
Steam Game Recommender

Jeong Dam (James) Lee

Goal

A new era of games has started since many people are quarantined. As a PC Master-race, Steam is one of the most valuable gaming platforms for everyone out there. With thousands of games to choose from, how can we decide on the right ones?



Steam

“Steam is a video game digital distribution service by Valve. It was launched as a standalone software client in September 2003 as a way for Valve to provide automatic updates for their games, and expanded to include games from third-party publishers.”

[https://en.wikipedia.org/wiki/Steam_\(service\)](https://en.wikipedia.org/wiki/Steam_(service))



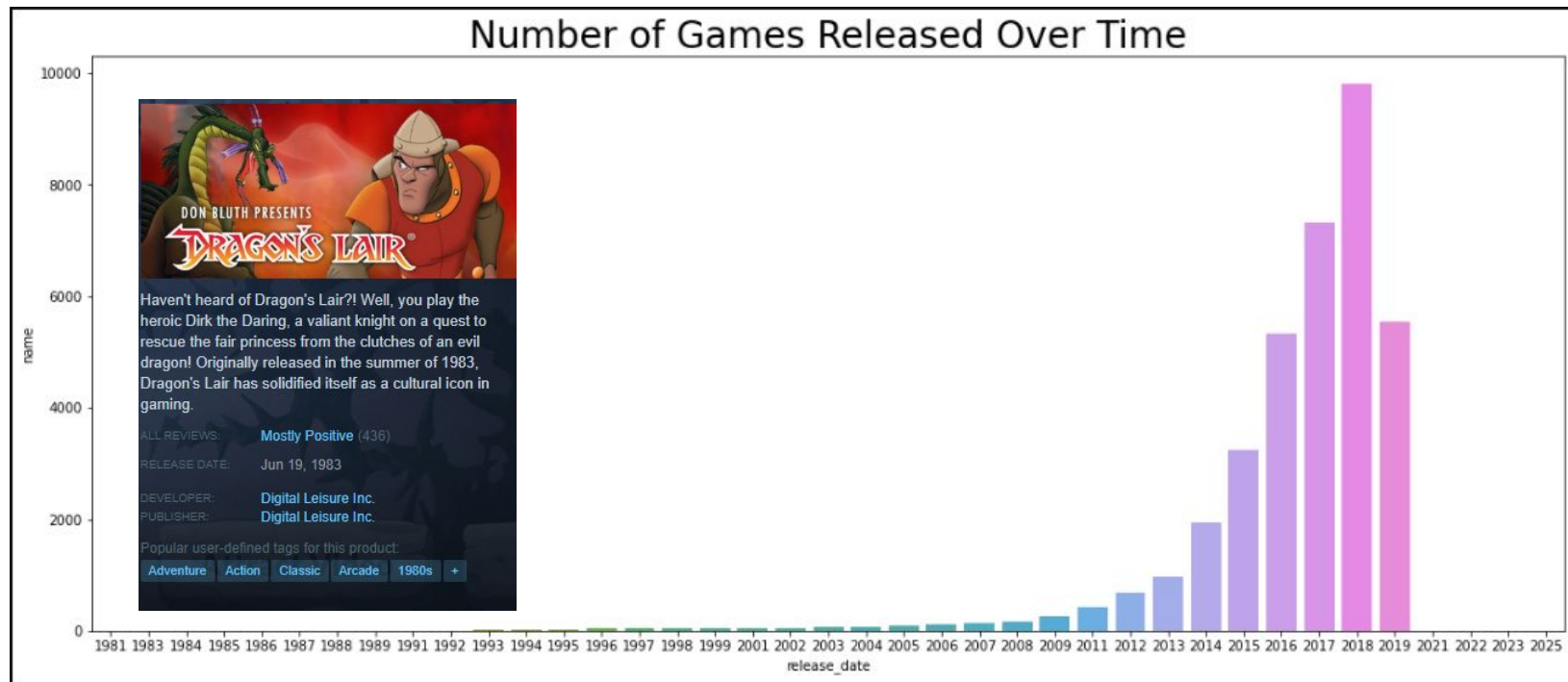
Data Collection

Kaggle and Steamspy.com

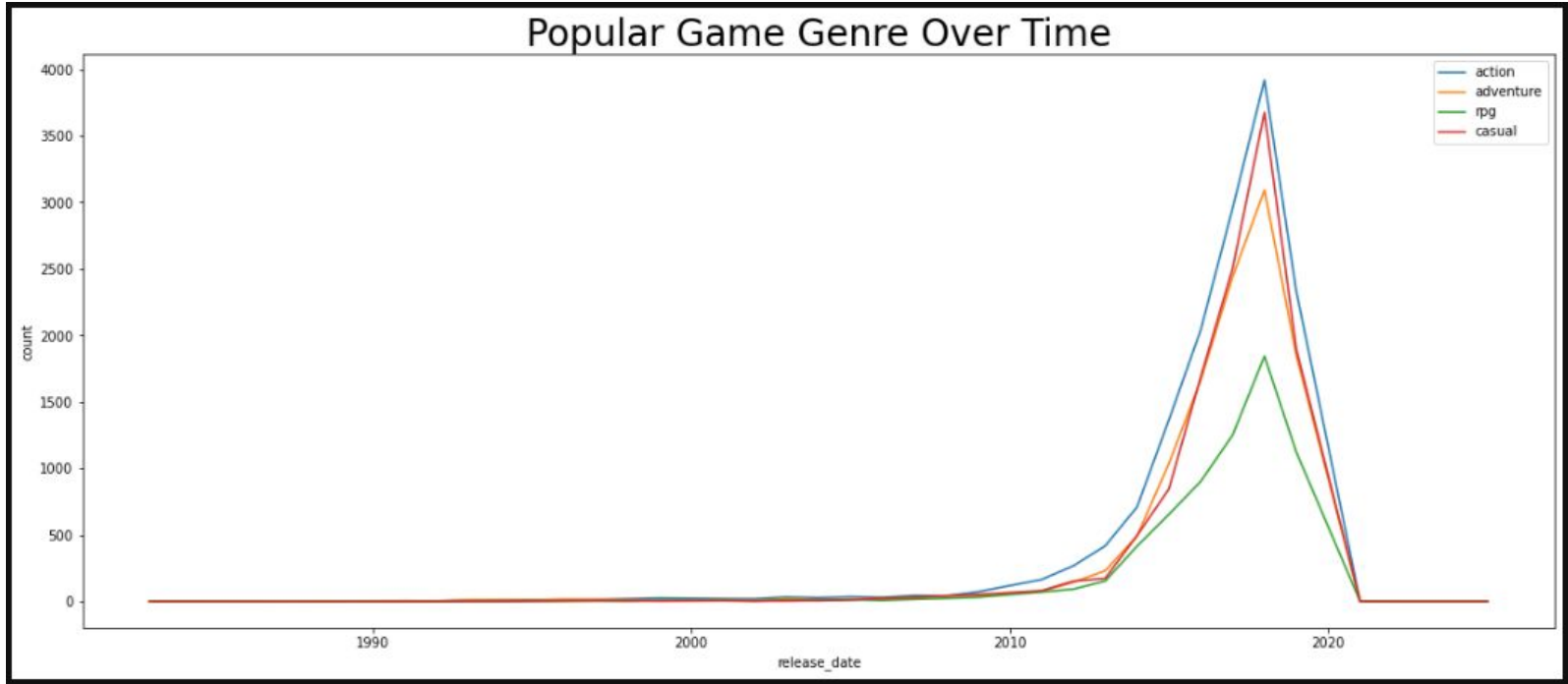
- **Kaggle**
<https://github.com/CraigKelly/steam-data>
- **SteamSpy API**

User friendly for all Steam games

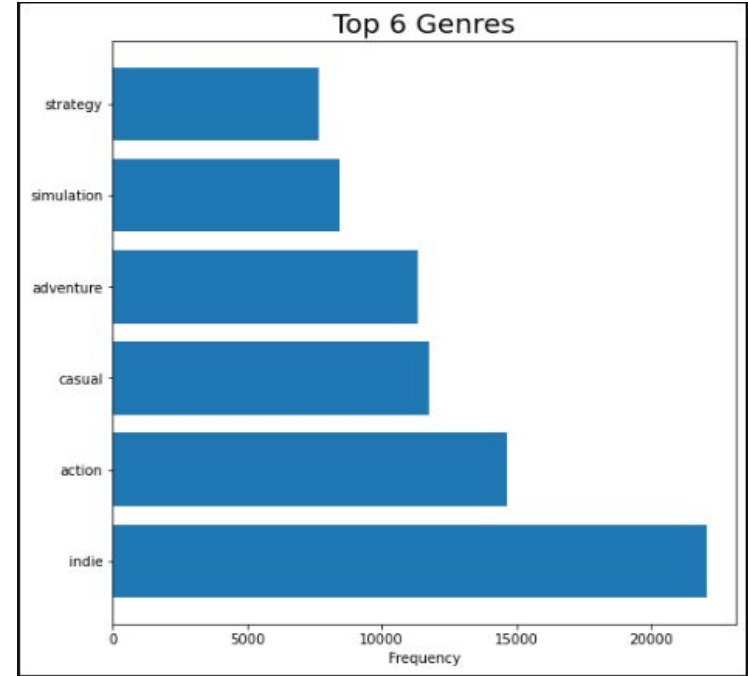
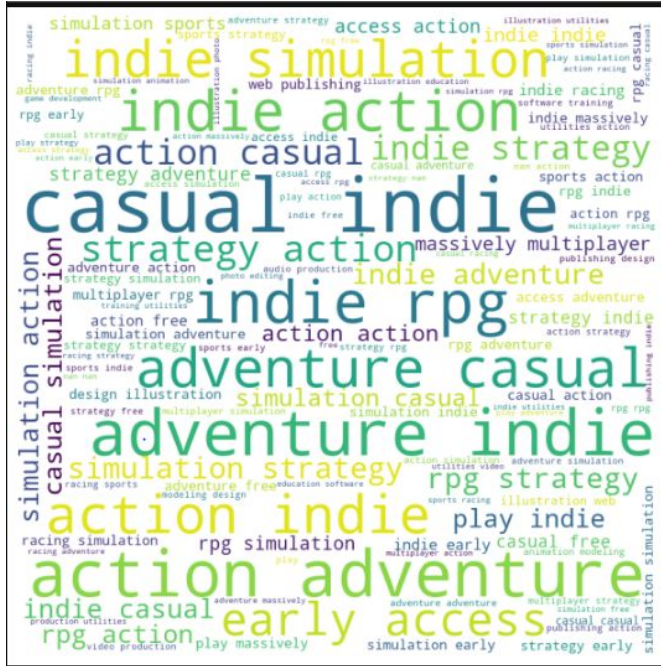
Games Released over the Years



