

Análise Exploratória

Código ▾

AUTOR

Mikael M Coletto

DATA DE PUBLICAÇÃO

5 de janeiro de 2023

Análise dos dados da Steam

Introdução

A ideia desta análise é ilustrar os conhecimentos adquiridos em R, começando com uma análise exploratória dos dados e tentando trazer alguns insights relacionados à jogos de sucesso no mercado. Para isso vou passar pelas etapas de elaboração do objetivo, trabalho no tratamento e limpeza dos dados, fazer análises de variáveis de interesse individualmente e depois cruzando algumas delas para tentar traçar algumas conclusões, que, por fim, podem sugerir alguns caminhos para empresas que pretendem lançar seus jogos terem sucesso na Steam.

O objetivo

Somos um estúdio de jogos e queremos analisar:

1. Qual gênero/categoria de jogo tem mais sucesso na steam, definindo sucesso como melhores avaliações;
2. Também queremos analisar qual faixa de preço é mais interessante para o lançamento desse novo jogo, como a faixa de preço se relaciona com o número de vendas e com a categoria.

Análise exploratória das bases de dados

Inicialmente irei fazer uma análise exploratória de algumas bases de dados e suas variáveis que podem ser interessantes para determinar sucesso de um jogo, a faixa de preço, a quantidade de vendas e relacionar essas questões aos gêneros e categorias. Foram feitas análises de três bases de dados, a primeira mais completa, e as duas seguintes mais focadas em reviews de usuários. São bases de dados extraídas do Kaggle que serão utilizadas para o estudo.

Base de dados - 1

Esta base de dados se encontra no Kaggle, através deste link:

<https://www.kaggle.com/datasets/fronkongames/steam-games-dataset>

Skimr

Usarei o Skimr para a base de dados completa e verificarei quais dados podem ser melhor investigados. Essa função faz um resumo geral das variáveis e é bem útil para uma primeira visualização.

► Código

► Código

Por sua complexidade e tamanho de análises, para a versão final deste relatório foi optada por não apresentar a saída da função `skimr`.

Limpezas e padronizações

Para padronizar os nomes das variáveis, será utilizado o pacote `janitor`, que é bastante útil na hora de limpeza e padronização do banco de dados.

► Código

Removendo Softwares e não jogos

Após a análise inicial, foi detectado que existiam nesta base de dados softwares que não eram considerados jogos, portanto eles serão excluídos para que não comprometam as futuras análises.

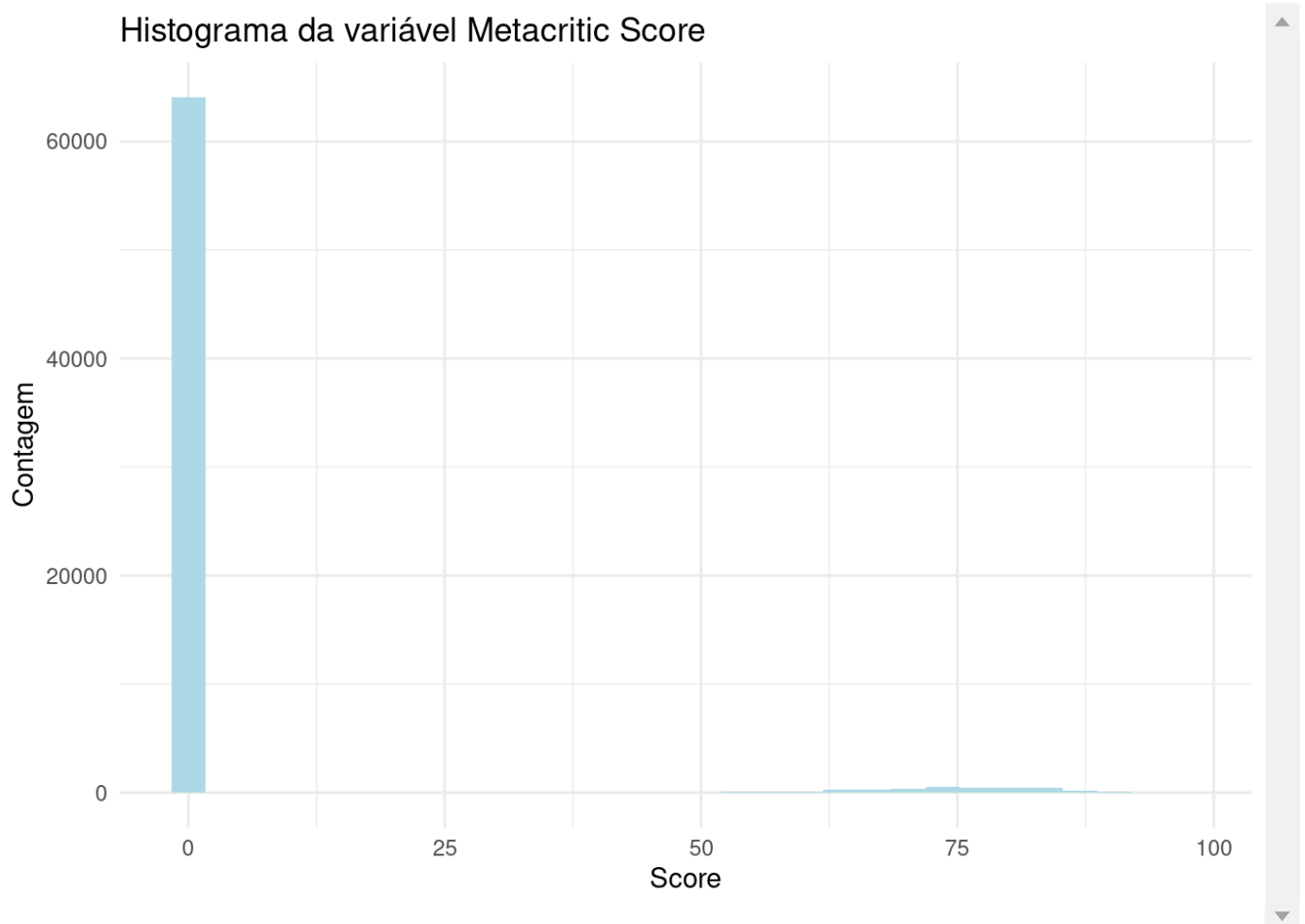
Vamos utilizar a variável `genres` e `tags` para identificar os gêneros que não são jogos.

► Código

Escolhendo as variáveis

E então, analisando as tabelas geradas pelo `skimr`, selecionei algumas variáveis que seriam analisadas mais profundamente. A começar pelas variáveis relativas a notas, críticas e avaliações dos jogos.

► Código



Pelo Histograma pude ver que a variável *Metacritic Score* (renomeada para *metacritic_score*) não é muito interessante, seu preenchimento foi visto como completo pelo skim, porém quase em sua totalidade as notas são "0", ou seja, não temos uma nota definida para avaliar os jogos.

Outras variáveis de avaliação como *Reviews*, *User Score*, e *Recommendations*, foram analisadas porém apresentaram pouco úteis, ou por terem um preenchimento muito baixo ou por envolverem outros tipos de dados (que não fossem uma avaliação direta).

Nova métrica de avaliações

As únicas métricas de avaliação nesta base que parecem interessantes, são as de avaliações positivas e negativas (as variáveis *Positive* e *Negative*).

Por isso, decidi criar uma nova variável, que mede a taxa de avaliações positivas e negativas de um jogo, chamada *overall_rate*.

► Código

```
[1] 35.29412 91.37931      NaN 100.00000 86.20690 63.97059 75.00000
[8]      NaN 92.68293 83.33333 73.53308 71.70732 100.00000 76.59574
[15] 95.23810 72.25673 100.00000      NaN 75.00000 67.77778 50.87719
[22] 78.50369 28.57143 87.30512 62.71186 26.47059 90.47619 82.75862
[29]      NaN 56.09756 66.66667 61.53846 60.84906 100.00000 77.77778
[36] 66.66667 70.54291 81.15942 37.67123 73.20819      NaN 77.46948
[43] 62.16216 50.00000 79.91676 40.00000 56.60377 87.50000 100.00000
[50]      NaN      NaN 50.00000 85.71429 80.23256 89.03251 10.52632
[57] 60.00000 100.00000 0.00000 75.46667 77.77778 56.45439 100.00000
[64] 89.52023 100.00000 100.00000 86.95652      NaN 96.96845      NaN
[71] 89.47368 100.00000 77.77778 40.00000      NaN 70.45455      NaN
[78] 100.00000 27.77778 0.00000 83.33333 83.33333 100.00000 33.33333
[85] 90.32258 76.11940 67.86885 100.00000      NaN 83.01887 89.70434
[92] 76.08696      NaN 96.49192 50.00000 100.00000 77.77778 100.00000
[99] 75.92593 91.13924
```

Analisando uma parte da base de dados com essa nova métrica, verifiquei que a tabela possuía dados *NaN* e valores de 100, respectivamente, resultados de divisões 0/0 e de jogos sem avaliação negativa. Os números *NaN* serão removidos, já que não serão úteis para avaliação dos jogos, os números 100 serão investigados mais a fundo.

Filtrando dados inválidos

Após as remoções anteriores, também achei pertinente fazer mais algumas filtrações na base de dados.

► Código

Os filtros foram definidos para que tenhamos uma base mais confiável. As regras para que um jogo se mantivesse na base são:

1. O jogo precisa ter pelo menos 20 avaliações positivas.
2. O jogo precisa ter pelo menos uma avaliação negativa (para que a avaliação não seja 100% positiva).

3. O jogo precisa ter pelo menos 50 mil usuários no caso de ser gratuito, e pelo menos 20 mil usuários para jogos pagos.

Deduplicação

Depois da análise do novo banco de dados filtrado, descobri mais um problema que precisava ser tratado, dados duplicados. Verificando pelo nome dos jogos, encontrei jogos duplicados, mas que possuíam IDs, avaliações e às vezes preços diferentes. Olhando com mais cuidado pelo próprio site da Steam e pelo site SteamDB ([SteamDB](#)), concluí que se tratavam de dados referentes à novas versões (jogo do ano, versão de luxo, etc), jogos que possuíam versões diferentes de modo Singleplayer e Multiplayer (como por exemplo nos jogos da série "Call of Duty") e mini-expansões, pacotes de skins (aparências para armas e personagens), pacotes de itens pagos como benefícios no jogos, etc. Portanto, decidi remover essas duplicatas, e usarei como base dados de número de usuários (o jogo base muito provavelmente vai ter mais usuários) e número de jogadores.

► Código

Então esta nova regra criada respeitará primeiro o número de jogadores, o jogo que possuir mais jogadores permanecerá na base, e segundo, em caso de empate no número de jogadores, será mantido o jogo com melhor avaliação.

Finalizando limpeza e transformação

E então, finalizando as filtragens, limpezas e transformações, tenho por fim uma base mais confiável para as análises que virão.

