

Lecture 23

Dot Maps

Dennis Sun
Stanford University
DATASCI / STATS 112

March 10, 2023



- 1 Review
- 2 Dot Maps
- 3 Storytelling with Dot Maps
- 4 Making Dot Maps in Python
- 5 Caution about Dot Maps
- 6 Reminders



- 1 Review
- 2 Dot Maps
- 3 Storytelling with Dot Maps
- 4 Making Dot Maps in Python
- 5 Caution about Dot Maps
- 6 Reminders



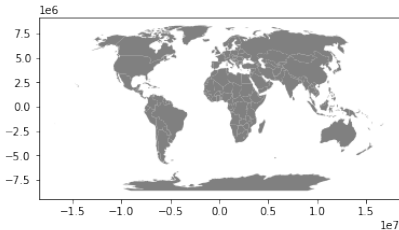
Making Maps

Last class, we learned to make maps using Geopandas.

```
import geopandas as gpd

gdf_countries = gpd.read_file(
    "/usr/local/lib/python3.9/dist-packages/geopandas/datasets/"
    "naturalearth_lowres/")
gdf_countries = gdf_countries.to_crs("ESRI:53030")

gdf_countries.plot(facecolor="gray")
```



Today: Adding data to a map.



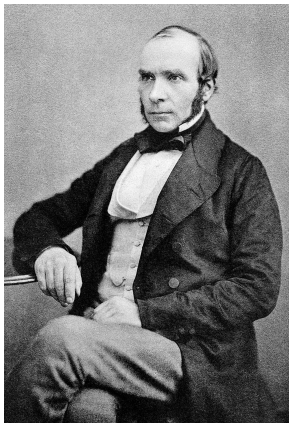
- 1 Review
- 2 Dot Maps
- 3 Storytelling with Dot Maps
- 4 Making Dot Maps in Python
- 5 Caution about Dot Maps
- 6 Reminders



Who is John Snow?



Jon Snow
(281–302 AC)



John Snow
(1813–1858)

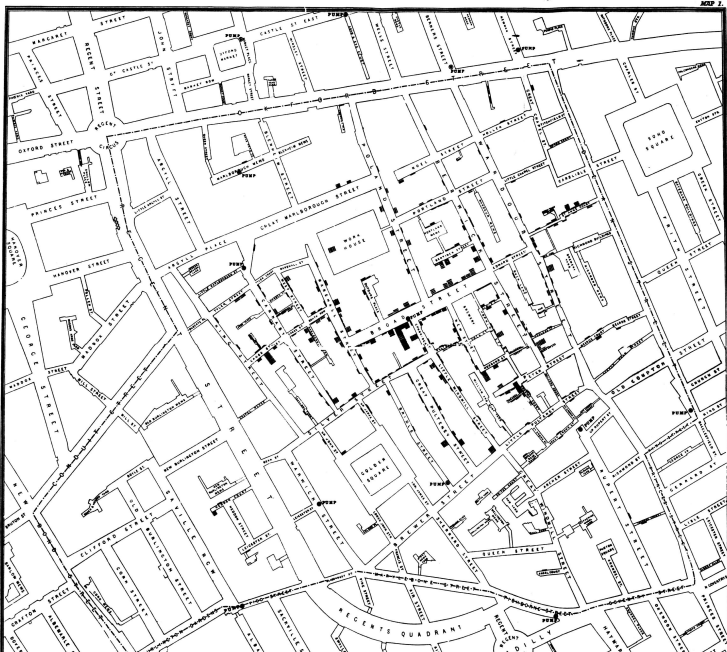


1854 Broad Street Cholera Outbreak

- In 1854, London was hit by a severe cholera outbreak.
- At the time, the cause of cholera was not known. There were two theories: the germ theory and the miasma theory.
- John Snow decided to investigate the cause, and he started by making a **dot map**.



