

Democratização do uso de Big Data: um relato de experiência em empresa de grande porte do ramo varejista

Almir J. Kania Jr.¹, Alinne C. Corrêa Souza¹

¹Coordenação de Engenharia de Software – Universidade Tecnológica Federal do Paraná (UTFPR)
85.660-000 – Dois Vizinhos – PR – Brazil

almirkania@gmail.com, alinnesouza@utfpr.edu.br

Abstract. *With the exponential evolution of technology, especially in data, it is necessary for companies to adapt to generate competitive advantages using their available data, based on a well-defined Big Data strategy. With the shortage of professionals in the field in the job market, the objective of measuring the democratization of access and knowledge about Big Data to employees in a large retail company proves to be efficient in this experience report, with more than 200 thousand accesses in dashboards and more than 1,500 training sessions carried out in the commercial department in two years.*

Resumo. *Com a evolução exponencial da tecnologia, em especial na área de dados, se faz necessária a adaptação das empresas para gerar diferenciais competitivos com a utilização de seus dados disponíveis, com base em uma estratégia de Big Data bem definida. Com a escassez de profissionais da área no mercado, o objetivo de medir a democratização do acesso e o repasse do conhecimento sobre Big Data para os colaboradores de empresa varejista de grande porte se mostrou eficiente nesse relato de experiência, com mais de 200 mil acessos em dashboards e mais de 1.500 treinamentos realizados no departamento comercial em dois anos.*

1. Introdução

Com o avanço contínuo da tecnologia, especialmente na área de dados, os métodos de coleta, análise e transformação de informações estão sendo cada vez mais utilizados devido sua rapidez e eficiência. Esse novo cenário permitiu o desenvolvimento de um novo método de trabalho nas empresas, além de possibilitar mais inovação e velocidade nas mudanças.

Dentre os diversos tipos de empresas, principalmente as varejistas, também estão lidando com um volume cada vez maior de dados que chegam em alta velocidade. Então, tem se tornado extremamente necessário saber lidar com essa grande quantidade de informações para transformar esses dados em vantagens competitivas [McAfee and Brynjolfsson 2012]. Essas empresas enfrentam um fluxo constante de informações, incluindo dados de transações, históricos de compras, dados de estoque e dados de clientes. A capacidade de coletar, armazenar, processar e analisar esses dados é fundamental para: (i) compreensão do comportamento do cliente; (ii) otimização de estoque; (iii) personalização e recomendações de compra; e (iv) previsibilidade da demanda.

Nesse contexto, destaca-se o conceito de *Big Data*, que se refere à gestão de complexos conjuntos de dados, com o intuito de evidenciar, por meio de processos

analíticos associados, tendências e conexões com potencial de contribuir para a estratégia de negócios [Gandomi and Haider 2015]. A tecnologia de *Big Data* desempenha um papel fundamental uma vez que permite aos empresários uma melhor compreensão sobre os seus negócios, ter previsões mais precisas e tomar as melhores decisões de criação e desempenho [McAfee and Brynjolfsson 2012]. Portanto, ao utilizar *Big Data* é possível dar inteligência aos dados e traduzi-los em vantagem competitiva para o negócio.

A utilização de *Big Data* tem se tornado, exponencialmente, um diferencial competitivo para as empresas, gerando valor para o negócio por meio da Inteligência Empresarial (IE). A IE consiste em um processo que visa transformar dados em sabedoria, através de um sistema informacional que tem como objetivo melhorar a posição competitiva [Sapiro 1993]. Segundo [Rezende 2002], o desenvolvimento e a sustentação de vantagens competitivas estão diretamente ligados à identificação correta das necessidades do negócio por meio da seleção assertiva das fontes de dados e do compartilhamento constante das informações geradas para o mais alto nível de direção das empresas.

