Mathematics

8.12 The student uses statistical procedures to describe data. The student is expected to:

(A) select the appropriate measure of central tendency or range to describe a set of data and justify the choice for a particular situation.

(B) draw conclusions and make predictions by analyzing trends in scatterplots.

(C) Select and use an appropriate representation for presenting and displaying relationships among collected data, including line plots, line graphs, stem and leaf plots, circle graphs, bar graphs, box and whisker plots, histograms, and Venn diagrams, with and without the use of technology.

Technology Applications

The student demonstrates knowledge and appropriate use of hardware components, software programs, and their connections. The student is expected to:

(1)(a) demonstrate knowledge and appropriate use of operating systems, software applications, and communicate and networking components.

(1)(c) demonstrate the ability to select and use software for a defined task according to quality, appropriateness, effectiveness, and efficiency.

(1)(f) perform basic software application function including, but not limited to, opening an application program and creating, modifying, printing, and saving documents.

The students uses appropriate computer-based productivity tools to create and modify solutions to problems. The student is expected to:

(7)(a) plan, create, and edit documents created with a word processor using readable fonts, alignment, page setup, tabs, and ruler settings.

Materials

Advanced Preparation:

- Student copies of **RoundandRound, BabyName,** and **WhatName** spreadsheets or copies available on a network
- Internet access to **http://www.ssa.gov/OACT/babynames/**
- Printer access
- Copies of **Round and Round** and **What's In A Name** worksheets for each student

For whole class demonstration:

- Transparencies 1 - 5 (2 copies of Transparency 4)
- Several hula hoops (ideally, one for every 2-3 students)

For each student:

- **Round and Round** worksheet
- Access to **RoundandRound** spreadsheet
- Access to **BabyName** spreadsheet
- **What's In A Name** worksheet
- Access to **WhatName** spreadsheet

ENGAGE

The Engage portion of the lesson is designed to create student interest in the concepts addressed. This part of the lesson is designed for whole class discussion/ demonstration.

1. Prompt a student to read **Transparency 1** and ask students to individually consider their prediction.
2. Prompt students to share their predictions with a neighbor before getting responses from the large group.

Facilitation Questions

- What data are you collecting and comparing for each student?
  We are looking for, and comparing, the number of revolutions on the first attempt and the number of revolutions on the second attempt.
- Once several students have recorded their predictions on the numbers of revolutions on the first attempt, what will you need to consider when making predictions about the second attempt?
  *Answers may vary, but you are looking for the fact that students recognize that, chances are, the change in the number of revolutions between attempts will not be exactly the same for all students—some will increase, some will not. Instead, they should look for trends that describe the change, such as "The number of revolutions for the second attempt **tends** to be _____."*
- Will your prediction be true for every student? Why or why not?
  No. Point out that this exemplifies how "real" data often does not fit into "clean" or exact patterns such as linear patterns. Instead, we have to look for any trends in the data.
- How many pieces of data would you need to make a prediction about the number of revolutions on a second attempt?
  Answers may vary, but you are looking for students to recognize that the more data you have, the better defined any correlation will be.

3. Show the table on **Transparency 2** that will be used to collect the data. Ask for 2-3 volunteers to demonstrate what will be recorded in the table. You may want to agree ahead of time as to whether or not each student will be allowed a practice attempt. (If hula-hoops are not available, you can substitute another event such as

paddle-ball, trash can basketball, etc. The goal is to choose an event where the trend in the data is not obvious or that you could make a case for multiple trends. For example, one might make the case that students would have more revolutions with the hula-hoop on their 2nd attempt because of the practice they got during the 1st attempt or that they would have fewer on their 2nd attempt because they were tired from the 1st attempt or that there would be no clear correlation.)

4. Pair the students in the class to collect the data for Transparency 2. Each student will take turns being the hula-hooper and the recorder. Record the results on Transparency 2.

Facilitation Questions

- If a student is able to complete 6 revolutions on the first attempt, what could happen on the second attempt? Why?
  The number of revolutions could be more than 6, less than 6, or the same as 6 on the second attempt.
- If that same student were able to make 10 revolutions on the second attempt (an increase of 4 over the first attempt), would this mean that the same will be true for the next student? Why?
  Possibly, but possibly not—we do not have enough data to make that prediction yet.

EXPLORE

The Explore portion of the lesson provides the student with an opportunity to be actively involved in the exploration of the mathematical concepts addressed. This part of the lesson is designed for groups of 2 students or individual investigation.

1. Prompt students (or pairs of students if there is not enough technology available) to open the spreadsheet **RoundandRound**.
2. Distribute the **Round and Round** worksheet. Have students follow the directions on the spreadsheet.

