

Investigating UFO sightings Before and After Covid

By John Cook

Introduction

UFO sightings really took off in the United States during the start of the cold war, a time period of stress and tension. Roswell happened in 1947 and Project Blue Book began in 1948. During the beginning of the Covid-19 pandemic, the Navy released videos of a UFO sighting they had. This begs the question, do we see more UFO sightings in the United States during stressful times? Covid could be a great modern test for this.

Introduction Continued

To answer this , I examined data of UFO sightings in the United States from the National UFO reporting Center from September 15th 2018 to September 15th 2021 in order to get data exactly 18 months before covid lockdowns started to 18 months after covid lockdowns started. Factors examined are numbers of sightings, most reported shapes, states with the most sightings, and report summary sentiment scores. I predicted that we would see more sightings after Covid lockdowns began because people would probably be under more stress than before Covid lockdowns.

Data Summary

15,814 sightings report records were gathered from September 15th 2018 to September 15th 2021 from a larger data set web scraped from the National UFO Reporting Center using the python libraries requests and BeautifulSoup. 5 columns were utilized from the data set:

- Date: Gives the date of when the UFO sighting happened
- State: State of where incident occurred
- Country: Country where incident occurred
- Shape: Reported shape of the UFO object seen
- Summary: A short text summary of the full report

Data Gathering and Cleaning

- In the full data set there are many badly formatted dates before the target date I wanted, before cleaning I needed to manually locate when 2018 starts and make a smaller data set only with 2018 and up dates. Then I turned all those dates into a datetime format from a string data type, and made a data frame dating from 09-15-2018 to 09-15-2021.
- The Date column also comes with a time stamp. I separated this by splitting on the space between the date and the time, assigned them to their own columns and dropped time as a column since it will not be used for this analysis.
- Records were filtered to show only sightings that happened in the United States.
- Columns not being used in the analysis were also dropped.

Data Gathering and Cleaning continued

- The text data in summary was pre-processed for sentiment analysis by taking out punctuation, apostrophes, and new lines. Stop words were eliminated and lemmatization was used to give better results. Lemmatization groups together inflected forms of a word so they can be analyzed as single items. For example: ate and eating would both be changed to eat.

Before Data Cleaning

Unnamed: 0		date	city	state	country	shape	duration	summary
0	0	3/3/22 08:01	Phoenix	AZ	USA	Egg	NaN	No sound weird looking though it was helicopter
1	1	3/2/22 21:40	Lino Lakes	MN	USA	Sphere	3 minutes	Three bright spheres moving across sky.
2	2	3/2/22 18:19	Indianapolis	IN	USA	Cone	7 mins	Heard a high pitched ringing and turned around...
3	3	3/2/22 18:18	lexington	NE	USA	Cylinder	2 minutes	While waiting for a locksmith I observed an ob...
4	4	3/2/22 18:10	paw paw	MI	USA	Cone	10 seconds	Driving to work and saw bright light slow movi...
...
137422	137422	NaN	Graham	WA	USA	Disk	2 minutes	Driving west on 224th looked right saw large c...
137423	137423	NaN	NaN	NaN	USA	NaN	NaN	((HOAX???)) I work as a security guard in marc...
137424	137424	NaN	Edmond	OK	USA	NaN	NaN	((HOAX???)) 30 second bright red glowing fireb...
137425	137425	NaN	Richland	WA	USA	Sphere	10 min	Multiple sightings
137426	137426	--	Monument	CO	USA	Changing	8-12 MINUTES	SPIDER/OCTOPUS SHAPED CRAFT APPRX 200-300 FT A...

137427 rows × 8 columns

After Data Cleaning

	date	state	country	shape	summary
1570	2021-09-15	IN	USA	Circle	rounnd white light moving sw ne stopped hover...
1571	2021-09-15	FL	USA	Light	evening dark partly cloudy listening live band...
1572	2021-09-15	CO	USA	Light	atmosphere entering tracer fire light
1573	2021-09-15	FL	USA	NaN	nuforc note witness provides detailed informa...
1574	2021-09-15	FL	USA	Changing	nuforc note witness provides detailed informa...
...
18461	2018-09-15	NJ	USA	Oval	saw bright oval light sky straight row flew a...
18462	2018-09-15	OH	USA	Circle	group bubblelike ufo seen ne ohio traveling nw
18463	2018-09-15	NY	USA	Triangle	saw triangular object green red line appeared ...
18464	2018-09-15	NJ	USA	Light	brilliant bright light brick nj
18465	2018-09-15	VA	USA	Light	multi color light light night sky temporary di...

