

Module 2: Synch Lecture

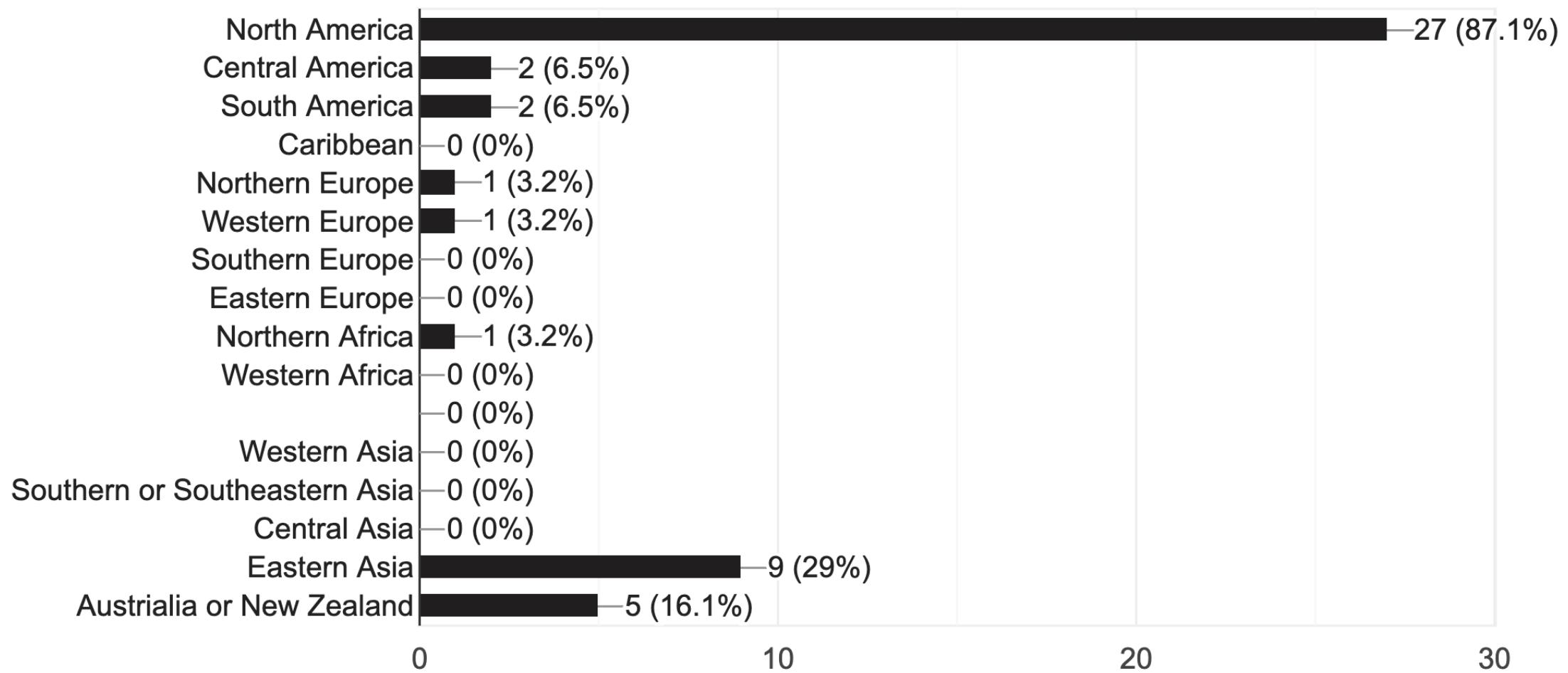
- 1) Housekeeping updates
- 2) Class survey
- 3) Meanwhile, make sure you have pulled
from the lecture repo

“Housekeeping” class updates

- Today you need to pull from the ‘live code’ github, which is generally where lecture materials live
- Homework is due at class time on Thursdays (9:35 AM CST) unless otherwise noted. We will try hard to be consistent about this.
 - organize your homework similar to the assignments repo, i.e. each week should be its own folder
- To clarify: you need to have signed up by today to meet with your TA, but the meeting can happen next week

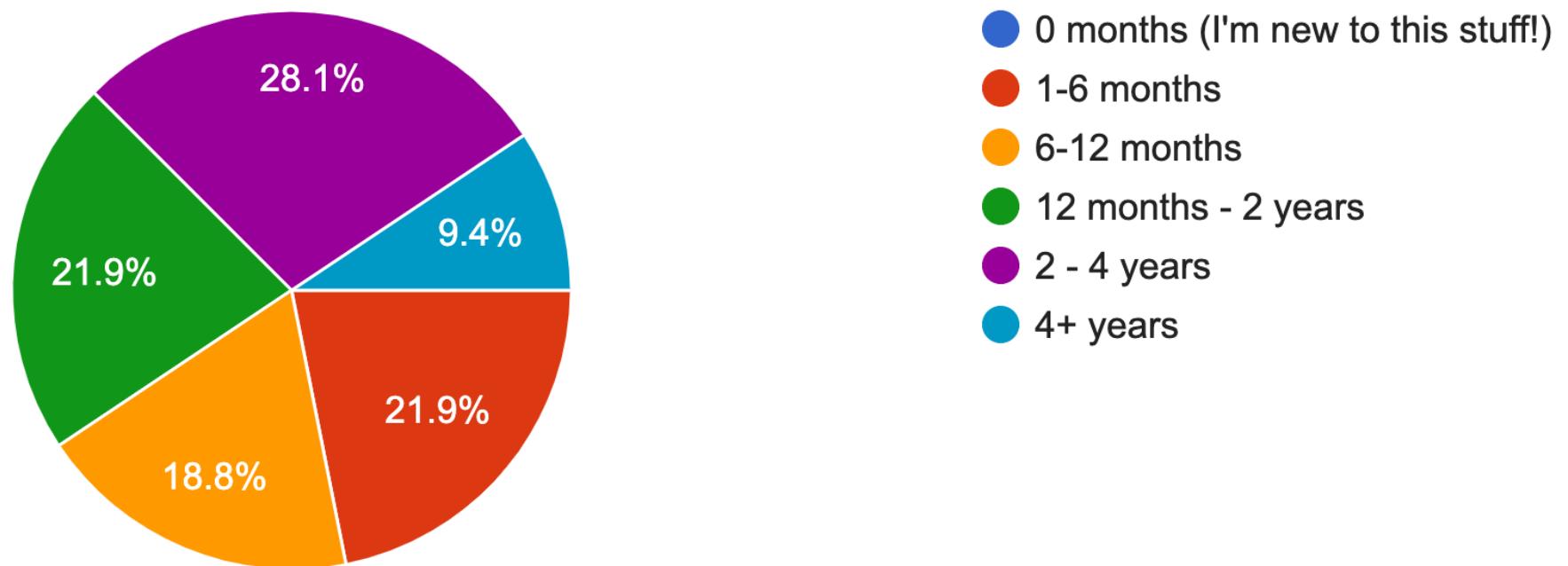
After you complete your Data Science MA, what region of the world would you MOST like to be employed (you can select more than one, but try to keep it simple)?

31 responses



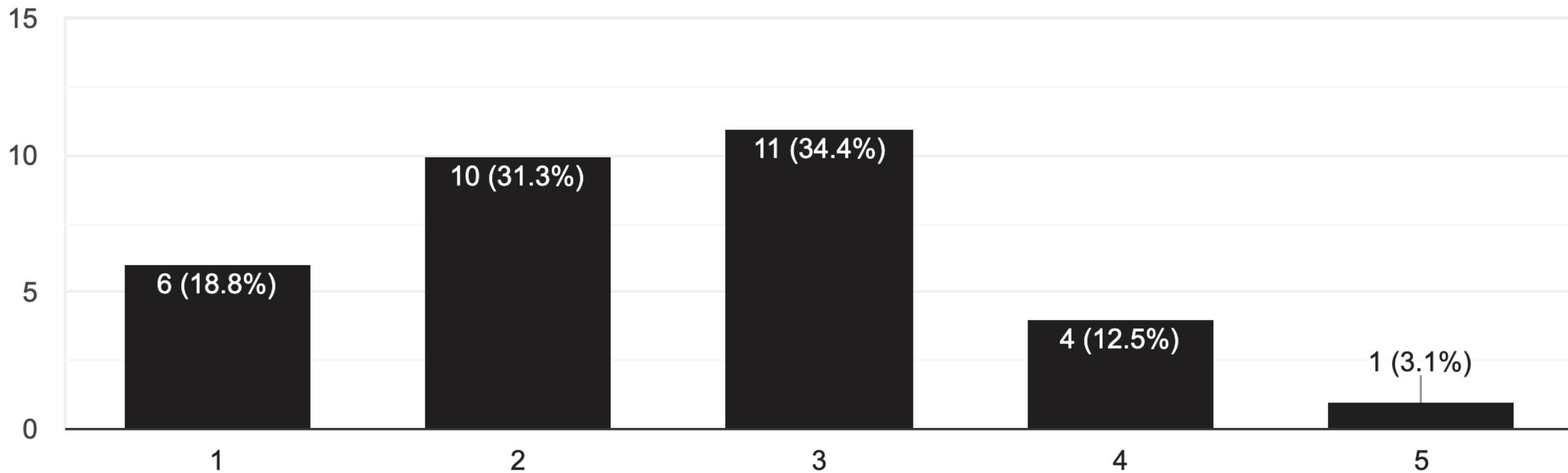
How many years of programming experience do you have (approximately)?

