Name: **Period :**

Multiple-choice. Circle the letter corresponding to the best answer choice.

1. Sixteen percent of the students in the College of Engineering are majoring in Civil
   Engineering, 20% in Chemical, 20% in Electrical, 35% in Biomedical, and 9% in
   Aeronautical. Which of the following could be used to represent the data?

   (A) a bar graph only.   (B) a pie chart only.   (C) a box plot only.
   (D) Either (A) or (B).
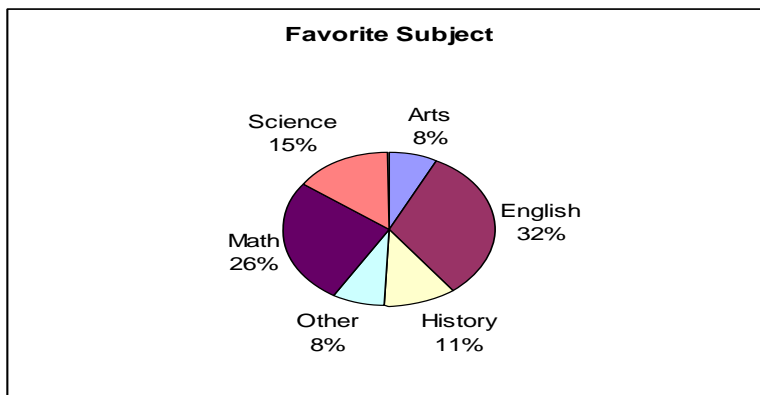
Favorite Subject

2. The pie chart above describes the distribution of favorite school subject for 300 high-school
   students. The number of students whose favorite subject is math is:

   (A) 26   (B) 45   (C) 60   (D) 78

3. If you want to make a graph to display the distribution of the salaries of the 1500 professors
   at a large university. The best choice is

   (A) a boxplot.
   (B) a histogram.
   (C) a pie chart.
   (D) a stemplot.

4. A common abuse of bar graphs is

   (A)   to use bars of equal width, so that we can't see differences among the bars.
   (B)   to replace bars by pictures and increase both height and width as the variable plotted
         increases.
   (C)   to use bars that don't touch each other.
   (D)   to stretch or squeeze the scale at the base of the bars, so that our eyes are misled.

5.

In a recent social research poll, 4800 randomly selected adults aged 18 to 26 were asked to record the number of hours they spent watching television and the number of hours they spent surfing the Web. Parallel boxplots of the data are given below.
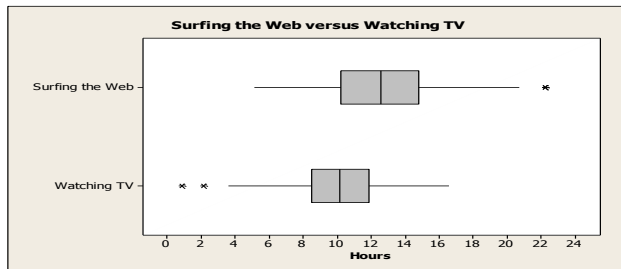
Based on the plots, which one of the following statements is FALSE?

(A) The range of the distribution of hours spent surfing the Web is larger than the range of the distribution of hours spent watching television.
(B) On average, adults aged 18 to 26 spent more time surfing the Web than watching television.
(C) The median number of hours spent surfing the Web is larger than the median number of hours spent watching television.
(D) The *IQR* for hours spent surfing the Web is smaller than the IQR for hours spent watching TV.

6. A company database contains the following information about each employee: age, date hired, sex (male or female), ethnic group (Asian, black, Hispanic, etc.), job category (clerical, management, technical, etc.), yearly salary. Which of the following list of variables are *all* categorical?
(A) age, sex, ethnic group.
(B) ethnic group, job category, yearly salary.
(C) sex, ethnic group, job category.
(D) yearly salary, age, job category

7. Here are the IQ scores of 10 randomly chosen fifth-grade students:

96   110   118   118   122   125   126   130   139   145

Which of the following statements about this data set is *False*?
(A) If the value 96 was removed from the data set, the median of the remaining 9 IQ scores would be higher than the median of all 10 IQ scores.
(B) If the value 96 was removed from the data set, the *IQR* of the remaining 9 IQ scores would be lower than the *IQR* of all 10 IQ scores.
(C) If the value 96 was removed from the data set, the standard deviation of the remaining 9 IQ scores would be higher than the standard deviation of all 10 IQ scores.
(D) If the value 96 was removed from the data set, the standard deviation of the remaining 9 IQ scores would be lower than the standard deviation of all 10 IQ scores.

