

UFO SIGHTINGS

Team **Uniquely Fun Observations**

Rachel Ng, Joyce Liao, Meredith Hu



HYPOTHESIS

There are more reported
UFO sightings around
major holidays.



BACKGROUND

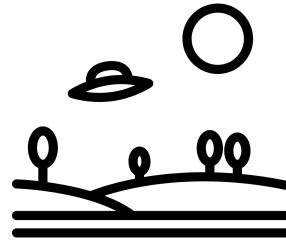
We thought it would be interesting to see if there was any pattern between sightings of UFOS. Our first thoughts were to check if there were any specific times of the year where UFOs are spotted at a higher frequency than the rest of the year.

World UFO Day (June 24th): marks when the first UFO identified in the U.S. by aviator Kenneth Arnold



RELATED ARTICLES

- The paper [**Social Psychological Correlates of Possible UFO Sightings**](#) by Troy A. Zimmer observes that there are **no significant correlations** between UFO sightings and cultural alienation, malevolent world view, and personal well being. The **strongest factor** shown in whether a person would report a UFO sighting is if they knew another person who reported a UFO sighting.
- The study [**Measured Personality Characteristics of Persons Who Claim UFO Experiences**](#) by Dr. June Parnell, concluded that people who report UFO sightings have a statistically significant greater chance (0.003) of endorsing unusual feelings, thoughts, and altitudes, to be suspicious or distrustful (.01), and to be creative, imaginative, or possibly schizophrenic (.034).
- The book [**The Gods Have Landed: New Religions from Other Worlds**](#) talks about people who report UFO sightings as having a tendency to be religious.



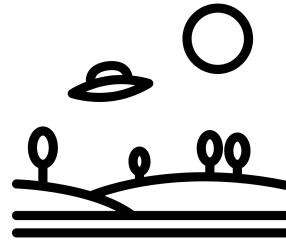
DATA PROFILE

[“UFO Sightings Around the World”](#) (Kaggle)

The original dataset had 80,333 data points.

After cleaning, we ended up with 71,773 data points, with sightings ranging from 1910 to 2014 from the following countries:

Country	Sightings
Canada	3791
Great Britain	2498
United States	64844
Total	71133

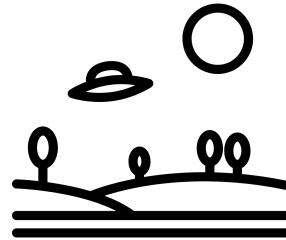


DATA PROFILE

CLEANING THE DATA

We created a SQLite database to clean our data

- Removed data entries with empty “country” fields.
- Fixed entries with incorrectly entered data fields
- Added leading zeros to all dates for easier sorting
- Removed countries without a significant number of data points and irrelevant columns from the original dataset



