Keyang Yan

CS 210 final report

7/16/2024

Relationship with Video Game Popularity and Pop Culture

**- Define Project:**

In today's era, there is no doubt that electronic games have become a very important part of our lives. Meanwhile, although electronic games are carried by computers or game consoles, they are essentially cultural products that are easily influenced by today's popular culture.

My project aims to analyze the trend of electronic games and their relationship with popular culture. The goal is to identify how various cultural phenomena, such as movies, music, and social media, affect the popularity of electronic games.

The ultimate goal of this project is to provide video game makers with a reference for content that they can succeed in the world of game production.

**-Strategic aspects:**

1. Learn about the impact of pop culture on the popularity of video games.

2. Identify the key cultural events that drove the surge in gaming popularity.

3. Develop predictive models to predict future trends based on cultural indicators.

**- Novelty and Importance:**

Novelty:

Nowadays, the game content on the market is often classified by game genre when it is produced, such as RPG or FPS, and the culture of its connotation is only a supplement to make up for the game genre. And my current project will provide a new direction for game makers to connect games with culture and find out the direction of the market.

Importance:

From the perspective of the game industry: game developers and marketers can create and promote their own games in terms of pop culture, so as to liberate themselves from the competition of their original "game genre", so as to obtain better market effect.

From a cultural point of view: let every pop culture add a new form of carrier, so that it can gain more attention and popularity.

From the consumer's point of view: when enjoying a culture, you can get new forms of entertainment in addition to music, film and literature, expand your own eyes and better enjoy the advantages brought by modern society.

Related Research:

An analysis of the popularity of pop culture and its related games during the same period.

A study on the impact of social media trends on the popularity of video games.

**-The problem of the data**

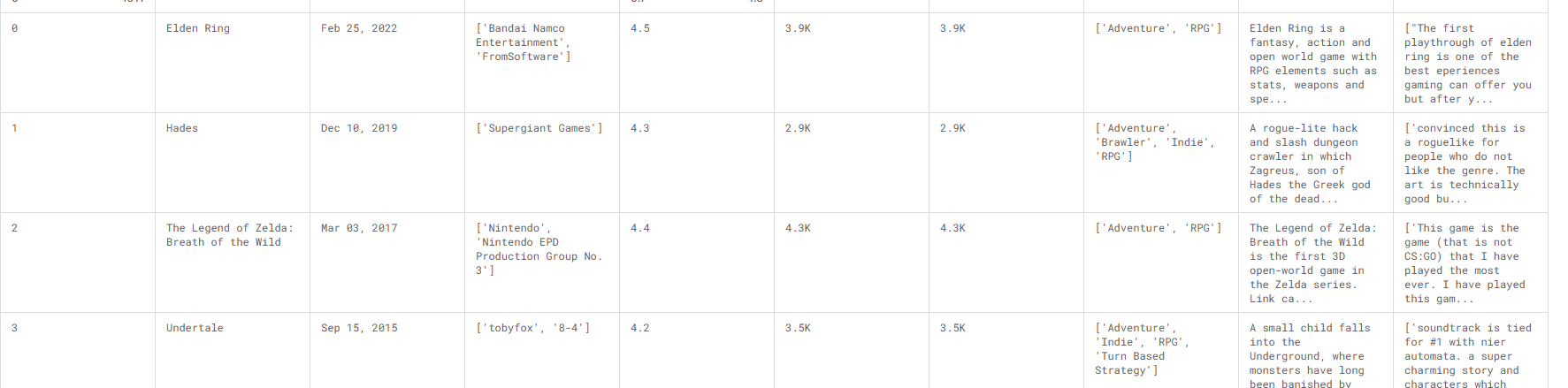
The problem with the data is that the data obtained from the Internet is often the data that the first author has refined and left over that they think is important, but this can cause a problem with my dataset: the refined data often does not represent some of the data information I need very well, and the rest of the information that I don't use is very complicated.

2. **Data Integration**:

During the 10 days of project preparation, I was not able to get some complete first-hand information in person, such as the games I want to discuss now, they often have some game genre generalizations, but the generalizations of these topics are not official summaries, but the competent experiences given by the authors of the datasets based on their own experiences.

