Final Project Plan - Analyze NYC Election Contribution Data

Data Sources
data.gov:
(multiple related csv files w/ at least one file containing >1000 records - 4pts)
2001 - Candidates & Contributions
2003 - Candidates & Contributions
2005 - Candidates & Contributions
2009 - Candidates & Contributions
2013 - Candidates & Contributions

data.gov
(CSV or JSON file you haven't used before with > 1000 records - 2pts)
2005 census - to get borough/area information

Wikipedia API
(Web API you haven't used before that requires no authorization - 3pts)
Retrieve candidate information:
Birthdate/Age
Political Affiliation
Years in Office

NYT API
(Web API you've used before - 2pts)
Search for articles on candidate either collectively or in given year

Total Challenge Score: 11? Or 9. Not exactly sure what the 2005 census counts as.

Create Database
Database will have 3 tables: Contributors, Candidates, and Boroughs

Contributors:
Will contain contributor Id, city, state, zip code, occupation (if applicable), amount donated, the candidate/recipient Id, and maybe the borough id (I honestly can't figure out if this is necessary since I have the city? I think I will include though to be safe. It seems to make the most sense.)

Candidates:
Will contain candidate/recipient Id, borough Id, and name

Boroughs:
Will contain borough Id and affiliated name

Function of Program
Allows users to search and sort through 7 sets of election contribution data to compare and contrast different contribution amounts to candidates over years. It also allows users to see how many contributions each candidate received and what those amounts were, as well as see the areas those contributions came from. Users will also have the ability to sort this information in a variety of ways, which will provide them with different information and context to consider while analyzing these contributions.

I originally wanted to analyze the contributors further, to see if they contributed multiple times or to multiple candidates. The dataset, however, was not set up in such a manner that would allow me to do this accurately.

Basic Ideas of functions are as follows:

Candidates
return candidate name, political affiliation, age, years in office, total contributions amount, num individual contributions, borough area
-default: returns all candidates sorted contribution amount
    can also sort by year and total contributions
-can also return candidates in specific election by choosing election year
-can also return candidates in specific area
    this can be done over all years or specific year
-can return candidates with specific political affiliation
-can choose specific candidate (by choosing number) and return individual donations to that candidate for predetermined year
    return city, state, donation amount of contributor
    can sort by donation amount ascending or descending/top-bottom

Contributions
returns contribution amount, contributor city, state, occupation if applicable, and recipient
default: returns 100 highest contributions from previous election year sorted by amount
    can also sort alphabetically by city
can also return all contributions for given election year, or all election years- default sorted by contribution amount
    can also sort by area, and highest/lowest contributions
can also return info on contributions MADE from people in a specific area

Maps
Bar graph - to compare candidate contributions in specific year
Scatter Plot - to display candidate location paired with contributor locations for specific year - hovering over individual dots reveals donation amount
Column graph - comparing total democratic vs republican contributions in a year
Scatter plot - reveal contributions made in a specific borough
Scatter Plot - reveal democratic and republican contributions for specific year throughout
Scatter Plot - contributions made within specific price range
Average contributions per party over 3 elections
Average contributions per borough over 3 elections?
(Not sure if I will be doing all of these. I have to see what works/makes sense).