

M112B Performance Task

Let's Get Weird, Let's Get Real

Modeling with Data: Two-Variable Measurement Data
Teacher Version

Design Notes

In this task, students will take on the role of a researcher and statistician to examine seemingly random and unrelated numeric data before examining numeric related data. In part one, students will first explore random data like their classmates' birth months and the number of pets they have. In part two, students will brainstorm a problem that they want to explore and collect data on two variables. For both parts, students will determine models of best fit and discuss what they notice and learn about correlations and causation.

Student-teacher consultations have been built in to scaffold the process. Consider using these structured check-in points as places where select students share their approach/findings with the class to support the collective learning of the class. This learning experience centers around the CCSSM modeling cycle as illustrated below:

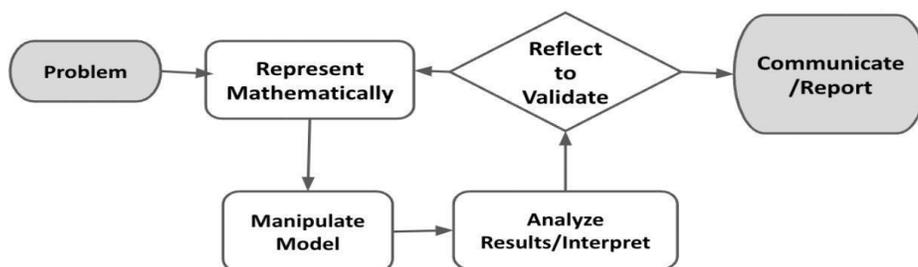


Figure 1. *Diagram of Modeling Process.* From “Progressions for the Common Core State Standards for Mathematics” [Draft manuscript], by CCSS Writing Team, 2019, p. 6.

Connections to Learning Principles

Students will:

- Analyze the relationships between random and related variables through multi-step data collection, analysis, and interpretation (LP1);
- Work independently and share their findings with others to deepen their understanding on correlated variables (LP2);
- Decide on the topic they would like to further investigate through data collection (LP4); and
- Utilize technology to create a model (LP7).

The following pages provide teacher versions of this task with guidance, suggestions, and comments in **red** and sample student responses in *blue italics*.

Student Task

The Modeling Cycle

Do you think your hand size predicts your math grade? Is there a relationship between the width of your earlobe and your shoe size? Have you measured how fast you can run or how high your friends can jump? How long can your teachers hold their breath, and does it depend on the numerical day of the week they were born? This task is divided into two parts. First, you will collect data on two seemingly unrelated variables and determine if there is a correlation. Then you will explore possible relationships between two variables for a topic of your choice.

Mathematical models are used to make sense of ideas and phenomena in our world. As illustrated in the following diagram, the modeling cycle is a process of collecting data about an idea or problem, creating models of that data, analyzing the story told by those models, and reporting the results to the appropriate audience.

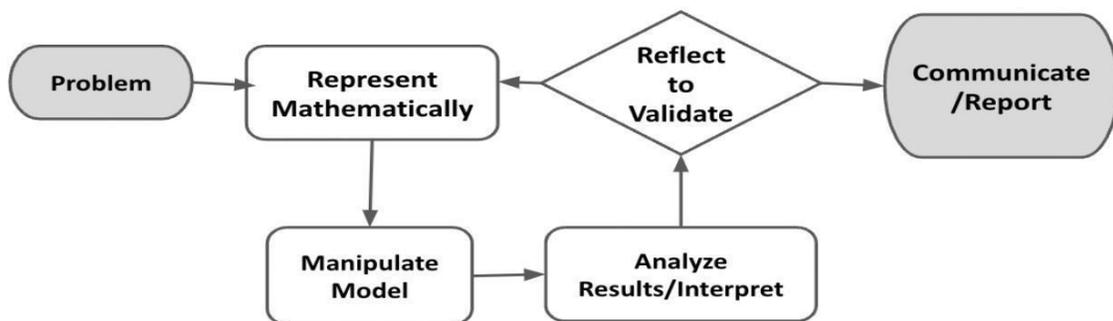


Figure 1. *Diagram of Modeling Process.* From “Progressions for the Common Core State Standards for Mathematics” [Draft manuscript], by CCSS Writing Team, 2019, p. 6.

In this task, you will be engaging with this mathematical modeling cycle to collect and analyze data, evaluate findings, and identify trends in two variables. You will use the models and your takeaways from this task to create a presentation, informational video, poster, or another creative product to communicate your hypothesis and findings. You will also interpret your findings and explain how your data and model can be used to benefit others.

