

# Data in the newsroom

**Noah Veltman**

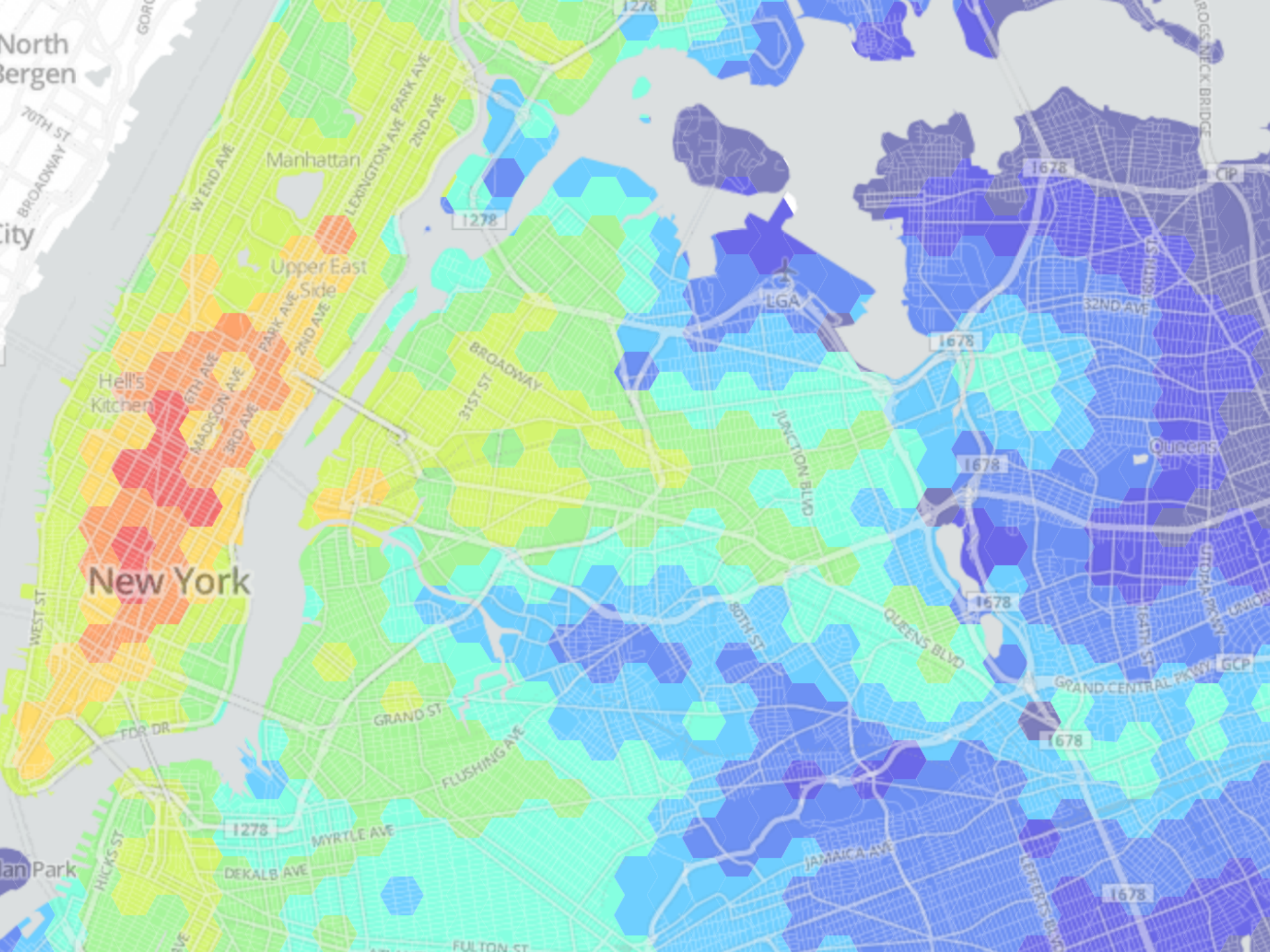
WNYC Data News Team

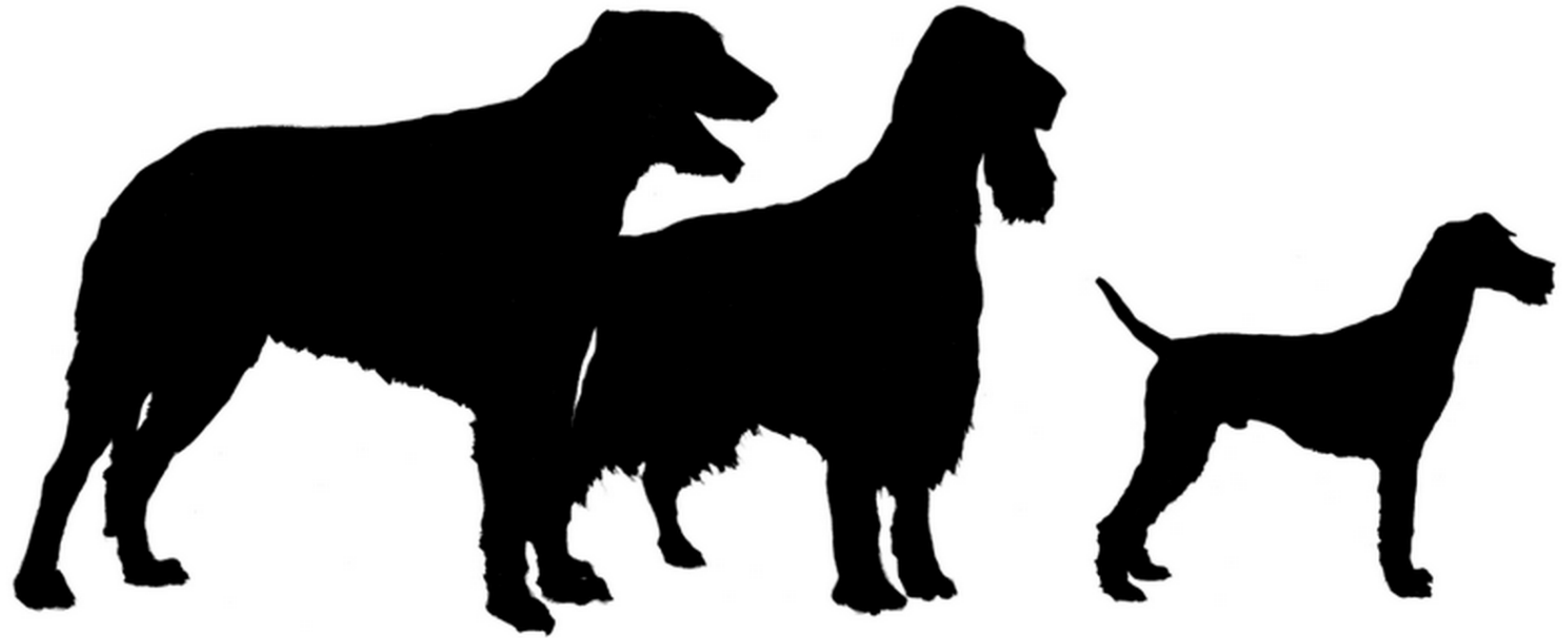
@veltman

noahveltman.com



**What we do**





irish wolfhound, irish setter, and irish greyhound

# Dogs of NYC

[Map](#)[Top 10 Dog Names](#)[Name Game](#)[Make a T-Shirt](#)[Tweet](#) 1,180

## WNYC Dogs of NYC | Map

Explore New York City's dog names and breeds by area, down to the zip code. Label size roughly corresponds with number of dogs with name in that area.

[Tweet](#) 

most common breeds

Lucy

Lucy

Rocky

Daisy



Victim:

Pedestrians ▼

Borough:

Brooklyn ▼

13  
deaths

20% of all deaths



**Marlene Baharlias, age 77**

March 18 Sheepshead Bay, Brooklyn

*Killed by a car parallel parking.*



**Roshard Charles, age 5**

March 16 Crown Heights, Brooklyn

*Killed by a double-parked minivan backing up.*

Street view



8:33 PM

289 Empire Boule

Roshard Charles v  
Charles was cross  
but later returned

Other Coverage:

[DNA Info](#)

[New York Post](#)

[New York Times](#)

Jurisdiction

71st Precinct, City Council District 35

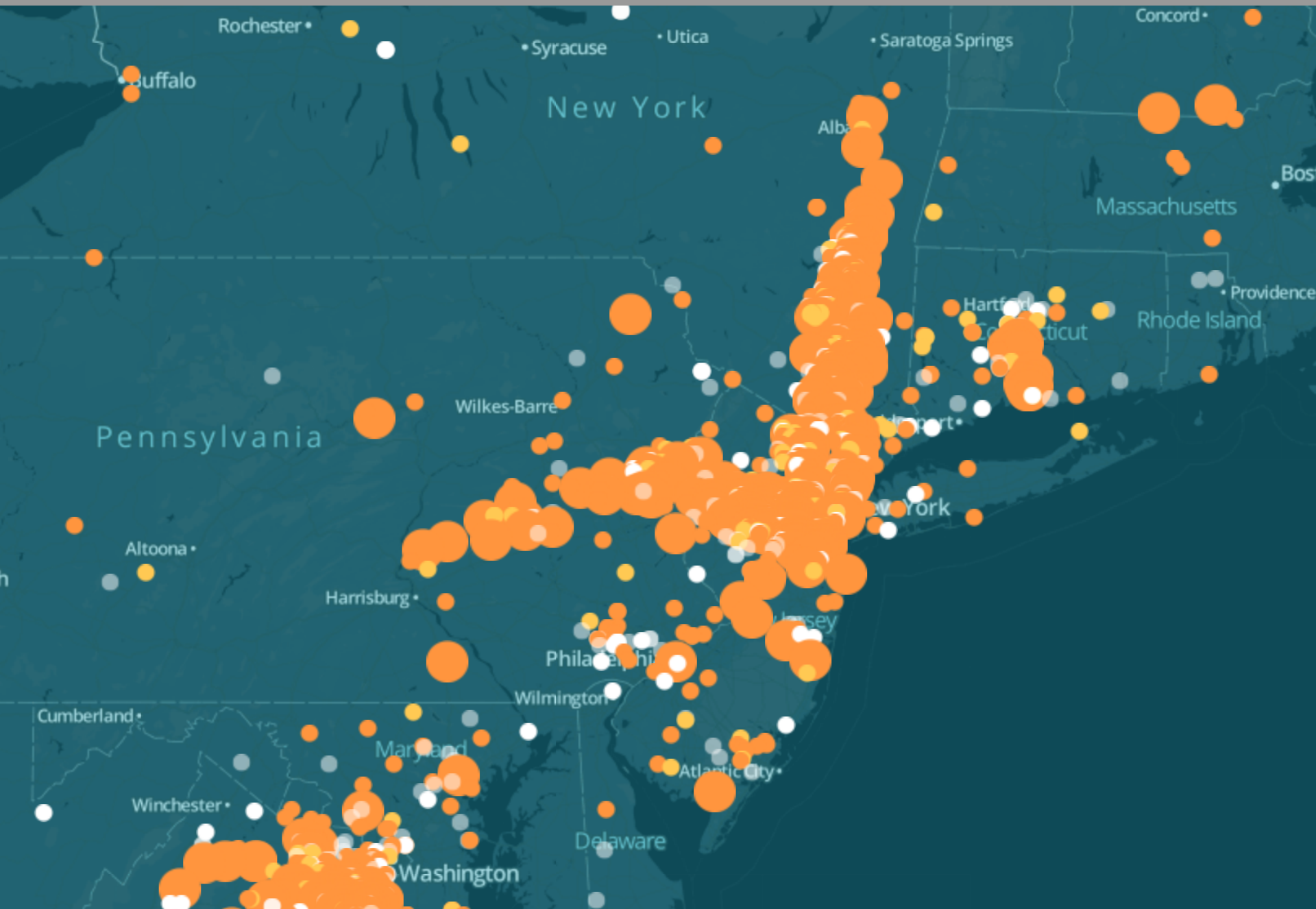
Vehicle

PLACE	POPULATION DENSITY	FACILITY TYPE	FEMA HISTORY	ENERGY AUDIT	TOTAL
-------	--------------------	---------------	--------------	--------------	-------