Base de dados - 2

Fonte

Esta base de dados se encontra no Kaggle, através deste link:

<https://www.kaggle.com/datasets/souyama/steam-dataset>

Visão inicial

Como os dados de avaliações na primeira base eram escassos e eu já havia encontrado outras bases de dados nas buscas que fiz, decidi analisar outras bases e verificar se alguma outra teria informações mais úteis com propósito de verificar dados de avaliação dos jogos.

► Código

```
[1] 1 -1
```

► Código

Porém, investigando a base, notei que ela era um pouco antiga, por não possuir títulos como o "Counter-Strike: Global Offensive", os jogos mais novos da série "Battlefield", entre outros, o que seria um problema para traçarmos conclusões no mercado atual.

Excluindo a base

E então, após contatar os problemas citados, decidi não utilizar esta base para a análise.

Base de dados - 3

Esta base de dados se encontra no Kaggle, através deste link:

<https://www.kaggle.com/datasets/antonkozyriev/game-recommendations-on-steam>

Após a exclusão da segunda base de dados, decidi verificar a última que havia encontrado de jogos da Steam.

► Código

Na análise inicial desta nova bases, concluí que ela possuía dados atualizados dos jogos, mas possui menos dados que a primeira base analisada. Não temos informações de línguas suportadas (tanto legendas quanto áudios), dados de pico de usuários, dados de DLC, dados de empresas (como sites, e-mail de suporte, imagem do jogo), dados do Metacritic e dados de tempo de jogo. Então utilizarei ela apenas para verificar a avaliação dos usuários (através das colunas `positive_ratio` e `user_reviews`).

Filtrando dados inválidos

Aplicando novamente os filtros utilizados para as bases anteriores, mas adaptando para as colunas desta nova base. Agora, utilizarei a variável `positive_ratio` para filtrar as avaliações positivas e `user_reviews` para o número de reviews. Os primeiros filtros definidos foram:

1. O jogo precisa ter pelo menos 20 avaliações positivas.
2. O jogo precisa ter pelo menos uma avaliação negativa (para que a avaliação não seja de 100%).
3. O jogo precisa ter pelo menos 50 mil usuários no caso de ser gratuito, e pelo menos 20 mil usuários para jogos pagos.

► Código

E para o terceiro filtro, a ideia seria fazer uma contagem através da tabela de reviews para verificar a quantidade do número de usuários por jogo. Que é a única tabela que possui os *IDs* de jogos e jogadores numa única tabela.

► Código

Porém, ao contar quantos jogos distintos haviam na tabela de reviews, tivemos o número de apenas 1900, ou seja, na base temos poucas informações sobre a quantidade de usuários segundo a tabela de reviews, portanto este terceiro filtro não será aplicado.

Deduplicação

O mesmo processo de deduplicação precisou ser feito na nova base.

► Código

Porém, ao contrário da base anterior, neste caso todos os jogos “duplicados” se tratavam de jogos diferentes, com diferentes avaliações, diferentes empresas e diferentes IDs, mas que possuíam o mesmo nome, portanto nenhum deles foi removido.

Adicionando coluna de tags através de outra base (também da base 3)

O próximo passo, após a verificação inicial da base 3, seria adicionar informações relacionadas às classificações dos jogos, nesse caso chamadas de “tags”. Esse processo envolveu ler uma segunda base de dados (ainda presente no mesmo grupo dos dados disponibilizados) que possuía as informações de metadados dos jogos. E então houve um processo de readequação, já que esta tabela em específico veio no formato json.

► Código

Removendo não jogos

Após a junção anterior, também removi as observações que se enquadravam nas categorias que definimos como não jogos.

► Código

Finalizando limpeza e transformação

E então, finalizando as filtragens, limpezas e transformações, tenho por fim a segunda base que utilizarei para as análises.

Diferenças entre base 1 e 3

Depois da seleção e limpeza das duas tabelas relacionadas a informações de jogos, decidi investigar as diferenças entre as bases de dados, para tentar entender qual seria o processo de ajuste necessário para a união delas.

Usei o `anti_join` como método para avaliar os jogos que estavam apenas em uma das bases e não na outra. Percebi que a base 1 excluiu muitos jogos devido aos filtros aplicados, principalmente os filtros relacionados a avaliações positivas e número de jogadores. Investigando os mesmos jogos na Base 3, que não possui informações do número de jogadores, cheguei à números bem mais altos de avaliação, e levando em conta o conhecimento prévio sobre esses jogos, interpretei como uma certa inconsistência nos dados da base 1.

► Código

Unindo as bases

Como comentado das discrepâncias entre as bases de dados, eu considerei algumas faltas de avaliações/jogadores como inconsistentes na base 1, e assim decidi utilizar a base 3 com a adição dos jogos não presentes nela, que estão presentes na base 1.

► Código

Também foi analisado novamente as duplicatas que foram geradas, e novamente foi decidido mantê-las, já que representavam, ou jogos que possuíam o mesmo nome, mas são jogos diferentes, ou são diferentes versões dos jogos (versão normal e versão “game of the year” ou um relançamento por exemplo).

Adicionando informações

Agora adicionaremos à base agrupada os dados referentes às informações do jogo, inicialmente adicionando através de outro arquivo presente nos arquivos da base 3.

► Código

Analizando dados

Agora que já uni as duas bases, e tenho uma quantidade mais razoável de jogos a serem analisados, precisarei definir qual será a forma de avaliar os estilos/gêneros/categorias dos jogos, por ser uma coluna presente em ambas as bases, a variável `tags` é uma ótima candidata, porém ela apresenta um número de possibilidades bastante grande e distinta, o que poderia ser prejudicial à análise. Já as outras variáveis `categories` e `genres` que são mais específicas, não estão presentes em boa parte dos dados. A análise será feita inicialmente nas três para posterior decisão de qual ou quais usar.

Categorias

► Código

Ou seja, na base de dados aproximadamente 60% das linhas não tem informação de categorias, a variável poderia ser utilizada, porém com algumas ressalvas.

Gêneros

► Código

Novamente, cerca de 60% da base não possui informações de gênero.

Categorias e Gênero

► Código

E, para efeito de comparação, podemos ver que os valores sem categorias e gênero tem o mesmo percentual (59%), o que faz bastante sentido já essas informações faltantes vem da primeira base que não possuía essas colunas, os valores que temos, os 40%, são dos jogos adicionados pela base 1.

Tags

► Código

Já no caso das tags, 100% das observações possuem alguma informação nas variáveis `tags` ou `tags_extra` (o que também faz sentido, já que unimos duas bases diferentes e uma delas possuía a coluna `tags` e a outra `tags_extra`).

Portanto, farei uma análise mais aprofundada desta variável para gerar uma única coluna que possa ser analisada.

Como uni duas bases de dados, precisamos tratar e unificar a nossa coluna que será definida como classificadora dos jogos. A base 1 possuía dados nas colunas `genres`, `categories` e `tags`, já a base 3 apenas em `tags`.

A coluna que se adequava melhor aos objetivos no caso da base 1, era a variável `tags`, já que possuía mais classificações para os jogos que a variável `genres`, normalmente possuindo tudo

que estava presente nela e mais algumas outras características. E a variável `Categories` estava mais associada ao modo de jogo e informações da Steam, como se possuía *Achievements*, conquistas nos jogos, se tinha suporte a controle pela Steam, etc.

Como alguns jogos estavam em ambas as bases de dados, por fim acabei ficando com duas colunas de classificação, e para não perder nenhuma classificação, achei interessante comparar as características e adicionar as faltantes, fazendo uma união de tudo que aparecia em ambas as variáveis.

Separação de tags

► Código

Analisando tags diferentes

Então, será criado uma nova coluna para as tags diferentes. É feita uma comparação, entre as duas colunas, e é mantido em um vetor o que temos de elementos iguais (`equal`), depois em outro vetor o que temos a mais em uma coluna (`diff1`) e no terceiro o que temos a mais na outra coluna (`diff2`).

Depois transformei novamente estes vetores em colunas para poder inspecionar como ficaram as novas colunas. E nesse momento verifiquei que por serem bases diferentes, em alguns casos existiam pequenas variações no nome da classificação, que foram ajustados para que não influenciassem na contagem final. Um exemplo destes casos era a categorização "Rogue like", que em uma base estava definida como "Rogue like" e na outra "Rogue-like", isso e mais algumas outras categorizações foram normalizadas.

► Código

Tags NA

Após a separação e tratamento de ambas os conjuntos de dados, observei um número grande de dados que não havia sido incluído em nenhuma base, nem no que possuía as tags iguais nem nos que possuíam as tags diferentes. Descobri que isso foi causado pela forma como o `R` trata `NA`, as colunas que possuem `NA` não são comparadas diretamente, então elas ficaram excluídas da junção anterior. Portanto, tratei de forma que também adicionasse estes casos de forma simples, copiando os dados de uma coluna à outra, para tratar de forma similar à que seria tratado o conjunto de dados que possuía as colunas iguais.

► Código

Visualizando tags mais frequentes

Após a organização e limpeza das tags, vou começar com uma visualização simples para identificar as tags mais utilizadas (e as menos utilizadas). Outro detalhe é que farei uma junção, agora da tabela com as tags diferentes e com as que defini como tags iguais.

► Código

tag	count
singleplayer	9632

tag	count
indie	9323
action	8605
adventure	7752
casual	5281
strategy	4669
simulation	4649
rpg	4623
multiplayer	4554
2d	4035

Organizando nova tabela

E então, para finalizar a organização das tags, nós vamos excluir as características dos jogos que vieram de ambas as bases, e vamos adicionar essa nova coluna tratada.

► Código

Criando base no formato long

Para futuras comparações, decidi criar dois tipos de bases de dados, uma no formato long, com cada linha possuindo uma única tag, e outra com todas as tags agrupadas em uma única coluna, unida por vírgula.

► Código

Investigando preços

Vou agora me debruçar sobre as faixas de preços e em seguida farei um cruzamento dos dados de preços e tags.

Utilizarei a base criada anteriormente, com a coluna de característica dos jogos ajustada (`tags`). E farei uma tabela com as informações descritivas relacionadas ao preço, para essa análise irei excluir os jogos gratuitos, para que não influencie tanto nas métricas.

► Código

metrica	valores
Mínimo	0.5
1° Quartil	5.99

metrica	valores
Mediana	10.99
3º Quartil	19.99
Máximo	109.99

► Código

Aqui já posso tirar algumas conclusões:

- As medidas de **tendência central** apontam para **faixas de preço** dos jogos pagos relativamente **baixa**, apesar de termos jogos mais caros (com o máximo sendo de \$110). Vamos analisar com mais cautela as categorias de preços dos jogos, já que será uma das bases de decisão na escolha do tipo de jogo a ser lançado.
- Também nos decis comparei a faixa de **preço do jogo lançado**, que será utilizada para as visualizações e análises, com a **faixa de preço atual**. Em geral vi uma **variação para baixo** nos decis quando comparamos o preço de lançamento e o preço atual, indicando que os **jogos tendem a cair de preço** com o tempo, o que é visto com frequência para quem acompanha os preços da Steam, **quando** os jogos começam a ficar **mais velhos**, é comum existirem **cortes de preços e promoções para** tentar **aumentar o número de vendas**.

