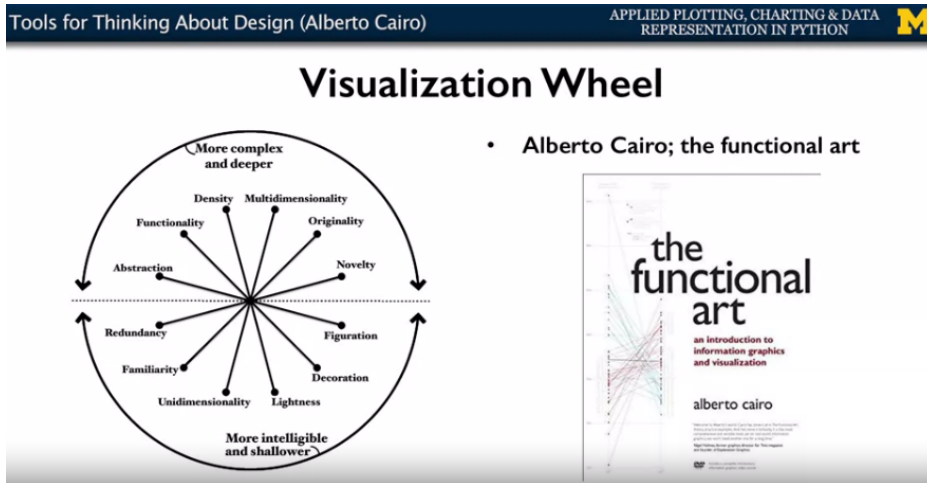
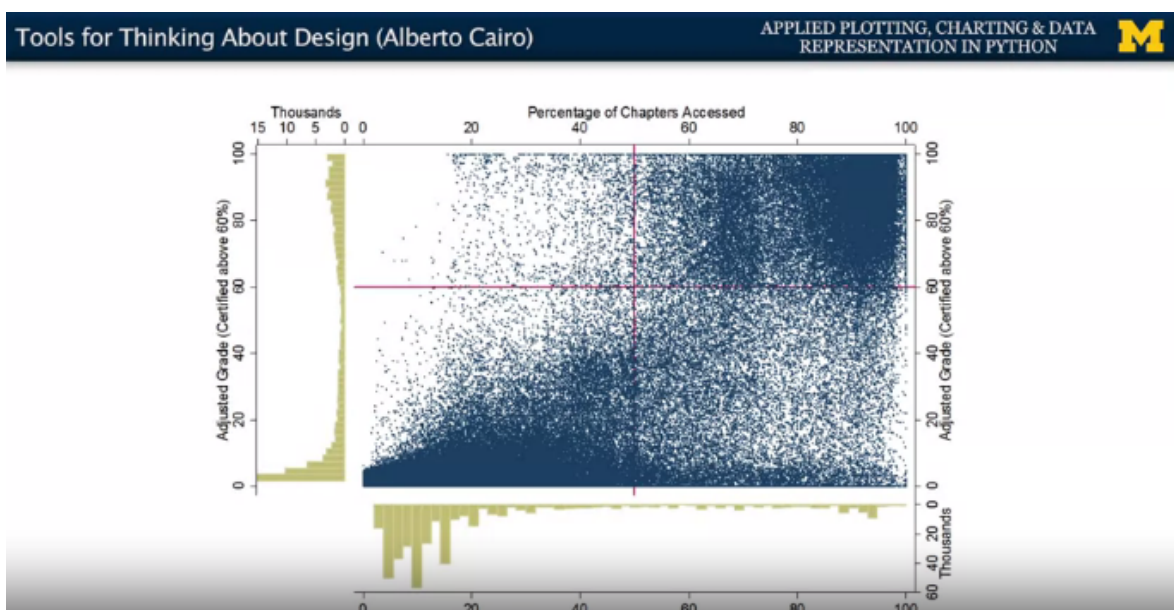


Week1

1. Tools for Thinking about Design (Alberto Cairo)



- Visualization Wheel Dimensions
 - Abstraction - Figuration
 - Boxes and charts(abstraction) or real-world physical objects (figuration)
 - Functionality - Decoration
 - No embellishments (functionality) or artistic embellishments (decoration)
 - Multidimensional - Unidimensional
 - Different aspects of phenomena (multidimensional) or single or few items of phenomena (unidimensional)
 - Novelty - Redundancy
 - Explaining each item once (novelty) or encoding multiple explanations of the same phenomena (redundancy)



- Regardless of whether this is a good figure or not, it would certainly be considered an example of a dense figure.

[illegible]

- It describes Napoleon's march into Russia in 1812.
- Broadly speaking, there are 5 different kinds of information being visualized in this graphic
- It includes location, direction, temperature, army size and dates.

Tools for Thinking About Design (Alberto Cairo)

APPLIED PLOTTING, CHARTING & DATA REPRESENTATION IN PYTHON

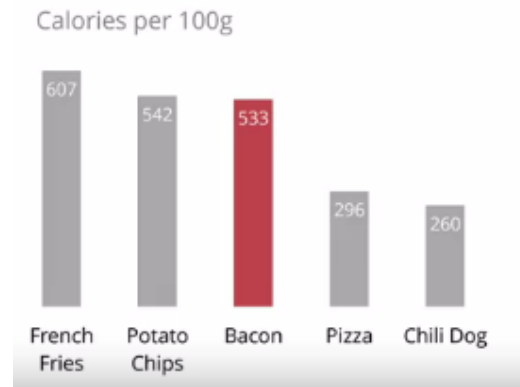
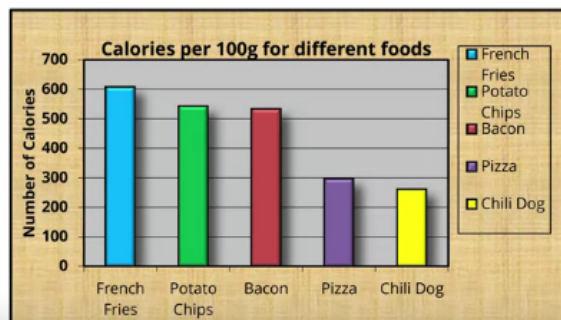
The wheel preferred by scientists and engineers

The wheel favored by artists, graphic designers, and journalists

- There are no rights and wrongs in the visualization wheel.
- It is one tool we have in order to better compare two different ways of visualizing information.

2. Graphical heuristics : Data-ink ratio (Edward Tufte)

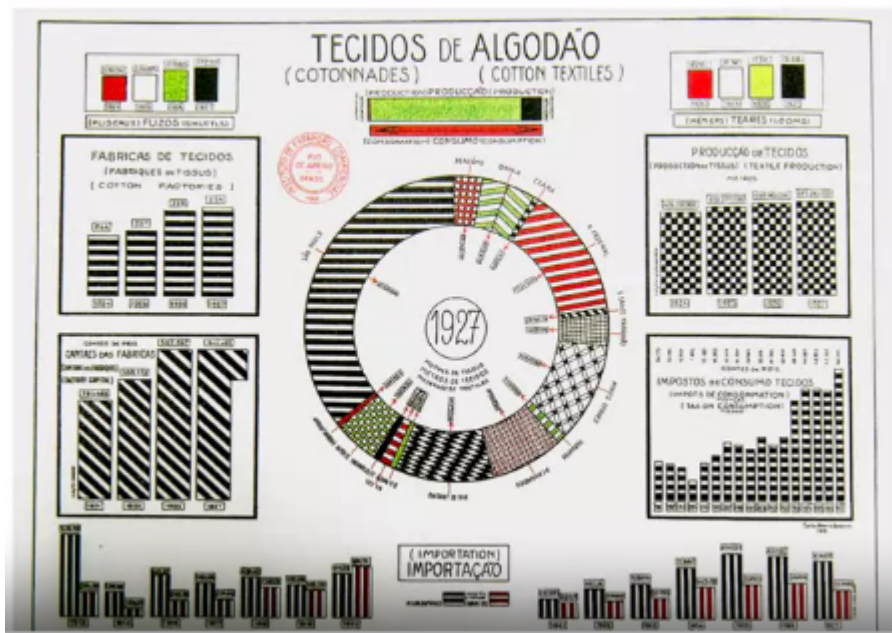
- heuristic : a process or rule that is meant to guide you in decision making