15814 rows × 5 columns

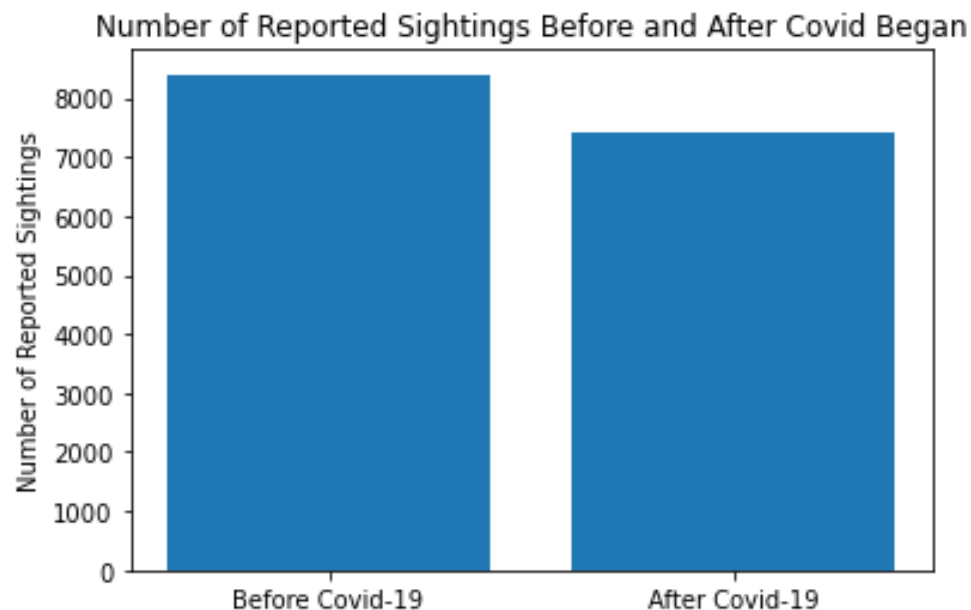
Comparing Number of Sighting Instances Before and After Covid-19

We see that the before Covid data set had about 1,000 more occurrences than the after Covid data set.

- People reported seeing less UFO sightings the period after covid lockdowns began.

Before Covid Count: 8398

After Covid Count: 7393



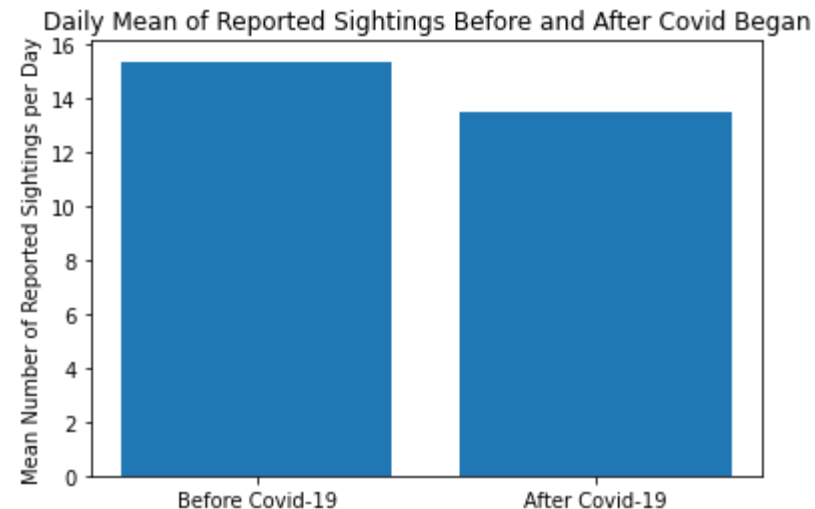
Comparing Number of Sighting Instances Before and After Covid-19: Daily

To really compare the number of report incidents, I decided to look at incidents per day because it gives a more quantitative view and can be better tested for significance than using one number count.

- The mean of daily covid sightings before covid is just about 2 sightings more than after covid.