WHAT HOBOKEN SHOULD HAVE BEEN AWARDED

Multi-Service Center Shelter Generator					
10	20	20	20	90	points
Police HQ Generator					
10	20	20	20	90	
Wallace School Generator					
10	20	20	20	90	

Seen:   ● Nymphs   ● New adults   ● Adults   ● 100+ adults



**What else I do**



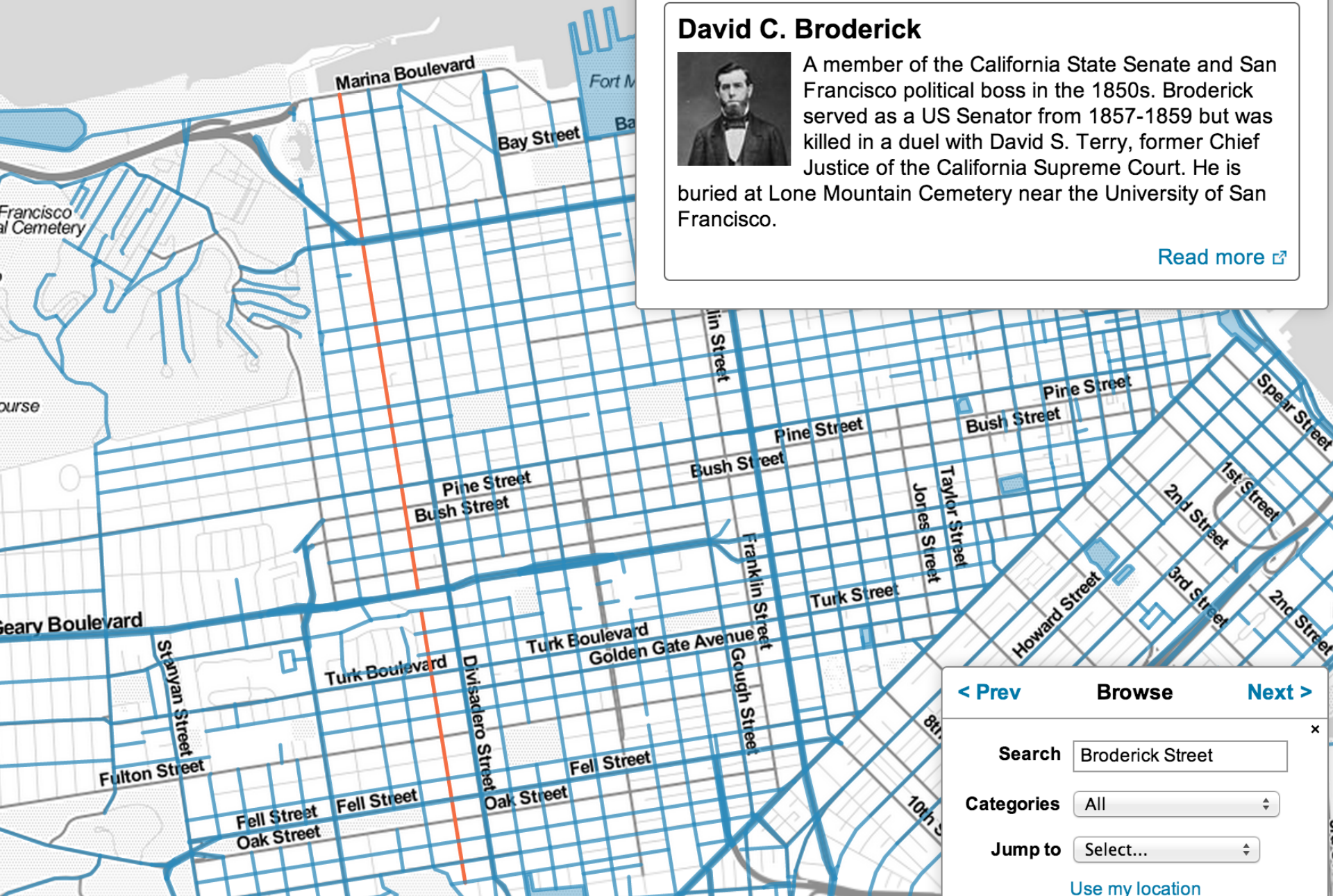
# Broderick Street

## David C. Broderick



A member of the California State Senate and San Francisco political boss in the 1850s. Broderick served as a US Senator from 1857-1859 but was killed in a duel with David S. Terry, former Chief Justice of the California Supreme Court. He is buried at Lone Mountain Cemetery near the University of San Francisco.

[Read more](#)



[< Prev](#) [Browse](#) [Next >](#)

Search

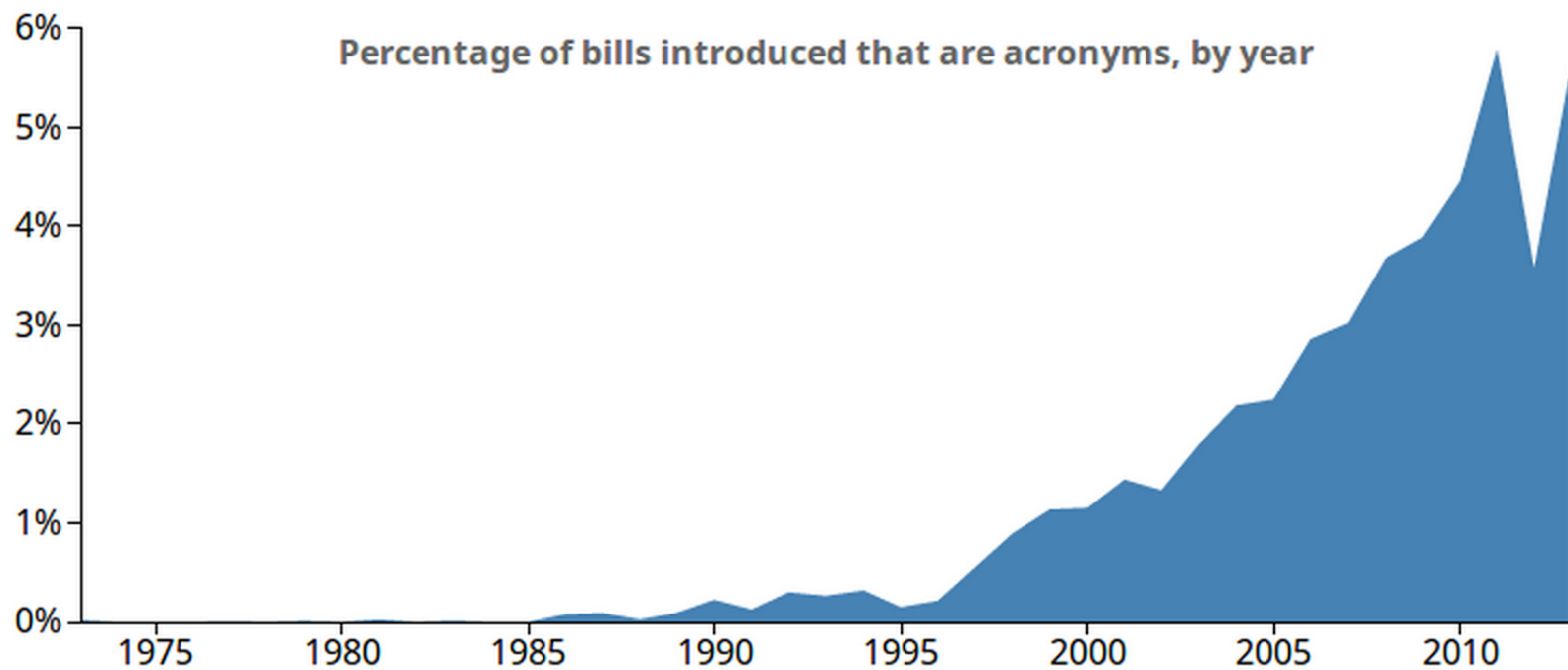
Categories

Jump to

[Use my location](#)



$$\text{crosswordiness}(w) = \frac{|\text{clues with answer } w|}{|\text{all clues}| \times \text{Google Books Ngram}}$$



## Highlights

### Most acronym-crazed legislators

The most prolific acronymmers in congress are:

- [Sen. Charles Schumer \(D-NY\)](#) - 42 acronym bills sponsored
- [Sen. Christopher Dodd \(D-CT\)](#) - 36 acronym bills sponsored
- [Sen. Robert "Bob" Menéndez \(D-NJ\)](#) - 34 acronym bills sponsored
- [Rep. Edward "Ed" Markey \(D-MA\)](#) - 31 acronym bills sponsored
- [Sen. Richard Durbin \(D-IL\)](#) - 30 acronym bills sponsored

The *percentage* leaders, with a minimum of 10 bills sponsored, are:

**What I really do**



Authority": "Y", "Dam Height (Ft.)": "29", "Max Discharge": null, "Latitude": "40.8869", "Fed Other  
Name": null, "Fed Regulatory": null, "Inspection Frequency": "0", "All Purposes": "Flood Control", "N  
", "Normal Storage": "21", "EAP": "NR", "Number Of Locks": null, "Max Storage": "80", "Submit Date"  
": "3.7", "Source Agency": "IA", "Foundation": "SK", "Jurisdictional Dam": "Y", "Inspection Authorit  
2.6458", "NID Height (Ft.)": "29", "Fed Inspection": null, "Number Of Separate Structures": "0", "Ri  
2", "Inspection\_Date": null, "Dam Length (Ft.)": "350", "State Reg Dam": "Y", "Outlet Gates": null,  
"Permitting Authority": "Y", "StateID": "IA03574000", "Other Structure Id": null, "Private\_Dam": "N"  
"Length Of Locks": "0", "Congressional Rep.": "David Loebsack (D)", "Fed Owner": null, "Fed Constr  
igner": "NRCS", "Owner\_Name": "DEAN CARL", "Section": "NW,SE,S27,T082N,R12W", "Primary Purpose": "Fi  
", "Dam Name": "CARL DAM", "Enforcement Authority": "Y", "Dam Height (Ft.)": "33", "Max Discharge":  
"0", "Url Address": null, "Other Dam Name": null, "Fed Regulatory": null, "Inspection Frequency": "  
03578", "Volume": "10100", "Width Of Locks": null, "Dam Type": "Earth", "Normal Storage": "20", "EAP  
(2013)", "Fed Funding": null, "Year Completed": "1999", "Surface Area": "2.3", "Source Agency": "IA",  
Fed Design": null, "Structural Height (Ft.)": null, "Longitude": "-92.2275", "NID Height (Ft.)": "33  
BUCKEYE CREEK", "Drainage\_Area": ".07", "Congressional District": "IA01", "Inspection\_Date": null,  
Type": "Private", "Spillway\_Width": "1", "NID Storage": "29", "Permitting Authority": "Y", "StateID"  
", "County": "BENTON", "Hydraulic Height (Ft.)": null, "Length Of Locks": "0", "Congressional Rep.":  
Party": "D", "State Reg Agency": "Iowa DNR"}{"Dam\_Designer": "NRCS", "Owner\_Name": "JOHN SELLS", "S  
r Small Fish Pond", "Fed Operation": null, "State": "IA", "Dam Name": "SELLS DAM", "Enforcement Auth  
3723", "Fed Other": null, "Core": "HK", "Year Modified": "0", "Url Address": null, "Other Dam Name":  
Fire Protection, Stock, Or Small Fish Pond", "NIDID": "IA03595", "Volume": "36200", "Width Of Locks":  
cks": null, "Max Storage": "309", "Submit Date": "02\\04\\2013", "Fed Funding": null, "Year Complete  
"Jurisdictional Dam": "Y", "Inspection Authority": "Y", "Fed Design": null, "Structural Height (Ft.  
null, "Number Of Separate Structures": "0", "River": "TR- LITTLE SIOUX RIVER", "Drainage\_Area": "1.  
t.)": "320", "State Reg Dam": "Y", "Outlet Gates": null, "Owner Type": "Private", "Spillway\_Width":  
her Structure Id": null, "Private\_Dam": "N", "Spillway\_Type": "U", "County": "CHEROKEE", "Hydraulic  
(R)", "Fed Owner": null, "Fed Construction": null, "Political Party": "R", "State Reg Agency": "IOW  
E,SW,S26,T071N,R24W", "Primary Purpose": "Fire Protection, Stock, Or Small Fish Pond", "Fed Operatio  
Y", "Dam Height (Ft.)": "22", "Max Discharge": null, "Latitude": "40.9154", "Fed Other": null, "Cor  
Fed Regulatory": null, "Inspection Frequency": "0", "All Purposes": "Fire Protection, Stock, Or Smal  
Dam Type": "Earth", "Normal Storage": "24", "EAP": "NR", "Number Of Locks": null, "Max Storage": "91

