# Describing Data Student Guide 

## Math 6 Unit 8 Accelerated

Part 2

## Glossary

| Term | Definition |
| :---: | :---: |
| experiment | An experiment is something you can do over and over again, and you don't know what will happen each time. <br> For example, each time you spin the spinner, it could land on red, blue, or green. |
| event | An event is a set of one or more outcomes in a chance experiment. For example, if we roll a number cube, there are six possible outcomes. |
| mean | The mean is one way to measure the center of a data set. We can think of it as a balance point. <br> For example, for the data set $7,9,12,13,14$, the mean is 11 . |
| mean absolute deviation (MAD) | The mean absolute deviation is one way to measure how spread out a data set is. Sometimes we call this the MAD. <br> To find the MAD, add up the distance between each data point and the mean. Then, divide by how many numbers there are. <br> For example, in this data set, the MAD is $\frac{4+2+1+2+3}{5}=\frac{12}{5}$ because there are 5 data points and each is a different distance away from the mean. <br> These travel times are typically $\frac{12}{5}$ or 2.4 minutes away from the mean. |

## desmos

Unit 7.8, Student Goals and Glossary

| median | The median is one way to measure the center of a data set. It is the middle number when the data set is listed in order. <br> For data set $A$, the median is 12 . <br> Set A: 7, 9, 12, 13, 14 <br> For data set $B$, there are two numbers in the middle. <br> Set B: 3, 5, 11, 12 <br> The median is the average of those numbers: $\frac{5+11}{2}=8$ |
| :---: | :---: |
| interquartile range (IQR) | The interquartile range is one way to measure how spread out a data set is. We also call this the IQR. <br> To find the interquartile range, we subtract the first quartile from the third quartile. <br> For example, the IQR of this data set is 20 because $50-30=20$. |
| outcome | An outcome of a chance experiment is one of the things that can happen when you do the experiment. For example, the possible outcomes of tossing a coin are heads and tails. |
| population | A population is a set of people or things that we want to study. For example, if we want to study the heights of people on different sports teams, the population would be all the people on the teams. |
| probability | The probability of an event is a number that tells how likely it is to happen. A probability of 1 means the event will always happen. A probability of 0 means the event will never happen. <br> For example, the probability of spinning red is $\frac{1}{4}$. |
| random | Outcomes of a chance experiment are random if they are all equally likely to happen. |

## desmos

Unit 7.8, Student Goals and Glossary

| representative | A sample is representative of a population if its distribution resembles the <br> population's distribution in center, shape, and spread. |
| :---: | :--- |
| sample | A sample is part of a population. For example, a population could be all the <br> seventh grade students at one school. One sample of that population is all the <br> seventh grade students who are in band. |
| sample space | The sample space is the list of every possible outcome <br> for a chance experiment. <br> For example, the sample space for tossing two coins <br> is: heads-heads, tails-heads, heads-tails, tails-tails. |
| simulation | A simulation is an experiment that is used to estimate <br> the probability of a real-world event. <br> For example, suppose the weather forecast says there <br> is a 25\% chance of rain. We can simulate this situation <br> with a spinner. If the spinner stops on red, it represents <br> rain. If not, it represents no rain. |

## Unit 8 Summary

| Prior Learning | Math 7, Unit 8 | Future Learning |
| :---: | :---: | :---: |
| - Statistical variability | - Probability | Math 8, Unit 6 |
| - Data distributions | - Sampling | - Scatter plots <br> - Lines of best fit |
| Math 7 <br> - Proportional reasoning (Unit 2) <br> - Percent change (Unit 4) | - Using center and variability to compare populations. |  |

## Probability

A probability is a number that represents how likely something is to happen.
If you flip a coin, the probability of the coin landing heads up is 0.5 .
The probability that the coin turns into a bunny is 0 .


When you repeat an experiment, the results get closer and closer to the probability.

If you flip a coin 100 times, it might land heads up 50 times, 49 times, 52 times, or maybe even 60 times. As you flip the coin more times, the fraction of heads gets closer to 0.5.


In an experiment, it is often helpful to know the sample space, a list of every possible outcome.

Here are two ways to represent the 12 possible outcomes of flipping a coin and then rolling a number cube.


Tree


## desmos

## Unit 7.8, Family Resource

## Sampling

Sometimes we want to know information about a group, but the group is too large for us to be able to ask everyone. It can be useful to collect data from a sample (some of the group) of the population (the whole group).

For example, we might want to know what percentage of Americans work from home. It would be too challenging to ask all Americans, so we can ask a smaller sample of
 working Americans.


