Assignment #2: Misleading Visualization
CS-GY 6313 B: Information Visualization

Given the ubiquity of visualizations in the news, especially in the digital age, it is crucial that you understand what makes a visualization honest or deceptive so that you can identify when someone is trying to mislead you with a visualization. In this assignment, you will identify a dataset of interest and make two visualizations from it. One visualization will be designed to be an honest, effective visualization that communicates its message earnestly, while the other is designed to be intentionally misleading without using obvious distortions or omissions.

Assignment (15 points)

You will create two static (single image) visualizations of your data set. One visualization should aim to communicate insights about the dataset earnestly and effectively, while the other visualization should aim to mislead the viewer to arrive at the wrong conclusions. You will also provide a writeup (no more than 3 pages, single spaced, not including images) that describes your design decisions and rationale.

We will consider an *earnest* visualization to be one for which:
- Data transformations are clearly and transparently communicated
- The visual encodings are appropriate and effective for the intended task
- The visualization is clear and interpretable by the general population
- The underlying data source (and any potential bias in the data) is clearly communicated

We will consider a *deceptive* visualization to be one which:
- The visual representation is intentionally inappropriate or misleading
- The existence or source of bias in the underlying data is unclear
- The data has been transformed or filtered in an intentionally misleading way
- Titles are skewed to intentionally influence the viewer's perception

For the earnest visualization, your goal is to be as clear as possible to allow your viewers to answer your intended question as easily as possible. For the deceptive visualization, your goal is to trick the viewer (including the course staff) into believing that the visualization is earnest and legitimate. **It should not be immediately obvious which of the two visualizations is the deceptive one.** Subtle ineffective choices in the design should require close and careful reading to be identified.

For the deceptive visualization, misleading strategies are fine but outright lying is not. For example, sketchy, unreliable or untrustworthy input datasets are discouraged, but

misleading omission, filtering, or transformation of trustworthy data records is fine. Deliberate lies in the title, axes, labels, or annotations is discouraged, but technically true/relevant but otherwise misleading text in the visualization is fine.

For both visualization designs, you should 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 are allowed 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

Whether the visualization is the earnest or the deceptive one should only be clearly stated in the write-up. **Do not name the images of the visualizations such that it is obvious which one is which.**

Data sources

You are required to use the same dataset for both visualizations. We have provided a couple of data sources that you can use if you would like to get started on the visualizations sooner rather than later. We have chosen datasets that cover politically-charged topics since these are the datasets for which misleading visualizations are most likely to proliferate. You are also free to choose your own dataset if you wish (here is a list of resources).

The World Bank Data, 1960-2018

The World Bank has tracked global human development by indicators such as climate change, economy, education, environment, gender equality, health, and science and technology since 1960. You can browse the data by indicators or by countries. Click on an indicator category or country to download the CSV file.

Greenhouse Gas Emissions, 1990-2019

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.). While the linked interface can be somewhat hard to navigate, you are free to

browse alternate themes from the panel on the left (such as education or health). You can download the data by selecting "Export" at the top of the chosen table.

Data: [Greenhouse Gas Emissions](#)

<u>DEA Pain Pills Database</u>

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.

Important Note: In order to download and use this data you may need to enter your email in order to access the Washington Post Data Access page.

Data: [Washington Post Data Access Page](#)

Extra Credit (3 points)

For extra credit, find a visualization online that you believe is deceptive (either intentionally or unintentionally). Write a separate report for this visualization, explaining why it is deceptive. You should include a screenshot of the visualization in this report, as well as a link to the source where you found it.

Grading

The assignment score is out of a maximum of 20 points (excluding extra credit). 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:

- Obvious identification of the earnest and deceptive visualizations.
- Ineffective visual encodings for your stated goal.
- Missing indication of the main analysis question.
- 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 Instructions

All assignments are to be completed individually (see syllabus for further details). The assignment is due on **Wednesday 10/16, by 11:59pm**. We will be discussing submissions in class, so be sure to avoid a late submission if you can.

You should submit .zip file containing the following components to Brightspace:
- A pdf of your written report (written using the provided LaTeX template)
  - We suggest using overleaf.com to write your reports
- All of the code necessary to recreate your visualization(s)
- Two image files (PNG or JPEG), one for each visualization. **Do not give them file names that make it obvious if the visualization is the earnest or deceptive one.**

Credit goes to the University of Washington's data visualization course for the assignment idea.