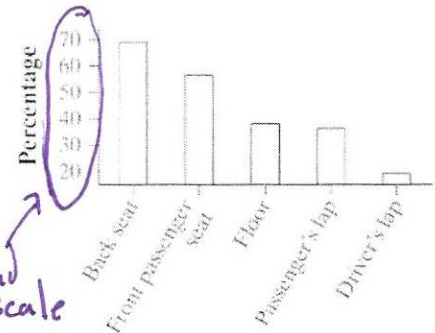**Part 1: Multiple Choice.** *Circle the letter corresponding to the best answer.*

E

Q   C   Q

1. You measure the <u>age</u>, <u>marital status</u> and earned <u>income</u> of an SRS of 1463 women. The number and type of variables you have measured is
   (a) 14563
   (b) four; two categorical and two quantitative.
   (c) four; one categorical and three quantitative.
   (d) three; two categorical and one quantitative.
   (e) three; one categorical and two quantitative. ⟵ circled

E

2. The bar graph below summarizes responses of dog owners to the question, "Where in the car do you let your dog ride?" Which of the following statements is (false?)
   (a) Some owners let their pets ride in more than one place in the car. True
   (b) A majority of owners allow their pets to ride in the front passenger seat. more than 50% True
   (c) The most common place dogs ride is in the back seat. True bad scale
   (d) The vertical scale of this graph exaggerates the difference between the percentage who let their dogs ride in the driver's lap *versus* a passenger's lap. True
   (e) These data could also be presented in a pie chart. ⟵ circled

Where does the dog ride?

C

3. Here are the ages of randomly selected people at a community playground.
   3, 10, 42, 25, 5, 8, 6, 12, 15, 4, 3, 21
   To make a stemplot of these scores, you would use as stems

   0
   1
   2
   3
   4

   (a) 0 and 1.
   (b) 1, 2, and 4.
   (c) 0, 1, 2, 3, and 4 ⟵ circled
   (d) 3, 10, 42, 25, 5, 8, 6, 12, 15, 4, 3, and 21
   (e) None of the above is a correct answer.

B

4. If a distribution is symmetrical, which of the following is true?
   (a) The mean must be less than the median.
   (b) The mean and median must be equal. ⟵ circled
   (c) The mean must be greater than the median.
   (d) The mean is either equal to or less than the median,
   (e) It's impossible to tell which of the above statements is true without seeing the data.

M $\bar{x}$
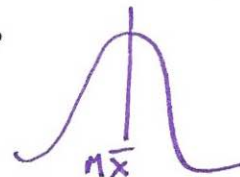
C

5. If a distribution is skewed to the right, which of the following is true?
   (a) The mean must be less than the median.
   (b) The mean and median must be equal.
   (c) The mean must be greater than the median.
   (d) The mean is either equal to or less than the median,
   (e) It's impossible to tell which of the above statements is true without seeing the data.

$\bar{x}$

A

6. If a distribution is skewed to the left, which of the following is true?
   (a) The mean must be less than the median. ⟵ circled
   (b) The mean and median must be equal.
   (c) The mean must be greater than the median.
   (d) The mean is either equal to or less than the median,
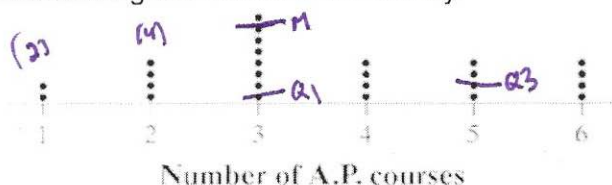   (e) It's impossible to tell which of the above statements is true without seeing the data.
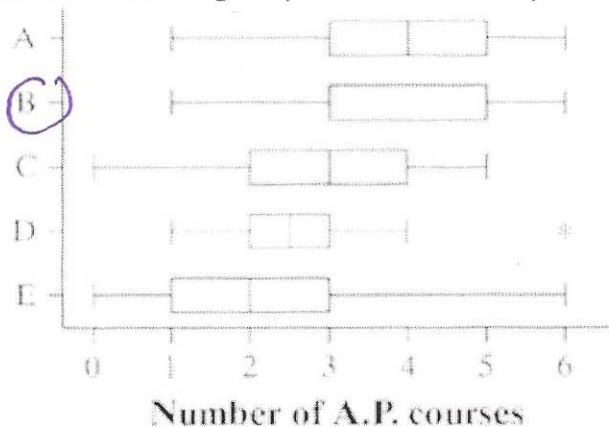
$\bar{x}$
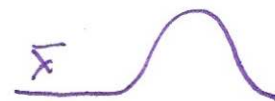
B 7. Mr. Williams asked the 26 seniors in his statistics class how many A.P. courses they had taken during high school. Below is a dot plot summarizing the results of his survey.