Facilitation Questions

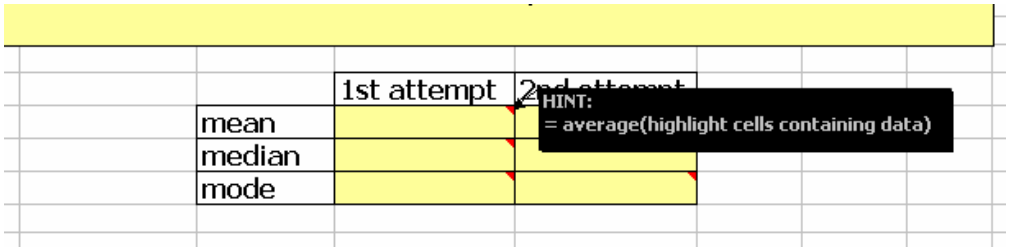
- How would you describe any trends that you might see in the scatterplot drawn from the data from your class?
  Answers may vary depending on the data collected. Look for statements such as "As the number of revolutions during the 1st attempt increases, the number of revolutions during the 2nd attempt (increases/decreases)." Or "Students tend to _____ ."

3. For part B, inform students these graphs were drawn based on fictitious data, and not their own, as they answer questions #1 – 7 on the worksheet.

Facilitation Questions

- Describe a possible scenario that would produce each of the three scatterplots.
  Answers may vary…
  Scatterplot A – Most students performed about the same on their first attempt as on their second attempt.
  Scatterplot B – Most students did considerably better on their second attempt than on their first attempt.
  Scatterplot C – Some students followed the explanation of the scenario for scatterplot A and some for scatterplot B…no clear pattern for the group as a whole.

NOTE: There are comments with "HINTS" in the cells requiring the formulas to calculate central tendency.



Facilitation Questions

- What do mean, median, and mode describe about any set of data?
  Answers may vary…
  Mean – the value of each data point should all data points be "evened out"
  Median – the value of the data point in the "middle" when considering the data points in numerical order (one-half are equal or greater than the median and one-half are equal or less than the median)
  Mode – the value of the data point that occurs more often than other data points
- Looking at your data, how do the mean, median, and mode for the 1st attempt compare to that in the 2nd attempt? What might this imply about the comparison of the number of revolutions in the 2nd attempt as related to those in the 1st attempt? *Answers may vary.*
- Have the students share their answer to #9. What are you looking for in the data when you try to determine the measure of central tendency that will best describe the data?
  Answers may vary.
  Mean – data is clustered with no outliers
  Median – most of the data is clustered except for one or more outliers
  Mode – if one piece of data appeared significantly more times than others

EXPLAIN

The Explain portion of the lesson is directed by the teacher to allow the students to formalize their understanding of the TEKS addressed in the lesson.

1. Once the students have completed their work on the spreadsheet **RoundandRound**, display **Transparency 3** to debrief. (Transparency 3 contains the same graphs that are on the RoundandRound spreadsheet.)
2. Guide the students in drawing trendlines (if possible) on the three original scatterplots using the spreadsheet's drawing toolbar.

Facilitation Questions

For each scatterplot (A, B, C)

- Use the drawing toolbar in the spreadsheet to draw a line that would include the data points if the number of revolutions on the second attempt for each student were the same as the first attempt. ($y=x$) Does this line "fit" the data? Why or why not?
  Answers may vary. A line y = kx (proportional) will best "fit" to scatterplot A because it follows a similar trend…as x increases, y increases. Line y = kx does not "fit" with scatterplots B or C because the data on those scatterplots does not fit a similar trend.
- For each scatterplot on the spreadsheet, is it possible to click on the red line below the scatterplot and place it on the scatterplot in such a way that it better exemplifies the relationships/trends in the data? (Click on the red line to move it and select "draw," "rotate," and "free rotate" to rotate the line.)
  Answers may vary. Minor adjustments (as compared to line y = x) may be made for scatterplot A, whereas the red trend lines for scatterplots B and C should be significantly different than line y = kx.

3. Select students to draw possible trendlines on the transparency. Discuss any differences in opinion. Use the trendlines to make predictions.

Facilitation Questions

- After drawing a trendline for scatterplot A, consider the points that would fall on or near the trendline we drew. As the number of revolutions made on the 1st attempt increases, what happens to the corresponding number of revolutions made on the 2nd attempt?
  They increase as well.
- What type of correlation (trend) is this?
  A positive correlation (trend)
- Based on this trendline, about how many revolutions would you expect students to make on the second attempt if they made 13 revolutions on their 1st attempt? What about if they had made 30 revolutions? 50 revolutions?
  Answers may vary slightly depending on how the trendline was drawn.
- After drawing a trendline for scatterplot B, consider the points that would fall on or near the trendline we drew. As the number of revolutions made on the 1st attempt increases, what happens to the corresponding number of revolutions made on the 2nd attempt?
  They increase.
- What type of correlation (trend) is this?
  A positive correlation (trend)
- Based on this trendline, about how many revolutions would you expect students to make on their second attempt if they made 13 revolutions on their 1st attempt? What about if they had made 30 revolutions? 50 revolutions?
  Answers may vary slightly depending on how the trendline was drawn.
- After attempting to draw a trendline for scatterplot C, why is it more difficult to draw a trendline on this scatterplot?
  Answers may vary. Students should note that there is no clear pattern in the number of revolutions on the 2nd attempt (increasing or decreasing) as the number of revolutions increases on the 1st attempt. Visually, the points do not cluster around any line, rather they are spread more randomly throughout the scatterplot.
- What type of correlation (trend) is this?
  There is no correlation (trend)
- Knowing we did not draw a trendline, about how many revolutions would you expect students to make on the second attempt if they made 13 revolutions on their 1st attempt? What about if they had made 30 revolutions? 50 revolutions?
  With no clear trend, it is impossible to make a prediction based on this data alone.

