

FERRAMENTAS E BIBLIOTECAS DE CÓDIGO PARA VISUALIZAÇÃO DE DADOS: UMA ANÁLISE DE FUNCIONALIDADES

Filipe Garcia¹

Maicon Bernardino Da Silveira²

Elder Rodrigues³

Resumo:

A geração de gráficos para a visualização de dados é um ponto chave para o entendimento de informações, e se torna um fator determinante para obter benefícios em um determinado domínio. Em diversos contextos e situações, para que a abstração de um problema, resultado, ou compreensão de ações, seja facilitada, os dados gerados por meio de um determinado processo, são transformados em gráficos, sejam eles de quaisquer tipo, padrão, interativo ou não. O presente trabalho tem como motivação, a análise de ferramentas de geração de gráficos, para que após a análise, seja possível realizar um comparativo entre todas as opções encontradas, e selecionar a que melhor se enquadra, para ser utilizada em um trabalho de conclusão de curso. A escolha será utilizada como base para o módulo de uma ferramenta de teste de desempenho, para gerar gráficos a partir do arquivo de dados coletados durante o monitoramento. O teste de desempenho possui atividades chave, sendo estas: modelagem/scripting; execução; monitoramento e análise dos resultados. Através da geração de carga, são realizados testes em um software, para determinar sua capacidade de desempenho, e durante os testes, são monitoradas todas as ações, que geram dados. Estes dados são utilizados para realizar a análise de resultados, onde são produzidos gráficos. Uma proposta onde gráficos interativos são dispostos durante a última etapa, seria de grande proveito, a fim de tornar a compreensão das informações mais simples e eficaz. Para a realização do estudo, foi realizado um mapeamento, a fim de encontrar ferramentas e bibliotecas que se encaixam no escopo do trabalho, para que através da escolha de cada uma delas, seja possível realizar um comparativo entre suas funcionalidades. Por meio do estudo realizado, é possível analisar, através dos resultados obtidos, as diferenças, vantagens e desvantagens de cada ferramenta, além do contexto onde seu uso é mais adequado. Ao término do estudo apresentado, os leitores serão capazes de entender o que são as ferramentas e bibliotecas de geração de gráficos para a análise de dados, para que servem, e qual solução usar, em um determinado cenário.

Palavras-chave: Ferramentas de visualização de dados; visualização de dados, geração de gráficos, teste de desempenho, ferramentas.

Modalidade de Participação: Iniciação Científica

FERRAMENTAS E BIBLIOTECAS DE CÓDIGO PARA VISUALIZAÇÃO DE DADOS: UMA ANÁLISE DE FUNCIONALIDADES

¹ Aluno de graduação. filipe.garcia1997@gmail.com. Autor principal

² Docente. maiconbernardino@unipampa.edu.br. Orientador

³ Docente. eldermr@gmail.com. Co-orientador

FERRAMENTAS E BIBLIOTECAS DE CÓDIGO PARA VISUALIZAÇÃO DE DADOS: UMA ANÁLISE DE FUNCIONALIDADES

1 INTRODUÇÃO

A geração de gráficos para a visualização de dados é um ponto chave para o entendimento de informações, e se torna um fator determinante para obter benefícios em um determinado domínio (TRAINA *et al.*, 2001). Em diversos contextos e situações, para que a abstração de um problema, resultado, ou compreensão de ações, seja facilitada, os dados gerados por meio de um determinado processo, são transformados em gráficos, sejam eles de quaisquer tipo, padrão, interativo ou não.

O presente trabalho tem como motivação, a análise de ferramentas de geração de gráficos, para que após a análise, seja possível realizar um comparativo entre todas as opções encontradas, e selecionar a que melhor se enquadra, para ser utilizada em um trabalho de conclusão de curso. A escolha será utilizada como base para o módulo de uma ferramenta de teste de desempenho, para gerar gráficos a partir do arquivo de dados coletados durante o monitoramento. O teste de desempenho possui atividades chave, sendo estas: modelagem/*scripting*; execução; monitoramento e análise dos resultados (WEYUKER *et al.*, 2000). Por meio da geração de carga, são realizados testes em um software, para determinar sua capacidade de desempenho, e durante os testes, são monitoradas todas as ações, que geram dados. Estes dados são utilizados para realizar a análise de resultados, onde são produzidos gráficos. Uma proposta onde gráficos interativos são dispostos durante a última etapa, seria de grande proveito, a fim de tornar a compreensão das informações mais simples e eficaz.

