Statistics and Probability

TABLE OF CONTENTS

Posing Questions and Gathering Data

1

Overview

Mathematical Focus

- Investigative questions
- Data investigations that address those questions
- Collection and organization of data

In this activity, students are presented with a pile of 100 pennies. They brainstorm questions they would like to answer about the pennies, for example, *How many pennies are in the pile? What is the oldest/newest penny? How many decades are represented? Which decade has the most pennies?* For each question posed, students make an initial prediction about the answer. As they begin to gather data on the pennies, they have an opportunity to revise their predictions. Finally, students discuss and explore different strategies for gathering and recording their data.

Preparation and Materials

- 100 pennies

You may want to gather the pennies needed for the activity from a variety of sources. This will increase the likelihood of having a range of dates.

Save the penny data for future activities. Consider gathering data on two or more different sets of 100 pennies. Students can then make comparisons between the sets. Keep the list of questions students pose at the end of this activity; you will use them again in Activity 2.

Activity

Posing Questions

1. **Brainstorm questions you could ask about a pile of pennies.**

   Give students 100 pennies in a pile. Their list may include questions such as these:

   - *How many pennies are there?*
   - *What is the date of the oldest penny?*
   - *What is the date of the newest penny?*
   - *Is there a penny from every year in the '90s? '80s? '70s?*
   - *How many decades are represented in the pile?*
   - *How many different years are there?*
   - *Which year has the most number of pennies?*

2. **Predict the answers to your questions.**

   Tell students that once they begin to gather data on the pennies, they will have a chance to revise their predictions.

Collecting and Recording Data

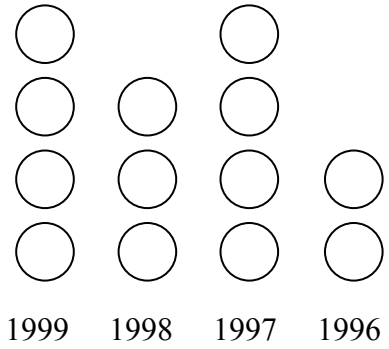
1. **Count the pennies.**

   Ask: *What strategies could you use for counting this pile of pennies?* Students' suggestions may include the following:

   - Count them one-by-one
   - Count by two's
   - Group by ten's and then count the groups of ten

2. **Use the dates on the pennies to organize them in some way.**

   Ask students to describe the strategy they use to organize the pennies. Discuss and compare different possible strategies for organizing the pennies and collecting the data. Some students may want to group the pennies by decade first—for example, all of the pennies from the 1990s—and then go back and sort by date within each decade. Other students may prefer to make a column of pennies for each date represented. Many strategies are possible.



As students work, encourage them to revisit the predictions and revise those predictions as necessary. Have them share discoveries they make as they gather data.

3. **Discuss and explore different strategies for recording the data about the dates on the pennies.**

   Explain that over the course of the next few sessions, they will use this data for different activities. Rather than sorting the pennies each time, it would be much easier to have a written record of the penny data. Talk about the advantages and disadvantages of each strategy. Strategies may include the following:

   - Making a list of all of the dates, either in random order, ascending order, or descending order.
   - Keeping a tally of each date, for example:

     2000 𝍸 |||
     1999 𝍸
     1998
     1997

   - Making a chart, for example:

Year	Number of Pennies
2000	8
1999	5
1998	
1997	

Answering Questions

1. **Describe the strategies you used and share the discoveries you made about the pennies.**

   Return to the list of questions they generated at the beginning of the activity and the predictions they made. Ask questions such as these:

   - *Which questions can you answer?*
   - *Were your first predictions on target?*

- *Are you surprised by anything you discovered? Why or why not?*

Discussion

1. **Think of other types of questions for which you could potentially gather and analyze data.**

   Talk with students about the ways people use data analysis to investigate information about the world. Sometimes the results of these investigations are used to make decisions. Here are some examples to prompt them, if necessary:

   - *What is the most popular ice cream flavor (TV program, sport, movie, game, food, playground equipment, etc.) among the students at my school?*
   - *Who in my class can run the fastest?*
   - *How many books are checked out from the library in a month?*
   - *What is the most popular type of book (e.g., biography, mystery, adventure) checked out from the library?*

2. **Discuss how data could be collected to answer these questions and how this information could be used.**

   Talk about the ways in which students might collect data to answer each question. Suggestions might include surveys, experiments, measurements, or observations. Discuss how information from these investigations might be used; for example, knowing that science fiction books are the most popular type of book checked out from the library could justify ordering additional new science fiction books, or it might indicate the need for a greater selection of books in other categories.

2 Representing Data

Overview

Mathematical Focus

- Interpretation of graphs
- Data represented in graphs
- Comparison of data representations

In this activity, students investigate different ways of representing data, using line-plots, bar graphs, and circle graphs. They begin by looking at and analyzing an example of each type of graph. Then, using the penny data they gathered in Activity 1, students create their own line-plots, bar graphs, and circle graphs. Students look for patterns in the data which are highlighted by the different representations. Finally, students extend their investigation by looking for examples of each type of graph in newspapers and magazines.

Preparation and Materials

- Student Page 1: Line-Plot
- Student Page 2: Grid Paper, several copies, or several sheets of quarter-inch grid paper
- Student Page 3: Bar Graph
- Student Page 4: Circle Graph
- Student Page 5: Blank Circle Graph
- Penny data from Activity 1
- Tape

Save the graphs created in this activity for use in Activity 3.

