

GERAÇÃO DE VALOR COM BIG DATA E A INFLUÊNCIA DA CULTURA DE DADOS NA TOMADA DE DECISÃO EM GRANDES EMPRESAS

GENERATING VALUE WITH BIG DATA AND THE INFLUENCE OF DATA CULTURE ON DECISION-MAKING IN LARGE COMPANIES

¹Henrique Hiluey Roriz PERIN.
²Guilherme GUILHERMINO NETO.
³Luiz Henrique Lima FARIA.
⁴Tiago José Menezes GONÇALVES.

¹Instituto Federal do Espírito Santo - IFES. E-mail: (hiluey.henri@gmail.com).
²Instituto Federal do Espírito Santo - IFES. E-mail: (guilherme.neto@ifes.edu.br).
³Instituto Federal do Espírito Santo - IFES. E-mail: (luizlima@ifes.edu.br).
⁴Instituto Federal do Espírito Santo - IFES. E-mail: (tiago.goncalves@ifes.edu.br).

Artigo submetido em 18/11/2024 e aceito em 16/12/2024.

Resumo

A presente pesquisa investigou como as organizações enfrentam o desafio de gerar valor a partir da complexidade do big data e analisou se empresas com uma cultura de dados consolidada apresentam maior eficiência no processo de tomada de decisão. Para alcançar esses objetivos, aplicou-se um questionário com 8 itens a 60 empresas selecionadas com base em seu modelo de estruturação de dados, divididas em dois grupos: Organizações Avançadas em Dados (OAD), que adotam práticas *data-driven*, e Organizações Iniciais em Dados (OID), que ainda não viabilizam essa prática. A análise utilizou representações gráficas para identificar desafios, benefícios e percepções sobre a governança e estratégia de dados. Testes estatísticos de proporcionalidade, como Qui-quadrado e Fisher, validaram se as OAD utilizam mais dados nos processos decisórios e se a capacidade de geração de valor está associada à maturidade técnica e analítica da empresa. Os resultados indicam que a consolidação de uma cultura de dados é fundamental para estabelecer decisões mais eficientes e atingir maior geração de valor com big data. Apesar disso, este estudo apresentou limitações, como amostra reduzida e ausência de relação com métricas financeiras. Como recomendação para futuras pesquisas, indicou-se ampliar a diversidade da amostra, incluir indicadores como ROI e faturamento, e explorar as diferenças entre decisões operacionais e estratégicas.

Palavras-chave:

Big data, Análise de big data, Cultura de dados, Data-driven, Tomada de decisão.

Abstract

This research investigated how organizations address the challenge of generating value from the complexity of big data and analyzed whether companies with a consolidated data culture demonstrate greater efficiency in decision-making processes. To achieve these objectives, a 8-item questionnaire was applied to 60 companies selected based on their data structuring models, divided into two groups: Data-

Advanced Organizations (OAD), which adopt data-driven practices, and Data-Beginner Organizations (OID), which have not yet implemented such practices. The analysis utilized graphical representations to identify challenges, benefits, and perceptions regarding data governance and strategy. Statistical proportionality tests, such as Chi-square and Fisher's exact test, validated whether OADs rely more on data in decision-making processes and whether their ability to generate value is associated with their technical and analytical maturity. The results indicate that establishing a data culture is essential for achieving more efficient decisions and greater value generation from big data. However, the research presented limitations, such as a small sample size and the lack of financial metrics. Future studies are recommended to expand sample diversity, include indicators such as ROI and revenue, and explore differences between operational and strategic decision-making processes.

Keywords:

Big data, Big data analytics, Data culture, Data-driven, Decision-making.

1 INTRODUÇÃO

A era moderna é marcada pela grande quantidade de dados circulantes nos meios digitais. A evolução tecnológica global tornou possível o compartilhamento de informações em velocidade instantânea. Prevê-se que o volume total de dados criados e consumidos em todo o mundo seja de 149 *zettabyte* (ZB) em 2024, quase duas ordens de grandeza acima de 2013 (Chauhan; Sood, 2021). O crescimento desenfreado dos dados enquadra-se no conceito de *Big Data* (BD), descrito em Sharda et al. (2019) por “6Vs”: volume, variedade, velocidade, veracidade, variabilidade e valor. O BD abrange tanto os dados estruturados quanto os não estruturados. Os conjuntos de dados estruturados são constituídos por tipos de dados definidos, facilmente organizados em bancos de dados relacionais. Os dados não-estruturados são aqueles que não são facilmente armazenados em formato relacional (Chauhan; Sood, 2021). Estas quantidades de dados exigem métodos complexos para a sua manipulação, bem como abordagens estatísticas intensivas e modelos de programação para extrair informações relevantes (Misra et al., 2022).