DATA PROFILE

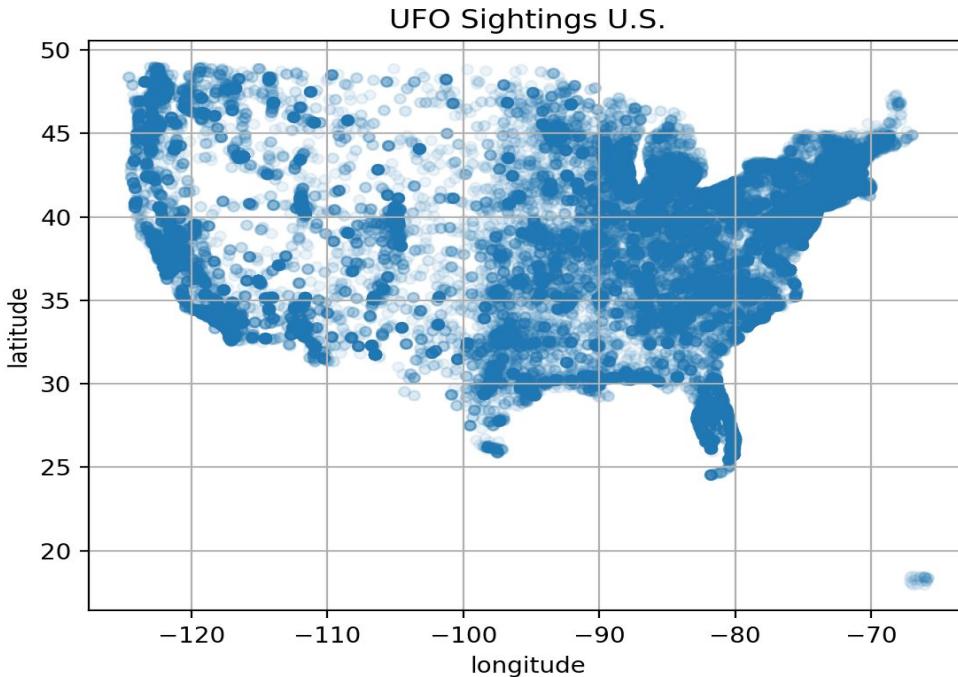
BIASES IN THE DATA

While processing the data, we recognized that the data was biased towards the US, Canada, and the UK.

To account for this, we removed countries that had a disproportionately small amount of data compared to those countries, to less heavily reflect the results of the bias.



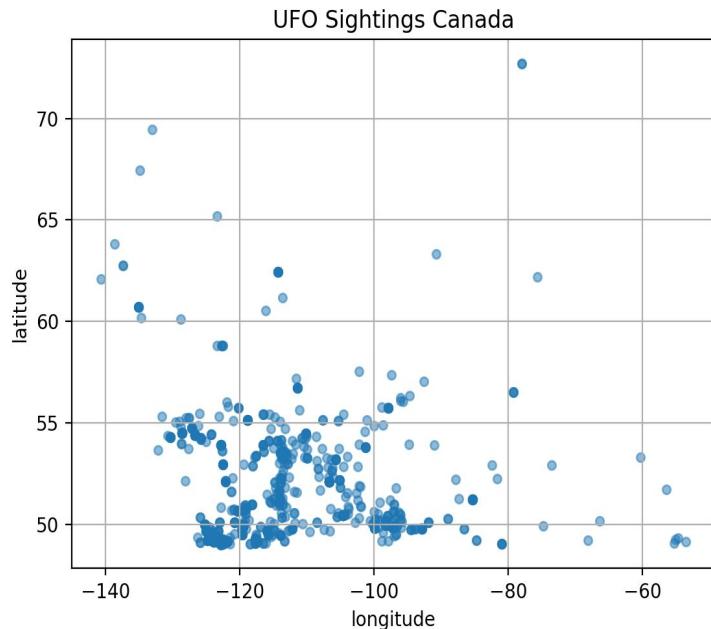
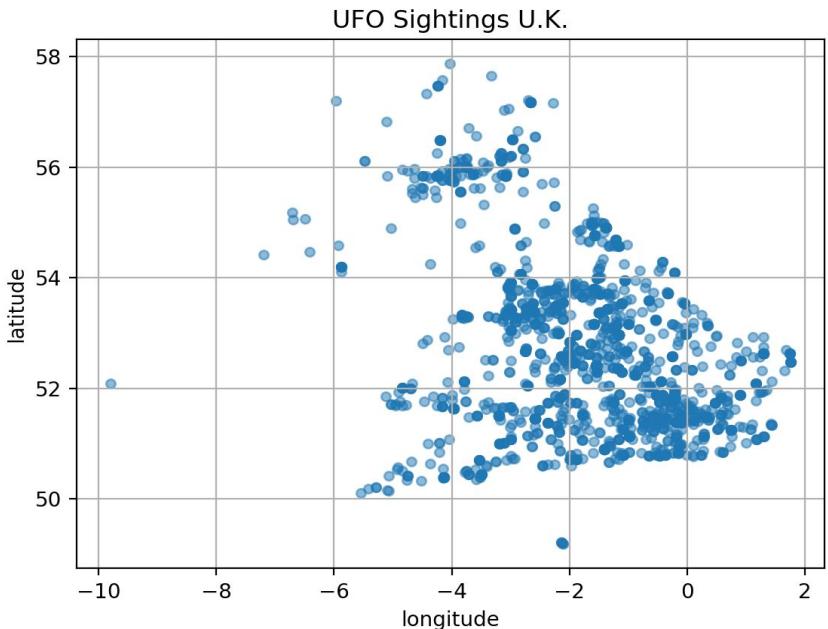
LOCATIONS OF SIGHTINGS (U.S.)