Before Covid Daily Mean:
15.352833638025594

After Covid Daily Mean:
13.49087591240876



Comparing Number of Sighting Instances Before and After Covid-19: Testing for Significance

```
Statistics=173661.500, p=0.000  
Different distribution (reject H0)
```

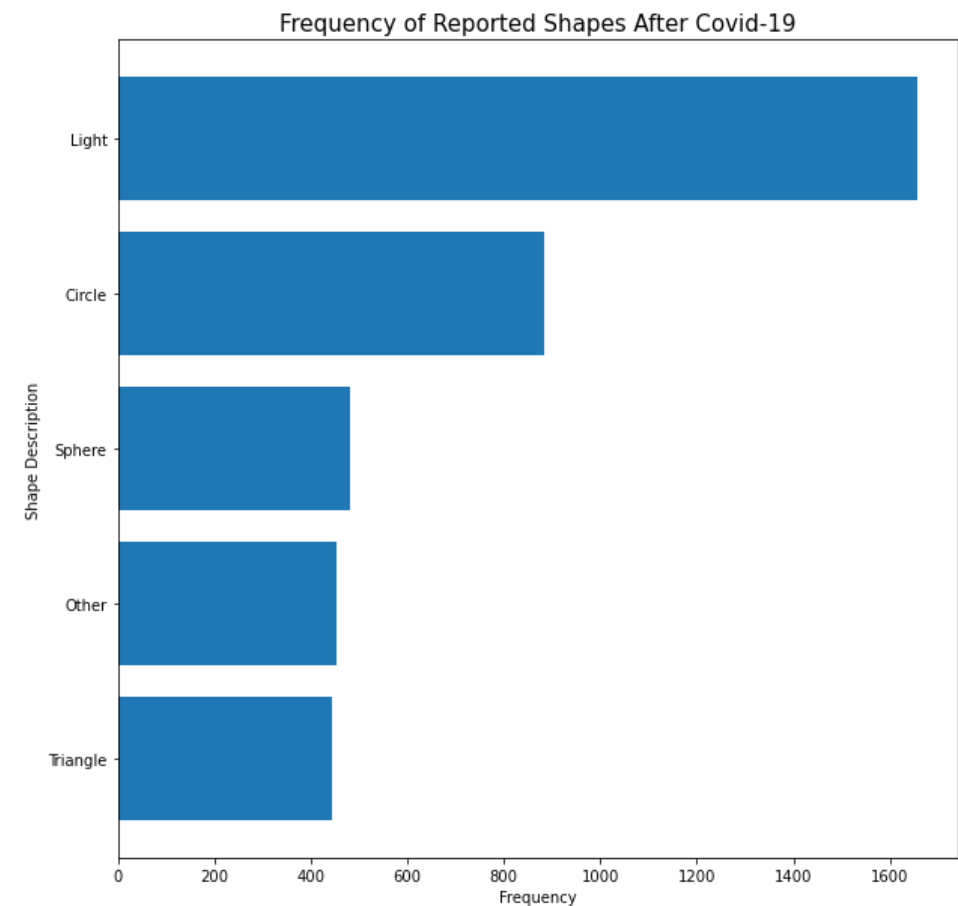
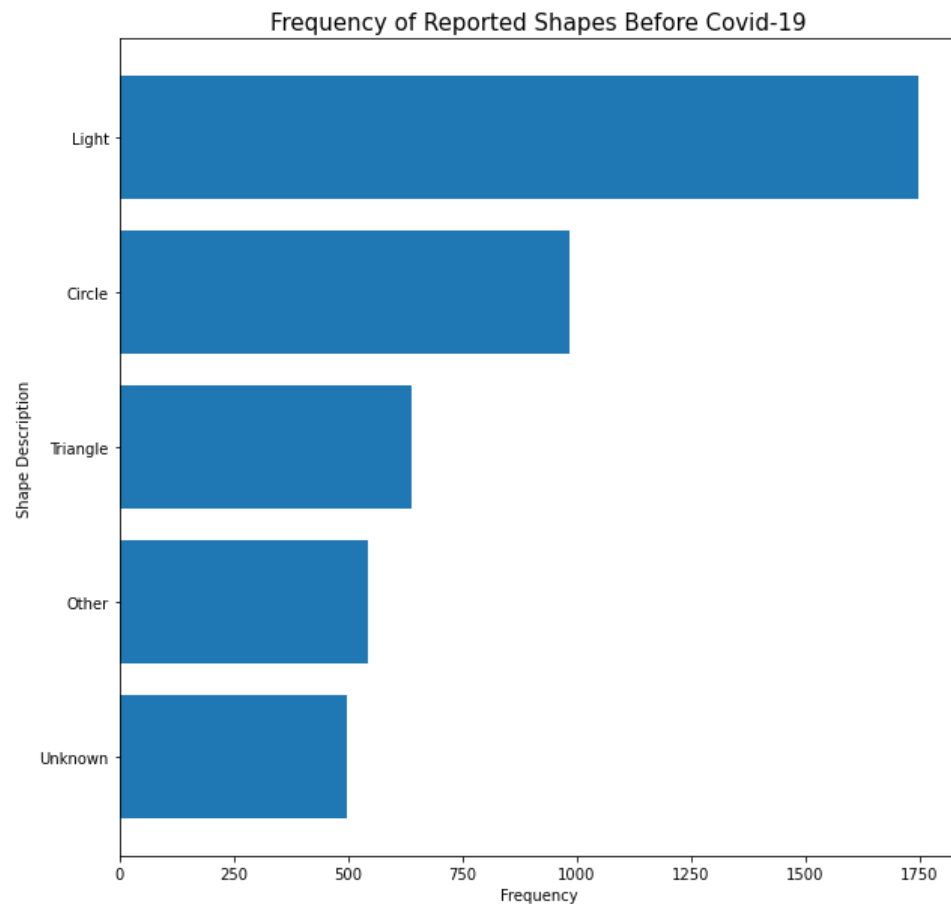
```
5.34918254858708e-06
```