Fenômenos recentes como a Internet das coisas (IoT - *Internet of Things*) e a evolução da Inteligência artificial (IA) refletem a ascensão da geração de valor através do *big data*. IoT pode ser definida como a capacidade de permitir uma comunicação robusta entre o mundo físico e o mundo digital a partir de um sistema condicionado por aprendizado de máquina e processamento de dados (Misra et al., 2022). Para Dwivedi et al. (2021), o campo da IA envolve o desenvolvimento de sistemas computacionais capazes de realizar tarefas que normalmente

requerem a inteligência humana, como a percepção sensorial e a tomada de decisão. A análise de *big data* (BDA - *Big Data Analytics*) combinado com a IA, tem o potencial de transformar diferentes áreas da manufatura, da saúde e da Inteligência de negócios (BI - *Business Intelligence*), oferecendo insights avançados num contexto preditivo (Alvarez et al., 2018). Empresas que aplicam eficazmente a análise de *big data* evidenciaram aumentos de 70% na expansão de novos mercados, elevação de 79% na satisfação dos clientes, e melhoria de 76% nas vendas (Akter et al., 2016).

Entretanto, Paradza e Daramola (2021), mostram que a real vantagem da análise de decisão dos negócios está, na verdade, atrelada ao quanto de valor as empresas conseguem extrair dos seus dados. Relatórios publicados pela Forrester em 2020 sugerem que entre 60 a 73% de todas as informações nunca são usadas para fins analíticos (Gualtieri, 2016) e, atualmente, cerca de 0.5% de todos os dados são analisados corretamente pelas empresas (Myler, 2017). A partir de estimativas, entre 80 a 85% dos dados organizacionais encontram-se em algum formato não estruturado (Sharda et al., 2019), um formato inacessível para propósitos analíticos. Nesse sentido, volume e variedade de dados obtidos nem sempre representam ganhos na perspectiva da tomada de decisão.

À medida que as organizações avançam cada vez mais ao *big data*, a expansão heterogênea dos dados, tanto em termos de quantidade como de variedade, exigirá a evolução da infraestrutura e da análise de dados para processá-los (Chauhan; Sood, 2021). Um relatório publicado pela NewVantage Partners em 2022 revela que 91% das empresas entrevistadas aumentaram o seu investimento em iniciativas de dados, 26% conseguiram estabelecer uma cultura de dados, porém apenas 19% efetivamente a incorporaram em suas rotinas (Davenport; Bean, 2022).

O presente estudo, a partir do entendimento de que em muitos casos as organizações são ineficientes na geração de informação útil com os dados, propõe fornecer a resposta para a seguinte lacuna: “Como as organizações podem enfrentar os desafios associados à extração de valor do *big data*, e de que maneira a cultura de dados influencia o sucesso na geração de ganhos e na melhoria do processo de tomada de decisão?”. Para que essa questão de pesquisa seja respondida, o seguinte objetivo geral deve ser alcançado: Analisar como as organizações lidam

com o desafio de extrair valor do amplo volume de dados disponíveis e verificar se a cultura de dados é um fator de sucesso para uma empresa atingir ganhos e melhorar sua tomada de decisão.

Este artigo é composto por cinco tópicos: 1 – Introdução. 2 – Referencial Teórico. 3 – Metodologia. 4 – Análise de Dados. 5 – Conclusão e Considerações Finais.

2 REFERENCIAL TEÓRICO

Este tópico tem por finalidade apresentar a conceituação basilar do estudo. Está dividido em duas temáticas. Dessa forma, foi possível tratar das especificidades do BD, bem como do processo decisório baseado em dados. São sob essas duas bases teóricas que os resultados analisados por este estudo produziram suas conclusões e considerações finais.

