How Access to STEM Can Save the World

Jennifer Sabourin, and Lucy Kosturko, SAS Institute Inc.

ABSTRACT

For every challenge we face on the planet, there is data that can help us find the solution. The next generation of problem solvers, "the digital natives", are poised to see issues such as poverty, hunger, gender equality, and climate change not as insurmountable but as solvable puzzles. With unprecedented access to technology and data, and an unquenchable thirst for digital connection, this generation holds incredible promise in solving social challenges that affect every population, every race, every gender. An interdisciplinary approach to education that incorporates science, technology, engineering, and math (STEM) and project-based learning is the best way to prepare these students for the world that awaits their contributions. GatherIQ™, a mobile and web app produced by SAS® in conjunction with their education software division, Curriculum Pathways, invites students to join the global quest to reach the United Nation's Sustainable Development Goals for 2030 through experiences in STEM and project-based learning. Using GatherIQ, students not only learn about the issues but combine their wits to address social challenges in their own back yards and around the world.

INTRODUCTION

Whether we are cognizant of it or not, everyone has experience turning data into actionable information. From life's small choices, like choosing where to pump gas or what to feed our families, to decisions that have impact on a global scale, data is all around us. Educators argue "every high-school graduate should be able to use sound statistical reasoning to intelligently cope with the requirements of citizenship, employment, and family and to be prepared for a healthy, happy, and productive life" (Franklin et al, 2007, p. 1).

Beyond data-driven decisions in our daily lives, the proliferation of data around the world necessitates an army of data scientists—individuals who can fluently collect, analyze, and disseminate innovation uncovered through the mass of data we generate every day. The answers to the world's pressing questions are waiting to be discovered by the next generation of data scientists.

Therefore, we look to institutions of formal education to prepare students for the data-rich world in which they live. Newly drafted education standards across the country promote the importance of statistics and data literacy through disciplines such as science, technology, engineering, and math (STEM). However, to truly develop the skills necessary for data literacy, students need opportunities to consider their data analysis *in context*.

Project-based learning, in which students apply core content to solve real-world challenges, creates an ideal venue for developing data literacy. Tools such as GatherIQ™ complement this pedagogical strategy by providing contextualization using the United Nation's Sustainable Development Goals (also known as SDGs, Global Goals) as a framework. GatherIQ's social networking and communications capabilities serve as a platform to connect localized student work to peers around the world banding together to solve our society's biggest challenges.

STEM AND DATA LITERACY

K-12 courses and projects in the areas of science, technology, engineering, and math (STEM) provide students with opportunities, methods, and strategies for higher order thinking skills such as critical analysis, creativity, collaboration, and communication. Within these contexts, students are encouraged to follow specific problem-solving processes, which serve as essential models of behaviors used by professionals around the world. Masquerading by names such as the Scientific Method, the Engineering Process, statistical problem solving (GAISE, 2016), Project-Based Inquiry Process, and many others, strategies for STEM problem solving focus on four primary areas. Students are taught how to ask high-quality questions, collect their own data, analyze their own findings, and disseminate results—all of which transcend across domains therefore preparing students to strategically solve their own problems.

At the heart of every STEM problem is data. While it can come in many forms, ranging from small choices to global decisions, data is how the world communicates innovation. Given the exponential proliferation of data, the importance of developing statistically literate citizens is paramount. In fact, 2019 marks the fourth year in which data scientists prevailed as the most in-demand professionals in the country according to Glassdoor (CBS News, 2019). Despite our personal careers, we citizens of the data-rich world use such information every day to make decisions about our personal lives making data literacy relevant to all students, not just those pursuing studies in STEM.

