

**Standard Deviation:** A measure of spread for symmetric data. The “average deviation” or distance from the mean. The symbol for the population standard deviation is  $\sigma$ . We use  $s$  to represent the standard

deviation of a sample. 
$$s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}}$$

**Example 1.** The table below shows a sample of 10 MTH 243 student’s heart rates, measured in beats per minute (bpm). The mean of this set is 72. Calculate the deviation and squared deviation by hand. Then calculate the variance and the standard deviation.

Heart Rate (in bpm)	Deviation from the mean	Squared Deviation
52		
65		
67		
68		
70		
73		
74		
78		
81		
92		
Sum of the Squared Deviations		

Variance:  $s^2 =$

Standard Deviation:  $s =$

**Example 2.** Comparing Standard Deviations

For a, b, and c, examine the two data sets. Without doing any calculations, which set has the higher standard deviation and why?

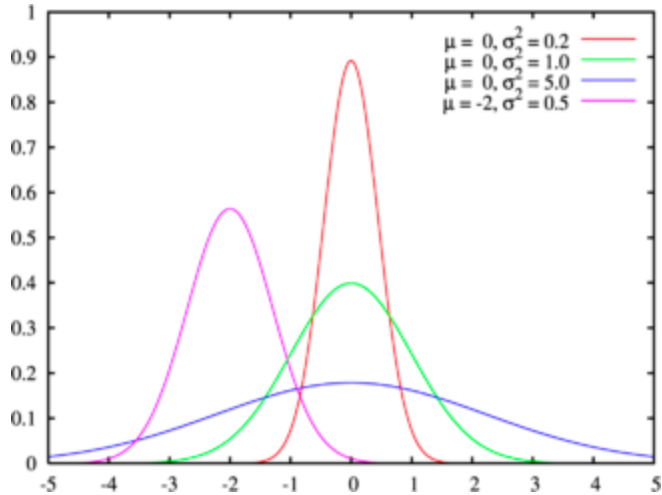
Set 1	Set 2
a) 4, 7, 7, 7, 10	4, 6, 7, 8, 10
b) 100, 140, 150, 160, 200	10, 50, 60, 70, 110
c) 10, 16, 18, 20, 22, 28	48, 56, 58, 60, 62, 70

a.

b.

c.

**Example 3. Center and spread.** Comparing Means and Standard Deviations graphically



a. Which distributions have the same mean?

b. Put the distributions in order from smallest to largest standard deviation.

Graph Source: [http://people.stern.nyu.edu/adamodar/New\\_Home\\_Page/StatFile/statdistns.htm](http://people.stern.nyu.edu/adamodar/New_Home_Page/StatFile/statdistns.htm)

**Relating Histograms and Boxplots for the Same Distribution**

**Class Activity 1.** Matching histograms to boxplots

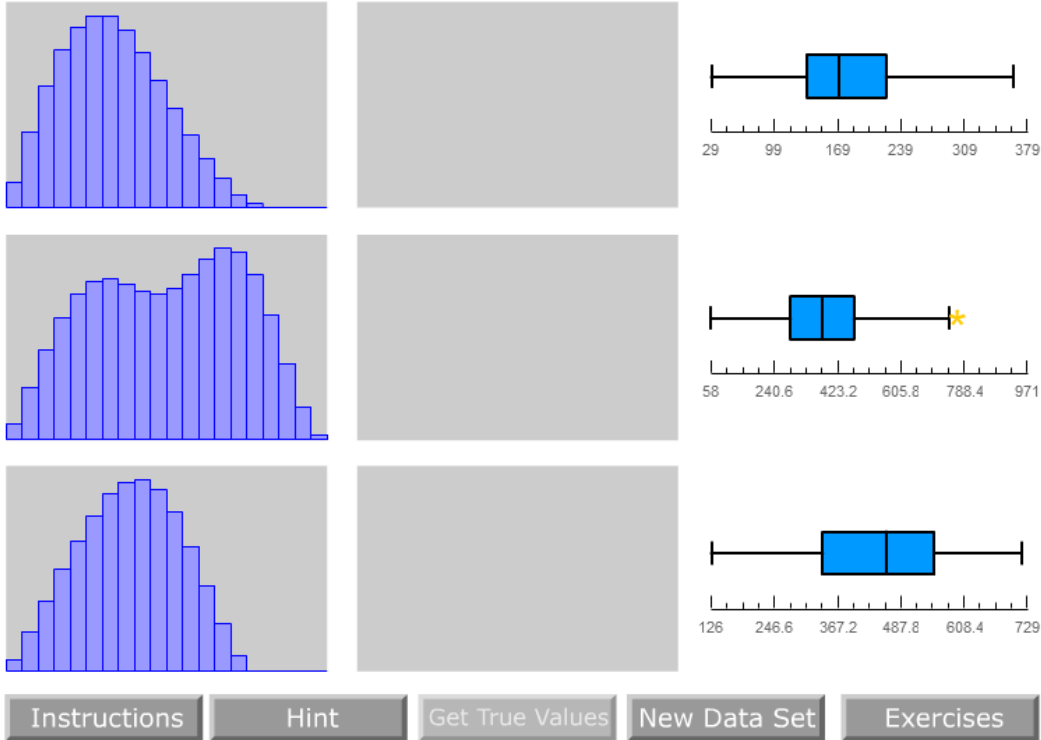
[http://higheredbcs.wiley.com/legacy/college/mann/0470444665/applets/applet\\_01\\_v4.html](http://higheredbcs.wiley.com/legacy/college/mann/0470444665/applets/applet_01_v4.html)

