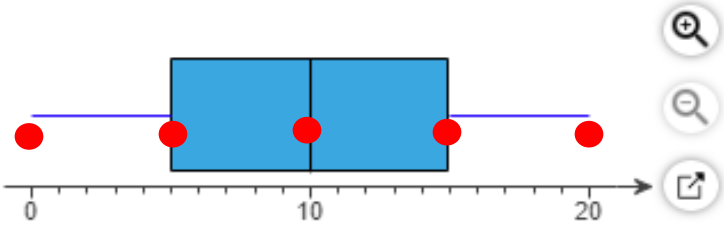
3.5 The Five-Number Summary and Boxplots MATH 241 THOMPSON

1. (a) Identify the shape of the distribution, and (b) determine the five-number summary. Assume that each number in the five-number summary is an integer.



Stat crunch GRAPH HISTOGRAM

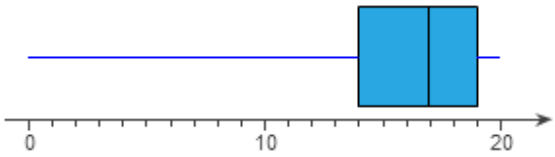
Hover over box to get
5 numbers:
lower, Q_1, Q_2,Q_3, upper

a. Choose the correct answer below for the shape of the distribution.

- A. The distribution is skewed right.
- B. The distribution is skewed left.
- ✔ C. The distribution is roughly symmetric.
- D. The shape of the distribution cannot be determined from the boxplot.

b. The five-number summary is 0, 5, 10, 15, 20. look at red dots

another example for #1



a. Choose the correct answer below for the shape of the distribution.

- A. The distribution is skewed right.
- ✔ B. The distribution is skewed left.
- C. The distribution is roughly symmetric.
- D. The shape of the distribution cannot be determined from the boxplot.

b. The five-number summary is 0, 14, 17, 19, 20.

In a boxplot, left whisker is substantially longer than the right whisker, the distribution is skewed left

2. After giving a statistics exam, Professor Dang determined the following five-number summary for her class results.

56 64 73 85 94

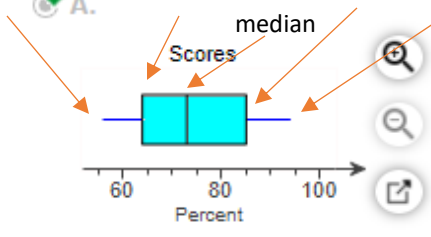
Use this information to draw a boxplot of the exam scores.

Stat crunch GRAPH HISTOGRAM

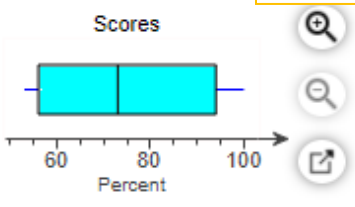
Hover over box to get
5 numbers:
lower, Q_1, Q_2,Q_3, upper
which are the points on the box plot

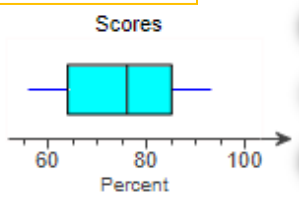
Choose the correct graph below.

lower ✔ A. Q_1 Q_2 Q_3 upper

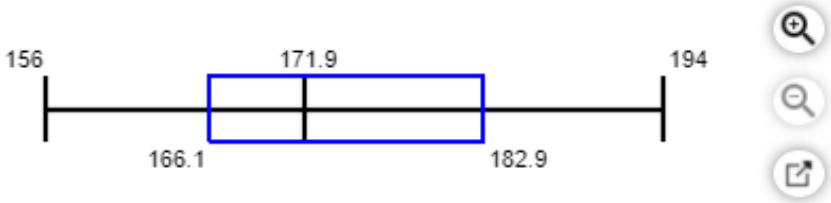


B.





3. The boxplot shown below results from the heights (cm) of males listed in a data set. What do the numbers in that boxplot tell us?



The minimum height is 156 cm, the first quartile Q_1 is 166.1 cm, the second quartile Q_2 (or the median) is 171.9 cm, the third quartile Q_3 is 182.9 cm, and the maximum height is 194 cm.
(Type integers or decimals. Do not round.)

4. Data was gathered on the number of blue M&M's in 2-pound bags of M&M's.
The five-number summary is 21, 38, 45, 55, 77.

