How to NOT Map a Disease – COVID in California

GEOG 55 - Critical Analysis in GIS – Daniel Elpa Palana'25 – Summer 2023

INTRODUCTION

During the COVID-19 pandemic, mapping became an essential tool to understand the spread of the virus. Studying where cases are rising among the population is necessary to ensure that the resources and essentials are set in place to treat illnesses and prevent mortality from rising. However, maps are also tools that can be used to **lie**. In his book *How to Lie with Maps*, Mark Monmonier presents how nearly every element in a map can be used to distort and lie about the information it is actually trying to convey. In a situation where life and death are on the line, it is important to understand how maps are manipulated for conveying information.

This project is an exploration of how different elements of a map can distort data. I am applying this in the scenario of mapping the rate of infections for an infectious disease. The goal is to determine whether there is a map that can be distorted to the point it could change the overall severity of the disease is spreading. I used the state of California during the COVID-19 Pandemic as an example.

BACKGROUND AND METHODS

The COVID case data was provided by the California Department of Health and Human Services. Data is taken from the recording period of January 1st 2021 to January 8th 2021, when the state underwent a surge of cases and deaths.

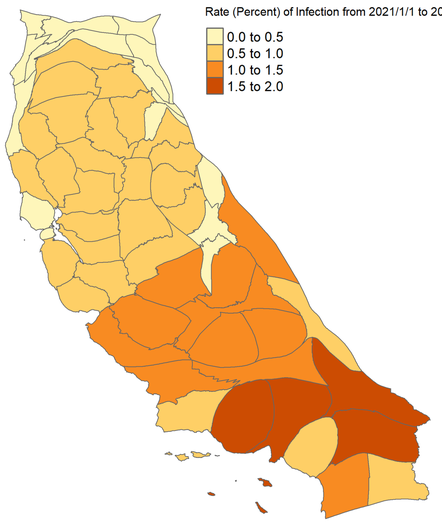
Percentage Rate of Cases was the ratio of dividing total number of cases during recording period from the population at risk, which was the total population above 18, multiplied by 100. Cases per Capita was multiplying the ratio by 100,000.

Maps were created in R using the tmap package with Polygon Census data provided by the tidycensus package.

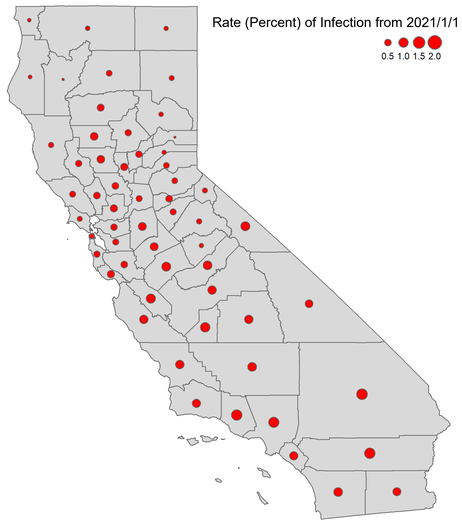
REFERENCES

- California Department of Public Health. *COVID-19 Time Series Metrics by County and State*. California Health and Human Services Agency. https://data.chhs.ca.gov/dataset/covid-19-time-series-metrics-by-county-and-state
- Monmonier, M. S. (2018). *How to lie with maps* (Third edition). Chicago; London: The University of Chicago Press. Intro, Ch. 1-3; p. 1-44.
- Data.census.gov

