

U-F-OH NO

CSC369 Project

Members: Anna Barone, Ethan Chan, Cameron Storton, Braden Wicker

Date: 15 March 2020

Project Overview

We set out to determine whether there is any correlation between the releases of science fiction movies, or movies of other genres, and the number of UFO sightings while the movies are in theaters. We estimated that a movie would stay in theaters for 30 days.

Program Description

The program first parses the movie dataset and the UFO sighting dataset. From the movie dataset, it extracts, for each movie, its release date, popularity, and list of genres. From the UFO sighting dataset, it extracts each date there were sightings and how many sightings occurred on that date.

After this, it collects the total UFO sightings within a month of release of each film, duplicates that data for each of the film's genres, normalizes the popularity by genre, and finally averages each genre's number of sightings weighted by the normalized popularity of that genre.

Hadoop was necessary for this project because the datasets used were relatively large. The UFO dataset was 13.28 MB and contained 80,000 reports, and the movies dataset was 32.85 MB and had metadata for 45,000 movies. Hadoop allowed for efficient parsing of all this data, taking an average of 16.810 seconds.

Experimental Results

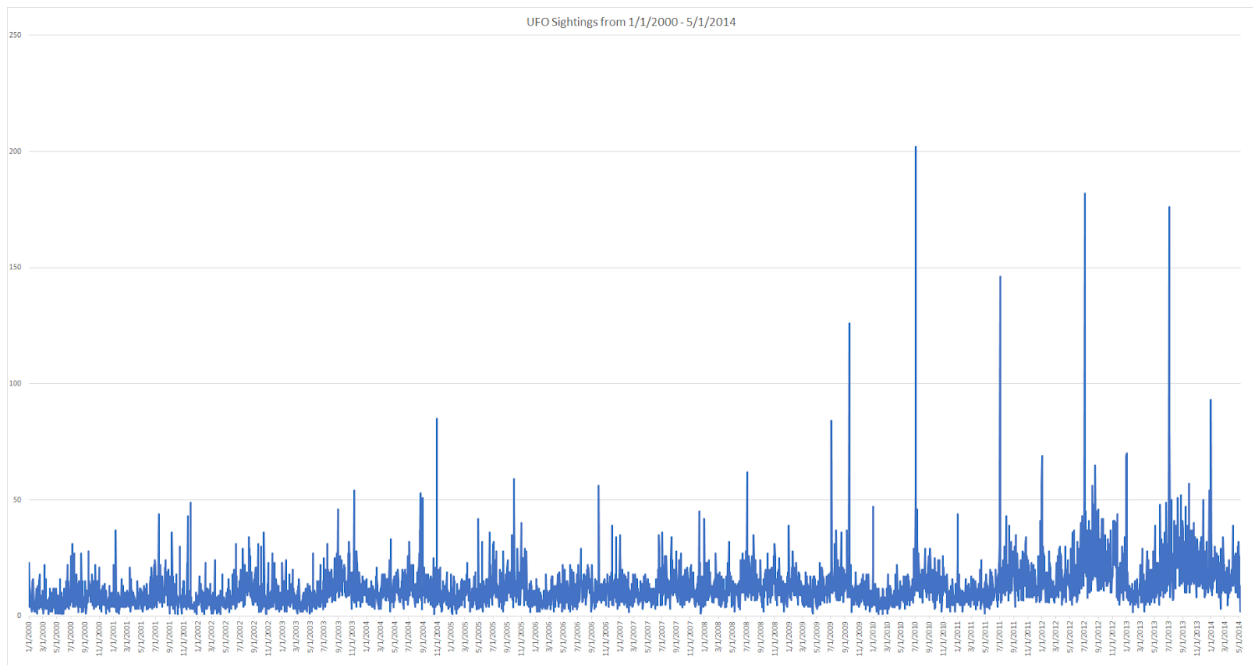


Figure 1: UFO Sightings Per Day

We first plotted the UFO sightings from 2000 to 2014, shown in **Figure 1**. From this data, we saw that there was an average of 12.4 UFO sightings per day and the highest peak was 202 sightings on a single day.