DEVELOPING DATA LITERACY

The Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K-12 Curriculum Framework (Franklin et al., 2007) suggests "instructional programs from pre-kindergarten through grade 12 should enable all students to formulate questions…select and use appropriate statistical methods to analyze data, develop, and evaluate inferences and predictions that are based on data, and understand an apply basic concepts of probability." (p. 5) To achieve this, the organization (GAISE, 2016) offers six recommendations for statistics education:

1. Teach statistical thinking. Statistics is not just a collection of tools and equations used to generate numbers and graphs. Instead, it is a multi-step problem-solving process where students must identify problems, formulate questions, and critically evaluate data to see if it addresses the initial goals. Students should be given the opportunity to take part in each of these steps. This often involves presenting students with open-ended problems and supporting students as they identify the questions, variables, data, and methods that are most appropriate.

2. Focus on conceptual understanding. Rather than focusing on the procedural steps and calculations needed to analyze data, instructors should focus on the underlying concepts that drive those techniques. Students should be able to articulate what they are doing and why, not just how.

3. Integrate real data with a context and a purpose. The context associated with a data set is just as important as the data itself. It is important for identifying bias and recognizing which questions can and cannot be answered by the data provided. Students should have opportunities to collect data on their own, or, if presented with complete data sets, understand how and why that data was collected.

4. Foster active learning. Students should play an active role in the learning process and should not just be receiving information in the form of a lecture. Opportunities to talk with other students and discuss their current understanding or questions about the material are important activities for fostering engagement and

synthesizing knowledge. Students should practice asking questions and making decisions throughout the statistical process and should not be simply following explicit sets of instructions.

5. Use technology to explore concepts and analyze data. Advances in technology have changed the speed and strategies used to analyze data and this should be reflected in student course work. Along with recommendation 2, the focus is not on being able to compute statistical measures but in understanding when and why they should be used. As such, technology is a great resource for accomplishing the task of the actual calculations as well as visually exploring and interpreting results.

6. Use assessments to improve and evaluate student learning. Assessments of student learning should be based on the recommendations above. They should evaluate students' understanding of the underlying concepts and motivations as well as how they move through the steps of the statistical process. Assessments should also be timely, so that students and instructors can recognize gaps in understanding and address them.

A common theme among the committee's recommendation is context. The report argues, "data are not just numbers, they are numbers with a context" (Franklin et al., 2007, p. 7) and it is this distinction that separates statistics—a discipline that revolves around storytelling and finding meaning—from mathematics. A good statistician not only needs to know the math, but also be an expert in the context, making statistics a truly interdisciplinary field. Therefore, developing data literacy requires practicing *in context*, and the strategy of project-based learning lends the perfect pedagogical lens.

PROJECT-BASED LEARNING

Project-based learning (PBL) has its roots in problem-based learning, building on a strong orientation on real-world problems. PBL has a reputation as a highly effective instructional strategy juxtaposing knowledge creation with real-world problem solving, critical thinking, creativity, collaboration, and communication skills. While complex in development and integration, the robustness of PBL experiences often involve skills and objectives that cut across the content areas and satisfy a wealth of curriculum standards.

In addition, scaffolded PBL experiences at school provide students with the opportunity to practice the transfer classroom content to real-world situations—a fundamental priority of education. Furthermore, such authentic learning activities have additional motivational benefits including student ownership of the learning experience (Savery & Duffy, 2001), student autonomy to highlight their own academic strengths (Holm, 2011), and direct and concrete contextualization of purpose (Savery, 2006).

With respect to the latter, studies have shown the level of contextualization plays an important role in student ownership and motivation. In a recent study by Kosturko and colleagues, student performance and on-task activity were found to be significantly higher when engaged in a project for a local, familiar context, as opposed to generic or hypothetical scenarios (2017).

However, the benefits of PBL should not overshadow the complexities and nuances associated with classroom integration. Designing and deploying PBL activities demands an instructor's 1) value for the constructivist nature of PBL, 2) specialized understanding of the student population, 3) careful consideration for provided resources and curricular aids, and 4) working, flexible knowledge of scaffolding methods (Savery, et. al., 2006; Holm, 2011). Because of these factors, preparing for PBL lessons takes a considerable amount of time and effort, which is inconsistent with today's over-scheduled school days meaning instructors often shy away from this valuable instructional method.