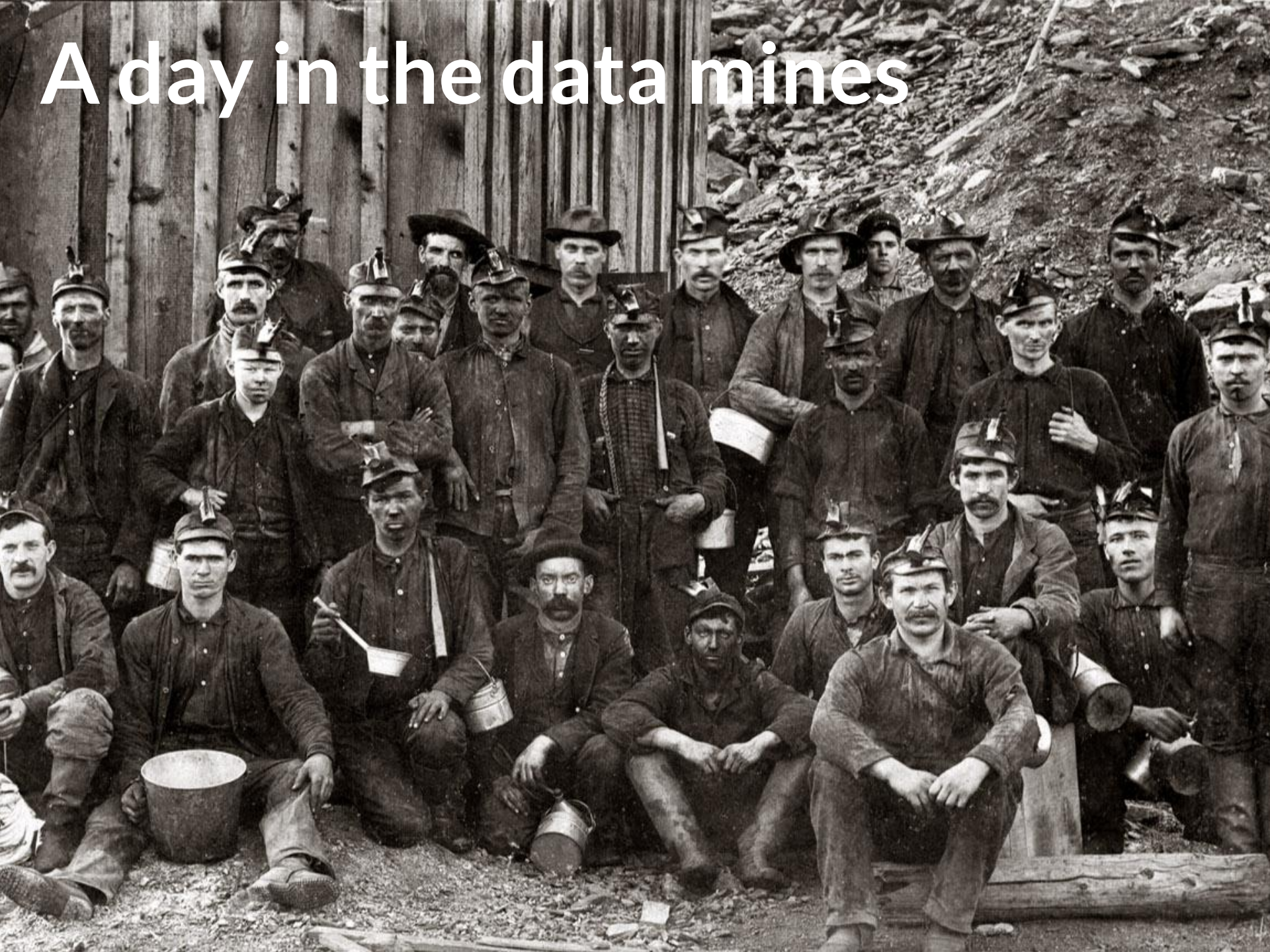
A photograph of an iceberg floating in the ocean. The visible tip of the iceberg is small and jagged, while the submerged portion is much larger and more complex in shape, illustrating the concept of hidden data. The water is a deep blue, and the sky is a lighter blue with some wispy clouds.

Fun little visualization

Fifty hours of data work



# A day in the data mines







**Proxy data**



# The Big Mac index

Select base currency:

US dollar



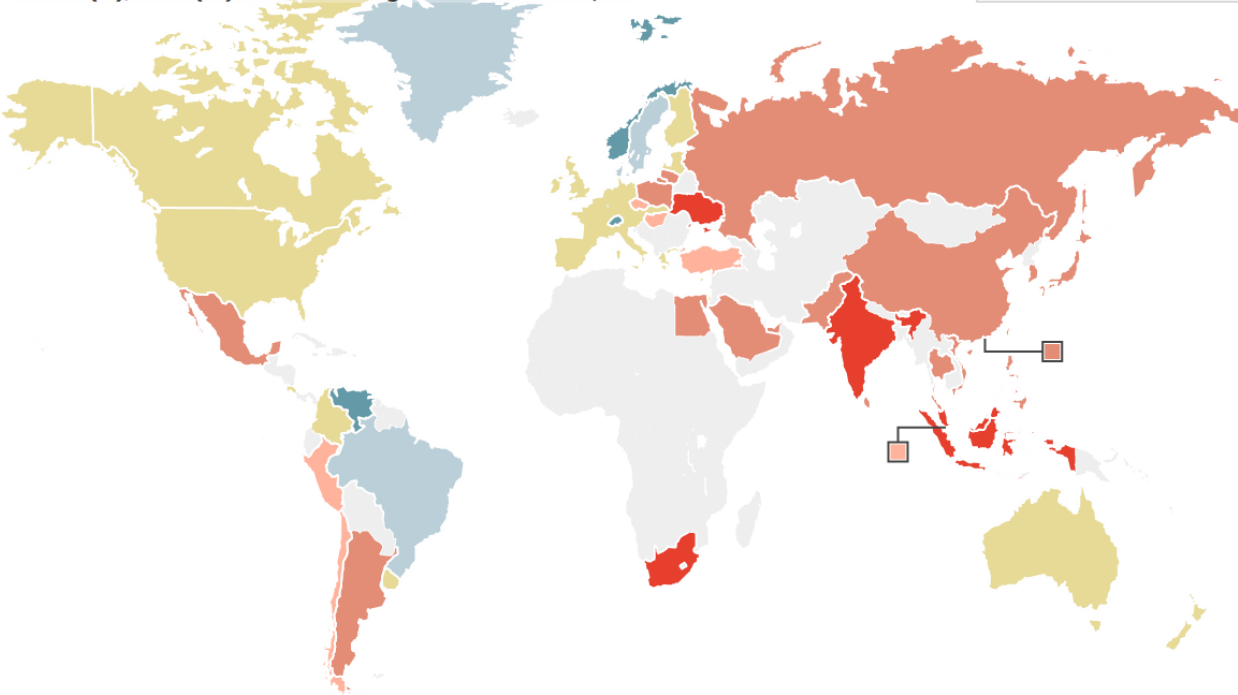
Raw index

Adjusted index

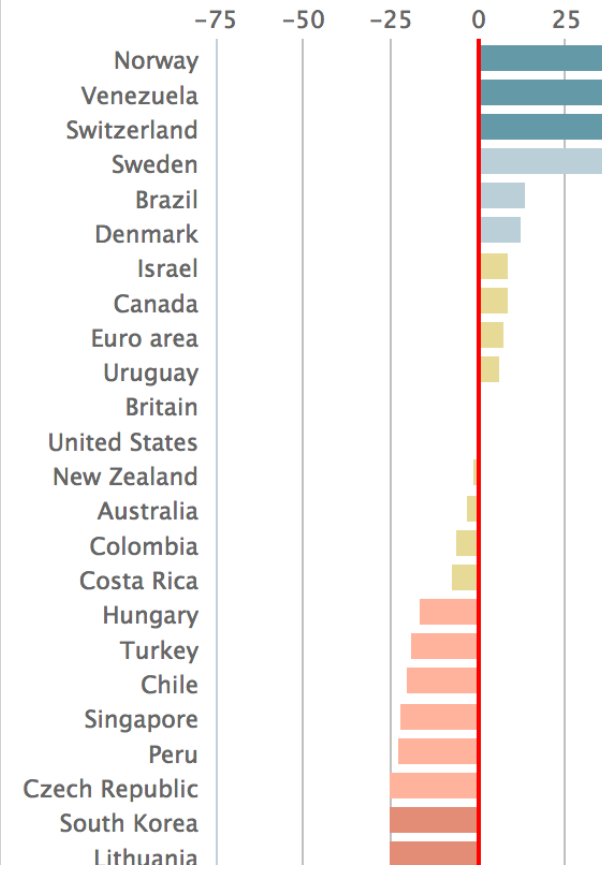
## Raw index

Under(-)/over(+) valuation against the dollar, %

Zoom to

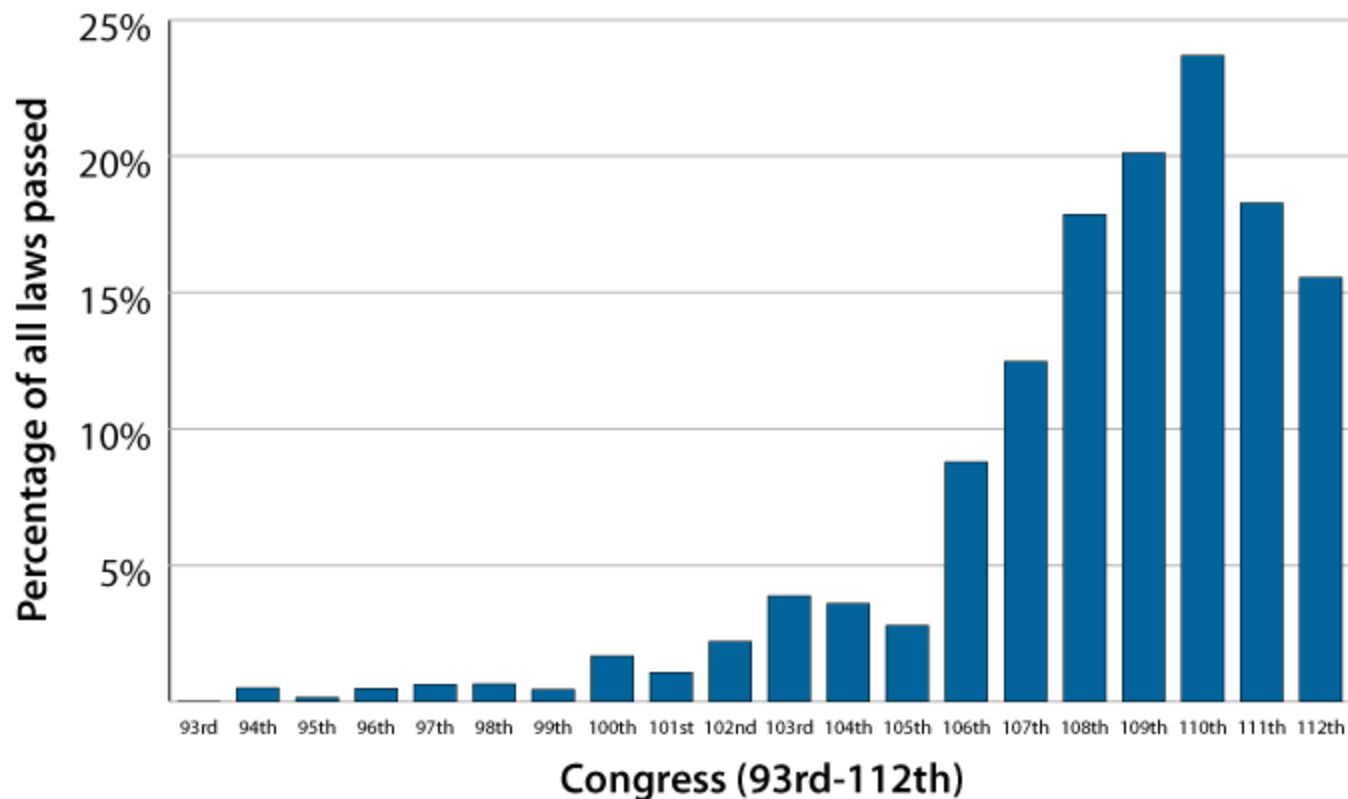


January 2014



## Federal laws passed for the renaming of post offices

*by United States Congress, as a percentage of all laws passed*

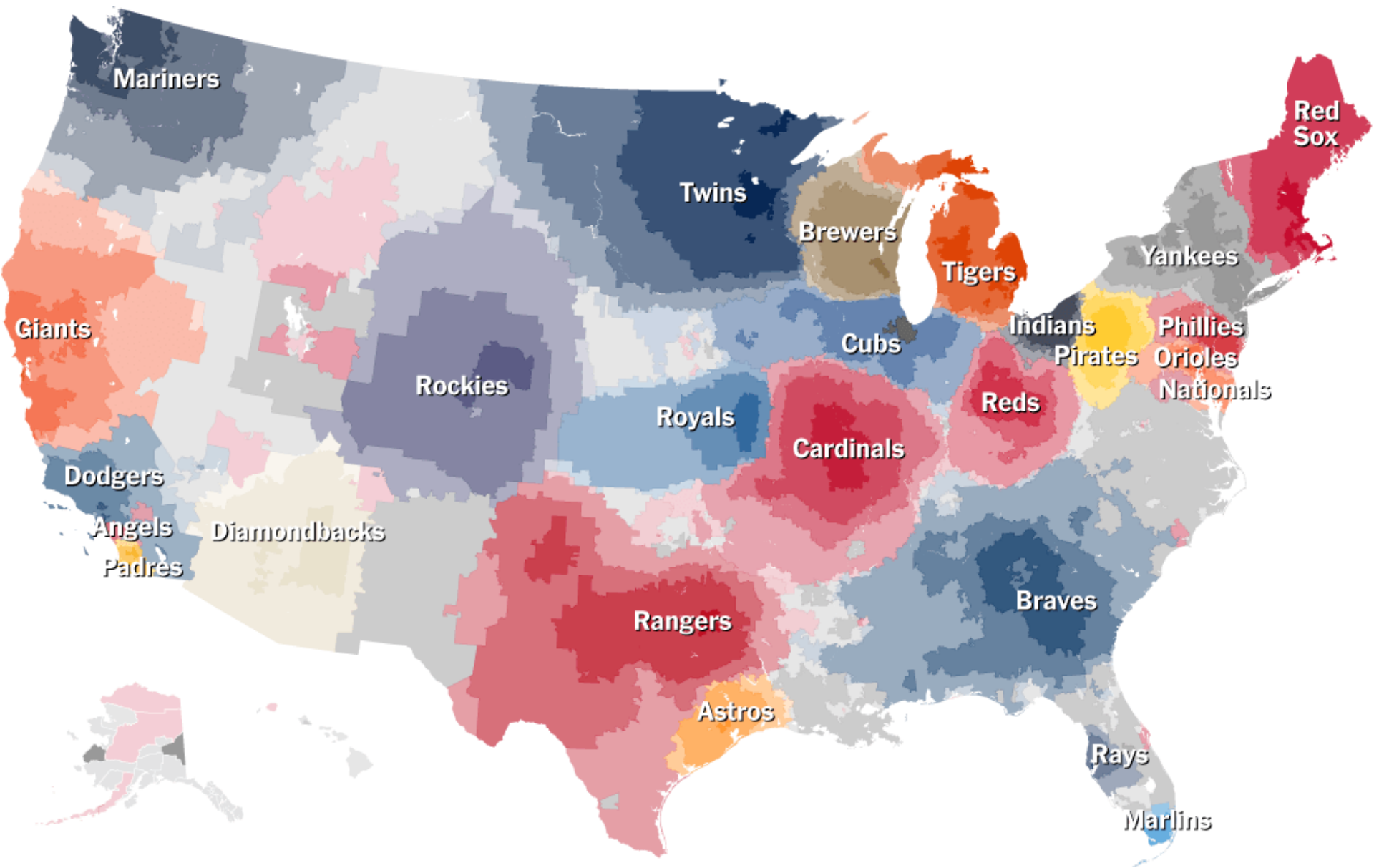


H.R.514 : To designate the facility of the United States Postal Service located at 16150 Aviation Loop Drive in Brooksville, Florida, as the "Sergeant Lea Robert Mills Brooksville Aviation Branch Post Office".

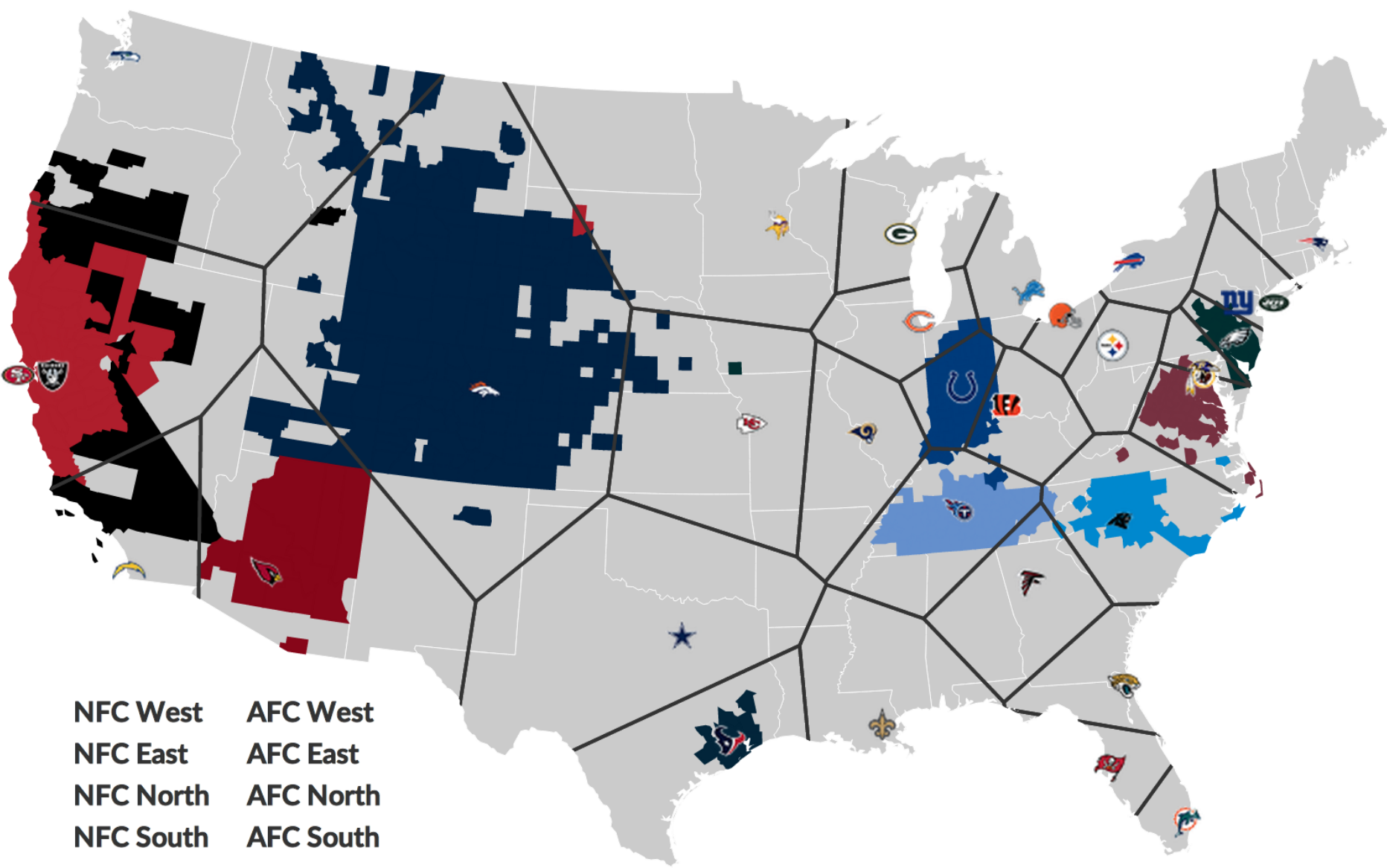
H.R.577 : To designate the facility of the United States Postal Service located at 3903 South Congress Avenue in Austin, Texas, as the "Sergeant Henry Ybarra III Post Office Building".

H.R.988 : To designate the facility of the United States Postal Service located at 5757 Tilton Avenue in Riverside, California, as the "Lieutenant Todd Jason Bryant Post Office".

