Statistics & Data Analysis Review Packet

Mean, Median, Mode, and Range

Mean (also called the average): sum of the values divided by the number of values

Median: the middle value when the values are written in numerical order
~ If a data set has an even number of values, you average the two middle values to find the median.

Mode: the value that occurs most often
~ You can have one, none, or many modes.

Range: the difference between the greatest value and the least value

Example 1: Find the mean, median, mode, and range of the data set.

$$90, 112, 105, 118, 96, 128, 110, 133$$

Mean: __111.5__ $\dfrac{90+112+105+118+96+128+110+133}{8} = \dfrac{892}{8} = 111.5$

Median: _111__ $90, 96, 105, 110, 112, 118, 128, 133 \Rightarrow 90, 96, 105, 110, 112, 118, 128, 133 \Rightarrow$

$$\dfrac{110 + 112}{2} = \dfrac{222}{2} = 111$$

Mode: __none__ $90, 96, 105, 110, 112, 118, 128, 133$ (no data piece occurs more than once)
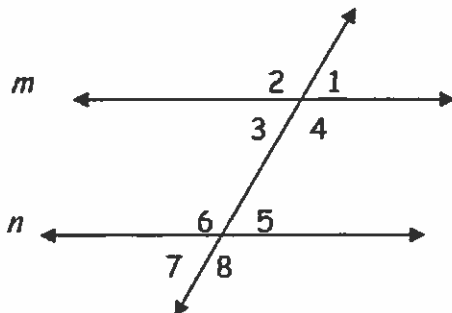
Range: __43___ $133 - 90 = 43$

Find the mean, median, mode(s), and range of the data.

1) $85, 96, 72, 88, 95, 80, 86$

2) $300, 367, 300, 307, 410, 300, 200, 280, 300, 240$

Angle Relationships

4) **In the diagram below, $m \parallel n$. If $m\angle 1 = 54°$, find the missing angle measures.**

3) **Use the diagram to the left to answer a – d.**

a) Identify the 4 pairs of corresponding angles.

b) Identify the 2 pairs of alternate interior angles.

c) Identify the 2 pairs of alternate exterior angles.

d) Name 1 pair of supplementary angles.

Creating a Box-and-Whisker Plot

Box-and-whisker plot: A data display that organizes data values into 4 groups.
* Each of the 4 groups represents 25% of the data.
  - The entire box represents about 50% of the data.
  - Each whisker represents about 25% of the data.
* The **interquartile range** is the difference between the upper and lower quartile.
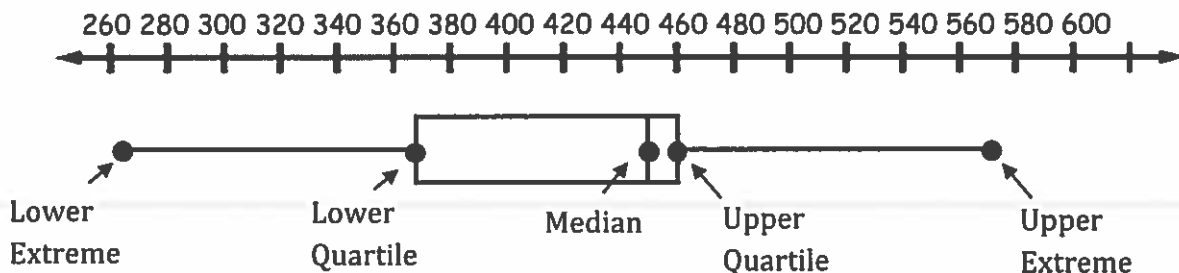
Making a Box-and-Whisker Plot:

- Order the data from least to greatest. Find the median to split the data into upper & lower half.

- Find the median of the upper half and the median of the lower half.
  • The median of the lower half is called the **lower quartile.**
  • The median of the upper half is called the **upper quartile.**

- Identify the least data value and the greatest data value.
  • The least data value is called the **lower extreme.**
  • The greatest data value is called the **upper extreme.**

- Draw a number line and plot the median, quartiles, and extremes.
  • Draw a box from lower quartile to upper quartile & draw a vertical line through the median.
  • Draw a horizontal line (a "whisker") from the edge of the box to each extreme.

