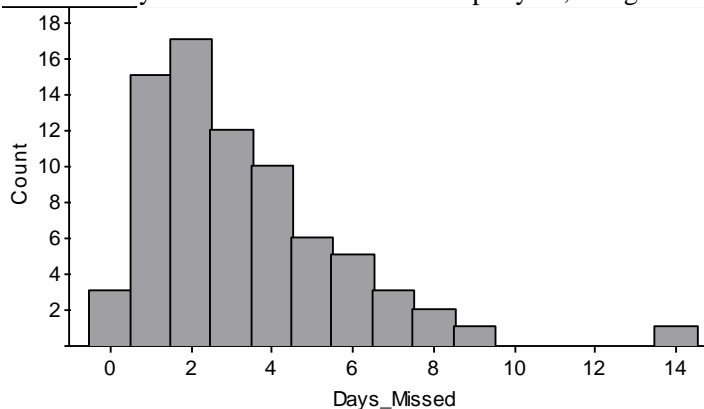Name _____

AP Statistics
Part 1 Review Test 2

1. You have a set of data that you suspect came from a normal distribution. In order to assess normality, you construct a normal probability plot. Which of the following would constitute evidence that the data actually came from a normal distribution?

   (a) A strongly linear relationship between the data and their standardized values.
   (b) A bell-shaped (normal) relationship between the data and their standardized values.
   (c) A random scattering of points when the standardized values are plotted against the data.
   (d) A strongly non-linear relationship (with no outliers) between the data and their percentiles.
   (e) A uniform relationship between the percentiles and the standardized values.

2. The cost of glass cleaner is nicely described by a Normal model with a mean cost per ounce of 7.7 cents with a standard deviation of 2.5 cents. What is the z-score of Windex with a cost of 10.1 cents per ounce?

   (a) 0.96
   (b) 1.31
   (c) 1.94
   (d) 2.25
   (e) 3.00

3. A large company has offices in two locations, one in New Jersey and one in Utah. The mean salary of the office assistants in the New Jersey office is $28,500. The mean salary of office assistants in the Utah office is $22,500. The New Jersey office has 128 office assistants and the Utah office has 32 office assistants. What is the mean salary paid to the office assistants in this company?

   a. $22,500
   b. $23,700
   c. $25,500
   d. $27,300
   e. $28,500

4. In the northern U.S., schools are sometimes closed during winter due to severe snowstorms. At the end of the school year, schools have to make up for the days missed. The following graph shows the frequency distribution of the number of days missed due to snowstorms per year, using data from the past 75 years.


Which of the following should be used to describe the center of the distribution?

   (a) Mean, because it is always the best measure.
   (b) Median, because the distribution is skewed.
   (c) IQR, because it excludes outliers and includes only the middle 50 percent of the data.

(d) First quartile, because the distribution is left skewed.

(e) Standard deviation, because it is unaffected by outliers.

5. A resident of Auto Town was interested in finding the cheapest gas prices at nearby gas stations. On randomly selected days over a period of one month, he recorded the gas prices (in dollars) at four gas stations near his house. The box plots of gas prices are as follows:
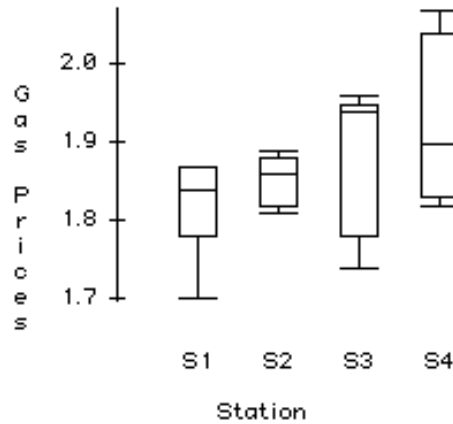
Which station has more consistent gas prices?

(a) Station 1
(b) Station 2
(c) Station 3
(d) Station 4
(e) Cannot be determined

6. A small kiosk at the Atlanta airport carries souvenirs in the price range of $3.99 to $29.99, with a mean price of $14.75. The airport authorities decide to increase the rent charged for a kiosk by 5 percent. To make up for the increased rent, the kiosk owner decides to increase the prices of all items by 50 cents. As a result, which of the following will happen?

a. The mean price and the range of prices will increase by 50 cents.
b. The mean price will remain the same, but the range of prices will increase by 50 cents.
c. The mean price and the standard deviation of prices will increase by 50 cents.
d. The mean price will increase by 50 cents, but the standard deviation of prices will remain the same.
e. The mean price and the standard deviation of prices will stay the same.

