UFOU

Easy and convenient access to UFO sighting data

Christopher de Freitas

Anastasia Gonzalez

Conner Soule

Introduction

There are 70 sextillion stars in the universe[1], 400 quintillion of which are considered "sun-like" stars.[2] If just 1% of the "sun-like" stars contained an Earth-like exoplanet orbiting it, then 100 * 10^{18} plants in the universe potentially sustain life, leaving the Milky Way alone 8.8 billion potentially habitable planets.[3] Italian physicist Enrico Fermi was the first to formally publish the obvious question: "Where is everybody"?[4] It seems like the universe should be teeming with life, but there's very little empirical evidence pointing to the existence of life outside Earth.

Our mathematical certainty that life *must* exist beyond our planet is chilling in the face of our lived experience as the sole conscious species in the universe. As such, we should aim to hold onto any evidence pointing to extraterrestrial life, whenever possible. Convenient access to UFO sighting data is essential to better understand our universe and prepare ourselves for life outside of our galaxy. Our project scrapes UFO sighting data from the National UFO Reporting Center (NUFORC) and compiles it into one database. Next, analysis of the data in the database is used to find commonalities in patterns and sightings, decomposing thousands of reports into easily digestible statistics. Finally, visualization of the data is displayed to the user on a web page in simple graphs and charts providing users the information they need to be on the lookout for the next UFO event.

---

[1] Lazo M. Manojlović, "Photometry-based estimation of the total number of stars in the Universe," Appl. Opt. 54, 6589-6591 (2015)

[2] Ethan Siegel, Forbes, "Astronomers Debate: How Many Habitable Planets Does Each Sun-Like Star Have?", Tue Nov 12 2019, https://www.forbes.com/sites/startswithabang/2019/10/01/astronomers-debate-how-many-habitable-planets-does-each-sun-like-star-have/#7fa90a3e253a

[3] Prevalence of Earth-size planets Erik A. Petigura, Andrew W. Howard, Geoffrey W. Marcy Proceedings of the National Academy of Sciences Nov 2013, 201319909; DOI: 10.1073/pnas.1319909110

[4] Seth Shostak, SETI, "Fermi Paradox | SETI Institute", April 19 2019, https://www.seti.org/seti-institute/project/fermi-paradox

Project Description

The two largest UFO reporting centers are the Mutual UFO Network (MUFON) and the National UFO Reporting Center (NUFORC). Together, they've compiled thousands of unique UFO sightings from around the world dating all the way back to the 15th century. This data is valuable, but inaccessible. NUFORC organizes their data into neverending HTML tables in small, illegible fonts. MUFON locks most of its data behind a paywall, providing detailed information only for a small handful of cases each year. This site is intended to dramatically improve both the accessibility and intelligibility of these reports.

Although it would be invaluable to collect and cross-reference data from both MUFON and NUFORC, the work required to scrape data from MUFON would be a bit beyond the time constraints (and budget) for this project. NUFORC, on the other hand, provides thousands of easy to access, tabularized reports. A small **web scraping program** will be running alongside the web server, checking at least once per day for updates to NUFORC's web reporting site (http://www.nuforc.org/webreports.html). The program will scrape any updates it finds on the site and update the database used by the site proper.

Each NUFORC report consists of a date occurred, date reported, date posted (the date it was actually posted on the NUFORC site), location, shape of the UFO, duration, and description. A **Microsoft SQL Server database** will serve as our central, stable repository for these reports. Our backend model will maintain the same schema as NUFORC: a single Report table as the heart of the database, with each report containing all of the above information. Indexes will be added to various columns and column combinations on the table, lengthening the time it takes for the scraper to insert or update data, but limiting query time for users.

There may need to be some regularization done on the back end on data entry, particularly for the duration field. When filing a report, “Duration” is just a text box to be filled, so there is little to no consistency in the formatting or units between reports. Queries will be made in LINQ rather than EF syntax to support the super-scaling of this database in the future.

The **computational analysis** will include determining ways to display the information stored in the database in a graphical format. Additionally, means, medians, and modes will be computed to find correlations between UFO shape size and UFO sighting location. The evaluated data will be stored in the database so that recomputation does not happen on every refresh. It will also only compute when new or updated information is stored into the database to avoid superfluous calculations. Lastly, most of the computations will be closely related to the needs of the visualization on the front-end.

