

Section 3.4

Measures of Position and Outliers

Objectives

- Determine and interpret z-scores
- Interpret percentiles
- Determine and interpret quartiles
- Determine and interpret the interquartile range
- Check a set of data for outliers

The mean height of males 20 years or older is 69.1 inches with a standard deviation of 2.8 inches. The mean height of females 20 years or older is 63.7 inches with a standard deviation of 2.7 inches. Data is based on information obtained from National Health and Examination Survey.

Who is relatively taller?

Kevin Garnett whose height is 84 inches

or

Margo Dydek whose height is 86 inches

z-score
above or
below the mean
X - individual

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

μ - Pop Mean
 σ - Pop S.D.

Kevin

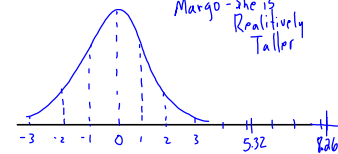
$$z = \frac{84 - 69.1}{2.8}$$

$$z = 5.32$$

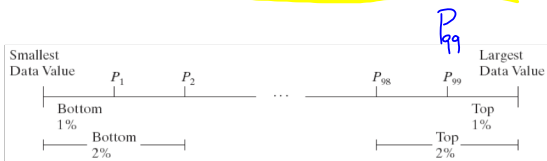
Margo

$$z = \frac{86 - 63.7}{2.7}$$

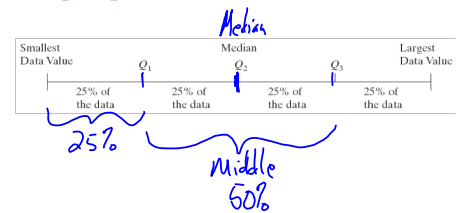
$$z = 8.26$$



The k th percentile, denoted, P_k , of a set of data is a value such that k percent of the observations are less than or equal to the value.



Quartiles divide data sets into fourths, or four equal parts.



A group of Brigham Young University—Idaho students collected data on the speed of vehicles traveling through a construction zone on a state highway, where the posted speed was 25 mph. The recorded speed of 14 randomly selected vehicles is given below:

20, 24, 27, 28, 29, 30, 32, 33, 34, 36, 38, 39, 40, 40

Find and interpret the quartiles for speed in the construction zone.

Min	Q_1	Median	Q_3	Max
20	28	32.5	38	40

 → L_1 , 1-Var Stats

Middle 50%
28 - 38

What % speed more 38mph
25%

The interquartile range, IQR, is the range of the middle 50% of the observations in a data set. That is, the IQR is the difference between the third and first quartiles and is found using the formula

$$IQR = Q_3 - Q_1$$

$$IQR = 38 - 28 = 10$$

↑
Spread
When data skewed/outliers

Suppose a 15th car travels through the construction zone at 100 miles per hour. How does this value impact the mean, median, standard deviation, and interquartile range?

	Without 15 th car	With 15 th car
Mean	32.1 mph	36.7 mph
Median	32.5 mph	33 mph
Standard deviation	6.2 mph	18.5 mph
IQR	10 mph	11 mph

Symmetric (handwritten) points to the first two rows. *Skewed* (handwritten) points to the last three rows.

Summary: Which Measures to Report		
Shape of Distribution	Measure of Central Tendency	Measure of Dispersion
Symmetric	Mean	Standard deviation
Skewed left or skewed right	Median	Interquartile range

Checking for Outliers:

Determine the fences. Fences serve as cutoff points for determining outliers.
 Lower Fence = $Q_1 - 1.5(IQR)$ $IQR = 39 - 28$
 Upper Fence = $Q_3 + 1.5(IQR)$

Upper fence = 39 + 1.5(11)
Above 55.5
is an outlier
100 mph speeder an outlier

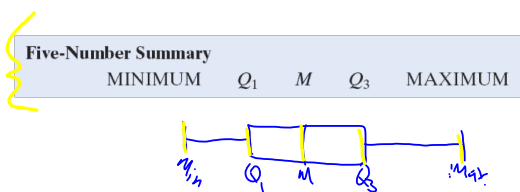
Section 3.5

The Five-Number Summary and Boxplots

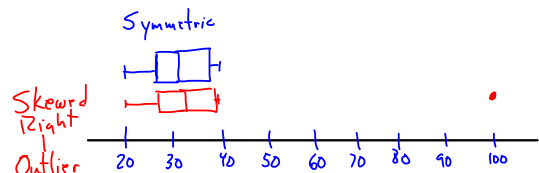
Objectives

1. Compute the five-number summary
2. Draw and interpret boxplots

The five-number summary of a set of data consists of the smallest data value, Q_1 , the median, Q_3 , and the largest data value. We organize the five-number summary as follows:



Draw a box plot of the previous data. Interpret what you see.



Use a boxplot and quartiles to describe the shape of a distribution.