32 responses



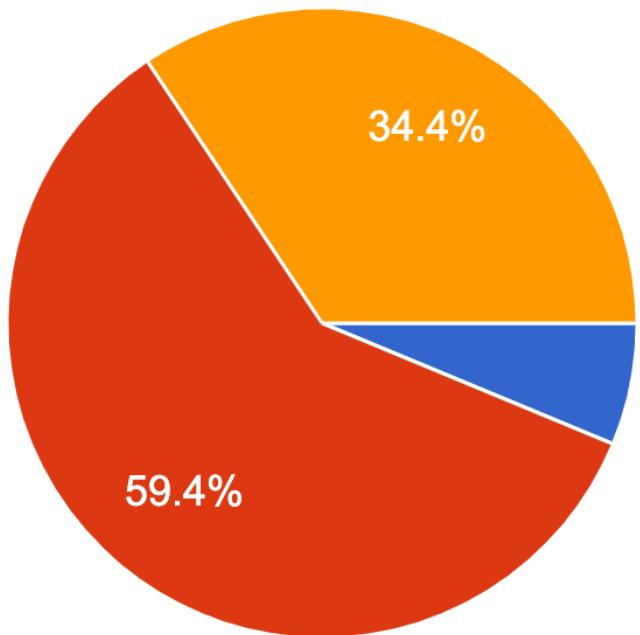
How confident are you with the R programming language?

32 responses



Tell me a bit about your background with statistics:

32 responses



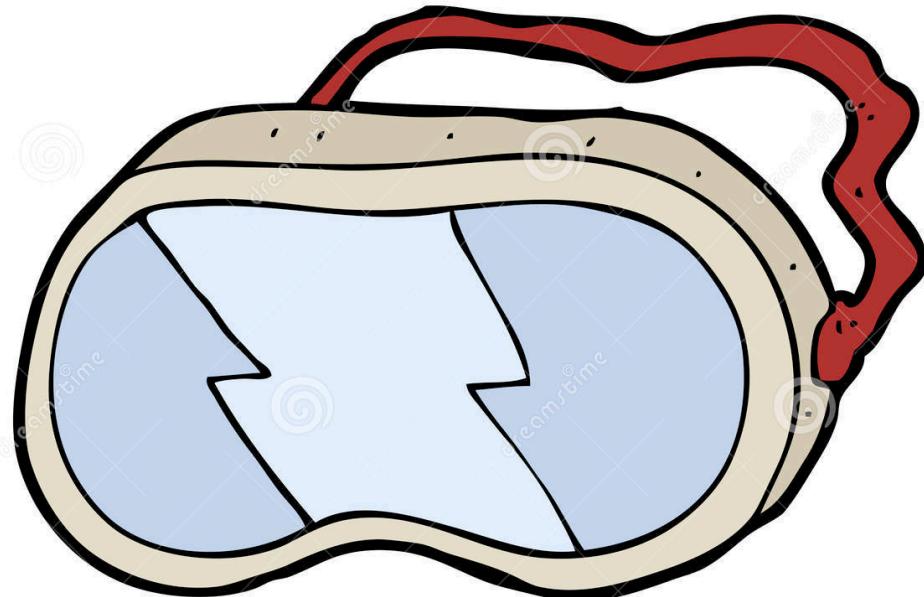
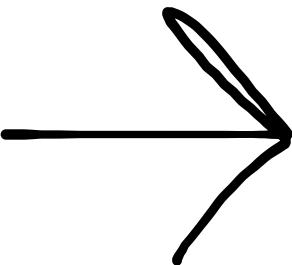
- Stats is a new area for me. I've not had much training or experience.
- Stats is an area I have some familiarity with basic concepts like correlation and summary statistics
- Stats is an area that was central to my previous work and I have high comfort with stats material

Module 2: Synch Lecture

PRINCIPLES OF VISUALIZATION



DATA VIZ GOGGLES



Edward Tufte

- Hates powerpoint
 - Thinks reports are the way to go
- BUT...
- Tufte uses a presentation format
 - Tufte typically presents his theories to large audiences (400+) in a one-way format

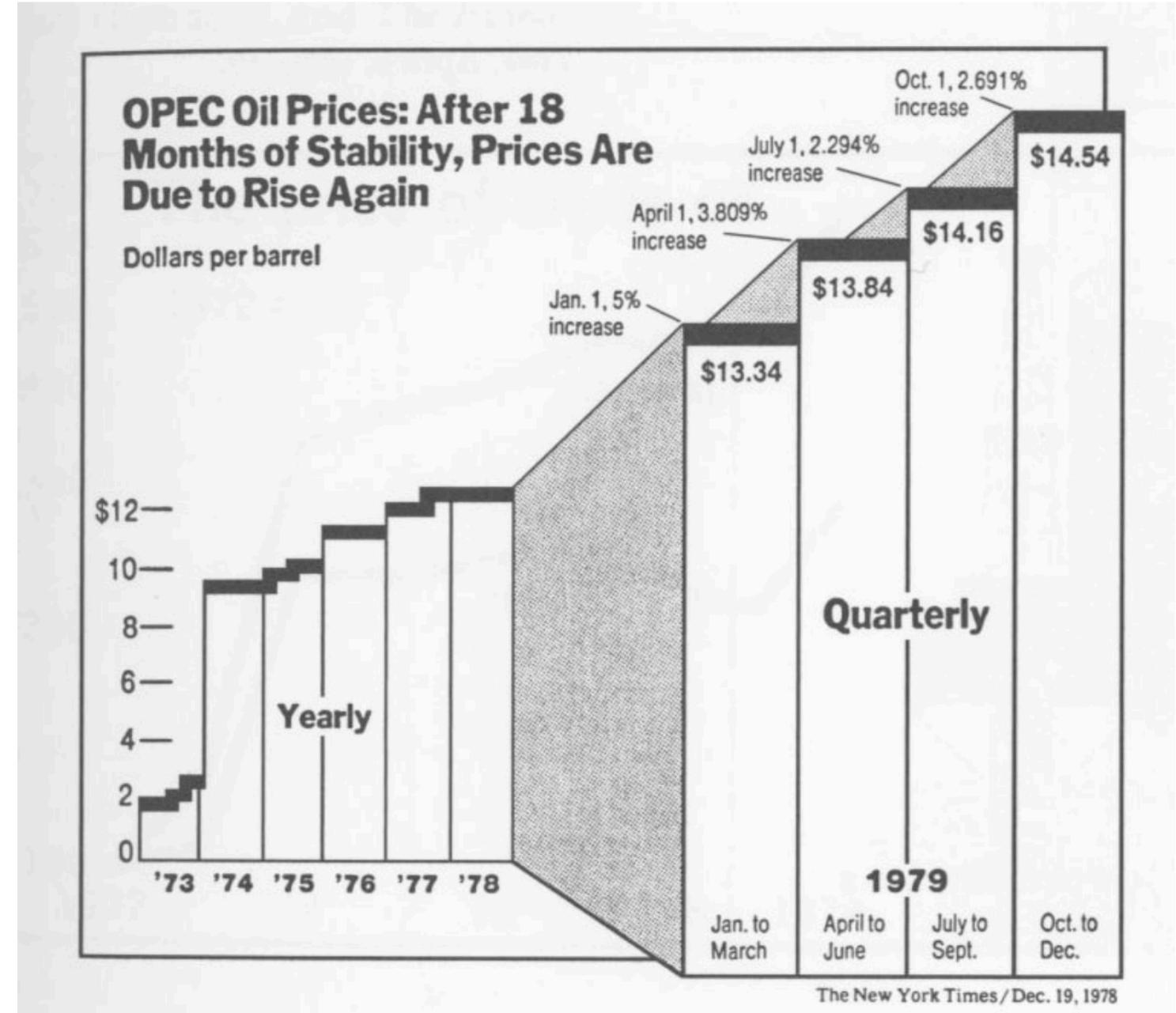


Principles of Graphical Integrity

1. Above all else show the data.
2. Maximize the data-ink ratio/erase non-data ink.
3. Revise and edit.

Above all else,
show
the data

Show data not design variation



Maximize the data ink ratio

- the proportion of a graphic's ink devoted to the non-redundant display of data information

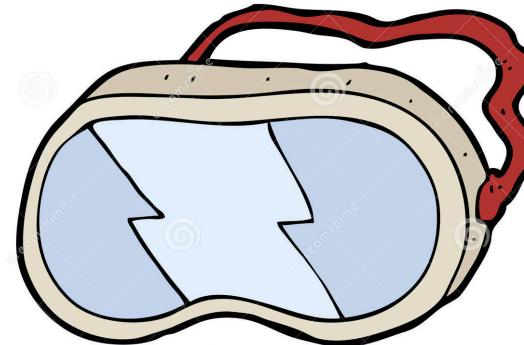
$$\text{Data ink ratio} = \frac{\text{data ink}}{\text{total ink}}$$

Context + “Chartjunk”

- **Vibrating chartjunk:** cross-hatching or patterns that distract the mind from the key insights
- **Grids:** make them a lighter gray, not black, if you are going to use them
- **Self-promoting graphics** where color schemes and patterns are introduced for artistic appeal rather than information content.

