11.1 Stem and Leaf Plots and Histograms

Stem and Leaf Plot is a simpler way to show data than just a list. There are 2 kinds: ordered and unordered.

Let’s say we look at the ages of people attending a community fundraiser.

51, 40, 42, 24, 33, 46, 27, 7, 9, 20, 30, 55, 56, 48, 28, 28, 51

The stems would be \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

It is easier to make an unordered one, but the ordered one is more useful. Remember – all stem and leaf plots must have a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

Mode (MOST common): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Median (middle): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

* The stem can be more than one digit, but the leaf must be one digit.
* If you are using a stem and leaf plot to show a decimal value, do not use a decimal point in the data display. Use the key to show that it is a decimal value.

You can also easily spot the high and low, which allows you to calculate the range (high – low).

You can also find the number of data points above (or below) a given point.

How many people are under the age of 30? \_\_\_\_\_\_\_\_\_\_\_\_\_\_

Frequency tables – these start out with tally marks. Tally marks are 4 vertical lines followed by the 5th diagonal line through the other 4. Most people recognize that as “5”. Sometimes they are a count of a specific number or sometimes they are a count of a range:

Madison asked 20 classmates how many songs they downloaded in the past month. The numbers are listed below. Make a frequency table of the data using intervals of 5.

12,12,12,2,6,7,18,19,7,4,13,13,7,8,9,4,10,14,14,14

|  |  |  |
| --- | --- | --- |
| Songs | Tally | Frequency |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

You can also make a histogram from the same data. A histogram looks much like a bar graph, however a histogram has numbers across the bottom. A bar graph has words (like colors, foods, animals, etc). Draw a histogram of the CD data above, using the same interval of 5 from the frequency table.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

If you are not given the range to use, you want 4 to 8 bars. Any less or more is not as effective in showing your data.

Which is the most common range of songs downloaded? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

How many students downloaded less than 5 songs? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

How many students downloaded more than 15 songs? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

HW: 11.1 (page 614) # 6, 8, 10- 11, 15-16, 30

#10: First interval is 1-3; #11: First interval is 101-120. You must make the rest of the intervals.

Pie Charts/Circle graphs (pages 618-619)

Yesterday we learned about bar graphs, frequency tables, and stem and leaf plots. These are very useful to show relations among different sets. A pie chart can take the same data and ***compares*** the ***parts*** of the data to the ***whole*** set.

Favorite colors of this class:

Blue red green yellow

Orange pink purple other

First we find the total number of students who answered: \_\_\_\_\_\_\_\_\_\_

Now find the ratios that compare the number of students who like a certain color to the total number of students and also turn that into a percent.

Blue red green yellow

Orange pink purple other

How many degrees are in a circle? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

So if everyone liked blue, we would use all 360o. Since not everyone picked blue, we set up a proportion and find the number of degrees needed. Round to the nearest degree.

Example: Let’s say 10 out of 83 students chose blue.

Convert that to a decimal. Then multiply by 360 (degrees in a circle).

Remember that 90o is a right angle. 180o is a straight line.

Blue red green yellow

Orange pink purple other

Using this circle, you could create a pie chart.

Most often, we are asked to interpret circle graphs. We do this by comparing the size of the sections that we see. On most circle graphs the percentage is given. Sometimes instead, the actual number is shown.

Go to your book, page 618. Review these example problems.

HW: pg. 619 # 1-4

11.2 Box-and-Whisker Plots

The RANGE of a set of data is the difference between the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ and the \_\_\_\_\_\_\_\_\_\_\_\_\_\_ number in the set.

The INTERQUARTILE RANGE is the range of the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ of the data.

Finding the interquartile range is a multi-step process.

1. Put the data in ascending order and find the median.

2. Now find the median of the upper half of the data.

This is called the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

3. Find the median of the lower half of the data.

This is called the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

4. The interquartile range (IQR) is the difference between the upper quartile and the lower quartile.

The upper quartile is the upper \_\_\_\_\_\_% of the data

The lower quartile is the lower \_\_\_\_\_\_\_% of the data

An \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ is an extreme value in a data set.

There is a mathematical way to tell if something is an outlier, but we are just going to look to see if it is very far above or below the data.

Here is the number of home runs that Babe Ruth hit each year of his career (from Algebra 2 with Trigonometry, Prentice Hall, page 692).

$$0, 4, 3, 2, 11, 29, 54, 59, 35, 41, 46, 25, 47, 60, 54, 46, 49, 46, 41, 34, 22, 6$$

Put these in order, and then find the upper quartile, lower quartile, interquartile range, and any outliers

Median: \_\_\_\_\_\_\_\_ Upper Quartile: \_\_\_\_\_\_\_\_\_\_\_ Lower Quartile: \_\_\_\_\_\_\_\_\_\_\_

 Upper Extreme: \_\_\_\_\_\_\_\_\_\_\_ Lower Extreme: \_\_\_\_\_\_\_\_\_\_\_

Now let’s do a Box and Whisker with our data.