Neste contexto, é necessário ter uma gestão de dados assertiva e profissionais habilitados para gerenciar de maneira inteligente as informações obtidas e o conhecimento gerado para que seja criado um diferencial competitivo a ser incorporado pelas empresas a partir de seus processos gerenciais. O presente artigo detalha o relato de experiência da utilização de *Big Data* e treinamentos específicos como estratégia em uma empresa de varejo a fim de mensurar a efetividade da democratização do uso de *Big Data*. Espera-se com esse resultado auxiliar a compreensão e fornecer mais autonomia aos analistas de negócios para criarem suas próprias análises e, conseqüentemente, gerar diferenciais competitivos para a companhia.

O restante do artigo está estruturado da seguinte forma. Na Seção 2 são descritos alguns trabalhos relacionados ao contexto; na Seção 3 é detalhado o relato de experiência; na Seção 4 são apresentados alguns resultados relacionados à democratização de dados na companhia, e por fim, na Seção 5 são apresentadas as conclusões e trabalhos futuros.

2. Trabalhos Relacionados

Com a finalidade de obter uma melhor visão e entendimento sobre o tema, nesta seção são apresentados os principais trabalhos relacionados.

Em [Félix et al. 2018] foram identificadas oportunidades e barreiras para a adoção de estratégias de *Big Data* no varejo virtual por meio da condução de um estudo de caso na empresa Magazine Luiza. De acordo com os resultados, foi possível concluir que a mudança de cultura demandou a quebra de uma série de paradigmas, desde o gerenciamento constante da expectativa das áreas de negócio, até a falta de profissionais capacitados no mercado, não se referindo apenas às capacidades técnicas, mas também à parte comportamental e aptidão para negócios. Para lidar com esse desafio, as equipes buscaram se qualificar internamente, somando os esforços a consultorias externas.

O trabalho desenvolvido por [de Menezes 2018] visou identificar as principais inovações e benefícios relacionados ao uso de *Big Data* em uma empresa varejista de grande porte. A autora conduziu uma entrevista semiestruturada com uma das gestoras da empresa e relatou que a companhia utiliza *Big Data* para incentivar a experiência digital, trazendo inovações para os seus consumidores e oferecendo ofertas exclusivas, listas

personalizadas e prêmios. Essas soluções tecnológicas criaram uma base de dados que possibilitou analisar e prever o comportamento dos clientes da companhia. Para avaliar os benefícios dessa estratégia, foi realizada uma análise do desempenho financeiro da empresa. Os resultados indicaram que as inovações geradas resultaram em uma recuperação financeira após um ano de diversos resultados negativos para o segmento de varejo da companhia, recuperando a tendência de vendas e faturamento no ano de 2017.

Ambos os estudos de caso avaliaram a efetividade da implantação de estratégias de *Big Data* com resultados inferidos por meio de relatos dos colaboradores das empresas, registrados em entrevistas. Os trabalhos também trouxeram características intrínsecas ao processo de transformação e implantação da estratégia de *Big Data*, desafios e impactos no negócio.

Por outro lado, neste trabalho foi utilizada uma medição da quantidade de treinamentos realizados pelos analistas de negócio sobre o tema e o levantamento do uso das plataformas de dados e *dashboards*, atendendo ao objetivo de mensurar a efetividade da democratização do acesso à tecnologia de *Big Data* dentro da companhia. A dificuldade de encontrar profissionais capacitados para operar dados em massa no mercado de trabalho, levantada pelos estudos de caso, tiveram soluções apresentadas nesse relato de experiência, como a habilitação de analistas de negócios para criarem análises e produtos de dados de forma autônoma, democratizando o acesso às informações.

3. Relato de Experiência

Este relato visa apresentar a experiência vivenciada com a utilização de *Big Data* e aplicação de treinamentos sobre o tema em uma varejista de grande porte no Brasil, a fim de mensurar a efetividade da democratização do seu uso. A empresa foi fundada em 1977, em Curitiba, mantendo sua sede administrativa na região metropolitana até hoje. Atua em mais de 1.700 cidades em todos os estados brasileiros e no Distrito Federal, além de mais de 50 países com lojas físicas e operações de mercado eletrônico. No varejo, a empresa opera com os formatos de venda direta, hipermercados, farmácias e lojas próprias e franqueadas.