PART 1: Random Relationships

Launch

There are different types of data that you can collect in the world. The most common types are categorical data and numeric data. Categorical data are descriptive or qualitative. Examples of categorical data include name, hair color, brand of sneaker, type of shampoo, school subject, favorite sport, and travel method to school, to name a few. In contrast, numerical data are quantitative — values that can be measured and placed into a logical order. Examples of numerical data can include height, birth month, age, number of pets owned, shoe size, birth order compared to siblings, average temperature on your birthday, SAT Math score, and how quickly you can run a mile, among others.

QUESTION 1



Think about these different types of data and any others that come to mind. Brainstorm some information that you are curious to learn about your classmates.

Answers will vary. Possible student responses:

- *Favorite color*
- *Grade in math class*
- *Hobbies*
- *Number of hours you sleep*
- *Number of hours spent studying*
- *Number of hours spent on social media*
- *Number of siblings*

The purpose of this question is to get students thinking about categorical and numeric variables that pertain to their classmates that they are interested in learning about. Consider beginning with a quick open discussion led by the teacher to help students begin the brainstorming process before having the students brainstorm independently or in small groups. As a class, you can ask the students what their favorite season or sport is and keep track of their responses on the board.

QUESTION 2



Now identify two numeric variables that are seemingly unrelated. You can choose variables from your list above or something new that came to mind.

Answers will vary. Possible student responses:

- Birth month and number of technological devices at home
- Height and TikTok views
- Number of hobbies and length of hair
- Shoe size and the number of pets
- Size of wrist and grade point average

The purpose of this question is for students to choose numeric variables that are random, and that you would not expect there to be any correlation. Encourage students to choose variables that are as unrelated as possible.

Consider engaging students in a whole group discussion about related variables versus unrelated variables. Related variables such as temperature outside and ice cream sales could possibly result in a positive correlation and have a causal relationship whereas that would not have been the case for temperature and choice of ice cream flavor.

Perhaps make it into a contest to see which students can find the highest correlation among data that are initially identified as the most random and possibly least related.

In the **Problem** and **Formulate** sections, you will analyze the data you collected to determine the possible relationship; then you will evaluate this relationship to identify a model of best fit and create a related equation. Specifically, consider linear, quadratic, and exponential functions as possible models for the relationship. The **Problem Formulation Checklist** can be used to ensure you have the necessary components before moving on to the remaining parts of the task.

Problem

Let's get weird. Can we find relationships by looking at two seemingly unrelated variables? Survey your classmates to gather information on the two variables that you chose.

QUESTION 3



What two variables did you choose to collect data on? How do you know your variables are random and seemingly unrelated?

Answers will vary. Possible student responses:

- Variables from above seemed fun to research. I am curious to know if there is a relationship between height and the number of pets a person has.
- I know they are unrelated because [variable 1] has nothing to do with [variable 2].

The purpose of this question is for students to clearly state the variables that they need to collect data on and explain their rationale for why they are unrelated.

ACTION 1

Conduct a survey to collect data from your classmates on the two variables that you chose.

Students will conduct their survey based on the guidelines that you set for data collection.

The purpose of this action is for students to collect data. Consider confirming student choice of random variables before data is collected. In addition, there will need to be some kind of “Class Data Collection Chart” to make data collection timely and manageable because every student needs to collect data on two variables from all their peers (see example below).

Data collection can be done electronically with the suggested chart below or on paper. The number of columns will depend on the unique variables that need to be collected. The number of rows will depend on the total number of students in your class. Electronically, this can be done with one Google document that is shared with each student. The linked Google document has a chart where the top row is for each student to type their variable and the left most column for each student to put their name. If technology is unavailable, a similar data chart can be created on the classroom board or students can circulate their own data chart to their classmates.

Sample Student Data Collection Chart

	<i>[Numeric variable here]</i>	<i>[Numeric variable here]</i>	<i>[Numeric variable here]</i>	<i>[Continue adding columns until you reach the total number of students in your class]</i>
<i>[Student name here]</i>				
<i>[Student name here]</i>				
<i>[Continue adding rows until you reach the total number of students in your class]</i>				

QUESTION 4



Graph the data. Describe the shape of the data. Do you see a relationship between the two variables of data? Why or why not?