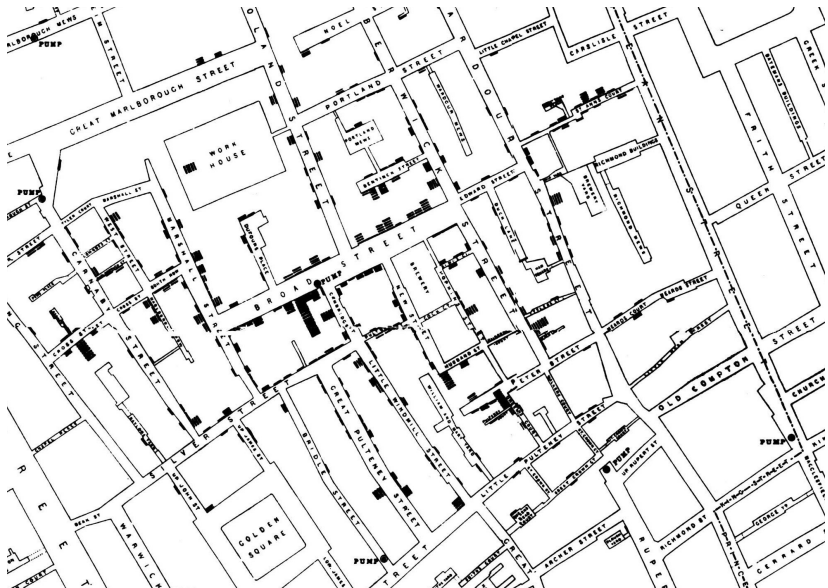
John Snow's Dot Map





Each "dot" (thin black box) represents a cholera case.

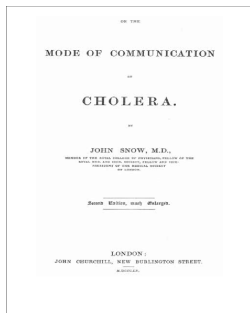




Snow observed that cholera cases centered around the Broad Street water pump.

Snow's Analysis

Snow followed up on the insight with careful, on-the-ground detective work.



There were only ten deaths in houses situated decidedly nearer to another street-pump. In five of these cases the families of the deceased persons informed me that they always sent to the pump in Broad Street, as they preferred the water to that of the pumps which were nearer. In three other cases, the deceased were children who went to school near the pump in Broad Street....



The End of the Story

In the end, Snow was able to build a strong case that the Broad Street pump was the source of the cholera outbreak.

The result of the inquiry, then, is, that there was been no particular outbreak or prevalence of cholera in this part of London except among the persons who were in the habit of drinking the water of the above-mentioned pump well....In consequence of what I said, the handle of the pump was removed on the following day.



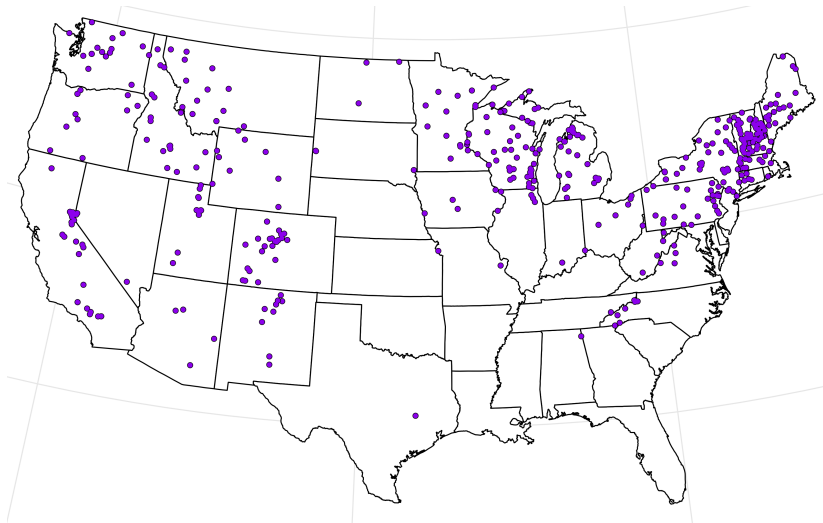
Snow's analysis established cholera as a waterborne disease and affirmed the germ theory of disease.



