Using Local Data for Personalized Algebra Learning in K-16 Mathematics

Dr. Shande King
NKing2@trevecca.edu
Trevecca Nazarene University

OVERVIEW

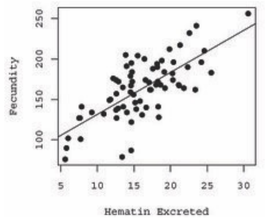
This presentation explains a project that positions students as the researchers to find data that they find interesting and applicable to their daily lives so that they may run basic analyses to research, pose, and answer questions about their lives and society around them. Incorporated in algebra classes both at the college algebra level and intended to equally be used in high school algebra courses, this project details the lesson planning and some student results of projects that help students recognize the types of data and the analyses possible with quantitative data using a least-squares regression line.

In personalizing the project to allow student autonomy in developing their own research questions, this mathematics-focused lesson can be multidisciplinary in nature so that students can see the value in and practical applications of data analyses in other fields, including but not limited to STEM fields such as science and technology. Students also begin to recognize how data can be used to help them pose questions of interest and thus perhaps discover inequities or injustices in the world around them, particularly within local contexts that impact them personally.

LESSON DESIGN

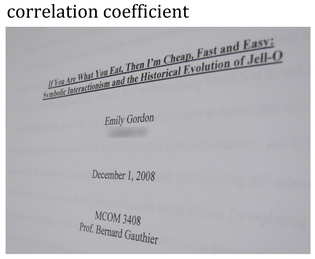
Identify data:

- Students were presented a real-life situation in a scatter plot, recognizing the need for two quantitative variables to run a linear regression model
- Then, they considered two quantitative variables relevant to an issue or research question that interests them
- They also find an article that considers the relationship between their variables



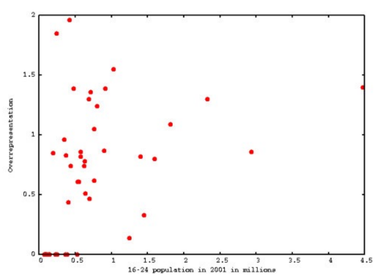
Data:

- The students are then asked to consider the validity of their data.
- They write a description of the raw data and the value of curve fitting a linear regression line to this data. This includes a discussion of the data itself and any shortcomings (i.e. lack of data points available, outliers, etc.)
- Mathematically, students address the strength of the line in discussing the correlation coefficient



Mathematics:

- Students create their own linear regression line
- Then, they write a research question, and answer it using their linear regression line. They must describe the algebraic components of the linear equation (slope, y-ihtercept) in context of the data.
- Students then write a reflection on the value and importance of data and analyzing the data with linear regression



Discussion:

- To conclude, students consider shortcomings of the model
- They discuss how a linear model may not be appropriate or how the data could be better (more data points, more specific data for the research question, etc.)
- Students asked to apply mathematical reasoning to support more advanced ways to use and analyze data (i.e. moving beyond linear regression)

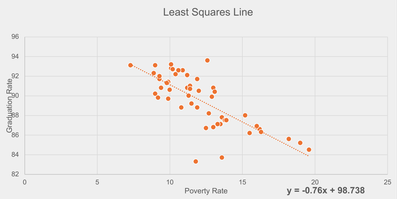
REVIEW OF THE LITERATURE & EXAMPLES

Recent trends in education have encouraged the use of data science and statistical education in mathematics classrooms across all grades, reaching down to elementary and middle school classrooms. Recent STEM education initiatives believe that the teacher holds great power to impact students to show an interest in data science and STEM education to encourage more students to enter the STEM work field (Sami, Sinclair, Stein, & Medsker, 2020). Educators position students in a place to recognize that data science is a "hands-on data analysis" (Donoghue, Voytek, & Ellis, 2021) that allows them to discover new truths about questions that they can answer. Focusing particularly on linear regression, data science allows students to see "beyond the data" (Lübke, Gehrke, Horst, & Szepannek, 2020) so that they can think clearly about the process of data generation. Students begin to consider the potential causal relationship between the two quantitative variables and gain a better grasp of real-world analytical contexts for mathematics with linear regression modeling through this project.

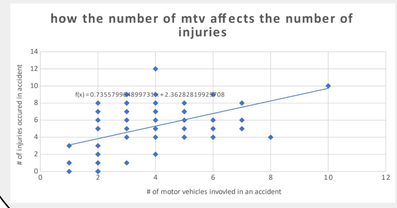
As intended by the research team, students sought local data that was valuable to them. This encouraged students to look both at large data sets open in the public domain via websites such as data.gov, github.com, kaggle.com, census.gov, and data.nashville.gov, a website for local data (in Nashville, Tennessee). The following examples of projects took local data and/or national data that informs students of an issue or research question that interests them.

