SIADS 591/592

Milestone I project proposal

version 2021.07.24.1.AW

**Instructions:** please make a ***copy*** of this template file (do not edit original).

## **Proposal Title:** replace ***this*** line with your proposal title.

## 1. Team members

Please list your team members (2-3 max).

* Person name 1 (uniqname)
* Person name 2 (uniqname)
* Person name 3 (uniqname)

## 2. Project summary

Summarize your proposed project in a few sentences.

#### What is your proposed project and why are you proposing it?

#### What are the question(s) you want to answer, or goal you want to achieve?

| [Please use this space for your response. You may expand or contract this box as needed.] |
| --- |

## 3. Datasets

#### Describe one primary dataset and at least one secondary dataset. If other secondary datasets will be used please describe them as well.

#### The proposed datasets should exhibit different features/columns and/or different access methods, e.g., \*.csv file, \*.json file, API retrieval, web scraping, etc. Different time periods, for example, with the same features/columns is not considered a different dataset. Remember, the focus of the project in this Milestone course is to give you the opportunity to practice your data manipulation skills, so feel free to challenge yourself.

#### If you're unsure if your data sets are "different enough" describe the datasets and request a review via the *#siads591\_[semester]\_001\_project* Slack channel.

#### **Please note:** all proposed datasets ***MUST*** be publicly available to all members of the class (students, instructors, course support personnel, etc.). Use of proprietary datasets for this project is ***not*** permitted.

## 3.1 Primary dataset description

Describe your primary dataset. How is the data collected and how will you access it? Please share what features in the dataset are relevant to your topic. At a minimum, include the following information:

#### Short description (i.e., 1-3 sentences) of its key features

#### Estimated size (in records and/or bytes)

#### Location (give the URL or other access method)

#### Format (CSV, JSON, etc.)

#### Access method (download, web scraping, API, etc.)

| [Please use this space for your response. You may expand or contract this box as needed.] |
| --- |

## 3.2 Secondary dataset(s) description

Describe your secondary dataset(s). How is the data collected and how will you access it? Please share what features in the dataset(s) are relevant to your topic and describe the data types you’re expecting. At a minimum, for each secondary dataset include the following information:

#### Short description (i.e., 1-3 sentences) of its key features

#### Estimated size (in records and/or bytes)

#### Location (give the URL or other access method)

#### Format (CSV, JSON, etc.)

#### Access method (download, web scraping, API, etc.)

| [Please use this space for your response. You may expand or contract this box as needed.] |
| --- |

## 3.3 [ ] Affirm: datasets are public.

Please check the above box to confirm that your primary and secondary datasets are accessible and available to your classmates and the instructional team.

## 4. Cleaning and manipulation

How will you join your primary and secondary datasets? What cleaning and manipulation challenges, if any, do you anticipate?

| [Please use this space for your response. You may expand or contract this box as needed.] |
| --- |

## 5. Analysis

Describe any analyses you plan to undertake. For each, please give the technique or approach and briefly explain what you expect to learn from it.

| [Please use this space for your response. You may expand or contract this box as needed.] |
| --- |

## 6. Visualizations

Describe in 1-3 sentences at least **two** data visualizations that you plan to create. Include the chart type (e.g. bar chart, scatterplot, SPLOM, etc.) as well as the variables (features) you intend to plot.

| [Please use this space for your response. You may expand or contract this box as needed.] |
| --- |

## 

## 7. Ethical considerations

Does your choice of data raise any ethical issues? If so, briefly describe the concern and how you plan to mitigate it.

| [Please use this space for your response. You may expand or contract this box as needed.] |
| --- |

## 8. Contributions

Indicate the contribution that each team member will make to the project.

| [Please use this space for your response. You may expand or contract this box as needed.] |
| --- |

### 

## Changelog

(2021.07.24.1.AW) Adjust title, number sections, simplify section headings, edit text