# Data Science & ML Course

## Lesson #11 Exploratory Data Analysis VI

Ivanovitch Silva

October, 2018

# Agenda

- Case Study #1 - John Snow Map
- Case Study #2 - Open Data Natal

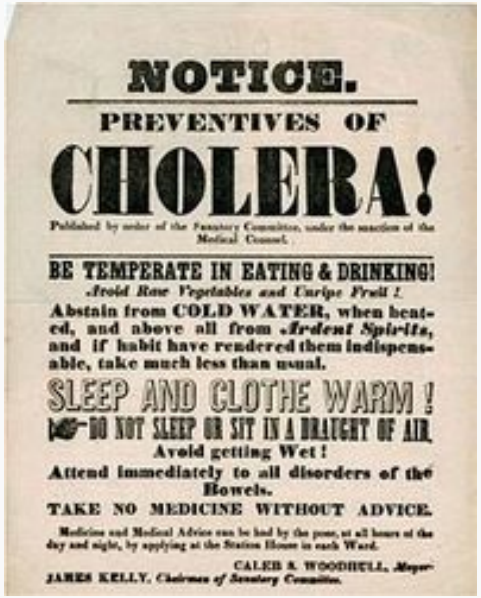
# Update from repository

```
git clone https://github.com/ivanovitchm/datascience2machinelearning.git
```

Or ....

```
git pull
```

# Dr. John Snow



NOTICE.

PREVENTIVES OF

CHOLERA!

Published by order of the Sanatory Committee, under the sanction of the Medical Council.

BE TEMPERATE IN EATING & DRINKING!

*Avoid Raw Vegetables and Unripe Fruit!*

Abstain from COLD WATER, when heated, and above all from *Ardent Spirits*, and if habit have rendered them indispensable, take much less than usual.

SLEEP AND CLOTHE WARM!

DO NOT SLEEP OR SIT IN A DRAUGHT OF AIR.

Avoid getting Wet!

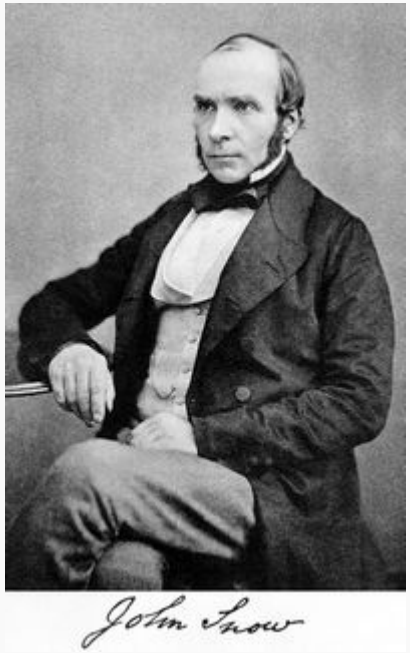
Attend immediately to all disorders of the Bowels.

TAKE NO MEDICINE WITHOUT ADVICE.

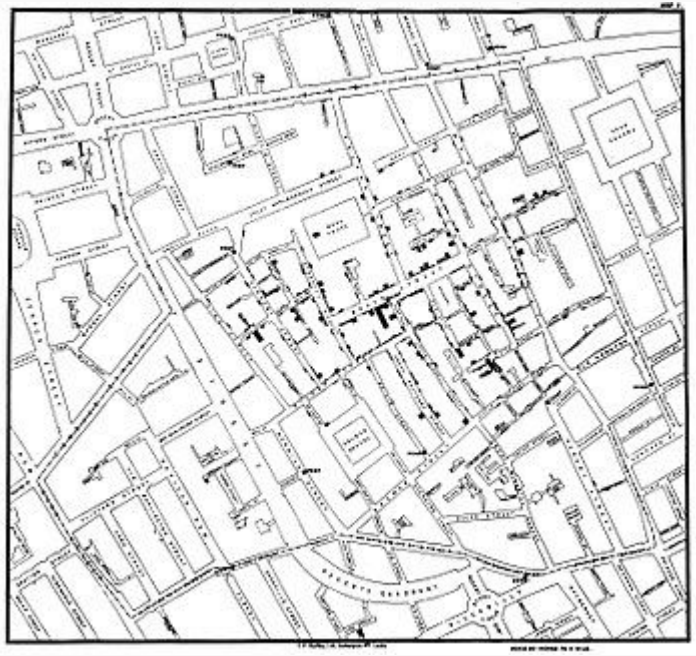
Medicine and Medical Advice can be had by the poor, at all hours of the day and night, by applying at the Station House in each Ward.

CALEB S. WOODHULL, *Mayor.*

JAMES KELLY, *Chairman of Sanatory Committee.*



John Snow

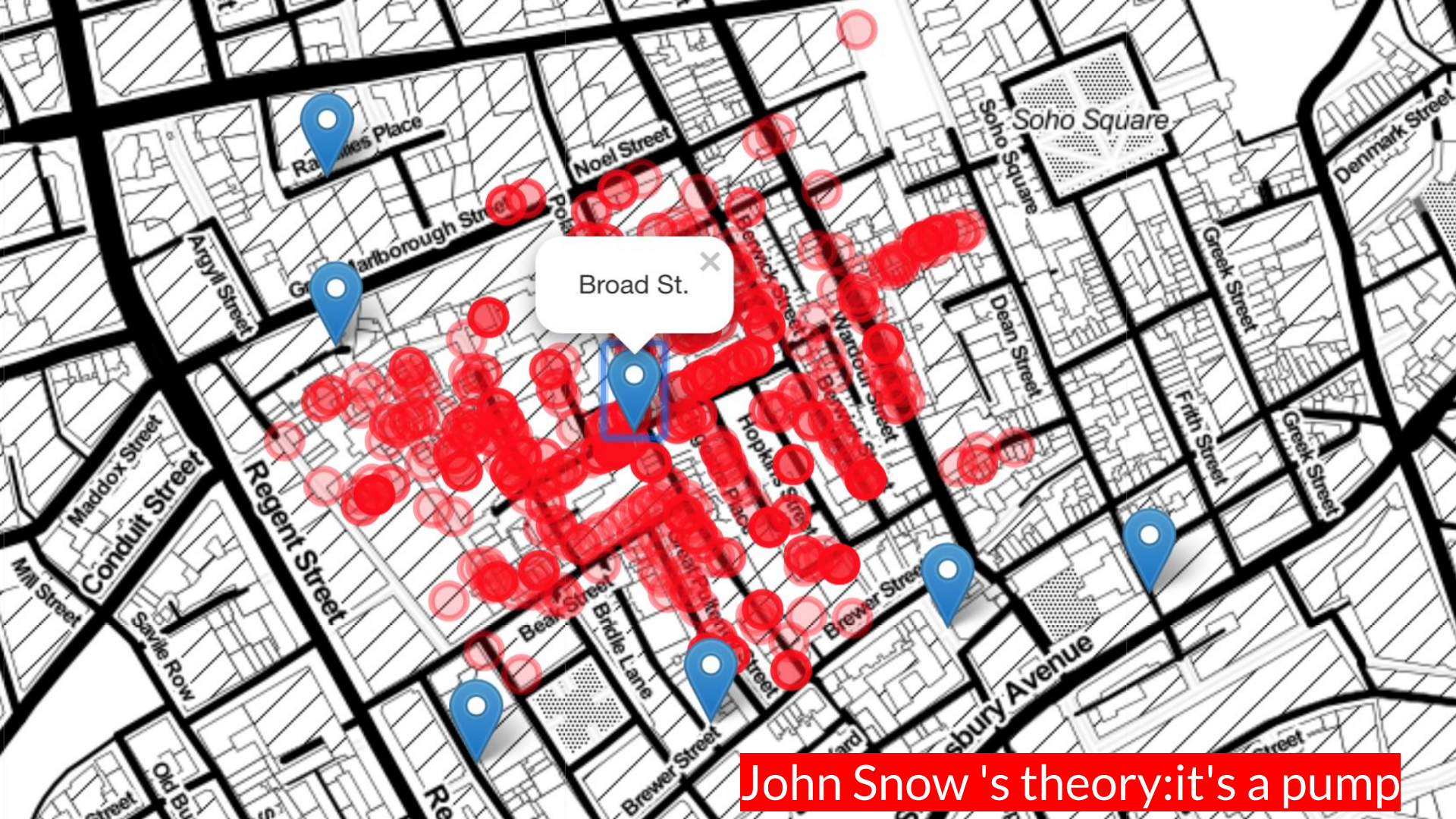


# Cholera Attacks

| | 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 |

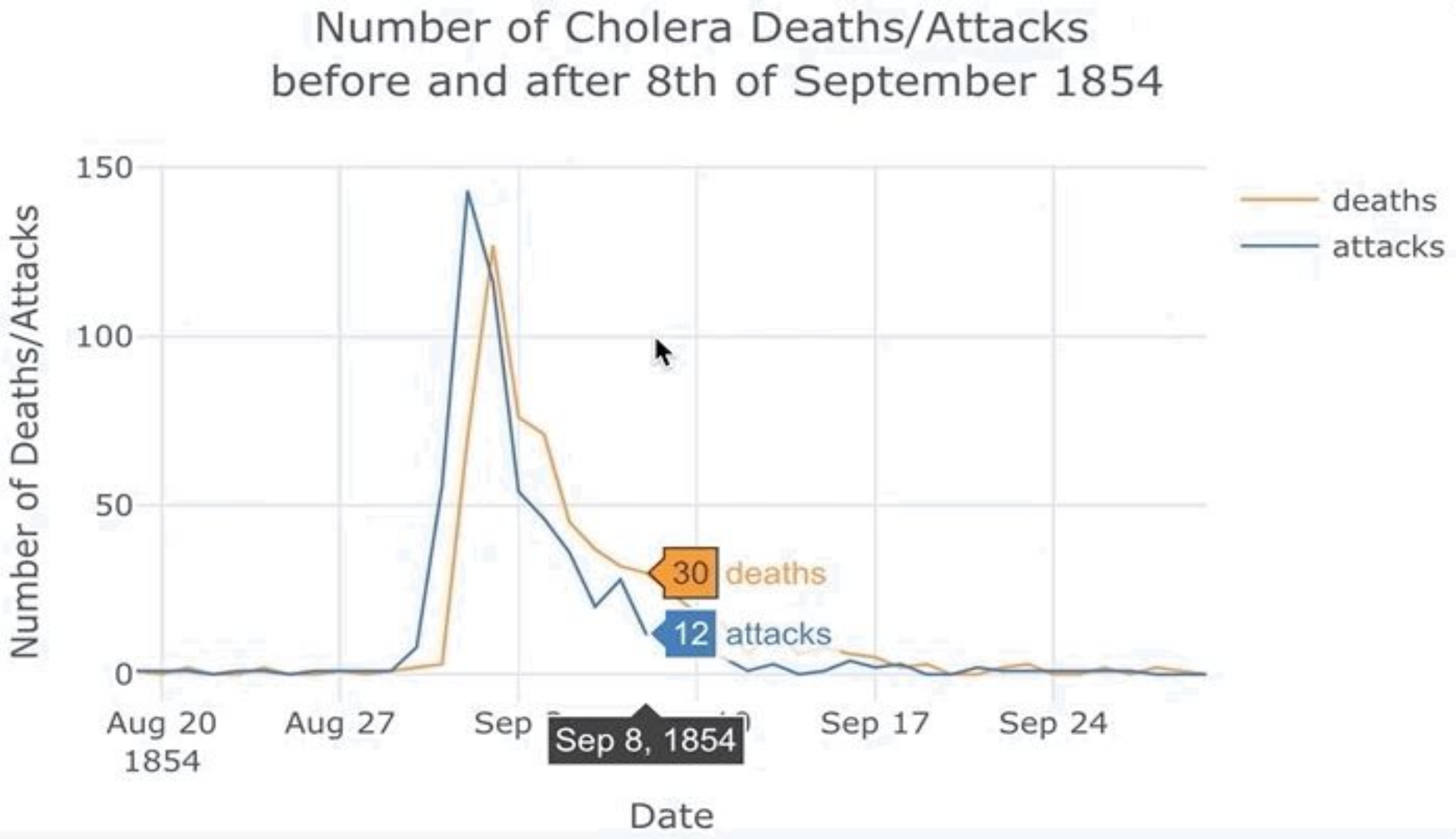
# The Ghost Map



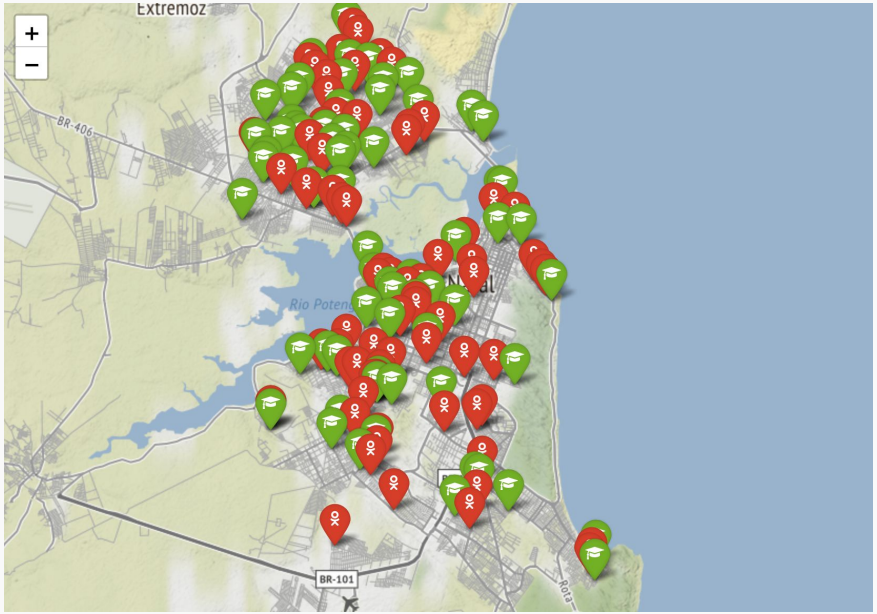