Things you can tell from a box and whisker plot:

The median, the range of the middle 50% of the data points, the range of the upper or lower 25% of the data points

Things you CANNOT tell form a box and whisker plot:

The mean, the mode, the number of data points

Uses: SAT for choosing a college, used in keeping a process in control.

Turn to page 621. Look at example 2.

Half of the state parks had \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ or fewer visitors.

What percent of the parks had 18.2 million visitors or more? \_\_\_\_\_\_\_\_\_\_

What is the total number of visitors to the state parks? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

What percent visited California? \_\_\_\_\_\_\_\_\_\_\_\_\_\_

What is the mean number of visitors? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Is the mean or the median a better measure of central tendency? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Turn to page 622. Look at example 3.

Which city had the larger range of monthly temperatures? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Which city had the highest average monthly temperature? \_\_\_\_\_\_\_\_

The middle half of the month’s average temperature in Seattle fall between

\_\_\_\_\_\_\_\_\_\_\_ and \_\_\_\_\_\_\_\_\_\_.

The three coldest months in Boston (3 out of 12) have average temperatures between

\_\_\_\_\_\_\_\_\_ and \_\_\_\_\_\_\_\_\_.

HW: 11.2 (page 623) #8,9,10, 21-23

Changes in Data Values and Making Data Displays

Pages (626 – 629)

If I change a set of data points, how do the mean, median, and mode change?

Let’s start simple: 1, 2, 3, 3, 4, and 5

Mean \_\_\_\_\_\_\_ Mode \_\_\_\_\_\_\_\_\_\_ Median \_\_\_\_\_\_\_

What happens when I multiply each by 2?

Mean \_\_\_\_\_\_\_ Mode \_\_\_\_\_\_\_\_\_\_ Median \_\_\_\_\_\_\_

What happens if I add 5 to each?

Mean \_\_\_\_\_\_\_ Mode \_\_\_\_\_\_\_\_\_\_ Median \_\_\_\_\_\_\_

What happens if I add a new value? Let’s add 6.

Mean \_\_\_\_\_\_\_ Mode \_\_\_\_\_\_\_\_\_\_ Median \_\_\_\_\_\_\_

What if I take out a value? Let’s get rid of 1.

Mean \_\_\_\_\_\_\_ Mode \_\_\_\_\_\_\_\_\_\_ Median \_\_\_\_\_\_

Think through carefully when changing data sets.

Other displays you may encounter:

**Line plots (also called Dot plots)** are similar to bar graphs, except a number line is used to organize the data. It is very useful to use as you collect data.

Number of cousins: Just using x’s or dots.

----+------+------+------+------+------+------+------+------+------+------+------+------+------+-----+-

 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14

**Bar graphs** are very similar to histograms. Remember histograms are ALL NUMBERS. Bar graphs have words at the bottom.

We could do a bar graph of favorite colors. We would do a histogram of student heights.

Bar Graph: Histogram:



\*Notice this has categories on the bottom. \*This has a numerical range along the bottom.

How many students chose soccer as their favorite sport? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

What was the most common grade range on the test? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

A **line graph** is used to show a relationship between a set of data and time. Time would be on the BOTTOM axis.



Between which two days was the biggest increase? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Which day had the highest sales? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Scatter plots** are similar. They show the relationship between two sets of data.

Scatter plots can show a positive correlation, a negative correlation or no correlation. We learned about these in Chapter 8.



Review:

What type of correlation is shown?

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

HW: page 627 # 1 – 3, page 629 # 1-5

11.3 Choosing an Appropriate Display

When given data in a table, sometimes we have to choose the appropriate way to display the data. Some things to remember when choosing:

* A \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ shows frequency of pieces of one set of data, along with the range.
* A \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ or (\_\_\_\_\_\_\_\_\_\_\_\_\_) shows the relationship between parts of a data set to the whole.
* A \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ is used when there is only one set of numbers (example – favorite colors).
* A\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ has two sets of data; one set of data must divide into intervals. This display looks like a bar graph.
* A \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ has two sets of data and shows relationships between the two sets. There can be a positive relationship, negative relationship, or no relationship.
* A \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ shows change over time.
* A \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ organizes one set of data by separating the last digit and grouping in rows.
* A \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ organizes one set of data into 4 quartiles.

Note that there are two types of data: categorical and numerical. An example of categorical data would be favorite colors and the number of students who picked each. Note that they are choosing a category. Numerical data would be number of students who are a certain height.

