1. Complete the following table with the appropriate symbols for each box:

	Mean	Proportion	Standard Deviation
Population parameter			
Sample statistic			

Multiple choice. Write the letter of the best available choice in the blank provided. (6 pts each)

2. _____ Data were collected on the number of text messages sent by each student in a large high school for one day. A boxplot of the data is shown below.
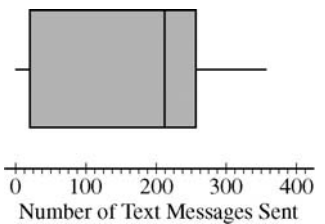
Based on the boxplot, which of the following statements is the most reasonable conclusion?

A) There are more students with data values below the median than there are students with data values above the median.

B) There are more students with data values between the first quartile and the median than there are students with data values between the median and the third quartile.

C) There are fewer students with data values between the first quartile and the median than there are students with data values between the median and the third quartile.

D) There are approximately the same number of students with data values between the first quartile and the minimum as there are students with data values between the third quartile and the maximum.

E) The data are less spread out between the first quartile and the median than between the median and the third quartile.

3. _____ For which of the following distributions is the mean greater than the median?

A)

0

B)

0

C)

0

D)

0

E)

0
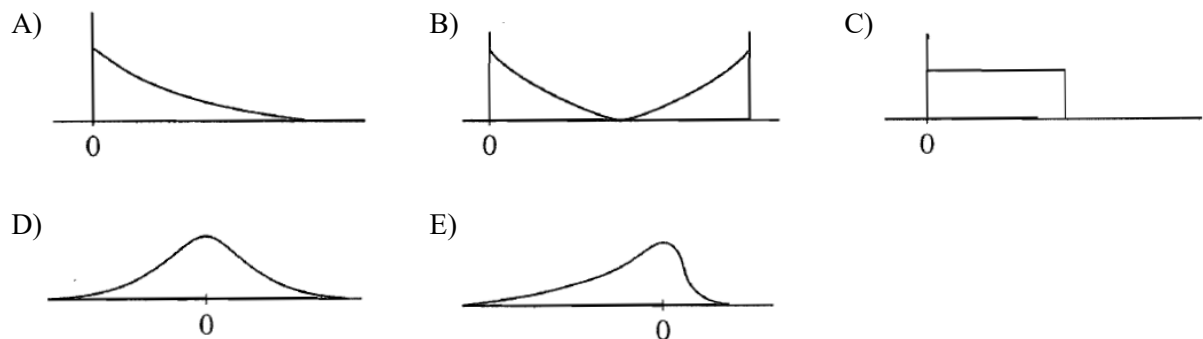
4. _____ Consider a set of positive values, at least two of which are not equal. Which of the following sample statistics will be changed when each value in this set is added to a constant whose value is greater than 0?
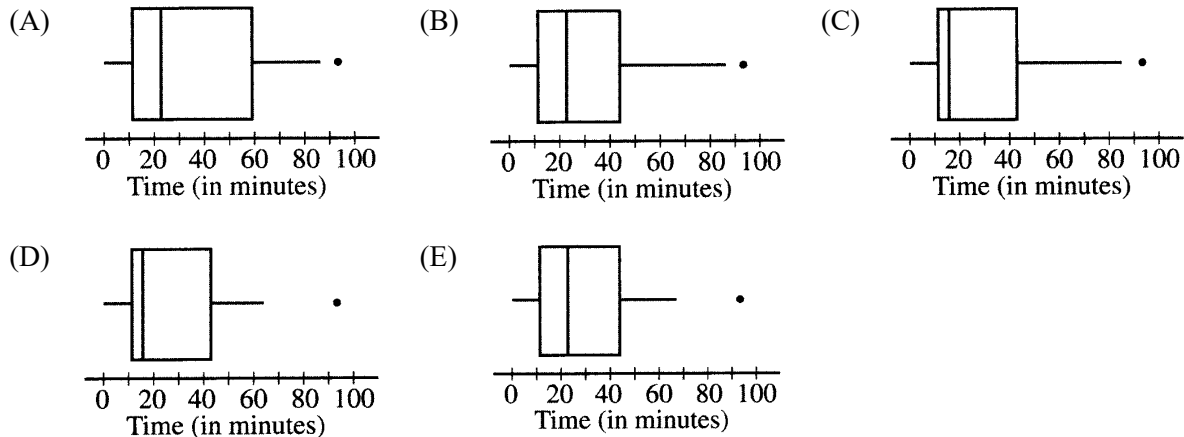
I. Mean   II. Median   III. Range

