

Exercise 2 - Deceptive Visualizations

Data, CDAD-UH 1001Q, Spring 2022

Assigned: February 27, 2022

Due: March 13, 2022

1 Earnest & Deceptive Visualizations

In this assignment, you will find a dataset and create two visualizations to communicate the data.

One visualization will be an *earnest representation* of the data, whereas the other will be a *deceptive visualization* that aims to mislead the viewer.

The design challenge is to mislead without using obvious distortions or omissions in the deceptive visualization.

You will also provide a short write-up (no more than 4 paragraphs) describing your design rationale for both visualizations.

1.1 What is earnest?

A clear, easy-to-interpret, effective visualization where

- The visual encodings are appropriate and effective for the intended task
- Data transformations are clearly and transparently communicated
- The underlying data source (and any potential bias) is clearly communicated

1.2 What is deceptive?

- The visual representation is intentionally inappropriate or misleading
- Titles are skewed to intentionally influence the viewer's perception
- The data has been transformed or filtered in an intentionally misleading way
- The existence or source of bias in the underlying data is unclear

You want to trick your peers into believing that the visualization is legitimate and earnest. They shouldn't be able to immediately tell the earnest from the deceptive one.

In part 2 of this exercise, we will share your visualizations with the class to provide peer feedback and if they can't tell which is which, then you are a master deception artist!

For both visualization designs, start by choosing a question you would like to answer. Design your visualization to answer that question either correctly (for the earnest visualization) or incorrectly (for the deceptive visualization). You may choose to address a different question with each visualization. Be sure to document the question as part of the visualization design (e.g., title, subtitle, or caption) and in your assignment write-up.

Your write-up should contain the following information:

- The specific question each visualization aims to answer

- A description of your design rationale and important considerations for each visualization
- include in your write-up the two visualizations in the context of the write-up

Your figures should

- Be clearly titled or annotated with what they hope to show
- Have clearly labeled axes or legends if used
- Include A reference to the datasets used
- Provide **no** label as to whether they are the earnest or the deceptive figure.

Turn in: writeup.pdf: A short 1-2 page writeup describing your designs with the figures included in the text. The two figures as separate pdfs, jpgs, or pngs randomly named as netID-figA.pdf/png/jpg or netID-figB.pdf/png/jpg etc.

2 Where to find data?

We recommend using one of these datasets or sources to get started quickly, but you are free to use any dataset of your choice, especially if you are familiar with it.

You must use the same dataset for both visualizations, but you may transform the data differently or choose to address a different question for each design.

The World Bank Data, 1960-2018. You can browse and download the data by indicators (<https://data.worldbank.org/indicator>), such as climate change, economy, education, technology, etc. or by countries (<https://data.worldbank.org/country>).

Greenhouse Gas Emissions, 1990-2019. This dataset (https://stats.oecd.org/Index.aspx?DataSetCode=AIR_GHG) compiled by the organization for economic co-operation and development (OECD) tracks the emissions of all participating countries by pollutant and by source (e.g. energy, agriculture, etc.)

DEA Pain Pills Database. 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. Note that this is a very large dataset with many different facets, so you may want to focus on a particular area or set of attributes of interest (<https://www.washingtonpost.com/national/2019/07/18/how-download-use-dea-pain-pills-database/?arc404=true>).

Other data sources to look at:

- <https://www.data.gov/> - U.S. Government Open Datasets
- <https://www.ipums.org/> - Integrated Census Survey Data from around the World

- <https://www.fec.gov/data/> Federal Elections Commission - Campaign Finance Expenditures
- <ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/daily/> NOAA Daily Global Historical Climatology Network Data
- <https://www.yelp.com/dataset> - Yelp Open Dataset
- <https://github.com/fivethirtyeight/data/> Data and Code behind the Stories and Interactives
- <https://github.com/BuzzFeedNews> Open-source data from BuzzFeed's newsroom
- <https://www.kaggle.com/datasets> Datasets for Kaggle contests
- More datasets curated by Mike Freeman <https://observablehq.com/@mkfreeman/teaching-datasets>

Grading

The assignment score is out of a maximum of 10 points. We will determine scores by judging the soundness of your visualization designs, the duplicity of your deceptive visualization, and the quality of the write-up. Here are examples of aspects that may lead to point deductions:

1. Obvious identification of the earnest and deceptive visualizations.
2. Ineffective visual encodings for your stated goal.
3. Missing indication of the main analysis question.
4. Missing or incomplete design rationale in write-up.

We will reward entries that go above and beyond the assignment requirements to produce effective (and deceptive) graphics. Examples may include outstanding visual design, effective annotations and other narrative devices, exceptional creativity, or deceptive designs that require the write-up in order to properly identify the misleading design components.

Submission

This is an individual assignment.

1. You will package your solution into a folder with the title: 'Ex2-netID-firstname'. (e.g. Ex2-aa175-azza is the name of the folder I would submit).
2. This folder will contain exactly **three** files: (a) an image file A (netid-figA .jpg, .png or .pdf), (b) an image file B (netid-figB .jpg, .png, or .pdf) a write-up file (.pdf that includes the figures inside it described in the context of the write-up).
3. **Zip this folder** and submit a .zip folder via DropBox at the following link: <http://bit.ly/Data-F22-Ex2>

We will not grade any submission that does not strictly follow the submission rules.