Mainly concentrated
on the coasts.

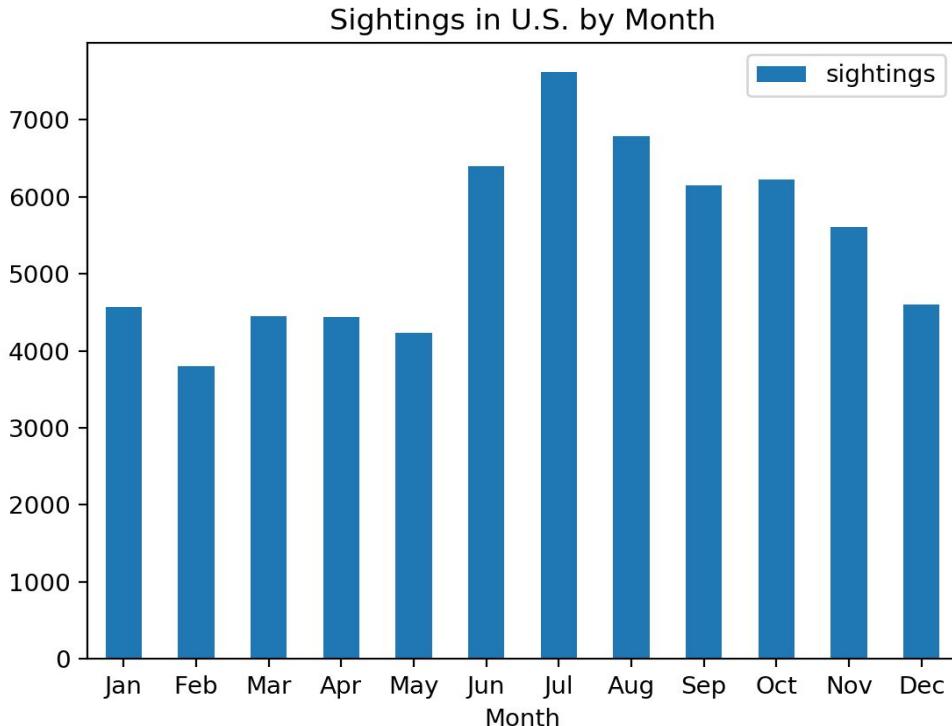


LOCATIONS OF SIGHTINGS (U.K. / CANADA)