The Elements of Graphing Data (1985)

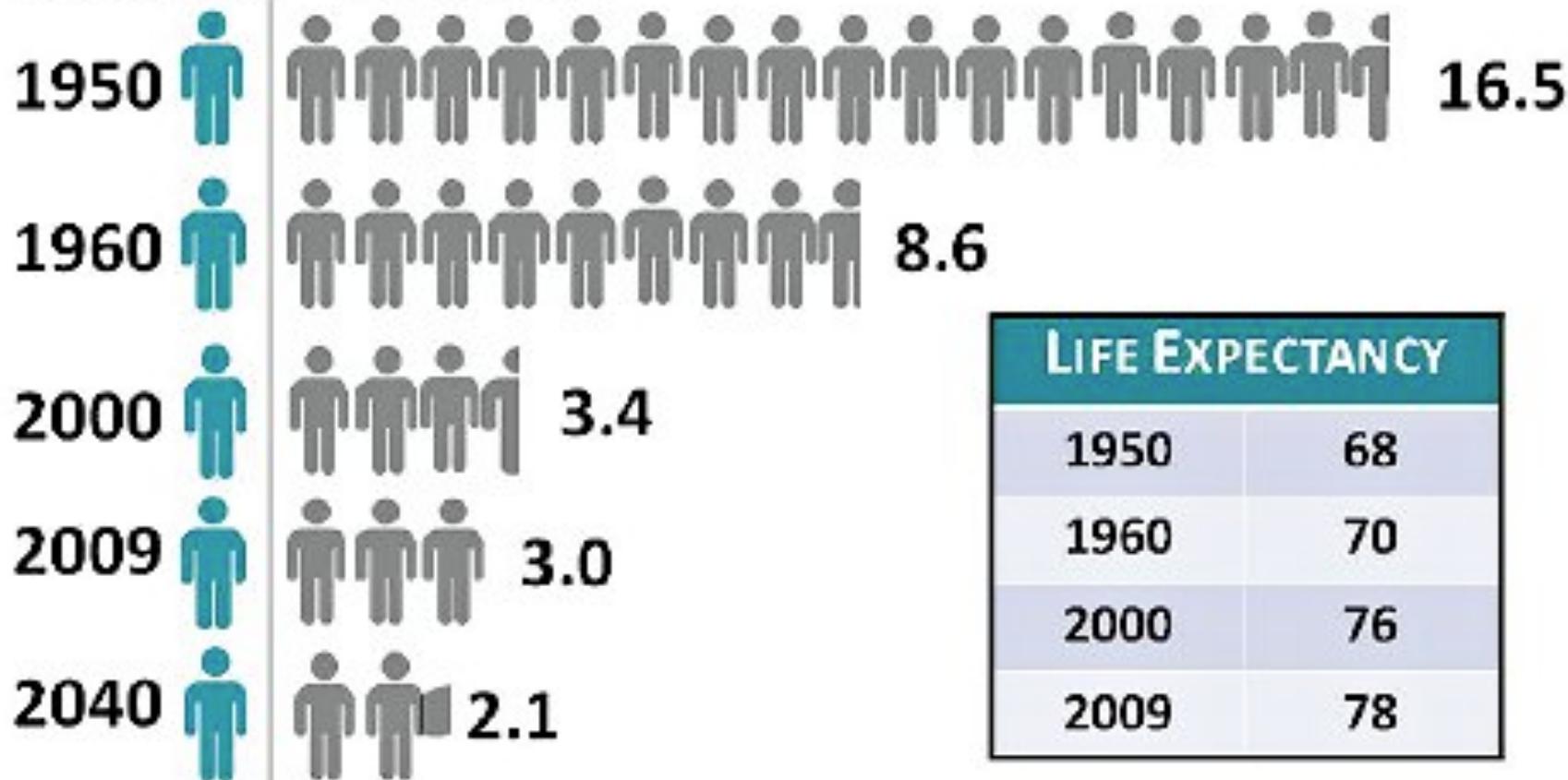
- Foundation of “Grammar of Graphics”
- Leland Wilkinson
- Preceded Tufte’s work
- Four major categories in the principles of graph construction:
 1. Clear vision
 2. Clear understanding
 3. Scales
 4. General strategy



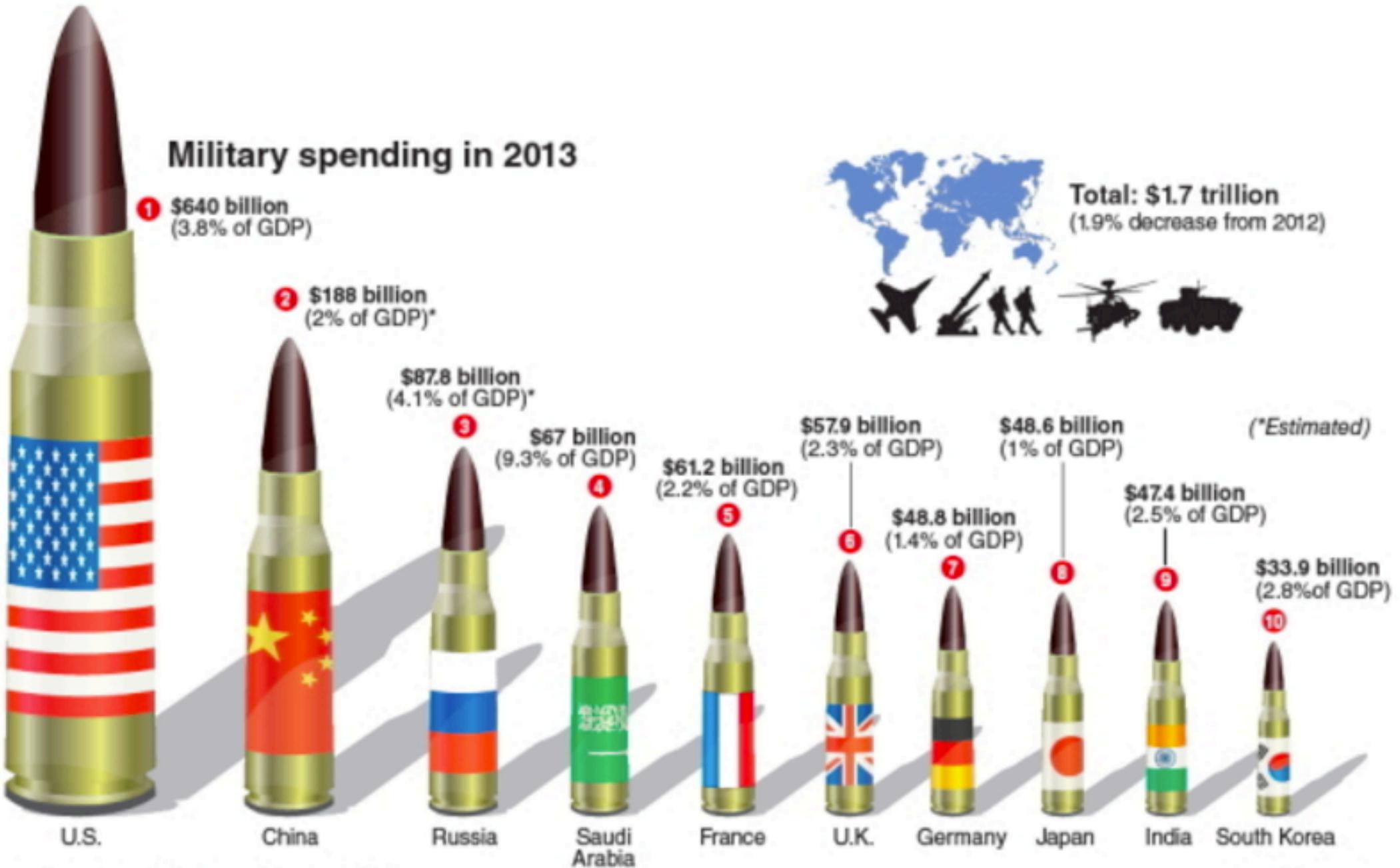
Chartjunk examples

TAXPAYERS SUPPORTING EACH SOCIAL SECURITY RECIPIENT

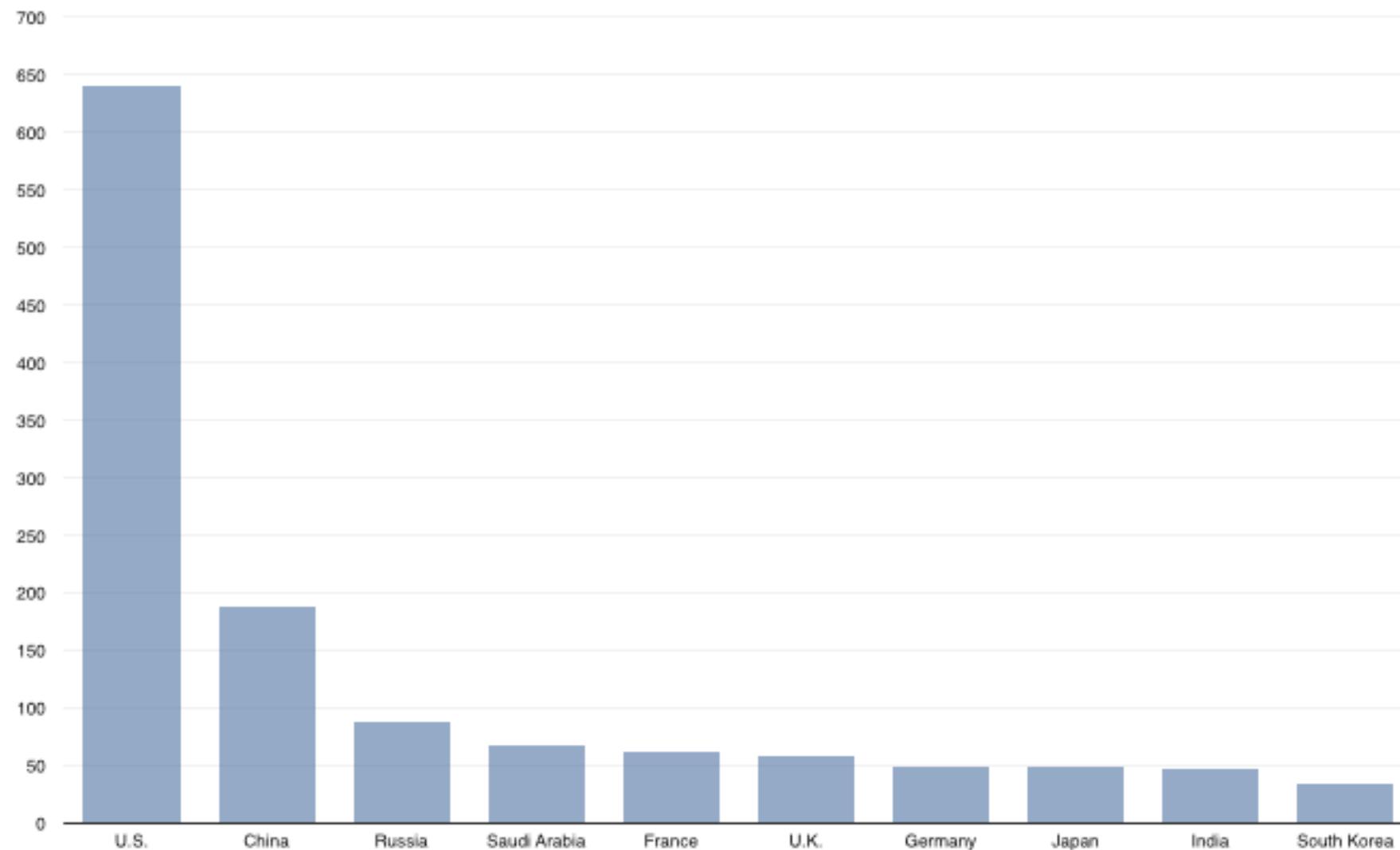
RECIPIENT WORKERS



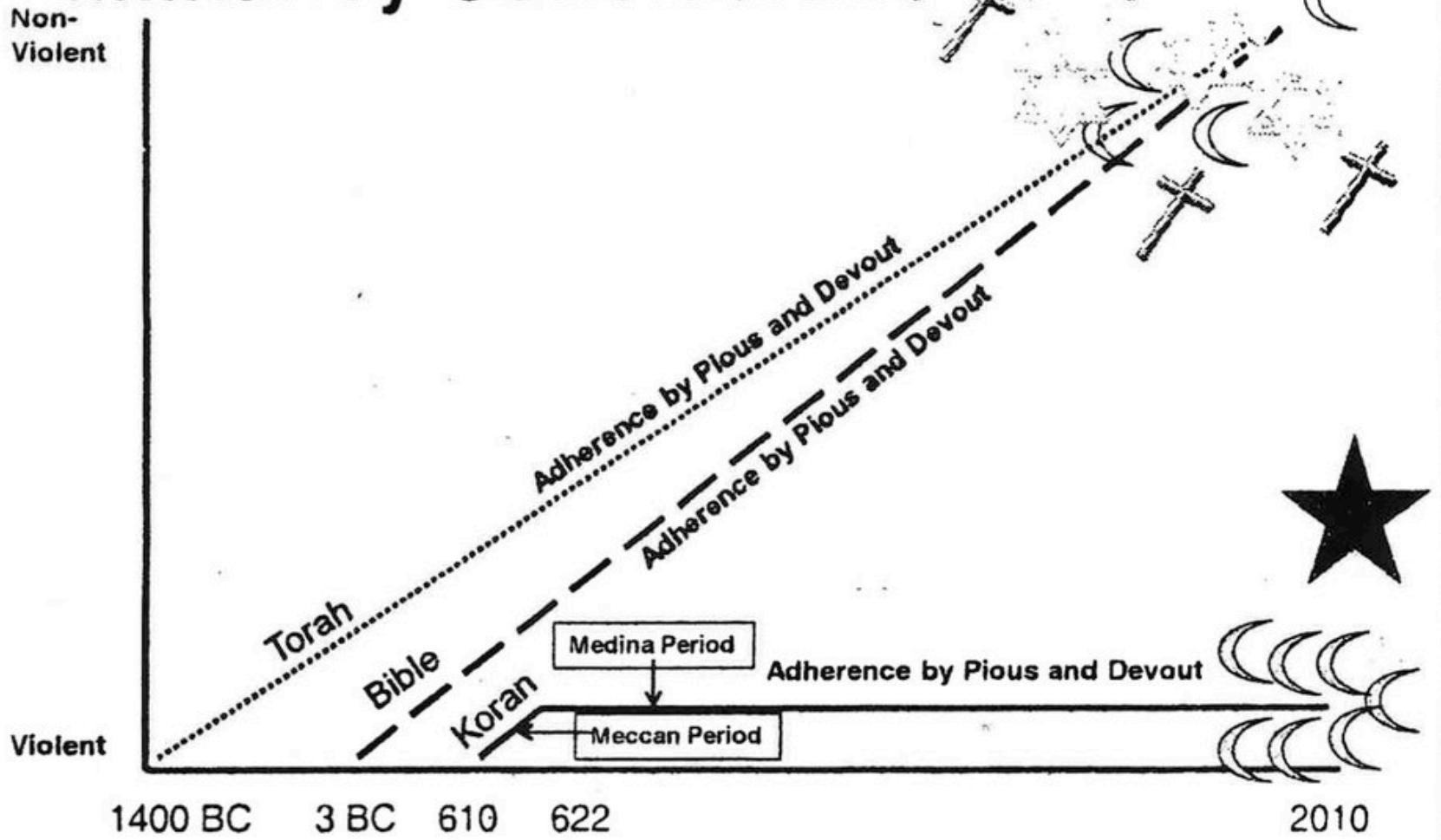
SOURCE: SOCIAL SECURITY ADMINISTRATION, THE 2010 ANNUAL REPORT OF THE BOARD OF TRUSTEES;
CENTERS FOR DISEASE CONTROL, U.S. LIFE TABLES



■ Military spending in 2013 (billion)



Militancy Considerations



SOUTHWEST BORDER APPREHENSIONS OCTOBER - APRIL



Source: U.S. Border Patrol

Happening
now

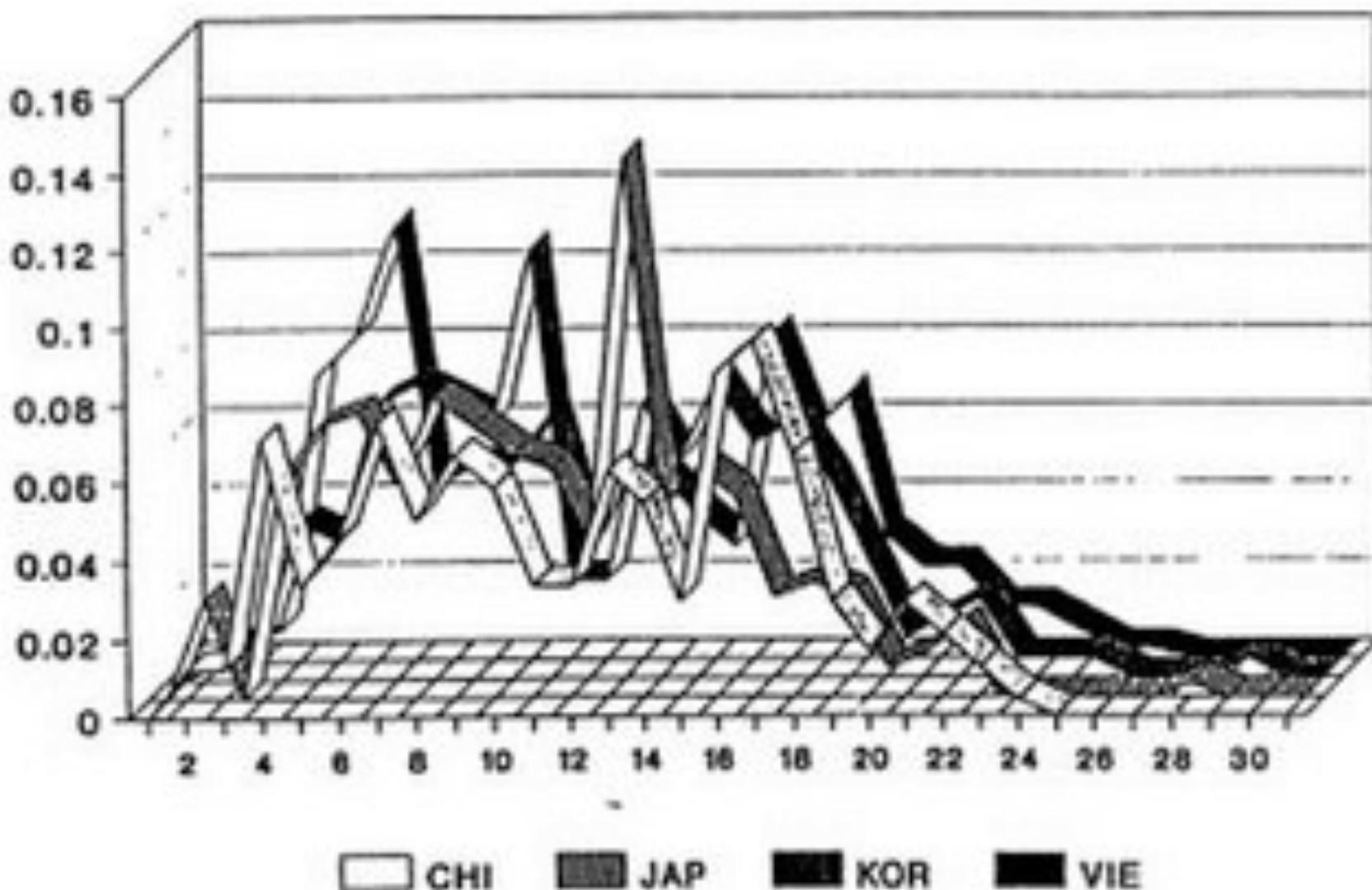
\$43 BIL IN FEB, DOWN 3.4% FROM JAN'S REVISED

NAS

35.18

B

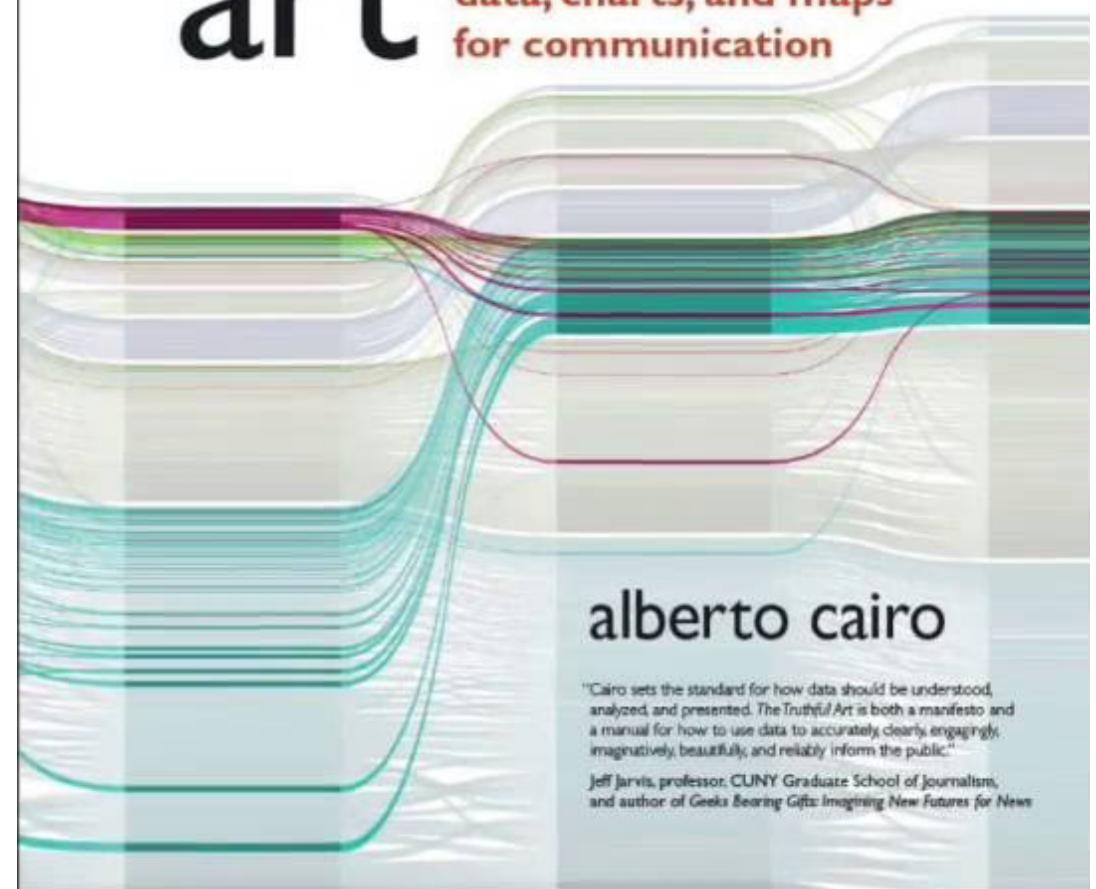
BINNED FREQUENCY DATA - D10S28 CHINESE, JAPANESE, KOREAN, VIETNAMESE



1. **It is truthful**, as it's based on thorough and honest research.
2. **It is functional**, as it constitutes an accurate depiction of the data, and it's built in a way that lets people do meaningful operations based on it (seeing change in time).
3. **It is beautiful**, in the sense of being attractive, intriguing, and even aesthetically pleasing for its intended audience—scientists, in the first place, but the general public, too.
4. **It is insightful**, as it reveals evidence that we would have a hard time seeing otherwise.
5. **It is enlightening** because if we grasp and accept the evidence it depicts, it will change our minds for the better.

the truthful art

**data, charts, and maps
for communication**



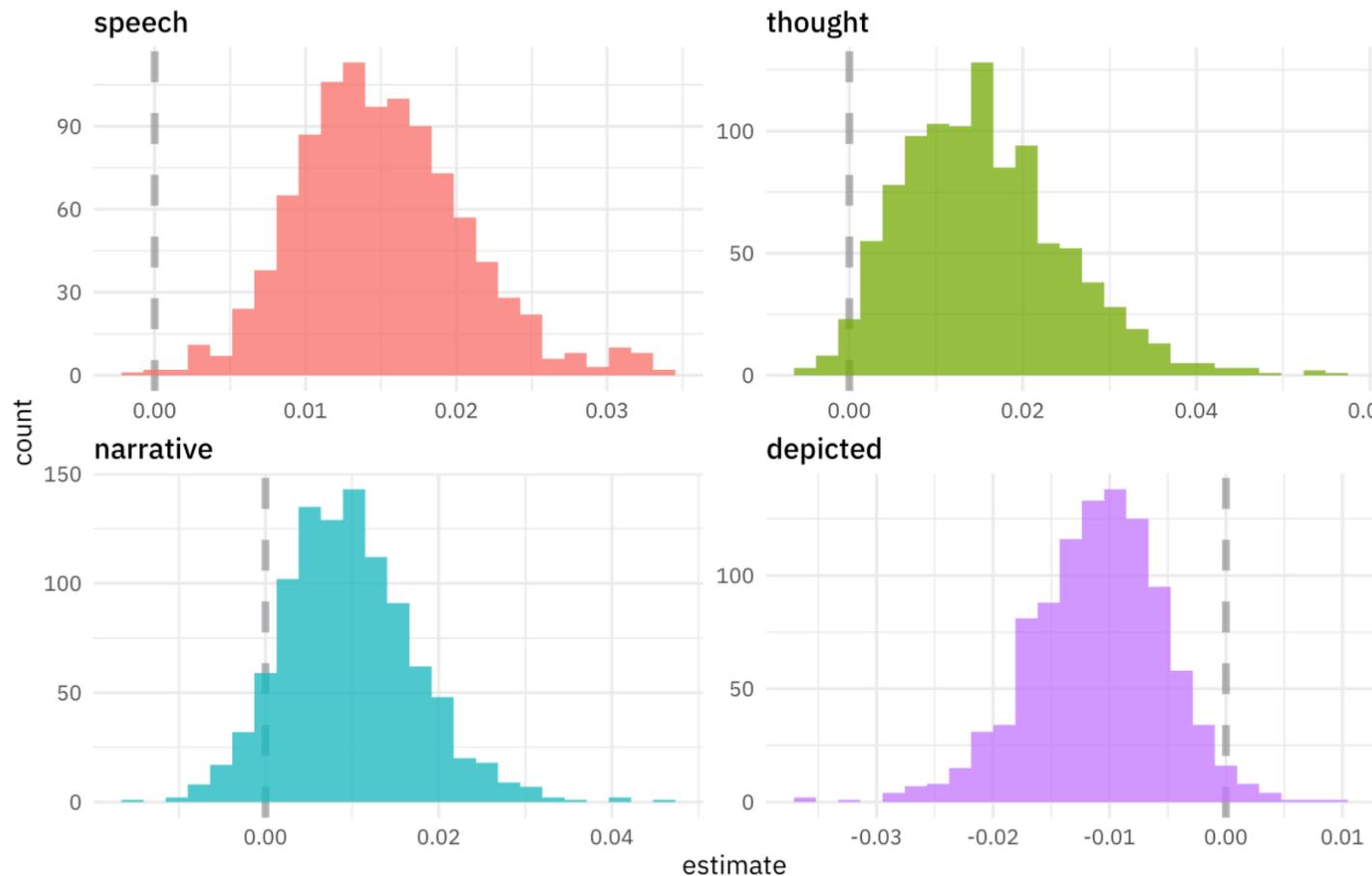
"Cairo sets the standard for how data should be understood, analyzed, and presented. *The Truthful Art* is both a manifesto and a manual for how to use data to accurately, clearly, engagingly, imaginatively, beautifully, and reliably inform the public."

Jeff Jarvis, professor, CUNY Graduate School of Journalism,
and author of *Geeks Bearing Gifts: Imagining New Futures for News*

Julia Silge (<https://juliasilge.com/blog/>)

Which Uncanny X-Men issues pass the Bechdel test?

Comparing the top 25 characters' speech, thought, narrative portrayal, and total depictions



Data from the Claremont Run Project

Your turn: Graphic Design Activity

1. Get into groups
2. Imagine how you would “rework” each chart/graphic from the Economist. Have one person share their screen.

Discuss:

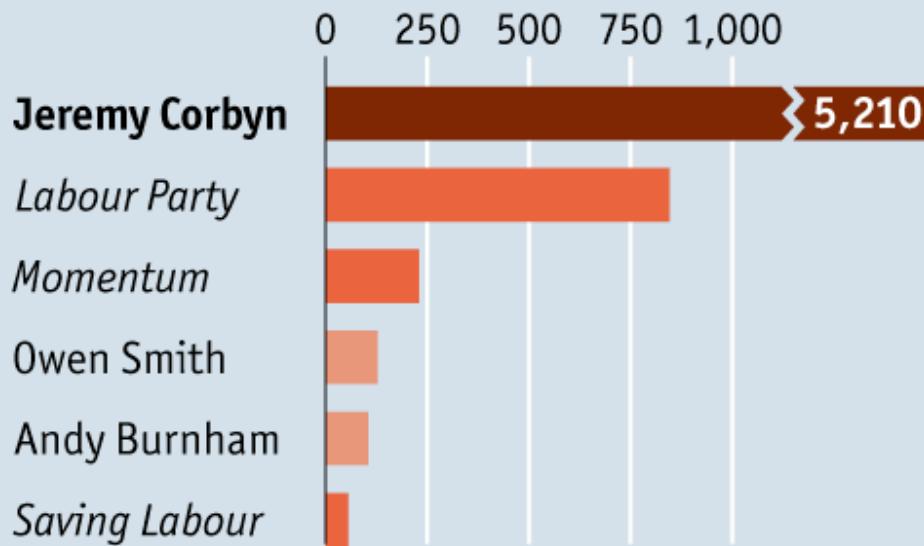
- Is it misleading?
- Is it confusing?
- What would you change?

3. Take notes
4. Once done, brainstorm a set of 3-5 “golden rules” for viz goggles of your own. Write them on a shared google doc/document or slack channel.

Original

Left-click

Average number of likes per Facebook post
2016

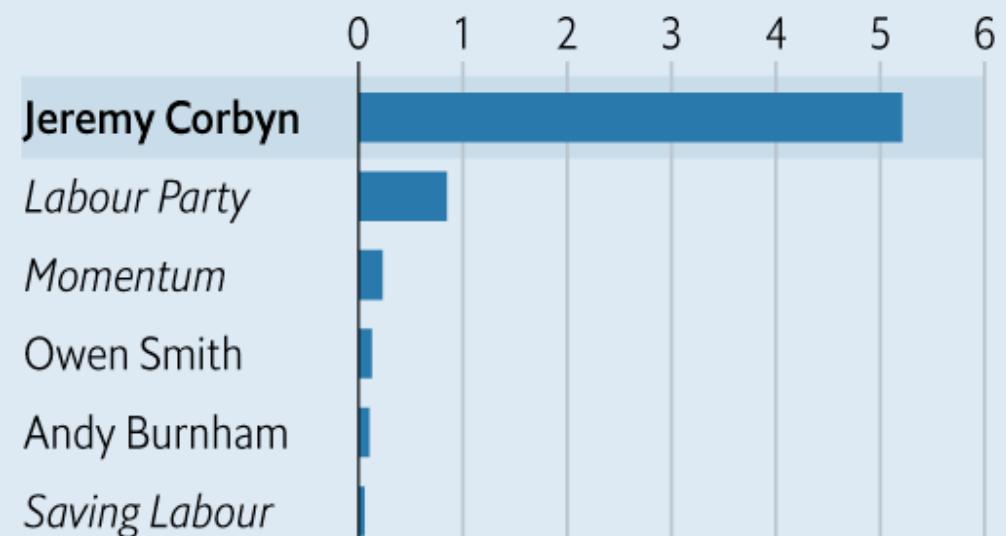


Source: Facebook

Better

Left-click

Average number of likes per Facebook post
2016, '000



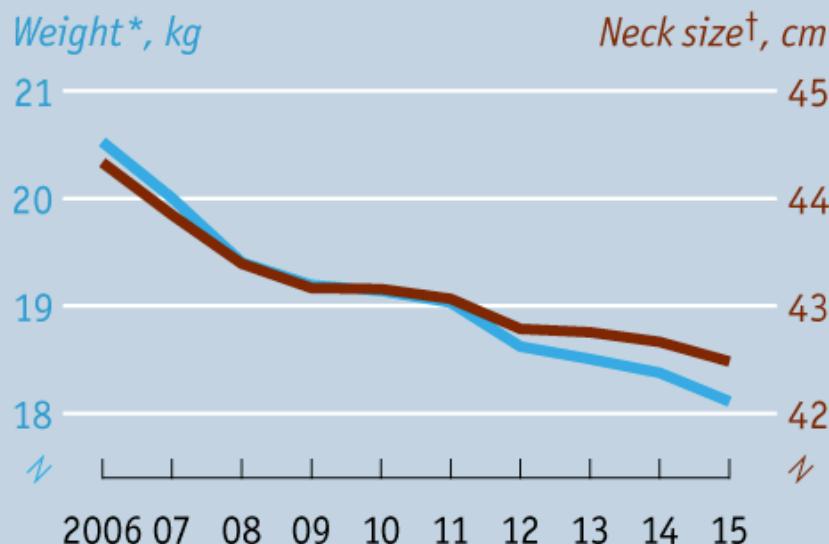
Source: Facebook

MISLEADING CHARTS (1)

Original

Fit as a butcher's dog

Characteristics of dogs registered with the UK's Kennel Club, average when fully grown



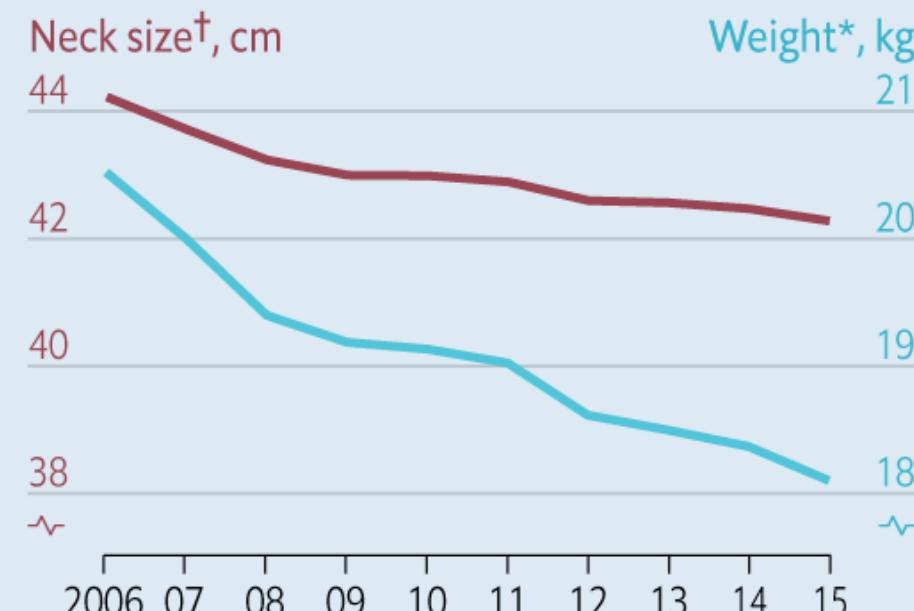
Sources: Kennel Club;
The Economist

*Where at least 50 are
registered per year †Where at
least 100 are registered per year

Better

Fit as a butcher's dog

Characteristics of dogs registered with the UK's Kennel Club, average when fully grown



Sources: Kennel Club;
The Economist

*Where at least 50 are registered per year
†Where at least 100 are registered per year

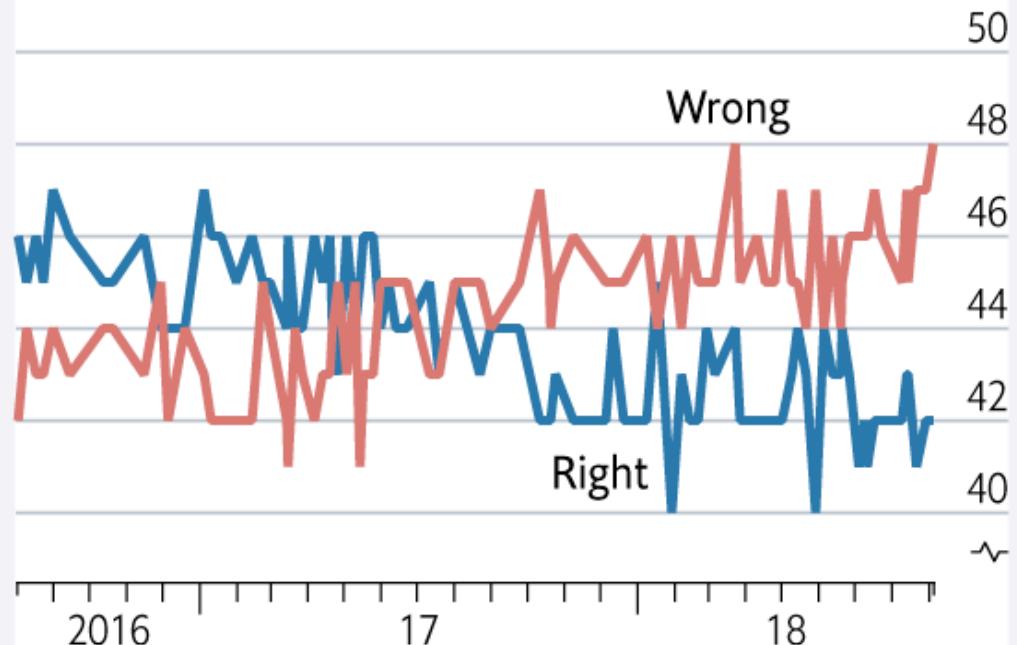
MISLEADING CHARTS (2)

Original

Bremorse

"In hindsight, do you think Britain was right or wrong to vote to leave the EU?"

% responding



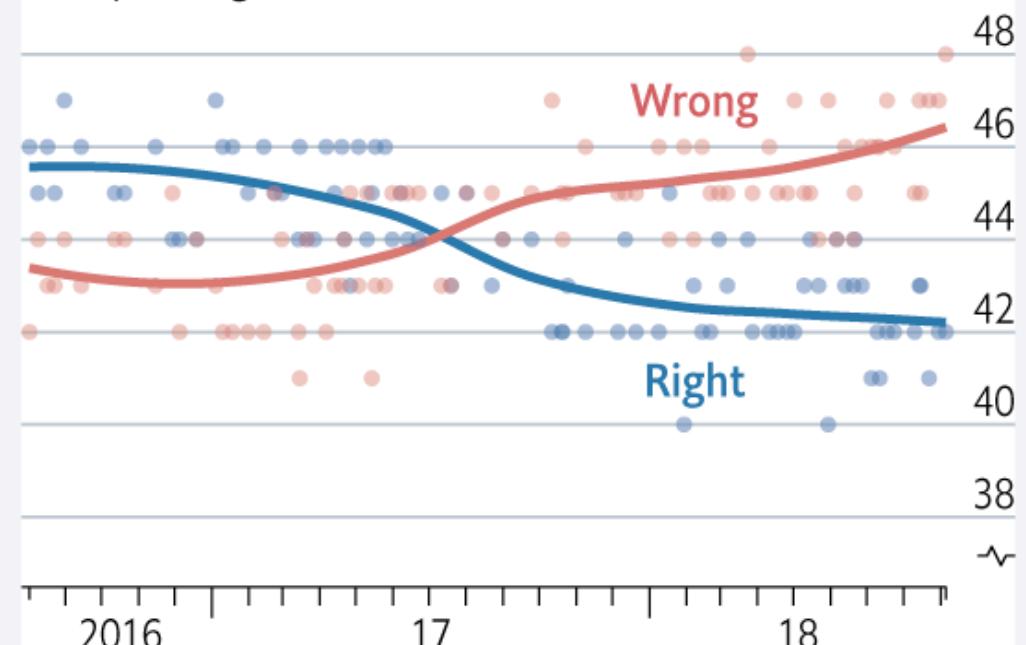
Source: NatCen Social Research

Better

Bremorse

"In hindsight, do you think Britain was right or wrong to vote to leave the EU?"

% responding



Source: NatCen Social Research

MISLEADING CHARTS (3)

Original



Better

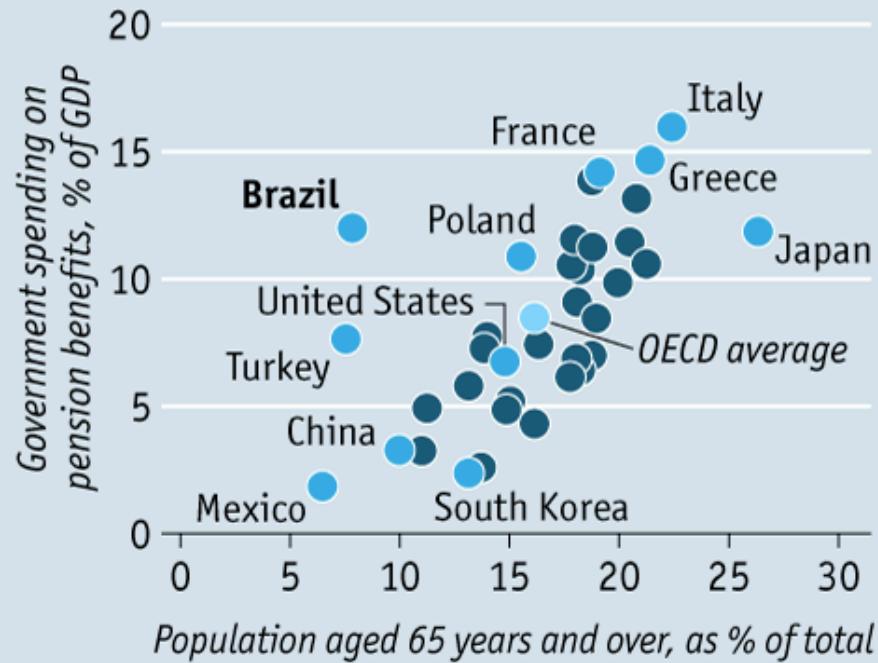


CONFUSING CHARTS (4)

