

Ferramentas de Baixo Custo para Integração e Geração de Dashboards de Visualização de Dados

Luander I. de Arruda¹, Marcos P. Mesquita²

¹Faculdade de Ciências Exatas e Tecnológicas (FACET) - Universidade do Estado do Mato (UNEMAT), Campus Universitário Jane Vanini
Av. São João, S/N, CEP 78216-060, Cavalhada, Cáceres, MT - Brasil

{luander.arruda, mpmesquita}@unemat.br

Abstract. *The volume of data generated by Open Government policies demands the development of effective tools for processing and extracting information from these large data sources. This article proposes a methodology that uses accessible software interfaces to integrate data and generate interactive visualization dashboards. Based on the microdata from the UNEMAT available in INEP-CENSUP, we developed a process that includes feature selection, processing procedures, automation scripts, and the creation of interactive charts. The tools were well-suited and provided high-quality visualizations. Performance was also satisfactory.*

Resumo. *O volume crescente de dados gerado pelas políticas de Governo Aberto demanda o desenvolvimento de ferramentas eficazes para tratamento e extração de informações dessas fontes massivas de dados. Este artigo propõe uma metodologia que utiliza interfaces de software acessíveis para integrar dados e gerar dashboards de visualização interativos. Com base nos microdados da UNEMAT disponíveis no INEP-CENSUP, desenvolvemos um processo que inclui a seleção de funcionalidades, procedimentos de tratamento, scripts de automação e criação de gráficos interativos. As ferramentas utilizadas demonstraram ser viáveis com qualidade de visualização e performance satisfatória.*

1. Introdução

A disponibilização de dados públicos por meio de iniciativas do Governo Aberto promove maior transparência, facilitando o controle social sobre as ações governamentais e incentivando a participação cidadã. Segundo Berberian, Mello e Camargo (2014), essa prática também possibilita a criação de novas informações e aplicações a partir dos dados divulgados. Marques (2024) destaca que os governos e instituições precisam adotar políticas e tecnologias adequadas para extrair *insights* valiosos desses dados.

Na educação, essa tendência é observada em diversos estudos. Por exemplo, Meijer (2009) demonstra que a publicação de dados sobre o desempenho de instituições de ensino na Internet tem trazido benefícios globais. No Brasil, o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) lidera a disponibilização de microdados educacionais, como os do Censo da Educação Superior (CENSUP), que fornecem dados importantes para políticas públicas e pesquisas educacionais. Ao lado do acesso a essas fontes massivas de dados, também está a necessidade de organizá-los para obtenção de informações relevantes. Nesse sentido, a visualização de dados desempenha um papel importante nesse processo, facilitando a compreensão dos dados e a identificação de padrões e tendências [Few, 2014] [Pentland, 2013].

Entretanto, criar um ambiente para visualização de dados enfrenta desafios significativos devido ao alto volume de informações e ao custo das ferramentas tradicionais de *Business Intelligence* (BI). Este estudo visou identificar e integrar algumas soluções gratuitas ou de baixo custo desenvolver um *dashboard*, com foco nos microdados da UNEMAT disponíveis no CENSUP. Esta pesquisa descreve a integração de ferramentas de código aberto, *frameworks front-end* e plataformas de desenvolvimento de código, buscando alternativas que garantam uma visualização de dados e reduzam custos operacionais para gestores e analistas de dados.

2. Metodologia

Inicialmente, foram coletados os dados do portal de dados abertos do INEP. O portal fornece microdados referentes ao CENSUP, organizados por ano desde 1995. Nesta pesquisa, trabalhamos com os dados de 2010 a 2023. O CENSUP apresenta dados consolidados sobre cursos e Instituições de Educação Superior (IES), acompanhados por uma descrição detalhada de seus metadados.

Após a coleta dos dados, o processamento, armazenamento em nuvem e apresentação dos dados são executadas em três etapas, descritas a seguir:

Etapa 1: Foram desenvolvidos *scripts* em *Python*¹ com a biblioteca *Pandas* para filtrar e manipular os microdados da UNEMAT. Entre as principais tarefas executadas, destaca-se o processo de padronização dos dados. Como resultado, os microdados já filtrados e padronizados foram concatenados por meio de um *script* para um único arquivo e convertidos do formato para *.xlsx*, para a redução do tamanho dos arquivos, otimizando o processamento. Os arquivos e *scripts* de padronização foram publicados no *GitHub*².

Etapa 2: O arquivo final foi carregado na plataforma *Google Sheets*³ para disponibilização dos dados na nuvem. Seguiu-se com a interface do *Google Sheets*, em conjunto com o *Google Apps Script*⁴, para desenvolver funções que de manipulação da planilha. Adicionalmente, foram criadas funções para integrar a planilha com aplicações externas por meio de requisições HTTP, retornando os dados no formato *JSON*. As funções desenvolvidas nesta etapa também estão disponíveis no repositório *GitHub*⁵.

