe size is 8.25 with a standard deviation 1.15 and the average male shoe size is 10 with a standard deviation of 1.5. *Do you have big feet?*

$$z_{shoe} =$$

Suppose Sharon wears a size 9 shoe and Andrew wears a size 9. *Does Sharon have big feet? Does Andrew?*

$z_{Sharon} =$ $\qquad\qquad$ $z_{Andrew} =$

In order to compare values that are measured using different scales, you must first **standardize** the values. The standardized values have no **units** and are called **z-scores**. Z-scores represent how far the value is above the mean (if **positive**) or below the mean (if **negative**).

Ex: z = 1 means the value is **one** standard deviation **above** the mean
z = -0.5 means the value is **one-half** of a standard deviation **below** the mean

The **larger** the z-score, the more unusual it is.

Standardized values, because they have no units, are therefore useful when comparing values that are measured on different **scales**, with different **units**, or from different **populations**.

Adding a constant to all of the values in a set of data adds the same constant to the measures of **center** and percentiles. It does not, however, affect the **spread**.

Example: Add 5 to each value in the given set of data (on the left) to form a new set of data (on the right). Then find the indicated measures of center and spread.

{5, 5, 10, 35, 45}.	**{10, 10, 15, 40, 50}.**
Center:	Center:
$\bar{x}$ =	$\bar{x}$ =
M =	M =
Mode =	Mode =
Spread:	Spread:
Range =	Range =
IQR =	IQR =
SD =	SD =

Multiplying a constant to all of the values in a set of data multiplies the same constant to the measures of **center** and **spread**.

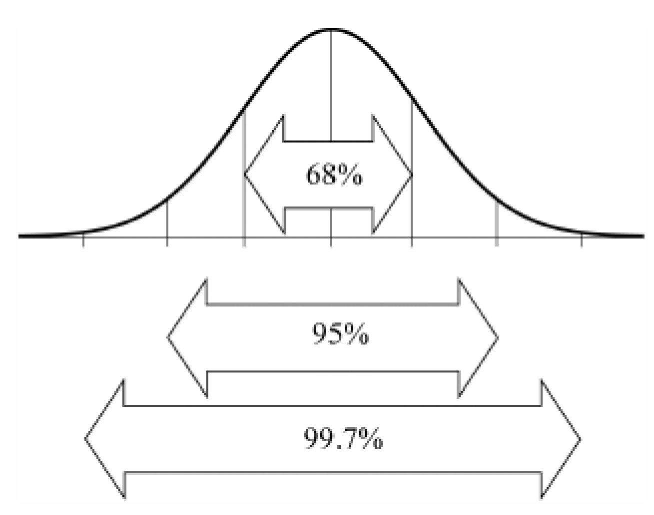
Example: Multiply each value in the given set of data (on the left) by 2 to form a new set of data (on the right). Then find the indicated measures of center and spread.

{5, 5, 10, 35, 45}.	**{10, 10, 20, 70, 90}.**
Center:	Center:
$\bar{x}$ =	$\bar{x}$ =
M =	M =
Mode =	Mode =
Spread:	Spread:
Range =	Range =
IQR =	IQR =
SD =	SD =

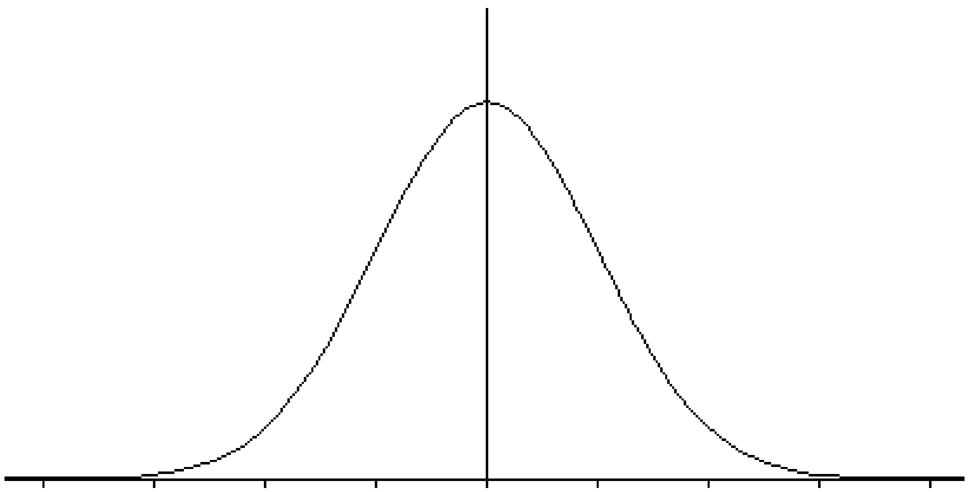
By standardizing values, we shift the distribution so that the mean is **0**, and rescale it so that the standard deviation is **1**. Standardizing does not change the **shape** of the distribution.

