

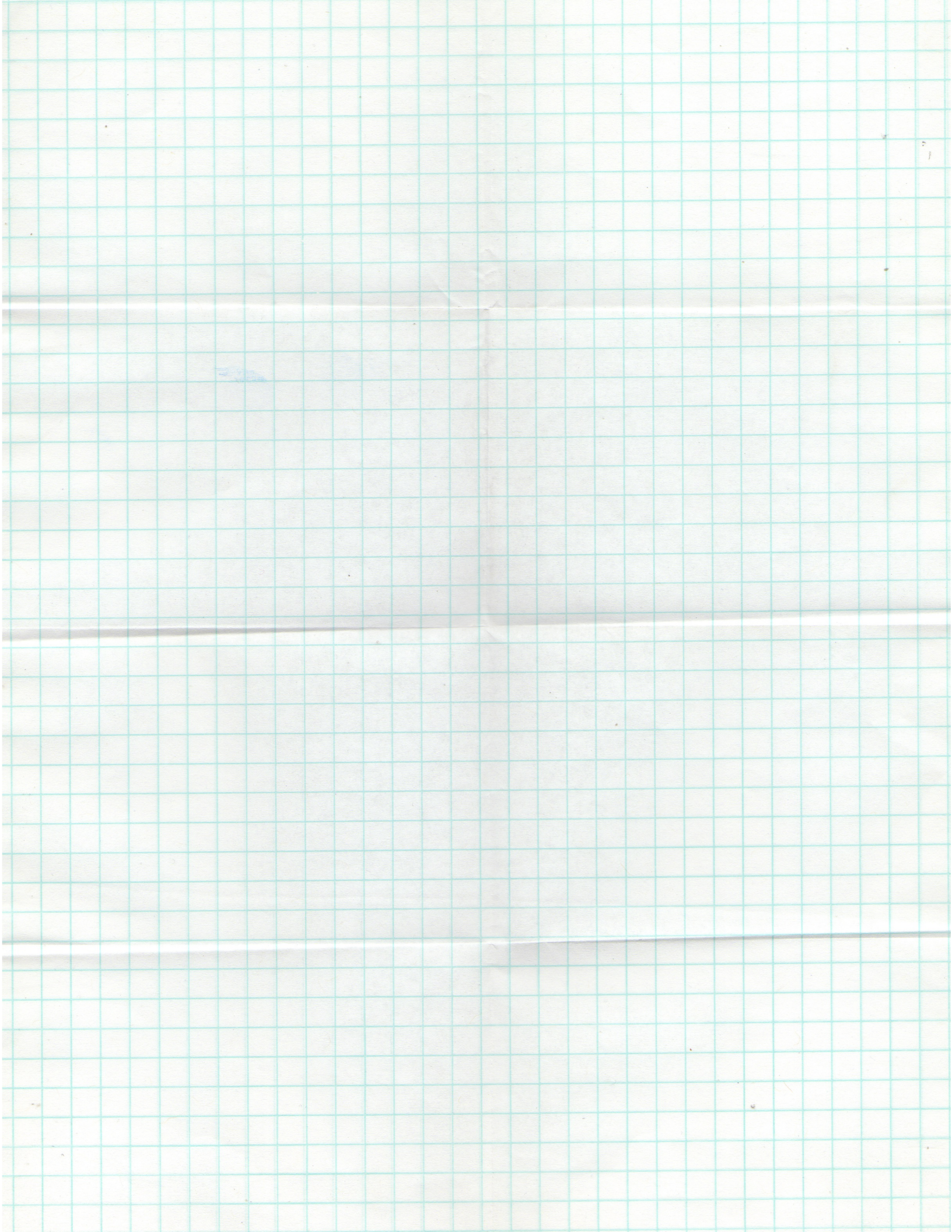
Data

Visualization

the great synthesizer



An activity book to help different
classes practice the skills they
need to succeed



Foreword

Acknowledgements

We would like to thank the University of Rhode Island, the US Census Bureau, NERR (Natural Estuarine Research Reserve), Brown University, the Stroud Water Research Center for the use of their data, programs, visualizations, and other support they have provided in the making of this book.

A special thanks goes out to NSF (National Science Foundation) and RI C-AIM (Rhode Island Consortium for Coastal Ecology Assessment Innovation and Modelling) for bringing all of us together.

Purpose

The purpose of this book is to present a number of activities for high school students related to understanding, analyzing and creating data visualizations. It is intended for use by high school teachers who are teaching data visualization concepts in their classrooms.

This book has a list of possible activities, lists of sources and data sets, and more information about the activities. Feel free to modify, expand, or use any of the activities as an inspiration in your classroom.

Authors

This book was written by faculty and students at Rhode Island College. They are Dr. Sally Hamouda, Dr. Anabela Maia, Lauren Cenedella, Destiny Gonzalez, Meghan VanSchalkwyk, and Matthew Spaulding. You can contact us at: shamouda@ric.edu.

Sponsor Information

This material is based upon work supported in part by the National Science Foundation under EPSCoR Cooperative Agreement #OIA-1655221.

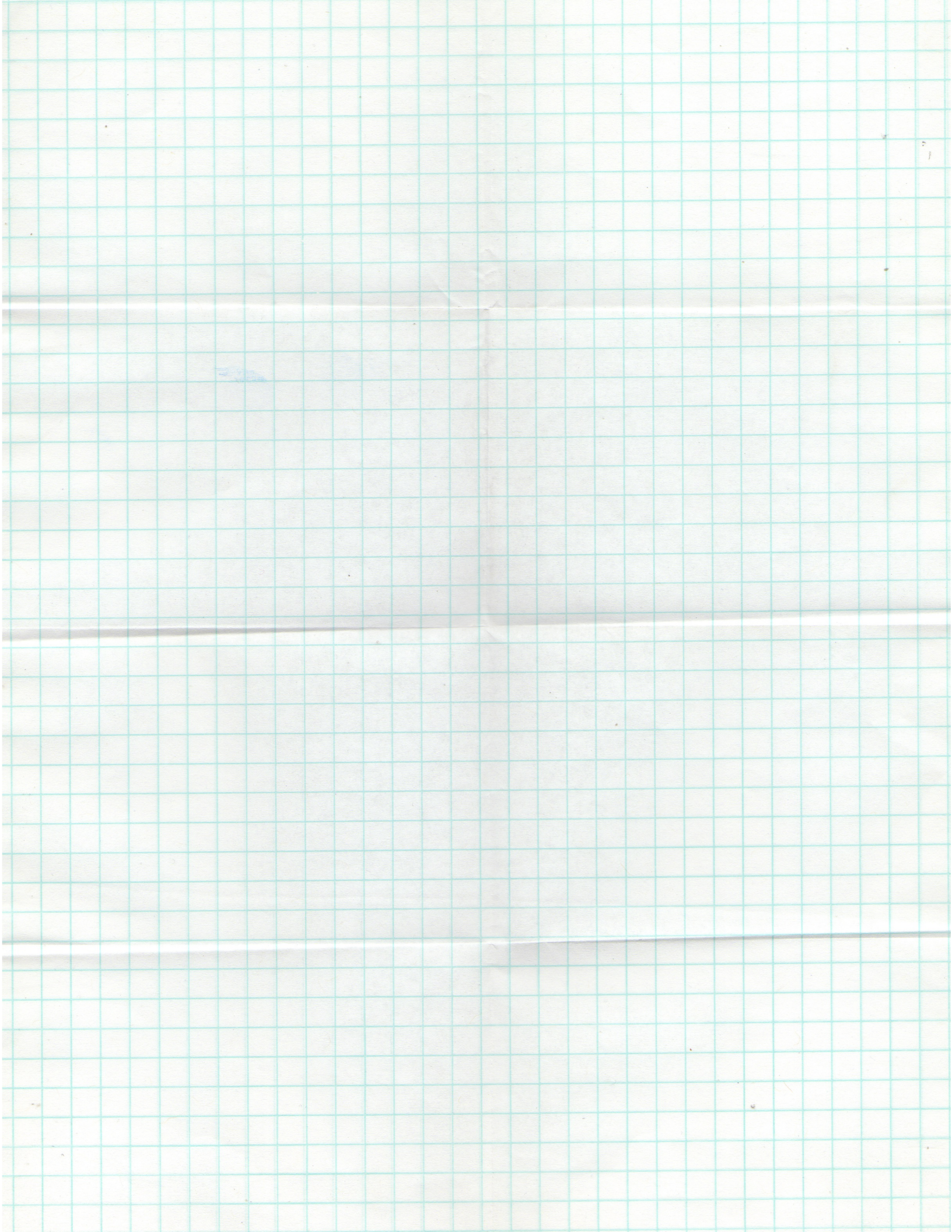
Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.



Table of Contents

Foreword	i
Computer Science	2
Code.org’s: Good vs. Bad Visualizations	3
The Ethics of Visualizations: The Challenger Space Shuttle Disaster	5
Back to Basics With Bootstrap	7
Fishy Data (Cleaning Data)	11
How Are You Feeling?	13
Creating a Graph With Jupyter Notebook.....	16
English/Literature	20
What’s Trending Now?.....	21
Graphing Grammar	23
Social Networking – Who Knows Who.....	25
Complete the Infographic	28
Visualize a Theme.....	30
History of a Text	32
Mathematics	34
Analyze an Unfamiliar Visualization	35
Let’s Sample Some...	38
So Many People!	41
Geometry Jam – Understanding Scale	43
Algebra Exposition.....	44
Science	46
Analyze a Visualization	47
Graphing Over Time	49
Model My Watershed	51
Navigating NERR Data	54
Combining Data Sets	57
Social Studies	62
National Geographic’s Marshallese Stick Charts	63

Voting: Then and Now.....	64
Torres Strait Tallies – A Chain of Achievements	66
Mapping America	68
Land of the Favored	71
Miscellaneous.....	72
What Should I Be Eating?	73
What’s So Good About It?.....	75
Self-Portrait in Data.....	77
Taking a Look at Treemaps.....	78
Wearable Data.....	84
Appendix	86
Data Sets and Links.....	87
Links to more information about activities.....	88
Works Cited	89



Computer Science

Code.org's: Good vs. Bad Visualizations

Categories: Computer Science

Yields: Discussion, Visualization (optional)

Summary: Using Code.org to show and discuss the differences between good and bad visualizations.

This activity will require a free Code.org account for either the teacher, or all of the students to make a virtual classroom. To get to the data visualizations, click on “Course Catalog” in the top navigation bar.

1. On the new page, scroll down until you see this chart (*figure 1*). Where it says “High School,” click on “CS Principles” --> “View Course.”
2. Scroll down until you see the section for “Post AP – Data Tools (‘19-’20)” and click “Go to Unit.”
3. On the unit page, go to “Lesson 2” (*figure 2*). It will have Collections “A” and “B” that you can assign to students if they have accounts, or you can pull them up on a projector and go through the visualizations as a class.
4. On the right side of the “Lesson 2” box, click “View Lesson Plan.”

Recommended Code.org courses [View my recent courses >](#)

Courses from Code.org for students in grades K-12 and professional learning for teachers.

Elementary school					Middle school			High school						
K	1	2	3	4	5	6	7	8	9	10	11	12		
													CS Principles	
													CS Discoveries	
													CS Fundamentals	
													Pre-reader Express	CS Fundamentals: Express
Professional Learning for all grade levels												Learn more		

Figure 1. The interactive Code.org course catalog guide.

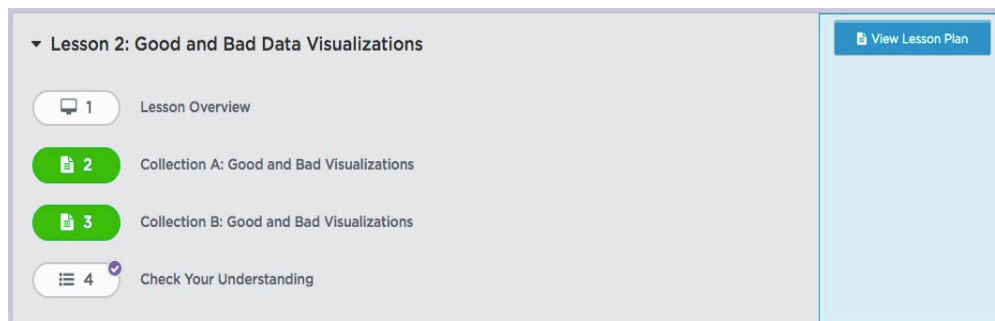


Figure 2. Lesson 2 (Good and Bad Data Visualizations) from Code.org.

5. On the right side of the “Lesson Plan” page, there is a section called “Preparation.”
6. This section has the link to the scorecard worksheet students can use to rate each visualization. It also has sections that ask students to choose the best and worst visualizations out of the collection and explain their reasoning. Optionally, they can also redesign the worst visualization to make it more coherent and explain why their way works better.

Alternatively, you could use the slides in the Day 1 PowerPoint presentation we used in the workshop. It can be found by using [link #1](#) in the appendix’s “Data Sets and Links.”

The Ethics of Visualizations: The Challenger Space Shuttle Disaster

Categories: Computer Science, Ethics

Yields: Discussion, Visualization (optional)

Summary: Using the provided figure to discuss what is wrong and why. Opportunity to discuss why people sometimes make misleading visualizations.

Activity adapted from: www.adam-p-jones.com/why_is_visualization_important/

This exercise requires an image of the O – Ring damage report that was used before the Challenger Space Shuttle Disaster (see *figure 1*). You can find a digital copy by using [link #2](#) in the appendix’s “Data Sets and Links” section. Give the students time to analyze the image. What do they see wrong with it? Why does this make the visualization confusing to try to understand? Why did it fail to delay the launch of the Challenger shuttle despite the engineers realizing there was a risk of O – Ring failure?

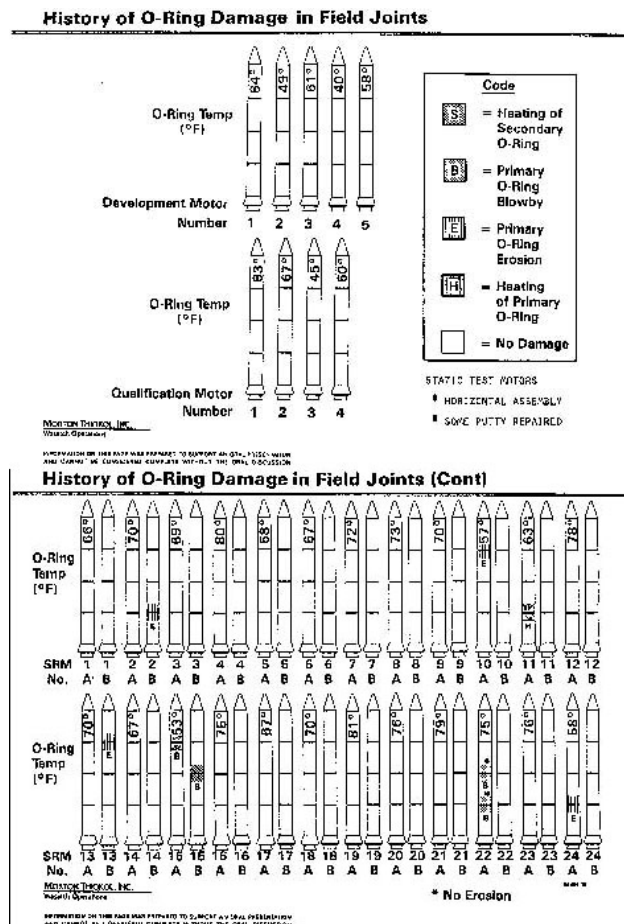


Figure 1. NASA's O - Ring damage visualization. Link to digital version in appendix. Source: Jones

Optionally, given the information about O – Ring damage, have the students create their own visualization.

Student Questions and Answers:

1. What problems do you see with the visualization?

Answer: Visual is too cluttered, extra information is included, which makes it very difficult to read.

2. Why does this make the visualization confusing to try to understand?

Answer: It is not organized by temperature, which would clearly link temperature to O-Ring failure.

3. Why did it fail to delay the launch despite the engineers realizing there was a risk of O-Ring failure?

Answer: Since it is hard to read the low temperature was not seen as enough of a threat to delay the launch.

4. How would you organize a visualization to display the same information?

5. How should they have presented it to be clear and compelling? Remember, public pressure was on management to launch so a very clear argument must be made to not launch.

Answer: Present the same information as a bar (or similar) chart with no extra details to make it very clear and compelling.

Back to Basics With Bootstrap

Categories: Computer Science

Yields: Visualizations (digital)

Summary: Working with different data sets to make different visualizations.

Everyone needs to start somewhere, and this activity is meant to help students that need an introduction (or refresher) on creating visualizations from data. I used the Music, US Income, and RI Schools data sets from the [Bootstrap link](#) (#3 from the appendix's "Data Sets and Links") to create a bar chart, pie chart, line graph, and scatter plot.

1. Starting with the Music data set, I decided to clean the data by adding a column with data validation. Right-click on column J and insert a column to the right of it and call it "genre."
2. Right-click on cell K2, scroll down and click on "Data validation."
3. In "Cell range," add ":K994" to the end (see *fig. 1*).
4. In "Criteria," click the dropdown menu and choose "List of items" and choose the genres you think are common (keep it simple and small!). I chose classical, country, cultural, dubstep, jazz, pop, r&b, rock, and other/unknown. Also be sure to check the "Show dropdown list in cell" option (see *fig. 1*).
5. Click save and we can begin to filter the data.
6. Select cell H, then click "Data" in the top toolbar, and then choose "Create a filter."
7. To make the list more manageable, click the dropdown arrow next to the column header called "familiarity." Choose "Filter by condition," "Greater than or equal to," and put 0.9. We will only have to categorize the songs that 90% of the public is familiar with.
8. After doing that, we can make a pivot table. Click "Data" again but choose "Pivot table" on a new sheet.
9. Add genre to "Rows" and uncheck "Show totals."
10. Add genre to "Values."

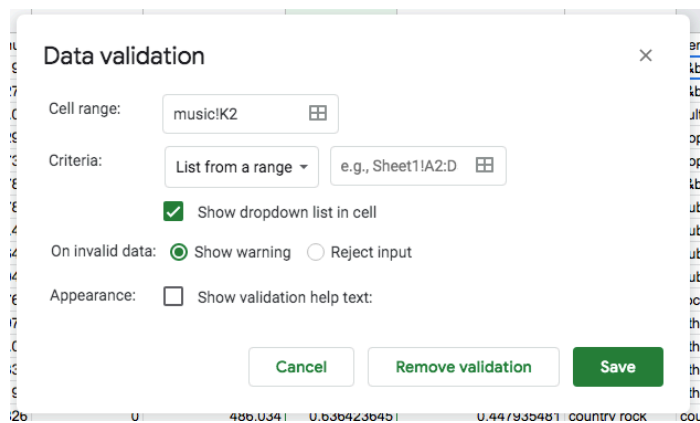


Figure 1. Excel's data validation tool. Used here to restrict a column's data.

11. Add the same conditional filter from the sheet to “Filters.”
12. With cell A1 selected on the pivot table, click “Insert” in the top toolbar and choose “Chart.”
13. Now that you have a chart, you can customize it as you please. I was able to make a bar and pie chart from this one pivot table (*fig. 2 and 3*).

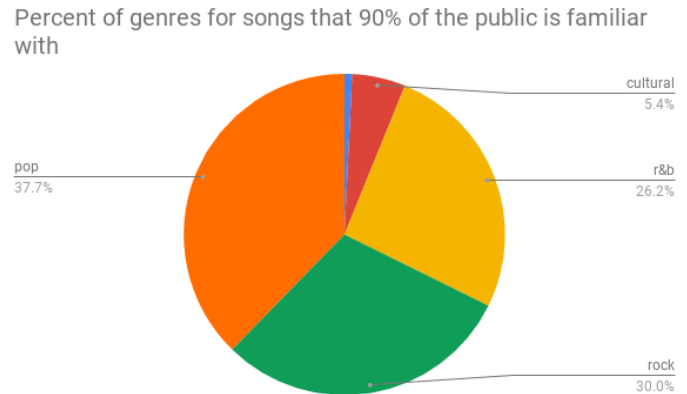
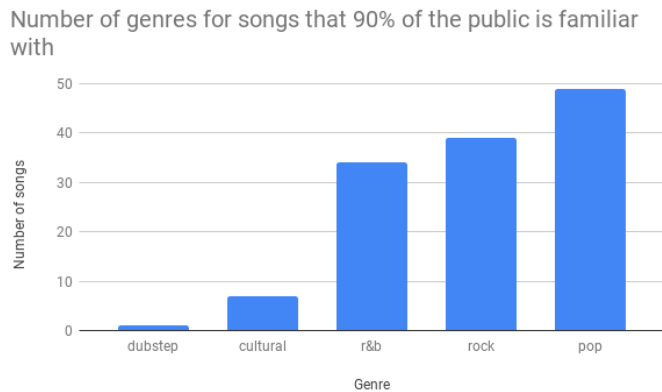


Figure 2. A bar chart of the number genres for the most popular songs. Figure 3. A pie chart of the most popular songs' genres.

1. Next, with the US Income data set, there is no cleaning needed.
2. Simply select columns A, L, and M, and then click “Insert,” and “Chart.”
3. It should come up with a line chart comparing the median and mean income of households over time. With this visualization, you can take the time to discuss why mean is so far above median (a few households with a ridiculous amount of money skews the data), and why you should always consider what is being graphed. My line graph, after customizing it, is seen below in figure 4.

