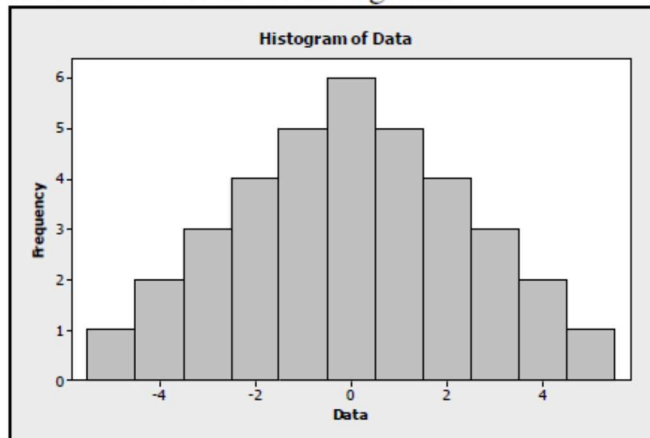


Intro Stats Test A – Data Analysis – Part I

Name _____

1. School administrators collect data on students attending the school. Which of the following variables is quantitative?
A) class (freshman, sophomore, junior, senior) B) grade point average
C) whether the student is in AP classes D) whether the student has taken the SAT
2. Which of the following variables would most likely follow a Normal model?
A) family income B) heights of singers in a co-ed choir
C) weights of male elephants D) scores on an easy test
3. A professor has kept records on grades that students have earned in his class. If he wants to examine the percentage of students earning the grades A, B, C, D, and F during the most recent term, which kind of plot should he make?
A) boxplot B) timeplot C) dotplot
D) pie chart E) histogram
4. Which is true of the data shown in the histogram?



- I. The distribution is approximately symmetric.
II. The mean and median are approximately equal.
III. The median and IQR to summarize these data summarize the data better than the mean and standard deviation.
- A) I only B) III only C) I and II
D) I and III E) I, II, and III
5. Two sections of a class took the same quiz. Section A had 15 students who had a mean score of 80, and Section B had 20 students who had a mean score of 90. Overall, what was the mean score for all of the students on the quiz?
A) 84.3 B) 85.0 C) 85.7
D) none of these E) It cannot be determined.
 6. It is permissible to violate the area principle if
A) the percentages do not add to 100%. B) we need a flashy display to make a point.
C) we do not care about being true to the data. D) None of the above.

7. Suppose that a Normal model described student scores in a history class. Parker has a standardized score (z -score) of +2.5. This means that Parker
- A) is 2.5 points above average for the class.
 - B) is 2.5 standard deviations above average for the class.
 - C) has a standard deviation of 2.5.
 - D) has a score that is 2.5 times the average for the class.
 - E) None of the above.

8. The advantage of a stem-and-leaf display over a dotplot is that a stem-and-leaf display
- A) satisfies the area principle.
 - B) shows the shape of the distribution better than a dotplot.
 - C) preserves the individual data values.
 - D) A stem-and-leaf display has no advantages over a dotplot.

9. The five-number summary of credit hours for 24 students in an introductory statistics class is:

Min	Q1	Median	Q3	Max
13.0	15.0	16.5	18.0	22.0

From this information, we know that

- A) there are no outliers in the data.
 - B) there is at least one low outlier in the data.
 - C) there is at least one high outlier in the data.
 - D) there are both low and high outliers in the data.
 - E) None of the above.
10. Which of the following data summaries are changed by adding a constant to each data value?
- I. the mean
 - II. the median
 - III. the standard deviation
- A) I only B) III only C) I and II
D) I and III E) I, II, and III

11. The table shows whether students in an introductory statistics class like dogs and/or cats.

		Like Dogs		
		Yes	No	Total
Like Cats	Yes	194	21	215
	No	110	10	120
Total		304	31	335

- What is the marginal distribution (in %) of “liking dogs”?
 - What is the conditional distribution (in %) of “liking dogs” for students who like cats?
 - What kind of display(s) would you use to show the association between “liking dogs” and “liking cats”? (Just name the graph(s).)
 - Do “liking dogs” and “liking cats” appear to be independent? Give statistical evidence to support your conclusion.
12. A local plumber makes house calls. She charges \$30 to come out to the house and \$40 per hour for her services. For example, a 4-hour service call would cost $\$30 + 4(\$40) = \$190$.
- The table shows summary statistics for the past month. Fill in the table to find out the cost of the service calls.