Activity

Line-Plots

1. **Tell what you know about graphs.**

   Ask students these questions to help generate what they remember about graphs:

   - *What are graphs used for?*
   - *Who uses them?*
   - *What are some different types of graphs you've seen?*
   - *Where have you seen graphs used?*

   Explain that in this activity students will explore and create three different types of graphs—line plots and circle graphs—using the penny data they collected in the last activity. line-plots, bar graphs, and circle graphs.

2. **Discuss the features of a Line-Plot graph from Student Page 1.**

   Ask questions such as these:

   - *What is the title of the graph?*
   - *What do you think this graph shows?*
   - *What are the different parts of this graph?*
   - *What are the numbers below the horizontal line?*
   - *What do the X's represent?*

   (This line-plot shows the number of pizzas delivered from Tony's Pizza on each day of the month of April. The horizontal line is a segment of a real number line, showing all of the dates in the month of April [not just the ones on which pizzas were sold]. The X's above each date indicate the number of pizzas sold on that date.)

   Ask students to describe any patterns they notice in the data. (Students might observe that more pizzas seem to be delivered on Fridays and Saturdays than on other days.)

3. **Create a line-plot for the penny data you collected in Activity 1.**

   Give students several copies of Student Page 2: Grid paper or several sheets of quarter-inch grid paper that they can tape together to create the graph. Together, walk through the steps of creating the graph. Explain that the first step is to determine the smallest and largest values in the data set. Ask: *What are the smallest and largest values in your data set?* (This will be the date of the oldest penny and the date of the newest penny.) Have students create a number line with these two numbers at either end and all the years in between filled in, even the ones for which there are no pennies. Explain that the next step is to place an X on the number line to represent each data point, i.e., each penny found for that date.

4. **Look for patterns in the penny data, using the line-plot you created.**

Ask:

- *What patterns do you notice in the data?*
- *What do the patterns tell you about the penny data?*
- *Does this graph give you a different picture of the data than the record you made in Activity 1? Why or why not?*

Bar Graphs

1. **Compare a Bar Graph from Student Page 3 with the Line-Plot from Student Page 1.**

   Explain that this bar graph also shows the number of pizzas delivered from Tony's Pizza on each day of the month of April. Ask:

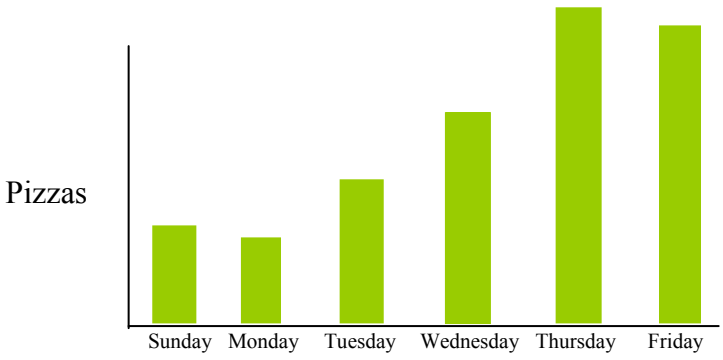
   - *In what ways are the graphs similar? In what ways are they different?*
   - *Is it easier to see patterns in one graph than the other? Why or why not?*

2. **Create a bar graph that shows the total number of pizzas sold on each day of the week for the month of April.**

   Have students describe how they would create a bar graph to show this information. Ask:

   - *How would you label the axes?*
   - *How would you determine the height of each bar?*

   Their graph may look similar to the one below.

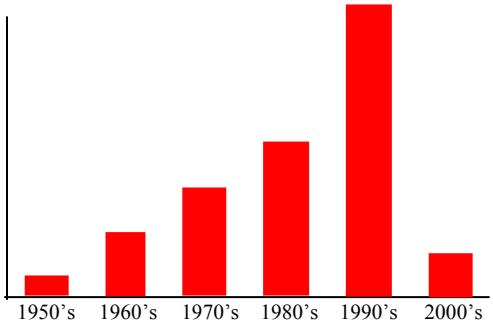


3. **Compare the information that the different graphs show about Tony's Pizza.**

4. **Construct a bar graph of the penny data.**

   Give students several copies of Student Page 2 or sheets of quarter-inch grid paper that they can tape together to create the graph. Discuss possible categories for the horizontal axis. Students may want to label the horizontal axis with individual years, or they may want to label it with

decades and have each bar represent all of the pennies found for that decade, as in the example below.



5. **Compare your bar graph with the line-plot of the penny data.**

   Ask questions such as these:

   - *In what ways are the two graphs similar?*
   - *In what ways are they different?*
   - *Which one gives you a better snapshot image of the data?*

Circle Graphs

1. **Compare and contrast the bar graph with a circle graph for Tony's Pizza found on Student Page 4.**

   Like the other bar graphs, the circle graph shows the number of pizzas delivered from Tony's Pizza on different days of the week for the month of April. Ask questions such as these:

   - *How are the graphs similar?*
   - *How are they different?*
   - *Which gives you a better picture of the data? Why?*

2. **Discuss how you would create a circle graph for the penny data.**

   Ask: *If you were to represent every year for which a penny was found, how many wedges would you have in your circle? What if you were to show every decade?* Discuss whether students will show every year or only the decades. Students may want to create two circle graphs and explore both strategies for displaying their data.

3. **Create a circle graph for the penny data.**

   Give students a copy of Student Page 5: Blank Circle Graph to use as a template for the circle graph. Point out that the circle is lightly divided into 100 equal-size wedges.