4. Use **Transparency 4** to debrief the data in relationship to the mean or median.

5. Draw in the mean lines and discuss the characteristics (in relationship to the mean) of the pieces of data in each of the four resulting quadrants. (Students can use the drawing toolbar to draw the lines on their spreadsheet as you draw them on the transparency.)

Facilitation Questions

- Another way of looking at the data, other than a trendline, is to look at it in relationship to a central tendency such as mean or median. Look at scatterplot C where it was difficult to draw a trendline. What is the mean(average) number of revolutions made on the 1st attempt?
  8.6667 (Draw in a vertical line at 8.6667 on the x-axis.)
- About how many data points fell below the mean? above the mean? What does this say about the data from those students?
  6 below…3 above… Rationale may vary…should include a discussion about outliers and/or the spread of the data.
- What is the mean(average) number of revolutions made on the 2nd attempt?
  10.44 (Draw in a horizontal line at 10.44 on the y-axis.)
- About how many data points fell below that mean? Above the mean? What does this say about the data from those students?
  4 below…5 above… Rationale may vary…should include a discussion about outliers and/or the spread of the data.
- When you look at both mean lines, the data points divide into 4 groups. Describe the characteristic of each group.
  Below the average on both attempts, below the average on the 1st attempt and above the average on the 2nd, above the average on the 1st attempt and below the average on the 2nd, above average on both attempts
- Why do you think the number of data points varies from group to group?
  The lines were drawn using the mean values. Outliers may "pull" the line away from the center.

6. Use the second copy of Transparency 4 to draw in the median lines and discuss the characteristics (in relationship to the mean) of the pieces of data in each of the four resulting quadrants. (Students can use the drawing toolbar to draw the lines on their spreadsheet as you draw them on the transparency.)

Facilitation Questions

- If you were to do the same for the median lines, how do you think the data will be spread among the 4 groups?
  Because the medians are the midpoints, the data should be evenly spread between the groups.
- Why might you want to look at the data in this manner?
  Answers may vary. This is another way (other than trend lines) to communicate the relationship between the number of revolutions on the corresponding 1st and 2nd attempts.

ELABORATE

The Elaborate portion of the lesson provides an opportunity for the student to apply the concepts of the TEKS within a new situation. This part of the lesson is designed for groups of 2 students or individual investigation.

1. Direct students (or pairs of students if there is not enough technology available) to open the spreadsheet **BabyName**. Have students record their answers on notebook paper if the option of printing their work is not available.
2. Read the "Given" and the "Question" and have students turn to a neighbor and share their thoughts before sharing with the large group.

Facilitation Questions

- Read the "Given" statements and tell me what that means in your own words.
  Answers may vary. Look for paraphrasing that connects "popularity" of names with frequency and the concept of ranking.
- Why might the popularity of certain names vary over time?
  Answers may vary. Students might consider factors such as culture, famous figures, etc.
- Read the "Question" and turn to your neighbor and share your thoughts. (Pause) Do you think there will be a difference in the change in popularity of boy names versus girl names? Why or why not?
  Answers may vary. Accept all answers for now.

3. Read through the directions. Make sure students can access the data website (**http://www.ssa.gov/OACT/babynames/**) or print and have hard copies of the data available if using the Internet is not an option. NOTE: Using technology to search on the Internet is much more efficient than searching on paper.
4. Have students complete spreadsheet.

Facilitation Questions

- (After locating the top ten boy names for 1965…) Look at the data source and explain again how certain names make this list and others do not.
  Answers may vary. Students should determine that the data comes from counting the number of times a particular first name was put on applications for Social Security cards for newborns. Ex. Since "Michael" is ranked first, this means that there were more newborns with the first name of "Michael," according to the information parents gave on their baby's Social Security card application, than any other first name.

5. Use **Transparency 5** to debrief the activity.

Facilitation Questions

- What was the only central tendency not calculated in the activity?
  Mode
- Why do you think mode was not included?
  Since the names each have a unique rank (understanding that it would be highly unlikely that two names would occur exactly the same number of times), there will be no mode.
- When answering #9, what characteristics in the data made you choose to draw the lines for the mean or for the median?
  Answers may vary. Looking for some discussion of clustering of data and/or outliers.
- (Transparency 5 – Draw in the lines based on the discussion from the previous question, then draw in a line representing the ranking from 1965.) Describe the relationship among the three lines.
  Answers may vary. Students should note that the central tendency line for the rankings of boy names is much closer to the line representing the rankings in 1965…a visual demonstrating how the popularity of those boy names has remained somewhat steady in comparison to the girl names of similar rankings.
- Look at the lines you drew on the scatterplot. Would you say, based on the data you have, that you could better predict the change in popularity of a boy name or a girl name?
  Answers may vary. Since the points representing boy names are more clustered around that line, chances are the popularity of a boy name will change less than that of girl names with similar original rankings.

