

Documentação do Trabalho Prático II

Modelo básico de visualização para os dados
de *download* de aplicativos da *Google Play Store*.

André Fernandes La Rocca Teixeira
Visualização de Dados - 2018/02
Departamento de Ciência da Computação
Universidade Federal de Minas Gerais

Introdução

O objetivo deste trabalho é criar uma visualização de dados que permita a análise de dados relevantes sobre o número de *downloads* de aplicativos na *Google Play Store*.

A análise aqui presente trata da quantidade aproximada do número de instalações de aplicativos móveis projetada em uma série de parâmetros, como: categoria, tempo desde a última atualização e tamanho do aplicativo.

Analisar esse tipo de dado é relevante para entender quais são os aplicativos preferidos da loja, ou seja, entender onde os usuários pretendem gastar mais dinheiro ou investir a maior parte do seu tempo. A compreensão do comportamento do usuário na loja de aplicativos é importante para direcionar investimentos, realizar promoções, divulgações e tomar outras decisões estratégicas.

Os dados foram retirados do *Kaggle*^[1], e filtrados de forma a conter somente aplicativos com mais de dez milhões de instalações totais.

Visualização Desenvolvida

O presente trabalho tem como objetivo analisar os dados de número de instalações de aplicativos móveis a partir da *Google Play Store*. Os dados foram cruzados com outras informações para permitir a comparação entre elas.

O gráfico de barras foi o escolhido para representar esta distribuição, pois permite a comparação visual de valores através do tamanho de suas barras. Quatro conjuntos de dados foram aplicados ao gráfico, cada um referente a um dos parâmetros escolhidos.

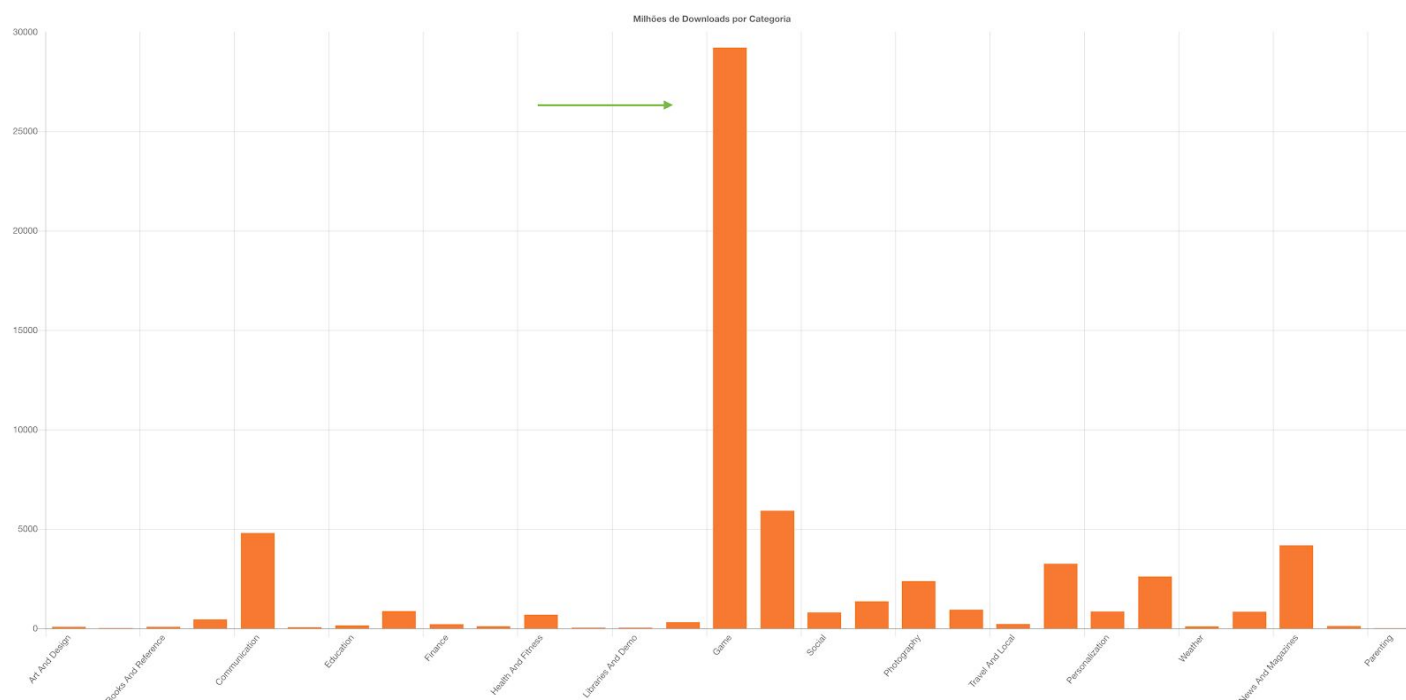
Desta forma, é possível visualizar a diferença da quantidade de *downloads* por categoria, quantidade de *downloads* por tamanho do aplicativo e quantidade de *downloads* por tempo desde a última atualização.

