

Rationale

The purpose of this module is to introduce students to basic concepts within data science while also providing an introductory activity for the instruction of related topics contained in the Missouri Learning Standards. This module will allow students to experience logarithmic scales and why they are beneficial when graphing certain types of data.

This module serves as a starting point for instruction related to the following Missouri Learning Standards:

- A2.SSE.A.4 - Understand why logarithmic scales are used, and use them to solve problems.
- A2.SSE.A.1 - Develop the definition of logarithms based on properties of exponents.
- A2.SSE.A.2 - Use the inverse relationship between exponents and logarithms to solve exponential and logarithmic equations.
- A2.SSE.A.3 - Use properties of logarithms to solve equations or find equivalent expressions.

MoExcel Data Science Standards

- MoExc1: **Identify** issues, problems, questions, or claims that can be addressed using large datasets.
*The expectation is that students be able to **identify** statements, claims, or questions that can be refined into testable hypotheses.*
- MoExc2: **State** data-driven investigative questions.
*The expectation is that students be able to **state** investigative questions based on quantitative data.*
- MoExc3: **Construct** visual representations of real-life data from publicly available datasets and **describe** patterns observed.
*The expectation is that students are familiar with large datasets of publicly available data that allow users simple but rich manipulation of bivariate data and **describe** patterns that result from purposeful manipulation of the information.*
- MoExc4: **Suggest** and **discuss** the possible interactions among data.
The expectation is that students can provide and consider alternative explanations to the relationships (or lack thereof) among data.
- MoExc5: **Identify** and **discuss** potential factors that can influence relationships between the independent and dependent variables.
*The expectation is that students reflect on the complexity of real-life problems and consider it when attempting analyses or problem-solving. This includes identifying and accounting for different forms of control variables (intervening, confounding, or antecedent). Discussion of the differences among control variables is **not** expected.*
- MoExc6: **Interpret** real-life data by using patterns and relationships among data.
The expectation is that students are able to construct stories that provide plausible explanations for relationships that have been identified among data.

Standards for Mathematical Practice

Standard#:	Standard:
MP1	Making sense of problems and persevere in solving them.
MP2	Reason abstractly and quantitatively.
MP3	Construct viable arguments and critique the reasoning of others.
MP4	Model with mathematics.
MP5	Use appropriate tools strategically.
MP6	Attend to precision.
MP7	Look for and make use of structure.
MP8	Look for and express regularity in repeated reasoning.

Prior Knowledge & Possible Misconceptions:

Prior Knowledge:

This module assumes that previous instruction has covered the Algebra I standards, specifically:

- A1.IF.B.4 Interpret the parameters of a linear or exponential function in terms of the context.
- A1.IF.B.1 Using tables, graphs and verbal descriptions, interpret key characteristics of a function that models the relationship between two quantities.
- A1.CED.A.1 Create equations in one variable and use them to model and/or solve problems.
- A1.CED.A.2 Create and graph exponential equations in two variables.

Possible Misconceptions:

1. When examining a logarithmic scale for the first time, students may instinctively call it an “exponential scale.” While not technically wrong, it will be important to note that while the scale values increase exponentially, we use a logarithm to designate the value’s distance from the origin.
2. Students may struggle to grasp when it is appropriate to use a logarithmic scale versus a linear scale. It should be emphasized that logarithmic scales do not universally produce superior graphs, but often perform better than linear scales when there is a significant skew to the data. For example, many small values and a few large values.

Algebra 2 Data Science Math Module

Example: 1 (This can be a teacher's demonstration or computer lab class.)

Question: Does income level affect happiness?

Data: [Gapminder Tools](#)

Discussion Outline:

Begin lesson by asking students the following discussion questions:

1. Do you think income affects happiness?
2. Does a higher income yield more happiness (or less) and why?

Show students the Gapminder graph (and allow some time to look at the graph) and ask:

- What does each bubble represent?
- What are the two variables on the graph? What is the X-axis and what is the Y-axis? What are the scales on the X- and Y-axis?
- Many countries are on the left side of the graph while a few are more in the middle. What does this indicate?
- Can you see a trend? Does the graph do a good job of showing the trend? Does the graph utilize its space effectively?
- Could we change the graph to better display the trend? If so, how?

Note: Students will likely respond with suggestions like “zoom in” or “change the window” (which might be valid ways to address the issue), but the objective will be to change the scale on the income axis from a linear scale (which doesn't work well because of the disparity in income levels) to a logarithmic scale.

On Gapminder this can be achieved by selecting the “Income” drop-down menu on the X-axis and toggling the scale to “log” or by opening this link: [Gapminder Tools](#)

- How does the graph look now? Is the trend any clearer? Does it match your initial assessment of the relationship between income and happiness?
- Describe the scale on the X-axis. What is the pattern in the scale? Why does this scale do a better job of displaying the data than the linear scale?

This opens the door to discussing logarithmic scales and ultimately logarithms. From here you could point out that while the scale is growing exponentially (which the students should ideally be able to identify on their own), the graph now features a smaller, more relevant portion of the X-axis than that of the linear scale and also allows for the lower income countries to be spaced in a manner more conducive to answering the initial question (i.e. do you think income affects happiness?).

Logarithmic scales are often used when plotting data with a large range and/or significant growth/decay. Students will see logarithms and logarithmic scales in situations dealing with pH levels, sound intensity, earthquake intensity, etc.

Recalling their previous knowledge of exponentials, you can challenge students to produce the exponential function/model for the scale on the X-axis ($y=1000 \cdot 2^{(x-3)}$, $x>0$) and emphasize to them that each unit on the X-axis (instead of increasing by a fixed amount) now increases by a factor of 2 (and scaled by 1000). This is how logarithmic scales behave.

Students may inquire as to why it is called a “logarithmic scale” as opposed to an “exponential scale” or you could pose this question to them. This provides a great opportunity to explain that logarithms are exponents as well as a starting point for lessons about the formal definition of logarithms, the inverse relationship between exponential functions and logarithmic functions, etc.

Note: Alternatively, this lesson could take place after the normal logarithm instruction takes place and tie into what they already learned.

Student might ask:

- Why are the bubbles different sizes? Population.
- Why are the bubbles different colors? World regions.

Option for after the discussion:

1. Have students read this article: <https://www.nytimes.com/2020/03/20/health/coronavirus-data-logarithm-chart.html> This allows students to see how a logarithmic scale can be used on the y-axis and reemphasizes why logarithmic scales are important when telling the story about certain data sets.
2. Challenge students to think of or find other types of data where a logarithmic scale may be better suited than a linear scale. Using Google Sheets (or by hand), have students create a graph with a logarithmic scale and a graph with a linear scale to see if the story of the data set is different.
 - ◇ In Google Sheets, a linear scale can be changed to a logarithmic scale by opening the Chart Editor for a graph and then selecting Customize-Vertical Axis-Log Scale.