

# Uncharted Territory: An Introduction to Making Clear and Accurate Maps

By: Ayanthi Gunawardana

R-Ladies Meetup

May 21, 2020

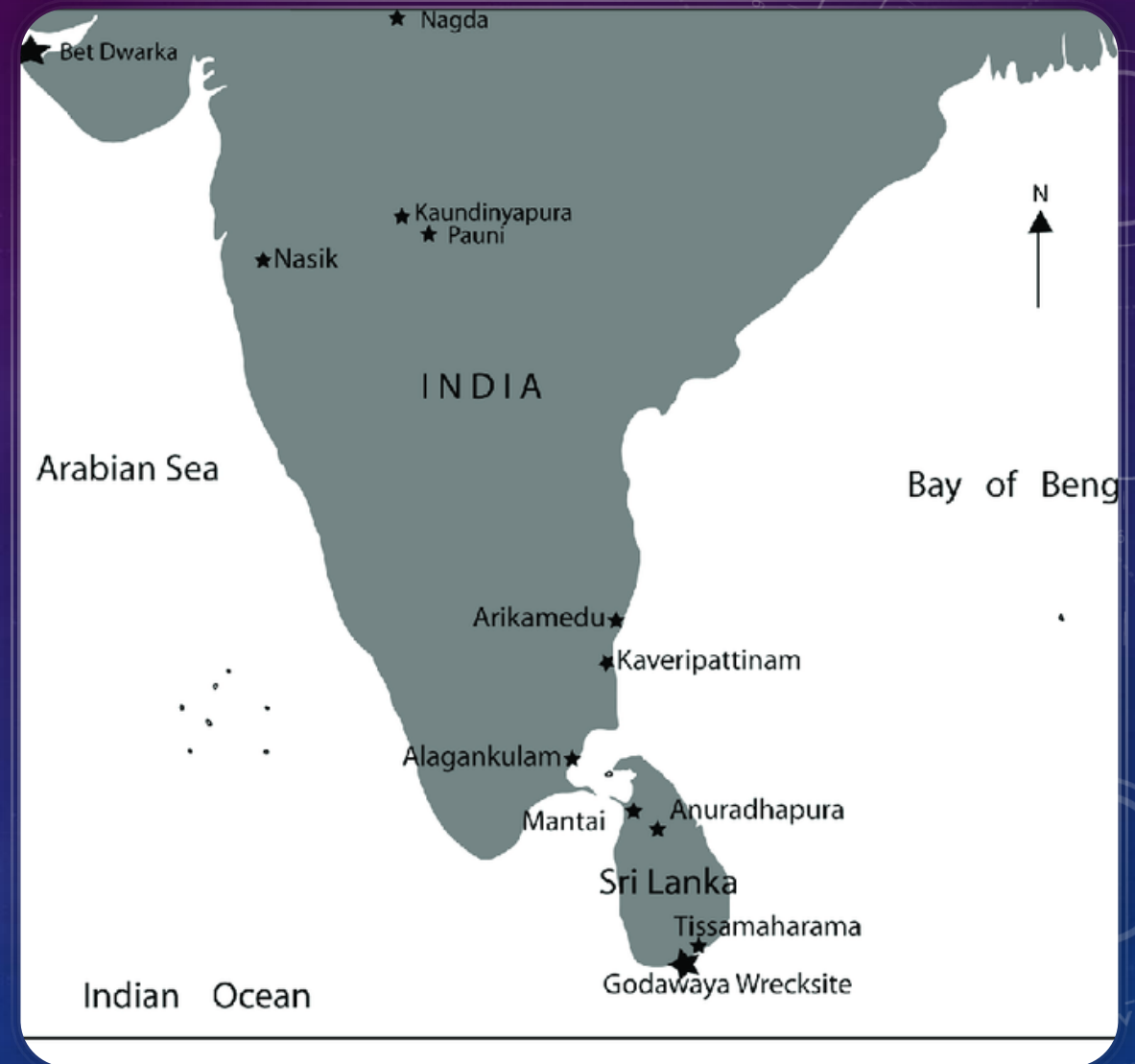


[https://images-na.ssl-images-amazon.com/images/I/51iF%2BpUk3wL\\_AC\\_.jpg](https://images-na.ssl-images-amazon.com/images/I/51iF%2BpUk3wL_AC_.jpg)

# Overview

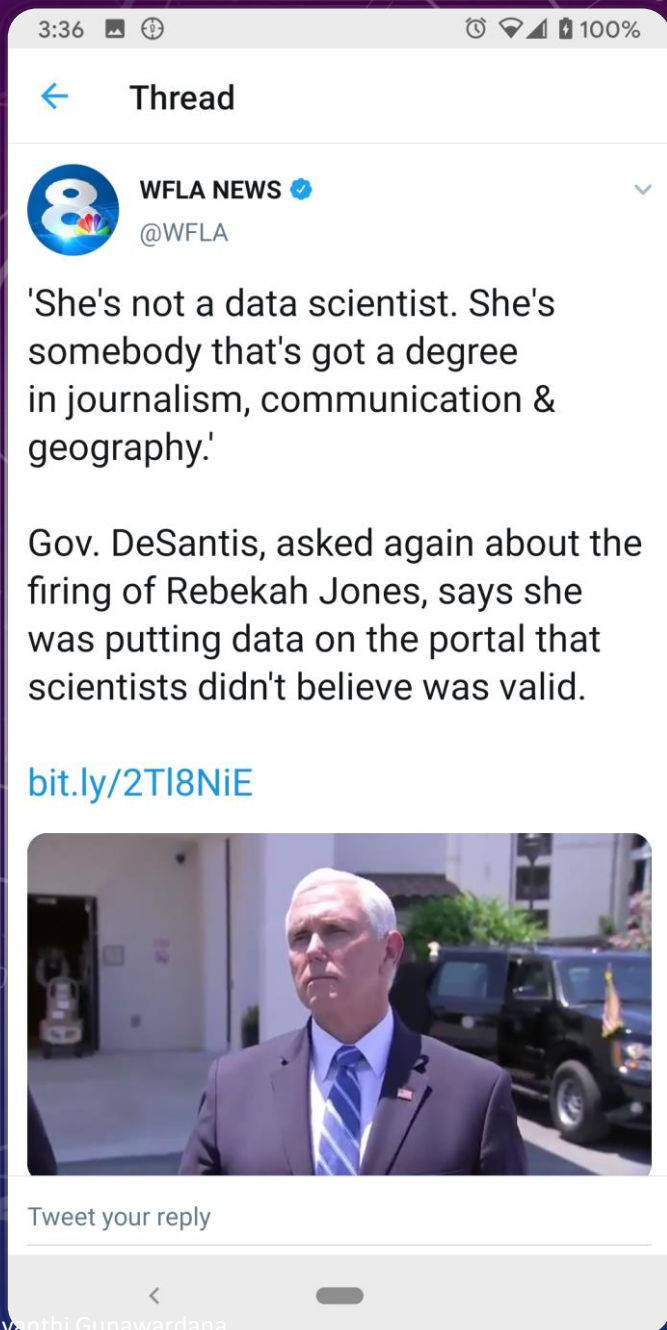
- Mapping Terminology
- Data Types
- Coordinate Systems and Projections
- Visual Map Elements
- Bad Map Samples

One of the many reasons I like maps is because my parents are from a country (Sri Lanka) that is frequently left out of conventional maps!