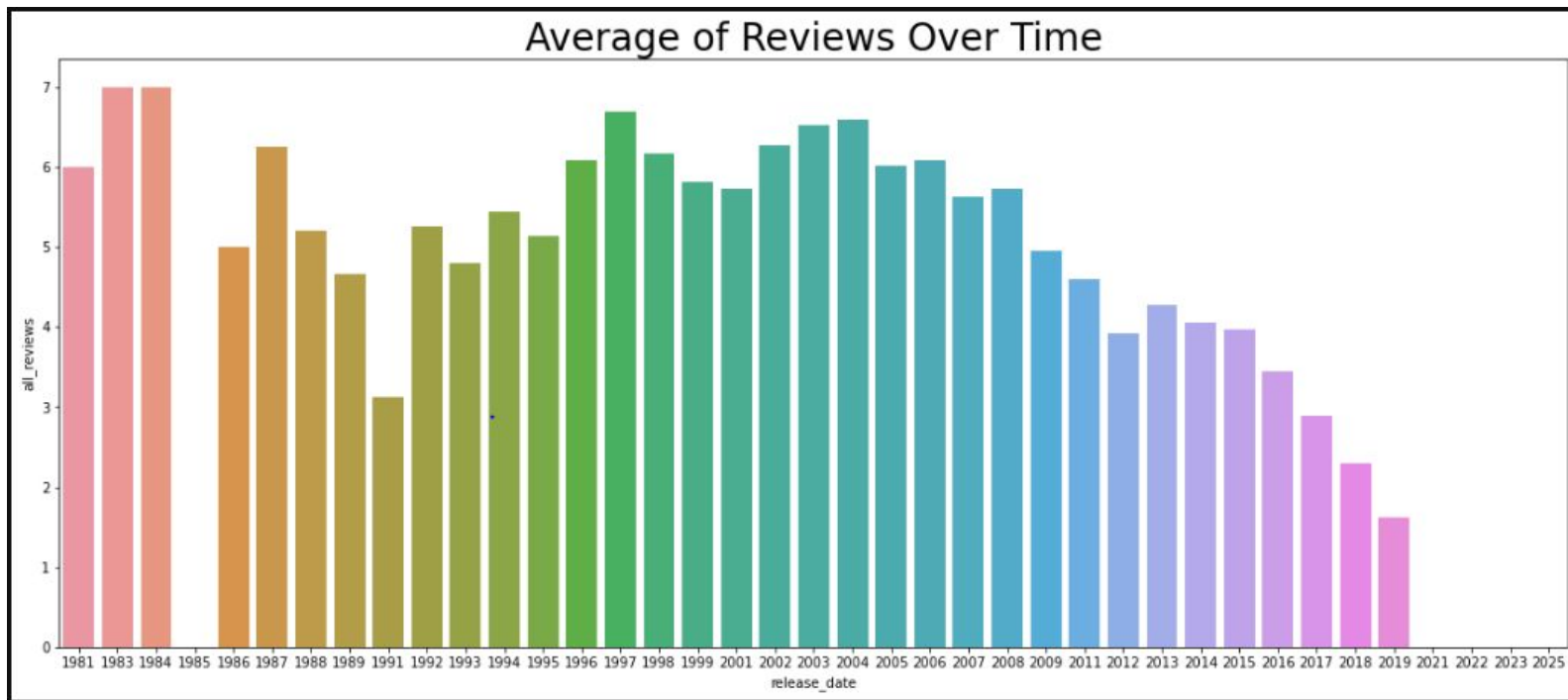
Genres of Games Released



Most common Genres



Average Reviews Over Time





Recommender

There are a few ways to create a recommender

→ **Title**

Using the context in the title of the game

→ **Collaborative Filtering**

Using "texted" description or reviews

→ **Feature Based**

Using the features that the game contains

Feature Based

- Tags/Genre
- Publisher



Call of Duty is back, redefining war like you've never experienced before. Building on the Call of Duty 4®: Modern Warfare engine, Call of Duty: World at War immerses players into the most gritty and chaotic WWII combat ever experienced.

RECENT REVIEWS: **Very Positive** (570)

ALL REVIEWS: **Very Positive** (26,737)

RELEASE DATE: Nov 18, 2008

DEVELOPER: **Treyarch**

PUBLISHER: **Activision**

Popular user-defined tags for this product:

Zombies **World War II** **FPS** **Multiplayer** **Action** +



CD PROJEKT®

CV and Tfidf

CountVectorizer on Genres - Word frequency

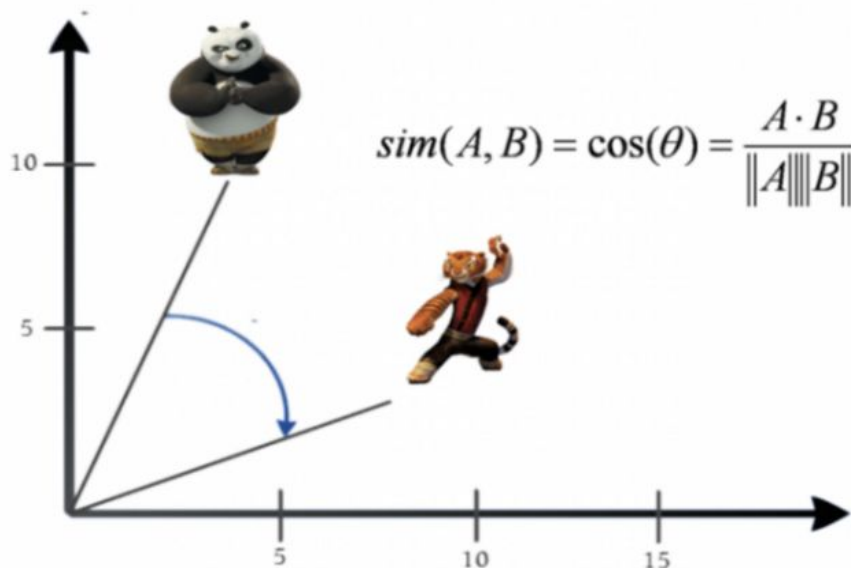
TfidfVectorizer on Publisher - Numeric score of each publisher

Cosine Distance

Similarity between two vectors calculated by the angle between them

Different type: Euclidian Distance - similarity between two vectors calculated by the straight line distance between the two points.

Cosine Similarity