In a single bag, there are typically 45 blue M&M's, with the middle 50% of the bags containing more than 38 but less than 55 blue M&M's.

Only one-fourth of the bags were observed to have fewer than 38 blue M&M's, and the most number of blue M&M's found in a bag was 77.

5. Which of the following is NOT a value in the 5-number summary?

Choose the correct answer below.

- ✔ Mean

6. Provided below is a simple data set for you to practice finding descriptive measures. For the data set, complete parts (a) through (c).

1, 2, 4, 6, 8, 1, 2, 4, 6, 8

a. Obtain the quartiles.

Q_1 = 2

Q_2 = 4

Q_3 = 6

Open a new window for Stat crunch
click Open sttcrunch
enter the data in column 1
GRAPH – BOX PLOT
⊗ Draw boxes horizontally
Hover over box

b. Determine the interquartile range.

The interquartile range is 4. (Type an integer or a decimal. Do not round.)

c. Find the five-number summary.

1, 2, 4, 6, 8
(Type integers or decimals. Do not round. Use ascending order.)

7. Listed below are the measured radiation absorption rates (in W/kg) corresponding to 11 cell phones. Use the given data to construct a boxplot and identify the 5-number summary. Create your boxplot by hand first, and then select the one which matches the one you created.

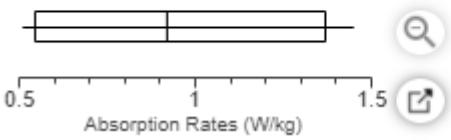
1.28	1.01	1.44	0.54	0.84	1.12	0.58	0.92	1.45	0.51	0.78

The 5-number summary is .51, .58, .92, 1.28, and 1.45, all in W/k
(Use ascending order. Type integers or decimals. Do not round.)

Which boxplot below represents the data?

Open a new window for Stat crunch
click Open sttcrunch
enter the data in column 1
GRAPH – BOX PLOT
⊗ Draw boxes horizontally
Hover over box

A.



B.

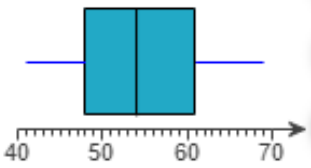
8. The data represent the age of world leaders on their day of inauguration. Find the five-number summary, and construct a boxplot for the data. Comment on the shape of the distribution.

67	44	56	44
63	51	55	46
61	43	46	51
50	52	48	

The five-number summary is 43, 46, 51, 56, 67.

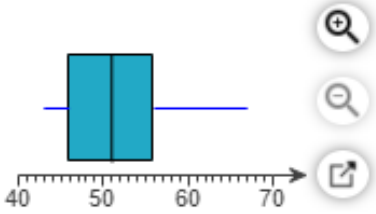
Choose the correct boxplot of the data below.

A.



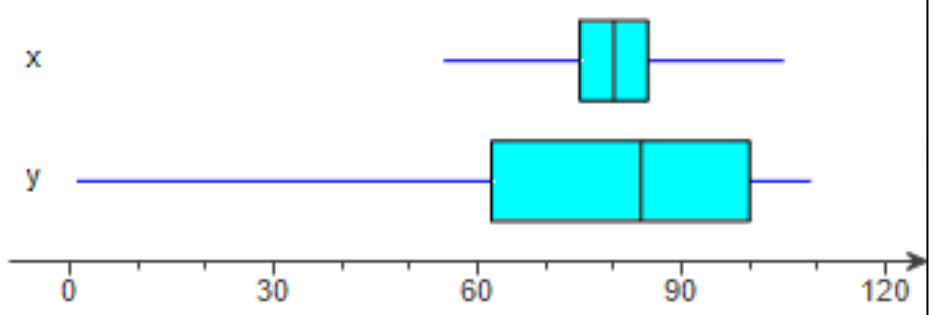
Open a new window for Stat crunch
click Open sttcrunch
enter the data in column 1
GRAPH – BOX PLOT
⊗ Draw boxes horizontally
Hover over box

C.



B. The distribution is skewed to the right.

9.



Dispersion is the degree to which the data are spread out. A measure of dispersion that is based on quartiles is the interquartile range, the range of the middle 50% of the observations in a data set. The more spread a set of data has, the higher the interquartile range will be.

(a) What is the median of variable x?