3.1. Descrição das plataformas e dos métodos utilizados

Na empresa varejista é utilizada a plataforma *Google Cloud Platform (GCP)*¹ para o armazenamento dos dados e das aplicações da empresa. A empresa realizou a implantação dessa plataforma a fim de centralizar a quantidade gigantesca de dados gerados diariamente pelos seus sistemas e democratizar o acesso a eles por meio de uma gestão facilitada de perfis e controle do uso de dados sensíveis.

Para que os usuários das diferentes áreas da empresa tivessem capacidade de acessar esse repositório de maneira inteligente, utilizando de boas práticas e ferramentas pré-definidas pelo time de Governança de Dados, foram desenvolvidos e disponibilizados de maneira ampla uma série de treinamentos técnicos com foco em habilitar os analistas de negócio a acessar e prover diferenciais competitivos para a companhia por meio de análises e criação de *dashboards* para divulgação de seus resultados. Esses treinamen-

¹Consiste em uma suíte de computação em nuvem com ferramentas de gerenciamento modulares criada pelo Google em 2008.

tos foram publicados em uma plataforma interna de cursos, que foi desenvolvida com o auxílio de um fornecedor terceirizado.

Tanto a plataforma de dados como a solução onde os treinamentos são disponibilizados possuem um módulo de análise integrado onde é possível acompanhar a adesão, a quantidade de acessos e detalhes sobre os produtos de dados desenvolvidos. Assim, foi possível realizar a medição dos ganhos de tal estratégia na área comercial da empresa, que é o escopo deste artigo.

A coleta de dados foi realizada diretamente dessas plataformas, consultando o histórico de janeiro de 2022 a outubro de 2023, utilizando como variáveis a quantidade de treinamentos realizados pelos analistas de negócio, a quantidade de páginas de *dashboards* visualizadas e a quantidade de usuários que acessaram tanto os *dashboards* como a plataforma de dados em si. Médias simples e análises gráficas foram utilizadas para ajudar na visualização e entendimento dos resultados.

Os *dashboards* que tiveram suas visualizações medidas foram desenvolvidos na ferramenta *Google Data Studio* e possuem as mais variadas aberturas e temas com foco no varejo e na operação comercial, desde a análise detalhada das vendas realizadas e suas previsões, controle de estoque, atingimento de metas e comportamento de clientes. Por fim, foram realizadas entrevistas semiestruturadas junto ao time do departamento comercial e também de governança de dados com o intuito de coletar dados qualitativos a respeito do uso das ferramentas, ganhos intangíveis e contextualização sobre a estratégia da empresa em si.

3.2. Etapas

A estratégia desenvolvida em relação ao uso de *Big Data* na empresa foi dividida em cinco etapas: (1) criação da governança de dados; (2) disponibilização dos acessos e das ferramentas necessárias para consumo e análise dos dados; (3) produtização de dados estratégicos para a companhia na plataforma de *Big Data*; (4) realização de treinamentos necessários para que os colaboradores de diversas áreas consigam consumir as informações e desenvolver análises e produtos de dados com impacto no negócio; e (5) implantação de produtos de dados dentro das áreas e medição contínua da adesão à estratégia desenvolvida.

A **Etapa 1** (Criação da Governança de Dados) visa compartilhar boas práticas, gerenciar documentações sobre os dados disponíveis e controlar o acesso a dados sensíveis. Assim, foram criados processos para garantir a governança de dados da companhia, presentes no GCP. A segmentação de ambientes e perfis de desenvolvimento, a popularização dos meios de comunicação sobre dados e a massificação do catálogo de dados foram realizados com o intuito de dar suporte à estratégia de *Big Data* planejada pela empresa a partir do ano de 2021.

