Data visualization

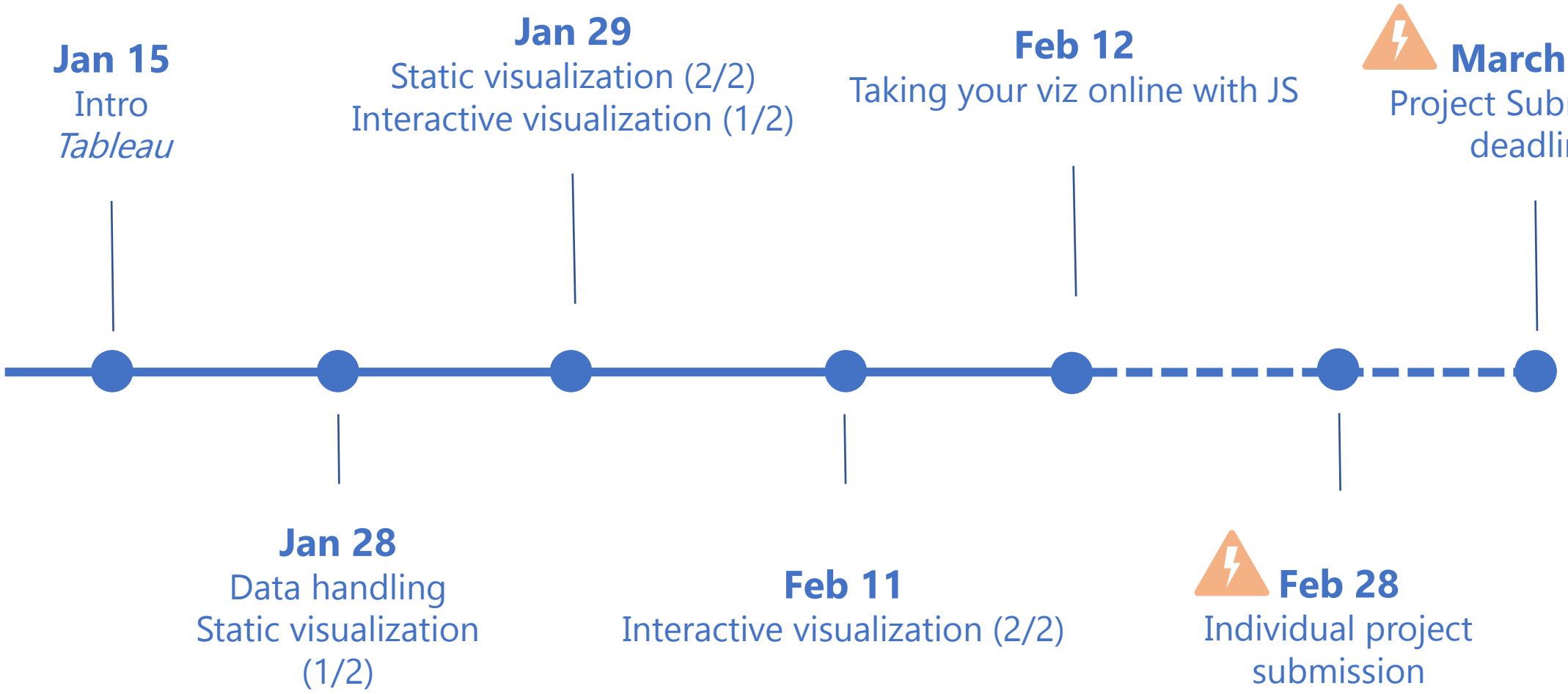
Msc Finance & Big data

Ano Kuhanathan

Kavala Capital & Neoma

Jan 15, 2021

1. Course overview
2. Setup & Some tips
3. What is Data viz and why does it matters?
4. Skills, Workflow & Tools
5. Some principles
6. Getting familiar with Tableau



Course Grading
50% Individual Project – due on Feb 28
50% Final Project (teams of 3-4) – due on March 12

1. Course overview
2. Setup & Some tips
3. What is Data viz and why does it matters?
4. Skills, Workflow & Tools
5. Some principles
6. Getting familiar with Tableau

- Install Anaconda on your computer
  - If you don't, you'll have to install manually python, a notebook and many basic libraries (NumPy, Pandas etc.)

- We will mostly work within a notebook
  - You can choose Jupyter (included in Anaconda) or Zeppellin

- Create & leverage a GitHub
  - Make all your projects open source
  - Include the link on your CV / LinkedIn

- Stack Overflow is your friend
- Don’t copy/paste too much code - we can tell

- Pay attention - viz is wrongly thought as the poor cousin of Data Science

1. Course overview
2. Setup & Some tips
3. What is Data viz and why does it matters?
4. Skills, Workflow & Tools
5. Some principles
6. Getting familiar with Tableau

Why do we need visualization ?

Only about 5% of our cognitive activity is conscious

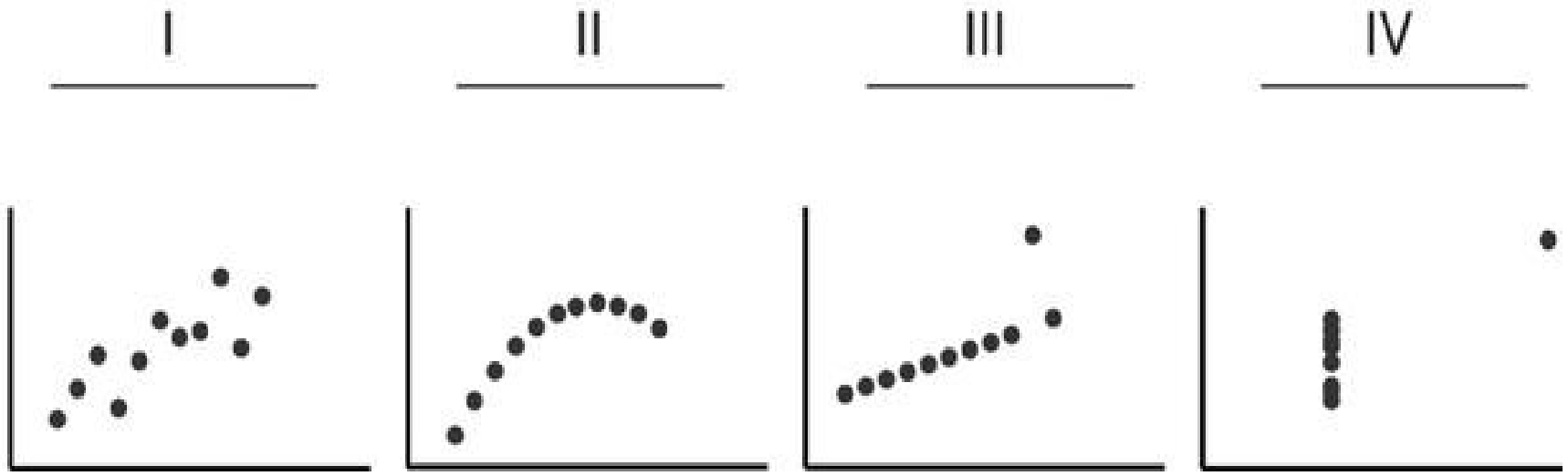
green	blue	yellow	blue
blue	red	yellow	red
yellow	yellow	green	red
yellow	green	blue	yellow
green	red	blue	green
blue	yellow	blue	red

Further reading :Malcolm Gladwell, "Blink"

Why do we need visualization ?

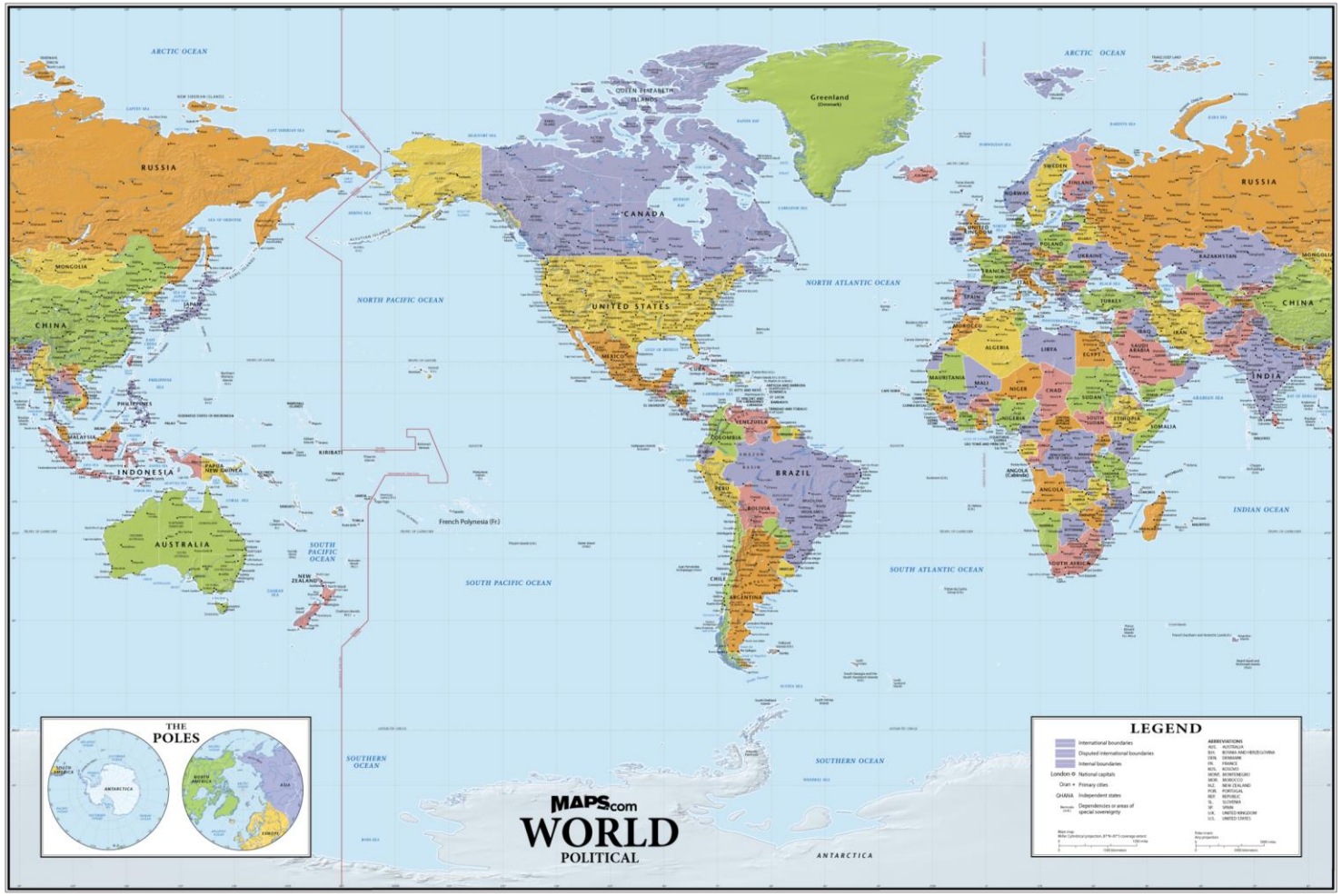
I		II		III		IV	
x	*y*	*x*	*y*	*x*	*y*	*x*	*y*
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.10	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.10	4	5.39	19	12.5
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89

Why do we need visualization ?



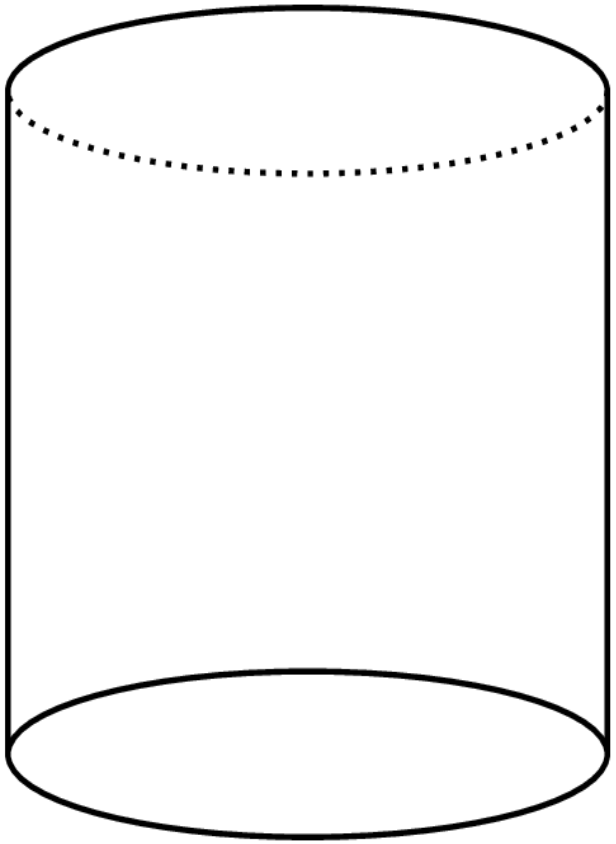
It’s never (or rarely) about the data…

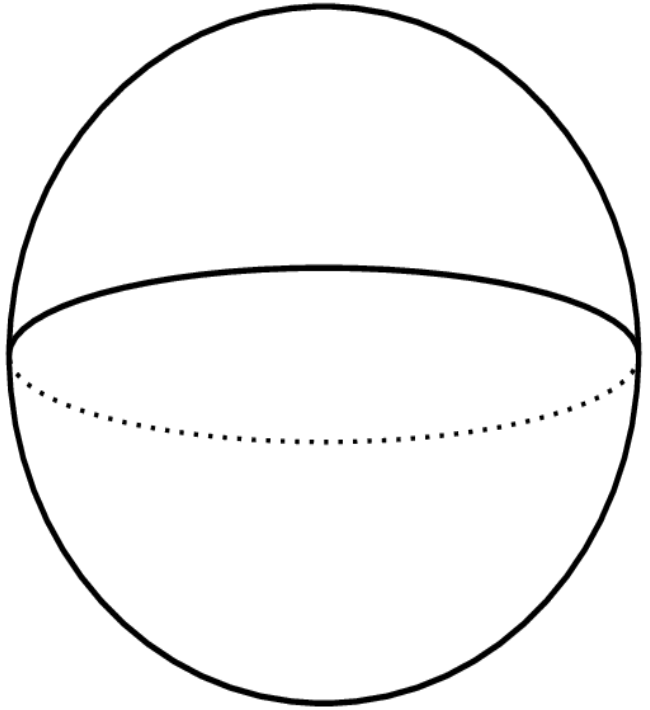
It’s never (or rarely) about the data…



It’s never (or rarely) about the data…

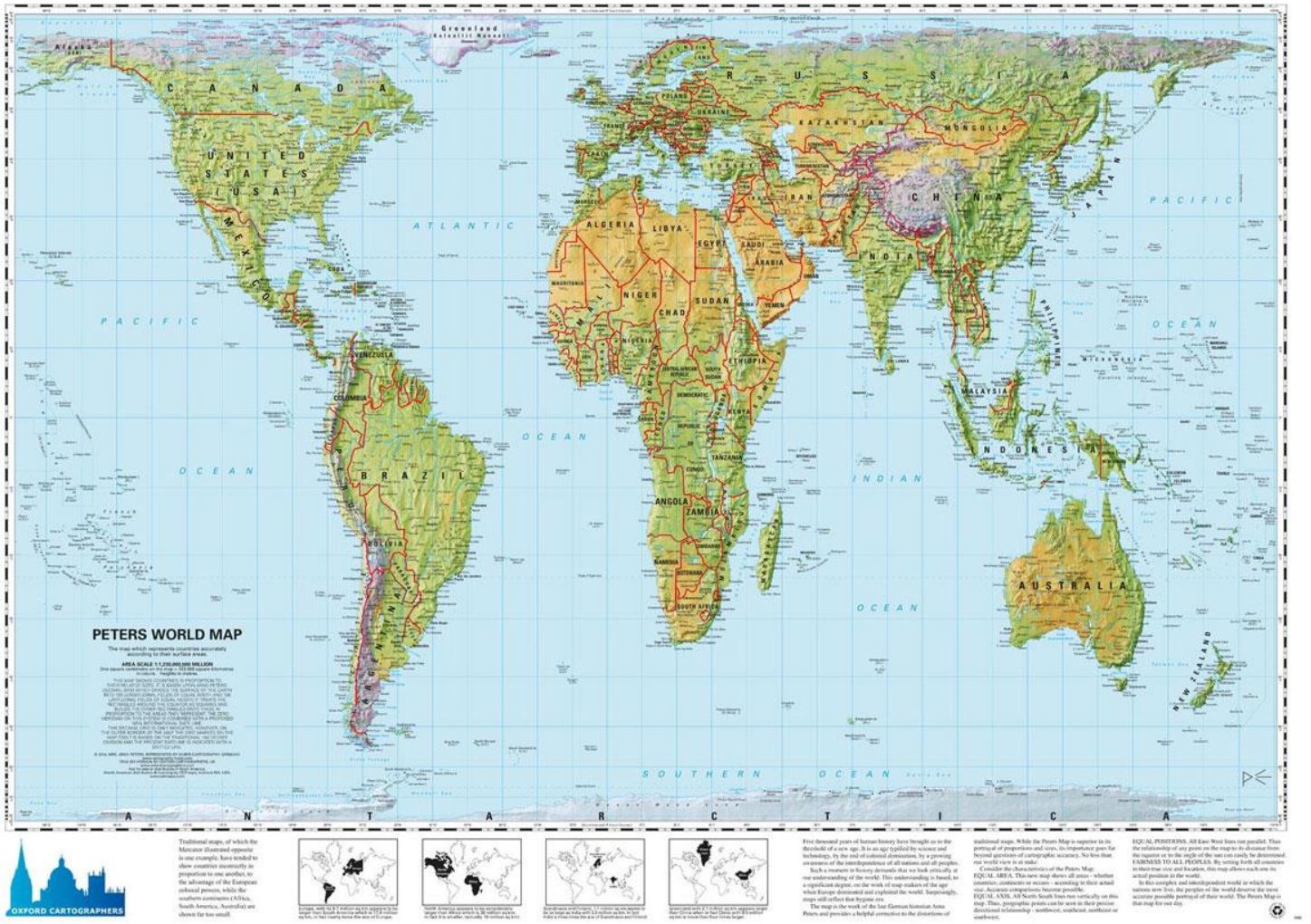
Easy to represent in 2D





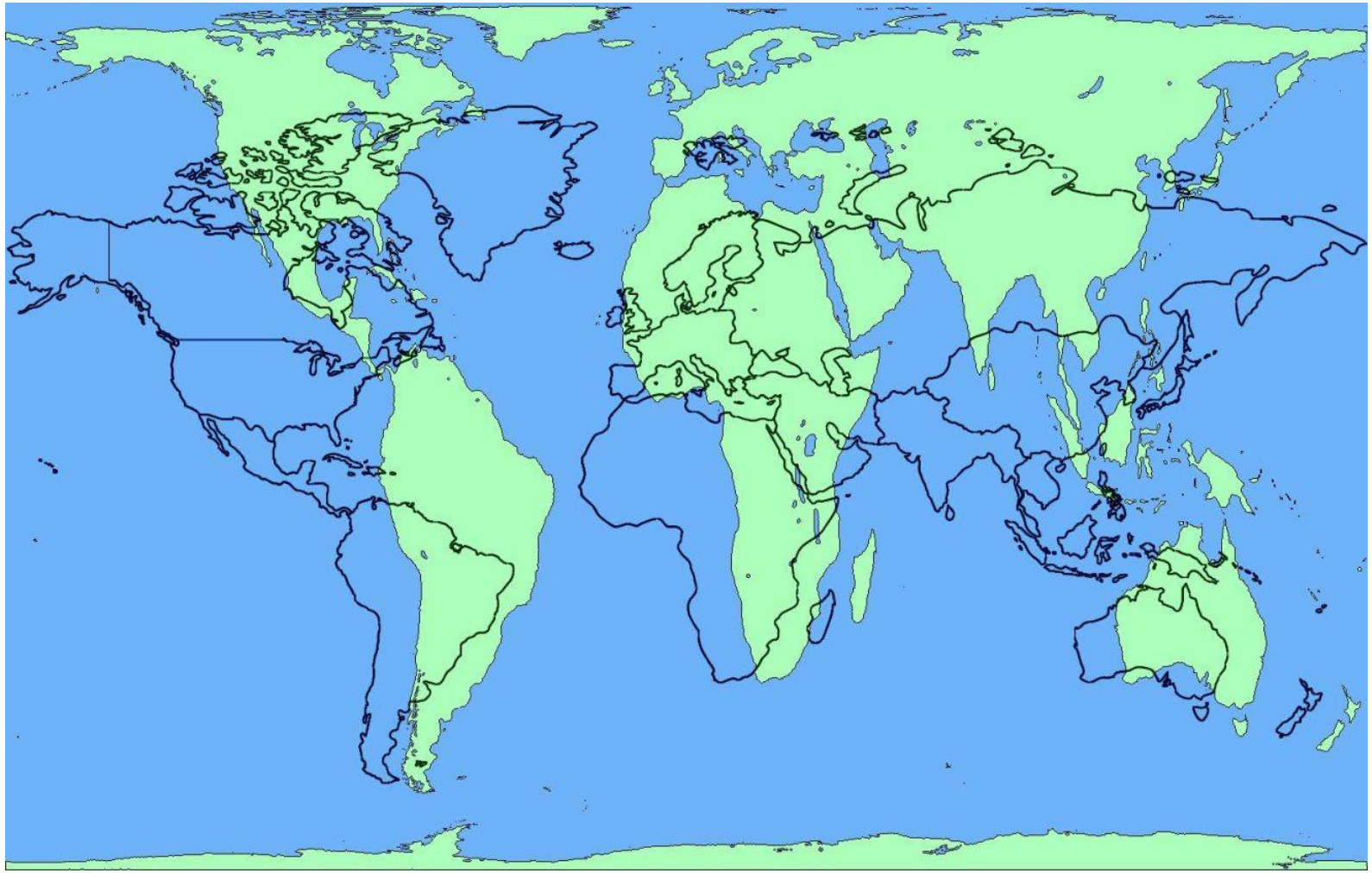
Tougher nut to crack

It’s about the story…



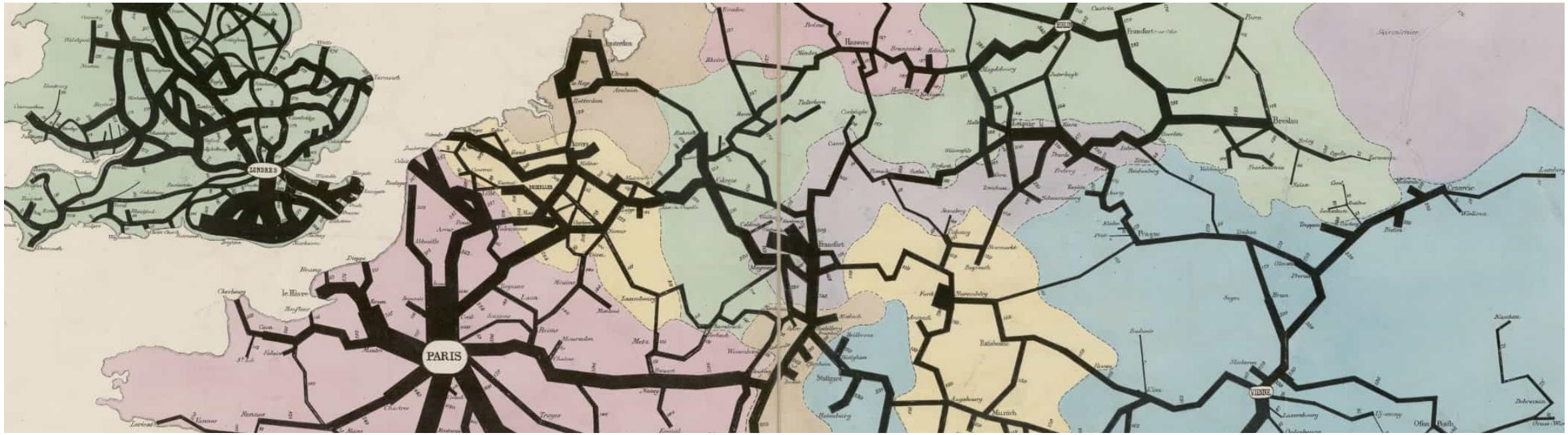
It’s about the story…

Peters vs Mercator



Visualization is not a new field

From 1840 Minard's flows visualization

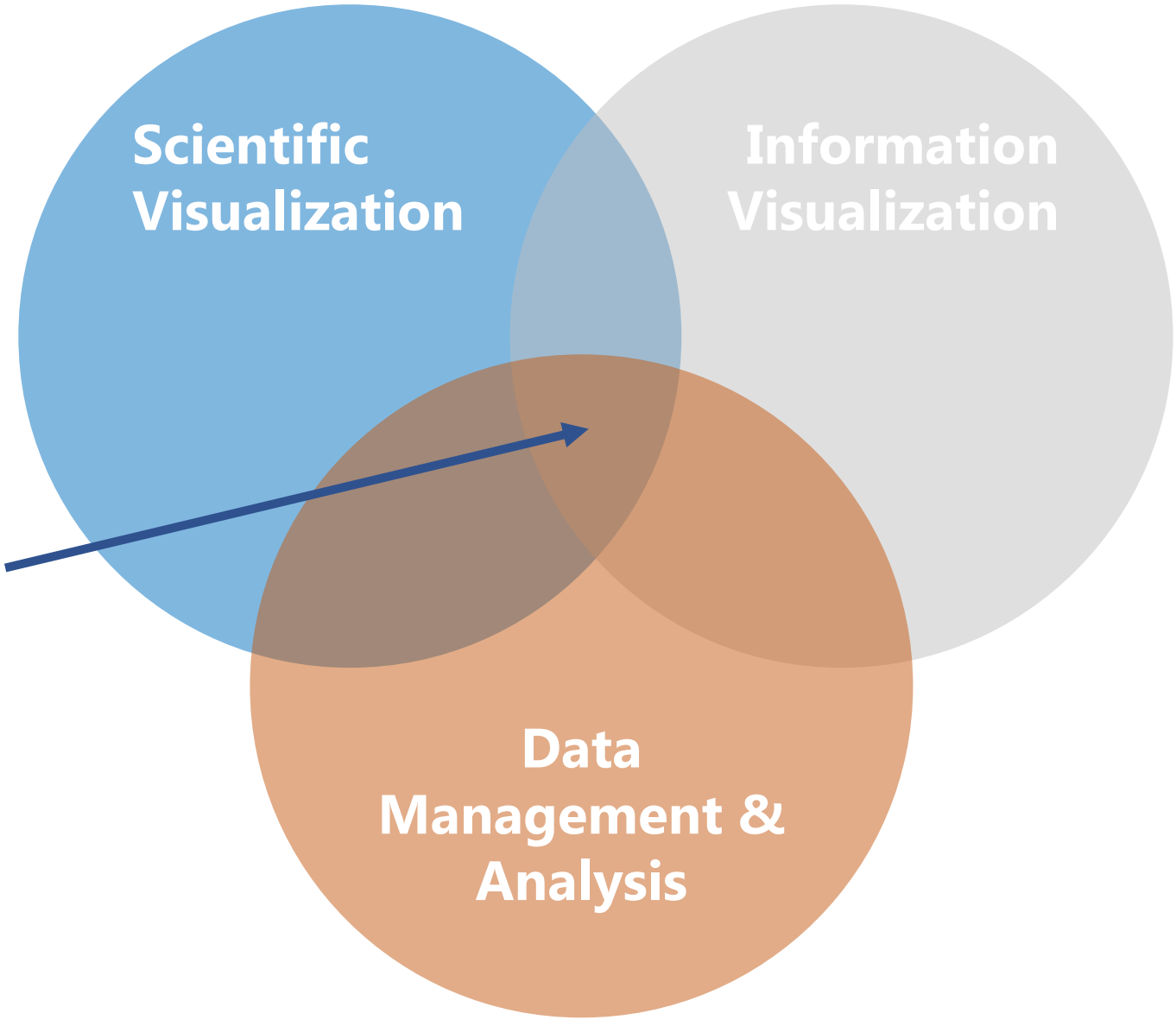


European railways Minard's flows map, 1846

Viz helped solve famous problems

1850's London Cholera Epidemic

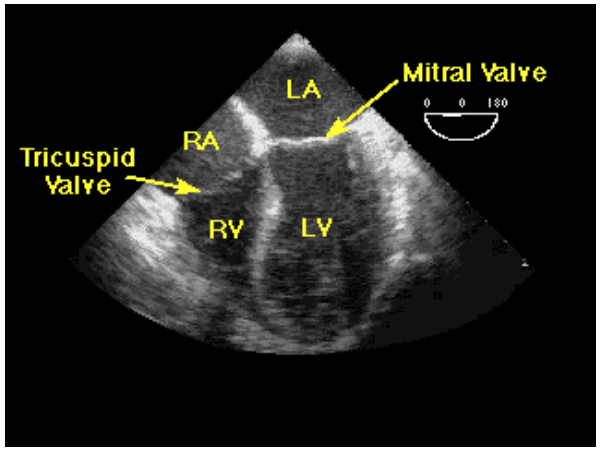
John Snow (1854)

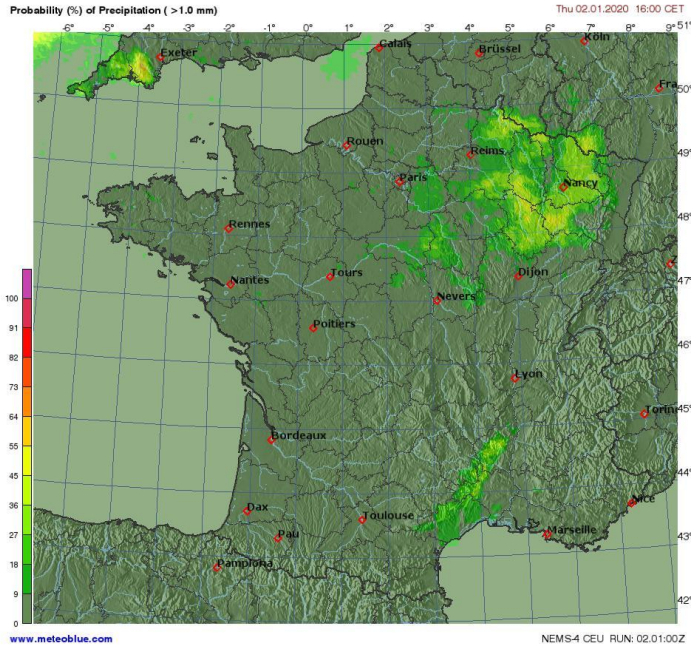

Scientific Visualization
Information Visualization
Data Management & Analysis

Data Viz

Scientific visualization

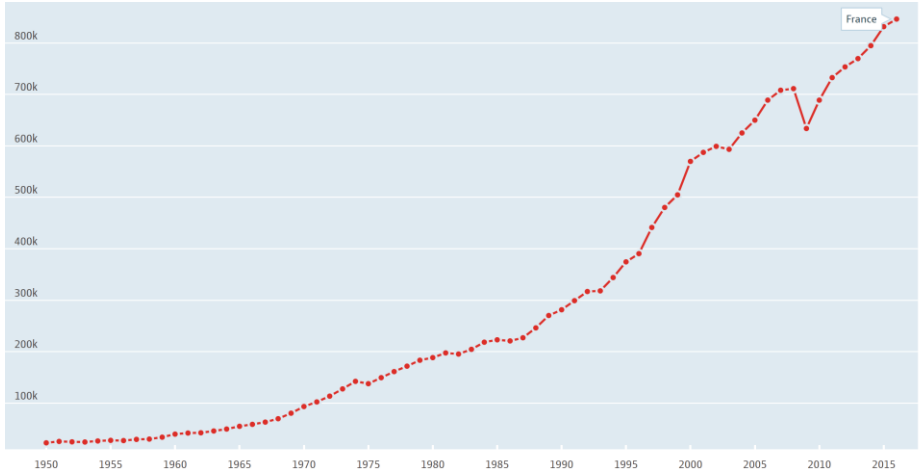
- Scientific phenomena x Visualization
- Spatial representation is given and obeys norms





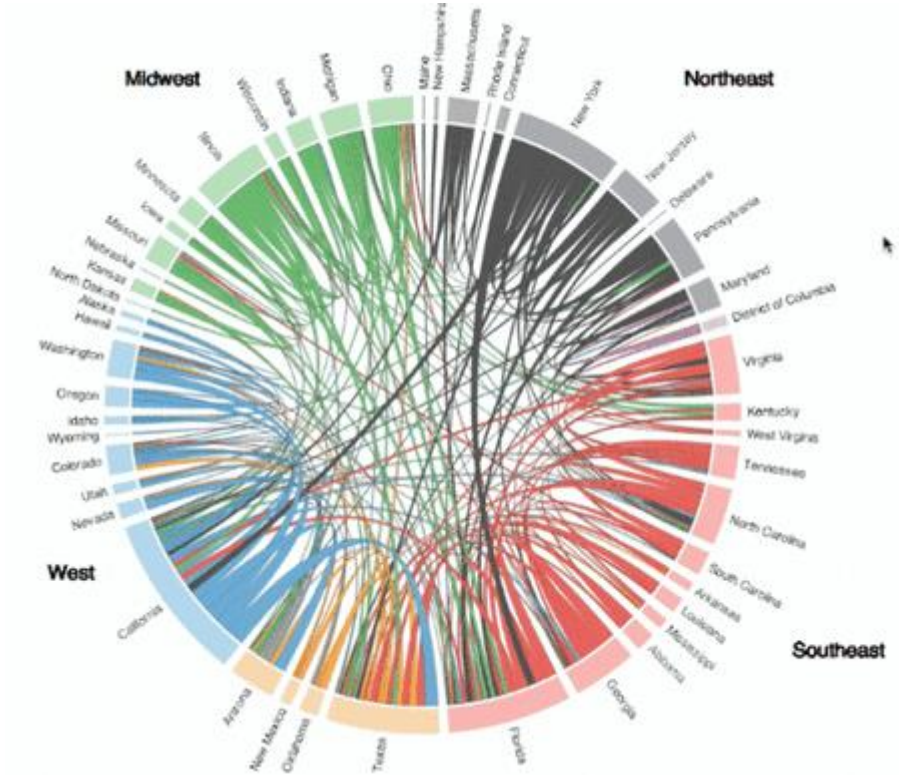
Information visualization

- Any information x Visualization
- Spatial representation is chosen



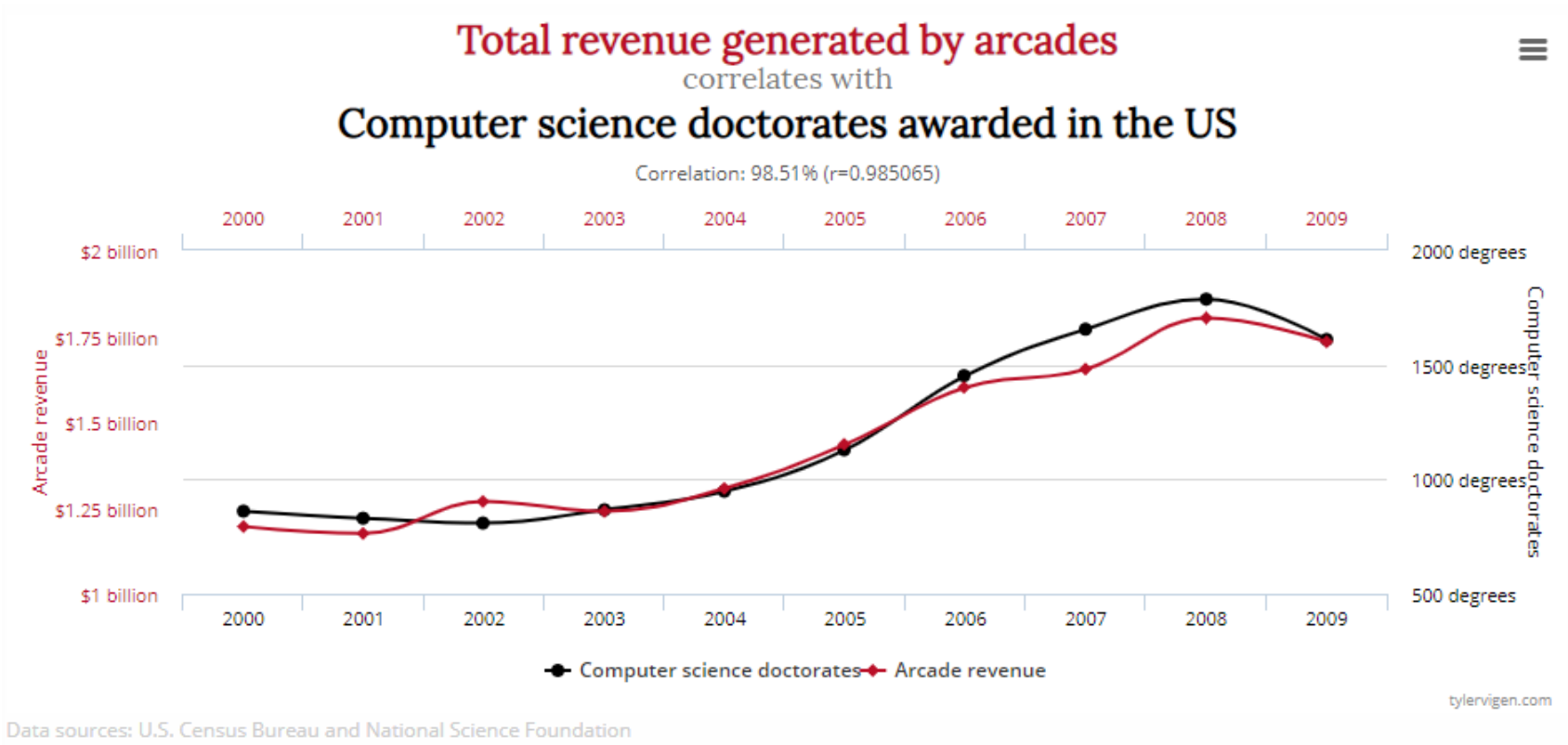
Data visualization

- Any information x Visualization
- Spatial representation is chosen
- Analytical approach
- Interactive interface



Great power comes with great responsibilities

As Data Analysts/Scientist you have a duty of scientific honesty/truthfulness



Ethics

- It’s okay to make mistakes
- It’s not okay to manipulate data
- “Would an expert approve my approach ?”
- “Would I approve it if someone else was presenting this ?”

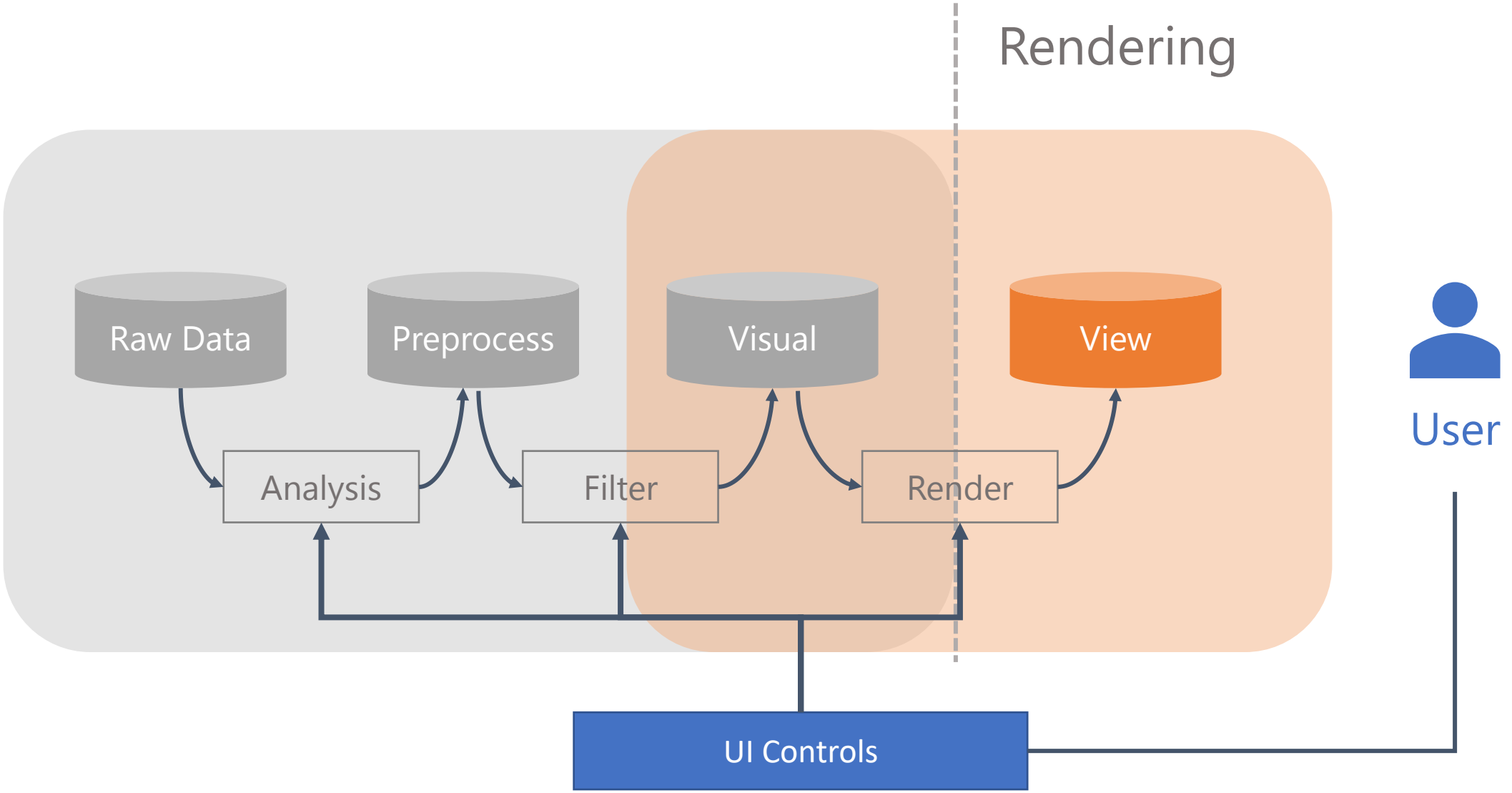
1. Course overview
2. Setup & Some tips
3. What is Data viz and why does it matters?
4. Skills, Workflow & Tools
5. Some principles
6. Getting familiar with Tableau

What do you need to master?

- Mathematics ✓
- Statistics ✓
- Data handling – *We'll go through some basics*
- Data analysis – *We'll go through some basics*
- Data science – *We'll go through some basics*
- Coding (Python, JS, HTML) – *You should already be on the right track*
- Proficiency in visualization solutions – *You've come to the right place !*
- Aesthetics/Design – *We'll try to get you on the right track*

Self assessment

1. Go to **menti.com**
2. Use the code **72 83 53 5** & Log in


Rendering
Raw Data
Preprocess
Visual
View
Analysis
Filter
Render
User
UI Controls

What can help us achieve great viz ?

BI Software

- Tableau
- Power BI
- Qlik
- FineReport
- etc.

Javascript

- D3.js
- HighCharts
- Echarts
- etc.

Python

- Matplotlib
- Seaborn
- Plotly
- etc.

1. Course overview
2. Setup & Some tips
3. What is Data viz and why does it matters?
4. Skills, Workflow & Tools
5. Some principles
6. Getting familiar with Tableau

How many 8s can you count ?

3	1	8	5	0	0	6	4	7	7	0	6
6	4	3	7	0	2	6	8	9	3	6	1
4	2	6	5	2	4	8	4	0	3	7	5
8	8	6	0	5	2	5	3	4	4	8	4
0	4	1	7	6	7	7	1	2	0	9	4
1	9	4	9	6	3	9	3	6	9	1	3
1	5	8	5	2	6	0	1	2	6	3	1
2	0	2	7	9	0	3	2	2	0	9	3
5	7	7	6	7	3	5	8	7	5	6	6
5	7	5	8	4	5	0	4	5	2	6	4
8	1	3	6	9	6	6	6	5	3	8	1
9	6	8	4	9	0	7	4	5	0	8	0

How many 8s now ?

3	1	8	5	0	0	6	4	7	7	0	6
6	4	3	7	0	2	6	8	9	3	6	1
4	2	6	5	2	4	8	4	0	3	7	5
8	8	6	0	5	2	5	3	4	4	8	4
0	4	1	7	6	7	7	1	2	0	9	4
1	9	4	9	6	3	9	3	6	9	1	3
1	5	8	5	2	6	0	1	2	6	3	1
2	0	2	7	9	0	3	2	2	0	9	3
5	7	7	6	7	3	5	8	7	5	6	6
5	7	5	8	4	5	0	4	5	2	6	4
8	1	3	6	9	6	6	6	5	3	8	1
9	6	8	4	9	0	7	4	5	0	5	0

How many 8s now ?

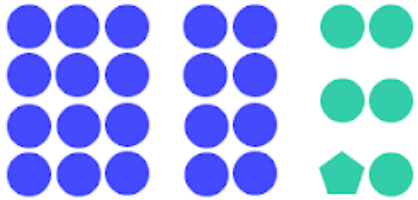
3	1	6	5	0	0	6	4	7	7	0	6
6	4	3	7	0	2	8	8	8	8	6	1
4	2	6	5	2	4	8	8	8	8	7	5
9	3	6	0	5	2	8	8	8	8	4	4
0	4	1	7	6	7	7	1	2	0	9	4
1	9	4	9	6	3	9	3	6	9	1	3
1	5	4	5	2	6	0	1	2	6	3	1
2	0	2	7	9	0	3	2	2	0	9	3
5	7	7	6	7	3	5	0	7	5	6	6
5	7	5	3	4	5	0	4	5	2	6	4
3	1	3	6	9	6	6	6	5	3	5	1
9	6	4	4	9	0	7	4	5	0	5	0

How many 8s now ?

8	8	8	8	8	8	8	8	8	8	8	8
6	4	3	7	0	2	6	3	9	3	6	1
4	2	6	5	2	4	1	4	0	3	7	5
5	0	6	0	5	2	5	3	4	4	0	4
0	4	1	7	6	7	7	1	2	0	9	4
1	9	4	9	6	3	9	3	6	9	1	3
1	5	6	5	2	6	0	1	2	6	3	1
2	0	2	7	9	0	3	2	2	0	9	3
5	7	7	6	7	3	5	4	7	5	6	6
5	7	5	7	4	5	0	4	5	2	6	4
0	1	3	6	9	6	6	6	5	3	7	1
9	6	6	4	9	0	7	4	5	0	5	0

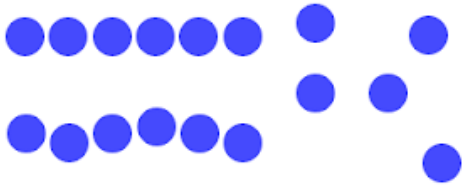
Gestalt Principles

1. PROXIMITY
Elements which are close together seem to be a group



2. CLOSURE
The human brain ignores gaps and tries to understand the bigger context

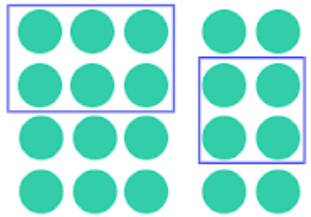
5. CONTINUITY
Elements which are ordered in a line or curve seem to be a group



6. FIGURE & GROUND
The human brain instinctively recognizes if something is in the fore- or background

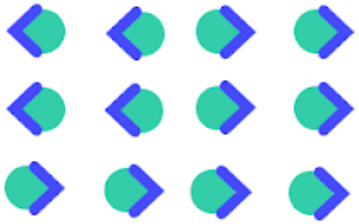
3. SIMILARITY
Elements which look similar seem to be a group

4. COMMON REGION
Elements which are close together seem to be a group



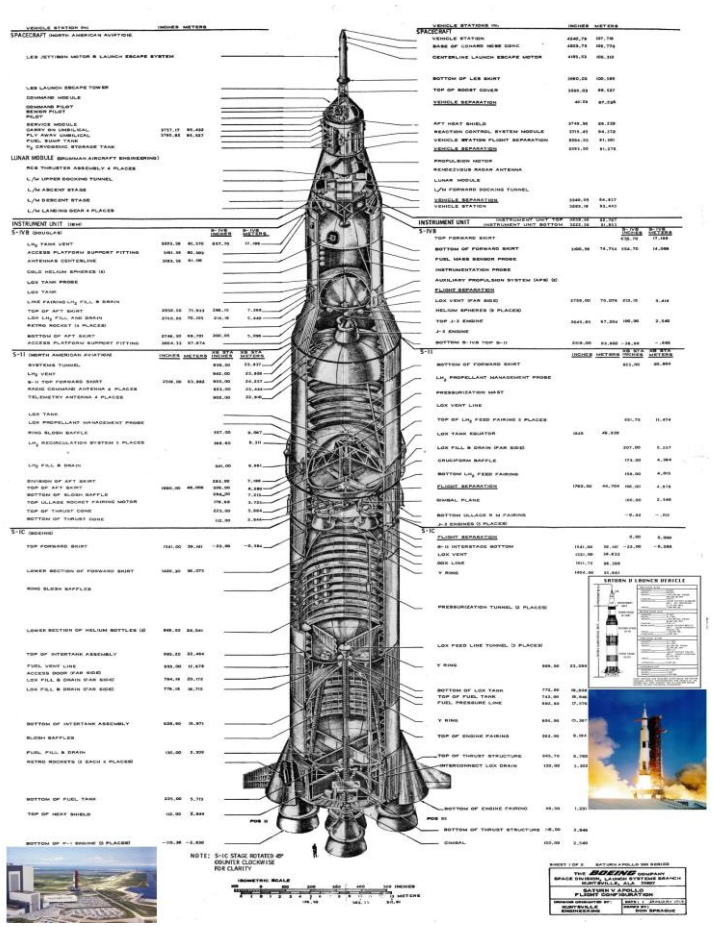
7. SYMMETRY
Symmetric elements give the human brain the feeling that everything is ordered

8. COMMON FATE
Elements which move in the same direction seem to be a group



Source : https://uxcam.com/blog/8-gestalt-principles-design-infographic/

Know your audience



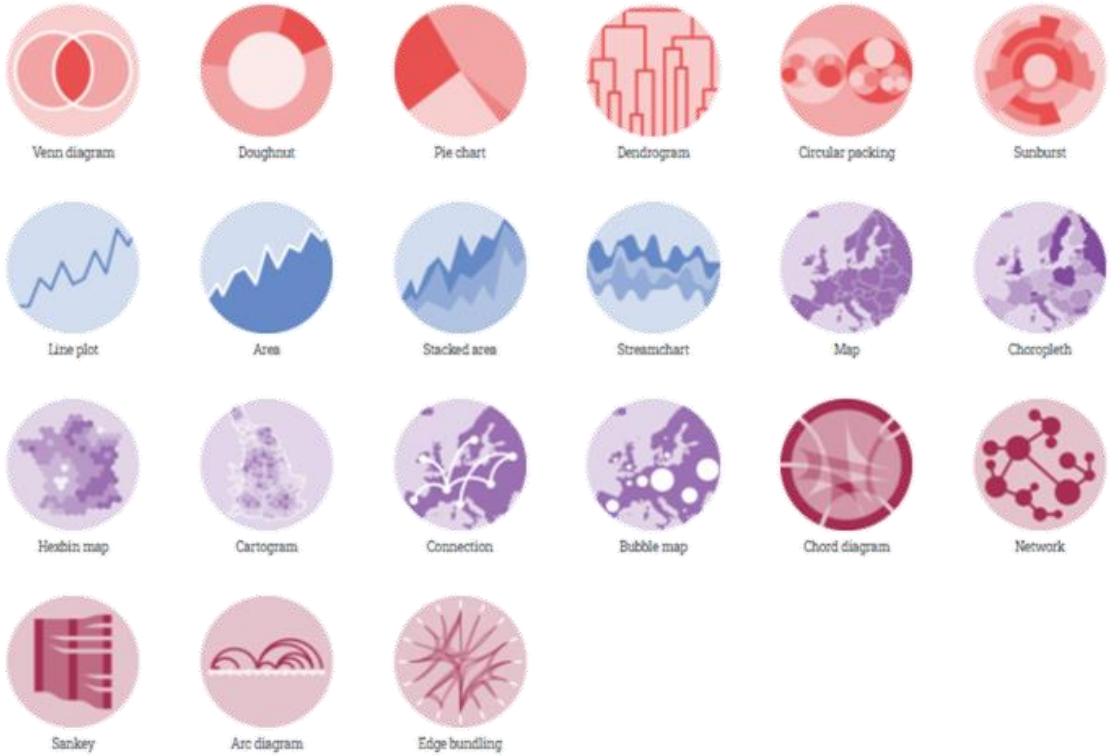
Experts

Non Experts

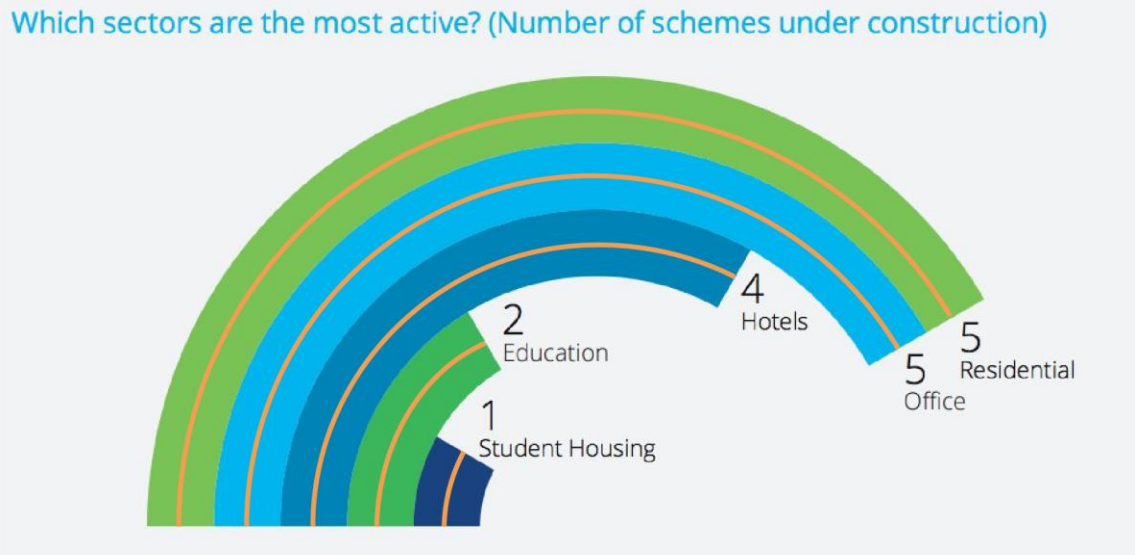
Aesthetics

- Aesthetics comes 2nd
- **Data is always first**
- But if your visualization is messy or unreadable it says: **I don't care if *you* don't understand**

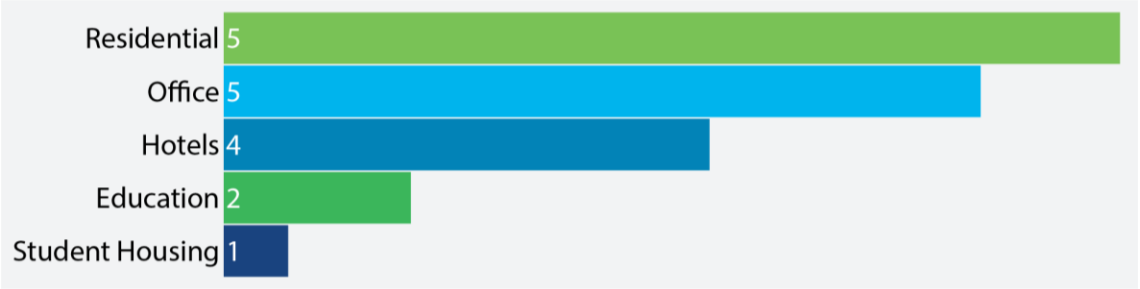
Know your graphs



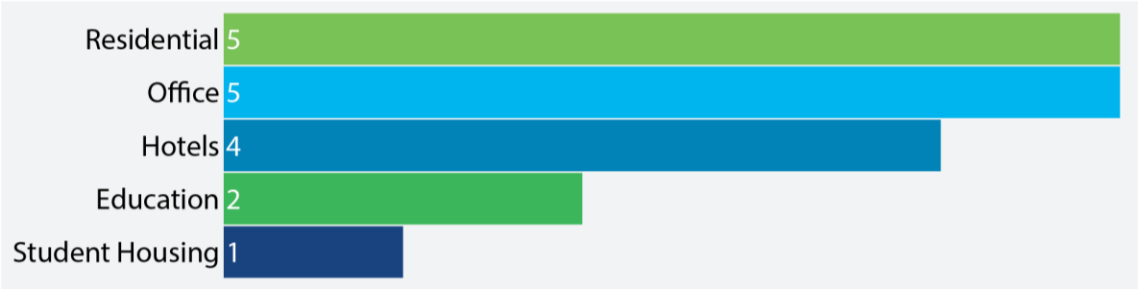
Caveats



ARC LENGTH PROPORTIONS

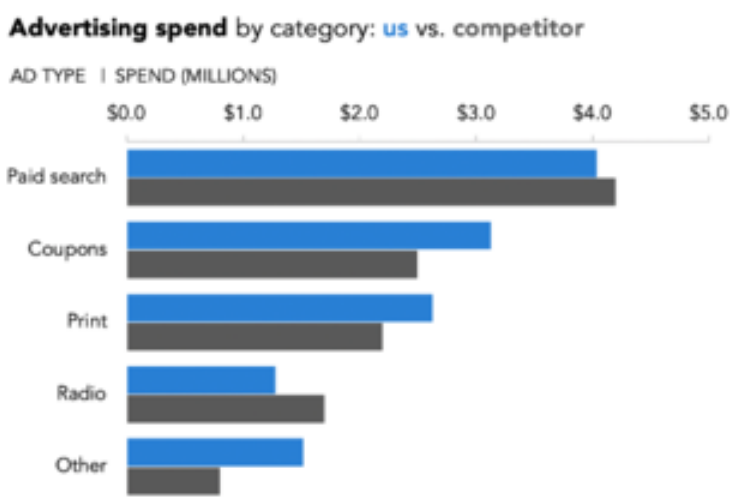


CORRECTED BAR LENGTH PROPORTIONS

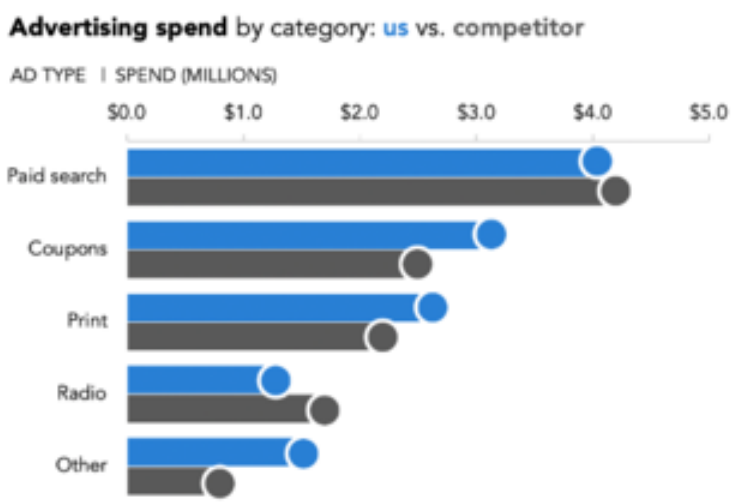


Simple is better

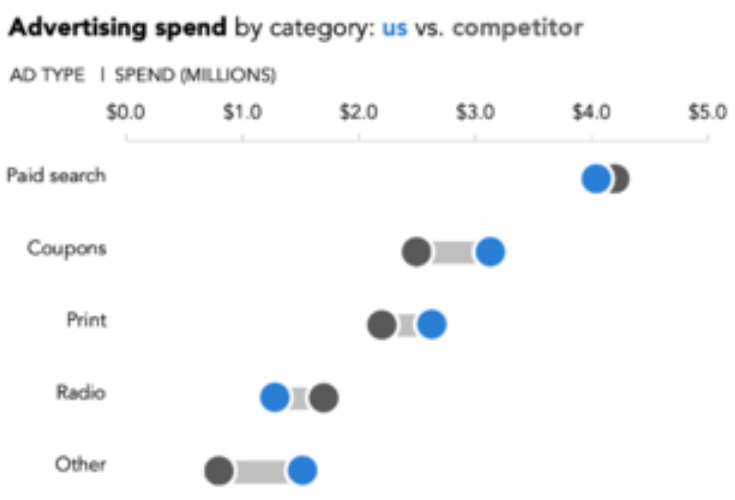
A dual-series bar chart…



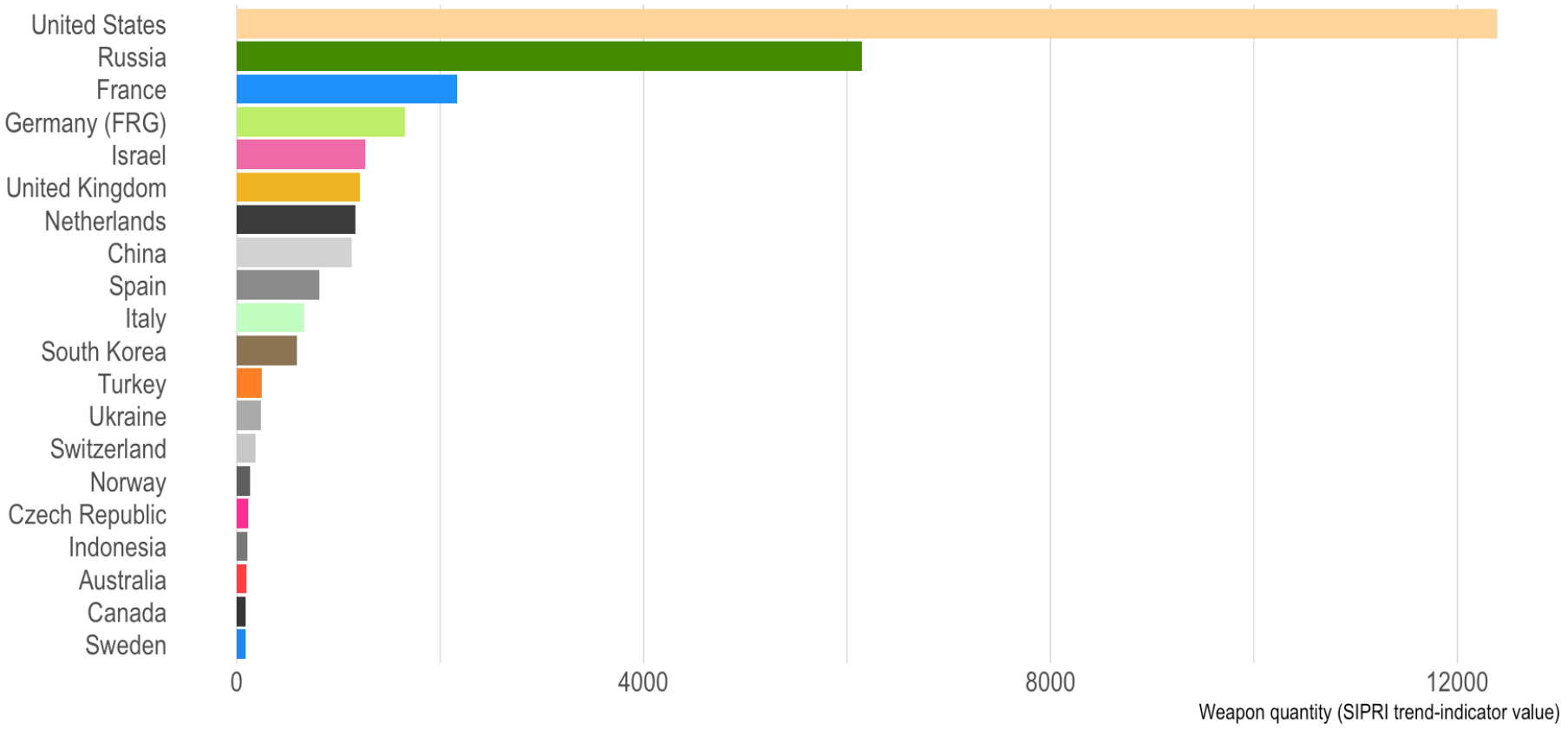
Encode end points with dots.



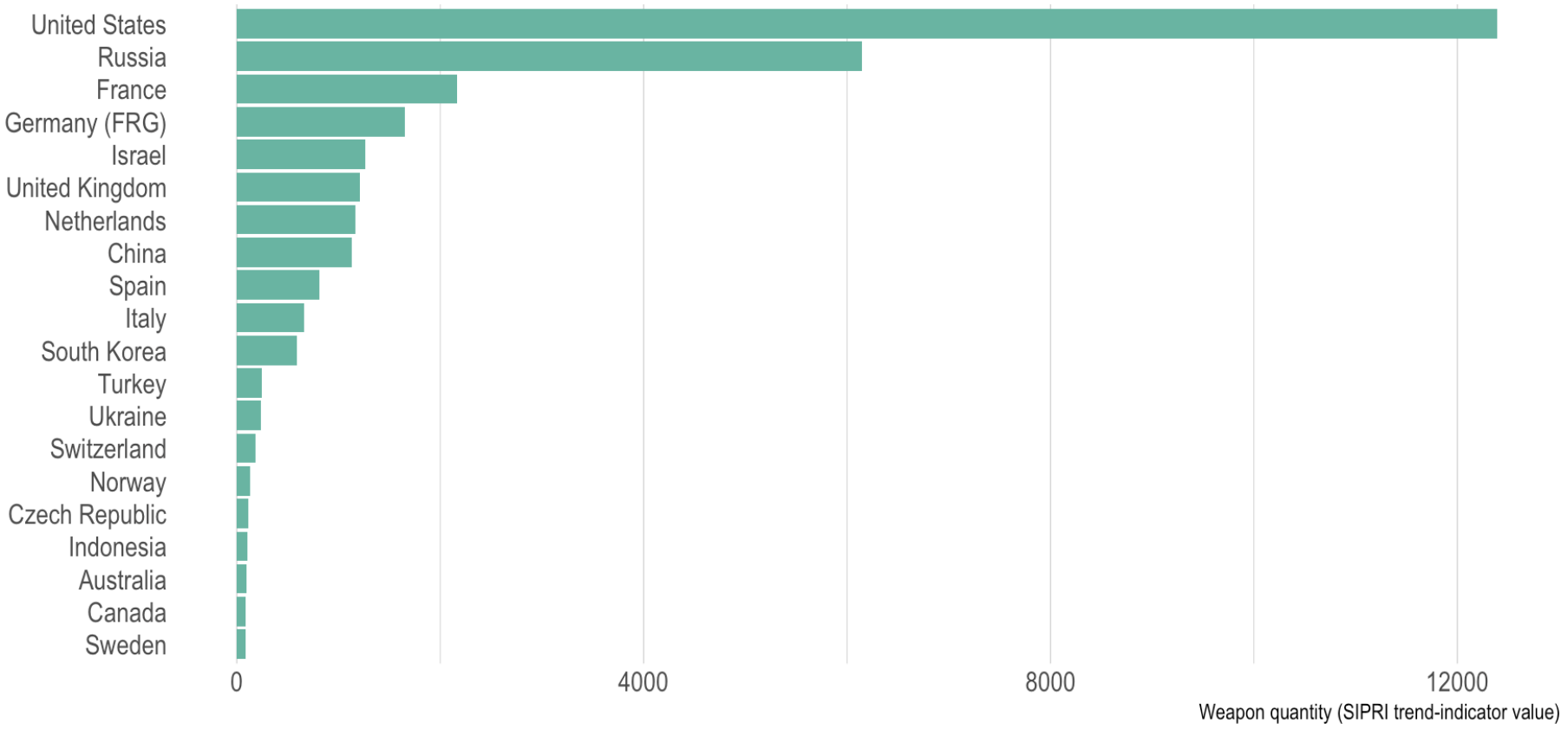
Collapse, remove bars, connect.



