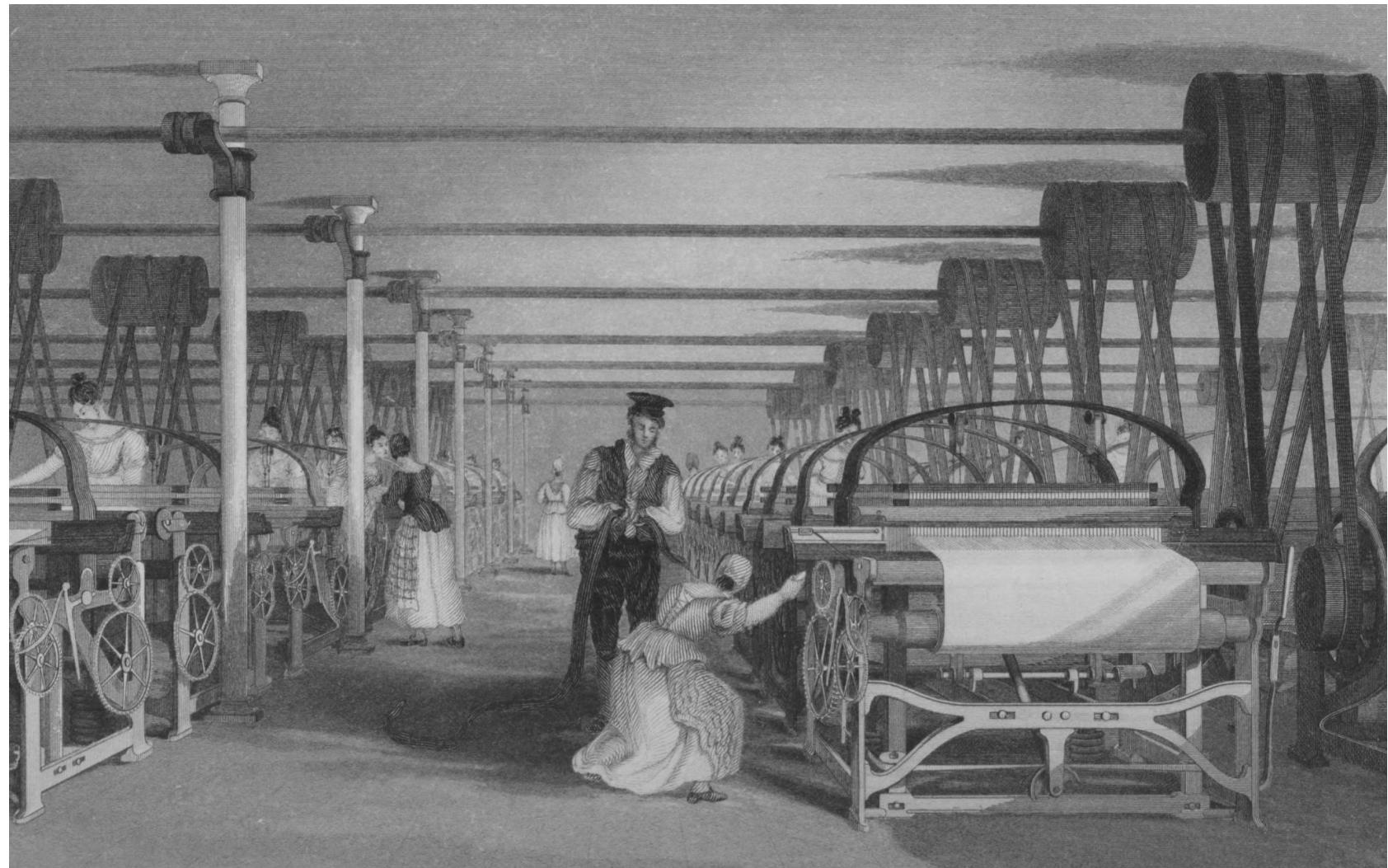
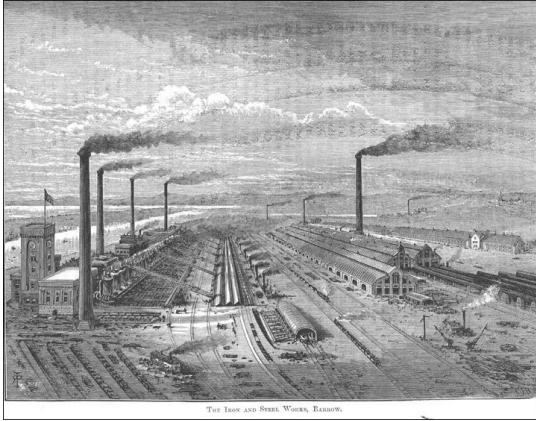


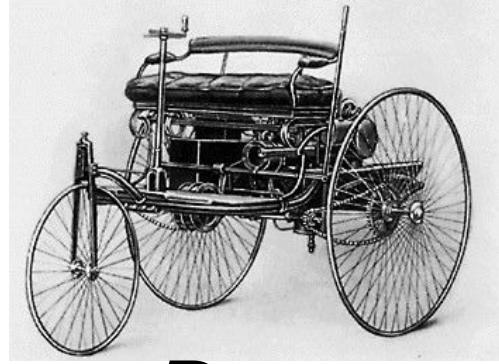
Why visualization?

1854, London





Industrial revolution
(1760~1840)



Benz
Patent-Motorwagen
(1885)

