Feedback for
**Visualization Part Only**

# Visualization Q&A

The Human Factor

*Benjamin Bach*

April 2019

Edinburgh University

*bbach@inf.ed.ac.uk*

# Outline

1. Foundations
2. Task and Scenarios **> TUTORIAL I**
3. Visualization Techniques I
4. Visualization Techniques II
5. Design **> TUTORIAL II**
6. Sketching session (bring data)
7. Presentation and Storytelling **> TUTORIAL III**
8. Interaction
9. Vis tools
10. Evaluation
11. general Q&A

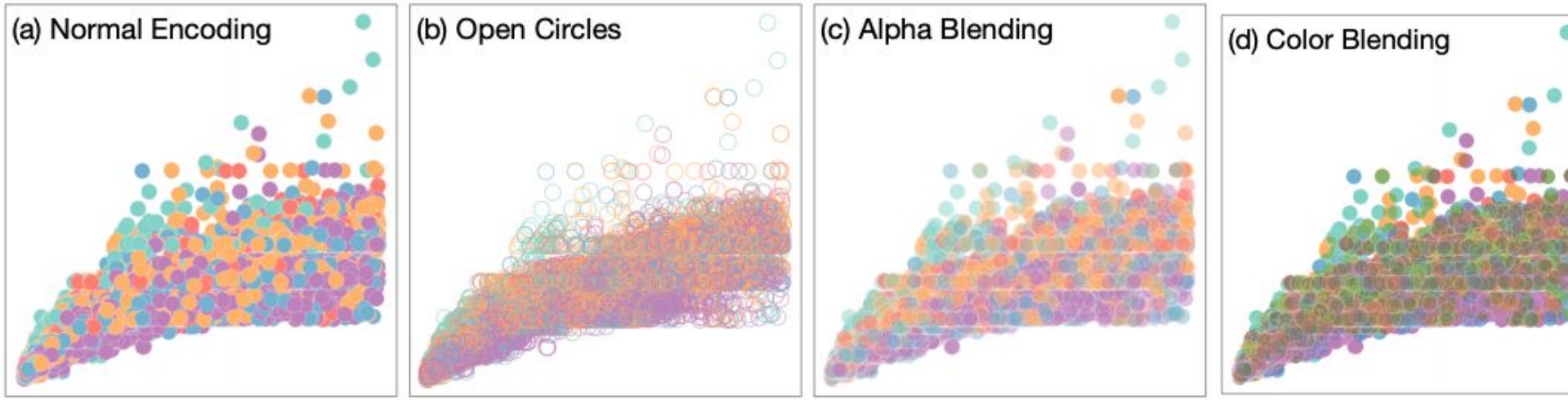
Feedback for
**Visualization Part Only**

1. **Are we allowed to use infographic template websites such as canvas or vengage for the project?**

# 2. Can we use data from kaggle.com?

https://www.kaggle.com/datasets

# Overplotting + big data



# Best practises for large data sets

1. Brute-force
2. Chose the important
    a. Exploratory data analysis
    b. Multiple views
3. Interaction
4. Animation
5. Aggregation

# What is big data?

1. Volume
2. Variety
3. Velocity
4. Veracity

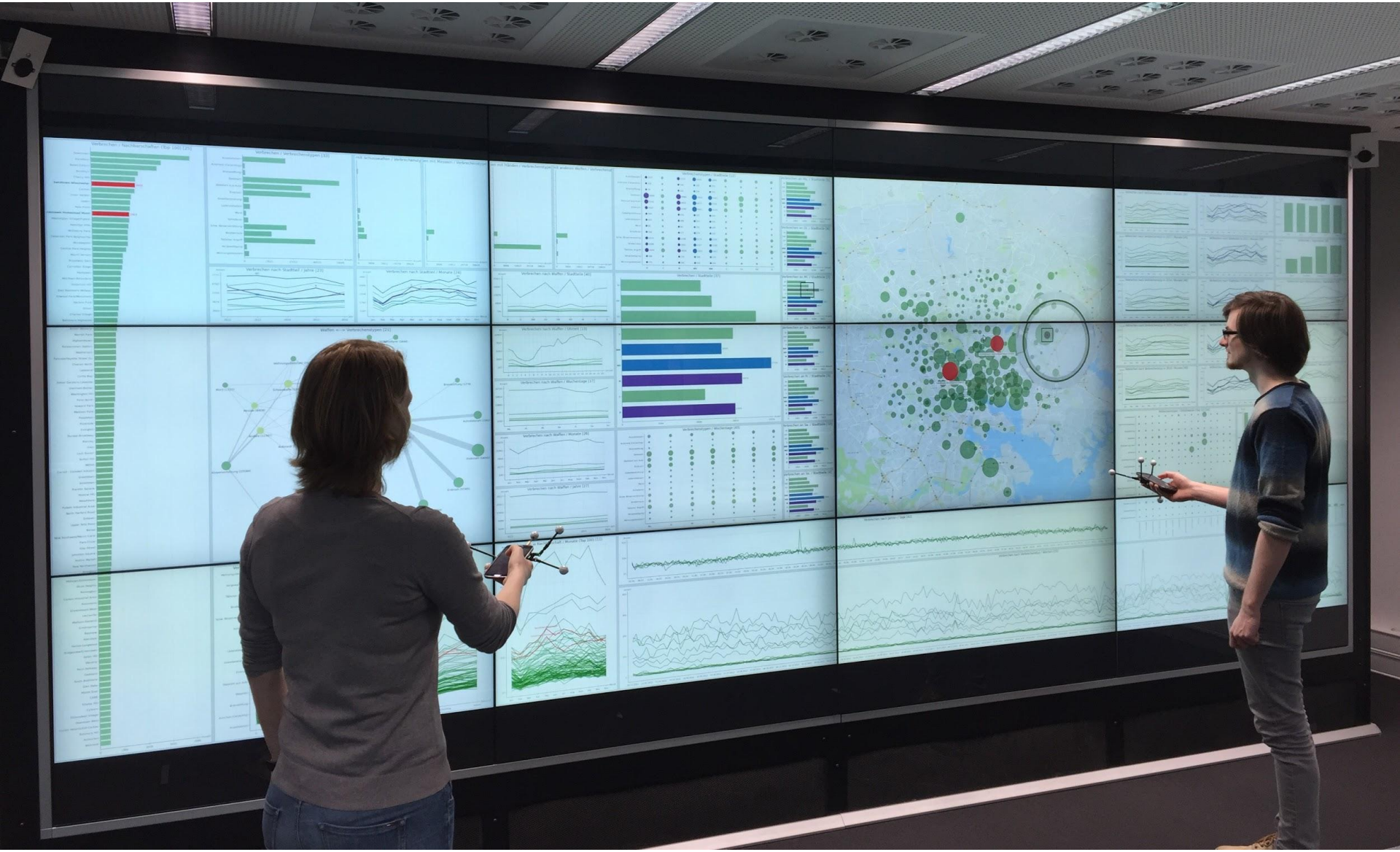
| PASSING STATS | | | | | |
|---|---|---|---|---|---|
| SEASON | TEAM | GP | CMP | ATT | CMP% |
| 2005 | GB | 3 | 9 | 16 | 56.3 |
| 2006 | GB | 2 | 6 | 15 | 40.0 |
| 2007 | GB | 2 | 20 | 28 | 71.4 |
| 2008 | GB | 16 | 341 | 536 | 63.6 |
| 2009 | GB | 16 | 350 | 541 | 64.7 |
| 2010 | GB | 15 | 312 | 475 | 65.7 |
| 2011 | GB | 15 | 343 | 502 | 68.3 |
| 2012 | GB | 16 | 371 | 552 | 67.2 |
| 2013 | GB | 9 | 193 | 290 | 66.6 |
| 2014 | GB | 16 | 341 | 520 | 65.6 |
| 2015 | GB | 16 | 347 | 572 | 60.7 |
| 2016 | GB | 16 | 401 | 610 | 65.7 |
| 2017 | GB | 6 | 128 | 193 | 66.3 |
| **Career** | | **148** | **3,162** | **4,850** | **65.2** |

