Paul Rosen

paul.rosen@utah.edu
@paulrosenphd
https://cspaul.com

www.sci.utah.edu

THE UNIVERSITY OF UTAH

White Hat / Black Hat Visualization

The Problem

- It is tempting to think of data and data visualization as a neutral actor. An emphasis on a minimalist aesthetic — particularly through the use clean, precise geometric lines — lends an air of objective, transparent reporting that masks visualization's persuasive power.

- Given the growing ubiquity of visualization as a medium for recording, analyzing, and communicating data, we have a responsibility to examine how our design choices can influence the way a visualization is read, and what insights a reader walks away with.

Part 1: Select A Dataset

- The DEA Pain Pills Database (https://www.washingtonpost.com/national/2019/07/18/how-download-use-dea-pain-pills-database/?arc404=true). The Washington Post has published a significant portion of a database maintained by the Drug Enforcement Administration (DEA) that tracks every opioid from their manufacturer through to distributors and into pharmacies in towns and cities across the United States. This is an enormous database, and you can choose to work with it at any level of detail (e.g., state-wide, individual counties, or national summaries).

- Greenhouse Gas Emissions since 1990 (https://stats.oecd.org/Index.aspx?DataSetCode=AIR_GHG). The Organization for Economic Co-operation and Development (OECD) has compiled data for the emissions of all participating countries broken out by the pollutant (e.g., carbon monoxide, methane, etc.) and by different sources (e.g., energy, agriculture, etc.). The linked interface can be a little difficult to use, but you can access various slices of the data by either choosing alternate themes in the left-hand side menu or by customizing the pollutants and variables in the dropdown menus in the main view.

- Gender Equality Indicators since 1960 (https://github.com/ZeningQu/World-Bank-Data-by-Indicators/tree/master/gender). The World Bank tracks a number of different measures, including fertility rate, literacy, employment and ownership of businesses, and wages, to study the extent of gender equality around the world. The linked dataset curates a smaller subset of the overall set of gender indicators which you are welcome to use as well.

Part 2: White Hat Visualization

- With any tool, create a white hat visualization of the dataset
- A white hat visualization is
  - clear and easy to interpret for the intended audience (often the general population);
  - any data transformations (e.g., filters, additional computations, etc.) are clearly and transparently communicated; and
  - the sources of the data, including potential bias, is communicated.

Part 3: Black Hat Visualization

- With any tool, create a black hat visualization of the dataset
- A black hat visualization will exhibit one or several of the following characteristics:
  - The visual representation is intentionally inappropriate, overly complex and/or too cluttered for the audience;
  - Labels, axes, and legends are misleading;
  - Titles are skewed to intentionally influence the viewer's perception;
  - The data has been transformed, filtered, or processed in an intentionally misleading way; or
  - The source and provenance of the data is not clear to the viewer.

Part 4: Present your visualizations

- Pick one of your visualizations to present to the group
- Everyone must guess white hat or black hat