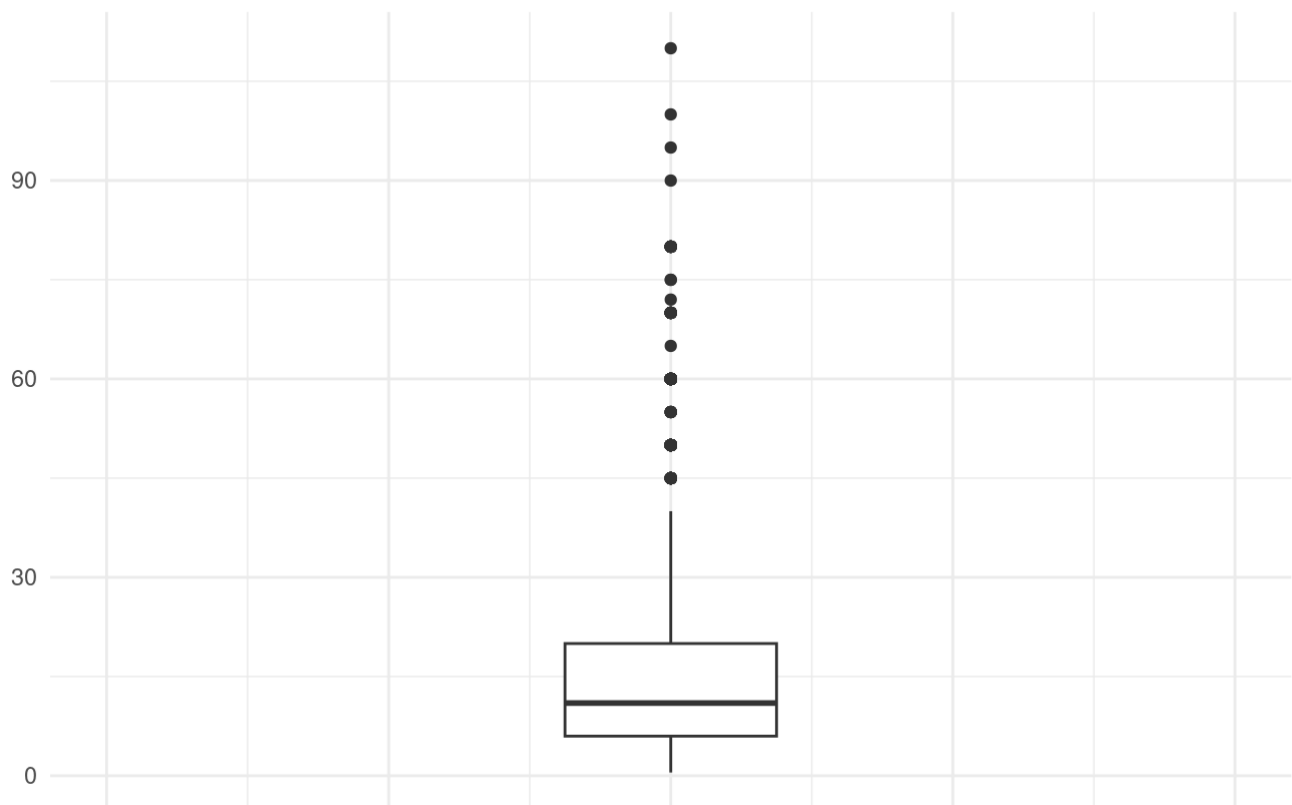
Visualizações das faixas de preço

Vou utilizar um gráfico para ilustrar as diferentes faixas de preços dos jogos. Para isto, escolhi um boxplot que apresenta algumas informações interessantes.

► Código

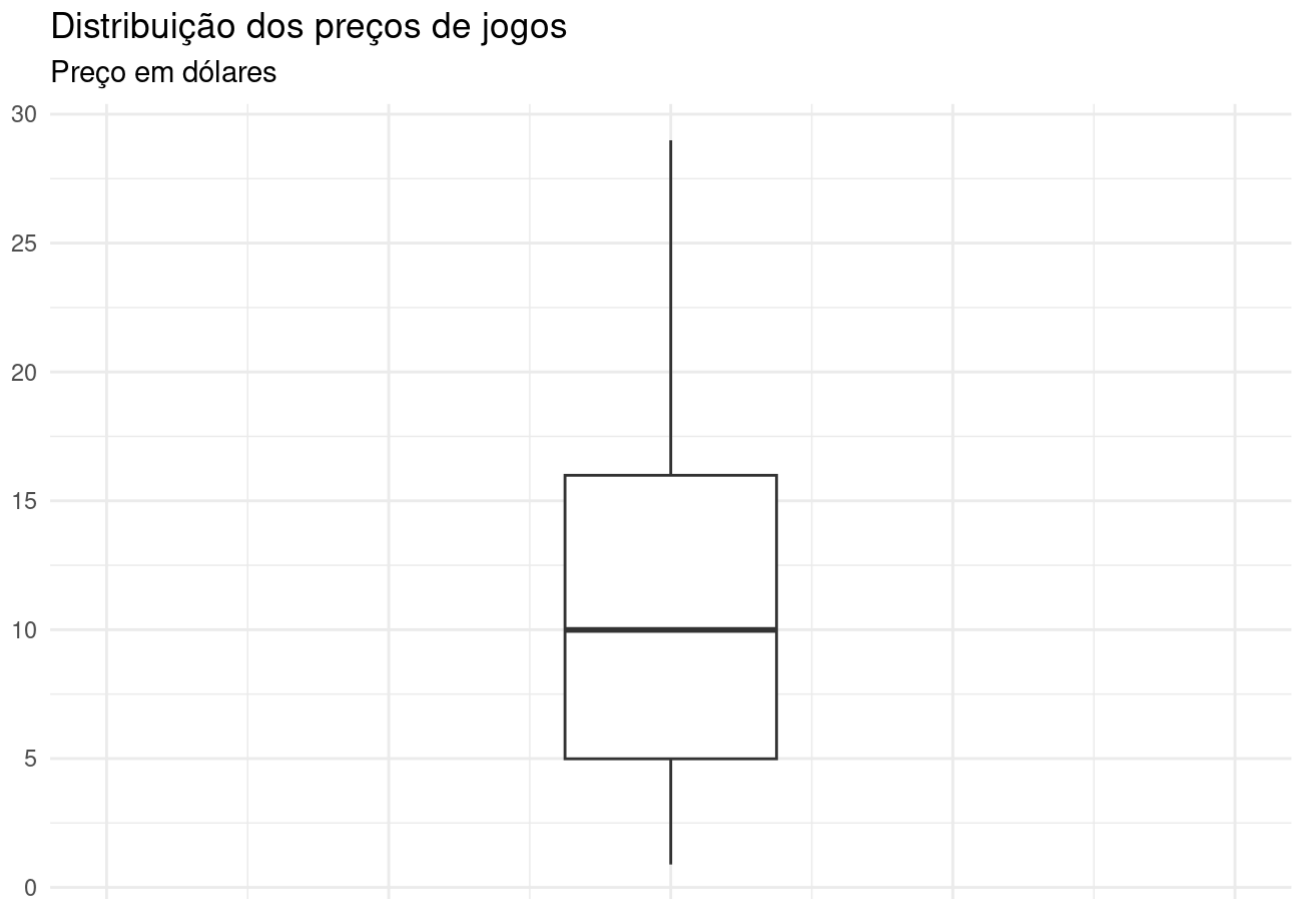
Distribuição dos preços de jogos

Preço em dólares



E aqui podemos ver que ele classificou muitos jogos como outliers, em faixas superiores à \$45, portanto resolvi tirar os jogos do primeiro e último decis, de acordo com seu preço. Depois de remover os 10% inferiores e superiores:

► Código



Observei então uma distribuição com menos valores discrepantes, tendo sua **mediana** em torno dos **\$12**, **mínimo próximo de zero** e **máximo próximo de \$30**, com **1º quartil em \$5** e **3º quartil em \$15**.

Investigando a hipótese da possibilidade de descartar estes jogos para a análise, inicialmente eu defendo a exclusão dos mais caros, pois costumeiramente são os jogos mais famosos e conhecidos, que já tem seu público definido e podem se dar ao luxo de cobrar mais caro. Já os jogos mais baratos, imagino que sejam jogos menores e que não se encaixam na nossa proposta de um jogo que fosse lucrativo, muitas vezes também são jogos mais antigos que foram tendo cortes de preços para continuarem vendendo. De qualquer forma, investigaremos nos dados ambas as hipóteses.

Investigando os 10% mais caros

► Código

Verifiquei muitos jogos considerados “AAA”, jogos de grandes editoras e publicadoras, como The Witcher 3, Cyberpunk 2077, Red Dead Redemption 2, e alguns um pouco menores mas que fizeram bastante sucesso como Rust, DayZ, Deep Rock Galactica. Com preços variando de \$29.99 até \$109.99. Citando como exemplo de jogos mais caro em seu preço original o TEKKEN 7, lançado em 2017 e com preço de lançamento de \$109.99, interessante de se notar que o preço atual (em maio/2023) do jogo é de \$16.49.

Investigando os 10% mais baratos

► Código

Já na análise dos preços mais baratos, verifiquei que os jogos que possuem mais reviews na categoria dos mais baratos, são os jogos gratuitos, porém são jogos que possuem uma estratégia diferente, e que tem crescido bastante em popularidade, os jogos gratuitos com microtransações. Jogos que costumam fazer receita através da venda de itens cosméticos dentro do próprio jogo, ou algum sistema de assinatura que garanta benefícios (cosméticos ou facilitação para o jogador).

