

## CAPP30122 – Project Deliverable #2

### python\_parser\_tongues

#### **TEAM MEMBERS (Team: python\_parser\_tongues):**

Ken Kliesner (kenkliesner)

Gabriela Ayala (mariagabriela)

Fabian Araneda Baltierra (baltierra)

Stefan Manuel (stefmanuel)

## **Deliverable #2: Project Check-In**

**Please come back here in week 7. We will provide a link to provide additional information about your project progress. We will also require additional information about your data sources and team member responsibilities.**

**Please submit to your group repository a single document named `proj_d2.pdf` that provides an update to your current status of your project. You will answer the same questions from the previous survey but in written form. Make sure to answer the following questions:**

- 1. Please provide 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.**

Project Modeling Covid-19 New Vaccinations vs US Political Party Affiliation. The goal would be to analyze correlation between vaccination status (first dose, fully vaccinated, or boosted) and US general election data. We will review both sets of data at the county-level and run a regression on vaccination rates to identify all relevant control and explanatory variables (ie, possible controls: state density, income, race). There will be an interactive data visualization component with a US State/County Map including components from each data source.

- 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.**

Vaccination Status by State (API):

*Gabriela (3/1/2022)*

<https://data.cdc.gov/Vaccinations/COVID-19-Vaccinations-in-the-United-States-County/8xkx-amqh>

2020 Presidential General Election Data - Republican (Web Scraping):

*Ken (3/6/2022)*

<https://uselectionatlas.org/RESULTS/party.php?year=2020&type=national&no=2&f=1&off=0&elect=0>

2020 Presidential General Election Data - Democrat (Web Scraping):

*Ken (3/6/2022)*

<https://uselectionatlas.org/RESULTS/party.php?year=2020&type=national&no=1&f=1&off=0&elect=0>

Additional Control Data (Downloaded):

*Stefan (3/6/2022)*

*TBD: working on finding*

## CAPP30122 – Project Deliverable #2

### python\_parser\_tongues

- 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

- i) Retrieve Vaccination Data from API  
(1) Gabriela (3/1/2022)
- ii) Write code for web scraping the US Election Data by County/State  
(1) Ken (3/6/2022)
- iii) Collect Control Dataset for Regression Models  
(1) Stefan (3/6/2022)
- iv) Run Regressions and Create Simulation Models  
(1) Stefan (3/6/2022)
- v) Mapping Visualization  
(1) Fabian (3/6/2022)
- vi) Implement User Interface  
(1) All (3/10/2022)

- 4. Include any additional information you wish to provide about your project.**

N/A

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

For web-scraping, do you recommend we used Beautiful Soup or Selenium? We were thinking BS4 because it is static data based from a home page and then visits different links based on state and county.

Other than that, we request nothing but your mercy. 😊