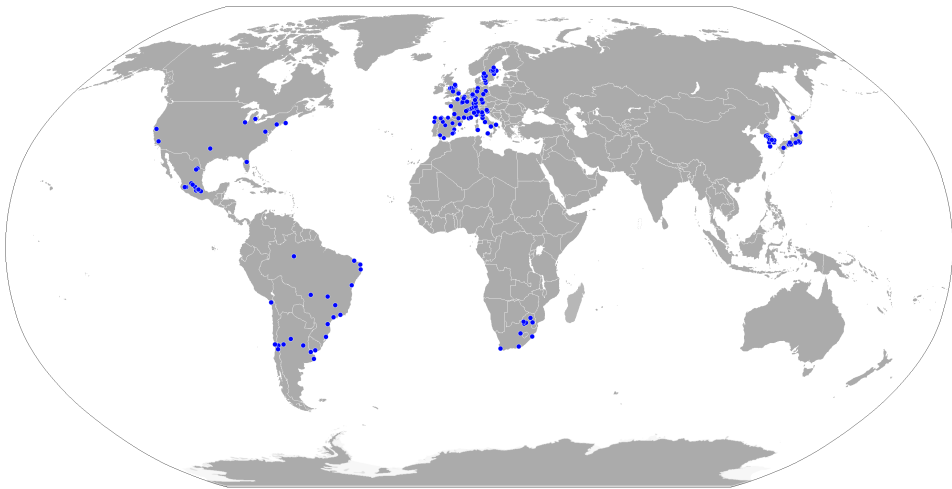
- 1 Review
- 2 Dot Maps
- 3 Storytelling with Dot Maps**
- 4 Making Dot Maps in Python
- 5 Caution about Dot Maps
- 6 Reminders



What data is being plotted here?



What data is being plotted here?



- 1 Review
- 2 Dot Maps
- 3 Storytelling with Dot Maps
- 4 Making Dot Maps in Python**
- 5 Caution about Dot Maps
- 6 Reminders



Reading in Earthquakes Data

We will work with a dataset of earthquakes worldwide on June 4, 2018.

```
import pandas as pd
data_dir = "https://dlsun.github.io/pods/data/"
df_quakes = pd.read_csv(data_dir + "earthquakes.csv")
df_quakes
```

	time	latitude	longitude	...	status	locationSource	magSource
0	2018-06-05T17:51:13.660Z	19.407833	-155.282837	...	automatic	hv	hv
1	2018-06-05T17:46:26.600Z	35.378333	-117.858333	...	automatic	ci	ci
2	2018-06-05T17:46:24.020Z	38.803665	-122.740837	...	automatic	nc	nc
...
538	2018-06-04T18:24:37.410Z	-7.055000	123.203900	...	reviewed	us	us
539	2018-06-04T18:20:04.548Z	37.160100	-117.552900	...	reviewed	nn	nn
540	2018-06-04T18:10:35.980Z	46.873000	-112.521167	...	reviewed	mb	mb

541 rows x 22 columns



Converting the DataFrame to a GeoDataFrame

We want to plot the earthquakes data on a map, which requires converting the `DataFrame` to a `GeoDataFrame`.

```
gdf_countries = gpd.read_file(
    "/usr/local/lib/python3.9/dist-packages/geopandas/datasets/"
    "naturalearth_lowres/")
gdf_quakes = gpd.GeoDataFrame(
    df_quakes,
    geometry=gpd.points_from_xy(df_quakes["longitude"],
                                df_quakes["latitude"])
).set_crs(gdf_countries.crs)
gdf_quakes
```