3. **Scalability**:

When I discuss these things about the culture of the game, there are often a lot of points of ambiguity. For example, the 2022 game of the year "Elden Ring", which has been widely praised in the game industry, its most attractive is the challenging nature of the game, but its cultural background is also an element that players talk about, written by the well-known author "George · R· Written by R. · Martin, the apocalyptic context of a "world of decline from prosperity and death" is related to the real world of transition from prosperity to conflict and death after the Covid-19 pandemic that ended around 2022, but it is not well reflected in this table. Elden Ring is because I've played it, so I know a little bit about it, and I know how it relates to pop culture, but I can't understand all the problems in the game data over the past few decades, so the relationship between culture is not a perfect reflection.



4. Data Privacy & Security:

Compliance with regulations: Complying with data privacy laws (e.g., GDPR, CCPA) requires strict data management practices.

Data breaches: Protecting sensitive data from unauthorized access and leakage is critical.

Data Governance:

5. Lack of standardization:

When I process the data, the logic I use must be different from the direction that the source author wants to study, so there are often some key deviations, which results in the omission of some important information, and I cannot find the missing information from the source author without first-hand information.

6. Real-time data processing’s latency:

This statistic spans 24 years, and many of these wanted information are often influenced by other factors, and even cultural popularity can be arbitrarily modified over time, so my research can only be used as a second-hand reference.

- Mathematical models

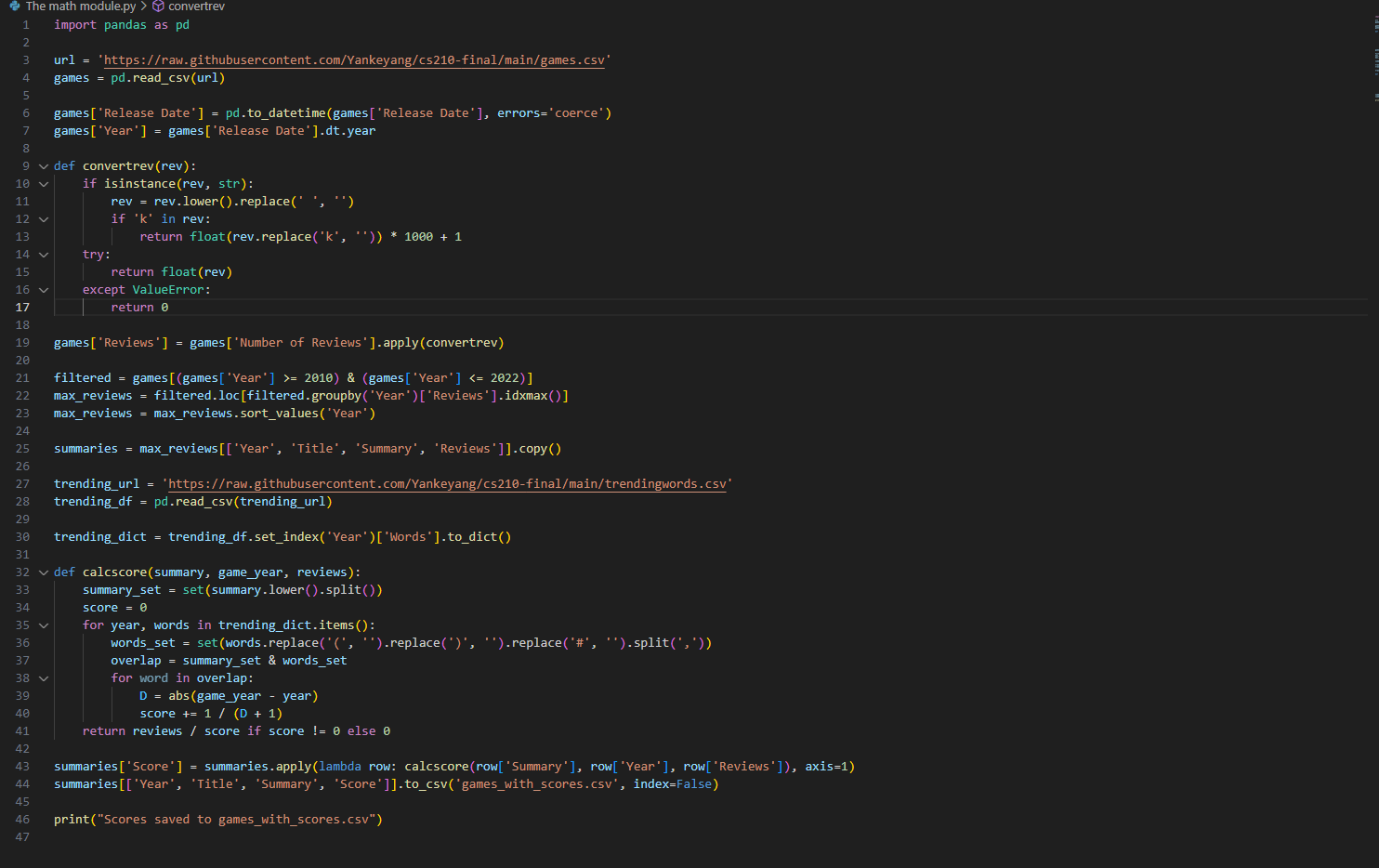
1.We can design a formula that combines these factors and represents the rating of the game's popularity in relation to the trend words P is the popularity of the game, expressed by the number of reviews. O is the coincidence degree of trend words, expressed by the number of trend words contained in summary. D is the difference between the year of the trend words and the year the game was released. It is shown that the relationship between the two is inversely proportional, that is, the higher the value of S, the less influenced by trend words.