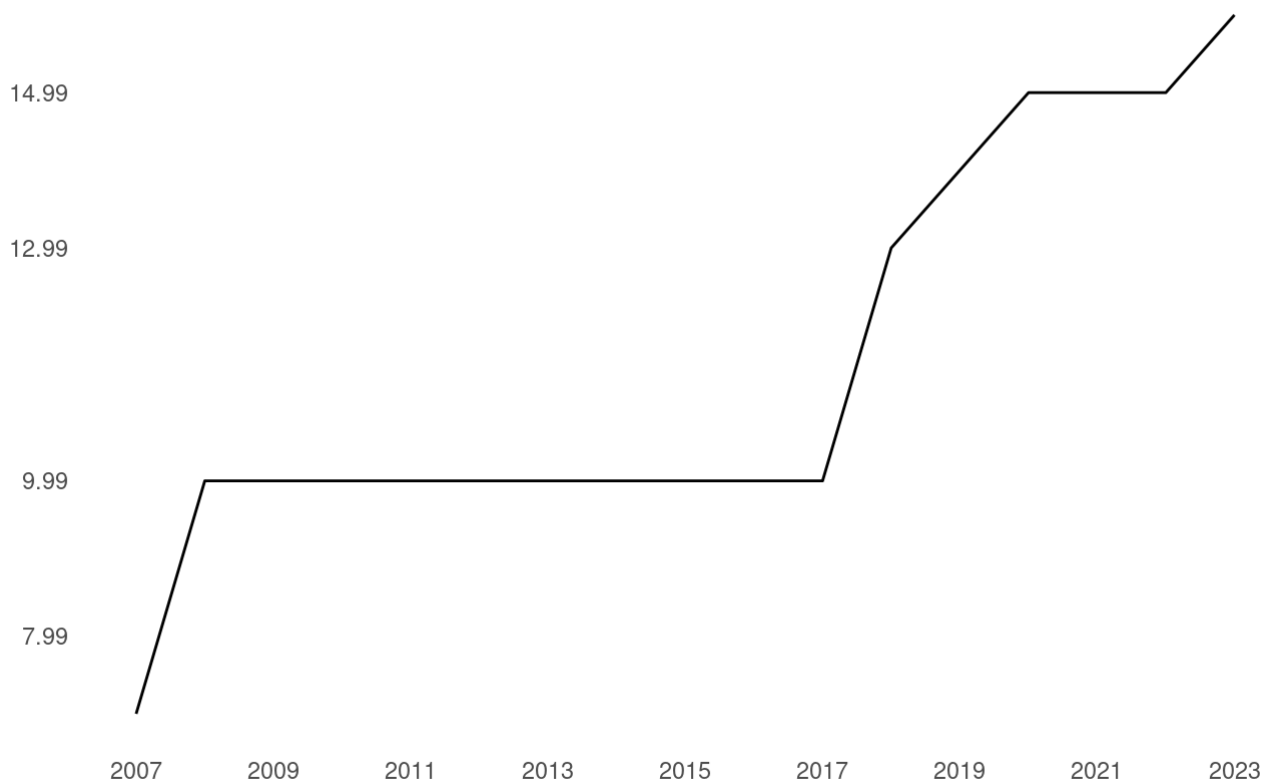
Evolução de preços

Outra coisa interessante que notei, é que os **jogos mais antigos eram lançados por preços mais baixos**, parece ter havido um **aumento de preço médio dos jogos**, o que pode explicar em partes a quantidade de jogos nas faixas mais baratas. Para isso, vamos tentar observar a variação de preço ao longo dos anos, com o intuito de entender as faixas de preços atuais e definir a melhor faixa de preço para o lançamento de um jogo.

► Código

Evolução da mediana dos preços ao longo dos anos

Preços



Apesar de já ter verificado um **aumento na mediana e na média de preços**, por ter acompanhado esse mercado por alguns anos tive a impressão de que o resultado não refletia exatamente os aumentos observados por mim, então decidi me debruçar mais sobre alguns exemplos de jogos considerados “triple A”.

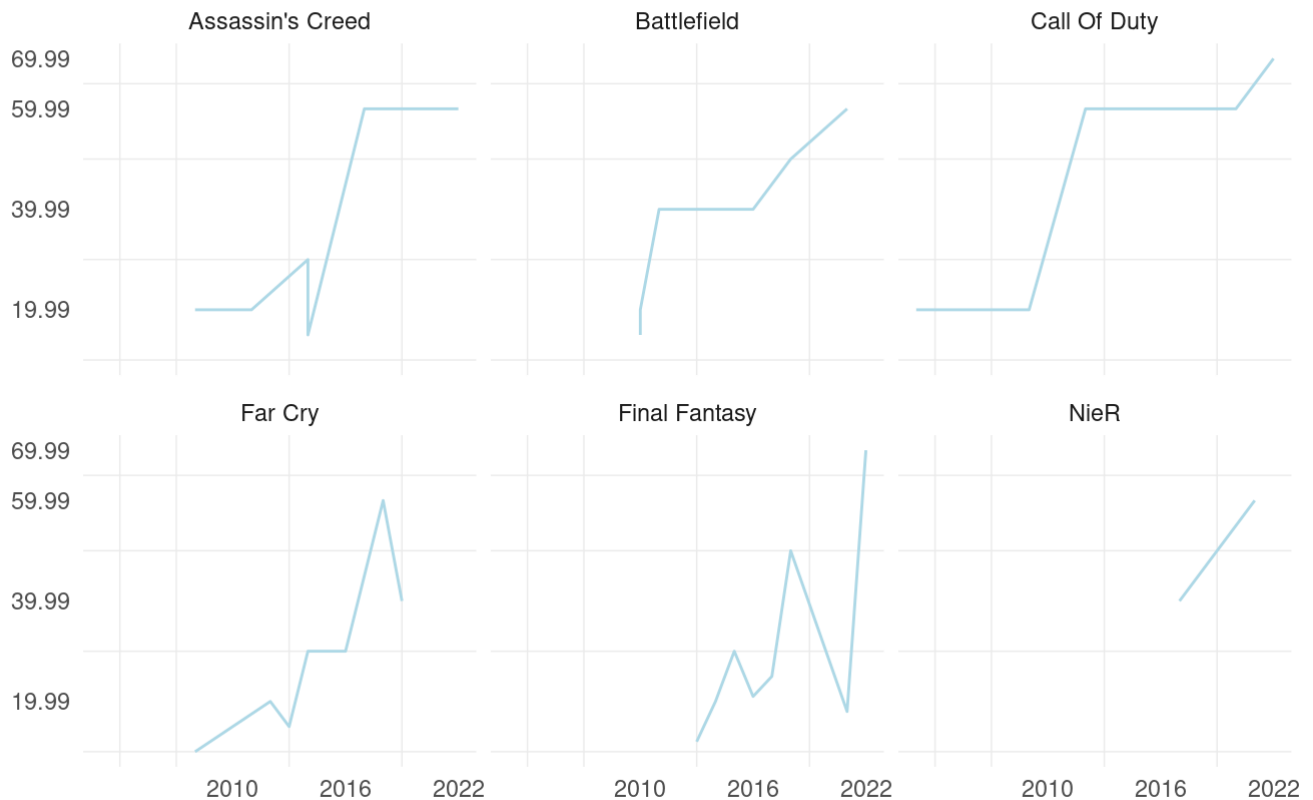
Seguindo alguns exemplos de **jogos que são grandes franquias e são lançados quase que anualmente**, pude observar bem esta tendência de aumento no preço. Vamos usar como exemplo os jogos Assassin’s Creed, Battlefield, Call Of Duty, Far Cry, Final Fantasy, NieR que são franquias já muito conhecidas e que estiveram em alguns períodos no catálogo da Steam (alguns dos jogos não estão presentes de forma integral por figurarem apenas em lojas próprias das

empresas). Removi os Remasters nos jogos que possuíam, que são jogos que são relançados geralmente com pequenas alterações gráficas e acabam tendo preço mais baixo, por ser um jogo mais antigo e no caso do Final Fantasy, como existem muitos jogos menores sendo relançados, mantive apenas o preço mais alto do ano. As franquias Battlefield e Call Of Duty também tiveram que ser corrigida em seus anos de lançamento, pois muitos jogos foram adicionados com sua data de lançamento incorreta.

► Código

Evolução dos preços ao longo dos anos

Preços



Confirmei então **essa tendência de aumento**, mesmo que em alguns casos não seja uma tendência clara, em geral **os preços dos jogos considerados "AAA" subiram** desde os \$19.99 nos anos de 2010 até os \$59.99 e até \$69.99 em alguns casos até o ano de 2022.

Investigando jogos bem avaliados

Mais avaliações positivas

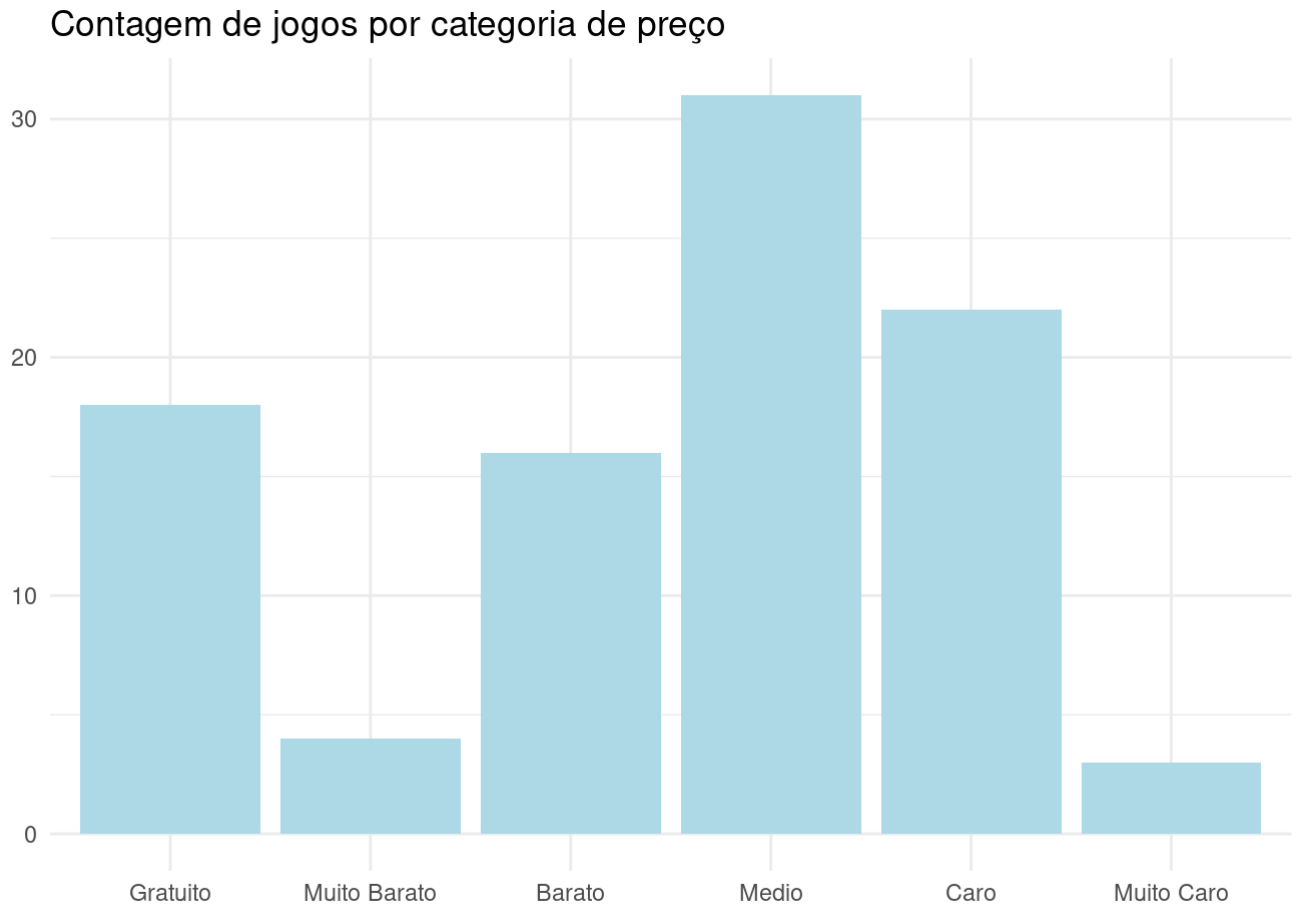
Primeiramente vou analisar os jogos que tem mais avaliações positivas, decidi utilizar os 100 jogos mais bem avaliados da base para tentar trazer algum insight. Além disso, faremos um filtro de ano, para que tenhamos jogos mais atuais.