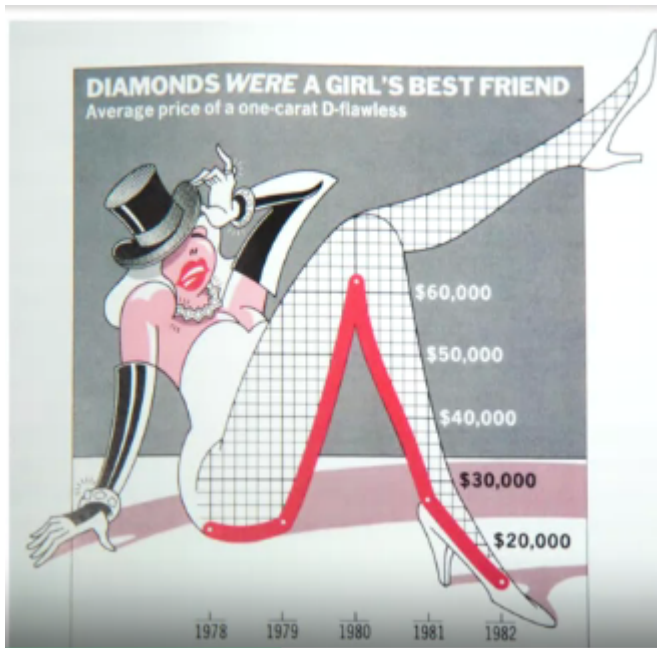
- Dark Horse Analytics : <https://www.darkhorseanalytics.com/>

3. Graphical heuristics : Chart junk (Edward Tufte)

- He suggested there are 3 kinds of chart junk.
 - Unintended optical art
 - excessive shading or patterning of chart features



- Patterns here make the human eye jump and cause visual fatigue
 - it's called moire patterns.
- Grid
 - He suggested the grid is both unnecessary as data ink.
 - Direct labeling of data is another great way to reduce this form of chart junk.
- The duck



- Sometimes it may be easier to recall the details of charts using the ducks.
- Does this mean that you should use embellishments in your charts?
- You should have, at least, some thoughts from designers as well as from a user study. It's worth not only reflecting on the principles you use and the results you are sharing, but also the process by which you came up to create the graphics.
- [Useful Junk?: The Effects of Visual Embellishment on Comprehension and Memorability of Charts.](#)

4. Graphical heuristics : Lie Factor and Spark Lines (Edward Tufte)

- Lie factor : the size of an effect shown in the graphic divided by the size of the effect actually in the data
- In this graphic, various barrels of oil are shown, showing the price of oil over six years.
- It's unclear to the viewer what the size of one barrel means in relation to the others.
- Are the barrels different sizes, or do they just appear

- **Lie Factor**
 - *If you consider volume, the lie factor is 9.4 times the stated prices*

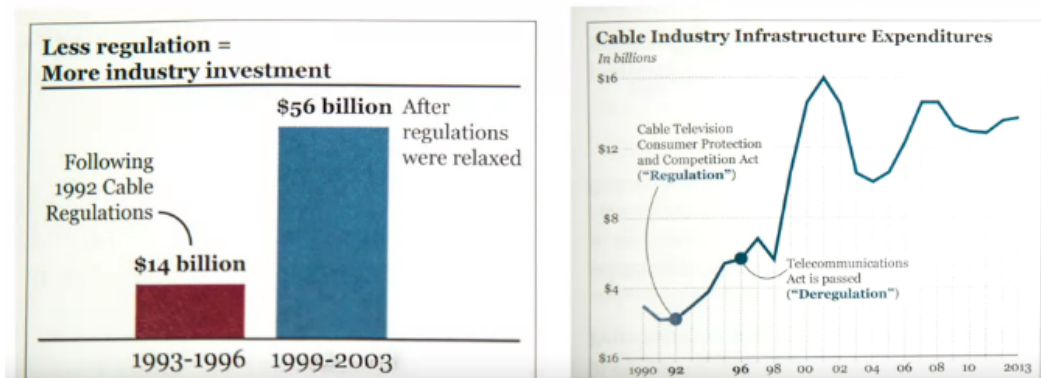


5. The Truthful Art (Alberto Cairo)

- The qualities of a good visualization

1. Truthful

- Be aware of your actions when cleaning, summarizing and manipulating data and ensure you are not.
 - misleading yourself
 - misleading your audience
- Skepticism is important as a data scientist.



