

What is Visualization

Data Management and Visualization



SoftEng
<http://softeng.polito.it>

Version 2.4.0
© Marco Torchiano, 2022








This work is licensed under the Creative Commons Attribution–NonCommercial–NoDerivatives 4.0 International License.

To view a copy of this license, visit

<http://creativecommons.org/licenses/by-nc-nd/4.0/>.


You are free: to copy, distribute, display, and perform the work

Under the following conditions:

-  **Attribution.** You must attribute the work in the manner specified by the author or licensor.
-  **Non-commercial.** You may not use this work for commercial purposes.
-  **No Derivative Works.** You may not alter, transform, or build upon this work.
 - For any reuse or distribution, you must make clear to others the license terms of this work.
 - Any of these conditions can be waived if you get permission from the copyright holder.

Your fair use and other rights are in no way affected by the above.

Topics

- Visualization literacy
 - ◆ Visual perception
 - ◆ Graph design
- Visualization skill
 - ◆ Tool:  **+ a b | e a u**
 - ◆ Practice with different visualization problems and graph types

Exam [5 points + 1.5* Theory]

- Assessment
 - ◆ Question [0.25*]
 - ◆ Data [1.25*]
 - ◆ Visual
 - Proportionality [0.75]
 - Utility [0.75]
 - Clarity [0.5]
- Redesign [0.25* + 1.25]

* Multiple choice questions

Definition

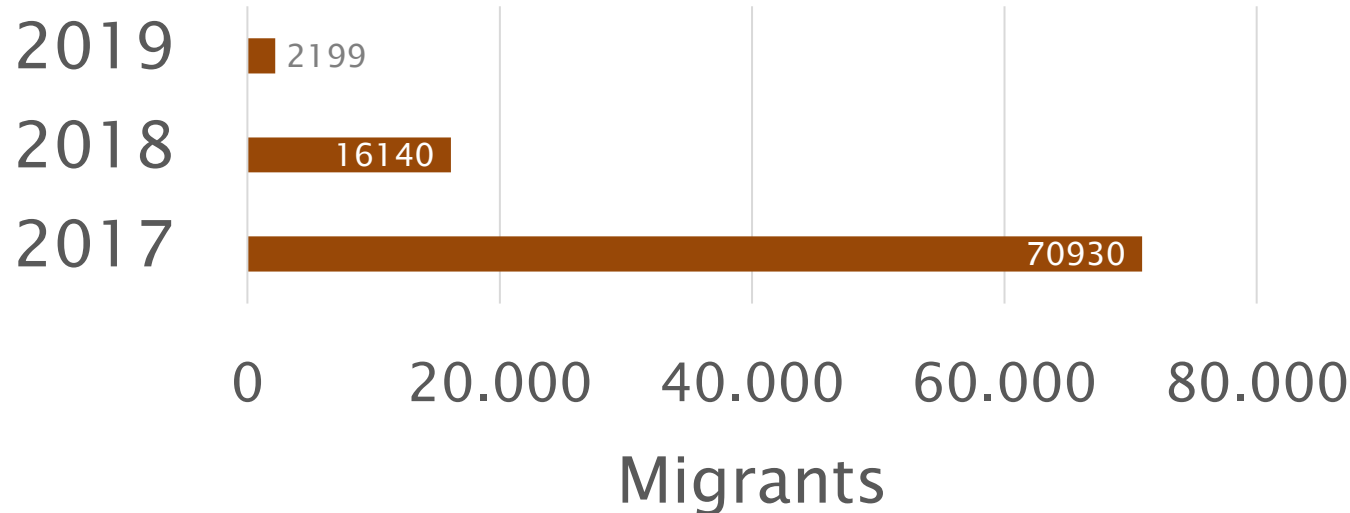
Visualization:

Usage of visual features to
encode data in order to
convey useful information



WHY VISUALIZATION?

Migrants arrived in period January – June



http://www.interno.gov.it/sites/default/files/cruscotto_statistico_giornaliero_19-06-2019.pdf

The **accidents at work** happened and reported to Inail in first quarter 2019 have been 131 thousand (109 thousand at work and 22 thousand while traveling), on the rise by 1,7% (+2 thousand reports) with respect to first quarter 2018

https://www.istat.it/it/files/2019/06/NotaTrimestrale-Occupazione-I_2019.pdf

Motivation

Information retrieval

- After 3 days
 - ◆ Text alone: 10%
 - ◆ Text + visuals: 65%

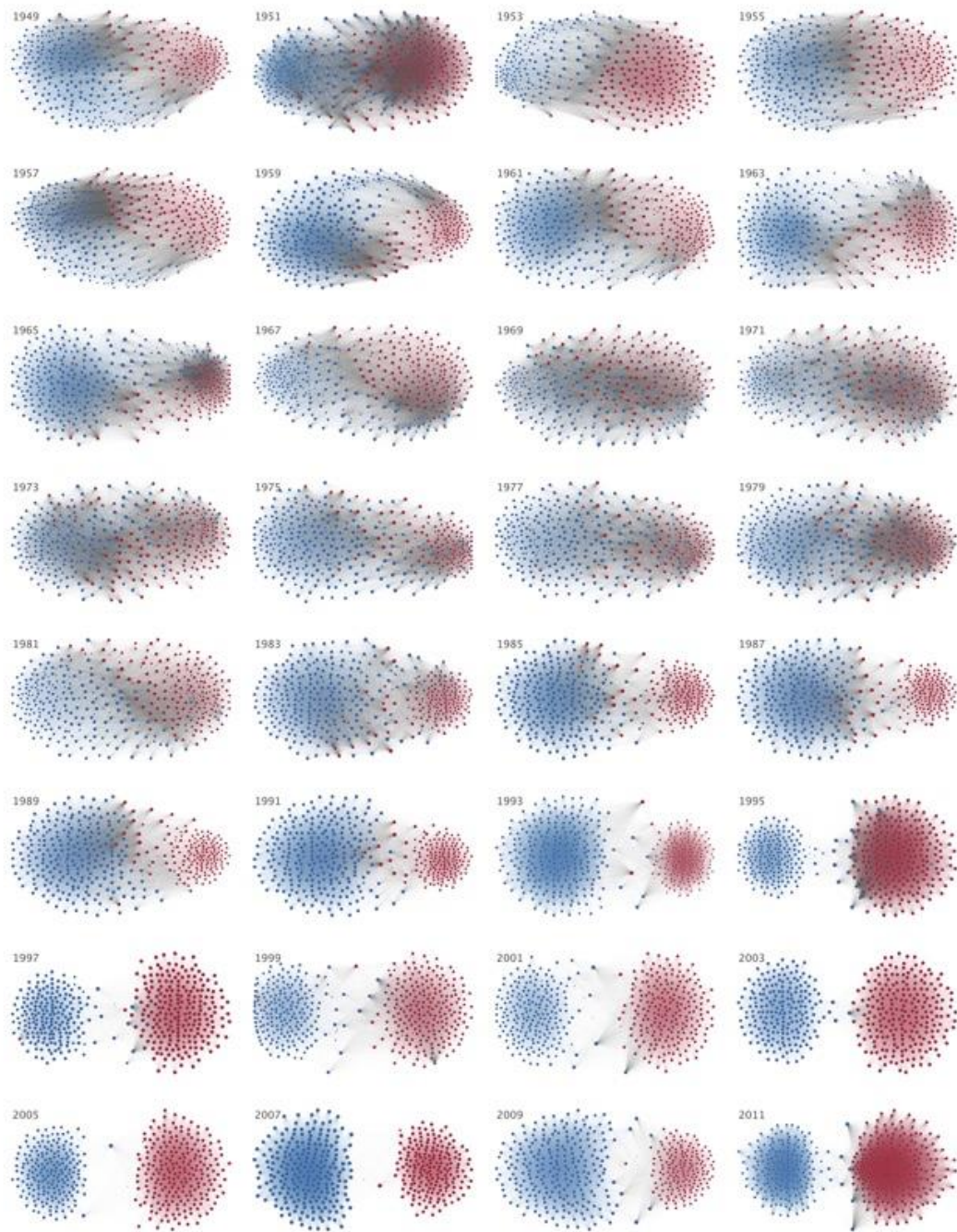
[John Medina, Brain Rules, 2008]

Motivation

Information retrieval

Information density

- In principle every single pixel in an image could encode a datum
 - ◆ Screen (1024x768) ~ 1 M pixels
 - ◆ 1 M characters ~ 250 pages



Motivation

Information retrieval

Information density

Information context

Visualization compares multiple values and puts the information into context.

A single number means nothing.

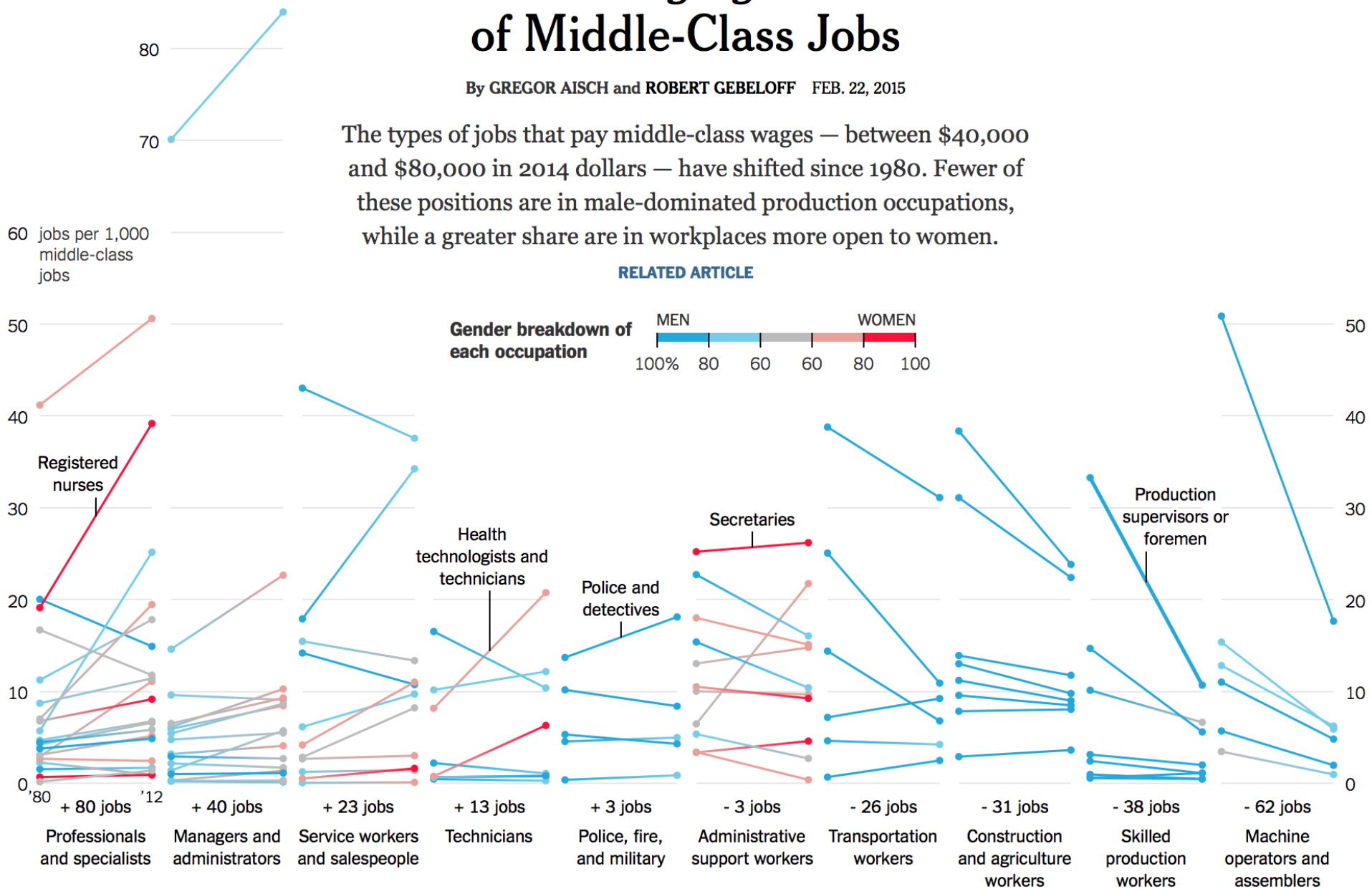
[Randy Krum presentation at Malofiej 23 (March 2015)]

