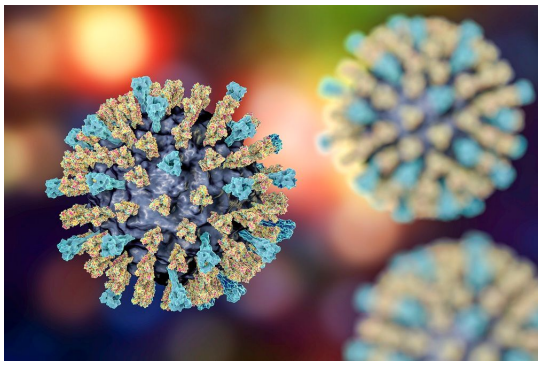
A Curious Case of the Measles

The Physical and Digital Spread of the Measles



Jessica Eisemann, Erin Harned, Kendall Loh, and Sanjna Shenoy

Introduction

Our final project visualizes the physical and virtual spread of measles and measles concern amongst citizens of the United States. We used a variety of data sources including data from the Center for Disease Control (CDC), state Department of Health websites, Twitter, county-level vaccination reports in New York State, and available demographic data from the Department of Agriculture.

Initial Project Plan

Overview

Upon formation of our group, we had initial discussions on topic ideas. We decided that, given our backgrounds and interests, a data visualization project using healthcare data would be appropriate. Our initial ideas included creating some sort of risk metric for US citizens for a particular disease, using available data to create useful visualizations related to Planned Parenthood locations and current policy, and mapping physician shortages in rural areas. Due to the rising coverage of measles cases in the news and the escalating case counts, we decided to focus on measles in the US. This topic was further narrowed as we discovered what types of measles datasets were publically available.

Data

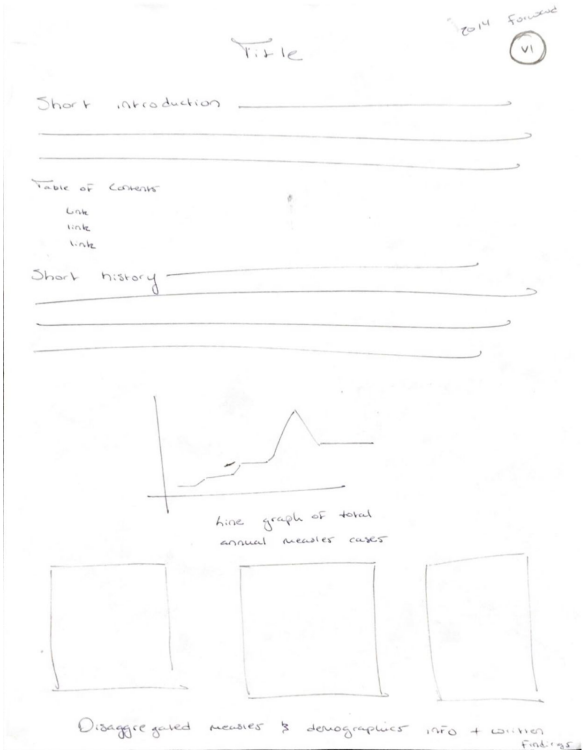
The team took several weeks to explore what was available to better determine the scope of the project. The first datasets we located for this project were as follows.

- CDC's National Notifiable Diseases Surveillance System (NNDSS) for incidences of measles in the United States over time (state-level)
- Non-Medical Vaccine Exemption claims (county-level)
- Google search data of measles symptoms (county-level)
- World Health Organization (WHO) statistics on measles (country-level)
- US Census Bureau's American Fact Finder's demographic data (state- and county-levels)