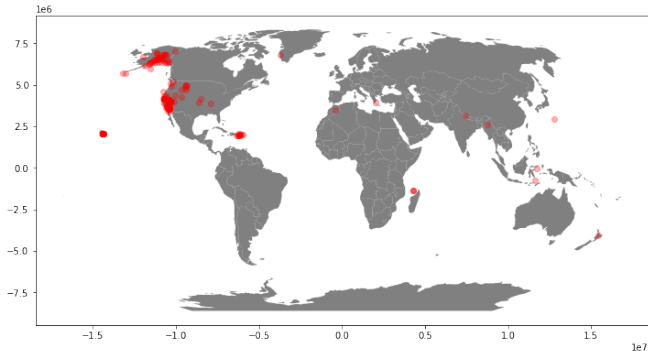
	time	latitude	longitude	...	locationSource	magSource	geometry
0	2018-06-05T17:51:13.660Z	19.407833	-155.282837	...	hv	hv	POINT (-155.28284 19.40783)
1	2018-06-05T17:46:26.600Z	35.378333	-117.858333	...	ci	ci	POINT (-117.85833 35.37833)
2	2018-06-05T17:46:24.020Z	38.803665	-122.740837	...	nc	nc	POINT (-122.74084 38.80367)
...
538	2018-06-04T18:24:37.410Z	-7.055000	123.203900	...	us	us	POINT (123.20390 -7.05500)
539	2018-06-04T18:20:04.548Z	37.160100	-117.552900	...	nn	nn	POINT (-117.55290 37.16010)
540	2018-06-04T18:10:35.980Z	46.873000	-112.521167	...	mb	mb	POINT (-112.52117 46.87300)

541 rows x 23 columns



Plotting the Data on Top of the Map

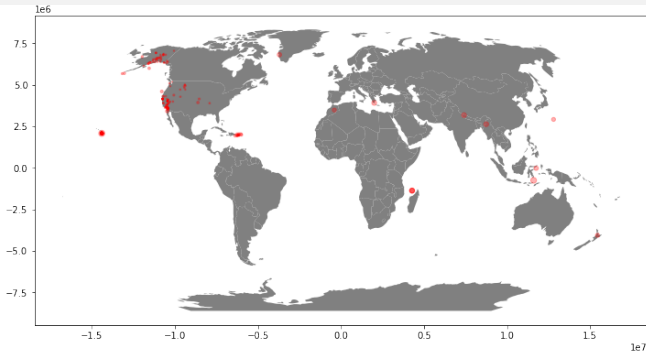
```
ax = gdf_countries.to_crs("ESRI:53030").plot(  
    facecolor="gray", figsize=(12, 12))  
gdf_quakes.to_crs("ESRI:53030").plot(  
    ax=ax, color="red", alpha=0.3)
```



Representing Another Variable by Size

The dot map above treats every earthquake the same, no matter how small. What if we wanted to indicate larger earthquakes by larger points?

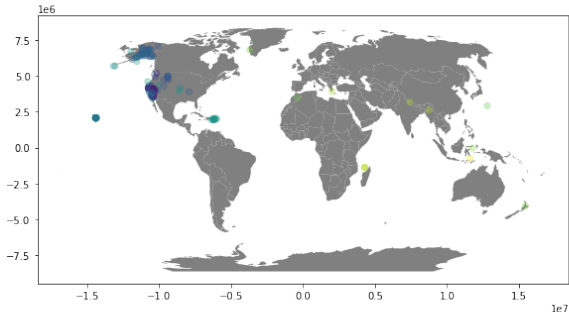
```
ax = gdf_countries.to_crs("ESRI:53030").plot(
    facecolor="gray", figsize=(12, 12))
gdf_quakes.to_crs("ESRI:53030").plot(
    ax=ax, color="red", alpha=0.3,
    markersize=2 ** gdf_quakes["mag"])
```



Representing Another Variable by Color

We can also represent a variable by color.

```
ax = gdf_countries.to_crs("ESRI:53030").plot(  
    facecolor="gray", figsize=(12, 12))  
gdf_quakes.to_crs("ESRI:53030").plot(  
    ax=ax, alpha=0.3,  
    column="mag", legend=True)
```



