Lab 2 – Fall'21

SI649/EECS548

Due 9/14 at 9pm. To turn in: one PDF document with your answers (two visualizations and two short explanation paragraphs).

This week's lab will be quite a bit shorter because we don't have lecture. We're going to have you (1) prepare your Python environment for the labs that start next week (please follow the instructions in the video), and (2) find us a good visualization and create an ugly one, (3) if you want to get ahead, please watch the video on the Grammar of Graphics. It will make Lab 3 shorter.

Installing Python and Altair

We've created a short video for you to watch on how to do this. You can get away with just using Google Colab for the first couple of weeks, but we'd encourage you to install and use Python locally now. You'll need it for the later labs.

Colab: To use Google Colab, simply go to: http://colab.research.google.com/. You don't need the professional account, but please explore the site. The video explains some things you can do, but you can find additional resources on the web (as a warning, many of them are far more advanced than we need).

Local Python/Jupyter/Altair:
If you're confused about any of these steps or want to know more about them, please watch the video.

1) We recommend you install the Anaconda distribution of Python for this. If you have it, go to step 2. If you don't, download and install from here: https://www.anaconda.com/products/individual
2) Once installed, we suggest creating and using a custom environment for this class. This will help maintain a clean setup with only the packages you need. To do this, type the following in your command prompt:

```
conda  create --yes -n si649f21
```
Which will create an environment called si649f21. Once this is done, activate the environment (you will need to do this every time you open your command prompt):
```
conda activate si649f21
```

3) At this point we can start installing packages. If you are using conda, you'll mostly be using the conda package manager. Otherwise you will use the "pip" command which should have come with your Python installation. Let's start by installing jupyter (our notebook environment):

```
conda install jupyter
```

4) Next, we want the latest version of altair so we will use a repository called conda-forge:

```
conda install -c conda-forge altair vega_datasets
```

5) You may as well install streamlit and altair_saver at this point (we might need to come back later in the semester to update these as I can't predict which version is the latest in the repository). Both of these need to be installed through pip (don't ask):

```
pip install  streamlit altair_saver
```

6) At this point you should be able to type "`jupyter notebook`" to start coding (and test your setup).

Good Vis and Ugly Vis

Your actual assignment this week is to find one visualization (a "good vis" and to create one "the ugly vis").

1) Find us a (single!) really good visualization. There are many good visualizations out on the web. Start with one of the usual suspects (Economist, New York Times, Wall Street Journal, Financial Times, Fivethirtyeight, etc.) and see if you can identify a visualization you feel is excellent. For this activity please pick a *static*, *single*, visualization (small multiples like Figure 2 are probably ok). There is no need to find anything super-fancy. A simple line chart like the example below will work (Figure 1).

Figure 1

Write a paragraph describing what makes the visualization good. You may not have the right vocabulary yet (that's what the rest of the class is about) but try to isolate why you think the visualization is effective.

Please do not look for examples in visualization papers, vis blogs, vis textbooks, etc. We want to know *your* thoughts.
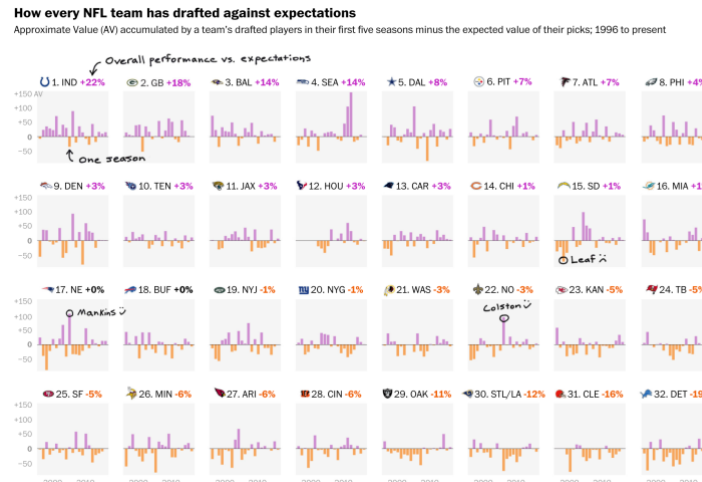

How every NFL team has drafted against expectations
Approximate Value (AV) accumulated by a team's drafted players in their first five seasons minus the expected value of their picks; 1996 to present

Figure 2

https://www.washingtonpost.com/graphics/sports/nfl-draft-values/

2) Watch our video on ugly visualizations. Then, produce an ugly visualization… like really ugly… the uglier the better. This does not mean too simple or not usable. We mean ugly (see Figure 3). You are welcome to use any dataset and produce this through any tool (Tableau, Excel, whatever). ***Describe your thinking about why this is "ugly" and why it might not be an effective visualization (if you think it*** is ***effective, that's fine, tell us why).*** Again, you might not have the vis "language" to talk about this yet. Use your own words.
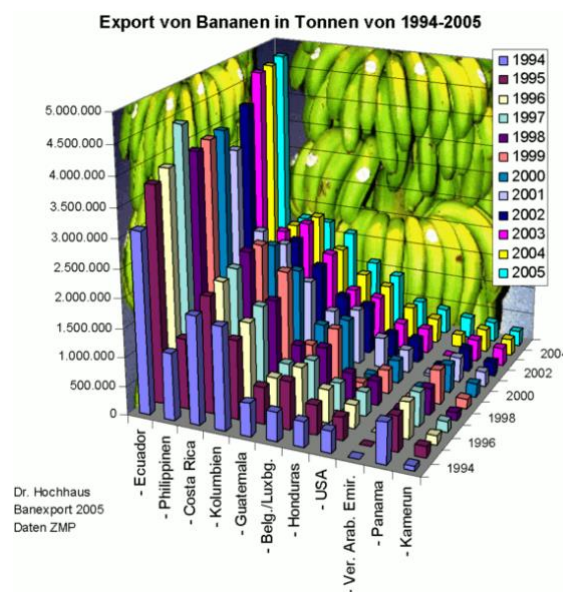

Figure 3