# Brute force

# Brute force

1. Toolkits (deck.gl, mapbox, processing)
2. OpenGL (Unity)
3. Large screens

# Multi-resolution walls

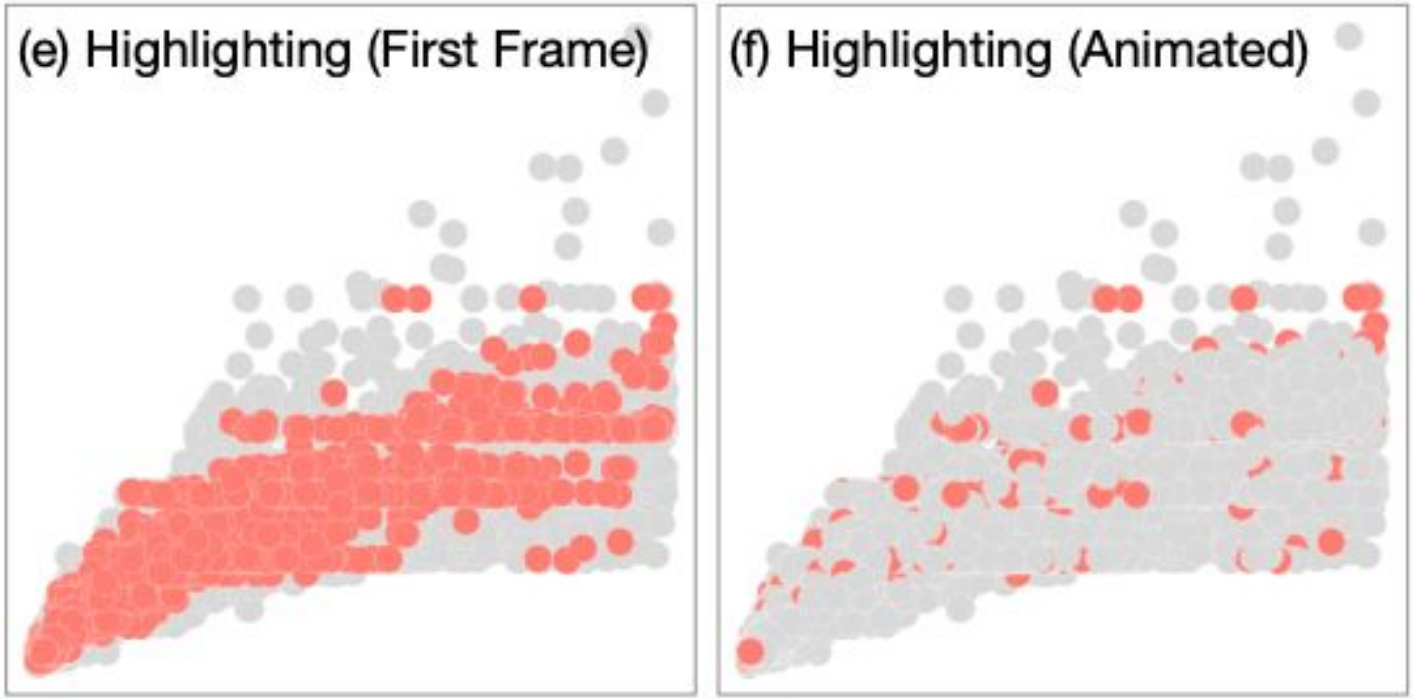


# Interaction + Animation

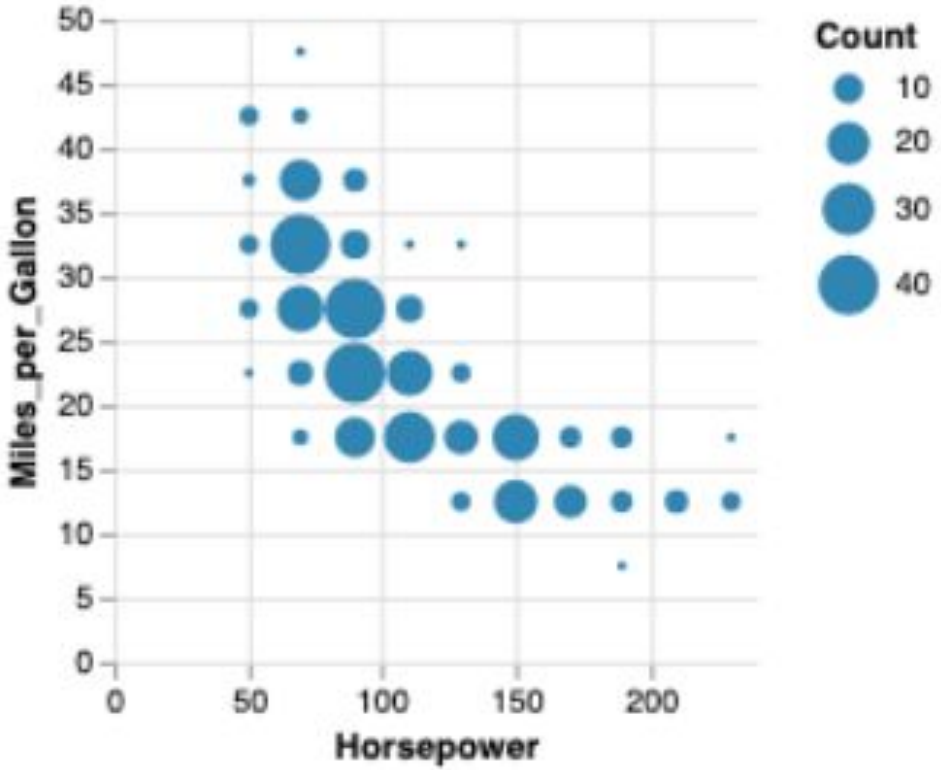
https://www.gapminder.org/

http://vistorian.net

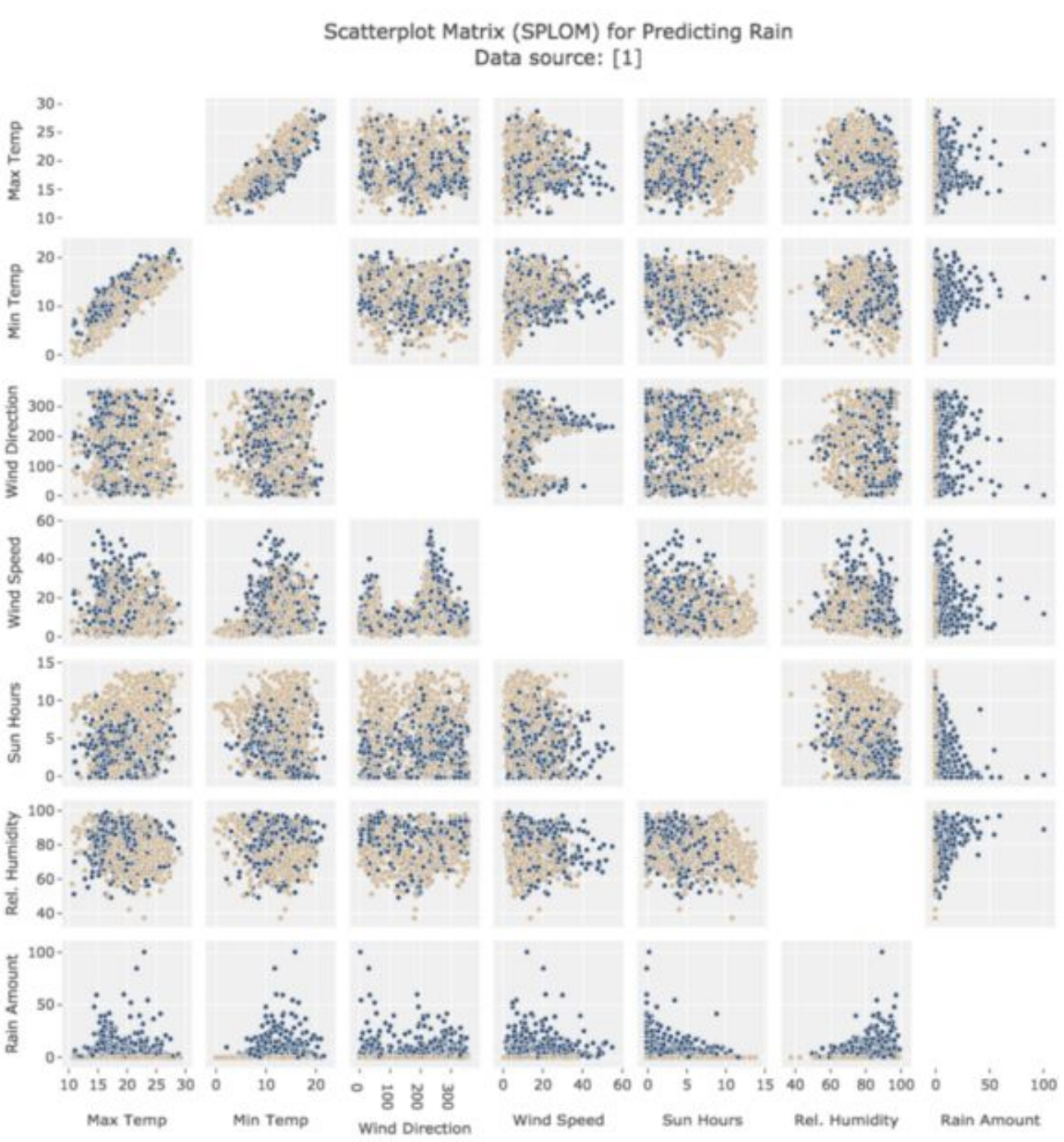
# Interaction + animation




http://vgl.cs.usfca.edu/animated-sploms/


Count
10
20
30
40
Miles_per_Gallon
50
45
40
35
30
25
20
15
10
5
0
0
50
100
150
200
Horsepower

# Multiple views



# Examples

# Infographic

## SMS

PARTICIPANTS

87

2% of all communication participants

TOP PARTICIPANTS

| | |
|---|---|
| Olga Bell | 40,662 |
| Joshua Gorchov | 458 |
| Ryan Case | 368 |
| Twitter | 348 |
| Michael L Orenstein | 292 |
| Matthew Grzywinski | 222 |

MOST SMS

February 17

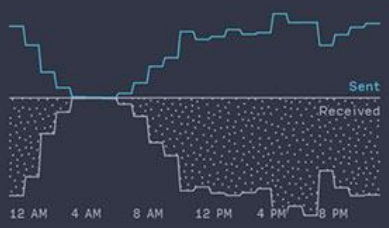
586 messages

DAYS W/O SMS

7

2% of year

HOURLY SMS




SENT SMS

18,131

Average 50/day

RECEIVED SMS

25,895

Average 71/day

TOTAL MONTHLY SMS ×1,000

SENDERS

84

97% of participants

RECIPIENTS

59

68% of participants

FAILED SMS

FIFTEEN

LONGEST GAP IN SMS

2 days, 18 hours

Beginning February 8 at 6:12 PM

## Email

PARTICIPANTS

3,744

66% of all communication participants

TOP PARTICIPANTS

| | |
|---|---|
| Authorize.net | 1,899 |
| Big Cartel | 1,185 |
| Nicholas Felton | 916 |
| Olga Bell | 735 |
| Trulia | 661 |
| Flickr | 481 |

MOST EMAIL

March 13

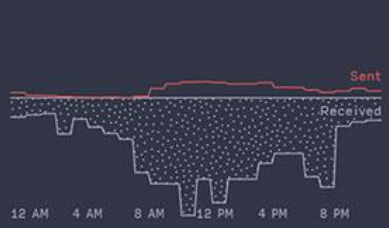
1,074 emails

LEAST EMAIL

August 4

23 emails

HOURLY EMAIL




SENT EMAIL

3,852

Average 11/day

RECEIVED EMAIL

27,917

Average 76/day

TOTAL MONTHLY EMAIL ×1,000

SENDERS

3,657

98% of participants

RECIPIENTS

950

25% of participants

EMAIL ADDRESS ANALYZED

NICHOLAS@FELTRON.COM

LONGEST GAP IN EMAIL

7 hours, 8 minutes

Beginning August 23 at 8:22 PM

## Facebook Messages

PARTICIPANTS

336

6% of all communication participants

TOP PARTICIPANTS

| | |
|---|---|
| Justin Shaffer | 289 |
| Cooper Liska-Smith | 279 |
| Edan Rosenberg | 225 |
| Roddy Lindsay | 212 |
| Joey Flynn | 147 |
| Ryan Case | 137 |