Number of A.P. courses

min = 1
$Q_1 = 3$ } Same value
med = 3
$Q_3 = 5$
max = 6

Which of the following boxplots is a correct representation of the same distribution?

Number of A.P. courses

B 8. A survey was designed to study how business operations vary according to their size. Companies were classified as small, medium, or large. Questionnaires were sent to 200 randomly selected businesses of each size. Since not all questionnaires in a survey of this type are returned, researchers decided to investigate the relationship between the response rate and the size of the business. The data is shown in the table below:

Size	Response	No Response	Total
Small	125	75	200
Medium	81	119	200
Large	40	160	200

Small company response rate = $\frac{125}{200}$ = 62.5%

large company response rate = $\frac{40}{200}$ =

Which of the following conclusions seems to be supported by the data?
(a) There are more small companies than large companies in the survey.
(b) Small companies appear to have a higher response rate than medium or big companies.
(c) Exactly the same number of companies responded as didn't respond.
(d) Small companies dislike larger companies
(e) If we combine the medium and large companies, then their response rate would be equal to that of the small companies.

C 9. A small school has 10 employees, one of whom is the principal and the others are teachers. Suppose the principal makes $100,000 per year and the other employees make between $40,000 and $50,000 per year. One day, the principal wins an award and is given a $20,000 raise. Which of the following describes how the school's mean and median salaries would change?

$\frac{20,000}{10}$ = $2000 increase

(a) The mean and median would both increase by $2,000.
(b) The mean would increase by $12,000 and the median would not change.
(c) The mean would increase by $2,000 and the median would not change.
(d) The median would increase by $2,000 and the mean would not change.
(e) The mean would increase by $2,000, but we cannot determine the change in the median without more information.

Adding 20,000 to the total earned by 10 people will increase mean by $\frac{20,000}{10}$.
Since principal already had max salary, increasing it will not change median.

C 10. The mean salary of all female workers is $35,000. The mean salary of all male workers is $41,000. What must be true about the mean salary of all workers?

The mean of all workers will be a weighted average of the salaries of both male and female workers. This average must be between $35,000 and $41,000

(a) It must be $38,000
(b) It must be larger than the median salary
(c) It could be any number between $35,000 and $41,000.
(d) It must be larger than $38,000
(e) We don't have enough information to draw any conclusion about the mean salary of all workers.

___E___ 11. The mean mercury concentration in sampled cans of tuna is 0.285 ppm with a standard deviation of 0.300ppm. Which of the following is a correct interpretation of standard deviation?
   (a) The mean amount of mercury in all other cans of tuna will be within 0.300 ppm of the mean.
   (b) All cans of tuna have between 0 and 0.585 ppm of mercury in them.
   (c) About half the cans of tuna have between 0 and 0.585 ppm of mercury in them
   (d) The difference in the mean and the median amount of mercury is 0.300 ppm
   (e) The distance between the ppm of mercury in each can and the mean ppm is, on average, about 0.300ppm