Student graphs will vary. Possible student responses:

- *The data have no apparent shape because they show no pattern.*
- *The data points are scattered, so there is no real shape.*
- *There is no relationship between the two variables because they are independent of each other.*
- *There is some type of relationship because they kind of look linear, but I'm not sure.*

The purpose of this question is for students to use a scatter plot to represent two-variable measurement data. Then they describe the shape of their data and note any visible patterns or relationships between the data such as positive correlation, negative correlation, or no correlation.

Formulate

QUESTION 5



Using **Desmos**, determine a model of best fit for this data. Would any other model possibly fit? If no model fits this data, explain why not.

Students should attempt to fit their data to a linear, quadratic, or exponential model. You should see them sketching lines, parabolas, or curves. They may also decide to use technology such as Desmos or graphing calculators to create an equation for the line of best fit if they notice a linear association. Possible student responses:

- *I tried fitting my data to a function, but nothing made sense.*
- *I don't think any model will work because the points are just scattered, which makes sense because these variables are unrelated.*
- *A linear model kind of works for this data, but it is not highly correlated.*

The purpose of this question is for students to attempt to fit their two-variable measurement data to either a linear, quadratic, or exponential function. Given that the two variables are unrelated, there is a high chance that students will be unable to fit their data to a linear, quadratic, or exponential model.

QUESTION 6



Do you believe this model is a good fit for the data that you have collected? Why or why not? How strong is the correlation for the model that you chose?

Answers will vary based on variables that students chose and whether there is an actual correlation and depending on if the data fits linearly or with a quadratic or exponential function.

Possible student responses:

- *This model is a decent fit because the correlation coefficient is between 0.65 and 0.85.*
- *There is no relationship between the two variables because they are independent of each other.*

The purpose of this question is for students to assess the fit of their function using observations, correlation coefficients, residuals, or other tools. Again, this may not happen given that the variables are unrelated. It is important that students recognize why some variables in the real world may not be related and thus may not be able to be modeled by linear, exponential, or quadratic functions.

Problem Formulation Checklist

Requirement	Teacher Notes: For each Requirement, what are some specific things that came up in that part of the task?
Two seemingly unrelated variables have been selected.	(Teachers, please add your notes here.)
The data points for these two variables have been accurately plotted, and the shape of the graph has been clearly described.	(Teachers, please add your notes here.)
Where possible, I identified a model and created an equation for relating the two variables. If not possible, I gave a thorough explanation as to why a model does not fit.	(Teachers, please add your notes here.)
General Teacher Notes:	(Teachers, please add your notes here.)

PART 2: Real World Relationships

Launch

Keisha has noticed that the number of views on her TikTok videos are not surpassing 500. She wonders if there is a relationship between the length of her videos (which can be long) and the number of views. She identifies her problem as: TikTok video views are stuck between 400 and 500. She brainstorms possible variables that could be contributing to this: length of video, video content, age of content creator, and number of followers. What variables may be correlated? What questions can Keisha pose for this problem?

Similar to Keisha, you may have your own questions. Are high school students spending too much time on social media? How do sleep patterns impact your test scores? How can I improve my batting average? In the second part of this task, you and your classmates will act as researchers and statisticians to answer a question of interest. This task will give you an opportunity to explore possible relationships between two seemingly related numerical variables for a topic of your choice. This can be education, gaming, social media, sports, or anything else you are interested in!

Problem

QUESTION 7

Brainstorm a question that you have or that maybe you have heard your classmates or teachers discuss. Create a hypothesis that clearly states the two numeric variables for the question that you want to investigate. For example, a hypothesis can be that taller baseball players have a higher batting average because they have a longer reach. In this example, your two numeric variables would be the height of the baseball player and that player's batting average.

What is your hypothesis? Why do you think these variables are related?

Answers will vary. Possible student responses:

- *What is the average weekly coffee consumption among high school students, and does this negatively impact their grades? My hypothesis is that the more coffee you drink, the lower your grades are because caffeine increases anxiety.*

The purpose of this question is to get students thinking about a problem they have (similar to the example in the Launch) and/or a question they want to answer. Consider beginning with an open discussion about the Launch where students share variables that they think are correlated and pose possible questions that Keisha could ask. This should help students begin their brainstorming process.

In Part 2, it is very important for students to have choice when determining their questions but use your discretion when reviewing their questions and hypotheses. It may be necessary to name for your students that their peers do not have to share personal information for data collection purposes. Also be aware that the variables they choose should seemingly have some type of relationship, unlike Part 1.