Este tipo de análise permite visualizar dados como: preferência dos usuários entre os tipos de aplicativos oferecidos, se os usuários preferem aplicativos mais completos (maiores) ou aplicativos mais simples, se a manutenção do aplicativo tem influência em seu número de *downloads*.

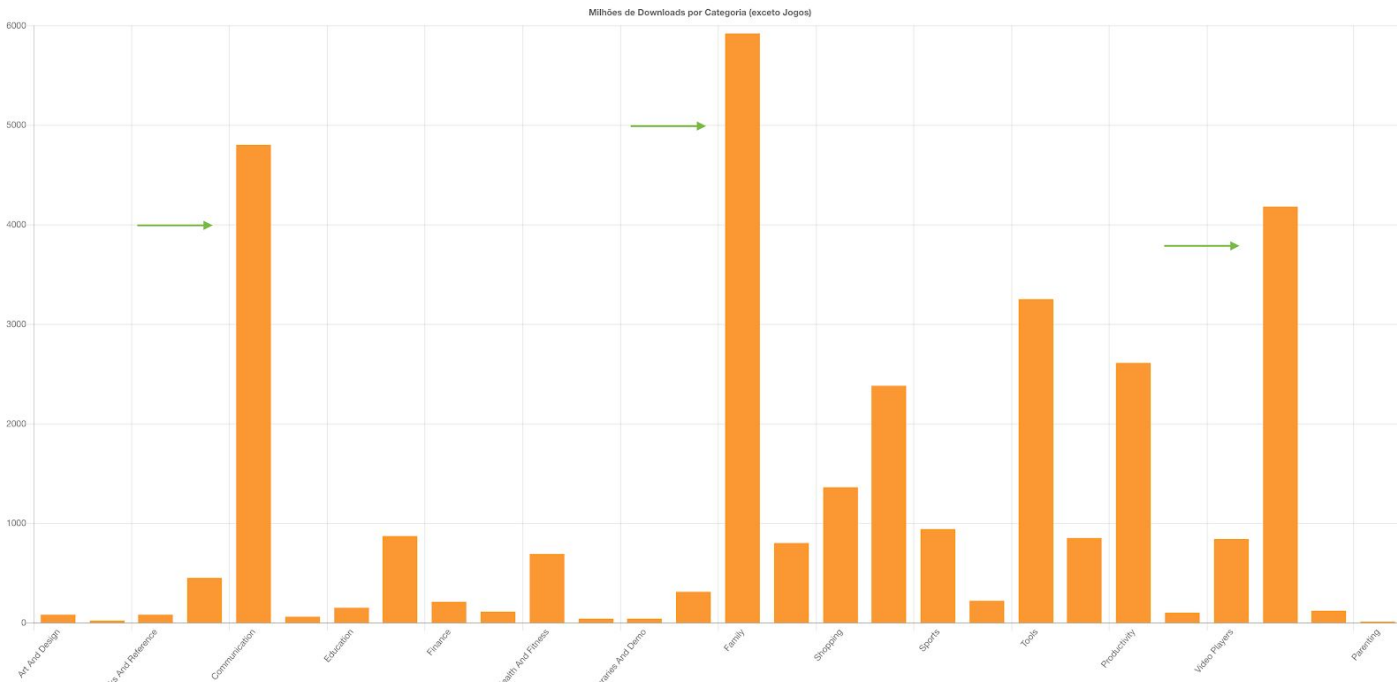
Análises dos Dados

Diversas conclusões podem ser observadas através da visualização, como por exemplo:

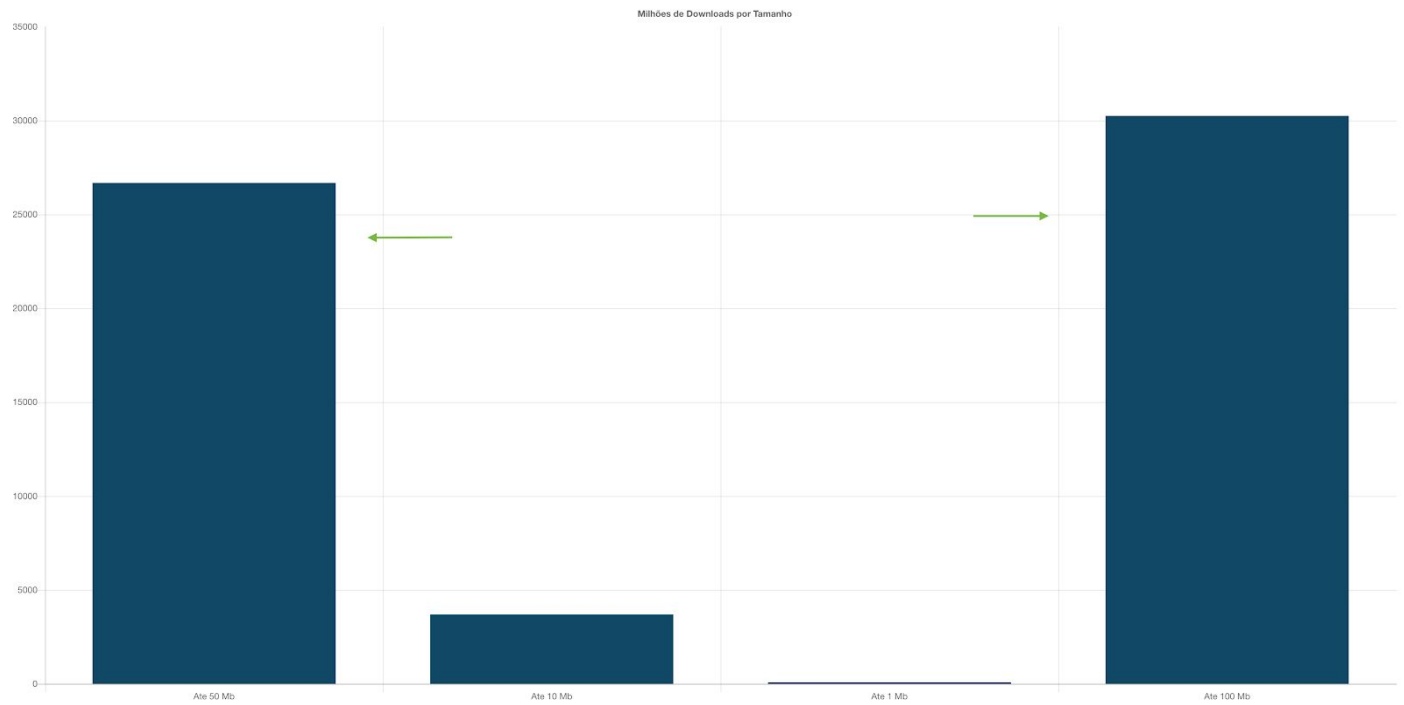
- A categoria de jogos é, de longe, a preferida pelos usuários. Isso explica o seu perceptível crescimento. Pelo gráfico, é possível perceber que a categoria, com mais de 29 bilhões de *downloads*, chega a atrapalhar a comparação entre outras categorias. Portanto, é possível escolher outra visualização de categorias, desconsiderando os jogos. Para um investidor, por exemplo, essa visualização permite compreender que a grande maioria dos usuários gasta bastante tempo em seus jogos, e isso pode ser uma opção de bom retorno financeiro. O fato de usuários investirem tempo em um aplicativo significa que verão anúncios, gastarão dinheiro, entre outras coisas.