MOST MESSAGES

November 7

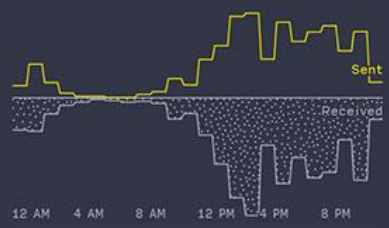
225 messages

DAYS W/O MESSAGE

92

25% of year

HOURLY FACEBOOK MESSAGES




SENT MESSAGES

2,057

Average 6/day

RECEIVED MESSAGES

2,454

Average 7/day

TOTAL MONTHLY FACEBOOK MESSAGES ×100

SENDERS

326

97% of participants

RECIPIENTS

190

57% of participants

PARTICIPANTS FROM "OTHER" INBOX

FIFTY-FOUR

LONGEST GAP IN MESSAGES

7 days, 20 hours

Beginning May 15 at 11:33 PM

## Mail

PARTICIPANTS

1,397

25% of all communication participants

TOP PARTICIPANTS

| | |
|---|---|
| New Yorker | 48 |
| Merrill Edge | 25 |
| Merrill Lynch | 17 |
| Facebook | 14 |
| Wired | 13 |
| Con Edison | 12 |

MOST MAIL

June

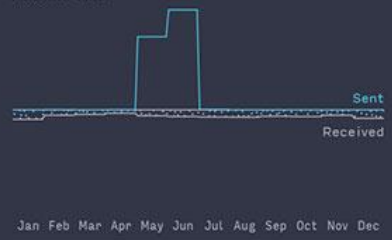
732 pieces of mail

LEAST MAIL

April

24 pieces of mail

MONTHLY MAIL




SENT MAIL

1,192

Average 3/day

RECEIVED MAIL

527

Average 1/day

DAILY MAIL

SENDERS

218

16% of participants

RECIPIENTS

1,188

85% of participants

ADDRESSES USED TO RECEIVE MAIL

SIX

LONGEST GAP IN MAIL

8 days

Beginning April 13

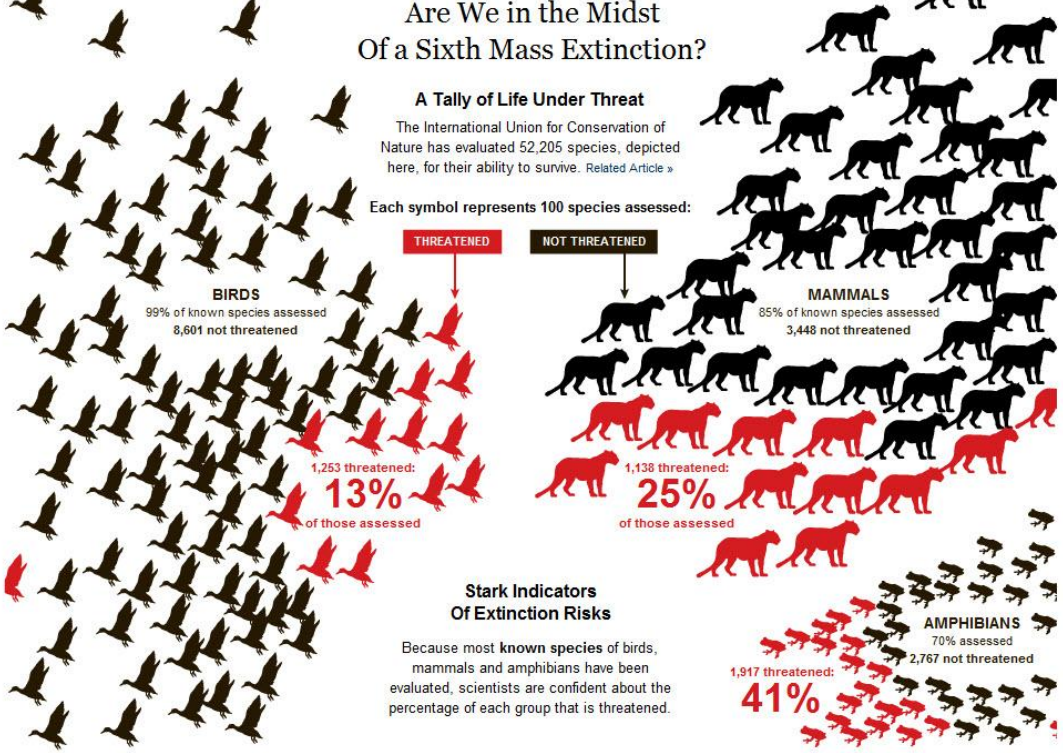
# Are We in the Midst Of a Sixth Mass Extinction?

## A Tally of Life Under Threat

The International Union for Conservation of Nature has evaluated 52,205 species, depicted here, for their ability to survive. Related Article »

**Each symbol represents 100 species assessed:**

THREATENED | NOT THREATENED



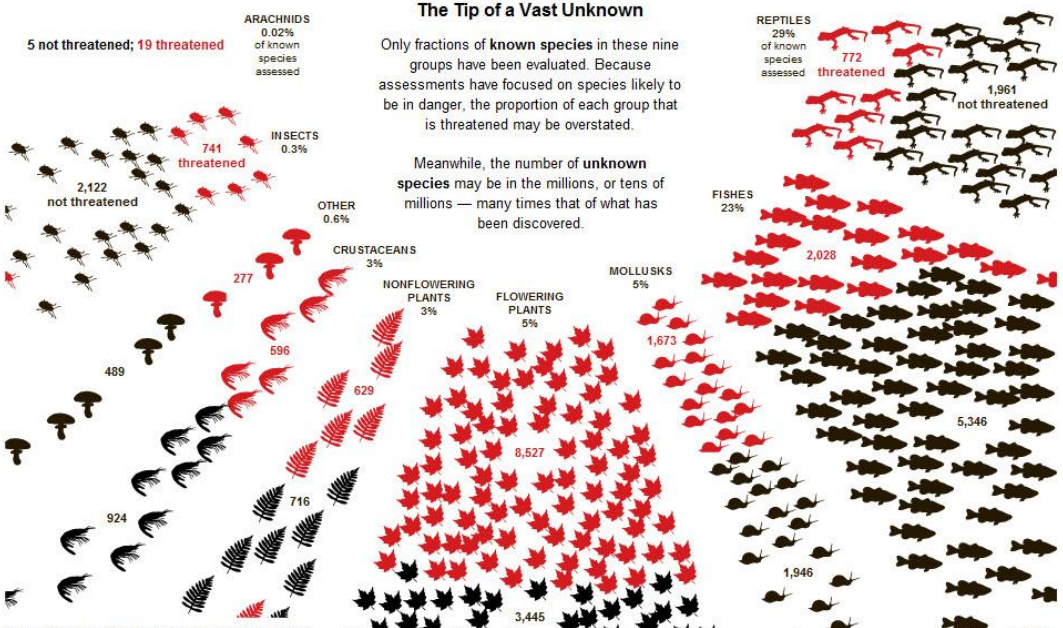

## Stark Indicators Of Extinction Risks

Because most **known species** of birds, mammals and amphibians have been evaluated, scientists are confident about the percentage of each group that is threatened.

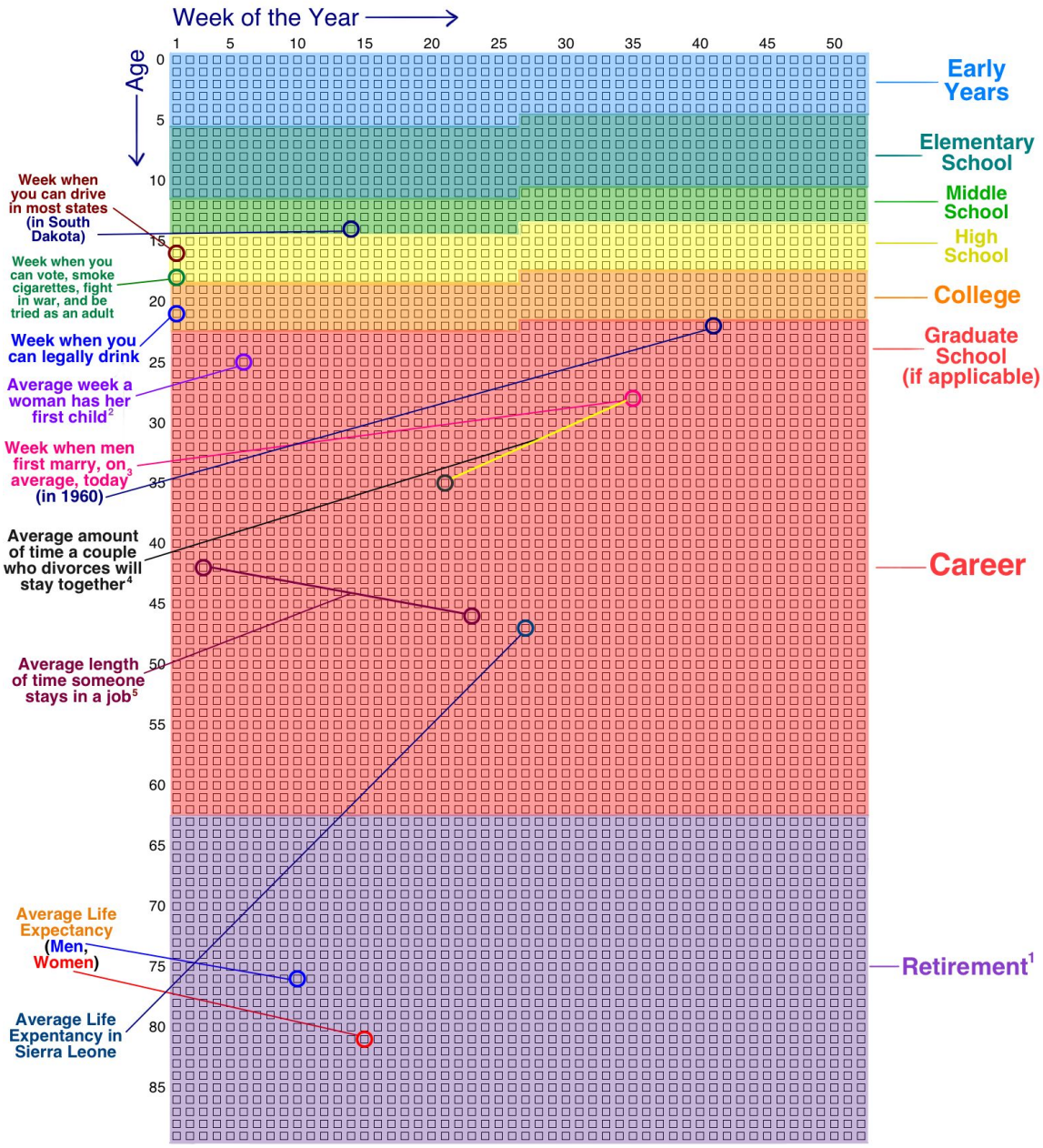
## Other Threatened Life: The Tip of a Vast Unknown

Only fractions of **known species** in these nine groups have been evaluated. Because assessments have focused on species likely to be in danger, the proportion of each group that is threatened may be overstated.