Better support tools and scaffolds are required to alleviate the demands of effective PBL. Such tools should scaffold the entire PBL process from asking relevant, measurable, and contextualized questions to publicly disseminating findings. Such is the goal of GatherIQ.

GATHERIQ

The objective of GatherIQ is to enable the next generation of problem solvers with the tools they need to make informed decisions and take action in order to help reach the Global Goals by 2030. GatherIQ's unique contribution is to provide users with an opportunity to better understand the Goals through the lens of data and then take personal action motivated by story and in partnership with community.

As a resource for students and global citizens as a whole, GatherIQ (Figure 1) introduces users to each of the 17 goals for a better world. First, by educating them on the individual goals through story and data analysis. Second, by providing interactive quizzes and challenges that encourage a change in behavior in support of each goal. Third, by providing a starting point for project-based learning and group challenges that can multiply the effects of individual actions as a community.

Available in April 2019 on iOS and web, GatherIQ v3 is a free app supported by SAS® and Curriculum Pathways®, SAS' free educational software supporting over 4.5 million teachers and students.
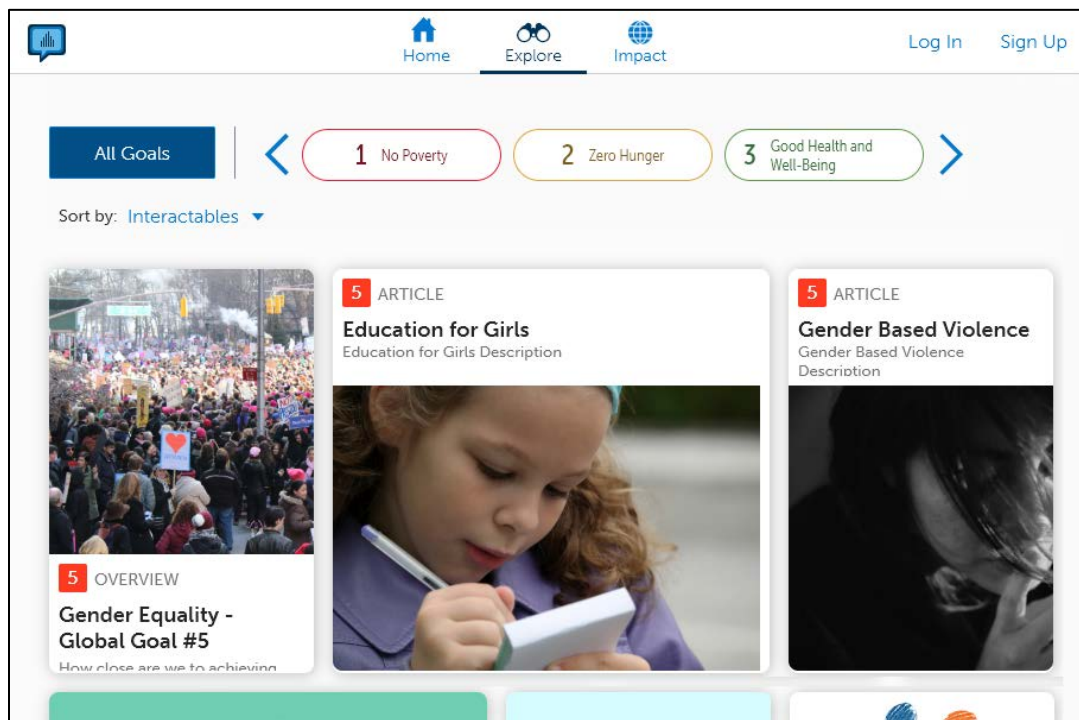

Figure 1. GatherIQ

SAMPLE ACTIVITY

To better illustrate how GatherIQ and PBL can support data literacy and STEM education, consider the following scenario. As presented in Table 1, the scenario uses the recommendations from the GAISE report (2016) and the statistical analysis process (Franklin et al., 2007) as the underlying framework.