- 1 Review
- 2 Dot Maps
- 3 Storytelling with Dot Maps
- 4 Making Dot Maps in Python
- 5 Caution about Dot Maps**
- 6 Reminders



Bad Dot Map #1

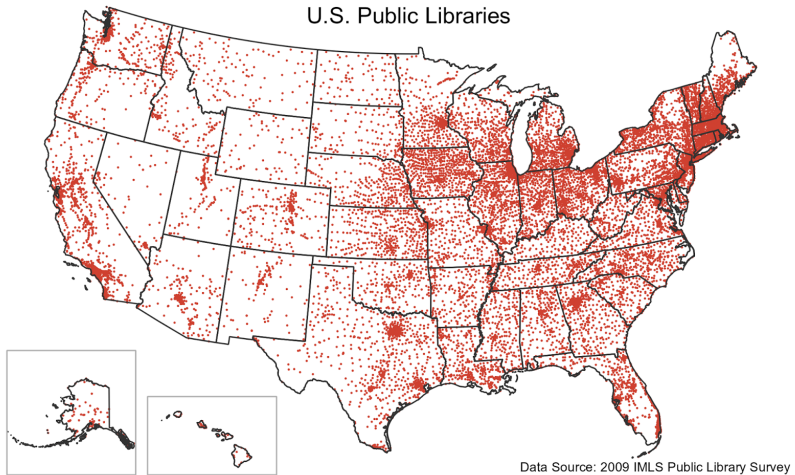


Unemployment: The number of job losses in the U.S. chronicled in this stunning image

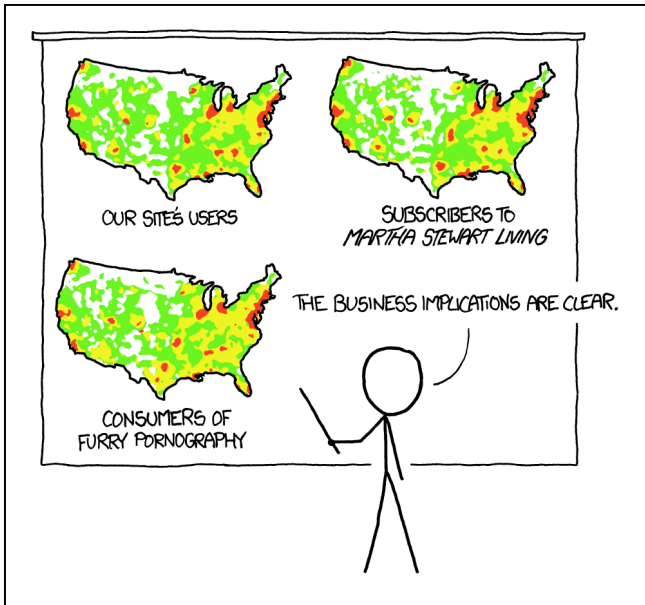
—The Daily Mail, June 19, 2012



Bad Dot Map #2



Beware of dot maps that are just population maps.



PET PEEVE #208:
GEOGRAPHIC PROFILE MAPS WHICH ARE
BASICALLY JUST POPULATION MAPS

- 1 Review
- 2 Dot Maps
- 3 Storytelling with Dot Maps
- 4 Making Dot Maps in Python
- 5 Caution about Dot Maps
- 6 Reminders**



Reminders

- Graded Exam 2 was handed back in section yesterday. If you didn't get yours, please see your TA.
- If you have any questions about the grading, please come talk to me.
- Don't forget to sign up for a final project presentation here: [link to form].
- Assignment 7 due tonight! (Sorry, I had forgotten to create a submission on Gradescope, but it's there now.)
- Optional Assignment 8 due next week. It replaces your lowest assignment grade.
- A look to next week...
- If you are interested in having lunch with our guest speakers after class next Wednesday, please fill out [this form] by Sunday.