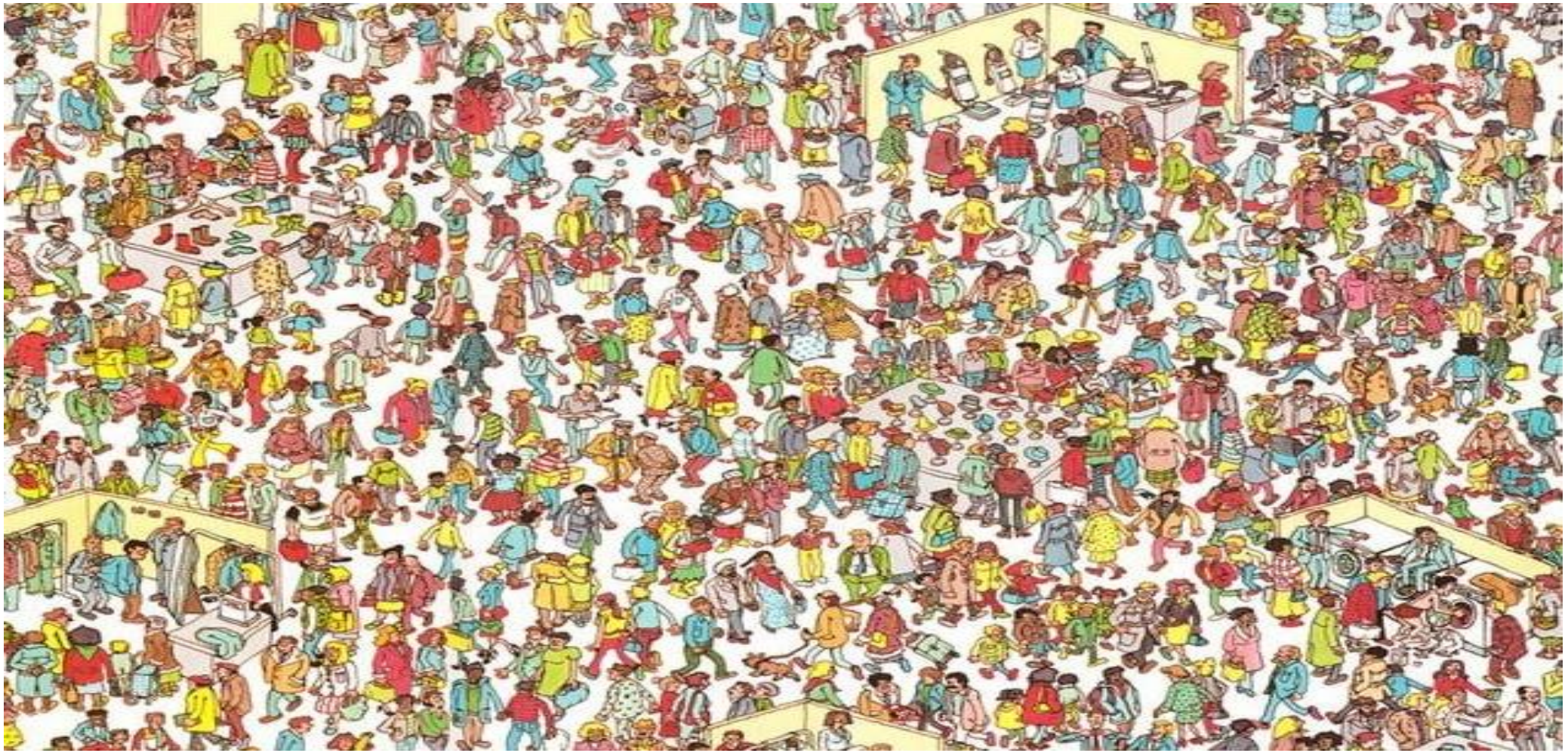
Be mindful with colors



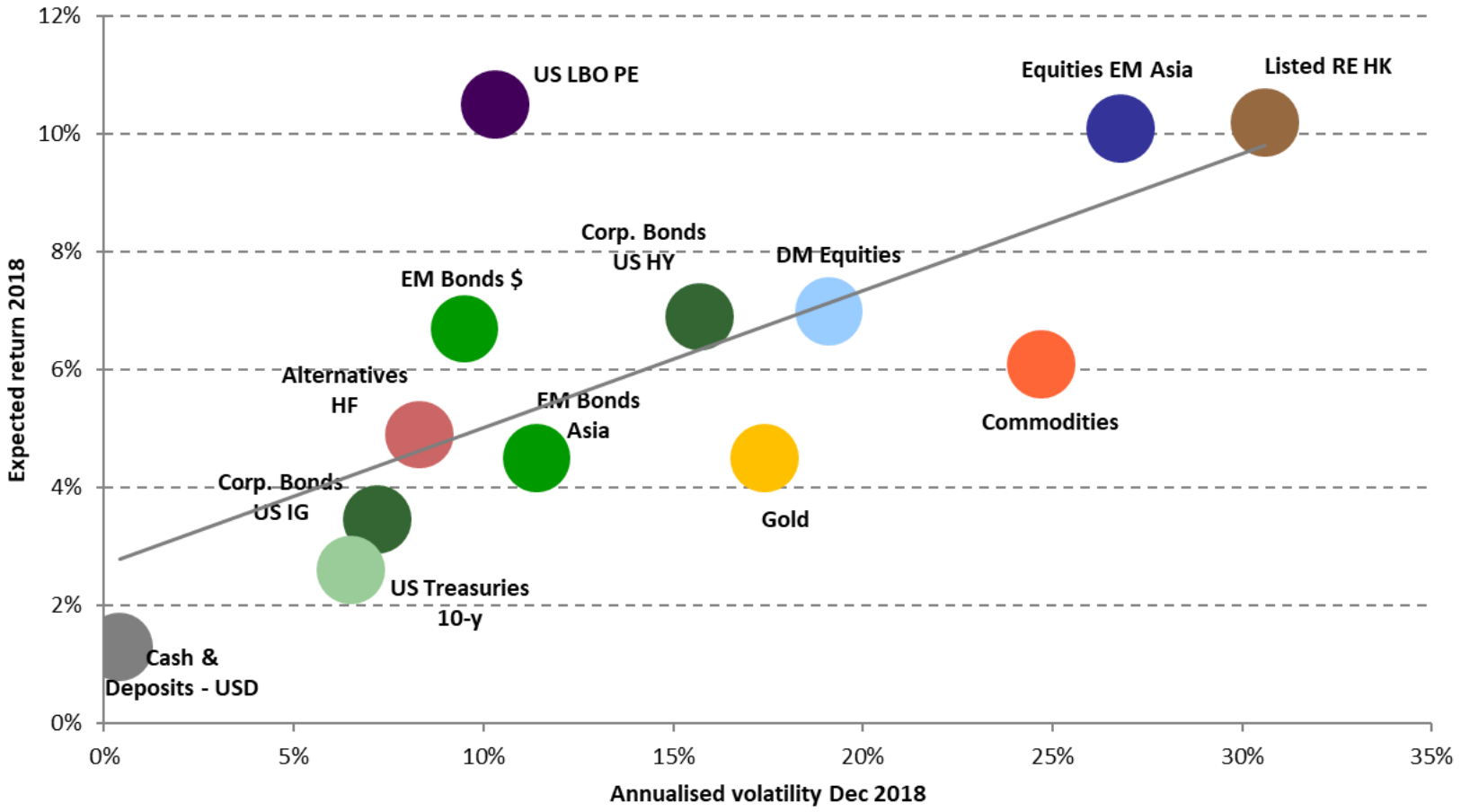
Be mindful with colors



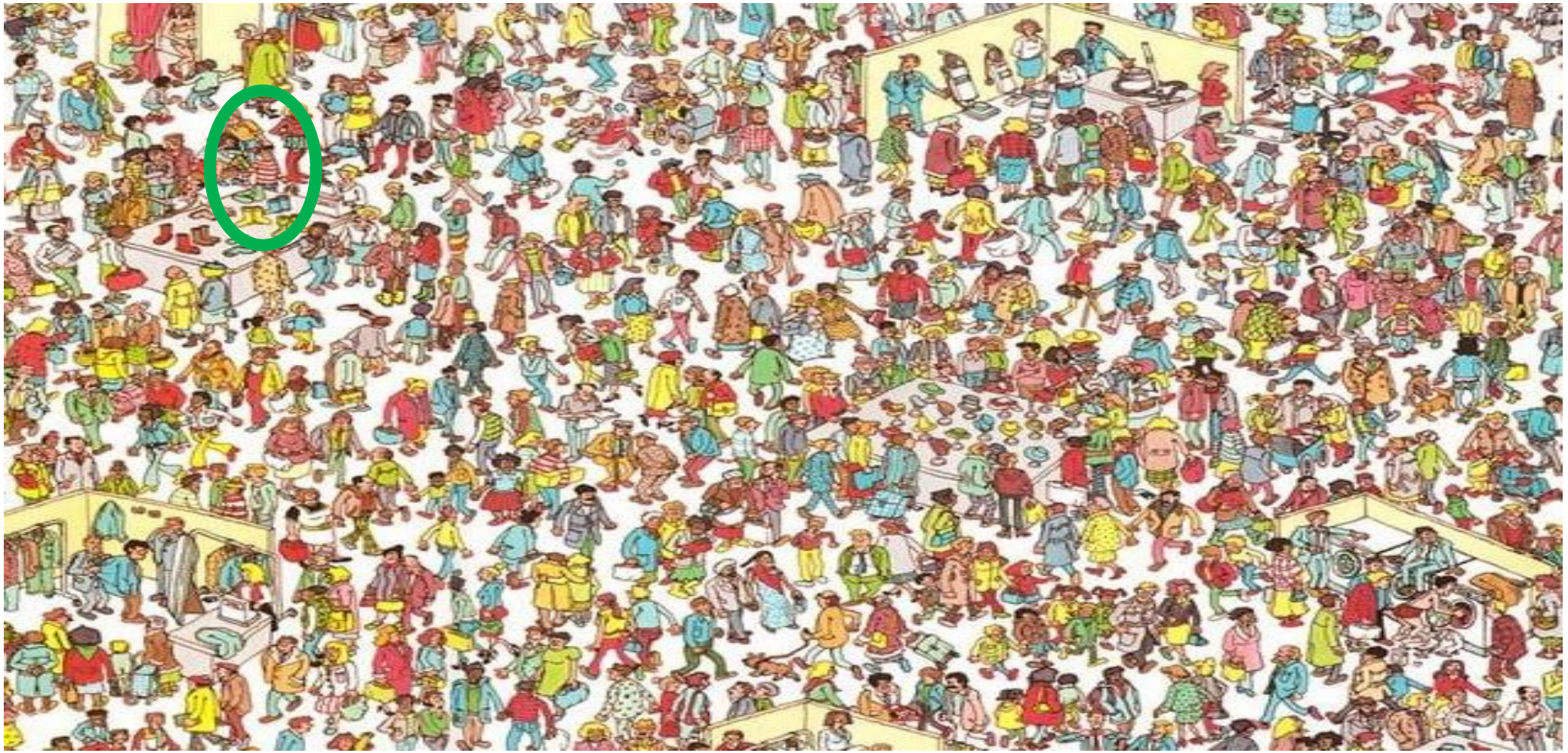
Guide your audience



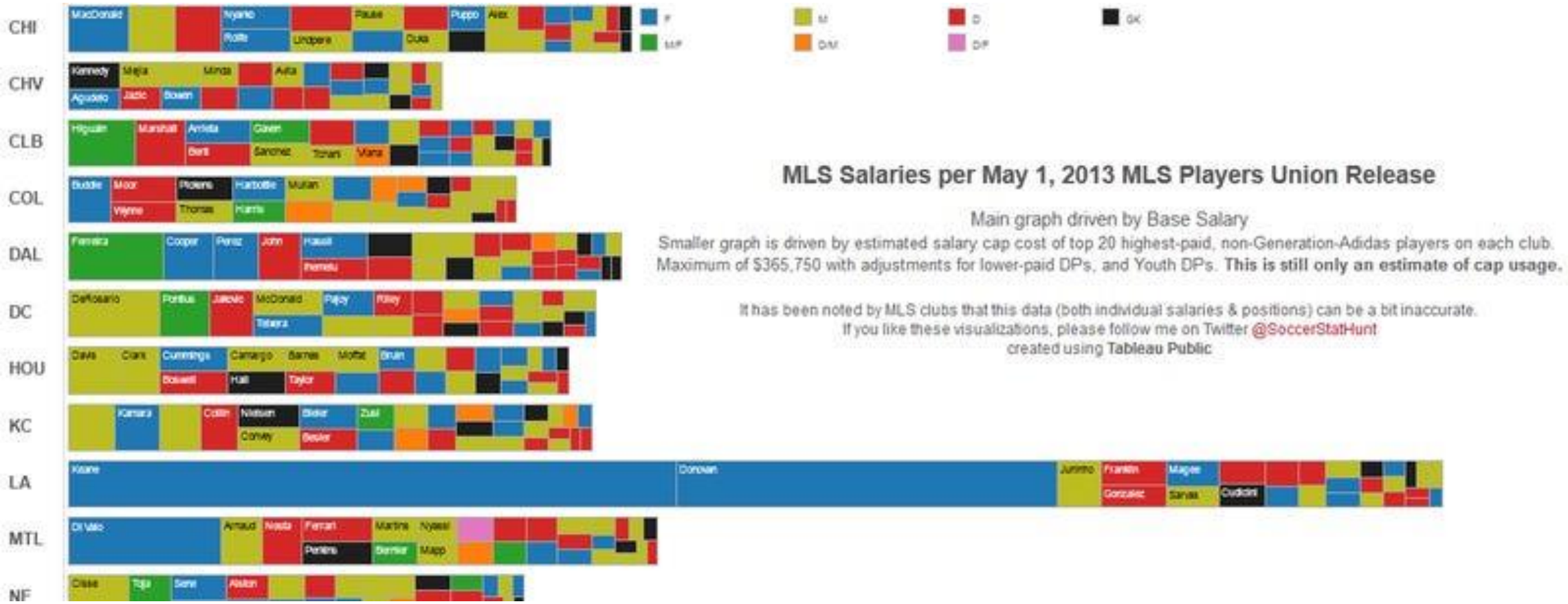
Guide your audience



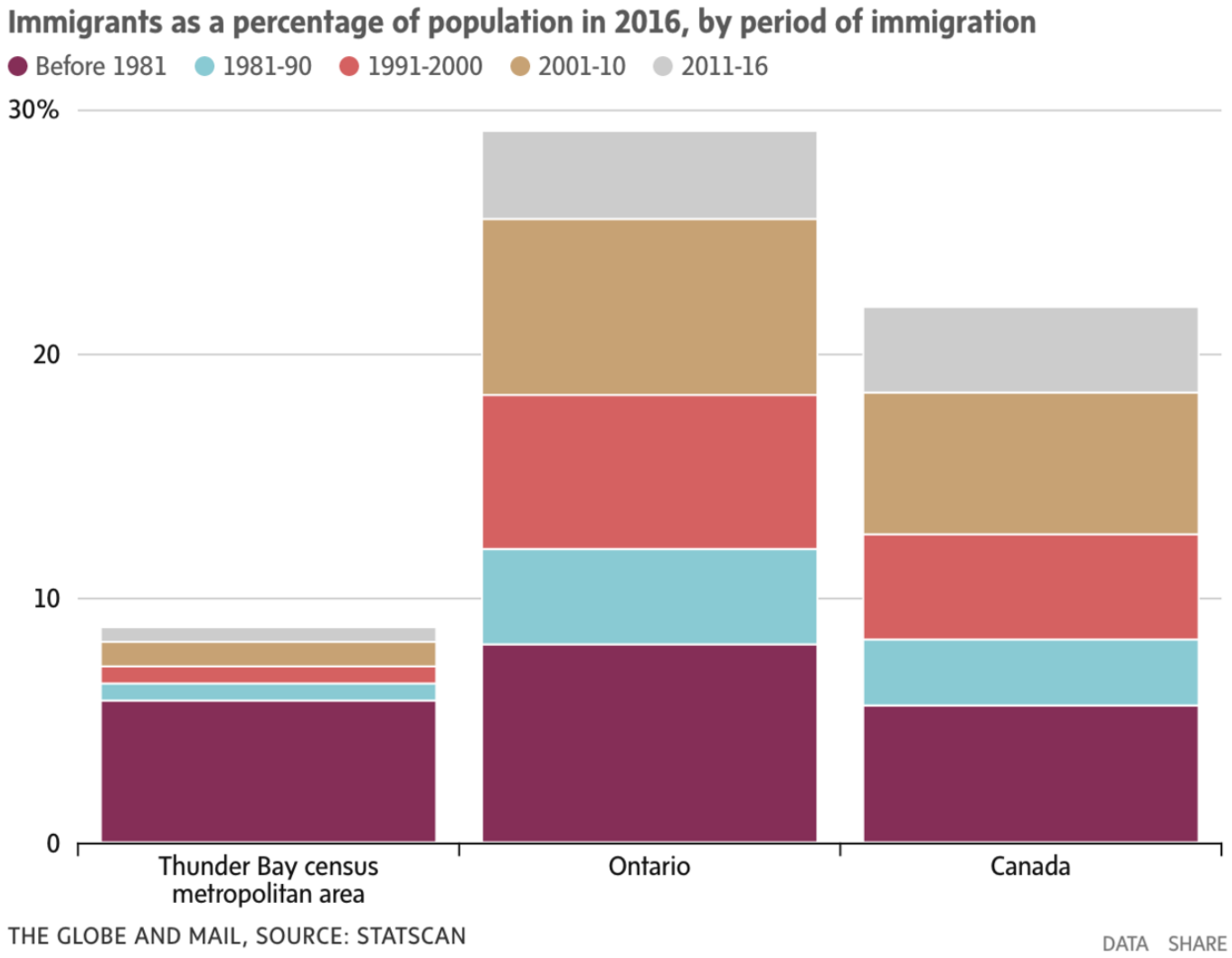
Guide your audience



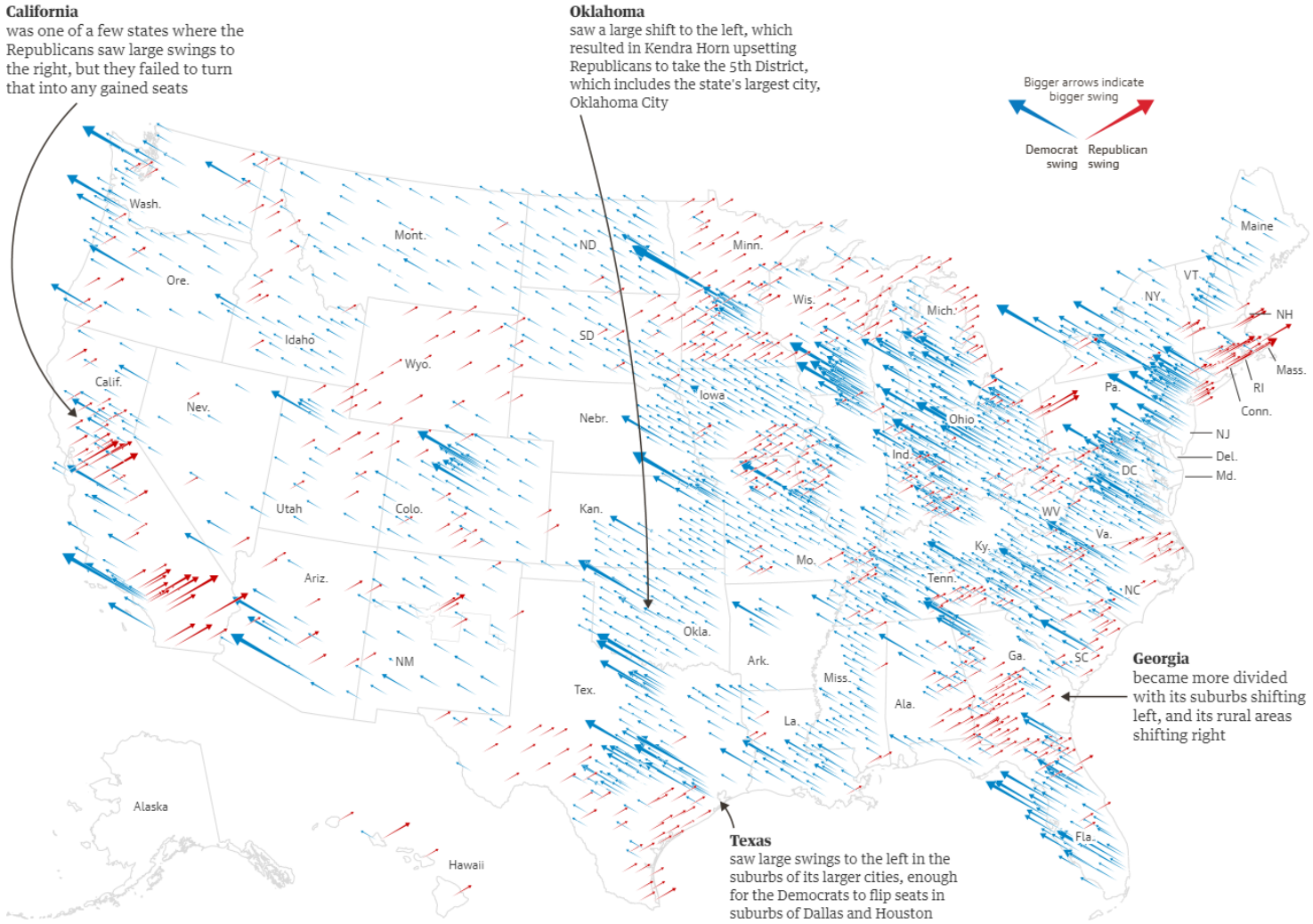
1 graph = 1 message



#Fail

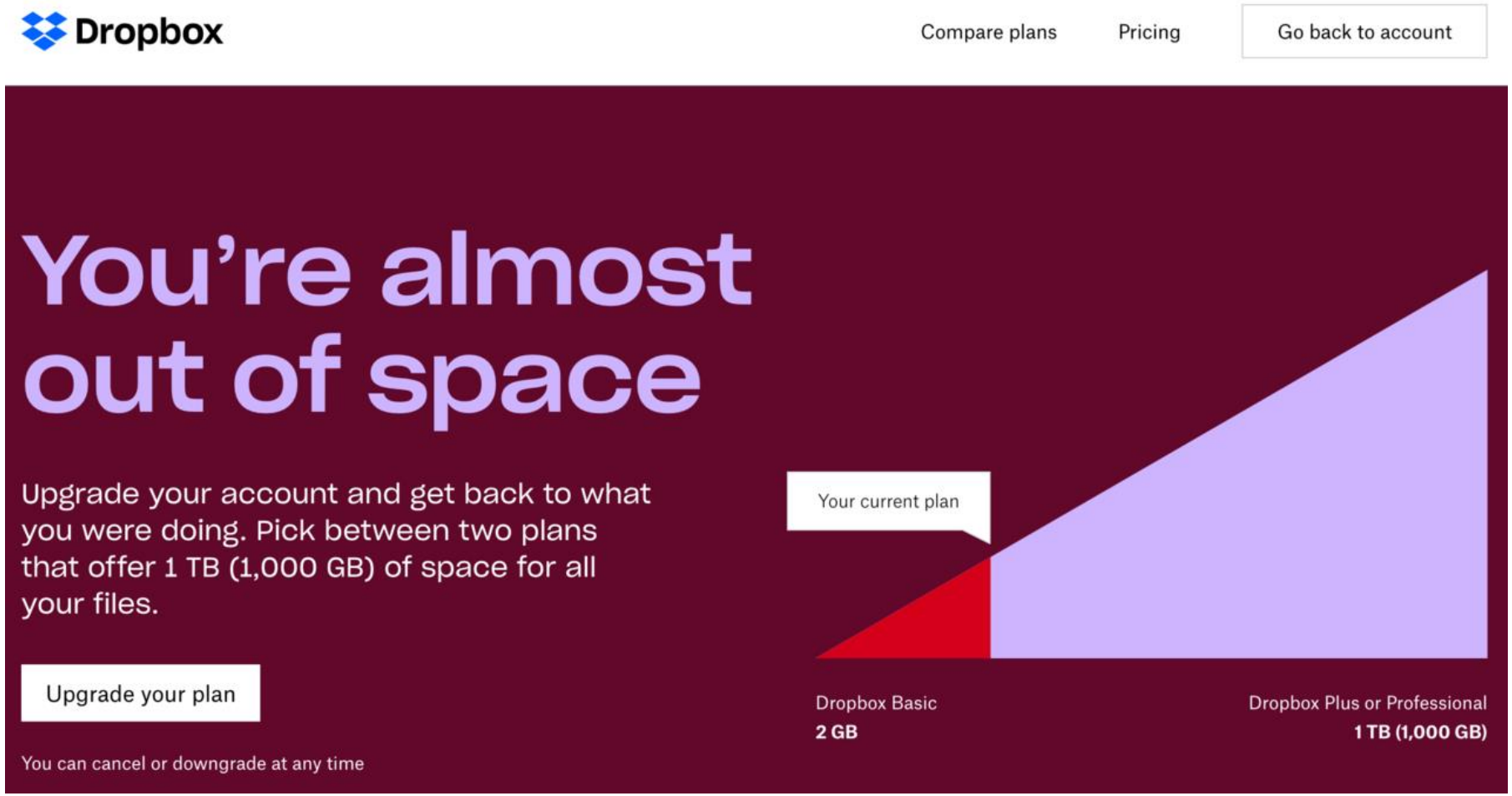

Immigrants as a percentage of population in 2016, by period of immigration
Before 1981
1981-90
1991-2000
2001-10
2011-16
30%
20
10
0
Thunder Bay census metropolitan area
Ontario
Canada
THE GLOBE AND MAIL, SOURCE: STATSCAN
DATA
SHARE