1854, London



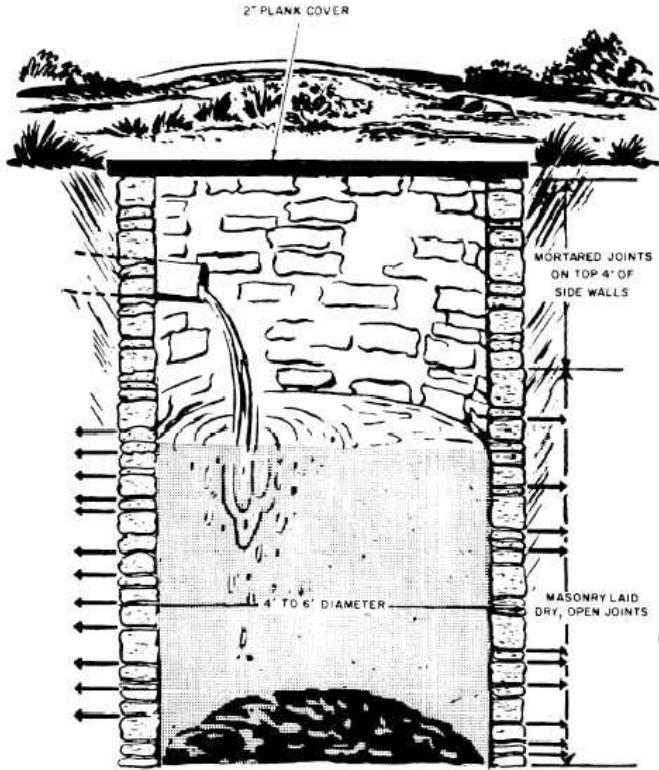
Urbanization



*but no
infrastructure*

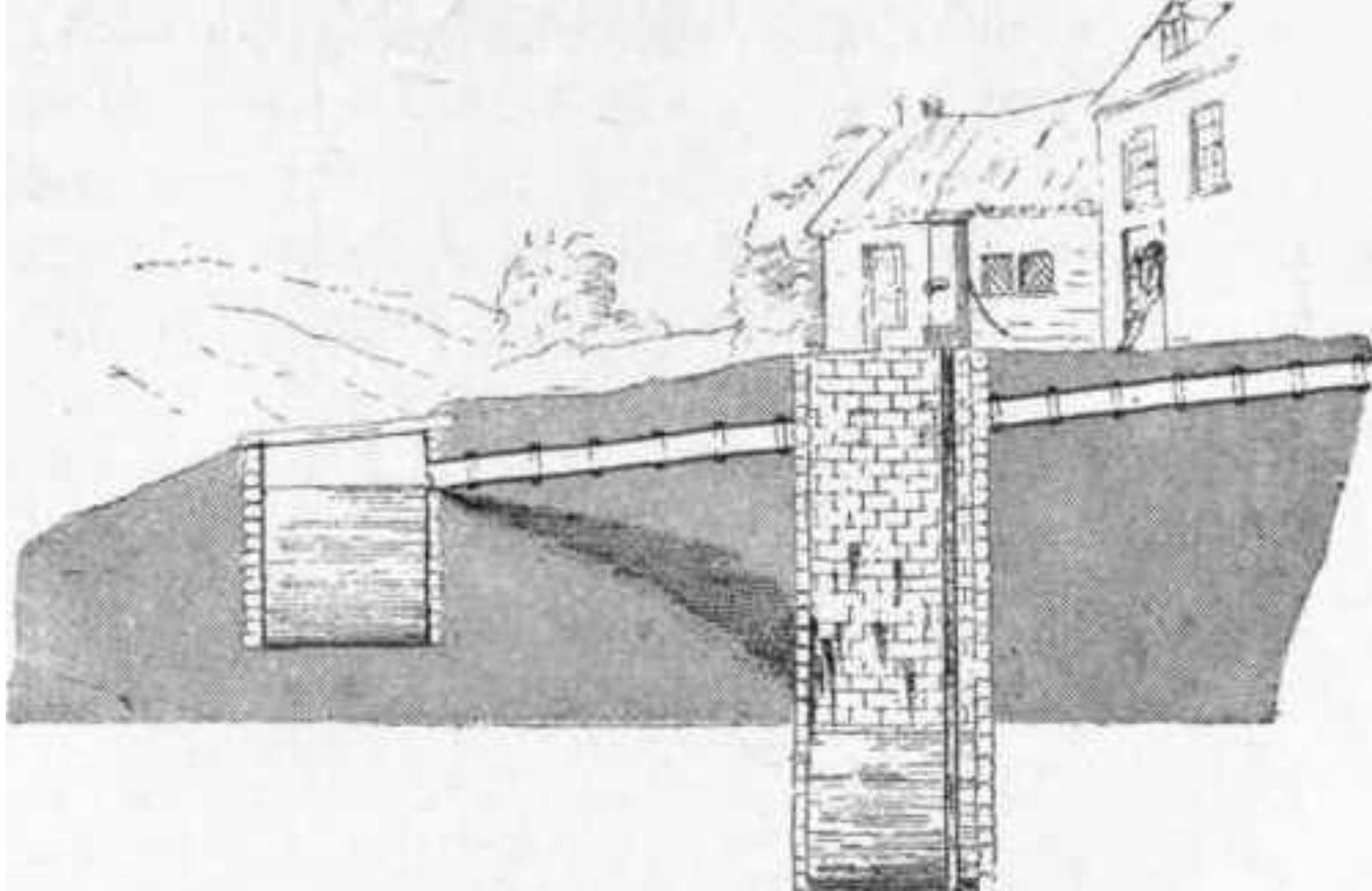
“Nightmen”





“Night soil”

Cesspool



1854 London Cholera outbreak

*127 died in three
days,
616 died at the end.*

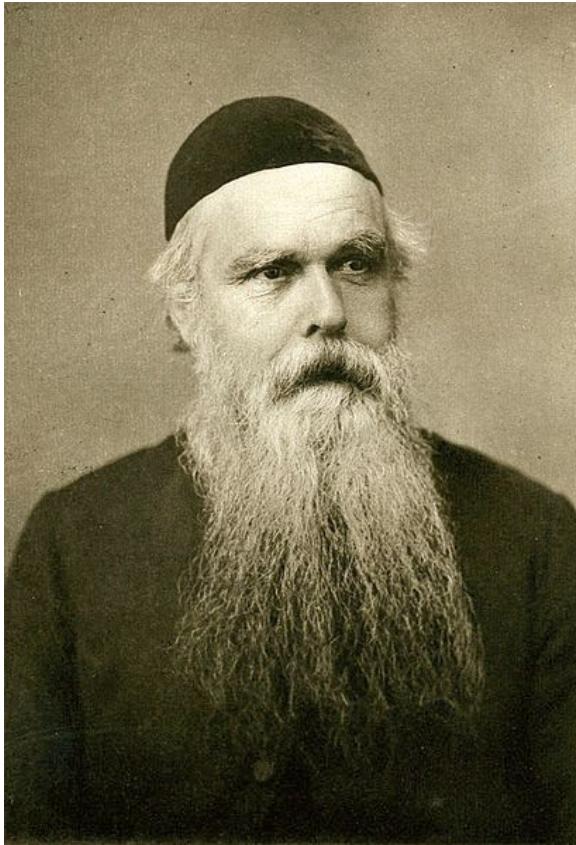


“Miasma (bad air) theory”

VS.

“Germ theory”





Miasm a!

Rev. Henry Whitehead

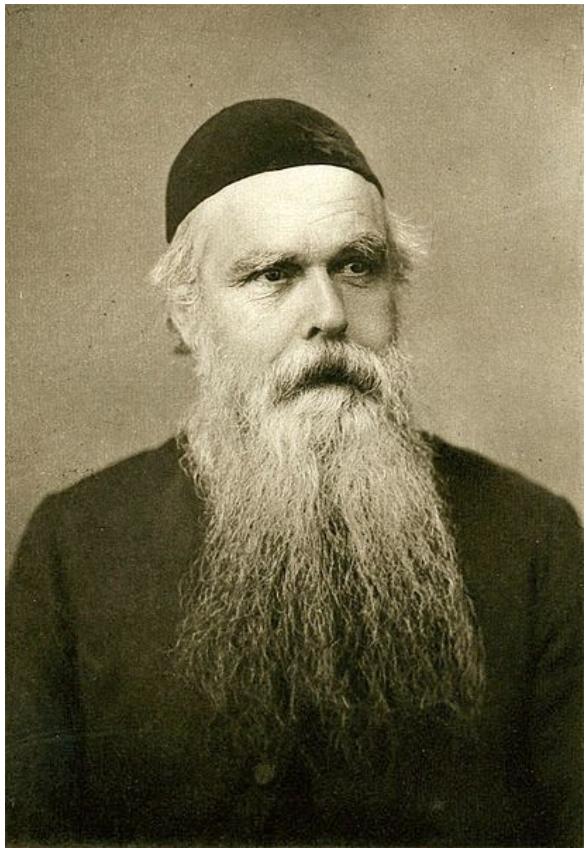


John Snow

Germs

!