Meanwhile, the number of **unknown species** may be in the millions, or tens of millions — many times that of what has been discovered.



# The Life of a Typical American




waitbutwhy.com

# EBOLA
## BY THE NUMBERS

The Ebola outbreak in West Africa continues to rage, with the number of people infected roughly doubling every 3–4 weeks. More than 8,000 people are thought to have contracted the disease, and almost half of those have died, according to the World Health Organization. Although these estimates are already staggering, the situation on the ground means that not all cases and deaths are being reported, so the true extent is likely to be much greater.

Outside of Africa, a health-care worker in Texas has become infected while treating a patient who was hospitalized in Dallas after travelling from Liberia and who has now died. And a nurse in Madrid has contracted the virus after caring for a missionary who had become infected while caring for patients in West Africa. Health-care workers remain one of the groups at highest risk of exposure: by 8 October, 416 had become infected and 233 had died.

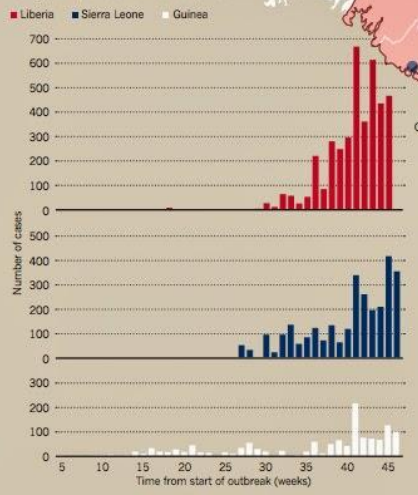
The spread beyond the epicentre of Guinea, Liberia and Sierra Leone remains limited. Apart from the people in Dallas and Spain, only two other exported cases are known: one in Nigeria and one in Senegal. A man who travelled to Lagos from Liberia sparked a further 19 cases in Nigeria, but that outbreak was curtailed by the swift actions of the authorities in tracing and monitoring those who had contact with the infected man. Similar public-health measures stopped further cases in Senegal after an infected man travelled from Guinea to Dakar.

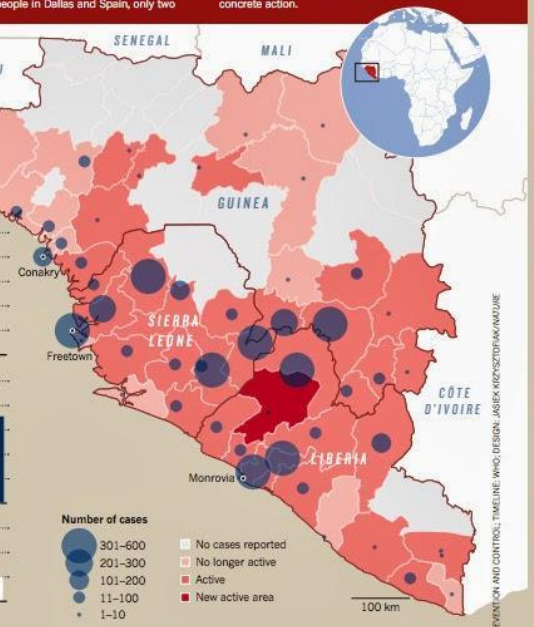
Within the epicentre, authorities have made some progress in slowing transmission — but the disease is resurgent in places where it had seemed under control, such as in Conakry, Guinea's capital.

Meanwhile, the estimated cost of fighting the disease is spiralling upward. UN secretary-general Ban Ki-moon warned on 9 October that "at least a 20-fold surge in assistance" was needed to confront the outbreak. But "things will get worse before they get better", he warned. Just how much worse will depend on the international community — which has been widely criticized for its belated response, and its slow translation of pledges into concrete action.

### A RISING TOLL

The number of Ebola cases continues to rise because control measures in the outbreak area are insufficient. But it could drop quickly if the international community and affected countries manage to implement an effective response.

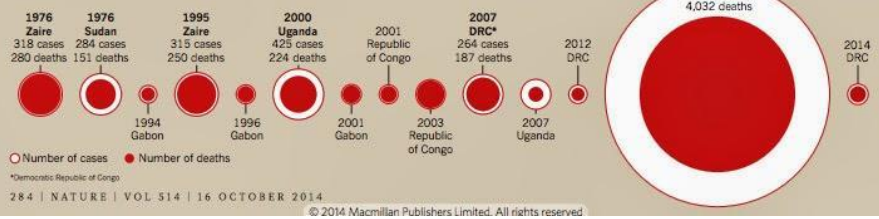



MAP: WHO; CASE PLOTS: EUROPEAN CENTRE FOR DISEASE PREVENTION AND CONTROL; TIMELINE: WHO; DESIGN: JASJEK KRZYSZTOFIAK/NATURE

### UNPRECEDENTED SIZE

The current outbreak dwarfs the largest historical outbreaks in Africa, which were rural and relatively easy to control. Ebola has now spread to dense urban areas, where control is harder to achieve.






© 2014 Macmillan Publishers Limited. All rights reserved

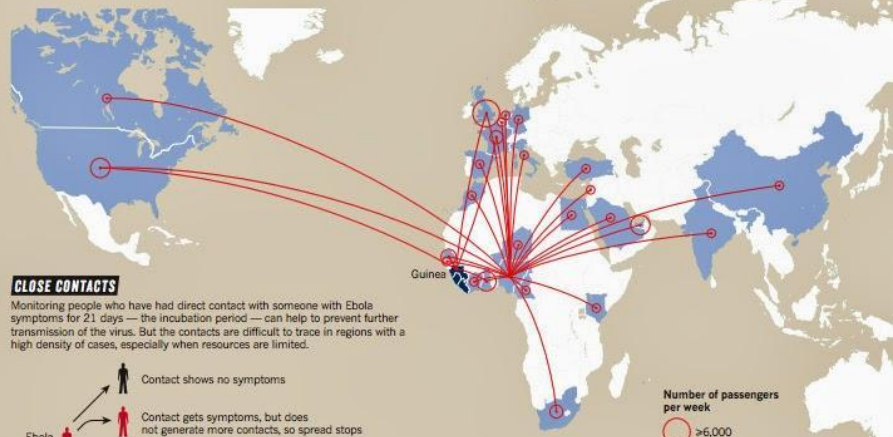
### TRANSMITTING DISEASE

Ebola is spread by contact with an infected person's bodily fluids, but is less contagious than many common diseases, such as mumps and measles. In the current outbreak, each person with Ebola will infect 1–2 other people.

TRANSMISSION: EBOLA RESPONSE TEAM *N. ENGL. J. MED.* HTTP://DOI.ORG/NVW (2014); OTHERS: GO.NATURE.COM/ACTQ4R

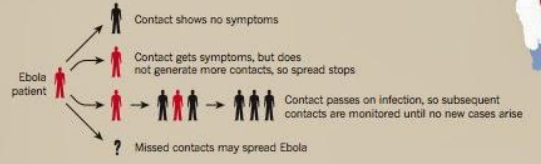
### GLOBAL REACH

Modelling of historical air-passenger volumes and flight networks can point to international destinations where a traveller with Ebola might end up.




### CLOSE CONTACTS

Monitoring people who have had direct contact with someone with Ebola symptoms for 21 days — the incubation period — can help to prevent further transmission of the virus. But the contacts are difficult to trace in regions with a high density of cases, especially when resources are limited.