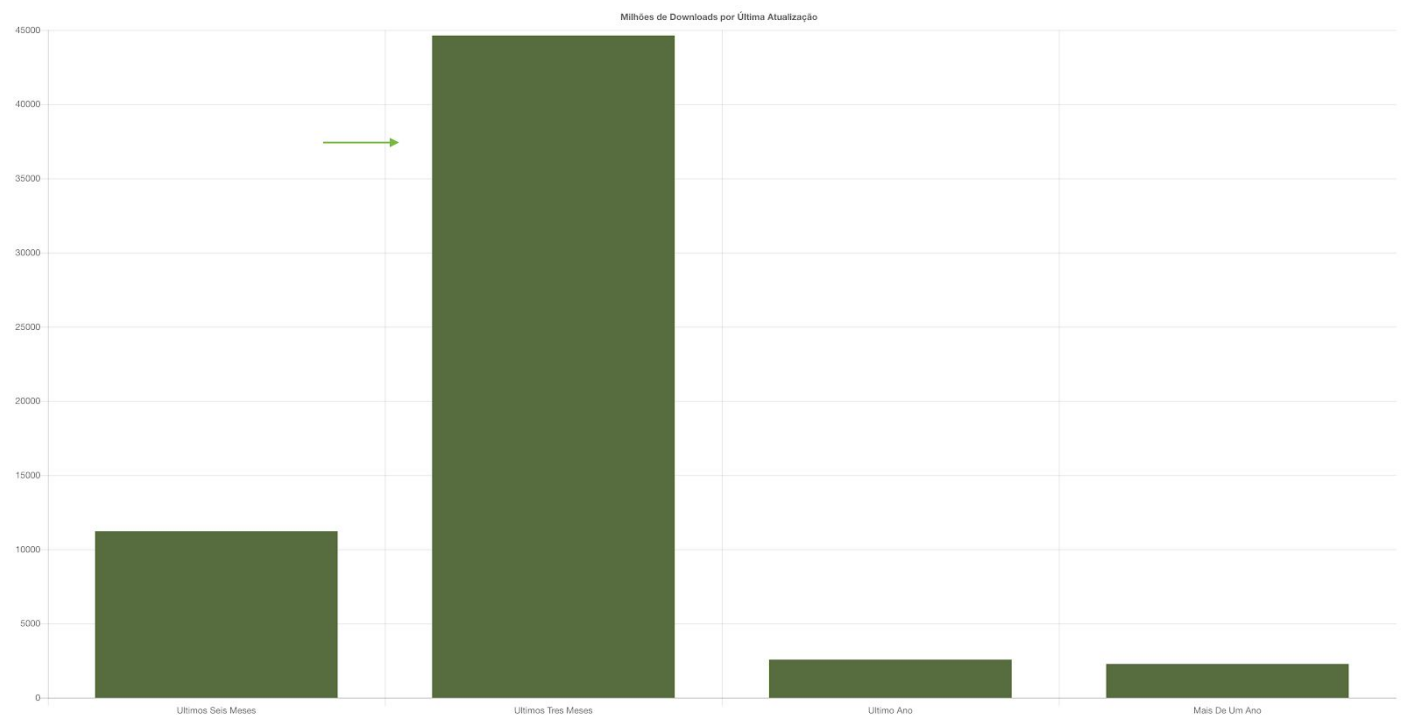
- Sabemos, porém, que o mercado de jogos não é fácil. Poucos jogos dominam o mercado, controlando grande parte dos usuários desse nicho. Portanto, uma outra visualização foi criada, para permitir a comparação entre outras categorias, que não jogos. Nessa visualização, é possível perceber que as alternativas mais relevantes aos jogos são: Comunicação, Família e Notícias. Essa é uma boa notícia para os investidores: vários aplicativos se enquadram nessas categorias, e podem significar excelentes retornos financeiros. Na parte de Comunicação, temos os famosos aplicativos mensageiros, que podem ajudar a acumular informações valiosas sobre seu público. Em relação à categoria Família, os aplicativos voltados para o público infantil representam uma fatia relevante do número total de *downloads*, já que as crianças de hoje em dia são fascinadas com tecnologia e aprendem cedo a desfrutar disso. A evolução tecnológica fez com que as mídias evoluíssem e, portanto, as notícias evoluíram também. Outra importante fatia do mercado se resume aos aplicativos de notícias, que substituem jornais e revistas e, em alguns casos, até a televisão.