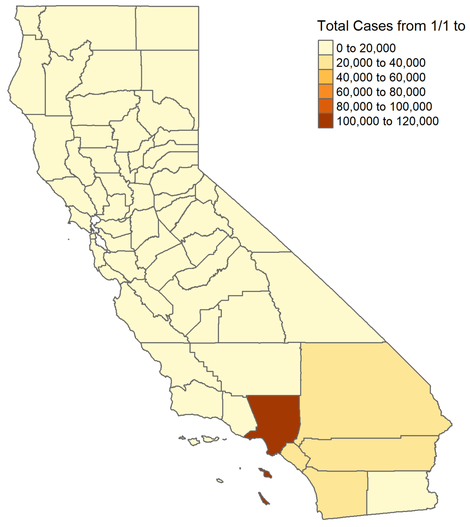
MAPS AND ANALYSIS



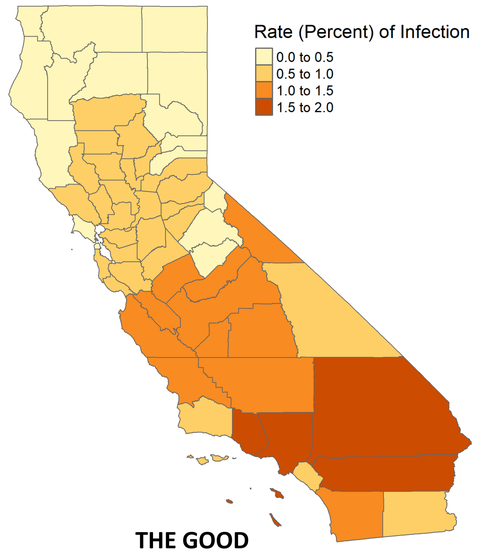
(1)



(2)



(3)



THE GOOD MAP

WHY IS THIS GOOD?

- Using **Choropleth** for measuring infection **intensity**
- **Classification breaks** are reasonable and does not cause **overgeneralization**
- **Simple**, only the **colors** convey the necessary information.
- Takes into account the background population for calculating the rates.

WHAT IS WRONG WITH THESE MAPS?

(1) Shapes and Sizes

- **Map (1)** is a cartogram. While the data is not distorted by the map, the state of California itself is distorted. **Cartograms** are maps that represent geographic data values by **distorting their shape and size** according to values that are being measured (Monmonier, 17). In this case, a county with a higher rate would undergo greater distortion.
- **Map (1)** is problematic because it misinforms the **shapes and sizes of counties**. With the shapes unrecognizable, it can be difficult to discern which county is which.
- Map looks **too complex** to understand. Audiences must deal with both **colors and area distortion** to interpret. That is excessive and unnecessary.

(2): Area Symbol vs. Choropleth

- **Map (2)** is communicating the intensity of COVID using **Area Symbols**. However, this type of data visualization is also problematic. Monmonier highlights the type of data visualization can also be an issue for maps. In this case, graduated symbols is a measurement of **magnitude**, not **intensity** which is what a rate measures (22-23). The result is that the map is **not effective**.
- With the range of rates being very small, the sizes of the symbols do not vary much. It is **difficult to determine** the highest rates of infection on this map. It could very well be most counties.

(3): Mismatch and Overgeneralization

- **Map (3)** is displaying the total the number of cases, or the overall magnitude of the data. A graduated symbol would be more fitting seen in **Map (2)**.
- One final variable a maps can distort is through **content generalization**. This is done through map classifications, especially in Choropleth maps. **Map (3)** shows how class breaks magnify/minimize the spread of a disease (40). Breaks distort the message of a Choropleth because it changes the patterns on the map.
- Data shows few counties have over 20,000 total cases. This gives the interpretation of Los Angeles County (dark red) as the only county undergoing infections and that the rest of California is safe.
- Important to note that LA County has the largest population at 12 million, which leads to high total cases

DISCUSSION AND CONCLUSION

The examples offered in this project only scratch the surface on the many forms of how a map can be distorted. Other elements that could be further distorted that have been brought up by Monmonier include:

- **Color**
- **Symbol Type**
- **Geometry Generalization**

Beyond the elements of a map, the disease data deserves its own analysis. The scale at which the data is gathered and mapped is at the county level and that is problematic. Some counties can be very large, like Los Angeles and San Bernadino. They mask the true extent of a disease at smaller scales like cities. Unfortunately, more accurate data will come at the cost of privacy, and that is another issue for the future.

Maps are not objective tools. They can be distorted to the author's choice. However, that does not mean we should disregard them completely. **By understanding how the elements of a map can be distorted, we can overcome those biases and understand the world.**