The Changing Nature of Middle-Class Jobs

By GREGOR AISCH and ROBERT GEBELOFF FEB. 22, 2015

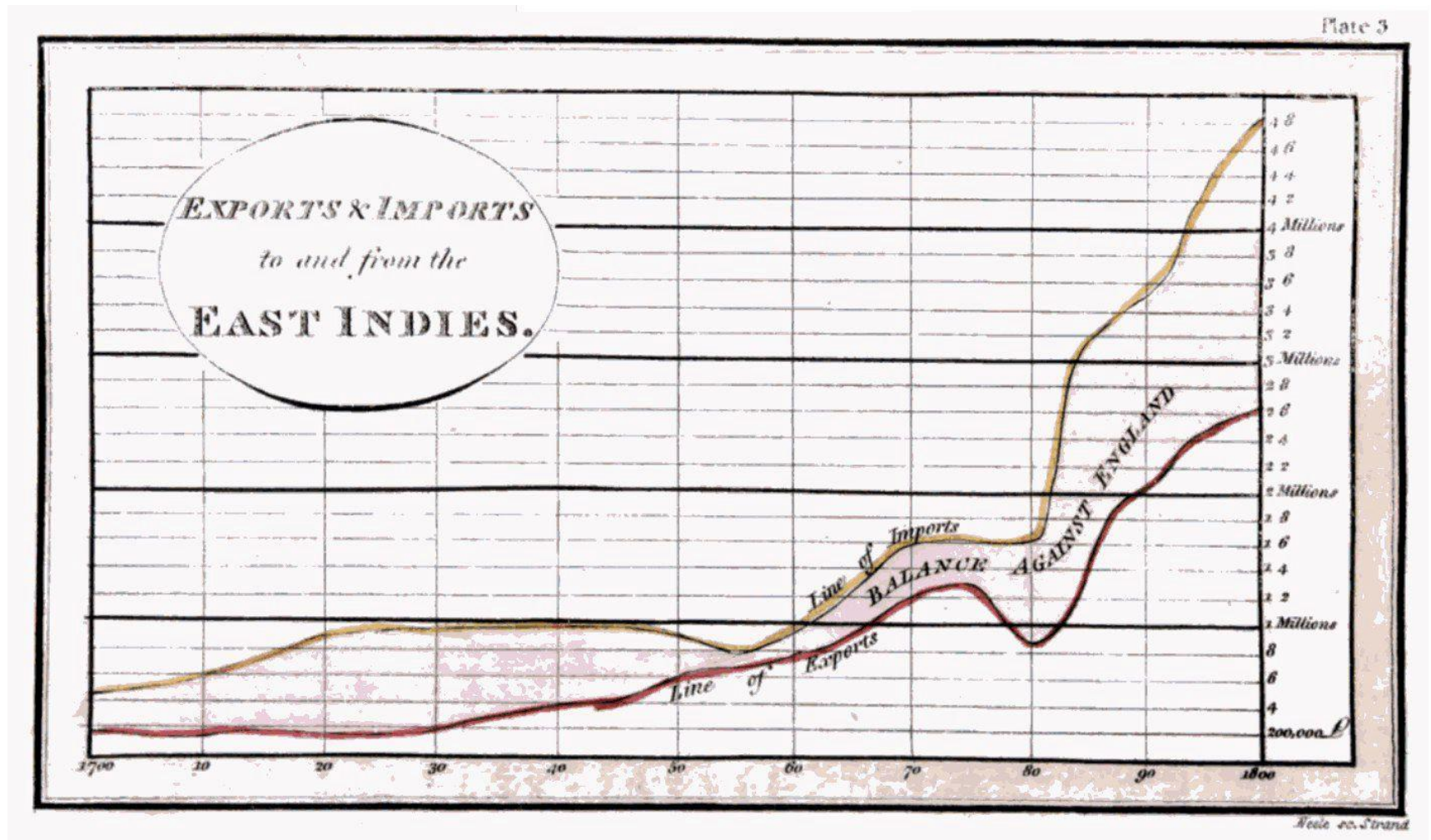
The types of jobs that pay middle-class wages — between \$40,000 and \$80,000 in 2014 dollars — have shifted since 1980. Fewer of these positions are in male-dominated production occupations, while a greater share are in workplaces more open to women.

RELATED ARTICLE



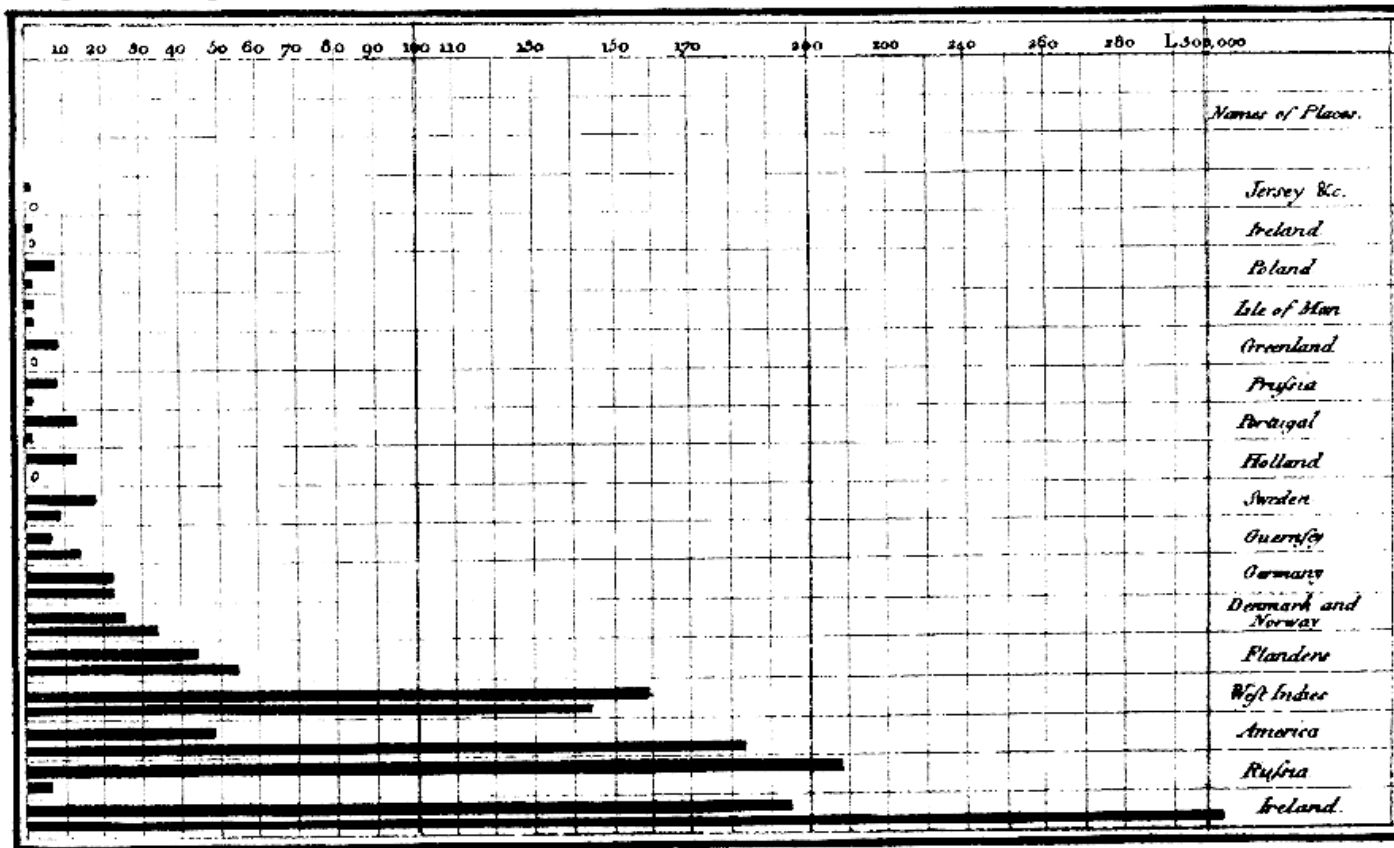
HISTORY

William Playfair



W.Playfair, The Commercial and Political Atlas, London 1786

Exports and Imports of SCOTLAND to and from different parts for one Year from Christmas 1780 to Christmas 1781.



The Upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Ribbed Lines Imports.

Published as the Act directs, June 7th 1786 by W^m Playfair

Printed by J. Smith, Strand, London.

W.Playfair, The Commercial and Political Atlas, London 1786

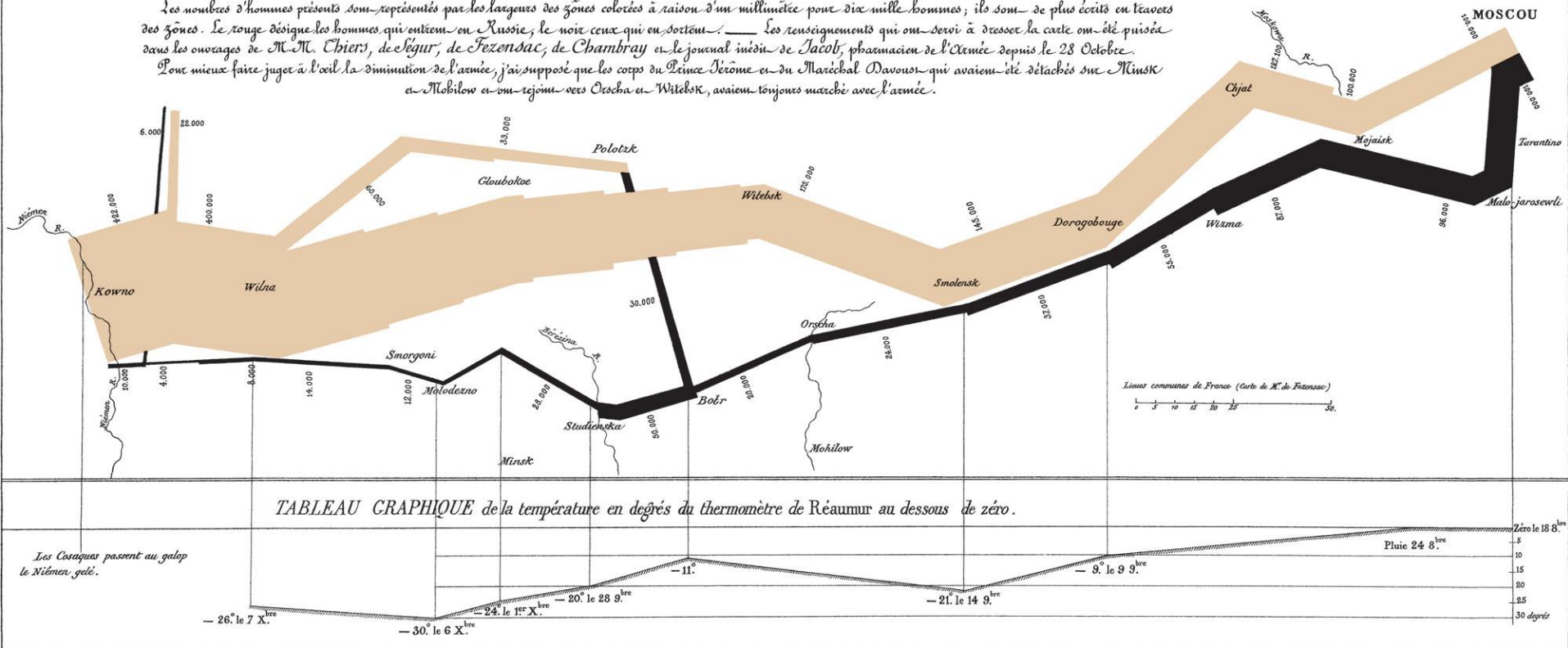
Charles Joseph Minard

Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.

Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite. Paris, le 20 Novembre 1869.

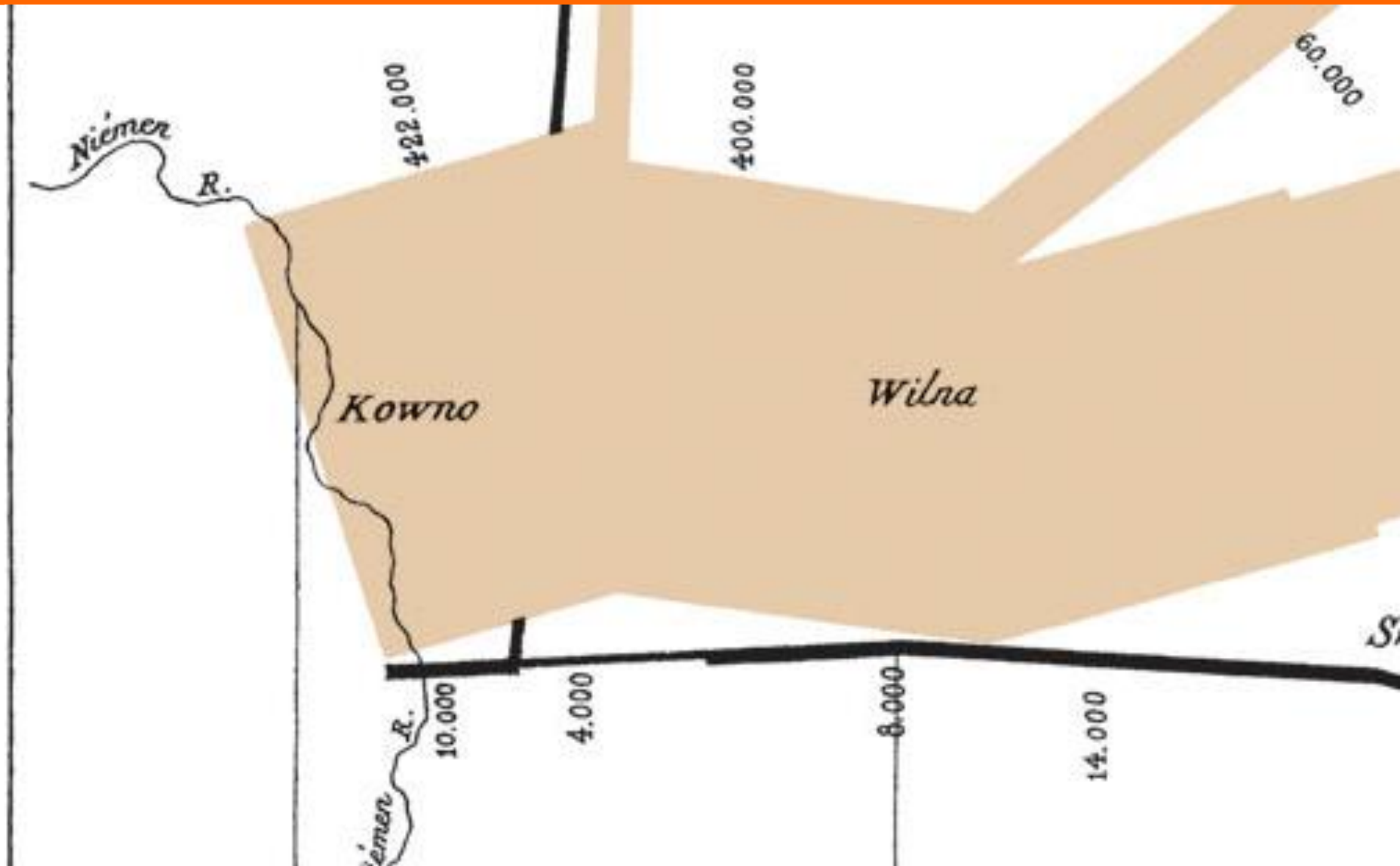
Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Ségur, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk et Mohilow et qui rejoignent vers Orscha et Witebsk, avaient toujours marché avec l'armée.

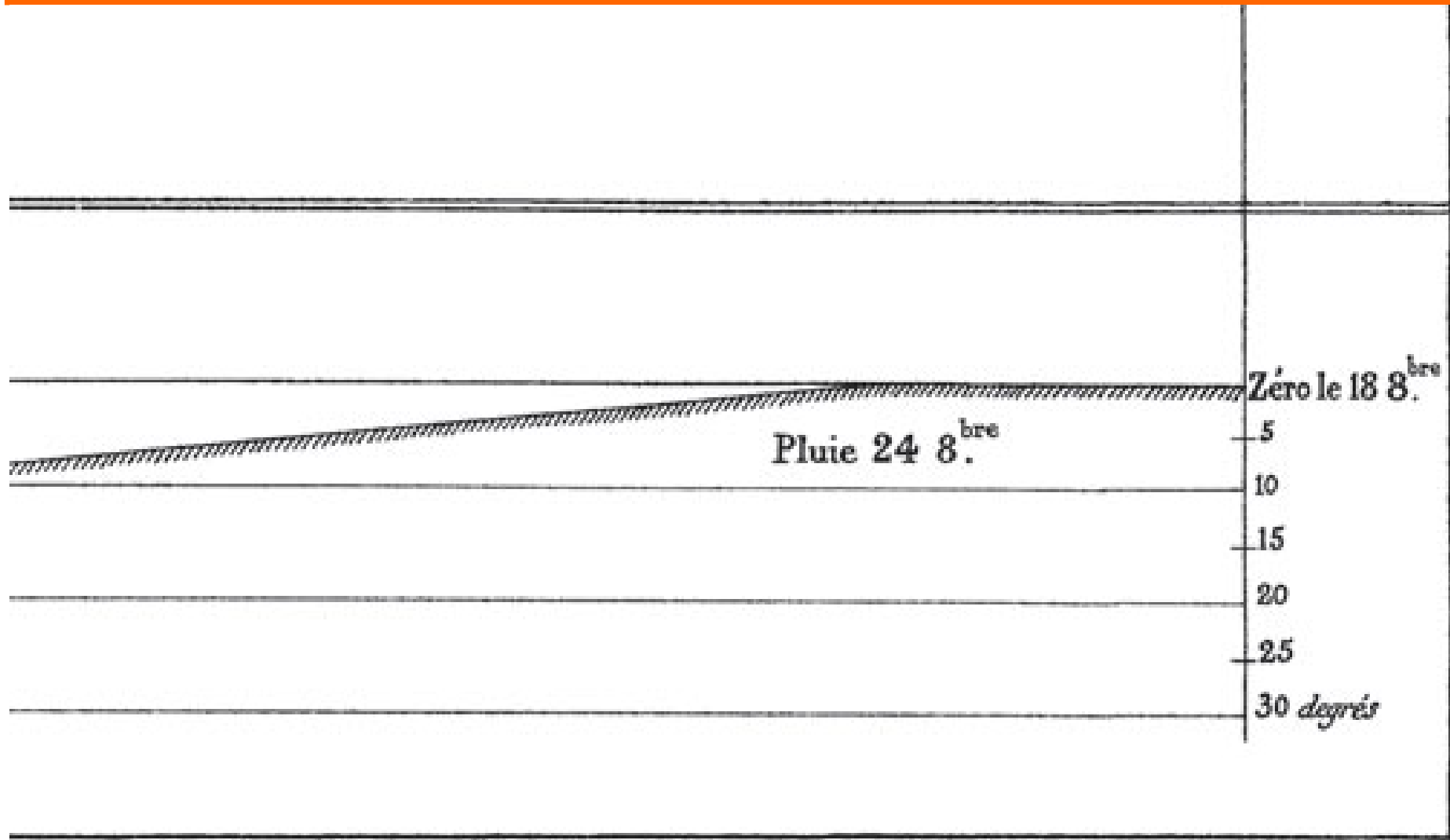


https://en.wikipedia.org/wiki/Charles_Joseph_Minard

Numbers and direction

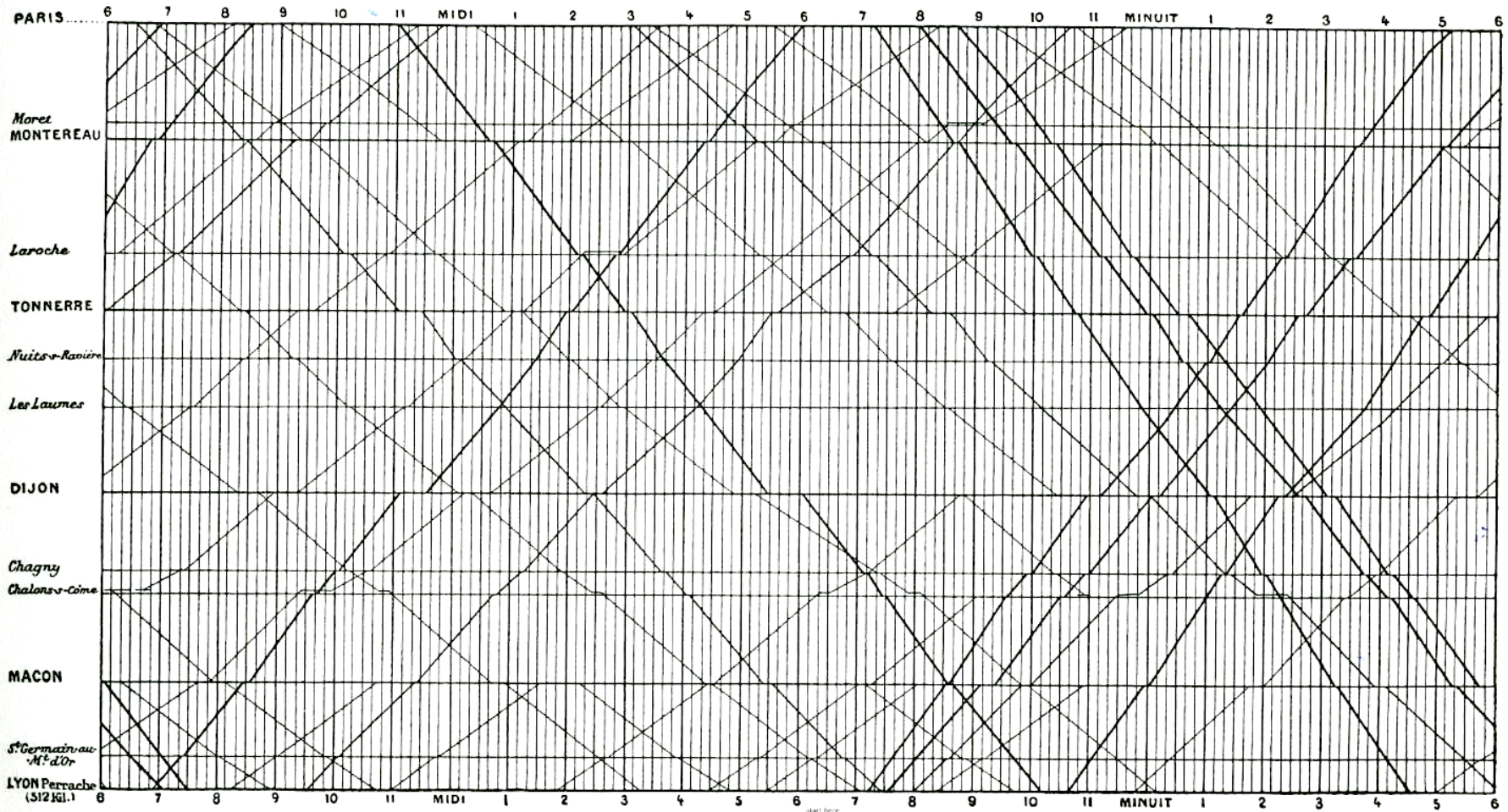


Temperature



Imp. Lith. Regnier et Dourdet.

