

Multiple Choice
Ch. 20 Review

- ① D
- ② D
- ③ A
- ④ C
- ⑤ D
- ⑥ B
- ⑦ C
- ⑧ B
- ⑨ A
- ⑩ E
- ⑪ ✓
- ⑫ B
- ⑬ D
- ⑭ D
- ⑮ B
- ⑯ C
- ⑰ A
- ⑱ E
- ⑲ C
- ⑳ D
- ㉑ C
- ㉒ C
- ㉓ B
- ㉔ C
- ㉕ B
- ㉖ A
- ㉗ A

- ㉘ E
- ㉙ E
- ㉚ C
- ㉛ D
- ㉜ A
- ㉝ B
- ㉞ B
- ㉟ E
- ㊱ C
- ㊲ D
- ㊳ B
- ㊴ D

1. New York Yankee Roger Maris held the single-season home run record from 1961 until 1998. Here are Maris's home run counts for his 10 years in the American League:

15 28 16 39 61 33 23 26 8 13

- a) Using your calculator, find Maris's mean number of home runs (\bar{x}) and the standard deviation (s).

$\bar{x} = 26.2, s_x = 15.52$

- b) Determine if this set of data has outliers. Show all work!

$IQR = 33 - 15 = 18$ 61 is an outlier

$(1.5)(18) = 27$

$33 + 27 = 60$ no upper outliers

$15 - 33 = -18$ no lower outliers

- c) Then use your calculator to find \bar{x} and s for the 9 observations that remain when you leave out the outlier. How does the outlier affect the values of \bar{x} and s ? Should the standard deviation be used here to describe the spread of the data?

$\bar{x} = 22.33$ $s_x = 10.14$ Outlier increases \bar{x} and s . Since there are outliers, it would not be wise to use standard deviation as a measure of spread.

2. In the following scenarios would you prefer the five-number summary OR \bar{x} and s as a brief numerical description?

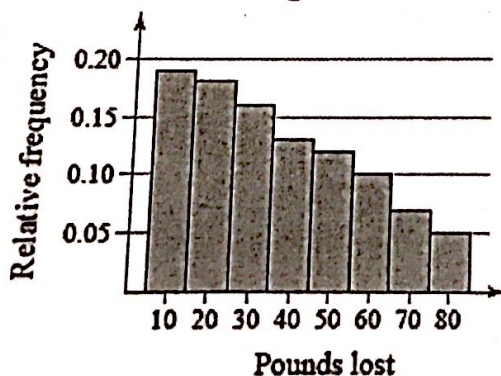
- a) Weight's of cats in the United States