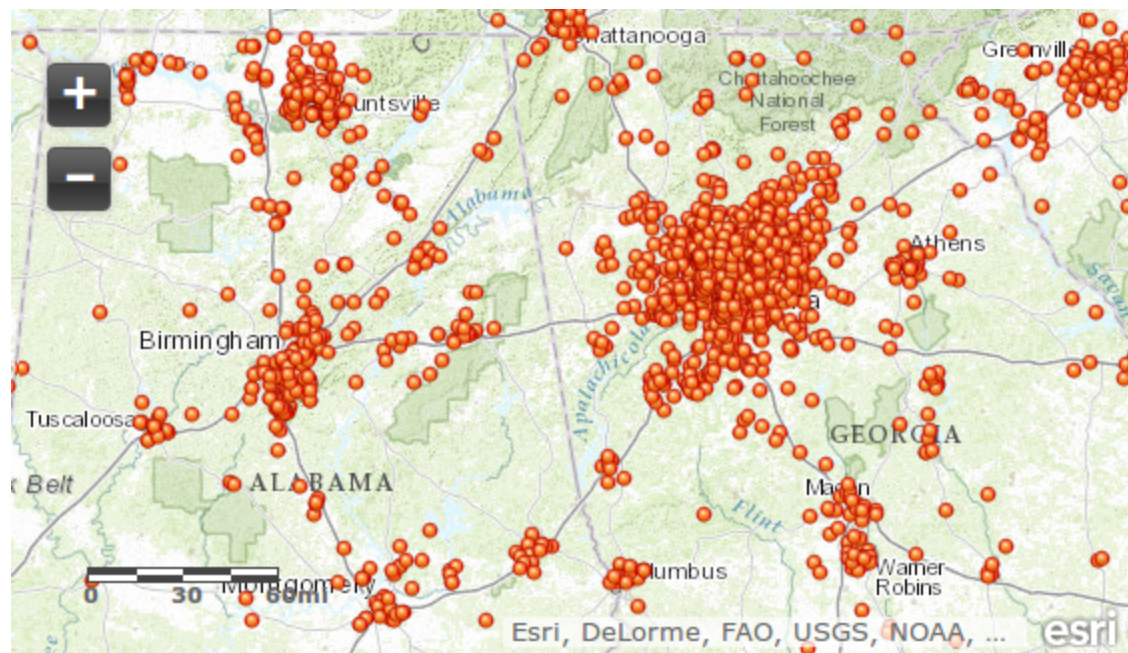
H.R.1402 : To designate the facility of the United States Postal Service located at 320 South Lecanto Highway in Lecanto, Florida, as the "Sergeant Dennis J. Flanagan Lecanto Post Office Building".





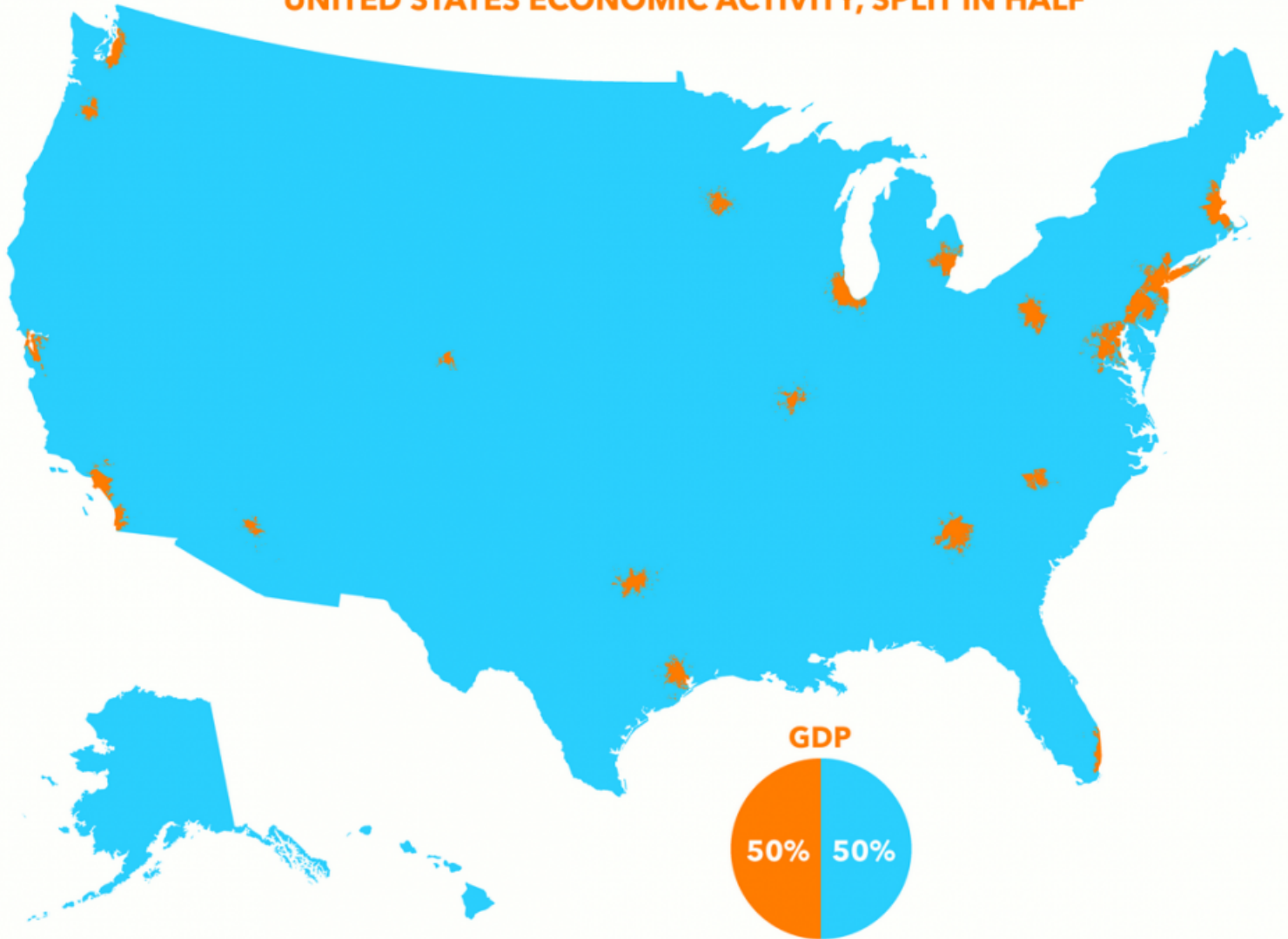


NFC West	AFC West
NFC East	AFC East
NFC North	AFC North
NFC South	AFC South



ESRI

## UNITED STATES ECONOMIC ACTIVITY, SPLIT IN HALF



*A. Veltchikov*

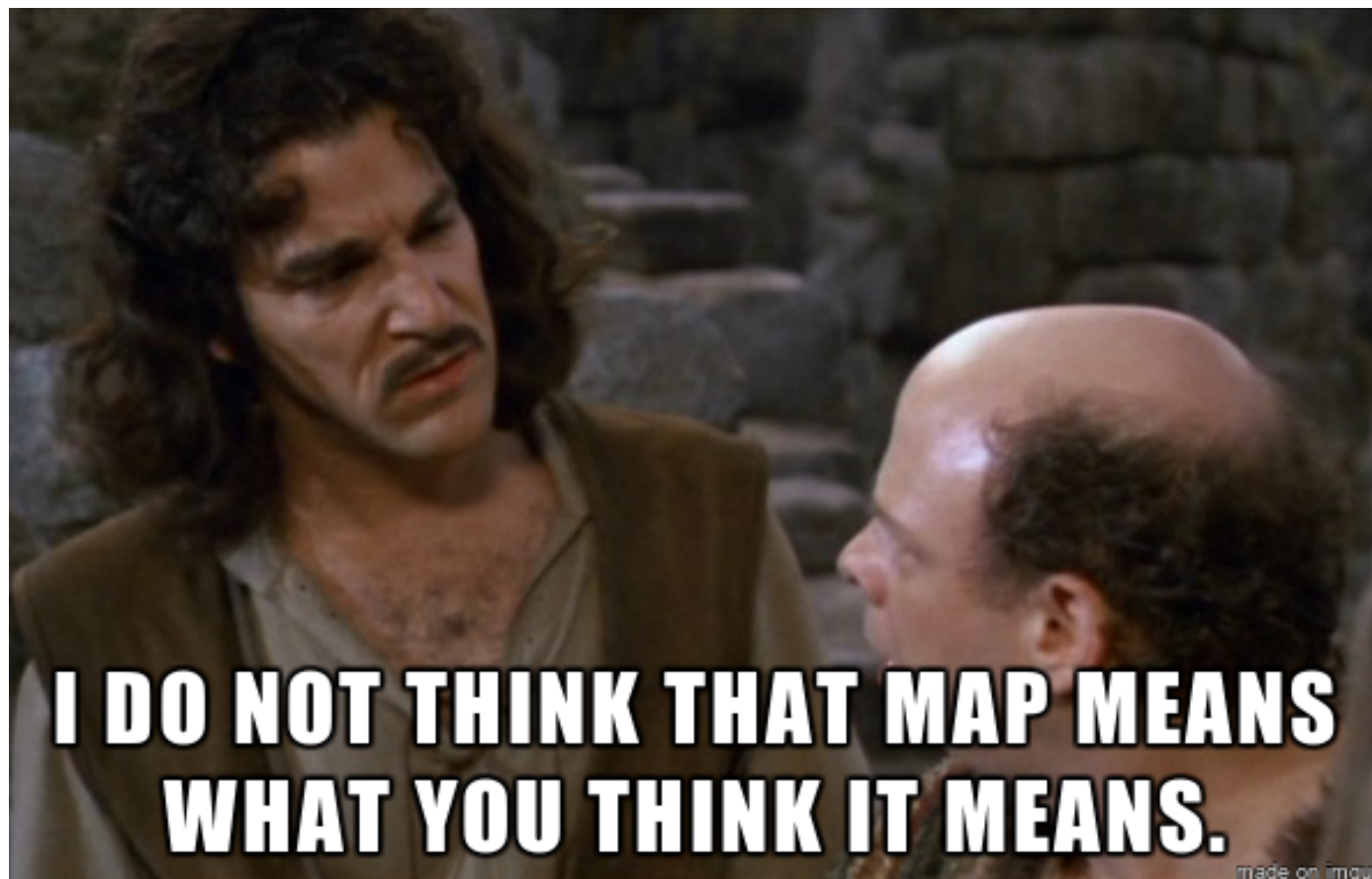
/u/a trubetskoy





Flowing Data/Runkeeper





HHISPANIC	HI	HIAP	HIAPANIC
HIASPANIC	HICPANIC	HIDP	HIDPANIC
HIHSP	HIIIIIISPANIC	HIIIISP	HIIISP
HIKSPANIC	HIOSP	HIOSPANIC	HIOSP[ANIC
HIPANICX	HIPS	HIPSANIC	HIPSNAIC
HIS	HIS0P	HIS;ANIC	HISANIC
HISAPANIC	HISAPNIC	HISDP	HISDPANIC
HISO	HISOANIC	HISOPANIC	HISP
HISP-ANIC	HISP.	HISP.\	HISP3
HISPA NIC	HISPAANIC	HISPABIC	HISPABNI
HISPACNI	HISPAIC	HISPAICI	HISPAINC
HISPAMIC	HISPAN	HISPAN IC	HISPANAIC
HISPANCI	HISPANCIS	HISPANI	HISPANI C
HISPANIC	HISPANIC/LATIN	HISPANIC/LATINO	HISPANIC/LATNIO
HISPANIC2	HISPANIC3	HISPANICA	HISPANICC
HISPANICF	HISPANICJ	HISPANICLE	HISPANICMJJJJJ
HISPANICQ	HISPANICS	HISPANICV	HISPANICX
HISPANIC`	HISPANID	HISPANIF	HISPANIFC
HISPANIKC	HISPANINC	HISPANINC.	HISPANIOC
HISPANIUC	HISPANIV	HISPANIVC	HISPANIX
HISPANNIC	HISPANOC	HISPANOIC	HISPANSIC
HISPAQNIC	HISPASNIC	HISPC	HISPI
HISPINIC	HISPNAIC	HISPANANIC	HISPNIC
HISPP	HISPPANIC	HISPQ	HISPS
HISPSP	HISP[	HISP[ANIC	HISP\
HISSPANIC	HISWP	HIS[	HIS[ANIC
HIUSPANIC	HIWPANIC	HJISP	HJISPANIC
HISPANTIC	HISPANIC	HSIP	HSIPANIC
HISPANII	HISPANIC	HUSPANIC	HYISP
OTHER HISP	OTHER HISPAANIC	OTHER HISPANIC	OTHER HISPANIK