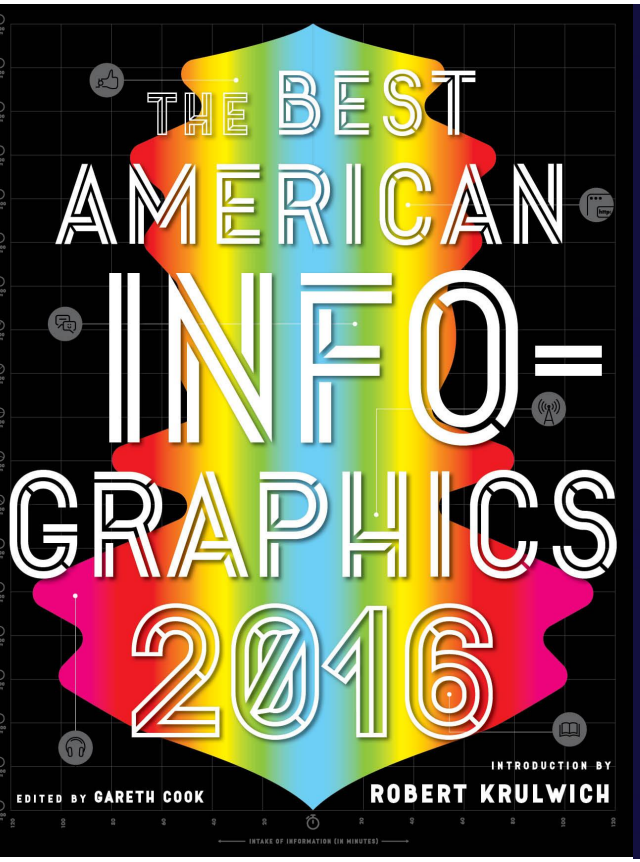


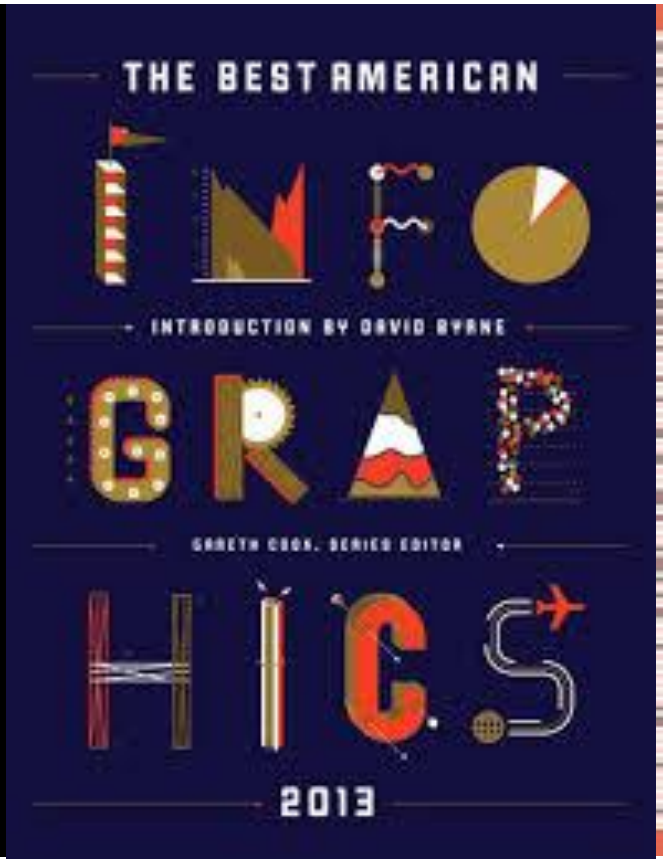

### FINANCIAL AID

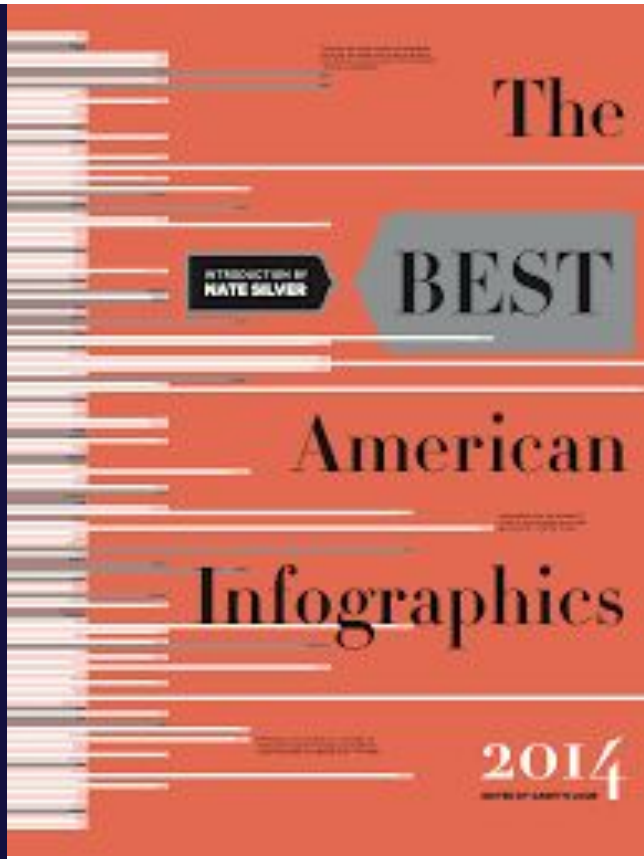
If outbreaks are allowed to grow, they become more difficult and costly to control. In April, the World Health Organization estimated that it would cost US$4.8 million to contain the Ebola outbreak, but by September that figure had ballooned to almost $1 billion. Experts say that the total cost of ending this outbreak is likely to be higher still.

FLIGHT MAP: GO.NATURE.COM/WDVP2U; FINANCIAL AID: WHO AND EBOLACRISISRESPONSE.ORG

# Best American Infographics



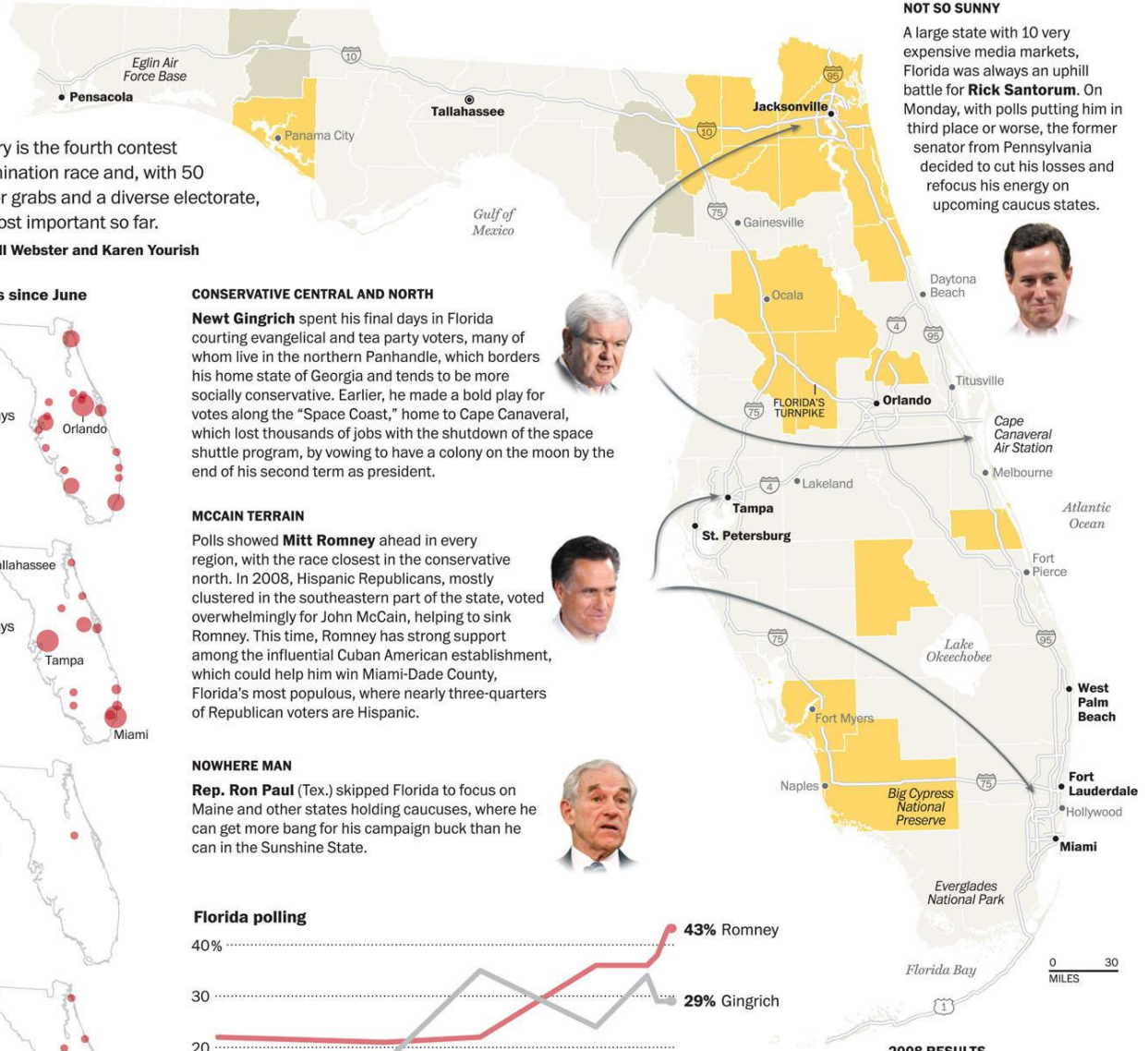



# Political Geography: Florida

Florida's primary is the fourth contest of the GOP nomination race and, with 50 delegates up for grabs and a diverse electorate, arguably the most important so far.

**By Ted Mellnik, Bill Webster and Karen Yourish**




## Campaign events since June

### CONSERVATIVE CENTRAL AND NORTH



**Newt Gingrich** spent his final days in Florida courting evangelical and tea party voters, many of whom live in the northern Panhandle, which borders his home state of Georgia and tends to be more socially conservative. Earlier, he made a bold play for votes along the "Space Coast," home to Cape Canaveral, which lost thousands of jobs with the shutdown of the space shuttle program, by vowing to have a colony on the moon by the end of his second term as president.

### MCCAIN TERRAIN



Polls showed **Mitt Romney** ahead in every region, with the race closest in the conservative north. In 2008, Hispanic Republicans, mostly clustered in the southeastern part of the state, voted overwhelmingly for John McCain, helping to sink Romney. This time, Romney has strong support among the influential Cuban American establishment, which could help him win Miami-Dade County, Florida's most populous, where nearly three-quarters of Republican voters are Hispanic.

### NOWHERE MAN



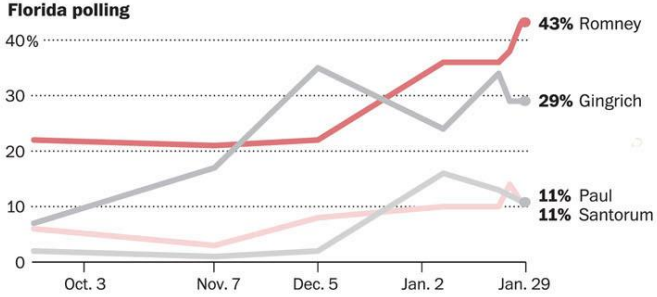
**Rep. Ron Paul** (Tex.) skipped Florida to focus on Maine and other states holding caucuses, where he can get more bang for his campaign buck than he can in the Sunshine State.

### NOT SO SUNNY

A large state with 10 very expensive media markets, Florida was always an uphill battle for **Rick Santorum**. On Monday, with polls putting him in third place or worse, the former senator from Pennsylvania decided to cut his losses and refocus his energy on upcoming caucus states.



## Florida polling




### 2008 RESULTS

| Romney | McCain | Huckabee |
|---|---|---|
| 31% | 36% | 13% |

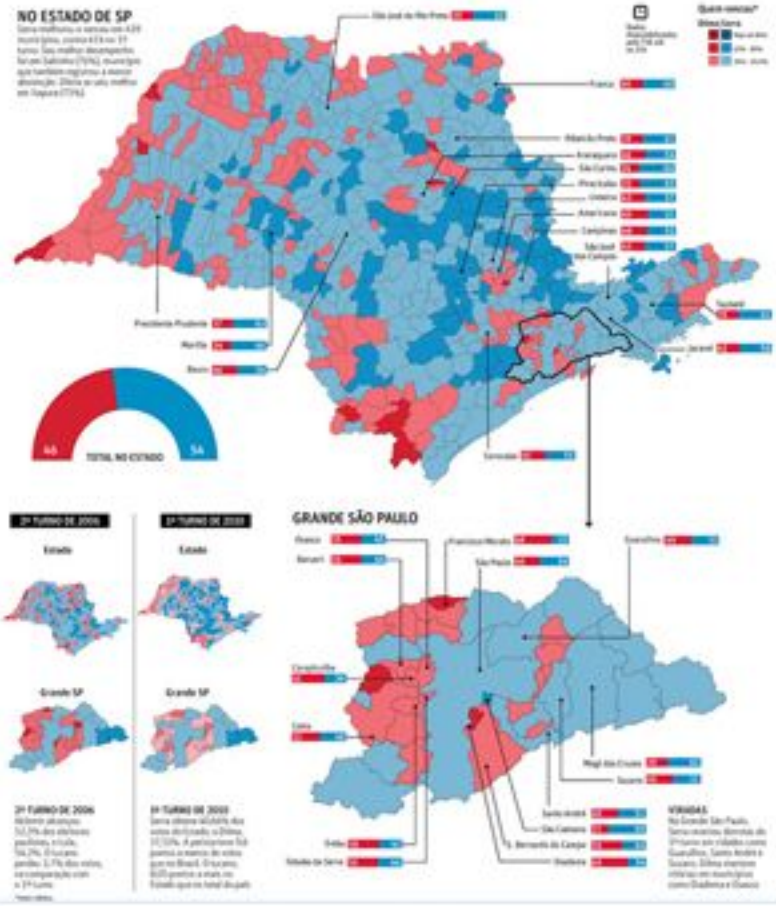
Romney lost to John McCain but had a fairly strong showing.

