IMD0905 - Data Science I
Lesson #18 - Visualization [End]

Ivanovitch Silva
November, 2018

Update the repository

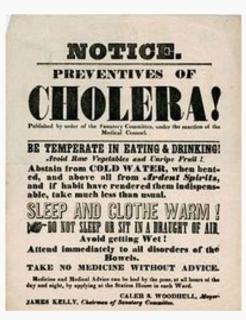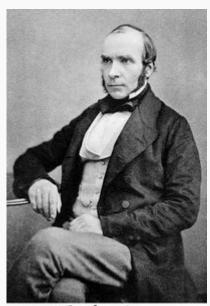git clone https://github.com/ivanovitchm/IMD0905_datascience_one.git

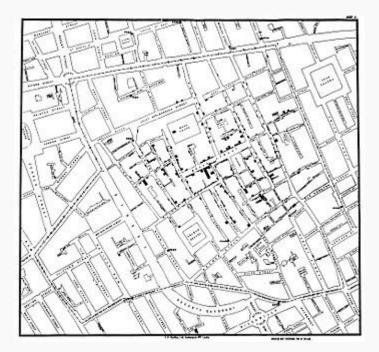Or ....

git pull

Agenda

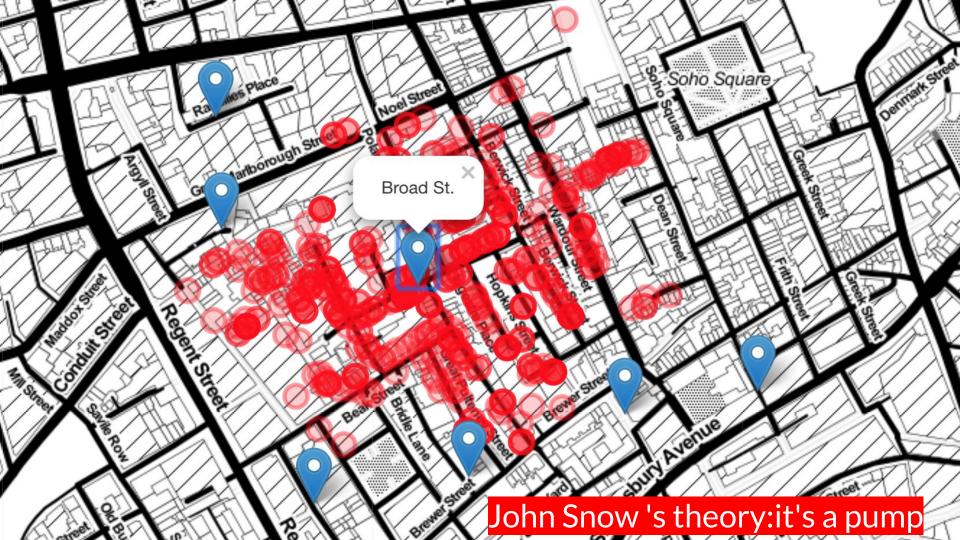- Case Study #1 - John Snow Map

Dr. John Snow

Cholera Attacks

	Death	X coordinate	Y coordinate
0	1	51.513418	-0.137930
1	1	51.513418	-0.137930
2	1	51.513418	-0.137930
3	1	51.513361	-0.137883
4	1	51.513361	-0.137883

	death_count	x_latitude	y_longitude
count	489.0	489.000000	489.000000
mean	1.0	51.513398	-0.136403
std	0.0	0.000705	0.001503
min	1.0	51.511856	-0.140074
25%	1.0	51.512964	-0.137562
50%	1.0	51.513359	-0.136226
75%	1.0	51.513875	-0.135344
max	1.0	51.515834	-0.132933

The Ghost Map

Broad St.

John Snow 's theory:it's a pump

He was looking for anomalies now (we would now say "outliers in data") and found two in fact where there were no deaths.

MISTAKING CAUSE FOR EFFECT.

Boy. "I SAY, TOMMY, I'M BLOW'D IF THERE ISN'T A MAN A TURNING ON THE CHOLERA."

Number of Cholera Deaths/Attacks
before and after 8th of September 1854

References

1. https://www.kaggle.com/daveianhickey/how-to-folium-for-maps-heatmaps-time-analysis
2. http://nbviewer.jupyter.org/github/python-visualization/folium/blob/master/examples/Plugins.ipynb
3. https://github.com/helmutd/datacamp/blob/master/Recreating%20John%20Snow's%20Ghost%20Map.ipynb
4. https://www.coursera.org/lecture/python-for-data-visualization/introduction-to-folium-CpjW0
5. https://pythonhow.com/web-mapping-with-python-and-folium/