Na **Etapa 2** (Disponibilização dos acessos e das ferramentas necessárias), para que fosse possível utilizar o GCP com todo o seu potencial, foram disponibilizados os acessos aos analistas de negócio à várias ferramentas do *Google Workspace* como o *Google Big Query* (serviço que suporta consultas usando a linguagem *Structured Query Language - SQL*), *Google Data Studio* (ferramenta on-line que permite a criação de relatórios e painéis informativos personalizáveis a partir de banco de dados) e *Google AppSheet* (aplicativo que fornece uma plataforma de desenvolvimento de *software* sem código), pos-

sibilitando a coleta, transformação e a análise dos dados disponíveis no GCP. Além disso, foi criada uma estrutura para solicitação de leitura às tabelas e ambientes disponíveis dentro da plataforma, criando uma gestão efetiva de acessos.

Na **Etapa 3** (Produtização de dados estratégicos), para que os dados estivessem disponíveis na plataforma, desde a implantação do GCP no ano de 2021, o time de Engenharia de Dados do departamento de Tecnologia da Informação (TI) foi responsável por migrar os dados de sistemas e bancos de dados legados para dentro do GCP. Também foram realizados refinamentos e transformações em cima dos dados brutos dos sistemas com o intuito de facilitar a leitura e a democratização das informações nas demais áreas da companhia.

Na **Etapa 4** (Realização de treinamentos) foram desenvolvidos treinamentos de cunho técnico com foco em dados e nas ferramentas disponibilizadas, desde a sua abordagem no nível ético e teórico até o detalhamento das ferramentas disponíveis para extração, transformação, carregamento, análise e visualização de dados. Esses treinamentos foram disponibilizados por meio de conteúdo multimídia (vídeos e apresentações) a partir do ano de 2022 em ferramenta interna de forma ampla, com uma forte comunicação institucional e metas ambiciosas de adesão de colaboradores, que passam por uma avaliação para validação dos conhecimentos adquiridos, recebendo um certificado de conclusão para cada módulo. Para acompanhar essa adesão, foram criados indicadores e *dashboards* de controle com o intuito de reconhecer as áreas com mais treinamentos realizados e medir a efetividade e o retorno do investimento realizado.

Por fim, a **última etapa** (Etapa 5 - Implantação de produtos de dados) consiste na implantação dentro de cada área de negócio, já que a governança de dados foi criada, os colaboradores treinados e o acesso facilitado à plataforma de dados e suas ferramentas de análise. Focando exclusivamente no departamento comercial de lojas próprias da empresa varejista, a primeira onda de democratização dos dados se deu durante o ano de 2022 pelo desenvolvimento de *dashboards* por um time mais especializado dentro do departamento, facilitando, assim, o compartilhamento das informações para o time operacional e administrativo das vendas realizadas, controles de estoque, atingimento de metas e comportamento de clientes.

Posteriormente, dentro do próprio departamento, estimulou-se a adesão aos treinamentos disponibilizados e encorajamento ao uso cada vez mais frequente das ferramentas com o intuito de possibilitar mais autonomia aos colaboradores para fazerem suas próprias análises e também saberem como aproveitar melhor os *insights* disponíveis nos *dashboards* já criados. Todo esse movimento de fortalecimento da cultura de dados dentro do departamento pôde ser medido pelo time especializado em dados no departamento por meio dos *logs* de consumo em *bytes* dentro da própria plataforma de dados, pela visualização de páginas de *dashboards* e pela adesão aos treinamentos disponibilizados para toda a empresa.

4. Resultados

Segundo relatado em entrevista semiestruturada, os principais frutos da estratégia de *Big Data* na área comercial da empresa se deram pela mudança de comportamento desde a alta liderança do departamento, passando pelos times que acompanham a performance de processos e chegando até o time de campo que está em contato constante com o cliente final,

realizando a venda dos produtos. Ao consumirem uma massa de dados cada vez maior com uma velocidade incomparável aos métodos anteriores, a análise de informações contidas nos sistemas de ponto de venda se tornou um diferencial estratégico ainda maior, facilitando as tomadas de decisão com impacto no negócio.