If our sample is some employees at a grocery store, we may get very different results than if our sample is some employees of a technology company. Neither sample is representative of the population of all working Americans.

A sample selected at random is most likely to be representative of the population because it has the chance to include all kinds of working adults.

We can use samples to estimate information about a population and to make comparisons.
For example, the Minty Fresh company collects data from a sample of gum buyers to see how many minutes it takes for their mint gum to lose flavor compared to other brands.
These box plots show that the median time for the sample of Minty Fresh Gum is longer than the sample of other brands.


However, there may not be a difference between two populations because there is a lot of overlap between the two samples, and because the difference in the medians is less than one IQR (interquartile range).


If you are interested in learning more about how these statistical measures are calculated, see the Statistics Summary.

## desmos

## Unit 7.8, Family Resource

## Try This at Home

## Probability

DesPops makes a mystery lollipop. Each lollipop is one of these flavors: strawberry, lime, pineapple, or mango, but you don't know which flavor you have until you try it.
1.1 What is the probability of getting a strawberry flavored lollipop?
1.2 Would it be surprising to buy 5 lollipops and none were strawberry? Explain your thinking.
1.3 Hailey says that if she buys 200 lollipops, exactly 50 of them will be mango. Explain to Hailey what is incorrect about her reasoning.

A game uses cards that say "forward" or "backward" and a spinner numbered from 1 to 5 . On their turn, a player picks a card and spins the spinner to find out which way and how much to move.
2.1 How many different outcomes are possible?
2.2 What is the probability that the next player will get to move their piece forward 5 spaces?
2.3 What is the probability that the next player will have to move their piece backward an odd number of spaces?

## Sampling

A city council wants to know how many buildings in the city have lead paint, but they don't have enough time to test all 100000 buildings in the city. They want to test a sample of buildings that will be representative of the population.
3.1 Describe a way to pick a sample that is likely to be representative of the population.
3.2 Describe a way to pick a sample that is not likely to be representative of the population.

Each year Brielle measures a random sample of her tomato plants.
Here is her data for the past two years.
4. Brielle is nervous that her garden this year isn't as healthy as last year's. Do you agree? Use at least one piece of evidence to support your claim.

Tomato Plant Heights<br>Last Year<br>Mean: 55 inches<br>MAD: 7.8 inches<br>This Year<br>Mean: 50.5 inches<br>MAD: 8 inches

## desmos

## Unit 7.8, Family Resource

## Solutions:

$1.1 \quad \frac{1}{4}$ (or equivalent)
1.2 No. Explanations vary. With a sample of only 5 lollipops, it isn't surprising. The more lollipops you buy, the closer the fraction of strawberry lollipops will get to the probability.
1.3 Responses vary. Even though about 25\% of your lollipops will be mango, we do not know exactly how many will be. It might be 50 lollipops, slightly more, or slightly less. There is even a small chance it will be way more or less than 50.
2.1 10 possible outcomes
$2.2 \quad \frac{1}{10}$ (or equivalent)
$2.3 \quad \frac{3}{10}$ (or equivalent)
3.1 Responses vary.

- Put the addresses of all the buildings into a computer and have the computer select 50 addresses randomly from the list.
- Put all of the street names in a bag. Pick several streets randomly and then test all the houses on that street.


### 3.2 Responses vary.

- Testing all the same type of buildings (like all the schools or all the gas stations).
- Testing buildings all in the same location, such as the buildings closest to city hall.
- Testing all the newest buildings or all of the oldest buildings.
- Testing a small number of buildings, like 5 or 10.

4. No, she does not need to be nervous.

Explanations vary. Even though this year, the mean height of her sample is less than it was last year, the difference is not that big. This year's plants are 4.5 inches shorter than last year's, which is less than the MAD. This means it's more likely that she just happened to end up with smaller plants in her sample, but that the populations are not that different.
$\qquad$

## My Notes

Alisha wants to know how many minutes 7 th graders at her school spend on their cell phone each day.

1. What is the population for Alisha's question?
2. What is a sample Alisha could use to help answer this question?

Alisha asked 50 random 7th graders how many minutes they spend on their phone per day and calculated a sample mean of 65 minutes.

3. Why would Alisha decide to collect a sample to answer her question rather than use the population?

## Summary

I can explain what a sample is and when it is useful.
I can compare the means of samples to the mean of the population.
$\qquad$

My Notes
Alisha wants to know how many minutes 7th graders at her school spend on their cell phone each day.

1. What is the population for Alisha's question?

The population is all the 7th graders at Alisha's school.
2. What is a sample Alisha could use to help answer this question?

A sample could be all the 7th graders in Alisha's first class.

