

Atividade 05 - Resultados Finais

1 - Tema:

SteamDE: Análise de *reviews* de jogos na plataforma Steam.

2 - Equipe:

- Daniel de Viveiros Inácio, 1923820, tredeneo, daniel.060798@alunos.utfpr.edu.br, BSI, UTFPR;
- Eduardo Darrazão, 1906399, eduponto21, darrazao@alunos.utfpr.edu.br, BSI, UTFPR;
- <https://gitlab.com/eduponto21/steam-de>

3 - Objetivo:

O objetivo deste trabalho é investigar as seguintes perguntas referentes a *reviews* de jogos na plataforma Steam:

1. Jogos com características similares possuem intersecção de tópicos abordados?
2. Como os *reviews* se comportam quando estão - ou não - recomendando o jogo analisado? Tem alguma relação com a quantidade de horas jogadas?
3. Qual é a relação entre a aceitação do jogo (medida pela porcentagem de recomendação dos *reviews*) e a quantidade de horas jogadas? Jogos com mais recomendações tendem a ser jogados por mais tempo?

4 - Dados e Modelos:

Para responder às perguntas e testar as hipóteses levantadas, foram utilizados dados de reviews de jogos coletados da plataforma Steam, disponíveis no Kaggle¹. Os dados contém *reviews* (majoritariamente em inglês) de jogadores de 48 jogos diferentes na plataforma Steam entre os dias 20/12/2010 até 16/02/2019, onde para cada *review*, temos os dados: data da postagem, título do jogo, *review* em si, se o autor do *review* recomenda ou não o jogo, quantas horas o autor jogou o jogo, quantas pessoas marcaram o *review* como engraçado, quantas pessoas marcaram o *review* como útil, e se o jogo estava em *early access* ou não. Os dados citados foram pré-processados para remover todos os *reviews* que não estavam em inglês, e removidos os dados referentes a jogos que possuíam 10 comentários ou menos (foi removido também o software Wallpaper Engine, que apesar de ter muitos comentários, não é de fato um jogo), sobrando assim 24 jogos contendo 398.591 *reviews* ao todo.

Com o intuito de analisar tópicos de jogos similares, bem como delimitar as análises de sentimentos, recomendação e aceitação dos diferentes *reviews*, criamos 7 agrupamentos de jogos, sendo eles:

¹ <https://www.kaggle.com/datasets/luthfim/steam-reviews-dataset>

1. **Singleplayer:** jogadores jogam sozinho usando apenas os sistemas do jogo em que não é necessário conexão de rede
 - a. factorio, euro truck simulator 2, slay the spire, subnautica, the elder scrolls v skyrim special edition, rimworld, resident evil 2;
2. **Coop:** jogadores jogam em grupo para vencer/alcançar um objetivo, geralmente podendo ser jogado sozinho
 - a. astroneer, don't starve together, divinity original sin 2 definitive edition, monster hunter world, stardew valley, terraria, left 4 dead 2;
3. **Multiplayer:** jogos em que o foco está em jogar com outros jogadores, jogando sozinho ou em grupo contra outros jogadores e/ou sistemas do jogos que requer conexão de rede(nos casos em que não jogam contra outros jogadores o mundo é online e compartilhado)
 - a. ark survival evolved, garry's mod, dead by daylight, sid meier's civilization vi, rocket league, human fall flat, insurgency sandstorm, rust, grand theft auto v, playerunknown's battlegrounds;
4. **FPS:** *first person shooter* são jogos de tiros em primeira pessoa, aqui juntamos alguns que não necessariamente são em primeira pessoa
 - a. grand theft auto v, playerunknown's battlegrounds, left 4 dead 2, resident evil 2;
5. **RPG:** *role playing game* são jogos que o jogador interpreta/controla um personagem que foi customizado por ele, necessariamente com alguns sistemas de escolher os ataques e os atributos dos personagens
 - a. divinity original sin 2 definitive edition, monster hunter world, terraria, slay the spire, the elder scrolls v skyrim special edition;
6. **Survival:** jogadores precisam gerenciar muito bem os recursos, com alguns sistemas sendo, peso dos itens, fome do personagem, com penalidades fortes ao jogador morrer, alguns casos o jogador perde tudo que ele estava carregando, em outros o personagem é perdido e precisa iniciar um novo
 - a. ark survival evolved, rust, astroneer, don't starve together, stardew valley, rimworld, subnautica;
7. **Competitive:** necessitam de rede para serem jogados, e os jogadores competem um contra o outro seja sozinho e/ou em grupos
 - a. rocket league, dead by daylight, insurgency sandstorm, playerunknown's battlegrounds;

4.1 - Análise de Tópicos:

Para realizar a análise de tópicos foi utilizado o BERTopic², utilizando os seguintes algoritmos para cada camada:

- **Embeddings:** all-MiniLM-L6-v2 (biblioteca SentenceTransformer);
- **Redução de dimensionalidade:** UMAP, vizinhos = 15, componentes = 5, distância mínima = 0, métrica = cosseno;
- **Clusterização:** HDBSCAN, tamanho mínimo = 15, métrica = euclidiana, método de seleção de cluster = excesso de massa (eom), dados de predição = verdadeiro;

² <https://maartengr.github.io/BERTopic/>

- **Tokenizador:** CountVectorizer (biblioteca scikit-learn), com *stop-words* em inglês;
- **Representações dos tópicos:** c-TF-IDF;
- **Ajuste fino da representação:** KeyBERTInspired;

Com o modelo preparado, foi realizada a análise de tópicos para cada jogo separadamente, garantindo assim que a desproporção entre quantidade de *reviews* entre eles não afetasse os tópicos gerados. O próximo passo foi a modelagem em uma estrutura de grafos para encontrar comunidades de tópicos similares entre diferentes jogos, e para tal, cada tópico de cada jogo foi representado por um nó, e foram adicionadas arestas entre os nós com o atributo peso sendo o coeficiente de similaridade entre os dois nós respectivos.

Para realizar o cálculo da similaridade entre os tópicos, foi feito uma combinação para cada par de jogos único, e foi gerado a similaridade de todos os tópicos deste dois jogos, utilizando uma matriz de similaridade a partir dos *embeddings* dos tópicos (fornecidos pelo modelo do BERTopic), com a métrica sendo a similaridade do cosseno. Feita a matriz, para cada tópico do jogo A, foi selecionado até 5 tópicos do jogo B com as maiores similaridades (em caso de jogos com poucos tópicos, a quantidade de tópicos selecionados é ajustada de acordo com a quantidade disponível). Isto resultou em uma lista contendo itens no formato (jogoA_topicoX, jogoB_topicoY, similaridade), os quais foram utilizados para adicionar as arestas no grafo. O grafo resultante possui 3.087 nós e 106.250 arestas.

