Data Visualization
This class isn’t mainly about how to **DESIGN** visualizations. It’s about how to **REASON** about visualization.
Anybody can learn visualization
for the same reasons that anybody can learn to read and write
1. Defining visualization
2. The elements of a visualization
3. How visualization lies
4. Visualization for communication: essential principles
I. Defining visualization
2. The elements of a visualization
3. How visualization lies
4. Visualization for communication: essential principles
A visualization is a representation designed to enable exploration, discovery, or communication.
<table>
<thead>
<tr>
<th>YEAR</th>
<th>TEMP</th>
<th>YEAR</th>
<th>SIGMA 1</th>
<th>SIGMA 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0.0659</td>
<td>100</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1001</td>
<td>-0.1241</td>
<td>1001</td>
<td>0.240347</td>
<td>0.480694</td>
</tr>
<tr>
<td>1002</td>
<td>-0.1208</td>
<td>1002</td>
<td>0.240346</td>
<td>0.480692</td>
</tr>
<tr>
<td>1003</td>
<td>-0.1801</td>
<td>1003</td>
<td>0.240347</td>
<td>0.480694</td>
</tr>
<tr>
<td>1004</td>
<td>-0.0711</td>
<td>1004</td>
<td>0.240347</td>
<td>0.480694</td>
</tr>
<tr>
<td>1005</td>
<td>-0.1334</td>
<td>1005</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1006</td>
<td>-0.0644</td>
<td>1006</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1007</td>
<td>0.0042</td>
<td>1007</td>
<td>0.240347</td>
<td>0.480694</td>
</tr>
<tr>
<td>1008</td>
<td>-0.1288</td>
<td>1008</td>
<td>0.240347</td>
<td>0.480693</td>
</tr>
<tr>
<td>1009</td>
<td>-0.0296</td>
<td>1009</td>
<td>0.240347</td>
<td>0.480693</td>
</tr>
<tr>
<td>1010</td>
<td>0.1187</td>
<td>1010</td>
<td>0.240347</td>
<td>0.480694</td>
</tr>
<tr>
<td>1011</td>
<td>-0.1252</td>
<td>1011</td>
<td>0.240346</td>
<td>0.480692</td>
</tr>
<tr>
<td>1012</td>
<td>-0.1634</td>
<td>1012</td>
<td>0.240347</td>
<td>0.480694</td>
</tr>
<tr>
<td>1013</td>
<td>-0.0791</td>
<td>1013</td>
<td>0.240347</td>
<td>0.480694</td>
</tr>
<tr>
<td>1014</td>
<td>-0.1120</td>
<td>1014</td>
<td>0.240347</td>
<td>0.480693</td>
</tr>
<tr>
<td>1015</td>
<td>-0.1146</td>
<td>1015</td>
<td>0.240346</td>
<td>0.480692</td>
</tr>
<tr>
<td>1016</td>
<td>-0.1206</td>
<td>1016</td>
<td>0.240346</td>
<td>0.480692</td>
</tr>
<tr>
<td>1017</td>
<td>-0.0815</td>
<td>1017</td>
<td>0.240346</td>
<td>0.480692</td>
</tr>
<tr>
<td>1018</td>
<td>-0.2031</td>
<td>1018</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1019</td>
<td>0.0305</td>
<td>1019</td>
<td>0.240347</td>
<td>0.480693</td>
</tr>
<tr>
<td>1020</td>
<td>0.1098</td>
<td>1020</td>
<td>0.240347</td>
<td>0.480694</td>
</tr>
<tr>
<td>1021</td>
<td>0.0244</td>
<td>1021</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1022</td>
<td>-0.0743</td>
<td>1022</td>
<td>0.240347</td>
<td>0.480693</td>
</tr>
<tr>
<td>1023</td>
<td>-0.0323</td>
<td>1023</td>
<td>0.240347</td>
<td>0.480693</td>
</tr>
<tr>
<td>1024</td>
<td>-0.0434</td>
<td>1024</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
<tr>
<td>1025</td>
<td>-0.1575</td>
<td>1025</td>
<td>0.240346</td>
<td>0.480693</td>
</tr>
</tbody>
</table>
Michael E. Mann, Raymond S. Bradley, and Malcolm K. Hughes
Intergovernmental Panel on Climate Change (IPCC), Third Report, 2001
A golden age of visualization in news media
Popularizing visualization: Hans Rosling
https://www.youtube.com/watch?v=hVimVzgtD6w
www.gapminder.org
1. Defining visualization
2. The elements of a visualization
3. How visualization lies
4. Visualization for communication: essential principles
Layers of a visualization (from the bottom-up):