7. Which of the following are true statements?

I.   The standard deviation is the square root of the variance.
II.   The standard deviation is zero only when all values are the same.
III. The standard deviation is strongly affected by outliers.

(a) I and II
(b) I and III
(c) II and III
(d) I, II, and III
None of the above gives the complete set of true responses.

8.

A random sample of 25 birthweights (in ounces) is taken, yielding the following summary statistics:

Variable	N	Mean	Median	TrMean	StDev	SE Mean
Birthwt	25	129.40	129.00	128.35	17.41	3.48

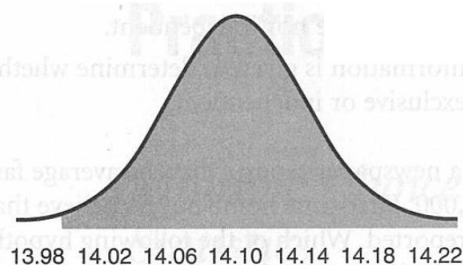| Variable	Minimum	Maximum	Q1	Q3
Birthwt	96.00	187.00	119.50	135.50

What can be said about the number of outliers for this data set?

(A) 0
(B) At least 1
(C) No more than 1
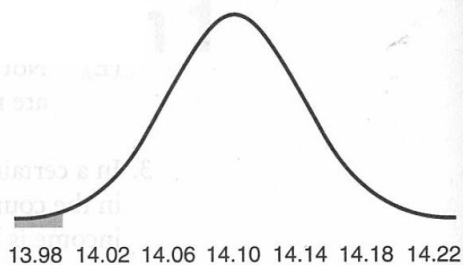(D) At least 2
(E) No more than 2

9.

Weights for a box of cereal are normally distributed with a mean of 14.10 oz and a standard deviation of 0.04 oz. Which of the following illustrates the probability of selecting a box with at *least* the advertised weight of 14 oz?
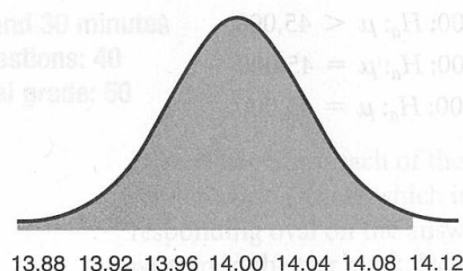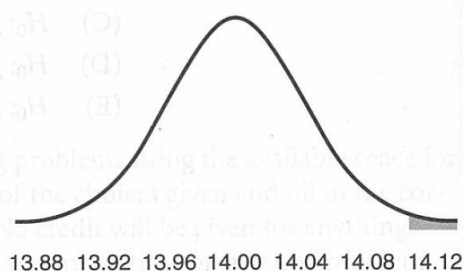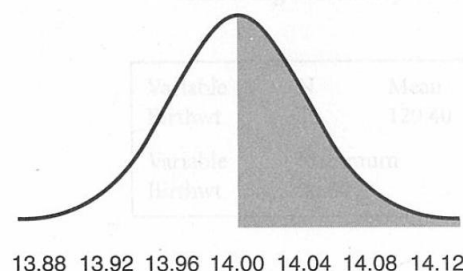
(A)

13.98 14.02 14.06 14.10 14.14 14.18 14.22

(B)

13.98 14.02 14.06 14.10 14.14 14.18 14.22

(C)

13.88 13.92 13.96 14.00 14.04 14.08 14.12

(D)

13.88 13.92 13.96 14.00 14.04 14.08 14.12
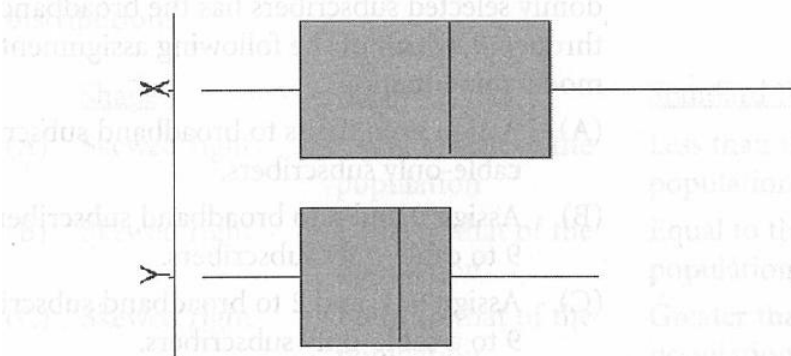
(E)

13.88 13.92 13.96 14.00 14.04 14.08 14.12

10.