\bar{x}, s summary

- b) Scores on a really easy test

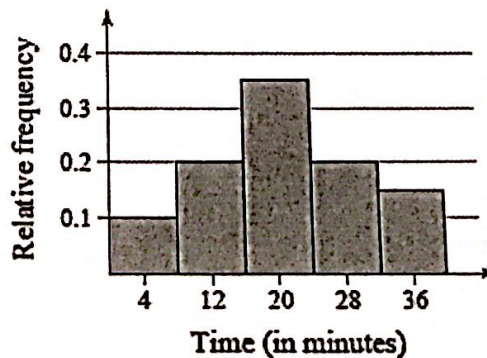
5 # summary

- c) **Weight Loss**



5 # summary

- d) **Waiting Time in a Dentist's Office**

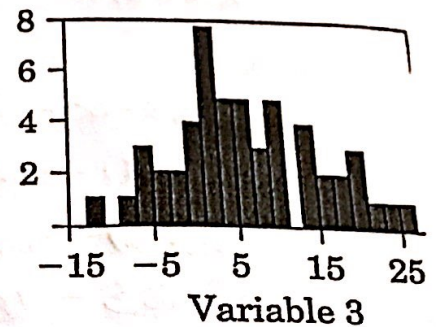
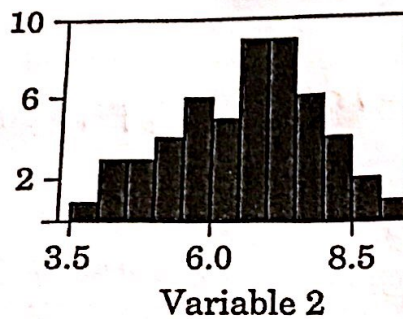
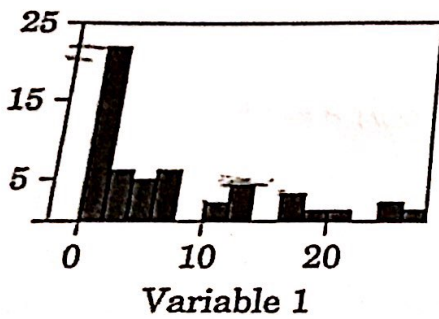


\bar{x}, s

3) You read that the mean income of U.S. households in 2007 was \$67,609. The median household income in 2007 was \$50,233. Explain why the mean is so much higher than the median household income.

There are a few houses that cost more that are outliers which raises the price of \bar{x} .

4) Match the summary statistics with the histograms. Explain how you made your decision.



- a) mean=6.6, median = 6.8, standard deviation = 1.3, variable = #2
 b) mean=6.6, median = 6.0, standard deviation = 8.65, variable = #1
 c) mean=6.6, median = 3.75, standard deviation = 7.4, variable = #3

5) The table below gives the distribution of grades earned by students taking the AP Calculus AB and AP Statistics exams in 2012.

No Data

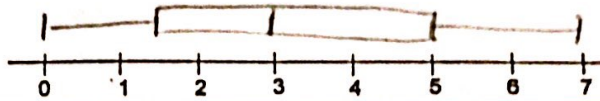
a) Make an appropriate graphical display to compare the grade distribution for AP Calculus AB and AP Statistics.

Find the 5-number summary and draw the boxplot for each data set.
Note: NEVER include the actual median in the data when finding the quartiles.

1. $n = 13$ 0 0 1 2 2 2 3 3 4 5 5 5 7

5-Number Summary:

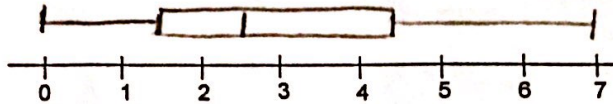
(...0, 1.5, 3, 5, 7...)



2. $n = 12$ 0 0 1 2 2 2 3 3 4 5 5 7

5-Number Summary:

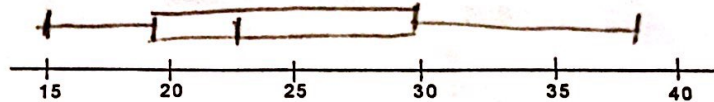
(...0, 1.5, 2.5, 4.5, 7...)



3. $n = 11$ 15 18 19 20 21 23 26 29 30 32 38

5-Number Summary:

(...15, 19, 23, 30, 38...)



4. $n = 10$ 53 56 65 68 72 78 85 86 88 93

5-Number Summary:

(...53, 65, 75, 86, 93...)



OUTLIERS

An outlier is an extreme value in a set of data. It is much smaller or much larger than the other values in your list. To find outlier(s), first find the interquartile range: $IQR = Q_3 - Q_1$. Then, mathematically, an outlier is defined as

$$\text{Outlier} < Q_1 - 1.5 IQR \quad \text{or} \quad \text{Outlier} > Q_3 + 1.5 IQR$$

Determine if this set of data has outliers:

0 16 18 19 19 20 20 21 22 25 27 29 33 37 45 54

Determine the five-number summary: (...0, 19, 21.5, 31, 45...)

Use the mathematical definition of an outlier to determine if this set of data has an outlier.

$$\text{Outlier} < Q_1 - 1.5 IQR \quad \text{or} \quad \text{Outlier} > Q_3 + 1.5 IQR$$

What are outliers for this data set?54.....