Project Deliverable 2
Evelyn Siu
Cole von Glahn
Carolyn Vilter

1. Please provide a one paragraph description of the goals of your project. You can list the same description from the previous deliverable or provide new details about aspects that have changed since Week 4.
    - Goal 1: Map the partisanality (% red / % blue) of the US by county in a heatmap using Bokeh & Geopandas.
    - Goal 2: Allow users to enter a county of interest. Display partisanality and demographics (education level, immigration status, race, income, etc.) for that county.
    - Goal 3: Allow users to enter demographic criteria of interest. Display resulting county data through a Django interface.
    - Store and query all data via a SQL database.

2. For each source of data that you expect to use, please list the source of data, who will be responsible for collecting data from that source, and a date by which you expect the work of gathering the data from that source to be complete.
    1. Election data – Via Harris professor Anthony Fowler. Democratic/Republican vote totals by county, entire US, for every presidential election from 1984 - 2016 (our focus will be 2016). Carolyn. Already secured.
    2. NHGIS Tiger Line shapefiles – Via IPUMS (census data aggregation) website. County level, entire US. Carolyn. Already downloaded. Carolyn also implementing an API approach to accessing the data - expecting to complete by 3/5.
    3. Census Bureau American Communities Survey – Via data.census.gov. County-level summary data, entire US. Various demographic variables: educational attainment, immigration status, race, income, etc. Expecting to finish choosing variables and joining tables by 2/25.

3. Please give a brief sketch of the work that needs to be done to complete your project (other than data collection), include a description of which team member(s) will be responsible for completing this work and the expected timeline for completion

    Tasks:
    - Database will be built and populated using SQL. Core tables include: presidential vote share by county, demographic information by county.
    - Heat map needs to be created using Bokeh & GeoPandas. Shape files are merged with elections data to allow the election data (rvotes / dvotes) to be surfaced when the user hovers over the county.
    - Web interface needs to be created using Django. Bokeh map will be embedded into Django, and there will be text fields that the user can enter into that will initiate the query to the SQL database (similar to PA3).

    Responsibilities and Timelines:
    - Carolyn is merging demographic data sources and doing other data cleaning an manipulation, including implementing a measure of county demographic similarity (3/5).
    - Cole is developing the database and query structures (3/5) as well as linking the Django front-end to the SQL back-end (3/12).
    - Evelyn is building the Bokeh/Geopandas map and embedding the Bokeh map into the Django interface (3/11).

    - All teammates are responsible for putting everything together (requirements.txt, install.sh/ file, QA/testing) (3/13).
4. Include any additional information you wish to provide about your project.
    - Would like to be able to update the map in real-time based on the counties of interest to the user, but am worried it would be impractical because of the long map load-time (user would have to wait ~1 minute between searches).
    - Also would like to give a heads up that the map takes a long time to load.

5. If you need additional feedback from me then please make sure to include that in your report.