What kind of data pre-processing have you done on the server side?

All of the data has been preprocessed on the server side. For the dashboard I've utilized Python's pandas library to obtain the following data:

1. The total amount of donations (not including refunds) for each party.
2. Monthly donation statistics, specifically the total amount of money earned through donation each month, for each party.
3. The interquartile range (i.e. 25th percentile, 50th percentile, and 75th percentile) for each candidate in terms of donation amounts.
4. The total distinct donors for each candidate.
5. The leading party (w/ respect to total donation amount) in each of the different states in the given data
6. The monetary support from the unemployed for each of the parties.

What transformations on the data do the client-side d3 apply to generate the plots?

The data has mostly been preprocessed such that very little to no processing is required to generate the data on the client-side. Other than selecting the data from the given csv files, no true transformations were necessary. This was done so that the processing on the client side would be extremely fast and computation would not hinder the performance of the dashboard.

What is the intent of each visualization? What are the right visualization techniques to present this information effectively? (e.g., which type of chart or graph is used? why?)

I will describe my response to the third and fourth questions for each of the following six visualizations in my dashboard for cohesiveness in my response.

1. <u>Donation Percentages Donut Chart:</u>
   With the donut chart, I attempted to show the relative percentages of the total donation amount that each party received. Ideally, I wanted the viewer to be able to make visual comparisons between donation portions of parties in a simple manner. The donut chart provided this. Using the ring of the donut chart, I was able to fill in a fraction of the donut based on the fraction of the total donations a party received. The inner circle of the donut chart allowed me to augment the display of information through textual labels which show actual percentages.

2. <u>2011 Donation Trends Area Chart:</u>
   With the stacked area chart, I attempted to show the trends for each party for donation amounts each month. Specifically, I tried to achieve two tasks:
   - Show the rise/fall of monthly donations for each party month to month.
   - Visually compare between parties the relative increase/decrease in donation amounts as well as total amount for a specific time period

The areas in the area chart allowed me to achieve both. Utilizing the visually computable areas in the chart, one can compare the two areas and obtain information about changes in donation months for a specific party itself or relative to other parties as well.

3. Unemployment Support Donut Chart:
   This donut chart attempts to focus on the unemployed. In the given data, I've defined unemployed as anyone who has "NOT EMPLOYED" in either the occupation or employer column. Through this visualization, I intend to show whom the unemployed support by the money they have donated to the parties. Comparing the percent of total money donated by the unemployed for each party gives a potential perspective from those who are struggling or looking for opportunities if they want change in the White House or not. I've utilized a donut chart to display differences in percentages, because the donut chart is resistant to scaling effects on the display and comparison of the fractions of the total donated money for each party.

4. Total Donors per Candidate Bar Chart:
   This bar chart attempts to display the number of distinct people that have donated for each candidate. I utilize the same assumptions used for distinct donors as in assignment 1. The visualization allows for a different perspective on the data as the number of donors for a candidate can show support that analysis by the total donation amount hides. The bar chart was effective for this visualization, as the horizontal bar allows the viewer to gauge the absolute number of donors and relative disproportions for each candidates based on bar size and axis value.

5. Donation Amounts Interquartile Range Chart:
   The interquartile range chart attempts to show the "midspread" (specifically the 25^{th} percentile, the median, and the 75^{th} percentile) of the donation amounts data. The chart intends to allow the viewer to make hypotheses or predictions about candidates based on spread of donations such as class of donors (i.e. middle class, upper class) for each candidate. The interquartile range chart could have been replaced with box plots, as it displays the statistical dispersion in donations that this chart attempts to show. However, for the dashboard, I felt the simplicity of the interquartile range, which only shows the "midspread" of the data, was more visually suitable.

6. Leading Party (by Donation Amount) US Choropleth Map:
   The choropleth map attempts to display to viewers the current leading party in terms of total donation amounts received in each state. This provides one perspective as to which of the three parties is the most popular in each state, although only in monetary terms. The choropleth map simplifies the concept of displaying preference for each of the fifty states by coloring in each state with a party color. The map enables the viewer to quickly locate a specific state and the party preference in that state in comparison to having to possibly sift through 50 state labels for the same information. In addition, the map allows the viewer to make interpretations about specific regions of support for the party in the United States. Thus, the map's main strength is efficient geographic display of party preference.