HOW DO FAMILIES LIVE AROUND THE WORLD? LESSONS LEARNED FROM A STATISTICAL LITERACY INTERVENTION BASED ON THE "DOLLAR STREET PROJECT" IN UNDERGRADUATE-LEVEL STATISTICS AND ECONOMICS COURSES

Emma Bojinova and Anushka Karkelanova
University of Connecticut, USA
University of Kentucky, USA
anushka.karkelanova@uky.edu

This study is based on a statistical literacy intervention in two different undergraduate courses at two US public universities based on the "Dollar Street Project" of the Gapminder Foundation. The goal of this innovative intervention was to expose students to real-world data and teach them how to apply their knowledge of statistics to conduct research, produce a professionally written report, and tell a meaningful story based on data. Students found this project relevant and reported an improved ability to analyze real-world data using statistical methods. Given the importance of understanding statistical concepts and the growing interest in data science, this study can be used as an example of how college courses can provide hands-on activities to increase students' statistical literacy, develop their analytical skills, and create knowledgeable and data-driven members of our society.

INTRODUCTION

The ultimate goal of statistics education is to create a statistically literate society in which people can use, analyze, and interpret data in a meaningful way. Because the study of statistics provides students with tools and ideas that allow them to evaluate quantitative information in the world around them, they should be able to use sound statistical reasoning to intelligently cope with the requirements of citizenship, employment, and family and be prepared for an insightful and productive life. As outlined in Bailey et al. (2013) researchers and statistics educators consistently suggest that students will learn statistics more effectively by conducting projects through which they actively engage in a broad spectrum of tasks integral to statistical inquiry, in the authentic context of a real-world application. The importance of student projects is also emphasized in the 2005 and 2016 Guidelines for Assessment and Instruction in Statistics Education (GAISE) reports (Halvorsen, 2010; Carver et al., 2016). According to the GAISE's recommendations, instruction in statistics education should focus on statistical problem solving, using real data, active learning, and appropriate use of technology. Statistical problem solving is defined as a four-step process: formulating questions that can be addressed through the collection and analysis of data, collecting and organizing appropriate data, analyzing the data, and interpreting results.

This study is based on a pilot statistical literacy intervention in two different undergraduate courses: introductory statistics (259 students) and environmental economics (79 students) taught at the University of Kentucky and at the University of Connecticut respectively. We incorporated the six recommended steps by the GAISE-2016 report into our statistical literacy intervention. Our goal was to encourage students to collect real-world data and teach them how to analyze the data by applying their knowledge of important statistical concepts. Students worked in groups and analyzed datasets collected from the Gapminder's "Dollar Street Project" (https://www.gapminder.org/dollar-street) that shows how families live at different income levels in different countries around the world. The "Dollar Street Project" visualizes the world as a street ordered by income. The poorest live to the left and the richest to the right. The "Dollar Street Project" uses photos for each unit of observation (family) and clicking on a photo provides more information about that family. This makes the dataset more interesting and visually appealing to young adults, thus moving away from a typical dataset with just numbers and descriptions.

Collaborative work provides students with a better chance of understanding the material by learning from one another, as well as leads to improved quality of their submitted assignments. Teamwork tends to increase student confidence in obtaining correct answers, which in turn improves students' study habits (Carvalho, 2008). Collaborative work can also put pressure on students to try to understand and learn the material, so they do not appear unknowledgeable of the course concepts when discussing them with peers. Instructional methods that involve students in projects or activity-based

methods are recognized as a highly valuable pedagogical approach to teaching statistical concepts (Ben-Zvi & Garfield, 2008; Scheaffer et al., 2013).

Undergraduate research and collaborative projects are considered a high-impact educational practice (HIP) defined by the American Association of Colleges and Universities. Kuh (2008) suggests that assessment of student involvement in HIP makes it possible to assess students' cumulative learning. Educational research suggests that HIP increases rates of student retention and student engagement. Findings recommend that active learning and collaborative learning as well as undergraduate research have broad-reaching positive effects across multiple learning outcomes, such as critical thinking, need for cognition, and intercultural effectiveness (Kilgo at al., 2015).