The median of variable x is 80 . (Round to the nearest integer as needed.)

(b) What is the third quartile of variable y?

The third quartile of variable y is 100 . (Round to the nearest integer as needed.)

Variable y—the interquartile range of variable y is larger than that of variable x.

Skewed left—the median is right of center in the box and the left whisker is longer than the right whisker.

10. The following data represent the dividend yields (in percent) of a random sample of 28 publicly traded stocks. Complete parts **(a)** to **(c)**.

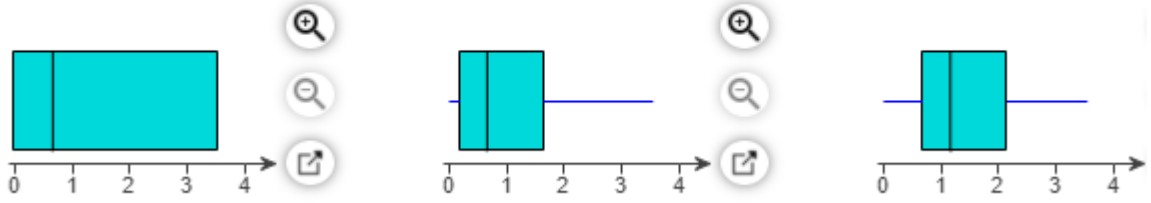
1.67	0.22	0.05	0	0.14	2.8	0.56
0.92	0.01	2.45	0	1.51	2.86	0.48
0.47	3.54	1.11	0.41	1.31	2.97	0
0.37	0	0.76	1.35	2.21	0.28	1.6

(a) Compute the five-number summary.

The five-number summary is 0 , .18 , .66 , 1.64 , 3.54 .
(Round to two decimal places as needed. Use ascending order.)

(b) Draw a boxplot of the data.

A. B. C.



(c) Determine the shape of the distribution from the boxplot.

A. The distribution is skewed to the right.

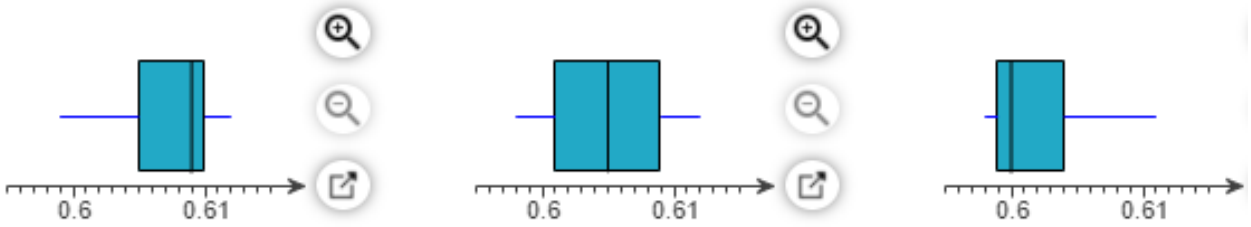
11. The following data represent the weight (in grams) of a random sample of 13 medicine tablets. Find the five-number summary, and construct a boxplot for the data. Comment on the shape of the distribution.

0.599	0.601	0.611	0.598
0.600	0.600	0.602	0.604
0.599	0.604	0.600	0.606
0.598			

The five-number summary is .598, .599, .6, .604, .611.
(Use ascending order.)

Choose the correct boxplot of the data below.

A. B. C.



A. The distribution is skewed to the right.

12. The data to the right represent the number of chocolate chips per cookie in a random sample of a name brand and a store brand. Complete parts **(a)** to **(c)** below.

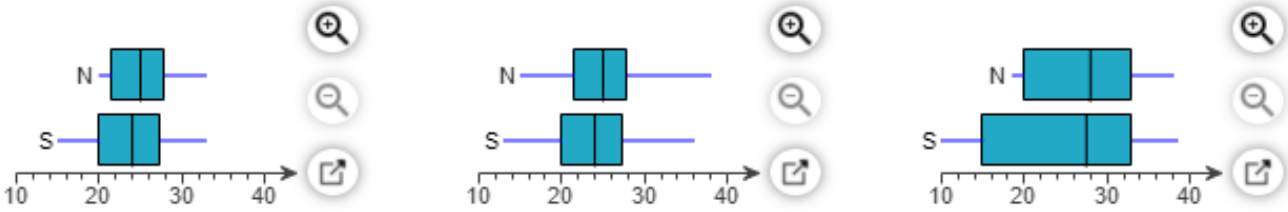
Full data set

Name Brand			Store Brand		
22	20	26	33	28	28
29	25	33	19	17	22
22	23	27	24	15	21
30	25	20	24	23	26
21			27		

Stat crunch GRAPH – BOX PLOT

⊗ use fences to identify outliners
⊗ Draw boxes horizontally
may graph STORE on top NAME on bottom in here

A. B. C.



(b) Does there appear to be a difference in the number of chips per cookie?

