DATA SCIENCE
TRUMAN STATE UNIVERSITY

5TH GRADE MATH MODULE: OUTLINE

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 create line graphs from data sets and explore trends in data to answer questions.

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

Math:

- 5.DS.A.1 - Create a line graph to represent a data set, and analyze the data to answer and solve problems.
- 5.DS.A.2 - Create a line plot to represent a given or generated data set, and answer questions and solve problems, recognizing the outliers and generating the median.

Science:

- 5.ESS2.C.1 - Describe and graph the amounts and percentages of water and fresh water in various reservoirs to provide evidence about the distribution of water on Earth.

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.

TRUMAN STATE UNIVERSITY

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 5th grade standards, specifically:

5.NF.A.1 Compare and order fractions and/or decimals to the thousandths place using the symbols >, = or <, and justify the solution.

5.GM.C.1 Define a first quadrant Cartesian coordinate system. a. Represent the axes as scaled perpendicular number lines that both intersect at 0, the origin. b. Identify any point on the Cartesian coordinate plane by its ordered pair coordinates. c. Define the first number in an ordered pair as the horizontal distance from the origin. d. Define the second number in an ordered pair as the vertical distance from the origin.

5.GM.C.2 Plot and interpret points in the first quadrant of the Cartesian plane.

Possible Misconceptions:

1. Students may confuse line graphs and line plots, specifically when it comes to what can be determined from them. Be sure to highlight the features and uses of line graphs as they are generated during the following lesson and then contrast them with line plots when that instruction occurs.

5th Grade Data Science Math Module

Example: Lake Mead Water Levels and the Effect of Population

Question: Will Lake Mead Become a "Dead Pool"?

Data: Lake Mead Data using data from Bureau of Reclamation

This can be a teacher demonstration or students can explore the visualization on their own device.

Discussion Outline:

Initial intuition questions before we consider any data:

1. In what ways is water important?
2. Lake Mead and many other reservoirs in the American Southwest are losing water. Why might this be?
    a. This may be an appropriate time to show students this website from NASA which allows them to visualize what is happening to the water levels of Lake Mead.
3. What do you think it means for a reservoir to become a "dead pool"?

Show students the following video about Lake Mead: Lake Mead At Risk Of Becoming A "Dead Pool"

Encourage them to take notes and watch for the significant water levels mentioned in the video.

Show students the following video for more background on the lake and why it is drying up: What Happened to Lake Mead?

After the videos discuss the following questions:

1. Why is Lake Mead important to the residents of the American Southwest?
2. What are some reasons water levels in Lake Mead are dropping?
3. What is the water level at Tier 1? What does this mean for the residents of the Southwest?
4. What is the water level at Tier 2? What does this mean for the residents of the Southwest?
5. What is the water level at Benchmark 3? What does this mean for the residents of the Southwest?
6. What is the water level at dead pool? What would it mean for Lake Mead to become a dead pool?

Now open the Lake Mead Data *in Google Sheets. The 'Water Level' sheet shows the monthly water level (in feet) of Lake Mead from 2000 (when the level started dropping) to present. The measurements are taken at the end of each month.*

Let students view the data as time goes on. Do they notice a trend? (Let students know that it is sometimes difficult to see a trend when viewing raw data.) Has the water level dropped below Tier 1, Tier 2, or Benchmark 3 at any point?

Select the month and water level columns and have Google Sheets create a line chart with the data. Consult this Google Sheets Tutorial *for assistance.*

*Do they notice a trend now? (Here is an updated chart that can also be viewed with measurements from before 2000: **Lake Mead Water Levels*)

7. What level is Lake Mead at currently?
8. Based on the graph, do you think Lake Mead is in danger of becoming a dead pool? If so, when do you think it will occur?

The 'Population' sheet *contains population data for Las Vegas and Los Angeles (two cities that rely heavily on Lake Mead for water and electricity) over the same time period. Create a line graph for one or both cities.*

9. How does the behavior of the population graph(s) relate to the behavior of the water level graph?
10. What are some things that could be done in order to prevent Lake Mead from becoming a dead pool?

Option for after the discussion:

Provide students graph paper and the most recent 6 or more months of water levels. Guide them in creating their own properly scaled/labeled line graph. Have them use their graphs to predict where the water level would be in 6 months and 12 months. Discuss what their predictions mean for Lake Mead and the communities in the Southwest.

Suggestions for Unit Integration

Throughout the unit, refer back to the **Water Level dataset** to reinforce and relate to the concepts that are being taught. For example:

Line Plots:

The data contained on the 'Water Level' sheet can also be used to instruct students on the creation of line plots and their uses. For example, using the most recent months of data, an axis with 1040, 1050, 1060, 1070, 1080 can be created and then students could determine how many months had water levels that fell in the 1040 range (all levels between 1040 and 1049) in the 1050 range, etc. Once created, direct instruction or guiding questions related to finding the mean water level over that time period (or months with outlier water levels) from a line plot can be delivered.