The Normal model:

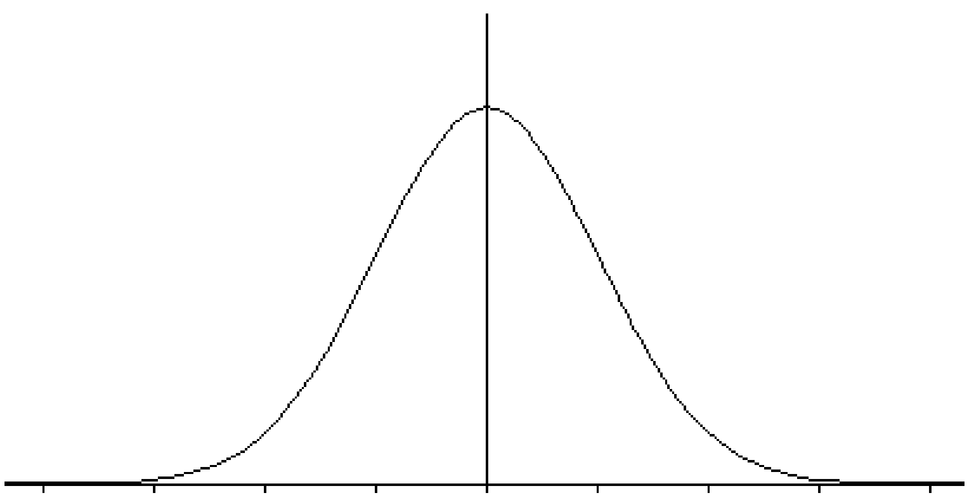
- is **symmetric** and **bell-shaped**.
- follows the **68-95-99.7 Rule**
  - About **68%** of the values fall within **one** standard deviation of the mean.
  - About **95%** of the values fall within **two** standard deviations of the mean.
  - About **99.7%** (almost all) of the values fall within **three** standard deviations of the mean.



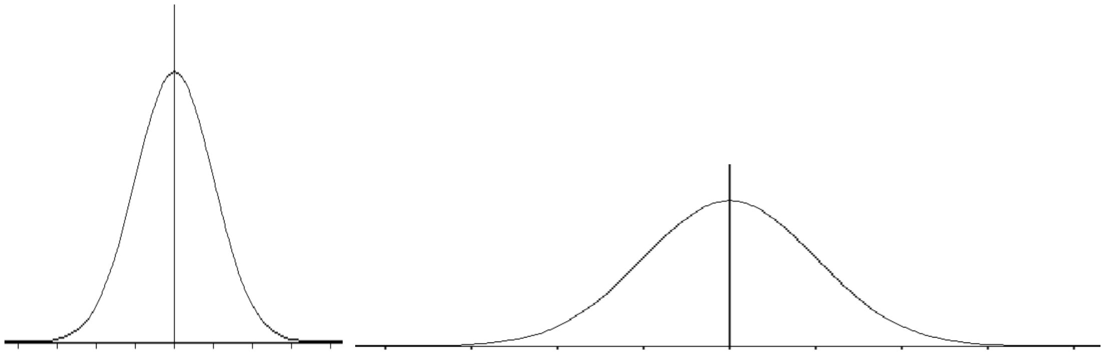
The standard Normal model has mean **0** and standard deviation **1**.



The Normal model is determined by **sigma** and **mu**. We use the Greek letters sigma and mu because this is a **model**; it does not come from actual **data**. Sigma and mu are the **parameters** that specify the model.



The larger sigma, the **more** spread out the normal model appears. The inflection points occur a distance of **sigma** on either side of **mu**.



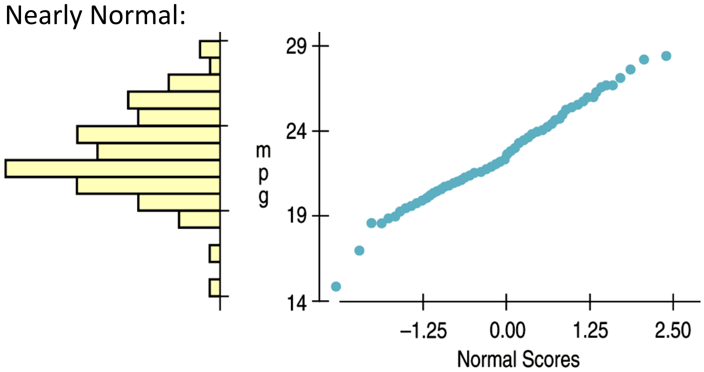
To standardize Normal data, subtract the **mean** (**mu**) and divide by the **standard deviation** (**sigma**).

$$z = \frac{y - \mu}{\sigma}$$

To assess normality:

- Examine the **shape** of the histogram or stem-and-leaf plot. A normal model is **symmetric** about the mean and **bell-shaped**.
- Compare the mean and median. In a Normal model, the mean and median are **equal**.
- Verify that the **68-95-99.7 Rule** holds.
- Construct a **normal** **probability** **plot**. If the graph is linear, the model is Normal.

Nearly Normal:



Skewed distribution: