

ISSUES: DATA SET

Visualizing a Disease Outbreak Using ESRI Story Maps

Sophie T. Breitbart^{1*} and Emily Weigel²

¹Department of Ecology & Evolutionary Biology, University of Toronto, Toronto, ON M5S 3B2, Canada

²Department of Biological Sciences, Georgia Institute of Technology, Atlanta, GA 30332

*Corresponding author: Sophie T. Breitbart (sophie.breitbart@mail.utoronto.ca)



Sneezing is a form of direct disease transmission via droplet spread.
By James Gathany (CDC Public Health Image library ID 11162) [Public domain], via Wikimedia Commons.

THE ECOLOGICAL QUESTION:

How can we harness geographic information systems (GIS) to understand how the magnitude and spatial distribution of disease changes over time?

TIEE

Teaching Issues and Experiments in Ecology - Volume 15, February 2019

ECOLOGICAL CONTENT:

Disease ecology, viral infections, spatio-temporal spread of disease, geographic information systems (GIS), epidemiology

WHAT STUDENTS DO:

Students transform tabular epidemiological data into a series of dynamic maps to narrate a disease outbreak spatially and verbally.

STUDENT-ACTIVE APPROACHES:

[Brainstorming](#), [Think-pair-share](#), Group work, Student presentation

SKILLS:

- Format, clean, manipulate with functions, and practice quality control of raw data
- Create ESRI Story Map and learn basic GIS skills to organize spatio-temporal data
- Synthesize principles of disease ecology to analyze and predict areas of high risk to public health
- Practice clear communication of scientific data by developing and justifying strategy to handle outbreak

ASSESSABLE OUTCOMES:

At the end of this activity, students will demonstrate their ability to understand basic principles of disease transmission, convey real epidemiological data as visual stories, and develop a plan to handle a “real” outbreak. Instructors may evaluate students’ knowledge and comprehension of the basic principles of disease transmission by analyzing how students’ initial and final hypotheses changed throughout the exercise, as well as by assessing the introductory section of their presentations. Students will demonstrate their ability to apply their understanding of disease ecology to a real situation by reporting how an outbreak evolved over time, expressed through dynamic maps in their presentations. Instructors may assess the students’ presentations for both Story Map accuracy and ability to synthesize information in order to create a plan to handle the outbreak. The “Guidelines for Student Assessment” in the Faculty Notes section explicitly link the initial skills outcomes with the learning objectives by providing several ways for students to demonstrate their learning of the skills.

SOURCES:

- Georgia Tech Stamps Health Services. 12/6/2017. Approved for on-campus use and distribution.
- State of California, Department of Public Health. (2017). Infectious Disease Cases by County, Year, and Sex. Available from California Open Data Portal Website: <https://data.ca.gov/dataset/infectious-disease-cases-county-year-and-sex>
- State of California, Department of Public Health. (2018). West Nile Virus Cases, 2006-present. Available from California Open Data Portal Website: <https://data.ca.gov/dataset/west-nile-virus-cases-2006-present>

ACKNOWLEDGEMENTS:

We thank Georgia Tech Stamps Health Services for the dataset that inspired this lab and the QUBES network for valuable feedback on its design. We thank Mitchell Irving, Lindsay S. Miles, Samreen Munim, Madeleine Oman, and Felipe Torres Vanegas, as well as our many students for

their feedback on the lab. We also thank two anonymous reviewers for providing helpful comments on earlier drafts of the manuscript.

OVERVIEW OF THE MODULE

Ecology is a science that examines biotic and abiotic interactions within an environment. One subclass of these interactions directly addresses the spread of infectious disease, which has large implications on epidemiology. In this field, disease and disease-causing agents are investigated at a population level to understand disease dynamics, in particular disease distribution and spread. Studies in this field are directly responsible for identifying key factors of disease and how to prevent its spread.

In relation to human health, mathematical epidemiology predict the establishment and spread of human pathogens. This is done through modeling -- an important skill for understanding ecological phenomena. For most of these models, a set of assumptions are made that allow one to abstract the patterns and path of disease through space and time to answer questions about disease spread.

In this exercise, students will deepen and extend their understanding of disease ecology and the Susceptible-Infected-Recovered (SIR) epidemic model by portraying an epidemiological outbreak with a series of dynamic maps. Instructors may tailor the exercise by choosing from a variety of diseases, scales, and outbreak locations. The students will convert raw disease incidence data into their own ESRI Story Maps, illustrating how the intensity and geographic prevalence of the outbreak changed over time. First, the students will perform quality control measures on the input incidence data spreadsheet and format it as a typical GIS input file. After incorporation into the ESRI Story Map platform, students will explore elements of map design by crafting a verbal and cartographic narrative of the outbreak. Ultimately, the students will use their models to devise a plan to minimize future disease outbreaks and will communicate their plans to the community (their class).

Resources for Instructors

Diekmann, O. and A. P. Heesterbeek. 2000. *Mathematical Epidemiology of Infectious Diseases*, Wiley.

Jackson, A. 1989. Modeling the Aids Epidemic. *Notices of the American Mathematical Society* 36:981-983.

Keeling, M. 2004. *The Mathematics of Diseases*, <http://plus.maths.org>.

Kermack, W. O. and A. G. McKendrick. 1927. A Contribution to the Mathematical Theory of Epidemics. *Proc. Roy. Soc. Lond. A* 115:700-721.

Murray, J. D. 1989. *Mathematical Biology*, Springer-Verlag.

DATA SETS

- [Original Norovirus dataset](#)

TIEE

Teaching Issues and Experiments in Ecology - Volume 15, February 2019

- Intermediate Excel files used to generate ultimate Excel files that are uploaded to the Story Map interface
 - [Step 1](#)
 - [Step 2](#)
 - [Step 3](#)
- Excel files that are uploaded to the Story Map interface (for instructors wanting to shorten the exercise)
 - [Week 1](#)
 - [Week 2](#)
 - [Week 3](#)
 - [Week 4](#)
 - [Week 5](#)
 - [Week 6](#)
 - [Week 7](#)
 - [Week 8](#)

STUDENT INSTRUCTIONS

Introduction

Ecology is a science that examines biotic and abiotic interactions within an environment. One subclass of these interactions directly addresses the spread of infectious disease, which when applied to human health has come to be known popularly as epidemiology. In this field, disease and disease-causing agents are investigated at a population level to understand disease dynamics, in particular disease distribution and spread. Studies in this field are directly responsible for identifying key factors of disease and how to prevent its spread.

In relation to human health, mathematical epidemiology models the establishment and spread of human pathogens. This is done through modeling: an important skill for understanding ecological phenomena. For most models, including these, a set of assumptions are made that allow one to abstract the patterns and path of disease through space and time to answer questions about disease spread.

Compartment Models of Disease: The SI and SIR Models of Disease

Models often use 'compartments'- these are essentially 'bins,' or categories, in which portions of an overall population are placed in order to create a baseline mathematical framework.

In the most basic models, we can split the population into two categories:

1. Susceptible to the infection of the pathogen (S, or 'susceptibles')

TIEE

Teaching Issues and Experiments in Ecology - Volume 15, February 2019

2. Infected by the pathogen (I, or 'infecteds')

For some models, we instead split the population into three groups by adding a third category:

3. Recovered or immune (R, or 'recovers')

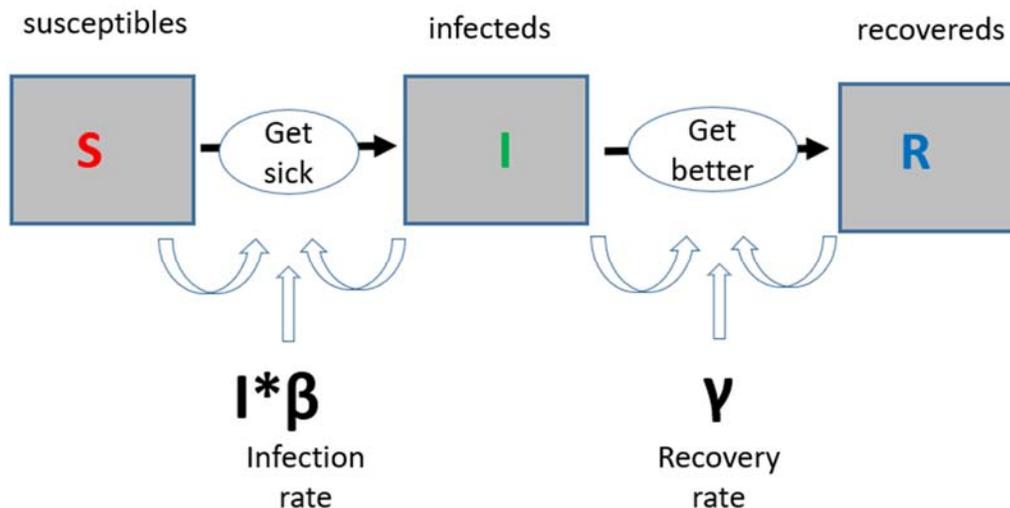


Figure 1. The basic framework of the SIR Model (after Kermack and McKendrick 1927). Boxes indicate compartments and arrows the transition rates between compartments. Between S and I, the transition rate is $I * \beta$ (the probability of getting the disease in a contact between a susceptible and an infectious subject). Between I and R, the transition rate is γ (the rate of recovery).

SIR Model: A Closer Look at the Assumptions

Just like all other mathematical models, this basic model makes a few fundamental assumptions. Among them are the following:

- We assume no individuals enter our population (from birth or immigration) and no deaths occur (unless specifically as a result of our modeled disease).
- The only way to “leave” the susceptible compartment is to become infected.
- Everyone in the population is equally likely to be in each of our compartments. This means a chronically ill person, an infant, and the elderly have the same chance of getting sick as everyone else within the population.

- Everyone also is assumed to be heterogeneously-mixed (that is, sick people do not necessarily “cluster”- there is no spatial component inherent in this model.
- We also assume a fixed transfer between each compartment, although the exact rate of conversion from S to I and I to R may be differ between the two conversion types.

What potential practical issues do you see with these assumptions? Can you think of other assumptions not listed here?

SIR Model: Testing Hypotheses and Creating Realistic Scenarios

The phenomenological assumptions made in constructing your model allow you to build a framework to which you can investigate the effects of particular parameters, as well as predict and calculate directly the prevalence and duration of an epidemic. This information can be used to determine the most effective interventions to prevent disease spread, such as strategic vaccinations to specific populations or methods like quarantine.

Although the SIR model is one of the simplest compartmental models, it is reasonably predictive for human-to-human infectious diseases where recovery confers lasting resistance, such as measles, mumps and rubella. As spatial dynamics or infection times necessitate, this model is modified to include parameters that allow for infections to spread by proximity and for large boom-and-bust like dynamics as infection times shorten.

For some infections, such as influenza and norovirus, the concept of recovery is not biologically meaningful (at least for long). This is because there is no long-term immunity conferred by a previous infection. This means individuals, unless removed from our model by death, continuously cycle between infected and susceptible again. These types of outbreaks leave populations ever-vulnerable to issues of disease outbreak.

Your task in lab will be to examine epidemiological data and study the spatial and temporal dynamics of disease outbreak on campus. You will use what you learn from these models to devise a plan to minimize future disease outbreaks on campus, and to communicate to your plan to the campus community.

Objectives

At the end of this activity, you will be able to:

1. List and relate basic principles of disease transmission
2. Convey real epidemiological data as visual stories
3. Develop a plan to handle a “real” outbreak

Lab Instructions- Georgia Tech Dataset

Step 1: Remove Irrelevant Data

1. Open norovirus data spreadsheet, “Original_Norovirus_Data”, in Excel. Name tab “Norovirus_Incidences”.
 - a. Select all data and sort the Residence column alphabetically. The entire table, not just the Residence column, is sorted; i.e. the rows remain the same but their order reflects the alphabetical order of the Residence column.
 - b. Delete “Off Campus” entries.

Step 2: Assign Recognizable Locations to Sites

2. Copy all data into new tab called “Housing_Library”.
 - a. Oftentimes, databases need cleaning before using. There are often duplicate entries, entries with different spellings, and null entries. Make sure that all residences are spelled correctly.
 - b. Add addresses of residences in new column, “Address”.
 - i. Make sure addresses include commas after the street and city
 1. Example: 555 8th St NW, Atlanta, GA 30318
 - ii. List of dorms here: <https://housing.gatech.edu/housing-options>. You will have to search online for the Greek house addresses.

Step 3: Create Separate Spreadsheets Detailing Weekly Patient Counts

3. Switch back to “Norovirus_Incidences”
 - a. Sort by Residence. Copy/paste the addresses from the “Housing_Library” tab next to the “Residence” column.
 - b. Sort by Date of First Visit. Add new column, “Week”. Then, assign each entry a week number, with Week 1 spanning from Sunday, October 8- Saturday, October 14.

TIEE

Teaching Issues and Experiments in Ecology - Volume 15, February 2019

- c. Delete the “Date of First Visit” column. Rename “Patient Number” → “Patient”. When importing .csv files into ArcGIS software, it’s important to keep column names short and without spaces or special characters. If necessary, add underscores to separate words.
- d. Select the “Residence”, “Address”, and “Week” columns and create a PivotTable in order to find the number of patients belonging to each residence per week. If you are unfamiliar with PivotTables, it will be helpful to search for online resources for insight into why they are helpful and how to make them. Use the screenshot below to help you arrange the PivotTable settings as such:

The screenshot shows an Excel spreadsheet with a PivotTable. The main data table has columns: Patient, Residence, Address, Week. The PivotTable is set up with Row Labels as 'Residence' and Values as 'Count of Residence'. The PivotTable Field List on the right shows 'Residence', 'Address', and 'Week' checked. Annotations with orange boxes point to these settings: 'Residence and Week fields are checked.', 'Week, then Residence fields are Row Labels box.', and 'Count of Residence is in Values box.'

- e. Copy each week’s residence and count of residence into new tabs (8 in total), with only two columns in each week’s tab: “Residence” and “Count”. It is important that these tabs summarize the number of patients admitted each week, as well as their residences. Add a new column, “Address”, to match the address with each residence (utilize the “Housing_Library” tab).
- f. Once you’ve separated each week into different tabs, save the tabs as their own .csv files (ex. “Norovirus_Data_Week_1”). You should have eight .csv files in total for eight weeks of data.

Step 4: Bring Tabular Data into ArcGIS Online and Designate Patient Size Symbology

4. Create a free ArcGIS Online account here:
<https://www.arcgis.com/home/createaccount.html>
 - a. From the home page, click the “Map” button on the upper left corner of the screen.
 - b. Again, look to the upper left corner of the screen to press “Add” → “Add Layer From File”. Add each week’s .csv file. Make sure that the address field is identified as an “Address or Place”. The locations should appear on the Georgia Tech campus.
 - i. If you get an error, edit the .csv file, checking that everything is spelled correctly, there aren’t spaces in the column title names, and save again. Re-import; it should work this time.
 - c. For Step 1, choose “Count” as the attribute to show. For Step 2, select “Counts and Amounts (Size)” for the symbology option.
 - i. Choose your own symbol color, shape, and transparency.
 - ii. Find the size of the largest symbol for each week. Do this by first finding the residence with the highest count out of all weeks. This week will have the largest symbol size. Find the maximum symbol size for each week by finding the highest patient count per residence for each week, then sizing it proportionally.
 1. For example: In week 4, Residence “A” had 10 patients, the highest count for any residence of any other week. This week’s maximum symbol size will be 30 (3x the max patients; 3 is an easy multiplier that will show up well on a map. Feel free to experiment with this multiplier).
 2. Several residences had only one student visit the health center per week. For these weeks, the minimum symbol size will be 3 (3x the min patients).

TIEE

Teaching Issues and Experiments in Ecology - Volume 15, February 2019

Fill out the following table for help determining symbology scaling.

Example:

Week	1	2	3	4	5	6	7	8
Highest Count Per Week	2	7	8	10** (Highest)	6	3	1	2
Lowest Count Per Week	1	1	3	4	3	2	1* (Lowest)	1
Multiplier	3	3	3	3	3	3	3	3
Maximum Symbol Size	6	21	24	10x3=30	18	9	3	6
Minimum Symbol Size	1	1	9	12	9	6	1x3=3	3

Week	1	2	3	4	5	6	7	8
Highest Count Per Week								
Lowest Count Per Week								
Multiplier								
Maximum Symbol Size								
Minimum Symbol Size								

d. Save layer as its own file.

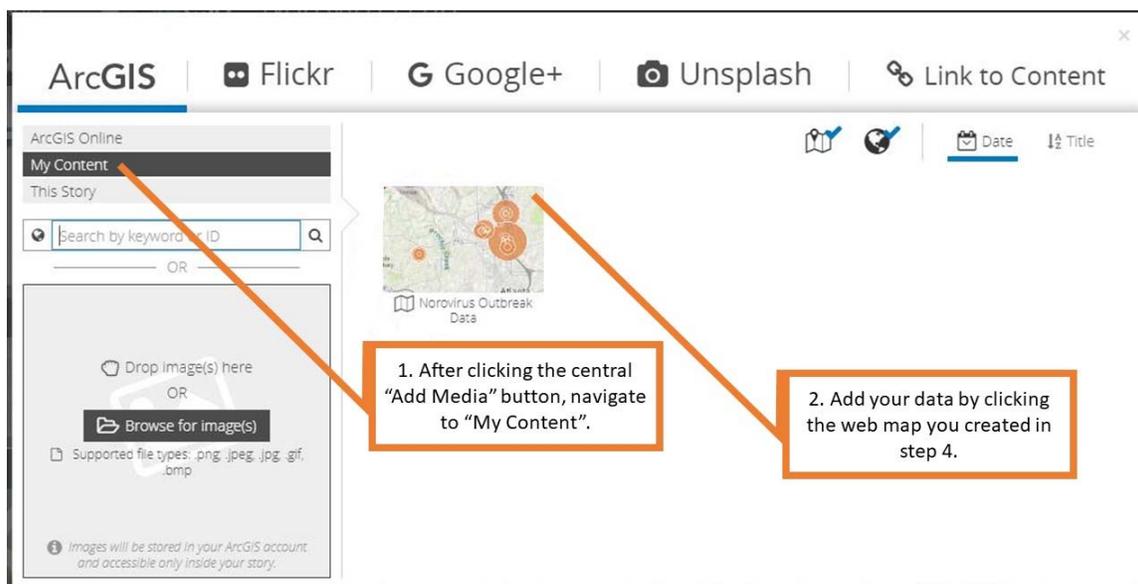
Step 5: Visually Convey Temporal and Spatial Aspects of Tabular Data

5. Create your story map. Navigate here: <http://storymaps.arcgis.com/en/my-stories/>
 - a. Click "Create Story".
 - b. Choose "Cascade".
 - c. Add title and your name.
 - d. Add Text or Media slides if you'd like. The norovirus data will be conveyed with the Immersive slide.

TIEE

Teaching Issues and Experiments in Ecology - Volume 15, February 2019

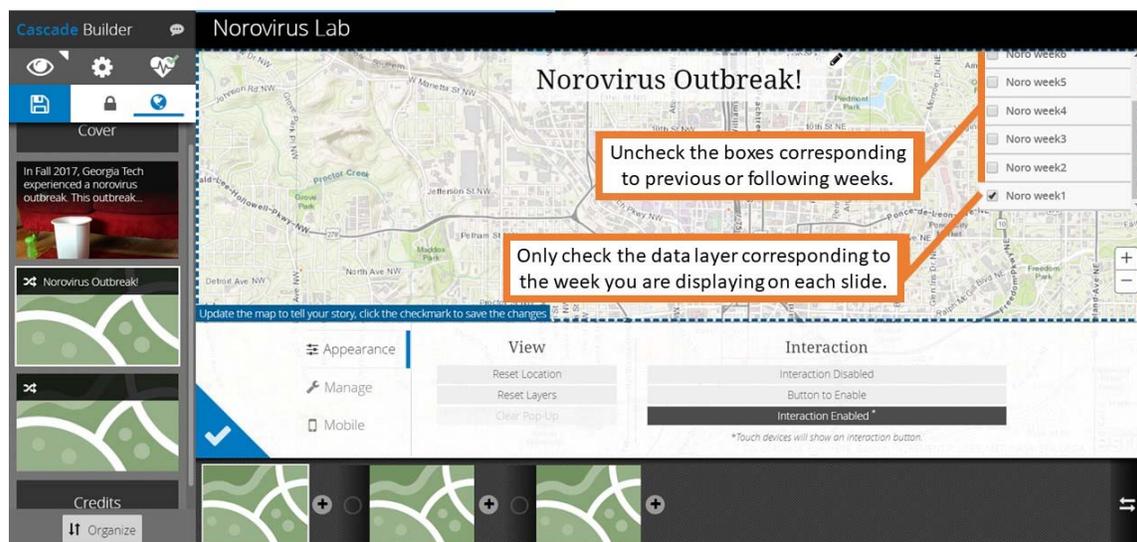
- e. After you've gotten to the Immersive slide, add titles and text to explain the backstory to the norovirus outbreak. Then, press the center "Add Media" button to add the layers with each week's norovirus data.



- i. At first, all the layers saved in the ArcGIS Online map will be superimposed. You will turn off all weeks except the one you are showcasing by unchecking the other weeks' layers. Do this by clicking the gray pencil immediately right of the vertical slides on the screen's left side. Then, uncheck every week's layer except for week 1:

TIEE

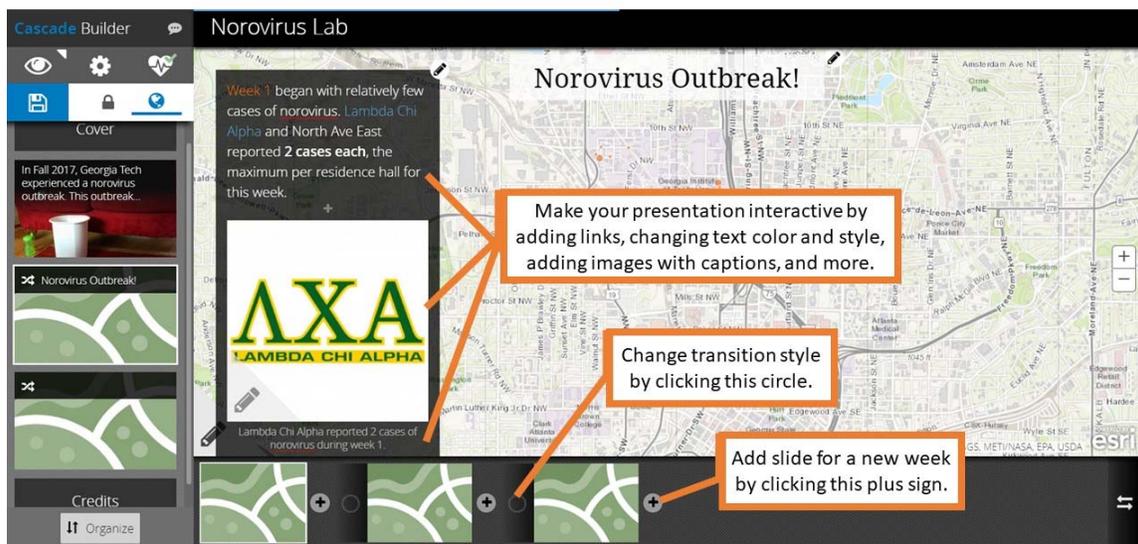
Teaching Issues and Experiments in Ecology - Volume 15, February 2019



- ii. Check the blue triangle's check mark (located where the pencil was) to save the change.
 - iii. Duplicate the slide component (the horizontal slides on the bottom of the screen) by pressing the button on its upper right corner. Repeat this process, creating as many slides as the number of weeks in the outbreak. Show each week's norovirus data.
 - iv. Add transitions if you wish- a circle on the slide component will appear after you add a second slide component and change the week.
- f. Play with the styles and settings for the story map. You can add links to the narrative font, change its colors and styles, add media, make it dissolve between slide components, move and zoom in/out of the map... personalize it as you wish!
 - g. Be sure to save as you go!

TIEE

Teaching Issues and Experiments in Ecology - Volume 15, February 2019



Next steps:

Your narrative of the outbreak should include when the outbreak originated, where and the pattern of spread across campus. Look carefully at the areas where the outbreak begins to spread rapidly. What can you say about those physical spaces? Ask yourselves, what additional information might I need to know about the virus and how it spreads in order to interpret this data? Use what you've learned about patterns of disease spread with your Story Map to devise a plan to minimize future disease outbreaks that you will communicate to the community (your class).

Lab Instructions- Other Datasets

Step 1: Remove Irrelevant Data

Open data spreadsheet. Select subset of dataset so that only records for one disease are present (Microsoft Excel's Sort & Filter tool is helpful for this and the following steps). Delete columns with irrelevant data. Alphabetize entries in location column and delete any rows with locations that are irrelevant or erroneous (ex. sites in Nevada for a Oregon-based study). Delete any summary rows that duplicate entries, such as those that contain "California" in the location-based "County" column. If data contains another variable such as sex, decide whether the final dataset's population will include data from one or both sexes.

Then, delete the duplicate rows so that you are only left with, for example, only female entries, or only total (combined male and female) entries. Then, delete the “Sex” column. At this point, there should be only columns for three variables: population size (number of infected patients), location (city, state, country, etc.), and time unit (day, week, year, etc.)

Step 2: Assign Recognizable Locations to Sites

Oftentimes, databases need cleaning before using. There are often duplicate entries, entries with different spellings, and null entries. Make sure that all locations are spelled correctly.

If identifying locations with addresses, create one new column that lists the address of the sites with commas after the street, city, etc. such as with 555 8th St NW, Atlanta, GA 30318. If using coordinates instead of addresses, create two columns that list the latitude and longitude of the sites in the World Geodetic System 1984 (WGS84) coordinate system.

If the existing location data uses larger scales such as town or state scales, choose a landmark that can function as the location’s center, such as a town hall or state capitol building, and input the address or latitude and longitude in the appropriate columns. Only address (one column) OR latitude and longitude (two columns) are required.

Step 3: Create Separate Spreadsheets Detailing Periodic Patient Counts

Create individual .csv files for each unit time (day, week, month, etc.) and fill in each unit time’s corresponding data. Use the Sort function to sort dates. Each .csv must contain at least two columns: one with a count of infected patients, and one with the address of the site (if using latitude/longitude to designate address, there will be three columns.) It may be helpful to make a PivotTable to summarize the number of patients per time period per location (see image below for reference).

TIEE

Teaching Issues and Experiments in Ecology - Volume 15, February 2019

The screenshot shows an Excel spreadsheet with a PivotTable. The data table has columns: Patient, Residence Address, and Week. The PivotTable is set up with 'Week, then Residence' as Row Labels and 'Count of Residence' as the Values field. The PivotTable Field List on the right shows 'Residence' and 'Week' checked under 'Choose fields to add to report:'. The Row Labels box contains 'Week' and 'Residence', and the Values box contains 'Count of Residence'.

It is fine to retain the column with the original location names. Delete the unit time column, since it is accounted for in the separate .csv titles. Make sure each .csv contains a header row with column names. Keep column names short and without spaces or special characters. If necessary, add underscores to separate words.

Step 4: Bring Tabular Data into ArcGIS Online and Designate Patient Size Symbolology

Create a free ArcGIS Online account here:

<https://www.arcgis.com/home/createaccount.html>. From the home page, click the “Map” button on the upper left corner of the screen. Again, look to the upper left corner of the screen to press “Add” → “Add Layer From File”. Add each week’s .csv file. Make sure that the address field is identified as an “Address or Place”. If using coordinates instead of an address, after pressing the “Add” → “Add Layer From File” buttons, the sites should show up automatically. Check that the site locations appear where they ought to. If you get an error, edit the .csv file, checking that everything is spelled correctly, there aren’t spaces in the column title names, and save again. Re-import; it should work this time.

For Step 1, choose “Count” as the attribute to show. For Step 2, select “Counts and Amounts (Size)” for the symbology option. Choose your own symbol color, shape, and transparency. Find the size of the largest symbol for each time unit (day, week, year, etc.). Do this by first finding the location with the highest count out of all time units. This time units will have the largest symbol size. Find the

TIEE

Teaching Issues and Experiments in Ecology - Volume 15, February 2019

maximum symbol size for each time unit by finding the highest patient count per location for each time unit, then sizing it proportionally.

For example: In week X, location Y had 10 patients, the highest count for any location of any other week. This week's maximum symbol size will be 30 (3x the max patients; 3 is an easy multiplier that will show up well on a map. Feel free to experiment with this multiplier). For locations with one weekly report of illness, the minimum symbol size will be 3 (3x the min patients). Fill out the following table for help determining symbology scaling, then save the layer as its own file.

Example:

Week	1	2	3	4	5	6	7	8
Highest Count Per Week	2	7	8	10** (Highest)	6	3	1	2
Lowest Count Per Week	1	1	3	4	3	2	1* (Lowest)	1
Multiplier	3	3	3	3	3	3	3	3
Maximum Symbol Size	6	21	24	10x3=30	18	9	3	6
Minimum Symbol Size	1	1	9	12	9	6	1x3=3	3

Week	1	2	3	4	5	6	7	8
Highest Count Per Time Unit								
Lowest Count Per Time Unit								
Multiplier								
Maximum Symbol Size								

TIEE

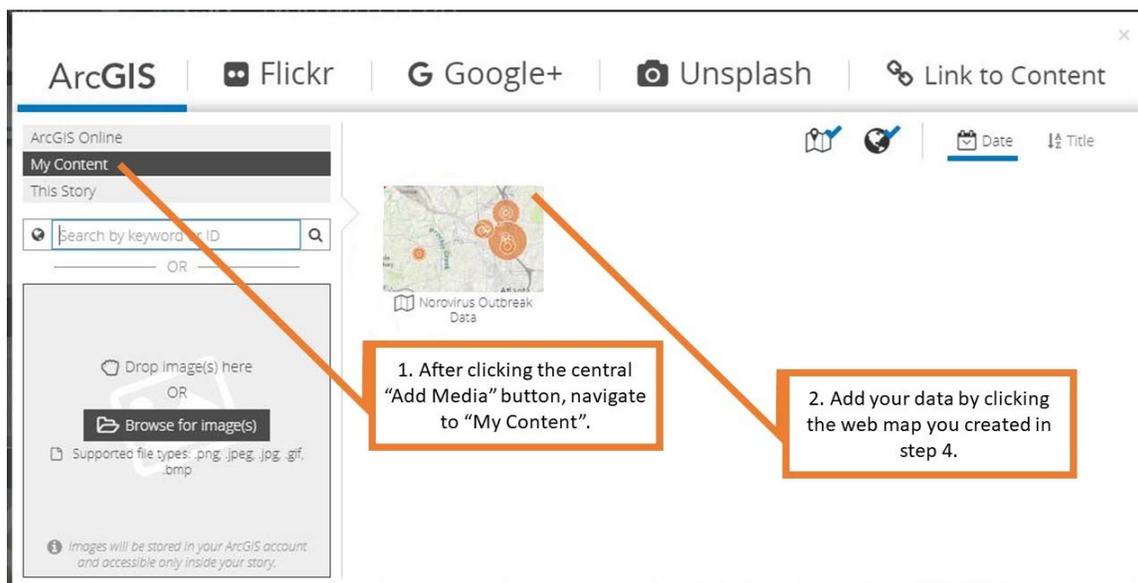
Teaching Issues and Experiments in Ecology - Volume 15, February 2019

Minimum Symbol Size								
------------------------------------	--	--	--	--	--	--	--	--

Step 5: Visually Convey Temporal and Spatial Aspects of Tabular Data

Create your story map. Navigate here: <http://storymaps.arcgis.com/en/my-stories/>. Click “Create Story”. Choose “Cascade”. Add title and your name. Add Text or Media slides if you’d like. The health data will be conveyed with the Immersive slide. After you’ve gotten to the Immersive slide, add titles and text to explain the backstory to the outbreak. Then, press the center “Add Media” button to add the layers with each week’s data.

At first, all the layers saved in the ArcGIS Online map will be superimposed. You will turn off all time units except the one you are showcasing by unchecking the other time units’ layers. Do this by clicking the gray pencil immediately right of the

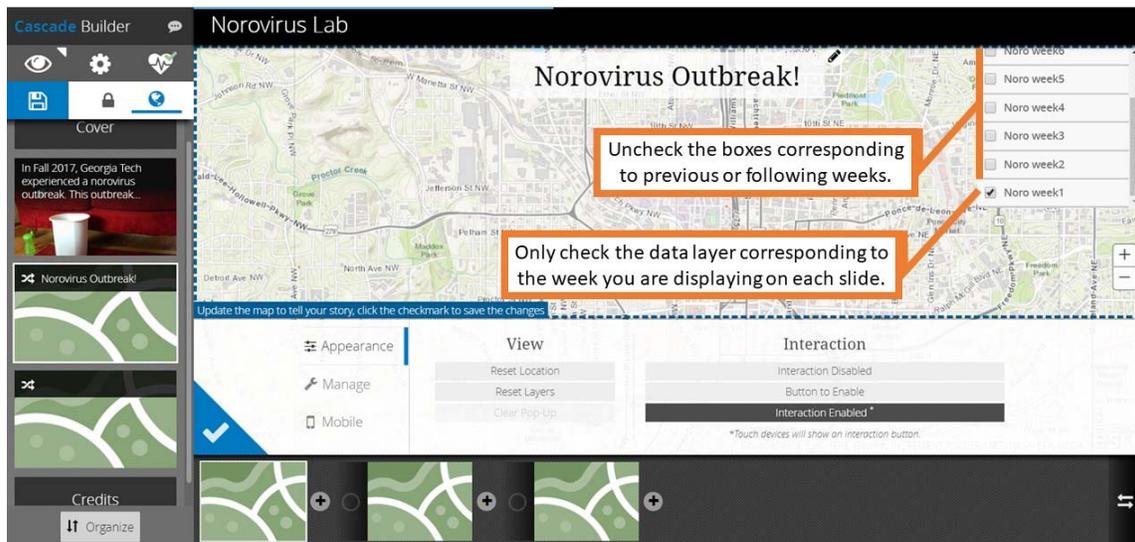


vertical slides on the screen’s left side. Then, uncheck every time unit’s layer

TIEE

Teaching Issues and Experiments in Ecology - Volume 15, February 2019

except for time unit 1:

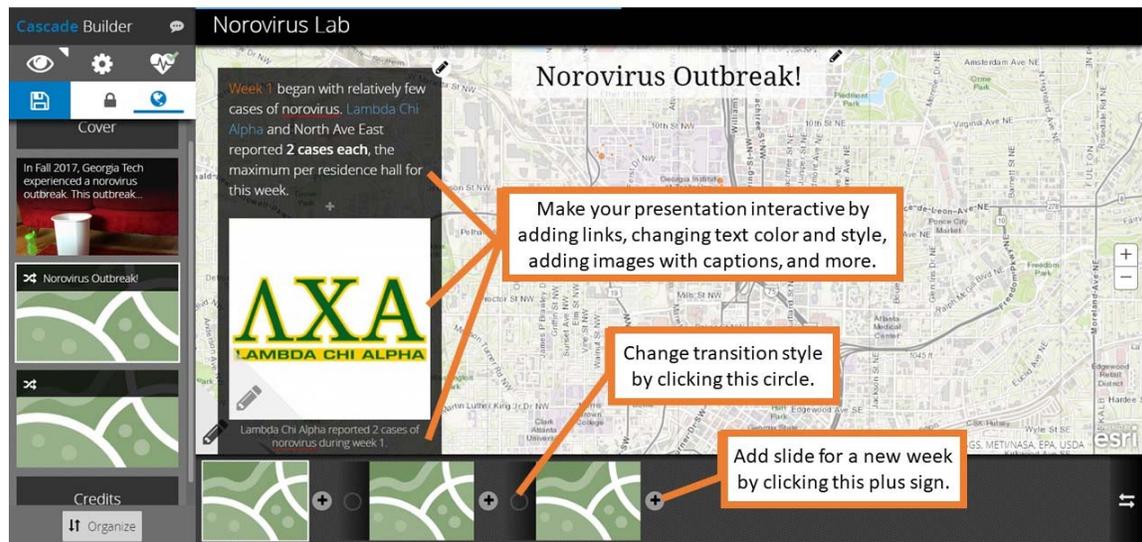


Check the blue triangle's check mark (located where the pencil was) to save the change. Duplicate the slide component (the horizontal slides on the bottom of the screen) by pressing the button on its upper right corner. Repeat this process, creating as many slides as the number of time units in the outbreak. Show each time unit's data.

Add transitions if you'd like- a circle on the slide component will appear after you add a second slide component and change the time unit. Play with the styles and settings for the story map. You can add links to the narrative font, change its colors and styles, add media, make it dissolve between slide components, move and zoom in/out of the map... personalize it as you wish!

TIEE

Teaching Issues and Experiments in Ecology - Volume 15, February 2019



Be sure to save as you go!

Next steps:

Follow your professor's instructions to craft a verbal narrative of the outbreak. Use your model to devise a plan to minimize future disease outbreaks that you will communicate to the community (your class).

NOTES TO FACULTY

Comments of Implementing this exercise

Students begin by working individually or in pairs to view the dataset and reflect on the data presented. As an opening exercise, they are asked to individually brainstorm and record their initial thoughts on how viruses spread (this may be done generally or with respect to a particular type of virus with which the students may be more familiar, e.g. flu viruses, HIV, etc.). As a think-pair-share, students then pair up with a partner to append or amend their notes on viruses, and finally combine pairs to make a conceptual model of the elements involved in virus spread. Ultimately, students develop these models to act as their own hypotheses regarding how disease spreads.

Next, instructors conduct a mini-lecture to describe the basic SIR model of disease (see supplementary powerpoint), and groups are tasked to review their notes and determine factors not included in the basic SIR model. Students will generate many ideas such as age-related susceptibility, quarantine-related timing

TIEE

Teaching Issues and Experiments in Ecology - Volume 15, February 2019

issues, and spatial elements of disease spread (the main focus of this exercise; if students do not arrive here on their own, guide them in this direction). We advise generating a class list on the board. Also, Nichol et al. (2010) may be used as prelab reading material.

Then, have students individually open the dataset on their local computers. Although we used a campus outbreak data, you may use any data for which you have spatial and temporal information. After looking at the dataset, have the students discuss in small groups what information they have and how they could use it. Ultimately, students need to answer the question, “How could I use this information to discover and support a strategy to prevent disease spread for the population in this area?”.

Once they answer the “why this data is useful” question, the students then are prepared to work through the lab to create their Story Maps. Students may choose to do this activity as individuals or as pairs, but we do not suggest it as small groups; students need to wrestle with the data directly. After the maps are prepared, students are expected to revisit their original conceptual models and make any revisions (e.g., amendments, additions) based on elements discovered in working with real data. Note that their models and examination of the data as to how viruses spread play directly into their “action plans” in role-playing health assessment personnel pitching a strategy to protect the population.

In the student instructions, we carefully provide all of the steps and necessary details to clean and work with the Georgia Tech dataset (note that how to do those steps, i.e., the specific manipulations, ‘clicks’, etc., is often version-specific to a software and can be found via search engines). We also provide detailed guidance for accomplishing the exercise with other appropriate datasets.

At the instructor's discretion, given their students' skill sets and the available time in class, instructors may wish to leave some steps out; this allows for greater student inquiry and engagement and can lead to interesting discussions about when and how data manipulation decisions are made (see Faculty Notes section below, “Suggestions for Shortening Exercise”, for more details).

Based on iterations of this lab with varying degrees of explicit instruction, we recommend at least some scaffolding around scaffolding around steps 3d (i.e., creating a PivotTable) and 4c (i.e., designing Story Map symbology) to ensure the exercise can be completed (see Faculty Notes section below, “Challenging Aspects of Exercise”, for more details).

Comments on Shortening Exercise

TIEE

Teaching Issues and Experiments in Ecology - Volume 15, February 2019

Based on the skill level of students, the exercise may take anywhere from 45 minutes to 2+ hours to complete. The following suggestions may be used if instructors wish to shorten the exercise:

1. Omit a portion of the in-class exercise
 - a. If instructors wish to skip teaching about data cleaning and formatting, they may provide students with the 8 attached CSV files (“Norovirus_week1”, “Norovirus_week2”, etc.) and start at step 4 of the instructions.
 - b. Instructors may assign a portion of the exercise as prelab prep-work and invite students to come to class having completed certain steps.
2. Review skills associated with steps
 - a. Instructors may allocate in- or out-of-class time for familiarizing students with the exercise’s steps. Instructors could present an in-class exercise trial to offer the most help to students. More independent approaches include presenting how to accomplish specific tasks involved in the exercise, such as creating a PivotTable in Excel, with hypothetical or actual exercise data. For the most hands-off approach, instructors could provide students with a list of skills they should study in order to prepare for the exercise. See section below, “Challenging Aspects of Exercise”, for identification of steps that may be most challenging to students and how to overcome them.

Comments on Challenging Aspects of Exercise

While teaching self-reliance by encouraging students to answer their own questions (i.e. Googling how to cut/paste in Excel) has significant value, instructors may preface this exercise by reviewing common issues encountered by students. We would recommend covering tasks such as:

1. In Excel:
 - a. Cell formatting, including sorting all data by one column, copying cell contents to several others such as by dragging, deleting duplicate rows
 - b. Creating and manipulating PivotTables
 - c. Saving .xlsx file tabs as separate .csv files
2. In Story Map interface:
 - a. How to navigate website

- b. Adjust symbol scaling

Comments on Student Assessment

Here we provide a few examples of targeted assessments to the objectives of our lesson.

Objective 1: “List and relate basic principles of disease transmission”:

1. **Have students list the basic elements of an SIR model.** These lists should, at minimum, include the 4 basic elements: a population (1) that consists of susceptible (2), infected (3), and recovered (4) individuals.
2. **Have students draw relational diagrams between S, I, and R groups within a population.** These models should have, at minimum, the 4 elements as above and indicate with arrows the movement of individuals between compartments.
3. **Have students relate in a paragraph of text the S, I, and R groups within a population.** This is essentially converting the ball-and-stick model to words. Answers should, at a minimum, include the 4 elements as above and describe the movement of individuals between compartments.

Objective 2: “Convey real epidemiological data as visual stories”

Given that this objective is the crux of the lesson while students may approach this lesson with varying levels of prior knowledge and ability, we provide sub-objectives below, each of which can be stacked on the prior sub-objective for deeper demonstration of knowledge. Alternatively, these elements may be used holistically by conversion into a rubric with each sub-objective an element of the overall rubric to master.

1. **Have students work with real data by associating physical locations and outbreak numbers in time and space.** At a minimum, students should show mastery in formatting data (1) in a consistent fashion with respect to (2) location, (3) time, and (4) patient counts. Should students stop here, graphs may be made comparing numbers of disease incidences between several locations. All data is to be accounted for, including incidences where data may be dropped for a justified reason (e.g., not enough detail, ambiguous reference, etc.).
2. **Have students convey temporal and spatial data using ArcGIS.** Students show mastery here by placing their data into an ArcGIS storyline, which shows the dynamic spatial and temporal patterns of the disease outbreak. All data is to be accounted for. To accomplish this, students should (1) establish a storyline, (2) feed in their data, (3) examine their

data to be sure it is correctly associated in time and space with a given outbreak.

3. **Have students annotate their data to tell a convincing story by leveraging the storyline features of ArcGIS.** Students show mastery here by adding annotation to the frames of the maps to explain how the disease is changing over time (1) and space (2). Students may make stylistic choices here to add media to enhance their descriptions, change fonts, etc., but key here is that the annotation is present and accurate.

Objective 3: “Develop a plan to handle a ‘real’ outbreak”:

These are the most difficult, but perhaps the most rewarding, of the objectives for the students. For maximal impact, we suggest that students master both of the above objectives before moving to this third objective.

1. **Have students examine what they have learned about the outbreak on the example dataset to identify key factors for outbreaks.** This can be spatial or temporal information they extract about the pattern of sickness. For example, ‘hotspots’ or clustered timing of sickness may exist. Answers are variable but should relate to the patterns seen in the data.
2. **From the factors the students have identified, students should then investigate current approaches to addressing outbreaks of this contagion.** Students should include (1) etiology/vector of disease spread, (2) the known infectivity, (3) current best practices to avoid infection, (4) current best practices to avoid spread once infected, and (5) vulnerable populations or conditions, such as children under 2 or the elderly.
3. **Students should then relate the patterns of spread on campus to the disease information to create a plan to limit the disease spread.** Their plans should include (1) the timing of their intervention, (2) the target population, (3) the actions individuals should take, (4) the actions officials should take, (5) any risks/unknowns that might mitigate the outcome/effectiveness of their plan, and (6) a well-reasoned communication strategy likely to be effective on campus.
4. **Advanced students can be asked to draw relational diagrams between S, I, and R groups within a population that now incorporate components of their mitigation strategy.** These models should have, at minimum, the four elements as above, arrows that indicate the movement of individuals between compartments, and compartments/arrows that modify the basic SIR model based on their intervention. For example, an intervention may choose to separate the most vulnerable of the susceptible populations so as not to come into contact with the rest of the

population; this could be shown by dividing the susceptible population into two boxes with only one box connected to the I (and R) components of the model. As a general guide, dividing populations in space (e.g., quarantining) or adding physical structures to fragment the population (e.g., handwashing stations) will usually mean compartments are split into sub-compartments, and changing frequencies of contact or increasing speed of recovery will change the rates (arrows) between compartments. Note that development of a complete, spatially-explicit model is likely outside of the scope for most students completing this lab. The goal here is to get the students to practice converting a real-life strategy into a mathematical expression (equation or variable) that could be used for modeling.

COPYRIGHT STATEMENT

The Ecological Society of America (ESA) holds the copyright for TIEE Volume 13, and the authors retain the copyright for the content of individual contributions (although some text, figures, and data sets may bear further copyright notice). No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the copyright owner. Use solely at one's own institution with ***no intent for profit*** is excluded from the preceding copyright restriction, unless otherwise noted. Proper credit to this publication must be included in your lecture or laboratory course materials (print, electronic, or other means of reproduction) for each use.

To reiterate, you are welcome to download some or all of the material posted at this site for your use in your course(s), which does not include commercial uses for profit. Also, please be aware of the legal restrictions on copyright use for published materials posted at this site. We have obtained permission to use all copyrighted materials, data, figures, tables, images, etc. posted at this site solely for the uses described at TIEE site.

GENERIC DISCLAIMER

Adult supervision is recommended when performing this lab activity. We also recommend that common sense and proper safety precautions be followed by all participants. No responsibility is implied or taken by the contributing author, the editors of

TIEE

Teaching Issues and Experiments in Ecology - Volume 15, February 2019

this Volume, nor anyone associated with maintaining the TIEE web site, nor by their academic employers, nor by the Ecological Society of America for anyone who sustains injuries as a result of using the materials or ideas, or performing the procedures put forth at the TIEE web site, or in any printed materials that derive therefrom.