2.1 Describing Location in a Distribution

HW: p. 105 (1, 5, 9-15 odd, 19-23 odd, 31, 33-38)

Measuring Position

Percentiles

- A common way to measure ______________ within a distribution is to tell what ____________ of observations fall ________ the value in question.
- The **pth percentile** of a distribution is the value with *p* percent of the observations _____ ______ it.

To find an individual's percentile:

Percentiles should be __________ numbers, so make sure to always round to the nearest _____________.

Example: *Wins in Baseball*

The stemplot shows the number of wins for each of the 30 Major League Baseball teams in 2009. Find the percentiles for the following teams:

a) The Colorado Rockies, who won 92 games.

b) The New York Yankees, who won 103 games.

c) The Kansas City Royals and Cleveland Indians, who both won 65 games.

5	9
6	2455
7	00455589
8	0345667778
9	123557
10	3

Key: 5|9 represents a team with 59 wins.

Cumulative Relative Frequency Graphs

Cumulative relative frequency graphs, or _________, provide a _______________ tool to find percentiles in a distribution.

To graph:

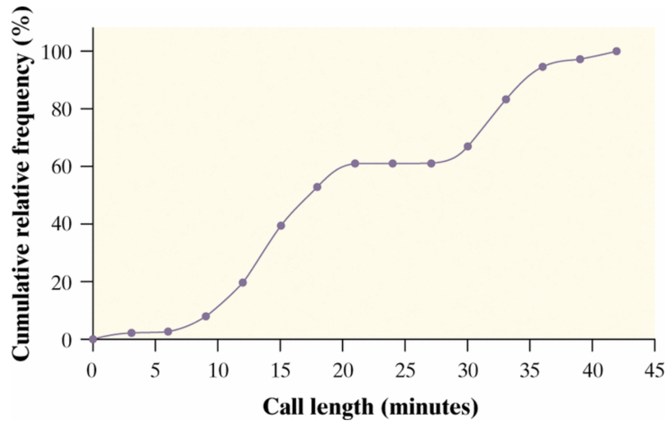
- Start at 0.
- Plot each cumulative relative frequency.
- Connect each point with a segment.
- Label your axes!!!!

Check Your Understanding, *p. 89*

1. Mark receives a score report detailing his performance on a statewide test. On the math section, Mark earned a raw score of 39, which placed him at the 68th percentile. This means that
   a) Mark did better than about 39% of the students who took the test.
   b) Mark did worse than about 39% of the students who took the test.
   c) Mark did better than about 68% of the students who took the test.
   d) Mark did worse than about 68% of the students who took the test.
   e) Mark got fewer than half of the questions correct on this test.

2. Mrs. Munson is concerned about how her daughter's height and weight compare with those of other girls of the same age. She uses an online calculator to determine that her daughter is at the 87th percentile for weight and the 67th percentile for height. Explain to Mrs. Munson what this means.

The graph displays the cumulative relative frequency of the lengths of phone calls made from the mathematics department office at Gabalot High last month.



3. About what percent of calls lasted less than 30 minutes? 30 minutes or more?

4. Estimate Q_1, Q_2 and the IQR

z-scores

- In order to accurately describe a distribution, we must consider ________ and ___________.
- The ________________ value, or *z*-score, of an observation takes into account both of those measures.
- The *z*-score tells us how many standard deviations ___________ or ___________ the mean a particular observation falls.
- It allows us to describe the _______________ of an individual in a distribution, and also allows us to ______________ ________________ of individuals in different distributions.

Some Tips

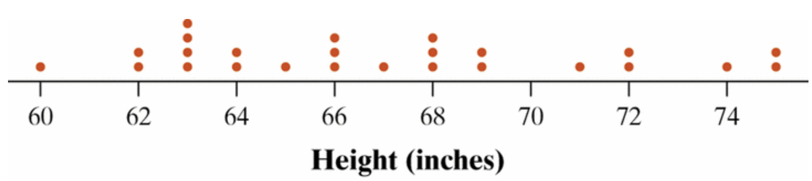
- A *z*-score is _____ measured in the ________ ________ as the variable.
- A *z*-score is ___________________. If it is positive, it means the observation is above the average. If it is negative, the observation is below the average.
- When _______________________ *z*-scores from different distributions, it is important that the distributions be roughly the _________ ______________.

How to find a *z*-score

If *x* is an observation from a distribution that has a known mean and standard deviation, the **standardized value** of *x* is:

Check Your Understanding, *p. 91*

Mrs. Navard's statistics class has just completed the first three steps of the "Where Do I Stand?" Activity (page 84). The figure below shows a dotplot of the class's height distribution, along with summary statistics from computer output.



