# Project 2 Proposal: UFOs

## Project Team

* Blake Bormes
* Steve Hewitt
* Tymon Silva

## Github Repository

UC-Berkeley-I-School/Project2\_Bormes\_Hewitt\_Silva

## Primary Dataset

[National UFO Reporting Center Web Reports (nuforc.org)](http://www.nuforc.org/webreports.html) (UFO.csv)

## Initial Plots, Figures, and Tables

* Time series showing number of sightings over time
  + Line Chart by year
* Bar chart of sightings by hour of day
* Bar chart of sightings by day of the week
* Heat map of UFO by state
* Bar chart with months/year and number of sightings
  + Distribution by month
* Bar chart showing distribution of the different shapes of UFO sightings
* Table describing top 10 sightings
* Table describing top 10 cities
* Histogram or Grouped bar chart for duration of UFO sighting
* Histogram of difference between Date\_Time and Posted\_Date

## Variables

* UFO Data
  + City, State, Shape, Duration, Date\_Time, Date\_Posted
* Census
  + Population by State (?), City (?), Year

## Potential Insights

* How did the pandemic impact number of sightings?
* Are there different shapes in different parts of the country?
* Do sightings go up after a release of an alien movie at the box office?
* Does weather/climate impact location and/or number of sightings? (possible)
* Where and when was/is the best place to see a UFO?
* What is more likely: sighting an UFO or being struck by lightning?

## Supplemental Datasets

* [Historical Weather Data](https://www.ncdc.noaa.gov/data-access/land-based-station-data/land-based-datasets/global-historical-climatology-network-ghcn) (possible)
* [GIS](https://github.com/awesomedata/awesome-public-datasets#gis) (possible)
* Movie Data (possible) – Top 10 grossing movies of all time with aliens

## Final Report

Executive Summary

* Key findings and insights

Introduction

* Goals, Objectives, Expected Impact, Background
* Overview of key variables and list of questions to answer
* Data Set used (include Link)

Methodology

* Analysis method / Design Process
* Data prep and cleaning
  + remove records where Date\_Time > Posted\_Date (data viz)
  + Remove non US
* Assumptions (Check correct position)

Analysis/ Research Findings

* Univariate Analysis
* Crosstabs

Conclusion/Recommendations

* Key findings insights
* Big takeaways
* Recommendations (What is more likely: sighting an UFO or being struck by lightning?)