- The left figure has mistakes.
2. Functionality
 3. Beauty
 4. Insightful
 5. Enlightening
- A combination of the previous four, but with a social ethical responsibility
 - Choosing topics ethically and wisely - casting light over relevant issues - matters a lot.

https://d3c33hcgiewev3.cloudfront.net/7yQEXL8gR8ykBFy_ILfMbg_8b3084c2d72342b9bb885f988fca5a27_BookChapterLIES.pdf?Expires=1633564800&Signature=PCd0Z2txfp~LfAd1vWTtHZNiSM SOZ7QO4YnE6aCmE2Eu5leV0Ez59MsSOkWjdOpiGAKSnYSEUOed1OtycmECQbsXGLJS3axH QGH9fsRHJGDDVPfbV0kowj9J8O-r-1xZ2VjE6GDO6~1zLjbOiBr0xWu~Yj9fBI4xr6okovAw_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A

1. Hiding relevant data to highlight what benefits us



FIGURE 1: TWO VERSIONS OF A FICTIONAL GRAPHIC. THE FIRST ONE DISPLAYS JUST THE AVERAGE CHANGE, HIDING THE FACT THAT THE TEN MARKETS IN WHICH THIS COMPANY OPERATES BEHAVED VERY DIFFERENTLY BETWEEN 2012 AND 2013. THAT REALITY IS SHOWN ON THE SECOND VERSION OF THE GRAPHIC. THE AVERAGE OF ALL THE VALUES ENCODED IN THE BAR GRAPH IS INDEED 12%

But see what happens when we offer a more nuanced picture of what hides behind that figure (man on the right): Among the ten markets in which this company operates, five improved, but the other five shrank. Why would you hide this important bit of information from customers or investors? That was a rhetorical question, in case you didn't notice. The first "infographic" is good "link bait". The second one is true.

2. Displaying too much data to obscure reality

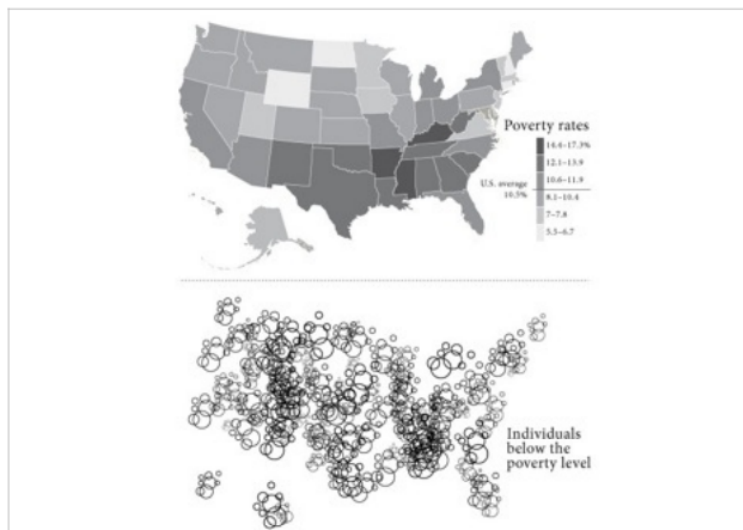


FIGURE 2: CHOROPLETH MAP AND PROPORTIONAL SYMBOL MAP. DATA AND LOCATIONS ARE NOT REAL.

