Data Scientists in the Classroom

Author's Name: Kieu Pham **Coach Name:** Debra Dimas
Host Organization: Lockheed Martin **ETP Type:** Classroom
Subject/Grade: Biology/9th

Abstract (~150 words)

Many science teachers have had the experience of students attempting to generate the "correct" data that they think the teacher wants. This unfortunately leads to a common practice of students copying data from others or even fabricating it themselves. The goal of this ETP is to push students to accept their data as is and shift their focus from generating "correct" data to deeply analyzing the raw data and understanding its implications. The skills can then be applied to any lab report where students need to analyze their data/determine trends.

"Analyzing and Interpreting Data" is also one of the eight Science and Engineering Practices outlined by the Next Generation Science Standards. The ability to turn data sets into graphs as well as identify significant patterns/trends is an essential skill for success in any science course. The purpose of this ETP is to teach students digital data analysis skills as well as create a final slideshow where they share their graphs and major trends they identified. The ETP will be designed to be distance-learning friendly, so will rely heavily on Google Suite but can easily be implemented in a traditional classroom with access to devices for each student.

Focal Content & Supporting Practices

HS-LS4-3: Apply concepts of statistics and probability to support explanations that organisms with an advantageous heritable trait tend to increase in proportion to organisms lacking this trait. [Clarification Statement: Emphasis is on analyzing shifts in numerical distribution of traits and using these shifts as evidence to support explanations.] [*Assessment Boundary: Assessment is limited to basic statistical and graphical analysis. Assessment does not include allele frequency calculations.*]

Science and Engineering Practices:

<u>Analyzing and Interpreting Data</u>
Analyzing data in 9–12 builds on K–8 experiences and progresses to introducing more detailed statistical analysis, the comparison of data sets for consistency, and the use of models to generate and analyze data.

- Apply concepts of statistics and probability (including determining function fits to data, slope, intercept, and correlation coefficient for linear fits) to scientific and engineering questions and problems, using digital tools when feasible. (HS-LS4-3)

21st Century Skills and Applications (1 - 2)

Problem Solving
- Accurately identifies, breaks down, and analyzes problem
- Produces thorough analysis of effects and tradeoffs for various alternatives

Communication and Collaboration:
- Written and oral communications are clear, correct, and concise
- Employs media and technologies that align with purpose and enhance communication

This ETP develops students' ability to problem solve by giving them another tool to be able to visualize/analyze data as well as identify relevant trends. Communication/collaboration skills are also developed, as the final assessment is a slideshow where students will determine how best to display the information as well as communicate their findings to the class.

Measurable Objective(s)

1. Students will collect data on trait frequency in different environments and apply data analytics on Google sheets in order to determine the distribution of traits in a population and identify important trends of trait frequency depending on a specific environment.
2. Students will create a presentation on Google slides in order to display appropriate graphical representations of their data and communicate important trends of trait frequency depending on a specific environment as well as infer to why these frequencies may change.

Formative Assessment(s)

Students will need several skills in order to successfully complete the summative assessment. The following formative assessments will measure if students have mastered a skill after being taught the lesson. Google forms (Quiz mode) and Google docs (an assignment posted via Google classroom) will be used for assessments.

1. Selecting the correct graph (pie, bar, or line) based on a given data set.
2. Applying formulas on Google sheets to determine the sums and percentages of a data set. Answer Key
3. Knowing how to create various charts of data sets using Google sheets. Answer Key

Summative Assessment(s)

Students will combine all of the skills learned in this data analysis unit to collect data on trait frequency in different environments, apply data analytics, and create a presentation on Google slides to communicate important trends of trait frequency depending on the environment.

1. Link to Assessment/Rubric (posted on Google classroom as an assignment)
   Exemplar Data Sheet, Exemplar Presentation

Fellowship Description (300-500 words)

Lockheed Martin is an American company that employs 110,000 people worldwide. The talented engineers and scientists who work at Lockheed Martin engineer and manufacture numerous products in various fields such as aerospace, defense, arms, security, and other advanced technologies. Lockheed Martin is the world's largest defense contractor with the majority of its revenue coming from military sales.

My position at Lockheed Martin for this summer fellowship is Production defect root cause analysis (RCA) expert. I am working with the Quality Assurance team to produce metrics for the THAAD (Terminal High Altitude Area Defense) program in order to better understand productivity and causes for design changes. Quality Engineers at Lockheed Martin apply data analysis to see patterns and understand critical phenomena buried deep in large data sets using graphical analysis. These insights allow program or business leaders to see and understand variables of interest in order to make investment decisions for research funds, to manage hidden risks, and to maximize product quality based on tradeoffs for various alternatives.