[https://www.researchgate.net/profile/Rasika\\_Muthucumarana/publication/261016208/figure/fig1/AS:717197353689089@1548004582193/Map-of-India-and-Sri-Lanka-with-the-locations-of-major-archaeological-sites.png](https://www.researchgate.net/profile/Rasika_Muthucumarana/publication/261016208/figure/fig1/AS:717197353689089@1548004582193/Map-of-India-and-Sri-Lanka-with-the-locations-of-major-archaeological-sites.png)





Geography is a form of data science,  
no matter what the haters think

# Mapping Terminology

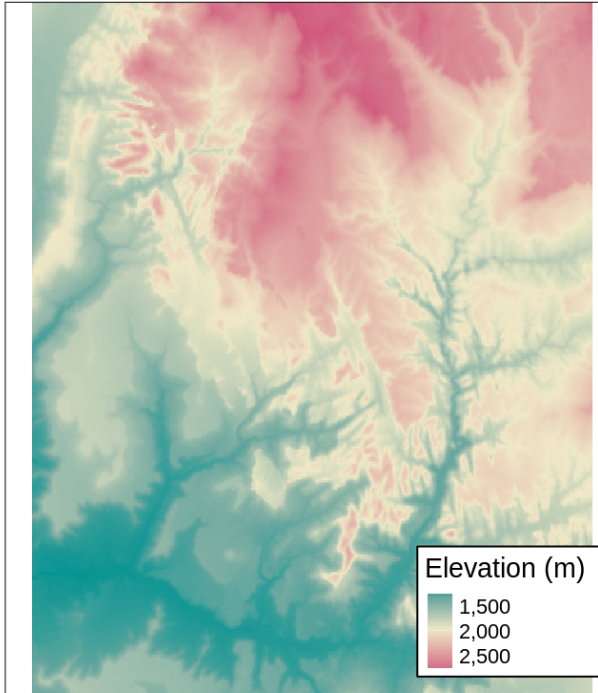
- Geographic Information System (GIS):
  - a framework for gathering, managing, and analyzing data. It analyzes spatial location and organizes layers of information into visualizations using maps and 3D scenes. GIS reveals deeper insights into data, such as patterns, relationships, and situations—helping users make smarter decisions.
  - (<https://www.esri.com/en-us/what-is-gis/overview>)
- Cartography:
  - the science or practice of drawing maps
- Shapefile:
  - Vector data format for GIS software.
  - Represents points, polygons, or lines; contains attribute and geospatial information about each feature
  - As GIS becomes more open source, other formats are becoming more popular, like Geopackage, GeoJSON, KML



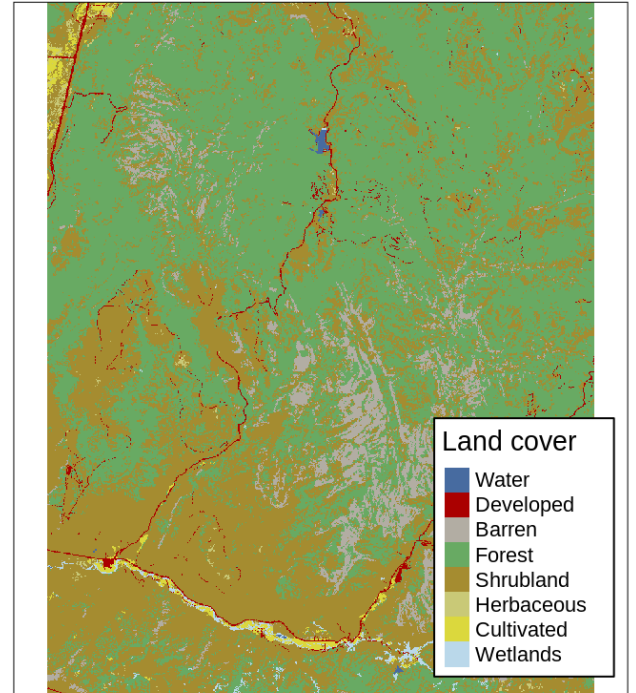
# Vector Data vs. Raster Data

- Vector Data
  - Consists of points, lines, and polygons
  - Each feature type needs to be stored separately
- Raster Data
  - Array of cells (pixels) used to represent real-world objects
  - Each pixel has its own value
  - Example: orthography
    - aerial images of land that have been geometrically corrected so the scale is uniform
  - Example: digital elevation model (DEM)
    - 3D representation of the Earth's surface using Lidar, radar, and other remote sensing tools

