Exploratory Visual Analysis with Tableau

The Federal Election Commission (FEC) maintains a database of the financial activity of Political Action Committees (PACs), including funds raised, money disbursed to committees & candidates, and operating expenditures. See the FEC site for a full listing of available data sets.

For this activity, we'll be looking at the *operating expenditures* of the 2020 election cycle: where committees and candidates are spending their money.

Download the dataset (~50MB zip file).

This dataset consists of three tables: the operating expenditures, the political committees, and the candidates. Follow these per-table links for more details about the columns each table contains.

Open Tableau, and import the tables like so:

expenditures.txt	——————	committees.txt	——————	candidates.txt

The diagram Tableau depicts is its version of a data join. Essentially, we're saying that for every expenditure, we want to find a matching political committee; and, for each political committee, we want to find the candidate they're spending on behalf of.

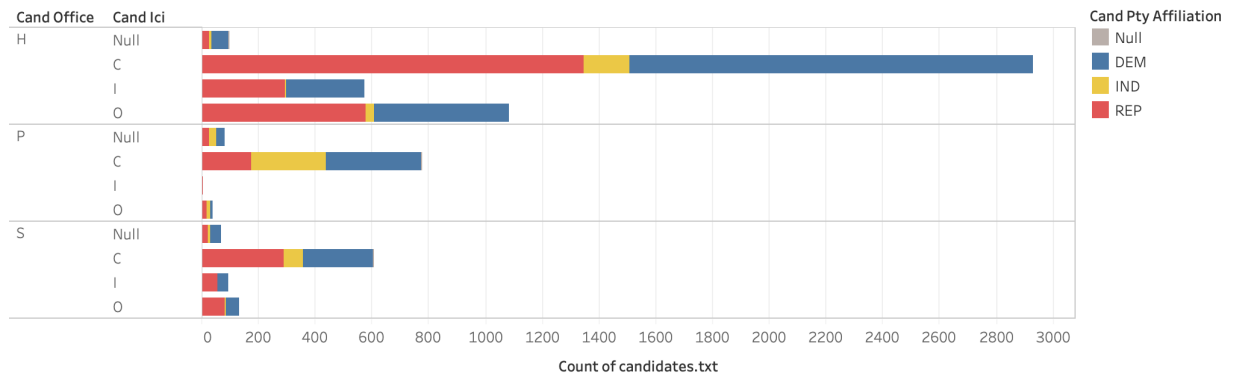**Work through this activity in your breakout room.**
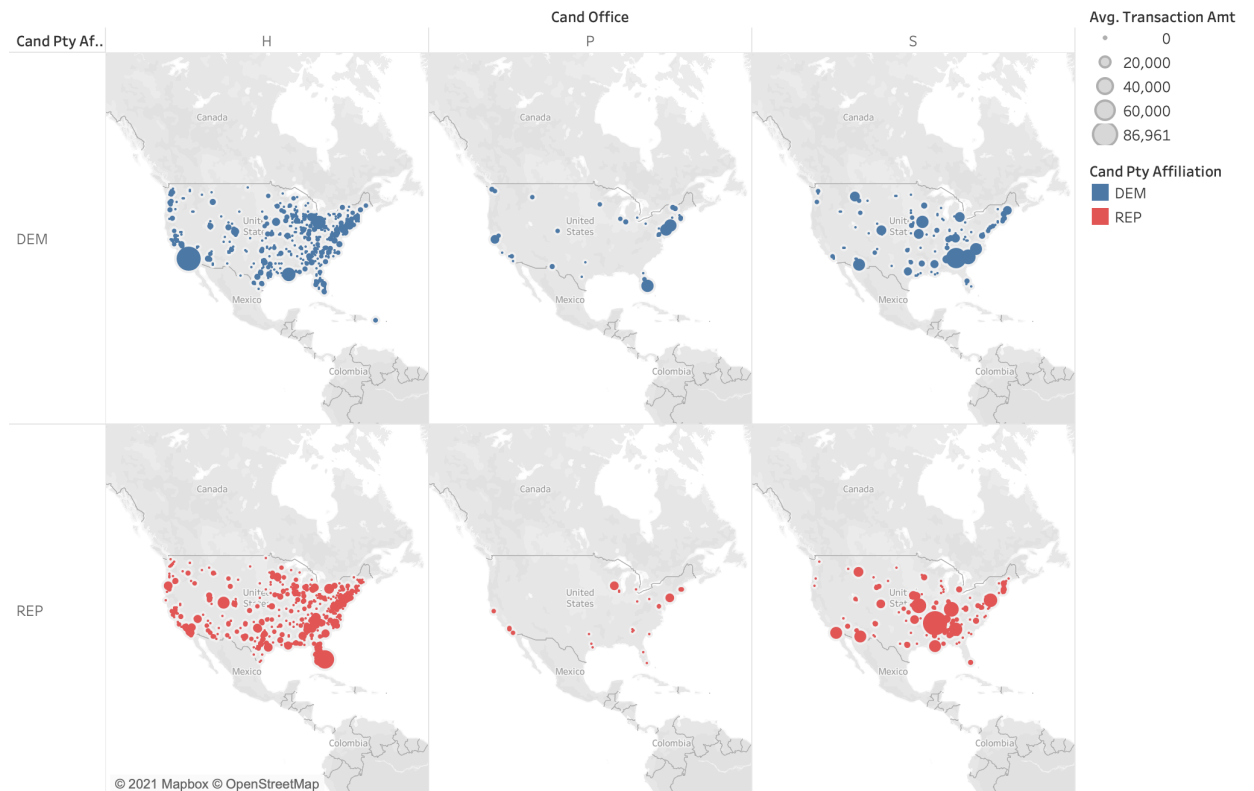
Table of Contents

Warm Up Activity (~10 mins)

This is an extremely large dataset, with many rows and columns. To familiarize ourselves with the dataset, let's first warm up by getting a sense of the shape of our data:

1. Distribution of Candidates by Office, Incumbency, & Party

2. Average spending across the country by candidate office

Assess Data Quality (~10 mins)

Now that we're a little bit more familiar with Tableau, let's start our exploratory visual analysis journey. Being the good data scientists that we are, we'll start by examining the quality of our data. Things to look for include potentially incorrect/surprising values, duplicates, data entry errors, etc. Looking for missing values is also a great place to start but, as you'll notice, there are a *lot* of missing values so try to dig deeper and see if you can characterize what these missing values may represent.

Feel free to explore a range of visual forms (i.e., you are not restricted to exploring only univariate summaries though they present a solid start).

For each visualization that gives you an interesting insight:
1. Take a screenshot (or use the *Worksheet > Copy > Image...* menu item).
2. [Add a slide to this slide deck](#), with a 1-sentence caption describing what you learned about the data quality from the visualization.
3. Remember to sign your captions with your names and your Kerberos usernames.

Explore a Question (remaining time)

Phew! Now that we have a better sense of the shape and quality of our dataset, we can dive deeper to explore a particular hypothesis we may have. Pick **one** of the following questions (or pose your own question under a new header) and use Tableau to construct visualizations to as thoroughly answer the question as possible in the alloted time.

As before, for each visualization that gives you an interesting insight, copy and paste it [into the slide deck](#) under the appropriate section and add a caption. What did you learn from the visualization? What do you want to explore next as a result? Remember to sign off with your name and Kerberos!