3.3-3.4: Summary Statistics: Measures of Center and Spread

Group Activity

1. Halloween Data. A data scientist has been counting the number of trick-or-treaters that come to his house every year. We will analyze the shape, center and spread of this data. Source: https://www.dataplusscience.com/HalloweenData.html

a. On the top grid, make a histogram of the number of trick-or-treaters using a bin-width of 100 trick-or-treaters. On the bottom grid you will make a boxplot, but not yet☺. Follow the steps in order. Label your axes.

Histogram of Trick-or-Treaters

Year	Number of Trick-or-Treaters
2008	492
2009	542
2010	726
2011	869
2012	673
2013	391
2014	454
2015	747
2016	822

Boxplot of Trick-or-Treaters

b. What is the shape of the histogram? If you are not sure yet, compare the mean and the median first.

c. Find the mean, median and mode, include units.

d. Find the 5-number summary, IQR and range, including units.

e. Use the 5-number summary to draw a boxplot on the second grid above. Make your horizontal scale match your histogram scale.
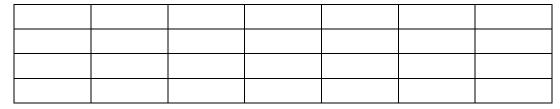
f. Do you think there are any outliers? Why or why not?

Comparing Distributions

2. Below are fictitious student test scores from a Math 105 midterm in two different classes. You will be making a boxplot for each to compare their distributions.

Class 1:   72, 86, 65, 99, 86, 71, 55, 86, 92, 73, 95, 71 points

Class 2:   75, 94, 82, 81, 69, 71, 85, 92, 88, 78, 73, 65, 66 points

a. Find the mean, 5-number summary, IQR and range for each class, including units.

    Class 1:   Class 2:

b. Draw the boxplot for each class using the same scale.

Boxplot for Class 1

Boxplot for Class 2

c. What is the shape of the data for each class? How can you tell?

Calculating Standard Deviation, *s*

d. Using your means rounded to one decimal place, find the standard deviation for each class, including units. The variable n refers to the number of data values.

Class 1:

Mean = _____, n = _____

Test Score (points)	Deviation from the mean	Squared deviation
72		
86		
65		
99		
86		
71		
55		
86		
92		
73		
95		
71		
Sum of the squared deviations (numerator)		

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

$$= \sqrt{\frac{}{ - 1}}$$

Class 2:

Mean = _____, n = _____

Test Score (points)	Deviation from the mean	Squared deviation
Sum of the squared deviations (numerator)		

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

$$= \sqrt{\frac{}{-1}}$$

e. Write a few complete sentences summarizing the four characteristics of the distribution of class 1 (shape, center, spread and unusual features).

f. Which class did better on the test? Use the vocabulary and values for center and spread in your answer.