Étienne-Jules Marey

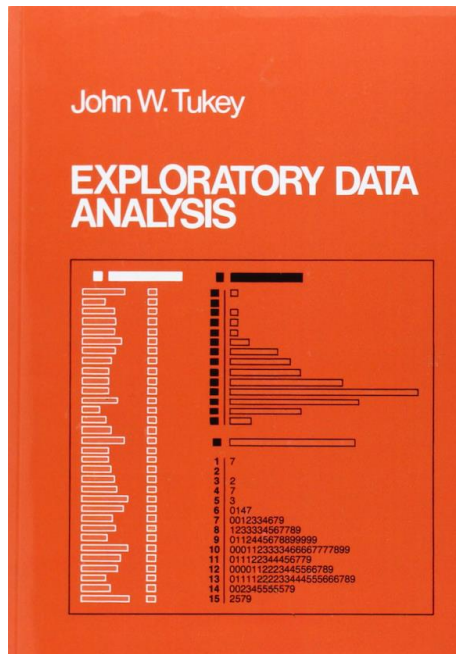


La Méthode graphique dans les sciences expérimentales et principalement en physiologie et en médecine, 1885
<https://archive.org/details/lamthodegraphiq00maregoog>

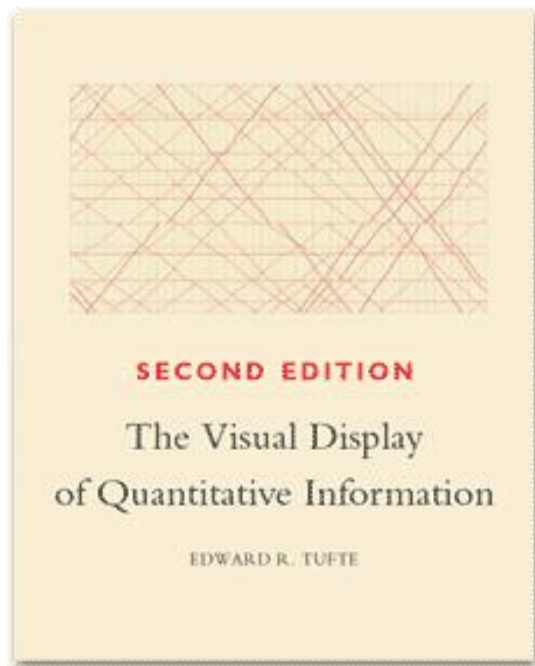
XX Century

- <http://www.datavis.ca/milestones/>

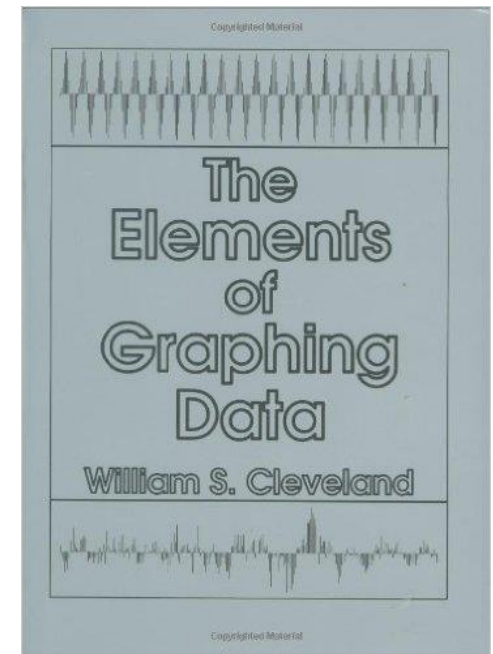
1977



1983



1985



INFORMATION VISUALIZATION

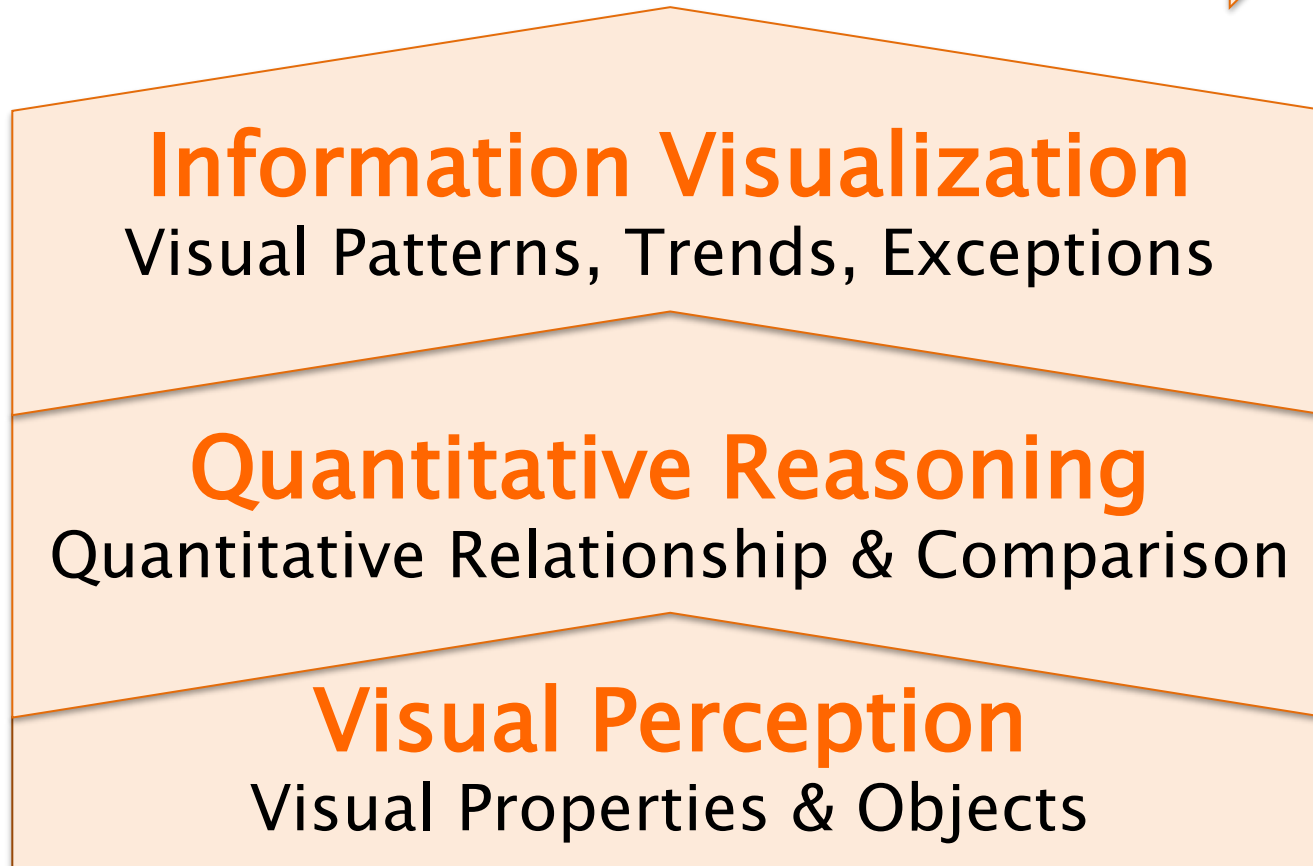
Information visualization

The use of computer-supported,
interactive, visual representations
of abstract data to amplify
cognition

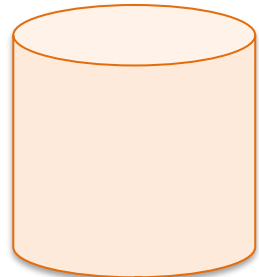
Readings in Information Visualization: Using Vision to Think.
S.K.Card, J.D.Mackinlay, and B.Shneiderman, Academic Press, 1999

Overview

Understanding  Decisions



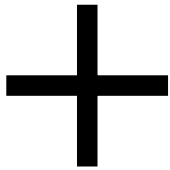
Data



 Representation/Encoding

Quantitative message

- Quantitative values
 - ◆ Express measures



- Categories
 - ◆ Identify what entities the values refer to
 - ◆ Define groups of entities

Understanding tasks

- Variation within quantitative measures
 - ◆ Distribution
 - ◆ Deviation
 - ◆ Correlation
- Variation within category
 - ◆ Ranking
 - ◆ Part-to-whole
 - ◆ Time
 - ◆ Space
- Multivariate

Visualization instruments

- Tables
 - ◆ Textual information
- Graphs
 - ◆ Visual information

Tables

- Main features
 - ◆ Data arranged in rows and columns
 - ◆ Data encoded as **text**
- Strengths
 - ◆ Easy **look-up** of values
 - ◆ Precise values
 - Allow selected comparisons
 - ◆ Several units of measure are possible

Graphs

- Main features

- ◆ One or more **axes** delineate the display area where values are shown
- ◆ Values encoded as **visual** objects in relation to axes
- ◆ Axes provide **scales**
 - Assign values and labels to visual objects
 - Both categorical and quantitative

- Strengths

- ◆ Overall shape of data (holistic)

Graphs

- Show
 - ◆ Trend
 - Pattern of change over time
 - ◆ Comparison of subsets
 - Overall
 - Spot similarities and differences
 - ◆ Highlight exceptions
- Display relationships among multiple quantitative values by giving them shape

In general

Use tables to

Look up individual values

Compare individual values

Precise values are required

There is more than one unit of measure

Use graphs to

Focus on the shape of values

Reveal relationships among multiple values

EXAMPLES

Good and Poor visualization

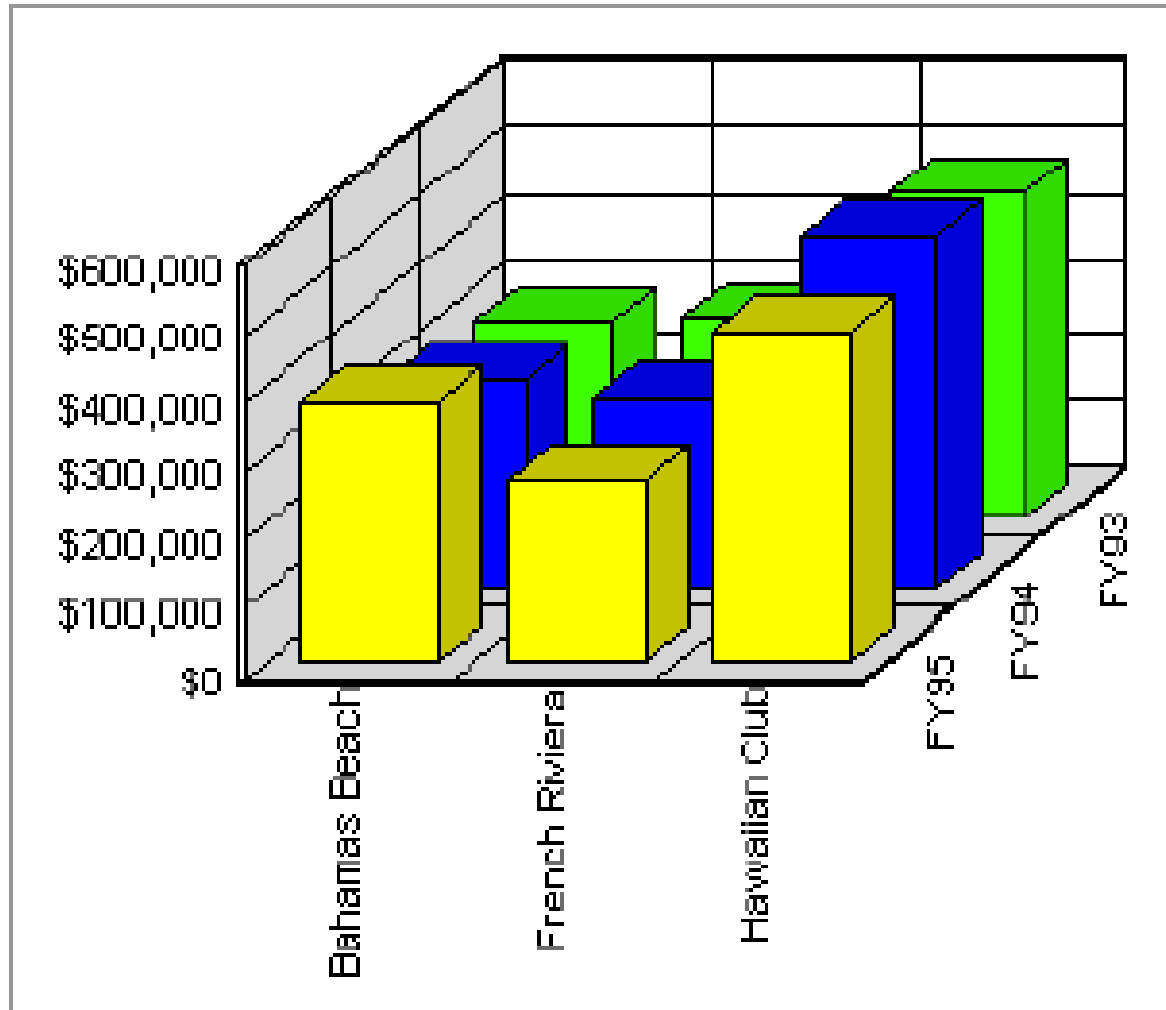
- Like good writing, good graphical displays of data communicate ideas with clarity, precision, and efficiency.
- Like poor writing, bad graphical displays distort or obscure the data, make it harder to understand or compare, or otherwise thwart the communicative effect which the graph should convey.

Friendly, Michael, and Daniel J. Denis. (2001)

"Milestones in the history of thematic cartography, statistical graphics, and data visualization."