Sources for polls: Quinnipiac University polls of likely primary voters and registered Republicans from December and earlier

THE WASHINGTON POST

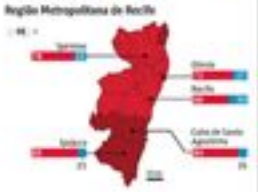
# O voto no **detalhe**

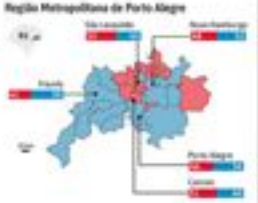
Vitória de Serra em **São Paulo** pouco ajuda a minimizar vantagem de Dilma pelo país; petista repete **maioria no Rio**, onde venceu em toda a região metropolitana, e ganha de novo em **Minas Gerais**

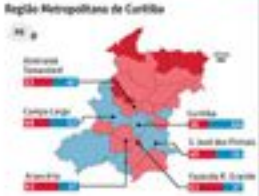


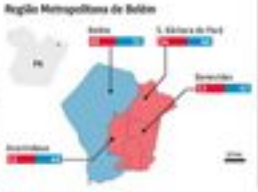

## Serra consegue viradas em cidades importantes de SP

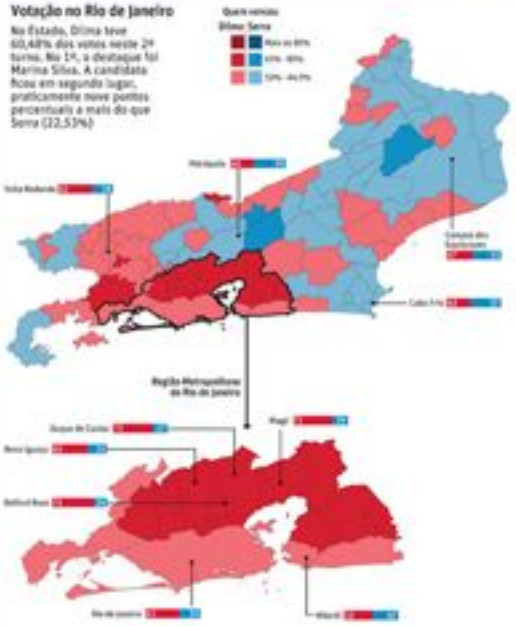
Tucano ganha em Guarulhos, Campinas e Santo André; ele teve mais votos que a petista em 440 dos 645 municípios do Estado

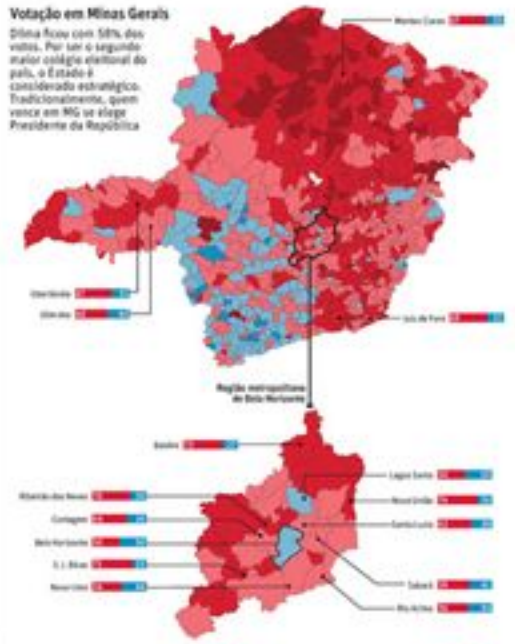




# Infographic in a nutshell

- Image-driven / image first
- Detailed
- Labeling and annotation
- Story though picture
- Exploratory: reader-driven

# Report



**PRESIDENTE 40**
**ELEIÇÕES 2010**

**CÂMARA**
Hoje **cada Estado** elege no mínimo **8** e no máximo **70** deputados. Em **1890**, no início da República, cada Estado podia eleger no mínimo **2** e no máximo **37** deputados. O total de deputados era de **205**, contra os **atuais 513**.

"O partido Verde em Minas é um partido que já está no governo. Tenho certeza que o PV em Minas terá uma posição favorável [a Serra]. Vamos conversar com eles, sempre com muita calma, para mostrar ao PV que a proposta do PSDB é a melhor para o Brasil" ANTONIO ANASTASIA (PSDB), governador reeleito de Minas

# Dilma tem mais votos no interior, e Lula, no litoral

## Em relação a 2006, ex-ministra supera o presidente em 63% dos municípios

**José Serra teve votação inferior à de Alckmin em 85% das cidades quando este perdeu a eleição para Lula**

UIRÁ MACHADO
DE SÃO PAULO

Dilma Rousseff (PT) teve neste ano um percentual de votos mais baixo que o de Lula (PT) em 2006, mas a ex-ministra superou a votação do presidente em cerca de 63% dos municípios brasileiros.

A votação da ex-ministra no primeiro turno foi mais diluída e mais interiorizada que a de Lula. O presidente, no ano em que foi reeleito, teve mais apoio em praticamente todo o litoral.

Algumas capitais nordestinas mostram como o voto em Lula foi mais concentrado. Em Fortaleza (CE), ele teve 68% dos votos, e Dilma, 51%. Em Natal (RN), o placar foi de 47% a 36%, em Salvador (BA), 68% a 54%, e em João Pessoa (PB), 59% a 42%.

Dilma avançou em praticamente todo o Centro-Oeste. Com ela, o PT também reconquistou espaço no Sul do país e venceu em um Estado da região (Rio Grande do Sul), algo que não acontecera quatro anos atrás.

O desempenho do PSDB nos municípios também tem diferenças significativas, embora os resultados por Estados não seja tão diferente.

Em 2006, Geraldo Alckmin (PSDB) havia sido melhor que Lula em dez Estados. Neste ano, José Serra (PSDB) superou Dilma em oito – nenhum em que o partido já não tivesse vantagem.

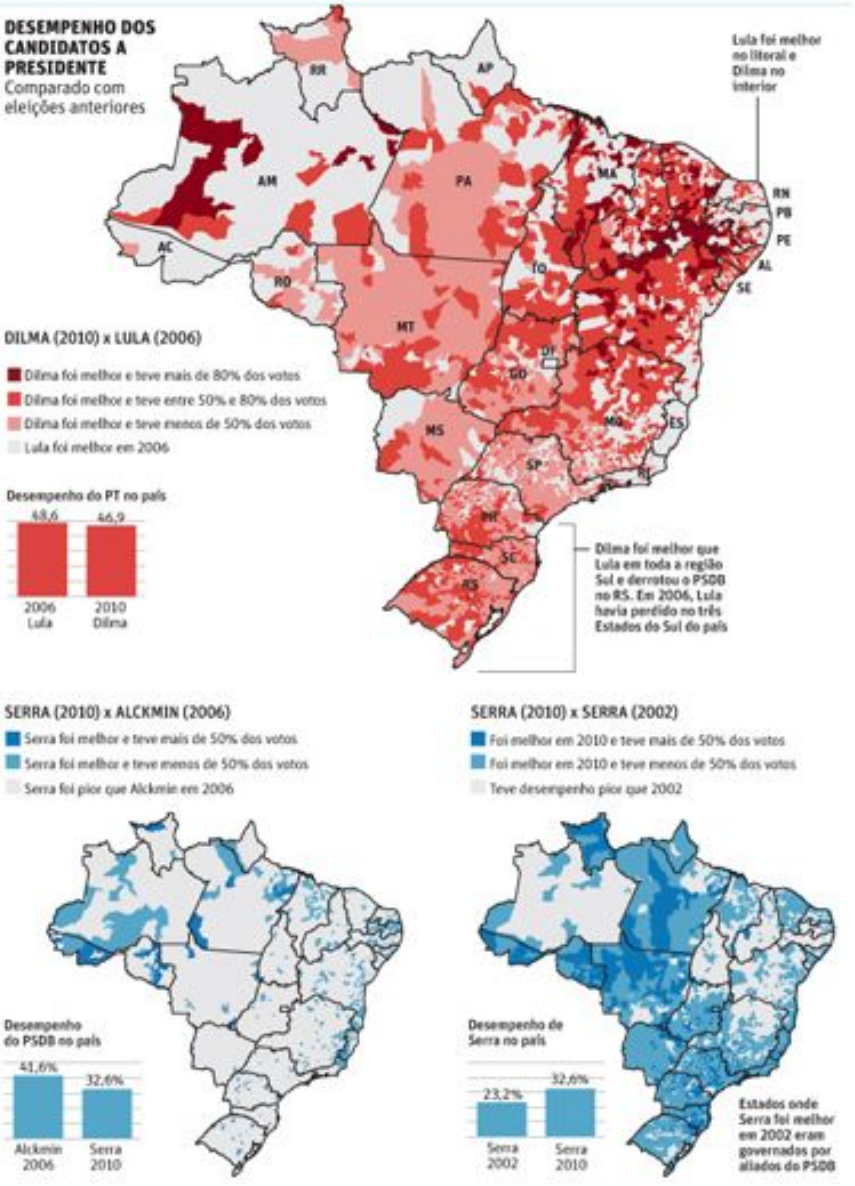
Quando considerados os municípios, porém, vê-se que Serra teve votação inferior à de Alckmin em cerca de 85% das cidades.

Em 2006, Alckmin chegou a 41,6% dos votos, patamar acima do Serra, que ficou com 32,6%.

Serra neste ano, por outro lado, melhorou significativamente sua votação em relação a 2002, quando foi derrotado por Lula (o petista teve 46,4% dos votos).

Naquele ano, o tucano ficou com 23,2% dos votos. O crescimento de Serra em comparação com 2002 não ocorreu apenas em Estados que eram então governados por partidos da base aliada do ex-presidente Fernando Henrique Cardoso (PSDB).

Rio Grande do Norte, Paraíba, Pernambuco e Piauí estavam nas mãos do PMDB (à época aliado do PSDB). Bahia, Tocantins, Maranhão e Amazonas tinham governadores do PFL (hoje, DEM), partido do vice de FHC.