Example #1:

- **Research Question**: What is the relationship, if any, between a state's percentage of peoples in poverty and that state's high school graduation rate?
- **Result**: In a statistical research report done in 2013 by Procedia: Social and Behavioral Sciences, there is data supporting the findings of this project. The data was also run through a Pearson Correlation analysis, producing a statistically significant negative correlation between the two rates.



Example #2:



- **Research Question**: Does the number of motor vehicles involved in an accident affect the number of injuries occurred in Nashville?
- **Result**: The slope on the graph is small, however the linear trendline shows an increase in some injuries to the rising number of vehicles involved. This is confirmed by the increase in injuries as the number of cars involved in a collision increases by an article by the Association for the Advancement of Automotive Medicine.

STUDENT AND TEACHER REFLECTIONS

"It [was] neat and interesting to see something that applies to my life in the future." -Student

"I found this topic interesting and I was excited to do my homework." -Student

Several students found it challenging to determine which variable is independent and which is dependent, but as they figured it out, it informed their ability to create a research question and thus also to find and use the linear regression line.

"It was rewarding to see students having to struggle to understand how to apply the math beyond a word problem in the textbook. When they created the research question, you could see the having to understand how to define the independent and dependent variables."

"I am excited about not only implementing this into my classroom, but also being able to extend this and the lessons learned from this project to other concepts in my classroom. How can we find other mathematical models in data around us? Is a quadratic regression more appropriate for certain data sets? What makes data reliable and valid? How do we find evidence to support our findings? These were questions that the local data pushed in student understanding of data and algebra!"

DISCUSSION, IMPLICATIONS, AND RECOMMENDATIONS

Discussion:

The overall sentiment of this project was highly favorable with the real-life application that students took away from their finished products. They recognized that they could actually find data that was interesting to them, but then they extended to having to write their own question and then analyze real data that might either support or refute their initial ideas. Teachers were pleased to see student excitement for and application of linear regression in their algebra classes, and they also wanted to be able to extend this type of excitement for application to mathematical concepts beyond a linear equation model.

Implications:

For future teaching:

- Introduce a real-life situation with data that they find collectively as a class to discover how to write a research question and find a linear regression line through an inquiry-based, task-based group-worthy lesson
- Encourage students to find data and an article that supports their data before beginning the entire project
- Have students present the data in some format (asynchronous or synchronous), such as a symposium, so that they may also learn how to present data in addition to analyzing it

For future research:

- Interview students and teachers in various school settings to compare and contrast the impact of data learning and using local data in different types of schools and student populations
- Extend to more advanced mathematics
- Extend to lower-level mathematics courses for the impact of data science understanding in all grade levels

Recommendations:

- Use inquiry-based learning for the sake of student understanding of the mathematics from the beginning of the unit/project. Engage students with a question or topic to consider, then allow student discovery through discussion and collaboration. This helps solidify their knowledge of the variables so that they can produce and create their own project with their own data that they personally use and is applicable to the algebra class.
- Use technology that is accessible to the students that still allows them to find the linear regression line while still understanding how to find the line. Using Desmos was recommended for real-time visual understanding of the linear regression line plotted against a scatter plot. It is also encouraged to understand what the least squares visual looks like, as well as the value of the correlation coefficient.
- Give students a chance to work with each other so that collaboratively they can learn from each other's mistakes about the mathematics as well as see the value in each other's personally-chosen mathematics topics.

REFERENCES

Donoghue, T., Voytek, B., & Ellis, S. (2021). Teaching creative and practical data science at scale. *Journal of Statistics and Data Science Education, 29*(SUP1), S27-S39. Retrieved from https://doi.org/10.1080/10691898.2020.1860725

Lubke, K., Gehrke, M., Horst, J., & Szepannek, G. (2020). Why we should teach causal inference: Examples in linear regression with simulated data. *Journal of Statistics Education, 28*(2), 133-139. Retrieved from https://doi-org.utk.idm.oclc.org/10.1080/10691898.2020.1752859

Sami, J., Sinclair, K., Stein, Z., & Medsker, L. (2020). Data science outreach educational program for high school students focused in agriculture. *Journal of STEM Education: Innovations and Research, 21*(1), 18-24.

*The author would like to thank Julie Steimer and Joanna Presley, algebra teachers at the Christian Academy of Knoxville, for the inspiration to pursue this project.

Presented at ICRSME 2022 Virtual Conference