Dirty data



South Korea  
Korea  
S. Korea  
Republic of Korea

Cahuenga

Cahaunga

Caheinga

Cahenga

Caheunga

Cahienga

Cahlenga

Cahoengia

Cahubnga

Cahue

Cahu





## < Meet Sooty >

(4 of 11)

**Name:** Sooty

**Age:** Five to 10 years    **Sex:** Male

**Breed/colour:** Short-haired black moggie/farm cat

**Character:** Friendly, tolerant and active, but somewhat aloof and calm. Doesn't fight. Lives with sibling Socks

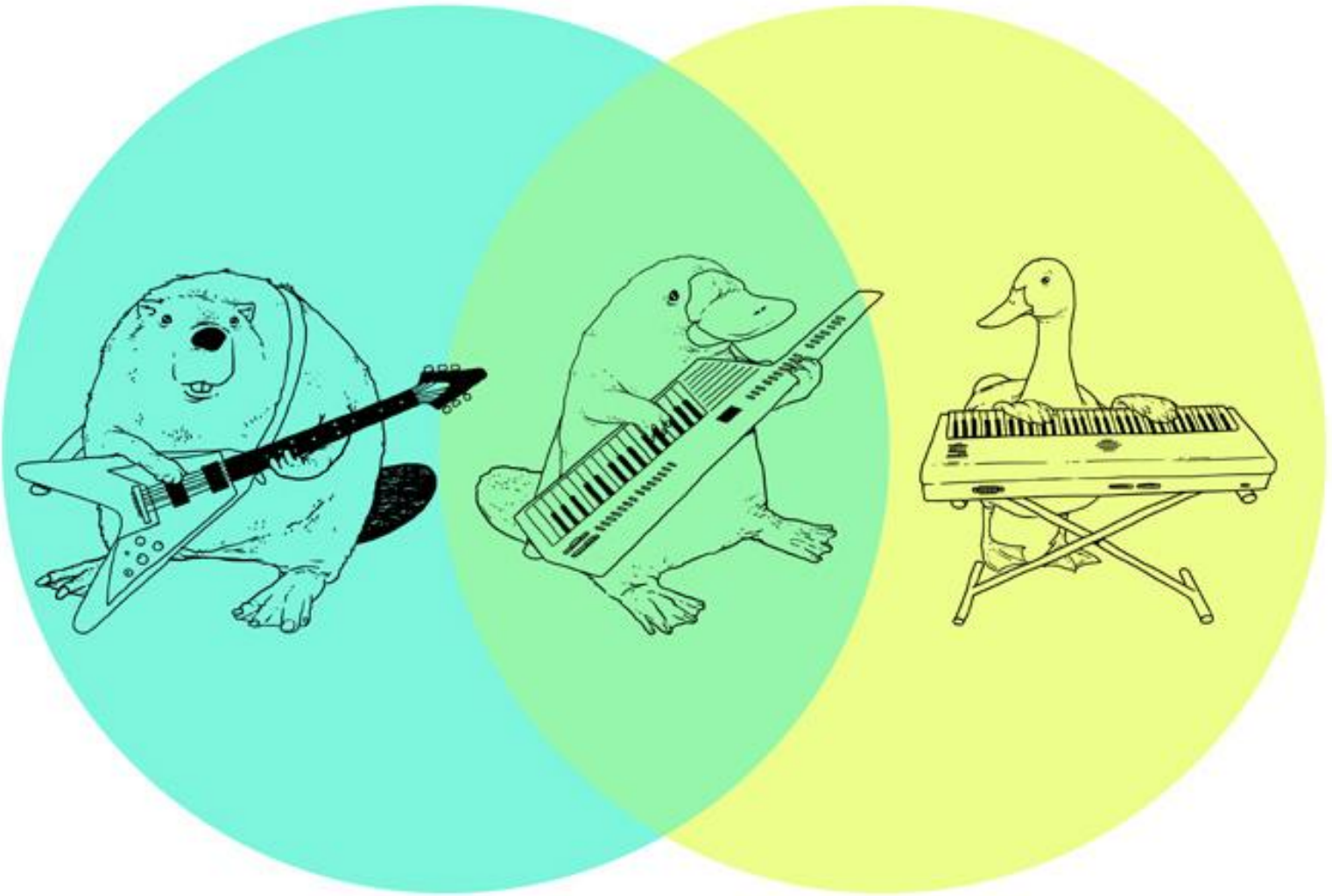
**Roaming:** Roams 186m (600ft) from home and covers nearly three hectares (seven acres)

**Prey:** Brings home small mammals a couple of times a week

**Expert view:** "Sooty has the largest roaming range of our 10 cats and liked to visit some trees on farmland next to some water. His house is on the edge of the village and so he has easy access to the countryside," say researchers.