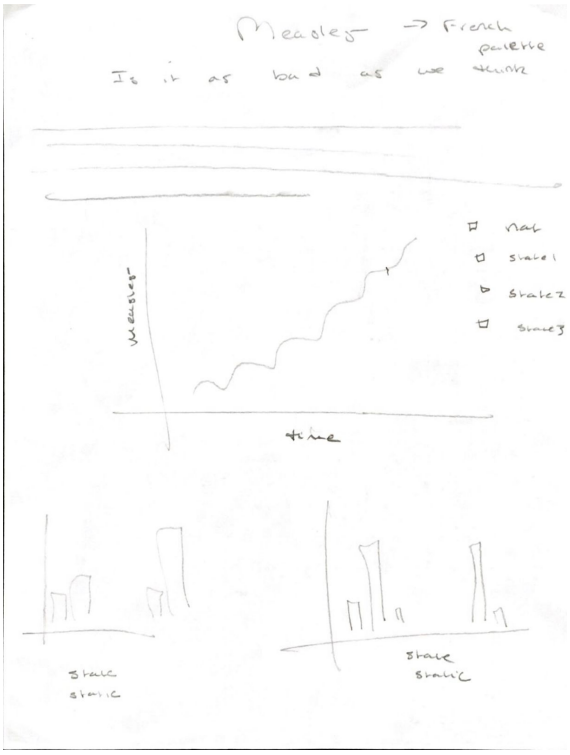
Mock Ups

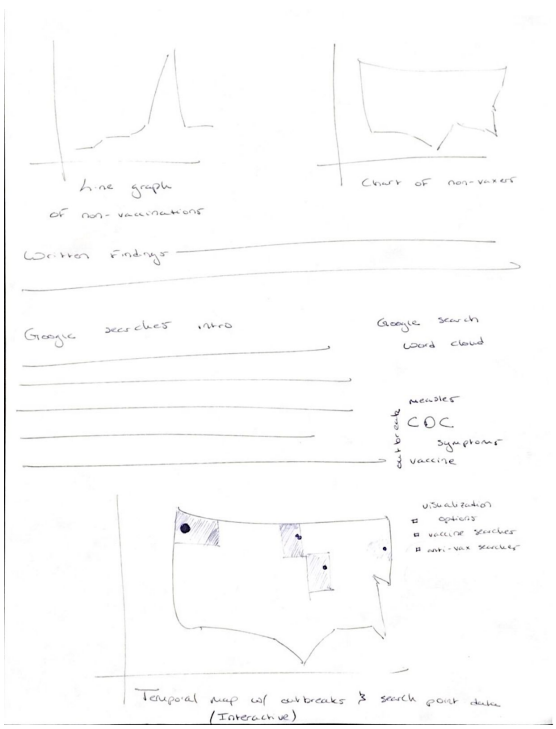
As the project focus shifted and additional datasets were identified, the layout of our webpage progressed through several different versions. At the initial stage we envisioned a single page which, as the user scrolled down, would walk him or her through the history of measles and progressively narrow the viewer's attention to more granular measles details.

Version 1 Layout (single page)



Version 2 Layout (single page)





Intermediate Webpage

Overview

In the interest of telling a story with our data and visualizations, the team decided to take a step-wise approach. We first oriented the viewer to the history of measles in the United States, then provided a brief overview of the current outbreak in the US, before finally narrowing the focus to the state of New York for further analysis. Ultimately, we identified the New York City schools most susceptible to the measles outbreak.

Our motivation for focusing on New York State and New York City schools was generated by current events. In late March 2019 Rockland County of New York State banned unvaccinated children from public spaces, including schools and transit. On April 9, 2019, the New York City Department of Health and Mental Hygiene followed suit by declaring a vaccination mandate in four zip codes. Through our county-level measles case counts, school immunization records, and Twitter data, we evaluated how well the Department of Health selected those areas under the mandate as well as New Yorkers' digital sentiments toward measles.

Data

Our initial data sets varied in geographic level, particularly between state and county levels. To tell a more consistent story about how measles cases have spread across the US, we needed a common geographic reference and chose county-level because it would be feasible to find measles and vaccination records specific to that level and would also allow for a more tailored analysis given the smaller geographic size (as compared to a state-level analysis, which would have diluted our findings).

- Measles case counts assembled from the California, Michigan, New Jersey, New York, Oregon, and Washington state Department of Health websites

We supplemented our initial datasets with demographic data provided at the county level by state in .csv files from the Department of Agriculture. We cross-referenced and mapped the demographic data with the initial measles outbreak data in order to see if any obvious trends appeared.