Example 1: The data below represents the numbers of oranges a local farmer collected from one of his orange tress for the last 9 years. Create a box-and-whisker plot for the data.

$$572, 452, 457, 460, 360, 407, 380, 458, 264$$

- 264, 360, 380, 407, 452, 457, 458, 460, 572   * Median: 452
- Lower half of data: 264, 360, 380, 407   * Lower Quartile: $\frac{360+380}{2} = 370$
  Upper half of data: 457, 458, 460, 572   * Upper Quartile: $\frac{458+460}{2} = 459$
- Lower Extreme: 264
  Upper Extreme: 572

5) The number of hours 12 middle school students spend doing homework in a month is listed here:

22, 9, 25, 18, 30, 28, 19, 32, 27, 20, 35, 27

Make a box-and-whisker plot of the data.

6) Find the range and interquartile range of the data above.

Interpreting Data Displays

Use the line graph to answer these questions.

7) True or false? Rainfall increased each month.

8) Which month had the greatest Amount of rain?

9) True or False? The amount of rainfall was about the same in May and July.

10) About how much more rain was there in August than April?

Average Rainfall in Erie, PA
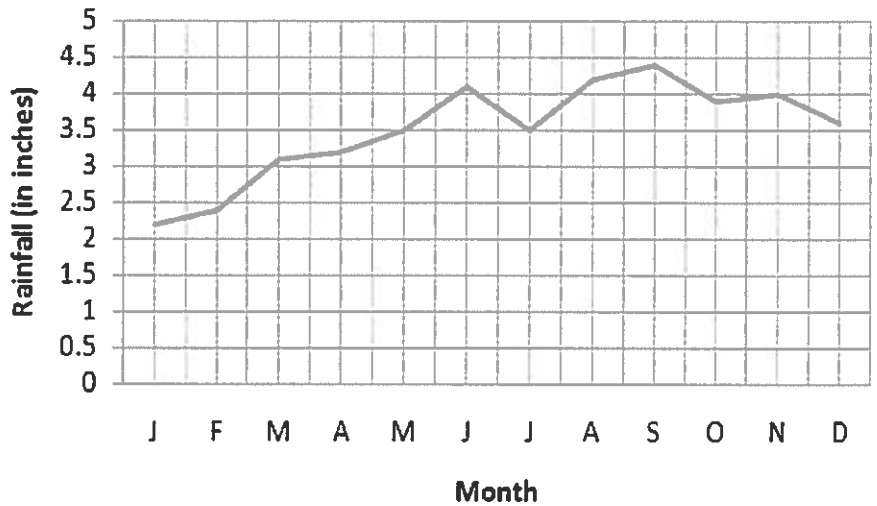
The stem-and-leaf plots below show the prices of pairs of sneakers at 2 shoe stores. Answer the following questions about the plots.

Store A	Store B
3 │ 0 2 5 6 9	3 │ 0 0 5 5 9
4 │ 1 1 5 5 9	4 │ 0 0 2 3 3 5 5 5
5 │ 0 3 4 5 8	5 │ 4 4 5 8
6 │ 0 0 2 5	6 │ 5 9

Key: 3 │ 0 = $30

11) What is Store A's approximate mean?

    a) $ 3.68   b) $47.89   c) $39.57   d) $45.63

12) Which store has the lower median price?

13) In which range are more than half of the sneakers at Store B priced?

   a) $30-39   b) $40-59   c) $50-69   d) $40-49

The two box-and-whisker plots below compare the amount of calories in all menu options at two local restaurants. Use it to answer the following questions.
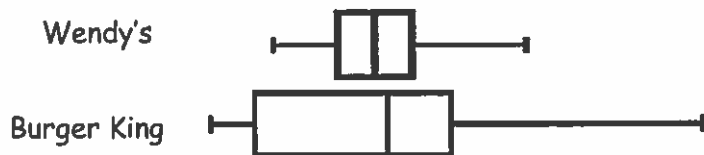
Wendy's

Burger King

14) Which statement best describes the two plots.
    a) The median number of calories is less at Burger King than Wendy's.
    b) About 25% of the Burger King food has less calories than all of the Wendy's food.
    c) All Burger King food has more calories than all Wendy's food.
    d) The interquartile range of calories is larger at Wendy's than at Burger King.

The box and whisker plot below represents the math midterm exam scores. Answer the following questions using it.

Exam Scores

50   60   70   80   90
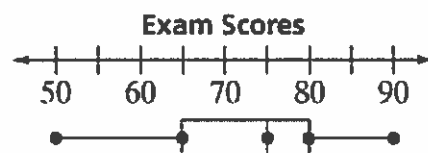
15) What percentage of students scored above a 65%?
    a) 50%   b) 25%   c) 35%   d) 75%

16) Which statement best describes the plot.
    a) The interquartile range is 15.   b) The median is 65.
    c) The upper extreme is 80.   d) The lower quartile is 75.

The histogram below shows the number of weeks into the season each hurricane occurred in 1997-2000. Use it to answer the following questions.

Hurricanes 1997-2000

Hurricanes

12
10
8
6
4
2
0

6--8   9--11   12--14   15--17   18--20   21--23   24--26
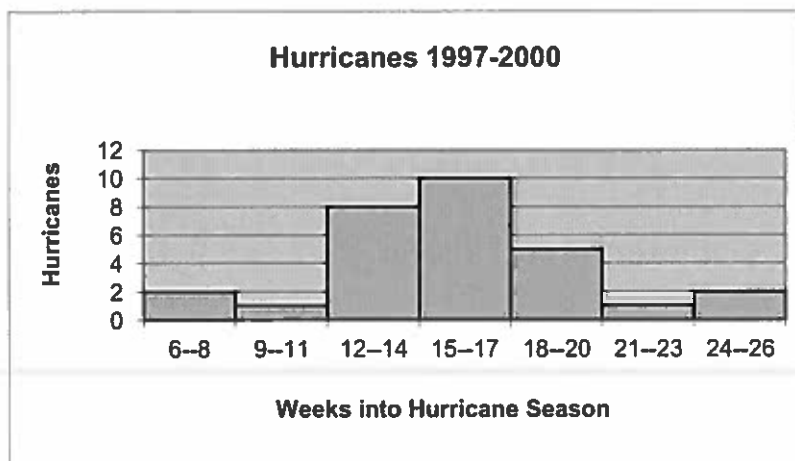
Weeks into Hurricane Season

17) Which interval has the greatest frequency?

18) About what percent of hurricanes occur in the first 14 weeks of the season?
    a) 28%   b) 14%   c) 62%   d) 38%

19) On average, tornadoes usually occur within the first 10 weeks of their season. Which statement best compares the tornado season with the hurricane season?

   a) On average, hurricanes occur earlier in their season than tornadoes.
   b) All tornadoes occur before hurricanes.
   c) On average, tornadoes occur earlier in their season than hurricanes.
   d) On average, tornadoes and hurricanes occur right around the same time in their respective seasons.

Student's Juice Preference

Use the circle graph to answer the questions below.

20) If 150 people were surveyed, how many said they preferred Juice B?

21) If 300 people were surveyed, how many more people preferred Juice A than Juice C?

Juice D 8%
Juice C 30%
Juice A 44%
Juice B 18%
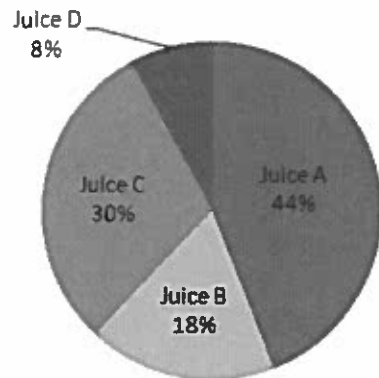
Ratios, Rates & Proportions

22) **Write the ratio is in simplest form.** a) 4 to 18   b) 7 : 9   c) $\frac{39}{13}$

23) **If a telemarketer can make 8 calls in 1 hour, how many calls can they make in 1 day?**

Solving Proportions: To solve proportions, we cross multiply to generate an algebraic equation so that we can solve for the unknown.

