**Proposal for Project 3: Visualizing Data**

**Group Members:**

Daniel Garza, Norma Espinosa, Meina Bian, Adam Freedman, Angele Yazbec

**Proposed Datasets:**

World Happiness Report (obtained from Kaggle)

* <https://www.kaggle.com/datasets/mathurinache/world-happiness-report>
* Years Pre-COVID: 2018-2019
* Years During/Post-COVID: 2020-2021
* 165 countries are included
  + Obtain coordinates for the countries
  + Merge with the World Happiness Report data

**Extract-Transform-Load**

* Clean and transform dataset in Python using pandas
* Export data as a JSON
* Load dataset into MongoDB (pre-COVID and post-COVID databases)

**Visualization Tools**

* Using HTML and JavaScript (GeoMapping, Leaflet, Plotly)
  + Different markers to show happiness across different factors (e.g. healthy life expectancy, social support, freedom to make life choices), and different calendar years.
  + Bar/scatter/line charts show the trends over time
* Potential Interactions:
  + Different map layers for the years
  + Drop-down menus to view data from different years/top 10 countries/bottom 10 countries
* Will use new library/libraries (exact one(s) TBD)

**Hypothesis Testing** (time permitting)

* Conduct following t-test
  + Hypothesis: Countries reported higher happiness ratings before COVID than during/after COVID.

**Create presentation**

* Slide deck
  + Highlights data munging and ETL
  + Highlights key visualizations and findings
* Presentation
  + Telling the story! Letting the visualizations do the talking
  + If time permits: embed HTML link in presentation

**Task break-down:**

* Daniel
  + Obtaining coordinates for the countries
  + Merging with Kaggle datasets
  + HTML
  + Assist with visualizations as needed
* Adam
  + Extracting datasets from Kaggle
  + Loading and cleaning in Pandas
  + Creating JSON
  + Assist with visualizations as needed
* Norma
  + HTML and GeoMapping
  + PowerPoint
* Meina
  + HTML and visualizations
* Angele
  + Loading into MongoDB
  + Hypothesis test
  + Assist with visualizations as needed