Alisha asked 50 random 7th graders how many minutes they spend on their phone per day and calculated a sample mean of 65 minutes.

3. Why would Alisha decide to collect a sample to answer her question rather than use the population?

Responses vary. It will take a lot of time and effort to ask all of the 7th graders.

## Summary

I can explain what a sample is and when it is useful.
I can compare the means of samples to the mean of the population.
$\qquad$

My Notes

1. In your own words, explain what it means for a sample to be representative of a population.
2. Match each headline with the sampling method that most likely produced it.

Headline
One Quarter of Working
Americans Spend Time
Working From Home!
Time Workt Americans Spend
From Home!
From Home!

## Sampling Method

A. Ask all 100 employees at one technology company.
B. Ask all the employees at 100 random grocery stores.
C. Call random phone numbers until you ask 100 people.
3. Which sampling method above is most likely to produce a representative sample? Explain your thinking.

## Summary

$\qquad$

My Notes

1. In your own words, explain what it means for a sample to be representative of a population. Responses vary.

- There is data in your sample from many parts of the population. For example, it is not only people from one grade or class.
- A sample is representative if it looks like a smaller version of the population. If there are half dogs and half cats in the population, then there should be about half dogs and half cats in the sample.

2. Match each headline with the sampling method that most likely produced it.


## Headline

C. One Quarter of Working Americans Spend Time Working From Home!
A. Most Americans Spend Time Working From Home!
B. Almost No One Works From Home!

## Sampling Method

A. Ask all 100 employees at one technology company.
B. Ask all the employees at 100 random grocery stores.
C. Call random phone numbers until you ask 100 people.
3. Which sampling method above is most likely to produce a representative sample? Explain your thinking.
C. Explanations vary. A and B could both be biased because employees at a technology company or at grocery stores aren't representative of all working Americans.

## Summary

$\qquad$

My Notes
Cameron bought a bag of White Flower Seed Mix and is curious how many flowers of each type there are. Cameron planted 25 seeds, and these were the results.

| Flower Type | Count | Percentage |
| :---: | :---: | :---: |
| Daisy | 14 |  |
| White Zinnia | 5 |  |
| Aster | 6 |  |
| Total | 25 |  |



1. Complete the table with the percentage of each flower type.
2.1 Estimate how many of the 600 seeds in the bag will be asters. Organize your calculations so others can follow them.
2.2 What could you do to be more confident in your estimate?

## Summary

$\qquad$

My Notes
Cameron bought a bag of White Flower Seed Mix and is curious how many flowers of each type there are. Cameron planted 25 seeds, and these were the results.

| Flower Type | Count | Percentage |
| :---: | :---: | :---: |
| Daisy | 14 | $56 \%$ |
| White Zinnia | 5 | $20 \%$ |
| Aster | 6 | $24 \%$ |
| Total | 25 | $100 \%$ |



1. Complete the table with the percentage of each flower type.
2.1 Estimate how many of the 600 seeds in the bag will be asters.

Organize your calculations so others can follow them.
Calculations vary.

- $24 \%$ of $600=0.24 \cdot 600=144$ asters
- $25 \cdot 24=600$ and $6 \cdot 24=144$ asters
2.2 What could you do to be more confident in your estimate? Responses vary.
- You could plant more samples of 25 seeds and see if the results are similar between the samples.
- You could plant a larger sample, like 100 seeds. A larger sample might be more accurate.


## Summary

$\qquad$

My Notes
The Minty Fresh company collects data to see how many minutes it takes for their mint gum to lose flavor compared to other brands.


Based on this data . . .
1.1 Why might someone believe this headline?
1.2 Why might someone not believe this headline?
2. How many IQRs is the difference between the medians?
3. Do your calculations make you believe the headline more or less? Explain your thinking.

## Summary

I can use measures of center and the variability of two samples to decide if two populations are very different.

I can use a measure of variability to explain the difference between measures of center.
$\qquad$

My Notes
The Minty Fresh company collects data to see how many minutes it takes for their mint gum to lose flavor compared to other brands.


Based on this data . . .
1.1 Why might someone believe this headline?

Responses vary. The median of Minty Fresh is higher than other brands. Also, it's more consistent.
1.2 Why might someone not believe this headline?

Responses vary. There are lots of times where the other brands of gum last longer than Minty Fresh.
2. How many IQRs is the difference between the medians?

The difference between the medians is $\frac{67-57}{25}=\frac{10}{25}$ or about 0.4 IQRs.
3. Do your calculations make you believe the headline more or less? Less

Explanations vary. If there really was a big difference, I would expect the difference between the medians to be more than the IQR. The difference isn't even one IQR, so the difference between Minty Fresh and other gums isn't that big.