8. The following dotplots show the average high temperature (in degrees Fahrenheit) for a sample of tourist cities from around the world for both January and July. What is one statement that can be made with certainty from an analysis of the graphical display?

Temperatures in Tourist Cities

Dot Plot
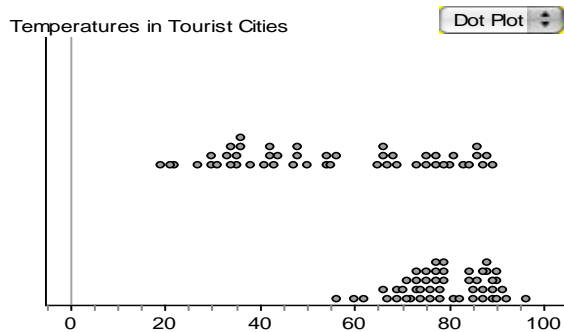
(A) Every city has a higher average high temperature in July than in January.
(B) The distribution of temperatures in July is skewed right, while the distribution of temperatures in January is skewed left.
(C) There appear to be outliers in the average high temperatures for January and July.
(D) There is more variability in average high temperatures in January than in July.

9. In a poll of 200 randomly selected students, males and females were asked their favorite type of pet. The comparative bar chart shows the results.
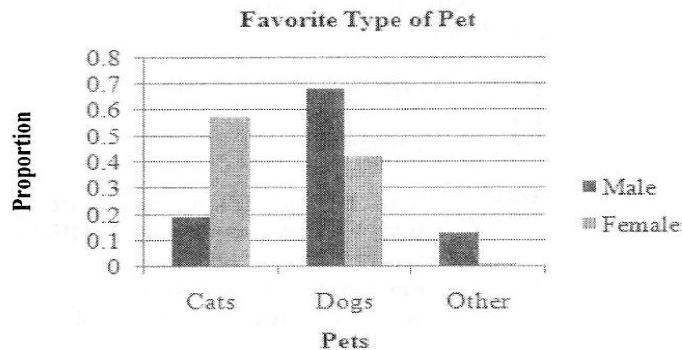
Which statement below is true?

(A) Almost half the dog lovers are female.
(B) The number of male dog lovers in the data set is greater than the number of female dog lovers.
(C) The percentage of females who prefer cats is higher than the corresponding percentage for men.
(D) About 30% of the students in the sample who prefer cats are male.

10. One way economists measure the health of the real estate market is by counting "housing starts," or the number of permits issued for construction of new homes. Below is a graph displaying housing starts (in thousands) in the United States from 2006 to 2009.
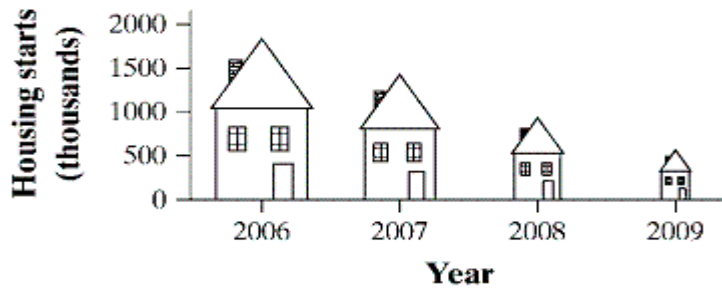
What is the principle weakness of this graphical presentation of data?