Para isso, defini alguns filtros para os jogos:

1. Ter sido lançados de 2019 até os dias de hoje, para que tenhamos um mercado observado mais próximo dos dias atuais.
2. Possuírem pelo menos 50000 avaliações.

► Código

► Código



O primeiro gráfico não trouxe nenhuma conclusão, existem jogos baratos que foram lançados e fizeram sucesso, como Phasmophobia, Valheim e Raft, existem jogos que são gratuitos e lucram através de itens cosméticos como o Apex Legends, Lost Ark e Destiny 2 (que foi pago e se tornou gratuito), e existem jogos caros que foram lançados e fizeram sucesso, como CyberPunk 2077, Elden Ring, Red Dead Redemption 2, Forza e Sekiro. Em resumo, os **jogos com maior número de avaliações positivas não parecem ter um padrão relacionado à faixa de preço.**

Mais avaliações positivas

Decidi então observar os jogos que possuíam mais avaliações positivas. E para isso também definimos filtros similares.

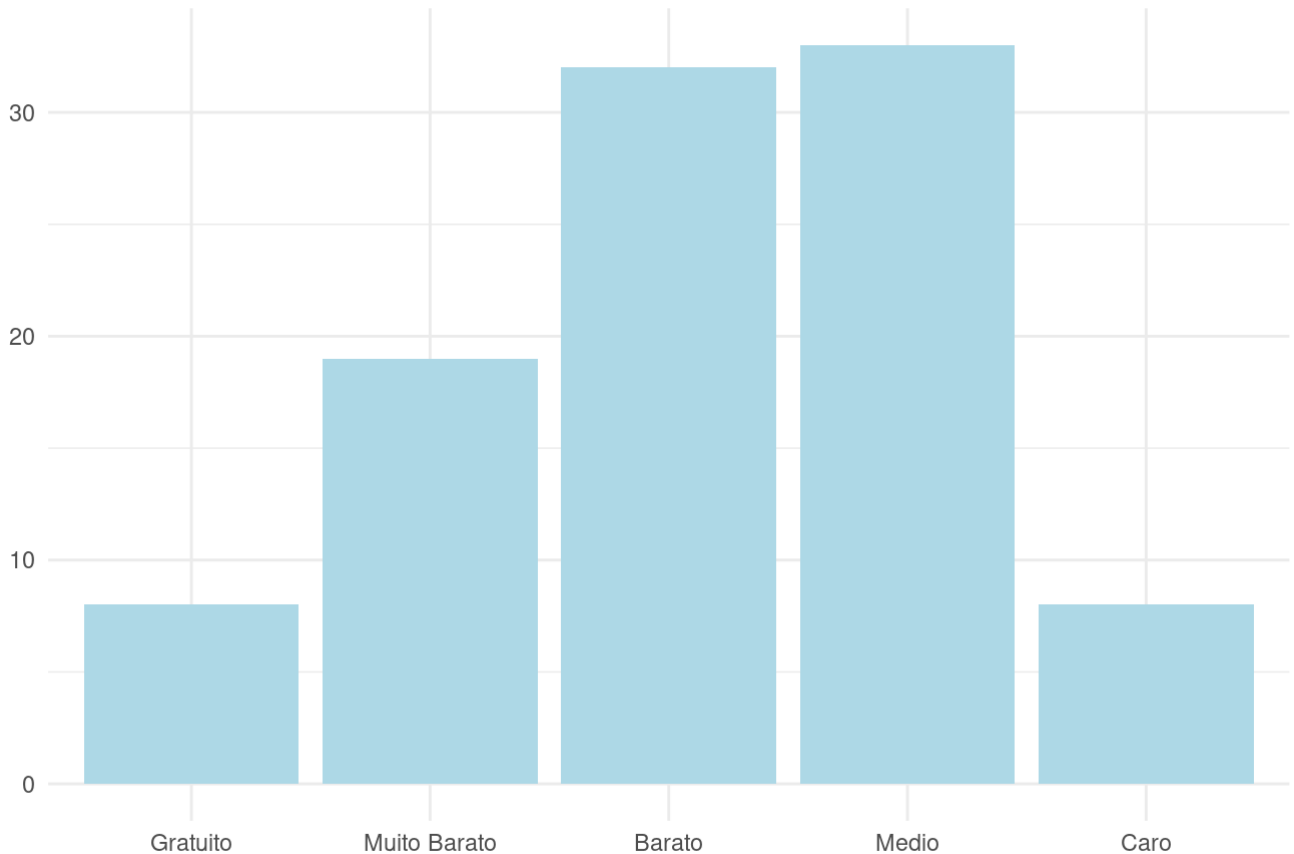
1 - Também terão que ser lançados recentemente, de 2019 até os dias de hoje.

2 - Possuírem pelo menos 10000 avaliações.

► Código

► Código

Contagem de jogos por categoria de preço



Dessa forma, agora ranqueando pelos jogos mais bem avaliados, observei uma **concentração maior** nas **faixas de preço mais baixas**, dos **muito baratos até os médios**, ainda com alguma representação de jogos caros, mas sem jogos muito caros como era visto anteriormente. Comparando com o gráfico anterior, vemos **uma diminuição grande de jogos gratuitos e jogos caros**, praticamente metade em ambos os casos, e um **crescimento grande dos jogos muito baratos e baratos**, os **jogos médios se mantiveram** mais ou menos com a mesma proporção, e os jogos muito caros não apareceram no segundo gráfico.

O que **pode indicar uma faixa de preço** interessante para o lançamento de um jogo seja tanto esse aumento de **jogos muito baratos e baratos**, quanto essa constância dos jogos de **preço médio**. Porém, vamos tentar analisar outros aspectos usando análises mais complexas.

Análise de variáveis combinadas

Investigando variações das tags em categorias de preços

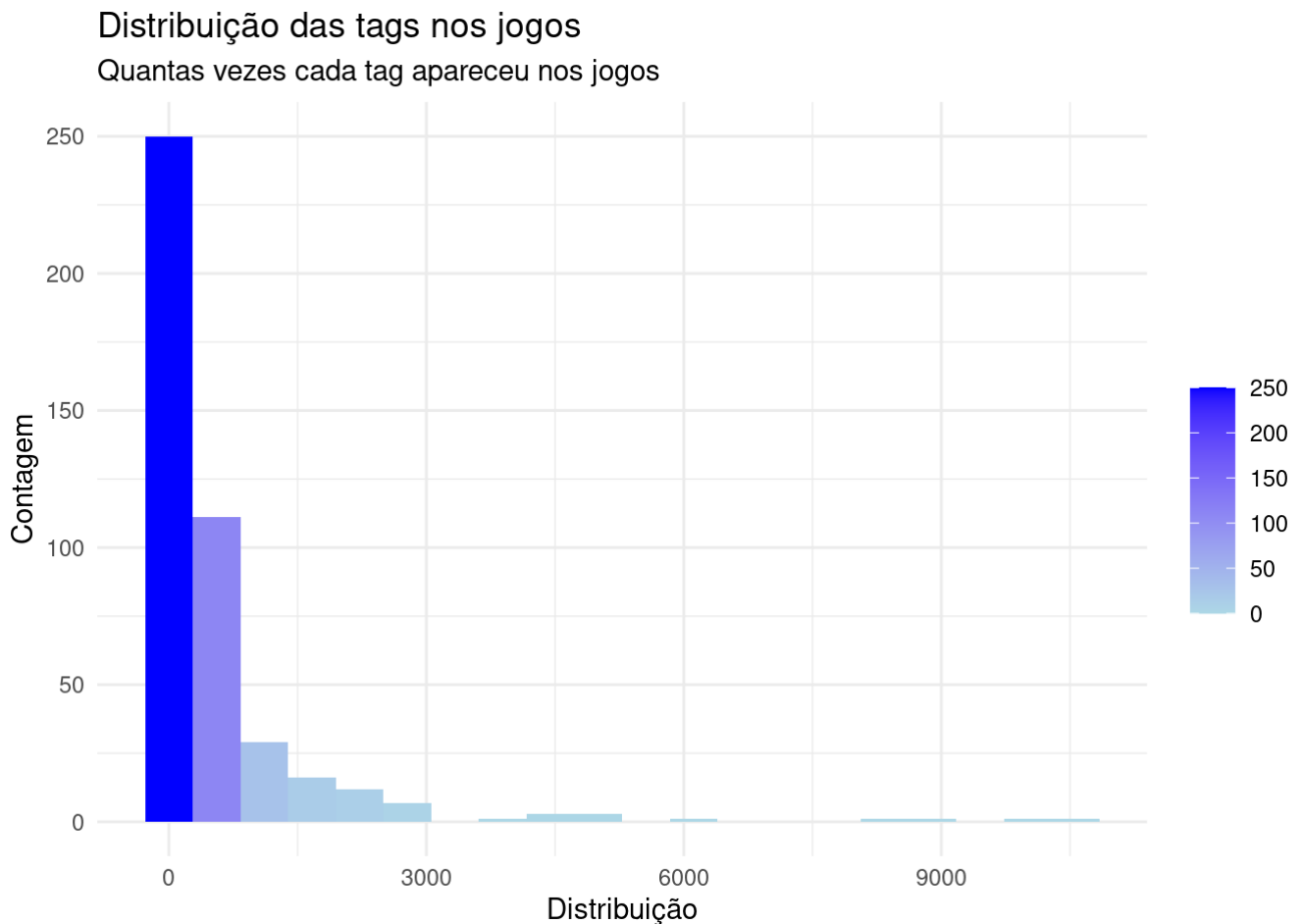
Tags e preços

Agora vou cruzar as informações relacionadas às tags e preços.

Vou fazer um agrupamento entre as tags dos jogos e uma recategorização para a variável de preços, visando facilitar a comparação entre os jogos. Essa recategorização irá replicar cada jogo separando as tags, portanto teremos ao final um número de linhas muito maior que o número de jogos da nossa base inicial, porque cada jogo pode possuir um número de 1 à 35 tags.

O número de **tags distintas** na base de dados é de **437**, e para cada tag, a **média** de vezes em que ela aparece nos jogos é de **594** e a mediana é de **203**, e pela diferença entre média e mediana percebemos que a **distribuição** das tags nos jogos é bastante **irregular**, analisaremos então esta diferença de outra forma.

► Código



Utilizando o histograma é possível verificar quantas vezes cada tag apareceu, e temos uma grande parte distribuída entre as primeiras barras do histograma, que foi separado por faixas, porém podemos observar algumas barras muito menores que aparecem mais vezes, ou seja, tags que aparecem em muitos jogos. Algumas dessas mais frequentes inclusive estão presentes em mais de metade dos **18356 jogos**.