- County-level education data by state
- County-level unemployment data by state

The recent measles emergency in New York provided the opportunity to gauge how measles concern was spreading as well. The goal was to investigate whether the outbreak was causing panic in the community. In addition, the vaccinations mandate was controversial. In order to tell a complete story of how people were responding to the news it was necessary to include the opinions of those on both sides of the issue.

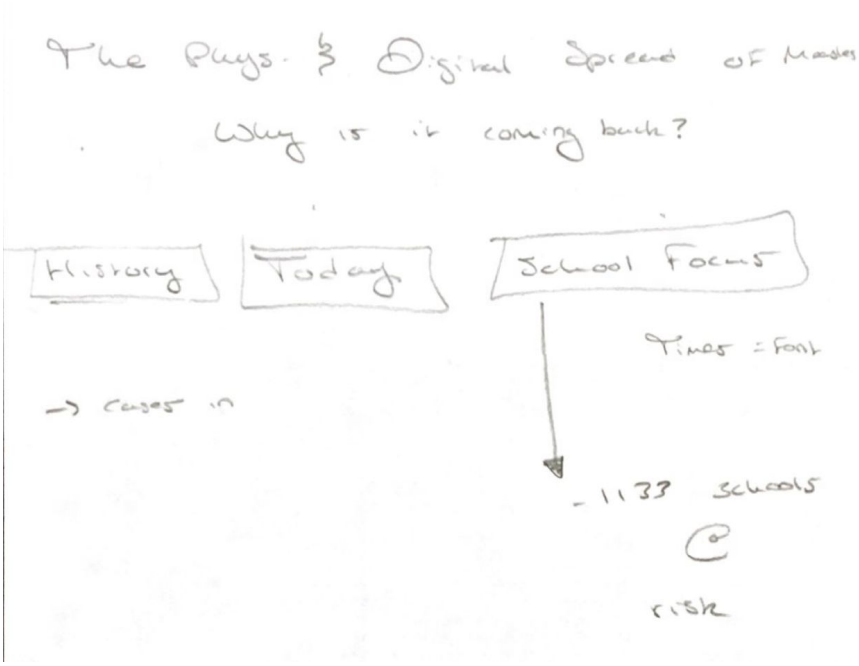
- Google Search Index data on relevant terms (i.e. “measles symptoms”, “symptoms of measles”, “measles rash”, “measles pictures”)
- Tweets from the Twitter Search API that include #measles in New York and New York City, from April 7, 2019 to April 28, 2019.
- Tweets from the Twitter Search API in the United States that include anti-vaccination and pro-vaccination hashtags including, #VaccineInjury, #VaccineFailure, #VaccineHarm, #VaccinesCauseAutism, #VaccinesWork, #GetVax, from April 6, 2019 to April 28, 2019.

Finally, we looked at immunization at a granular level by focusing on school-level immunization surveys collected by the New York Department of Health, beginning in the 2012-2013 school year. The School Immunization Survey records aggregate data from schools in New York State regarding the immunization status of all the students attending school. All schools, excluding New York City public schools, have reported the immunization status of all students in grades kindergarten through 12.

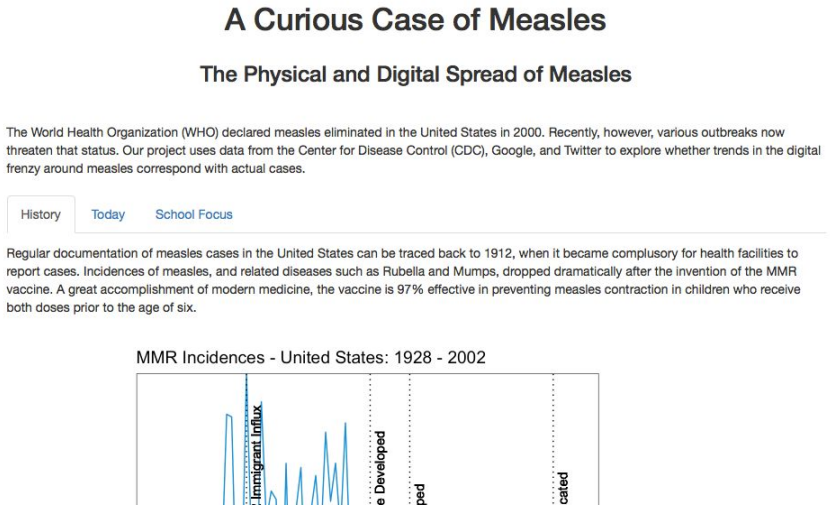
- School immunization surveys covering medical exemption rates, religious exemption rates, measles immunization rates, and location for schools in New York state, from the (2012-2013) school year to the (2017-2018) school year