## Summary

I can use measures of center and the variability of two samples to decide if two populations are very different.

I can use a measure of variability to explain the difference between measures of center.

## desmos 目

Unit 7.8, Lesson 15: Notes
Name $\qquad$


Here are statistics based on a sample from Atlanta, Georgia.
3. How does the asthma rate in Montgomery compare to the asthma rate in Atlanta?

## Asthma Rates

Mean: $\approx 115$
MAD: $\approx 9.2$

Median: 116
IQR: 10

## Summary

I can compare two groups by taking random samples, then calculating and interpreting the statistics.

Note: The data is derived from Centers for Disease Control and Prevention, National Center for Chronic Disease and Health Promotion, Division of Population Health, Atlanta, GA based on BRFSS 2018 or 2017, Census 2010 population counts or census county population estimates of 2018 or 2017, and ACS 2014-2018 or ACS 2013-2017.
$\qquad$


I can compare two groups by taking random samples, then calculating and interpreting the statistics.

## desmos 目

$\qquad$

My Notes
The letters PROBABILITY are in a bag. In an experiment, one letter is picked at a time and returned to the bag.

1. What is the sample space for this experiment?
2. Describe how to determine the probability of picking a $Y$ from the bag on the next pick.

Write the probability of each event happening on the next pick.
3. Picking the letter $P$ from the bag. $\qquad$
4. Picking the letter B from the bag. $\qquad$
5. Picking the letters B or I from the bag. $\qquad$
6. Picking the letter $S$ from the bag. $\qquad$
7.


Label the pieces in the bag so that it has these probabilities on the next pick:

- The probability of picking a $P$ is $\frac{1}{4}$.
- The probability of picking a $B$ is 0 .
- The probability of picking a G is equal to the probability of picking an R.

Summary

I can determine the probability of an event using its sample space.
I can compare probabilities written as fractions, decimals, and percentages.
$\qquad$

My Notes
The letters PROBABILITY are in a bag. In an experiment, one letter is picked at a time and returned to the bag.

1. What is the sample space for this experiment?

The sample space is the letters $P, R, O, B, A, B, I, L, I, T, Y$.
2. Describe how to determine the probability of picking a $Y$ from the bag on the next pick.

Responses vary. Count the letters in the sample space (11) and see how many are a $\mathbf{Y}(1)$. The probability is $\frac{1}{11}$.

Write the probability of each event happening on the next pick.
3. Picking the letter $P$ from the bag. $\frac{1}{11}$
4. Picking the letter $B$ from the bag. $\frac{2}{11}$
5. Picking the letters B or I from the bag. $\frac{4}{11}$
6. Picking the letter $S$ from the bag. 0
7.


Label the pieces in the bag so that it has these probabilities on the next pick:

- The probability of picking a P is $\frac{1}{4}$.
- The probability of picking a $B$ is 0 .
- The probability of picking a $G$ is equal to the probability of picking an R.

Summary

I can determine the probability of an event using its sample space.
I can compare probabilities written as fractions, decimals, and percentages.

## desmos 目

$\qquad$

My Notes
There are 6 blocks in a bag. The table shows results from 100 picks.

| Block <br> Color | Number <br> of Picks |
| :---: | :---: |
| Purple | 32 |
| Red | 68 |

1. Based on these results, how many of the 6 blocks are red?

Explain your thinking.
2.1 Design a bag where the probability of picking a green block is $60 \%$.
2.2 About how many times out of 50 picks
 do you expect to pick a green block?

Explain how you know.

## Summary

I know that sometimes outcomes of an experiment are not equally likely.
I can use proportional reasoning with data from a repeated experiment to make predictions.
$\qquad$

My Notes
There are 6 blocks in a bag. The table shows results from 100 picks.

| Block <br> Color | Number <br> of Picks |
| :---: | :---: |
| Purple | 32 |
| Red | 68 |

1. Based on these results, how many of the 6 blocks are red?

## 4 blocks

Explain your thinking.
Explanations vary. $\frac{68}{100}$ is $68 \%$, and $68 \%$ of 6 blocks is about 4 blocks.
2.1 Design a bag where the probability of picking a green block is $60 \%$.

Bags vary.

- 3 green blocks and 2 red blocks

- 6 green blocks and 4 blue blocks
2.2 About how many times out of 50 picks do you expect to pick a green block?

About 30 times.
Explain how you know. Explanations vary. Since 60\% of the blocks are green, I would expect to pick a green block close to $60 \%$ of the time. $60 \%$ of 50 is 30 times.