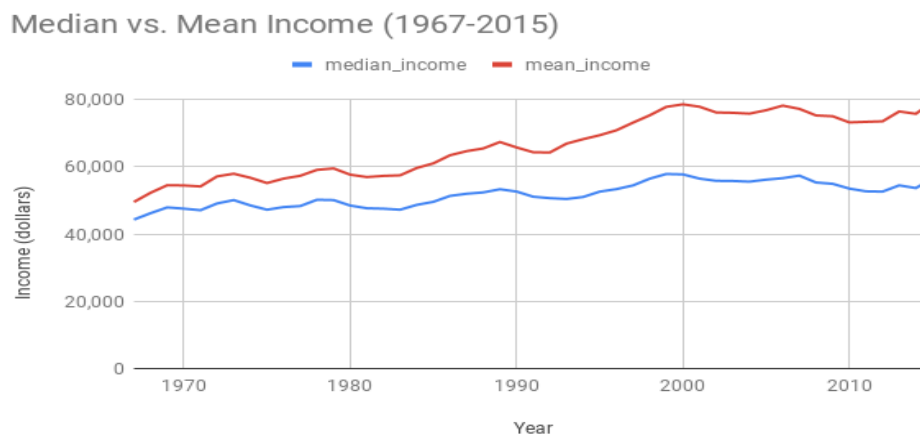


Figure 4. A line graph comparing the median and mean income of U.S. citizens over time.

Finally, with the RI Schools data set, you can play around with trends and correlations which are perfectly represented by scatter plots. For this example, I wanted to see if an increasing number of students correlated with the percentage passing English/language arts.

1. I added a filter to column P, and clicked “Sort A --> Z,” which ordered the data from smallest student body to the largest.
2. I chose to make a chart of the 10 smallest schools, so I selected ranges P1:P11 (total students) and C1:C11 (percent passing) and made a chart.
3. In the chart editor, you want to change the X- and Y-axes, so click on each and choose the correct series (total students on X-axis and percent on Y). This is shown in figure 5.
4. On the series tab, click the 3 dots and choose “Add labels.”
5. Click on the little box (see bottom right corner in *fig. 5*) to add the data range of A1:A11 (the school names).
6. Now in the “Customize” tab of the chart editor, scroll down to “Horizontal axis” and set the min and max bounds (0 and 200).
7. Go up to “Series” to show the trendline and R^2 value.
8. Customize the chart as you please. My visualization is shown below in figure 6. As you can see, Google Sheets has some trouble making the data labels visible.

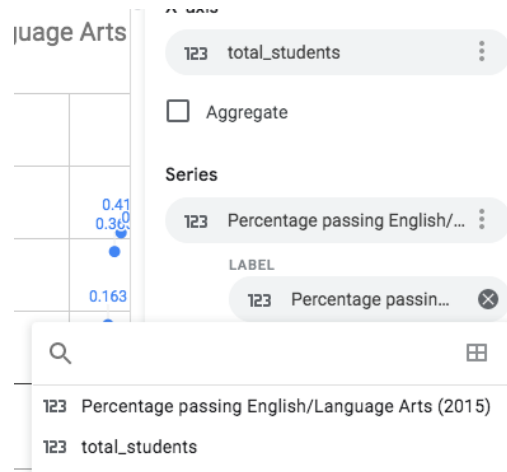


Figure 5. The chart editor showing the correct X- and Y-axis data. Also shows the “Add labels” box.

Percentage Passing English/Language Arts (2015) vs. Total Students

The 10 schools with the smallest student body.

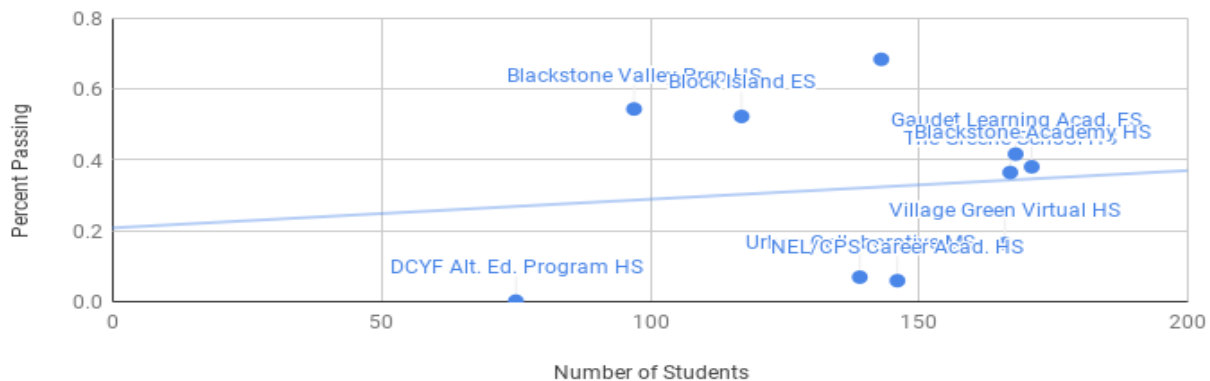


Figure 6. A scatter plot of percent passing vs. number of students. Also shows the trendline and R-squared value.

Student Questions and Answers:

1. What type of chart(s) did you choose for the music data and why?
2. Do you think the type of graph you chose displays the data clearly? If it doesn't, why not? How could it be improved?
3. For the US income data why is the mean so high?
Answer: The mean is so high due to people with very high income skewing the data.
4. What does that tell you about how outliers can affect data?
Answer: It shows that outliers can have a big impact.
5. Do you see any correlation between class size and passing English/Language arts? Why do you think that is?
Answer: The R^2 value is extremely small, so there is little to no correlation. However, you can hypothesize that the smaller the class size the easier it will be for the teacher to help students in the class, resulting in more students passing.
6. Do you think there would be a correlation between passing math and class size? Why do you think there would or would not be a correlation?
7. If you graph the data with math were you correct about there being a correlation?

Fishy Data (Cleaning Data)

Categories: Computer Science, Science

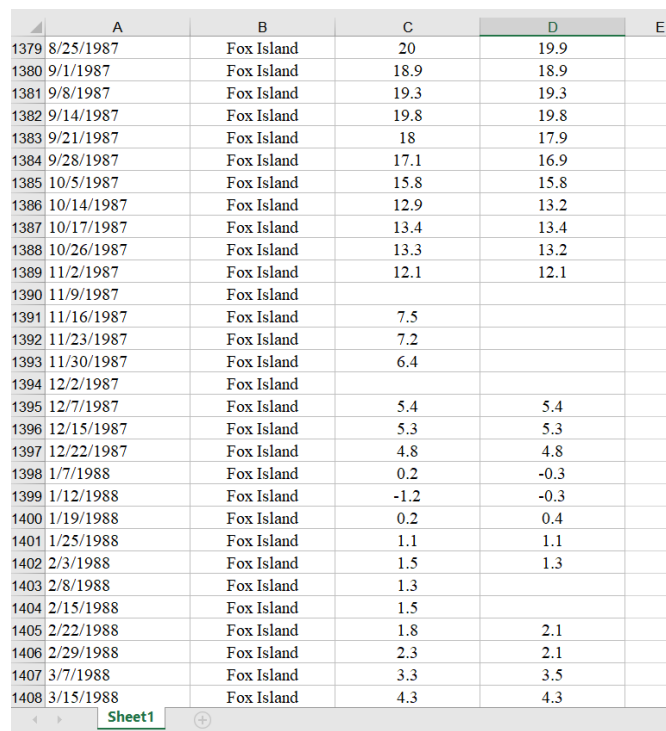
Yields: Clean data, Visualization (digital)

Summary: Discussing how uncleaned data can change a visualization. Students then clean the data.

Using the URI water temperature data from Fox Island ([link #4](#) in the appendix) have the students view the data. (Note: the excel sheet is 2,900 cells long, choosing a subset of the data that has some missing values that need to be cleaned is recommended for less experienced students)

Ask the students if they think the missing values would impact a visualization. How would it make an impact and what can be done about the missing data? Ask the students what other issues you may find in real world data that is not found in this data set. (Duplicate data, outliers, missing headers, etc.)

Now have the students clean the data and prepare it for a visualization. This will depend on the programming language or tool the students are using. Figure 1 is an example of some of the data with missing values.



	A	B	C	D	E
1379	8/25/1987	Fox Island	20	19.9	
1380	9/1/1987	Fox Island	18.9	18.9	
1381	9/8/1987	Fox Island	19.3	19.3	
1382	9/14/1987	Fox Island	19.8	19.8	
1383	9/21/1987	Fox Island	18	17.9	
1384	9/28/1987	Fox Island	17.1	16.9	
1385	10/5/1987	Fox Island	15.8	15.8	
1386	10/14/1987	Fox Island	12.9	13.2	
1387	10/17/1987	Fox Island	13.4	13.4	
1388	10/26/1987	Fox Island	13.3	13.2	
1389	11/2/1987	Fox Island	12.1	12.1	
1390	11/9/1987	Fox Island			
1391	11/16/1987	Fox Island	7.5		
1392	11/23/1987	Fox Island	7.2		
1393	11/30/1987	Fox Island	6.4		
1394	12/2/1987	Fox Island			
1395	12/7/1987	Fox Island	5.4	5.4	
1396	12/15/1987	Fox Island	5.3	5.3	
1397	12/22/1987	Fox Island	4.8	4.8	
1398	1/7/1988	Fox Island	0.2	-0.3	
1399	1/12/1988	Fox Island	-1.2	-0.3	
1400	1/19/1988	Fox Island	0.2	0.4	
1401	1/25/1988	Fox Island	1.1	1.1	
1402	2/3/1988	Fox Island	1.5	1.3	
1403	2/8/1988	Fox Island	1.3		
1404	2/15/1988	Fox Island	1.5		
1405	2/22/1988	Fox Island	1.8	2.1	
1406	2/29/1988	Fox Island	2.3	2.1	
1407	3/7/1988	Fox Island	3.3	3.5	
1408	3/15/1988	Fox Island	4.3	4.3	

Figure 1. A preview of the raw data from the Fox Island water temperature data set.

Student Questions and Answers:

1. How do you think missing values would affect a visualization?

Answer: Missing values may lead to incorrect conclusions.

2. What can be done about the missing data points?

Answer: They can be filtered out but you have to realize your visualization has missing data.

3. What other issues do you think you might find in a data set?

Answer: Other issues include missing column titles, data from different sources may not match up, spreadsheet formatting, human input error, time and date issues (no data was collected on the day listed for example).

4. What effect do you think they would have and how would you address the issues?

Answer: It makes data easy to misinterpret and requires extensive cleaning.

How Are You Feeling?

Categories: Computer Science

Yields: Discussion, Visualization (digital)

Summary: Students generate data by taking anonymous surveys, then clean the data. Special focus on the “countif” function to search unorganized data and count occurrences of a given parameter.

Times are tough and uncertain for everyone right now, and even when the threats of a pandemic aren't looming every time you go outside, it's always nice to watch out for each other. This activity has students creating data through daily anonymous surveys, and at the end they have a better understanding of how everyone is feeling – how their classmates are holding up away from everyone.

For the survey project, I made three separate surveys: the introduction to get data that's easy to manipulate (multiple choice, Yes/No, rating questions), the morning survey to ask about stress levels and provide a de-stressor (my example has a relaxing YouTube video), and the evening survey to ask about stress levels (a replica of the morning survey but without the de-stressor).

Once these are made, you can actually set up an automatic email to remind students to take the survey. In Google Calendar, I made an event called “Fill out survey” and added another account as a guest with the survey link in the description (*fig. 1*). You will add all of your students' accounts as guests, so they get the reminder. You can set the event to repeat each morning with the morning survey, and set another event to repeat each night with the evening survey.

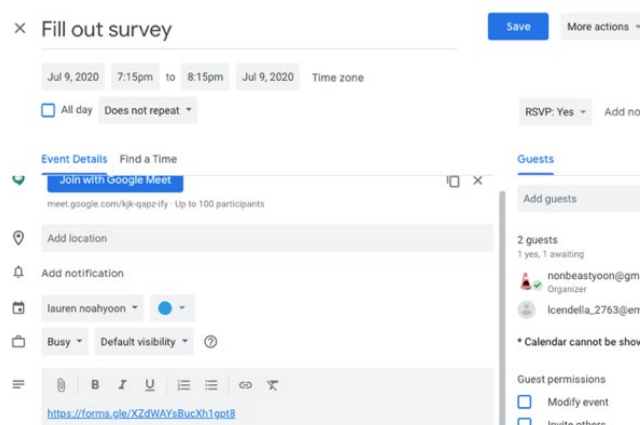


Figure 1. An example of what a Google Calendar reminder to take a survey might look like.

Looking at the introduction survey data, students will have three challenges to clean the data. They will have to separate the timestamp column, make a data validation column for the types of entertainment, and use the “countif” function and wildcard characters to count how many instances of a particular de-stressor there are. The first challenge is explained in the science activity “Navigating NERR Data,” and the second is explained in

the computer science activity “Back to Basics With Bootstrap.” For the last challenge, go to [link #1](#) in the “Links to more information about activities” section of the appendix. It will provide a good overview on how to use this function, but I will also explain how it is used in my cleaned data.

The last question allows users to pick more than one answer, meaning there will be several pieces of data in one cell instead of spread out individually.

1. We need to make a separate mini table of each option and its respective count so we’ll need to use the “countif” function.
2. The first parameter of the function needs a range, so I selected the column with unsorted data.
3. The second parameter needs a logic statement in between quotes. I wanted to put in the option by itself, but that doesn’t work. You have to add an asterisk (a wildcard character) so that the function looks for any instance with a space, no space, or comma before or after it.
4. You do not need to put the whole option between quotes, just something that is unique to it. For example, I have the options “Going outside” and “Going for a drive around town” as things that make you relaxed. I can keep “*Going outside*” in the function because it is short, but I can just put “*Going for*” as the parameter for the other option since it is a unique phrase found in no other option.

Figure 2, below, shows the data, the mini table I created, and the function that counts for instances of “Music.”

=COUNTIF(I2:I101, "*Music*")

H	I	J	K	L
Which pet do you prefer:	What are things that help		Thing	Count
Dog	Music, Going outside, Jus		Music	2
Cat	Going outside, Shopping,		Going outside	2
birds	Music, Going for a drive a		Just chilling around the h	1
	other		Going for a drive around t	1
			Shopping	1
			Hanging out	0
			Scented candles	2
			Playing with your pets	1
			Other	1

Figure 2. A sample data set generated by using the introduction survey. With cell L2 highlighted, the “countif” function and its parameters are shown.

These are links to each part of each survey. By going to each one, you will be able to take the survey but not see the results. You can create your own version of the surveys

to suit your class' needs. You can also find them by using [link #1](#) in from "Data Sets and Links."

Introduction Survey: <https://forms.gle/XR7HULrLrB5ExygFA>

Morning Survey: <https://forms.gle/EgJTLurL6PNjvxRt7>

Night Survey: <https://forms.gle/YVXHpeGqmbbXTsQV9>

Student Questions:

1. What did you learn about the process of collecting data through the surveys?
2. What would you change or improve about how you collected data from surveys?
3. What was the most common de-stressing tactic your classmates used?
4. Are there any days of the week or times when your classmates are the most stressed? The least stressed?
5. What did you learn about data, in terms of how to collect it, clean it, and analyze it?

Creating a Graph With Jupyter Notebook

Categories: Computer Science

Yields: Visualization (digital), Discussion

Summary: Using a graphing tool that requires programming.

Reference used in creating this activity: www.geeksforgeeks.org/bar-plot-in-matplotlib/

This activity requires Jupyter Notebook. This can be accessed either through installation or through the browser. To access the web version, visit the Jupyter website (<https://jupyter.org/try>) and select “Try Classic Notebook.” The web version includes common libraries such as matplotlib, numpy, and pandas. Libraries like seaborn and plotly must be installed manually, in which case Anaconda can be used for convenience.

Have students begin by creating an initial graph using hard-coded data. Provide them with a basic structure for the code if necessary and encourage them to experiment and change the values of variables, or customize the graph to, for example, change the color.

Once the students can display a graph and they’ve been given time to experiment, have them find or use a data set—CSV files are best. If the data set they use is too large, have them use a subset of the data so there’s only two types of values to compare (in the case below the two types are the date and the corresponding Apple stock price).

On the following page is an expected output (*fig. 2*) of the code from figure 1:

```
import pandas as pd
from matplotlib import pyplot as plt

data = pd.read_csv("https://raw.githubusercontent.com/plotly/datasets/master/2014_apple_stock.csv")
data.head()
df = pd.DataFrame(data)

date = df['AAPL_x']
price = df['AAPL_y']
fig = plt.figure(figsize=(20, 7))
plt.plot(date, price)
plt.title("Apple Stock Price")
plt.xlabel("Date")
plt.ylabel("Price")
plt.tick_params(labelbottom=False)
plt.show()
```

Figure 1. Code that reads a CSV file and generates a graph on Jupyter.


```

import pandas as pd
from matplotlib import pyplot as plt

data = pd.read_csv(r"https://raw.githubusercontent.com/plotly/datasets/master/2014_apple_stock.csv")
data.head()
df = pd.DataFrame(data)

date = df['AAPL_x']
price = df['AAPL_y']
fig = plt.figure(figsize=(20, 7))
plt.plot(date, price)
plt.title("Apple Stock Price")
plt.xlabel("Date")
plt.ylabel("Price")
plt.tick_params(labelbottom=False)
plt.show()

```

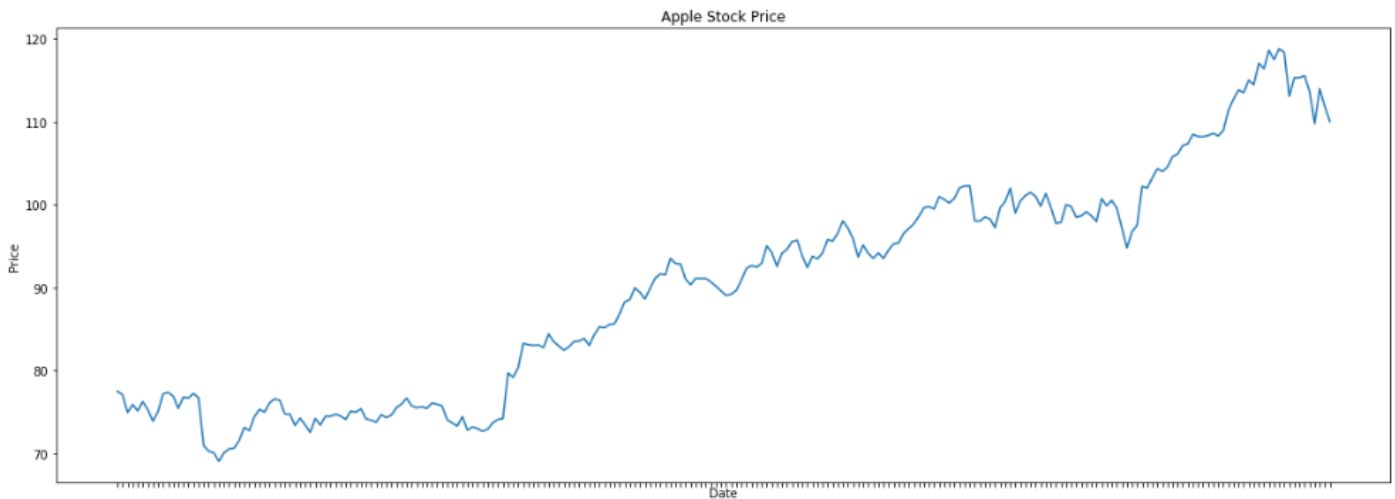


Figure 2. The resulting graph from figure 1's code.

Discuss any problems they had or any problems they can foresee. Maybe there was too much data for the graph and so the tick labels became squished and unreadable. Is removing the labels entirely, as seen in the picture, best? Or is there a better solution such as segmenting those values, in this case dates, into months or some longer time frame? Perhaps there were other issues as well.

Also discuss the benefits of creating a visual. Maybe the data they chose to use was large and so detecting a trend by reading the data alone became difficult, but seeing the visual representation allowed them to instantly find something interesting in the data.

