

TEAM TENACIOUS DATA

PROJECT PROPOSAL



TOPIC: UFO SIGHTINGS

Inspiration: We selected the topic of UFO sightings because of the perspective-broadening nature of the topic and the extensive data available. Sightings of unidentifiable objects have been reported as early as the 1940's. The first well known UFO sighting occurred in 1947 near Mount Rainier in Washington (Shostak, S. (2023, April 21). Unidentified flying object. Britannica <https://www.britannica.com/topic/unidentified-flying-object>). The topic raises great questions about why so many people have reported similar sightings around the world.

LINKS TO SIMILAR EDAS

- <https://www.kaggle.com/code/sagarsy2050/ufo-sightings/notebook>
- <https://towardsdatascience.com/are-we-alone-in-the-universe-data-analysis-and-data-visualization-of-ufo-sightings-with-r-42d0798679c3>
- <https://www.cs.ubc.ca/~tmm/courses/547-17F/projects/hayley-theodore/report.pdf>
- <https://www.kaggle.com/code/moosecat/world-ufo-sightings>
- <https://www.kaggle.com/code/brsdincer/meteorite-landings-analysis-all-eda-theory>

DATASET

Dataset: [UFO Sightings around the world | Kaggle](#)

This is a 13.71MB CSV dataset. It is comprised of eleven columns and over 80,000 rows. The data contained therein includes dates and times of the sightings spanning over 100 years, as well as the date the sighting was documented, city/state/country locations as well as latitude and longitude information of sightings, the shape of the unidentified object seen, the length of the encounter, and a text description.

RESEARCH QUESTIONS:

- 1) Are there trends in the locations of sightings?
- 2) If so, do those trends extend to UFO shapes sighted, or length of the encounter?
- 3) Is there any significant relationship between sightings and particular times or dates?
- 4) Are there any changes in sighting frequency or locations over time?

LINEAR/LOGISTIC REGRESSION:

Linear/Logistic regression

With our EDA, we hope to predict geographic areas where future UFO sightings will be likely. Can we predict an increase in sightings based on weather events? and will similar shapes be reported at the time of the sighting?

VISUALIZATION TYPES:

The types of visualizations we plan to use for our analysis include one or more line graphs (to examine any changes in sightings over time), bar graphs (to investigate trends in sighting location data), violin plots (perhaps to evaluate the duration of encounters), and hopefully a geomap or two using an API.

UNIQUE AND CREATIVE COLOR PALETTE:

#09225d

#21c0ab

#808080

ROLES/RESPONSIBILITIES:

Roles	Responsibilities
Data Cleaning	Group
Research question 1	Ann
Research question 2	Alex
Research question 3	Carlos
Research question 4	Abi
Linear Regression	Group

