**Transforming data**

Transforming converts the original observations to another scale. Transformations can affect the shape, center, and spread of a distribution

**Effect of Adding (or Subracting) a Constant**

* Measures of center \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* Shape \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* Measures of spread \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Example: A math test has a mean of 60 and a standard deviation of 12. The teacher wants the class average higher so he adds 15 points to every grade. What is the mean and standard deviation of the new grades?

**Effect of Multiplying (or Dividing) by a Constant b**

* Measures of center \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* Shape \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* Measures of spread \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Example: A math test has a mean of 60 and a standard deviation of 12. The teacher wants the class average higher so he multiplies every grade by 1.25. What is the mean and standard deviation of the new grades?

1. Suppose the class took a 40-point quiz. Results show a mean score of 30,

median 32, IQR 8, standard deviation 6, min 12, and Q1 27. (Suppose

YOU got a 35.)

What happens to each of the statistics if…

• I decide to weight the quiz as 50 points, and will add 10 points to every

score. Your score is now 45.

• I decide to weight the quiz as 80 points, and double each score. Your score

is now 70.

• I decide to count the quiz as 100 points; I'll double each score and add 20

points. Your score is now 90.

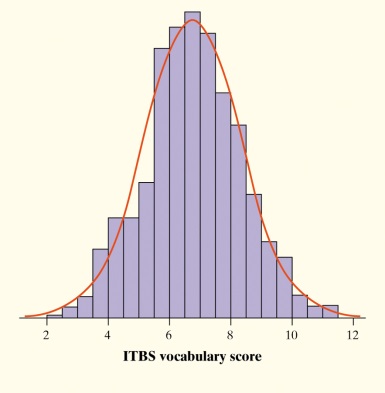
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Statistic | Original (y) | y+10 | 2y | 2y+20 |
| Mean | 30 |  |  |  |
| Median | 32 |  |  |  |
| IQR | 8 |  |  |  |
| Standard Dev | 6 |  |  |  |
| Minimum | 22 |  |  |  |
| Q1 | 27 |  |  |  |
| Your score | 35 |  |  |  |

**Density Curves**

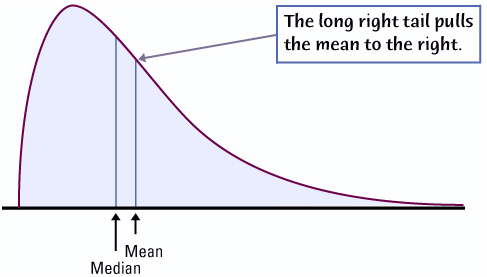
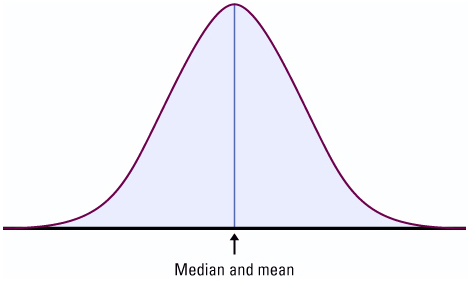
A **density curve** is a curve that

* + is always on or above the horizontal axis, and
  + has area exactly 1 underneath
  + the area under the curve is the proportion of all observations that fall in that interval.

Sometimes the overall pattern of a large number of observations is so regular that we can describe it by a smooth curve.

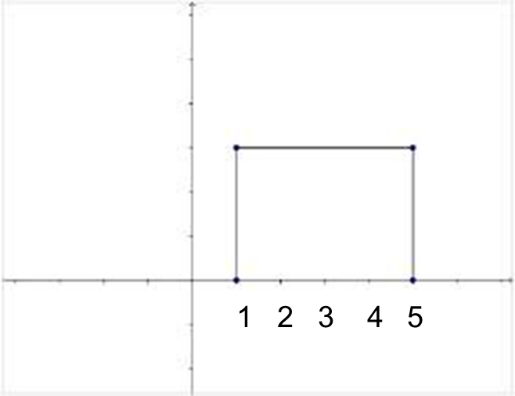


* Density Curves come in many different shapes; **symmetric, skewed, uniform, etc.**
* The area of a region of a density curve represents the % of observations that fall in that region.
* The median of a density curve cuts the area in **half.**
* The mean of a density curve is its **“balance point.”**



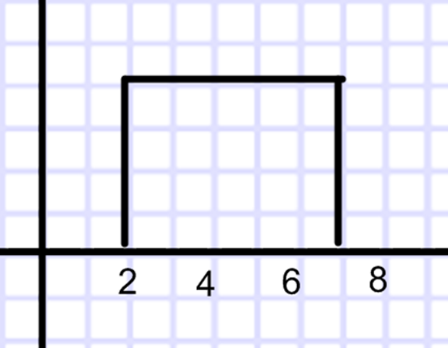


**Uniform Density Curves**

****

The height of the uniform density curve is:

1. What proportion of values are less than 2?
2. What proportion of values are greater than 3.5?
3. What proportion of values are between 2.5 and 4?



The height of the uniform density curve is:

1. What proportion of values are less than 3?
2. What proportion of values are greater than 3.5?
3. What proportion of values are between 2.5 and 5?