A) I only   B) II only   C) III only   D) I and II   E) I, II, and III

5. _____ The histogram below displays the times, in minutes, needed for each chimpanzee in a sample of 26 to complete a simple navigational task.
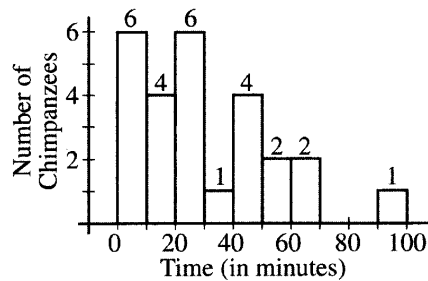
It was determined that the largest observation, 93, is an outlier since $Q_3 + 1.5(Q_3 - Q_1) = 87.125$. Which of the following boxplots could represent the information in the histogram?

(A)

0   20 40 60 80 100
Time (in minutes)

(B)

0   20 40 60 80 100
Time (in minutes)

(C)

0   20 40 60 80 100
Time (in minutes)

(D)

0   20 40 60 80 100
Time (in minutes)

(E)

0   20 40 60 80 100
Time (in minutes)

6. _____ From 1999 to 2000, there were 66 deaths attributed to "crowd rush" at rock concerts. Two of the statistics that are used to describe the center of a distribution are mean and median. Likewise, both standard deviation and IQR (interquartile range) are used to describe spread. Given the histogram to the right, which of the following is a true statement for this distribution of deaths attributed to crowd rush?
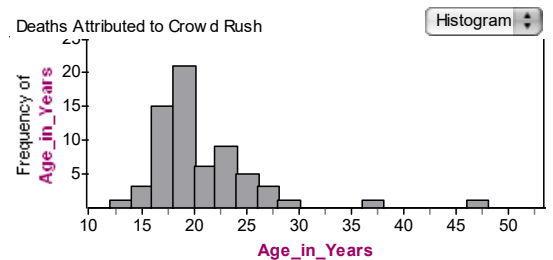
A) The mean is a better measure of center than the median, and the standard deviation is a better measure of spread than the IQR.

B) The median is a better measure of center than the mean, and the IQR is a better measure of spread than the standard deviation.

C) It is appropriate to use either the mean or median to describe the center of this distribution, as it is also appropriate to use either the standard deviation or the IQR to describe the spread.

7. _____ A local ice cream shop hand scoops each of its ice cream cones. The cones vary in weight such that the quartiles and median weights are 5.6, 9.1, and 7.2 ounces, respectively. The mean weight is 6.45 ounces, and the standard deviation is 1.2 ounces. Is the distribution symmetric, skewed to the left, or skewed to the right? Explain.

A) Skewed to the right, mean lower than median.

B) Symmetric, mean lower than median.

C) Skewed to the left, mean higher than median.

D) Skewed to the right, mean higher than median.

E) Skewed to the left, mean lower than median.

8. _____ The ages (to the nearest year) of people in a college biology class are as follows:

Age	18	19	20	21	22	23	24	25	32
Frequency	14	110	200	200	90	35	15	2	1

Note: This chart shows that there are 14 students who are age 18, 110 students who are age 19, etc.

What is true about the median age?