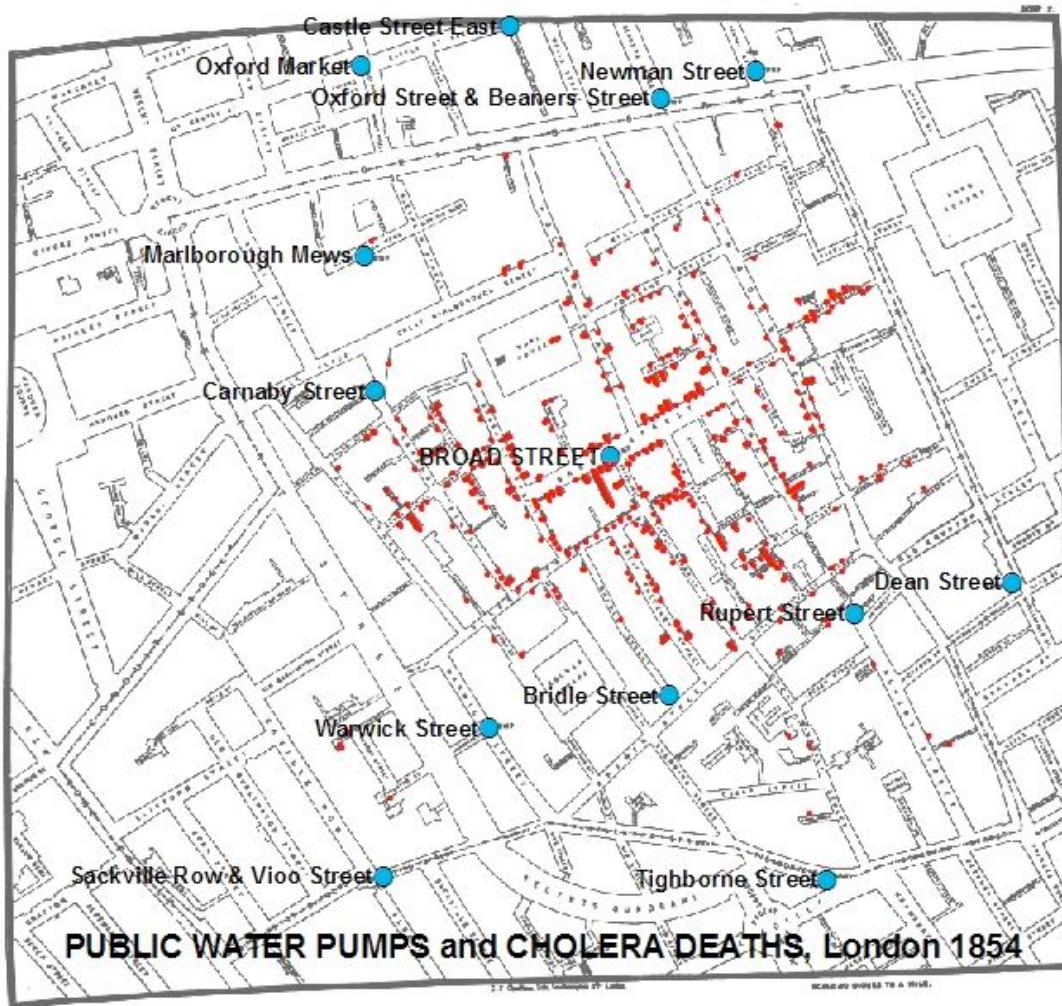




What he (they) did
changed the course of **human race.**

What did they do?



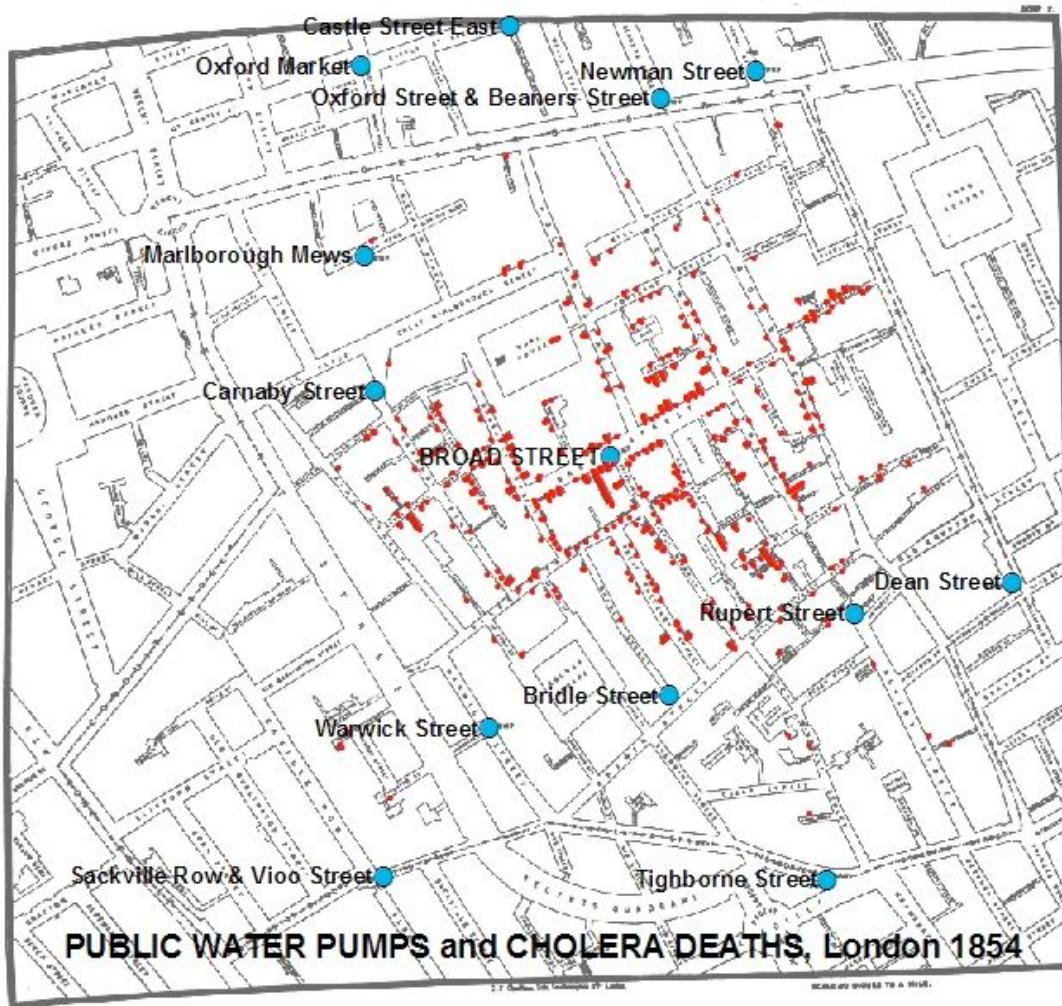


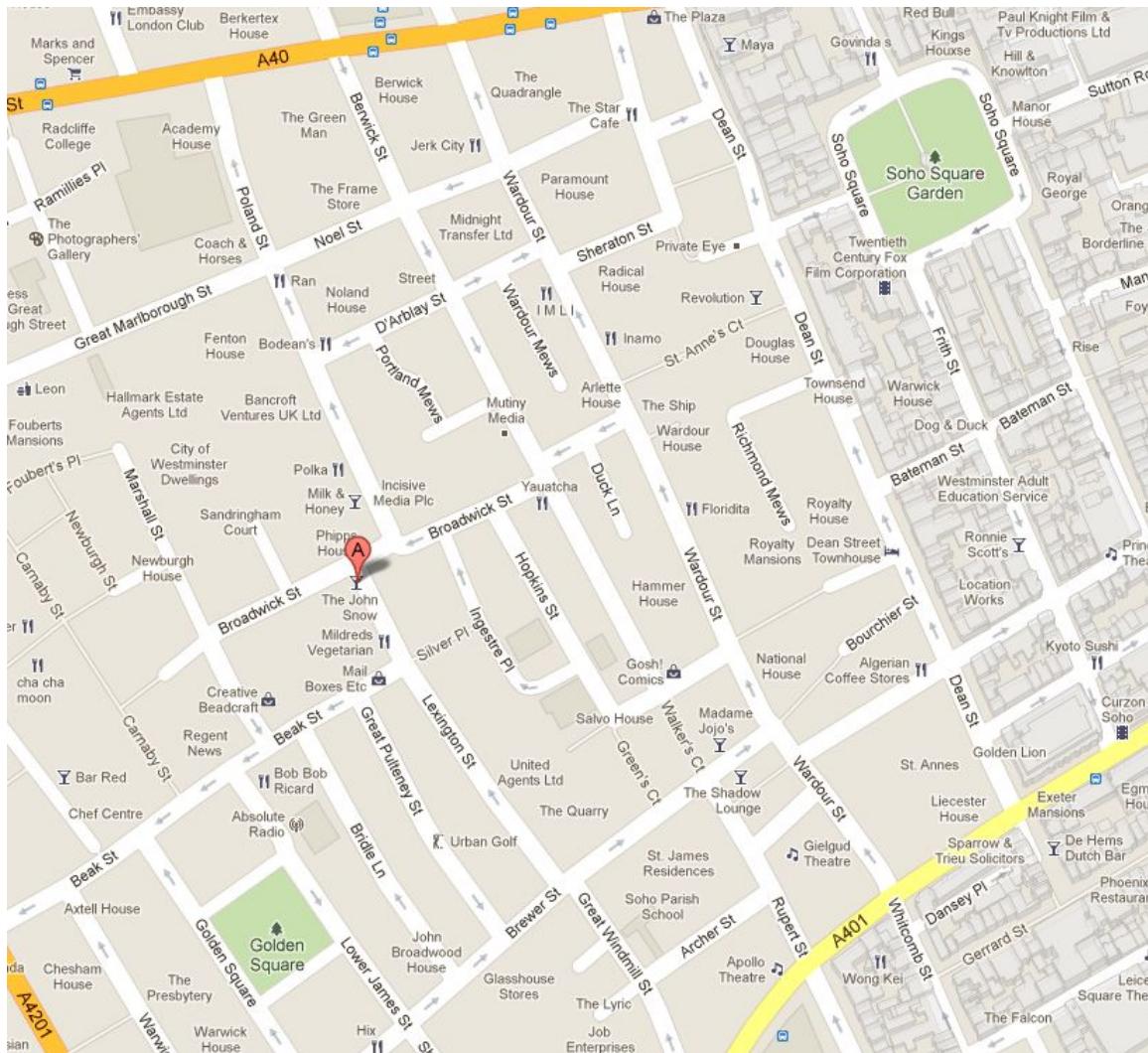
On proceeding to the spot, **I found that nearly all the deaths had taken place within a short distance of the [Broad Street] pump.**

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...*

...