To test for significance, I used the Mann-Whitney U test which is a non-parametric test that compares the medians of 2 independent samples. I chose to use a non-parametric test because it assumes very little about the population distribution.

We reject the null hypothesis which means in this sample, the number of daily instances was significantly more before covid lockdowns began than after the lockdowns began.

Comparing Reported Shapes

The most common UFO shapes that people reported seeing during their encounters are compared.



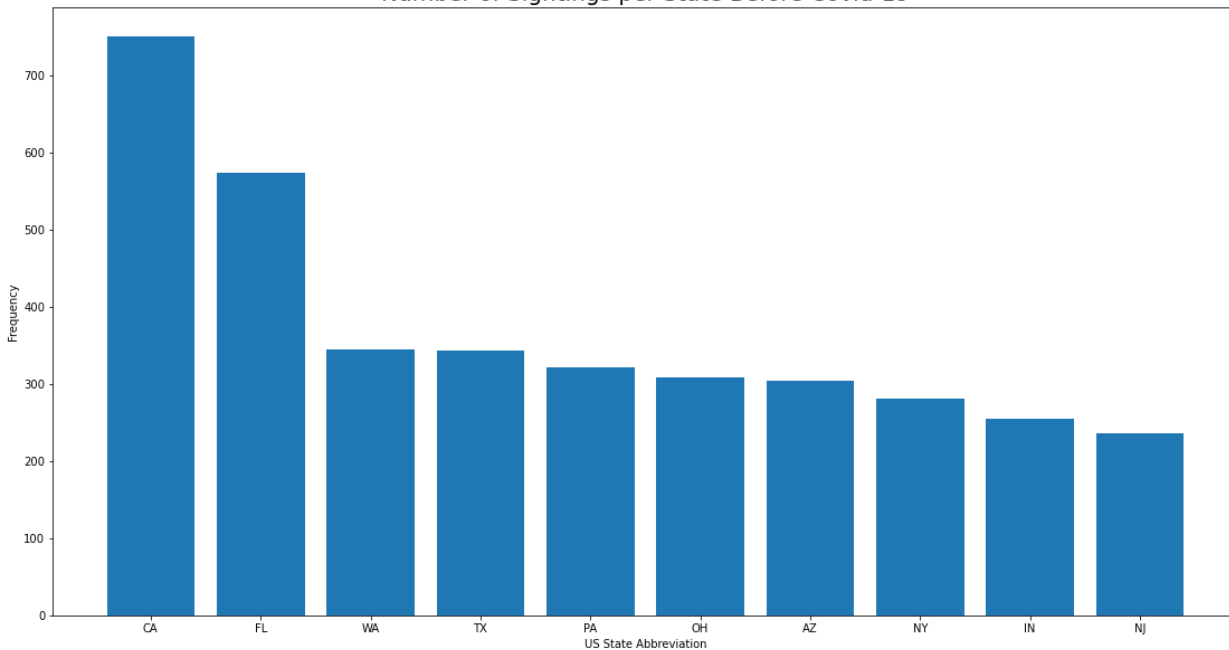
Comparing Reported Shapes Continued

- Light is the most reported shape for both groups.
- Circle is the second most reported shape for both groups.
- The third most reported shape for reports before Covid-19 lockdowns was triangles, after was spheres.
- “Other” shapes retained 4th place for both groups.
- Before Covid-19, the 5th most reported shape was unknown, after covid-19 it was triangles.
- Given these results, it seems that living during the Covid-19 pandemic did not seem to affect reported shapes. The differences are small and do not seem to indicate a discernable pattern.

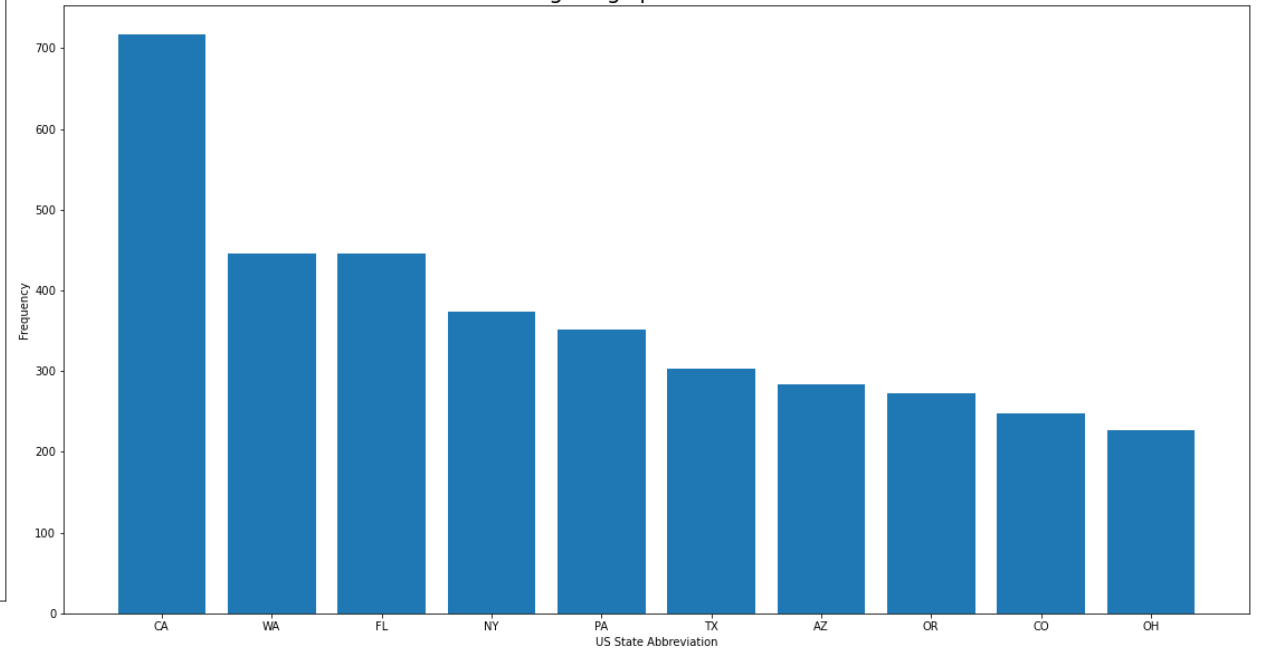
Comparing Sightings per State

The Top 10 states with the most sightings were compared from before to after covid-19 lockdowns began.

Number of Sightings per State Before Covid-19



Number of Sightings per State After Covid-19



Comparing Sightings per State Continued

- For the most part, the states seem to be in a similar order between groups.
- They both have California having the highest number of sightings followed by Florida, and Washington.
- New Jersey Uniquely shows up on the before covid chart while Colorado uniquely shows up on the after covid chart.
- New York has about 50 more reported sightings after covid-19 lockdowns started than before. Washington also rose by about 100 sightings after covid-19 lockdowns. This pattern goes against the expectation that the number of sightings should be lower for the after Covid-19 period.