O próximo passo foi encontrar comunidades de tópicos similares e gerar uma visualização das mesmas para poder analisá-las, e para tanto os seguintes procedimentos foram adotados:

1. Para cada um dos 7 grupos de jogos, foi gerado um subgrafo que contém somente tópicos referentes aos jogos respectivos;
2. Removidas as arestas que o atributo peso fosse menor que 0.9 (esse valor foi definido após análise dos resultados, testando diferentes valores para diferentes grafos, e esse valor foi unânime na qualidade dos resultados obtidos dos diferentes agrupamentos);
3. Executado o algoritmo de detecção de comunidades Louvain (levando o peso das arestas em consideração);
4. Removidas as comunidades com apenas um tópico;
5. Para cada comunidade, é gerado uma nuvem de palavras a partir das palavras dos tópicos daquela comunidade e suas respectivas relevâncias.

Com isso foi possível analisar os tópicos similares entre diferentes agrupamentos de jogos e gerar uma visualização dos mesmos.

4.2 - regressões e correlações

Para realizar a análise sobre a relação e influência foi utilizado regressão e correlação entre *features*:

- **hour_played (análise individual):** quantidade de horas jogadas pelo jogador no momento e que ele escreveu a análise;

- **hour_played (análise agrupada)**: quantidade de horas mediana que os jogadores jogaram o jogo;
- **recommend**: binário se o jogador recomenda ou não o jogo;
- **polaridade**: valor que indica o sentimento da análise, de -1 (negativo), até 1 (positivo);
- **percent_recommend**: indica a porcentagem de jogadores que recomendam o jogo.

Algoritmos utilizados:

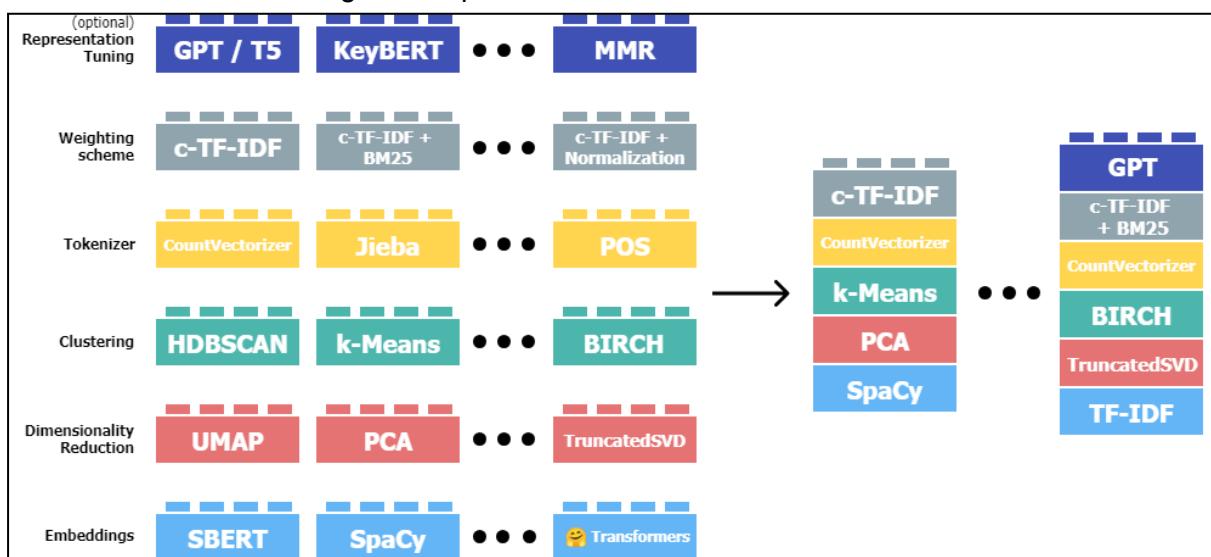
- **Regressão logística**: logit da biblioteca statsmodels;
- **Regressão linear**: ols da biblioteca statsmodels;
- **Correlação**: pandas para cálculo e seaborn para visualização;
- **Análise de sentimentos**: biblioteca textblob com o algoritmo NaiveBayesClassifier;

A questão 2, foi respondida usando os dados que foram filtrados e corrigidos sem qualquer modificação neles, e assim foi aplicada regressão usando hour_played como variável resposta, e calculamos a matriz de correlação. Para a questão 3, foi necessário agrupar os jogos:

- Para polaridade agrupamos utilizando a média;
- Para hour_played agrupamos utilizando a mediana (como não tem limite máximo existem alguns jogadores extremos que jogaram muito puxando a média para cima);
- Para percent_recommend foi agrupado a quantidade que recomenda e calculado a porcentagem que representa do total.

5 - Trabalhos correlatos e Fundamentos:

Para realizar a análise de tópicos foi utilizado o BERTopic, que é uma técnica de modelagem de tópicos baseado em Transformers e c-TF-IDF (*Contextual Term Frequency-Inverse Document Frequency*), que permite criar tópicos com boa interpretabilidade e manter as palavras mais relevantes de cada tópico. A estrutura do BERTopic é baseada em 6 camadas, cada uma com uma função específica, permitindo assim utilizar diferentes algoritmos para cada camada.



Exemplo da estrutura em camadas do BERTopic.

Com o intuito de calcular a similaridade entre diferentes tópicos, foi utilizado a similaridade do cosseno entre os *embeddings* dos tópicos. Os *embeddings* são representações vetoriais de palavras que capturam informações semânticas e relacionamentos entre as palavras, as quais são amplamente utilizadas em várias tarefas PLN. As técnicas para construí-las são projetadas para mapear palavras em um espaço vetorial, onde a proximidade entre os vetores reflete a similaridade semântica entre as palavras correspondentes. Ou seja, palavras que são usadas em contextos semelhantes terão representações vetoriais mais próximas.