4. **Compare and contrast the 3 graphs you have created for your penny data.**

Ask:

- *What are the features of each graph?*
- *Is any one of the graphs more informative than the others?*

5. **Use the three graphs to answer questions about the penny data.**

   Ask students:

   - *Which decades had the most pennies?*
   - *Which year had the most pennies?*
   - *What was the oldest penny? The newest penny?*
   - *How many pennies were from the 1970s?*
   - *Were there more pennies from the 1970s or the 1980s?*
   - *Which graph made it easier to find the answers to each of these questions?*

Discussion

1. **Discuss ways to display the data for each of the questions posed at the end of Activity 1.**

   For each question, discuss possible ways of displaying the data, using line-plots, bar graphs, or circle graphs.

2. **Look for examples of each type of graph in newspapers and magazines.**

   Ask: *Which type of graph seems to be the most frequently used? Why do you think this is?*

3 Interpreting and Evaluating Data

Overview

Mathematical Focus

- Data interpretation using exploratory analysis
- Development and evaluation of inferences, based on data

In this activity, students analyze and interpret the penny data they collected and represented in Activities 1 and 2. They describe the characteristics of their data by identifying important features, such as range, gaps, clusters, and outliers. Students use the median to describe the center point of their data and investigate its significance. After interpreting their penny data, students discuss how data collection methods can influence the data set. They consider the concept of representativeness of a sample within the context of the penny example, asking: *Is this set of 100 pennies representative of all pennies?*

In analyzing and comparing the three graphs, students consider which aspects of the data are highlighted or obscured by each representation.

Preparation and Materials

- Student Page 2: Grid Paper, or quarter-inch grid paper
- Penny data and graphs from Activities 1 and 2
- Student Page 6: Interpreting the Data, two copies
- Photocopy of a short paragraph from a book

Interpreting Data

1. **Use terms from Student Page 6 to describe your penny data set and graphs.**

   Student Page 6 lists several terms that can be used to describe the shape and important features of a data set. Discuss each of the terms and have students use them to describe their own data set.

   - *Range*: The range is the interval between the smallest and largest value in the data set. In this case, the range would be the interval between the date of the oldest penny and the newest penny.

   Have students describe the range of their data set and record it on Student Page 6.

   - *Outliers:* These are isolated data points.

   Have students look for outliers in their data set. For example, if they had a really old coin in their collection—say, a penny from the '40s or the '50s and most of the others were from the '70s, '80s, '90s, the older coin would be an outlier. Have students list any outliers on Student Page 6. Ask: *Why might it be important or useful to look at the outliers in a data set?*

   - *Clusters:* A cluster is a pile-up of data at one point on the number line.

   Ask what clusters tell you about the data.

   - *Gaps:* Dates with no X's, or "holes" in the data, are referred to as gaps.

   Ask: *Are there any gaps in your data? If you added 100 more pennies to the data set, do you think that some of the gaps would be filled in?* Have students describe any clusters and gaps in their data on student page 6.

2. **Find the median of the penny data by ordering the data from earliest date to latest date and then moving in from each side until you reach the middle.**

   Tell students that the median is the center point of the data; it divides the data into upper and lower halves. Explain that when the data set has an odd number of cases, there will be a single value for the median. If there is an even number of cases, the median is usually considered to be the halfway point between the two middle points. Have students draw a vertical line on the line-plot so that equal numbers of X's (50) are on each side of the line.

3. **Use the terms on Student Page 6 to describe and interpret the data in the Tony's Pizza line-plot on Student Page 1.**

4. **Collect data on the number of times different letters of the alphabet appear in a paragraph.**

   Give students a copy of a paragraph from a book.

5. **Make predictions about what the data will look like in a line-plot.**

   Ask questions such as these:

   - *Will the data be horizontal? What would that mean?* [There were equal numbers of each type of letter]
   - *Will there be gaps? Clusters? Where might these be? Why?*

6. **Display your data using a line-plot.**

   Give students Student Page 2: Grid Paper or a sheet of quarter-inch grid paper. Have students use the letters of the alphabet as the categories along the horizontal axis.

7. **Describe the characteristics of your data.**

   Ask questions such as these:

   - *Where is the data clustered? Around the vowels? Consonants? Is that what you would expect?*
   - *What else can you learn from the data?*

Evaluating Data and Making Inferences

1. **Reflect on the characteristics of the penny data and suggest why those characteristics might exist.**

   Return to the penny data line-plot and the description of the data written on Student Page 6. Ask questions such as these:

   - *Why do you think there are clusters in the '80s and 90s?*
   - *How can you explain the outliers?*

2. **Think about factors that affect the representativeness of a sample.**

   Ask:

   - *If you were to gather another set of 100 pennies and collect data on that set, do you think you would find exactly the same thing? Would you see the same trends, for example, clusters in the '80s and '90s,and outliers in the '50s and 60s?*
   - *In what ways might the method of data collection impact the nature of the data set? For example, what if you got the pennies from a piggy bank that a 70-year-old man had as a child?*

3. **Evaluate the data in the Tony's Pizza and letters-in-a- paragraph line-plots in the same way you evaluated the penny data line-plot.**

Encourage students to make inferences about the data, including factors that may affect the representativeness of the sample.

4 Exploring Probability

Overview

Mathematical Focus

- Basic ideas of chance and probability
- A language to discuss informal notions of probability

In this activity, students explore probability concepts as they create a probability line, and use informal language to describe the likelihood of various events occurring. The probability line ranges from *Impossible* to *Certain* and includes benchmark points in between. Students determine where such events as *The probability that the sun will come up tomorrow,* and *The probability that it will rain tomorrow* fall on the line. As numbers are added to the probability line, students begin to use fractions, decimals, and percentages to describe the likelihood of an event occurring. Finally, using the penny data from Activity 1, students explore the probability of drawing a penny of a given date from their collection of 100 pennies.