A. Continuous data



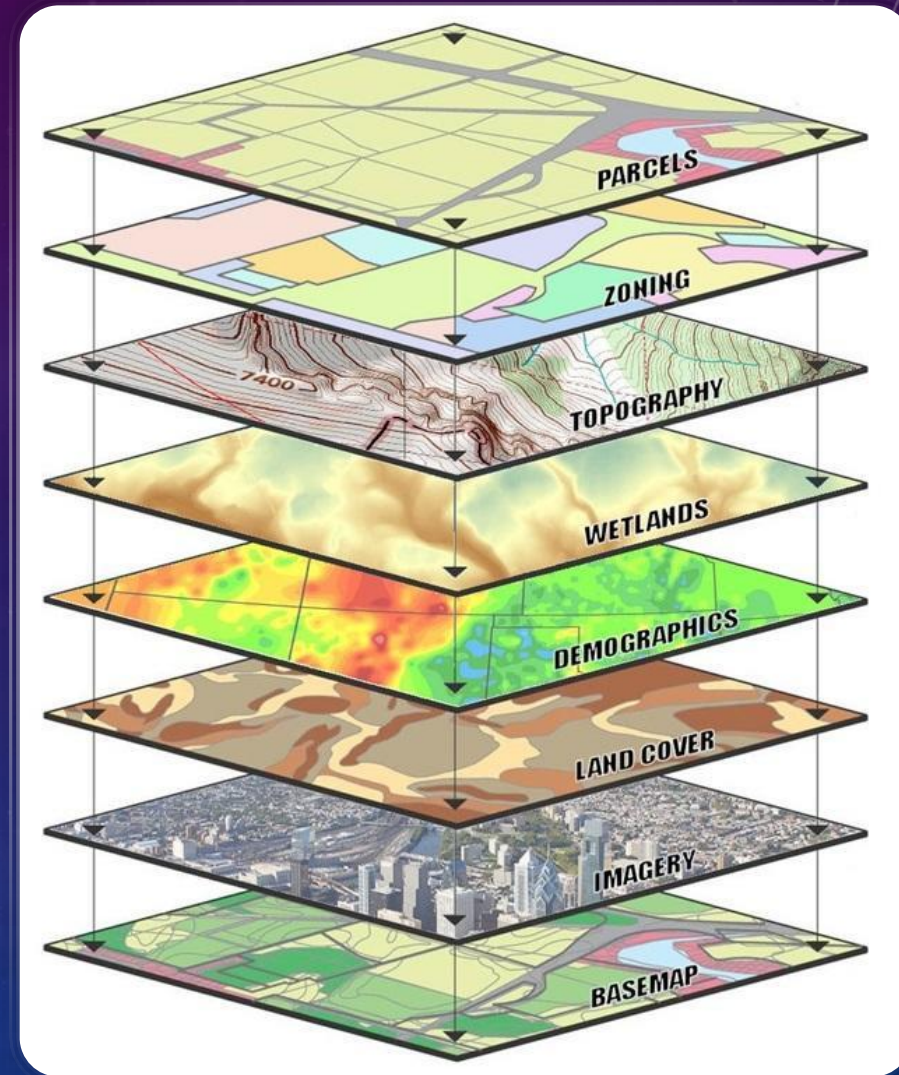
B. Categorical data



<https://geocompr.robinlovelace.net/spatial-class.html#raster-data>

# Data Representation

- Different types of data are represented in different layers
- Using these layers in a series of maps allow us to see the “whole picture”
- Spatial information
  - Geometry
  - Location
- Non-spatial information
  - Attribute information for each feature
  - Example: population in a specific census tract; population itself is not spatial, but is aggregated to a spatial entity

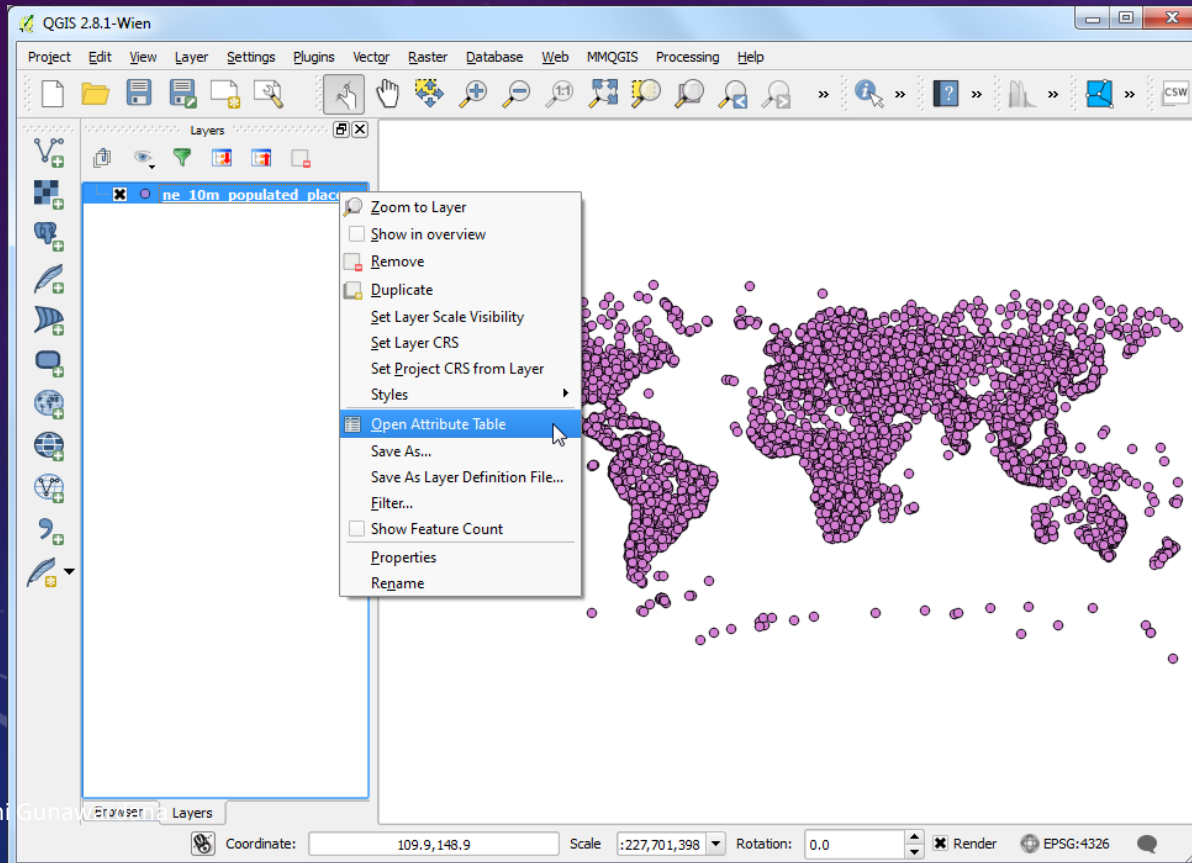


<https://www.usgs.gov/media/images/gis-data-layers-visualization>



# Attribute Table

- The back end of a shapefile has an attribute table, which contains data represented by the geography
- Each column represents a different type of data; a shapefile can only show one column at once
- You can create new columns/data from existing columns by doing calculations in GIS or R



Attribute table - ne\_10m\_populated\_places\_simple :: Features total: 7322, filtered: 7322, selected: 0

	scalerank	natscale	labelrank	featurecla	name	namepar	namealt
0	10	1	8	Admin-1 capital	Colonia del Sacra...	NULL	NULL
1	10	1	8	Admin-1 capital	Trinidad	NULL	NULL
2	10	1	8	Admin-1 capital	Fray Bentos	NULL	NULL
3	10	1	8	Admin-1 capital	Canelones	NULL	NULL
4	10	1	8	Admin-1 capital	Florida	NULL	NULL
5	10	1	8	Admin-1 capital	Bassar	NULL	NULL
6	10	1	8	Admin-1 capital	Sotouboua	NULL	NULL
7	10	1	7	Admin-1 capital	Medenine	NULL	NULL
8	10	1	7	Admin-1 capital	Kebili	NULL	NULL
9	10	1	7	Admin-1 capital	Tataouine	NULL	NULL
10	10	1	7	Admin-1 capital	L'Ariana	NULL	NULL
11	10	1	7	Admin-1 capital	Jendouba	NULL	NULL
12	10	1	7	Admin-1 capital	Kasserine	NULL	NULL
13	10	1	7	Admin-1 capital	Sdid Bouzid	NULL	NULL
14	10	1	7	Admin-1 capital	Siliana	NULL	NULL
15	10	1	7	Admin-1 capital	Mahdia	NULL	NULL
16	10	1	7	Admin-1 capital	Monastir	NULL	NULL
17	10	1	7	Admin-1 capital	Zaghuan	NULL	NULL
18	10	1	5	Admin-1 capital	Tay Ninh	NULL	NULL

Show All Features

[https://www.qgistutorials.com/en/docs/working\\_with\\_attributes.html](https://www.qgistutorials.com/en/docs/working_with_attributes.html)