Student Questions and Answers:

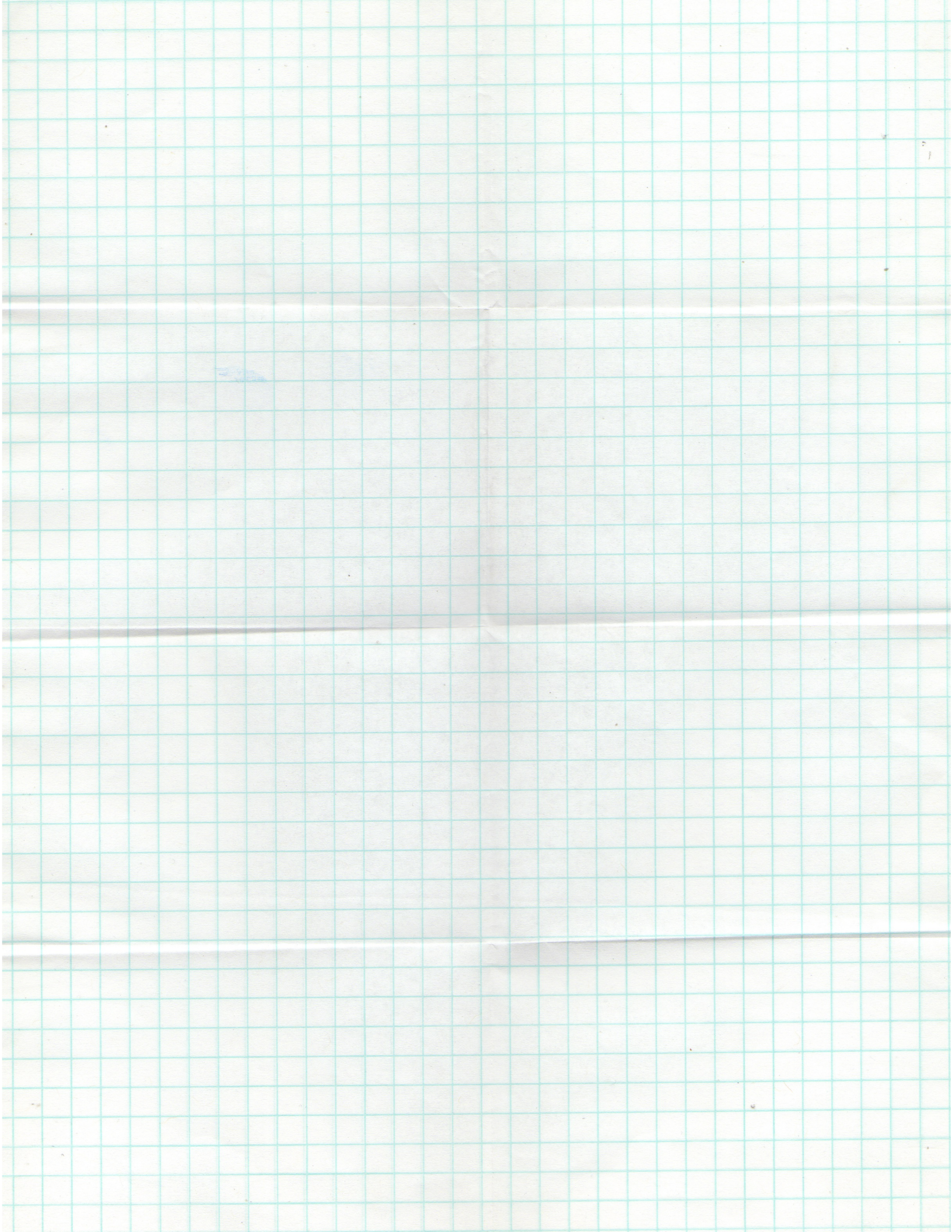
1. Are there any problems you encountered?

Answer: Problems they may experience or foresee may include issues with the code, inexperience with Jupyter Notebook, or with finding a data set.

2. If you got your graph working with your data and everything works fine, can you foresee any problems with other data sets?

3. Are there any benefits you see in representing data visually, compared to the raw CSV file?

Answer: Benefits to representing data include easily see trends and all of data, and it being easier to draw conclusions based off a visual than a CSV file of numbers.



English/Literature

What's Trending Now?

Categories: English, Social Science, Psychology

Yields: Discussion

Summary: Shows students how small changes in word choice can produce different visualizations.

Go to trends.google.com and type in two different topics: adopting dogs vs adopting cats. (Can do a poll for favorite animals amongst the class but be prepared for lower numbers) See what the comparable trends are (*fig. 1*) and what are the varying differences between states in the United States. What kind of map/graph is this the most like?

Now there are different ways that you can phrase this.

“cat adoption near me” and “dog adoption near me”; “buy dog near me” and “buy cat near me”; “animal shelter with cats” and “animal shelter with dog.”

See how the search levels fluctuate for different phrasings.

See how the same topic is being looked up but the phrasing makes a difference. This makes an important difference when you are doing research. Keep this in mind when you are doing research for certain topics.

*Can do any time frame or state but keep consistent.

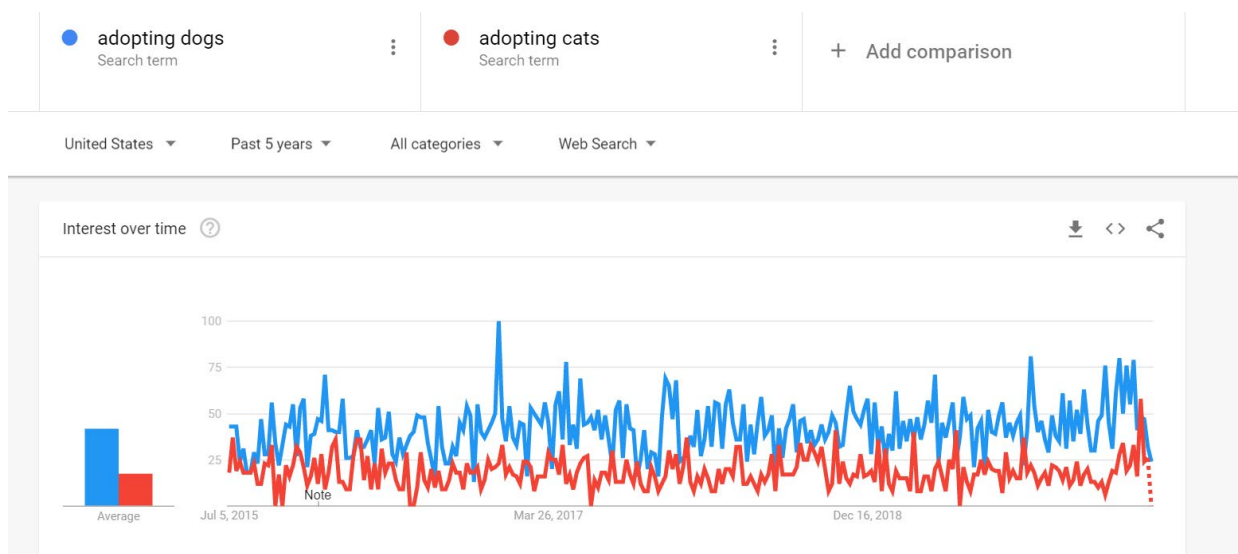


Figure 1. A Google Trends depiction of the popularity of searching "adopting dogs" and "adopting cats."

Student Questions and Answers:

1. What kind of map/graph is this most like?

Answer: It's a line graph.

2. What does this tell you about the potential to alter a visualization based on the search terms that are being compared?

Answer: Comparing different terms can alter the data to get you a different result.

3. Do you think a different kind of graph would be easier to understand? Why and what kind of graph?

Answer: This type of graph is fine, possible the graph would be clearer over less time. Or students may say a bar graph.

4. (Optional): have students pick two different time frames and do a comparison.

Graphing Grammar

Categories: English/Literature

Yields: Discussion, Writing Assignment

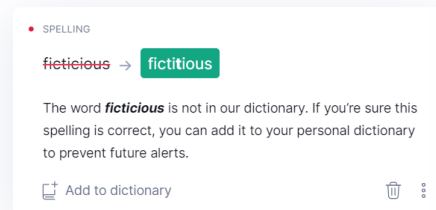
Summary: Using Grammarly to understand that visualizations can come from anywhere and still serve the same function – conveying information.

1. Using Google Chrome, download the Grammarly Extension.
2. Type two paragraphs about any given topic. Use sentences of varying lengths and colorful language. Use a minimum of 100 words.
3. See how Grammarly interprets sentence structure and gives suggestions.

Grammarly is a data visualization tool that works a little differently than how we are used to seeing Data Visualization. It uses everyday language to show us how we can improve our language skills and writing skills overall based on the generalities of the English language. This extension does many things such as correct spelling errors and suggest different verbiage for word placement to make sentences flow better (*fig. 1*). I encourage students to use this data visualization tool to upgrade their writing skills in various places whether it be work, school or personal matters. This is a simple example of a data visualization tool in the form of a common extension. It also has some premium features that can be paid for such as compound punctuation and word choice (*fig. 2*).

Add a title. We hope you're well!

Hello. My name is Destiny and today I will be advising you how to use the Grammarly extension on Google Chrome. I hope that you all come up with some interesting topics to write about and share with your peers in individual groups and tell stories, whether they are fictitious or truthful.



B I U | H1 H2 | ↻ | ☰ ☷ | ✕

52 words ↕

Figure 1. An example of Grammarly recommending a change to this paragraph.

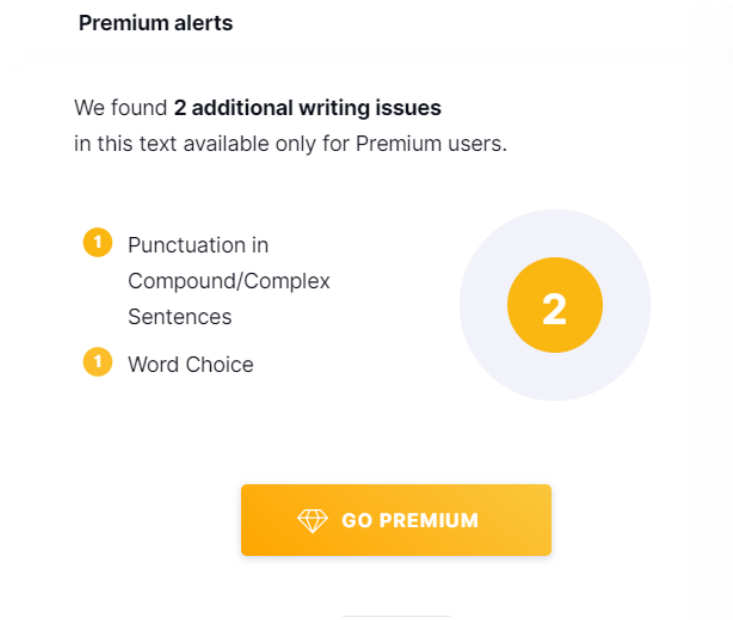


Figure 2. An example of Grammarly explaining how its premium service offers additional insights.

Student Questions and Answers:

1. How is this a data visualization tool?

Answer: Visually shows you where you can improve your grammar.

2. Do you see the potential of this being useful to you as a student in the future?

3. What do you notice about your writing style that you did not notice before?

4. Do you notice anything teachers have commented on before?

5. What, if anything, do you think you could learn from a person just by looking at their grammar?

Answer: May be able to learn about a person's attention to detail, education levels and maybe personality.

Social Networking – Who Knows Who

Categories: English/Literature

Yields: Discussion, Visualization (physical or digital)

Summary: Using a node diagram to track character relationships.

Activity adapted from: dataphys.org/list/harry-potters-social-network/

This activity introduces the concept of a node diagram. You can find a well-written description here at datavizcatalogue.com/methods/network_diagram.html. By using flourish.studio/, you will be able to create node diagrams for the characters and their relationships of any book the class is currently reading.

For example, by assigning a link value (based on the strength of their relationship) figure 1 shows the characters in *Romeo and Juliet*.

1. To begin, make an account with the Flourish website. You can sign in with your Google account, or you make a new account.
2. On the My Projects page, click “New visualization,” and then scroll all the way down until you see “Network graph” section. Choose the “Directional graph” option.
3. You will see a visualization that looks like the one in figure 2. Click the “Data” button and you will see 2 tables for the links and points.

Romeo and Juliet Characters

● Romeo ● Juliet ● Friar Lawrence ● Mercutio ● The Nurse ● Tybalt ● Capulet ● Lady Capulet ● Montague ● Lady Montague ● Paris ● Benvolio ● Prince Escalus ● Friar John ● Balthasar ● Sampson ● Gregory ● Abram ● The Apothecary ● Peter ● Rosaline ● The Chorus

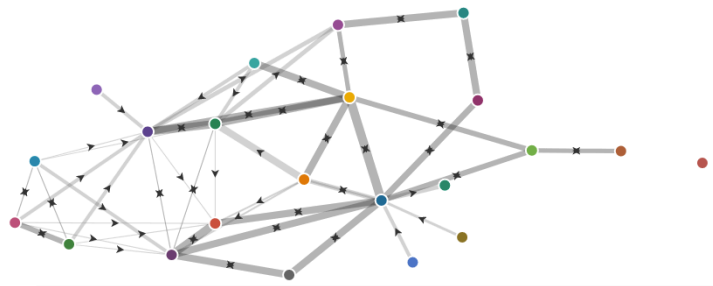


Figure 1. A node diagram of the relationships between characters in *Romeo and Juliet*.

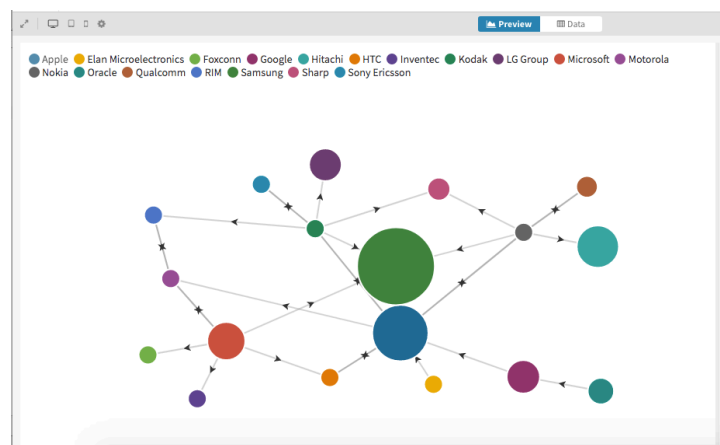


Figure 2. The default node diagram when choosing the network graph template.

4. Delete all of the data by clicking the dropdown arrows next to the labels “Links” and “Points,” and choose clear sheet.
5. You can then select the table headers (who, suing, who, etc.) and delete those, as well.

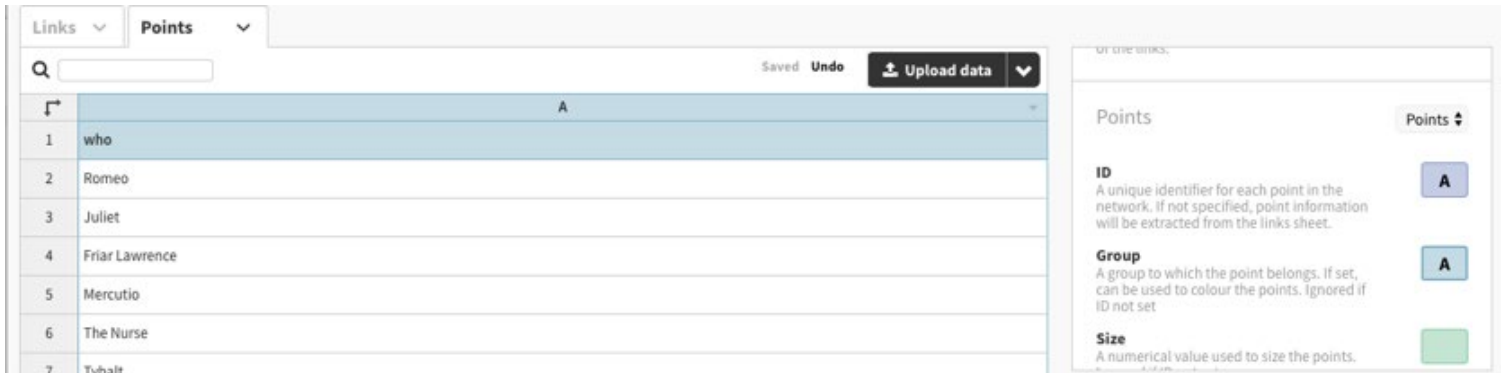


Figure 3. The correct "Points" table and side panel. Each row in the table is a character, and "ID" and "Group" in the panel are both set to "A."

6. Now we will set up our own tables. We want each point in the visualization to be 1 person, so we only need 1 column in the “Points” table.
7. To delete any extra columns, right click on their letter (A, B, C, etc.) and choose “Remove column.”
8. We can label the A column “who” and start filling in the characters (*fig. 3*).
9. On the right side of the screen there will be a list of options. Scroll down to the “Points” section and make sure “ID” and “Group” are both A, so that each point will have the name you put in and its own color (*fig. 3*).
10. For the links, choose a scale to define the strength of the relationship. For my visualization, I had 1-10 representing true hate, hate, merchant, servant, acquainted, friend, good friends, related/familial love, married, and true love.
11. The “Links” table will require 3 columns. To add a column, right click on the B column and choose “Insert column right.”
12. Label A as “who,” label B “to who,” and C “relationship.”

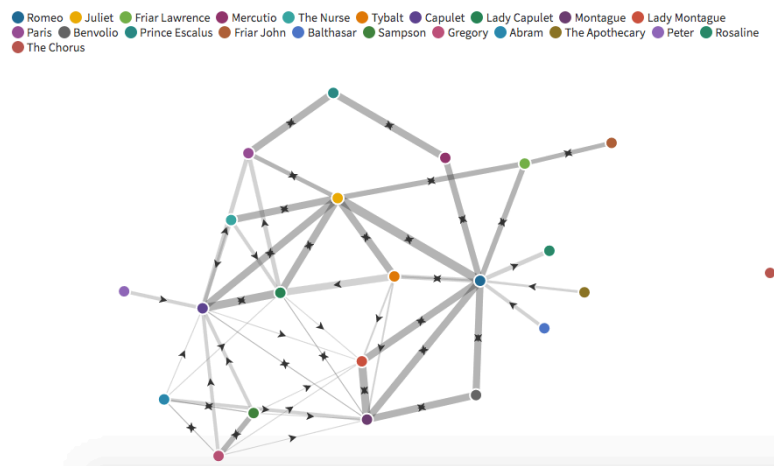


Figure 4. The same diagram as figure 1, but without a title.

13. On the options panel on the right, source should be A, target should be B, and value should be C.
14. After filling in all the relationships (I found SparkNotes' list of characters to be very helpful in this part), your chart should look like the one in figure 4.
15. You can add a title by going to the "Preview" tab, scrolling down to the "Header" section, and adding one.

Student Questions:

1. What is easier to understand in the graph compared to reading the book?
2. What is immediately clear about the relationships the main character has with the other characters?
3. Did you notice anything unexpected or that you did not realize before seeing the graph? What was it?

Complete the Infographic

Categories: English/Literature

Yields: Discussion, Visualization (digital)

Summary: Students complete an unfinished infographic provided by the teacher as they read a novel.

Activity adapted from: www.easel.ly/blog/infographics-for-teaching-classic-literature/

Based on the current book being read in class the teacher will make a visualization (of any sort) showing the characters, plot, themes or other relevant elements to the book. Provide an incomplete version of the visualization to the class and have them fill it in as the book is read. You can read more about this idea by using [link #2](#) from “Links to more information about activities.”

Optional: Show the students a visualization as an example and have them create their own. This can be done individually or in small groups.

Take a look at the example infographic on Piktochart with this link:

<https://create.piktochart.com/output/47671322-complete-the-infographic>

You can create your own personalized infographic by choosing a template (I chose the Monthly Project Status Report) and changing what is there to meet your needs. I added a blank graph for students to mark their reading (you would have to change the X-axis depending on how long you will cover the book) (*fig. 1*) and added a table for students to mark important events in each chapter (*fig. 2*).



Figure 1. A blank graph of pages read vs. day. It can be used in an incomplete infographic for students to track their progress.

Event Road Map

For each chapter, choose one event or quote that you think is important to the novel's progression or message. Provide its type (event, quote, theme, etc.) and a brief description.

	<i>Page Number</i>	<i>Description</i>
Chapter 1	<input type="text"/> Type:	
Chapter 2	<input type="text"/> Type:	
Chapter 3	<input type="text"/> Type:	

Figure 2. A blank road map students could use as they progress in a novel. It allows students to decide what they think is important in each chapter.

Student Questions:

1. What can you already learn or guess from the incomplete visualization you have been provided?
2. As you start to fill it in what do you notice about the book that you would not have seen without the visualization?
3. Did you find the visualization more helpful than the same information would have been in written form?

Visualize a Theme

Categories: English/Literature

Yields: Discussion, Spreadsheet, Visualization (digital)

Summary: Students track a word, theme, and/or phrase throughout the novel and visualize their findings.

Activity adapted from: www.edutopia.org/blog/homer-high-tech-data-visualization-james-earle

As a class have the students pick a theme or recurring event, such as every time a character is happy, or every time something is successful or fails. Track every instance in a spreadsheet, and at the end of the book create a visualization based on the data. This can be done in small groups or as a class. How the students organize their spreadsheets to best track the data and what type of visualization they make is up to them. You can read more about this idea and how others have implemented it by using [link #3](#) from “Links to more information about activities.”

As an example, I visualized themes and words from the first chapter in *The Great Gatsby*. I made a simple spreadsheet (*fig. 1*) since the data set was small, but complex tracking will result in a more complex sheet. You could include what chapter, page number, or quote the instance occurs in a collaborative spreadsheet.

JX	A	B	
1	Word/Theme	Frequency	
2	Gatsby	12	
3	Desires	9	
4	Assumptions	5	
5	Daisy	30	
6	Society	6	
7			
8			

Figure 1. A simple table for the number of occurrences of a word or theme in chapter 1 of The Great Gatsby.

Using Infogram, I chose a simple template and added a word chart. Depending on what students want to show, they will end up using different graphs and should be able to explain their choices. Students should also add a caption to their project as to what the visualization is showing if their title doesn't fully explain it.

My undecorated work ended up as shown in figure 2, but you can let the students get as creative as they want.