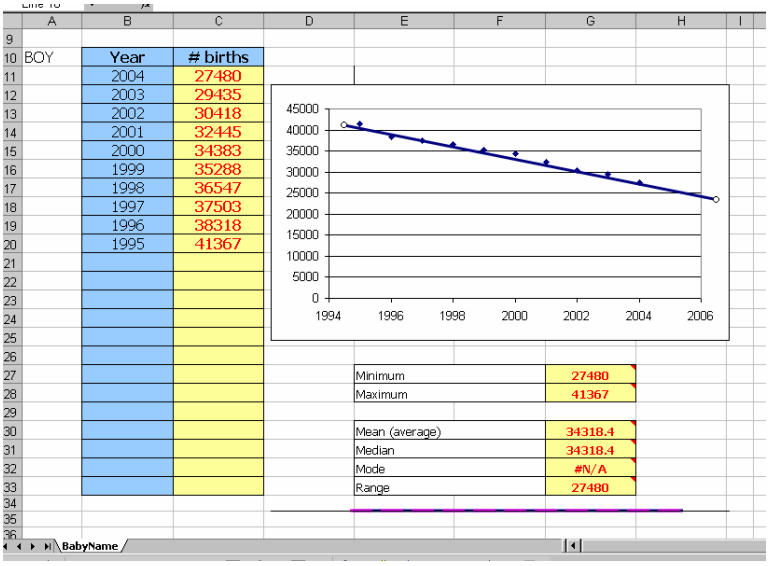
Facilitation Questions

- Think back to your response to the "Question" at the beginning of the activity. Given the additional information you now have, do you need to revise or elaborate on your initial thoughts?
  Answers may vary. In general, the selection of names for girls tends to be more susceptible to varying trends than names for boys. In other words, based on the data for the names we researched, the popularity of a particular name for a girl is much more likely to change over time; whereas the popularity for a particular name for a boy is more likely to remain somewhat steady. Just a thought…This could be related to the custom many have of designating males as the ones who will carry on the family name.
- How are the trends in this data similar to or different from the trend you saw in your data from the hula hoop experiment?
  Answers may vary. For example, if there were no clear trend between the numbers of revolutions made on the 1st and 2nd attempts, students might see similarities between that data and the data for the girl names.

EVALUATE

The Evaluate portion of the lesson provides the student with an opportunity to demonstrate his or her understanding of the TEKS addressed in the lesson.

1. Distribute **What's In A Name?** activity sheet to each student.
2. Upon completion of the **What's In A Name?** activity sheet, the teacher should use a rubric to assess student understanding of the concepts addressed in this lesson.



Answer may vary somewhat…around 20,000 people with the top ranking boy name in 2010
Justifications may vary…Based on the data in this scatterplot, there is a negative correlation between the year and the number of people with the top ranking boy name. As the years increase, the number of people with the selected name decreases. If this trend were to continue, you would expect around 20,000 people with the top ranking boy name in 2010. Students could also justify their prediction with one of the statistical measures such as finding a range and extending it to future years.

Answers and Error Analysis for selected response questions:

Question Number	TEKS	Correct Answer	Conceptual Error	Conceptual Error	Procedural Error	Procedural Error	Guess
1	8.12(B)	A	B	C			D
2	8.12(B)	C	A	B			D
3	8.12(A)	C	A		B		D
4	8.12(B)	C	A	B			D

EXPLORE

Round and Round

Open the **RoundandRound** spreadsheet.

A. Input your class data from Transparency 2. (Use the table that starts in row 6.)

Sketch the resulting scatterplot.

Will vary depending on data

B. For each statement, choose the scatterplot(s) that best represents the situation.

___*A, B, C*____1. After the 1st attempt, most students were able to increase the number of revolutions on their 2nd attempt.

____*none*___ 2. After the 1st attempt, most students made fewer revolutions on their 2nd attempt.

____*A*____ 3. The number of revolutions on the 1st attempt is about the same as the number of revolutions on the 2nd attempt.

____*C*_____ 4. There is not a strong relationship between the number of revolutions made in the two attempts.

____*B*_____ 5. Most students did considerably better on their 2nd attempt than on their 1st attempt.

___*vary*____ 6. Based on the data you have from your class, which scatterplot would look most like yours? Explain.

7. Now that you have analyzed possible scenarios for scatterplots A, B, and C, write a statement that describes the relationship between the 1st attempt and 2nd attempt for your class.
Answers may vary.

8. Use formulas to calculate the mean, median, and mode of the data for the 1st attempt and for the 2nd attempt. (Use the table in rows 50-53.) Record the results below.
Answers may vary.

	1st attempt	2nd attempt
mean		
median		
mode		

9. Which measure of central tendency best describes the number of revolutions made on the 1st attempt and 2nd attempt? Explain your choice.
Answers may vary.