To further elaborate, THAAD is an anti-ballistic missile defense system produced by Lockheed Martin. When THAAD is being produced, errors may sometimes occur in which changes must be made (Rapid Engineering Changes). This may happen because of a variety of reasons, such as design error/analysis error (technical errors such as incorrect dimensions), customer-directed changes (modifications requested by the customer), internal production improvement changes (to reduce cost or schedule), and more. The engineers then need to report the reason why they must make changes.

The Quality Assurance team I am working with then analyzes the data using Excel and Tableau to generate metrics or charts to better understand the data. They can then present this data to departments or program managers to identify the root cause of changes that are being made during the production process. This helps departments become more efficient as well as identify areas of strength and improvement. For instance, if a large portion of the changes are due to design errors, this indicates that designs require more review before being submitted for production.

The skills used in this summer fellowship include analyzing large data sets using Excel (specifically Pivot Tables), producing clear presentations using PowerPoint based on the data, and presenting the implications of the analysis to a team. These skills are used in many different careers, including but not limited to: quality assurance engineers, data scientists, human resources representatives, financial analysts, sales managers, accountants, and administrative assistants. These skills are valuable in a plethora of fields and are valuable to teach to our future workforce.

Fellowship Connection to School/Classroom (300-500 words)

My ETP is designed to specifically emphasize the skills used during my summer fellowship. This includes data analysis, producing metrics and charts, and creating clear presentations that discuss the implications of those metrics that can be shared with a team. The skills used by the Quality Assurance engineers at Lockheed Martin are the exact same skills my students will refine and use to do their summative assessment on leafhopper population trends. The ETP is designed to be distance-learning friendly, but can easily be taught in a traditional classroom as long as each student has access to a device. The data collection is also based on a simulation slideshow, but can easily be revised to be placed at the end of any other type of physical "wet lab" that requires extensive data (preferably quantitative) collection. The purpose is to be able to produce metrics at the end of any type of data collection as well as identify significant trends and their implications.

The ETP transfers those skills to Biology content (population changes in a Leafhopper population). Google Sheets/Slides is also used instead of Excel/PowerPoint, as the two produce similar results and my school provides Chromebooks for each student with access to Google suite. Just like Quality assurance engineers presented their metrics and emphasized important implications of the analysis, my students will practice these same skills by producing charts of population growth and writing about the implications of that population growth.

Before starting this unit, I will share with my students my personal experience over the summer and how I used these same skills in order to do my work. I will also emphasize that data analysis/presentation skills are used daily amongst the Quality Assurance team at Lockheed Martin. These skills are also present and pivotal to numerous other jobs, such as financial analysts or data scientists. All of these jobs make a great living and can be a fulfilling future career for my students.

In conclusion, I want to emphasize to my students how data analysis skills were used by engineers during this summer fellowship as well as how these skills apply to numerous other high-paying, fulfilling careers. This will hopefully give my students real-life context and answer the age-old question students always have: "Why are we learning this?"

Host Organization Engagement (~100 words)

I will have an introductory assignment where students explore what work Lockheed Martin does through introductory materials and produce a flipgrid (~1 minute) responding to the prompt: What type of work does Lockheed Martin do and how might graphs/charts be essential to their work?

Throughout the summer, I participated in many "data dives" at Lockheed Martin. During these Skype meetings, quality engineers at Lockheed Martin would present graphs and charts about engineers' productivity. The lead quality engineer would then ask follow-up questions to really push the quality engineers to understand what the data implies and give suggestions on new ways to splice the data to find new implications.

Throughout my years teaching science, I have emphasized to my students how to display data they already know the trend of. For example, they generally know that the trend is either increasing or decreasing before even graphing the data. I want to push them next year to make implications based on the data rather than just display data. This ETP is an introduction to that skill, by teaching basic graphing skills, but later on the school year I plan to implement a "discussion scaffold" based on data that students can have with their lab groups inspired by the questions Lockheed Martin engineers frequently asked during their data dives. Examples include but are not limited to: In what ways can the data be displayed differently? What do the trends seen in the graphs imply? How can we use this data to make changes to future experiments? However, to get to that point, they must have basic graphing skills down first. So this entire ETP can be seen as a "baby step" towards the advanced graphical analysis skills Lockheed Martin engineers employ in their work. Later on, I will invite my mentors to

Skype in to a "data dive" of classroom data and have them help us dig deeper into the data we collected and push our understanding of the data.

Instructional Plan (This is the bulk of your ETP and may take several pages.)
Plan something that can be done remotely.

Day #	Students will be able to...	Materials Needed
1	Select the best graph (bar, line, or pie chart) to use for a set of data.	Access to devices for each student with Google suite (sheets and slides), Internet access, Google classroom
2	Find the sum and percentages of data using Google sheets.	
3	Generate bar, line, and pie charts using Google sheets.	
4	Collect data on leafhopper populations and create graphs.	
5	Create a presentation with graphs and write about the implications of the data.	