S=

Explain:

n is the number of overlapping words, O is the number of occurrences of the 𝑖 overlapping word (usually 1), and D is the year difference of the 𝑖 overlapping words.

Here is the Mathematical model code



*For example, suppose there are two overlapping words, the first one is one year apart, the second is not, p is 3000, and in the other case there is only one overlapping word, only one year apart, the result is:*

*When n=2 ，d1=1 d2=0. P=3000 S=2000*

*When n=1, d1=1, P=6000.*

This result shows that the more the overlap with popular words, which we derive here as pop culture, the shorter the distance between the time interval, the smaller the value of S, which indicates that the higher the overlap between pop culture and video games, the proportional relationship.

-experiment prediction

All the data values obtained from the data, after plugging into the formula, the higher the coincidence degree of the game, the smaller the value of S, indicating that the greater the influence of popular culture.

**- Hypothesis:**

I hypothesize that the higher the degree of similarity between the sales of video games and the cultural products of the same period, the higher the sales, and that the relationship is proportional, and that cultural events significantly affect the popularity of related video games.

-Data processing

(A) Data collection/Data Storage/ Data Integration:

Get free working data resources from the <https://www.kaggle.com> website. There are two data collections that are currently planned to be used:

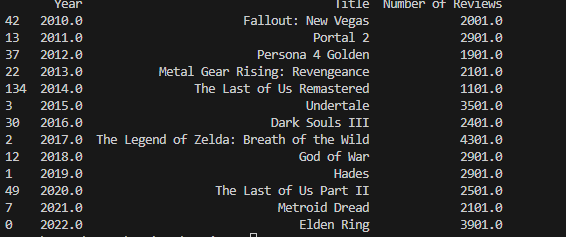
[Popular Video Games 1980 - 2023 🎮 (kaggle.com)](https://www.kaggle.com/datasets/arnabchaki/popular-video-games-1980-2023)

1. The popularity of the hottest game every year.

Now, we're going to take the most popular video games from 2010 to 2022 as examples from the dataset and use them as references.

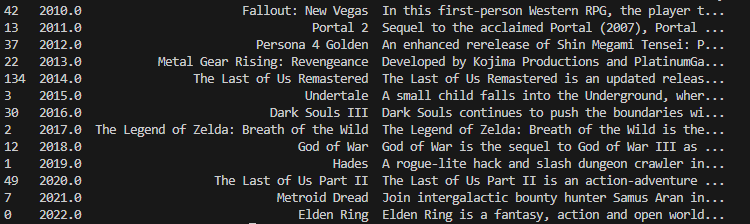
At the same time, we don't have direct access to the amount of investment and sales of each company in the game. Therefore, we will roughly use the number of player reviews on well-known gaming platforms as an evaluation.

HERE IS THE COLOUM OF VALUE OF P.



2. Now we've pulled out a summary of the keywords that commentators have given to the game each year, and those statistics will help us compare them to popular words.