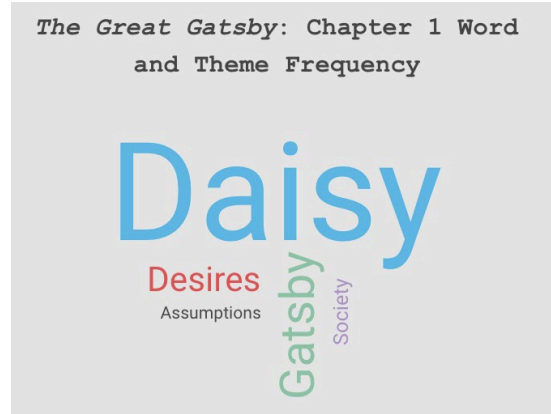


Figure 2. A word chart based off of the data in figure 1's table.

Student Questions:

1. Why did your group or the class pick the theme that they did?
2. How will you organize your spreadsheet to be clear and easy to read but still contain all the information necessary?
3. How will you organize your visualization?
4. Which is easier to understand, the spreadsheet or the visualization?
5. Is all of the data still displayed in the visualization?

History of a Text

Categories: English/Literature

Yields: Discussion, Visualization (physical or digital)

Summary: Students research a novel and create an infographic of their findings.

Providing context for a novel always enhances the experience of reading it. This activity provides an example infographic made with Infogram on what students could make when asked to research the context of a novel (see *fig. 1*). I researched *The Catcher in the Rye* and made a timeline of its relevance over time. You can view my infographic here: <https://infogram.com/history-of-a-text-1h7z2l3mgejy2ow>

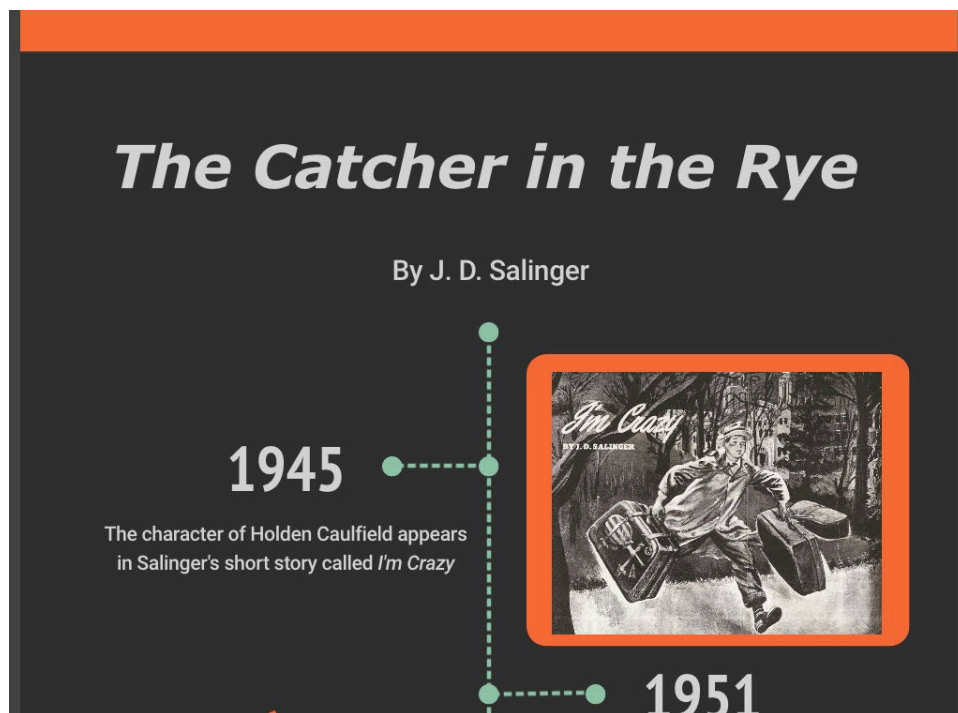


Figure 1. A snapshot of the infographic about *The Catcher in the Rye*.

Students can include information about the author's inspiration or life at the time their book was written, what was happening when the book was released, the impact the novel had (e.g. Sinclair's *The Jungle* led to food safety laws), or any other interesting facts. Instead of using Infogram, students could also make a PowerPoint/Google Slides presentation, or a physical poster/tri-fold board.

Student Questions:

1. What do you think was the inspiration for the book?
2. What were some relevant current events that were happening as the book was being written or around the same time it came out?
3. Did the book have an impact when it was first published?
4. Was the book controversial when it was published? If it was, explain why.
5. What did you find interesting when you began research about your text?

Mathematics

Analyze an Unfamiliar Visualization Average Test Grade

Categories: Math

Yields: Discussion, Visualizations (digital)

Summary: Finding the mean, median, mode, and range of different visualizations.

Imagine you are taking a math test and as a bonus question your teacher gives you a graph (see *fig. 1*). The question is as follows. “Using only this graph, find the mean, median, mode and range of the test scores for the last exam.” How would you approach? What is the mean, median, mode and range?

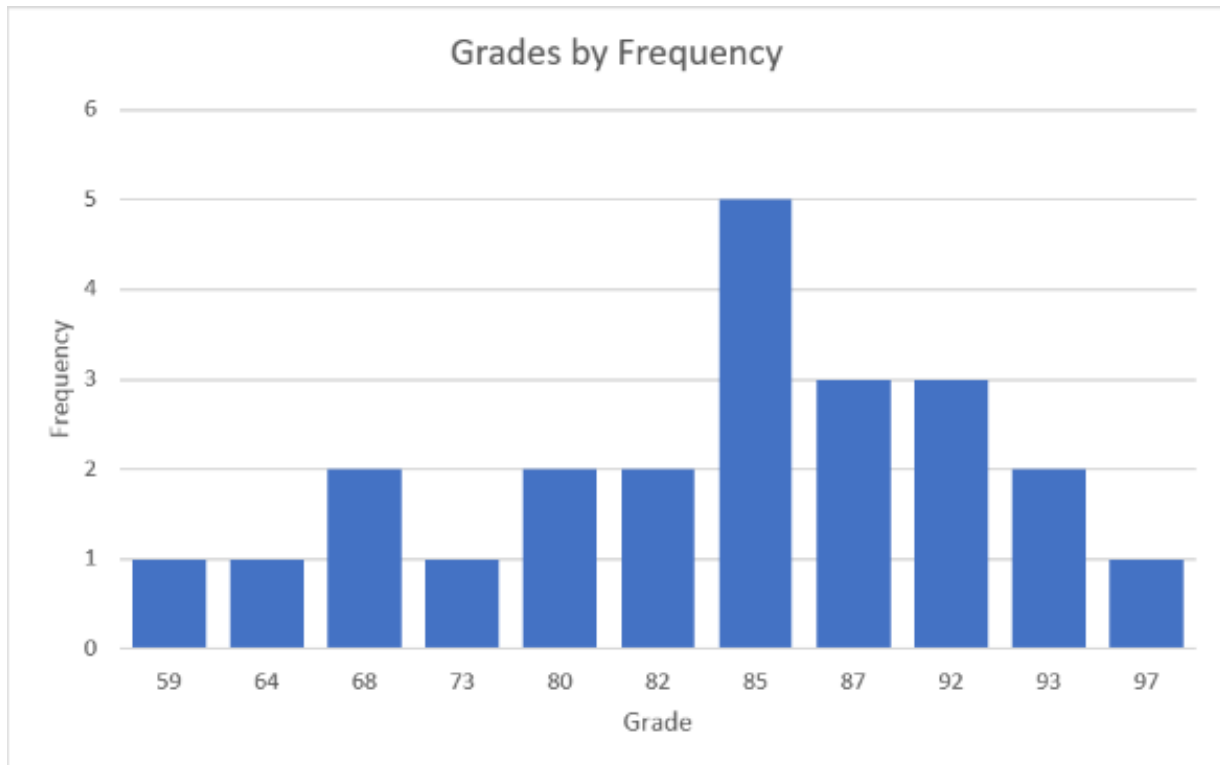


Figure 1. A bar graph of the frequency of grades.

Having knowledge of data visualization is something that can help with this, all that you need is to know how to take an average! It may not seem difficult, but this is a simple use of data visualization. Instead of having a list of scores that may seem like an overwhelming amount of information, the information is listed concisely in an ordered chart.

Student Questions and Answers:

1. Given the following graph can you determine the mean, median, mode and range of the data? Hint: you will need to find the average.

Answer: Mean: 82.65, Median: 85, Mode: 85, Range: 38

2. What do you see that is familiar in this visualization? What is unfamiliar?

Answer: Students will know a bar chart but may not be familiar with using it to get mean, median, mode and range.

3. What, if anything, confuses you about it?

4. Why do you think the creator chose to display the information in this manner?

Answer: Chose to display it this way to clearly show all information, creator may not be interested in the mean, mode, etc.

5. Do you think this is an effective way to display all of the information? Why or why not?

Answer: It does effectively display all information.

Regional Average Snowfall Winter 2019

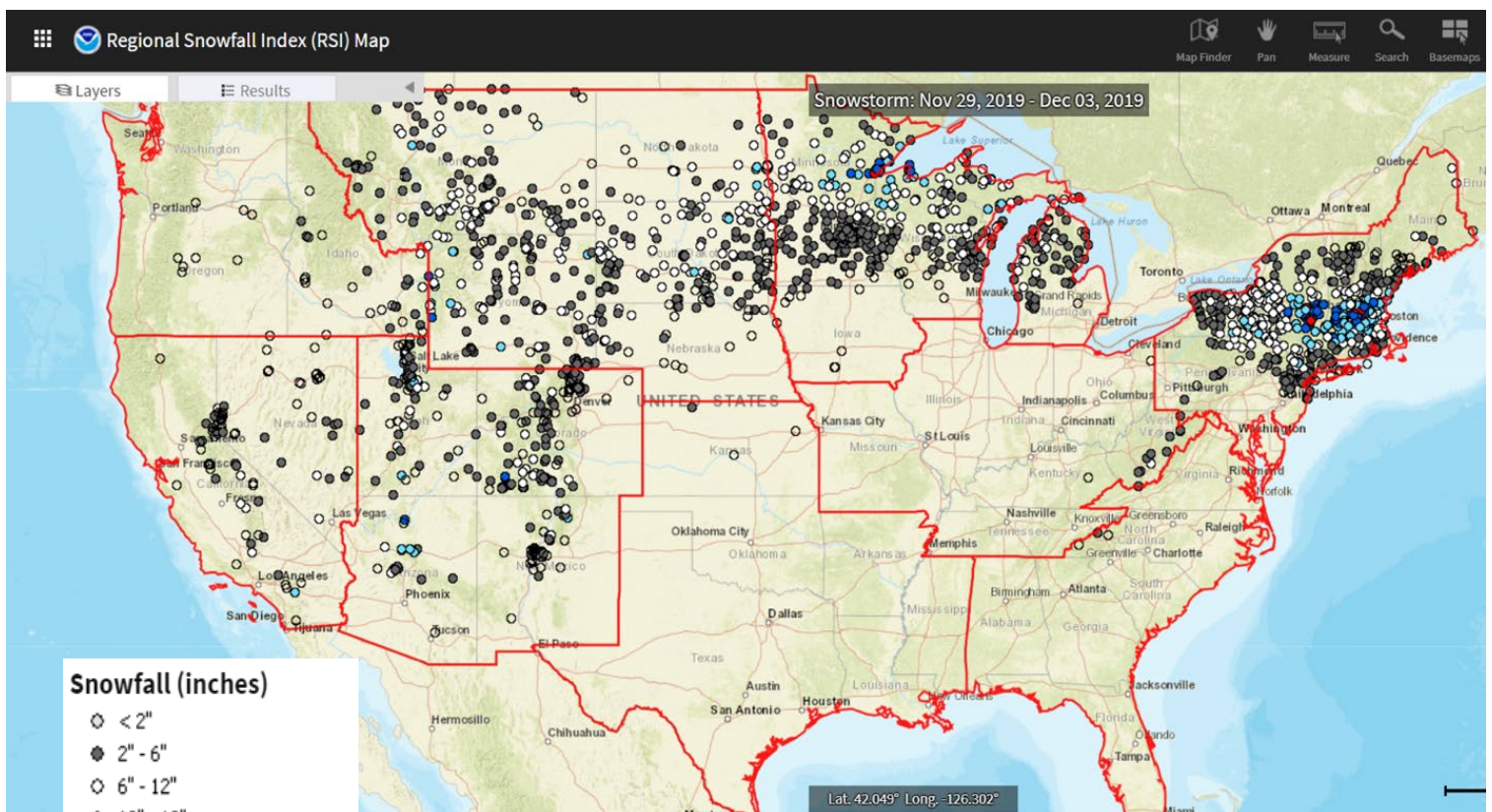


Figure 2. A map depicting snowfall in 2019 in different places with colored dots representing the amount of snow.

Using the above image (*fig. 2*) from the National Center for Environmental Information, answer the following questions. [Link #5](#) from “Data Sets and Links” brings you to the digital map.

Student Questions and Answers:

1. How would you go about finding the mode?

Answer: Must first find the average to find the mode. In this case you would have to estimate the average. Since most of the data points are dark gray or white, that leaves the average most likely in the middle at 6”.

2. Find the mode for the regions with Phoenix and Oklahoma City in inches.

Answer: Phoenix: 2” – 6” (up for debate), Oklahoma City: < 2”

3. Do any of the regions contain all the ranges of inches in the legend?

Answer: Yes, the New England region has all of the ranges.

4. What do you notice about the areas?

Answer: New England seems to have the highest dot concentration per area but many others have a lot scattered around (e.g. the region with Nebraska and North Dakota).

5. What color dot occurs the most often?

Answer: Dark gray or white

Let's Sample Some... Pizza!

Categories: Math

Yields: Discussion, Visualizations (digital)

Summary: Understanding how sample size impacts distribution and creating bell curve charts.

Below I have listed the prices of large cheese pizzas around Rhode Island as of 6/19/2020 (before tax and delivery). **Teacher's please let students know: Per the Central Limit theorem any sample with a size greater than thirty will create a normal distribution.** Create a graph of these thirty+ places using Excel. Here's a refresher on how to graph distributions – <https://www.extendoffice.com/documents/excel/2404-excel-template-bell-curve-chart.html> - bell_curve_chart

On TI calculator.

1. STAT --> EDIT --> put numbers in a list
2. STAT --> CALC --> 1-Var Stats
3. Make sure "List" is the list where you entered the data, and choose "Calculate."

Xbar (x with a bar on top) = the avg of the numbers

$$\text{Standard Error of } \bar{X} = \frac{s}{\sqrt{n}}$$

We will be using S_x : because that is the sample standard deviation. $S_x=s$; where n = number of elements in list/set.

Table 1. A list of pizza places and the cost of pizza. (n=32)

Place	Cost
Big Tony's	11.35
Tommy's Pizza	15.10
Providence Coal Fire Pizza	21.00
Sicilia's	11.99
Caserta Pizza	17.50
Napolitano's	16.00
Pizza J	15.00
Twin's Pizza	16.85
Tom's Kitchen	12.45
Federal Hill Brick Oven	13.25
Antonio's	12.25
Warren House of Pizza	12.90
Piezonis	12.49
Papa Gino's	13.99
Balasco's Pizzeria	15.45
Pomodoro Pizzeria	13.00

Place	Cost
Cantazaro's Pizzeria	18.99
Nino's Pizza	14.50
Uncle Tony's	15.20
Sako's Pizza	11.99
Tony's Pizza Palace	12.50
Maria's Pizza and Wings	12.50
Boulevard Pizza	11.99
Domino's	10.99
Vivaldi's Pizza	10.95
Tomaselli's Pizza	13.49
Papa John's	11.99
DePetrillo's Pizza and Bakery	13.50
East Providence House of Pizza	10.59
Pizza Roma	10.99
Rosa Mia Pizzeria	16.00
Nice Slice	16.00

Student Questions and Answers:

1. Does it fit a normal curve?

Answer: Yes

2. What was the standard error of the mean?

Answer: Standard error of mean: $2.47 / 5.66 = .4364$

3. Do you think you would get the same result if you looked at pizza prices across the United States?

4. Find the mean of the sample and plug into TI-84 (if available). If not use this standard deviation calculator, put in comma separated list, and make sure to select that it is a sample here: <https://www.mathsisfun.com/data/standard-deviation-calculator.html>

Answer: Based on mathsisfun.com: Mean = 13.835625, Standard Deviation (S_x) = 2.470287873, these are rounded to 2 decimal places

... Heights!

Height is normally distributed. We will take a sample of size less than 30. We will use women's heights which has a mean μ of 63.7. Say we have a group of women with the heights listed below, create a graph. Is it still normally distributed? Heights (in inches): 65, 48, 60, 52, 62, 60, 61, 63, 64, 60, 62, 63, 60, 62, 66, 67, 62, 63, 66, 63, 65, 63.

Student Questions and Answers:

1. Find the mean of the sample and plug into TI-84 (if available)

Answer: \bar{X} = 61.6818, S_x =4.3357, Standard error of the mean =
 $4.3357/4.6904 = .92438$

Optional:

1. Measure the heights of everyone in the class. Do you still get a normal distribution?

Answer: Heights of everyone in the class will only result in a normal distribution if there are more than 30 students in the class.

2. Why do you think you got that result?
3. Does gender make a difference? Number of data points?

So Many People!

Categories: Math, Data

Yields: Visualization (digital)

Summary: Letting students explore the Census Bureau to create visualizations.

Activity inspiration from: www.census.gov/programs-surveys/sis/activities/math/census-in-counties.html

According to the U.S. Census Bureau there are certain numbers that they have recorded here (<https://www.census.gov/schools/facts/> and in [link #6](#) from “Data Sets and Links”) and I would like for you to use them and pick 7 of your FAVORITE states (or territory). Create a bar graph of the populations in the year 2010 compared to the year 2018 in whatever program that you feel comfortable with (Excel online is perfectly acceptable and may be the easiest). Make sure to insert your chart with accurate numbers and your links with a list of states that you have chosen! Best of luck!

I have attached an example (*fig. 1*), please do not use ALL of the same states!

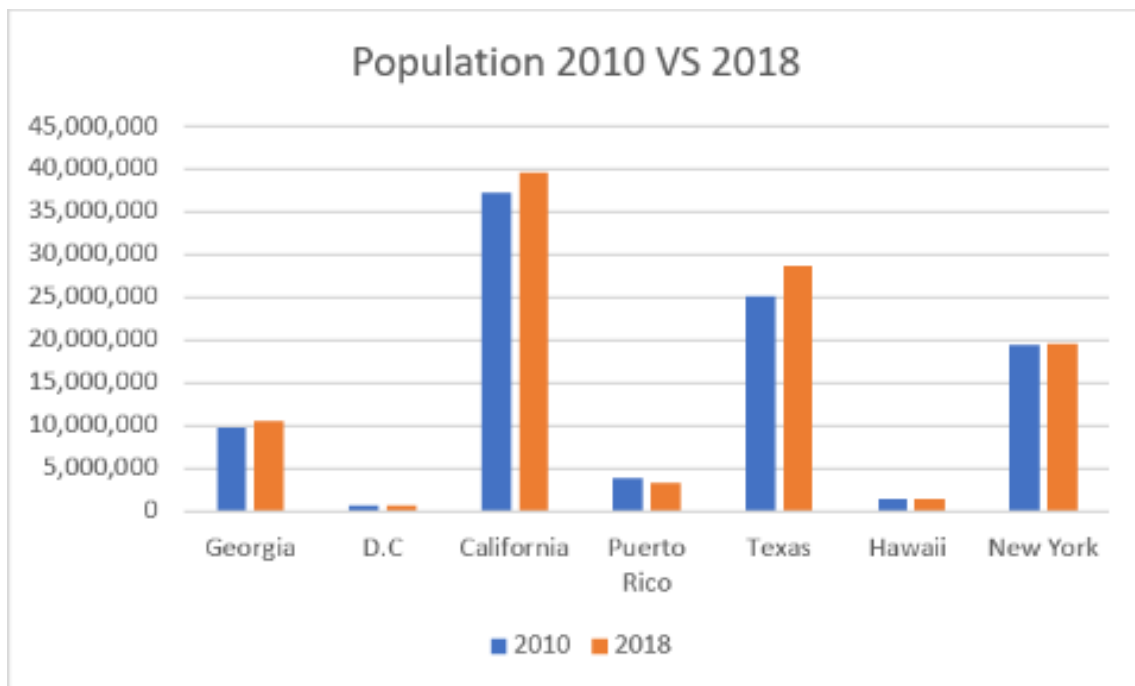


Figure 1. A bar graph comparing states' populations in 2010 vs. 2018.

Student Questions and Answers:

1. Do you see the population increasing or decreasing in each state?
2. If your seven states are close to each other geographically are they all decreasing or all increasing?
3. If some are increasing and some are decreasing what factors do you think would cause states near each other to have different population change?

Answer: Causes of population change may be natural disaster, availability of jobs, politics, and taxes/cost of living.

4. If you also graph 2014 is the change consistent in each state?

Geometry Jam – Understanding Scale

Categories: Math

Yields: Discussion, Visualization (physical or digital)

Summary: Using Margaret Ann Fazekas' activity to demonstrate the importance of measurements and scales. Students create a scale model by researching dimensions of a given building/place, or use dimensions provided by the teacher.