Transparency 1

How many revolutions of a hula hoop can you achieve in one attempt?

Will this prediction change for a second attempt? Why or why not?

How might we gather data to test our predictions?

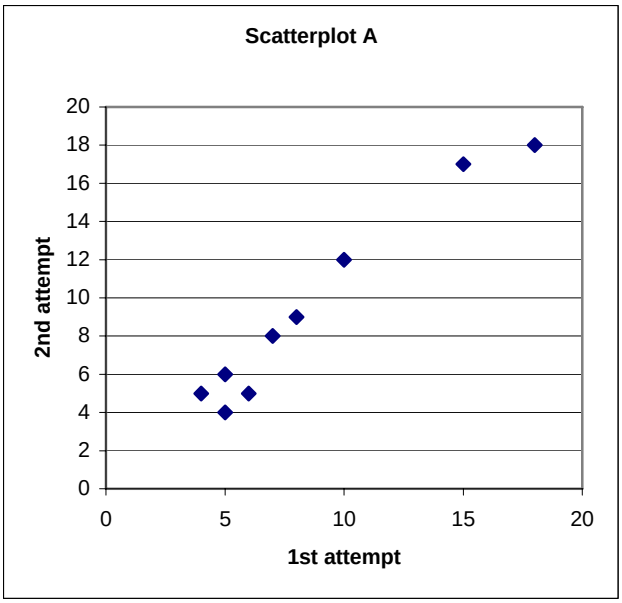
Transparency 2

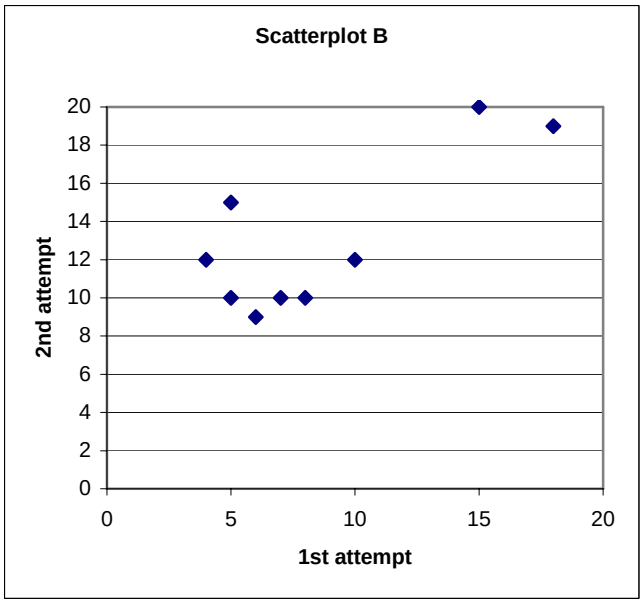
Student	# Revolutions on 1st Attempt	# Revolutions on 2nd Attempt