Now you try matching these histograms to the box-and-whisker plots:

**Relating Histograms and Box-and-Whisker Plots**


- Instructions
- Hint
- Get True Values
- New Data Set
- Exercises

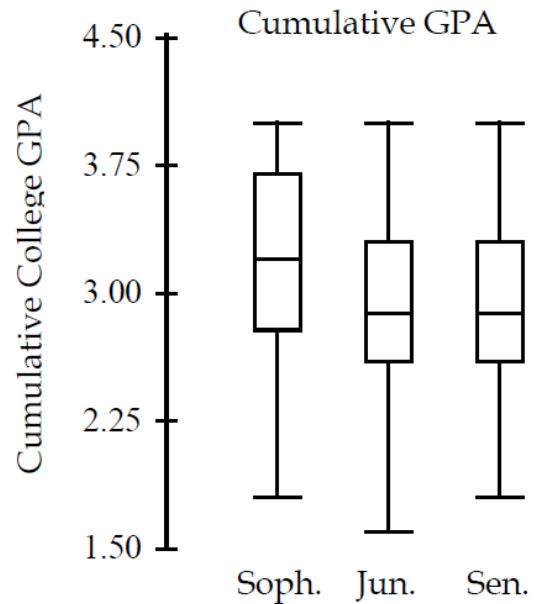
## Relating Histograms and Box-and-Whisker Plots



## Comparing Different Groups with Side-By-Side Boxplots

**Example 4.** The side-by-side boxplots show the cumulative college GPAs for sophomores, juniors and seniors taking an intro stats course.

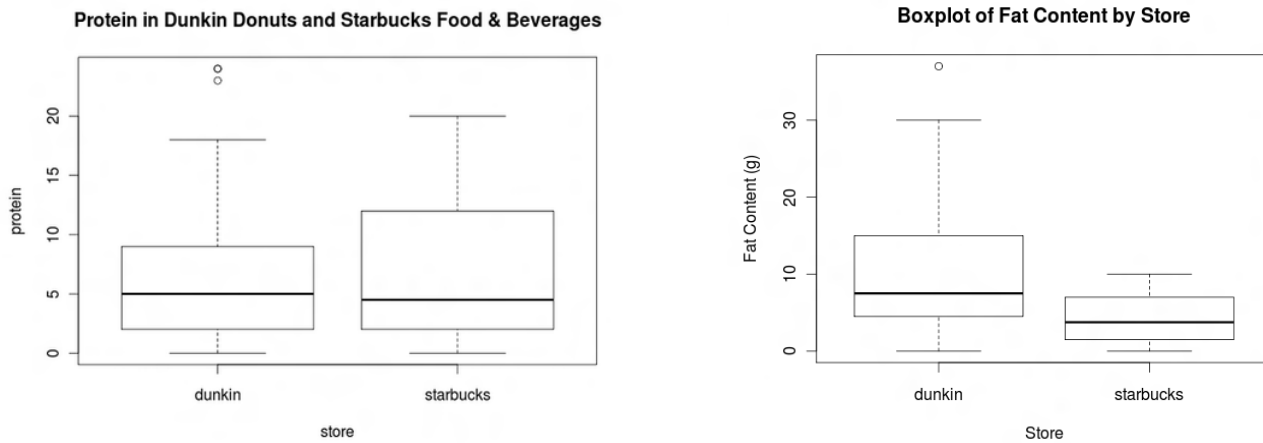
- Which class (sophomore, junior, or senior) had the lowest cumulative college GPA? What is the approximate value of that GPA?
- Which class has the highest median GPA, and what is that GPA?
- Which class has the largest range for GPA, and what is it?
- Which class has the most symmetric set of GPAs? The most skewed set of GPAs?



**Example 5.** Comparing side-by-side box plots

Data was collected on the nutritional content in the food and beverages served at Starbucks and Dunkin Donuts. The study shows data collected from the menus of each establishment. Data about the calorie, fat, carbohydrate, protein and fiber content of 20 items from each store was collected. Ten beverages and ten food items were chosen and recorded.

Source: <http://statsbls.weebly.com/nutrition-content-in-starbucks-and-dunkin-donuts.html>



a. Compare the distributions (shape, center and spread, and unusual features) of the protein in food and beverages at Starbucks and Dunkin Donuts.

b. Compare the distributions (shape, center and spread, and unusual features) of the fat content in the food and beverages at Starbucks and Dunkin Donuts.

## Activity 2. Summarizing Distributions

The number of chocolate chips was counted in each of 34 Keebler Chocolate Chip cookies. The data are listed below. Use GeoGebra to calculate the summary statistics and make a stacked histogram and boxplot with the same scale. Write a paragraph with complete sentences to summarize the distribution including all four characteristics.

29	31	25	32	27	31	30	29	31	26	32	33	32	30
33	29	30	28	32	35	37	31	24	30	30	34	29	27
24	38	37	32	26	30								

**Activity 3.** Use our class data for the variable "How many credits do you take per term." Use GeoGebra to calculate the summary statistics and make a stacked histogram and boxplot with the same scale. Write a paragraph with complete sentences to summarize the distribution including all four characteristics.