    A) It must be 20
    B) It must be 21
    C) It must be 22
    D) It could be any number between 19 and 21
    E) None of the above

9. _____ The ages of people attending the opening show of a new movie are summarized in the cumulative relative frequency graph (ogive) shown. Which of the following statements is/are true?
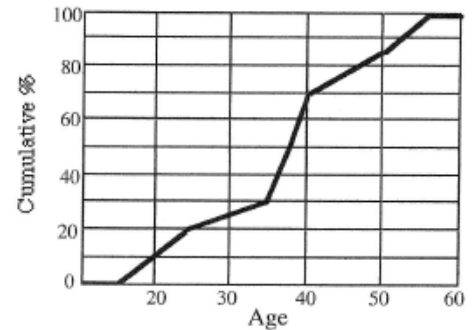
    I. The median age of people attending this movie is about 38.
    II. About 70% of the people attending this movie are have an age of at least 40 (age 40 or older).
    III. The IQR of ages of people attending this movie is about 13 or 14.

   A) I only   B) II only   C) III only   D) I and II   E) I and III

10. _____ A teacher is teaching two AP Statistics classes. On the final exam, the 20 students in the first class had a mean score of 92 while the 25 students in the second class had a mean score of only 83. If the teacher combines the classes, what will the overall mean final exam score be?

    A) 87   B) 87.5   C) 88   D) 85
    E) There is insufficient information to answer this question.

11. _____ A teacher is teaching two AP Calculus classes. On the final exam, the 20 students in the first class had a median score of 92 while the 25 students in the second class had a median score of only 83. If the teacher combines the classes, what will the overall median final exam score be?

    A) 87   B) 87.5   C) 88   D) 85
    E) There is insufficient information to answer this question.

12. _____ Three brands of candy pieces—X, Y, and Z—are made in many colors. Shaela bought one bag of each brand and counted the number of pieces of each color. The graph to the right shows the relative frequency distribution of colors for each bag. Which of the following statements must be true?
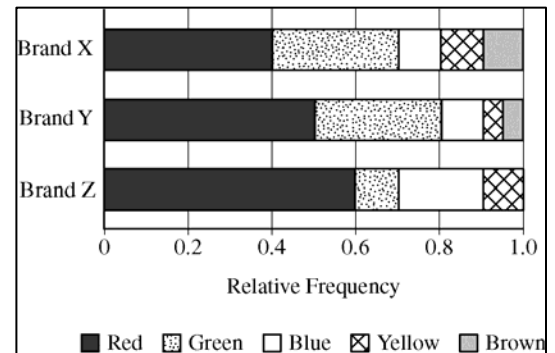

    A) For Brand X, there were more green candy pieces than red candy pieces in the bag.
    B) For Brand Y, there were more red candy pieces than green candy pieces in the bag.
    C) There were more green candy pieces in the Brand X bag than were in the Brand Z bag.
    D) There were the same number of blue candy pieces in the Brand X bag as were in the Brand Y bag.
    E) The number of blue candy pieces in the Brand Z bag was equal to the sum of the number of blue candy pieces in the other two bags.

13. _____ Of the following dotplots, which represents the set of data that has the greatest standard deviation?

A)
```
*  *  *  *  *  *  *  *  *  *
0                          10
```

B)
```
*****              *****
0                          10
```

C)
```
***        ****        ***
0                          10
```

D)
```
****        **        ****
0                          10
```

E)
```
**      ** ** **      **
0                          10
```

14. _____ The pulse rate for each person in a sample of 20 men and 20 women was recorded. The boxplots below summarize the pulse rates for the men and the women in the sample.
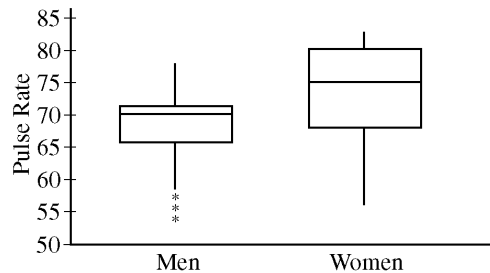
Which of the following statements about the people in the sample must be true?

A) There are more people between the first and third quartiles for women than there are between the first and third quartiles for men.

B) The person with the lowest pulse rate is a woman.

C) At least half of the women had higher pulse rates than three-fourths of the men.

D) More than half of the men had lower pulse rates than three-fourths of the women.

E) If a man and a woman were randomly selected from the 40 people, the man would have the lower pulse rate.