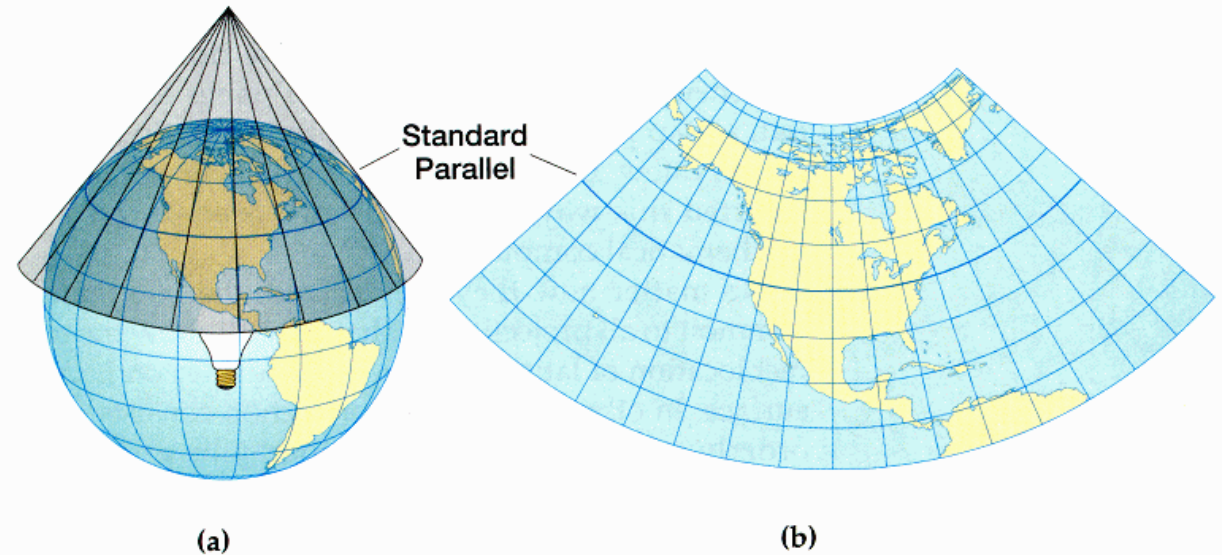


# Geographic Coordinate System vs. Projected Coordinate System

- “Maps are scale models of reality” – Mark Monmonier
- A map is a visual representation of reality, and is inherently distorted
  - You’re taking a piece of the globe and flattening it to create a map
- A geographic coordinate reference system describes where a location is on Earth
  - Since the Earth is round, this is measured in angular units, usually degrees
- A projected coordinate system is a geographic coordinate reference system that has been flattened
  - When using a projected coordinate system, you can measure distances using a straight line

# Conic Projection

- The point where the cone touches the globe is known as the standard parallel
- Since a cone is wider than it is tall, the map becomes more distorted as you move north or south of the standard parallel

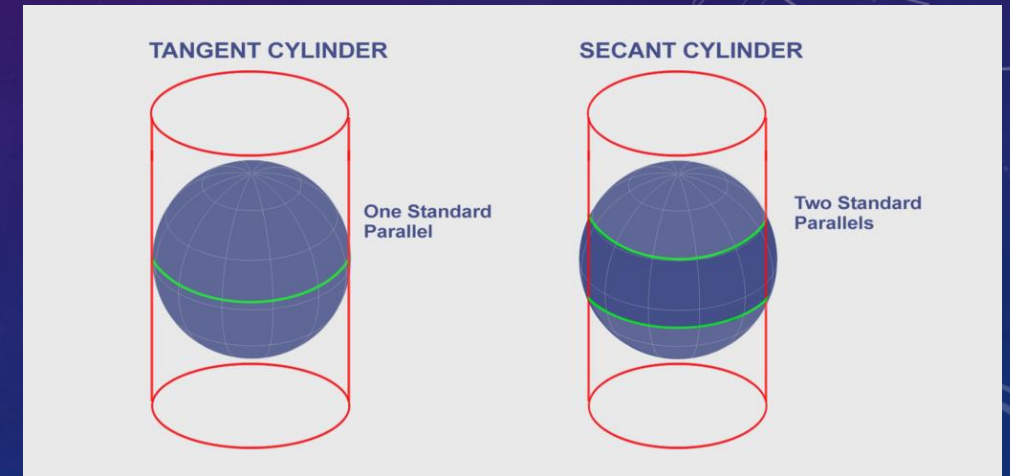


**Figure 2-8** (a) The projection theory, as illustrated by a globe with a light in its center, projecting images onto a cone. (b) The resultant map is called a conic projection.

<http://faculty.kutztown.edu/courtney/blackboard/Physical/05Project/conproj.html>

# Cylindrical Projection

- Commonly used for world maps
- The globe is represented as a cylinder that is tangent or secant to a selected line, and then flattened
- Tangent vs. secant (<http://geokov.com/education/map-projection.aspx>)
  - Tangent: cylinder's circumference touches the reference globe's surface;
  - Secant: cylinder intersects the globe; At the place where the cylinder cuts through the globe two secant lines are formed.
- The scale is constant at the tangent and secant lines, but the farther away you get from these lines, the more distorted the map becomes

