Name_____

SHORT ANSWER. Write your answer in the space provided or on a separate sheet of paper.

Provide an appropriate response.

1) Define the terms population, sample, parameter and statistic. How does a census compare to a sample?

2) Distinguish between categorical and quantitative data. Give an example for each.

Identify the sample and population. Also, determine whether the sample is likely to be representative of the population.

3) An employee at the local ice cream parlor asks three customers if they like chocolate ice cream.

4) 100,000 randomly selected adults were asked whether they drink at least 48 oz of water each day and only 45% said yes.

Provide an appropriate response.

5) Histograms and Pareto charts are both bar charts. What is the significant difference between the two?

6) Suppose you are comparing frequency data for two different groups, 25 managers and 150 blue collar workers. Why would a relative frequency distribution be better than a frequency distribution?

7) Describe at least two advantages to using stemplots rather than frequency distributions.

8) The median of a data set is always/sometimes/never (select one) one of the data points in a set of data. Explain your answer with brief examples.

9) A company advertises an average of 42,000 miles for one of its new tires. In the manufacturing process there is some variation around that average. Would the company want a process that provides a large or a small variance? Justify your answer.

10) Marla scored 85% on her last unit exam in her statistics class. When Marla took the SAT exam, she scored at the 85 percentile in mathematics. Explain the difference in these two scores.

Solve the problem.

11) A student earned grades of A, C, A, A, and B. Those courses had these corresponding numbers of credit hours: 2, 5, 3, 2, 3. The grading system assigns quality points to letter grades as follows: A = 4, B = 3, C = 2, D = 1, and F = 0. Compute the grade point average (GPA) and round the result to two decimal places.

12) The mean salary of the female employees of one company is $29,525. The mean salary of the male employees of the same company is $33,470. Can the mean salary of all employees of the company be obtained by finding the mean of $29,525 and $33,470? Explain your thinking. Under what conditions would the mean of $29,525 and $33,470 yield the mean salary of all employees of the company?

Use the empirical rule to solve the problem.

13) The amount of Jen's monthly phone bill is normally distributed with a mean of $59 and a standard deviation of $8. What percentage of her phone bills are between $35 and $83?

Solve the problem.

14) If the standard deviation of a set of data is zero, what can you conclude about the set of values?

15) The data set below consists of the scores of 15 students on a quiz. For this data set, which measure of variation do you think is more appropriate, the range or the standard deviation? Explain your thinking.

90	90	91	91	89
90	89	91	91	90
60	90	89	90	91

1) A population is the complete collection of all elements. A sample is a subset of elements drawn from a population. A parameter is a numerical measurement describing some characteristic of a population. A statistic is a numerical measurement describing some characteristic of a sample. A census is the collection of data from every element in a population; a sample is a subset of a population.
2) Qualitative data can be separated into categories that are distinguished by nonnumeric characteristics. Quantitative data consist of numbers representing counts or measurements. Examples will vary.
3) Sample: the 3 selected customers; population: all customers; not representative
4) Sample: the 100,000 selected adults; population: all adults; representative
5) Answers will vary. Possible answer: Histograms convey quantitative information about shapes of distributions. Pareto charts convey comparative information about relative standing of categorical data.
6) Answers will vary. Possible answer: A relative frequency distribution is better for comparison between groups whose numbers are different, since ratios are readily comparable.
7) Answers will vary. Possible answer: The shape of a distribution can readily be seen. The plot can be drawn quicker, since class width need not be calculated.
8) The general answer is that the median is *sometimes* in a data set. For example, the median of the data set 1, 2, 3, 4, 5 is 3.
9) A small variance is preferred, since this measure denotes consistency in the lifetime of the tires. Given small variation, buyers would get useful mileage from those tires around 42,000. Large variation would indicate that some buyers could have their tires wear out many miles short of 42,000, whereas others might get good use out of many miles past 42,000.
10) Marla's score of 85% on her statistics exam tells us that Marla knew 85% of the content on that exam. Marla's percentile score of 85 tells us that her score was better than 85% of the scores of examinees on that test.
11) 3.13
12) In general, the mean salary of all employees of the company cannot be obtained by finding the mean of $29,525 and $33,470 because each of these means typically is obtained by averaging a different number of salaries for male and female employees. The mean of $29,525 and $33,470 will yield the mean salary of all employees of the company only if the number of female employees is equal to the number of male employees.
13) 99.7%
14) All values are identical.
15) For this data set, the range is very misleading. The range depends only on the smallest and largest values and the remainder of the data contributes nothing to the range. In this case, the smallest value is an outlier. Thus even though all the values except one lie between 89 and 91, the range is 31. The standard deviation, while it will also be affected by the outlier, will be less misleading, as it depends on every piece of data.