___B___ 12. Mrs. Settle took a sample of scores from a Statistics Partner Quiz. The five number summary for the scores was $\underset{min}{\ } \underset{Q_1}{\ } \underset{med}{\ } \underset{Q_3}{\ } \underset{Max}{\ }$
   $\underset{min}{\ }$ 0, 7, 9, 10, 10
   About what percent of the scores are between 0 and 7?
(a) 50%   (b) 25%   (c) 35%   (d) 70%   (e) 95%

Use for problems 13-14: Forty students took a statistics exam having a maximum of 50 points. The score distribution is given in the following stem and leaf plot: $Q_1 = 23$ $med = \dfrac{(31+33)}{2} = 32$

```
0 | 28
1 | 2245 Q₁
2 | 01333358889     1|4 is a score of 14
3 | 001356679
4 | 22444466788     Q3 = 44
5 | 000  Q3
```

$IQR = 44 - 23 = 21$

___B___ 13. Josh got the lowest score, what was his score?
   (a) 28   (b) 2   (c) 0   (d) 5   (e) 0.28

___D___ 14. The interquartile range for the distribution of exam scores is
   (a) 23 to 44   (b) 32 to 44   (c) 9   (d) 21   (e) 11

___B___ 15. The mean of four people in a room is 30 years. A new person whose age is 55 years enters the room. The mean age of the five people now in the room is
   (a) 30   (b) 35   (c) 37.5   (d) 40   (e) Cannot be determined
                                                                          $\underset{min}{\ } \underset{Q_1}{\ } \underset{Med}{\ } \underset{Q_3}{\ } \underset{Max}{\ }$
___B___ 16. The five number summary for time spent on the internet each day (min) is 7, 30, 46.5, 77, 151. If 300 students participated in the survey, about how many people spent between 46.5 and 77 minutes on the internet?
   (a) 30.5   (b) 75   (c) 25   (d) 100   (e) Cannot be determined

(15) $\bar{x} = \dfrac{\Sigma x_i}{n}$ $\dfrac{120+55}{5} = \dfrac{175}{5} = 35$ years   (16) 25% of 300

$30 = \dfrac{\Sigma x_i}{4}$ .25(300) = 75

$\Sigma x_i = 120$

Part 2: Free Response

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

17. During the early part of the 1994 baseball season, many fans and players noticed the number of homeruns being hit seemed unusually large. Here is the dotplot and numerical summary for the number of homeruns hit by American League and National League teams in the early part of the 1994 season.
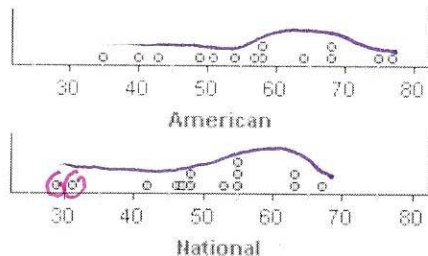
Variable	n	Mean	StdDev	Min	Q1	Median	Q3	Max
American	14	56.93	12.69	35.00	49.00	57.50	68.00	77.00
National	14	50.14	11.13	29.00	46.00	50.50	55.00	67.00

a.) Determine if there are any outliers in each distribution. <u>Show your work</u>.

American: $IQR = 68 - 49 = 19$

$1.5(IQR) = 1.5(19) = 28.5$
lower $Q_1 - 28.5 = 49 - 28.5 = 20.5$
upper $Q_3 + 28.5 = 68 + 28.5 = 96.5$

outliers < 20.5 or outliers > 96.5
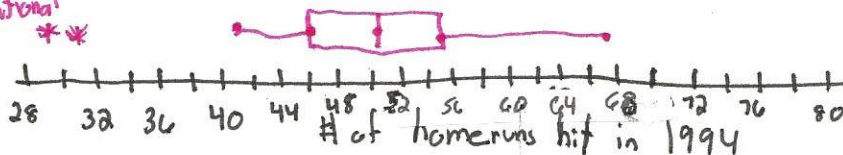no outliers

National: $IQR = 55 - 46 = 9$

$1.5(IQR) = 1.5(9) = 13.5$
lower $Q_1 - 13.5 = 46 - 13.5 = 32.5$
upper $Q_3 + 13.5 = 55 + 13.5 = 68.5$

outliers < 32.5 or outliers > 68.5
There are 2 outliers 29 and ≈31.

b.) Draw parallel boxplots of these two distributions. Be sure to <u>label the plots</u> and <u>provide a scale</u>.