Which displays allow for categorical data? \_\_\_\_\_\_\_\_\_\_\_\_\_\_ and \_\_\_\_\_\_\_\_\_\_\_\_\_\_

Give an example that would be appropriate for:

a line plot \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

a circle graph \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

a bar graph \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

a histogram \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

a scatter plot \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

stem and leaf \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

box and whisker \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

line graph \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Data displays can also be misleading, sometimes on purpose. A person can use large increments on an axis, use a break in the axis or make different widths. Sometimes these are appropriate; sometimes they are misleading.

(See book, page 631)

Look at example 2 (page 631) – the same data is displayed both ways. How do they differ?

Which do you like better?\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

What data display would show each individual score? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Look at the misleading displays at the bottom of 631.

* Broken y-axis, large increments, and different bar widths can all be misleading.

Sometimes you choose which average you want (mean, median, or mode). Sometimes one of those averages does not really represent the data but it shows what you want to show.

Example: You want someone to believe the salaries are high where you work.

The hourly rates for the 10 people in the office are

$10, $10, $12, $22, $9, $11, $10, $11, $9, $10

Would you use mean, median, or mode to entice them to work with you?

HW: 11.3 Practice A #1-12

Data Distributions

Pg. 634 A

The \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ is an average of how much the data values differ from the mean. It is another measure to show how the data is dispersed. It’s advantage over the range and the interquartile range is that is uses all the values when it is calculated.

Step 1: Find the mean of your data.

The following are the heights (in inches) of 11 WNBA players.

66, 66, 66, 71, 72, 73, 74, 74, 76, 77, 77

Mean: \_\_\_\_\_\_\_\_\_\_\_\_

Step 2: Find the difference between the mean and each value. Your final answer will be the absolute value of the difference.

|  |  |  |
| --- | --- | --- |
| Height | Deviation from Mean | Absolute Deviation |
| 66 |  |  |
| 66 |  |  |
| 66 |  |  |
| 71 |  |  |
| 72 |  |  |
| 73 |  |  |
| 74 |  |  |
| 74 |  |  |
| 76 |  |  |
| 77 |  |  |
| 77 |  |  |

Step 3: Find the mean of the absolute deviations.

The mean absolute deviation is about \_\_\_\_\_\_\_\_\_\_\_ inches.

Find the mean absolute deviation of the height (in inches) of 16 NBA players.

73, 75, 76, 77, 78, 79, 79, 79, 80, 81, 81, 82, 82, 82, 84, 84

Mean: \_\_\_\_\_\_\_\_\_

|  |  |  |
| --- | --- | --- |
| Height | Deviation from Mean | Absolute Deviation |
| 73 |  |  |
| 75 |  |  |
| 76 |  |  |
| 77 |  |  |
| 78 |  |  |
| 79 |  |  |
| 79 |  |  |
| 79 |  |  |
| 80 |  |  |
| 81 |  |  |
| 81 |  |  |
| 82 |  |  |
| 82 |  |  |
| 82 |  |  |
| 84 |  |  |
| 84 |  |  |

Mean absolute deviation: \_\_\_\_\_\_\_\_\_\_\_\_

HW: pg. 634 B #1-3 (skip 3a)

11.4- Collecting Data

One way to collect data about a group is to conduct a survey.

The entire group that you want information about is called a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

It is usually very difficult to survey every member in a population, so instead we can use a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_. This is where you collect data from part of a population.

Sampling Methods:

* \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_: Every member of a population has an equal chance of being selected.
* \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_: A rule is used to select members of a population.
* \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_: The population is divided into distinct groups. Then members are selected from each group.
* \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_: Only members of the population who are easily accessible are selected.
* \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_: Members of the population can select themselves by volunteering.

Choose which sample method is used.

1. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_- A school newspaper interviews every 5th student that enters the building about a proposed new gymnasium.
2. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_- A cable company emails 750 residents under 30 years old and 750 residents 30 years and older to determine the most watched stations.
3. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_- During a flight, a newspaper reporter surveys 10 passengers seated near him about airport security.
4. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_- A magazine invites its readers to email their answers to a questionnaire.
5. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_- A study wants to learn about the sleep habits of Americans. It uses a random digit dialing process to call houses to take part in the survey.

A \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ sample is not representative of the population. The best way to avoid this is to use random sampling. However, this method usually takes more time, effort, and money.

Identifying potentially biased samples:

A city council wants its residents to decide if they should build one of the following projects:

* Build a new library
* Build a new city hall
* Build a new park

Because they can’t survey every resident, it decides to survey a sample. Do you think the following methods could result in a biased survey? Explain.

A. Survey the city council members and their families. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

B. Survey people at a local park. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

C. Survey people at a local mall. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Biased Questions:

The questions in a survey should be free from your own opinion.

Tell if the following questions are biased. If so, rewrite it and remove the bias.

1. Don’t you agree that planting more trees will make our beautiful downtown area even better?

 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. Are you willing to pay higher taxes to support the building of a new stadium, even though we already have a stadium?