Statistical Analysis Process	Action	Recommendation					
		1	2	3	4	5	6
Formulate questions	Explore data		●			●	
	Read narratives	●		●		●	
	Localize problem	●		●	●		
	Research problem			●			
	Identify project			●			
Collect data	Design a plan to collect data	●		●	●		
	Use the plan to collect data	●		●	●		
Analyze data	Organize results	●		●	●	●	
	Analyze findings	●	●	●	●	●	
	Visualize data	●	●	●	●	●	
Interpret results	Present project to authentic audience	●		●	●		●
	Share experience with GIQ community				●		●

Table 1. Aligning the Illustrative Scenario to the Statistical Analysis Process and the Recommendations for Statistics Education.

Mr. Samson is a 7th grade teacher and is interested in doing a PBL activity surrounding water conservation with a focus on data literacy. As part of the project-based learning paradigm, he first must introduce students to an authentic problem. Using the Global Goals for Sustainable Development as a backdrop, he has students explore issues related Goal 6: Clean Water and Sanitation. Students learn more about this issue by visiting GatherIQ and drilling down to focus on content related to water. As they get a better understanding of the underlying concerns, students are encouraged to share their knowledge through social media and to test their understanding using built-in quizzes available as part of GatherIQ. At the end of exploring GatherIQ, students should be familiar with the issues surrounding clean water and have a personal investment in doing their part to help solve the global goal.

Next, Mr. Samson will guide the students through the statistical analysis process. He will first encourage students to use what they learned in GatherIQ as well as prior knowledge and experience to identify possible variables that impact water availability and cleanliness in their own community. He points his students to relevant lessons in Curriculum Pathways for standards-aligned material on topics such as Earth's Water, Water Budgets, and Stream Ecology (Curriculum Pathways, 2019). He will encourage them to formulate questions that can be answered by collecting data. For example, students might ask, "what is the most effective way for members of our community to reduce their water consumption?" Mr. Samson will ask students what variables affect water consumption in their community. Students might suggest that the time of year, lawn watering, frequency and duration of showers, and whether water is turned off while brushing teeth are factors that affect water consumption.

Given these class-generated ideas, students will then design a research and data collection procedure that will help them to identify what behaviors individuals can change in order to reduce water conception. Students should recognize that time of year, for example, is not a variable they can control and therefore should not be included in their investigation. They

also decide that not all classmates live in homes with lawns, and this variable is difficult to control. They then formalize how shower duration and water used while brushing teeth can be minimized in their instructions to participants. They devise a method for determining impacts on water use – specifically asking participants to bring in water bills from before and after implementing changes. All these activities involve students being active participants in the statistical process. They are formulating their own questions and developing appropriate ways to answer them. They are guided by a knowledgeable instructor but otherwise take ownership of the project and the problem-solving process.

Once students have designed their study, they must implement it. In this example, students are randomly assigned to one of 4 conditions 1) reduce shower time for every household member by half, 2) do not leave water on while brushing teeth, 3) both 1 and 2 and 4) no change in activities. Students also choose to collect data on how many people live in each household along with their age and genders.

At the end of the month, students gather all their data and synthesize it into a single digital file. Mr. Samson guides the students in identifying the questions they can now ask with their data and what statistical process they should use to find the answers. Students want to know:

- Which recommendation was most effective on its own?
- Did the recommendations conserve water compared to no change in behavior?
- Is following both recommendations better than following just one?
- Does the efficacy of the recommendation change based on the demographics of the household?

With a plan in place, the students carry out the analyses using digital statistics tools. They then discuss and interpret their results. Overall, students find that reducing shower time was significantly more effective at conserving water. They found that there was reduced water intake in even the control condition, which they attribute to those students still being cognizant of water waste. There was marginal improvement when participants used both recommendations and behavioral changes were particularly effective in larger households with older family members.