- A visualização de número de *downloads* por tamanho do aplicativo nos mostra um dado curioso. Diferente do esperado, os aplicativos mais baixados são, também, os mais pesados. Isso é claramente influenciado pela categoria de jogos, já que esses aplicativos tendem a ocupar bastante espaço na memória. Ainda assim, os números surpreendem. Os aplicativos com mais de 50 MB possuem cerca de 57 bilhões de downloads. Isso pode significar que, com a evolução dos aparelhos celulares, a preocupação com a memória interna do dispositivo não seja mais tão importante.



- Finalmente, a visualização de número de *downloads* por tempo desde a última atualização. Não é surpreendente o fato de que os usuários consomem muito mais os aplicativos que são frequentemente mantidos e atualizados por seus desenvolvedores. As atualizações periódicas significam correções de problemas e novas funcionalidades que, geralmente, agradam aos usuários. Alguns aplicativos ainda resistem, sem atualizações. Mas a supremacia dos aplicativos de manutenção mais frequente é indiscutível.



Ferramenta

A ferramenta escolhida para construção da visualização foi a biblioteca *Chart.js*^[1]. Esta biblioteca foi escolhida pois permite a criação de gráficos e outros tipos de visualização em um nível um pouco mais alto de abstração.

A *Chart.js* conta com diversas opções de gráficos prontos e é completamente personalizável. Ou seja, é possível utilizar como base o que a biblioteca oferece e, ainda, implementar o que mais for necessário para que a visualização fique completa.

Além disso, a biblioteca possui código aberto, é mantida pela comunidade, e oferece todos os seus serviços gratuitamente.

Conclusão

Foi possível criar uma visualização interativa^[2] que permite tirar diversas conclusões sobre o comportamento dos consumidores de aplicativos na *Google Play Store*. Os objetivos iniciais, portanto, foram atingidos.

Pontos importantes sobre esse comportamento puderam ser interpretados a partir da visualização, como a preferência por aplicativos mantidos por seus criadores, e a supremacia da preferência pela categoria de jogos.

Para acessar a visualização, basta acessar a referência [2], ou abrir o arquivo *index.html*, disponível no pacote, utilizando o navegador de sua preferência. Para interagir, basta clicar nos itens da legenda para alternar a visualização de cada um dos dados. O código também está disponível no repositório^[3] do *GitHub*.

Referências

[1] <https://www.kaggle.com/lava18/google-play-store-apps>

[2] <https://andrelarocca.github.io/play-store-apps/>

[3] <https://github.com/andrelarocca/play-store-apps>