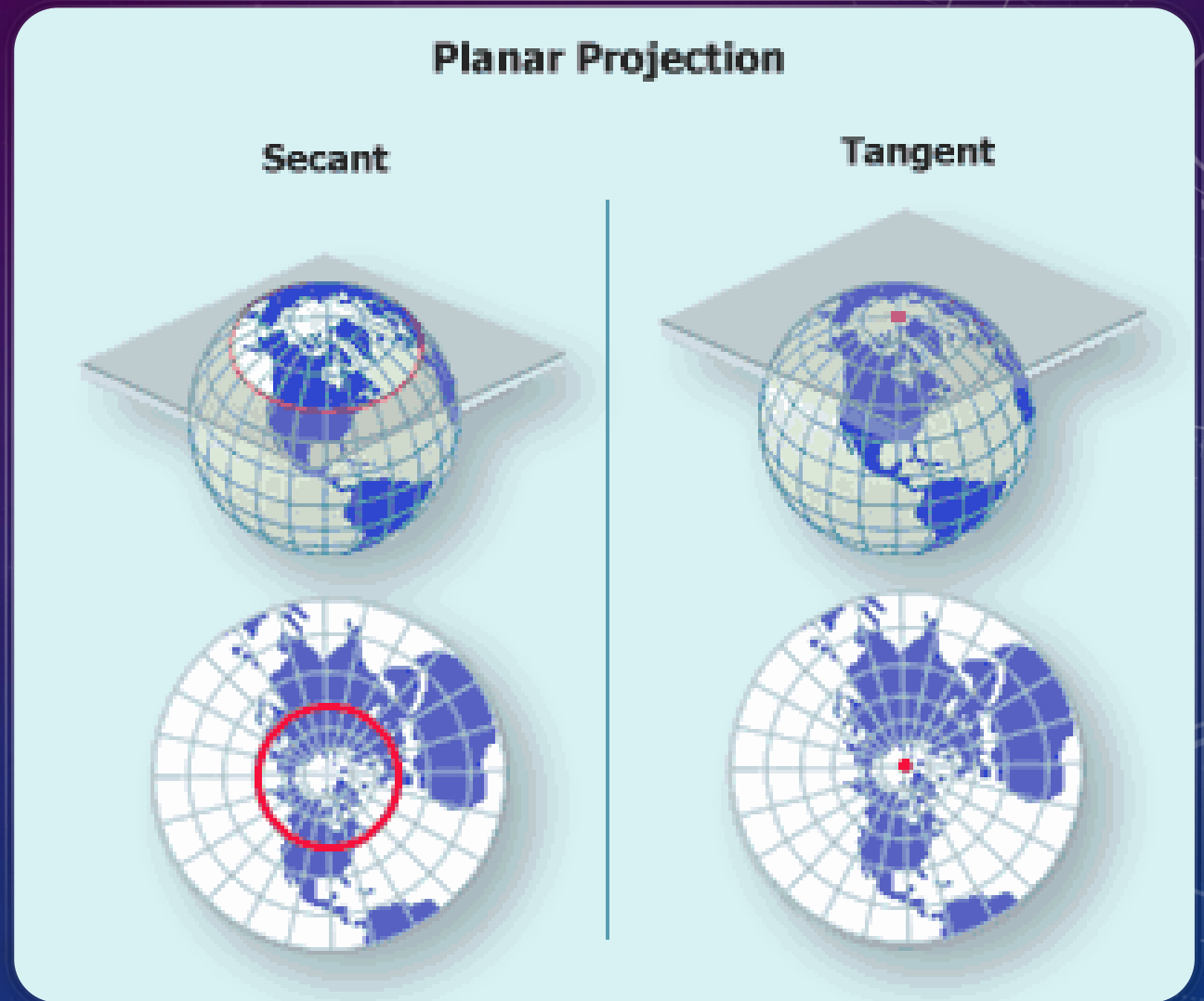


<https://gisgeography.com/cylindrical-projection/>



# Planar Projection

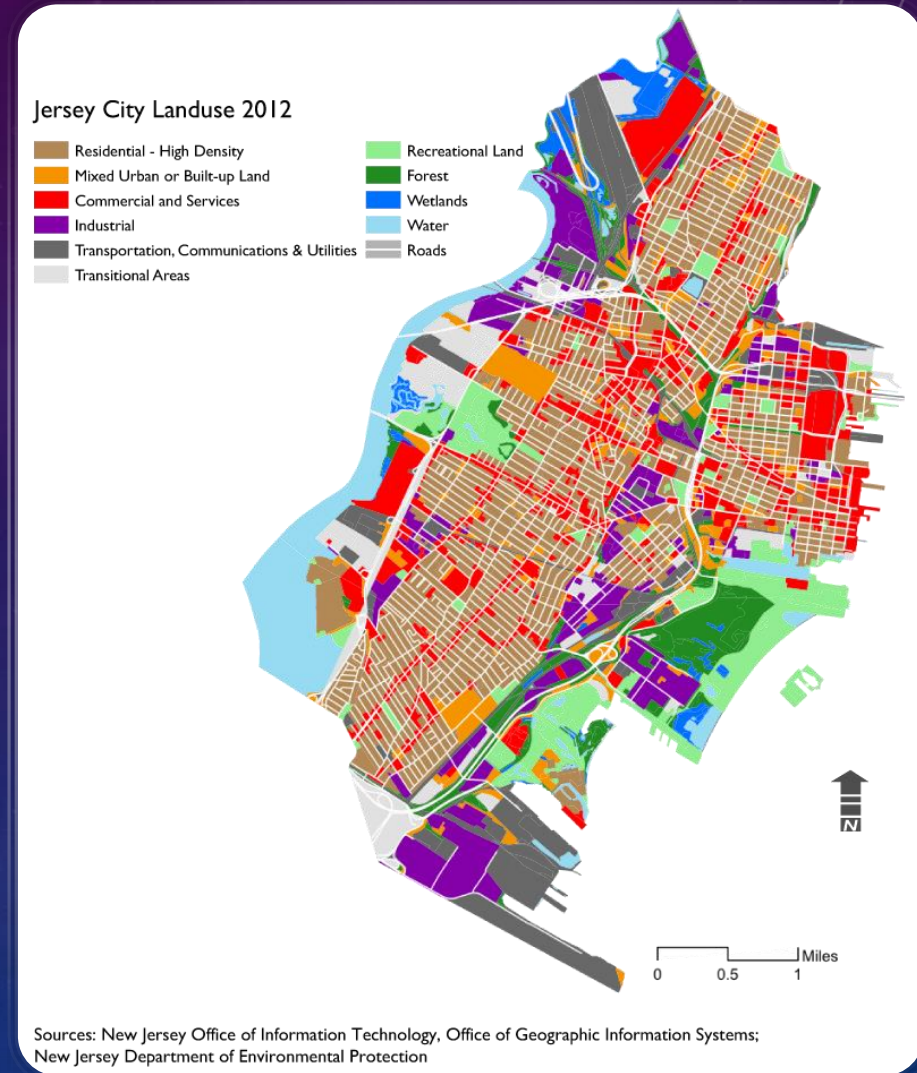
- Commonly used to highlight a specific area
- This plane is tangent to the globe at a single point
- Good for preserving distances
- However, distortions increase as you increase the size of the map and move away from the center



USGS

# Visual Map Elements

- Critical Elements
  - Map
  - Legend
  - Title
  - Data sources
  - Labels
- Optional (but recommended)
  - North arrow
  - Scale bar



<http://njbikeped.org/jersey-citys-new-citi-bike-stations-how-to-benefit-all-neighborhoods-equally/>

# Best Practices in Visualizing Geospatial Data

- Always normalize your data
  - Unequal geographic boundaries, areas, and populations can make data seem concentrated when it's just a reflection of the population
- Label critical elements only
  - Over-labeling makes a map hard to read
- Think critically about your color palette
  - Does your organization have its own palette?
  - Accounting for color-blindness
  - Industry standards
- Organize your shapefiles and document all analyses
  - You can add new fields to the back end of your shapefile, but if you're using dedicated GIS software, you may not remember what tools you used
    - Ex: how did I calculate population by area?
- Scale
  - Make sure you have a scale bar to show what area your map is covering



# What Do We Do With All This Data?

We can use any combination of this data to make a good map, but using too much data or the wrong type of data will ruin a map!

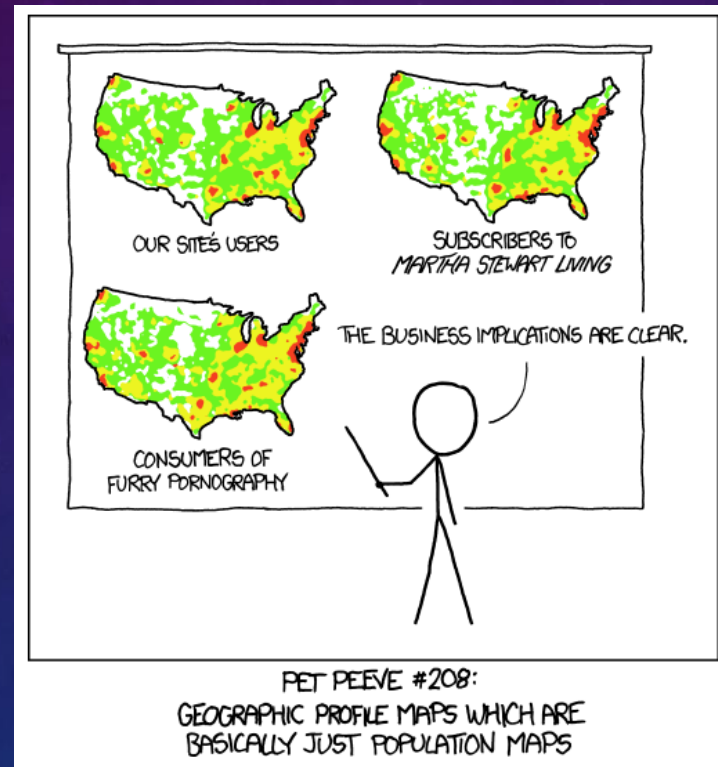
The following slides contain examples of bad and/or misleading maps and cartography practices.