Here is the summary of the Games and using for calculating the overlap.



(B). Modification: Based on the original model analysis, the natural language analysis of Twitter tweets is indeed a bit difficult.

Therefore, I used the statistics of hot words in Wikipedia for each year as my data source, and used pandas as a tool to build a csv table of the top five hot words in terms of the number of searches per year, as follows.

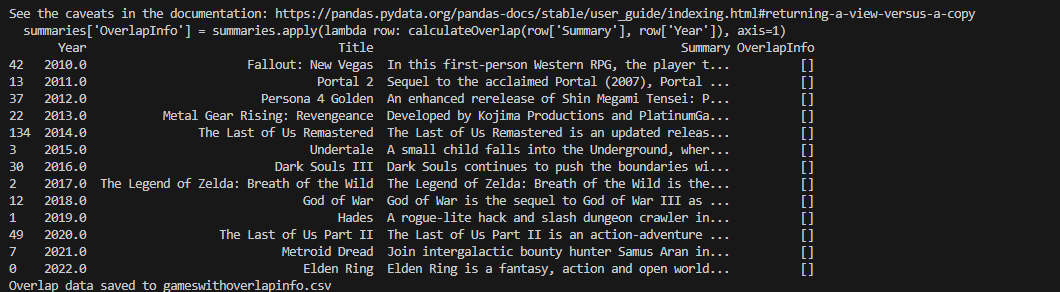
([Lists of Merriam-Webster's Words of the Year - Wikipedia](https://en.wikipedia.org/wiki/Lists_of_Merriam-Webster%27s_Words_of_the_Year))



(The above is the table I have listed after the statistics, including the popular words from 2010 to 2022)

(C). Now we need to analyze the data mentioned earlier, first we need to determine the year, and analyze the summary of the game content with the players in the games, here I have written an overlap to compare the summary of the games with the content in the table of trend words.

And calculate the year of the difference (**D**)



The results here show that our output is that there are no overlapping trend words elements.

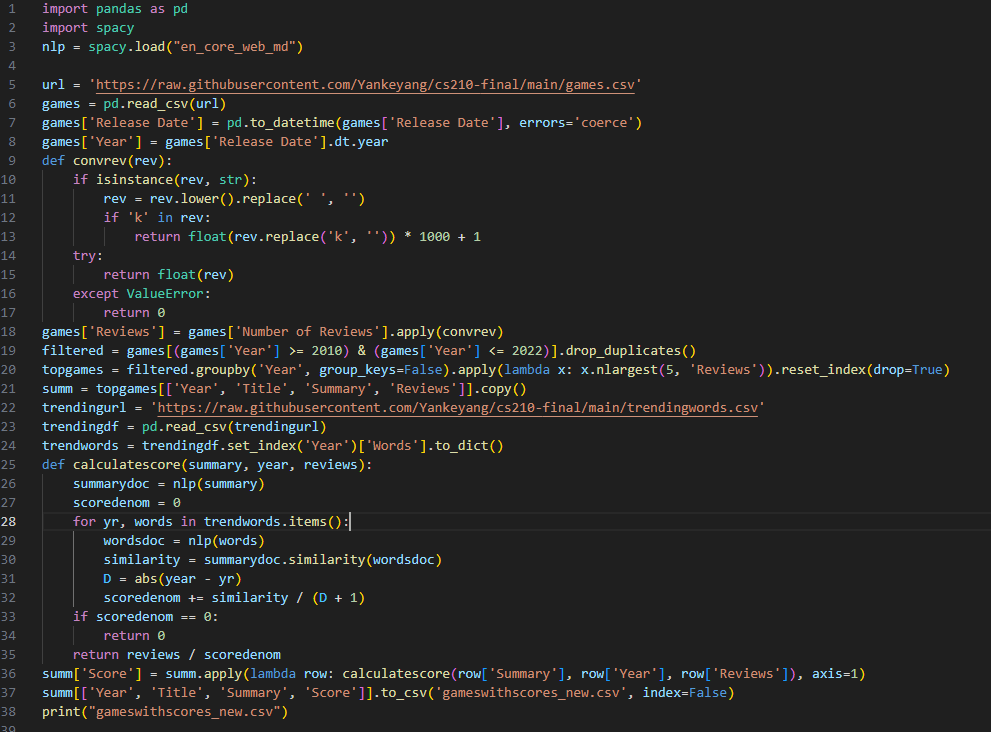
**The problem why we fail in this experiment:** I think the problem here is that the summary given by the player does not contain a complete representation of the cultural symbols that the game contains, so the experiment fails here.