American

National
** **

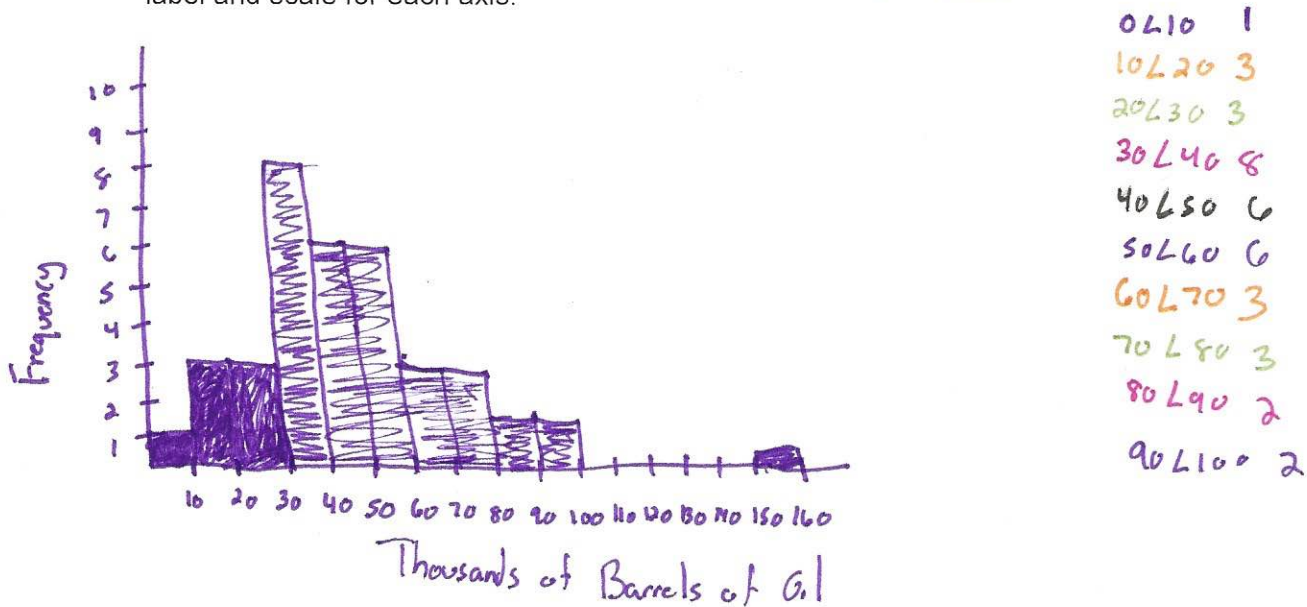28 32 36 40 44 48 52 56 60 64 68 72 76 80
# of homeruns hit in 1994

c.) Write a few sentences comparing the homeruns for each league.

The American League has more homeruns and more variability than the National league. The American league had a greater median of 57.5 compared to the National League's median of 50.5. The American League had a larger $IQR = 19$ (larger variability) compared to the National League, whose $IQR = 9$. The American League is more symmetrical than the National League which is skewed right (when ignoring the outliers found in part A)

18. How much oil wells in a given field will ultimately produce is key information in deciding whether to drill more wells. Here are the estimated total amounts of oil recovered from 38 wells in the Devonian Richmond Dolomite area of the Michigan basin, in thousands of barrels. The data is provided in ascending order.