## Summary

## desmos 目

$\qquad$

My Notes
Here is a graph of the results of 100 rolls of a number cube.
Fraction of Rolls That Are Twos (So Far)



1. What is the probability of rolling a 2 with a standard number cube?
2. Based on these results, what is the probability of rolling a 2 with this number cube?
3. Explain how you know this number cube is unfair.
4. Describe or draw what this number cube could look like.

## Summary

I can decide whether or not something is fair based on the results of a repeated experiment.
I can use the results from a repeated experiment to approximate the probability of an event.
$\qquad$

My Notes
Here is a graph of the results of 100 rolls of a number cube.

Fraction of Rolls That Are Twos (So Far)



1. What is the probability of rolling a 2 with a standard number cube?

$$
\frac{1}{6}
$$

2. Based on these results, what is the probability of rolling a 2 with this number cube?

$$
0.5 \text { or } \frac{1}{2}
$$

3. Explain how you know this number cube is unfair.

Explanations vary. If the number cube were fair, the graph would be getting closer to $\frac{1}{6}$. This graph got closer to $\frac{1}{2}$, which means that rolling a 2 is more likely than it should be.
4. Describe or draw what this number cube could look like. Responses vary.

- Half of the sides are 2 s . The other half are 3 s .
- The cube has the numbers 1,2,2,2,5, and 6 on it.


## Summary

I can decide whether or not something is fair based on the results of a repeated experiment.
I can use the results from a repeated experiment to approximate the probability of an event.
$\qquad$

## My Notes

Here is a game involving a spinner and a fair coin.

1.2 What is the probability of 1.3 Describe your strategy. getting an odd number and tails?
1.1 How many outcomes are in the sample space?

A new game is played with a spinner and a number cube.

2.1 Make a tree, table, or list to represent the sample space.
2.2 What is the probability of getting an " $A$ " and an even number?

Summary

I can write out the sample space for a multistep experiment using a list, table, or tree diagram.
I can calculate the probability of a multistep event.

## desmos目

$\qquad$

My Notes
Here is a game involving a spinner and a fair coin.

1.2 What is the probability of getting an odd number and tails?
$\frac{2}{6}$ (or equivalent)
1.1 How many outcomes are in the sample space?

## 6

1.3 Describe your strategy.

Responses vary. I made a list of all the outcomes. You can roll a 1 or 3 with tails, which is 2 out of the 6 total outcomes.

A new game is played with a spinner and a number cube.

2.1 Make a tree, table, or list to represent the sample space.
Responses vary.

|  | 1 | 2 | 3 | 4 | 4 | 5 | 6 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| A |  |  |  |  |  |  |  |
| B |  |  |  |  |  |  |  |

2.2 What is the probability of getting an " $A$ " and an even number?
$\frac{3}{12}$ (or equivalent)

## Summary

## desmos 目

$\qquad$

My Notes
Brianna designed a simulation to help her estimate the probability of seeing her three favorite animals awake when she visits the zoo.


1. Describe how Brianna could use these spinners to estimate the probability that at least two of her favorite animals will be awake when she visits the zoo.

The table shows the results of 300 experiments with the spinners.

| Experiments with . . . | Count | Percentage |
| :--- | :---: | :---: |
| No animals awake | 12 | $4 \%$ |
| 1 animal awake | 171 | $57 \%$ |
| 2 animals awake | 105 | $35 \%$ |
| 3 animals awake | 12 | $4 \%$ |

2. Estimate the probability that at least 2 of Brianna's favorite animals will be awake when she visits the zoo.

## Summary

I can use a simulation to estimate the probability of a multistep real-world event.
I can connect real-world situations and the probability tools I could use to simulate those situations.
$\qquad$

My Notes
Brianna designed a simulation to help her estimate the probability of seeing her three favorite animals awake when she visits the zoo.


1. Describe how Brianna could use these spinners to estimate the probability that at least two of her favorite animals will be awake when she visits the zoo.

Responses vary. Briana could run a lot of experiments and record how many of the experiments had at least two spinners land on an open eye.

The table shows the results of 300 experiments with the spinners.

| Experiments with ... | Count | Percentage |
| :--- | :---: | :---: |
| No animals awake | 12 | $4 \%$ |
| 1 animal awake | 171 | $57 \%$ |
| 2 animals awake | 105 | $35 \%$ |
| 3 animals awake | 12 | $4 \%$ |

2. Estimate the probability that at least 2 of Brianna's favorite animals will be awake when she visits the zoo.

39\%

## Summary

I can use a simulation to estimate the probability of a multistep real-world event.
I can connect real-world situations and the probability tools I could use to simulate those situations.