Mock Ups

In our final layout, we elected to divide our visualizations onto separate pages so viewers could move through the history, present day, and school-specific measles findings progressively.



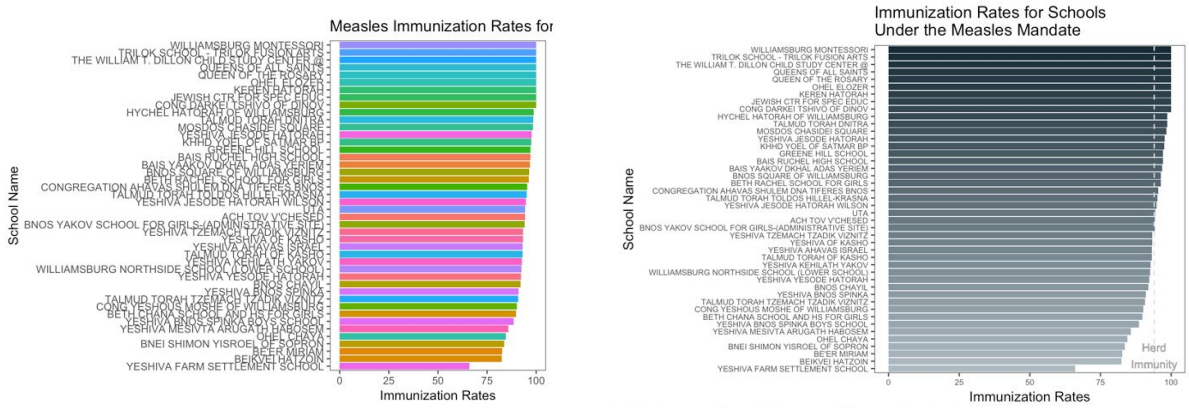
Final Layout (multiple tabs)



Intermediate Web Page (featuring multiple tabs)

Adjustments

Feedback received after presenting our webpage in class prompted a number of cosmetic edits to ensure images were visually consistent in terms of layout and color scheme. Several visualizations were also adjusted to improve interpretability. This included changing the plot of Google search terms over time into an interactive plot so users of the site could clearly see the trend of each search term.



Original and Final Measles Mandate Plots (left and right)

Additionally, we adjusted the last part of our school-focused analysis. We originally identified 14 schools in the top five zip codes with the lowest immunization coverage. However, this list was inclusive of schools with low immunization coverage due to medical exemptions. Since medical exemptions are not optional, we refined our final analysis to differentiate between medical and religious exemptions. The top 40 schools at highest risk of measles contraction are now presented in an interactive data table that users can explore on their own.

Final Webpage

Overview

Our final webpage was divided into three pages: History, Today, and School Focus. The history page provides a broad overview of measles in the U.S. and New York. The today page informs users of the outbreak rates in the last five years before delving into a text analysis of what New Yorkers are saying on Google and Twitter. The final page investigates which schools are most at-risk for measles cases and compares this to the schools affected by New York City's vaccine mandate. Together, these three pages give viewers a brief, but comprehensive history of the measles in the US and make the case for a review of NYC public health policy.