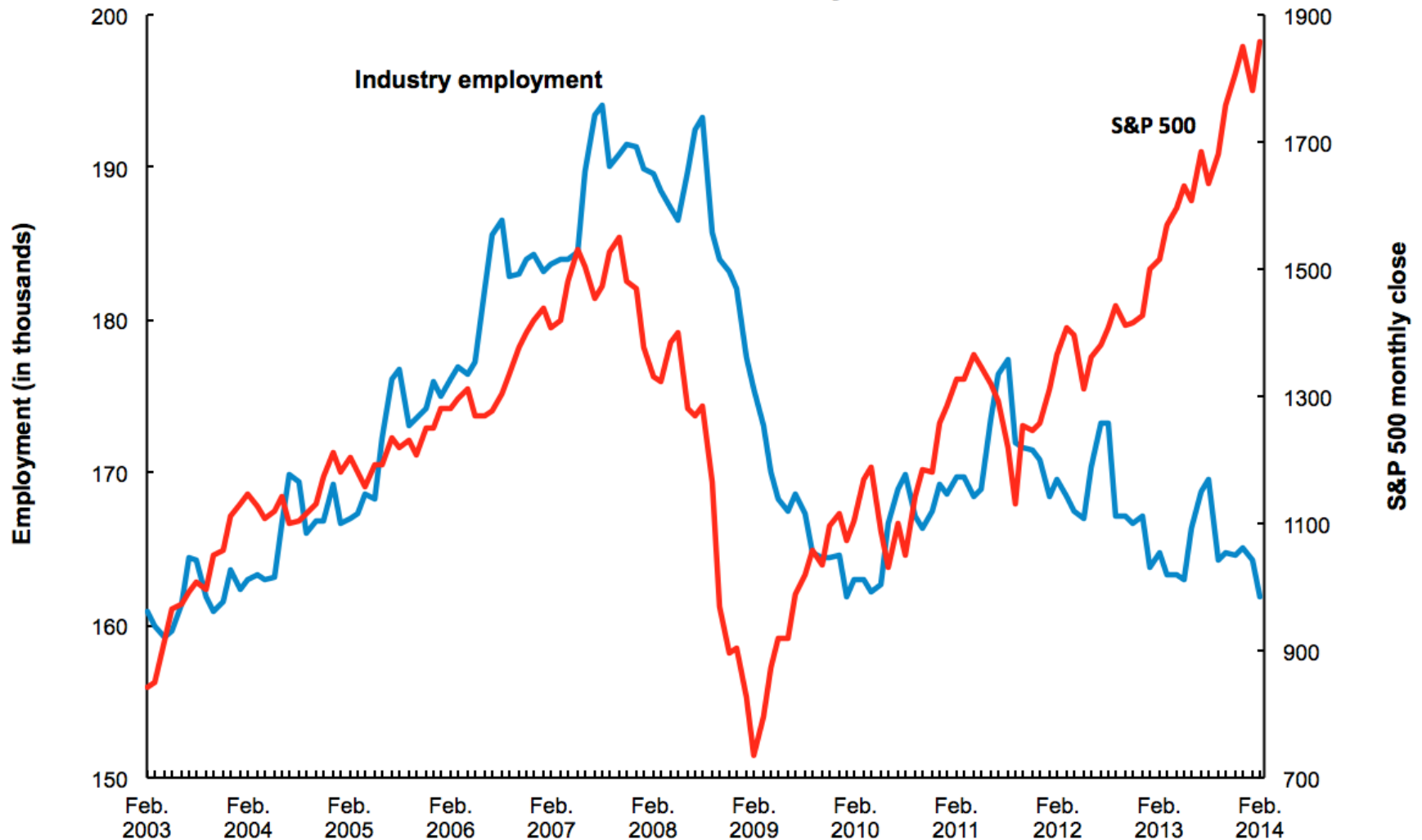
# Just plain wrong data



**Imperfect taxonomies**

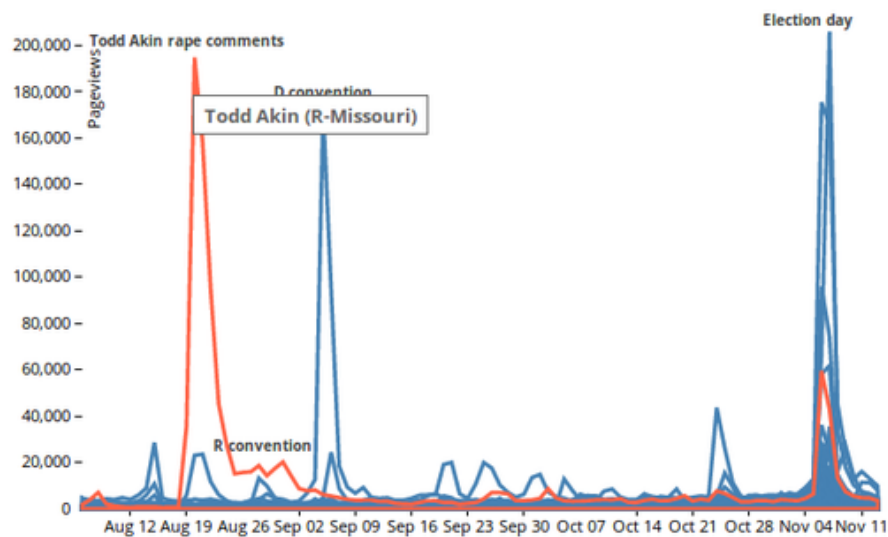
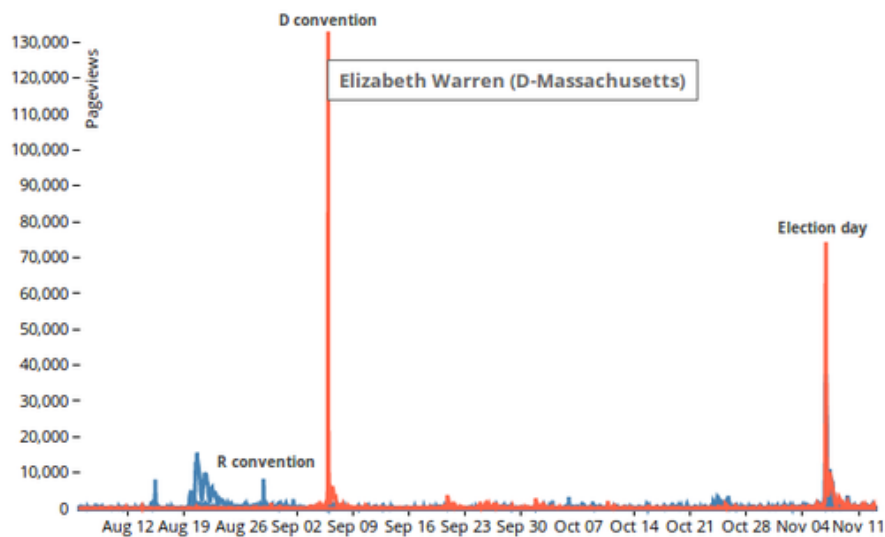
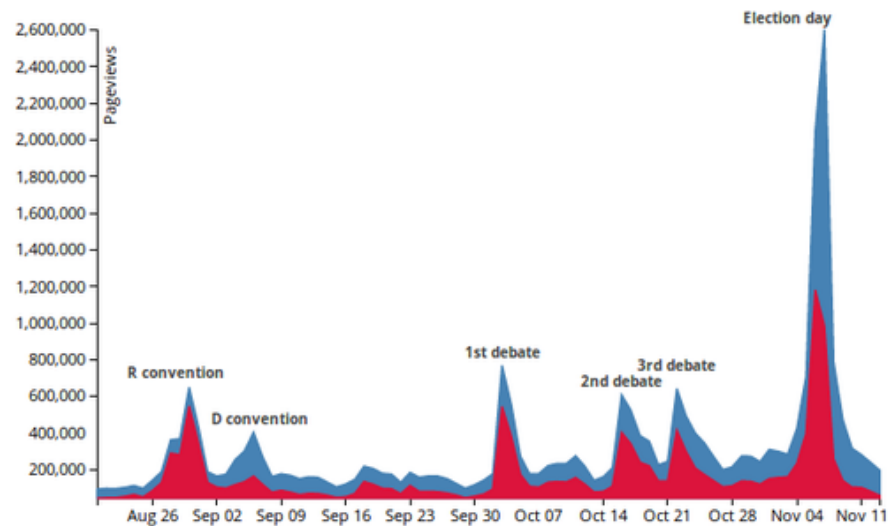
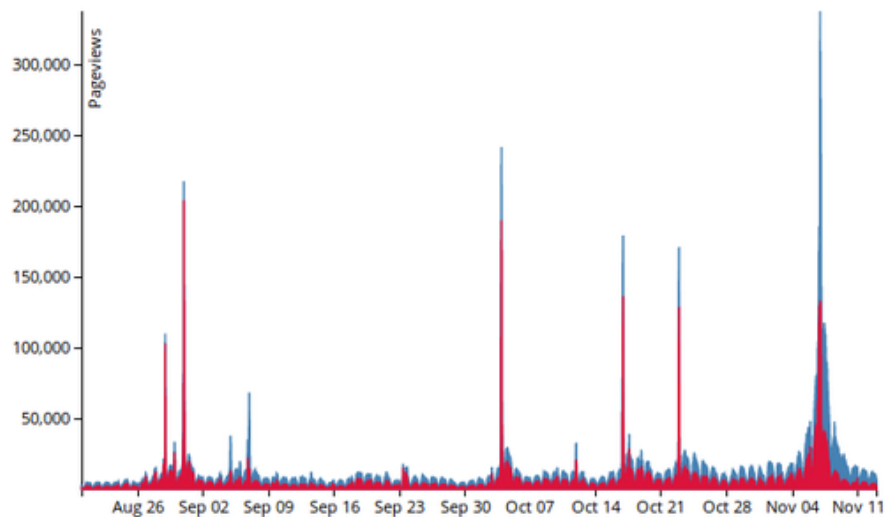


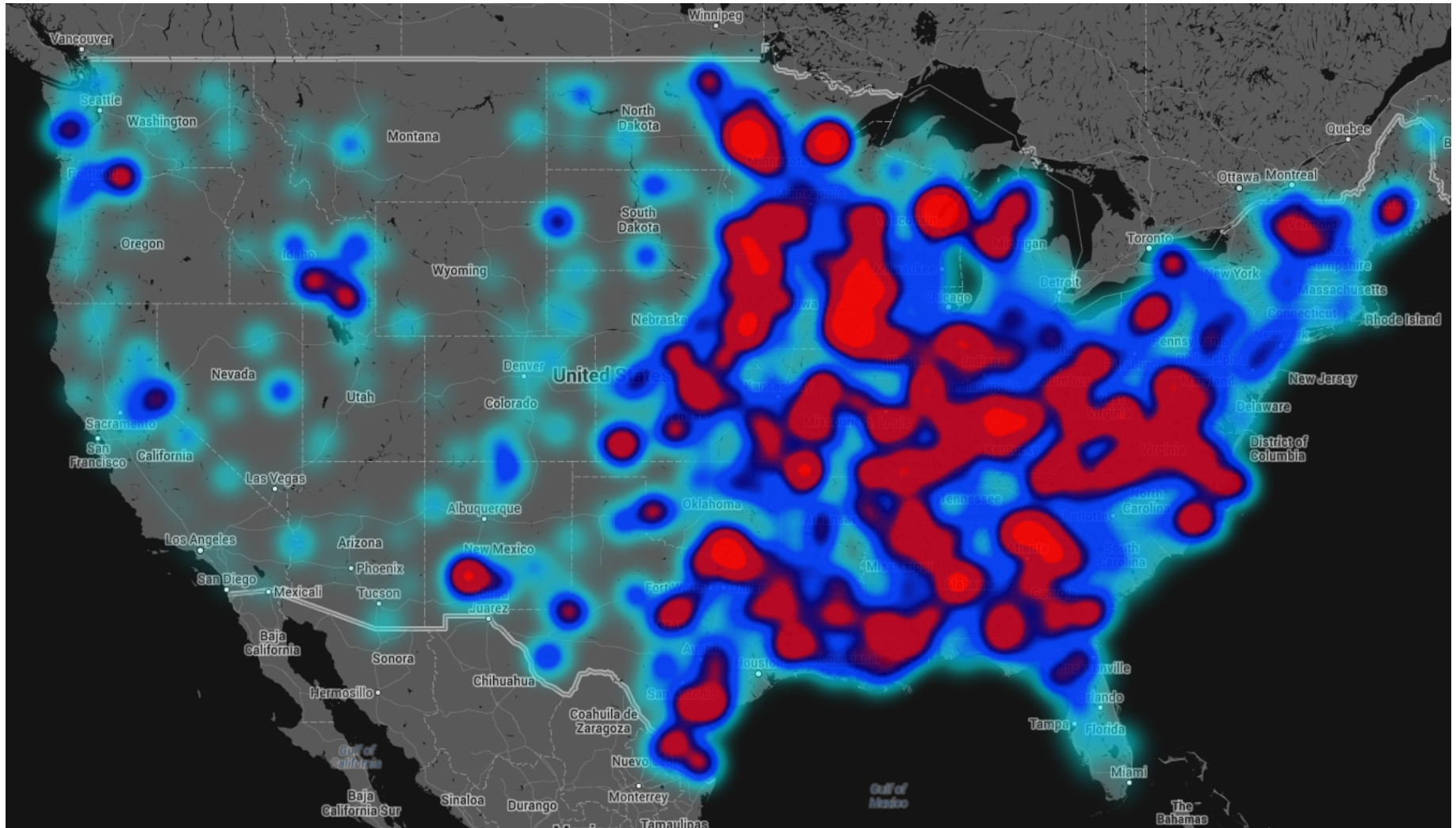
**Employment in the securities industry in New York City and the S&P 500, February 2003 - February 2014**



**Scale and aggregation**





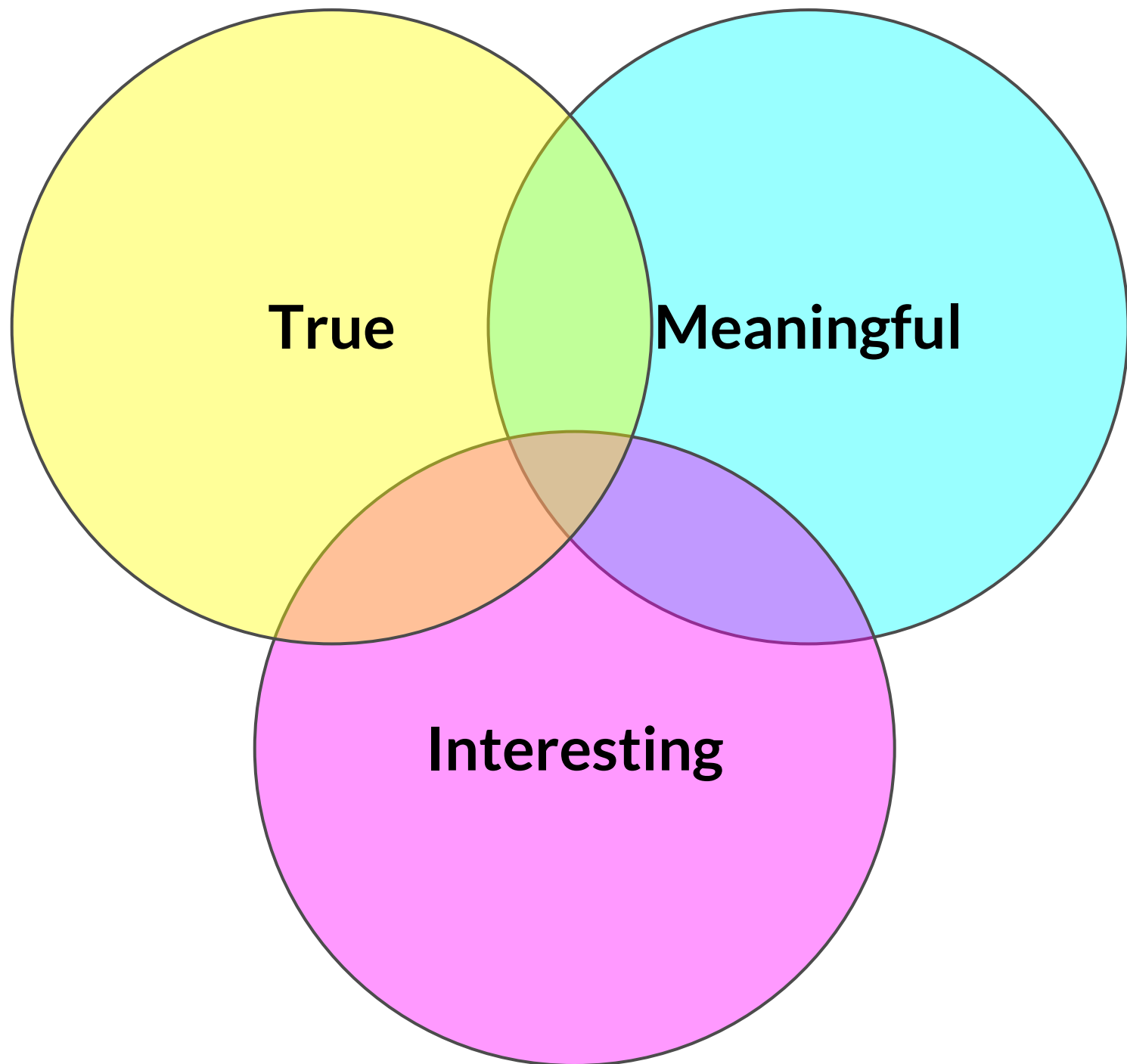


Humboldt State University

**Data doesn't exist in a vacuum.**

It's easier than ever to make  
something **true** and **dumb**.