A. Yes. The name brand appears to have more chips per cookie.

the median for name brand is higher

(c) Does one brand have a more consistent number of chips per cookie?

A. No. Both brands have roughly the same number of chips per cookie.

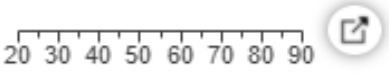
B. Yes. The name brand has a more consistent number of chips per cookie.

the distribution is smaller

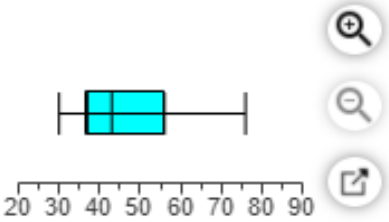
13. Use the same scale to construct boxplots for the ages of the best actors and best actresses from the accompany data sets. Use the boxplots to compare the two data sets.

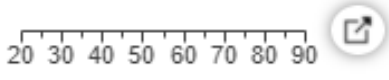
Click on the icon to view the data sets.

Actors

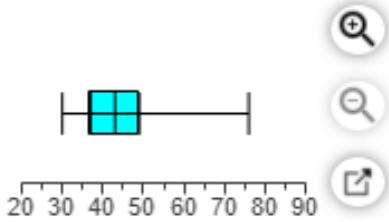


C.



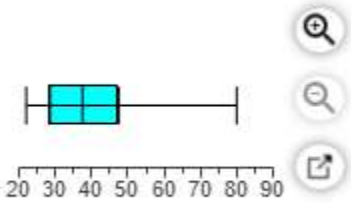


D.

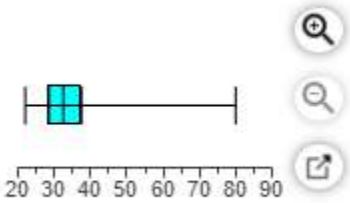


Actresses

C.



D.



Although actresses include the oldest age, the boxplot representing actresses shows that they have ages that are generally lower than those of actors.

14. Which of the accompanying boxplots likely has the data with the larger standard deviation? Why?

Click the icon to view the boxplots.

Choose the correct answer below.

A. Boxplot I likely has the data with the larger standard deviation because the IQR is smaller than that of Boxplot II, which likely results in a larger standard deviation.

B. Boxplot II likely has the data with the larger standard deviation because the boxplot appears to have a greater spread, which likely results in a larger standard deviation.

15.

In a boxplot, if the median is to the left of the center of the box and the right whisker is substantially longer than the left whisker, the distribution is skewed right .

16. (This is a reading assessment question. Be certain of your answer because you only get one attempt on th question.)

For a distribution that is symmetric, the left whisker is the same length as the right whisker.

17. (This is a reading assessment question. Be certain of your answer because you only get one attempt on this question.)

For a distribution that is skewed right, the median is left of center of the box.

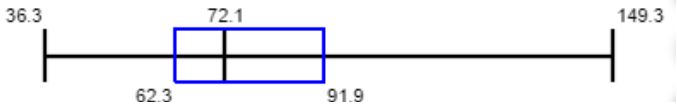
18. (This is a reading assessment question. Be certain of your answer because you only get one attempt on this question.)

For a distribution that is skewed left, the left whisker is longer than the right whisker.

19. Refer to the boxplots available below that are drawn on the same scale. One boxplot represents weights of men, and the other boxplot represents weights of women. Which boxplot represents weights of women? Explain.

Click the icon to view the boxplots.