Sentiment Analysis

- An off the shelf, pre-trained model called textblob was chosen for the sentiment analysis because it uses one scale of -1 (negative polarity) to 0 (neutral polarity) to 1 (positive polarity), which makes categorizing sentiment and comparing sentiment scores very simple to do.
- After getting the sentiment score, I categorized sentiment scores lower than 0 as negative, 0 as neutral, and higher than 0 as positive.

	date	state	country	shape	summary	sentiment	sentiment_category
9527	2020-03-14	TX	USA	Rectangle	cube shaped object silently glide neighborhood...	0.000000	neutral
9528	2020-03-14	MT	USA	Light	mysterious light sky looked like something sho...	0.250000	positive
9529	2020-03-14	WI	USA	Other	sky clear observed white starlike object weste...	0.033333	positive
9530	2020-03-14	PA	USA	Sphere	bright circle star like object	0.700000	positive
9532	2020-03-14	MN	USA	Light	white light moving fast medilla minnesota spa...	0.200000	positive
...
18461	2018-09-15	NJ	USA	Oval	saw bright oval light sky straight row flew a...	0.325000	positive
18462	2018-09-15	OH	USA	Circle	group bubblelike ufo seen ne ohio traveling nw	0.000000	neutral
18463	2018-09-15	NY	USA	Triangle	saw triangular object green red line appeared ...	-0.100000	negative
18464	2018-09-15	NJ	USA	Light	brilliant bright light brick nj	0.666667	positive
18465	2018-09-15	VA	USA	Light	multi color light light night sky temporary di...	0.400000	positive

8398 rows × 7 columns

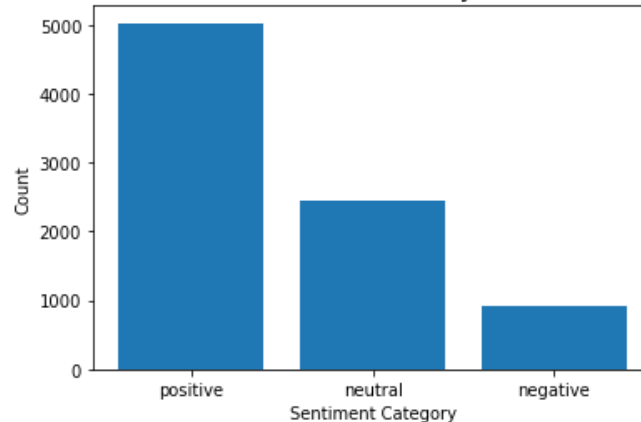
Sentiment Analysis: Sentiment Categories

- Report summaries with positive sentiment were the majority.
- Positive sentiment summaries were around 5 times higher than negative sentiment summaries.
- Positive summaries were 2 times higher than neutral before covid and close to that for after covid summaries.

Count of Before Covid
Summaries by Sentiment
Category:

positive	5029
neutral	2451
negative	918

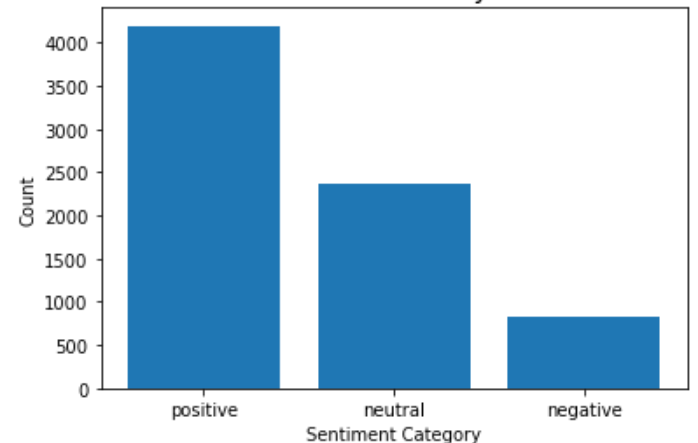
Count of Before Covid Summaries by Sentiment Category



