What is Visualization

Data Management and Visualization



Version 2.2.1 © Marco Torchiano, 2020





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.

Solution 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 problem and graph types



Exam [6 points]

- Theory [1]
- Assessment
 - Question [0.25]
 - Data [1.25]
 - Visual
 - Proportionality [0.75]
 - Utility [0.75]
 - Clarity [0.5]
- Redesign [0.25 + 1.25]



Definition

Visualization:

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





WHY VISUALIZATION?



Migrants arrived in period January - June



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

big:ftwwwstat.it/it/files//2019/06/NotaTrimestrale-Occupazione-I_2019.pdf

Information retrieval

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





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





http://www.mamartino.com/projects/rise_of_partisanship/index.html

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)]





http://www.nytimes.com/interactive/2015/02/23/business/economy/the-changing-nature-of-middle-class-jobs.html

HISTORY



William Playfair



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



5	ю	20	3	o_4	to 1	50	6	0	7º.	8	0 0	20	1	0 11	0	<u> </u>	<u>00</u>		150	- т	17))			0	10	ю.		•	260	<u>,</u>	280	Ļs	0	000
		i i i i i i i i i i i i i i i i i i i	 							-								-			-						-							0.00	Vames of Places.
						1				-			-					-		-					5. m.c.	_				-+	-†			-	Jersey &c.
		٦	-			-		•				1-			/				1	1						I					-				treland
Ĩ	i i	Ì				1			T				1				1																		Poland
-	-	-															Ì		T	Ţ								and a							tole of Man
0									1			1		• • •																					Areenland.
	•	-				1													T																Prysia
	÷								1			1							-	_														1	Bra igal
0																	1		1				ļ		1									-	Holland
		-							1				_			ļ	Ļ	ļ		_	_		ļ., ,		 										Awaden
									1 -							! 	ļ	<u> </u>	+	-	_			÷	ļ		1		-					-	Querticy
	d,	-							-										4.																Denmark and
						ļ			and the second							-			-	and the second second				+											Flandere
				1				-				1					Υ.,		+	Ē				; <u> </u>											West Indees
- 100 C												Ī			Ĩ			T	T	+			Ľ,	,						i					America
-				-	1										1			-			Ĩ													_	Rufna
	÷	-						÷	\$			i.																							treland.

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

The I'pright divisions are Ten Thousand Pounds each . The Black Lines are Exports the Ribbed lines Imports . Billion a de Automo Sone 1ª get to Men Martin

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



Charles Joseph Minard



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



Numbers and direction





Temperature

mannananananananananananananananananana	Zéro le 18 8 ^b



Étienne-Jules Marey



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



XX Century

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





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







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

Use graphs to

Look up individual values

Focus on the shape of values

Compare individual values

Reveal relationships among multiple values

Precise values are required

There is more than one unit of measure



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 pie chart

52,4%



3D Perspective worsen readability

Depth add noise



Pie chart (original)



SoftEng.

Bar chart (redesign)





Meaningless Data





TheUpshot

FOLLOW US: 🧗 💓 🔊 GET THE UPSHOT IN YOUR INBOX

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



Peak

Plays are expressed as a percentage of the maximum hourly plays for the leading song.

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



http://www.visualisingdata.com/2011/08/data-visualisation-stories-from-brazil-by-alberto-cairo/



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

201920182017

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 70930 20.000 40.000 60.000 80.000 **Migrants** 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 70930 20.000 40.000 60.000 80.000 Migrants The ratio of number of migrants

in 2018, with respect to 2017 is



Migrants arrived in period January - June





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

1k, 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?

1k, 50k, <mark>100k</mark>, 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/b00wgq0l
 - <u>https://www.youtube.com/watch?v=jbkSRLYSojo</u>





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