<http://www.datavis.ca/milestones>

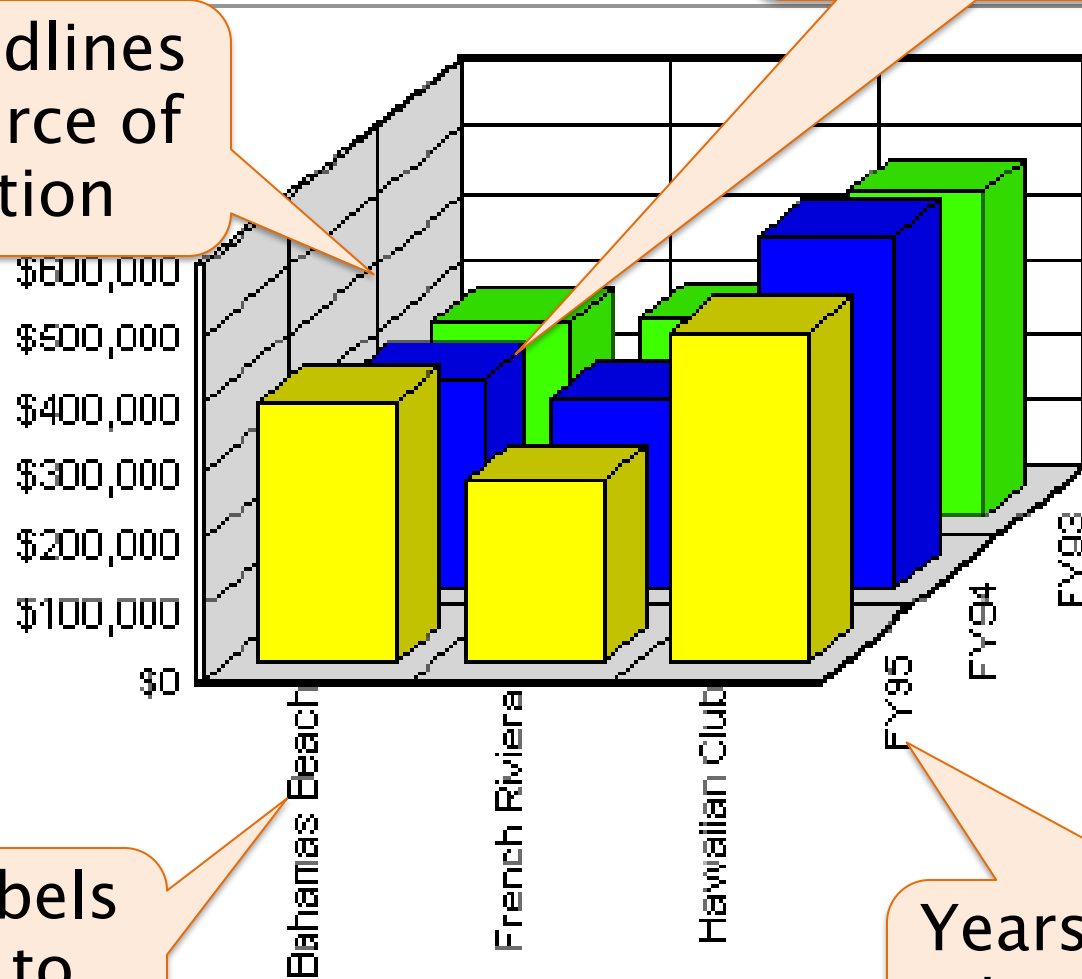
A bar graph



A **bad** bar graph

Heavy gridlines
are a source of
distraction

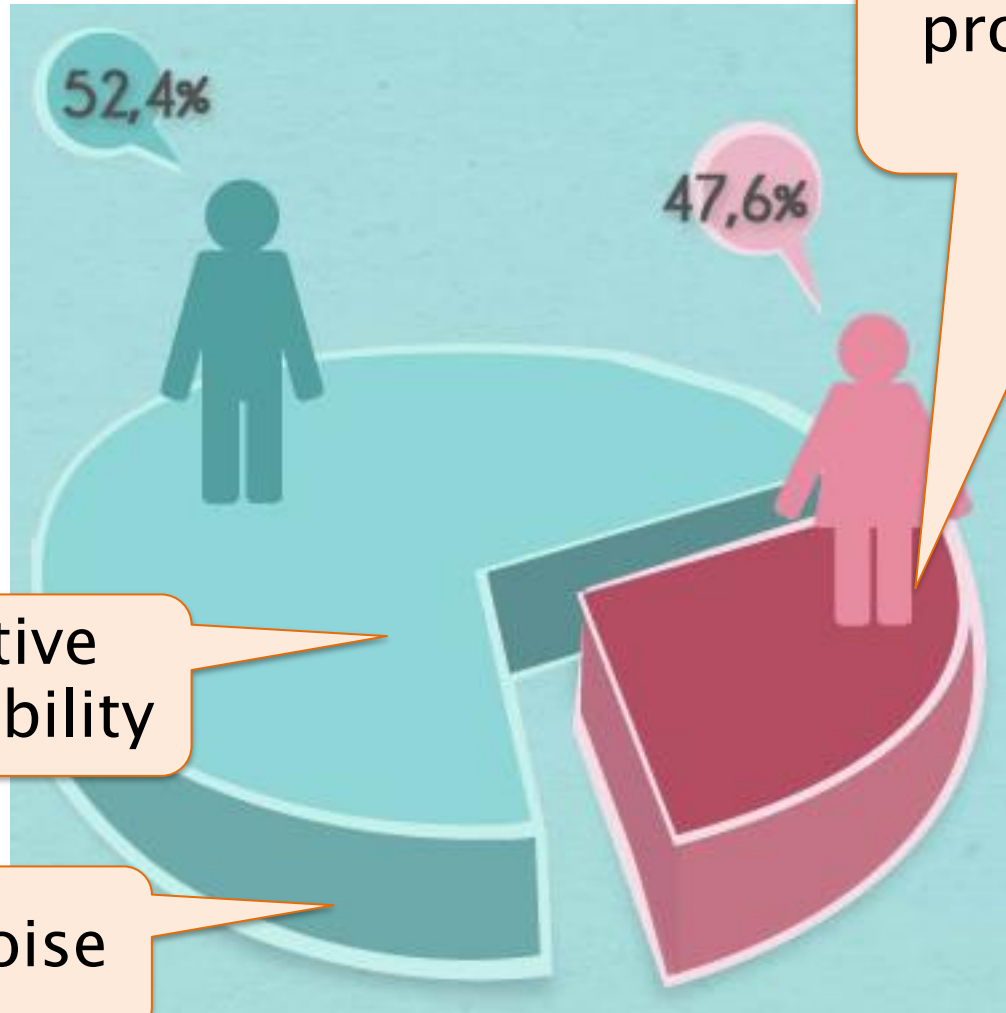
3D bars are
impossible to
read



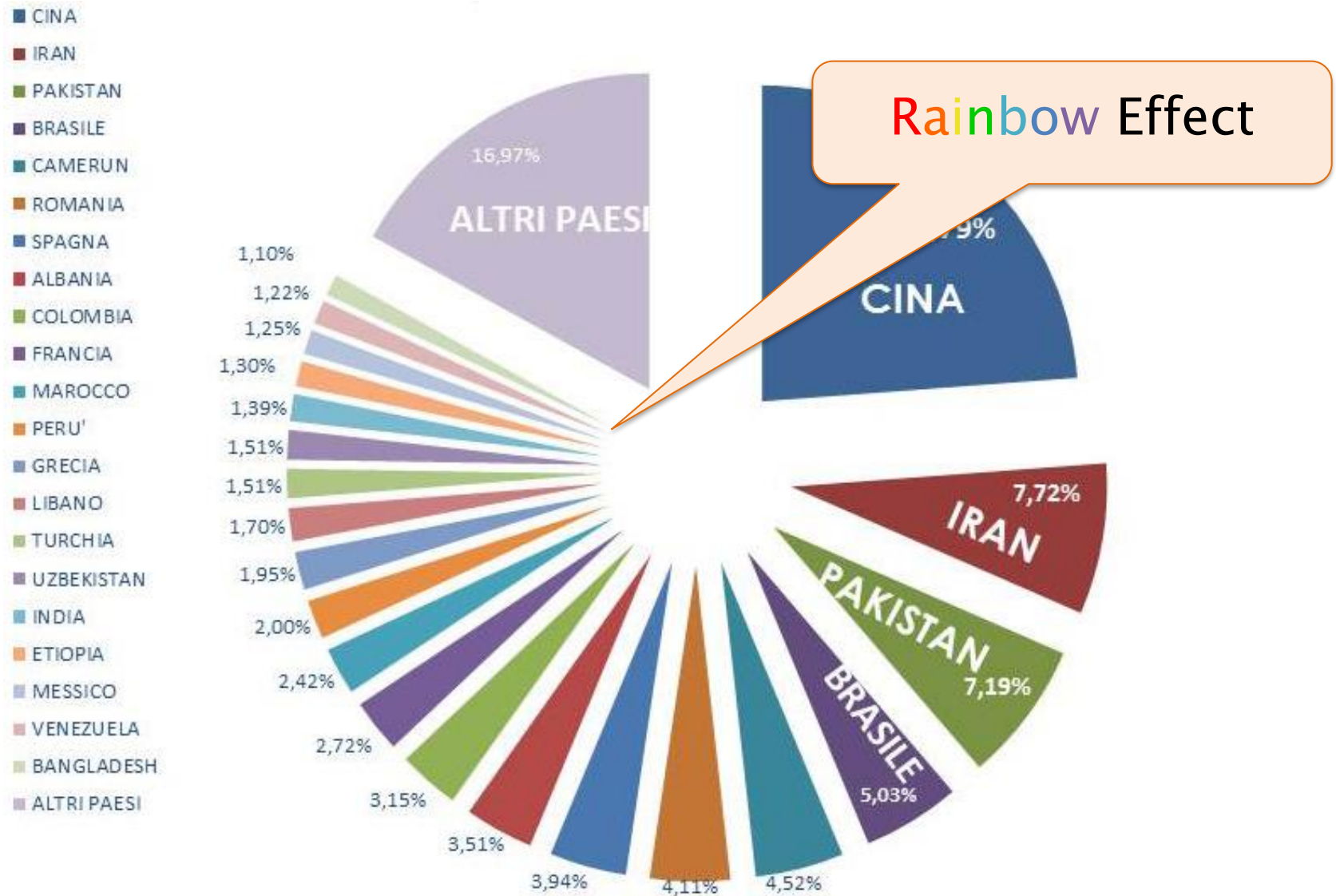
Vertical labels
are hard to
read

Years run counter-
intuitively from
back to front

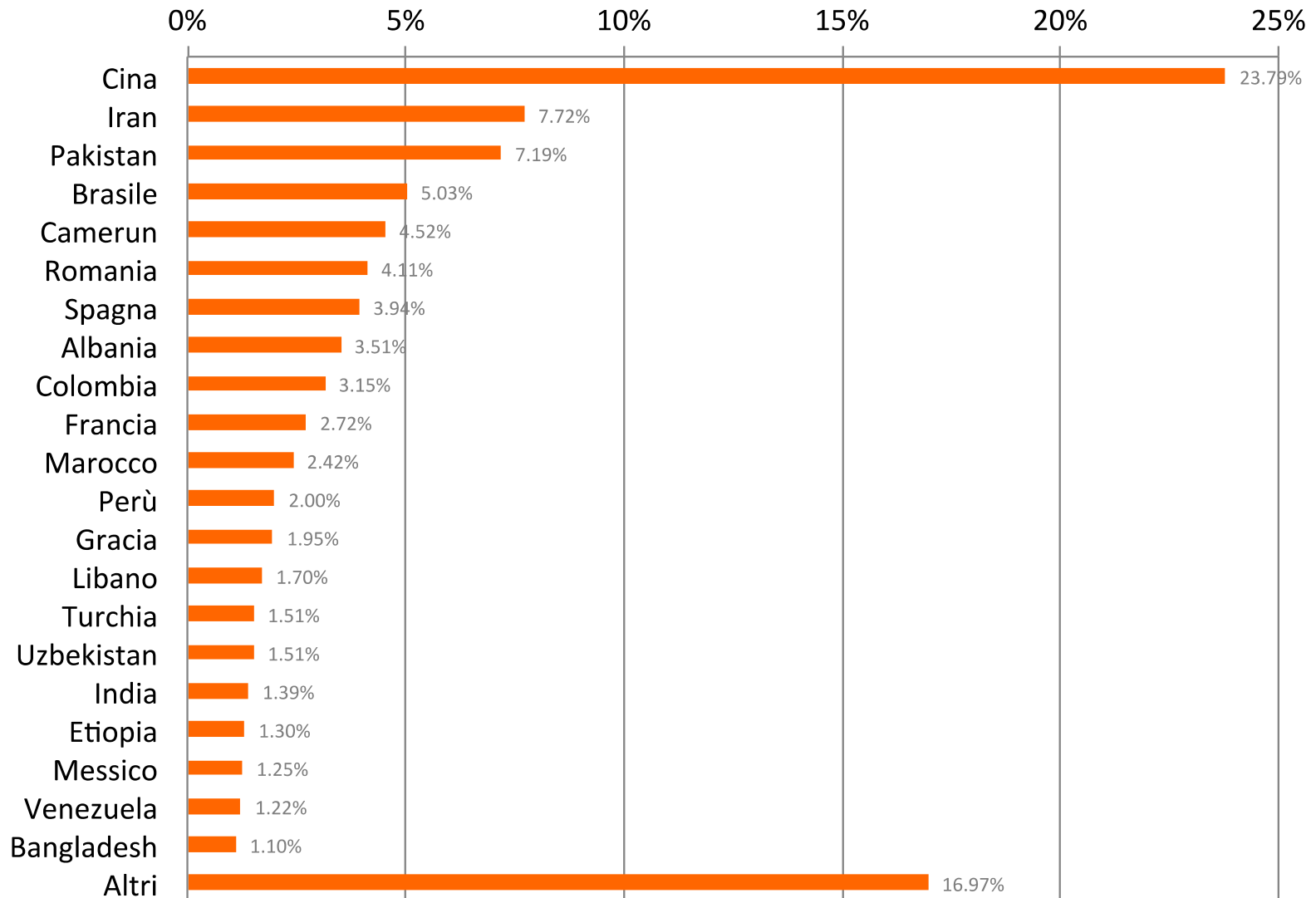
A pie chart



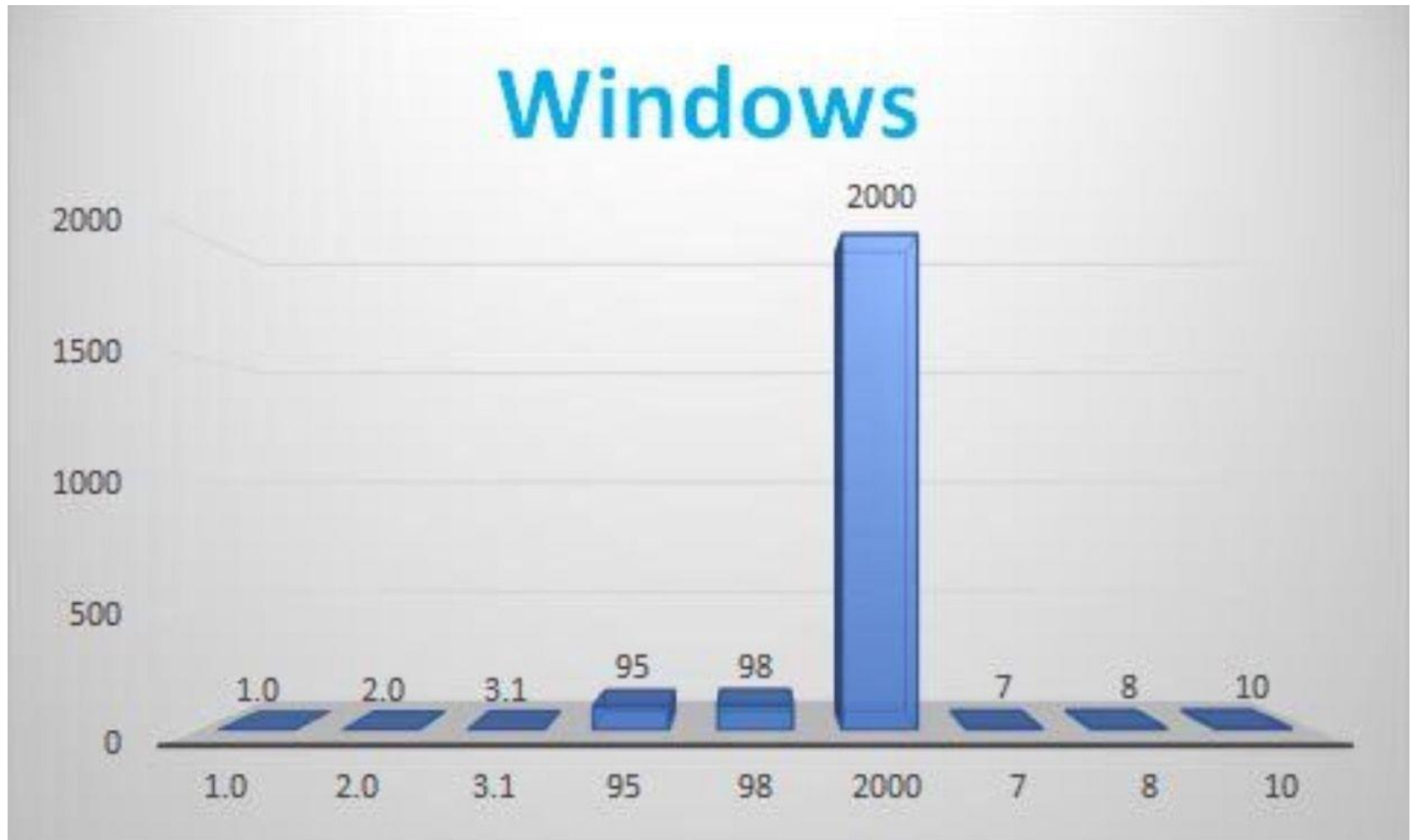
Pie chart (original)



Bar chart (redesign)