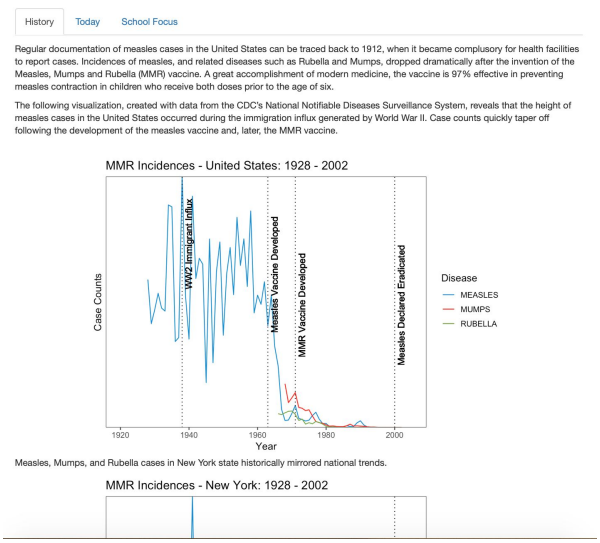
Data

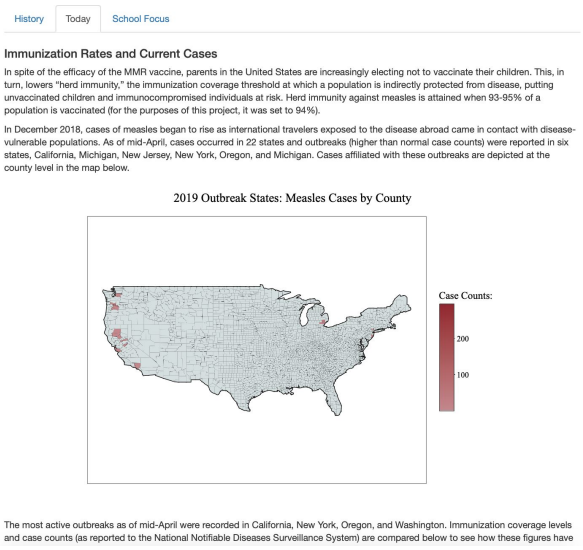
The final sources from which the content for our webpage was drawn are as follows.

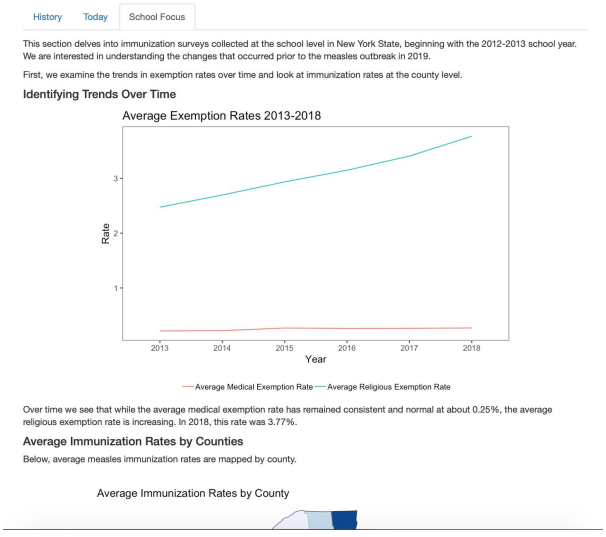
- Center for Disease Control's National Notifiable Disease Surveillance System
- Center for Disease Control's Measles, Mumps, and Rubella (MMR) Immunization Coverage
- State Department of Health websites (California, New York, Oregon, and Washington)
- Department of Agriculture Economic Research Service
- Google

- Twitter Search API
- New York State Department of Health School Immunization Survey

Site







Conclusions

In conclusion the measles outbreak has led to serious conversations by officials and New Yorkers on vaccination. We found that New Yorkers are concerned about the current measles outbreak. They recognize they are at risk of contracting the illness and some are starting to panic. Many are channeling this worry into a vaccination debate. Emotions are running high on both sides as each believe they are in danger. The Health Department has been forced to act on this issue and weigh in on the vaccinations debate via the vaccinations mandate in specific zip codes. However, we found that half of the schools in these zip codes have reached herd

immunity. The zip code with the lowest average immunity rate across schools was not identified by the Health Department. As a result, we recommend that the NYC Health Department reconsider its methodology when reviewing current and future policy. We believe our methodology better identifies at-risk schools in NYC.