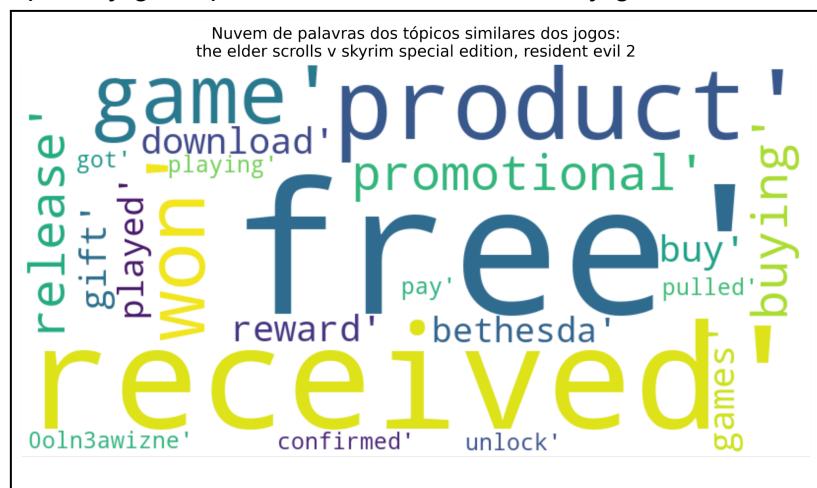
6 - Resultados:

Com o propósito de investigar se jogos com características similares possuem intersecção de tópicos abordados, analisamos as nuvens de palavras geradas a partir das diferentes comunidades encontradas em cada grupo. Cada nuvem de palavra contém no título quais são os jogos referentes à nuvem gerada, pois apesar de o grafo de cada grupo conter todos os jogos daquele grupo, as comunidades não necessariamente possuem tópicos de todos os jogos.

Outro ponto é que muitas dessas nuvens, mesmo que em grupos diferentes, retratam tópicos muito similares, como a qualidade do jogo em si e a diversão, utilizando palavras como jogar, bom, legal, melhor, divertido, etc. Portanto será relatado somente as nuvens que trouxeram aspectos diferentes desses, e por consequência, mais característico dos grupos em questão.

6.1 - Singleplayer:

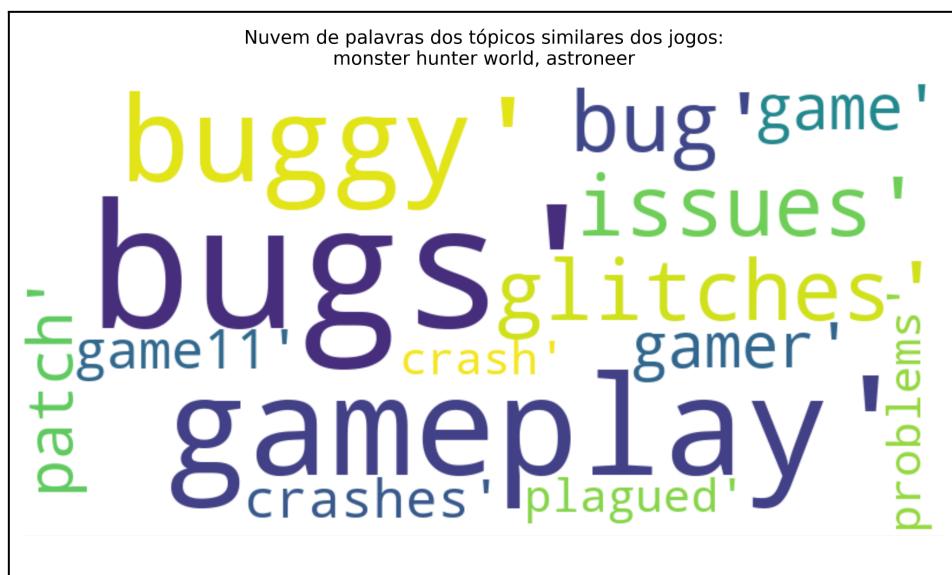
Neste grupo foram geradas 4 comunidades, das quais duas são interessantes: a primeira retrata o aspecto de ter recebido algo de graça, seja o jogo como presente, ou algum conteúdo dentro do jogo, demonstrando que possivelmente esteja jogando e comentando sobre o jogo por tal motivo. A segunda retrata aspectos inerentes aos jogos deste grupo, como a qualidade da *playthrough* (a jornada do início ao fim do jogo), ser viciante e ‘sem fim’: aspecto de poder jogar repetidas vezes e/ou continuar jogando.

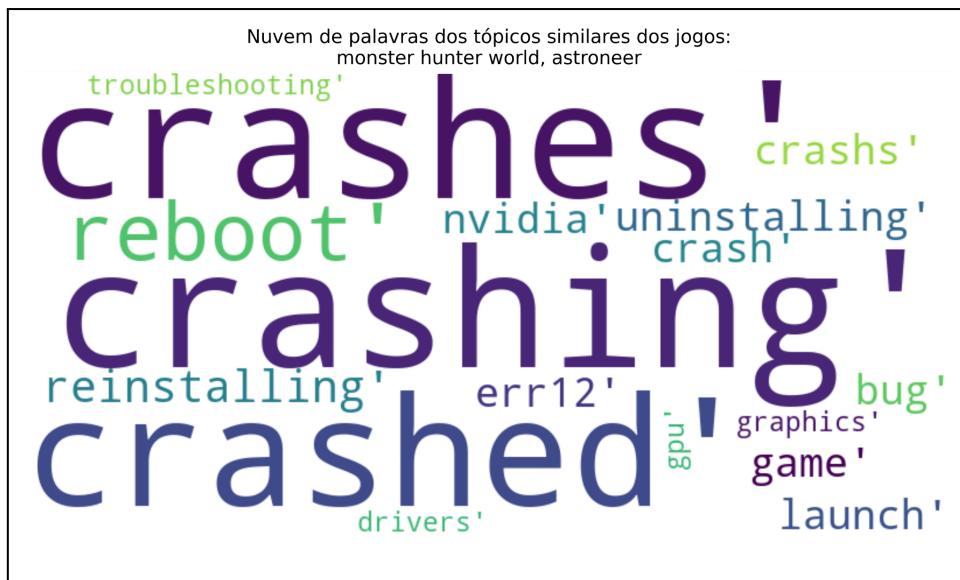




6.2 - Coop:

Foram geradas 13 comunidades, onde várias apresentam pequenos pontos porém relevantes, como o prazer em poder jogar com amigos/comunidade, ser um desafio passar pelas dificuldades dos jogos com amigos, ressaltando como é bom poder jogar cooperativamente (pois muito desses jogos também podem ser jogados sozinho) e vemos novamente uma nuvem similar ao 6.1 retratando receber conteúdo gratuitamente. Entretanto 4 se destacam: a primeira e a segunda abordam problemas encontrados com o jogo, como falhas, problemas, *bugs* e *crashing* (jogo fechar inesperadamente). A terceira relata a questão de *farm* e *grind*, que é um aspecto onde os jogadores precisam gastar muito tempo para conquistar certos recursos/equipamentos dentro dos jogos, e a quarta traz palavras como construção, exploração, e Minecraft, que é um jogo muito popular com esses recursos mas que não está no grupo, ou seja, fazendo uma comparação com este famoso jogo.