Preparation and Materials

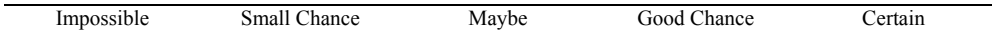
- Penny data from Activity 1

Activity

Developing an Informal Language for Discussing Probability

1. **Discuss the chance of various events occurring, using informal terms to describe the degree of chance.**

   Talk with students about how probability can be used to describe the likelihood of events. Draw a probability line and label it with the following words: Impossible, Small Chance, Maybe, Good Chance, Certain.



   Begin by listing some events, such as the ones below, and asking students to say where on the probability line these events should be placed. List each event below the appropriate label.

   - *What is the chance that the sun will come up tomorrow?*
   - *What is the chance that, while you are sleeping, your bed will turn to chocolate?*
   - *What is the chance that it will snow today?*
   - *What is the chance that the temperature will reach at least 70°F today?*
   - *What is the chance that you will have homework tomorrow night?*

   Ask students to generate additional events to add to the probability line.

Adding Numbers to the Probability Line

1. **Answer questions about the probability of various events.**

   Begin by asking students the following questions:

   - *If the weather person said there was a 50-percent chance of rain today, where would we put that on our probability line? Why?*
   - *If there was a 25-percent chance of rain, where would we put that on the probability line?*
   - *If your teacher said there was an 80-percent chance that you will have a math test tomorrow, where would you put that on the probability line?*

2. Determine what fraction, percent, or decimal should be written above "Maybe," "Small Chance," and "Good Chance," on a probability line.

Create a new probability line using the same informal terms. Write the numbers 0 and 1 at either end of the line. Explain that sometimes people use numbers instead of words to describe the chance of an event occurring. On this new line, 1 means the same thing as 100 percent, and 0 means the same thing as 0 percent. "Impossible" has a probability of 0, and "Certain" has a probability of 1. Here is one example:

0% or 0/100 0	25% or 25/100 .25	50% or 50/100 .5	75% or 75/100 .75	100% or 100/100 1
Impossible	Small Chance	Maybe	Good Chance	Certain

3. Think about the probability of pulling pennies from your group of pennies with various dates on them.

Have students look at their penny data. Ask:

- *If you put the 100 pennies in a bucket, reached in and pulled one out, what is the chance of pulling out a penny from 1991? (If there are eight pennies from 1991, students could write that as 8/100, .08, or 8%.) Where would the chance of this event (i.e., drawing a penny from 1991) go on the probability line?*
- *What is the chance of pulling out a penny from the '80s? Where would this go on the probability line?*
- *What is the chance of pulling out a penny from 1974?*
- *What is the chance of pulling out a penny from the '60s?*
- *What is the chance of pulling out a penny from the '60s or 70's?*
- *What is the chance of pulling out a penny from the '70s, 80s or 90s?*

Games of Chance

5

Overview

Mathematical Focus

- Basic notions of chance and probability

In this activity, students use games as a way to explore concepts of chance. Through a coin-toss activity, students learn that what happens with one toss cannot possibly influence the outcome of the next toss. The misconception that results will be "evened out" is clarified; students learn that each toss is independent of the other tosses and has an equally likely chance of coming up heads or tails. Students further explore concepts of chance as they play the Sums Game. On discovering that the original game is unfair, students must use data they have collected to revise the game and make it fair.

Preparation and Materials

- Student Page 7: Coin-Toss Tally
- Student Page 8: Sums Scorecard
- A coin
- Two dice

Activity

Coin Toss

1. **Predict whether a coin will show heads or tails when flipped.**

   Begin by flipping a coin and covering it with your hand. Ask: *Do you think the coin came up heads or tails?* Show students the result. Flip the coin two more times, each time having students predict whether the outcome will be heads or tails. After three flips, ask:

   - *Of the three flips, how many were heads? How many were tails?*
   - *The next time I flip the coin, do you think it is more likely to be heads or tails? Why?*

Teaching Tip

Many students will suggest that if two or three heads have come up in a row, the next toss is sure to be tails to "even things out." A goal of the coin-toss activity is explore and clarify this misconception. What happens with one toss cannot possibly influence the outcome of the next toss. Each toss is independent of the other tosses and has an equally likely chance of coming up heads or tails. The misconception that results will be "evened out" is known as the "gambler's fallacy."

2. **Toss a coin 50 times and keep track of the results on Student Page 7: Coin-Toss Tally to determine which side of the coin is more likely to come up in a toss.**

   Before beginning, ask students to predict which side of the coin is likely to come up more often. Have students toss a coin 50 times and record the results on Student Page 7. After 25 tosses, ask students if they would like to change their predictions.

   Ask students how many times they would have to toss to make it an even number of heads and tails. Ask when they toss the next one what are the chances it will be heads.

   Explain that what happens in one toss cannot possibly influence what happens in the next toss. With each toss, there is an equally likely chance of getting heads or tails.

Sums

1. **List all the possible sums that can be made by rolling two dice.**

Ask: *When I roll two dice, how many sums are possible? Is it possible to roll a sum of 1? 2? 3?* The possible sums are 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, and 12.

Sums Game

Goal: To have the highest point total at the end of the game.

Players: 2

Materials: 2 dice; copies of Student Page 8: Sums Scorecard

Instructions: Assign one player to receive a point each time that a sum of 2, 3, 4, 5, or 6 is rolled and the other player to receive a point each time that a sum of 7, 8, 9, 10, or 11 is rolled. No one gets a point if 12 is rolled. Roll the dice 50 times, each time awarding a point to the appropriate player. Keep track of the sums and points on Student Page 8. The player with the highest number of points after 50 rolls wins. After 25 rolls, stop and ask: *Do you think this game is fair? Why, or why not?*

2. Create a line-plot that shows the frequency with which different sums occur.

The categories along the horizontal axis of the line-plot should be the different possible sums. For example:

					x					
			x		x	x				
	x		x	x	x	x	x	x		
2	3	4	5	6	7	8	9	10	11	12

3. Give a possible explanation for why the Sums Game is not fair as written.

Ask: *Which sums occur most often? Why do you think this is?* Have students list all the ways to make each sum.