3. Using graphic forms in inappropriate ways (distorting the data)

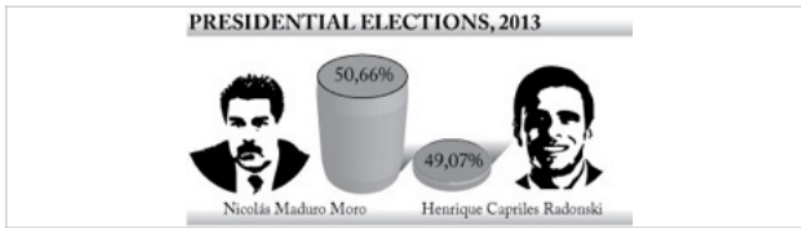


FIGURE 3: PRESIDENTIAL ELECTION RESULTS IN VENEZUELA, BASED ON A GRAPHIC BY VENEZONALA DE TELEVISION. NOTICE THE TRUNCATED Y-AXIS WHICH GREATLY DISTORTS THE DIFFERENCE BETWEEN THE PERCENTAGES OF VOTE.

When adding a zero-baseline, the picture becomes much more interesting —and discouraging if you're a fan of Maduro's (FIGURE 4). I've heard designers defend graphics like that one arguing that viewers could just read the numbers. Right but, then, what are the damn bars for? Are they just ornaments? And I could add that most people won't have enough time to make up for the distortion if the chart stays on the TV screen for just a few seconds.

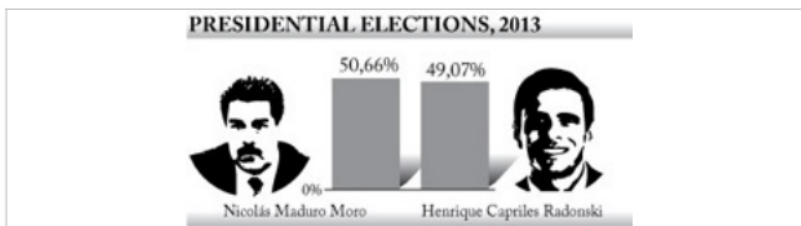


FIGURE 4: AN ALTERNATIVE VERSION OF THE PREVIOUS GRAPHIC IN WHICH A 0-BASELINE HAS BEEN ADDED, AND THE 3D EFFECT HAS BEEN REMOVED.

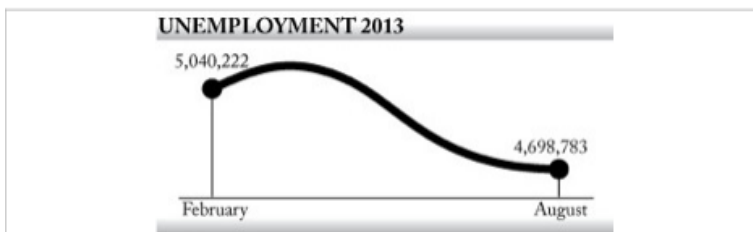


FIGURE 5: THE TRUNCATED Y-AXIS IS NOT THE ONLY PROBLEM HERE. BASED ON A GRAPH BY RTVE (SPAIN)

7 of 13

To understand why the designer is lying, you need to know that Spain's job market follows regular seasonal variations. A notable chunk of the country's wealth comes from the millions of tourists who visit it every year. Thus unemployment tends to increase during the winter and falls sharply during the summer, something that becomes clear when 12-month data are displayed (FIGURE 6.) The most newsworthy bit of information revealed by this chart is that unemployment was higher in August 2013 than in July 2012, something that TVE didn't mention. Maybe a smart government-appointed managing editor thought that she wouldn't be able to use the right headline if data were correctly depicted.