Being able to measure accurately is an important skill no matter where you end up in life, and this activity serves as one way to help students improve their skills. Read Margaret Ann Fazekas' "Baseball Project" activity from her thesis (link below and as #4 from "Links to more information about activities") to get a better understanding of where you could go with this idea of a scale/blueprint activity.

https://digitalcommons.lsu.edu/cgi/viewcontent.cgi?article=1538&context=gradschool_theses

To make this activity work for different grade or skill levels, you can change the output or data collection aspects. Instead researching dimensions and making a poster board presentation, you could have younger students make a scale model of a given house in games like Minecraft or the Sims. They would then take screenshots of their model in order to present their work. Older students can go through the process of collecting data, choosing a scale that fits on paper or a poster, and accurately using tools to create the model.

This is also a great opportunity to introduce the concept of significant figures, accuracy vs. precision, conversion factors, and general measurement etiquette. After going to the link below, scroll down to section five to see the rules and examples of significant figures.

<https://wou.edu/chemistry/courses/online-chemistry-textbooks/foundations-general-organic-biological-chemistry/chapter-1-measurements-chemistry/>

Student Questions:

1. Before this activity have you ever used a compass, geometry triangle or other tools before?
2. How familiar are you with the metric system of measurements?
3. What did you learn about the importance of accurate estimates?
4. What did you learn about scale and proportions?

Algebra Exposition

Categories: Math

Yields: Discussion, Visualizations (physical or digital)

Summary: Students are given a math topic and have to create an infographic to help their peers study. Suggested to use before an exam.

Whether you're preparing students for a mid-term or final, having an interactive study guide is a great way to convey material in a different way to facilitate different memory formation. Assign easy-to-visualize topics to small groups, and have them create a short presentation. For example, you could have students present or share an infographic about quadratic equations, linear equations, or systems of equations. With systems of equations (see *fig. 1*), students can provide the associated vocabulary, typical problems, one example problem with guided learning, and visualizations. Students can even make a Google Form of questions related to their topic for peers to fill out and check their work.

You can take a look at an example of a presentation students could make here:

<https://infogram.com/algebra-exposition-1hdw2jogj9le6l0>

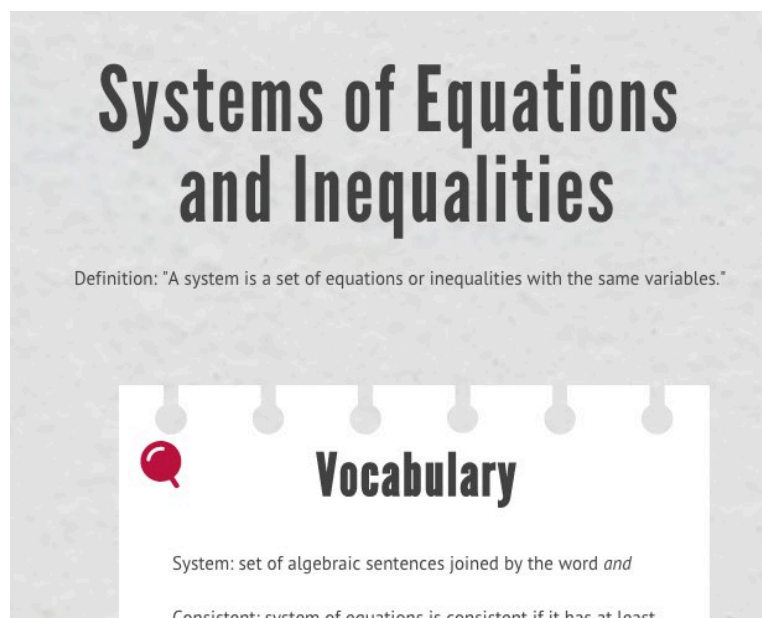


Figure 1. A snapshot of the example infographic students could make with this activity.

While this activity is focused on algebra concepts, it can work for any subject involving graphing and visualizing.

Student Questions:

1. Did you find a visual study guide to be more useful than a standard written one?
2. Did you find the presentations easier to learn from than the book or just looking at text?
3. Do you think it would be worth it to you to make your own visualization as a study guide?
4. Were the sources your classmates listed useful to you?

Science

Analyze a Visualization

Categories: Science

Yields: Discussion, Visualization (optional)

Summary: Students analyze a visualization about the oxygen requirements of different sea creatures.

From http://www.watershedcounts.org/marine_water_quality.html (or link #7 from “Data Sets and Links”), scroll down (about mid-way down the page) until you see the three thumbnails of visualizations. Have the students analyze the visualization titled “Dissolved Oxygen Needs” (*fig. 1*).

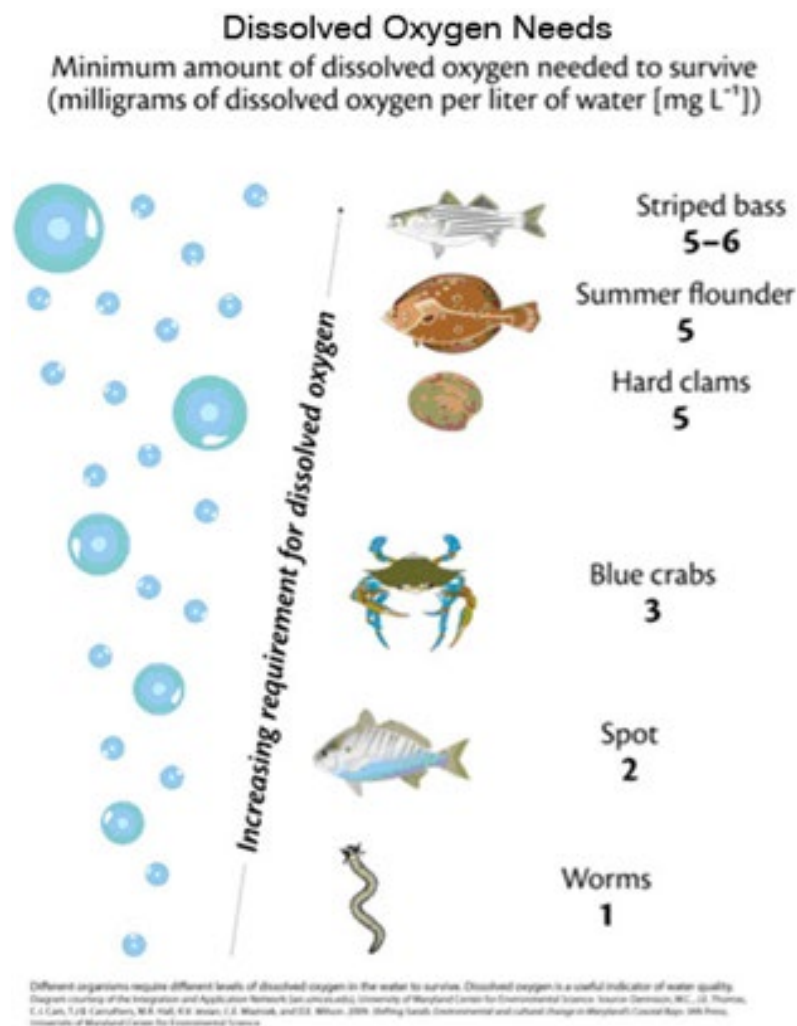


Figure 1. A visualization about the oxygen needs for different sea creatures. Source: “Marine Water Quality”

Student Questions and Answers:

1. Do you find the following visualization to be effective?

Answer: Not a very effective visualization.

2. What problems do you see?

Answer: Problems may include unnecessary air bubbles to the left, not clear summer flounder and hard clams need the same amount of oxygen, and size/width of each creature is different, which may change perception of how much oxygen is needed.

3. Why do you think the air bubbles to the left were included?

4. Was it a good decision to include the air bubbles?

Answer: Not a good decision to include the air bubbles, make the visualization unnecessarily cluttered.

5. What other ways could the same data be represented?

Answer: A bar graph would have been a better choice.

Graphing Over Time

Categories: Science

Yields: Discussion/Analysis, Visualization

Summary: Provide students with data or have them find it themselves. Students choose data points and graph their change over time.

Go to the URI data sets listed in the appendix's "Data Sets and Links" section (#4).

1. Download the Fox Island and Whale Rock 2006 to 2017 data set, an example of which is shown below in figure 1.
2. Choose two or more categories for your students to graph. You may also have them choose their own categories.
3. You will want to only select some of the data due to the amount of data provided. An example would be only 2010 to 2017, a date every 6 months. Dates that have all data available are best to avoid the need to clean the data.
4. Create a new Excel sheet to give to the students.

1	Date	Station	Surf Date	Station	Surface Temperature	Surface Salinity	Surface Depth	Surface Dissolved Oxygen Percent	Surface Dissolved Oxygen Concentration	Bottom Temperature
2	M/D/Y		degr M/D/Y		degrees C	parts per thousand (ppt)	meters	percent (%)	milligrams/liter (mg/L)	degrees C
3	4/3/2006	Fox Island	7.72 4/3/2006	Fox Island	7.72	30.12	0.554	109.8	10.76	7.58
4	4/3/2006	Whale Rock	6.31 4/3/2006	Whale Rock	6.31	31.43	0.507	104.1	10.45	6.06
5	5/11/2006	Whale Rock	11.06 5/11/2006	Whale Rock	11.09	27.7	0.484	102.8	9.49	10.02
6	5/16/2006	Fox Island	11.84 5/16/2006	Fox Island	11.84	26.94	0.014	ND	ND	11.38
7	5/16/2006	Whale Rock	11.16 5/16/2006	Whale Rock	11.16	29.78	0.306	95.9	8.72	10.58
8	5/22/2006	Fox Island	12.96 5/22/2006	Fox Island	12.99	28.77	0.454	ND	ND	12.06
9	5/22/2006	Whale Rock	11.44 5/22/2006	Whale Rock	11.44	30.54	0.339	100.2	9.02	9.03
10	5/29/2006	Fox Island	16.38 5/29/2006	Fox Island	16.38	27.48	0.651	ND	ND	11.52
11	5/29/2006	Whale Rock	12.52 5/29/2006	Whale Rock	12.52	30.29	0.355	105.1	9.26	9.7
12	6/4/2006	Fox Island	17.23 6/4/2006	Fox Island	17.23	25.49	0.171	100.9	8.32	14.14
13	6/16/2006	Fox Island	19.31 6/16/2006	Fox Island	19.31	23.5	0.523	180.1	14.44	14.16
14	6/16/2006	Whale Rock	16.97 6/16/2006	Whale Rock	16.97	28.07	0.347	ND	ND	12.68
15	6/19/2006	Fox Island	18.91 6/19/2006	Fox Island	18.91	26.19	0.204	122.9	9.77	14.81
16	6/28/2006	Fox Island	20.25 6/28/2006	Fox Island	20.25	26.5	0.406	ND	ND	16.54
17	6/28/2006	Whale Rock	18.64 6/28/2006	Whale Rock	18.64	29.21	0.683	103	8.09	16.93
18	6/30/2006	Fox Island	20.46 6/30/2006	Fox Island	20.46	25.95	0.377	ND	ND	18.74
19	6/30/2006	Whale Rock	18.64 6/30/2006	Whale Rock	18.64	29.2	0.421	ND	ND	17.05
20	7/10/2006	Fox Island	21.54 7/10/2006	Fox Island	21.54	24.8	0.513	ND	ND	17.01
21	7/10/2006	Whale Rock	18.36 7/10/2006	Whale Rock	18.36	29.13	0.47	ND	ND	15.93
22	7/17/2006	Fox Island	22.91 7/17/2006	Fox Island	22.91	26.47	0.295	96.1	7.09	19.55
23	7/17/2006	Whale Rock	20.23 7/17/2006	Whale Rock	20.23	29.46	0.209	88.5	6.73	17.03
24	7/24/2006	Fox Island	21.94 7/24/2006	Fox Island	21.94	27.84	0.207	101.9	7.59	20.51
25	7/24/2006	Whale Rock	20.93 7/24/2006	Whale Rock	20.93	30.15	0.397	66.7	4.99	17.13
26	7/31/2006	Fox Island	24.01 7/31/2006	Fox Island	24.01	28.24	0.241	97.3	6.97	20.72
27	7/31/2006	Whale Rock	22.91 7/31/2006	Whale Rock	22.91	29.42	0.431	68.8	4.99	16.24
28	8/7/2006	Fox Island	23.93 8/7/2006	Fox Island	23.93	29.17	0.967	86.3	6.16	19.46
29	8/7/2006	Whale Rock	22.77 8/7/2006	Whale Rock	22.77	30.82	0.518	109.6	7.9	17.87
30	8/14/2006	Fox Island	21.86 8/14/2006	Fox Island	21.86	30.34	0.866	ND	ND	20.87
31	8/14/2006	Whale Rock	19.94 8/14/2006	Whale Rock	19.94	31.79	0.706	ND	ND	17.25
32	8/23/2006	Fox Island	22.06 8/23/2006	Fox Island	22.06	30.42	0.433	ND	ND	21.03
33	8/23/2006	Whale Rock	20.54 8/23/2006	Whale Rock	20.54	31.48	0.715	93.2	6.07	16.98

Figure 1. The Fox Island and Whale Rock raw data.

Student Questions and Answers:

1. What is the best way to graph the data? (Bar chart, line graph, etc.)

Answer: Line graph for time-based data is best.

2. Individually or in small groups have them graph the data. Do they see any trends?
3. If there are trends what are they and why do they think they are there?
4. Compare and contrast the visualizations and have the students discuss the differences and if one type of graph is clearer and easier to read.

Model My Watershed

Categories: Science

Yields: Discussion, Visualizations (digital suggested)

Summary: Students use the watershed app to look at school districts in Rhode Island and visualize different data sets. They can also use the app to model different land conditions.

Going to the watershed app ([link #8](#) in “Data Sets and Links”), you will see a page about Model My Watershed and the link to run it.

1. Run the browser application, click “Get Started,” and select the area you want to look at.
2. In this example, I looked at different school districts in Rhode Island, so I zoomed the map in and chose school districts under the “Select boundary” option.
3. Choosing any school district will bring up two panels and a number of options within each (*fig. 1*).

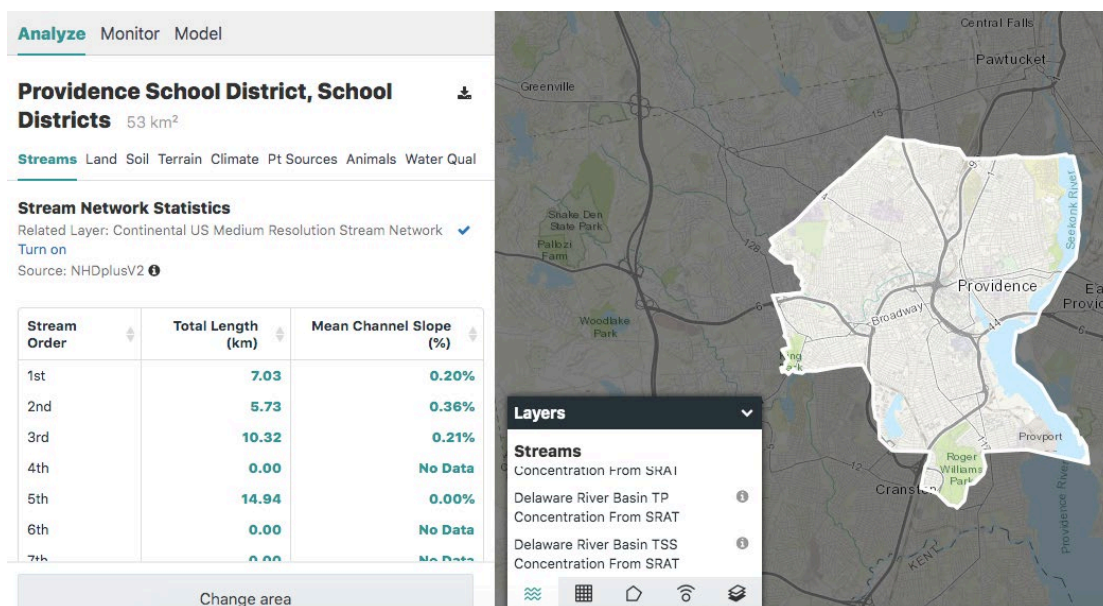


Figure 1. The Providence, RI school district and its data from the Model My Watershed app.

The left panel’s “Analyze” tab lets you look at different types of data and download a CSV file for each one. The “Monitor” tab allows you to query for past or ongoing data collections, and “Model” lets you simulate natural and human-made impacts on the

environment. The “Layers” panel on the right lets you apply different streams, coverage grids, boundaries, observation stations, and basemaps to the map.

I downloaded land data for the North Smithfield, Providence, and Narragansett and compared the percent of land coverage by type. This is a great tool to let students look at their own school district, state or any other area they choose. Below (fig. 2, 3, and 4) are three visualizations I was able to make after downloading the three data sets and combining them into one workbook.

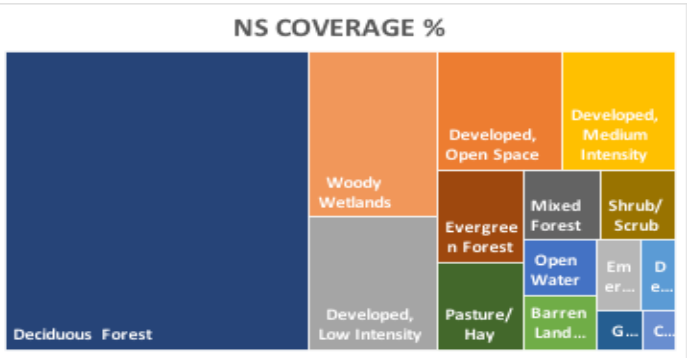
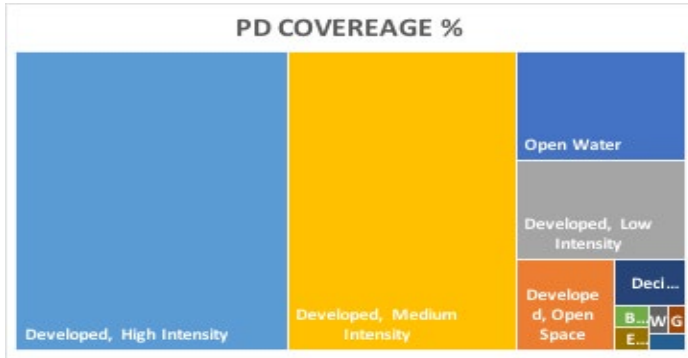


Figure 2. A treemap of Providence schools district's different land types.

Figure 3. A treemap of North Smithfield school district's different land types.

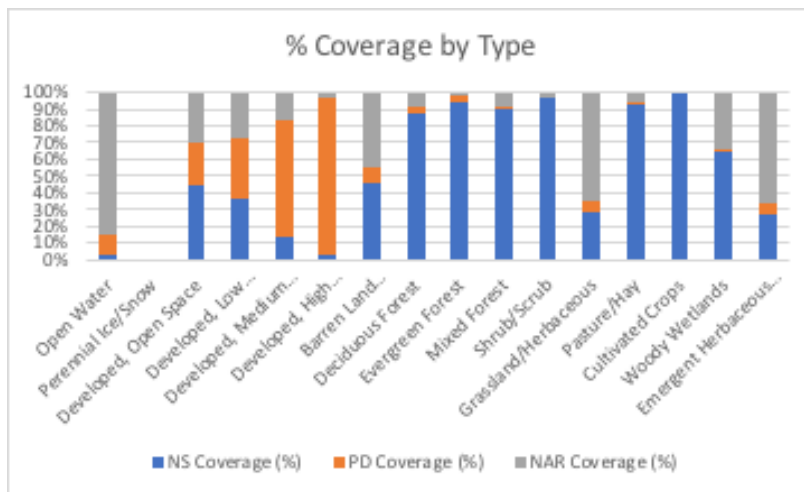


Figure 4. A 100% stacked bar chart comparing how much of each land type the North Smithfield, Providence, and Narragansett school districts have.

Student Questions and Answers:

1. Do the three graphs tell you the same information?

Answer: The top two graphs have the same information, they show the type of land for Providence and North Smithfield, the bottom graph has that information for all three locations.

2. What type of graph do you find easier to read and understand?
3. Are you familiar with the top type of graph (a treemap)?
4. Are you familiar with the bottom type of graph (100% stacked bar chart)? If students are struggling with that chart, tell them that each bar is essentially its own pie chart. It's harder to compare separate circles than bars that are right next to each other.
5. What type of land coverage seems to have the most impact on the amount of runoff?
6. What do you think is the best mix of land coverages that does not result in excess runoff?
7. Which location has the least amount of runoff? Which location has the most amount of runoff?

Navigating NERR Data

Categories: Science

Yields: Visualizations (digital)

Summary: Finding, cleaning, and visualizing data. Special focus on using the “Text to Column” tool to separate the date and timestamp.

The National Estuarine Research Reserve has a wealth of data available for the taking, but it does take a little getting used to. Using [link #9](#) in the “Data Sets and Links” section of the appendix brings you to a page to get data. The “Data Graphing and Export System” is intuitive to use, but the data and metadata gets emailed to you, so make sure students use an email that accepts correspondence from outside their network!

Opening up your CSV file, it will probably look something like figure 1. Reading the legend will explain the column names and what certain cells mean. In this example, the data revolved around air data, so I trimmed the file down to just the essentials: time stamp, air temperature and its data quality level, barometric pressure and its data quality level, and wind speed and its data quality level. **Tip: It’s always best to copy the data to another sheet in the same workbook and edit it there.**

1. Before we visualize anything, we must filter out data marked as questionable.
2. Click the box to the left of column A and above the first row to select everything, and then click “Data” in the top toolbar, and then “Filter.”
3. Every column should now have a filter, so go to each “Flag Code” column and uncheck any code that is below 0 (see *fig. 2*).

