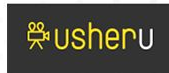


Movie Coast Guard

OvercookingSpaghetti

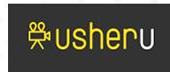


Problem

- Piracy in the movie industry has been steadily increasing
- There are not many tools available for tracking, control and statistics regarding it

Goals

- Gather information from torrents, including their peers, and subsequently their consumption patterns
- Make a comparison with relevant values such as the movie's production budget and box office
- Create a timeline

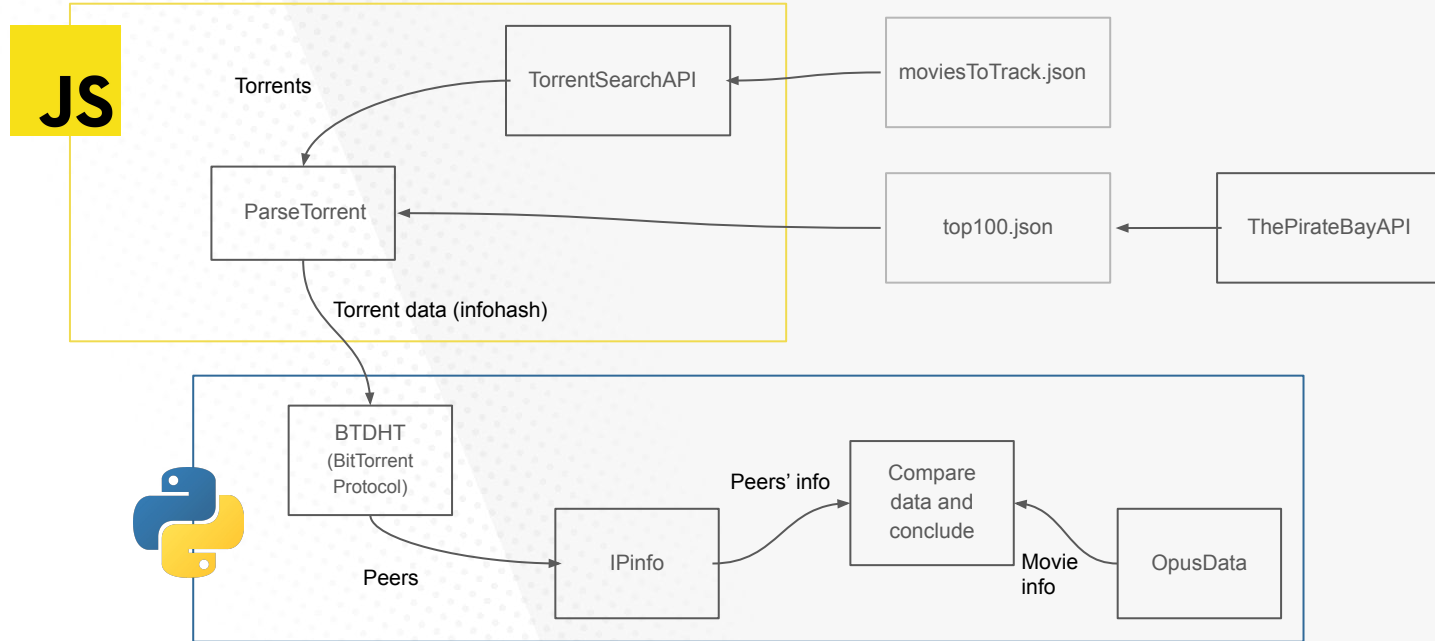


Approach

- Get the top 100 most torrented movies and extract relevant info (info hash, seeders and their IPs, geographical location)
- Gather relevant info from movies that the user provided
- Build a dataset with aforementioned information
- Cross reference with the information that was previously gathered and the OpusData dataset and try to answer the questions:
 - **What makes a movie more likely to be pirated?**
 - **What is the impact of piracy in the movie industry?**