Análise por faixa de preço

A análise será feita nas faixas de preço definidas, os jogos foram separados em **Gratuitos** (preço = 0), jogos **Muito baratos** (preço entre \$0.01 e \$10.00), jogos **Baratos** (preço entre \$10.01 e \$20.00), jogos de preço **Médio** (preço entre \$20.01 e \$40.00), jogos **Caros** (preço entre \$40.01 e \$60.00) e jogos **Muito caros** (preço de \$60.01 ou mais).

Para avaliar os segmentos, defini um número máximo de tags à serem analisadas por faixa de preço, utilizaremos as 20 tags que tem mais aparecem nos jogos por intervalo de preço.

Avaliando os jogos muito caros (Muito Caro)

► Código

Defini essa faixa de preço dos muito caros como jogos **acima de 60 dólares**, e vemos valores de média (3.41) e mediana (2) próximos, aliados ao desvio padrão que também não é alto (3.86), ou seja, temos **pouca variação** do número de **jogos por tag** nessa faixa de preço, indicando uma **distribuição mais uniforme** das categorias.

► Código

Categorias	Contagem Muito Caros
Singleplayer	25
Action	22
Simulation	21
Multiplayer	18
Adventure	12
Co-op	10
Military	10
RPG	10
Realistic	10
Anime	9

Observei que a maioria das tags possui poucas amostras, ou seja, na nossa base de dados e em geral na Steam, **poucos jogos são lançados acima dos \$60 independente da tag**, o que reflete na pouca quantidade de tags dentre os jogos.

Em comparação com a população de jogos, o tamanho da amostra para essa categoria de preço é pequena comparada à nossa base de dados, do total de **18356** apenas **0.27%** ou 50 jogos (um total de 521 tags) estão nesta faixa de valores.

[Avaliando os jogos caros \(Caro\)](#)

► Código

Já nos jogos caros, entre **40 e 60 dólares**, os valores de média (18.36) e mediana (6) possuem uma diferença maior, indicando uma **distribuição mais irregular**. O desvio padrão aumentou (33.75), junto com a diferença entre média e mediana **indica uma variação maior entre tags**.

► Código

Categorias	Contagem Caros
Singleplayer	296

Categorias	Contagem Caros
Action	229
Multiplayer	176
Adventure	174
RPG	162
Story Rich	159
Anime	147
Open World	120
Simulation	114
Great Soundtrack	109

O número de contagem por tag **já é consideravelmente maior** que a faixa anterior, ou seja temos uma quantidade de jogos maior nessa faixa de preços.

Nesta faixa temos **2.05%** dos jogos, ou 377 jogos (um total de 6131 tags) nesse segmento. Indicando um aumento significativo da quantidade de jogos, mas ainda uma proporção pequena na Steam.

[Avaliando os jogos medianos \(Medio\)](#)

► Código

Os jogos de preço médio, **entre 20 e 40 dólares**, vemos valores de média (75.31), mediana (28) e desvio padrão (138.98) novamente **aumentando sua variação**.

► Código

Categorias	Contagem Médios
Singleplayer	1376
Action	993
Adventure	942
Multiplayer	661
Simulation	653
Indie	645
Story Rich	586

Categorias	Contagem Médios
Strategy	586
RPG	577
Atmospheric	554

Os números na faixa de preço continuam crescendo, com **10.29%** dos jogos, ou 1888 jogos (um total de 31404 tags) nesse segmento, chegando à um valor mais expressivo de 10% da base de dados.

Avaliando os jogos “baratos” (Barato)

► Código

Nos jogos de preço baixo, **entre 10 e 20 dólares**, a variação aumenta temos de média 175.73, mediana 64.5 e desvio padrão 335.42. O que pode ser explicado pelo aumento do número de jogos e concentração de jogos em algumas tags.

► Código

Categorias	Contagem Baratos
Singleplayer	3149
Indie	2653
Adventure	2414
Action	2334
Simulation	1414
RPG	1393
Atmospheric	1317
Strategy	1300
2D	1281
Story Rich	1280

Os números de jogos para esse intervalo de preço são **25.48%** dos jogos, ou 4677 jogos (um total de 75211 tags) nesse segmento, mais um crescimento grande, agora com um quarto dos dados da nossa base.

Avaliando os jogos “muito baratos” (Muito Barato)

► Código

Para os jogos de preços **até 10 dólares**, classificados como “Muito Baratos” temos variações ainda crescentes mas mais próximos do segmento anterior, a média é de 209.64, mediana 69 e o desvio padrão 459.04.

► Código

Categorias	Contagem Muito Baratos
Indie	4495
Singleplayer	3798
Action	3319
Adventure	3129
Casual	2772
2D	2009
Simulation	1786
Strategy	1748
RPG	1556
Atmospheric	1549

Os números na faixa de preço crescem um pouco mais, com **38.76%** dos jogos, ou 7115 jogos (um total de 89724 tags) nesse segmento, ou seja, mais de um terço da nossa base. Se pensarmos nos dois intervalos mais baratas, temos mais da metade dos jogos cadastrados, sem dúvida onde os jogos estão mais concentrados.

[Avaliando os jogos gratuitos \(Free\)](#)

► Código

Na última faixa de preços, os **jogos gratuitos**, as variações diminuem. A média é de 133.25, mediana 42.5 e o desvio padrão 292.62.

► Código

Categorias	Contagem Gratuitos
Free to Play	2656
Indie	2373
Action	2058
Singleplayer	1920

Categorias	Contagem Gratuitos
Adventure	1731
Casual	1655
Multiplayer	1567
Strategy	1078
RPG	1048
2D	907

Agora temos uma pequena diminuição do número de jogos em relação ao segmento anterior, podendo ser causada tanto pelas filtragens, sendo uma delas a quantidade de avaliações e usuários por jogo, quanto a própria baixa popularidade nessa faixa de valor, com **23.15%** dos jogos, ou 4249 jogos (um total de 56763 tags) nesse segmento, ainda uma quantidade considerável.

Resumo das faixas de preço

Um ponto interessante e que foi observado foi de praticamente **dois terços**, que se encaixaram nos nossos filtros de avaliações positivas e números de usuários, estão nas faixas de preço até \$20 (de \$0-10 como “Muito Baratos” ou de \$10-20 como “Baratos”), e mais 25% de jogos gratuitos, são faixas de preços que exibem uma **grande diversidade de tags e classificações**. É um comportamento interessante, porque por serem **jogos baratos ou gratuitos**, contam com o **preço** como forma de **atrair usuários**, e podem pensar em **estratégias diferentes de acertar um público alvo**.

Visualizando top tags por faixa de preço

► Código

Gratuitos			Muito baratos			Baratos			
Categoria	N	Perc	Categoria	N	Perc	Categoria	N	Perc	Categoria
Free to Play	2656	17.5%	Indie	4495	18.8%	Indie	2653	16%	Action
Indie	2373	15.7%	Action	3319	13.9%	Adventure	2414	14.6%	Adven
Action	2058	13.6%	Adventure	3129	13.1%	Action	2334	14.1%	Simula
Adventure	1731	11.4%	Casual	2772	11.6%	Simulation	1414	8.6%	Indie
Casual	1655	10.9%	2D	2009	8.4%	RPG	1393	8.4%	Story f
Strategy	1078	7.1%	Simulation	1786	7.5%	Atmospheric	1317	8%	Strate

Gratuitos			Muito baratos			Baratos			
Categoria	N	Perc	Categoria	N	Perc	Categoria	N	Perc	Categoria
RPG	1048	6.9%	Strategy	1748	7.3%	Strategy	1300	7.9%	RPG
2D	907	6%	RPG	1556	6.5%	2D	1281	7.7%	Atmos
Simulation	899	5.9%	Atmospheric	1549	6.5%	Story Rich	1280	7.7%	Open '
Atmospheric	735	4.9%	Puzzle	1513	6.3%	Casual	1149	6.9%	First-P

Coloquei as informações separadas por categoria de preços, mostrando a proporção de cada tag nas faixas de preço, mas, pelo tamanho da tabela, decidi apresentar de forma diferente.

► Código

Utilizando a forma gráfica por linhas, por termos muitas tags, o gráfico ficou confuso, então achei melhor voltar ao formato de tabela, mas usando uma técnica diferente.

Trabalhando com tabela de destaque

► Código

Free	Muito Barato	Barato	Medio	Caro	Muito Caro
Free to Play	Indie	Indie	Action	Action	Action
Indie	Action	Adventure	Adventure	Adventure	Simulation
Action	Adventure	Action	Simulation	RPG	Adventure
Adventure	Casual	Simulation	Indie	Story Rich	Co-op
Casual	2D	RPG	Story Rich	Anime	Military
Strategy	Simulation	Atmospheric	Strategy	Open World	RPG
RPG	Strategy	Strategy	RPG	Simulation	Realistic
2D	RPG	2D	Atmospheric	Great Soundtrack	Anime
Simulation	Atmospheric	Story Rich	Open World	Third Person	Open World
Atmospheric	Puzzle	Casual	First-Person	Atmospheric	Third Person



Paleta de cores

Análises preço e tags

Podemos ver algumas tags de jogos que parecem fazer sucesso, lembrando que as tags não são excludentes, um jogo pode ter mais de uma tag, e em geral todos os jogos possuem muitas delas. Em relação ao modo de jogo, um jogo pode ser definido como sendo de único jogador (Singleplayer) e/ou multijogador (Multiplayer). E existem proporcionalmente mais jogos com modo singleplayer do que multiplayer, ainda que esse filtro esteja presente, provavelmente por conta do custo e da complexidade de se desenvolver um jogo multiplayer.

Singleplayer vs Multiplayer

Mesmo sendo **mais custosos**, muitos dos **jogos mais avaliados são multiplayer**, dos 20 mais avaliados, apenas não possuem modo multiplayer (The Witcher 3 e Cyberpunk 2077), são jogos que possuem entre 450 mil e 7 milhões de avaliações e em geral, são jogos que fazem bastante sucesso, ou seja, são muito vendidos (no caso dos pagos) e jogados. Além disso existem **diferentes modelos de negócio** dentro destes jogos mais avaliados, existem muitos **jogos gratuitos com vendas** de itens **dentro do jogo**, como o CS:GO, Dota 2, Apex, Destiny 2 e Warframe, que são gratuitos mas vendem benefícios dentro do jogo para quem quiser comprar. Existem **jogos com preços variados**, citando Terraria e Among Us como jogos **muito baratos**, Tom Clancy's Rainbow Six® Siege e ARK como jogos **baratos**, Rust e The Witcher 3 como jogos **médios** e Elden Ring e Cyberpunk 2077 como jogos **mais novos e caros**, a **única** categoria que **não** apareceu foi a de **muito caros**.

Porém, se **alterarmos a métrica para** os jogos mais **bem avaliados**, o cenário muda. Ao ordenarmos por taxa de avaliação positiva, os 20 **jogos mais bem avaliados** com nota máxima **são Singleplayer**, **mas** estes jogos **possuem um número** significativamente **menor de avaliações**, entre 300 e 11 mil **e consequentemente** um **número menor de vendas**, além disso são **jogos com preço mais baixos e jogos menos conhecidos**, e são, em geral, jogos ou baratos ou muito baratos.