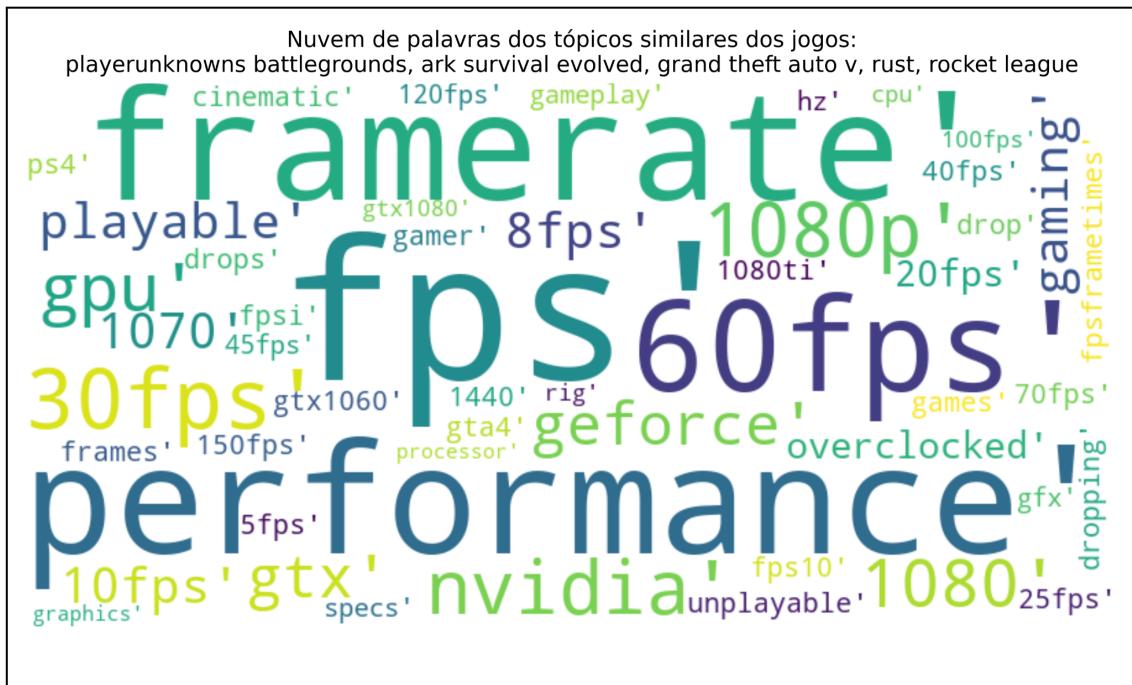


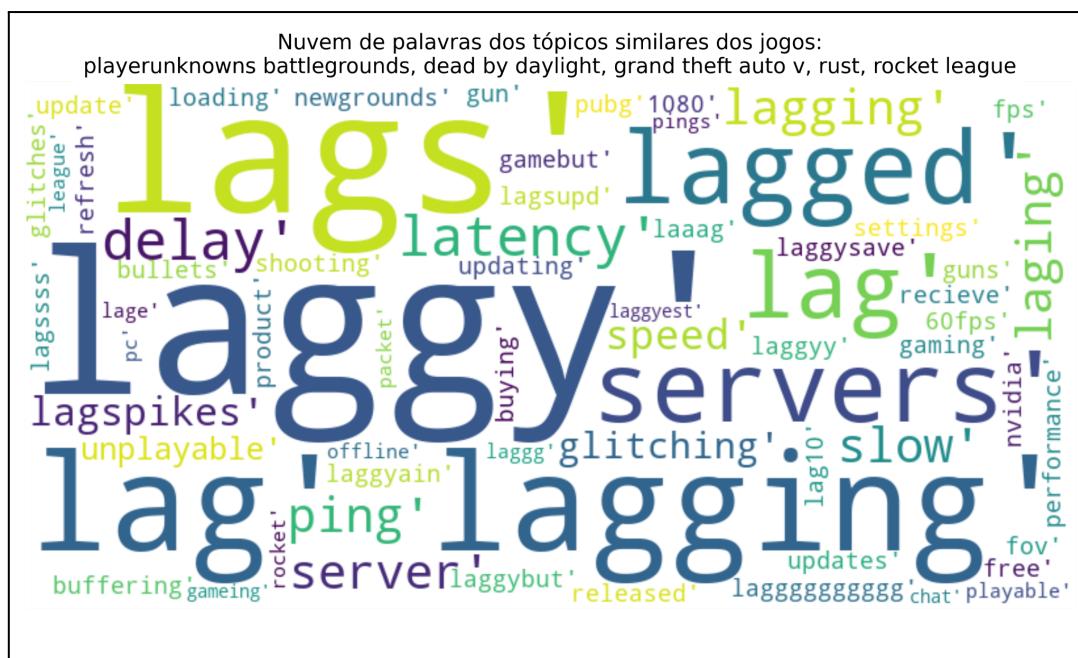


6.3 - Multiplayer:

Este é o grupo com mais comunidades (170), devido a popularidade/quantidade de comentários dos jogos. Com isso, foi possível encontrar diversos tópicos relevantes e coerentes com o grupo, abordando questões de performance (como *Frames Per Second*,

placas de vídeo populares, RAM, espaço em disco), questão da toxicidade da comunidade e banimentos, algo muito comum em jogos online onde pessoas que não se conhecem interagem entre si, problemas com o servidor (como dificuldade de encontrar partida, partidas desbalanceadas, *lag*, desconexões), abordando comunicação de voz, questões de jogabilidade multiplataforma (poder jogar com pessoas de diferentes consoles e/ou computador), microtransações e *pay-to-win* (compras dentro do jogo, vantagens para quem gasta dinheiro), tópicos de raiva e fúria (dada a competitividade muitas vezes existente e toxicidade), problema de *hackers* e *cheaters* (trapaceiros), e também tópicos *troll*, mais uma vez enfatizando a toxicidade da comunidade, retratando ódio, preconceito generalizado, drogas e afins. É observado também tópicos falando sobre *e-sports* (esportes eletrônicos), bem como a dificuldade dos jogos e suas curvas de aprendizado. Alguns exemplos são mostrados a seguir:





6.4 - FPS:

Esse grupo conta com 100 comunidades, porém por conter uma intersecção com multiplayer, muitos tópicos se repetem. O que mais se destaca aqui é tópicos sobre jogadores trapaceiros, microtransações e comunidade tóxica.