Researchers acknowledge the value of statistics in one's future life or career, but a lack of self-confidence among students studying statistics. Attitudes and beliefs concerning statistics represent a summation of experiences over time in the context of learning statistics. Students' negative attitudes toward statistics are an influential contributor to the low performance of students in statistics courses (Cashin & Elmore, 2005; Mills, 2004). Instructors should make sure that all hands-on activities and literacy interventions improve student learning. As discussed by Professor Cross from the University of California-Berkeley, "From the instruction that we provide, to the intellectual climate that we create, to the policy decisions that we make—all should start with the question, 'But will it improve students' learning?'" (Cross, 2005, p. 2).

In this paper, we provide empirical evidence for the claimed motivational and learning benefits of a learner-centered group project as a method of teaching statistical concepts. We discuss how the group project can be used to introduce most of the important statistical concepts through hands-on activities and strengthen the students' understanding, mastery and appreciation of the material covered in introductory level statistics and environmental economics courses. Our goal is to enhance the abilities of students at all levels to do statistical analysis, think critically, reason, and communicate well their findings.

METHODS

This study is based on implementing a group project in two different undergraduate level courses - introductory statistics and environmental economics. The experiment included students enrolled in STA 210 (Sections 3 and 5) at the University of Kentucky and ARE 1110 at the University of Connecticut in the spring semester of 2021. From now on, the two courses will be referred to as Stats and Econ. Stats is a core course required of all students who do not take calculus. The course was set up as two large lectures, comprising 259 students, meeting on Zoom and completely online. Econ is a general education course taken by students from many different majors and it was taught in a distance-learning format with regular class meetings via Blackboard Collaborate. Students' ages in Stats typically range from 17 to 50, with the majority of the group between 19 and 23 years of age. Students in Econ were between 19 to 22 years old.

For the statistical literacy intervention, a document with step-by-step instructions was created for students to follow in the three classes. All students had the same information and help materials needed for the project titled "What if everyone in the world lived on the same street?". Before beginning the work on the project, students were asked to watch a 15-minute TED talk by Anna Rosling, the creator of the "Dollar Street Project". One of our goals was to give students a chance to work with a global public dataset that is easy to understand and use, but still experience the different stages of working with data including data collection, cleanup, organization, and statistical analysis. The final product from the group project consisted of 3 parts: Data collection, written report, and presentation (e.g., Google Sheets or PowerPoint) posted for all students to see. Both instructors explained the project to their students and provided in-class and posted video instructions of how to collect data and complete the statistical analysis.

We asked students to collect data for two different populations (e.g., families living in Europe, Asia, Africa, or the Americas) from the Gapminder's "Dollar Street Project" (https://www.gapminder.org/dollar-street). Each group analyzed two non-random samples of families within a certain monthly income range. These samples were determined by the instructors to make sure that the different groups would have similar sample sizes. We provided instructions for collecting 12 different preset groups of two samples with income ranges and continents. The two Stats classes had 22 groups each and at most two groups had the same samples. All groups in the Econ course had

different samples to work with (12 groups). Our original plan was to have each sample consisting of 30-40 families. However, after assigning the project the data on the Gapminder website was updated and students ended up working with smaller sample sizes (about 25 families on average).

The instructors allowed students to pick their groups (max 6 people per group) and randomly assigned a group for the students who did not pick one to be a part of. We created step-by-step instructions for students explaining how to collect the data to obtain their preset samples from their populations of interest. The first part of the group project was to create an Excel or a Google Sheets file for all families in their two samples. Students were asked to record family name, continent, country, monthly income (dollar amount), number of adults living in the house, occupation of the head of the household (if available), number of kids, household size (1-,2-,3-bedroom, etc.), availability of electricity (yes, no, yes but interrupted, or n/a), availability of safe drinking water at home (yes, no, or n/a), toilet (indoor or outdoor), and dream of the family (e.g. buy a car, bicycle, etc.). The collected data were organized in a table with 12 columns. We asked each group to submit their collected dataset.