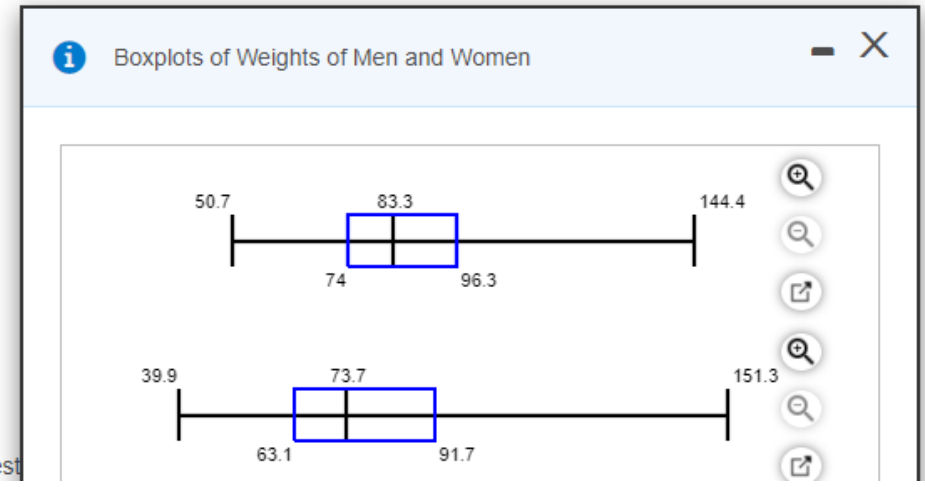
The top boxplot represents weights of women because it depicts weights that are generally lower.



Refer to the boxplots available below that are drawn on the same scale. One boxplot represents weights of men, and the other boxplot represents weights of women. Which boxplot represents weights of women? Explain.

Click the icon to view the boxplots.

The bottom boxplot represents weights of women because it depicts weights that are generally lower.



Quest

EXRA EXAMPLE:

The accompanying data represent the wait time (in minutes) for a random sample of forty visitors to an amusement park ride. Complete parts (a) and (b).

7	14	4	4	3	29	7	4	23	9
7	44	11	9	31	0	26	6	8	5
21	7	20	10	30	16	52	2	10	32
0	4	6	10	5	6	5	15	8	16

(a) Determine and interpret the quartiles.

By the quartiles, about 25 % of the wait times are Q_1 = 5 minute(s) or less, and about 75 % of the wait times exceed Q_1 minute(s); about 50 % of the wait times are Q_2 = 8.5 minute(s) or less and about 50 % of the wait times exceed Q_2 minute(s); about 75 % of the wait times are Q_3 = 18 minute(s) or less, and about 25 % of the wait times exceed Q_3 minute(s).

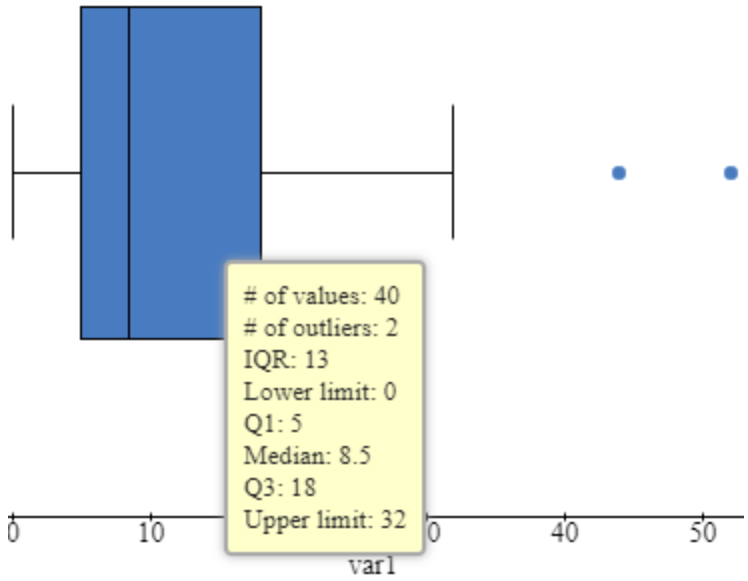
(Type integers or decimals. Do not round.)

(b) Does the data set have outliers? Select the correct choice below and, if necessary, fill in the answer box to complete your choice.

- A. The outlier(s) in the data set is(are) 44,52 .
  (Use a comma to separate answers as needed.)
- B. This data set does not have any outliers.

Stat crunch GRAPH HISTOGRAM

Hover over box to get
5 numbers:
lower, Q_1, Q_2,Q_3, upper
which are the points on the box plot



OUTLINERS: hover over DOT