Ações de melhoria puderam ser iniciadas, bem como a revisão e ajustes de processos do departamento. Com uma maior disponibilidade de dados, pôde-se realizar análises mais precisas acerca da qualidade do estoque das filiais, aumentar a previsibilidade da demanda com base em um histórico detalhado de vendas e no comportamento dos clientes em momentos chave como em datas festivas e períodos promocionais, medir ações de incentivo às vendas, entre outros. Todas essas análises puderam ser desenvolvidas com uma alta granularidade, segmentando as regiões territoriais, a árvore de produtos e a clusterização de clientes.

A quantidade de *dashboards* criados pelos especialistas de dados do departamento cresceram exponencialmente ao tempo em que as tabelas de dados eram produtizadas na plataforma oficial da empresa pela área de TI e, conseqüentemente, a adesão a esses produtos de dados foram aumentando. No período de 1º de junho de 2022 a 20 de outubro de 2023 foram registradas mais de 200 mil visualizações em *dashboards* (Figura 1), para mais de 6 mil usuários diferentes, alocados nos 27 estados brasileiros (Figura 2). Em média, os usuários permanecem por aproximadamente 2 minutos em cada página e consultam 3,7 páginas em cada sessão.

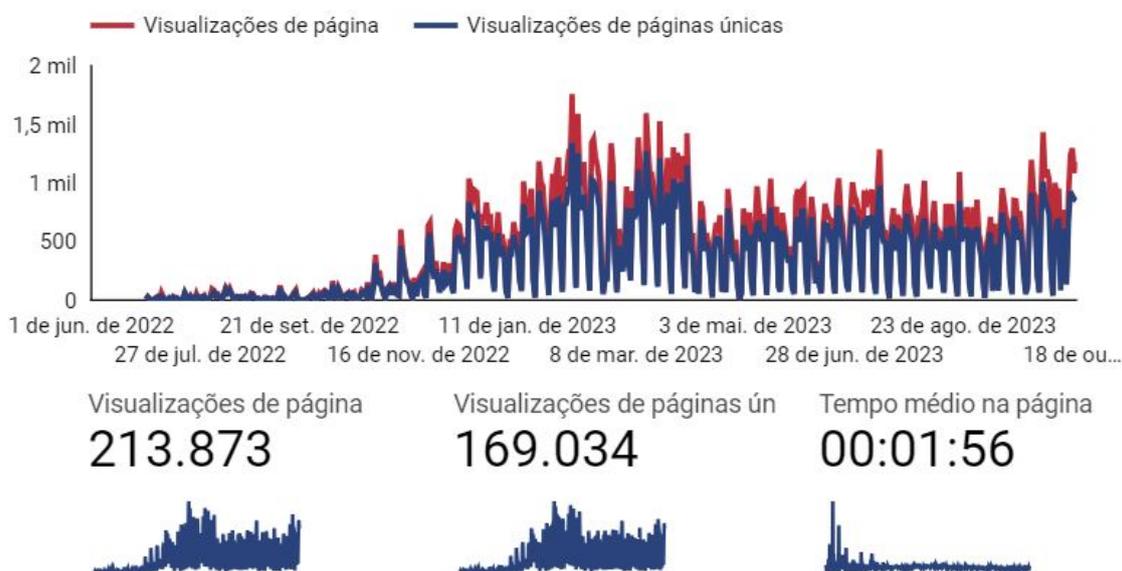


Figura 1. Visualizações de páginas de Dashboards registradas

Contudo, um dos desafios de longo prazo da estratégia de *Big Data* é fazer com que os colaboradores não sejam apenas usuários das informações criadas pelos especialistas de dados mas também desenvolvedores de análises, tabelas, *dashboards* e produtos de dados em geral. A adesão aos treinamentos citados anteriormente vem crescendo desde o ano de 2022, com números totais de mais de 1.500 cursos concluídos para 430 colaboradores diferentes. Tal ação permitiu uma maior autonomia no acesso e análise de dados

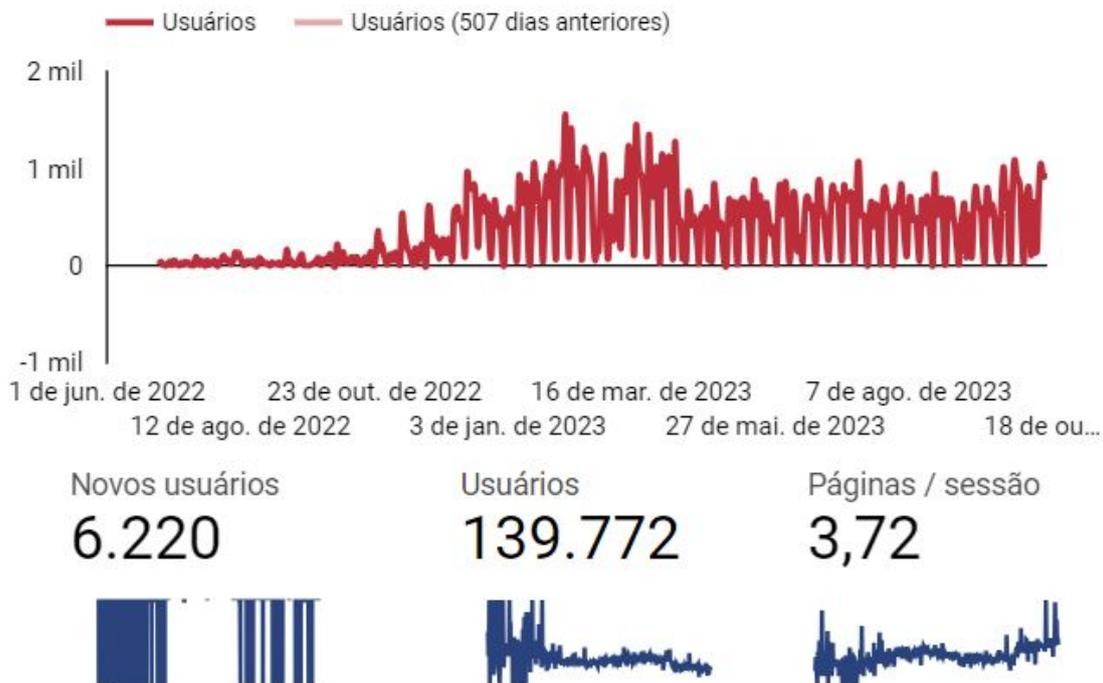


Figura 2. Quantidade de usuários impactados pelos Dashboards

pelos times de negócio, em especial nos times que realizam a avaliação da performance dos canais de venda e da execução das atividades de acordo com os processos e projetos implantados.

Analisando o gráfico temporal de treinamentos realizados (Figura 3), observa-se que a adesão teve um pico no mês de janeiro de 2023, graças às fortes ações de comunicação institucional e departamental e o incentivo por meio de premiações aos colaboradores que finalizassem os treinamentos, dada a importância do uso da plataforma de dados e das ferramentas atreladas à ela como diferencial competitivo para a empresa. Esse pico também pode ser correlacionado com as Figuras 1 e 2, dado que os maiores volumes de visualizações registradas se deram no mesmo período, mantendo-se com um índice estável logo após, reforçando que os *dashboards* não foram acessados somente no período de treinamentos. Um dos grandes desafios é aumentar a adesão e constância da realização dos treinamentos, dada a rotatividade natural de colaboradores da área e ao perfil mais voltado ao varejo que o departamento possui.