6.5 - RPG:

Entre as 7 comunidades geradas, nenhuma destaca algo exclusivo deste grupo, abordam questões de performance, qualidade percebida e problemas gerais.

6.6 - Survival:

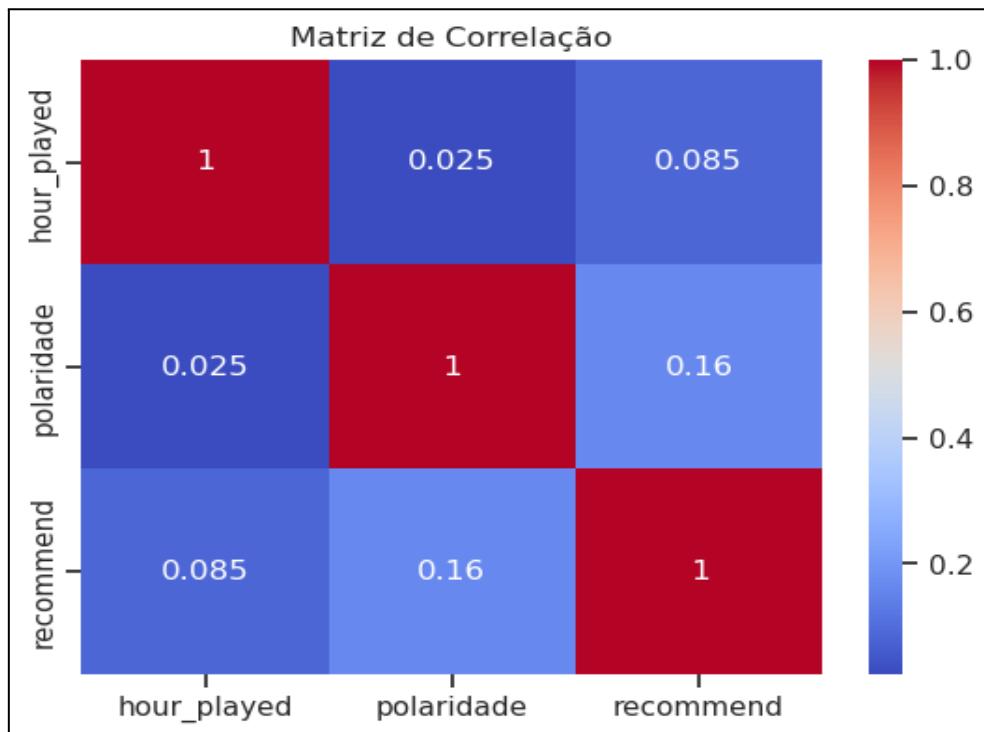
Criadas 12 comunidades, porém similar ao anterior, não houve nada muito específico deste grupo, foi muito similar às descobertas feitas no grupo Coop, mas pode se argumentar que os tópicos vistos novamente aqui sobre construção, exploração, e Minecraft fazem mais jus a este grupo, visto que é algo comum aos jogos de sobrevivência.

6.7 - Competitive:

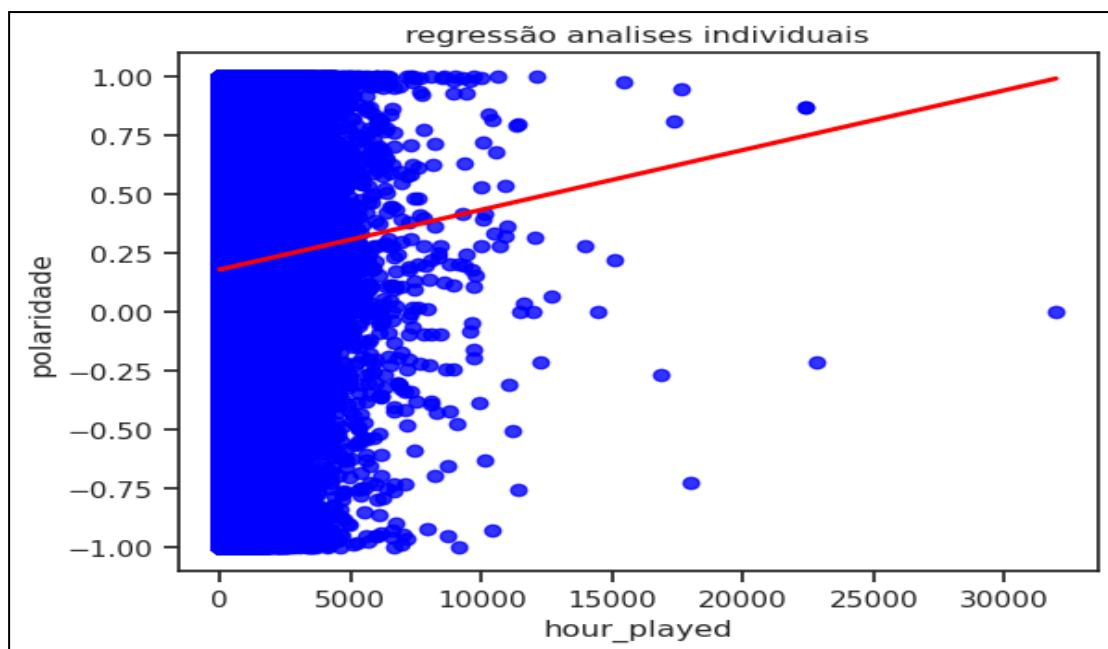
Com 120 tópicos, vemos muita similaridade com o grupo Multiplayer, dada a intersecção dos mesmos. Contudo o que é salientado neste grupo vai de acordo com o que normalmente é visto em cenários competitivos de jogos onlines, como a qualidade do *matchmaking* (construção da partida), toxicidade, problemas de servidor, viciante e raiva. Curiosamente o problema de jogadores trapaceiros não aparece aqui, mostrando que em jogos competitivos as empresas possivelmente tomam mais cuidados para evitar tal situação. É visto microtransações, mas não *pay-to-win*, mostrando que existem maneiras de capitalizar, mas sem fornecer vantagens a jogadores, mantendo assim a qualidade de um cenário competitivo.

6.8 - Outras análises:

Para responder as perguntas 2 e 3, utilizamos de regressões e correlações, a hipótese que jogadores que recomendam o jogo tem um sentimento mais positivo em comparação com os que não recomendam foi respondida com uma matriz de correlação.



Podemos notar que tanto quantidade de horas quanto polaridade (campo que representa sentimento da análise) tem correlação positiva com recomendação.



