

AVALIAÇÃO FINAL I

1. INTRODUÇÃO

O objetivo da análise deste trabalho está relacionado ao mercado de jogos eletrônicos. Ao estudar a viabilidade e potencial de lucro de um jogo novo, pode ser interessante olhar a resposta do público para cada gênero tanto de forma qualitativa, através de avaliações, quanto de forma quantitativa, através do número de vendas. Com isso, uma empresa consegue tomar a decisão de quais características devem estar em um *game* ou quais jogos devem obter um orçamento maior, por exemplo.

2. DESCRIÇÃO DOS DADOS

Os dados utilizados neste trabalho vieram de 2 *datasets* disponibilizados através do site kaggle. As fontes são:

- [Video Games Sale](#): Este *dataset* oferece a quantidade de vendas individuais das principais regiões do mundo para o mercado de jogos eletrônicos e do mundo todo para cada título. Os dados disponibilizados se referem a mais de 11 mil títulos lançados entre os anos de 1980 e 2017. Nesta tabela, ainda há alguns dados dos jogos como ano de lançamento e gênero;
- [IMDB Video Games](#): Os dados desta tabela indicam o *rating*, ou nota de avaliação dos usuários do IMDb (Internet Movies Database). Este *dataset* também indica a quais gêneros cada jogo pertence e de forma granular, assim, um jogo pode pertencer a mais de um gênero se fizer sentido para as características deste. Esta base possui quase 20 mil entradas de títulos lançados de 1952 até títulos que não foram lançados ainda, mas que já possuem uma entrada no IMDb.

Foram feitos dois gráficos para visualização de dados os quais serão descritos nas subseções seguintes.

2.1. PRIMEIRO GRÁFICO

No primeiro gráfico, foram utilizadas as colunas de ano de lançamento dos jogos, a quantidade de vendas globais e os gêneros da tabela do IMDb. O eixo Y do gráfico corresponde a quantidade de vendas globais dividida pela quantidade de jogos lançados de cada gênero. A quantidade de jogos lançados por gênero foi obtida ao agrupar os títulos por ano e utilizando a função “.sum()” da biblioteca Pandas para contar cada vez que um jogo faz parte de um gênero. A quantidade de vendas também foi obtida para cada gênero. O eixo X, por sua vez, corresponde aos anos dos lançamentos. As bases de dados foram mescladas a partir da coluna de nome e ano de lançamento. Portanto, caso um mesmo título esteja escrito de maneira diferente ou se os anos de lançamento de um mesmo título estiverem diferentes entre as tabelas, uma entrada pode estar repetida.

2.2. SEGUNDO GRÁFICO

Já neste gráfico, foram utilizadas apenas a coluna de *rating* (notas) dos usuários do IMDb e a de gênero da tabela de números de vendas dos jogos. Para este caso, foi utilizado o mesmo DataFrame mesclado com os dois *datasets* utilizado no gráfico anterior. Entretanto, a partir deste DataFrame, foi feito o agrupamento por gêneros iguais e feito uma média das notas através da função “.mean()” da biblioteca Pandas. Para o eixo Y, então, foi utilizado a nota de 1 a 10, enquanto no eixo X está cada um dos gêneros encontrados no *dataset*.

3. JUSTIFICATIVA DAS VISUALIZAÇÕES

No primeiro gráfico, foi utilizado uma linha de tempo em anos com séries em linhas para podermos observar se há alguma tendência de sucesso ou algum padrão no comportamento de mercado de determinados gêneros. Quanto à métrica do eixo Y utilizada, o objetivo seria eliminar o viés de que uma quantidade muito

grande de vendas de um gênero pudesse ser causada simplesmente por uma quantidade muito grande de lançamentos deste gênero.

No segundo gráfico, foi utilizado um gráfico de barras simples pelo motivo de ter métricas simples e com categorias bem discretas.

4. ANÁLISE DOS GRÁFICOS

Podemos observar no gráfico de vendas por gênero ao longo dos anos, que, especialmente dos anos 90 para a atualidade, o mercado parece estar bem variado e equilibrado com exceção do gênero de ação, aventura e *thriller*. Os dois primeiros, por sempre terem sido os gêneros mais populares e por isso podemos ver que aparecem na parte de baixo do gráfico. Como são gêneros com muitos exemplares, possui sucessos (de outra forma, as empresas não iriam apostar tanto neste gênero), porém, possuem muitos fracassos também que fazem com que a razão entre vendas e quantidade de jogos fique muito baixa. A terceira exceção se dá pelo motivo oposto às outras duas. Thriller não é um gênero muito explorado, com poucos títulos aparecendo ao longo dos anos. Por isso, em anos em que houve sucesso de algum título deste gênero, houveram picos na razão entre vendas e quantidade de jogos.

Já no segundo gráfico, observa-se que as médias de notas são bem próximas entre os gêneros, variando do valor mais baixo de 6,58 para o gênero “Misc” (que indica que o jogo pode ser algo como “um pouco de tudo”) até o valor mais alto de 7,79 para o gênero “Role-playing”. Novamente, pela questão da quantidade de lançamentos, os gêneros de ação e aventura ficam próximos (porém, acima) da média, que é de 7,17. Como destaques, além do Role-playing e Misc já citados, entre as notas mais altas, temos também o gênero de Estratégia e de tiro e, para as notas mais baixas, temos jogos de esportes e de quebra-cabeça.

5. LIMITAÇÕES E POSSÍVEIS EXPANSÕES

O *dataset* de vendas utilizado provavelmente foi construído a partir de dados públicos das empresas (ou vazados de alguma forma) e elas nem sempre disponibilizam estes dados e algumas vezes estes podem ser enviesados, devido a formas diferentes de contar as vendas. Também tem de se levar em consideração

que os preços dos jogos não são fixos e variam de desenvolvedora para desenvolvedora, o que pode influenciar na quantidade de vendas. Isso também significa que os dados apresentados não informam receita ou lucro dos gêneros. Outro ponto é o de que também seria interessante analisar as quantidades de vendas absolutas e quantidade absoluta de jogos lançados por gênero, separadamente, ao longo dos anos para se fazer uma análise melhor. Além disso, os dados já estão relativamente desatualizados, tendo dados úteis até apenas 2016 (há apenas algumas poucas entradas em 2017 e 2020) e eles podem ter mudado bastante, principalmente pós-pandemia.

À respeito dos dados sobre as notas do IMDb, a fonte destas notas poderia ser melhor. Muitas pessoas avaliam jogos no IMDb e há uma vasta base de dados sobre eles nesta plataforma, porém, como o nome sugere (Internet Movie Database), esta plataforma é mais dedicada a TV e Cinema. Para corrigir isto, estas notas poderiam ser obtidas de outras plataformas como o Metacritic, que é consideravelmente mais respeitado no mercado de *games* para avaliação da qualidade destes.