**Team Project One**

Green = done

1. Mode of communication: slack, email, phone
2. Jenny
   1. Email: mengjsun@gmail.com
   2. Cell: 317-238-9425
   3. Github username: yuanyuans29
3. Angie
   1. Email: gomeza@ucla.edu
   2. Cell: 323-637-1771
   3. Github username: agomez0
4. Eddie
   1. Email: erp\_perez@hotmail.com
   2. Cell: 323-378-7058
   3. Github username: navydoc77

**Project Ideas**

1. Idea
   1. Marijuana acceptance by age group
   2. Student loan data

# Project 1 Groups

- Group 1: Laura, Pavel, Chris

- Group 2: Dylan, Anna, Adriana

- Group 3: Charlotte, Stan, Bryan

- Group 4: Tom, Keegan, Jamuna

- Group 5: Trinidad, Judy, Warren, Ramone

- Group 6: Jesus, Ryan, Siyu

- Group 7: Jenny, Eddie, Angi

**Project Description**

## Checklist

- Do you have each other's Slack info?

- Do you have each person's Github usernames?

- One person from each group should set up a Git repo for the project.

- Grant write access to the repository to the people in your group.

- How will you come to an agreement on the topic?

- Write a draft proposal early; iterate on it until you can all agree on it.

- Check your data sources.

- Set milestones and due-dates EARLY. No last-minute all-nighters please!

## Finding Data

Once your group has written an outline, it's time to start hunting for data. You are free to use data from any source, but we recommend the following curated sources of high-quality data:

\* [data.world](https://data.world/)

\* [Kaggle](https://www.kaggle.com/)

\* [Data.gov](https://www.data.gov)

\* [Public APIs](https://github.com/abhishekbanthia/Public-APIs)

\* [Awesome-APIs List](https://github.com/Kikobeats/awesome-api)

\* [Medium APIs List](https://medium.com/@benjamin\_libor/a-curated-collection-of-over-150-apis-to-build-great-products-fdcfa0f361bc)

# Presentation Requirements

The presentation requirements for Project 1 are as follows.

Your presentation must:

\* [ ] Be at least 8-10 min. long

\* [ ] Describe the core message or hypothesis for your project.

\* [ ] Describe the questions you and your group found interesting, and what motivated you to answer them

\* [ ] Summarize where and how you found the data you used to answer these questions

\* [ ] Describe the data exploration and cleanup process (accompanied by your Jupyter Notebook)

\* [ ] Describe the analysis process (accompanied by your Jupyter Notebook)

\* [ ] Summarize your conclusions. This should include a numerical summary (i.e., what data did your analysis yield), as well as visualizations of that summary (plots of the final analysis data)

\* [ ] Discuss the implications of your findings. This is where you get to have an open-ended discussion about what your findings "mean".

\* [ ] Tell a good story! Storytelling through data analysis is no different than in literature. Find your narrative and use your analysis and visualization skills to highlight conflict and resolution in your data.

# Technical Requirements

The technical requirements for Project 1 are as follows.

\* [ ] Use Pandas to clean and format your data set(s)

\* [ ] Create a Jupyter Notebook describing the \*\*data exploration and cleanup\*\* process

\* [ ] Create a Jupyter Notebook illustrating the \*\*final data analysis\*\*

\* [ ] Use Matplotlib to create a total of 6-8 visualizations of your data (ideally, at least 2 per "question" you ask of your data)

\* [ ] Save PNG images of your visualizations to distribute to the class and instructional team, and for inclusion in your presentation

\* [ ] Optionally, use at least one API, if you can find an API with data pertinent to your primary research questions

\* [ ] Create a write-up summarizing your major findings. This should include a heading for each "question" you asked of your data, and under each heading, a short description of what you found and any relevant plots.

More data-gathering sites:

Public datasets and repositories:

https://github.com/awesomedata/awesome-public-datasets

https://docs.google.com/spreadsheets/d/1wZhPLMCHKJvwOkP4juclhjFgqIY8fQFMemwKL2c64vk/edit#gid=0

https://www.reddit.com/r/datasets/

https://data.world/

https://www.kaggle.com/

https://www.data.gov

https://github.com/abhishekbanthia/Public-APIs

https://github.com/Kikobeats/awesome-api

https://medium.com/@benjamin\_libor/a-curated-collection-of-over-150-apis-to-build-great-products-fdcfa0f361bc