	Content Standards	
	Adaptation also means that the distribution of traits in a population can change when conditions change. (HS-LS4-3)	
Timing	Day 1	Assessment
~ 1 hour	**Objective:** SWBAT select the best graph (bar graph, line graph, or pie chart) to use given a set of data. ***Focal question:*** *How do we best select graphs to display certain types of data?* **Student Materials:** 1. Pre-survey 2. Flipgrid assignment 3. Types of Graphs (Edpuzzle) 4. Types of Graphs (Practice Worksheet) *For teacher*: Answer Key **Teacher Instructions:** First, post the graphing pre-survey for students to do before they start on the assignments. Assign students the Flipgrid assignment to introduce the unit. In the assignment, students will learn what type of work is done at Lockheed Martin and make a prediction on how graphs and charts might be essential in their work. You can post this set of assignments on Google classroom with instructions for students to first complete the Edpuzzle and then try the practice assignment and turn in. Make sure to set the Edpuzzle so students can not skip through the video. After the majority of students have turned in the assignment,	Selecting the correct graph *Note* Answer key with explanations is already embedded in the quiz.

	post the link to the assessment (Google form) and collect data. When posting on Google classroom, make sure you change the assignment setting to "Make a copy for each student." Students who have not mastered the material can be invited to a Google/Zoom meet where you review the practice worksheet, then have students re-attempt the assessment. In the small group re-teaching, you can use Zoom polls to have students engage and measure their understanding. Just project the same questions as the practice worksheet and have students vote. If the percentage is low, review the general rules with students (percentages = pie charts, line graphs = shows trends, etc). Then have students go into breakout rooms (in groups of 2-3) for the next question and have them discuss the answer and why. During this time you can hop around into the rooms. This is the equivalent of an online "think-pair-share." Have the students vote again on the next few questions. If the percentage of right answers increases, have a student volunteer to explain why. If it stays the same, review and repeat this process.	
Timing	Day 2	Assessment
~ 1 hr	**Objective:** SWBAT find the sum and percentages of data using Google sheets. ***Focal question:*** *How can learning basic formulas on Google sheets speed up calculation time?* **Student Materials:** 1. Basic formulas on Google sheets (Edpuzzle) 2. Applying basic formulas (Practice Worksheet) + Answer Key **Teacher Instructions:** Similar to Day 1 Lesson, post the Edpuzzle and practice worksheet on Google classroom. The Edpuzzle is designed to walk them through the Guided Practice part of the spreadsheet then have them practice a few problems independently. When students turn in the assignment, post the assessment on Google classroom and give students feedback. When posting on Google classroom, make sure you change the assignment setting to "Make a copy for each student." **Note**: To check if students actually use the formula instead of calculating by hand, click on the boxes where there are percentages or sums. You should see a formula on the fx= box. (See picture here). Students who did not master the skill or use the formulas can be invited to office hours where you review the practice worksheet and have them re-attempt the assessment. *Tip:* have students share their screen and give them feedback. You can also have students work together in pairs using breakout rooms.	Applying formulas on Google sheets to determine the sums and percentages of a data set. Answer Key
Timing	Day 3	Assessment
~ 1 hr	**Objective:** SWBAT generate bar graphs, line graphs, and pie charts using Google sheets. ***Focal question:*** *How do we create graphs using Google*	Knowing how to create various charts of data sets using Google

	sheets? **Student Materials:** 1. Making graphs on Google sheets (Edpuzzle) 2. Making graphs on Google sheets (Practice Worksheet)+ Answer Key **Teacher Instructions:** Post the Edpuzzle and practice worksheet, then post the assessment last. If students are struggling with this piece, it would be best to schedule 1-on-1 meetings with the teacher. A great way to organize meetings is to set up a Calendly where students can sign up for available times to learn how to graph. Calendly will then automatically set a Zoom invite to both the teacher and student.	sheets. Answer Key
Timing	Day 3	Assessment
~ 1 hour	**Objective:** SWBAT collect data on Leafhopper populations and generate charts. ***Focal question:*** *How can we collect data and make charts using Google sheets?* **Student Materials:** 1. Leafhopper Final Assessment (all needed links are embedded in the instructions). **Teacher Instructions:** You should do a live meet where you explain the final assessment and rubric to your students as well as answer their questions. Students can type questions directly in the chat box for the teacher to take breaks and answer questions in intervals. You should also familiarize yourself with the Leafhopper slideshow and demonstrate how to count the bugs using the first example situation (Global Warming), then have students collect their own data for the next 3 situations. The exemplar materials are examples of a Level 3 on the rubric. Do not show students the exemplar, this is a resource for teachers to guide their grading. Spend this first day having students collect data and fill out the data table, and start making graphs.	Link to Assessment/Rubric (posted on Google classroom as an assignment) Exemplar Data Sheet (for teacher usage only)
Timing	Day 4	Assessment
~1 hour	**Objective:** SWBAT create a presentation with graphs and write about the implications of the data. ***Focal question:*** *How can we clearly display the meaning of data in a presentation?* **Student Materials:** 1. Leafhopper Final Assessment (all needed links are embedded in the instructions). 2. Graphing Post-survey **Teacher Instructions:** Students will copy and paste their graphs into this presentation and use the sentence frames to write about the implications of the data emphasizing adaptations. Based on the charts, they should be able to indicate the leafhoppers that were most adapted to the environment. You may have students turn in the slideshow to be graded for the content of the presentation, and presentation is not an assessment area of this module.	Link to Assessment/Rubric (posted on Google classroom as an assignment) Exemplar Presentation (for teacher usage only)