24) **Solve each proportion. Show your work.**

   a) $\frac{10}{25} = \frac{x}{10}$ b) $\frac{30}{2y} = \frac{35}{28}$

Probability & Odds

Probability: the likelihood an event will occur

Probability(event) = <u>favorable outcomes</u>
total possible outcomes

Odds in favor: <u>favorable outcomes</u>
unfavorable outcomes

* Probability and Odds are always written in simplest form so be sure to reduce your fractions if it is possible.

25) You have a bag containing colored blocks. There are 4 blue, 3 yellow, 2 green, and 1 orange. Use this information to find the probabilities listed below.

    a) Probability of green =

    b) P(blue or orange) =

    c) P(yellow) =

    d) Probability of not getting blue =

26) You randomly choose a letter from the word M-A-T-H-E-M-A-T-I-C-I-A-N-S.

    a) P(choosing a vowel)=

    b) Odds of choosing a vowel?

    c) P(choosing an M) =

    d) Odds of choosing an A?

27) You randomly choose a number between 1-50.

    a) P(multiple of 4) =

    b) Odds of choosing a multiple of 4?

    c) Odds of choosing an even number?

28) The table shows the population of Texas from 1900 to 2000. Which of the following is true based on the data?

    A) The population doubled from 1980 to 2000.

    B) The population in 1990 was 4 times the population in 1920.

    C) The population more than tripled from 1900 to 1960.

    D) The population in 1980 was 2 times the population in 1950.

Year	Population (thousands)
1900	3,049
1910	3,897
1920	4,663
1930	5,825
1940	6,415
1950	7,711
1960	9,580
1970	11,197
1980	14,229
1990	16,987
2000	20,852

Two Way Tables - Independent Practice Worksheet

Complete all the problems.

After a series of matches between a school's teams and their rivals, the school secretary analyzed the relationship of the number of wins and matches played. The results are summarized in a two way table below.

Sport	Boys Wins	Girls Wins
Volleyball	23	18
Cricket	40	10
Soccer	15	25

1. How many total wins did the school's Volleyball teams register?

2. How many more wins did the boys' teams have than the girls' teams?

The data is summarized in a two-way table for the number of boys and girls that regularly drink water, lemonade, or soda at lunch.

	Boys	Girls	Total
Water	45	32	77
Soda	50	38	88
Lemonade	42	32	74

3. What is the percentage of boys that regularly drink water?

4. What is the percentage of girls that regularly drink water?

5. What is the percentage of girls that regularly drink soda?

6. What is the percentage of boys that regularly drink soda?

7. What is the percentage of the students that regularly drink water?

8. What is the percentage of the students that regularly drink soda?

Name _____   Date _____

Below you will find an incomplete two ways table that shows the number of girls and boys that were passing Economics and Science. There are a total of 72 boys and 72 girls taking Economics. There are 78 boys and 60 girls taking Science.

Gender	Passing Economics	Failing Economics	Passing Science	Failing Science	Total
Boys	61		69		
Girls	67		53		
Total					

9. Calculate the percentage of boys' passing Economics.

10. Calculate the percentage of girls' passing Science.

11. Calculate the percentage of students passing Science.

12. Calculate the percentage of students passing Economics.

13. Complete the two way table above.

Name _____   Date _____

Complete all the problems.

Consider the following three Data Sets A, B and C.

A= {1, 2, 3, 4, 5}

B = {2, 2, 2, 2, 2}

C = {5, 7, 3, 11, 14}

1. Calculate the mean of each data set.

2. Calculate the standard deviation of each data set.

3. Which set has the largest standard deviation?

4. Is it possible to answer question "c" without calculations of the standard deviation?

The frequency Table is shown below.

Number of Children	frequency
20	4
40	2
30	6
80	5
60	3

5. Calculate the mean of the number of children of frequency.

6. Calculate the standard deviation of the number of children of frequency.

Consider the following three Data Sets A, B and C.

A= {7, 3, 2}

B = {6, 5, 13}

C = {8, 2, 2}

7. Calculate the mean of each data set.

8. Calculate the standard deviation of each data set.

9. Which set has the largest standard deviation?

10. Is it possible to answer question "c" without calculations of the standard deviation?

Scatter Plots of Linear Functions – Independent Practice Worksheet

Solve the problems. What are the equations of the trend line shown here?

1.

2.

3.

4.

5.

6.

7.

8.

9.

10.