Etapa 3: No *React*, o projeto incorporou bibliotecas importantes como *TailwindCSS*⁶, para estilização, e *Material-UI*⁷, que forneceu componentes baseados no *Material Design* do Google, usados para formulários de filtros para requisições. As bibliotecas *Axios*⁸ e *useQuery*⁹ foram utilizados para chamadas HTTP ao *endpoint* do projeto *Google Apps Script*, e a Biblioteca *ApexCharts*¹⁰ para gerar gráficos interativos e as visualizações dos microdados do CENSUP. O código do projeto foi armazenado no *GitHub*, facilitando o controle de versões e possíveis colaborações futuras. O projeto foi

¹ <https://www.python.org>

² <https://github.com/luanderilidio/scripts-inep-censup>

³ <https://developers.google.com/sheets>

⁴ <https://developers.google.com/apps-script>

⁵ <https://github.com/Luanderilidio/censo-unemat>

⁶ <https://tailwindcss.com/docs/>

⁷ <https://mui.com>

⁸ <https://axios-http.com/docs/intro>

⁹ <https://tanstack.com/query/latest>

¹⁰ <https://apexcharts.com>

implantado na Vercel¹¹, uma plataforma especializada em hospedagem de projetos *front-end*, que proporcionou uma implantação rápida, simplificada e gratuita, garantindo que a aplicação esteja disponível de forma estável e acessível para os usuários futuros.

3. Resultados

O painel gerado permite a exploração de informações como o número de ingressantes, matriculados, vagas e concluintes, conforme ilustrado na Figura 1. Além disso, os gráficos dessa mesma figura possibilitam a visualização das principais formas de ingresso dos estudantes. Outros elementos gráficos também podem ser explorados, com diferentes indicadores e trajetória dos dados ao longo dos anos. O painel oferece ainda diferentes tipos de filtros. O *dashboard* pode ser acessado pela url <https://censo-unemat.vercel.app/>.

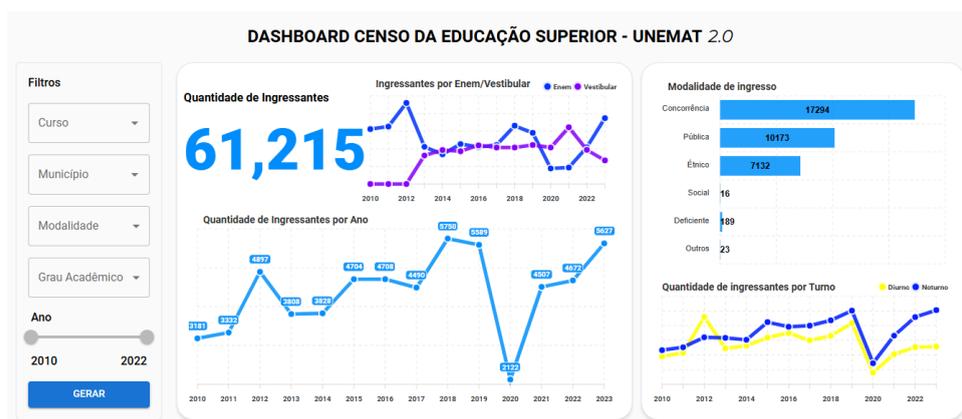


Figura 1 - Painel Dashboard inicial sobre quantitativos dos indicadores do CENSUP da UNEMAT

7. Considerações Finais

Descrevemos aqui ferramentas acessíveis que se mostraram adequadas para o desenvolvimento de um *dashboard* para visualização de grande volume de dados e mais especificamente, geração de indicadores relevantes para gestão do ensino superior. Detalhamos a metodologia aplicada e as funções para tratamento de dados que podem ser facilmente estendidas para outros domínios. Ferramentas para melhoria de performance e velocidade de geração dos gráficos estão sendo testadas.

4. Referências

- BERBERIAN, C.; MELLO, P.; CAMARGO, R. *Governo Aberto: a tecnologia contribuindo para maior aproximação entre o Estado e a Sociedade*. Revista do TCU, n. 131, p. 30–39, 1 jan. 2014.
- FEW, S. *Information Dashboard Design: Effective Visual Communication of Data*. [S.l.: s.n.].
- MARQUES, G. V. *DataHarbor: capacitando o engajamento cidadão por meio da visualização de dados abertos na governança municipal*. 2024. 72 p.
- MEIJER, A. Understanding Modern Transparency. *International Review of Administrative Sciences*, v. 75, n. 2, p. 255–269, 1 jun. 2009.

¹¹ <https://vercel.com>