Variable	n	$\bar{x}$	s_x	Min	Q_1	M	Q_3	Max
Height	25	67	4.29	60	63	66	69	75

1. Lynette, a student in the class, is 65 inches tall. Find and interpret her *z*-score.

2. Another student in the class, Brent, is 74 inches tall. How tall is Brent compared with the rest of the class? Give appropriate numerical evidence to support your answer.

3. Brent is a member of the school's basketball team. The mean height of the players on the team is 76 inches. Brent's height translates to a *z*-score of –0.85 in the team's height distribution. What is the standard deviation of the team members' heights?

Transforming Data

When we find *z*-scores, we are actually __________________ our data to a _________________ scale.

Sometimes we transform data to _________ between measurement _________. When we do this, it is important to know what happens to the ___________ and _____________ of the transformed distribution.

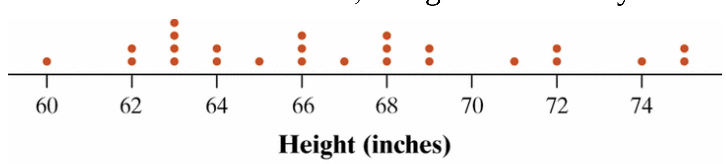
The **center** is affected by:

The **spread** is affected by:

*Shape and location of observations will remain __________________!

Example: *Taxi Cabs*
In 2010, taxi cabs in New York City charged an initial fee of $2.50 plus $2 per mile. In equation form, $fare = 2.50 + 2(miles)$. At the end of a month, a businessman collects all his taxi cab receipts and calculates some numerical summaries. The mean fare he paid was $15.45, with a standard deviation of $10.20. What are the mean and standard deviation of the lengths of his cab rides in miles?

Check Your Understanding, *p. 97*
The figure below shows a dotplot of the height distribution for Mrs. Navard's class, along with summary statistics from computer output.



Variable	n	$\bar{x}$	s_x	Min	Q_1	M	Q_3	Max
Height	25	67	4.29	60	63	66	69	75

1. Suppose that you convert the class's heights from inches to centimeters (1 inch = 2.54 cm). Describe the effect this will have on the shape, center, and spread of the distribution.

2. If Mrs. Navard had the entire class stand on a 6-inch-high platform and then had the students measure the distance from the top of their heads to the ground, how would the shape, center, and spread of this distribution compare with the original height distribution?

3. Now suppose that you convert the class's heights to *z*-scores. What would be the shape, center, and spread of this distribution? Explain.

Density Curves

When we have a __________ number of observations, we can describe the overall ___________ using a smooth __________ called a density curve.

A density curve:

- is always _____________ the horizontal axes
- has area exactly ____ underneath it.

No real set of data is exactly described by a density curve. It is an ______________________ that is easy to use and accurate enough for ___________________ use.

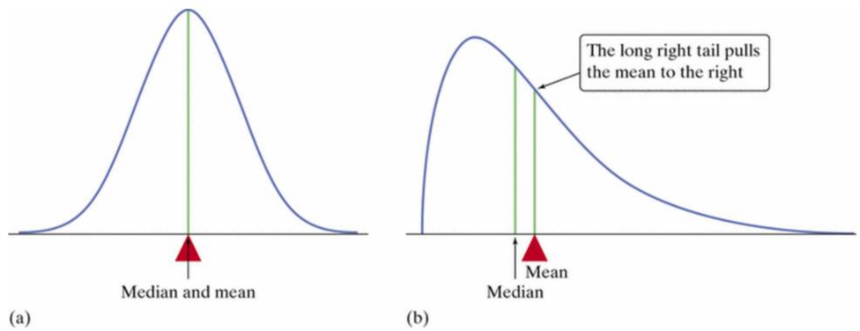
Under the Curve

Areas under a density curve represent proportions of the total number of observations.

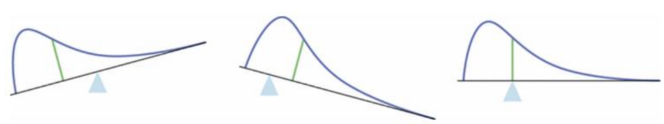
The median is the "equal area" point, because half of the area is to the left and half is to the right.

In a symmetric distribution, the mean and median will be the same.

In a skewed distribution, the mean will be pulled away from the median in the direction of the tail.



The mean is also the "balance point" of the distribution.



Check Your Understanding, *p. 103*

Use the figure shown to answer the following questions.

1. Explain why this is a legitimate density curve.

2. About what proportion of observations lie between 7 and 8?

3. Mark the approximate location of the mean and median.

4. Explain why the mean and median have the relationship that they do in this case.