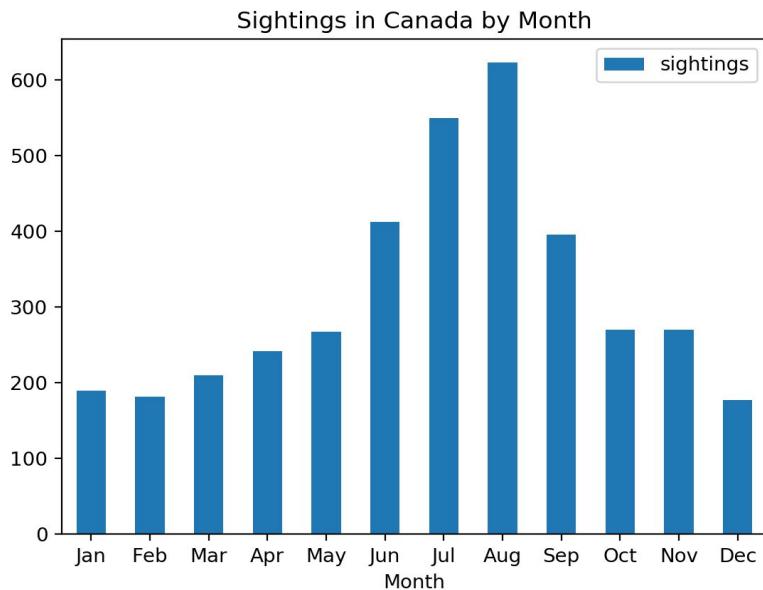
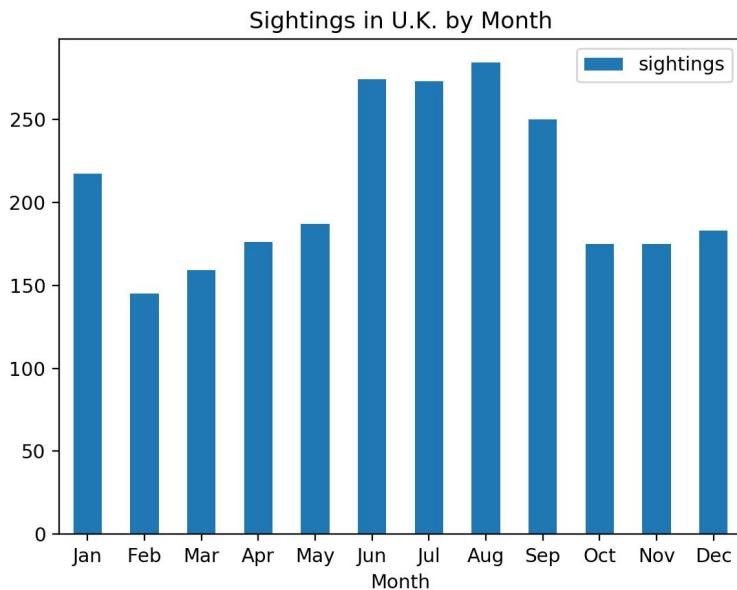
DISTRIBUTION OF SIGHTINGS BY MONTH (U.S.)



Number of sightings peaks during the month of July, with generally more sightings in the warmer months of the year.



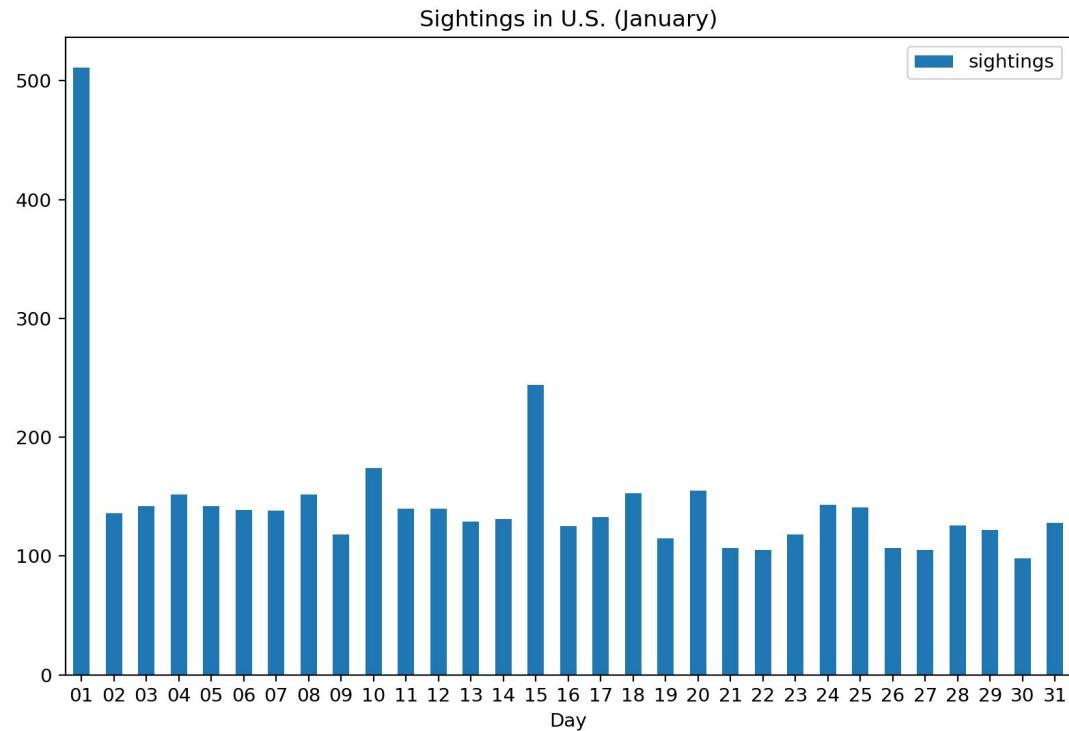
DISTRIBUTION OF SIGHTINGS BY MONTH (U.K. / CANADA)