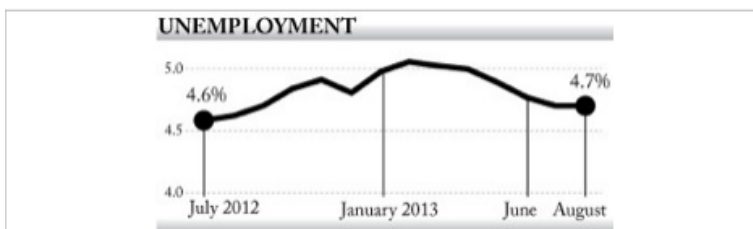


FIGURE 6: TWELVE MONTHS OF UNEMPLOYMENT RATES. BASED ON A GRAPH BY WWW.ELDIARIO.ES

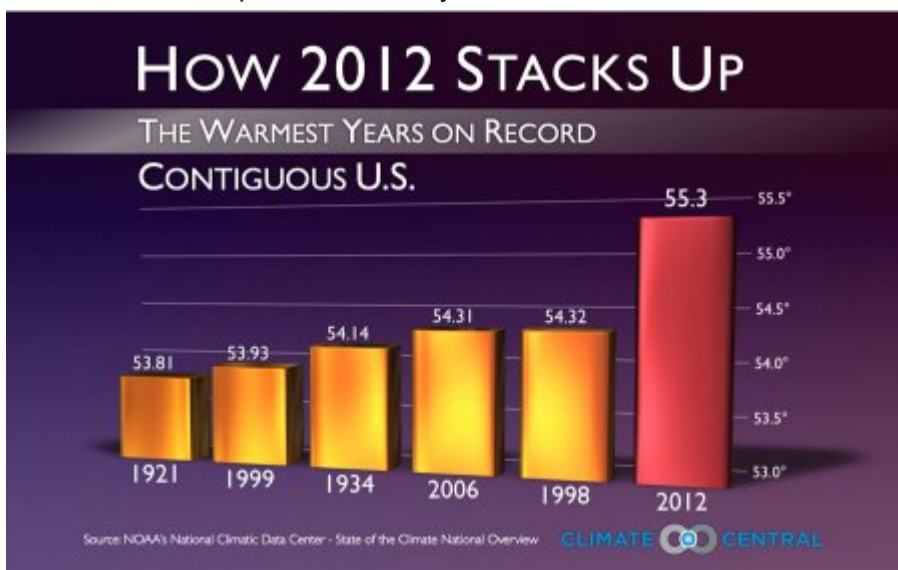
Assignment

Read Alberto Cairo's work, [Graphics Lies, Misleading Visuals](#)

Locate an example of a misleading visual that uses one or more of the mechanisms for misleading that Cairo outlines in his book chapter: (1) Hiding relevant data; (2) Displaying too much data and obscuring reality; (3) Distorting data through visual forms.

Please upload an image of this visual using a widely accessible graphic format (e.g., PDF, .jpg, .png)

How 2012 stacks up - the warmest years on record



This figure is provided by 'CLIMATE CENTRAL' which is an independent organization of leading scientists and journalists researching and reporting the facts about our changing climate and its impact on the public.

Briefly describe the context for the visual by addressing the following questions:

1. What is the source of the visual? (e.g., URL or bibliographic citation)
2. Who is the intended audience (i.e., decoders)? How do you know this?

1. <https://www.climatecentral.org/gallery/graphics/its-official-2012-is-hottest-us-year-on-record>

The article was published in Jan 10th, 2013.

2. The intended audience is the public because their purpose for reporting is providing the facts about climate change and impact on the 'public' as they mentioned on their website. The public can be anyone, like a researcher, student, politician who are interested in climate change and able to access this website.

1. **Identify the specific component(s) of the visual that is/are misleading**
2. **For each part(s) of the visualization that is/are misleading, identify the mechanism that is used: hiding relevant data to highlight what benefits us; displaying too much data to obscure reality; using graphic forms in inappropriate ways (distorting the data)**
3. **Explain how the mechanisms are used to mislead**

[X-axis order]

1) The order of year is not in chronical order but in ascending order of figure.

- mechnism used : using graphic forms in inappropriate ways (distorting the data)

- Explanation : It is possible there were ups and downs in temperature before 2012 but the figure shows bar in order of values so it looks like the temperature had been increased steadily.

2) The interval between years is not regular.

- mechnism used : hiding relevant data to highlight what benefits us

- Explanation : The interval should be regular but it picks up specific years with irregular intervals focusing on the increase of temperature. It cannot provide real changes as time went by.

[Ambiguous Y-axis unit]

3) It doesn't provide what Y-axis standards (whether the value is the average temperature of the year or stacked up temperature, which unit of the temperature is , where the area is limited in U.S or worldwide)

- mechnism used : hiding relevant data to highlight what benefits us

- Explanation : As we cannot know what this number represents, it's hard to guess this increase is meaningful or not. Not only that, due to the lack of information, we can't figure out if increasing changes happend only in the U.S or it's global issue.