The **visualization element** of the project will include simple graphs and charts describing data from the reports, and will support access to individual reports. For example, selecting a city, such as Roswell, will pull up and represent all UFOs sighted in Roswell visually, along with their shape. Clicking on a UFO will open the individual report filed for that sighting.

Time permitting, the following sections will also be implemented:
A **predictive analytics** component would be added, using regression analysis to predict when and where the next UFO sighting may be.

A **game** similar to Galaga or Space Invaders would be added in which the user could fire at and destroy various UFOs. The enemies would be based off the shapes and locations logged in the Reports table.

Team Member Contribution

Christopher will lead the web scraping off the National UFO Reporting Center (NUFORC) website where he will take the information and populate a database. If the time permits, Chris will apply a simple regression to predict the time and location of the next UFO sighting.

Anastasia will lead the backend computational statistical analysis of the data in the database. This analysis will be the glue between what is displayed on the front end of the website and what is in the database. When available, Anastasia will help Conner create graphs and models to visualize the data on the front end.

Conner will lead the front end visualization on the web page where the majority of the information will be from the backed computational analyses. At times, the data will come directly from the database. Additionally, Conner will make a simple game visualization of the data (such as space invaders or interactive map) if the time permits.

Team Member Contribution

Alpha phase

The research phase will consist of building mock-ups, understanding the flow of the final project layout, and deciding the features our project will support in the end.

Team Member	Tasks & Goals
Anastasia	• Research how to use the javascript library ploty for displaying of maps, charts, and graphs.
Chris	• Research of Web Scraping National UFO Reporting Center (NUFORC) • Organizing data from web scrape to be in usable format
Connor	• Research how to use javascript to support a game for users to play inside the browser.

Beta phase

The beta phase will consist of converting mock-ups, building basic pages, and begin supporting UFO data. Some basic hard coded game simulation will also be completed.

Team Member	Tasks & Goals
Anastasia	• Building the home screen and other important screens for easier navigation through the web page
Chris	• Organizing data from web scrape to be in usable format, specifically supporting UFO shapes, report dates, locations, etc.
Connor	• Able to play the basic game by importing the proper libraries and loading it in the browser with no major problems.

Production phase

The production phase will consist of finishing the front-end, adding maps and graphs to the UI, adding user support. Most of the phase will consist of the coding and finishing details.

Team Member	Tasks & Goals
Anastasia	• Building the home screen and other important screens for easier navigation through the web page • Build Identity support for users to save reports to their account
Chris	• Scraper adapted from a one-time use data gathering to a fully integrated tool • Built up customizability of the tool and allow admins some options when starting or stopping a crawl
Connor	• Build US map that displays the distribution of reports from the National UFO Reporting Center (NUFORC) • Link map to display detailed results of the reports based on state