Atualmente, existem abordagens de visualização de dados em teste de desempenho, que dispõe de diversas formas de visualização (LEE *et al.*, 2014), porém, ainda está longe de ser um padrão que permita ao usuário, ter várias opções de exibição, e também de interação com os dados, a fim de obter mais informações sobre uma informação específica. O conceito de visualização de dados de maneira interativa, já é antigo (BUJA *et al.*, 1991), existem diversas técnicas e propostas já estabelecidas. Apesar de ser um conceito antigo, alguns projetos tem se tornado realidade apenas no contexto atual, em que novas ferramentas têm sido apresentadas, possibilitando visualizar dados por meio de gráficos, e também de forma interativa (HUBER *et al.*, 2018).

A Seção 2 apresenta de forma sucinta qual foi o protocolo utilizado para a realização do estudo, e também quais os métodos e técnicas foram aplicados durante a análise das ferramentas. A seção expõe uma análise sobre quais as melhores formas de utilizar as ferramentas de teste de desempenho, e qual a motivação de realizar o uso delas. Posteriormente, a Seção 3 exhibe um comparativo entre todas as ferramentas analisadas, quais as vantagens e desvantagens de cada uma delas, e em quais condições o uso de cada uma é indicada. A Seção 4 dispõe das considerações finais sobre o estudo realizado, e descreve os trabalhos futuros. Ao término do estudo apresentado, os leitores serão capazes de entender o que são as ferramentas de geração de gráficos para a análise de dados, para que servem, e qual ferramenta usar, em um determinado cenário.

2 METODOLOGIA

Para a realização do estudo, foi realizado um mapeamento, a fim de encontrar ferramentas que se encaixam no escopo do trabalho, para que através da escolha de cada uma delas, seja possível realizar um comparativo entre suas funcionalidades, e também obter informações relevantes, a respeito de quais são os cenários, em que cada uma delas pode se tornar a melhor escolha. Para a execução do mapeamento foram utilizadas as bases de dados exibidas na Tabela 1 abaixo.

Fonte	Endereço	Tipo
ACM Digital Library	dl.acm.org	Híbrida
IEEE Xplore	ieeexplore.iee.org	Base Bibliográfica
ScienceDirect	sciencedirect.com	Base Bibliográfica
Scopus	scopus.com	Motor de Busca
Google Scholar	scholar.google.com	Motor de Busca

Tabela 1. Bases de dados.

Para a escolha de cada uma das ferramentas, foram estabelecidos cinco critérios para serem utilizados como base, em que para a seleção, a ferramenta deveria atender a pelo menos três critérios, do número total de critérios estipulados, sendo eles: 1- Visualização de dados a partir de um conjunto de informações. 2 - Disposição dos dados em diferentes tipos de gráficos e/ou tabelas. 3 - Visualização Interativa/Dinâmica dos dados. 4 - Facilidade da integração da ferramenta com outra plataforma. 5 - Qualidade da apresentação dos dados. Após o mapeamento realizado, foram selecionadas oito ferramentas, listadas na Tabela 2.

Ferramenta
JavaScript InfoVis Toolkit
Tableau
Plotly
RAWgraphs
D3.js
amCharts
Chart.js
Chartist.js

Tabela 2. Ferramentas selecionadas para o estudo.

A partir da lista de ferramentas selecionadas (Tabela 2), foram realizados estudos e análises em cada uma delas, a fim de obter informações a respeito das funcionalidades de cada uma delas. As investigações possibilitam que seja realizado um levantamento geral sob cada um dos artefatos de estudo, como o padrão de licença da ferramenta (código aberto, versão paga), operação via web e/ou desktop, linguagem e/ou bibliotecas utilizadas para a geração dos gráficos. Na Seção 3 será disponibilizado uma tabela, efetuando um comparativo entre as ferramentas, e em seguida, considerações gerais a respeito da usabilidade de cada uma delas.

3 RESULTADOS e DISCUSSÃO

Após a seleção das ferramentas dispostas na seção acima, foram realizadas diversas análises e comparações quanto às funcionalidades de cada uma delas. Os resultados obtidos, são frutos das análises referentes as ferramentas citadas na Seção 2. A Tabela 3 apresenta um comparativo em relação a algumas funcionalidades mais comuns das ferramentas de visualização de dados, para evidenciar a diferença de cada uma delas. A tabela não expõe questões de usabilidade, e sim características funcionais de cada uma delas, para que assim, seja possível realizar algumas considerações sobre a aplicação de cada uma delas em contextos específicos. Foi realizado um comparativo entre todas elas, para obter conhecimentos a respeito das funcionalidades, e a última coluna dispõe também, de um atributo “Recomendação de uso”, onde está exibido para qual tipo de gráfico, ou problema, a ferramenta será mais eficiente.

Ferramenta	Código Aberto	Ling. da Ferramenta	Web	Desktop	Gráficos Interativos	Recomendação de uso
JavaScript InfoVis Toolkit	✓	JavaScript	✓	✗	✓	Conexões e dependências
Tableau	✓	✗	✓	✓	✗	Gráficos para análise de tendências
Plotly	✗	JavaScript, Python, R	✓	✗	✓	Visualização de dados complexos
RAWgraphs	✓	JavaScript	✓	✗	✗	Gráficos para análise de dados em massa a partir de um arquivo
D3.js	✓	JavaScript	✓	✗	✓	Gráficos interativos e com animações
amCharts	✓	JavaScript	✓	✗	✓	Gráficos e mapas interativos

Chart.js	✓	JavaScript	✓	✗	✓	Gráficos minimalistas
Chartist.js	✓	JavaScript	✓	✗	✓	Gráficos responsivos e animados

Tabela 3. Comparativo entre as ferramentas.

Todas as ferramentas analisadas possuem contextos onde sua aplicação pode gerar resultados mais satisfatórios e eficientes, e também possuem cenários em que seu uso pode não ser tão proveitoso. A Javascript InfoVis Toolkit, possui diversos recursos para a criação de conexões e dependências, e conta também com alguns gráficos simples, pode ser muito bem utilizada para modelar dados que possuem diversas conexões, em contraponto, se trata de uma biblioteca de códigos, sendo assim, para utilizá-la, é preciso certo conhecimento prévio em programação. A Tableau, está disponível de diversas maneiras, com opções em Desktop, Cloud, e Web. Possui versões pagas e gratuitas, e apresenta diversas opções de gráficos e tabelas, não sendo necessário nenhum conhecimento prévio em programação, para utilizá-la, sua desvantagem se encontra na questão de não ser uma ferramenta de código aberto, e a versão completa, é paga. A biblioteca Plotly, dispõe de uma enorme variedade de gráficos, permitindo ao usuário trabalhar com diferentes linguagens de programação, a questão estética dos gráficos, é um diferencial. Em contraponto, sua versão paga é completa, e a gratuita, muito limitada, não possui todas as funcionalidades, e os gráficos criados se tornam públicos. A RAW *graphs* é extremamente fácil de ser utilizada, podem ser gerados gráficos, a partir de um arquivo ou planilha, todo o processo é realizado por meio do navegador, sendo disponibilizados em torno de dezesseis tipos de gráficos customizáveis. Deixa um pouco a desejar quanto a interatividade dos gráficos gerados, porém é totalmente gratuita. A D3.js é uma biblioteca JavaScript utilizada para a criação de gráficos. Para sua utilização, é necessário conhecimento prévio, e algumas noções de programação. Além de ser código aberto, sua interatividade junto as animações dos gráficos, são o principal destaque. A amCharts possui uma versão gratuita, e também versões pagas. Provê diversos tipos de gráficos, em excelente qualidade visual, seu destaque, fica em conta dos gráficos simples, e interativos, consegue cumprir muito bem sua proposta. O principal ponto negativo, fica também em questão de sua versão gratuita, não dispor de todas as funcionalidades. A Chart.js dispõe de dezenas de opções de gráficos, para serem utilizados. Para seu uso, é necessário um conhecimento prévio de programação. A biblioteca permite a criação de gráficos minimalistas e esteticamente limpos, com algumas funções de interatividade. Por fim, a Chartist.js permite a criação de gráficos responsivos e animados, utilizando também JavaScript. Para sua utilização, é necessário entender alguns conceitos de programação. Totalmente gratuita e livre de dependências, é utilizada por intermédio de um navegador, porém, a interatividade dos gráficos não é algo surpreendente, e pode evoluir.

