Warm-up

What are the mean and median of the set 4, 5, 5, 6, 6, 7, 7, 7, 9, 30?

* Without a calculator

$\bar{x} = 8.6 = \frac{86}{10}$

median = 6.5

AP Statistics

Chapter 1: Exploring Data

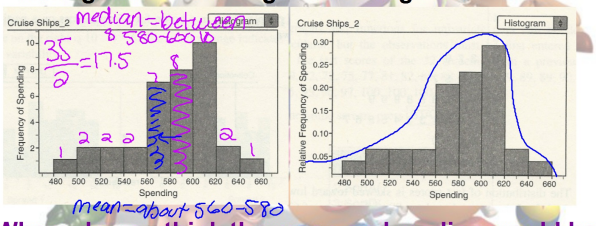
Day 3

HW: p. 70-74, #79, 81, 87, 98A, 107, 109

Histograms

Blue Green

- Used for continuous quantitative data, but may be used for discrete data that have a wide spread.
- The horizontal axis is broken into intervals. (bin)
- Histograms are also good for large data sets.



Where do you think the mean and median would be?

Mean

The mean is the average of a data set.

ch.1 mean — sum — individual observations

$$\bar{x} = \frac{\sum x_1}{n}$$

Total # of observations

- n = the number of observations in the data set
- x_1 represents individual observations counting from 1 to n.
- The summation symbol, $\sum$, is the command to add up the expression following it.

Measuring Spread

mean

* For data that are fairly symmetric and not affected by outliers, standard deviation and variance are useful measures of variability.

variance = (stand. dev)^2

* median

For data sets that have extreme values or skewness, the interquartile range would be a better measure of variability.

This is because standard deviation and variance are computed with the actual values of the data, like the mean.

Interquartile range, like the median, is not.

Variance

The variance is a numerical value used to indicate how widely individuals in a group vary.

If individual observations vary greatly from the group mean, the variance is big; and vice versa.

The variance is calculated with the formula:

$$s^2 = \frac{1}{n-1}\sum(x_1 - \bar{x})^2$$

Standard Deviation

The standard deviation is can be thought of as the

* typical distance an observation lies from the mean.

The more spread out the data is from the mean, the larger the standard deviation and vise versa.

The standard deviation is calculated with the formula:

Chapt. 7 Standard dev.

$$s = \sqrt{\frac{1}{n-1}\sum(x_1 - \bar{x})^2}$$

sum; mean; total # of observations; individual observations

Standard Deviation

Which data set would have the smallest standard deviation? Largest deviation?

A	B	C	D
$s_x = 1.13$, $\bar{x} = 2.42$	$s_x = 2.16$	$s_x = 2.14$	$s_x = 1.81$
	5	5.29	4.43
smallest?	largest		
2	2	2	2
2	3	4	4
2	4	4	4
5	5	5	4
2	6	7	4
2	7	7	5
2	8	8	8

$s^2 = \frac{1}{n-1}\sum(x_1 - \bar{x})^2$

Example

$s = \sqrt{\frac{1}{n-1}\sum(x_1 - \bar{x})^2}$

The average monthly rainfall in inches in Birmingham, England, is shown in the table below.

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2.3	1.9	2.1	1.8	2.2	2.2	2.0	2.8	2.2	2.1	2.5	2.6

$\bar{x} = 2.225$

Compute the variance and standard deviation of the monthly rainfall.

$S_x = \frac{1}{12-1}\sum (2.3-2.225)^2 + (1.9-2.225)^2 \ldots (2.6-2.225)^2 = 0.290$

$S^2 = (0.290)^2 = 0.084$