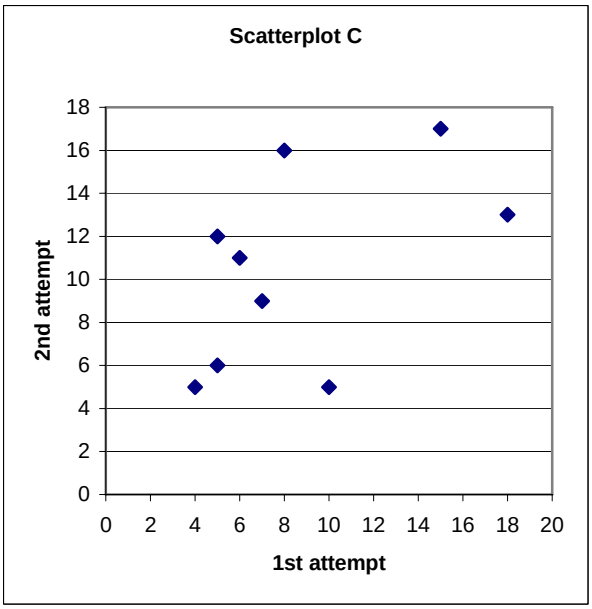
Transparency 3

1. If possible, sketch a trendline.

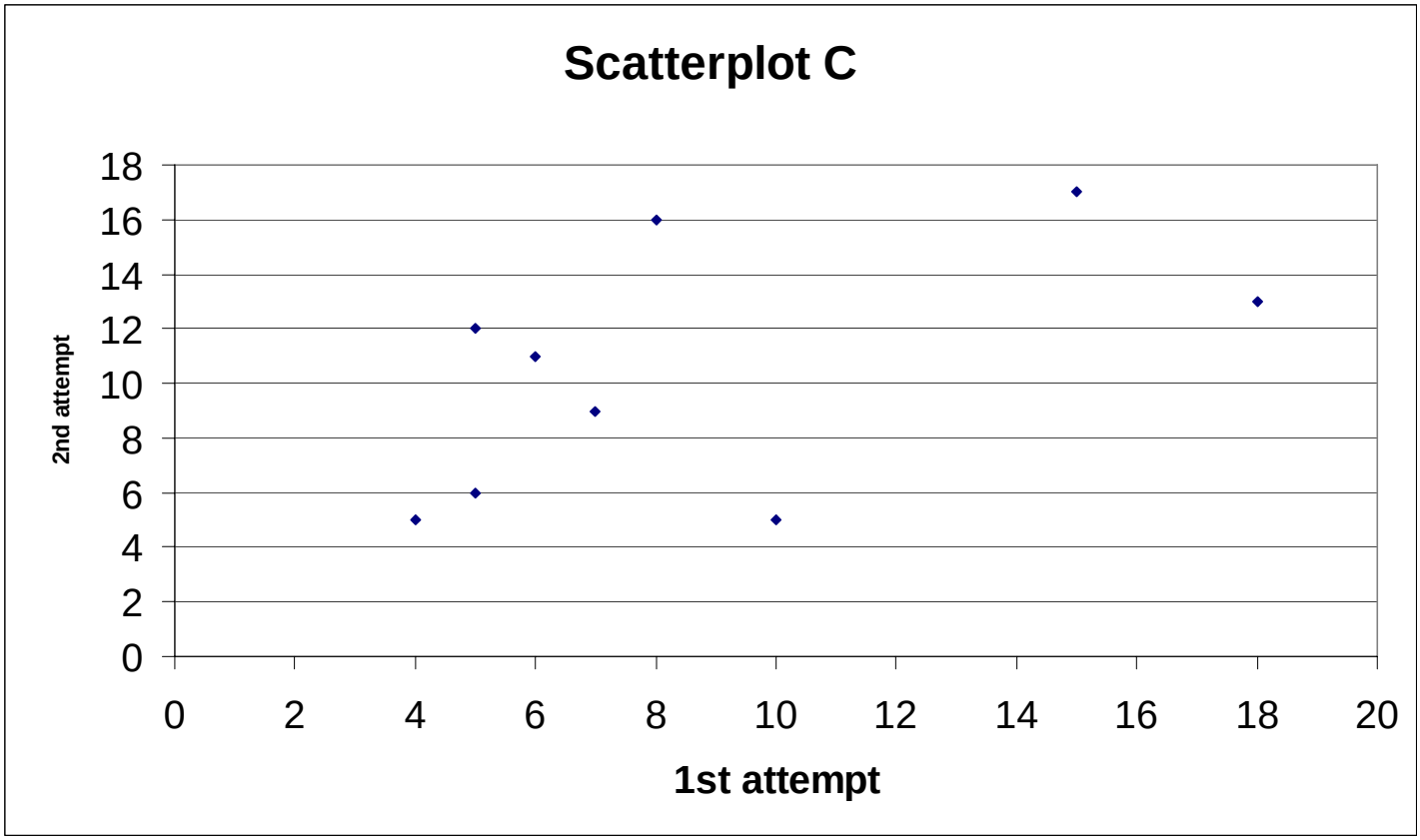
2. Predict the number of revolutions on the 2nd attempt if the number on the 1st attempt was 13...30...100.