4. Change the Sums Game rules so that it will be fair.

Have students explain how they changed the game and why they think the new version will be more fair. Play the new game a few times to test their theory.

Ideas for Data

6

Overview

In Activities 1–5, students used a set of 100 pennies to pose statistical questions and gather data, represent the data using different graphs, interpret and evaluate the data, and explore notions of probability. Similar investigations can be carried out, using different data. Suggestions for additional or alternate data explorations are given below, in the categories of: favorites, places, and things. Data explorations may also span the three categories.

Favorites

- *What is the most popular flavor of ice cream in my class?*
- *Among my friends, family, and/or classmates, what is the most popular field-trip location or excursion in the city?*
- *Is the favorite field-trip location the same as the one that has been visited most often?* (For example, the zoo might be the favorite spot, but, because of its distance from home or school, it may only have been visited once, whereas the natural science museum, because of its proximity to home or school, has been visited many times.)

Places

- *Of the places I have visited, which is the farthest from my home?*
- *Of the places I have gone this week, what method of transportation (e.g., bike, car, skateboard, school bus, subway, city bus) did I use most often?*

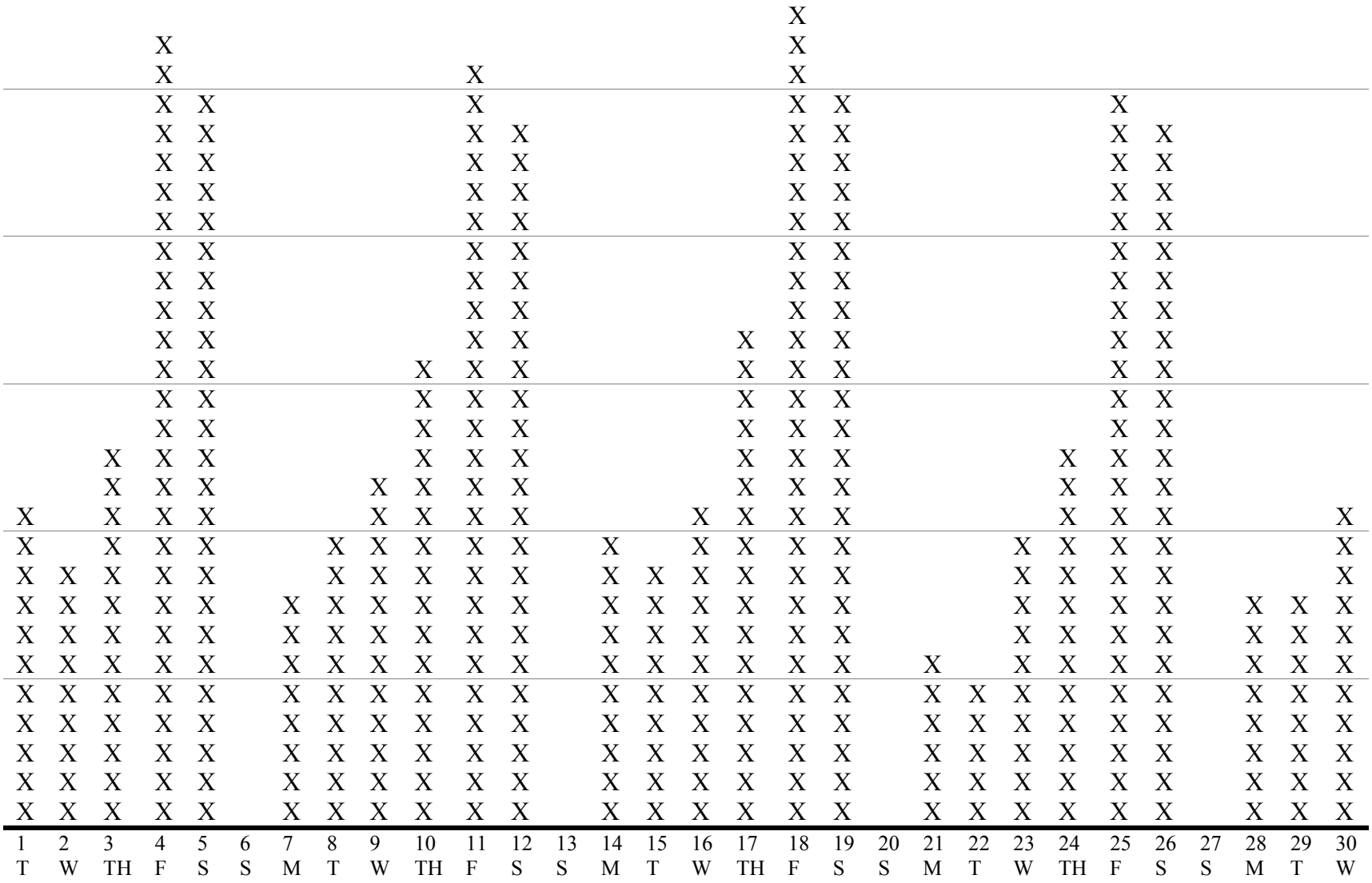
Things

- *How many different types of canned goods do we have in our kitchen? Which type do we have the most of? The least?*
- *In a bowl of 100 M&M's, how many M&M's are red? Green? Yellow? Blue? Orange? Dark brown? If I were to reach in and pull out one M&M, what is the likelihood that it would be green?*