Meaningless Data

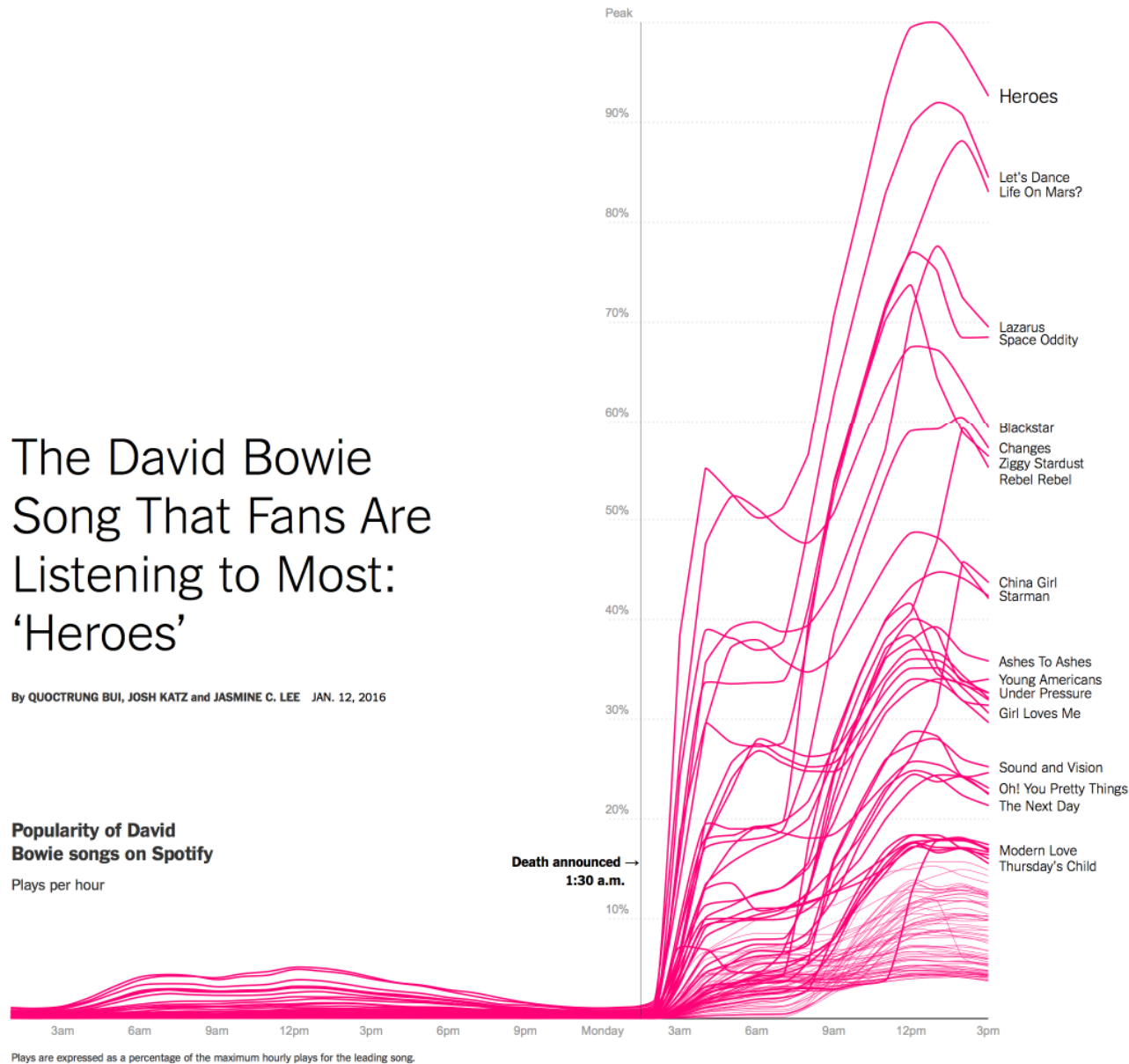


The David Bowie Song That Fans Are Listening to Most: 'Heroes'

By QUOCTRUNG BUI, JOSH KATZ and JASMINE C. LEE JAN. 12, 2016

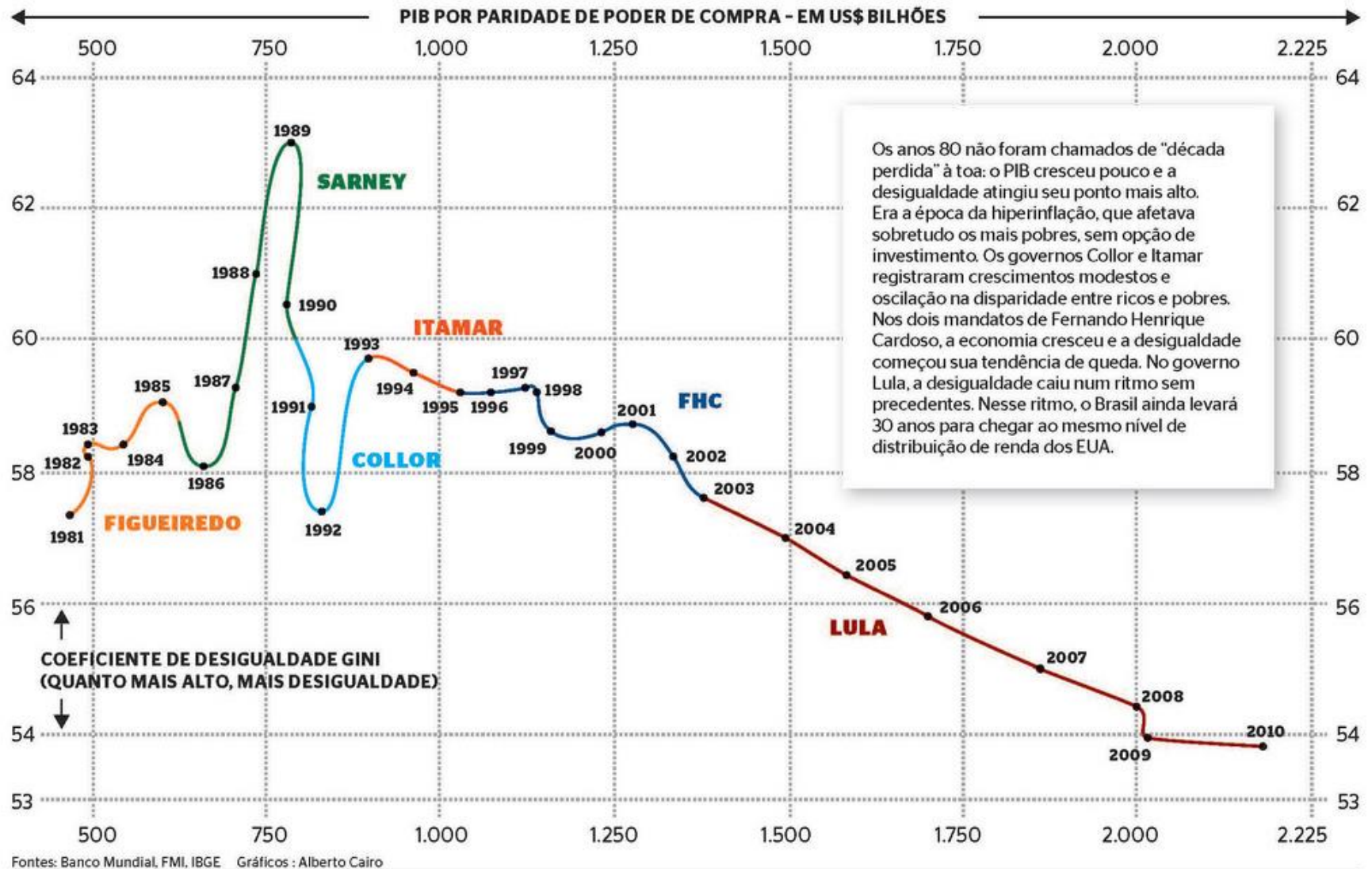
Popularity of David Bowie songs on Spotify

Plays per hour

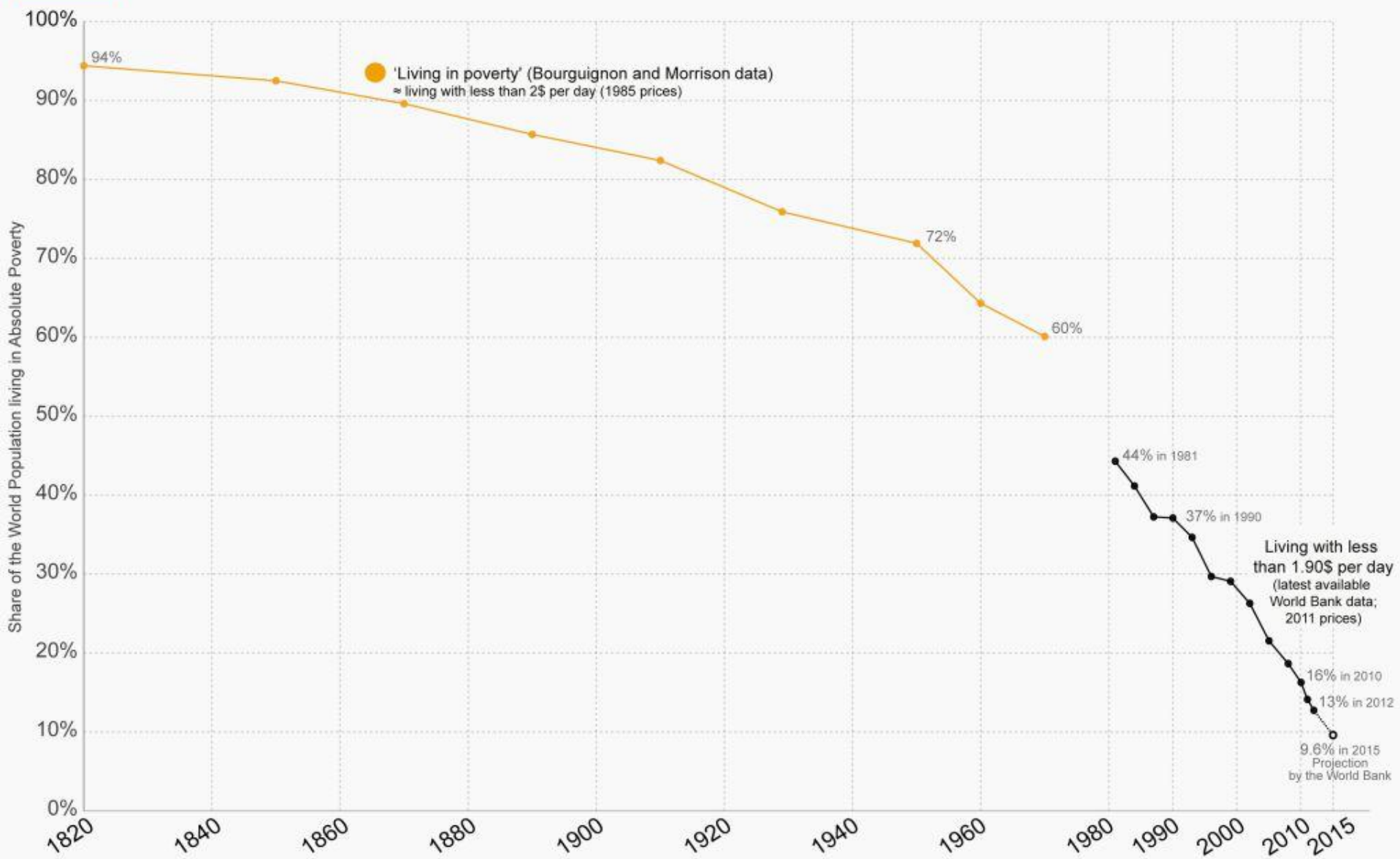


Quando o PIB cresce, nem sempre a desigualdade cai

O gráfico abaixo mostra o avanço do PIB comparado à evolução da desigualdade no Brasil desde 1980. Nem sempre o crescimento econômico levou a uma redução proporcional na disparidade de renda entre os mais pobres e os mais ricos



Our World in Data **Share of the World Population living in Absolute Poverty, 1820-2015 – by Max Roser**
All incomes are adjusted for inflation over time and for price differences between countries (1985-PPP before 1970; 2011-PPP after 1970).