— John Snow, letter to the editor of the *Medical Times and Gazette*







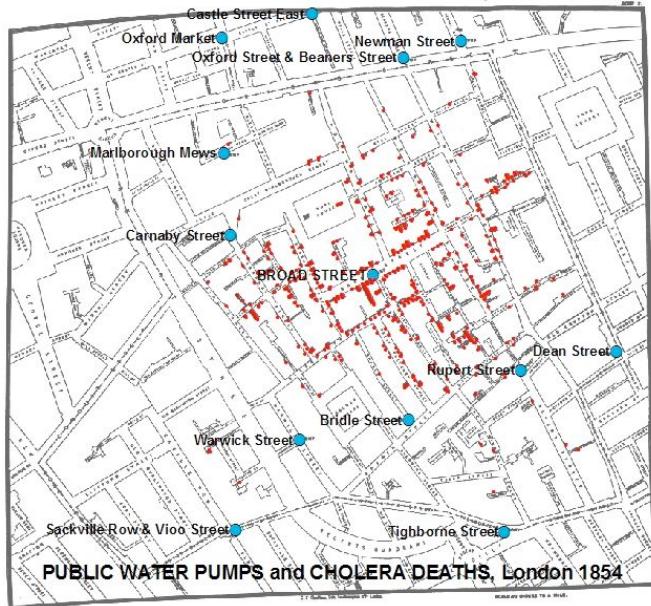


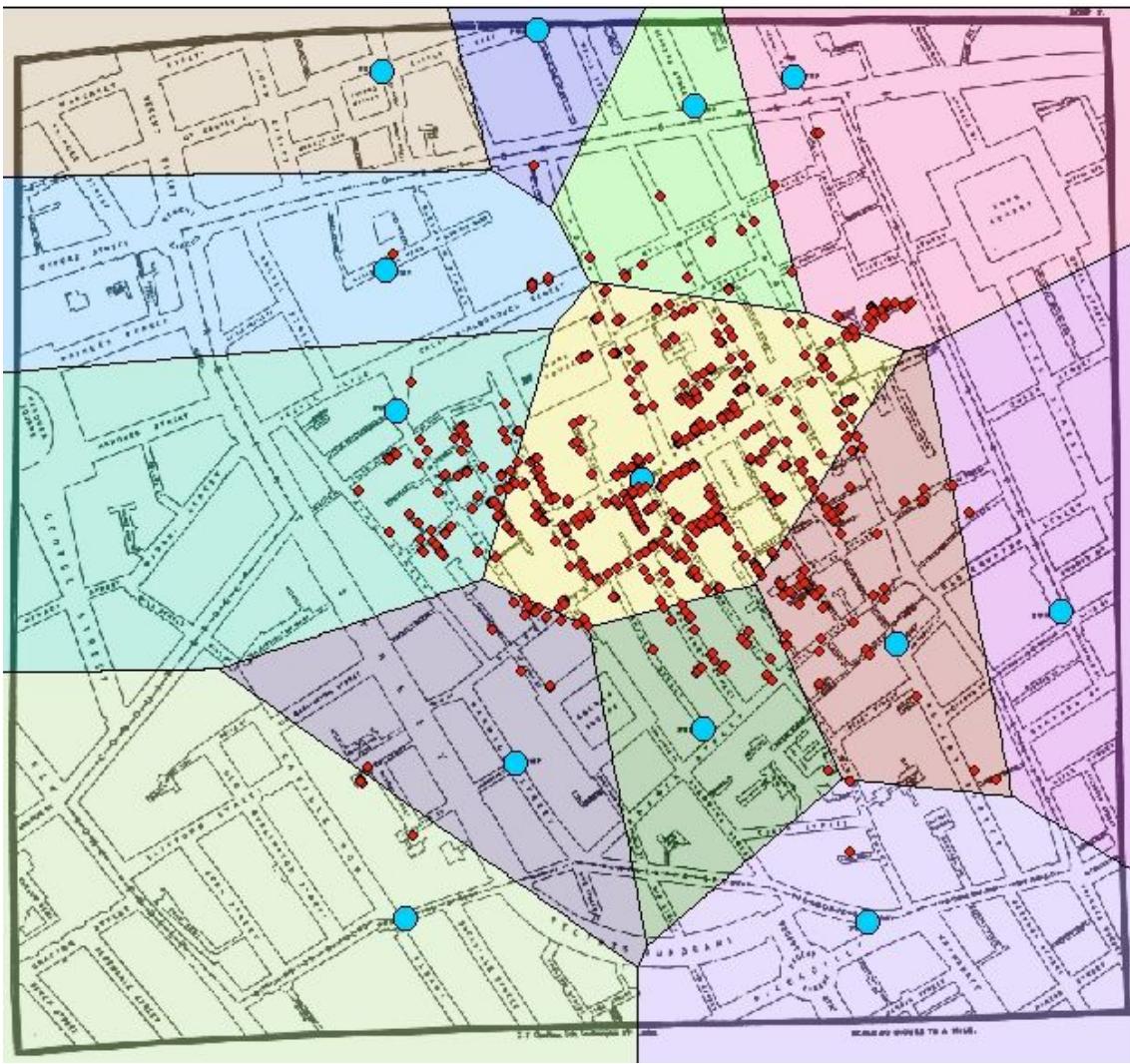
Now imagine...

Can you make
Dr. Snow's job easier?

**“The nearest pump
for any given
location”?**

How do you draw a map to reveal





“Voronoi diagram”

<https://blocks.roadtolarissa.com/mbostock/4060366>

<https://observablehq.com/@d3-hover-voronoi>

Can't we just use **numbers** and
statistics?

I		II		III		IV	
X	Y	X	Y	X	Y	X	Y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

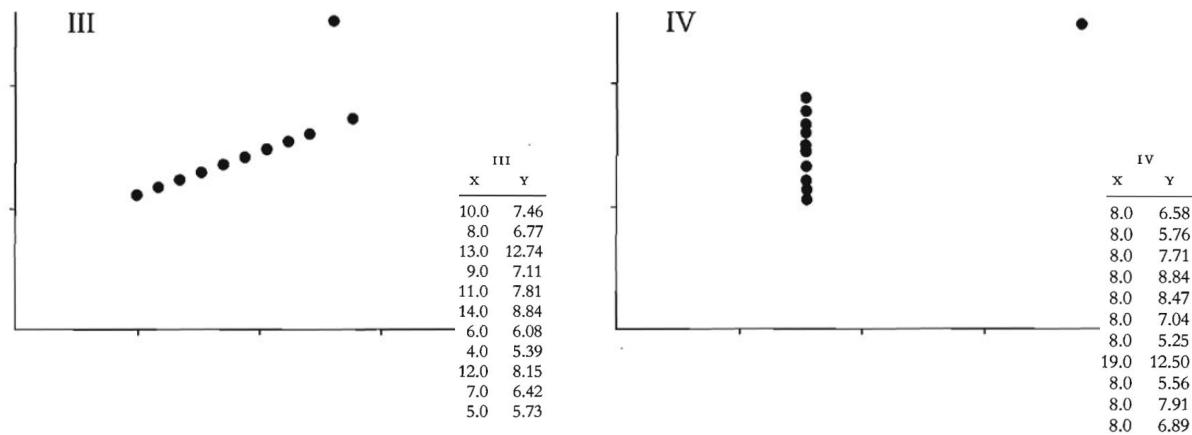
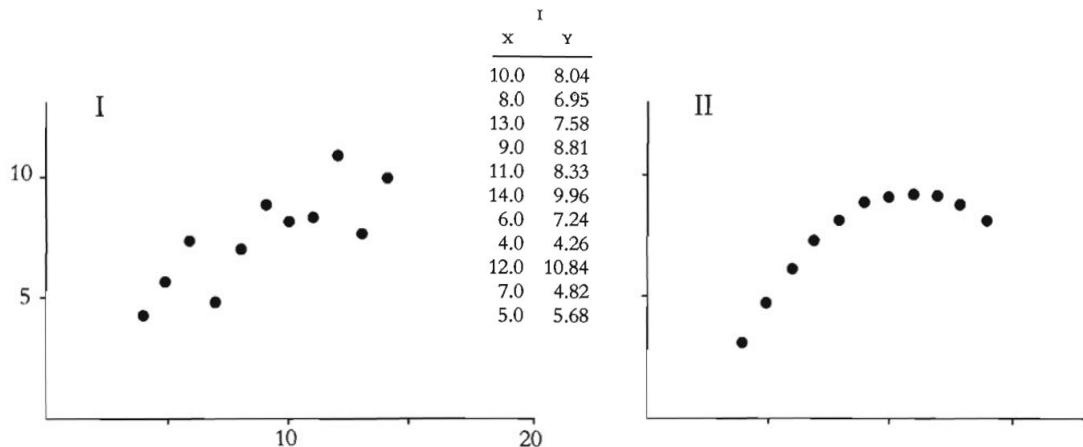
Download the attached file in “Module 1: [In-class Exercise] Why can't we simply use summary statistics?”

I		II		III		IV	
X	Y	X	Y	X	Y	X	Y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

“Anscombe's quartet”

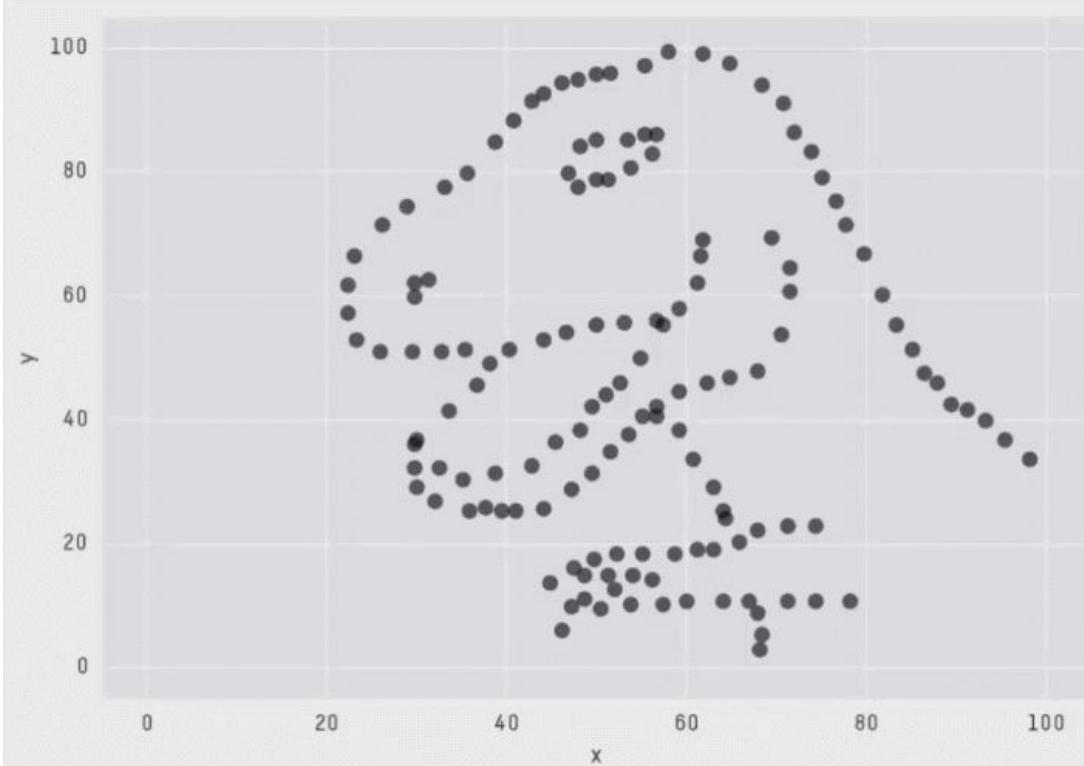


Property	Value	Accuracy	
Mean of x	9	exact	
Sample variance of x : s_x^2	11	exact	
Mean of y	7.50	to 2 decimal places	Francis Anscombe
Sample variance of y : s_y^2	4.125	± 0.003	
Correlation between x and y	0.816	to 3 decimal places	
Linear regression line	$y = 3.00 + 0.500x$	to 2 and 3 decimal places, respectively	
Coefficient of determination of the linear regression : R^2	0.67	to 2 decimal places	



“Same Stats, Different Graphs”

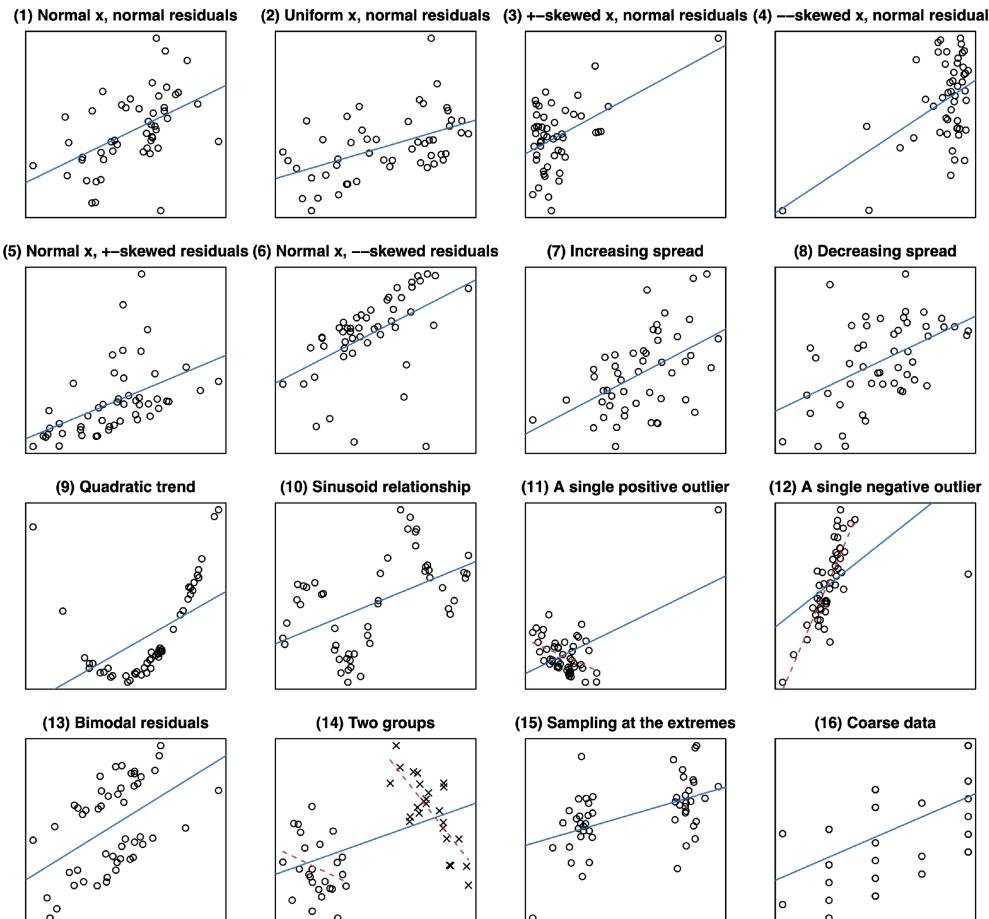
<https://www.research.autodesk.com/publications/same-stats-different-graphs>

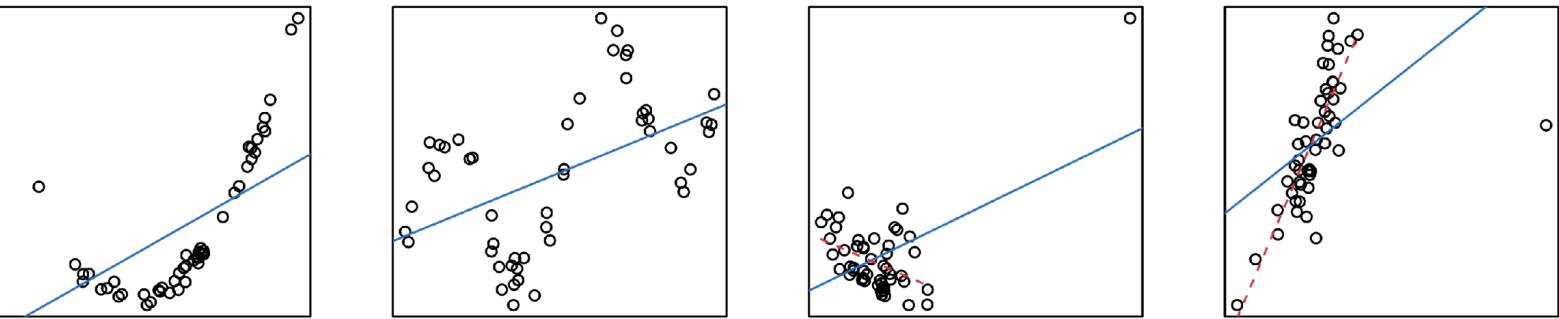
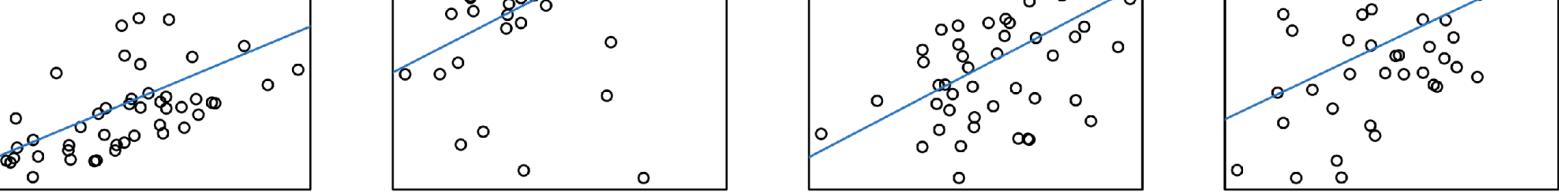


X Mean: 54.2659224
Y Mean: 47.8313999
X SD : 16.7649829
Y SD : 26.9342120
Corr. : -0.0642526

Correlation of 0.5?

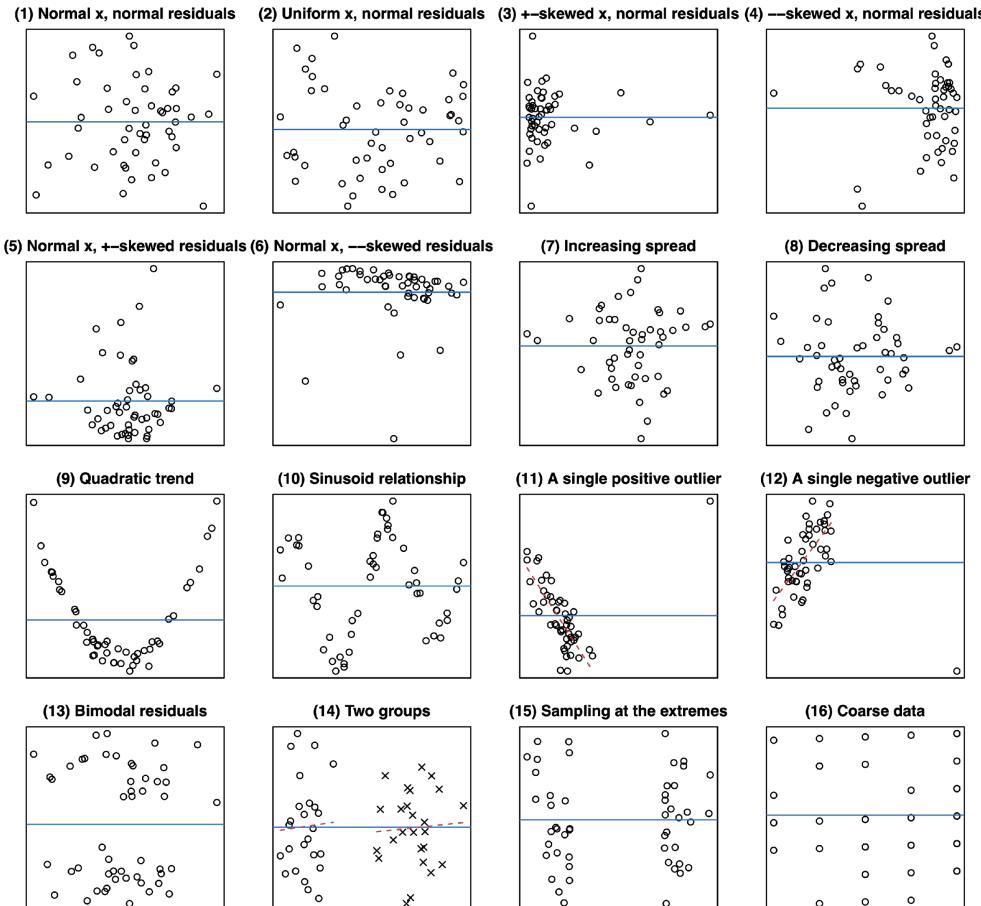
All correlations: $r(50) = 0.5$

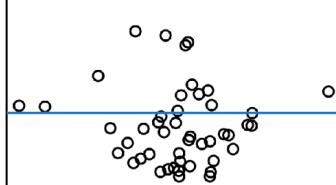




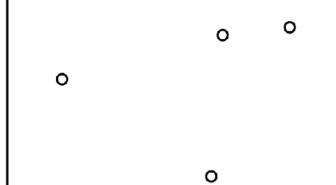
Zero correlation?

All correlations: $r(50) = 0$

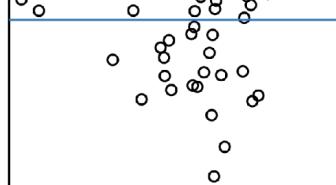




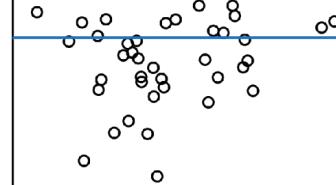
(9) Quadratic trend



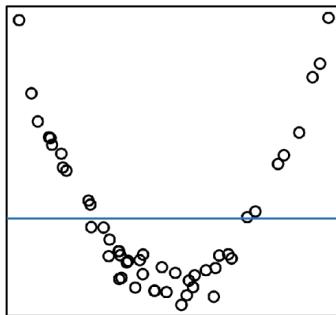
(10) Sinusoid relationship



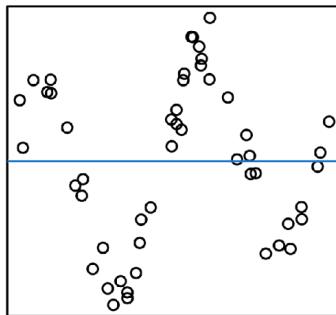
(11) A single positive outlier



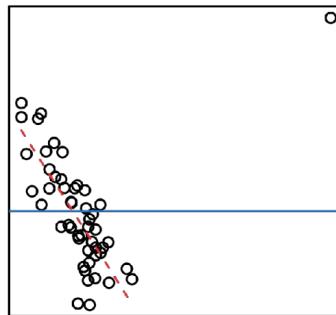
(12) A single negative outlier



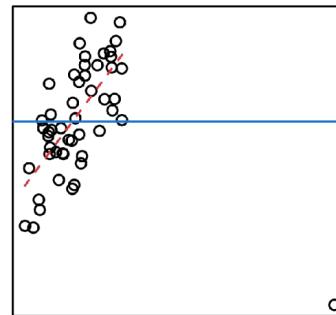
(13) Bimodal residuals



(14) Two groups



(15) Sampling at the extremes



(16) Coarse data

How many 5s?

192568719273163581623152957230
519263912701749619236102701375
069341629471037012639161

How many 5s?

192 $\textcolor{red}{5}$ 68719273163 $\textcolor{red}{5}$ 816231 $\textcolor{red}{5}$ 29 $\textcolor{red}{5}$ 723
0 $\textcolor{red}{5}$ 1926391270174961923610270137
 $\textcolor{red}{5}$ 069341629471037012639161

Our cognitive ability is limited.

**Visual aids free up
our mental capacity.**

Our visual system is a **massively parallel** pattern recognition machine.

Why now?

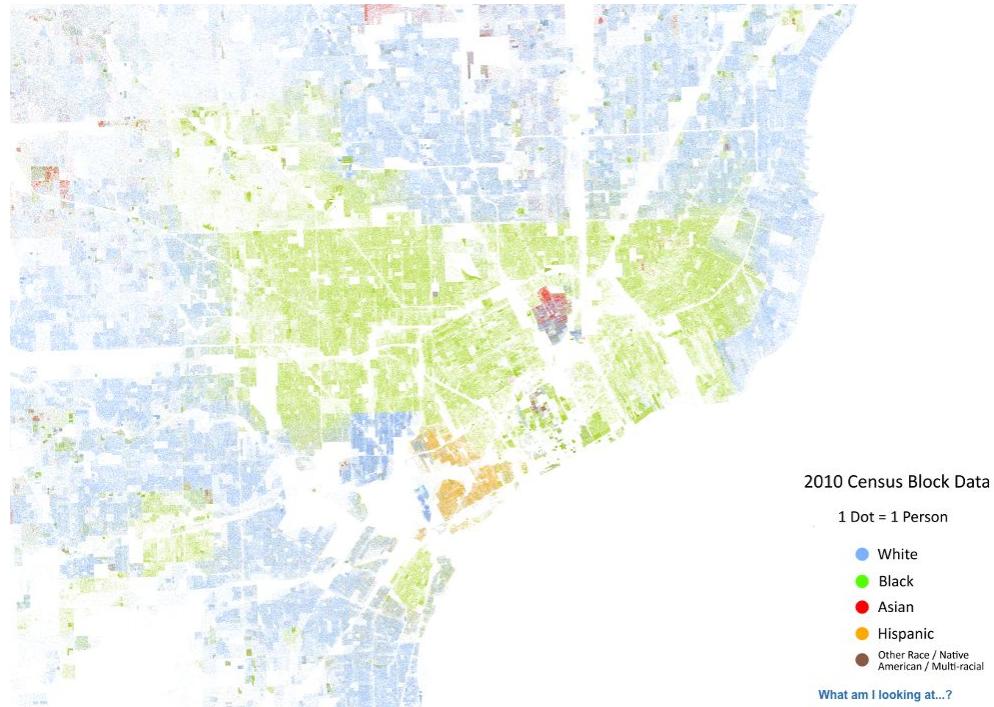
“big data”

Big data just means buying HDD

until you extract **insights.**

Census - Race and Ethnicity

<http://demographics.coopercenter.org/DotMap/index.html>



How?

Explore foundations and techniques with
Two foci:

Exploration & Explanation

E

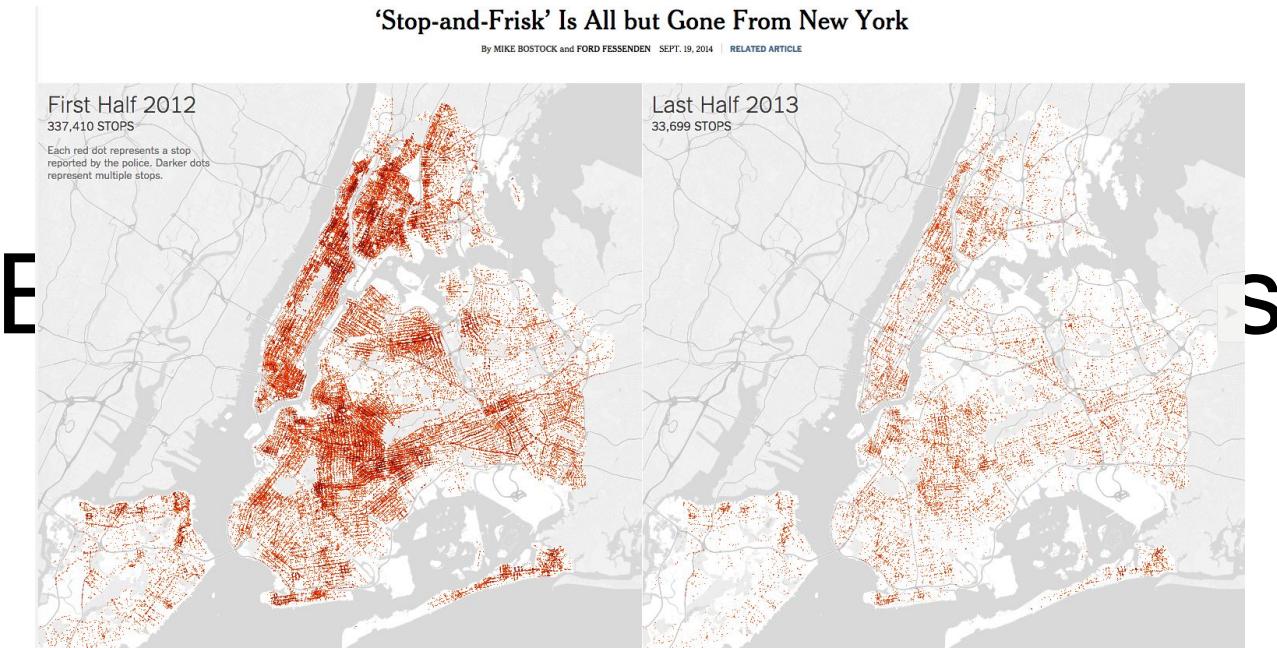
S

Set As... Filter... Merge...

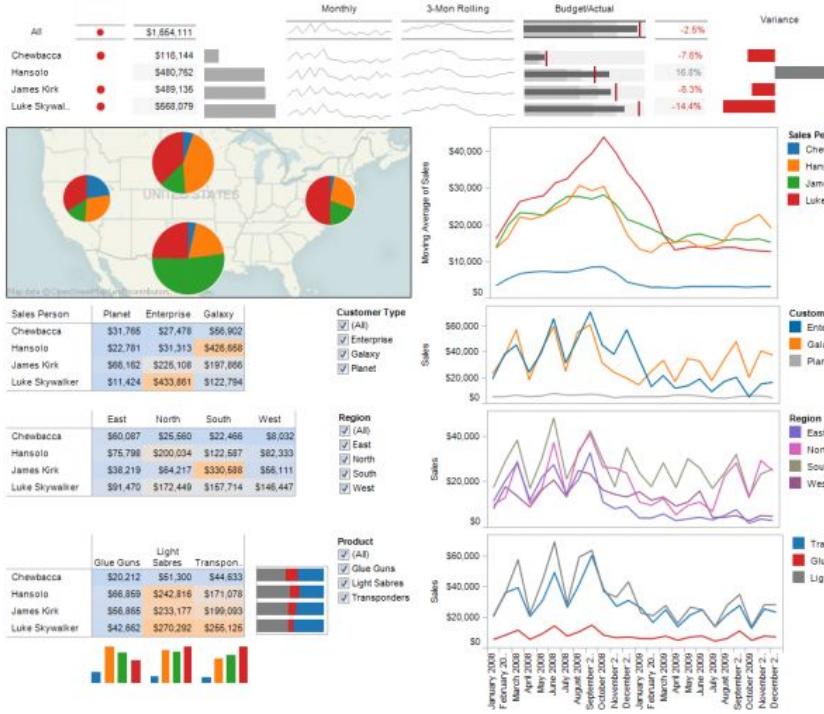
SampleID	Common_Name	Description	KeyHand	DigitHand	Hand	Individual
M9Akey217.1410...	keyboard	Akey	Left	NA	Left	M9
M9Bkey217.141032	keyboard	Bkey	Ambiguous	NA	Ambiguous	M9
M9Ckey217.141033	keyboard	Ckey	Left	NA	Left	M9
M9Dkey217.141034	keyboard	Dkey	Left	NA	Left	M9
M9Ekey217.141044	keyboard	Ekey	Left	NA	Left	M9
M9Enter217.1410...	keyboard	Ente	Right	NA	Right	M9
M9Fkey217.141065	keyboard	Fkey	Left	NA	Left	M9
M9Gkey217.141067	keyboard	Gkey	Left	NA	Left	M9
M9Hkey217.14109...	keyboard	Hkey	Right	NA	Right	M9
M9Ikey217.141066	human_skin	finger_tip	NA	Left	Left	M9
M9Jmid217.141043	human_skin	finger_tip	NA	Right	Right	M9
M9Kkey217.141040	keyboard	Kkey	Right	NA	Right	M9
M9Mid217.141043	human_skin	finger_tip	NA	Left	Left	M9
M9Mkey217.141060	human_skin	finger_tip	NA	Right	Right	M9
M9Nkey217.1410...	keyboard	Nkey	Right	NA	Right	M9
M9Okey217.1410...	keyboard	Okey	Right	NA	Right	M9
M9Pkey217.141035	human_skin	finger_tip	NA	Left	Left	M9
M9Rkey217.141036	human_skin	finger_tip	NA	Right	Right	M9
M9Skey217.141096	keyboard	Pkey	Right	NA	Right	M9
M9Tkey217.141010...	keyboard	Qkey	Left	NA	Left	M9
M9Rlnk217.141020	human_skin	finger_tip	NA	Left	Left	M9
M9Rlnk217.141080	human_skin	finger_tip	NA	Right	Right	M9
M9Skey217.141004	keyboard	Skey	Left	NA	Left	M9
M9Space217.1410...	keyboard	Space_bar	Ambiguous	NA	Ambiguous	M9
M9TSpace217.1410...	human_skin	finger_tip	NA	Left	Left	M9
M9Thrm217.1410...	human_skin	finger_tip	NA	Right	Right	M9
M9Vkey217.1410...	keyboard	Vkey	Left	NA	Left	M9
M9Wkey217.1410...	keyboard	Wkey	Left	NA	Left	M9
M9Xkey217.1410...	keyboard	Xkey	Left	NA	Left	M9
M9Ykey217.141029	keyboard	Ykey	Right	NA	Right	M9