2.1 BIG DATA

De acordo com Sharda et al. (2019), BD tornou-se um termo subjetivo. Isso porque “Big” refere-se a uma das premissas do BD que é apresentar alto volume mas não existe uma delimitação quantitativa do tamanho em GB ou TB que se estabelece o BD. O termo “*big data*” surgiu no ano de 1997 quando Michael Coxand e David Ellsworth, investigadores da Nasa, o utilizaram pela primeira vez em um estudo apresentado na conferência do Instituto de Engenheiros Eletricistas e Eletrônicos (IEEE) para explicar grandes conjuntos de dados (Wang; Kung; Byrd, 2018). Segundo Oussous et al. (2017), BD pode ser definido como grandes conjuntos de dados em crescimento, que incluem formatos heterogêneos: estruturados, não estruturados e semiestruturados. O BD é caracterizado por 6 variáveis descritivas a seguir: volume, variedade, velocidade, veracidade, variabilidade e valor (Sharda et al., 2019).

As tecnologias do BD são classificadas em diferentes níveis, incluindo: coleta, processamento, armazenamento e análise. Muitas vezes, a quantidade de informações ultrapassa a capacidade de aproveitamento (Oussous et al., 2017). Nesse aspecto, o processamento do BD difere do processamento de dados tradicionais. Em ambientes tradicionais, a abordagem envolve inicialmente a exploração dos dados, seguida pela concepção do modelo de processamento e pela criação da estrutura do banco de dados que vai armazenar os mesmos. Já o fluxo de arquitetura do BD se inicia com a coleta de dados provenientes de fontes internas ou externas, os quais são então armazenados na chamada “zona de destino”, uma

infraestrutura capaz de lidar com o volume e a variedade dos dados, frequentemente implementada por meio de um sistema de arquivos distribuídos. Após o armazenamento dos dados, as transformações são aplicadas para garantir sua eficiência e escalabilidade (Berisha; Meziu; Shabani, 2022).

As formas de BD compreendem dados em lote ou dados de fluxo dinâmico. Os dados em lote são armazenados de forma estática, enquanto os dados de fluxo são uma sequência contínua de filas de dados em tempo real. A fase de coleta e processamento fundamentam-se nos pilares do ETL (Extração, Transformação e Carregamento) (Berisha; Meziu; Shabani, 2022). O processo de ETL consiste em extrair os dados das fontes e carregá-los em destinos de armazenamento. Os dados passam por transformações, incluindo a limpeza e sua normalização, de acordo com o critério do esquema de destino. Os dados são carregados em um banco de dados operacional, também chamado de *data warehouse* (Bansal; Kagemann, 2015). No fluxo dinâmico, utilizam-se ferramentas que possam garantir a estabilidade do fluxo (Wang et al., 2020).

Em 2004, o Google introduziu o *MapReduce*, um modelo de programação distribuída amplamente utilizado para o processamento de grandes conjuntos de dados. O *MapReduce* possibilita aos usuários elaborar algoritmos complexos em conjuntos de dados extensos sem se preocupar com sincronização, tolerância a falhas ou disponibilidade (Wang et al., 2020). Um programa baseado em computação distribuída é um sistema de armazenamento que distribui os dados entre vários nós de uma rede, garantindo redundância e escalabilidade, permitindo o processamento eficiente de grandes volumes de dados (Sharda et al., 2019). O *Hadoop*, por sua vez, é um outro programa de arquivos distribuídos utilizado no processamento de dados, principalmente os não estruturados (Oussous et al., 2017). O sistema, composto pelo *Hadoop Distributed File System* (HDFS), é capaz de armazenar grandes volumes de dados não estruturados e semiestruturados, uma vez que não exige o formato relacional. Devido a essa característica, também é aplicável a fluxos de dados dinâmicos (Sharda et al., 2019).

De acordo com Wang et al. (2020), nos últimos anos houve uma mudança significativa nas técnicas de armazenamento de dados devido ao crescimento do BD. Entre as opções adotadas, destacam-se o NoSQL, NewSQL e outras linguagens de bancos de dados. Bancos de dados NoSQL podem lidar com dados não estruturados com alta disponibilidade e durabilidade

(Chauhan; Sood, 2021). Os bancos de dados relacionais oferecem armazenamento de tabelas que podem ser combinadas. Essas tabelas possuem um esquema fixo, usam Structured Query Language (SQL) para gerenciar dados e oferecem suporte a garantias de Atomicidade, Consistência, Isolamento e Durabilidade (ACID) (Microsoft, 2022). Os bancos de dados NoSQL normalmente não fornecem garantias ACID e referem-se a armazenamentos de dados não relacionais de alto desempenho. Já o NewSQL pode ser considerado uma tecnologia emergente de banco de dados relacional com escalabilidade distribuída do NoSQL, serviço SQL e garantia de transações ACID (Wang et al., 2020). Essas tecnologias podem ser aplicadas a sistemas baseados na computação em nuvem.

A evolução do BD se intensificou ainda mais nos últimos anos com o advento da computação em nuvem (Dwivedi et al., 2021). A computação em nuvem facilitou o armazenamento e o processamento de dados, fornecendo armazenamento quase ilimitado e facilidade de acesso a servidores de maior performance (Berisha; Meziu; Shabani, 2022). Diante das vantagens oferecidas pela computação em nuvem como a elasticidade, modelo de pagamento sob demanda e baixo investimento inicial, tornou-se uma opção viável para empresas armazenarem seus dados, principalmente diante da imprevisibilidade do volume de dados no BD (Berisha; Meziu; Shabani, 2022).

A computação em nuvem facilitou o desenvolvimento do *big data analytics*, o qual é uma metodologia criada para analisar os dados gerados no BD. O BDA também simplifica a coleta de dados operacionais e econômicos à medida que o escopo dos dados mais utilizados nas análises é gradualmente esclarecido (Mirarab; Mirtaheri; Asghari, 2019). O *big data analytics* surgiu como um dos fatores mais importantes na geração de insights para a tomada de decisões (Awan et al., 2021). Nesse sentido, destacam-se quatro tipos de análises principais: descritivas, diagnósticas, preditivas e prescritivas. Uma análise conduzida por Rialti et al. (2019) aponta que os efeitos positivos do BDA só existem ao se trabalhar com análises confiáveis. Foi observado que os analistas também podem se opor a usar uma tecnologia corretamente quando não são capazes de compreender os benefícios de sua aplicação (Rialti et al., 2019).

2.2 TOMADA DE DECISÃO

Mintzberg, Raisinghani e Théorét (1976) definem a decisão como um desenvolvimento do raciocínio com foco na escolha entre diferentes cursos de ação, ou seja, uma escolha entre alternativas. Na perspectiva de Nutt (1984), o processo decisório é constituído por um conjunto de atividades que começa com a identificação de um problema e termina com a ação. Essencialmente, para haver uma decisão nesse caso é preciso existir um problema.

Herbert Simon (1977) oferece uma visão sobre como os seres humanos encaram o desafio de tomar decisões em um mundo complexo e cheio de incertezas. Essa teoria introduziu o conceito de "racionalidade limitada", o qual reconhece as restrições inerentes às capacidades cognitivas e ao ambiente em constante mudança que influenciam as escolhas. De acordo com Shimizu (2001), é comum que as organizações se deparem com desafios na tomada de decisão. Esses desafios se caracterizam por sua amplitude e complexidade, envolvendo riscos e incertezas.

Mintzberg, Raisinghani e Theorêt (1976) realizaram um estudo de campo que consistiu na observação de vinte e cinco processos de tomada de decisão “não estruturada”. Este estudo deu origem à organização de um processo de tomada de decisão estratégica que descreve o processo decisório em situações de alta incerteza. Nas organizações, quando surgem obstáculos, é comum reconsiderar alternativas e até retroceder, tornando o processo decisório não linear e imprevisível. O modelo de processo para tomada de decisão de Mintzberg, Raisinghani e Theorêt (1976) descreve as seguintes etapas:

1. Identificação: Esta fase envolve a identificação de problemas a serem resolvidos ou oportunidades em uma determinada situação.
2. Desenvolvimento: Nesta fase, são criadas alternativas de solução para o problema.
3. Seleção: Nesta fase, um curso de ação é escolhido entre as alternativas desenvolvidas na fase de Desenvolvimento.
4. Avaliação e Escolha: as soluções são avaliadas para que seja selecionada a mais adequada. Pode-se utilizar três modos de avaliação e escolha:
 - 4.1 Julgamento: decisão tomada individualmente, usando procedimentos que às vezes não podem ser explicados.

4.2 Barganha: escolha feita por um grupo de indivíduos, cada um com suas próprias metas e julgamentos.

4.3 Análise: avaliação das alternativas por um técnico ou especialista, baseado em métodos matemáticos e dados. Normalmente resulta em um relatório comparativo das alternativas, que é fornecido aos responsáveis pela decisão para escolher um curso de ação.

5. Autorização: quando o decisor escolhe um curso de ação.

Ao entender os possíveis caminhos de avaliação e escolha do processo decisório proposto por Mintzberg, Raisinghani e Theorêt (1976), o presente estudo procura estabelecer o critério de avaliação da tomada de decisão sob a ótica analítica (explicado no tópico 4.3), devido a sua natureza racional e não intuitiva.

A melhor estratégia para a avaliação de uma decisão é o seu julgamento com base em dados. O processo analítico deve ser pautado em indicadores mensuráveis, na identificação de padrões e na busca por insights potenciais, visando desenvolver decisões estratégicas que maximizem o retorno de valor, principalmente no contexto do BD (Mirarab; Mirtaheri; Asghari, 2019). “Baseado em dados” (DD - *data-driven*) é a ênfase dada ao uso de dados como parte principal do processo decisório de uma organização (PUCPR, 2024). O *data-driven* surgiu com o advento da Indústria 4.0, a qual refere-se à transformação digital de diversos setores dos mercados industriais. Ela representa a chamada quarta revolução industrial na manufatura. A indústria 4.0 visa aumentar a eficiência na produção industrial a partir da conexão inteligente de processos e sistemas digitais em áreas como a logística e a cadeia de suprimentos (Misra et al., 2022). O *data driven* emerge nesse cenário mediante a alta combinação de dados provenientes das tecnologias IoT, inteligência artificial e o próprio *big data* (Dwivedi et al., 2021).

Diante da compreensão dos benefícios gerados pelo DD, algumas organizações adotaram a metodologia como uma base para a inteligência de negócio (BI). Isso porque as informações provenientes de experimentações com dados são valorizadas acima de opiniões. Dessa forma, assegura-se decisões holísticas a partir da análise do passado, da compreensão do presente e da predição do futuro (PUCPR, 2024).

Nesse sentido, a Tomada de Decisão Baseada em Dados (TDBD) se caracteriza pelo uso de dados para orientar decisões de negócios. A TDBD se inicia a partir da definição de um objetivo, o que envolve compreender a meta a ser buscada pela organização com uma decisão. Esse método orienta a escolha de indicadores-chave de desempenho (KPIs) e as métricas que serão analisadas para apoiar esses objetivos (Tableau, 2024). No contexto do big data, o grande desafio da TDBD reside no mapeamento das informações que de fato são relevantes para o negócio. Converter os problemas, metas e objetivos corporativos em dados torna-se complexo devido à vasta gama de possíveis escolhas disponíveis para cada situação (Paradza; Daramola, 2021). Esse processo de conversão só se torna viável com um nível adequado de centralização e normalização dos dados dentro das empresas. Com o ambiente estruturado, ainda assim os dados provenientes só terão valor se os seus consumidores estiverem preparados para recebê-los e entendido a importância de pautar suas decisões em dados (Rialti et al., 2018). Quando esse cenário se estabelece, o data-driven torna-se cultural dentro de uma organização, e em todo caso o DD será como uma premissa básica dentro do processo decisório (Awan et al., 2021).

A cultura de dados representa as convicções compartilhadas por pessoas que valorizam o uso de dados para aprimorar a tomada de decisões. Consequentemente, o DD é integrado às operações, à mentalidade e à identidade de uma organização. A cultura de dados capacita todos os membros da organização a buscarem uma abordagem orientada por fatos (Tableau, 2024). Uma pesquisa de mercado feita pela Barc e pelo Tableau em 2023 estruturou a cultura de dados em seis áreas de atuação:

- Liderança de dados: refere-se a indivíduos que coordenam e gerenciam os dados dentro de uma empresa. Essa liderança deve ser responsável pela governança dos dados, estratégias, arquitetura, tecnologias e confiabilidade (Tera, 2020).
- Estratégia de Dados: plano de longo prazo que define as tecnologias, os processos, as pessoas e as regras necessárias para potencializar o ganho de valor através dos dados em uma organização. A estratégia de dados descreve os objetivos da organização em relação aos seus dados (Amazon, 2024).
- Governança de Dados: abordagem fundamentada em princípios para gerenciar dados ao longo do seu ciclo de vida, desde a aquisição até o seu descarte. Abrange todas as

medidas tomadas para garantir a segurança, privacidade, disponibilidade e qualidade dos dados (Google, 2024).