Architecture



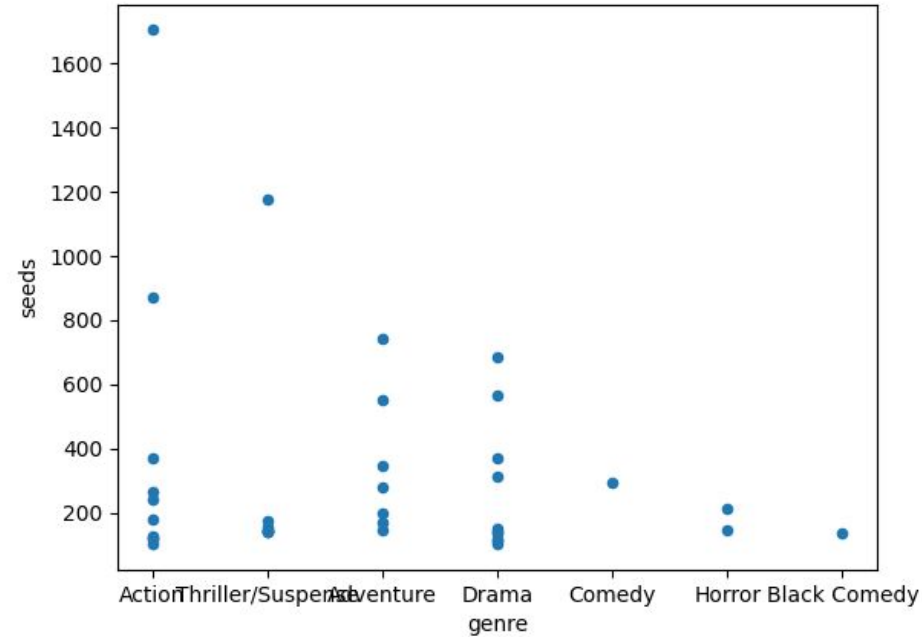
What makes a movie more likely to be pirated?

Genres:

- Action
- Drama

Peak seeders:

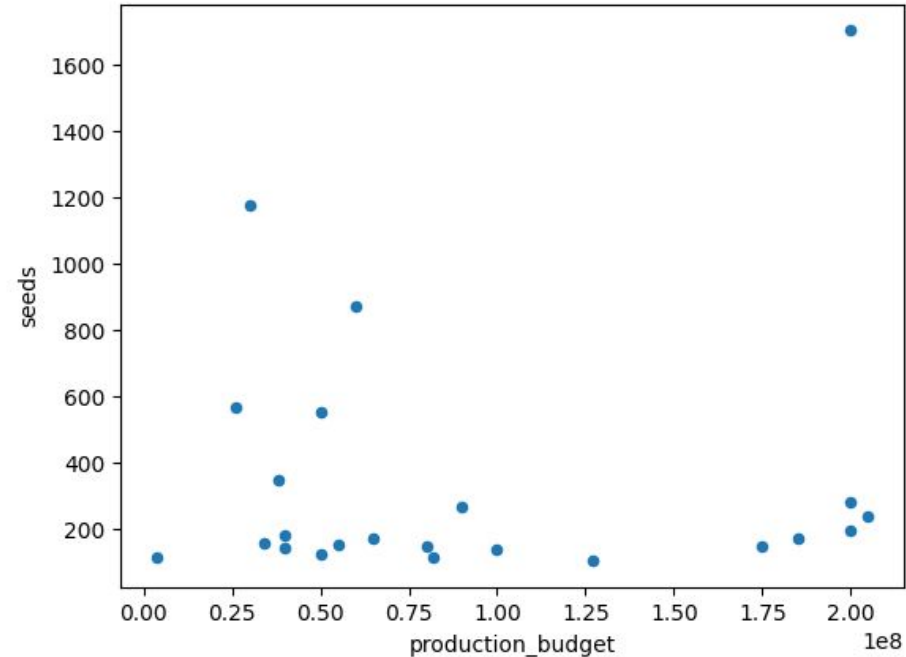
- 1600



What makes a movie more likely to be pirated?

Production Budget:

- Unclear whether it is relevant or not



What is the impact of piracy in the movie industry?

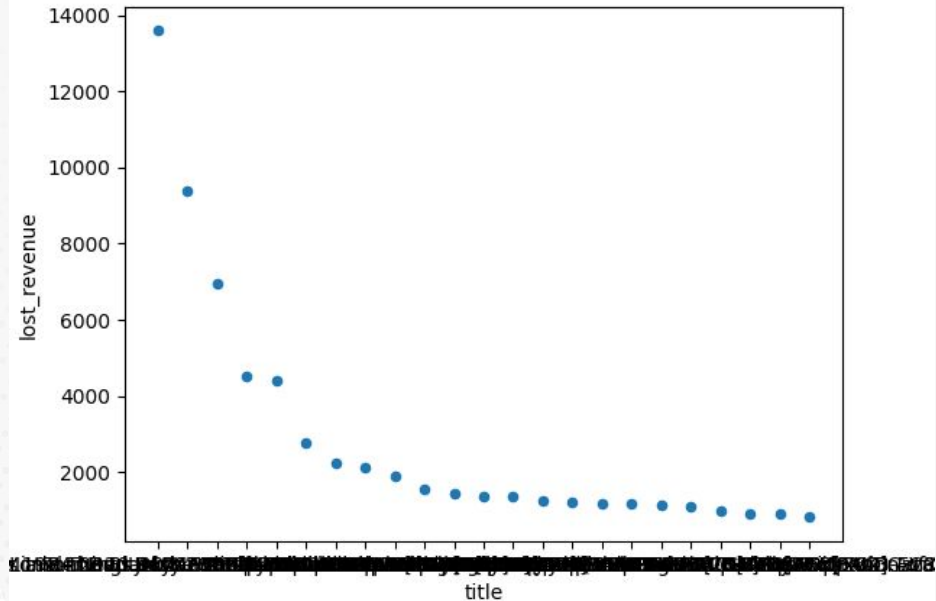
Avg. revenue lost:

- 2,794\$

Peak revenue lost:

- 14,000\$

Note: based on 2014 avg ticket price

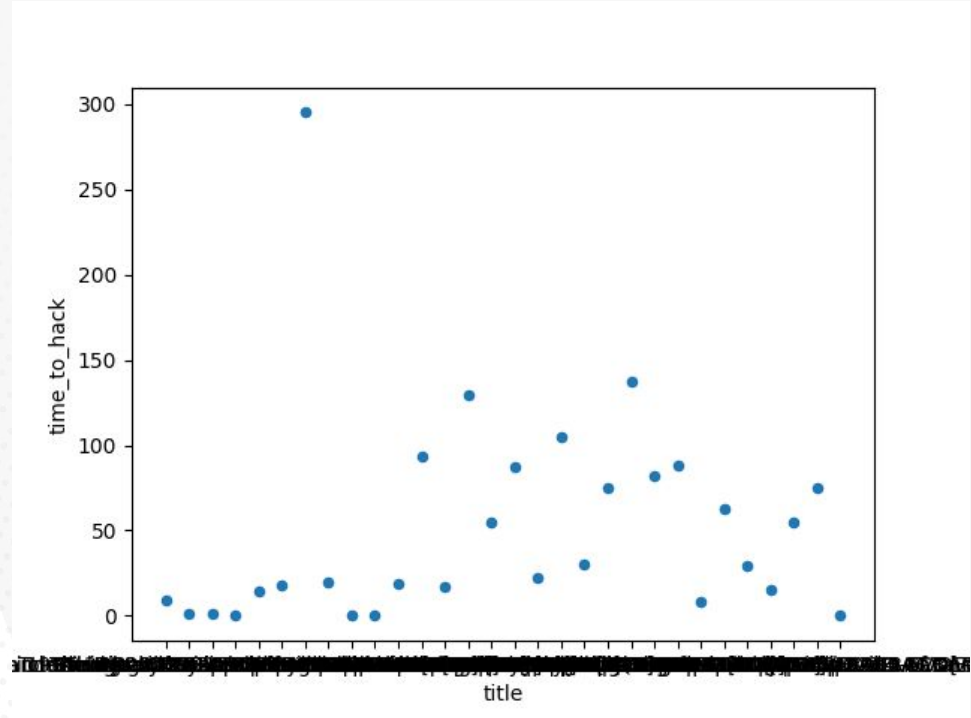


What is the impact of piracy in the movie industry?

Avg. Time to hack:

- 51 days

Note: it is very difficult to know the first upload date of a torrent



Possible Issues

- VPN
 - The conclusions may not be accurate due to the fact that some users might be using VPNs
- Downloaded != Viewed
 - There's no way to ensure that a user has actually seen what they downloaded. It is only possible to know if the peer finished downloading the torrent (thus becoming a seeder). Using the same logic, a single download can be shared with/seen by multiple users
- Counting the exact number of downloads
 - It is only possible to take into consideration users that are either downloading the torrent, or have finished downloading it and are still seeding it
- Average Movie Ticket Price
 - The database that we found has data from 2014
- Lack of data
 - Specially for recent movies



Future Work

- Periodically running server:
 - Make our code run periodically and store data about each torrent, in order to create a timeline
 - Count the number of distinct seeders of each torrent to get the total number of downloads
- Cross reference location:
 - Cross our data with the locations we found
 - Due to lack of time we could not measure downloads per location and, subsequently, their impact



OvercookingSpaghetti

- Andreia Gouveia
- Carlos Albuquerque
- Carlos Duarte
- Sofia Lajes

Thank you for your time!