Dep. Variable:	hour_played	R-squared:	0.001			
Model:	OLS	Adj. R-squared:	0.001			
Method:	Least Squares	F-statistic:	257.7			
Date:	Wed, 28 Jun 2023	Prob (F-statistic):	5.62e-58			
Time:	09:52:26	Log-Likelihood:	-3.0784e+06			
No. Observations:	398591	AIC:	6.157e+06			
Df Residuals:	398589	BIC:	6.157e+06			
Df Model:	1					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	358.0682	0.917	390.275	0.000	356.270	359.866
polaridade	25.4605	1.586	16.054	0.000	22.352	28.569

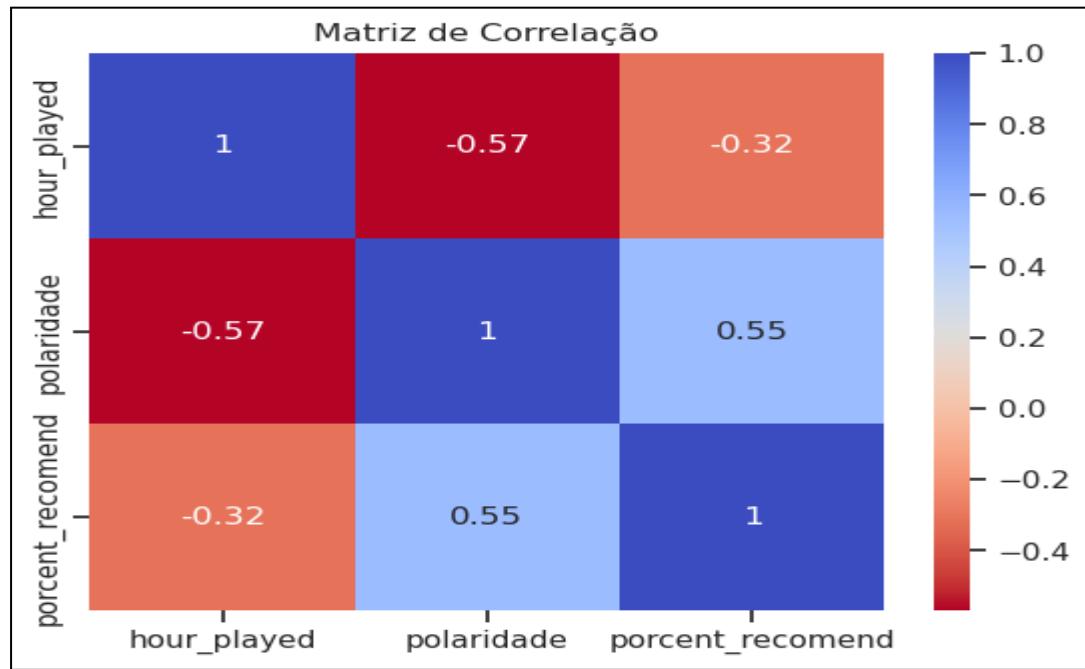
Com regressão linear simples podemos confirmar a hipótese que conforme aumenta o número de horas jogadas, o sentimento tende a ficar mais positivo. Isso pode ser justificado pelo fato de que os jogadores que não gostaram tanto do jogo tendem a abandonar o jogo e jogar cada vez menos trocando para um jogo que o agrade mais.

A hipótese de que jogadores com sentimentos mais positivos e com mais horas de jogo devem recomendar mais é respondida com regressão logística.

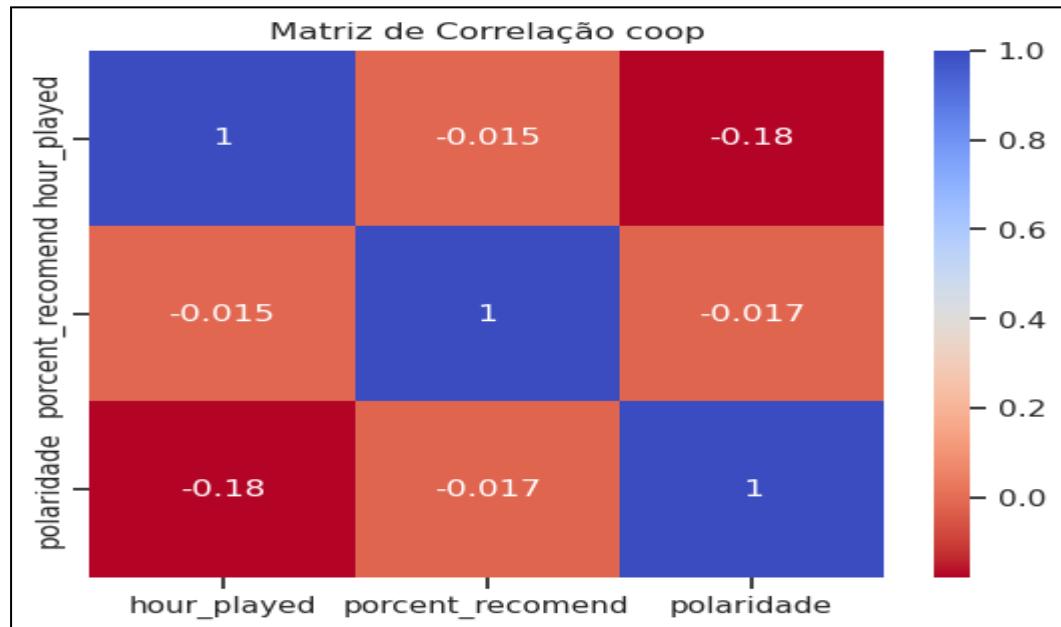
Current function value: 0.597821						
Iterations 5						
Logit Regression Results						
:						
Dep. Variable:	recommend	No. Observations:	398591			
Model:	Logit	Df Residuals:	398588			
Method:	MLE	Df Model:	2			
Date:	Wed, 28 Jun 2023	Pseudo R-squ.:	0.02840			
Time:	09:52:43	Log-Likelihood:	-2.3829e+05			
converged:	True	LL-Null:	-2.4525e+05			
Covariance Type:	nonrobust	LLR p-value:	0.000			
	coef	std err	z	P> z	[0.025	0.975]
Intercept	0.5759	0.004	129.983	0.000	0.567	0.585
hour_played	0.0004	8.36e-06	52.101	0.000	0.000	0.000
polaridade	0.6482	0.006	101.007	0.000	0.636	0.661

O resultado indica que para cada hora jogada aumenta em 0.0004% a chance da pessoa recomendar o jogo, e que a análise ter sentimentos positivos aumenta também a chance da recomendação dos jogos.