 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. How many times per month do you use the new library? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

HW: 11.4 (page 638) #5-14

11.5- Interpreting Data

Once you collect your data from your sample, you can use that data to make predictions about the entire population.

In a survey of 5000 randomly selected American households, 460 said that they watched This Is Us. Of all American households (105 million), about how many watch Duck Dynasty.

Step 1: Find the percentage of the sample that watch Duck Dynasty.

Step 2: Use that percentage to make a prediction as to how many total US households watch This Is Us. You can use either method to solve.

Margin of Error:

Due to the variations in responses from different samples, a survey should include a margin of error.

A survey of a random sample of voters predicts that Mr. Jenkins will receive 52% of the votes and Mrs. Williams will receive 48%. The margin of error is $\pm 3\%.$ Can you predict who will win the election?

Use the margin of error to find the interval of the percentage of votes that each candidate is likely to receive.

Mr. Jenkins: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Mrs. Williams: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Can you predict who will win? Why or why not? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

HW: 11.5 (page 646) #5-14, Challenge #15

Two-Way Tables

Pg. 647 A

A \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ shows observed frequencies for two categories of data collected from the same subjects. \*Subjects are the people or objects being measured in the study.

Example 1:

You randomly survey students in your school asking if they can wiggle their ears. The results of the survey are shown below.

Add the frequencies in each column and row.

|  |  |  |
| --- | --- | --- |
|  | **The student…** |  |
| **Gender** | **Can wiggle:** | **Cannot wiggle:** | Total: |
| Male | 24 | 161 |  |
| Female | 20 | 130 |  |
| Total |  |  |  |

 How many students can wiggle their ears? \_\_\_\_\_\_\_\_\_\_\_

 How many cannot? \_\_\_\_\_\_\_\_\_\_\_

 How many students were surveyed? \_\_\_\_\_\_\_\_\_\_\_\_\_\_

 What percent of the male students only can wiggle their ears? \_\_\_\_\_\_\_\_\_\_\_\_\_

 What percent of the female students only can wiggle their ears? \_\_\_\_\_\_\_\_\_\_\_

 What percent of the entire population surveyed can wiggle their ears? \_\_\_\_\_\_

Example 2:

A car dealership surveys customers asking what is more important when deciding to by an alternative-fuel vehicle.

16-40 years old:

 128- saving money

 144- helping environment

 48- money and environment

 More than 40 years old:

 170- saving money

 75- helping environment

 5- money and environment

Fill in the two-way table. Find the totals for each column and row.

|  |  |  |
| --- | --- | --- |
|  | **Survey response:**  |  |
| **Age:** | **Saving Money:** | **Helping Environment:** | **Money and Environment:**  | Total: |
| 16 to 40 years old:  |  |  |  |  |
| More than 40 years old:  |  |  |  |  |
| Total |  |  |  |  |

Fill in the percentages for each section.

|  |  |
| --- | --- |
|  | **Percent:**  |
| **Age:** | **Saving Money:** | **Helping Environment:** | **Money and Environment:**  |
| 16 to 40 years old:  |  |  |  |
| More than 40 years old:  |  |  |  |

Describe the relationship between age and response.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

HW: pg. 647B #1-3

Review for Chapter 11/Word Problems

Questions that I guarantee are on the test:

1. Mean, median, mode

2. Upper Quartile, Lower Quartile, Interquartile Range

Remember:

Upper Quartile (UQ) is the median of the Upper half of the data.

Lower Quartile (LQ)is the median of the Lower half of the data.

The Interquartile Range (IQR) is the (Upper Quartile – Lower Quartile)

3. Several questions asking if something is a positive, negative, or no relationship.

4. Several questions asking about which type of graph would be best to represent the data given

5. Question dealing with misleading graphs – are they misleading, how are they misleading?

6. Questions that ask you to look at one of the types of graphs we learned this chapter and answer questions about that graph.

Pre-Algebra Assignment:

1. Collect data that is interesting to you. Your data must have at least 20 numerical values.

If you can’t find enough values, see me before you begin your project.

Once you have your data, please turn it in to Mrs. Hines for review and approval.

2. Create the following to display your data:

 Stem-and-Leaf plot

 Box-and-Whisker plot

3. You can choose to display either:

 Line Graph or Scatterplot

 \*To create a Line Graph- You must find data that changes over time.

\*To create a Scatterplot- You must find bivariate data. (Example: List the quarterbacks with the most touchdowns and their ages.)

4. Use the Line Graph or Scatterplot that you created to make a prediction, inference, or conclusion.

5. Calculate the mean, median, mode, and range using a spreadsheet program (Excel or Numbers).

6. Which measure of central tendency would you use to describe your data? Why?

Total: 30 points

Note that 3 points will be subtracted for each day late.