Team Member Contribution

UI Design

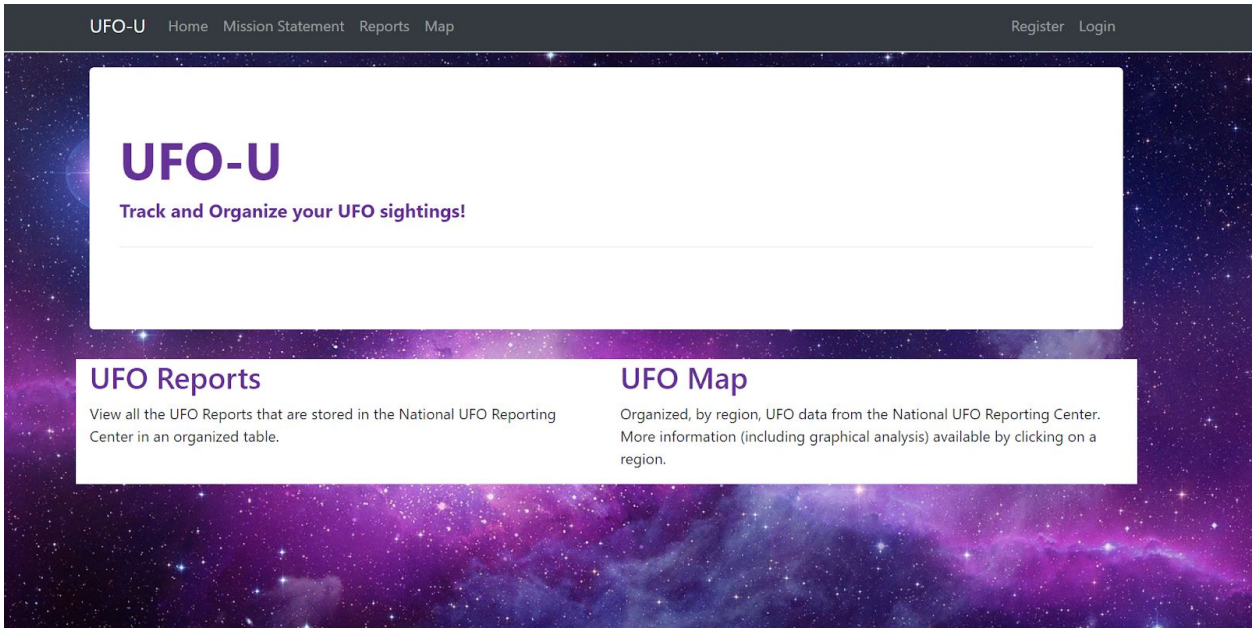


Figure 1: Home Screen

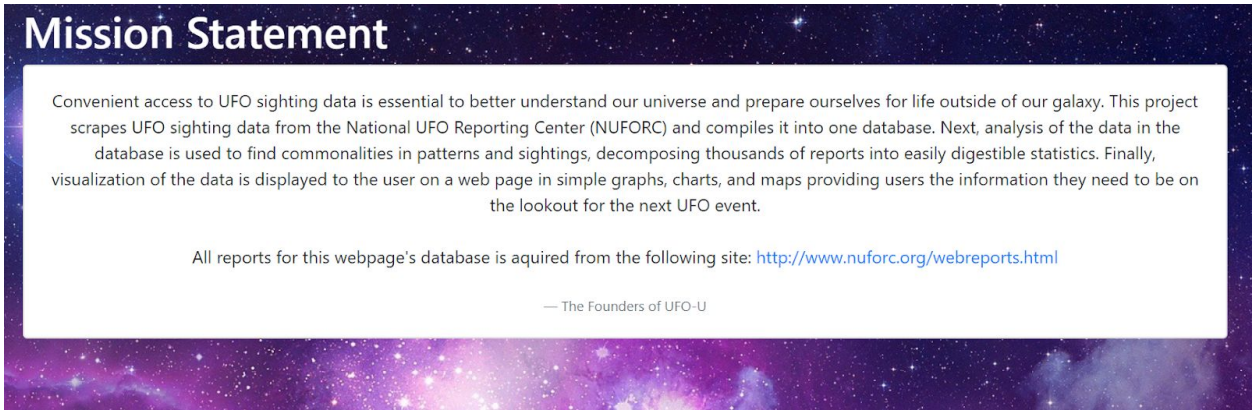