**We've figured out the data.**  
**Now what?**

# **Principles of designing interactives for the web**

Users are  
impatient, fat-fingered bums.

Don't make **scavenger hunts**.



Make it **personal.**

What's the lede?

**Build of the web, not just for it.**

**Edit ruthlessly.**

# Science vs. Journalism



Different **time scales**.

Different audiences.

Who **collects** the data?

Things **journalists** can learn  
from **scientists**

Norms of openness.



Norms of **reproducibility**.

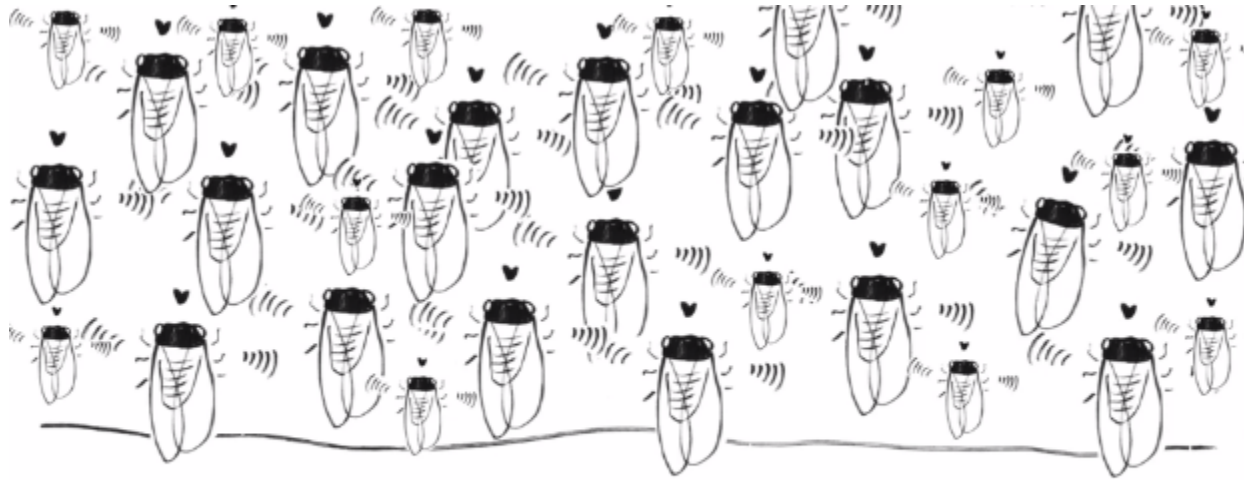
Norms of **uncertainty**.

**Build things that endure.**

Driven by **hypotheses.**

**Collecting** your own data.

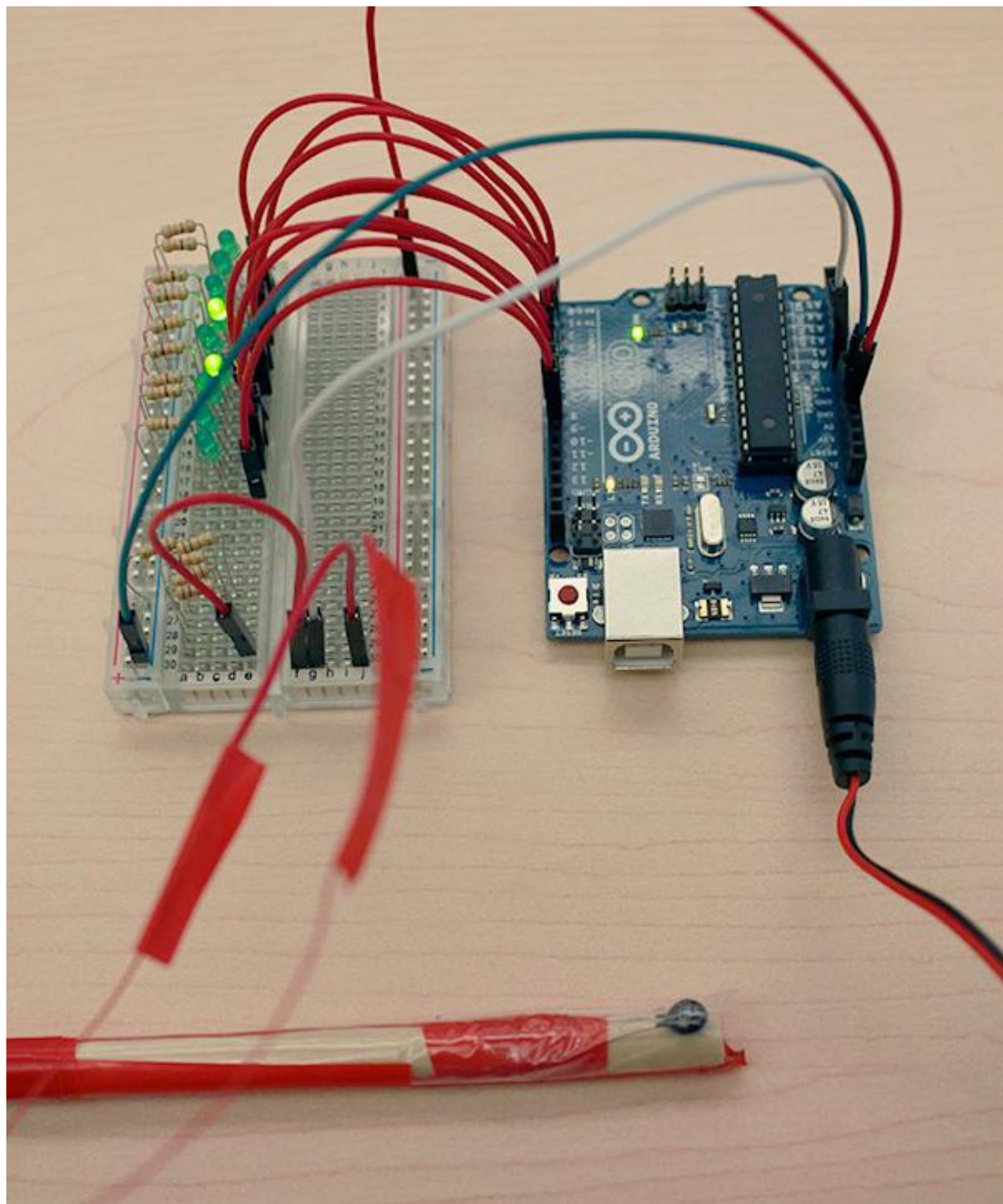




**Get ready for Swarmageddon**

**Citizen science**





My Sleep

Findings

Stories

Sleep Diary

Profile

My Progress

Fill in a wake time

If you mis  
still edit, c

Apr 27 - M

Apr 20 - A

Apr 13 - A

Apr 6 - Ap

Mar 30 - A

Last night

This morning



Bedtime

Wed Apr 30

Time I turned out the lights ⓘ

Select

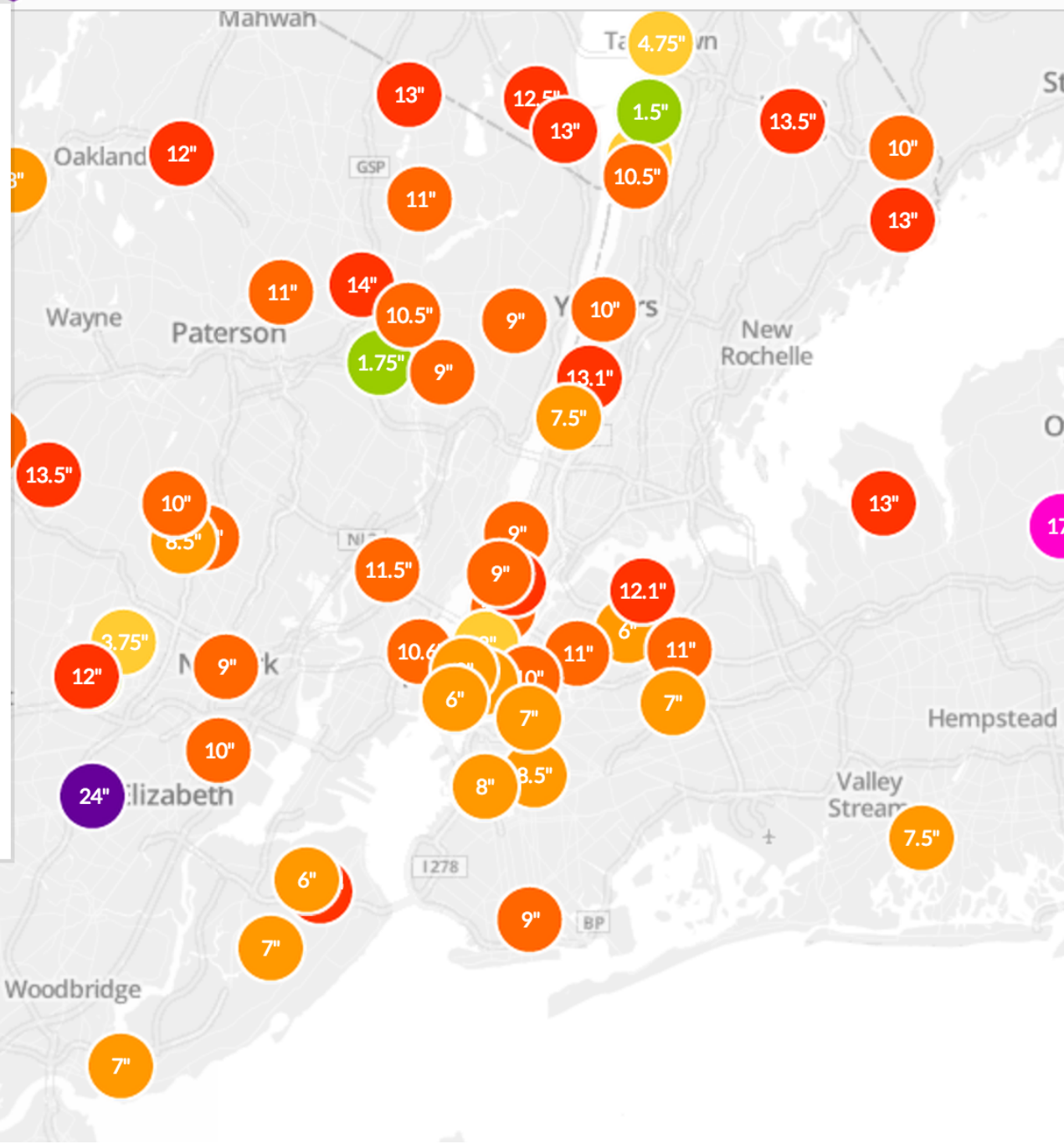


Wakeup

Thu May 1

Time I woke up for the day ⓘ

Select



# Citizen science lessons

Everyone is a **sensor**.



# Participation vs. Rigor

**Experts welcome.**

Make science **personal.**

**Make science a team sport.**

# Questions?

**Noah Veltman**

WNYC Data News Team

@veltman

noahveltman.com