# Report in a nutshell

- Text driven
- Story and info in the text
- Beginning, middle, end
- Figures referenced (a,b,c…)
- Annotations and labels in figure
-

GUYS LET'S JUST HAVE A CHAT ABOUT

# MASS EXTINCTIONS.

Cruxgaster fabarasae

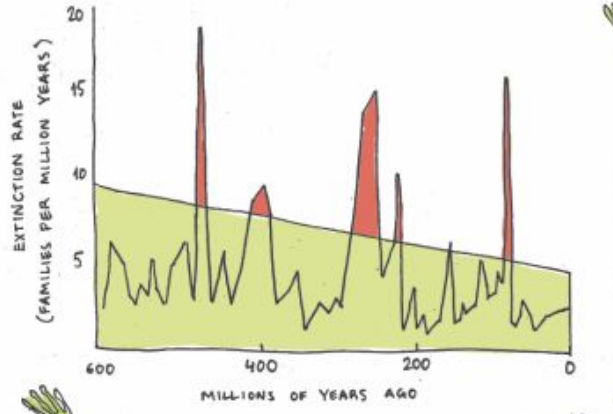
THERE'S GONNA BE A LOT OF GRAPHS AND TEXT AND SHIT SO JUST BE PREPARED TO FEEL LIKE THIS.

SERIOUSLY IF YOU DON'T NEED THIS IN YOUR LIFE YOU CAN SKIP TO PAGE 3.

LET'S START WITH A HISTORY LESSON.

HERE COMES THE FIRST GRAPH.

Source: University of California Museum of Paleontology's Understanding Evolution (http://evolution.berkeley.edu)




IT'S A BIT HARD TO READ BUT I'M GONNA TRY AND BREAK IT DOWN TO YOU. IT'S ABOUT MASS EXTINCTIONS IN THE PAST 600 MILLION YEARS.

THE Y-AXIS SHOWS US THE AMOUNT OF DEATH BASICALLY.

THE X-AXIS IS THE TIMELINE.

WE'RE TALKING LONG TIME PERIODS. 600 000 000 - 0 YEARS AGO.

THE GREEN AREA IS BACKGROUND EXTINCTION. IT'S THE RATE AT WHICH SPECIES DIE OFF NORMALLY, LIKE IT'S NO BIG DEAL.

THIS IS FINE

DON'T WORRY

THE BIG DEALS ARE THE SPIKES. I'VE MADE THEM RED SO YOU CAN FIND THEM MORE EASILY.

THOSE ARE THE MASS EXTINCTION EVENTS.

THAT'S WHEN MORE THAN 75% OF ALL SPECIES ON EARTH HAVE DIED OFF.

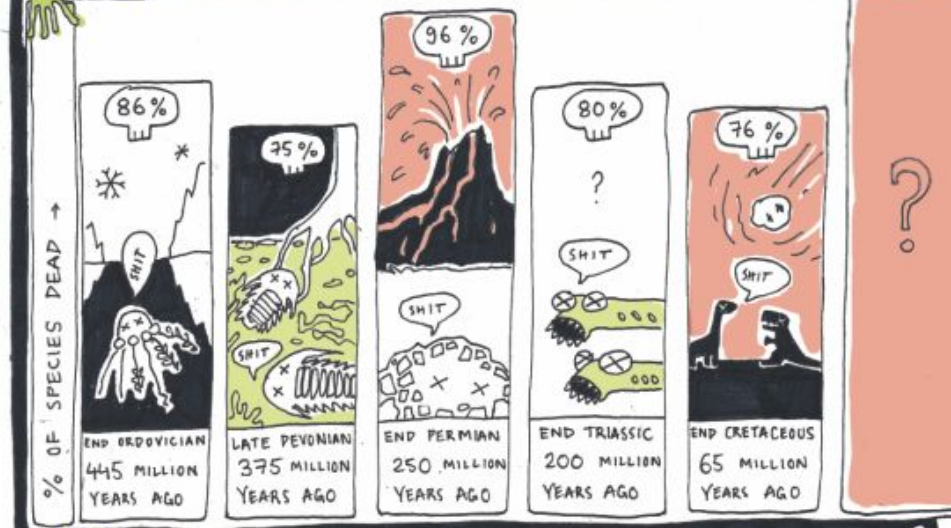
SO THAT'S HAPPENED 5 TIMES HERE.

LOOK HERE'S MY ATTEMPT TO VISUALISE THE BIG EXTINCTION EVENTS. YES I MADE IT MYSELF.

AND I'M A FROG.

THERE ARE NUMBERS HERE THAT AREN'T FROM THE PREVIOUS GRAPH. DON'T FREAK OUT. I USED A DIFFERENT SOURCE: Viviane Richter: "The Big Five Mass Extinctions" https://cosmosmagazine.com/palaeontology/big-five-extinctions

ESSENTIALLY IT'S THE SAME THING THOUGH.




IT'S A DATA COMIC INSIDE A DATA COMIC! HAHAHA

I'M VERY EXCITED ABOUT DATA COMICS

ANYWAY

IMAGINE IF THIS PORTION OF LIFE ON EARTH JUST DIED OFF.

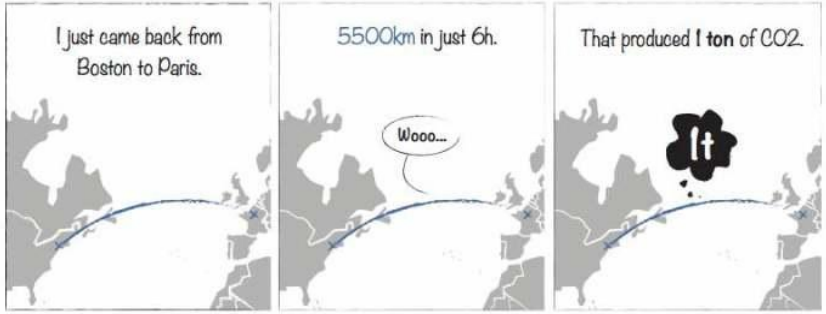
THAT'S BAD, RIGHT?!!?

WELL, A LOT OF SCIENTISTS SAY THAT WE'RE NOW ENTERING THE 6TH MASS EXTINCTION.

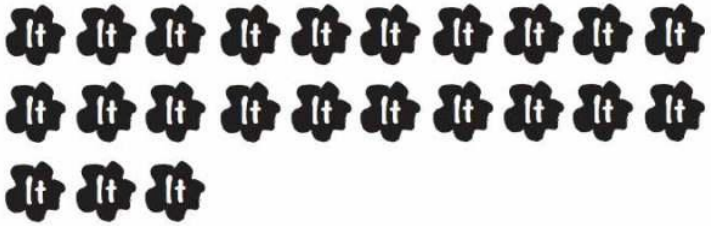
SEE THAT OMINOUS PANEL THERE? THAT'S WHERE WE ARE.

AND GUESS WHAT'S CAUSING IT THIS TIME? (TURN THE PAGE TO FIND OUT)

# CO Footprint

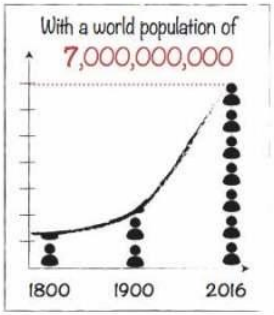


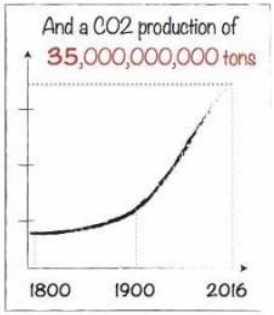
Which produced...

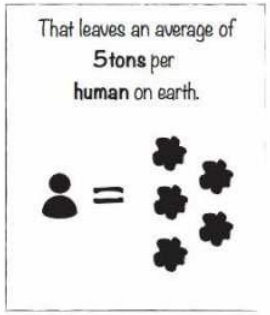


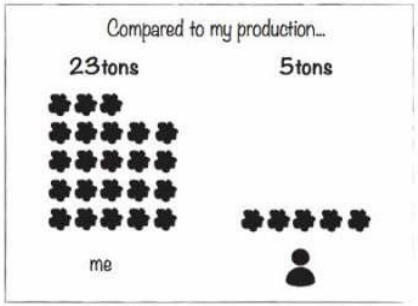

**23** tons of CO2.

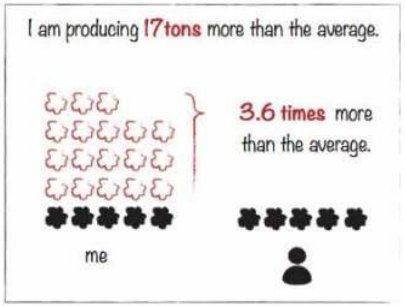
# **23** tons of CO2.