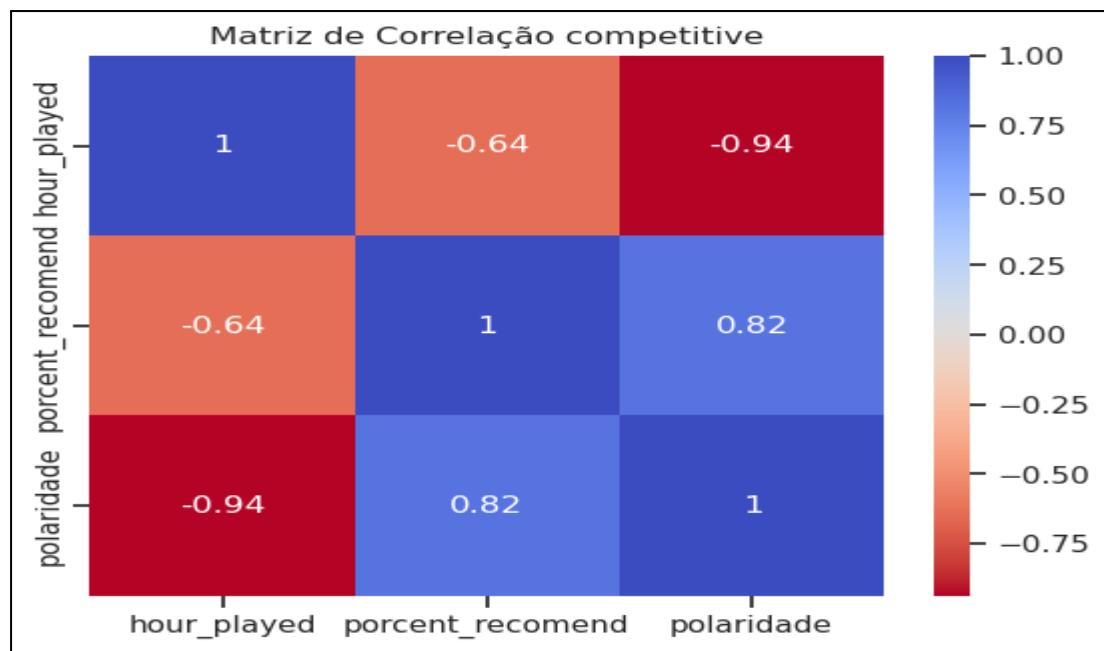
Para os resultados dos jogos agrupados, realizamos a correlação analisando todos os jogos juntos:



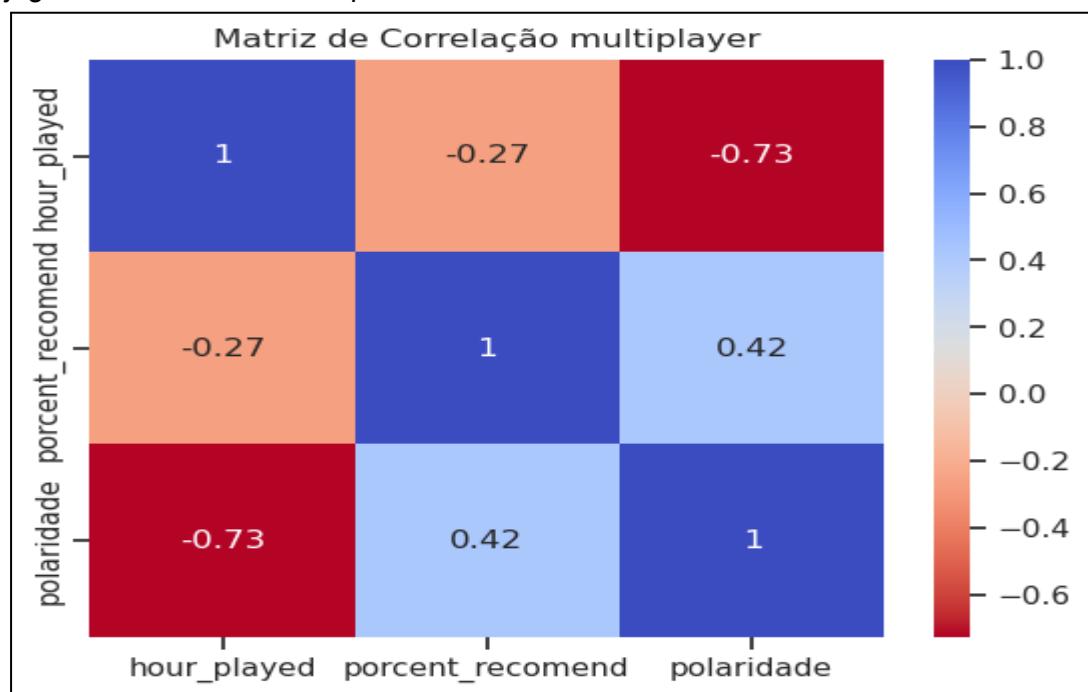
Onde percebemos que sentimentos positivos e recomendação tem correlação negativa com horas de jogo, algo que não era esperado. Separando os jogos por grupos podemos ver as situações que isso ocorre:



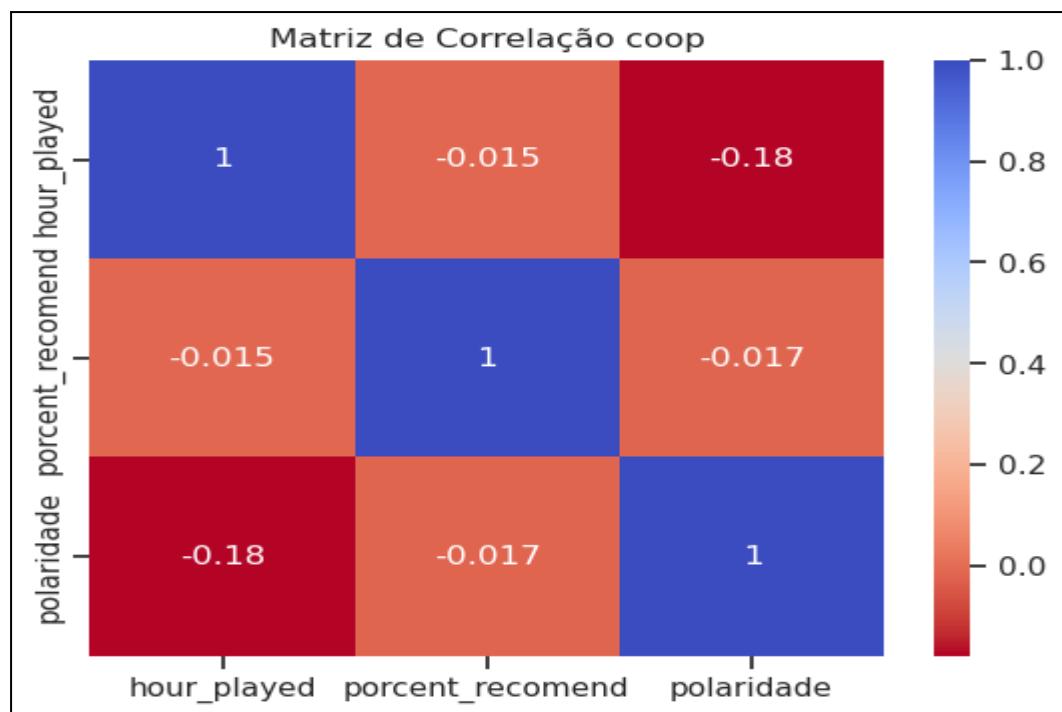
Jogos cooperativos tendem a ser mais repetitivos, por isso muitas horas não aumenta positivamente o sentimento, mas possui correlação baixa pois não chega a ser frustrante com o passar do tempo.