Line-Plot

1

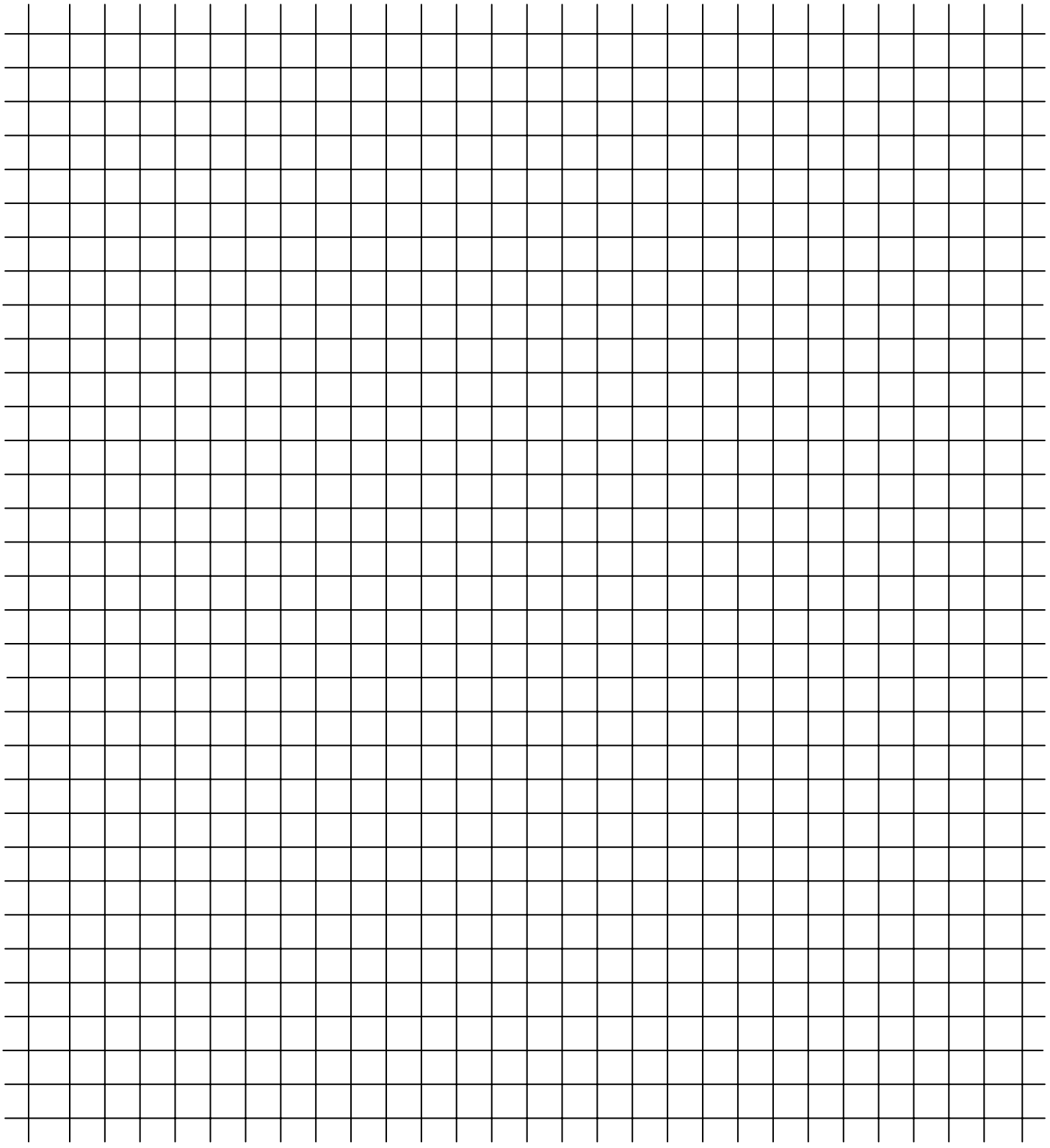
Tony's Pizza: Number of Large Pizzas Sold in April



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
T	W	TH	F	S	S	M	T	W	TH	F	S	S	M	T	W	TH	F	S	S	M	T	W	TH	F	S	S	M	T	W