*****LEGEND FOR FIELD NAMES LOCATED AT BOTTOM OF FILE*****

Station_Code	ISSWMP	m/d/y	hhmm	See Legend	See Legend	Frequency	See Legend	C	Flag Code	%	Flag Code	mb	
Station_Code	ISSWMP	Date	TimeStr	Historical	Provisional	Frequency	F_Record	A_Temp	F_A_Temp	RH	F_RH	BP	
narpmet	P	1/1/20	0:00	0	1	15		4.6	<0>		80	<0>	10
narpmet	P	1/1/20	0:15	0	1	15		4.7	<0>		80	<0>	10
narpmet	P	1/1/20	0:30	0	1	15		4.6	<0>		81	<0>	10
narpmet	P	1/1/20	0:45	0	1	15		4.4	<0>		82	<0>	10
narpmet	P	1/1/20	1:00	0	1	15		4.4	<0>		82	<0>	10
narpmet	P	1/1/20	1:15	0	1	15		4.3	<0>		82	<0>	10
narpmet	P	1/1/20	1:30	0	1	15		4.3	<0>		82	<0>	10
narpmet	P	1/1/20	1:45	0	1	15		4.4	<0>		81	<0>	10
narpmet	P	1/1/20	2:00	0	1	15		3.9	<0>		75	<0>	10
narpmet	P	1/1/20	2:15	0	1	15		3.7	<0>		72	<0>	10
narpmet	P	1/1/20	2:30	0	1	15		3.6	<0>		69	<0>	10
narpmet	P	1/1/20	2:45	0	1	15		3.4	<0>		68	<0>	10
narpmet	P	1/1/20	3:00	0	1	15		3.3	<0>		67	<0>	10
narpmet	P	1/1/20	3:15	0	1	15		3.2	<0>		68	<0>	10
narpmet	P	1/1/20	3:30	0	1	15		3.1	<0>		69	<0>	10
narpmet	P	1/1/20	3:45	0	1	15		3.1	<0>		69	<0>	10
narpmet	P	1/1/20	4:00	0	1	15		3	<0>		70	<0>	10
narpmet	P	1/1/20	4:15	0	1	15		3	<0>		68	<0>	10
narpmet	P	1/1/20	4:30	0	1	15		3	<0>		69	<0>	10
narpmet	P	1/1/20	4:45	0	1	15		2.9	<0>		69	<0>	10
narpmet	P	1/1/20	5:00	0	1	15		2.9	<0>		68	<0>	10
narpmet	P	1/1/20	5:15	0	1	15		2.9	<0>		67	<0>	10
narpmet	P	1/1/20	5:30	0	1	15		2.8	<0>		67	<0>	10
narpmet	P	1/1/20	5:45	0	1	15		2.7	<0>		67	<0>	10
narpmet	P	1/1/20	6:00	0	1	15		2.7	<0>		67	<0>	10

Figure 1. The raw data from one of the National Estuarine Research Reserve's data sets.

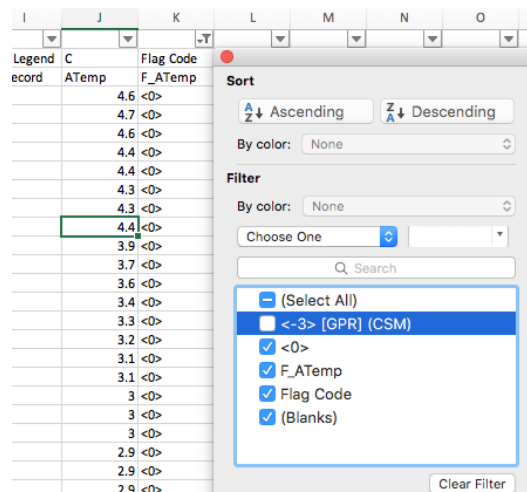


Figure 2. A filter to keep only data with acceptable flag codes.

4. The next problem you have to fix is changing the time stamp column into two separate columns for date and time. In Excel, click on column D, and choose "Insert."
5. Do this again so you have two blank columns next to the time stamp column (C).
6. Click on C so the whole column is selected, then click on "Data" in the top toolbar, and then click "Text to Columns."
7. Choose "Fixed Width" on the first page.
8. Make sure there is only one break (between the date and time) on the second page (fig. 3).
9. Change the column data format to date and make the destination the next column over (in this case D) on the last page (fig. 4).
10. Click "Finish" and you're back in the spreadsheet. Right-click on the column with the dates, choose "Format Cells...," and make sure they are set to "Date."

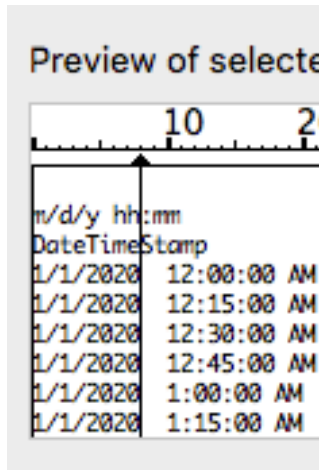


Figure 3. Page 2 of the "Text to Columns" tool. With the fixed width option, you choose where the text is broken up.

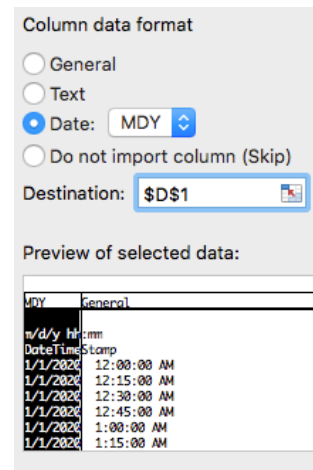


Figure 4. Page 3 of the "Text to Columns" tool. The highlighted column contains dates, so it is formatted as a date. The destination is the next column over.

You are now ready to make visualizations! My two examples (*fig. 5 and 6*) are air temperature over the course of 1 day and air temperature at 12 AM over the course of 7 months.



Figure 5. A line graph of air temperature over the course of a day.

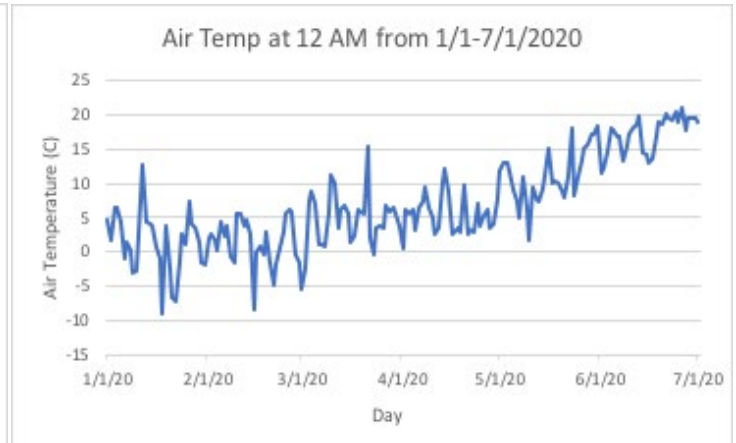


Figure 6. A line graph of air temperature over the course of 6 six months.

Student Questions and Answers:

1. What do you think would happen if you did not clean the data and simply tried to make a visualization from it?

Answer: Without cleaning the data missing data points or other issues with the data set could affect the visualization and result in an incorrect conclusion.

2. What other issues do you think could be in a data set that would affect a visualization?

Answer: Other issues could include missing or unclear column titles, spreadsheet formatting issues, methodology used to collect data changes, date and time mismatch (example: there was no data collection on the day listed), dates/temperatures /times/ other data from different data sets, sensors, or other sources do not match up.

3. What do the graphs show you that you did not see in the raw data?

Combining Data Sets

Categories: Science

Yields: Discussion, Visualization (digital suggested)

Summary: Students hypothesize a correlation, find data from multiple sources, and try to visualize it. Discuss findings and revise hypotheses. For example, this activity tries to find the correlation between volcano eruptions and earthquakes.

Sometimes the answers you need can't be found in one neat data set, which means you have to comb through multiple sets, databases, or even make your own experiments to get all of the necessary data. This activity serves as a starting point for teachers looking for an idea of how they want to present this topic, or it could be used essentially as is for students just beginning to learn about data science.

This activity looks at data from the Smithsonian Global Volcanism Program and the United States Geological Survey.

1. I downloaded data on eruptions in the United States from here: volcano.si.edu/search_eruption.cfm
2. I downloaded the earthquake data from here: <https://earthquake.usgs.gov/earthquakes/search/>
3. I set “Date & Time” to be this past decade.
4. I set “Geographic Region” to Conterminous U.S.
5. And I set “Output Options” to export the first 1000 data points as a CSV file.
6. After looking at both data sets, make a new workbook (*fig. 1*) to combine the two with the five columns you will need: ID, Date, Latitude, Longitude, and Type.
7. Copy and paste the necessary data, but to get the date from the earthquake data you will need to use the “Text to Column” option which is explained in “Navigating NERR Data.”

ID	Year	Latitude	Longitude	Type
Cleveland	2020	52.8	-169.9	Eruption
Semisopochnoi	2019	51.9	179.6	Eruption
Shishaldin	2019	54.8	-164.0	Eruption
Semisopochnoi	2019	51.9	179.6	Eruption
Great Sitkin	2019	52.1	-176.1	Eruption
Semisopochnoi	2018	51.9	179.6	Eruption
Veniaminof	2018	56.2	-159.4	Eruption
Great Sitkin	2018	52.1	-176.1	Eruption
Bogoslof	2016	53.9	-168.0	Eruption
Cleveland	2016	52.8	-169.9	Eruption
Pavlof	2016	55.4	-161.9	Eruption
Pavlof	2014	55.4	-161.9	Eruption
Pavlof	2014	55.4	-161.9	Eruption
Ahyi	2014	20.4	145.0	Eruption
Shishaldin	2014	54.8	-164.0	Eruption
Veniaminof	2013	56.2	-159.4	Eruption
Pavlof	2013	55.4	-161.9	Eruption
Mariana Back-Arc Segment at 15.5°N	2013	15.4	144.5	Eruption
Pagan	2012	18.1	145.8	Eruption
Kanaga	2012	51.9	-177.2	Eruption
Cleveland	2011	52.8	-169.9	Eruption
Pagan	2011	18.1	145.8	Eruption
Cleveland	2010	52.8	-169.9	Eruption
South Sarigan Seamount	2010	16.6	145.8	Eruption
Pagan	2010	18.1	145.8	Eruption
uu60402867	7/21/20	37.5	-113.5	Earthquake
uu60402868	7/21/20	38.4	-113.5	Earthquake

Figure 1. A combined spreadsheet of eruption and earthquake data.

8. The only difference is that you will use the delimited option instead of fixed width, and with "T" as the marker (see *fig. 2*).
9. For Type, you just mark the data as being an eruption or earthquake.

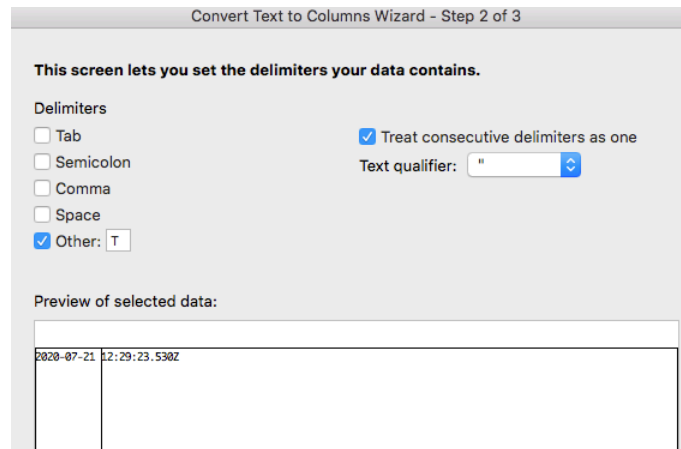


Figure 2. Page 2 in the "Text to Columns" tool with the delimiter option. It will separate each cell at the letter "T."

10. Once you have the workbook ready, go to <https://flourish.studio/> and sign in to make the necessary visualization.
11. Click new visualization, scroll down to the projection map section, and choose the world map option.
12. Go to the "Data" tab and delete all the values in the right panel under Regions (*fig. 3 and 4*).

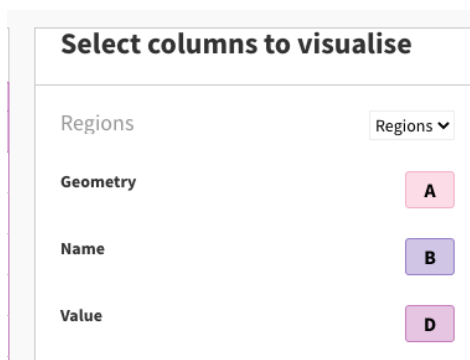


Figure 3. The side panel with data that needs to be changed.

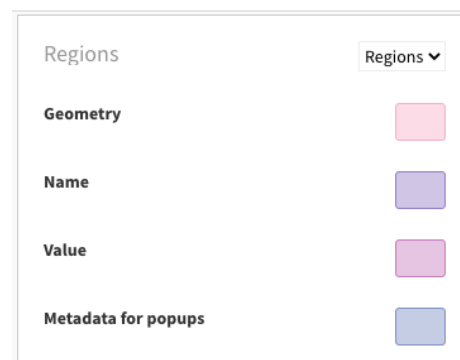


Figure 4. What the side panel should look like.

13. Go to the "Points" tab, and upload your combined workbook.
14. On the right panel, assign columns to the different options. It should look something like figure 5, which would result in the map (see *fig. 6*) also shown below.

Regions		Points			
Q <input type="text"/>					
	A	B	C	D	E
1	ID	Date	Latitude	Longitude	Type
2	Cleveland	2020	52.8	-169.9	Eruption
3	Semisopochnoi	2019	51.9	179.6	Eruption
4	Shishaldin	2019	54.8	-164.0	Eruption
5	Semisopochnoi	2019	51.9	179.6	Eruption
6	Great Sitkin	2019	52.1	-176.1	Eruption

Points

Name A

Longitude D

Latitude C

Colour category E

Figure 5. What the "Points" data tab and side panel should look like. The columns are ordered as ID, Date, Latitude, Longitude, and Type. The side panel is ordered as A, D, C, E, and B (not shown).

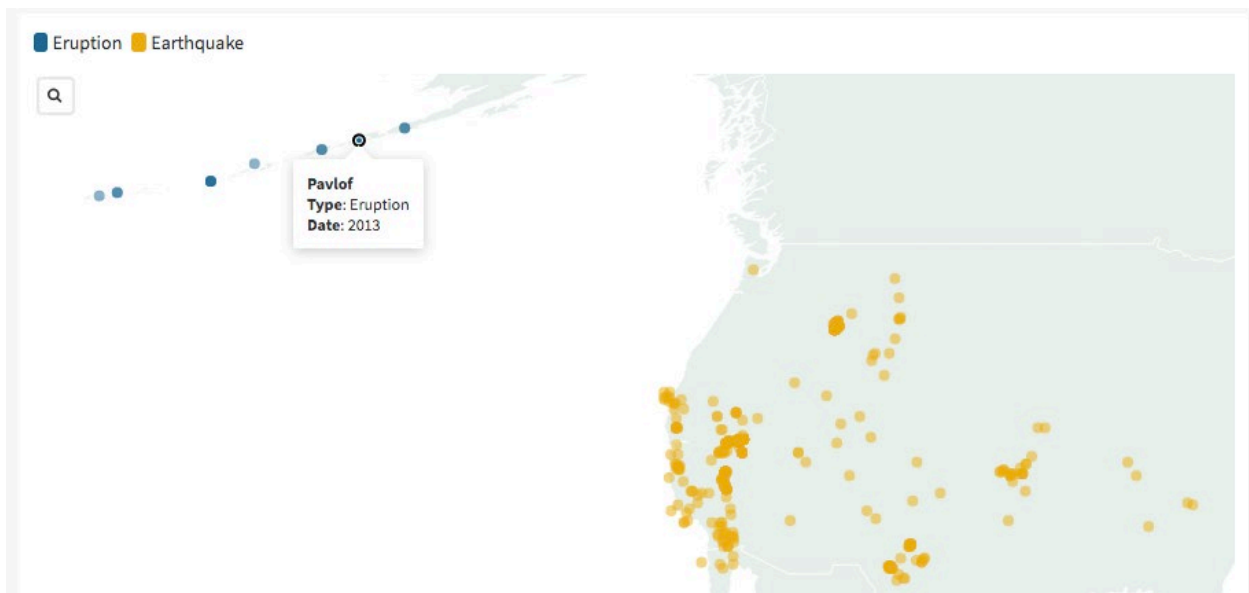


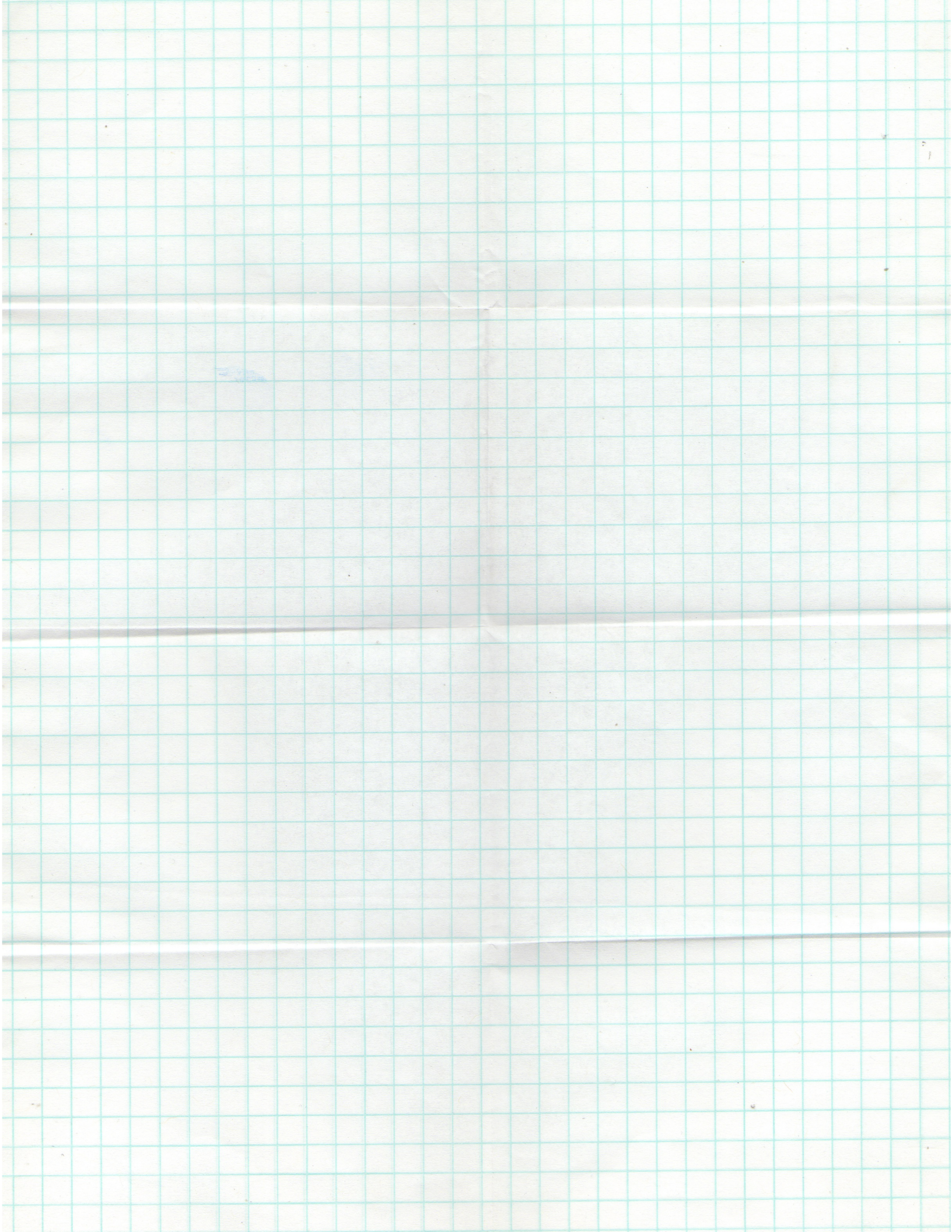
Figure 6. A snapshot of the map or eruptions in blue and earthquakes in yellow. The popup show the ID, Type, and Date.

Student Questions and Answers:

1. As you can see, there are a few problems you can discuss in class. There are more eruptions in the data set, but since they are from the same volcano they don't show up on the map. How can you improve this visualization to display that information?

Answer: (Not the only answers) Change it to a bubble map to indicate size of eruptions and earthquakes, or add labels for each eruption.

2. In terms of tectonic plates, the Alaska-Aleutian Megathrust and the Juan de Fuca are both subduction zones, but why is only the former resulting in volcanic activity?
3. Are volcanoes in the Cascade Range dormant, or are they gearing up for another period of activity? These are all questions you can discuss or have students research. Looking at data can be a starting point for discussion, not just the end result!



Social Studies

National Geographic's Marshallese Stick Charts

Categories: Earth Science, Oceanography, Geography, Physical Geography, Social Studies, World History

Yields: Discussion, Visualization (optional)

Summary: Using National Geographic's activity to learn about visualizing your surroundings in unconventional ways.