The second part of the group project was to create a written report based on the collected data and student answers to variety of questions including: "How many countries do you have in each of your samples for the two continents?", "How many of these countries are included in the UN's list of the least developed countries in the world?", "What does it mean for a country to be included in this list?", etc. We asked students to use Google Sheets, Excel, StatKey or another data analysis software to obtain descriptive statistics (mean, median, range, minimum, maximum, standard deviation, and coefficient of variation) for the variables income and number of kids per family in each of their samples. We asked them to make boxplots for the variables income and number of kids per family and provide a brief analysis of their data. Students had to discuss their findings related to the computed descriptive statistics (e.g., determine the sample with the higher mean and median income, with the higher variability based on the range and standard deviation, with the higher median and mean number of children). Students also created a bar graph for the number of children in the two samples (i.e., number of families in each sample who have 1 child, 2, 3 4, and 5 or more children) and had to report the maximum and the minimum number of children per family in each of their samples.

Students were also asked to study the correlation between family income and the number of children, to create a scatterplot of income and number of kids for each sample, and finally interpret their result. Calculating the coefficient of correlation was used to support and explain their findings. In addition, Econ students had to provide a more detailed explanation for the reasons of having a negative correlation between income and the number of kids per family based on the covered topics in the course related to fertility rates.

To help students apply inferential statistics, we asked them to find the proportion of families in each sample with available electricity in their home, the proportion of families in each sample with available drinking water, and the proportion of families with an indoor toilet. In addition, students had to briefly describe the typical living conditions of these families within the specified income range in each sample and look for similarities and differences, as well as discuss the typical dreams of these families (e.g., buy a car, a bigger house, etc.) and explain if there was something in the lives of these families that surprised them. Moreover, Stats students provided a description of the population, parameter, sample and statistics based on each of their populations and calculated and interpreted confidence intervals for the true proportion of individuals in each population of interest with more than 3 kids and confidence intervals for the average household size in each population of interest.

All groups submitted a professionally written report and a presentation on the course learning management systems for other students to see. Students were required to work as a team, split tasks among their group, and tell a meaningful and coherent story about the research they had done. The focus throughout this project was to teach students how to ask and answer scientific questions using statistical analysis. Data from our classes was used to evaluate student knowledge of statistical concepts and student perceptions about this project.

RESULTS

There were 259 students enrolled in the two sections of the statistics course and 244 students completed the survey (94.2% response rate). The environmental economics course had 79 students, but not all students worked on the Dollar Street project because it was given as an extra credit. Twenty-two students submitted the group project and 16 completed the survey (72.7% response rate). The table

below summarizes student responses to several questions from the survey administered upon completion of the project.

Table 1. Summary of Survey Responses

Survey Question	Strongly Agree and Agree		Neither Agree nor Disagree		Disagree or Strongly Disagree	
	Stats	Econ	Stats	Econ	Stats	Econ
I found the topic of the "Dollar Street Project" interesting and relevant.	79.1%	93.8%	16.8%	6.3%	4.1%	0.0%
The "Dollar Street Project" improved my ability to analyze real-world data using statistical methods.	87.7%	81.3%	9.8%	18.8%	2.5%	0.0%
The "Dollar Street Project" developed by Gapminder is one way to increase social awareness of the income inequality present across and within nations.	91.4%	100.0%	7.8%	0.0%	0.8%	0.0%
The "Dollar Street Project" helped me learn about the living conditions in other countries.	96.7%	93.8%	2.5%	6.3%	0.8%	0.0%

The majority of students found the group project interesting and relevant (greater than 79.1%). They thought that this was not a typical "boring" project in statistics education. Below are some comments of our students from the survey.