Figure 2: Mission Statement

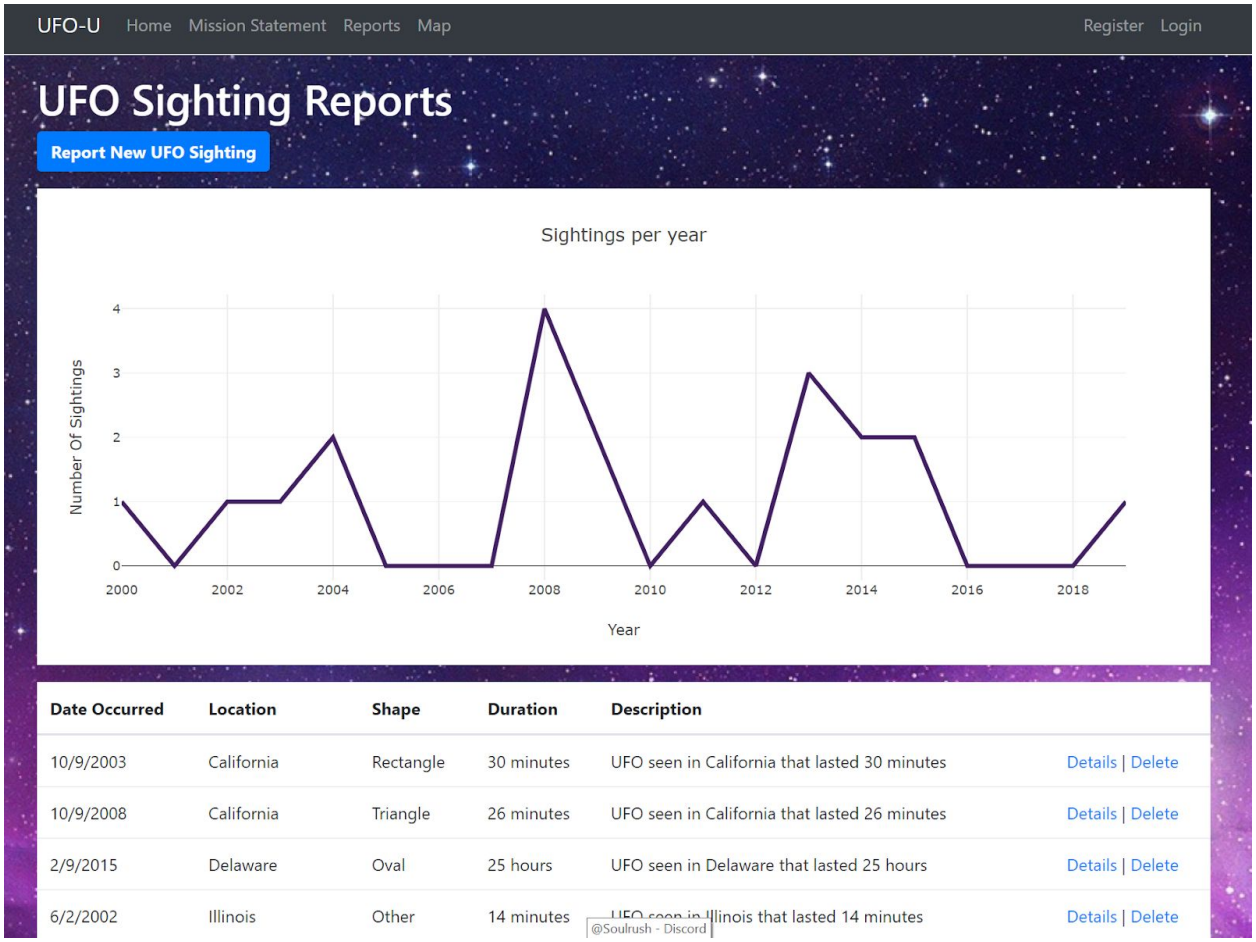
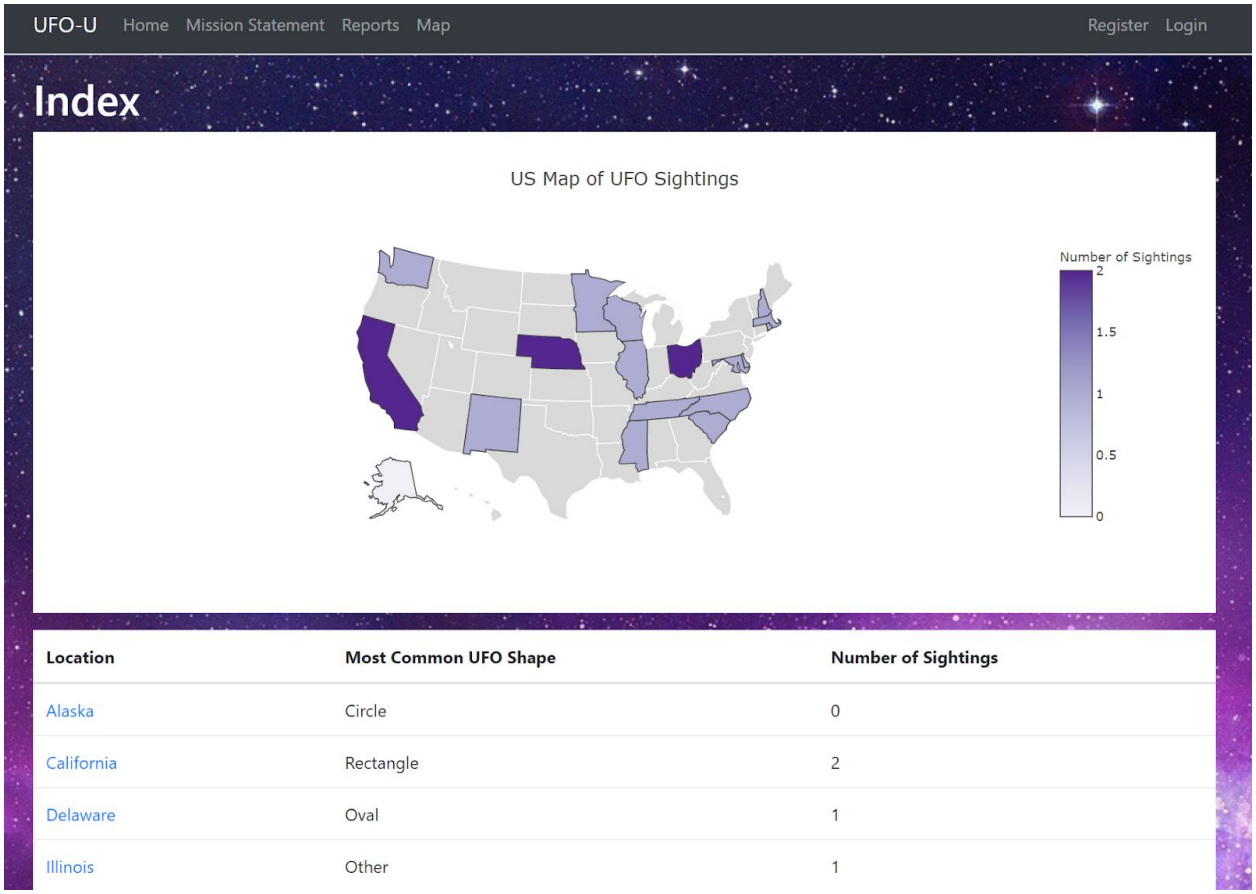