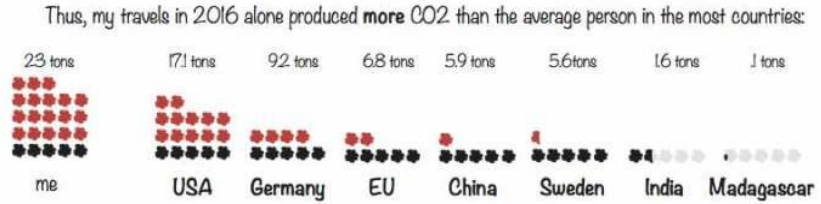


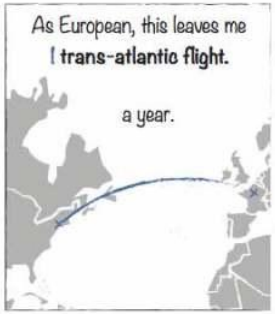



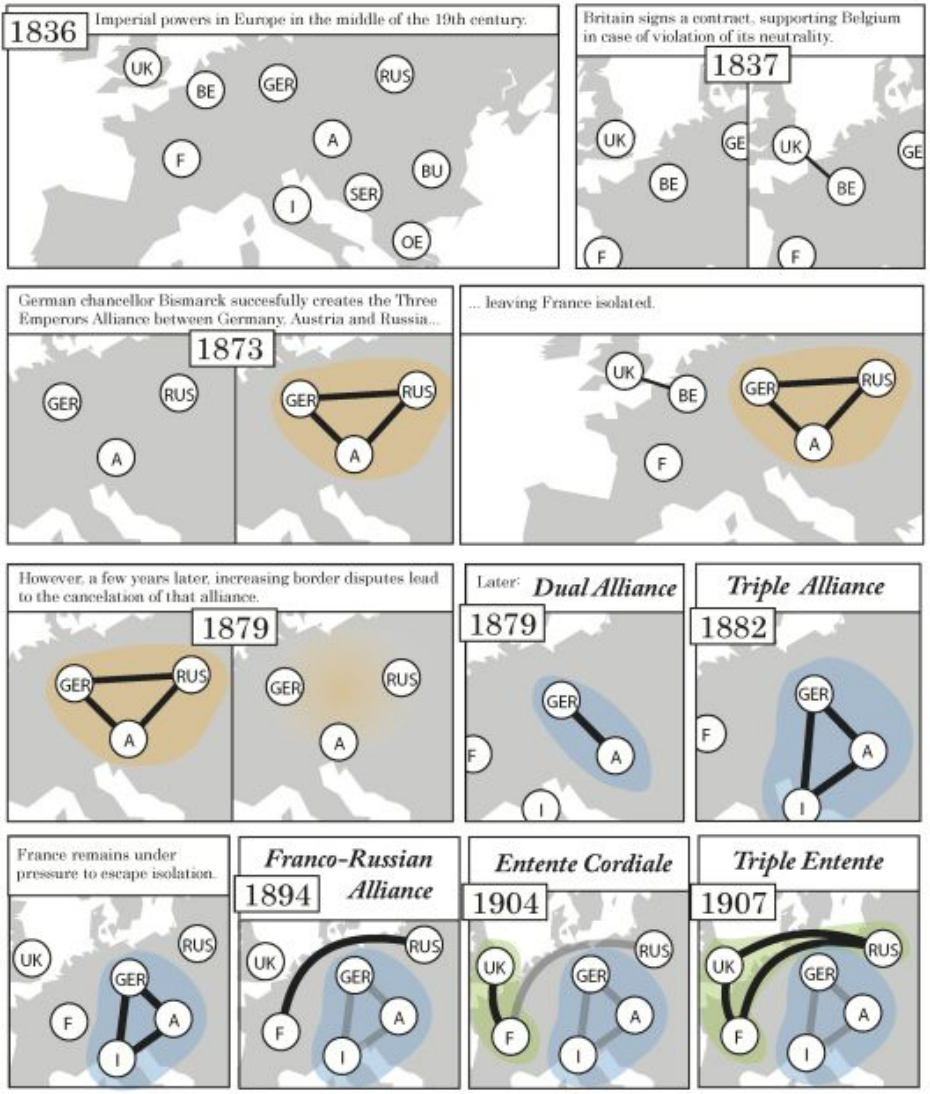



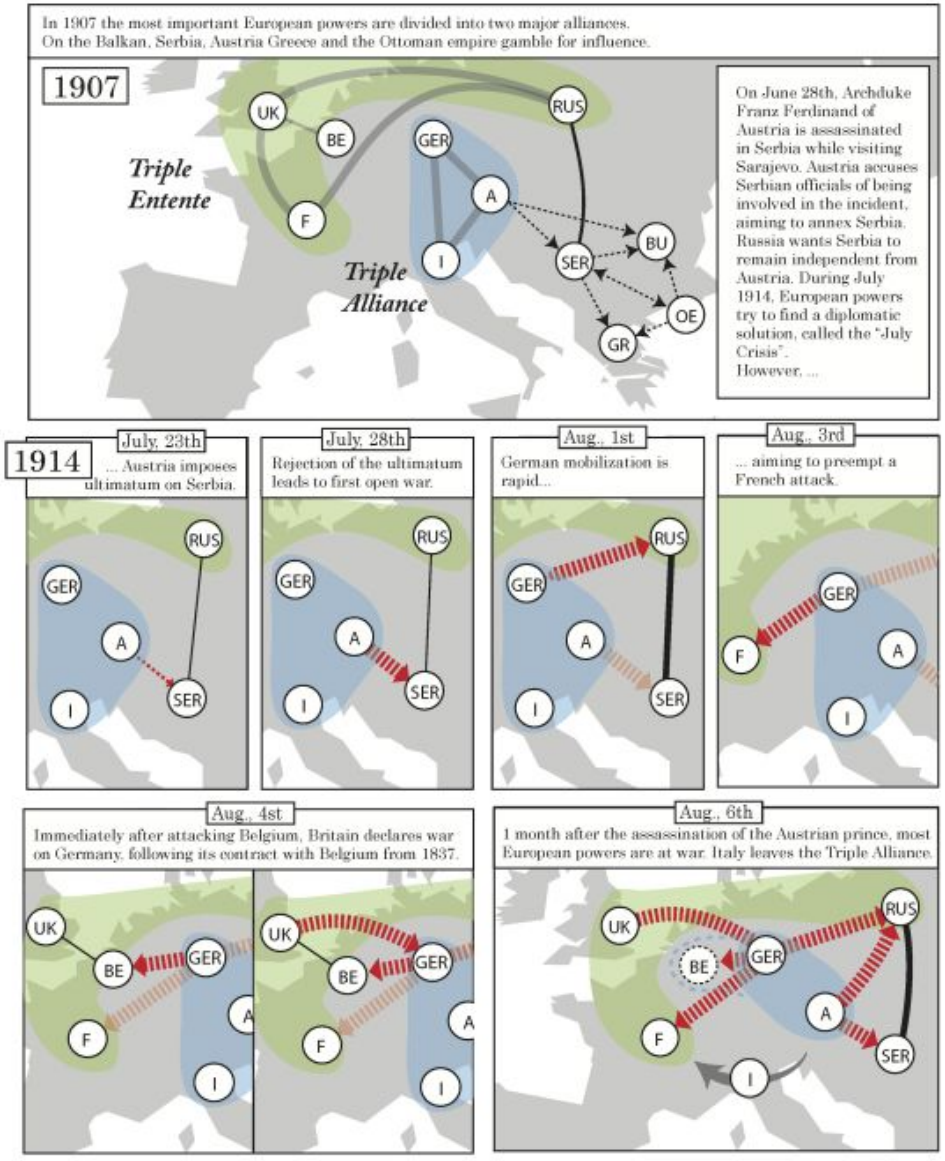





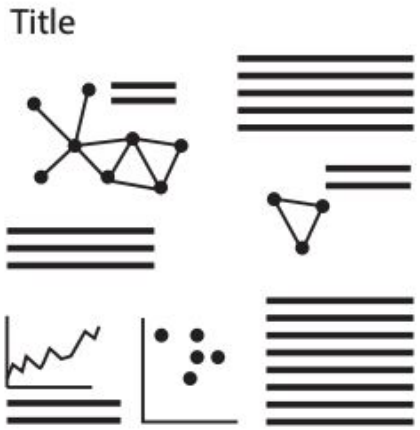
# European Alliances before World War I (1836-1914)



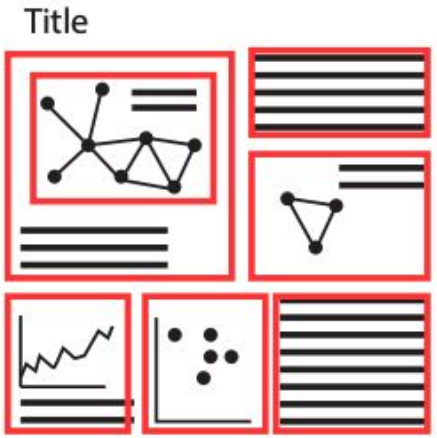



# Data comic in a nutshell

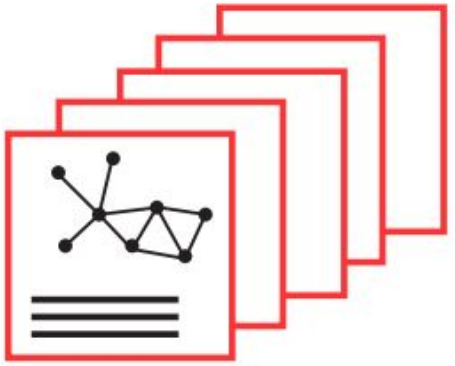
- Panel-driven
- Text and pictures are complementary
- One-panel, one–message
- Combines report + infographic




Infographic

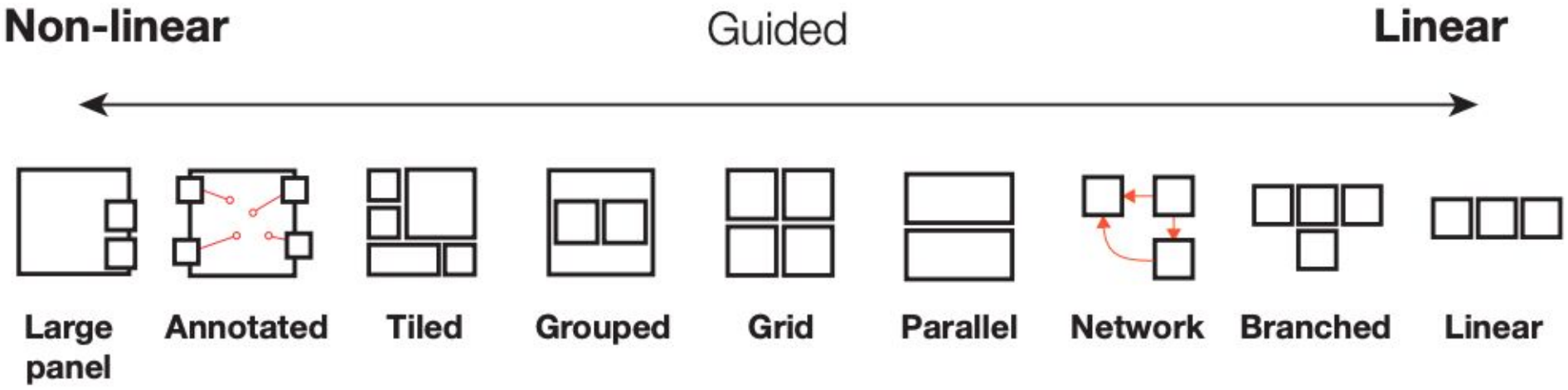



Panels on an infographic



Panels as frames (scenes) in a video/slideshow.

# Data comic in a nutshell



# Messages

- Take-home message
- Information insights
- Literacy messages
  a. Visualization literacy
  b. Data literacy
  c. Numeracy
  d. ....

# Style

- Everyone can draw!
- Adobe Illustrator
- https://vectr.com
- Keynote
- Google draw
- Affinity (50% Student discount)

**Visualization guidelines**

*https://wp.inf.ed.ac.uk/thehumanfactor/core-textbook-readings-for-each-week/500-2/*