3	22	35	43	49	57	70	92
13	25	35	44	60	59	70	98
15	31	37	45	50	63	74	157
19	33	37	46	53	65	86	
21	35	38	48	56	66	82	

a.) Construct a histogram for this distribution. Choose an appropriate bin width, and be sure to provide a label and scale for each axis.

10

Thousands of Barrels of Oil

0L10   1
10L20 3
20L30 3
30L40 8
40L50 6
50L60 6
60L70 3
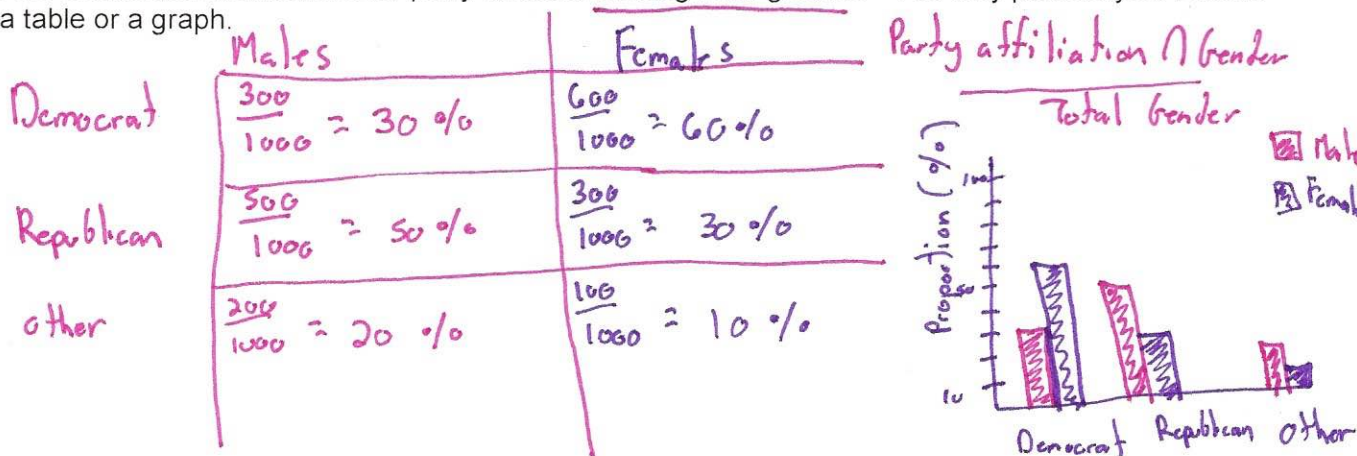70L80 3
80L90 2
90L100 2

b.) Based on your histogram, what numerical measures of center and spread would be best to use for this distribution? Explain your choice.

Since there is an outlier and graph is slightly skewed right, the median and IQR are the best measures of center and spread since they are resistant to outliers and skewness.

19. Does party affiliation differ among the genders? A review of voter registration in a small town yielded the following table of the number of males and females registered for each party

Party Affiliation	Male	Female	Total
Democrat	300	600	900
Republican	500	300	800
Other	200	100	300
Total	1000	1000	2000

given

a.) Calculate the conditional distributions for party affiliation among each gender. You may present your results in either a table or a graph.

Males

Democrat $\frac{300}{1000} = 30\%$

Republican $\frac{500}{1000} = 50\%$

other $\frac{200}{1000} = 20\%$

Females

$\frac{600}{1000} = 60\%$

$\frac{300}{1000} = 30\%$

$\frac{100}{1000} = 10\%$

Party affiliation ∩ Gender / Total Gender

Proportion (%)

☑ Male ☑ Female

Democrat Republican Other

b.) Discuss the relationship between gender and party affiliation in two or three sentences.

The appears to be an association between gender and party affiliation. The males have the highest proportion of Republicans (50%) while females have the highest proportion of Democrats (60%).

The secondary party affiliation is equal at 30%. (Democrats for males, Republicans for females.)

Both genders had the party affiliation "other" as the lowest proportion. (20% for males vs 10% for females)