This activity gives students a quick look at how some people in the South Pacific Ocean used to navigate the waters around their homes. National Geographic has a great explanation of these stick charts and a few questions to make sure students understand them.

National Geographic: www.nationalgeographic.org/media/micronesian-stick-chart/

Follow-up: You can have students make stick charts of their own. They can use whatever materials are usually around them, just like the Polynesians. The chart can be a map of whatever they want.

For example, a student making a map about "traffic" flow in a particular hallway at school could use pencils of different colors and lengths to show where people walk and where.

As another example, a student could make a map about where they usually spend time in their house. They could use straws as the frame since those are in most people's home, and different objects could represent points of interest. If they spend a lot of time in their kitchen, they could attach little pictures of food to the chart.

Anything works! Let students get creative showing patterns they see.

Student Questions:

1. What location and activity did you choose to make a map of?
2. Why did you choose that location?
3. What material did you make your map out of and why did you choose that material?
4. What did the other students make a map of and what materials did they use?
5. Do you see common patterns in the what your classmates mapped and what materials they used?

Voting: Then and Now

Categories: Social Studies, World History, Political Science

Yields: Discussion, Visualization (optional)

Summary: Using National Geographic's article about voting to show students the difference between democracies. Opportunity to talk about the state of democracy today.

Voting is an essential part of our country, and the idea of citizens having a voice extends all the way back to the 5th century B.C.E. The difference between Ancient Greece and modern America is the type of democracy – we have a representative democracy while they had a direct one. Read National Geographic's article about democracy, which can be found at www.nationalgeographic.org/encyclopedia/democracy-ancient-greece/.

1. After reading, create a ballot box. For example, cut a hole in the side of a shoe box.
2. Cover the hole by gluing one side of a tissue on the inside of the box.
3. Use 2 tokens to differentiate between “yes” and “no” votes. You could use black and white checker pieces, for example.
4. Check the votes by taking the lid off the box and counting.
5. You can count the law as passing any way you choose, so long as it's the same between direct and representative voting. For example, it could be simple majority or two-thirds majority.

1. For the direct democracy, randomly choose 25% of students to act as lawmakers.
2. That group will make a new law, propose a revised law, or present an already existing law in the United States.
3. All students must vote on the proposal using their tokens. If a student doesn't vote, a red “X” should be placed on their desk.

1. For the representative democracy, have students vote for representatives (again 25% of the class).
2. The representatives will vote on the same law presented by the direct democracy group.

Follow-up: Have students make visualizations of the results to see the difference, if any, more clearly. For example, in a class of 25 students, I made 100% stacked bar charts of data I made up (*fig. 1*).

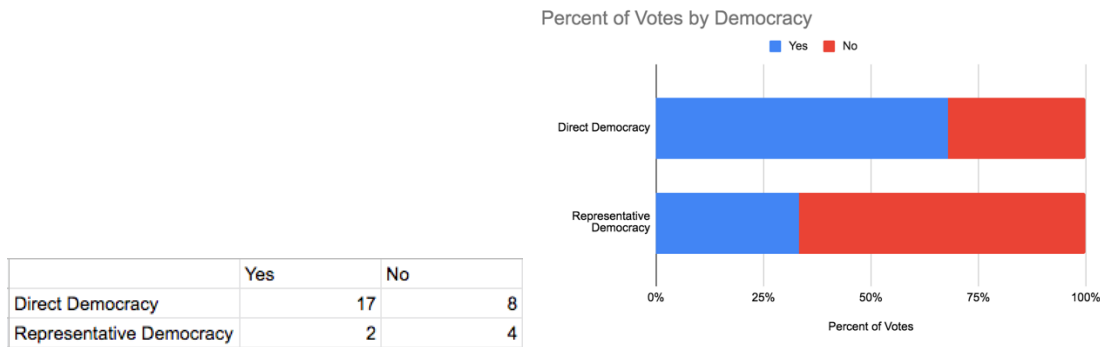


Figure 1. A table of votes and how it looks as a 100% stacked bar chart.

Student Questions:

1. Compare the voting results. Did the same law pass/fail between the democracies?
2. After completing the exercise on the direct voting vs representative voting what do you think are the pros and cons of each system?
3. How does gerrymandering impact the results in a representative democracy?
4. Also consider the Australian system for people who do not vote. Do you think that could work in America?

Answer: (Elaboration of the question) Australia has compulsory voting in state general elections, and failure to vote has different consequences. If your reason to not vote is valid, there is no consequence. An invalid reason results in a fine. Not giving a reason or not paying the fine means you may have your driver's license removed.

5. Discuss how it could be argued that your government and who is in power is a form of data visualization. Your vote is the raw data, and the resulting makeup of Legislative and Executive branches is the visualization.
6. Is this visualization always accurate to the raw data? For example, there have been 5 times where a candidate had the popular vote but lost the election because of the electoral college.
7. How do the systems in place work to represent, or misrepresent, the data given to it?

Torres Strait Tallies – A Chain of Achievements

Categories: Social Studies, World History, Health

Yields: Discussion, Visualization (physical or digital)

Summary: Students visualize something about themselves that they are proud of.

In the 19th century, Alfred Cort Haddon discovered that the Torres Strait Islanders used physical representations for social purposes. A “kupe” was a bundle of sticks tied to a string (see *fig. 1*) and could be used to tally a specific achievement, such as the number of turtles they caught.



Figure 1. A drawing of a kupe from the Torres Strait Islanders. Source: Dragicevic

Have students come up with a list of things they have done more than once and are proud of. For example, it could be the number of good deeds they do in a week, the number of times they participated in/won a competition, how many languages they know. Their kupe will represent one of these list items.

Next, have them design a kupe of their own. It can have a clear representation or one only they can decipher, similar to the Marshallese stick charts. For example, a digital kupe of good deeds could be a PowerPoint slideshow with each slide representing one deed. The picture on each slide could represent what they did, such as a rake if they raked leaves. For a personal kupe, students could use color-coded popsicle sticks tied together with string.

If students want to, they can discuss what they’re proud of and how their kupe represents it.

Student Questions:

1. What did you make your visualization out of?
2. Why did you choose that material?
3. What accomplishment did you visualize?
4. Why is that accomplishment important to you?

Mapping America

Categories: Geography, Social Studies, Political Science

Yields: Discussion, Visualization (digital)

Summary: Students explore the U.S. Census Bureau to find data categorized by state (e.g. populations) and then make a heat map.

Geographic heat maps are a great way to visualize a range of data that can be categorized by state or country. The U.S Census Bureau has years' worth of data that you can download to look at in Excel or Google Sheets, look at in tables, and view as geographic heat maps they created. This activity will walk you through the process of making your own.

1. Let's make a map of population by state based off of data from 2019. To begin, go to the Census Bureau website, find the search bar and type "population by state".
2. On the results page, scroll down until you find the "State Population Totals: 2010-2019" link (*fig. 1*).

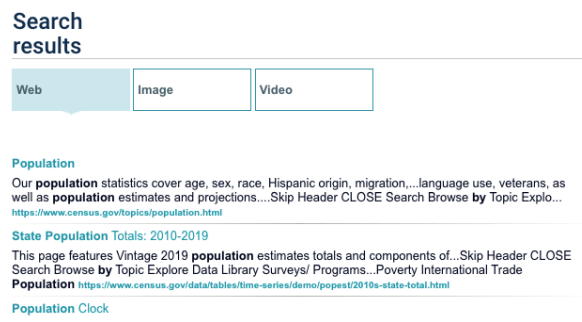


Figure 1. The results from searching "population by state" in the U.S. Census Bureau.

3. On the new page, scroll down until you find the "Datasets" section, and click on the Excel file link to download the data as a comma separated value (.csv) file (*fig. 2*).

Datasets

Population, Population Change, and Estimated Components of Population Change: April 1, 2010 to July 1, 2019 (NST-EST2019-alldata)

 File Layout [< 1.0 MB]

 Population, Population Change, and Estimated Components of Population Change: April 1, 2010 to July 1, 2019 (NST-EST2019-alldata) [<1.0 MB]

Figure 2. The Excel file link for population data.

4. Open up your Google Drive, and upload the file by clicking on New --> File upload, or by dragging the file into your browser window.
5. After that, create a new Google Sheets file.
6. In the top left corner of your new sheet click File --> Import. It will bring up your Google Drive, and you can click on the data set you just uploaded.

7. Selecting the set will bring up a list of options, most of which can be left alone. The only one that matters is if you want the data used on the blank Sheet 1 you have, or to create a Sheet 2 and use it there.
8. There will be a lot of data! Don't worry, we will only be focusing on 2 columns: NAME and POPESTIMATE2019.

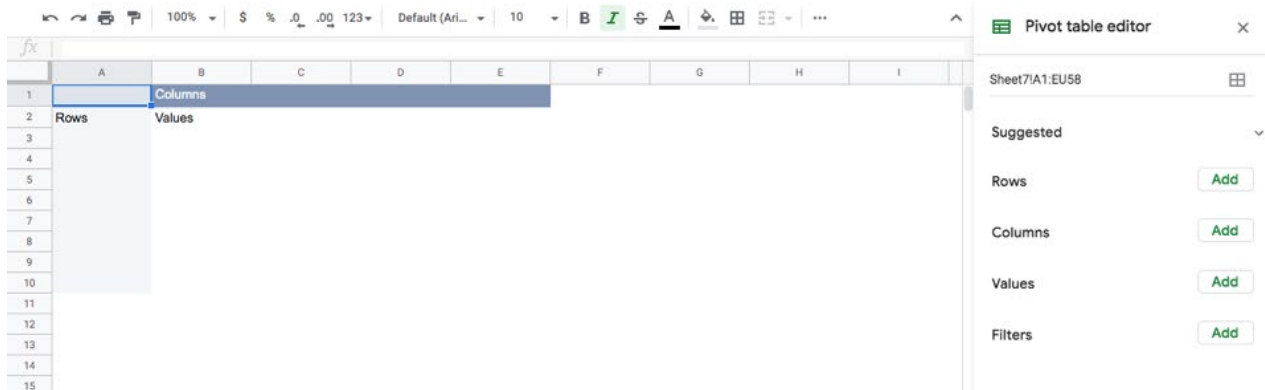


Figure 3. A generic pivot table with no customization.

9. Without anything selected, go to the top toolbar and click on Data --> Pivot Table --> Create (you don't have to worry about the options). Your screen should now look like figure 3.
10. Since we want to display 2019 populations, click on Values' Add --> POPESTIMATE2019.
11. Next, since we want those values organized by state, click on Rows' Add --> NAME.
12. Finally, since we don't want regions or national sums, click on Filters' Add --> NAME --> Showing all items.
13. Scroll through the list, or search, for Midwest Region, Northeast Region, Puerto Rico (the Google chart doesn't include it so we can't leave it to skew data), South Region, United States, and West Region. Uncheck them all, and you're now ready to make the map!
14. Highlight all the values from A1 to B52, go to the top toolbar and click Insert --> Chart.
15. In the Chart Editor, click "Chart Type" and scroll down until you see the "Map" section.
16. Click on the regular "Geo chart" option.
17. To change from a world view, click the Chart Editor's Customize --> Geo --> Region --> United States. Your chart should now look like figure 4.

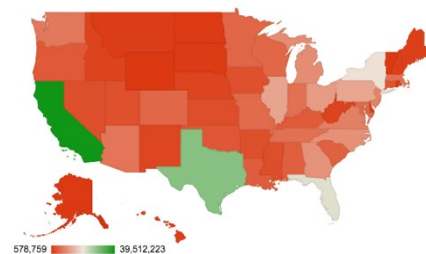


Figure 4. A heat map with the correct data, but no color correction.

18. You should also see color options for Min, Mid, Max, and No Value. By changing those, and generally adhering to the rule that colors should be a range of one shade, I made mine to look like figure 5. I used a Min of light pink, a Mid of light purple, and a Max of dark purple. I feel the pink helps differentiate the low and medium values, but feel free to experiment with your own charts! For more information, check you this YouTube video at www.youtube.com/watch?v=fTt6EjiXI2M.

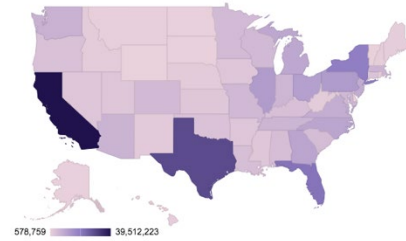


Figure 5. The same heat map as figure 4 but with customized colors.

Student Questions and Answers:

1. Looking at the maps everyone has made do some colors seem more effective at displaying the same data than others?
2. What other information would you want to include in a map of estimated population?

Answer: Information you may also want to include is demographics, percentage of population each state is of the whole US, city names, estimated number of people.

3. How would you make a map showing population change from 2010 to 2019?

Land of the Favored

Categories: Social Studies

Yields: Discussion

Summary: Students can explore the Pew Research Center, and are asked to answer some questions about a particular poll.

Pew Research has many interactive data visualization tools that are available for free to play with. One of the interesting ones is a global map of questions asked to those across the world. Amongst those questions is “Please tell me if you have a very favorable, somewhat favorable, somewhat unfavorable or very unfavorable opinion of the United States.”

The results can be found here at

<https://www.pewresearch.org/global/database/indicator/1>

Student Questions and Answers: Pick a year, any year, and answer the following questions:

1. Are there any years that have the same percentage as an answer?
Answer: Yes, multiple years.
2. What is the country that favors the US with the highest percentage?
3. Why do you think that country feels that way?

Make sure that you can also put the list in unfavorable, answer the following:

1. Who favors the US the least (in percentage)?
2. Are the countries with the highest confidence in the president and who favor the US the most the same in your selected year? What year is that?
3. Why do you think that is?

Miscellaneous

What Should I Be Eating?

Categories: Health

Yields: Discussion, Visualization (physical or digital)

Summary: Students look at the USDA's visualizations of a healthy diet and discuss the pros and cons of each. They then create their own visualization or improve an existing one.

Understanding what you should be eating at any given meal is always a little confusing. To make things easier, we have grown up looking at different visualizations of a healthy diet – the two most iconic being the pyramid and plate (see *fig. 1 and 2*). Each one has its pros and cons, and today we're going to discuss them. Bring up these images, or something like them, and present them to students.



Figure 1. The most recent diet visualization from the USDA. Source: "A Brief History of USDA Food Guides"

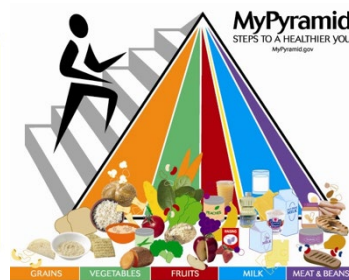


Figure 2. The second most recent diet visualization from the USDA. Source: "A Brief History of USDA Food Guides"

Tell students about the statistics of what your diet should actually look like. This website, www.center4research.org/myplate-new-alternative-food-pyramid/, details the facts the visualizations should be showing. Some points include getting exercise, varying protein sources, and making half of your grains whole grains. Discuss if either of the visualizations match the facts of a healthy diet.

Have students work in groups to come up with a diet visualization of their own. It can be an infographic with multiple data visualizations and facts, or one visualization focusing on a single message. Have students present their work explain their choices.

To help students get inspiration for their designs, have them look at www.choosemyplate.gov/eathealthy/brief-history-usda-food-guides. It shows the USDA's old food guides (see *fig. 3*) and explains the drawbacks that led to new and improved ones. Some of them are more of an infographic, some try to use the common chart styles we know about today (mainly pie charts), and some create new visualization styles (the modern pyramid and plate).

Ask the students what elements they like and do not like about the older food guides, as well as why they think the older guides were changed.



Figure 3. Three of the older food guides that are no longer in use. Source: "A Brief History of USDA Food Guides"

Student Questions and Answers:

1. What does each visualization tell you about your diet?
2. Which one is easier to understand in terms of real-life application?
3. Do either of them look similar to common forms of data representation (bar charts, line graphs, pie charts, etc.)?

Answer: The plate looks like a pie chart.

4. How effective do you find the following two visualizations about the USDA food guide?
5. Would you change anything about either of them?
6. What would you change and why?
7. Do either of the visualizations match the facts of a healthy diet?
8. What elements do you like and not like about the older food guides?
9. Can you see from the older food guides why the guides were changed?

What's So Good About It?

Categories: Business

Yields: Discussion, Visualization (physical or digital)

Summary: Students read an article about toys being data visualization and discuss their thoughts. Students create toy packaging based off of this idea, and can optionally look at toy sales over time and try to correlate it with rebranding.

Packaging is one of the most important aspects of a product. It conveys information to consumers about the product's abilities, the brand's stance on certain issues, or how great the product will make your life. All of this information, or data, has to be represented well in order to make a sale; that is exactly Michael Sandberg talked about in his "Datavis as Art: Are Toys Data Visualizations?" article. Taking a look through his blog post ([link #6](#) in the "Links to more information about activities" section in the appendix), you'll see he talks about the DIKW Pyramid (*fig. 1*) and how it relates to processing the information given on a toy's packaging.



Figure 1. The Data Information Knowledge Wisdom (DIKW) pyramid. It shows the hierarchy of knowledge processing in people. Source: Ellerton

Student Questions and Answers:

1. As a group, look at different toys and discuss what each one tells you in terms of data or information, and how that can become actionable decisions.

Answer: Toys may tell you about cost, batteries needed/included, intended age group, features/attributes of the toy, name of the toy, what other users thought of the toy/reviews, and who made the toy.

2. Do different types of toys emphasize different data?

Answer: Yes, different types of toys emphasize different data.

3. Without looking at any of the words can you still tell what the toy is and what age group it is aimed at? (Example, video game rated E versus one rated M)

After this, have students work independently to design packaging for a new toy they create, or one that already exists. When everyone is done, students can work in small groups to discuss each other's creations. Did the students convey the same information in their packaging as in the professionally made packaging? What visual strategies and tactics did they use to convince people to buy their toy?

Follow-up: You can also have students look at how different packaging correlates to sales over time. Be sure to emphasize that correlation is not causation! As an example, let's look at Barbie.

From the bar chart found at www.statista.com/statistics/370361/gross-sales-of-mattel-s-barbie-brand/, it shows that prior to 2015 Barbie's sales were declining. After 2015, though, the company started diversifying the Fashionista line to appeal to more people according to the article found at www.cnn.com/style/article/barbie-vitiligo-hairless-dolls-wellness-scli-intl/index.html. Not only is the product now appealing to people, but the packaging must convey the new outlook on giving kids a toy that looks like them. A quick Google search of "barbie wheelchair doll" yields results showing the packaging for different women in wheelchairs being prominently displayed, the Barbie logo, and a small infographic about the chair itself (fig. 2 and 3).



Figure 2. One of the new products from the Barbie Fashionistas line. Source: "Mattel Barbie and Wheelchair"



Figure 3. Another product from the Barbie Fashionistas line. Source: Konecky

Self-Portrait in Data

Categories: Art, Health

Yields: Discussion, Visualization (physical or digital)

Summary: Students look at examples of self-portraits based on data and then make their own self-portrait with data they generate.

If you ever find yourself doing a unit on self-portraits, try adding the option of letting students use data to define themselves! Take a look at the slideshow presentation “Art as Information, Information as Art” by Loren Madsen, a data artist. It can be found here at www.slideshare.net/lorenmadsen/information-as-art. The self-portrait part starts as slide 55.

After looking at how some people define themselves (as a years’ worth of receipts, the TV guide on the day they were born, different patterns of movement, etc.), how does it relate to conventional portraits you’ve seen before? Do feel these ones convey more, or less, information about a person than regular self-portraits?

If students take the option of visualizing themselves as data, but aren’t sure what to do, give them some examples. It could be the number of hours they played for different video games, the amount of different foods they consume or products they buy in a week, or what their daily schedule looks like. A simple chart could work, or they could make an infographic with multiple points of data. Infogram is one free source they could use to create a visualization with multiple data sets and some text.

Student Questions:

1. What data did you chose to make your self-portrait?
2. Do you think it shows more or less than a standard self-portrait?
3. What information did you choose to make your self-portrait?

Taking a Look at Treemaps

Categories: Business, Personal Finance

Yields: Visualization (digital)

Summary: Students find, clean, and visualize hierarchical data through treemaps. Students can explore different graphing tools and their different capabilities.

Being able to properly infer information about spending habits and then make data-driven decisions is important for anyone trying to balance their budget. It is also important in a work environment if you're tasked with figuring out their expenditures, sales, or other hierarchical information. Treemaps are perfect for displaying this type of data because categories can be nested inside of their parents. For example, say you have a category for food spending. You can have leaf nodes to show your two sub-categories for eating out and groceries. You can choose to show the nested categories, or just their sum/parent node (food spending). Figure 1 is an example of a visualization you might have to make in the future. We will be finding data and making something similar with Flourish and Infogram.