QUESTION 8

What is your plan for data collection — would you like to collect data on yourself? Would you like to collect data from your classmates or from a publicly available source? Cite your sources if you decide to use publicly available data.

Answers will vary, but students should choose to collect data on themselves, their peers, or a publicly available source.

The purpose of this question is for students to choose how they will collect data. In terms of the number of data points, provide students with a requirement such as 75% to 100% of the total number of peers in this class. Make sure that students clarify the units of measure they will use and are consistent. For example, if you are using a variable defined by volume, make sure the students are consistent in using liters, milliliters, cups, etc.

Possible websites students could use:

- Sports
 - [Sports News, Scores, Predictions and Analysis](#)
 - <https://sports.yahoo.com/>
- Education
 - [US Department of Education Data & Stats](#)
- Social Media
 - [Top Social Media Statistics And Trends Of 2023 - Forbes](#)
- Gaming & Technology
 - <https://explodingtopics.com>

ACTION 2

Conduct a survey to collect data from your classmates on the two variables that you chose.

Students will conduct their survey based on the guidelines that you set for data collection.

The purpose of this action is for students to collect data. Consider confirming student choice of their variables before data is collected. Given that data collected may be more personal for students, consider allowing students to use an anonymous Google Form to collect data. However, you may also decide that the “Student Data Collection Chart” from Part 1 is appropriate.

Formulate

QUESTION 9



Graph the data and describe the relationship. Identify a model of best fit for this data. Would any other model possibly fit? Why or why not?

Student graphs will vary. They should attempt to fit their data to a linear, quadratic, or exponential model. You should see them sketching lines, parabolas, or curves. They may also decide to use technology such as Desmos or graphing calculators to create an equation for the line of best fit if they notice a linear association. Possible student response:

- There seems to be a linear relationship between the data because as variable one increases, variable two increases. I drew a line of best fit, and it would not make sense to use a different model.*

The purpose of this question is for students to use a scatter plot to represent two-variable measurement data. Then they describe the shape of their data and note any visible patterns or relationships between the data such as positive correlation, negative correlation, or no correlation. Next, they use functions (linear, quadratic, or exponential) to determine which one best fits the data.

QUESTION 10



Do you believe this model is a good fit for the data that you have collected? Why or why not? How strong is the correlation for the model that you chose?

Answers will vary. Possible student responses:

- *A linear model works well for my data because I was able to draw a line of best fit and find the equation.*

The purpose of this question is for students to assess the fit of their function using observations, correlation coefficients, residuals, or other tools. At this point, students should clearly see a linear, exponential, or quadratic relationship that can be used to model their data and should clearly specify which model they chose and why.

Problem Formulation Checklist

Requirement	Teacher Notes: For each Requirement, what are some specific things that came up in that part of the task?
A hypothesis concerning the relationship between two variables has been formed.	(Teachers, please add your notes here.)
The data points for two variables have been accurately plotted, and the shape of the graph has been described.	(Teachers, please add your notes here.)
A model has been identified and an equation has been created for relating the two variables of choice.	(Teachers, please add your notes here.)
General Teacher Notes:	(Teachers, please add your notes here.)

In the **Compute** and **Interpret & Analyze** sections, you will use your equation to make predictions about future populations, evaluate your predictions against other factors, and consider the importance of and relationship among these variables in these kinds of analyses. The **Compute, Interpret, & Analyze Checklist** can be used to ensure you have the necessary components before moving on to the remaining parts of the task.

Compute

ACTION 3

Use your model (graph and/or equation) to predict information for future data. Predict information for at least five additional data points using the equation and model you created.

Answers will vary based on student models (graphs and/or equations).

The purpose of this action is for students to use their model to make predictions. Students should predict values that are not included in the class data.

Interpret & Analyze

QUESTION 11

How accurate do you think this prediction will be? Give a rationale for your explanation.

Answers will vary. Possible student responses:

- I believe my prediction is accurate based on the correlation coefficient of 0.89. This r value indicates a strong positive correlation.*
- I believe my prediction is accurate based on the pattern displayed in the scatter plot.*

The purpose of this question is for students to show their understanding of assessing the fit of a model and share their insights about analyzing two-variable measurement data and the limitations of curves fitted to data.