# Fun, But Chaotic



[https://www.reddit.com/r/terriblemaps/comments/fj7ih8/us state map marked with a tv show that takes/](https://www.reddit.com/r/terriblemaps/comments/fj7ih8/us_state_map marked with a tv show that takes/)

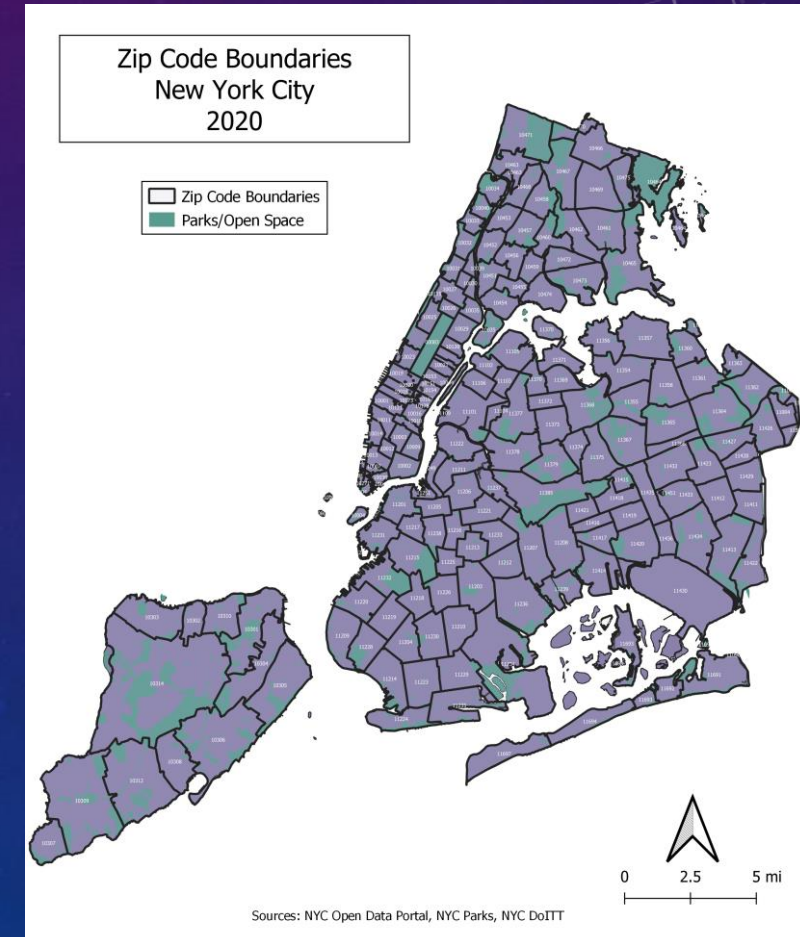
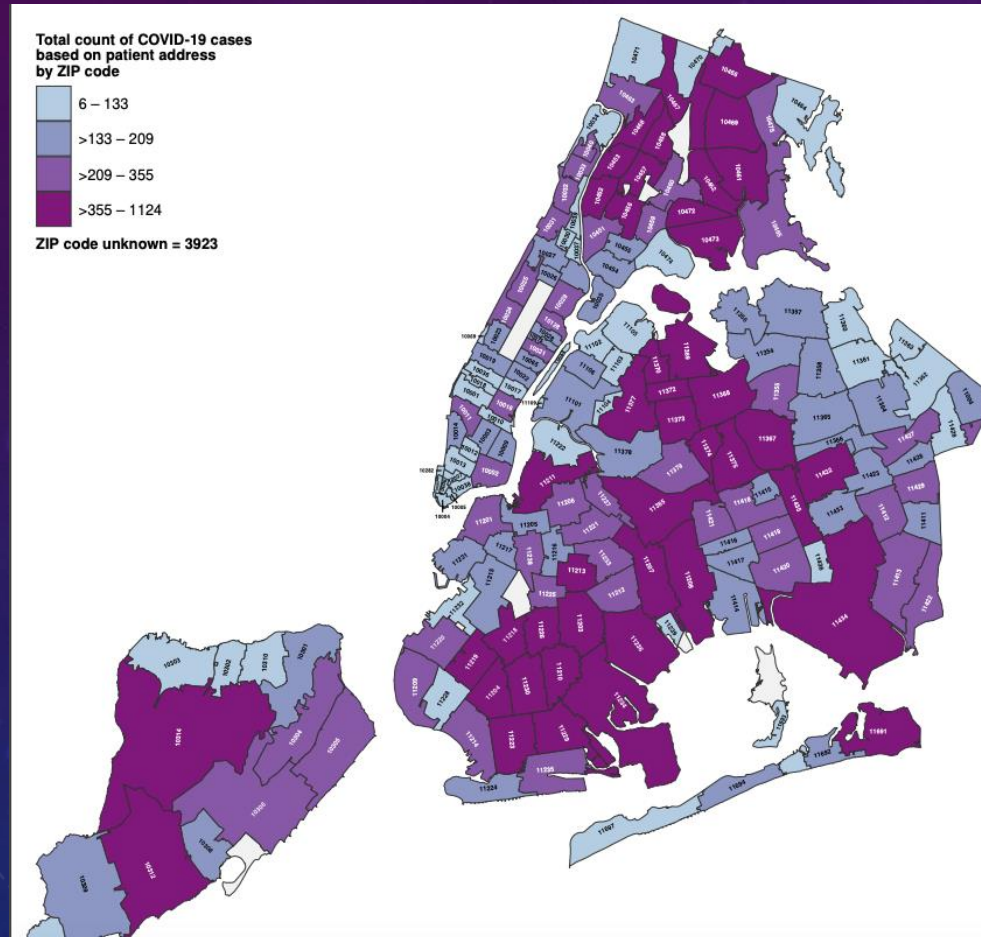
# Sometimes It's Just a Population Map



