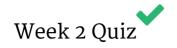
Week 2 Quiz

6/6 points (100%)

Quiz, 6 questions

**✓** Congratulations! You passed!

Next Item



1 / 1 points

6/6 points (100%)

Quiz, 6 questions

- 1.

Which of the below data sets has the **highest** standard deviation? You do not need to calculate the exact standard deviations to answer this question.

0,1,2,3,4,5,6
0, 25, 25, 25, 25, 25
0,1,1,1,1,1,2
0, 100, 200, 300, 400, 500, 600

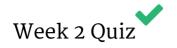
## Correct

This question refers to the following learning objective(s):

Note that there are three commonly used measures of center and spread:

- center: mean (the arithmetic average), median (the midpoint), mode (the most frequent observation)
- spread: standard deviation (variability around the mean), range (max-min), interquartile range (middle 50% of the distribution)

The dataset with the least repeated observations that are farthest from the center has the most variability, hence the highest standard deviation.



6/6 points (100%)

Quiz, 6 questions

2.

The distribution of housing prices in a country where 25% of the houses cost below \$350,000, 50% of the houses cost below \$450,000, 75% of the houses cost below \$1,000,000 and there are a meaningful number of houses that cost more than \$6,000,000 is most likely

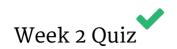
$\bigcirc$	symmetric
$\bigcirc$	left skewed
	uniform
0	right skewed

### Correct

This question refers to the following learning objective(s):

Identify the shape of a distribution as symmetric, right skewed, or left skewed, and unimodal, bimodoal, multimodal, or uniform.

There is a long tail on the right side of the distribution, yielding a right skew.

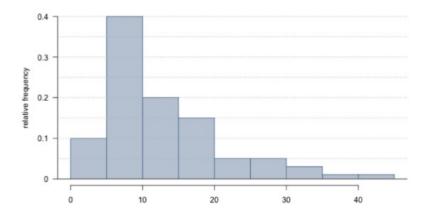


6/6 points (100%)

Quiz, 6 questions

3.

Based on the relative frequency histogram below, which of the following statements is supported by the plot?



The distribution is multimodal.
 It is not possible to estimate the median without knowing the sample size.

There are no outliers in the distribution.

The IQR of the distribution is roughly 10.

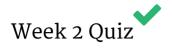
## Correct

This question refers to the following learning objective(s):

Use histograms and box plots to visualize the shape, center, and spread of numerical distributions, and intensity maps for visualizing the spatial distribution of the data.

Using the relative frequency histogram, we can tell that 10% of observations are below 5 (in the first bin), 40% are between 5 and 10, 20% are between 10 and 15, and 15% between 15 and 20. Therefore Q1 is in the second bin (between 5 and 10) and Q3 is in the fourth bin (between 15 and 20). This confirms that the IQR is roughly 10.

4 of 8 5 / 10 / 18 <sup>2</sup> 3:09 PM



6/6 points (100%)

Quiz, 6 questions

4

The midrange is defined as the average of the maximum and the minimum.

True or False: This statistic is robust to outliers.

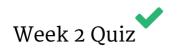
True
False

#### Correct

This question refers to the following learning objective(s):

Define a robust statistic (e.g. median, IQR) as a statistics that is not heavily affected by skewness and extreme outliers, and determine when such statistics are more appropriate measures of center and spread compared to other similar statistics.

If a distribution has outliers, these will be the maximum and the minimum. Therefore the midrange cannot be robust to outliers.



6/6 points (100%)

Quiz, 6 questions

5.

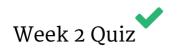
Phi Delta Kappa (PDK) is an international professional organization for educators that, in collaboration with Gallup, has been conducting polls on the public's attitudes toward the public schools since 1969. The following was one of the questions on the 2011 poll:

"Most teachers in the nation now belong to unions or associations that bargain over salaries, working conditions, and the like. Has unionization, in your opinion, helped, hurt, or made no difference in the quality of public school education in the United States?"

The respondents' answers broken down by party affiliation are shown below. Which of the following statements is **most justified** by these data?

	Republican	Democrat	Independent	Total
Helped	35	146	69	250
Hurt	197	78	178	453
Made no difference	52	105	90	247
Don't know / refused	6	12	4	22
Total	290	341	341	972

The results of the survey suggest that opinion on teachers belonging to unions or bargaining associations and political party affiliation appear to be independent.
 A histogram or a box plot would be useful for investigating if distribution of opinion on teachers belonging to unions or bargaining associations varies by political party affiliation.
 14% of Republicans and 58% of Democrats think that teachers belonging to unions or bargaining associations helped the quality of public school education in the United States.
 The results of the survey suggest a relationship between opinion on teachers belonging to unions or bargaining associations and political party affiliation.



6/6 points (100%)

Quiz, 6 questions

6.

Professors regularly give two versions of an exam. The professor may also provide summary statistics for each version. Suppose the following summary is provided:

	Number of exams	Mean	Median	SD
Version A	53	65.4	72	16
Version B	65	66.5	71	17

A student who took Version A says that he should get an extra point because his exam was harder as evidenced by the lower mean score for Version A, as shown by the mean score. Does the student have a good argument? Pick the best answer below.

We need to know the shape of the distribution for each version to determine if this argument is valid.
Yes. The difference in the exam scores means that there is a difference in difficulty between the versions.
We need to know the minimum and the maximum for each version to determine if this argument is valid.
No. The average scores are relatively close when considering the spread of the distributions. The difference might just be due to just chance.

# Correct

This question refers to the following learning objective:

Note that an observed difference in sample statistics suggesting dependence between variables may be due to random chance, and that we need to use hypothesis testing to determine if this difference is too large to be attributed to random chance. Set up null and alternative hypotheses for testing for independence between variables, and evaluate the data support for these hypotheses using a simulation technique.

The man scarce are off by only about 1 point when

Week 2 Quiz

Quiz, 6 questions

6/6 points (100%)