Project 3
Group 6
Project Proposal

Caleb Meinke

Adam Loux

Revano Harahap

Monica Mitry

Dataset Information:
- With a couple team members being baseball fans we chose to look into data about home runs and stadium locations.
- The chosen data set:
  https://www.kaggle.com/datasets/supremeleaf/mlb-hitting-leaders-2019-2022
- Since this data set was missing latitude and longitude we will also utilize this data set to help fill in some gaps.
  https://www.kaggle.com/datasets/logandonaldson/sports-stadium-locations?resource=download

Inspiration:
- While this dataset doesn't have any code attached with it we still find some inspiration of statistics we were all curious about.
  - What is the correlation between home runs with stadium locations?
- A secondary data set that we had considered is:
  https://www.kaggle.com/datasets/vivovinco/2023-mlb-player-stats
  - Due to this data being for one year we chose to use the other data set in order to have greater historical data to analyze.

Example Visualizations:
- Map of stadium location in correlation to home runs.
- Bar graph of baseball plays batting statistics
- Heat map

Color Palette:
- MLB branding colors
- https://coolors.co/f03a47-af5b5b-f6f4f3-276fbf-183059
- https://teamcolorcodes.com/major-league-baseball-logo-colors/

Roles and Responsibilities:
- Caleb:
  - Data Cleaning
  - Database Design
- Revano:
  - Flask API
- Adam:
  - Map
- Monica:
  - Plotly
  - Lead on Slides
- Each will take lead on certain areas, however, each team member will be working together to assist with all aspects of the project.