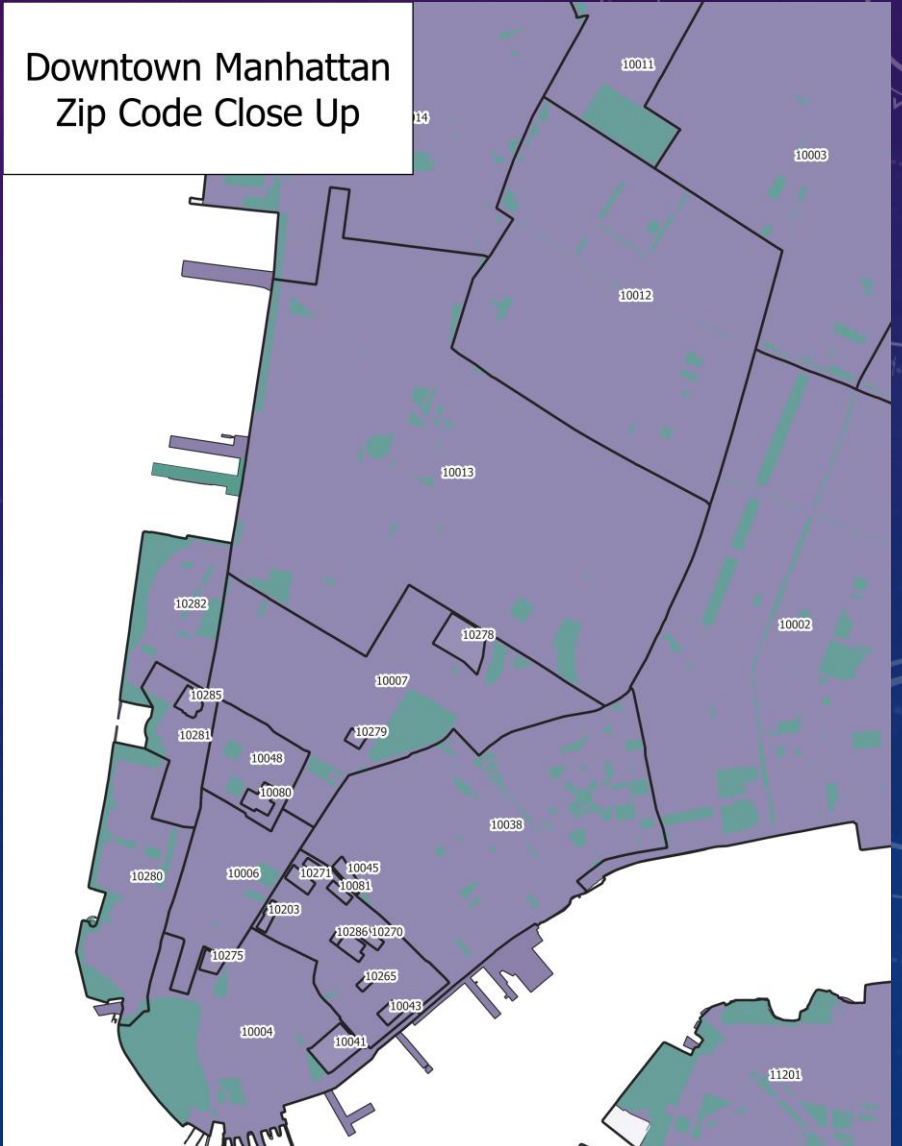
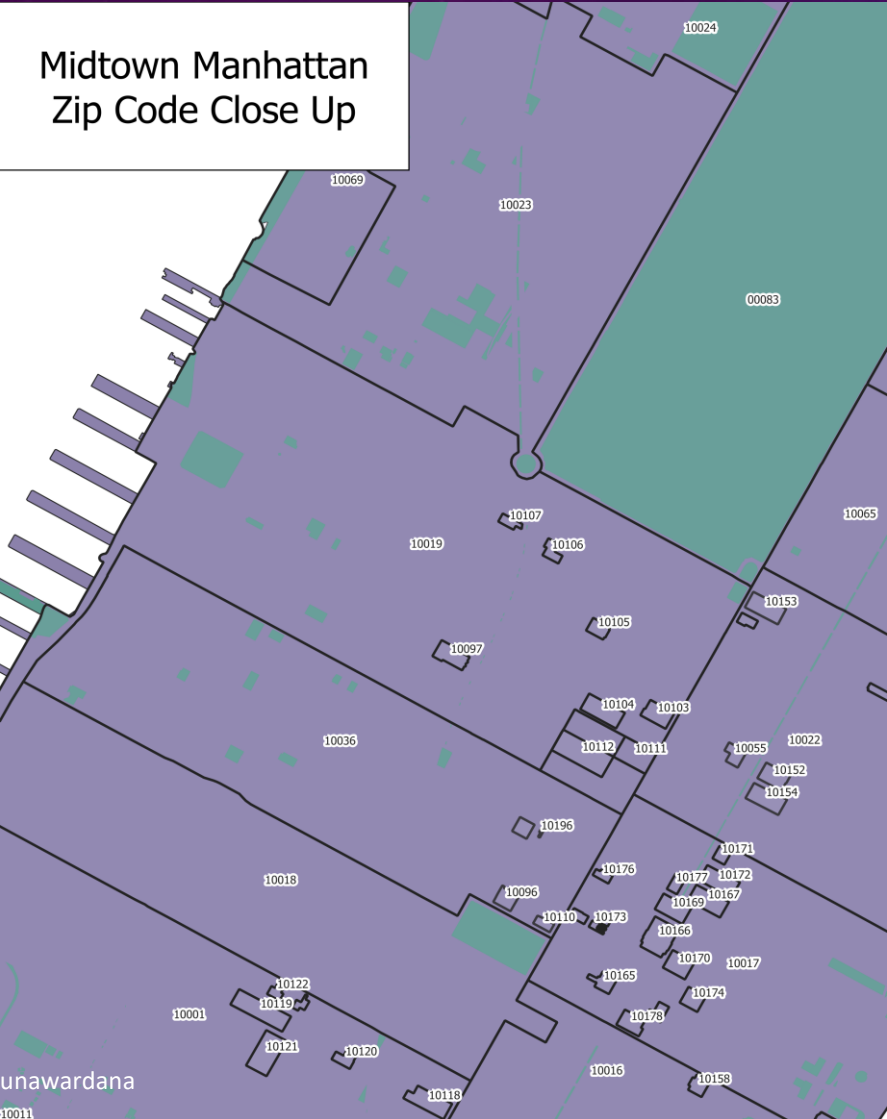
xkcd