John Snow 's theory:it's a pump

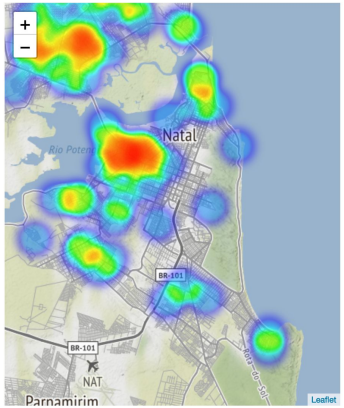
He was looking for anomalies now (we would now say "outliers in data") and found two in fact where there were no deaths.

*Boy.* "I say, Tommy, I'm blow'd if there isn't a Man a turning on the Cholera."

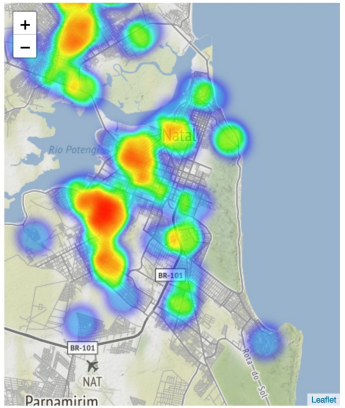

Number of Cholera Deaths/Attacks
before and after 8th of September 1854
150
100
50
0
Number of Deaths/Attacks
deaths
attacks
30 deaths
12 attacks
Aug 20
1854
Aug 27
Sep 17
Sep 24
Sep 8, 1854
Date

# Case study: open data Natal




Elementary School




Kindergarten

# 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/