- Alfabetização em Dados: capacidade de uma pessoa trabalhar e interpretar dados (Bange; Grosser, 2023). Segundo um estudo feito pela Forrester em 2021, 82% dos gestores esperam que seus funcionários saibam trabalhar com dados da forma correta.
- Acesso a Dados: relacionado à democratização dos acessos aos dados. As pessoas devem ter acesso às plataformas que disponibilizam os dados, limitadas pelo nível adequado de segurança do acesso (Bange; Grosser, 2023).
- Comunicação de dados: A comunicação de dados é a capacidade de trabalhar em equipe de forma conjunta e eficiente com dados, além de comunicar os resultados de maneira correta. Fornece uma base sólida para discussões e decisões (Barc, 2024).

3 PROCESSOS METODOLÓGICOS/MATERIAIS E MÉTODOS

Este tópico se destina a apresentar a classificação metodológica deste estudo, bem como, os instrumentos e procedimentos que nortearam a condução das análises de dados desenvolvidas para responder à questão de pesquisa. Além disso, há a apresentação das hipóteses formuladas para o alcance dos objetivos propostos.

3.1 TIPOLOGIA DA PESQUISA

A pesquisa que busca informações diretamente com um grupo de interesse é conhecida como survey. Este método foi escolhido pelo autor por ser eficaz em estudos exploratórios (Santos, 1999). O survey envolve a coleta de informações sobre características ou opiniões de um grupo específico considerado representativo de uma população-alvo. Normalmente, questionários são os principais instrumentos de investigação (Fonseca, 2002). Nesse sentido, o presente estudo utilizou abordagem quantitativa com caráter descritivo. O método quantitativo oferece uma visão ampla e generalizável dos resultados, além de possibilitar a identificação de padrões e tendências nos dados coletados. Este método é adequado para testar hipóteses e estabelecer relações causais entre variáveis (Gil, 2010).

3.2 INSTRUMENTOS E PROCEDIMENTOS

O instrumento de coleta de dados refere-se ao meio pelo qual as questões são apresentadas aos participantes da pesquisa e através do qual obtêm-se os resultados (Mattar, 2001). No presente estudo, foi aplicado um questionário, por ser uma ferramenta de pesquisa amplamente aceita e utilizada em levantamentos acadêmicos. De acordo com McDaniel Jr. e Gates (2005), além de possibilitar uma ampla gama de participantes, o questionário é uma ferramenta eficaz para identificar quem são os respondentes, por que eles adotam certos comportamentos e como tomam suas decisões.

De acordo com Gil (2010), o processo de elaboração do questionário consiste em traduzir os objetivos específicos da pesquisa em questões cuidadosamente formuladas. Esta transformação visa garantir que cada pergunta contribua diretamente para a obtenção das informações necessárias. O modo da aplicação do questionário se deu por meio online, o que possibilitou a liberação do acesso para empresas selecionadas por meio das diversas redes de comunicação disponíveis na internet. Entre elas, destaca-se o compartilhamento da pesquisa através das plataformas whatsapp e linkedin. O autor optou por utilizar a plataforma de criação de formulários Google Forms da Google.

O formulário foi elaborado com referência ao survey "BARC Data Culture Survey 23: How to Liberalize Data Access to Empower Data Users", desenvolvido pela BARC (Business Application Research Center) em parceria com a Tableau (Bange; Grosser, 2023). As perguntas e respostas do questionário são mostradas na seção de Resultados. A utilização de um instrumento já validado e reconhecido no campo de pesquisa confere maior robustez à metodologia aplicada (Gil, 2010). Para as possíveis respostas, foram elaboradas diferentes escalas de mensuração das variáveis de estudo para melhor atender ao escopo da pesquisa. Nesse contexto, foram utilizadas escalas de Likert de quatro ou cinco pontos com diferentes ancoragens e escalas personalizadas que melhor se adequam ao tipo de resposta esperada.

3.3 DEFINIÇÃO DA POPULAÇÃO E AMOSTRAS

A população desta pesquisa é composta por empresas de médio a grande porte que lidam com o BD e utilizam dados em suas operações diárias. A seleção da amostra de empresas foi realizada por meio de amostragem intencional, também conhecida como amostragem por julgamento. Esse tipo de amostragem é uma técnica não probabilística, na qual "a seleção dos

elementos da população para compor a amostra depende ao menos em parte do julgamento do pesquisador ou do entrevistador no campo" (Mattar, 1996, p. 132). De acordo com Aaker et al. (1995), o procedimento de amostragem por julgamento pode ser útil quando é necessário incluir um pequeno número de unidades na amostra. Diante desse cenário, o autor optou pelo método de amostragem por julgamento devido à limitação de tempo disponível para a execução de um método pautado em amostragem probabilística.