JANUARY UFO SIGHTINGS (U.S.)

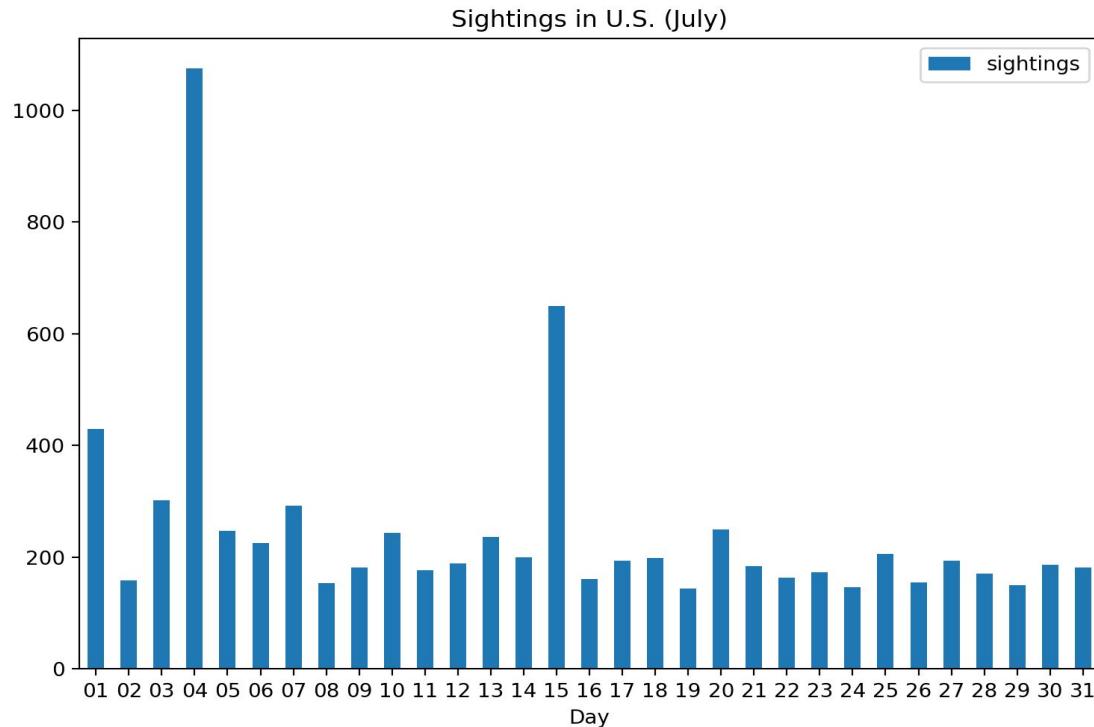
There is an unusually high number of sightings on **January 1st** (New Years).