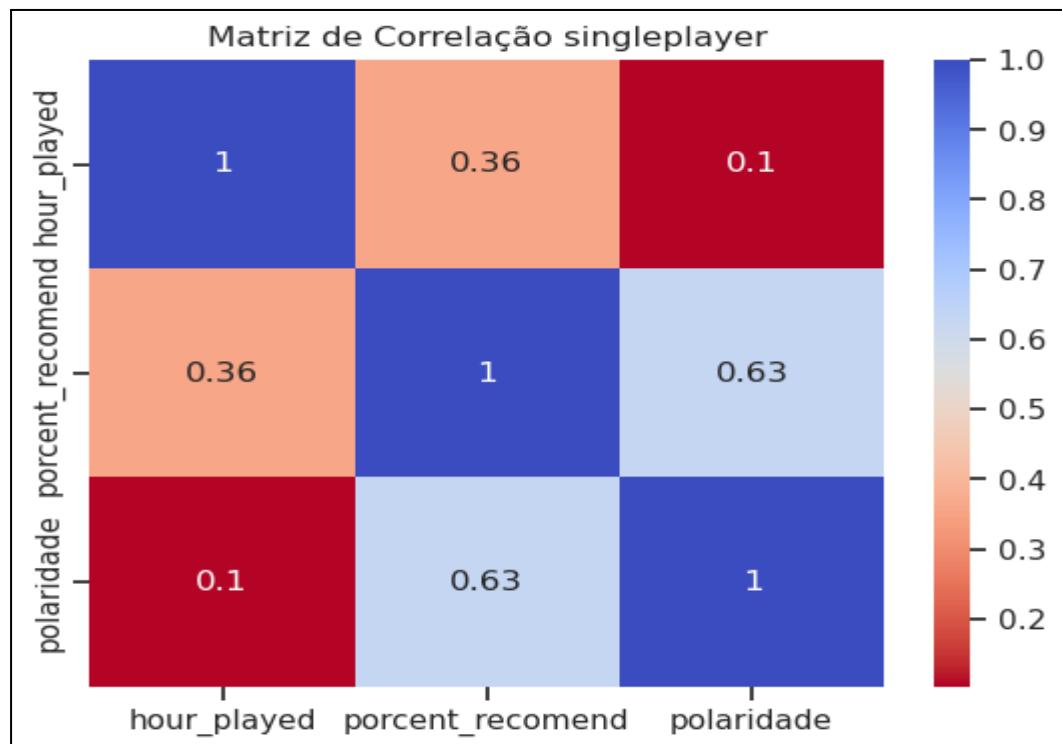
Jogos competitivos tendem a atrair jogadores com reações mais fortes, e as partidas tendem a ser ou frustrantes ou muito recompensadoras. Esse resultado era esperado, quanto mais a pessoa joga, menos ela vai recomendar o jogo, mas também quanto mais um jogo é recomendado mais positivo é o sentimento.



Jogos multiplayer tem características similares aos competitivos, geralmente atraindo até uma boa parte do mesmo grupo de jogadores para os jogos que temos os dados. Por terem um nível de competição menor, a correlação das horas jogadas fica um pouco mais positivamente correlacionada.



Jogos cooperativos, possui algumas similaridades com os multiplayer mas como há bem pouca competitividade, aumenta positivamente em relação às categorias anteriores.



Singleplayer é a única categoria onde todas as features tem correlação positiva com as outras, é a categoria de jogos com praticamente nenhum embate com outros jogadores, e atrai o público menos competitivo.

7 - Limitações e Trabalhos futuros:

7-1 Limitações

- **Quantidade de jogos:** a separação dos jogos por categorias poderia ter ficado melhor com mais jogos de mais gêneros, muitos jogos podem pertencer a diferentes gêneros, e por temos poucos jogos utilizamos gêneros bem genéricos, alguns mais específicos seria uma melhoria;
- **Interpretabilidade dos dados:** os resultados referentes aos tópicos é difícil de analisar, precisamos supor várias coisas para achar uma explicação;
- **Linguagem específica:** a comunidade de jogos tem uma maneira de se comunicar muito específica, os modelos pré-treinados não consegue captar uma parte do que é comunicado, tem muitas gírias e jargões que só jogadores usam, inclusive tem termos comuns em um gênero que em outro pode ter um significado diferente;

7-2 Trabalhos futuros

- **Seleção de tópicos não supervisionada:** os tópicos relevantes foram selecionados manualmente, precisou-se verificar cada um individualmente, seria interessante achar um algoritmo que facilite esse trabalho;
- **Diferentes maneiras de modelar o problema:** tínhamos uma ideia clara passo a passo do que fazer até uma parte do trabalho, porém após isso tínhamos muitas opção para seguir, e não com muita certeza que seria o resultado ideal pela falta de experiência. Seguimos um caminho com mais complexidade, todavia existem maneiras mais práticas e elegantes de resolver o problema, como projetar os reviews em si em um espaço 2D e agrupá-los com algoritmos como o K-means.

8 - Apêndices:

A seguir estão todas as nuvens de palavras de maneira reduzida.

8.1 - Singleplayer:



8.2 - Coop:



8.3 - Multiplayer:









8.4 - FPS:



8.5 - RPG:



8.6 - Survival:



8.7 - Competitive:

