Data Visualization Workshop 2

Presented by Simone Betito

Feb 2020

Theme: Using Data Viz to Make Informed Decisions

Goals:

- Get a basic familiarity with Tableau Public
- Be critical of data provenance
- Learn some data visualization best practices

Agenda

- Go Over Assigned Work
- Tableau Public Example📈
- Data Visualization Best Practices
- Data Quality ☕
- Break🙏
- Different Data Visualization Mediums 🎨🎼
- Storytelling With Data📕
- In-class activity💃

Assigned Work

- How did it go?
- Pitfalls? Learnings?
- Anyone want to present their viz?

Since 2004, Low and Middle Income Countries Produce the Most C02 Emissions
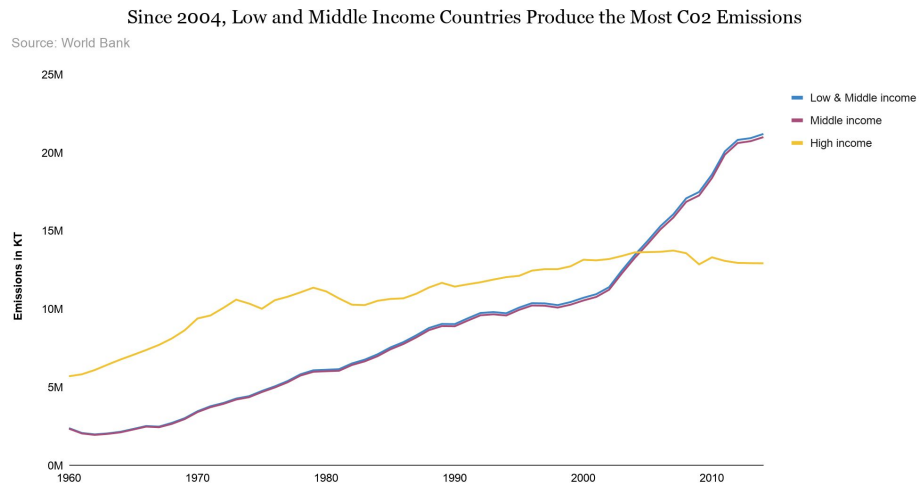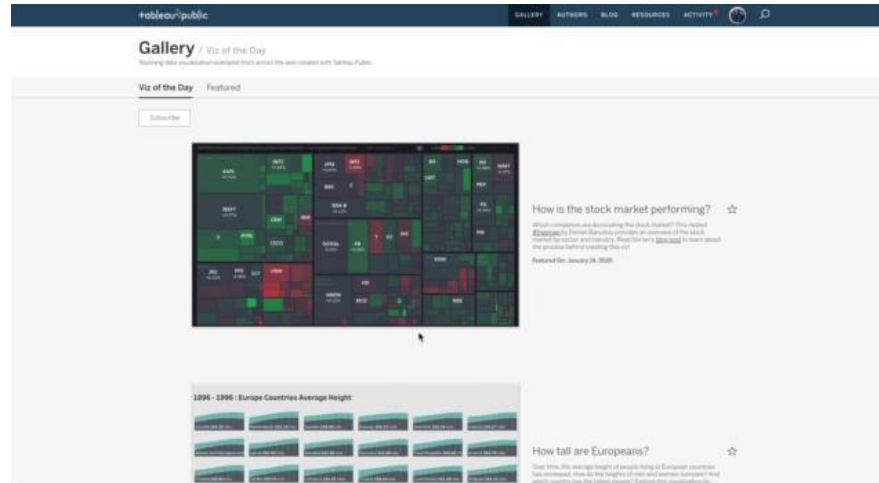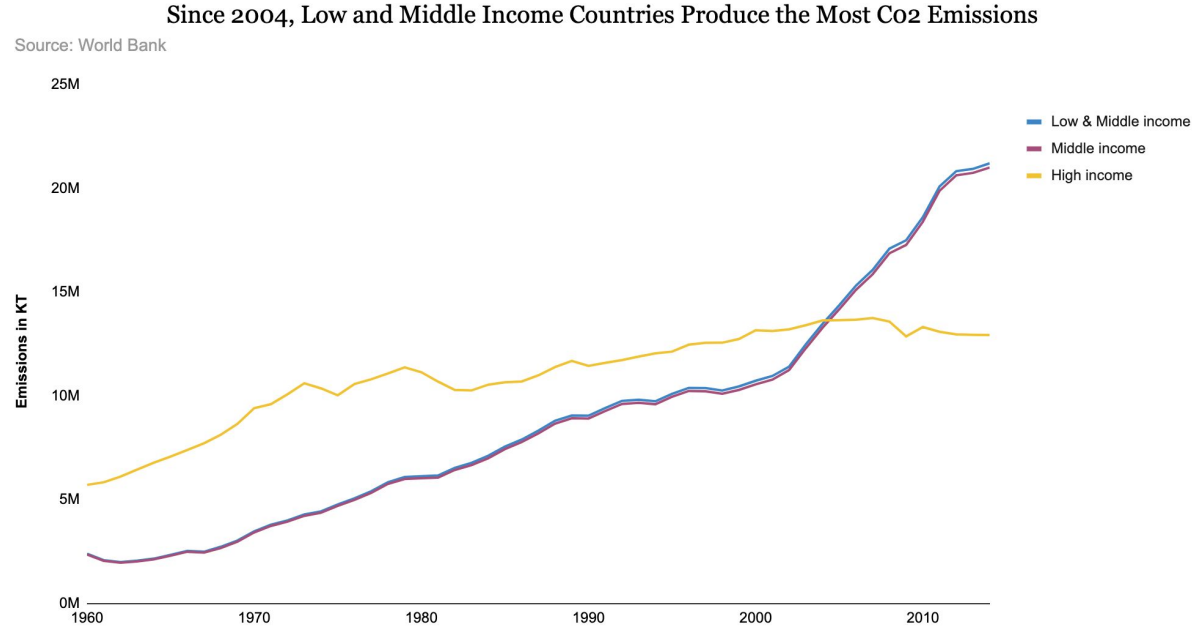
Source: World Bank

Tableau Public

Tableau Public

- Tableau Community (Makeover Monday), #VOTD
- Free
- Analyze & visualize data without code but more freedom than Sheets
- Good beginner resource: https://www.thetableaustudentguide.com/

What if we reproduced this in Tableau?

Since 2004, Low and Middle Income Countries Produce the Most C02 Emissions

Measures & Dimensions

- Common to have qualitative and quantitative measures named as this in BI tools (SAP, Looker, Tableau)
- Measures = Quantitative or numeric values, usually dependent (y)
- Dimensions are qualitative or categorical variables, usually independent (x)

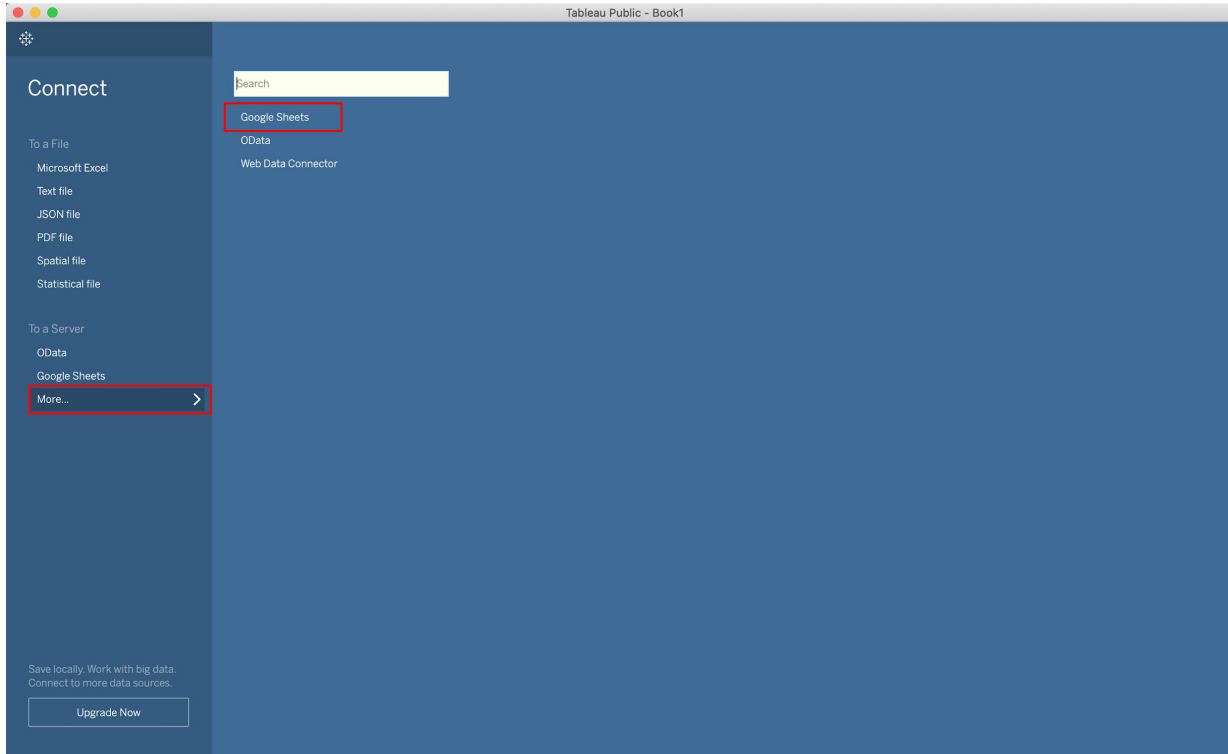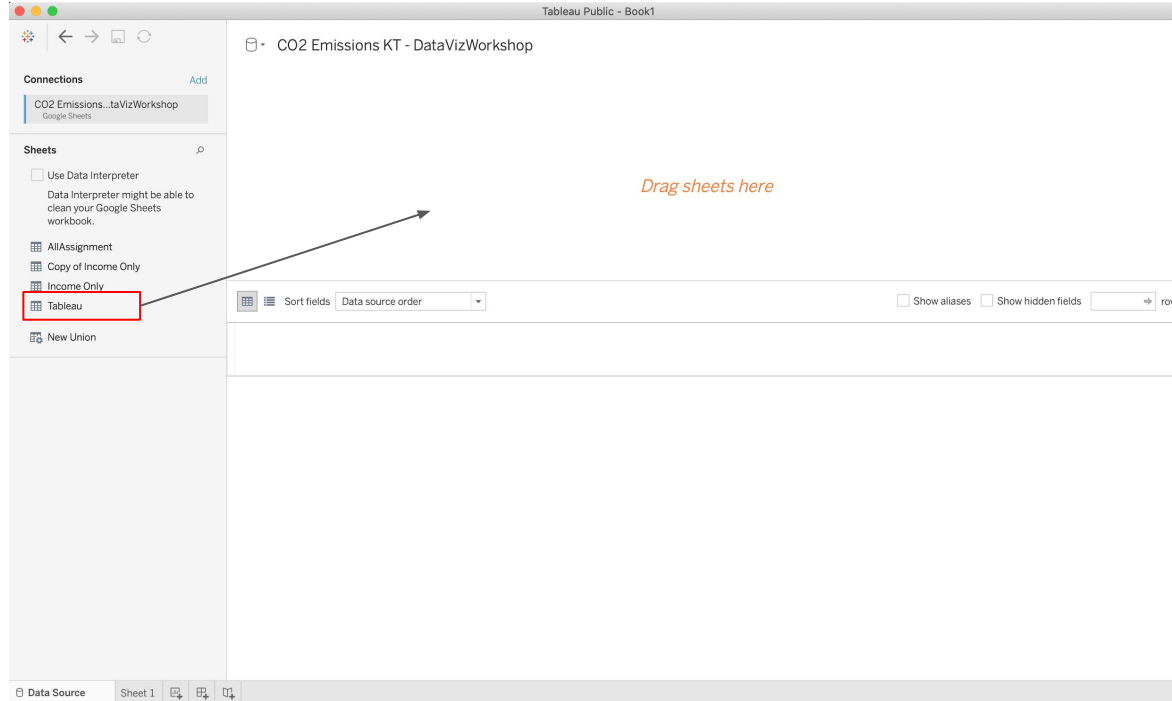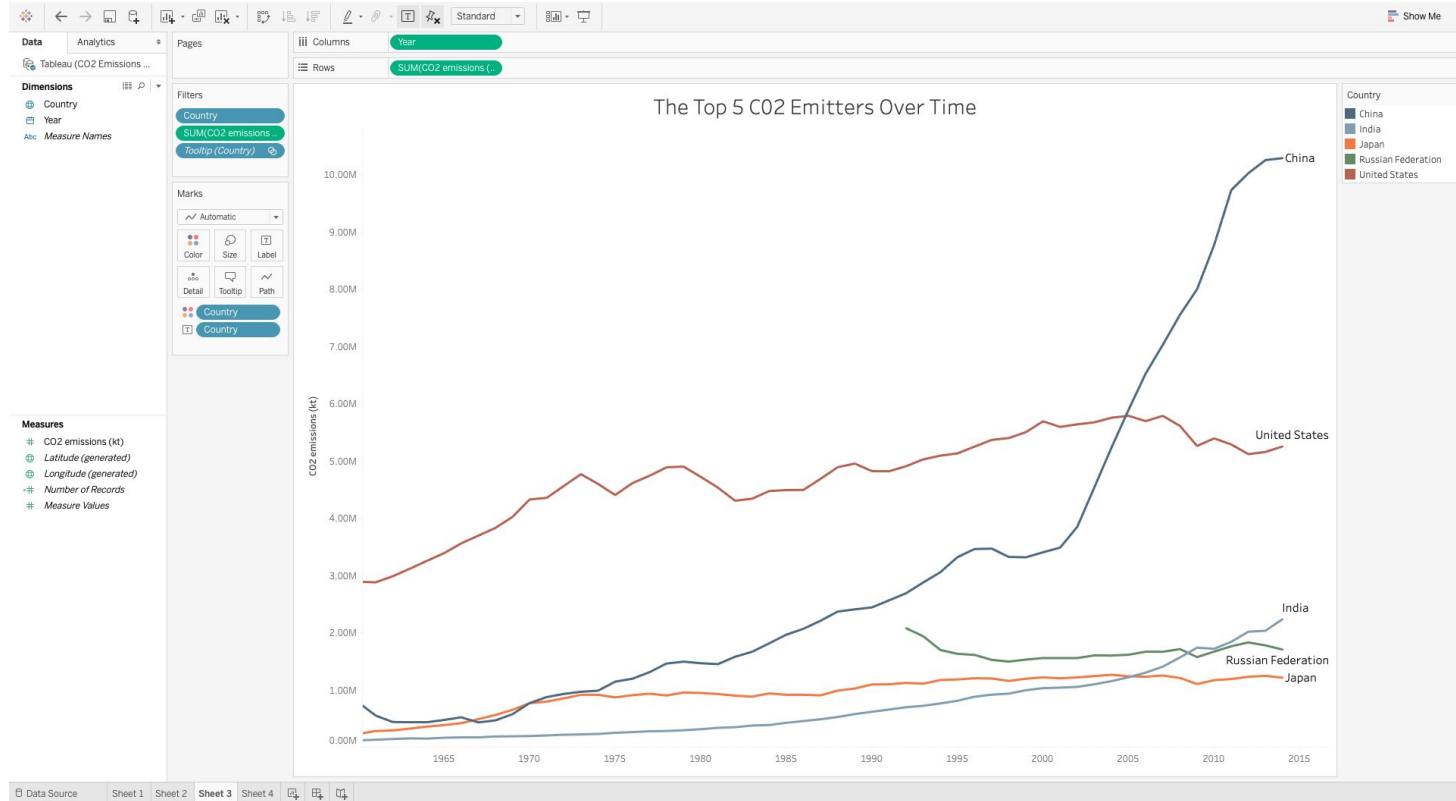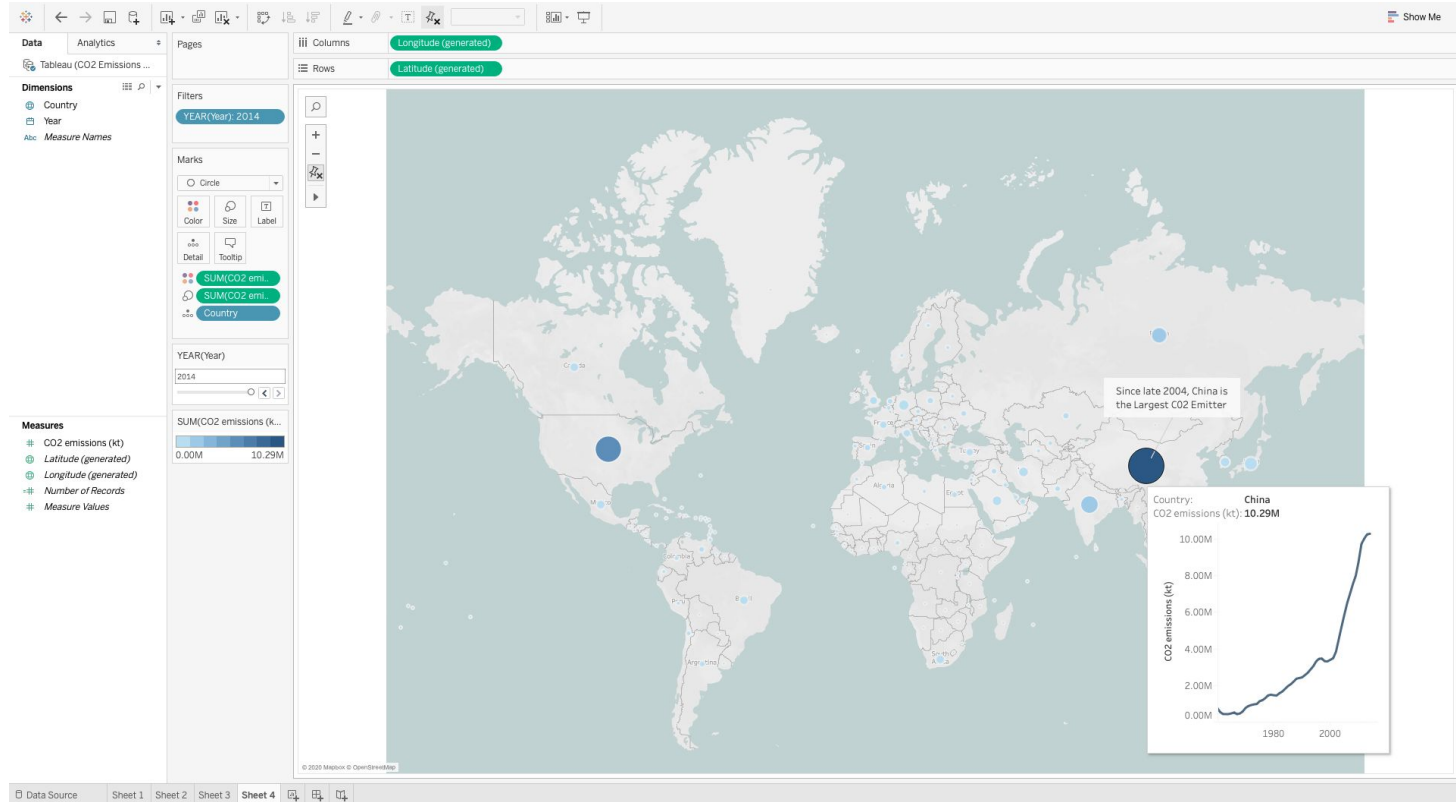| Measure or Dimension?	
Sales in US $	Measure
Months	Dimension
Order ID	Dimension

Tableau Data Source Page

Tableau Data Source Page

What if we reproduced this in Tableau?

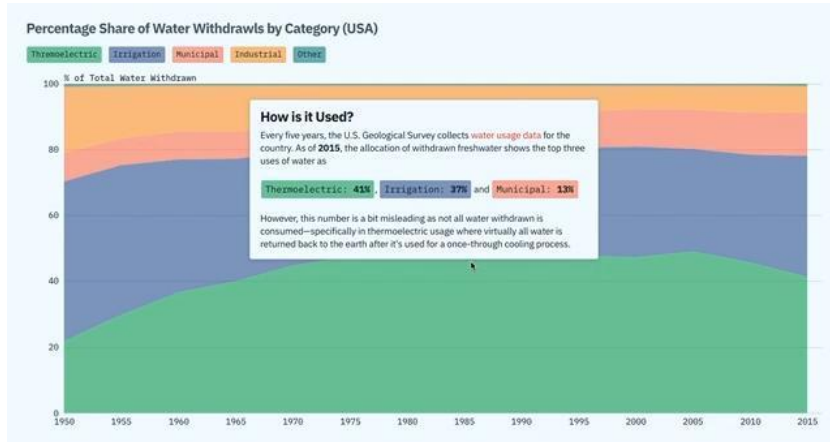# Other Way to Visualize in Tableau
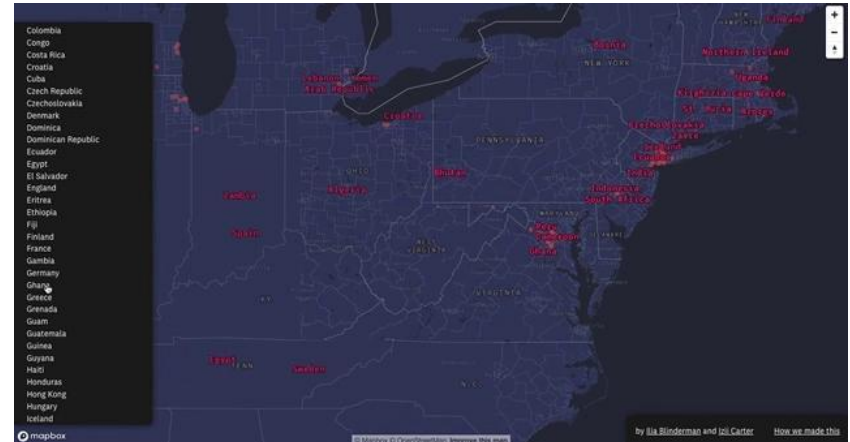
Data Visualization Best Practices

Data Visualization Best Practices

- Audience: Exploratory vs Explanatory
- Data: Pick the right chart for your data
- Design: contrast, colour blindness, comparing colours, semantics, gray
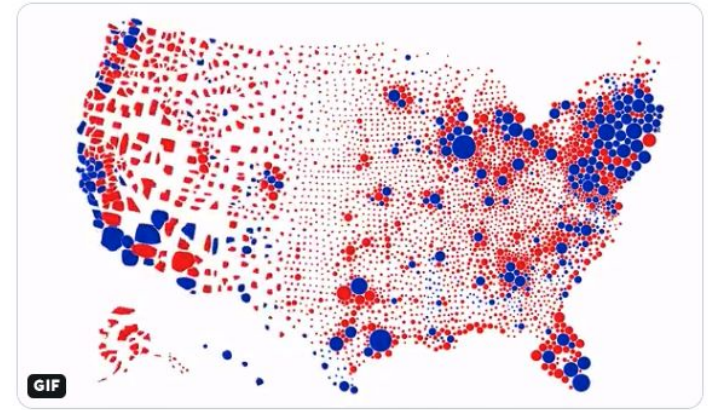- Iterative Process

Explanatory

Exploratory

Data Quality

Data Quality

- Are there nulls?
- Are the columns or rows summed?
- What's the source?
- How was the data collected?
- What are you deciding to show or hide?

- What might be a shady source?

Lara Trump @LaraLeaTrump · Sep 28, 2019
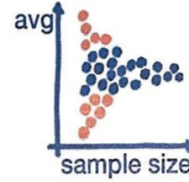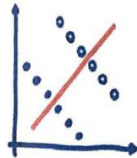Try to impeach this.

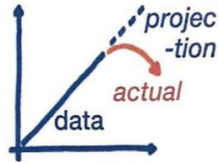# Data Quality

OUTLIER
skews average

VARIATION
the lower average has
a higher maximum

SMALL SAMPLES
are more skewed
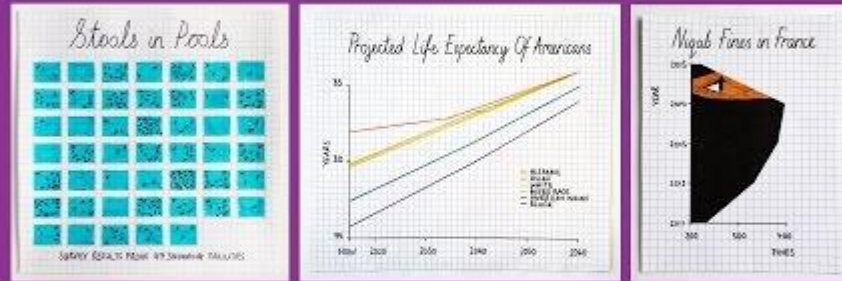by an outlier

AGGREGATION
reverses pattern

EXTRAPOLATION
assumes trend

NONEXISTENT
absurd summaries

Mona Chalabi - How to Spot A Bad Statistic

Mediums

Mediums

Left to right: Giorgia Lupi for & Other Stories, Jer Thorp's Herald/Harbinger, Adrien Segal data sculpture

Data Storytelling

Data Storytelling

"The interdisciplinary nature of data storytelling is quantitative and poetic, machine and human"
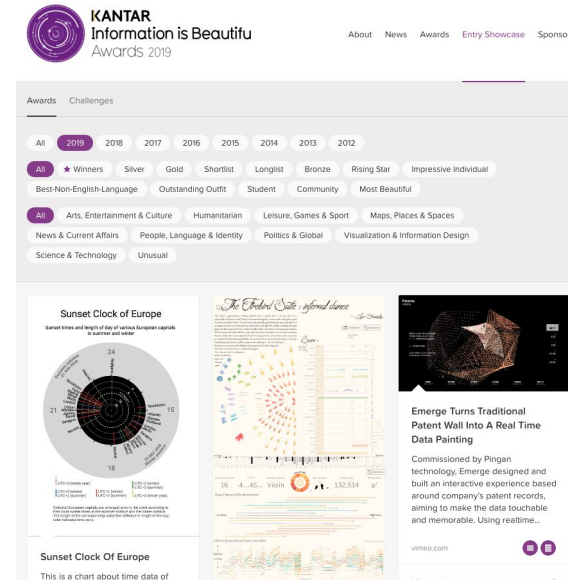
- What is the objective?/Call to Action?

- What do you want people to do with information?

- Do you want people to think differently? Drive Change?

- Impact Behaviour

In-Class Activity

- Go to https://www.informationisbeautifulawards.com/ -> entry showcase

- Choose a viz and jot down why you think it's effective
  - Think about:
    - Colours
    - Storytelling
    - Data Quality
    - Medium
    - Data Dimensions (size, position etc)
    - Chart Choice
    - Subject matter
    - Analysis

- Discuss in groups of two

Thank you!