Nesse caso, a seleção de amostras intencionais é realizada com base no discernimento do pesquisador. Para Aaker et al. (1995), na amostragem intencional escolhem-se elementos típicos e representativos para compor a amostra. Em alguns casos, é comum recorrer à escolha de profissionais especializados para definição das características da amostra (Kish, 1965). Após uma análise minuciosa, o processo de seleção das amostras foi dividido em dois grupos distintos de empresas:

- Empresas que já adotaram sistemas de informação que viabilizam a cultura *data-driven*: Este grupo inclui empresas que, além de lidar com BD, possuem uma arquitetura de sistemas de informação estabelecida e estrutura organizacional que promove a cultura orientada a dados. Também abrange empresas que estão em processo de implementação de uma cultura *data-driven*, ou seja, que estão desenvolvendo e integrando práticas e sistemas para tornar a tomada de decisões baseada em dados como parte de suas operações e estratégias.
- Empresas que ainda não adotaram sistemas de informação que viabilizam a cultura *data-driven*: Este grupo inclui empresas que, apesar de lidarem com BD, ainda não possuem sistemas de informação adequados para tal, além de não possuírem estrutura que promove a cultura orientada a dados ou são principiantes nesse quesito.

3.4 PROCEDIMENTOS DE COLETA E ANÁLISE DOS DADOS

Inicialmente, foi realizado um teste piloto de pesquisa com atribuição de 10 testes para validar a ausência de erros no processo de preenchimento do formulário e certificar a aderência das perguntas. As potenciais empresas que se adequaram às amostras previamente definidas foram selecionadas e, em seguida, a pesquisa foi disponibilizada. De forma a manter a confiabilidade das opiniões coletadas como representação da visão organizacional daquela empresa, foram escolhidos somente gestores de áreas ou diretores dessas organizações. Após a

conclusão da coleta de dados, iniciou-se o processo de tabulação dos resultados, que consiste na conversão das respostas em dados numéricos para análises (Samara; Barros, 2007). Esta etapa foi realizada com o auxílio de planilhas eletrônicas, permitindo a organização sistemática das informações coletadas. Em seguida, procedeu-se à limpeza dos dados, sendo essa uma etapa fundamental para garantir a qualidade e confiabilidade dos dados (Malhotra, 2012). Durante este processo, foram identificados e excluídos quatro questionários que apresentavam padrões de resposta inconsistentes ou que não atendiam aos critérios de inclusão estabelecidos para o estudo. Foram adotados 60 questionários, 30 para a amostra A e o restante para amostra B. A etapa seguinte compreende a análise exploratória dos dados, na qual para facilitar a visualização dos resultados, a amostra A, que inclui empresas que já implementaram sistemas de informação que sustentam uma cultura orientada por dados, será denominada “Organizações Avançadas em Dados” (OAD). Por outro lado, a amostra B, que abrange empresas que ainda não adotaram tais sistemas, será chamada de “Organizações Iniciais em Dados” (OID).

Em seguida, foi realizada a análise inferencial dos dados para validação das hipóteses da pesquisa. O método de inferência estatística, segundo Akter et al. (2016), se destaca no contexto da análise de dados por permitir conclusões sobre uma população a partir das amostras do estudo. Isso significa que os resultados obtidos podem ser generalizáveis para um grupo representativo. Testes estatísticos como o teste t de Student, o teste Qui-quadrado e a análise de variância são comumente utilizados nesse campo de estudo (Rehman et al., 2016). Um dos métodos mais relevantes de inferência estatística na engenharia é o teste de hipóteses, amplamente aplicado em problemas de tomada de decisão (Montgomery; Runger, 2018).