Statistic	Hours of Service Call	Cost of Service Call
Mean	4.5	
Median	3.5	
SD	1.2	
IQR	2.0	
Minimum	0.5	

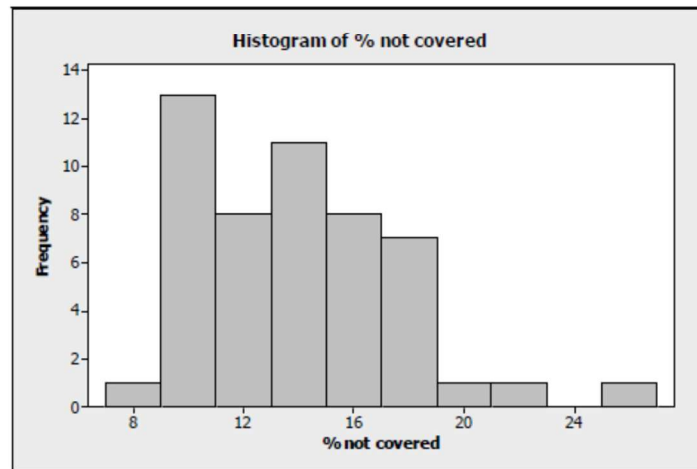
- This past month, the plumber had a service call that corresponded to a z -score of -1.50 . What was the z -score for the cost of that service call?

13. The *World Almanac and Book of Facts 2004* reported the percent of people not covered by health insurance in the 50 states and Washington, D.C., for the year 2002. Computer output gives these summaries for the percent of people not covered by health insurance:

Min	Q1	Median	Q3	Max	Mean	SD
7.9	10.8	13.4	16.7	25.8	13.9	3.6

a. Were any of the states outliers? Explain how you made your decision.

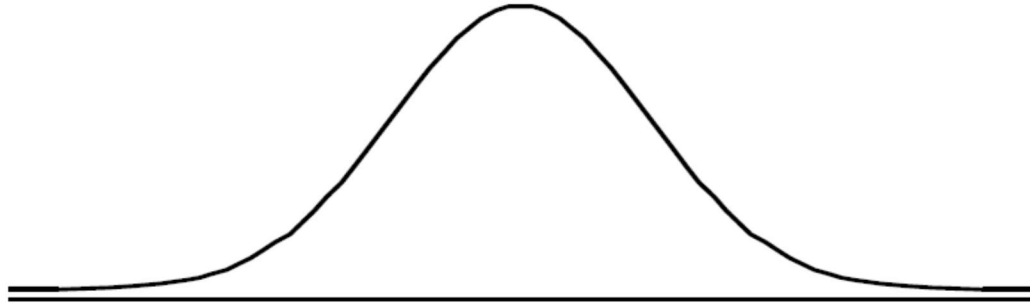
b. A histogram of the data is as follows:



Is it more appropriate to use the mean and standard deviation or the median and IQR to describe these data? Explain.

14. Costs for standard veterinary services at a local animal hospital follow a Normal model with a mean of \$80 and a standard deviation of \$20.

a. Draw and clearly label this model.



b. Is it unusual to have a veterinary bill for \$125? Explain.

c. What is the IQR for the costs of standard veterinary services? Show your work.

15. A machine that fills cans with soda fills according to a Normal model with mean 12.1 ounces and standard deviation 0.05 ounces.

a. If the cans claim to have 12 ounces of soda each, what percent of cans are under-filled?

b. Management wants to ensure that only 1% of cans are under-filled.

i. Scenario 1: If the mean fill of the cans remains at 12.1 ounces, what standard deviation does the filling machine need to have to achieve this goal?

ii. Scenario 2: If the standard deviation is to remain at 0.05 ounces, what mean does the filling machine need to have to achieve this goal?