(A) The data only shows housing starts for four years, which is not enough time to identify a meaningful trend.
(B) Data of this type should only be displayed in a pie chart.
(C) Using proportionally-sized pictograms exaggerates the difference between years.
(D) It is unclear which dimension of the house represents the number of housing starts for that year.

11. <u>Match</u> the summary statistics with the correct histograms.   **Variable Number**

(A) mean = 4.99, median = 3.13, standard deviation = 5.49   _____
(B) mean = 4.89, median = 4.83, standard deviation = 7.99   _____
(C) mean = 5.01, median = 6.87, standard deviation = 5.49   _____
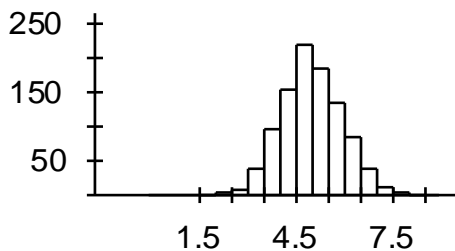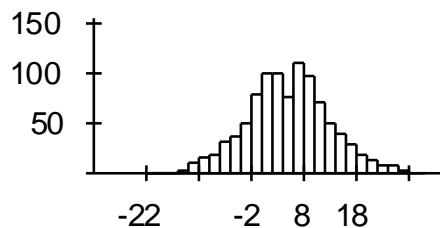(D) mean = 4.96, median = 4.93, standard deviation = 0.96   _____
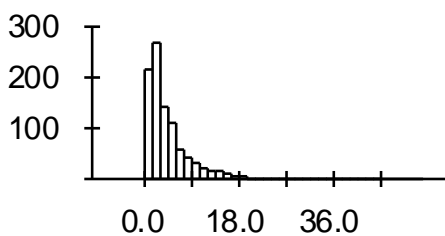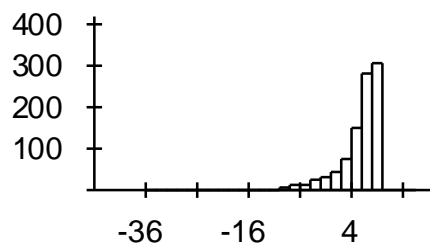

variable 1

variable 2

variable 3

variable 4

Part II: Short answer questions.

12. The manufacturer of exercise machines for fitness centers has designed two new elliptical machines that are meant to increase cardiovascular fitness. The two machines are being tested on 30 volunteers at a fitness center near the company's headquarters. The volunteers are randomly assigned to one of the machines and use it daily for two months. A measure of cardiovascular fitness is administered at the start of the experiment and again at the end. The table below contains the differences in the two scores (After - Before) for the two machines.

| Machine A		Machine B
	5	3, 5, 9
6, 1	4	2, 5, 7
9, 7, 4, 1, 1	3	2, 4, 8, 9
8, 7, 6, 3, 2, 0	2	1, 5, 9
5, 4	1	0
	0	2

(a) Write a few sentences comparing the distributions of cardiovascular fitness gains from the two elliptical machines.

(b) Which machine should be chosen if the company wants to advertise it as achieving the highest overall gain in cardiovascular fitness? Explain your reasoning.

(c) Which machine should be chosen if the company wants to advertise it as achieving the most consistent gain in cardiovascular fitness? Explain your reasoning.

13.   The following cumulative relative frequency graph (ogive) displays the distribution of the lifetime (in hours) of a AA battery.

Lifetimes of AA Batteries (hrs)