1. Scaffolding layer
   Features that support the content such as legends, scales, axes, etc.

2. Encodings layer
   The features that represent the data

3. Annotation layer
   Words on the visualization itself that explains data points or put them in context

4. The “me” layer
   Where is the reader in the data? Can readers find themselves? Can they create different scenarios? Can they even generate their own data?
Layers of a visualization (from the bottom-up):

1. Scaffolding layer
   Features that support the content such as legends, scales, axes, etc.

2. Encodings layer
   The features that represent the data

3. Annotation layer
   Words on the visualization itself that explains data points or put them in context

4. The “me” layer
   Where is the reader in the data? Can readers find themselves? Can they create different scenarios? Can they even generate their own data?
Track National Unemployment, Job Gains and Job Losses

By Andrew Van Dam and Renee Lightner

Winners and Losers: Job Gains and Losses

Track the number of sectors gaining or losing jobs each month. Boxes are shaded based on percentage change from the previous month in each sector’s payrolls.

https://graphics.wsj.com/job-market-tracker/
Track National Unemployment, Job Gains and Job Losses

Winners and Losers: Job Gains and Losses  Jump to National Unemployment

Track the number of sectors gaining or losing jobs each month. Boxes are shaded based on percentage change from the previous month in each sector’s payrolls.

https://graphics.wsj.com/job-market-tracker/
Layers of a visualization (from the bottom-up):

1. Scaffolding layer
   Features that support the content such as legends, scales, axes, etc.

2. Encodings layer
   The features that represent the data

3. Annotation layer
   Words on the visualization itself that explains data points or put them in context

4. The “me” layer
   Where is the reader in the data? Can readers find themselves? Can they create different scenarios? Can they even generate their own data?
Data visualization consists of mapping data onto properties of objects (commonly abstract shapes.) These properties are called “encodings”, “methods of encoding”, or “aesthetics”.

Encodings

Source: The Truthful Art
What encodings can you spot here?

The rate of violent crime was up 236% in Milwaukee between 1975 and 2015.

https://www.themarshallproject.org/2016/08/18/crime-in-context#.Q2NYkV3mN
Maps Show A Dramatic Rise In Health Insurance Coverage Under ACA
Layers of a visualization (from the bottom-up):

1. Scaffolding layer
   Features that support the content such as legends, scales, axes, etc.

2. Encodings layer
   The features that represent the data

3. Annotation layer
   Words on the visualization itself that explains data points or put them in context

4. The “me” layer
   Where is the reader in the data? Can readers find themselves? Can they create different scenarios? Can they even generate their own data?
The geography of voting — and not voting

By Ted Mellnik, Lauren Tierney and Kevin Uhrmacher  Oct. 23, 2018

People around the country can pass judgment on their government Nov. 6 in the first national election in two years — if they’re registered to vote, and cast a ballot.

Many will not, if recent history holds true. Only about 60 percent of U.S. adult citizens voted in the last national election, a turnout that ranked in the bottom third of the world's developed countries.

Layers of a visualization (from the bottom-up):

1. Scaffolding layer
   Features that support the content such as legends, scales, axes, etc.

2. Encodings layer
   The features that represent the data

3. Annotation layer
   Words on the visualization itself that explains data points or put them in context

4. The “me” layer
   Where is the reader in the data? Can readers find themselves? Can they create different scenarios? Can they even generate their own data?
Challenging readers before showing information

You Draw It: What Got Better or Worse During Obama’s Presidency

By LARRY BUCHANAN, HAEYOUN PARK and ADAM PEARCE  JAN. 15, 2017

Draw your guesses on the charts below to see if you’re as smart as you think you are.