Nas áreas mais voltadas para a administração de vendas, controle de processos, projetos e previsão de demanda, onde a maioria das análises críticas são realizadas, mais de 70 por cento dos colaboradores realizaram ao menos um treinamento voltado à estratégia de *Big Data* da companhia. Incluindo o time de vendas, que realiza o atendimento dos clientes em campo, esse número cai para 12 por cento, evidenciando a oportunidade nesses cargos.

Dado esse contexto, o consumo de dados tanto pelos acessos em *dashboards* como pela construção de análises utilizando a plataforma oficial de dados da empresa cresceu

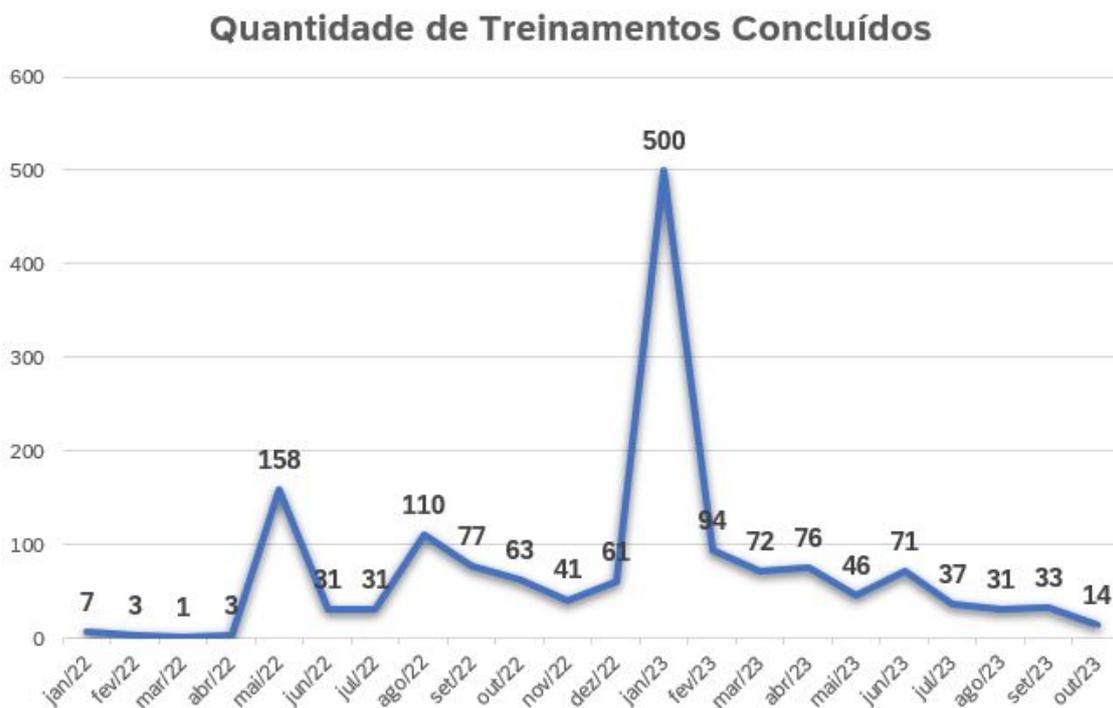


Figura 3. Quantidade de treinamentos realizados por mês

exponencialmente no departamento, chegando a uma média de 180 *terabytes* por mês no segundo semestre de 2023, o dobro da média do primeiro semestre do mesmo ano e quase três vezes maior que a média do ano anterior.

Além da área de *Big Data*, essa estratégia organizacional descentralizada, dando autonomia aos times de negócio também foi aplicada ao desenvolvimento de *softwares* na empresa. Em entrevista semiestruturada junto ao time de cultura organizacional, foi relatada a ferramenta *Google AppSheet*², conforme apresentada na Figura 4. Essa ferramenta consiste em uma solução *low-code* que permite aos colaboradores fazerem seus próprios aplicativos de forma online e colaborativa, ainda que não saibam desenvolver ou não tenham contato com qualquer linguagem de programação até o momento. A ferramenta possibilita uma jornada completa, desde a criação do banco de dados a ser utilizado, passando pelas ações a serem executadas no aplicativo e a customização de toda a experiência do usuário.