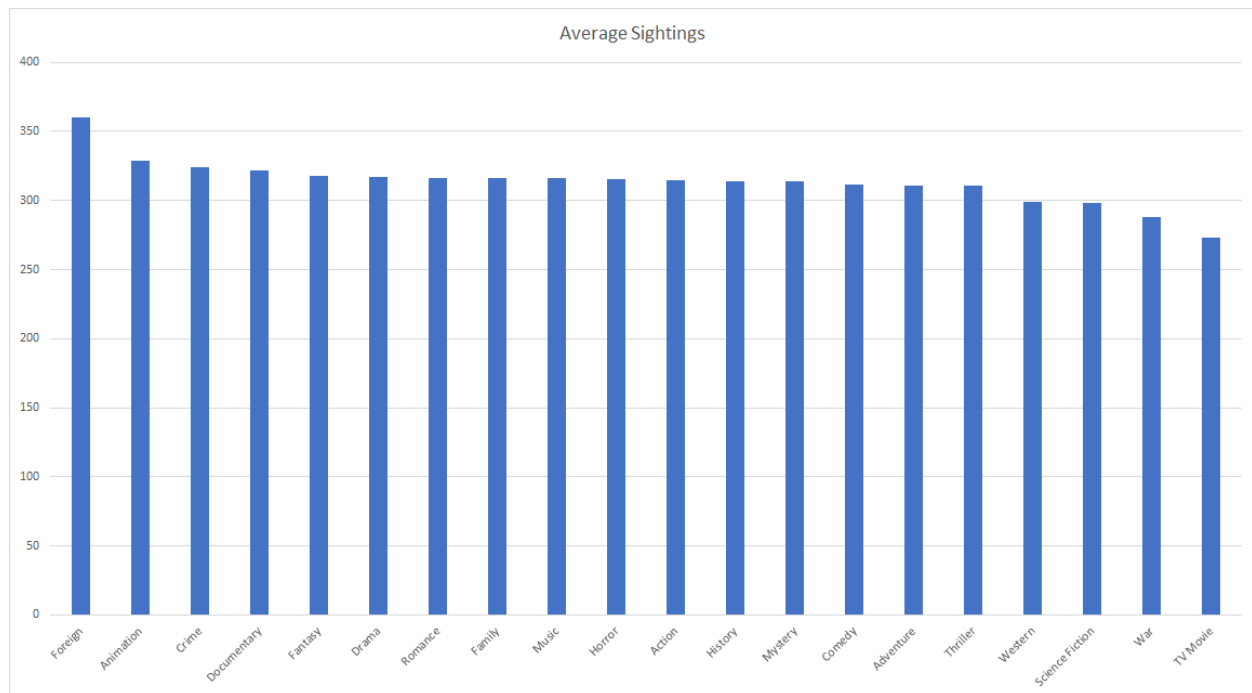


Figure 2 - Average Sightings Per Genre

Next was the average number of sightings within a month of release per movie for every genre listed in the Movies dataset, shown in **Figure 2**. If there was a correlation between science fiction and UFO sightings, we would expect the number of UFO sightings in the science fiction column to be much higher than all other genres. However, this is not true – foreign films have the highest average of UFO sightings within a month at 356. To verify that an average within the 300 range was reasonable, we looked at the previously stated average UFO sightings per day (12.4), so the average within a month – 370 sightings – checks out.

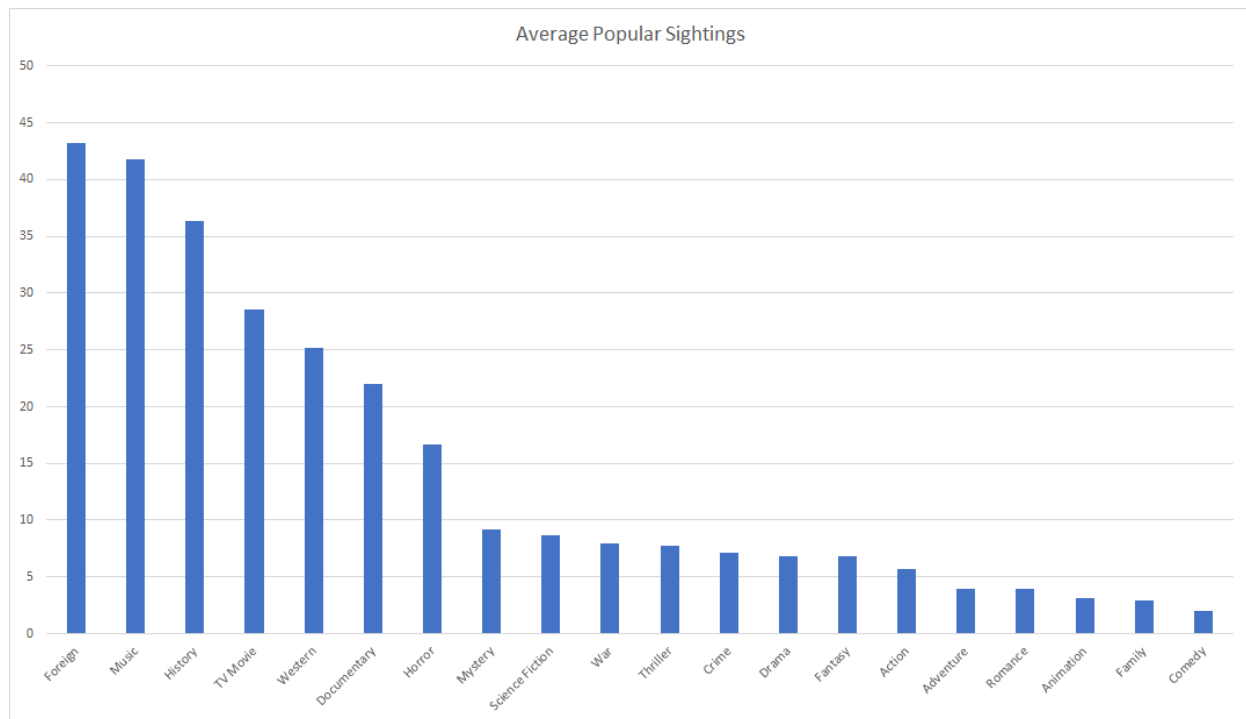


Figure 3 - Average Sightings Weighted by Popularity

Finally, in order to get more meaningful data we normalized each movie's popularity depending on its genre, shown in **Figure 2**. We made sure each genre was normalized since there's a chance that one genre's films could have been more popular than another's. This data, however, does not show much correlation either. Because each movie has multiple genres, certain subgenres appeared more than others. For example, many movies had genres Action and Music, or Animation and Music. Because of this, the actual genres that might have a slight correlation with UFO sightings are Western, Documentary, and Horror. Ultimately, the results found are not significant enough to conclude a correlation between movie genres and UFO sightings.