Data sources: 1820-1970 Bourguignon and Morrison (2002) - Inequality among World Citizens, In The American Economic Review; 1981-2015 World Bank (PovcalNet)
The interactive data visualisation is available at ourworldindata.org. There you find the raw data and more visualisations on this topic. Licensed under CC-BY-SA by the author Max Roser.

<http://ourworldindata.org/data/growth-and-distribution-of-prosperity/world-poverty/>

Migrants arrived in period January – June

2019

2018

2017

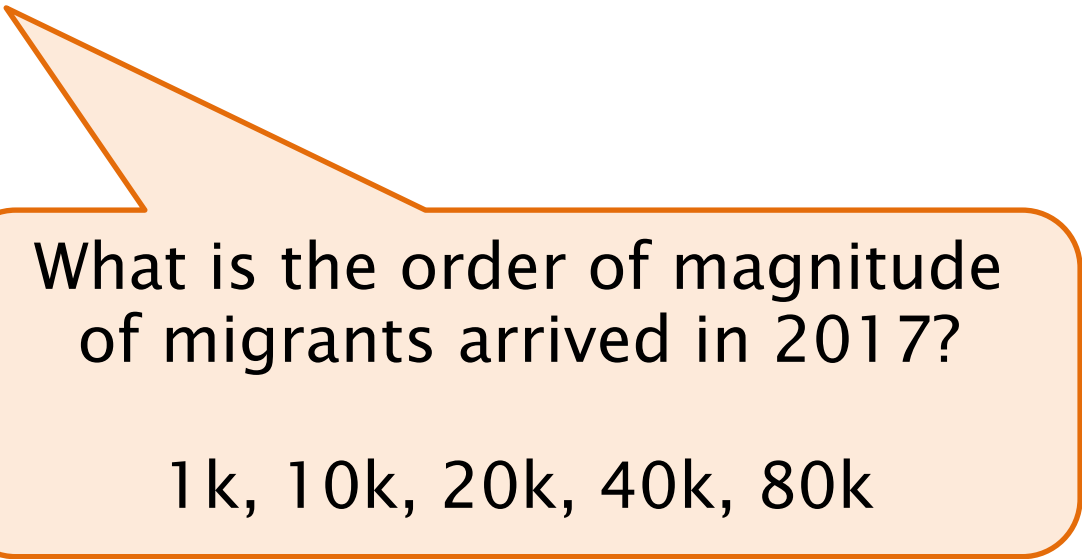
The accidents at work happened and reported

Migrants arrived in period January – June

2019

2018

2017



What is the order of magnitude
of migrants arrived in 2017?

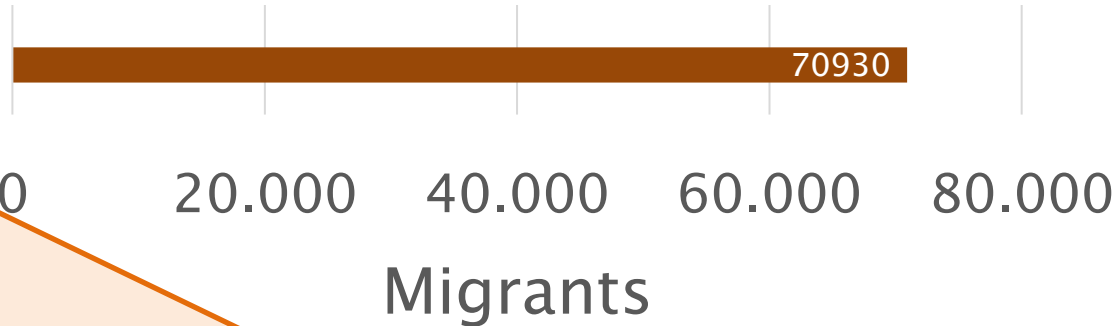
1k, 10k, 20k, 40k, 80k

Migrants arrived in period January – June

2019

2018

2017



What is the order of magnitude of migrants arrived in 2017?

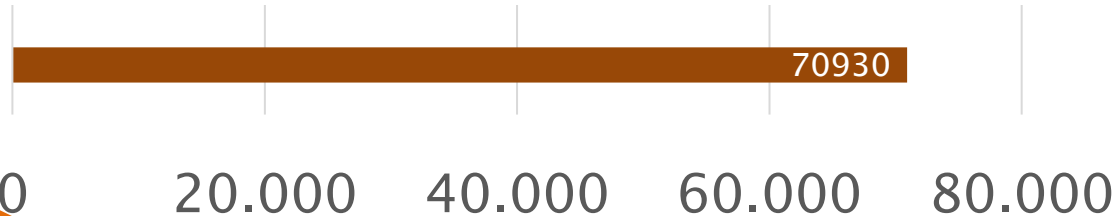
1k, 10k, 20k, 40k, 80k

Migrants arrived in period January – June

2019

2018

2017

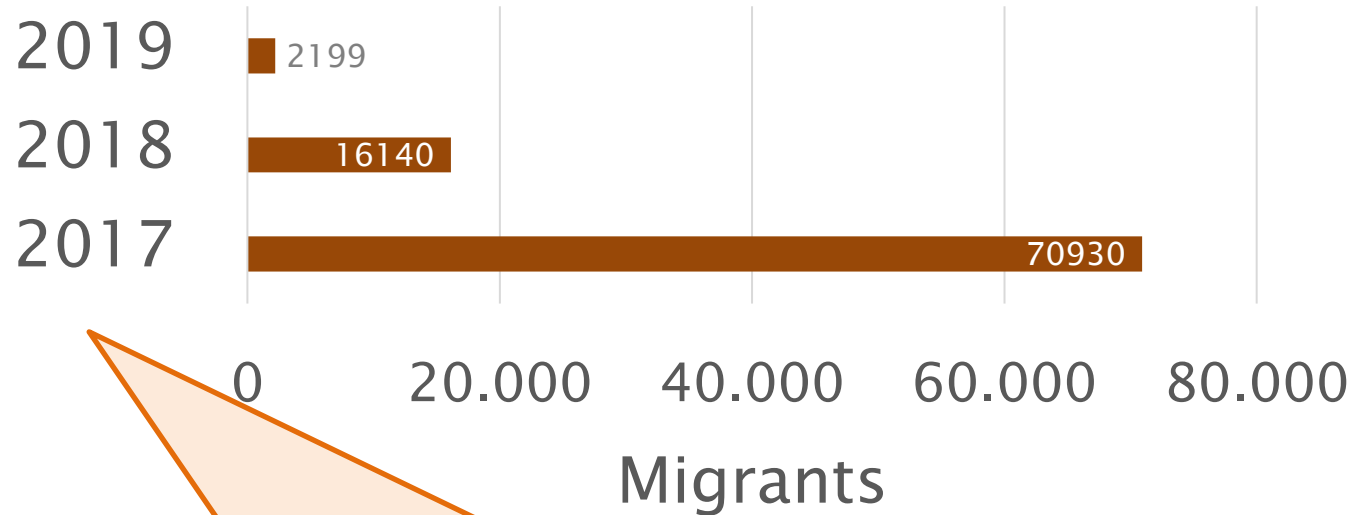


Migrants

The ratio of number of migrants
in 2018, with respect to 2017 is

1:1 , 1:2 , 1:4 , 1:10 , 1:20

Migrants arrived in period January – June



The ratio of number of migrants in 2018, with respect to 2017 is

1:1 , 1:2 , 1:4 , 1:10 , 1:20

What is the order of magnitude
of accidents in Q1 2019?

1 k, 50k, 100k, 200k, 500k

The **accidents at work** happened and reported
to Inail in first quarter 2019 have been

What is the order of magnitude
of accidents in Q1 2019?

1 k, 50k, **100k**, 200k, 500k

The **accidents at work** happened and reported to Inail in first quarter 2019 have been 131 thousand (109 thousand at work and 22 thousand while traveling),

With respect to Q1 2018 how much have changed accidents in Q1 2019?

-5k, -2k , ± 500 , +2k , +5k

The **accidents at work** happened and reported to Inail in first quarter 2019 have been 131 thousand (109 thousand at work and 22 thousand while traveling),

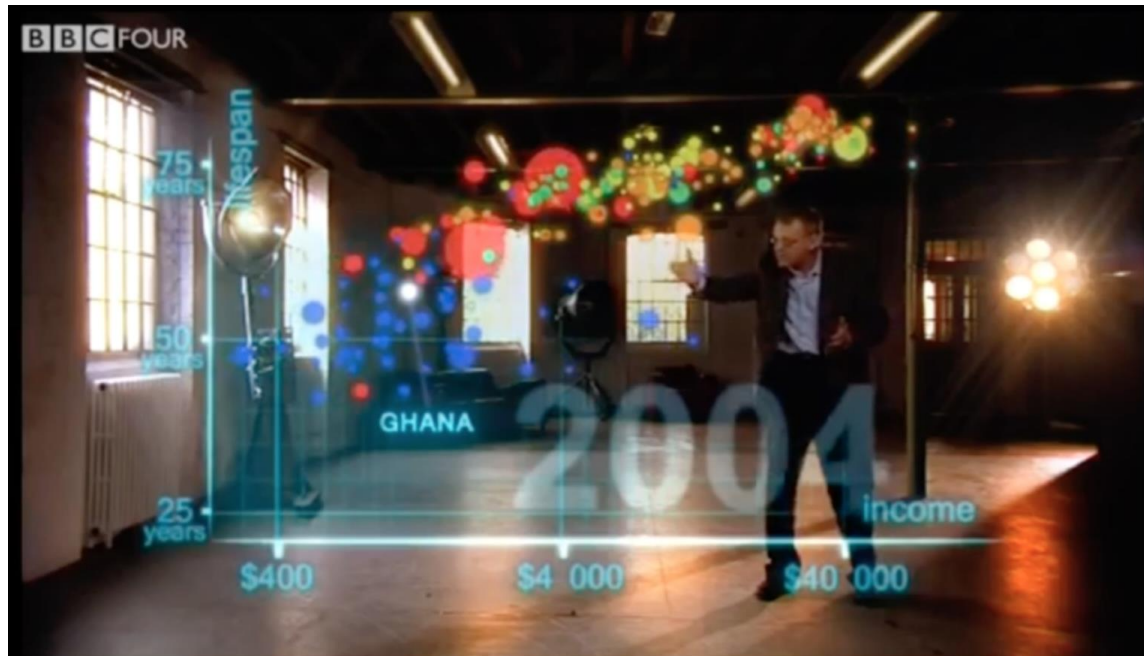
With respect to Q1 2018 how much have changed accidents in Q1 2019?

-5k , -2k , ± 500 , **+2k** , +5k

The **accidents at work** happened and reported to Inail in first quarter 2019 have been 131 thousand (109 thousand at work and 22 thousand while traveling), increased by 1.7% (+2 thousand reports) with respect to first quarter 2018

Hans Rosling (1948–2017)

- 200 Countries, 200 Years, 4 Minutes
 - ◆ The Joy of Stats – BBC 4
 - <http://www.bbc.co.uk/programmes/b00wgg0l>
 - <https://www.youtube.com/watch?v=jbkSRLYSojo>



References

- The picture superiority effect in associative recognition.
 - ♦ <https://www.ncbi.nlm.nih.gov/pubmed/18927048>
- John W. Tuckey, 1977. Exploratory Data Analysis, Pearson.
- Edward R. Tufte, 1983. The Visual Display of Quantitative Information. Graphics Press.
- William S. Cleveland, 1994, The Elements of Graphing Data, Hobart Press
- S.K.Card, J.D.Mackinlay, and B.Shneiderman. Readings in Information Visualization: Using Vision to Think. Academic Press, 1999