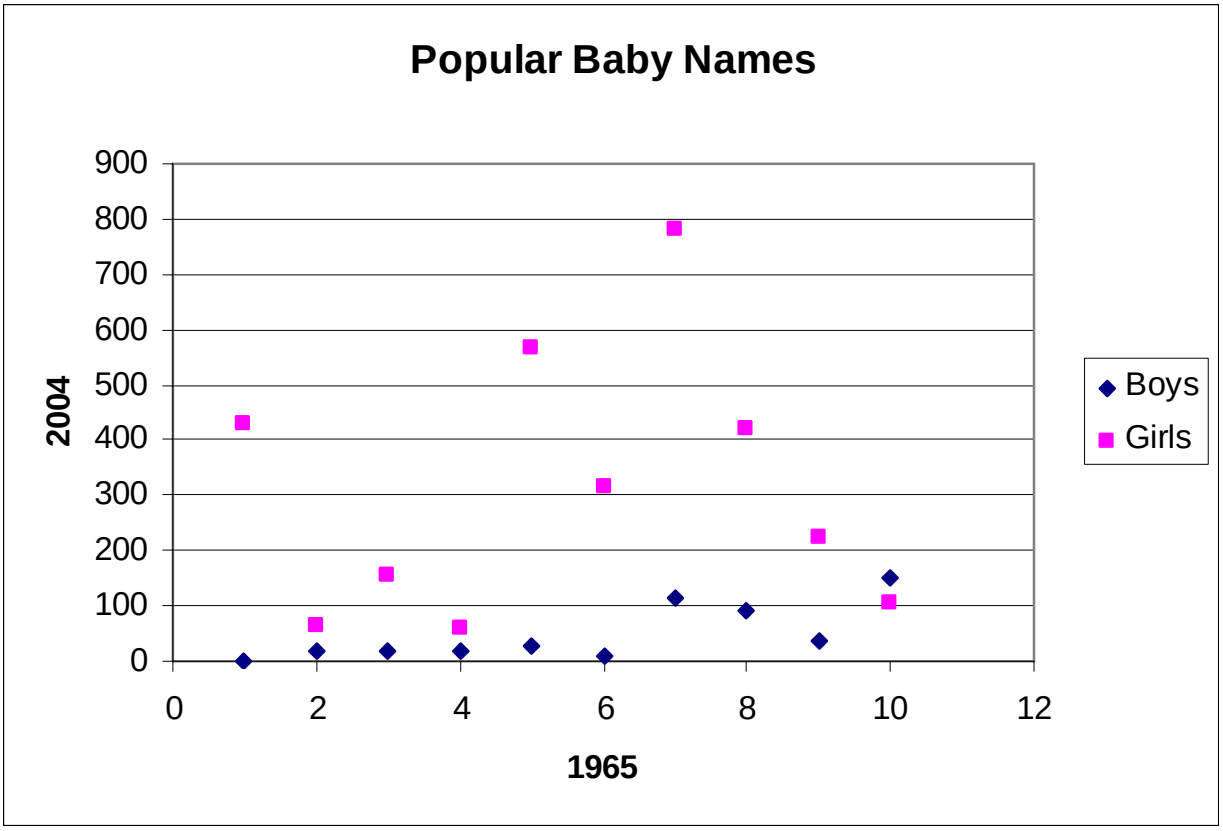


Transparency 4



	1st attempt	2nd attempt
mean	8.6667	10.44
median	9.0741	10.94

Transparency 5



	1965 — — —	2004 – Boys – – – – –	2004 – Girls .....
Mean	5.5	48.1	312.1
Median	5.5	23.5	269.5
Range	9	147	504

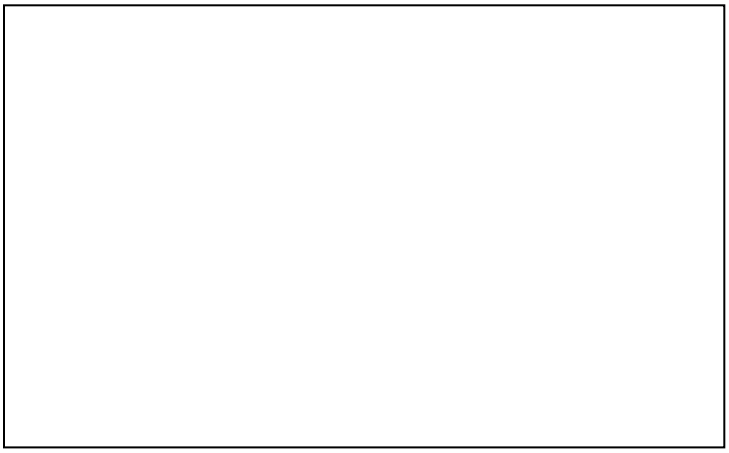
Student Name(s)_____ Date_____

Round and Round

Open the **RoundandRound** spreadsheet.

A. Input your class data from Transparency 2. (Use the table that starts in row 6.)

Sketch the scatterplot that resulted from entering class data.



B. For each statement, choose the scatterplot(s) that best represents the situation.

_____1. After the 1st attempt, most students were able to increase the number of revolutions on their 2nd attempt.

_____2. After the 1st attempt, most students made fewer revolutions on their 2nd attempt.

_____3. The number of revolutions on the 1st attempt is about the same as the number of revolutions on the 2nd attempt.

_____4. There is not a strong relationship between the number of revolutions made in the two attempts.

_____5. Most students did considerably better on their 2nd attempt than on their 1st attempt.

_____6. Based on the data you have from your class, which scatterplot would look most like yours? Explain.

7. Now that you have analyzed possible scenarios for scatterplots A, B, and C, write a statement that describes the relationship between the 1st attempt and 2nd attempt for your class.

8. Use formulas to calculate the mean, median, and mode of the data for the 1st attempt and for the 2nd attempt. (Use the table in rows 50-53.) Record the results below.

	1st attempt	2nd attempt
mean		
median		
mode		

9. Which measure of central tendency best describes the number of revolutions made on the 1st attempt and 2nd attempt? Explain your choice.

Student Name(s)_____ Date_____

What's In A Name?

Access the website **http://www.ssa.gov/OACT/babynames/**. In an earlier activity we compared the ranking of the top ten names of your parents' generation (1965) to the ranking of those names today to answer the question about how the popularity of names stands the test of time.

a. Consider the following set of questions.
   How has the number of people having the most popular boy name changed over the last 10 years? How many people do you predict might have the most popular name in 2010?

b. Open the **WhatName** spreadsheet and input the data to create a scatterplot.