	Dota 2	Counter-strike: Global Offensive	Team Fortress 2	Playerunknown's Battlegrounds
name				
Dota 2	0.000000	0.424803	0.458668	0.546461
Counter-strike: Global Offensive	0.424803	0.000000	0.301747	0.330651
Team Fortress 2	0.458668	0.301747	0.000000	0.420454
Playerunknown's Battlegrounds	0.546461	0.330651	0.420454	0.000000
Unturned	0.683641	0.645161	0.544617	0.544842
Warframe	0.570331	0.544842	0.496047	0.400000
Left 4 Dead 2	0.509501	0.371245	0.383608	0.412905
Grand Theft Auto V	0.818166	0.621234	0.588694	0.419619
Terraria	0.547089	0.650750	0.641431	0.683772
Brawlhalla	0.569452	0.609066	0.567156	0.618268

```
df["Playerunknown's Battlegrounds"].sort_values()[1:13]
```

```
name
Realm Royale           0.206783
Orion: Prelude          0.235907
Tom Clancy's Splinter Cell Blacklist  0.249806
Arma 3                  0.264785
Rise:30 Minutes To Extinction  0.269118
Spacelords              0.278289
Primal Carnage: Extinction  0.280853
Z1 Battle Royale        0.280853
God Mode                 0.282781
Infected Battlegrounds  0.283885
Darwin Project           0.286932
Primal Carnage           0.286932
```



For Future

Include different platforms and consoles for more variety.

Include Collaborative Filtering and Title based and have a filter system to see more options.

Find more features/ use different features.

Let's see the app.