Foram disponibilizados treinamentos específicos para essa ferramenta e já são mais de 100 soluções criadas na empresa como um todo para resolver problemas do dia a dia nas áreas de negócio. Assim, nesses casos, o time de tecnologia assume um papel mais consultivo, fazendo com que cada departamento consiga criar soluções de baixa complexidade de desenvolvimento de maneira rápida e eficiente, tornando os colaboradores mais produtivos e tecnológicos.

²Disponível em <https://about.appsheet.com/home/>

A smarter way to work, built by you.

Build apps

Build powerful mobile and desktop apps, customized to your business needs.

-  Capture rich data using forms, barcodes, locations, signatures, and photo
-  Customize features, UX, security, and branding
-  Manage users with role-based security, authentication, and usage reports

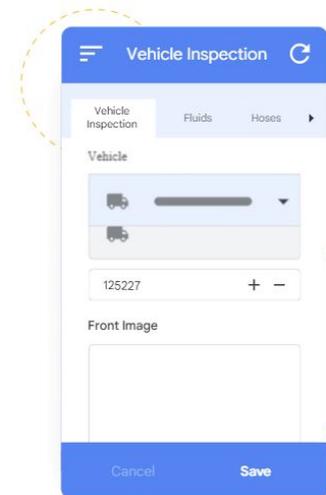


Figura 4. Apresentação da Ferramenta *AppSheet*

5. Considerações finais

Este trabalho relatou a experiência da democratização do acesso e conhecimento relacionado à análise de dados no departamento comercial de uma empresa varejista de grande porte. Dado o mercado escasso de profissionais com *background* de negócio que também tenham conhecimento na área computacional, a solução de formar colaboradores com esse perfil dentro da própria empresa para que seja possível atender às demandas específicas com maior rapidez se mostrou eficiente dado o aumento gradativo da criação e acessos aos *dashboards* com informações estratégicas, no aumento do consumo da plataforma oficial de dados da empresa e na quantidade representativa de treinamentos técnicos realizados pelos colaboradores da área de negócio.

Tal ação permitiu uma mudança geral de comportamento ao utilizar os dados para evidenciar ganhos em automatizações e melhoria de processos. Assim, vislumbra-se para os próximos anos a realização de treinamentos mais específicos, com escopo exclusivo para o departamento, focando nos tipos de dados, ferramentas e variáveis exclusivas da área.

Para trabalhos futuros, seria possível detalhar atividades e ganhos específicos atrelados aos processos de vendas, gestão de estoques e comportamento de clientes. Com a evolução constante de ferramentas *low-code* na empresa, também seria possível evidenciar diferenciais estratégicos dentro da companhia, detalhando algumas das mais de 100 soluções em *Google AppSheet* já criadas, correlacionando e/ou comparando com outras ferramentas de Inteligência Artificial com desenvolvimento facilitado.

Referências

de Menezes, J. S. (2018). Big data como estratégia de inovação em uma empresa de varejo.

- Félix, B. M., Tavares, E., and Cavalcante, N. W. F. (2018). Fatores críticos de sucesso para adoção de big data no varejo virtual: estudo de caso do magazine luiza. pages 112–126. *Revista Brasileira de Gestão e Negócios*.
- Gandomi, A. and Haider, M. (2015). Beyond the hype: Big data concepts, methods and analytics. pages 137–144. *International Journal of Information Management*.
- McAfee, A. and Brynjolfsson, E. (2012). Big data: The management revolution. pages 61–68. *Harvard Business Review*.
- Rezende, Y. (2002). Informação para negócios: os novos agentes do conhecimento e a gestão do capital intelectual. pages 120–128. *Ci. Inf., Brasília*.
- Sapiro, A. (1993). Inteligência empresarial: a revolução informacional da ação competitiva. pages 106–124. *Revista de Administração de Empresas*.