Grid Paper

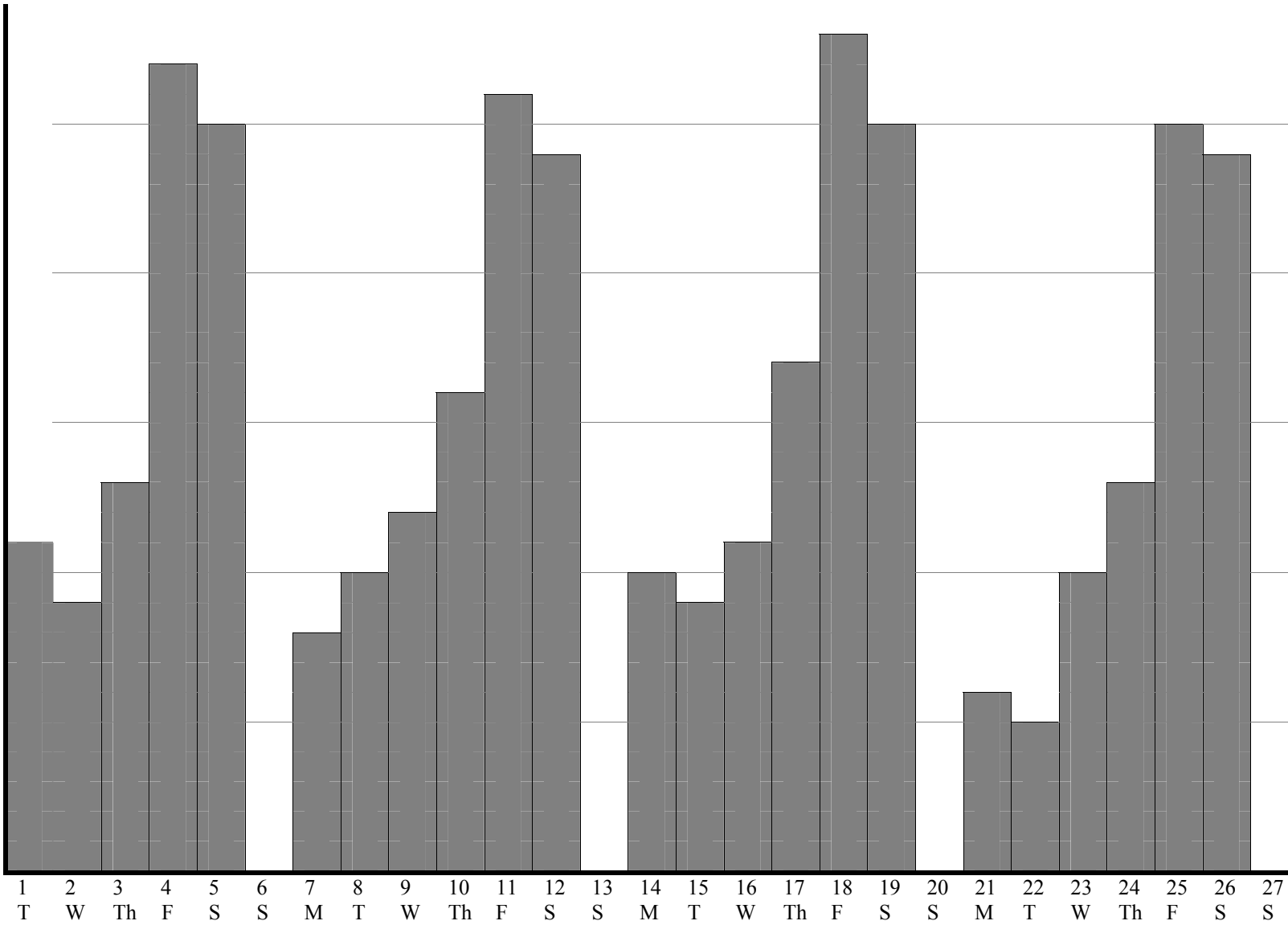
2



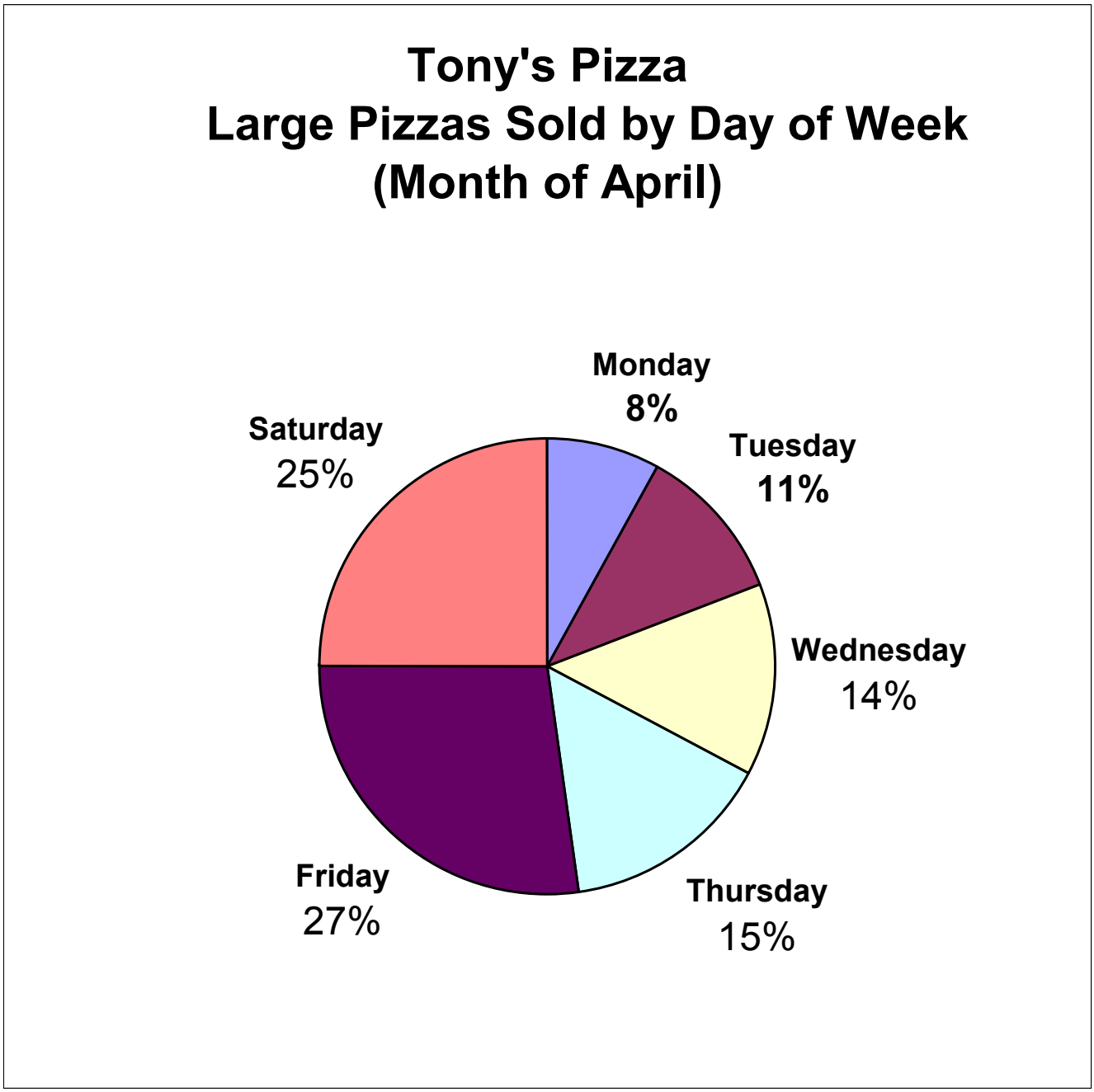
3

Bar Graph

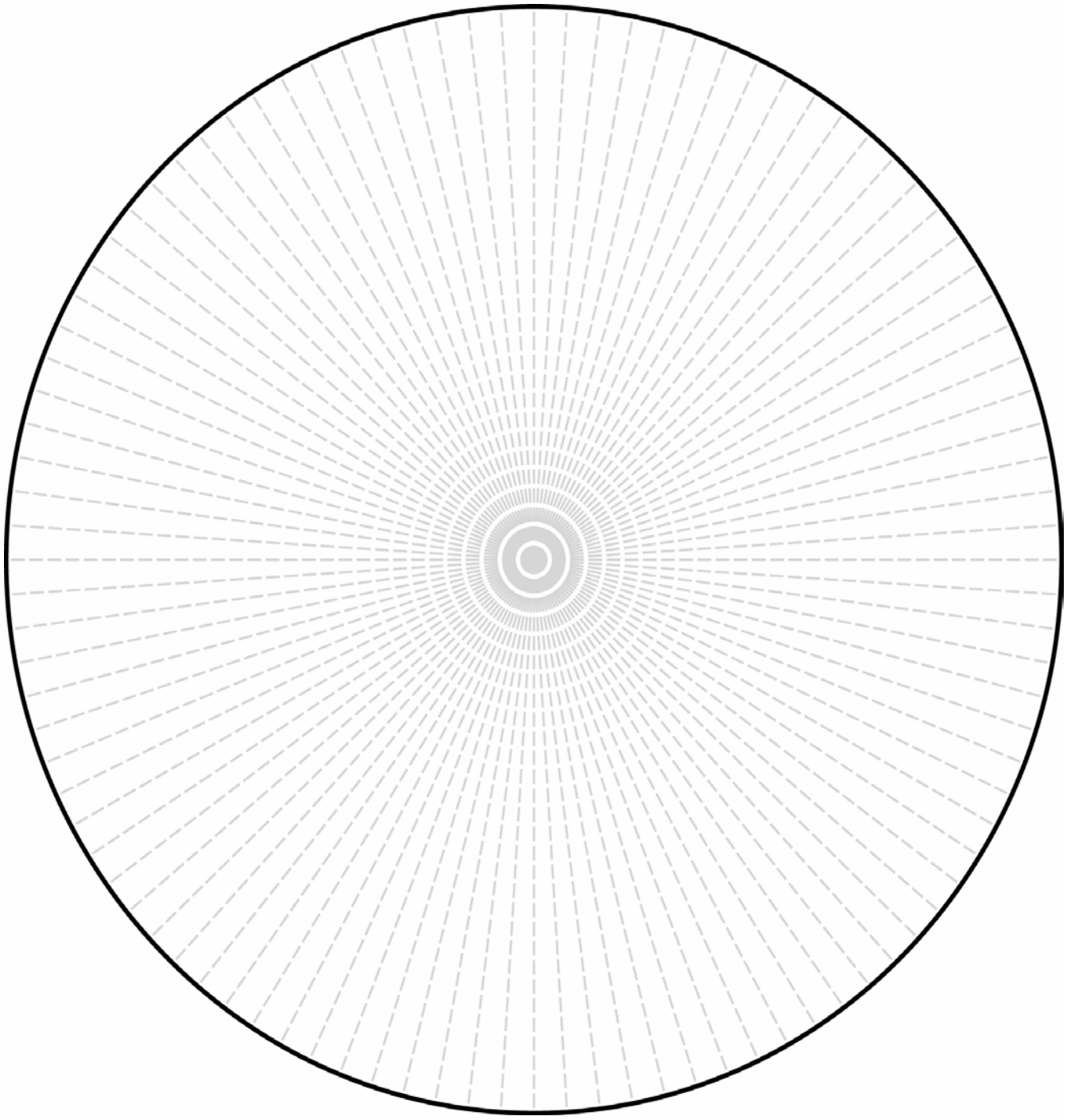
Tony's Pizza: Number of Large Pizzas Sold in April



Circle Graph



Blank Circle Graph



Interpreting the Data

RANGE: The range is the interval between the smallest and largest value in the data set. In this case, the range would be the interval between the date of the oldest penny and the newest penny.

OUTLIERS: These are isolated data points.

CLUSTERS: A cluster is a pile-up of data at one point on the number line.

GAPS: Dates with no X's or "holes" in the data, are referred to as gaps.

Coin-Toss Tally

Side	Tally	Total After 25 Tosses	Total After 50 Tosses
Heads			
Tails			

Sums Scorecard

Roll	Sum	If Sum = 2, 3, 4, 5, or 6, Point for Student(s)	If Sum = 7, 8, 9, 10, or 11, Point for Mentor
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			