Compute, Interpret, & Analyze Checklist

Requirement	Teacher Notes: For each Requirement, what are some specific things that came up in that part of the task?
The equation has been used to make predictions for future data points.	(Teachers, please add your notes here.)
Thoughtful, valid reasons have been provided for why the model will or will not be good at predicting future data points.	(Teachers, please add your notes here.)
General Teacher Notes:	(Teachers, please add your notes here.)

In the **Validate, Reflect, and Report** sections, you will identify ways to research and improve the accuracy and effectiveness of your equation; you will also create a product that appropriately demonstrates your conclusions. The **Validate, Reflect, & Report Checklist** can be used to ensure you have the necessary components before completing the task.

Validate & Reflect

QUESTION 12



Returning to your hypothesis, do the data you collected and the model (graph and/or equation) you created support your hypothesis? Why or why not?

Answers will vary.

The purpose of this question is for students to return to their original hypothesis and explain whether their model supports or rejects their hypothesis.

QUESTION 13



If you were to do this again, what would you change? Would you choose a different variable? Would you change your data collection method? Why or why not?

Answers will vary because students may decide to change a variable, amount of data collected, the method they used to collect data, or possibly the function they used to fit the data.

The purpose of this question is for students to reflect on this process and share things they would think about changing such as a variable, amount of data collected, the method they used to collect data, or possibly the function they used to fit the data.

Report

ACTION 4

Create a presentation, informational video, poster, or some other method to communicate succinctly what hypothesis you tested and what your findings were.

Presentation formats will vary. A possible format may be a student creating a pamphlet about the harms of excessive coffee consumption. The model created should support the student hypothesis/claim.

The purpose of this action is for students to share their hypothesis and findings in an organized way with their peers. Depending on the amount of time, this can be an extension, or students can share their findings in small groups like they did in Part 1. Consider hosting a gallery walk so that students can read about the various real-life relationships examined by their peers.

QUESTION 14



Based on your model and findings, how can the data you collected and the model you created be used to benefit others?

Answers will vary, but student responses should connect to the study they conducted.

The purpose of this question is for students to reflect on how their findings and model can benefit themselves or others. Given that students are choosing problems and/or questions that are important to them, we want them to see the value of research and statistics for real-world problem solving.

Validate, Reflect, & Report Checklist

Requirement	Teacher Notes: For each Requirement, what are some specific things that came up in that part of the task?
A clear, comprehensive summary connects all the variables studied and provides some possible actions to consider in tackling the chosen challenge.	(Teachers, please add your notes here.)
A valid reason has been provided for needing to share the data, model, and findings.	(Teachers, please add your notes here.)
General Teacher Notes:	(Teachers, please add your notes here.)

References

Exploding Topics. "Discover Exploding Topics." *Exploding Topics*, Exploding Topics,

<https://explodingtopics.com/> . Accessed 26 November 2023.

Frakt, Austin. "Sleep Can Be Good for Your Salary." *New York Times All Access The Upshot*, New York

Times, 4 November 2019, <https://www.nytimes.com/2019/11/04/upshot/sleep-can-be-good-for-your-salary.html>.

Google Classroom. "Correlation and Causation | Lesson." *Google Classroom*, Kahn Academy,

<https://www.khanacademy.org/test-prep/praxis-math/praxis-math-lessons/gtp--praxis-math--lessons--statistics-and-probability/a/gtp--praxis-math--article--correlation-and-causation--lesson#:~:text=Correlation%20means%20there%20is%20a,a%20correlation%20betwee> .

Accessed 26 November 2023.

"Sports, News, Scores, Fantasy Games." *Yahoo Sports*, Yahoo Sports,

https://sports.yahoo.com/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAK1HnC85YDRPYYOwyy5enyfFlzoh-7GZqg_ydv4KmLtmJCJhdt2rY_z0N3OJC-q-etBg-dccf2x_6GRPczlPY3lhFUsC9kzMmQyqBbUdIQlhZVfHj1ipQ09sCNZ8_oeWG2z2thiWeT3BS2Ecl6kTY . Accessed 26 November 2023.

U. S. Department of Education. "Data and Statistics." *Research and Statistics*, U. S. Department of

Education, <https://www2.ed.gov/rschstat/landing.jhtml?src=ft> . Accessed 26 November 2023.

Wong J.D., Belle. "Top Social Media Statistics And Trends Of 2023." *Forbes Advisor Business*, Forbes

Advisor, 18 May 2023, <https://www.forbes.com/advisor/business/social-media-statistics/#:~:text=In%202023%2C%20an%20estimated%204.9.5.85%20billion%20users%20by%202027> .