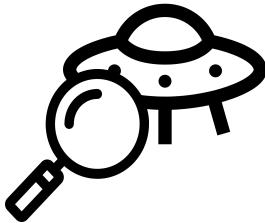




JULY UFO SIGHTINGS (U.S.)

There is an unusually high number of sightings on **July 4th** (Independence Day) and **July 15th**.

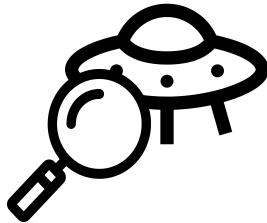




CLOSE INVESTIGATION: TWO SAMPLE T-TEST

To test whether the data was statistically significant, we ran a two-sample t-test for independent variables (holiday vs not holiday). We chose to use a t-test because the true standard deviation of the dataset (mean number of UFO observations recorded on one day of the year) is unknown.

It is a two sample, one tailed test because we are comparing two distributions and we want to find out if one distribution's mean is **higher** than the other's.

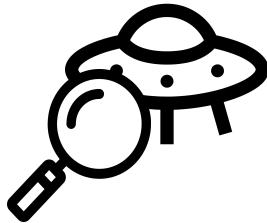


CLOSE INVESTIGATION:

TWO SAMPLE T-TEST

PROCESS

- We split the data into two distributions: one that included only UFO observations recorded on Independence Day and New Year's Day, and one that included only UFO observations recorded on every other day.
- All the data points were from the U.S., grouped by year and the number of data points was summed.
- The numbers were then divided by the number of days in each list .
- We then found the mean of each list.

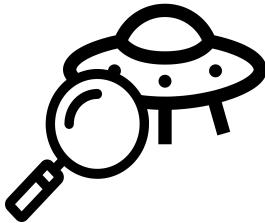


CLOSE INVESTIGATION: TWO SAMPLE T-TEST HYPOTHESES

The mean of the holiday list (u_1) was **7.552380952380952**
The mean of the non holiday list was so small, it divided out to **0**.

Null hypothesis: $u_1 = u_2$
Alternate hypothesis: $u_1 > u_2$

We chose alpha to be **0.01**.



CLOSE INVESTIGATION: TWO SAMPLE T-TEST

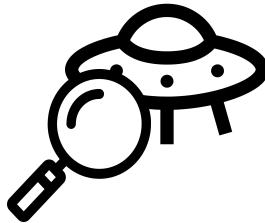
CALCULATING T

$$t\text{-statistic} = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}}$$

We then calculated the t-statistic.

$$t = 3.6491228650640073$$

$$df = \frac{\left(\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}\right)^2}{\frac{\left(\frac{s_x^2}{n_x}\right)^2}{n_x-1} + \frac{\left(\frac{s_y^2}{n_y}\right)^2}{n_y-1}}$$



CLOSE INVESTIGATION:

TWO SAMPLE T-TEST

P & CONCLUSION

We then calculated p (the probability than our t-statistic could be chosen by chance.

$$p = 0.00016637733467228522$$

Since p is **smaller than alpha**, we reject the null hypothesis. The result is **statistically significant**, and the mean of the holiday set is higher than the mean of the non-holiday set.

On average, people tend to report UFO sightings more frequently on Independence Day and New Year's Day.



CONCLUSIONS

- The results **support our hypothesis** because it was statistically proven that significantly more UFO sightings were reported on holidays compared to other days in the same month in the U.S.
- An **alternative interpretation** of the results is that people participate in outdoor activities more frequently during the warming months, making them more likely to observe the sky
- **Limitations**
 - Only two holidays were analyzed
 - Most of the data are from locations within the U.S.
 - Data from a wide range of years were looked at as a whole



FUTURE WORK

- Examine UFO sighting trends across different years
- Analyze UFO sighting data from more countries
- Focus on more holidays



**THANK YOU
FOR LISTENING**