- "I liked getting into a topic that I was not very familiar with and one that I would not typically associate with statistics. I think I learned a lot more since it was a real-world application of stats and not a theoretical example." (Stats)
- "The website was easy to navigate, and the goal of the assignment was actually impactful and not just a random project about marbles or something along those lines." (Stats)
- "I liked how the data involved real people and places and showed us candid pictures of how people live." (Econ)
- "I enjoyed that it was very different from anything I have ever done in school yet." (Stats)

Our students also thought that this project was an eye opener and helped them learn about the living conditions in other countries. It broadened their horizons and increased their social awareness about income inequality around the world and within countries.

- "I liked that we were able to use information from the real world and it also made me really aware of the world around me and how other people live. It made me really realize the severity of poverty in some counties and made me realize that we need to help others as much as we can." (Stats)
- "I enjoyed how I got to see perspectives of real family members across different parts of the world, also taught me to be more thankful for what I have. I wanted to thank you for coming up with a project that was learning, beneficial and fun!" (Stats)
- "I liked learning about living conditions in countries I have not seen before. It truly opened my eyes to how fortunate I am to be able to have things I take for granted, like running water, a toilet, and electricity." (Stats)
- "I really enjoyed this project because it was very humbling to see how some people live across the world. It deeply saddened me that some of these families, especially those with young children, have to live in such poor conditions where they do not even have access to clean drinking water in

their home. It made me realize how lucky I am. It also helped me to get a better perspective on the income inequalities throughout the world and how this is such a large issue." (Econ)

- "I liked how much new information I was able to learn from this project." (Econ)

The majority of students in the three courses found the group project very beneficial in terms of learning how to apply various statistical concepts to real-world data. They also shared with us that this statistical literacy intervention improved their ability to analyze real-world data using statistical methods.

- "We were able to use concepts that we have learned in class and apply them to a real-world situation." (Stats)
- "I learned how to use basic statistics to draw conclusions using the data I collected. I learned that people are more similar across cultures in how they live. People in the same income bracket are going to live extremely similar lives even if they live in different areas or regions." (Econ)
- "I liked looking at the statistical analysis of the people in different areas to learn about their income and how that affects the way they live." (Econ)
- "I liked that we had to come up with our own data to draw conclusions. Although collecting data was a bit tedious, I was able to understand what a research project and data analysis includes." (Econ)

Moreover, the group project allowed students to collaborate with their peers. Our courses were not in person due to the COVID-19 pandemic and the project provided them with an enjoyable learning experience and a real connection to their classmates.

- "I really liked that this project had to do with something that was beyond just statistics. I also loved feeling like I was almost in class again since covid has started because we were able to be in groups and work together for this project!" (Stats)
- "I loved the opportunity to work with other students in this class to analyze and collect data for our project. I enjoyed how we were able to use creativity in this project by putting our findings into a PowerPoint presentation." (Stats)
- "I enjoyed getting to analyze different aspects of how income can affect people in different parts of the world. I also enjoyed how we could combine geography, humanities, and economics with statistics." (Stats)

The six groups that completed the group project in Econ received a grade above 90%. They all put significant effort into finishing the different stages of the extra credit assignment. Students showed a good understanding of various statistical concepts, excellent statistical reasoning and writing skills, and a high level of critical thinking. All 44 groups in Stats class received a grade of 80% or above, with overall average score or 92 and median 90 (based on 259 students who completed the project). The instructor used a grading rubric that was posted and available to students from the beginning of the project. Students were given 4 weeks to complete the group project. Overall, all students worked hard, communicated well, and produced a high-level written report with all questions answered and very interesting visual presentation to summarize their findings and conclusions.