Todas as soluções analisadas possuem diferentes funcionalidades, questões técnicas, e usabilidade distintas. Como apresentado, existem diversas opções de ferramentas e bibliotecas, e cada uma delas pode ser mais adequada em um determinado cenário. Com as informações na coluna de “Recomendação de uso”, exibida na Tabela 3, é possível ter uma visão genérica, dos cenários mais adequados para a utilização de cada uma delas, e nas considerações acima, alguns detalhes quanto a usabilidade e outros recursos de cada uma delas.

4 CONSIDERAÇÕES FINAIS

Em vista dos resultados obtidos, através da comparação das ferramentas, da análise de suas funcionalidades, pode-se concluir que não existe uma ferramenta que irá ser a melhor opção para todas as situações e contextos possíveis, porém, podem ser escolhidas a ferramenta que mais se adequa a um cenário específico, como descrito na seção de Resultados. Todas as ferramentas possuem pontos positivos e negativos, portanto, a melhor escolha, depende de inúmeras situações e fatores, além dos requisitos de utilização. Para a escolha de uma ferramenta de geração de gráficos, deve-se empregar um tempo considerável para estudar e analisar fatores como quais os tipos de gráficos podem ser gerados, se eles são interativos ou não, padrões visuais e de visualização, e por fim, analisar qual tipo de gráfico mais se encaixa para a realização de uma determinada tarefa. A partir dos resultados coletados, pode ser feita uma observação perante a todas as opções descritas no estudo, e selecionar a solução que é a ideal para uma determinada finalidade. Por fim, como trabalho futuro, essa análise será aplicada, para a definição de qual ferramenta será escolhida, para ser utilizada, em um trabalho de conclusão de curso. Cabe ressaltar também que muitas abordagens novas estão surgindo recentemente na literatura acadêmica, a possibilidade de campo de estudo e desenvolvimento de soluções para geração de gráficos, é grande e promissora, e deve entrar cada vez mais em evidência no cenário da computação.

REFERÊNCIAS

- BUJA, Andreas et al. **Interactive data visualization using focusing and linking**. In: Proceedings of the IEEE Conference on Visualization. IEEE. p. 156-163, 1991.
- WEYUKER, Elaine J. et al (2000). **Experience with performance testing of software systems: issues, an approach, and case study**. In: IEEE transactions on software engineering, v. 26, n. 12, p. 1147-1156, 2000.
- HUBER, Timothy C. et al (2018). **Developing an Interactive Data Visualization Tool to Assess the Impact of Decision Support on Clinical Operations**. Journal of Digital Imaging, p. 1-6, 2018.
- LEE, Sungchul; JO, Ju-Yeon; KIM, Yoohwan. **Performance testing of web-based data visualization**. In: Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC). IEEE, p. 1648-1653, 2014.
- TRAINA, Agma JM et al. **Visualização de Dados em Sistemas de Bases de Dados Relacionais**. In: SBBD (Simpósio Brasileiro de Banco de Dados). p. 95-109, 2001.