# Mapping Raw Counts and Bad Boundaries

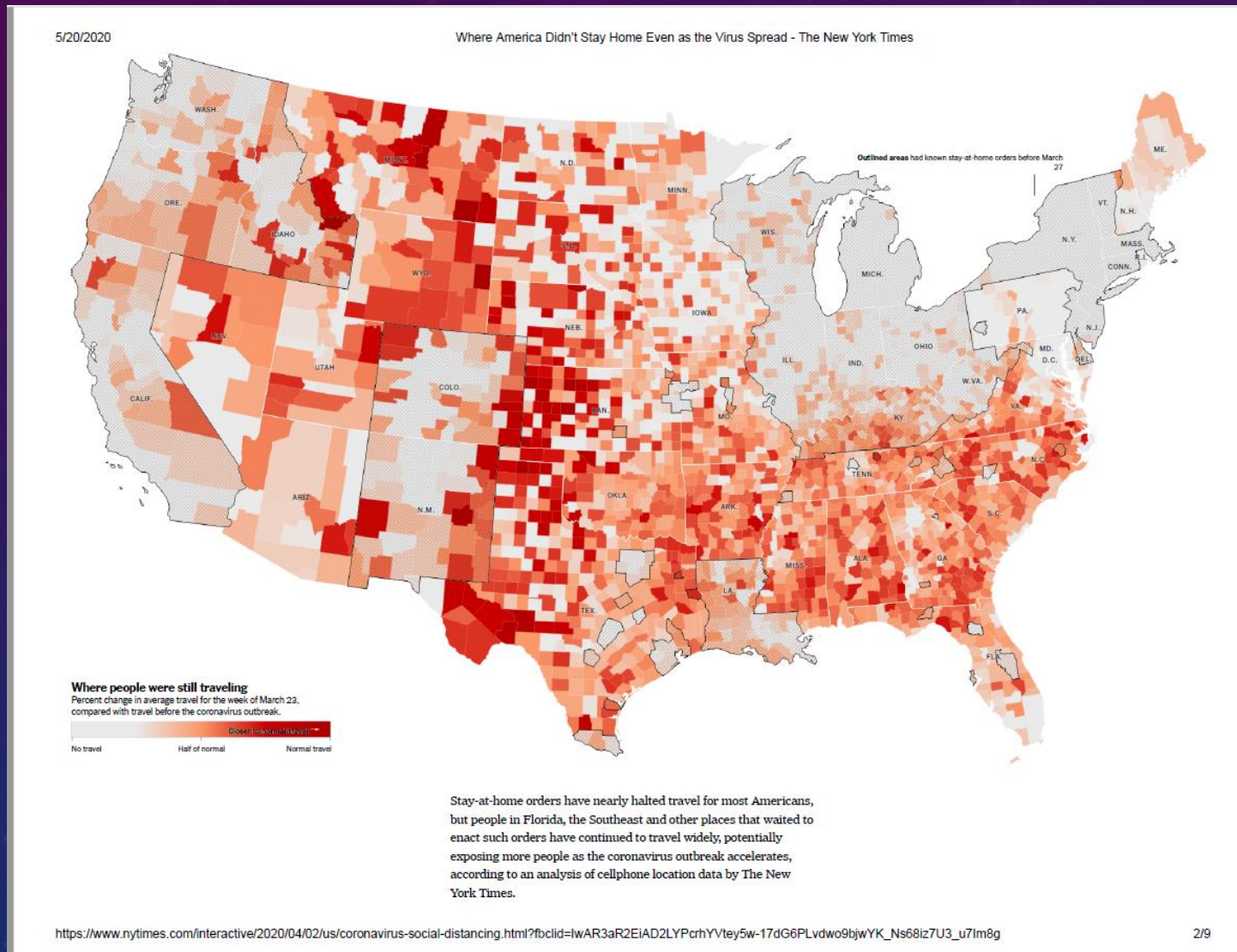


# The Chaos of Zip Codes



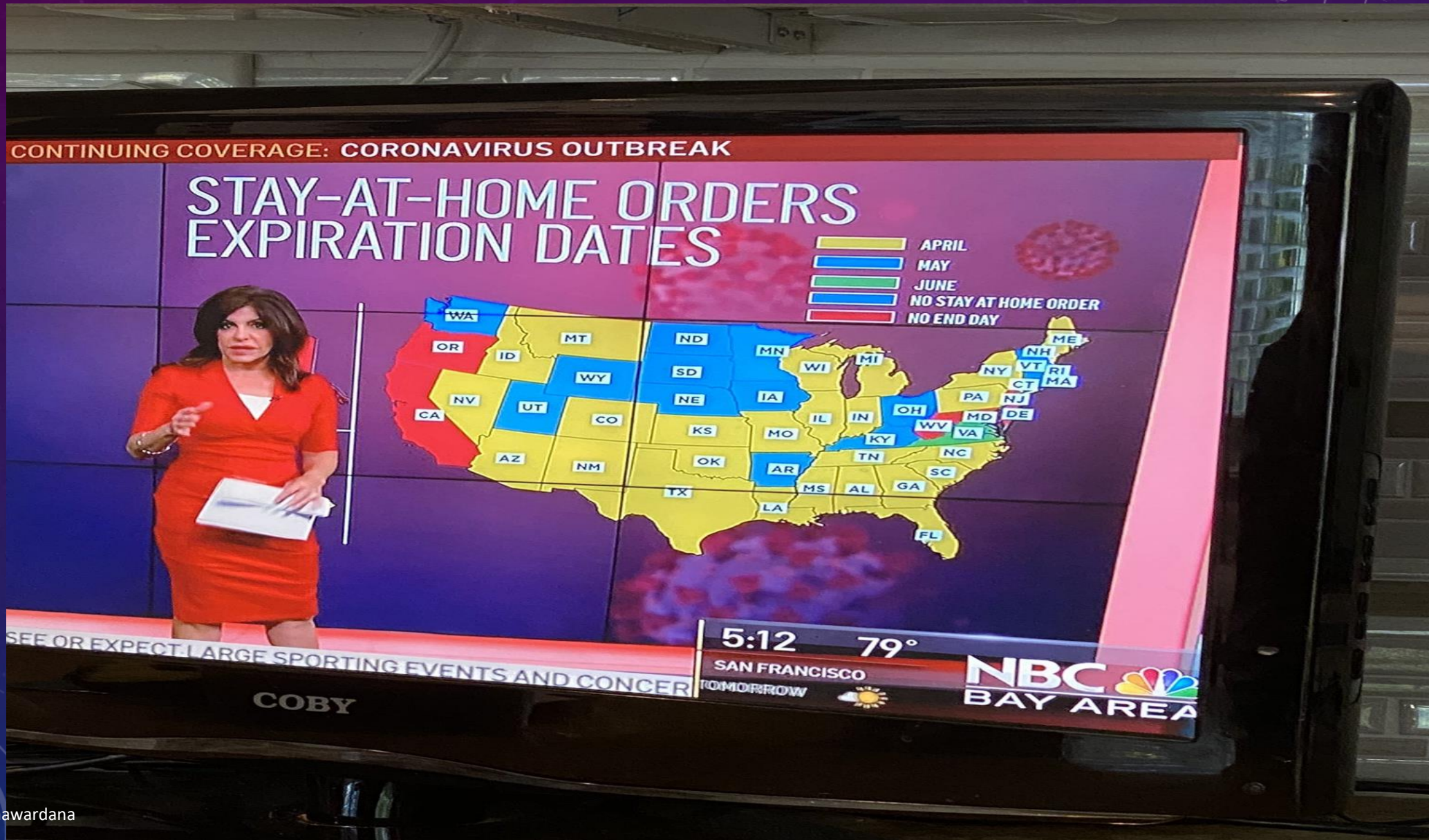


# Rural vs. Urban Areas





Just....BAD



# Useful Resources

- Books
  - “How to Lie With Maps” by Mark Monmonnier
  - “Geographic Information Systems and Science” by Longley, Goodchild, Maguire, and Rhind
- GIS Organizations/Companies
  - USGS – United States Geological Survey
    - <https://www.usgs.gov/>
  - ESRI – maker of proprietary GIS software
    - <https://www.esri.com/en-us/home>
  - QGIS – open source GIS software
    - <https://qgis.org/en/site/>
- NYC Open Data Portal – has several geospatial datasets to get you started with analysis
  - <https://opendata.cityofnewyork.us/>



# Thank You for Listening – Mischief Managed!