-**Modification (because I had not planned to go ahead with the completion, but saw the notice about NLP and started over to finish the project, so I kept the previous work).**

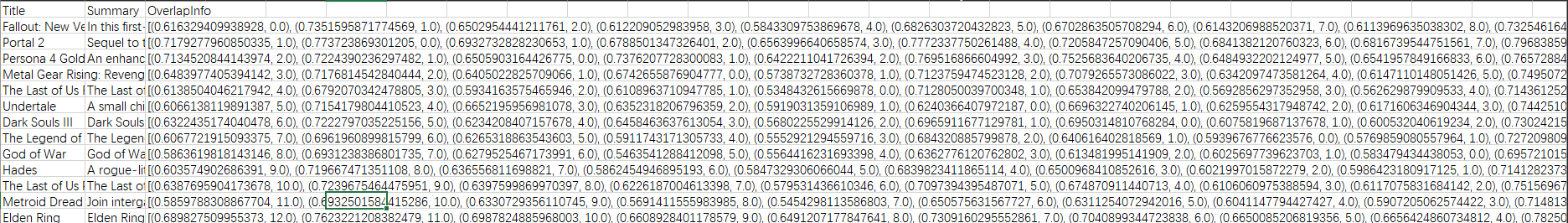
**1. Since the coincidence degree is not completely consistent, I will add natural language analysis based on the relationship between summary and trend words. Reformulate a mathematical formula:**

**Important update! Since similarity is calculated by division, the larger the value of S, the greater the impact of popular words on ranking, that is, the positive relationship.**

S=

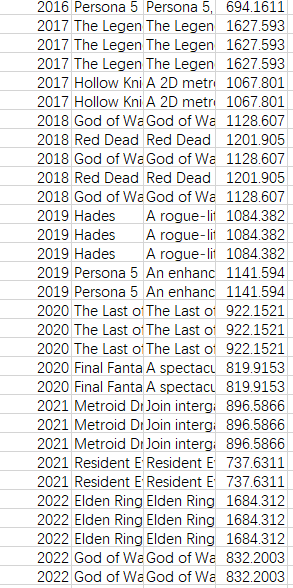


2.Similarity == Based on natural language similarity to buzzwords in summary's description of the game.

**Now we use the NLP language and the recommended model to analyze the similarity of various values, as shown in the following figure**.

Mathematical model code Final edition.

3. At the same time, because of the use of similarity, I find that if I want to conform to requirements and make a clearer comparison, I decide to use the top five S of each year for comparison as examples.



The proof is successful, and here are my results. Based on the previous ranking results, for example, elder ring is the game with the highest P value in 2022, it is predicted that its S value should also be the highest, and the experimental results show that it is.

Example: S of GOD OF WAR = 832.2003

S of Elden Ring = 1684.312.

And Elden Ring is the most popular game in 2022 and the winner of the TGA(The Oscars of the video game)’ game of year.

4. Summary:

This function successfully shows a positive correlation between popular cultural concepts and game sales, so that game companies can make successful games based on recent cultural trends.

- **Final summary:**

Limitations and shortcomings: The summary of each game is based on the player's subjective generalization, which may not be very accurate, subjective assumptions affect the accuracy of the results, and the player reference base is small (compared to the overall player).

Advantage: It has been successfully proven that cultural trends are a key factor in game sales, and the more popular culture-related games bring in more traffic and players.

Sources:

1. [PKU-TANGENT/nlp-tutorial: NLP新手入门教程 (github.com)](https://github.com/PKU-TANGENT/nlp-tutorial)() <https://github.com/PKU-TANGENT/nlp-tutorial>.
2. <https://en.wikipedia.org/wiki/Lists_of_Merriam-Webster%27s_Words_of_the_Year>
3. [Popular Video Games 1980 - 2023 🎮 (kaggle.com)](https://www.kaggle.com/datasets/arnabchaki/popular-video-games-1980-2023)