	Optional Have a few students present on Zoom for extra credit and give them feedback. Or use breakout rooms and have students present their slideshows in pairs and give each other feedback using the rubric before turning in the final assignment. Lastly have students fill out the graphing post-survey to compare data to the pre-survey. **Note** Line graphs are shown in the exemplar but if students decide to make pie charts of trait frequencies or bar charts comparing the generations, that is also appropriate for the data. As long as they can make the connection between increased trait frequency and adaptations using their data, they are showing mastery of both graphing and understanding of adaptations.	

In-Person Enhancements

Here are some ideas to make these lessons more engaging if implemented in a traditional classroom. Instead of using the EdPuzzles, teachers can present and walk students through each step in a traditional lecture format using the following slideshows:
[1] Pick the right graph! (Day 1 Lesson)
[2] Basic Formulas Practice (Day 2 Lesson)
[3] Graphing with Google Sheets (Day 3 Lesson)
Teachers can also project the spreadsheets and walk through the formulas/graphing with them. Instead of just turning in the slideshow for grading, teachers can have students present their slideshows to a seat partner and give each other feedback using the rubric. Have a few students volunteer to present for extra credit.
For the summative assessment, teachers can also replace the Leafhoppers data collection slideshow with their own lab activity as long as it requires students to collect some sort of quantitative data that can be graphed. For example, a simple data collection lab that also indirectly teaches density is to have students measure the mass and volume of different-sized objects. Teachers can cut a piece of balsa wood into various shapes and give each lab group a piece of wood. Or you can use marbles or clay balls of various sizes and have each group focus on one marble/clay ball. Have students measure the mass using a balance then measure volume using a graduated cylinder (displacement). Demonstrate how to do this. Have students add their data to the class set (post this spreadsheet on Google classroom). Then share the data with the class and have them graph volume on the X axis and mass on the Y axis. The slope is the density. You can still grade them based on the graph they made and how they organized their data table. You can then have them make a slideshow and present to their seat partner what the density of the unknown material is.

Supply List

A device (such as chromebooks) for each student with internet access and Google Suite (Sheets and Slides) is all that is needed for this ETP.

References

[1] Google Sheets function list - Docs Editors Help. (n.d.). Retrieved July 02, 2020, from https://support.google.com/docs/table/25273?hl=en

[2] Types of charts & graphs in Google Sheets - Docs Editors Help. (n.d.). Retrieved July 02, 2020, from https://support.google.com/docs/answer/190718?hl=en

[3] Polling for meetings. (n.d.). Retrieved July 16, 2020, from

https://support.zoom.us/hc/en-us/articles/213756303-Polling-for-meetings

[4] Enabling breakout rooms. (n.d.). Retrieved July 16, 2020, from https://support.zoom.us/hc/en-us/articles/206476093-Enabling-breakout-rooms

[5] Video tutorials on Calendly. (n.d.). Retrieved July 16, 2020 from https://help.calendly.com/hc/en-us/articles/360000234614-Video-Tutorials

[6] Getting Started on Flipgrid: Educators. (n.d.). Retrieved July 27, 2020 from https://help.flipgrid.com/hc/en-us/articles/360007460474-Getting-Started-Educators

[7] Getting Started on Flipgrid: Students. (n.d.). Retrieved July 27, 2020 from https://help.flipgrid.com/hc/en-us/articles/360007559273-Getting-Started-Students

Keywords

Digital graphing, google sheets formulas, data analysis, data interpretation, distance learning

Links to Files in this ETP

Google Forms

1. Graphing Pre-survey
2. Graphing Assessment Survey
3. Graphing Post-survey

EdPuzzles:

4. Types of Graphs (Edpuzzle)
5. Basic formulas on Google sheets (Edpuzzle)
6. Graphing with Google sheets (Edpuzzle)

All other links can be found in the following folders

1. Lesson Materials by Day
2. Assessments