Battery Life (in hours)
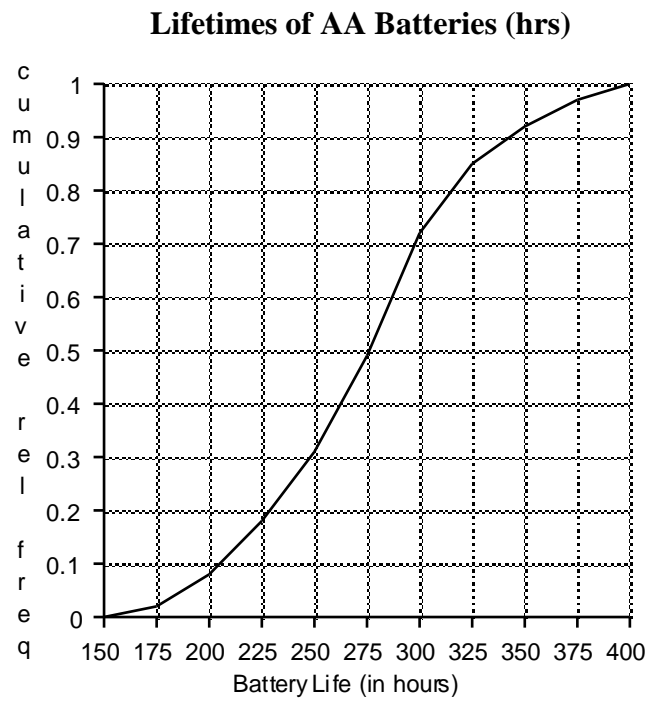
(a) Approximately what percent of the batteries will last fewer than 250 hours?

(b) Approximately what proportion of batteries will last 325 hours or more?

(c) What battery life would be needed so that the battery was at the 60th percentile?

(d) Determine the IQR of the distribution of battery lifetimes.

14. A group of 78 third-grade students in a Midwestern elementary school took a "self-concept" test that measured how well they felt about themselves. Higher scores indicate more positive self-concepts. A histogram and some summary statistics from Minitab for these students' self-concept scores are given below
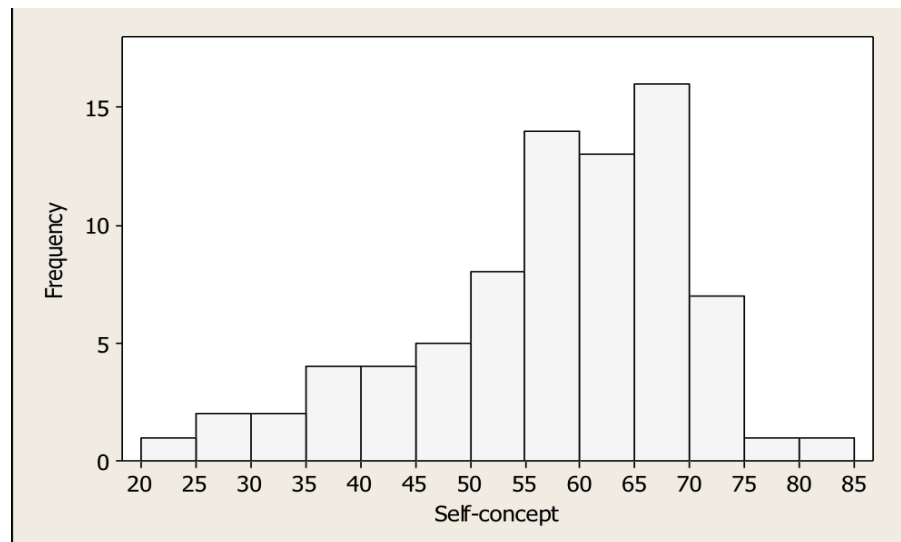
Descriptive Statistics: SelfConc

Variable	N	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
SelfConc	78	56.85	12.35	20.00	50.00	59.00	65.00	84.00

(a) Thirty-three students had self-concept scores higher than 62. One student in the group had a self-concept score of 62. Calculate and interpret this student's percentile and the z-score for this student.

(b) Describe the shape of the distribution.

15. Twenty-two males and 14 females in two statistics classes at a Dallas private school responded to the following question, "What is the fastest you have ever driven?" The summary statistics for the data is given below.

	Males	Females
Minimum	55	65
Lower Quartile	85	79.5
Median	92	85
Upper Quartile	105	92.5
Maximum	158	98

a) Use the same scale to draw parallel boxplots for the fastest driving speeds of males and females. [Note: there may be outliers in either distribution]. The three fastest male driving speeds were 158, 146, and 120. The three fastest female driving speeds were 98, 95, and 95.

b) Write a few sentences comparing the two distributions.