The following boxplots summarize two data sets, X and Y. Which of the following MUST be true?

(A) Set X and set Y have the same number of data points.
(B) The box of set X contains more data points than the box of set Y.
(C) The data in set X have a larger range than the data in set Y.
(D) About 50% of the values in set X are greater than about 75% of the values in set Y.
(E) The median of set X is less than the median of set Y.

11.

For a set of values, suppose the mean is 10 and the standard deviation is 2. If each value is multiplied by 9, what will be the mean and standard deviation for this new set of values?
(A) mean 10; standard deviation 2
(B) mean 10; standard deviation 18
(C) mean 90; standard deviation 2
(D) mean 90; standard deviation 6
(E) mean 90; standard deviation 18

12.

In this year's county mathematics competition, a student scored 40; in last year's competition, the student scored 35. The average score this year was 38 with a standard deviation of 2. Last year's average score was 34 with a standard deviation of 1. In which year did the student score better?
(A) The student scored better on this year's exam.
(B) The student scored better on last year's exam.
(C) The student scored equally well on both exams.
(D) Without knowing the number of test items, it is impossible to determine the better score.
(E) Without knowing the number of students taking the exam in the county, it is impossible to determine the better score.

13.

The lengths (in innings) of 25 randomly selected Little League baseball games were recorded, and a *cumulative* frequency histogram was created from the results. What is the best conclusion that can be made from the graph?
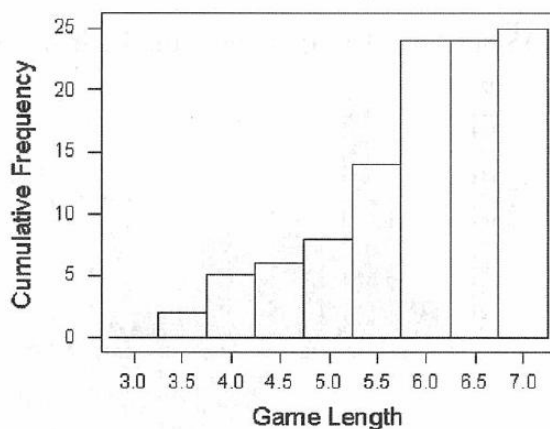
(A)   The median game length is 5 innings.
(B)   Fourteen games lasted 5.5 innings.
(C)   A majority of the games lasted 6 or more innings.
(D)   The distribution of game lengths is severely skewed left.
(E)   Games lasting more than 6 innings occurred least frequently.

14.

A random variable has a standard deviation of 1.3. A new variable is created by transforming the values of the random variable using the following rule: Multiply each value by 2 and then add 5. What is the value of the standard deviation for this transformed variable?

(A)   1.3
(B)   2.6
(C)   6.3
(D)   7.6
(E)   8.5

15.

For the dotplot shown, which of the statistical values listed would be least affected by the rightmost value?
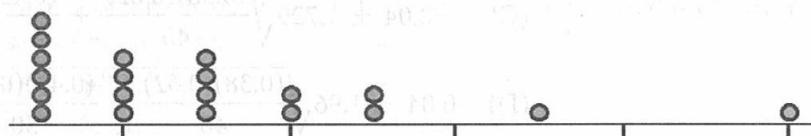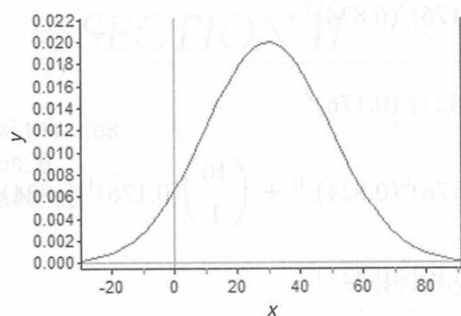

(A)   Standard deviation
(B)   Range
(C)   Variance
(D)   Mean
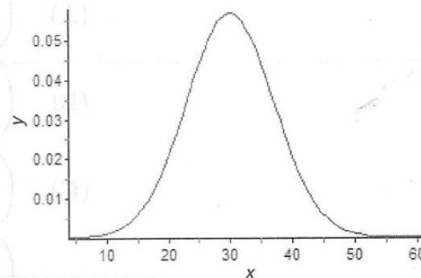(E)   Median

16.