#Nice

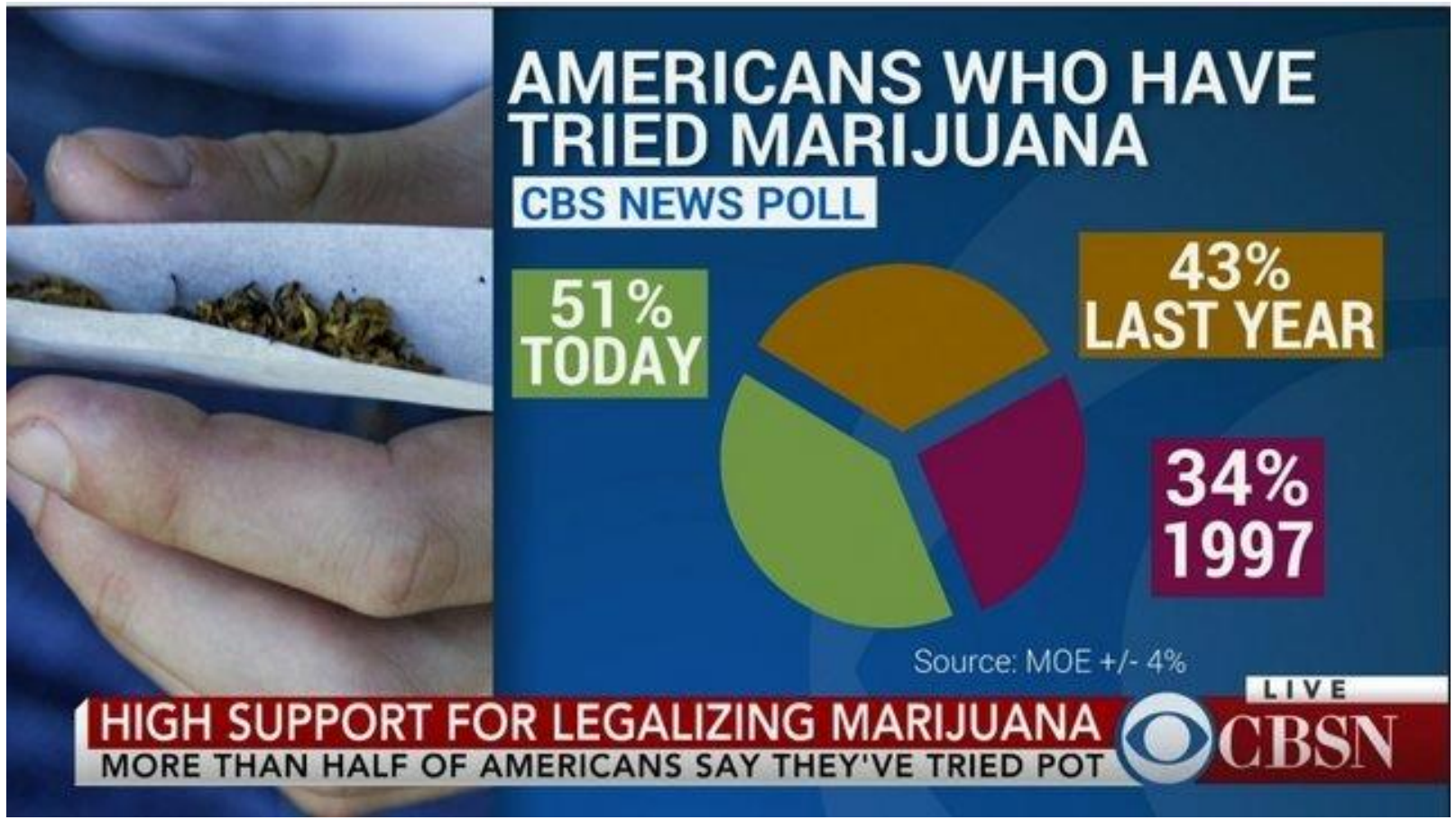


Source : https://www.theguardian.com/us-news/ng-interactive/2018/nov/07/blue-wave-or-blue-ripple-a-visual-guide-to-the-democrats-gains

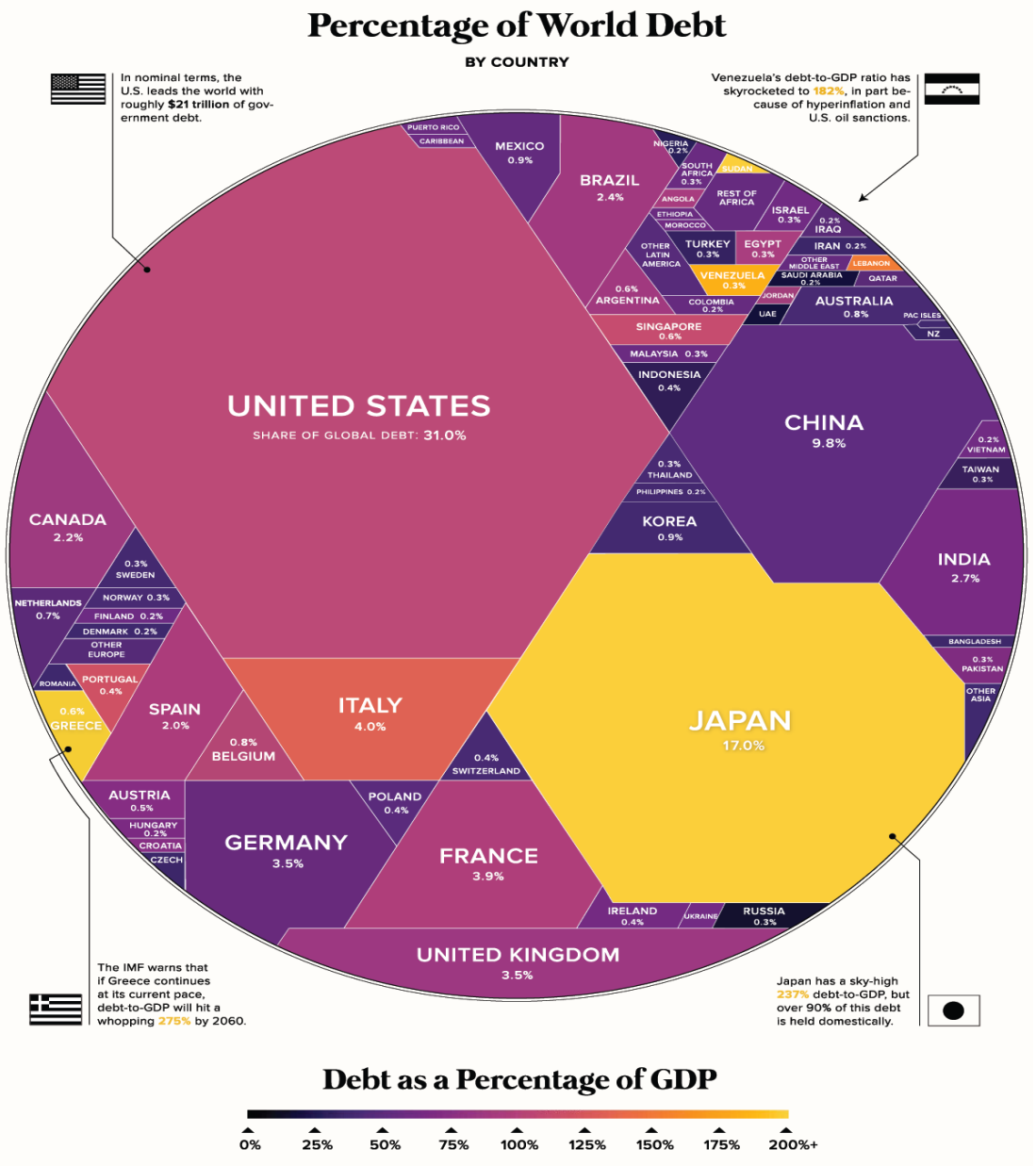
#Fail


Dropbox
Compare plans
Pricing
Go back to account
You're almost out of space
Upgrade your account and get back to what you were doing. Pick between two plans that offer 1 TB (1,000 GB) of space for all your files.
Upgrade your plan
You can cancel or downgrade at any time
Your current plan
Dropbox Basic
2 GB
Dropbox Plus or Professional
1 TB (1,000 GB)

#Fail


AMERICANS WHO HAVE TRIED MARIJUANA
CBS NEWS POLL
51% TODAY
43% LAST YEAR
34% 1997
Source: MOE +/- 4%
HIGH SUPPORT FOR LEGALIZING MARIJUANA
MORE THAN HALF OF AMERICANS SAY THEY'VE TRIED POT
LIVE
CBSN

#Nice


Percentage of World Debt
BY COUNTRY
In nominal terms, the U.S. leads the world with roughly $21 trillion of government debt.
Venezuela's debt-to-GDP ratio has skyrocketed to 182%, in part because of hyperinflation and U.S. oil sanctions.
UNITED STATES
SHARE OF GLOBAL DEBT: 31.0%
CHINA 9.8%
JAPAN 17.0%
GERMANY 3.5%
FRANCE 3.9%
ITALY 4.0%
UNITED KINGDOM 3.5%
SPAIN 2.0%
CANADA 2.2%
BRAZIL 2.4%
MEXICO 0.9%
INDIA 2.7%
AUSTRALIA 0.8%
KOREA 0.9%
SINGAPORE 0.6%
INDONESIA 0.4%
MALAYSIA 0.3%
THAILAND 0.3%
PHILIPPINES 0.2%
VIETNAM 0.2%
TAIWAN 0.3%
BANGLADESH
PAKISTAN 0.3%
OTHER ASIA
NETHERLANDS 0.7%
SWEDEN 0.3%
NORWAY 0.3%
FINLAND 0.2%
DENMARK 0.2%
OTHER EUROPE
PORTUGAL 0.4%
ROMANIA
GREECE 0.6%
BELGIUM 0.8%
SWITZERLAND 0.4%
AUSTRIA 0.5%
HUNGARY 0.2%
CROATIA
CZECH
POLAND 0.4%
IRELAND 0.4%
UKRAINE
RUSSIA 0.3%
PUERTO RICO
CARIBBEAN
NIGERIA 0.2%
SOUTH AFRICA 0.3%
SUDAN
ANGOLA
ETHIOPIA
MOROCCO
REST OF AFRICA
ISRAEL 0.3%
IRAQ 0.2%
IRAN 0.2%
OTHER LATIN AMERICA
TURKEY 0.3%
EGYPT 0.3%
OTHER MIDDLE EAST
LEBANON
SAUDI ARABIA 0.2%
QATAR
VENEZUELA 0.3%
JORDAN
UAE
ARGENTINA 0.6%
COLOMBIA 0.2%
PAC ISLES
NZ
The IMF warns that if Greece continues at its current pace, debt-to-GDP will hit a whopping 275% by 2060.
Japan has a sky-high 237% debt-to-GDP, but over 90% of this debt is held domestically.
Debt as a Percentage of GDP
0%
25%
50%
75%
100%
125%
150%
175%
200%+

#Nice

https://projects.fivethirtyeight.com/p-hacking/

https://projects.christianlaesser.com/travel-visa-inequality/

Where you can get some inspiration

fivethirtyeight.com

Visualcapitalist.com

NYTimes.com

Guardian.com

1. Course overview
2. Setup & Some tips
3. What is Data viz and why does it matters?
4. Skills, Workflow & Tools
5. Some principles
6. Getting familiar with Tableau

Build your own dashboard on Stock Prices

Data available here:

https://github.com/mscfinance-big-data-viz/session-01
https://simfin.com/data/bulk

Thank you