Figure 3: Reports with a Line Graph Showing Sightings Over the Years



Location	Most Common UFO Shape	Number of Sightings
Alaska	Circle	0
California	Rectangle	2
Delaware	Oval	1
Illinois	Other	1

Figure 4: Map showing sightings around the US

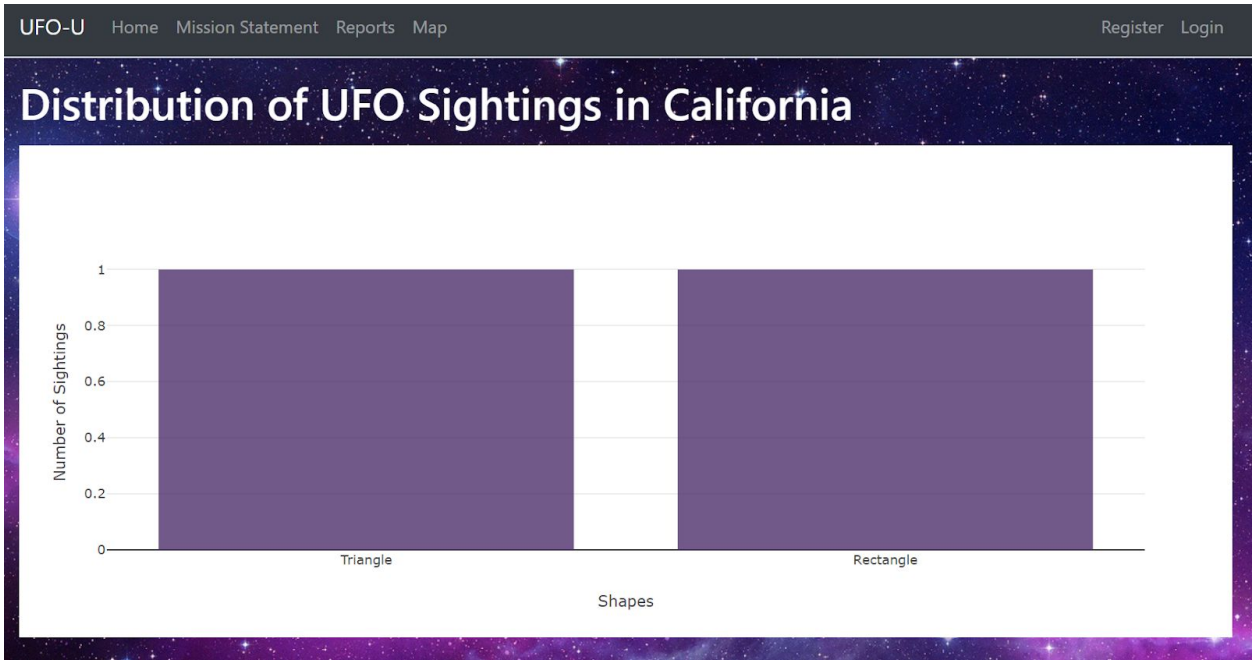


Figure 5: Graph showing distribution of shape sightings in a state

URL

ec2-3-84-192-66.compute-1.amazonaws.com