Under President Obama, the unemployment rate ...

Draw the line for the Obama years.

BUSH YEARS

OBAMA YEARS

Letting people see themselves in the data, and compare themselves to others

How demographics predict party affiliation
Two questions about race and religion divide the population into distinct categories; others about education, age and sexuality split them further. But with the right mix of personal details, people who are separated initially can converge near the bottom of this tree.

Most Democratic
The group most likely to be Democrats are black women older than about 30.

Meeting in the Middle
These groups were initially split by the question about religious importance. But they converged near the middle because of their education, gender and more.

Reliable Republicans
White, religious Protestants lean Republican, but those who are male and live in the South are even more likely to identify with the party.

The Atlas Of Redistricting

By Aaron Bycoffe, Ella Koeze, David Wasserman and Julia Wolfe

Letting readers create their own scenarios

https://projects.fivethirtyeight.com/redistricting/maps/
1. Defining visualization
2. The elements of a visualization
3. How visualization lies
4. Visualization for communication: essential principles
Common misconceptions

1. “A picture is worth a thousand words”
2. “Visualization is intuitive”
3. “The data should speak for itself”
Common misconceptions

1. “A picture is worth a thousand words”
2. “Visualization is intuitive”
3. “The data should speak for itself”
You may think that you’re reading this chart correctly, but maybe you aren’t.
The danger of aggregating data too much, and presenting just averages and other simplistic summaries

Most places are pretty safe, and have likely remained down here (these aren’t real data points)
The danger of aggregating data too much, and presenting just averages and other simplistic summaries

Some places are so far up that they skew the national rate

Most places are pretty safe, and have likely remained down here (these aren’t real data points)
The Affordable Care Act ("Obamacare") was passed on March 2010.

Thousands of people in the workforce
It’s easy to use charts to confirm what we already believe

Liberal pundit: “Obamacare isn’t bad for the job market”

Thousands of people in the workforce

The Affordable Care Act (“Obamacare”) was passed on March 2010
—but a chart shows only what it shows, and nothing else

Liberal pundit: “Obamacare isn’t bad for the job market”

The Affordable Care Act (“Obamacare”) was passed on March 2010

Stimulus package passed on February 2009

Thousands of people in the workforce
But what if Obamacare hadn’t passed?

Thousands of people in the workforce

The Affordable Care Act (“Obamacare”) fails to pass on March 2010

Alberto Cairo • University of Miami • www.thefunctionalart.com • Twitter: @albertocairo
But what if Obamacare hadn’t passed?

Thousands of people in the workforce

The Affordable Care Act (“Obamacare”) fails to pass on March 2010
1. Defining visualization
2. The elements of a visualization
3. How visualization lies
4. Visualization for communication: essential principles
Length or height

Position

Area

Angle/area

Line weight

Hue and shade

Encodings

Source: The Truthful Art

Figures represented in all these graphics: 22%, 25%, 34%, 29%, 32%
Always think about the purpose of a graphic

Figure 2 - Main nationalities of arriving migrants – 2016

Greece

Syria 47%
Afghanistan 24%
Iraq 15%
Pakistan 5%
Iran 3%
Palestine 1%
Algeria 1%
Others 4%
Visual vocabulary

Designing with data

There are so many ways to visualise data – how do we know which one to pick? Use the categories across the top to decide which data relationship is most important in your story, then look at the different types of chart within the category to form some initial ideas about what might work best. This list is not meant to be exhaustive, nor a wizard, but is a useful starting point for making informative and meaningful data visualisations.

ft.com/vocabulary


Alberto Cairo • University of Miami • www.thefunctionalart.com • Twitter: @albertocairo
Height, length, position are appropriate encodings to enable accurate judgments: Comparisons, relationships, change over time, etc.

Alberto Cairo • University of Miami • www.thefunctionalart.com • Twitter: @albertocairo
Other encodings, such as area, color, etc., aren’t great to enable accurate estimates, but may work well if the goal is to see the big picture.
When possible and appropriate, let people choose
1. The elements of visualization
2. How visualization lies
3. Visualization for communication: essential principles

THANK YOU!