Tags mais populares nos 100 jogos com mais reviews

Uma possível verificação que eu acho interessante de ser feita, são as avaliações das tags dos 100 jogos mais populares, definido pelo número de avaliações positivas no jogo, para verificar dentre esses que tem mais "sucesso", o que se destaca.

► Código

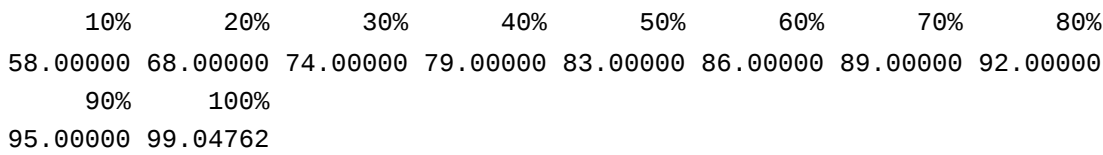
tags_all	count
Action	89
Multiplayer	73
Singleplayer	67
Adventure	57

tags_all	count
First-Person	55
Open World	54
Co-op	53
Atmospheric	43
FPS	40
Shooter	40

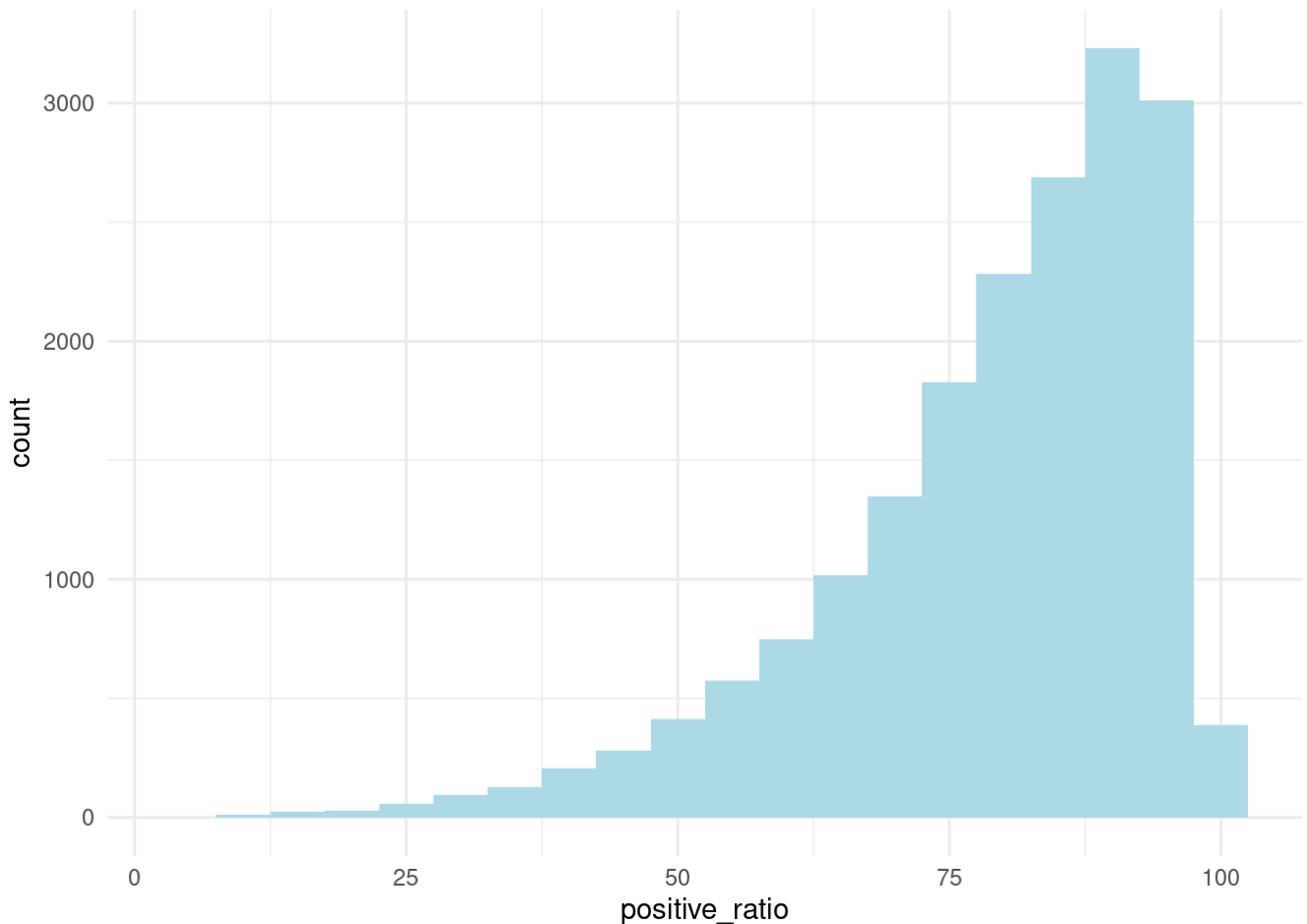
Primeiramente vale lembrar que cada jogo pode possuir um número ilimitado de tags, por isso vemos uma contagem tão alta de aparecimento delas. Também é interessante observar como a **tag Ação** passou à frente das demais, estando **presente em quase todos os jogos** (91%), e como no caso dos jogos do top 100, existem **mais jogos que se definem como Multiplayer do que Singleplayer** (Respectivamente, 74% e 69%), o que nos mostra que **boa parte dos jogos** desse top **possui ambos os modos**, tentando agradar diferentes públicos. Outras tags que valem a pena serem citadas como bastante presentes são a de **Cooperativo e Open World** ambos com 54%, que não são tão fortes na nossa tabela geral, e que **ganham mais destaque em jogos mais caros**, indicando possivelmente que são jogos mais complexos e mais custosos de serem feitos. A tag FPS aparece bastante também com 41% e Indie também tem seu destaque com 36%.

Testei novamente a base com os top 200 jogos ao invés de 100, e o resultado foi muito próximo, mostrando o que parece ser uma tendência real dos jogos mais populares.

► Código



► Código



Avaliando a taxa de avaliações positivas, que vai de 0 a 99 e é calculado com base na relação de opiniões positivas e negativas dos usuários, temos um índice bem alto, na maioria dos jogos. Com praticamente metade desses jogos tendo avaliações acima de 90, e concentração das avaliações entre 70-98. Alguns poucos jogos ficaram abaixo disso, e possivelmente são jogos que podem ser investigados para entender o que motiva essas avaliações tão negativas. O maior exemplo é o Battlefield 2042 com o pior índice de avaliações (35/100), Tale of Immortal (51/100) e PUBG: Battlegrounds (57/100). Em minha experiência, são jogos que realmente foram lançados com problemas e/ou tiveram problemas pós lançamento, portanto é justificável esse resultado baixo nas avaliações.

Tags em jogos com melhores avaliações

Outra análise interessante é a dos jogos mais bem avaliados, definido pela taxa de avaliações positivas, que pode não ser tão precisa pois acaba tendo jogos com um número de avaliações pequeno em comparação ao tópico anterior, mas pode trazer alguns insights em relação a tags de jogos que são bem avaliadas.

► Código

tags_all	count
Singleplayer	57
Indie	54
2D	45

tags_all	count
Adventure	40
Casual	40
Cute	34
Action	31
Story Rich	28
Pixel Graphics	27
Puzzle	25
Anime	23
Funny	21
Atmospheric	19
Female Protagonist	19
Simulation	18
Colorful	16
Multiple Endings	16
Psychological Horror	16
Great Soundtrack	15
RPG	15
Horror	14
Retro	14
Visual Novel	14
Exploration	13
Relaxing	13
Choices Matter	12
Comedy	12
Difficult	12
Mystery	12

tags_all	count
Point & Click	12
Family Friendly	11
Romance	11
Cartoony	10
Dating Sim	10
LGBTQ+	10
Stylized	10
Hand-drawn	9
Strategy	9
Minimalist	8
Platformer	8
RPGMaker	8
Soundtrack	8
Survival	8
Cats	7
Choose Your Own Adventure	7
Controller	7
Fantasy	7
First-Person	7
Interactive Fiction	7
Logic	7
Music	7
Old School	7
Sci-fi	7
2D Platformer	6
Arcade	6

tags_all	count
Emotional	6
Gore	6
Sexual Content	6
Top-Down	6
3D	5
Action-Adventure	5
Bullet Hell	5
Cozy	5
Dark	5
Drama	5
FPS	5
Historical	5
Narration	5
Physics	5
Shooter	5
Sokoban	5
Wholesome	5
Combat	4
Dog	4
Early Access	4
Education	4
Faith	4
Free to Play	4
Hidden Object	4
JRPG	4

tags_all	count
Magic	4
Nudity	4
Post-apocalyptic	4
Psychological	4
Puzzle Platformer	4
Rhythm	4
Score Attack	4
Survival Horror	4
World War II	4
1980s	3
1990's	3
Abstract	3
Cartoon	3
Collectathon	3
Conversation	3
Futuristic	3
Grid-Based Movement	3
Hentai	3
Level Editor	3
Management	3
Memes	3
Metroidvania	3
Nonlinear	3
Realistic	3
Shoot 'Em Up	3
Sports	3

tags_all	count
Stealth	3
Third Person	3
Thriller	3
Turn-Based Combat	3
Violent	3
Beautiful	2
Building	2
Bullet Time	2
Classic	2
Clicker	2
Comic Book	2
Cooking	2
Crafting	2
Creature Collector	2
Dark Fantasy	2
Dark Humor	2
Detective	2
Dungeon Crawler	2
Dystopian	2
Fast-Paced	2
Immersive Sim	2
Investigation	2
Life Sim	2
Linear	2
Local Co-Op	2
Military	2

tags_all	count
Moddable	2
Mouse only	2
Multiplayer	2
Mystery Dungeon	2
Party-Based RPG	2
Procedural Generation	2
Robots	2
Sandbox	2
Sequel	2
Space	2
Strategy RPG	2
Supernatural	2
Tabletop	2
Tactical RPG	2
Text-Based	2
War	2
3D Platformer	1
Aliens	1
Arena Shooter	1
Automobile Sim	1
Bikes	1
Board Game	1
CRPG	1
Character Customization	1
Cinematic	1

tags_all	count
Conspiracy	1
Crime	1
Cult Classic	1
Cyberpunk	1
Dark Comedy	1
Demons	1
Destruction	1
Dinosaurs	1
Dragons	1
Electronic Music	1
Epic	1
Experimental	1
Fishing	1
Hacking	1
Idler	1
Isometric	1
Job Simulator	1
Local Multiplayer	1
Lore-Rich	1
Lovecraftian	1
Martial Arts	1
Massively Multiplayer	1
Match 3	1
Medieval	1
Minigames	1
Motorbike	1

tags_all	count
NSFW	1
Narrative	1
Nature	1
Nostalgia	1
Open World	1
Otome	1
Parody	1
Perma Death	1
Philosophical	1
Photo Editing	1
Puzzle-Platformer	1
PvE	1
Remake	1
Replay Value	1
Resource Management	1
Side Scroller	1
Software	1
Solitaire	1
Spaceships	1
Surreal	1
Swordplay	1
Tactical	1
Team-Based	1
Time Management	1
Time Manipulation	1
Top-Down Shooter	1