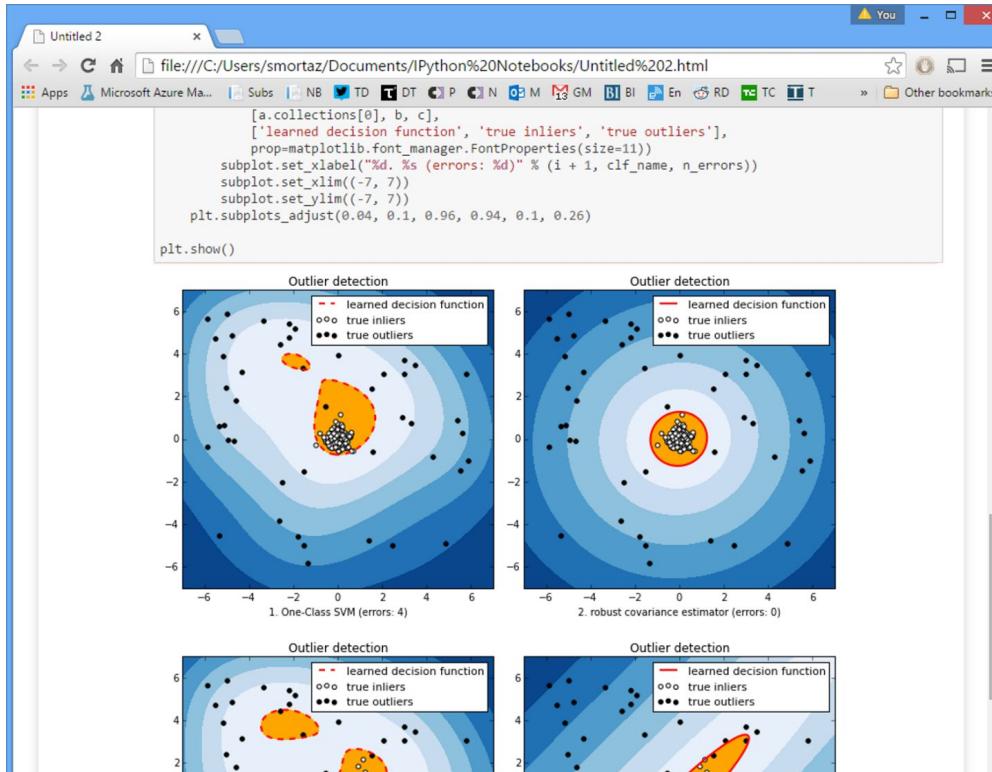


discover insights



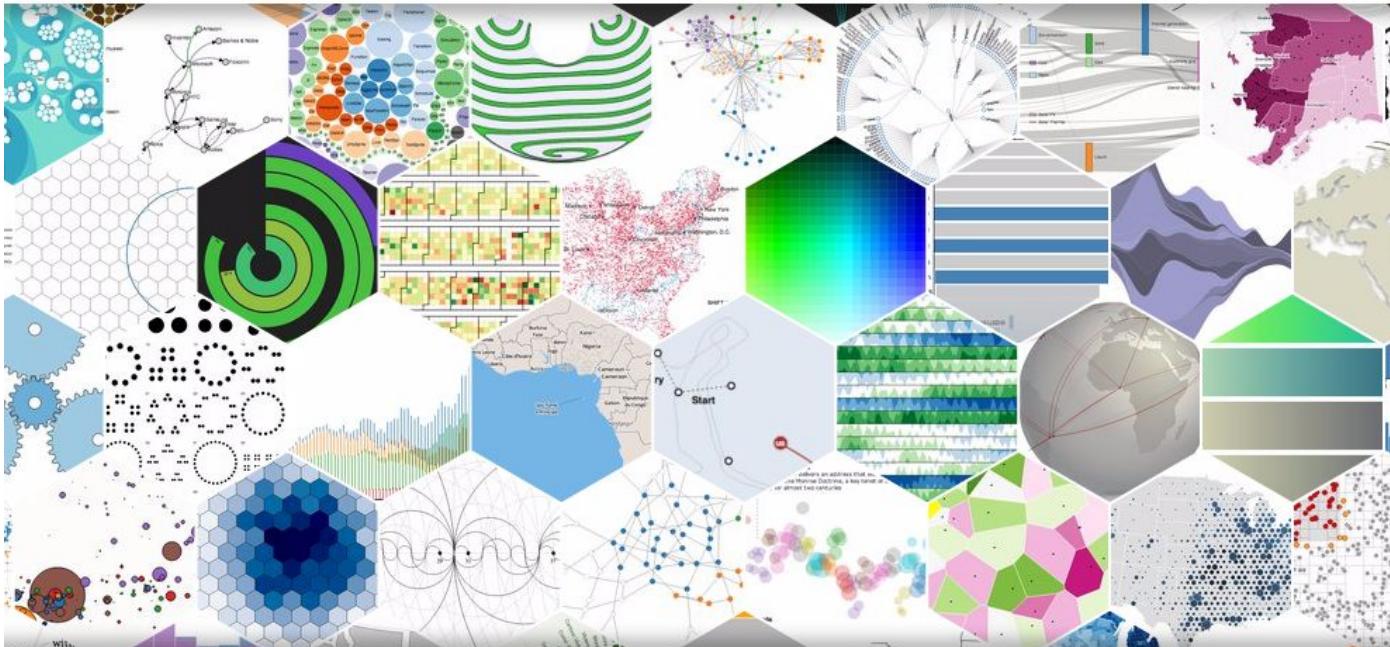
communicate the insights







Data-Driven Documents



D3.js is a JavaScript library for manipulating documents based on data. **D3** helps you bring data to life using HTML, SVG, and CSS. D3's emphasis on web standards gives you the full capabilities of modern browsers without tying yourself to a proprietary framework, combining powerful visualization components and a data-driven approach to DOM manipulation.

[See more examples.](#)

Homepage:

yongyeol.com/courses/2014F-d

Mailing list:

dviz-l@indiana.edu

For the next class

J. Heer et al. A Tour through the Visualization Zoo.

<https://queue.acm.org/detail.cfm?id=1805128>

J. VanderPlas, The Python Visualization Landscape.

<https://youtu.be/FytuB8nFHPQ>

Further readings:

<https://github.com/yy/dviz-course/blob/master/m01-intro/class.md>

Questions?