Count of After Covid
Summaries by Sentiment
Category:

positive	4194
neutral	2369
negative	830

Count of After Covid Summaries by Sentiment Category



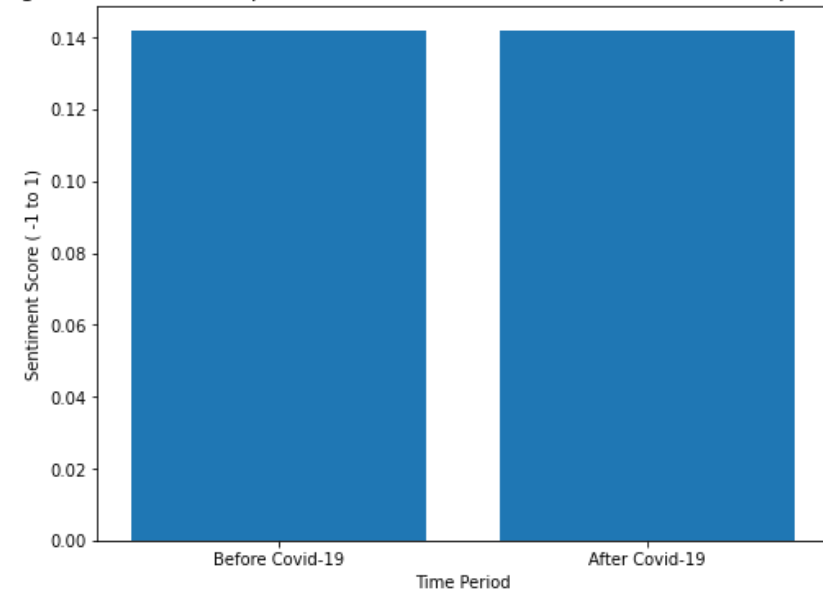
Sentiment Analysis: Comparing Average Sentiment

- Average sentiment of After Covid reports was about .01 lower than Before covid reports.
- These numbers seem to be very close.
- Average sentiment scores are close to 0 meaning that they are closer to neutral than a strong sentiment.

Before Covid mean of sentiment scores is: 0.14177733626496727

After Covid mean of sentiment scores is: 0.1333834429339673

Average Sentiment Comparison of Before Covid and After Covid Report Summaries



Sentiment Analysis: Hypothesis Testing

To test for significance, the Mann-Whitney U test was used again because it assumes very little about the population distribution.

```
Statistics=31998559.000, p=0.001  
Different distribution (reject H0)
```

- In this case the null hypothesis is rejected.
- There is evidence to suggest that sentiment is significantly lower in reports that were made after covid lockdowns began.
- Maybe living with the stress of the pandemic can explain why reports had higher negativity.

Overall Results

- There were significantly less daily reports during the Covid-19 pandemic than before in the United States. This was the opposite of what I predicted. It is possible that because people spent higher periods of time in the house, maybe they had less opportunities to see UFOs.
- The shapes of UFOs that people reported had very little differences.
- The number of UFOs in individual states had very little differences however, We do see a rise in incidences in Washington state and New York which goes against the overall pattern we discovered. On one hand, these states were known to have stricter lockdowns than other states, although California was too, and it had a lower number of incidents after covid began.

Overall Results Continued

- The number of report summaries with positive sentiment scores were around two times higher than neutral for before covid summaries and close to that for after covid summaries. In both time periods, positive summaries were around 5 times higher than negative summaries. According to these results, it looks like most sightings could have been positive experiences for people despite the time period.
- Most sightings had report summaries with positive leaning sentiment. People seemed to feel more positivity regarding what they witnessed. Very few reports reflected negative feelings towards their sightings.
- Sentiment from before Covid-19 lockdowns to after became significantly lower. This may be because pandemic in general did have many people stressed. This could have been reflected in people's incident reports.

Conclusion

There were some interesting finds and there was a lot that went well with this project although, there is always room for improvement. It is possible that the patterns seen in this data set could be uniquely attributed to this data only. Sentiment analysis could be improved if the full reports were used instead of just summaries of the reports. Lastly, another limitation to this study is that we do not have a foreseeable way to accurately measure stress over time. This project simply uses post covid years as a guess that stress levels were higher however, we do not know if those levels varied or even went down anytime during the post-covid data that was used.

Next Steps:

- Data from other UFO reporting services should be taken to see if there are similar patterns.
- Using a different event as a “Stressful event” such as the cold war to see if the same results would be found.
- To do another sentiment analysis on the data using full reports instead of summaries.
- It may be interesting to test out a different country than the United States.