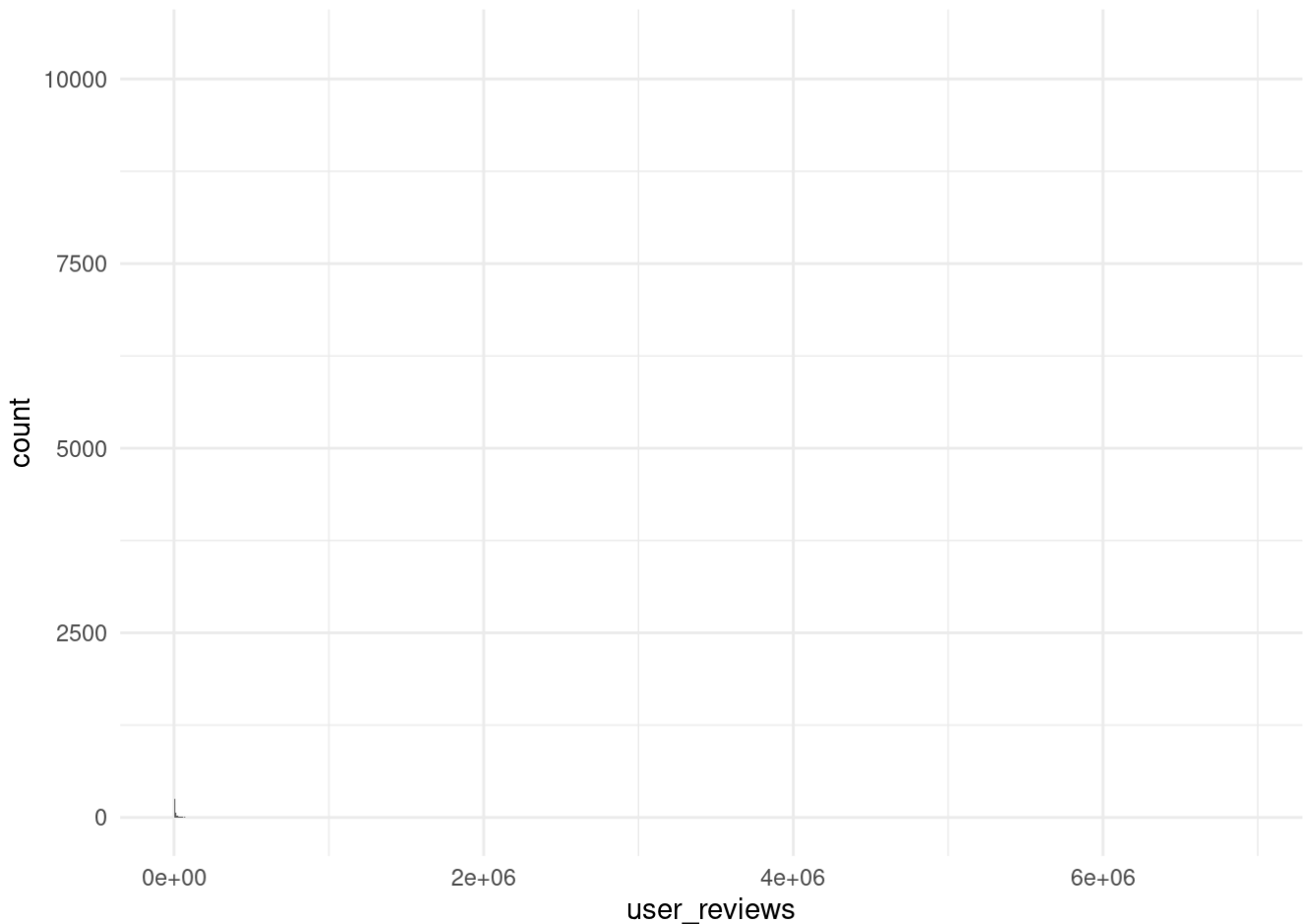
tags_all	count
Transhumanism	1
Turn-Based	1
Turn-Based Strategy	1
Turn-Based Tactics	1
Underground	1
VR	1
Vampire	1
Word Game	1
World War I	1
Zombies	1

Vemos uma diferença considerável nas tags comparado aos jogos com mais avaliações, agora mais em linha com os jogos gratuitos e baratos. A **tag Singleplayer** está presente em **67% dos jogos**, e a **tag Multiplayer** aparece em **apenas um jogo**, mostrando uma diferença enorme da classificação anterior. Outra tag que **ganha mais destaque** é a de **jogos indie**, com **quase o dobro de jogos** que a avaliação anterior. As **tags Action, Open World e Co-Op caem drasticamente**, com respectivos 32% (contra 91% da anterior), 2% (Contra 54% da anterior) e 2% (Contra 54%). Mostrando que as diferenças nessas duas formas de ranquear os jogos refletem em mudanças bem grandes dos jogos.

► Código

10%	20%	30%	40%	50%	60%	70%	80%
124.0	158.0	206.0	277.0	382.0	571.0	930.5	1823.0
90%	100%						
5165.5	6941137.0						

► Código



Outro ponto que acho importante citar, que podemos notar pelo histograma e pelos quartis, é que **aproximadamente dos terços** da nossa base **possui entre 120-1000 avaliações**, com apenas alguns poucos jogos que tem mais avaliações, e por conseguinte mais populares.

Jogos competitivos

Outra análise adicional que achei interessante de fazer, é sobre jogos com modos competitivos, são jogos que existe toda uma comunidade em volta, de pessoas que transmitem seus jogos através do youtube ou interagindo mais com seu público em lives na twitch ou youtube, gerando assim mais engajamento e aumentando a vida útil do jogo. Os jogadores também costumam se manter no jogo por mais tempo, por criar objetivos dentro do jogo de se melhorar e atingir níveis mais altos de ranking, mostrando melhores habilidades. Também são jogos que geralmente são jogados com amigos e grupos, o que tende a aumentar o tempo que um jogador permanece jogando. Porém, são jogos que também trazem seus problemas, principalmente ligados à balanceamento, por serem jogos que dependem de um público grande, é preciso que existam muitos jogadores para que todos possam jogar contra jogadores de seu nível. Além disso, o modelo de negócio, por depender de uma base de jogadores grande, tende a ser gratuita para que muitos possam jogar, com a venda de itens dentro do jogo, normalmente itens cosméticos, que não influenciam nas mecânicas. A questão da distância dos jogadores com os servidores também pode impactar negativamente a experiência do jogo. Além, é claro, dos jogadores que usam programas ilegais para obterem benefícios como mira automática, mais recursos no jogo, etc.

Vou retornar à lista dos 100 jogos mais populares, e analisar quantos deles tem modos competitivos.

► Código

Em análise através das tags, verificamos um total de 10 jogos competitivos, como a minha experiência de jogador achou um número baixo, decidi investigar os 100 jogos um à um, e cheguei num total de 16 jogos (+6 além dos que já haviam sido encontrados). O jogo mais bem avaliado deste top 100 que possui modo competitivo é o Counter Strike (sua primeira versão), que é um fenômeno único e ainda é um jogo muito jogado em suas versões mais novas, tanto que o CS GO também aparece neste top 100. Alguns jogos como Dota 2, Apex, Battlefields e Call of dutys, também apareceram e são jogos que tem seu modo competitivo bastante falado (atualmente).

Concluindo, mesmo totalizando 16 jogos, **não é um número tão relevante** se considerarmos a população de 100 jogos **e dado às complicações** que esta tag pode trazer, **a recomendação** seria de manter **apenas** com modo **Multiplayer casual**, ou seja, sem um ranqueamento, parecem ser mais bem aceitos atualmente.

Conclusão

Ao finalizar essa análise descritiva dos dados de jogos da Steam, pude chegar à algumas conclusões, principalmente após as análises referentes a preços e categorias.

A primeira delas é que **se existe orçamento** capaz de suportar o **desenvolvimento de jogos multiplayer**, a recomendação é que o faça, já que **são os jogos com maior número de jogadores e com mais avaliações**, a interação entre jogadores, quando bem feita, é algo que costuma prender os jogadores e mantê-los jogando por longos períodos de tempo. Nesse caso é preciso estudar o mercado, buscando jogos similares, e **identificar qual modelo de negócio é mais adequado**. Para um **novo estúdio ou nova franquia** os jogos que tem feito mais sucesso são os **jogos gratuitos** com formas de **transações dentro do jogo**, seja algum tipo de facilitação dentro do jogo ou itens cosméticos. A **vantagem dos jogos gratuitos** é que atingem um **grande público**, o que é vital para o desenvolvimento de uma comunidade de jogadores, e **dentro desta comunidade uma parcela poderá gastar dentro do jogo**. Lembrando que, em geral, são jogos que **vendem cosméticos ou acelerações de desempenho**, portanto uma questão **muito importante** é trabalhar de forma que **torne os itens** cosméticos e as acelerações de desempenho itens **desejáveis, porém não desbalanceados** a ponto de incomodar os jogadores que não gastam no jogo.

Já caso o **orçamento** seja **limitado**, jogos com **preços baixos** e que **foquem na experiência Singleplayer** se saem melhor, são jogos que podem ter um **orçamento menor ou estúdios independentes**. Os jogos Indie, que surgem por estúdios de desenvolvimento de jogos independentes costumam ter um gasto menor com desenvolvimento, frequentemente com equipes reduzidas e **tentam trazer algum diferencial**, podendo ser a **história** do jogo, o **design** ou até mesmo alguma **mecânica de jogabilidade**. Jogos como o FTL, Minecraft, Super Meat Boy e Torchlight são exemplos de jogos que começaram como indie e fizeram muito sucesso no mercado de jogos. Inclusive o Minecraft e Torchlight são jogos que perduram até hoje, o primeiro tendo sido comprado por uma gigante no mercado de tecnologias, a Microsoft, e o segundo com o crescimento do estúdio e podendo ser considerado já um estúdio de médio porte.

Ainda, **caso o estúdio possa investir**, porém **não quer se preocupar** com tanto **suporte à longo prazo** de servidores, na questão de hackers ou mesmo tenha foco em modos competitivos, existem **jogos Singleplayers** que tem se destacado no mercado, como o recente Hogwarts

Legacy que teve ótima crítica e tem sido bastante falado, e os jogos The Witcher 3 ou Cyberpunk que estão na lista dos 20 jogos mais bem avaliados, contudo, fica a ressalva de que esses jogos **não são jogos comuns**, The Witcher já é o terceiro jogo da série e foi crescendo em público com o tempo, e Cyberpunk é um jogo da mesma desenvolvedora e conseguiu construir uma campanha de marketing para se tornar um jogo muito aguardado. E o Hogwarts Legacy é um jogo que se utiliza de uma franquia muito famosa no mundo do cinema para atrair o público.