15. _____ Data on homes recently sold in a certain town included the area of the home, reported in square feet. The table below shows summary statistics of the reported areas, in square feet.

Mean	Minimum	Q1	Median	Q3	Maximum	Standard Deviation
1,754.14	1,656	1,704	1,758	1,806	1,843	61.0723

An auditor determined that an error was made in the reported areas and that all of the areas should have been 100 square feet greater than what was reported. The areas were corrected and new summary statistics were reported.

What are the interquartile range (IQR) and the standard deviation of the corrected areas?

A) IQR 102, standard deviation 61.0723

B) IQR 102, standard deviation 161.0723

C) IQR 202, standard deviation 61.0723

D) IQR 202, standard deviation 161.0723

E) IQR 187, standard deviation 61.0723

Free Response – Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

16. The following summary statistics are given for the daily temperature (in degrees Celsius) during a summer month in a nice coastal city:
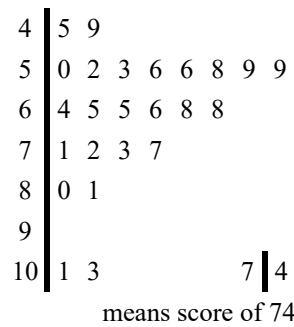
      mean = 29   median = 27   IQR = 6   standard deviation = 4   minimum = 25

In order to convert temperatures from Celsius to Fahrenheit, you first multiply the temperature by 1.8, and then add 32. If you convert each of the given statistics from Celsius into Fahrenheit, what will the new measures be?

      mean = _____   median = _____   minimum = _____

      IQR = _____   standard deviation = _____   variance = _____

17. The test grades from a recent exam at Podunk high school are displayed in the following stemplot:

```
 4 | 5 9
 5 | 0 2 3 6 6 8 9 9
 6 | 4 5 5 6 8 8
 7 | 1 2 3 7
 8 | 0 1
 9 |
10 | 1 3                    7 | 4
                         means score of 74
```

a) List the 5 number summary. (Label each value in the list. Example: median: _____)

b) Construct a box plot for the test grades. ***Show your work on the test for outliers.***

18. The following data is the home run count per season for Mark McGwire and Babe Ruth.

McGwire:	49	32	33	39	22	42	9	9	39	52
	58	70	65	32	3	29				

Ruth:	54	59	35	41	46	25	47	60	54	46
	49	46	41	34	22					

a) Construct a back-to-back stem and leaf plot for this data.

b) Write a few sentences comparing the distributions of home runs per season by Mark McGwire and Babe Ruth.

c) Based on the stem and leaf plot as well as your answer in part (b), who was the better home run hitter? Explain your reasoning.

19. AP Comic Design students at East Podunk High School and West Podunk High School are competing to see which school does better on the last major exam before Spring Break. The teachers for at both schools decide to give the same exam to their students at each school (each school has only one class period of AP Comic Design), and the results of their test scores are summarized in the parallel boxplots below:
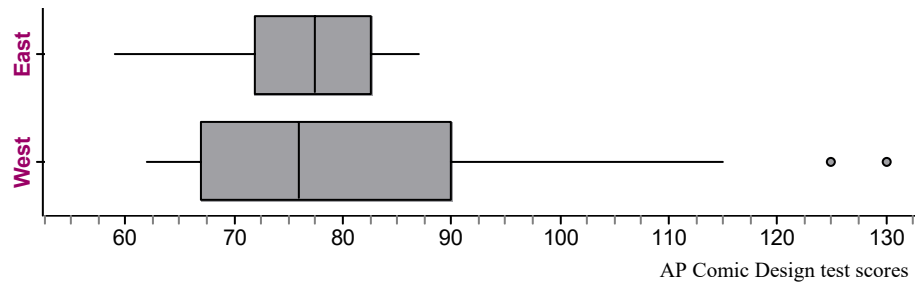
a) Write a few sentences comparing the distributions of test scores for the two schools.

b) Which school do you think performed better on this exam? Explain. *(Hint: Try to make a case for both schools!)*

20. Parker and Katrina asked 117 of their classmates about what major they intended to pursue in college. The results are summarized in the 2-way table below.