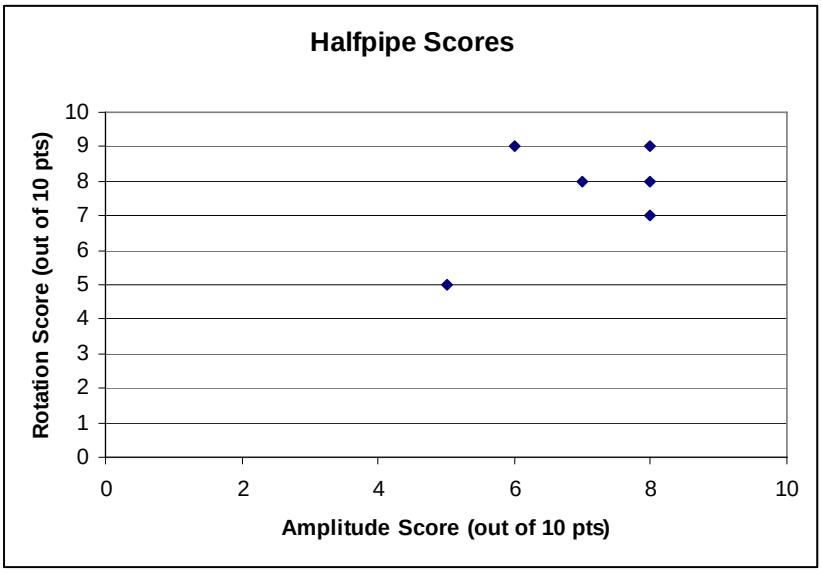
c. Draw a trendline if appropriate.

d. Calculate the mean, median, and range of your data.

e. Respond to the questions in part a. Justify your answers using the scatterplot, trendline, and/or statistical measurements to support your conclusions.

Student Name(s)_____ Date_____

1. The scatterplot below compares the score for amplitude (height) to the score for rotations (spins and flips) for six skateboarders at the weekend meet.



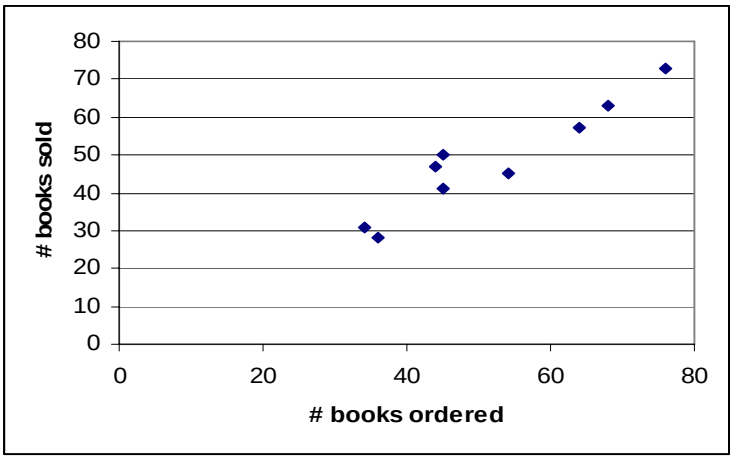
Which of the following statements would be supported by the scatterplot?

A. As the score for amplitude increases, the score for rotations tends to increase.
B. As the score for amplitude increases, the score for rotations tends to decrease.
C. As the score for rotations increases, the score for amplitude tends to decrease.
D. The score for rotations tends to be the same as the score for amplitude.

2. Which relationship, when graphed on a scatterplot, would NOT be described as having a positive trend?

A. The number of fans in a football stadium compared to the noise level of the stadium.
B. The amount of money earned babysitting compared to the number of hours spent babysitting.
C. The number of miles driven compared to the amount of gasoline in the tank.
D. All of the above relationships have a positive trend.

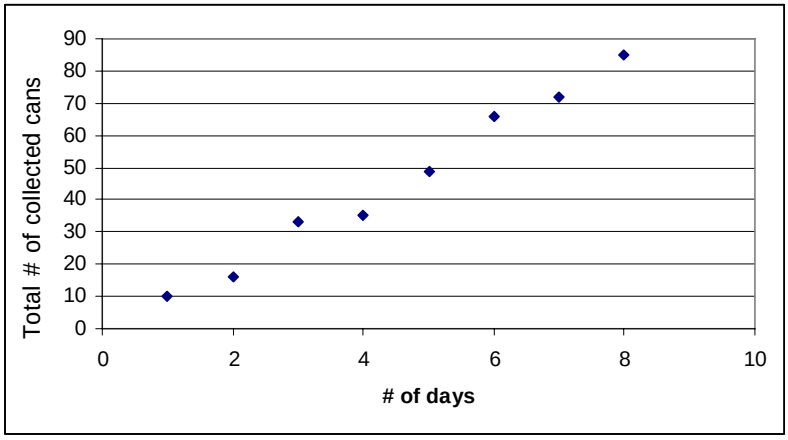
3. The following scatterplot compares the number of books ordered through the school fund raiser to the number of books that were sold.



If the mean number of books ordered is about 52, estimate the mean number of books sold based on the trends in data in the scatterplot.

A. greater than 55
B. between 50 and 55
C. between 45 and 50
D. less than 40

4. Ms. Smith's class is recording data about an aluminum can recycling project as shown in the scatterplot below.



At this rate, about how many days will it take to collect 150 cans?

A. 150 days
B. 20 days
C. 15 days
D. 10 days