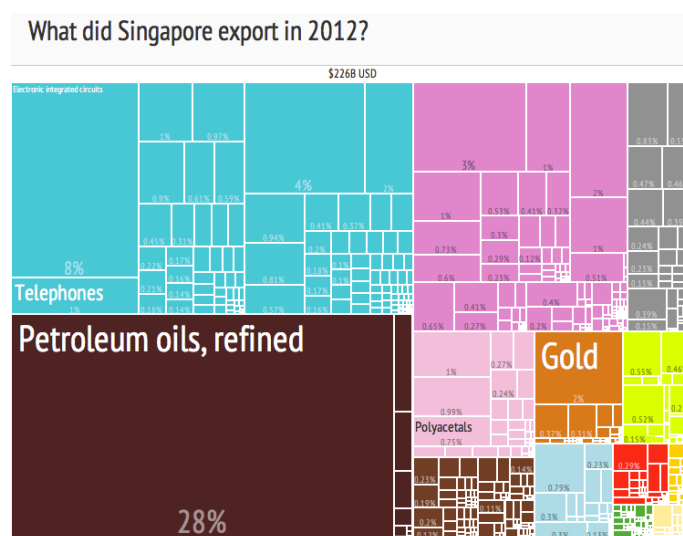


Figure 1. A treemap of Singapore's exports in 2012. Source: Silvermanaz

1. Go to wits.worldbank.org, hover over “Trade Stats” in the top left corner, and click “By Country” (fig. 2). I chose to look at North America, but students can look at whichever country they want and compare afterwards.
2. In the new screen, choose the second link (Trade Summary) in the “Quick Links” panel on the left. You will see a basic treemap that students should aim to replicate but style in their own way.
3. On the top right, click “Download” to generate a .csv file.
4. In the file, there will be data for several years but since we are just focusing on the 2018 data, you can hide the columns or delete the data.
5. Space out the columns so you can read what is in each cell, and then you can start filtering.
6. I wanted export data, so I added a filter to column D (Indicator Type) (fig. 3).
7. After this, the data used in the website’s treemap emerges (see fig. 4).



Figure 2. The World Integrated Trade Solution (WITS) website. The “Trade Stats” navigation menu is shown.

Partner	Product categories	Indicator Type	Indicator	2018
...	...	Export		
...	...	Export		
Canada	All Products	Export		
China	All Products	Export		
Germany	All Products	Export		
Japan	All Products	Export		
Korea, Rep.	All Products	Export		
Mexico	All Products	Export		
United Kingdom	All Products	Export		
United States	All Products	Export		
World	Animal	Export		
World	Capital goods	Export		
World	Chemicals	Export		
World	Consumer goods	Export		
World	Food Products	Export		
World	Footwear	Export		
World	Fuels	Export		
World	Hides and Skins	Export		
World	Intermediate goods	Export		
World	Mach and Elec	Export		
World	Metals	Export		
World	Minerals	Export		
World	Miscellaneous	Export		
World	Plastic or Rubber	Export		
World	Raw materials	Export		

Indicator Type	
Sort	
A Z ↓ Ascending	Z A ↓ Descending
By color:	None
Filter	
By color:	None
Equals	Export
Choose One	
<input checked="" type="checkbox"/> (Select All) <input type="checkbox"/> Development <input checked="" type="checkbox"/> Export <input type="checkbox"/> Import <input type="checkbox"/> Tariff <input type="checkbox"/> Trade Indicator	

Figure 3. The filter applied to the “Indicator Type” column. You can either filter by deselecting everything except “Export,” or you can filter by only allowing data that equals “Export.”

24	United States	All Products	Export	Trade (US\$ Mil)-Top 5 Export Partner	337828.64
27	World	Animal	Export	Export(US\$ Mil)	41077.09
28	World	Capital goods	Export	Export(US\$ Mil)	620831.69
29	World	Chemicals	Export	Export(US\$ Mil)	199201.63
30	World	Consumer goods	Export	Export(US\$ Mil)	551352.81
31	World	Food Products	Export	Export(US\$ Mil)	61564.69
32	World	Footwear	Export	Export(US\$ Mil)	2459.58
33	World	Fuels	Export	Export(US\$ Mil)	291995.68
34	World	Hides and Skins	Export	Export(US\$ Mil)	4893.82
35	World	Intermediate goods	Export	Export(US\$ Mil)	434106.48
36	World	Mach and Elec	Export	Export(US\$ Mil)	424267.31
37	World	Metals	Export	Export(US\$ Mil)	108613.89
38	World	Minerals	Export	Export(US\$ Mil)	19985.41
39	World	Miscellaneous	Export	Export(US\$ Mil)	224973.28
40	World	Plastic or Rubber	Export	Export(US\$ Mil)	97935.04
41	World	Raw materials	Export	Export(US\$ Mil)	309453.97
42	World	Stone and Glass	Export	Export(US\$ Mil)	93884.81
43	World	Textiles and Clothing	Export	Export(US\$ Mil)	30480.39
44	World	Transportation	Export	Export(US\$ Mil)	347745.47
45	World	Vegetable	Export	Export(US\$ Mil)	94364.52
46	World	Wood	Export	Export(US\$ Mil)	72160.94
47	World	All Products	Export	Exports (in US\$ Mil)	2115603.55
50	Goods	All Products	Export	Partner share(%) Top 5 Export Partner	14.17

Figure 4. The export data that will be used in the treemap. It has the product type and amount of money it made.

8. It will be easiest to save this highlighted data into a new workbook along with the column names. If you also want to include the exports' product share percentage, scroll down and copy that data into the new workbook.
9. Going to flourish.studio/, sign in with a Google account and choose the treemap option from the template list.
10. You will see an already made treemap, and going into the "Data" tab, you will see how they did it.
11. The panel on the right describes the different nesting, sizing, filtering, and popup options you can utilize.
12. Upload your new workbook, and then we can condense everything down into 3 columns (Product categories, 2018 Export (US\$ Mil), and Export Product share (%)).

13. In the panel on the right, nest with column A, size by column B, and add C to the popup info (fig. 5). Your visualization should look like figure 6.

	A	B	C
1	Product categories	2018 Export (US\$ Mil)	Export Product share (%)
2	Animal	41077.09	1.94
3	Capital goods	620831.69	29.35
4	Chemicals	199201.63	9.42
5	Consumer goods	551352.81	26.06
6	Food Products	61564.69	2.91

Figure 5. The correct "Data" tab. The table has an extra column for each product's share percentage. The right panel shows how "Size by" and "Info for popups" should be assigned.

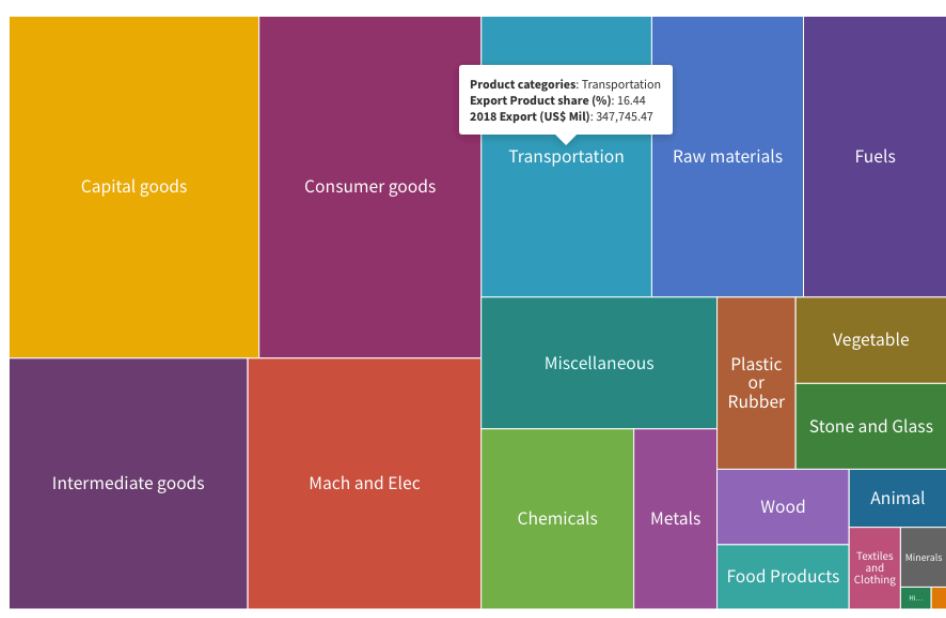


Figure 6. The resulting treemap and popup information based on the data and settings of Figure 5.

Putting the same data from Flourish into Infogram yields a slightly different visualization. Instead of a popup with the share percentage, it just makes the same treemap in a different tab (see fig. 8). While these visualizations are good, they still don't show the nesting capabilities that treemaps specialize in, so I will add some more generalized categories in Flourish to sort this data.

14. Taking a quick look at the product categories, I added a column to the left called "Category" and made "product categories" into "Product."

15. My categories were Fuel, Goods, Organic, and Other (not the best but this is just for demonstrating purposes). I organized the products and the treemap came out to look like figure 7.
16. The popups change to only show the lowest category information. Also, the nesting feature also does not work on Infogram.

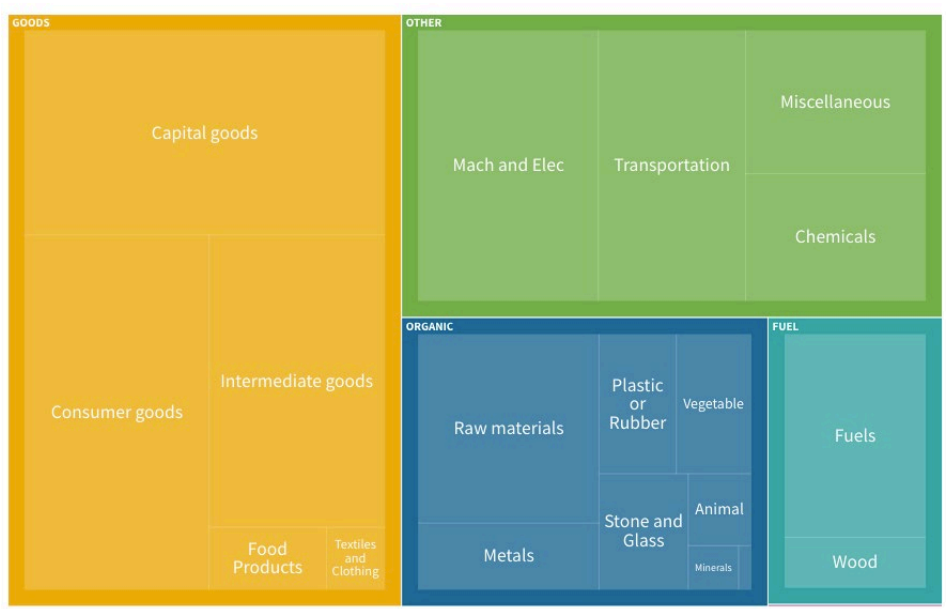


Figure 7. A treemap where the data has been categorized into broader categories. The visualization shows the hierarchical nature of treemaps.

Feel free to let students use whatever data they can find. If you want to take more of a personal finance approach to this activity, you could have students create a spreadsheet to track their spending by categories and sub-categories (e.g. an entertainment category and different streaming services as sub-categories). Either way, remember to add appropriate titles and metadata (where you got the data or how it was collected) (fig. 9)!

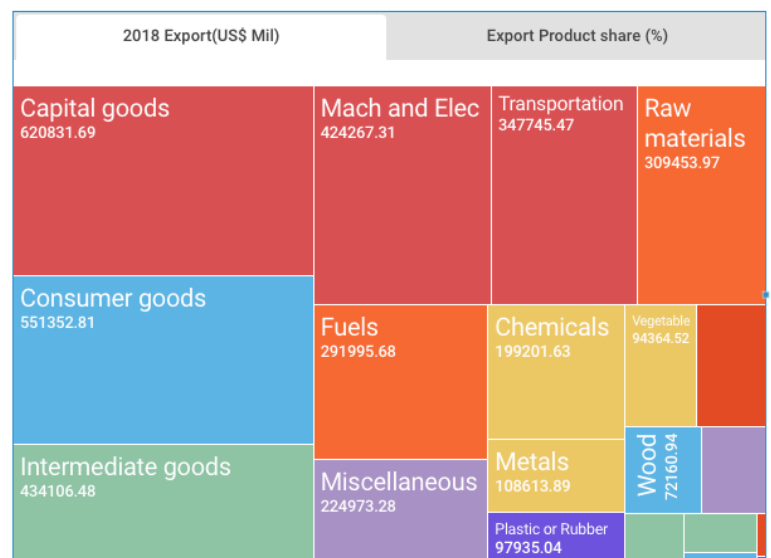


Figure 7. A treemap with the same data as Figure 5, but made in Infogram instead of Flourish. The share percentage column becomes a separate treemap instead of being used for popup information.

North America 2018 Exports
Data from the World Integrated Trade Solution



Figure 8. An example of a title and explanation of where the data came from for this treemap.

Student Questions:

1. Do you see anything in the treemap that you did not see when you were looking at the data.
2. Do you find treemaps to be useful, what other types of data could treemaps be used for?
3. Have you seen treemaps used before now?

Wearable Data

Categories: Art, Design, Home Economics

Yields: Discussion, Clothes

Summary: Students find data and make a physical visualization, preferably through clothes. Discussion on Laura Perovich's clothing line and Seung Lee's sleep blanket.

The goal of this activity is not to make clothes like the ones in figure 1, but instead something much simpler. You can convey data meaningfully with just a scarf, for example. So, we will be making a temperature scarf here to introduce the idea of data that you wear.



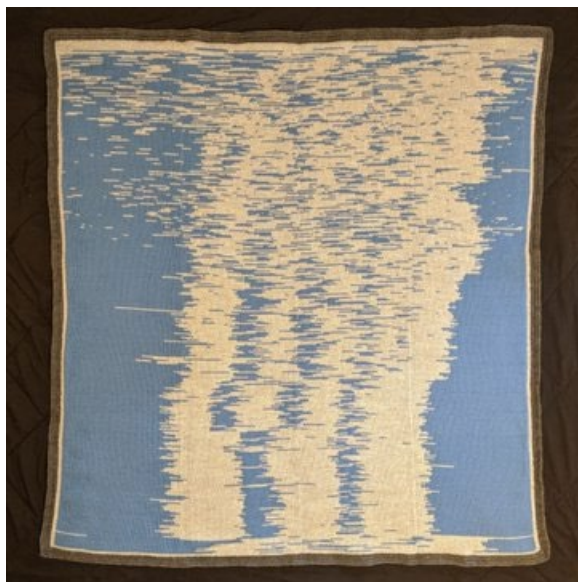
Figure 1. Laura Perovich's collection of clothes visualizing the chemicals found in different homes. Source: Perovich

1. Temperature data can be found here at www.timeanddate.com/weather/usa/providence/historic?month=1&year=2019. I set the location to Providence, but you can change that. You can also change what date you want to look at.
2. To start the project, choose a range of time you want the scarf to cover. It could be 1 day, 1 month, 3 months, or anything you think would work.
3. Next, choose a ratio of scarf rows to temperature point. For example, if doing 1 day and there are 26 recorded data points, you could have a ratio of 3 rows to 1 point. This would give you 78 rows of yarn for your scarf.
4. Look at the range of temperatures you have and create color codes for specific subsets of that range.
5. For example, looking at January 1st, 2019, the range of temperatures is 34-58. Since the difference is 24, you could have 6 colors to represent subsets of 4 (34-38, 39-42, etc.) or 4 colors to represent subsets of 6. To show change over time, using 6 colors

might be better with such a small range. Keep in mind that adding more colors to your scarf does mean buying more yarn, though.

6. Some color codes would be what you see on a regular weather map. Dark blue or purple for extremely cold or freezing, lighter blues for cooler temperatures, and yellows, oranges, and reds for warmer temperatures.

If you would like to see more of the designs shown at the beginning of this activity, you can check them out here at www.lauraperoovich.com/projects/dressedindata.html.



If you're more interested in the Home Economics side, you could make pieces for a child based on their data like Seung Lee did. He made a blanket out of his son's sleeping pattern over his first year of life (*fig. 2*). His Twitter thread describing the process can be found here at

<https://twitter.com/Lagomorpha/status/1149754592579600384>. If that link doesn't work, you can also look up "Seung Lee child blanket."

Figure 2. Seung Lee's sleep blanket. He tracked his son's sleeping pattern and made a blanket out of the data. Source: @Lagomorpha

Student Questions:

1. What data (personal data, public data) would you chose to make a wearable visualization?
2. Why did you choose the data set?
3. Is there a reason behind the colors you chose? If yes what is the reason?

Appendix

Data Sets and Links

1. The 2020 Data Visualization Best Practices Workshop and resources.
 - a. https://drive.google.com/drive/folders/1uzdPAscZDi77AazK0Lgwkl-NxZO_o9pm?usp=sharing
2. Article about visualizations and ethics, had a good image of the Challenger Space shuttle O – ring image.
 - a. https://www.adam-p-jones.com/why_is_visualization_important/
3. Bootstrap, from Brown University. Math, science, and computer science related.
 - a. <https://www.bootstrapworld.org/materials/spring2017/courses/data-science/resources/teachers/index.shtml>
4. Fish trawl, water temperature and related data from the University of Rhode Island.
 - a. <https://web.uri.edu/fishtrawl/data/>
5. Direct link to the snowfall totals map
 - a. <https://gis.ncdc.noaa.gov/maps/ncei/rsi>
6. The math activity “So Many People!” uses this link to Census Bureau data
 - a. <https://www.census.gov/schools/facts/>
7. Narragansett Bay water quality, links to other related sites as well as visualizations and research done on the water.
 - a. http://www.watershedcounts.org/marine_water_quality.html
8. A watershed modeling app from the Stroud Water Research Center.
 - a. <https://wikiwatershed.org/model/>
9. NERR – National Estuarine Research Reserve data, multiple data sets included.
 - a. <https://coast.noaa.gov/digitalcoast/data/nerr.html>
 - b. Or also here is a link to NERR data
 - c. <https://cdmo.baruch.sc.edu/get/landing.cfm>
10. Links provided by Dr. Peter Meyer from RIC’s Physical Science Department
 - a. U.S. Geological Survey: <https://www.usgs.gov/>
 - b. National Oceanic and Atmospheric Administration: <https://www.noaa.gov/education/resource-collections/data>
 - c. Smithsonian Institution Global Volcanism Program: <https://volcano.si.edu/>
11. Census bureau data has links to specific math related problems with visualizations.
 - a. https://www.census.gov/programs-surveys/sis/activities/math.Grades_9-12.html

Links to more information about activities

1. Use “countif” function to search for unsorted data within a column
 - a. <https://www.howtogeek.com/446398/how-to-count-data-matching-set-criteria-in-google-sheets/>
2. English literature ideas for visualizations, need to scroll a bit, activity based off the ideas under lesson one
 - a. <https://www.easel.ly/blog/infographics-for-teaching-classic-literature/>
3. Track events in a book in a spread sheet, make a visualization when done (the example used the Iliad and rage)
 - a. <https://www.edutopia.org/blog/homer-high-tech-data-visualization-james-earle>
4. Graduate school paper for the math activity “Geometry Jam”
 - a. https://digitalcommons.lsu.edu/cgi/viewcontent.cgi?article=1538&context=gradschool_theses
5. An online chemistry textbook that explains significant figures and how to properly measure (have to scroll down to section five)
 - a. <https://wou.edu/chemistry/courses/online-chemistry-textbooks/foundations-general-organic-biological-chemistry/chapter-1-measurements-chemistry/>
6. Michael Sandberg’s blog post about toys being data visualizations
 - a. <https://datavizblog.com/2017/06/25/dataviz-as-art-are-toys-data-visualizations/>
7. The Hubspot and Visage presentation that was summarized in the workshop
 - a. https://cdn2.hubspot.net/hub/53/file-863940581-pdf/Data_Visualization_101_How_to_Design_Charts_and_Graphs.pdf

Works Cited

- @Lagomorpha. “The Sleep Blanket[.] A visualization of my son's sleep pattern from birth to his first birthday. Crochet border surrounding a double knit body. Each row represents a single day. Each stitch represents 6 minutes of time spent awake or asleep.” *Twitter*, 12 Jul. 2019, 2:57 p.m., twitter.com/Lagomorpha/status/1149754592579600384.
- “A Brief History of USDA Food Guides.” *ChooseMyPlate*, U.S. Department of Agriculture, www.choosemyplate.gov/eathealthy/brief-history-usda-food-guides.
- “Marine Water Quality.” Narragansett Bay Watershed Counts, http://www.watershedcounts.org/marine_water_quality.html
- Dragicevic, Pierre. “1898 – Tallies Used as Social Displays on Pacific Island.” *List of Physical Visualizations and Related Artifacts*, 24 Nov. 2014, dataphys.org/list/tallies-used-as-social-displays-on-pacific-islands/.
- Ellerton, Shon. “DIKW Pyramids and Car Crashes.” *Medium*, Medium, 15 Oct. 2019, medium.com/ironkeel/dikw-pyramids-and-car-crashes-de77591f1491.
- Jones, Adam P. “History of O – Ring Damage in Field Joints.” Adam P. Jones, 27 Feb. 2019, www.adam-p-jones.com/why_is_visualization_important/.
- Konecky, Perri. “This Black Barbie in a Wheelchair Comes With a Ramp to Fit Inside the Dreamhouse (!!).” *Popsugar*, Group Nine Media Inc, 28 Jun. 2019, www.popsugar.com/family/Black-Barbie-Wheelchair-From-Fashionista-Collection-46328126.
- “Mattel Barbie Doll and Wheelchair.” Kmart, Transform SR Brands LLC, 30 Aug. 2019, www.kmart.com/mattel-barbie-doll-and-wheelchair/p-A049022178.
- Perovich, Laura. “Dressed in Data.” *Laura Perovich*, <http://www.lauraperovich.com/projects/dressedindata.html>.
- Silvermanaz, Gordon. “What did Singapore export in 2012?” Wikipedia, 18 Jul. 2014, en.wikipedia.org/wiki/File:2012_Singapore_Products_Export_Treemap.png.