Gender

Major		Female	Male	Total
	Liberal Arts	12	23	35
	Business	15	16	31
	Science	23	28	51
Total		50	67	

a) Create a graphical display which shows the association between choice of major and gender for the students in this survey.

b) For this group of people, does there appear to be an association between "gender" and "field of study"? Write a few sentences describing what this graph reveals about this association.
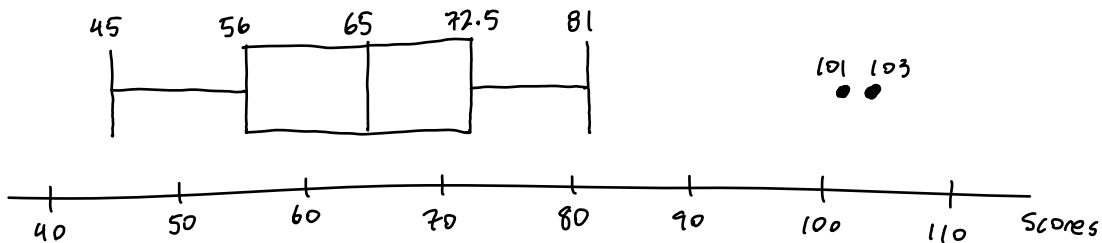
ANSWERS

1. Check your notes?

2. D   3. A   4. D

5. E   6. B   7. E

8. B   9. E   10. A   11. E   12. B

13. B   14. C   15. A

16. mean = 84.2   median = 80.6   minimum = 77
    IQR = 10.8   s.d. = 7.2   variance = 51.84

17. a) Min: 45   Q1: 56   Median: 65   Q3: 72.5   Max: 103
    b) (Make sure to show the work for the outlier test!)

18. a)

   b) **In general**, McGwire had the better single season of 70, but also had more variation in
         the number of home runs hit than did Ruth.
      Center: Babe Ruth had a higher mode (in the 40's) than did McGwire (in the 30's).
      Shape: The distribution for number of home runs hit per season by Ruth is unimodal and
         approximately symmetric. McGwire's distribution is also unimodal, but skewed
         to the higher number of HRs.
      Spread: McGwire's distribution goes from 3 – 70 (range of 67), which has more
         variability than Babe, who went from 22 – 60 (range of 38).
      Unusual Features: Mark McGwire's distribution has a gap in the 10's, with 3 possible
         low outliers (possibly due to injury). No unusual features for Babe Ruth's
         distribution.

   c) Answers may vary here, but should justify based on the STEMPLOT and response to part (b) only.
   (For instance, outside knowledge of baseball history should <u>not</u> be used to justify your response)

19. a) **CENTER:** The distribution of test scores on this Comic Design at East Podunk has a higher median than West Podunk (about 78 to 73).
SPREAD: The scores at West Podunk are more spread out (62 to 130, range of 68) than at East Podunk (60 to 88, range of 28). West Podunk's distribution also has a larger IQR than East Podunk.
SHAPE: While the shapes are difficult to compare with boxplots, the distribution of scores for East Podunk appears slightly left-skewed, while the scores for West Podunk appear skewed to the right.
OUTLIERS: West Podunk has 2 high outliers at around 126 and 130, while there are no outliers for East Podunk.

b) Answers may vary, but should make comparisons between each groups' min, Q1, median, Q3, and max.

20. a) A graph that compares PROPORTIONS should be used. One possible example:
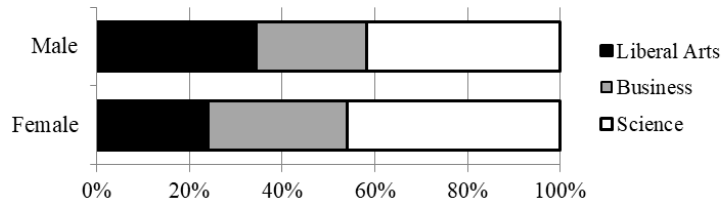
b) Yes; Student should proceed to COMPARE the PROPORTIONS (not counts) of each field of study between the two genders in context.