CONCLUSION

This statistical literacy project focused on the entirety of the statistical process, from data collection to drawing the appropriate conclusions from the statistical analysis and to the process of preparing a professionally written report and a presentation. Through this process, students were exposed to several effective learning techniques such as using real and interesting dataset, small group discussions, collaboration with peers, and active student-oriented learning. It emphasized major conceptual understanding of course content and helped students improve their communication and writing skills, as well as learn to work as part of a team. The main goals of the project were not only for students to become more effective users and communicators of statistics, but also for them to see how statistics can be used and applied effectively to inform on important matters globally. Our motivation for incorporating this statistical literacy intervention into our courses is the firm conviction that students learn statistics by doing statistics and that doing statistics means organizing data, using appropriate statistical methods for analyzing the data, and presenting results using proper statistical language that is easy to understand.

As discussed in the results section, students found this project impactful, interesting, and engaging. We note that the results of this study may not be generalized to the population of all

undergraduate students or to other content domains. Future research will focus on analyzing whether this statistical literacy intervention has improved student learning and understanding of statistical concepts. We were not able to use an experimental design with this pilot study, but we plan to incorporate this in the future with an experimental and a control group. Finally, it is worth mentioning the lack of pretest data as a limitation of the present study. A pretest-posttest research design is highly desirable to further investigate the effects of this group project on student learning.

The foundational position of the study is that when students can visualize a statistical concept in action, a deeper level of understanding occurs. Our pilot study results provide evidence for the benefits of this statistical literacy intervention. With the "Dollar Street Project", abstraction becomes a reality. Students are able to place global poverty and inequality in the context of material culture. Instead of charts to illustrate global poverty, the Dollar Street idea brings humanity, color, and life, providing students with a stronger understanding of what income inequality means around the world.

REFERENCES

Aliaga, M., Cobb, G., Cuff, C., Garfield, J., Gould, R., Lock, R., ... & Witmer, J. (2005). Guidelines for assessment and instruction in statistics education: College report. *Alexandria, VA: American Statistical Association*.

Bailey, B., Spence, D. J., & Sinn, R. (2013). Implementation of discovery projects in statistics. *Journal of Statistics Education*, *21*(3).

Ben-Zvi, D., & Garfield, J. (2008). Introducing the emerging discipline of statistics education. *School Science and Mathematics*, *108*(8), 355-361.

Carvalho, C. (2008). Collaborative Work in Statistics Classes: Why do it? In *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI/IASE Study 18 and 2008 IASE Round Table Conference.*

Carver, R., Everson, M., Gabrosek, J., Horton, N., Lock, R., Mocko, M., Rossman, A., Roswell, G. H., Velleman, P., Witmer, J., & Wood, B. (2016). Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report 2016. Retrieved from https://commons.erau.edu/publication/1083

Cashin, S. E., & Elmore, P. B. (2005). The Survey of Attitudes Toward Statistics scale: A construct validity study. *Educational and Psychological Measurement*, *65*(3), 509-524.

Cross, K. P. (2005). What Do We Know about Students' Learning and How Do We Know It? Research & Occasional Paper Series: CSHE. 7.05. *Center for Studies in Higher Education*.

Gapminder. (2021, May 29). Retrieved May 29, 2021, from www.gapminder.org/dollar-street

Halvorsen, K. T. (2010, July). Formulating statistical questions and implementing statistics projects in an introductory applied statistics course. In *Proceedings of the 8th International Conference on Teaching Statistics* (Vol. 8).

Kilgo, C. A., Sheets, J. K. E., & Pascarella, E. T. (2015). The link between high-impact practices and student learning: Some longitudinal evidence. *Higher Education*, *69*(4), 509-525.

Kuh, G. D. (2008). Excerpt from high-impact educational practices: What they are, who has access to them, and why they matter. *Association of American Colleges and Universities*, *14*(3), 28-29.

Mills, J. D. (2004). Students' attitudes toward statistics: Implications for the future. *College Student Journal*, *38*(3), 349-362.

Scheaffer, R. L., Witmer, J., Watkins, A., & Gnanadesikan, M. (2013). *Activity-based statistics: student guide*. Springer Science & Business Media.