The final step in the project is for students to disseminate their findings. They use their data and interpretations to design a pamphlet (Figure 2) encouraging families in their community to reduce their shower times. They pull motivating data and narratives from their initial GatherIQ research and use their own analyses to back up their recommendations. They print and disseminate their brochure locally to the students in their school. Mr. Samson also uses social media to share the PBL activity and his students' work with the team that curates GatherIQ content. Since part of GatherIQ's mission is to share activities such as this one, it becomes a new tile as part of GatherIQ, encouraging other educators to use Mr. Samson's lesson as an example for developing their own data-centric PBL activities.
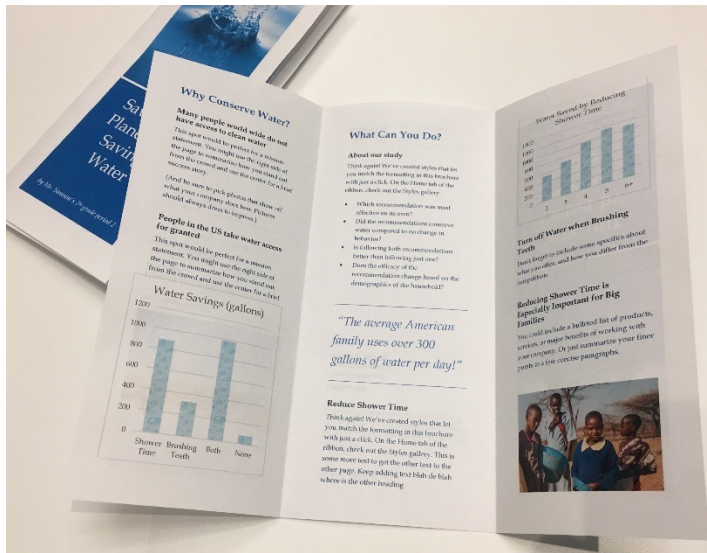
Figure 2. Sample Brochure to Disseminate Findings

CONCLUSION

Project-based learning is an effective and compelling strategy for providing context around statistics to emphasize important concepts to students. However, PBL is time-consuming to develop and implement. Consequently, resources such as GatherIQ, which scaffold steps of the problem-solving process and provide teachers with inspirational activities, are of incredible value in promoting the use of PBL in K-12 classrooms. These activities focus on real-world problem-solving using the data literacy skills students will require as 21st century learners, employees, and citizens and prepare today's students to have impact on the global stage.

REFERENCES

CBS News. 2019. The best jobs in America for 2019. https://www.cbsnews.com/news/the-50-best-jobs-in-america-for-2019-according-to-glassdoor/

Curriculum Pathways. 2019. Global Goals Strategy Page. https://www.curriculumpathways.com/portal/#/pd/strategies/22

Curriculum Pathways. 2019.

GAISE College Report ASA Revision Committee. 2016. Guidelines for Assessment and Instruction in Statistics Education College Report 2016, http://www.amstat.org/education/gaise.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. 2007. Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K-12 Curriculum Framework. Alexandria, VA: American Statistical Association.

Holm, M. 2011. PROJECT-BASED INSTRUCTION: A Review of the Literature on Effectiveness in Prekindergarten. River Academic Journal, 7(2), 1-13.

Kosturko, L., Hoffmann, K., Sabourin, J., Bissette, R., Crossland, C., Nietfeld, J. 2017. Authentic Purpose and Audience: Utilizing Local Contexts to Enhance Project-Based Learning. In Proceedings of the 2017 Annual Meeting of the American Educational Research Association.

Savery, J. R. 2006. Overview of problem-based learning: Definitions and distinctions. Interdisciplinary Journal of Problem-Based Learning, 1(1), 9-20.

Savery, J.R., & Duffy, T.M. 1995. Problem-based learning: An instructional model and its constructivist framework. In B. Wilson (Ed.), Constructivist learning environments: Case studies in instructional design (pp. 135-148). Englewood Cliffs, NJ: Educational Technology Publications.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Jennifer Sabourin
Jennifer.Sabourin@sas.com

Lucy Kosturko
Lucy.Kosturko@sas.com