Original

Brazil's golden oldie blowout

Latest available

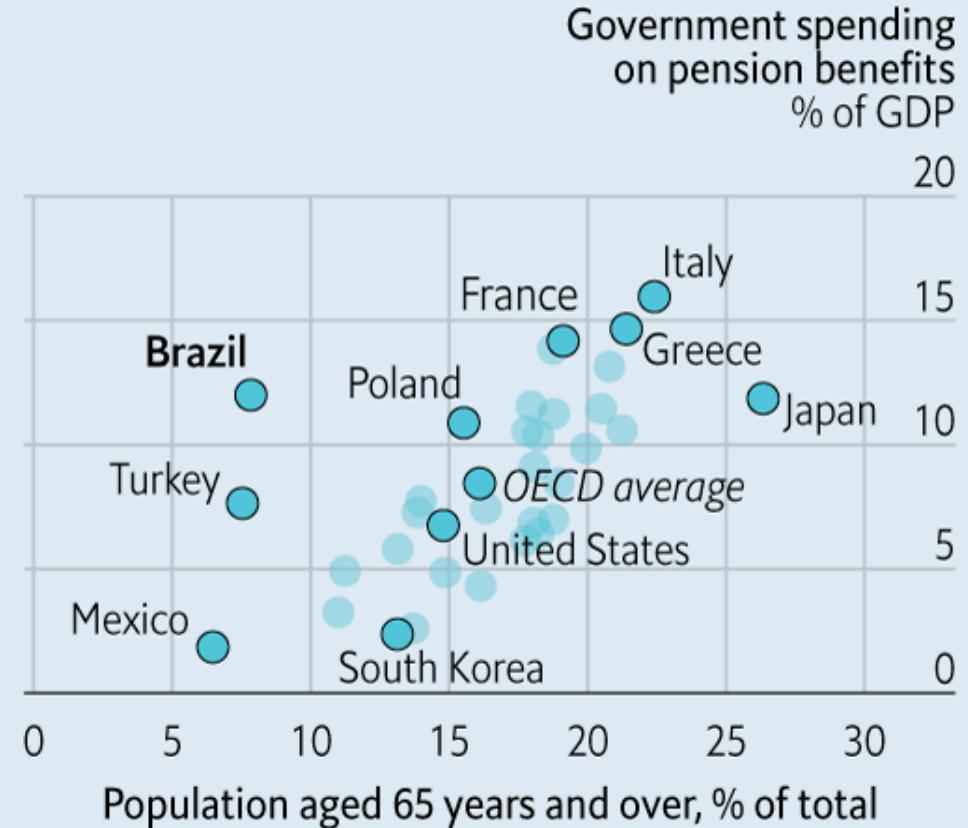


Sources: OECD; World Bank; Previdência Social

Better

Brazil's golden oldie blowout

Latest available



Sources: OECD; World Bank; Previdência Social

CONFUSING CHARTS (5)

Original

Surfeit of surpluses

Selected euro-area countries, €bn



Budget balances



Sources: Eurostat; National statistics

Current-account balances



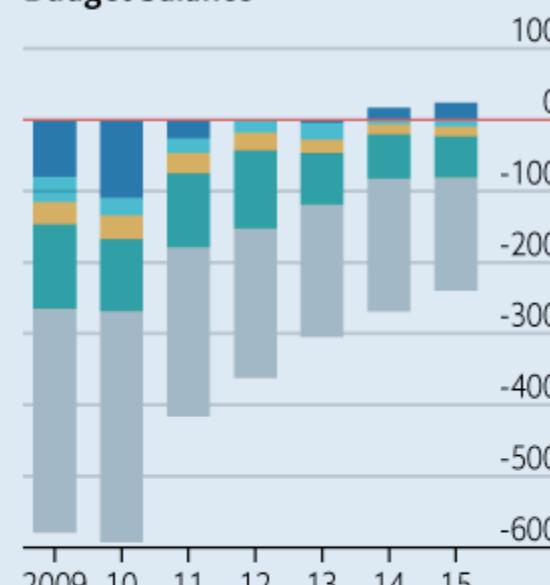
Better

Surfeit of surpluses

Euro-area, €bn



Budget balance



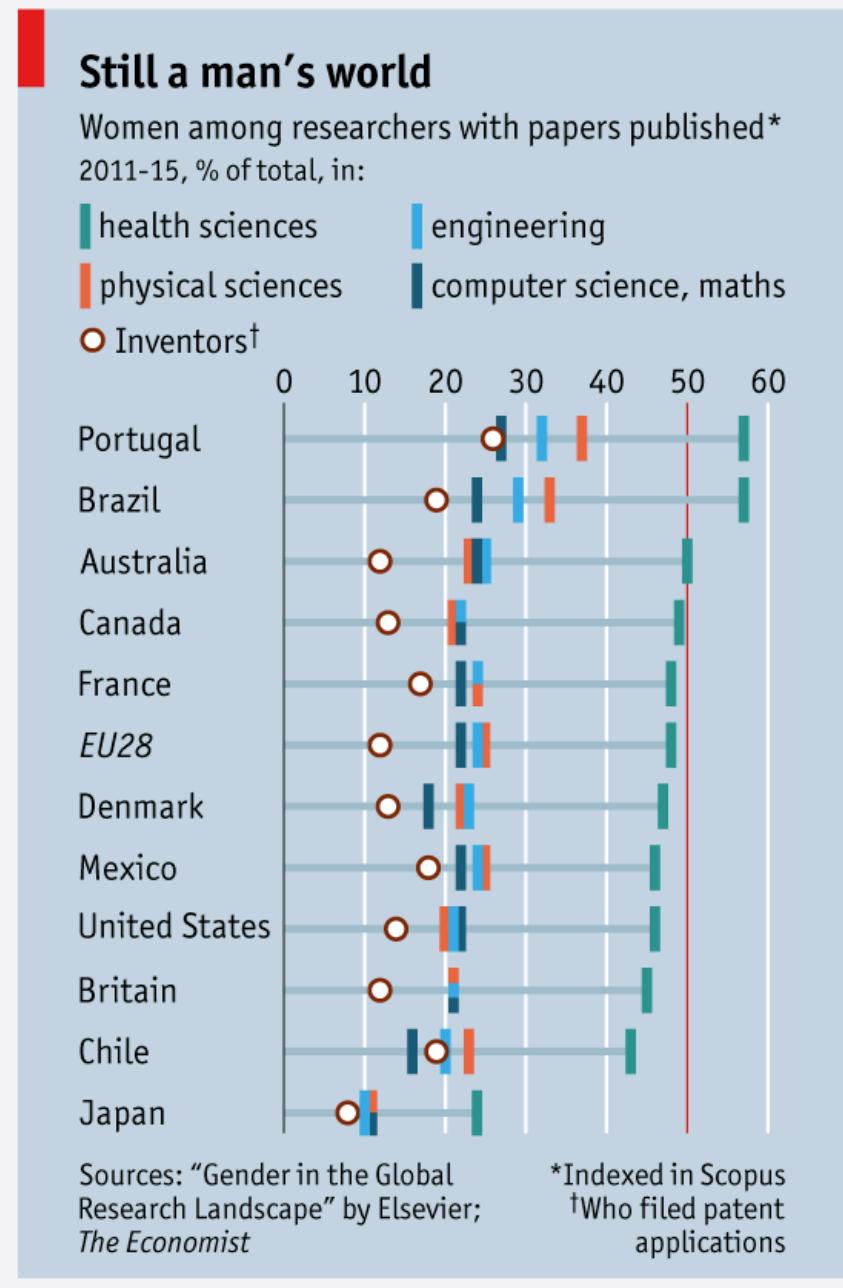
Source: Eurostat

Current-account balance



Charts that conceal their point (6)

Original



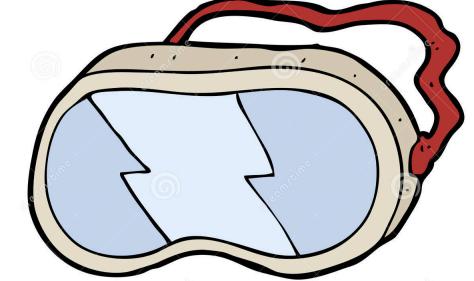
Better



(7)

What would you program
your goggles to look for?

What are some of your
'golden rules'?



In groups, discuss each graphic from the economist.

What would you change?

How could it be improved?

Then as a group decide on 3-5 'golden rules' to
remember about data visualization best practices