Which of the following distributions has a mean of 30 and a standard deviation of 7?
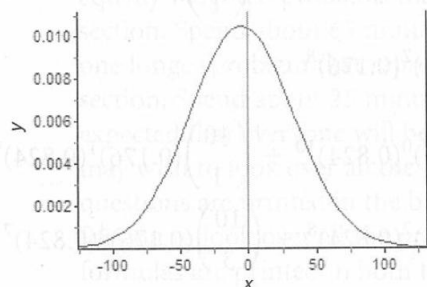
(A)

(B)

(C)

(D)
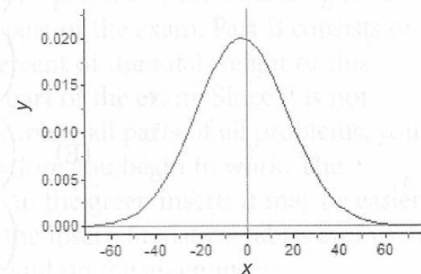
(E)
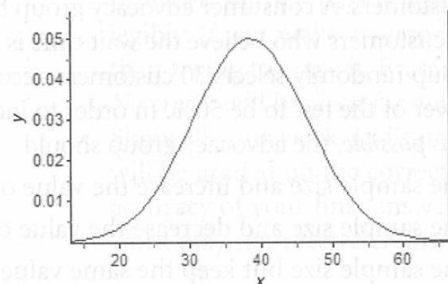
17.

Suppose that 25% of horses live over 23.4 years, while 85% live less than 25.2 years. Assuming the ages of horses are normally distributed, what are the mean and standard deviation for the life expectancy of horses?

(A) mean 20.14; standard deviation 4.86

(B) mean 22.690; standard deviation 1.052

(C) mean 22.690; standard deviation 4.97

(D) mean 24.110; standard deviation 1.052

(E) Not enough information is given to find the mean and standard deviation.

For Hospital A, the average waiting time (time between walking in the door and seeing a doctor) in the emergency room is 135 minutes with a standard deviation of 45 minutes. For Hospital B, the average waiting time in the emergency room is 90 minutes with a standard deviation of 22.5 minutes. In which hospital are you more likely to wait less than 45 minutes? Assume the distributions of waiting times are normally distributed.
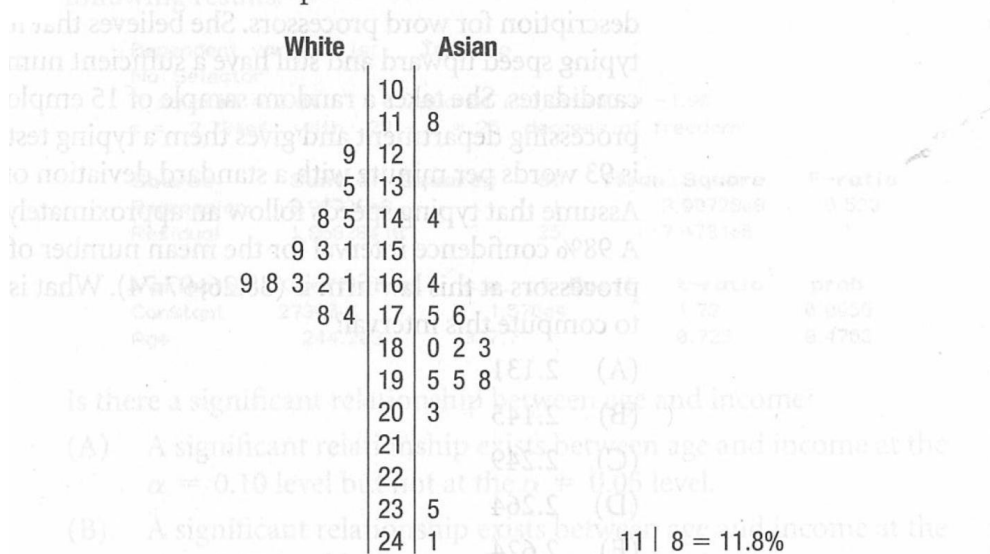
(A) Hospital B, because the average wait-time is only 90 minutes, rather than 135 minutes in Hospital A.

(B) Hospital A, because with twice the standard deviation of Hospital B, it has twice the spread.

(C) Neither, because for both hospitals, the probability of waiting less than 45 minutes is 2.275%.

(D) Neither, because for both hospitals, the probability of waiting less than 45 minutes is 97.725%.

(E) It is impossible for any patient to wait less than 45 minutes in either hospital unless the patient is in critical condition.

19. In the accompanying display, which has the larger mean and which has the larger standard deviation?
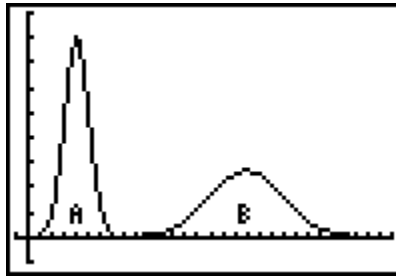
(a) Larger mean, A; larger standard deviation, A
(b) Larger mean, A; larger standard deviation, B
(c) Larger mean, B; larger standard deviation, A
(d) Larger mean, B; larger standard deviation, B
(e) Larger mean, B; same standard deviation

20.

The percentages of children living below poverty levels for white and Asian children in the United States from 1987 to 2000 are given in the back-to-back stemplot.

White		Asian
	10	
	11	8
9	12	
5	13	
8 5	14	4
9 3 1	15	
9 8 3 2 1	16	4
8 4	17	5 6
	18	0 2 3
	19	5 5 8
	20	3
	21	
	22	
	23	5
	24	1

11 | 8 = 11.8%

Which of the following is a statement that can be made from an examination of the back-to-back stemplot?

(A)   For every year from 1987 to 2000, the percentage of white children living in poverty is lower than the percentage of Asian children.

(B)   If the outliers are removed from the distribution of Asian poverty-level percentages, the range for the distribution of white children will be larger than that for Asian children.

(C)   The mean poverty-level percentage for both white and Asian children is less than the respective median value.

(D)   The poverty-level percentages for both white and Asian children have increased over time.

(E)   There is more variability in the poverty-level percentages for Asian children than for white children.

1.

Suppose that the weights of a name-brand cereal vary normally with mean $\mu = 11.13$ oz and standard deviation $\sigma = 0.08$ oz. The advertised weight is 11 oz. For the equivalent generic brand with an advertised weight of 11 oz, the weights vary normally with mean $\mu = 11.15$ oz and standard deviation $\sigma = 0.16$ oz.

a. For each of the brands, find the probability that the weight of a box of cereal will be less than the advertised weight.

b. If you believe strongly in getting what you pay for, that is, you want to make sure you get at least the weight advertised on the box, which brand of cereal are you more likely to buy and why?

c. If you wanted to get the most for your money, that is, you would like to get much more than the advertised weight, which brand of cereal are you more likely to buy and why?

2. Five hundred randomly selected middle-aged men and five hundred randomly selected young adult men were rated on a scale from 1 to 10 on their physical flexibility, with 10 being the most flexible. Their ratings appear in the frequency table below. For example, 17 middle-aged men had a flexibility rating of 1.

Physical Flexibility Rating	Frequency of Middle-Aged Men	Frequency of Young Adult Men
1	17	4
2	31	17
3	49	29
4	71	39
5	70	54
6	87	69
7	78	83
8	54	93
9	34	73
10	9	39

(a) Display these data graphically so that the flexibility of middle-aged men and young adult men can be easily compared.

(b) Based on an examination of your graphical display, write a few **sentences** comparing the flexibility of middle-aged men with the flexibility of young adult men.

3.

Male and female life expectancies (in years) are given for 15 regions of the world.

Male: 51.1, 76.0, 65.2, 63.8, 70.3, 61.3, 64.5, 65.8, 64.1, 74.9, 76.3, 76.9, 67.0, 74.2, 77.5

Female: 53.2, 83.0, 72.2, 71.5, 74.6, 62.4, 69.5, 70.3, 73.9, 81.5, 82.3, 82.7, 72.2, 79.9, 84.1

a. Create a back-to-back stemplot to compare male and female life expectancies.

b. Create parallel boxplots to compare male and female life expectancies.

c. Based on an examination of your graphical displays, write a few sentences comparing the life expectancies of men and women.

d. Give one advantage to using the back-to-back stemplot over the parallel boxplot.

e. Give one advantage to using the parallel boxplot over the back-to-back stemplot.