Conforme Devore (2019, p. 296), "uma estatística de teste é uma função dos dados amostrais utilizada como base para decidir se a hipótese nula deve ser rejeitada". Outro elemento fundamental na análise de hipóteses é o valor-p, que indica a probabilidade de se observar nos dados disponíveis um resultado da estatística de teste tão ou mais contraditório em relação à hipótese nula (Devore, 2019). A decisão final depende da comparação entre o valor-p e um nível de significância pré-definido, o que determina a aceitação ou rejeição da hipótese nula (Devore, 2019). Dada a natureza qualitativa das variáveis coletadas no questionário, optou-se pela aplicação do teste de proporções Qui-quadrado e teste exato de Fisher.

4 RESULTADOS E DISCUSSÃO

Neste tópico, são apresentadas as análises a respeito da aplicação do método sobre a amostra de dados. A apresentação da análise de dados, deste estudo, está dividida em duas partes. Na primeira parte, são apresentadas análises descritivas dos dados. Na segunda parte, são apresentados os resultados obtidos pela aplicação da inferência estatística.

4.1 ANÁLISE EXPLORATÓRIA DOS RESULTADOS

4.1.1 Processo decisório

O gráfico 1 ilustra as respostas das organizações à pergunta: “O processo de tomada de decisão na sua empresa baseia-se em dados ou na intuição?”.

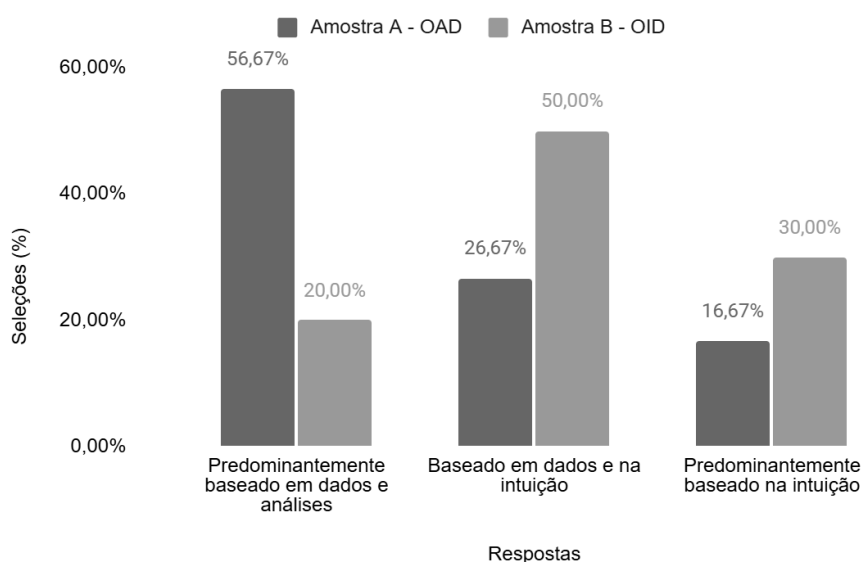


Gráfico 1: Diferenças na tomada de decisão entre OADs e OIDs

Fonte: elaborado pelo autor.

Conforme demonstrado, o grupo das Organizações Avançadas em Dados, possui predominância para a tomada de decisões baseada em dados, com 56,6% das organizações relatando que suas decisões são fundamentadas em dados e análises. Uma proporção menor (26,6%), indica uma abordagem híbrida, combinando dados e intuição, enquanto 16,67% relatam uma dependência predominante da intuição no processo decisório. Em contraste, o grupo das OID apresenta um perfil decisório diferente. Neste grupo, apenas 20% das

organizações reportam a tomada de decisão predominantemente baseada em dados e análises. Em grande maioria, o grupo das OID apresenta uma abordagem baseada em dados mais intuição.

4.1.2 Capacidade de assimilação do big data

A segunda pergunta do questionário visa avaliar o grau de assimilação das organizações em relação às variáveis volume, variedade e velocidade no contexto do BD. Nesse contexto, a tabela 1 fornece os resultados obtidos à indagação: “Sob a perspectiva do alto volume, variedade e velocidade dos dados no contexto do *big data*, escolha a opção que melhor representa a percepção da sua empresa.”

Assimilação do <i>Big Data</i> e suas principais variáveis			
Respostas	Amostra A - OAD	Amostra B - OID	Total geral
Não conseguimos assimilar o <i>big data</i> e nem extrair valor dos dados.	4	9	13
Não temos dificuldade de assimilar o <i>big data</i> e conseguimos extrair valor dos dados.	20	5	25
Temos dificuldade de assimilar o <i>big data</i> mas conseguimos extrair valor dos dados.	6	16	22
Total geral	30	30	60

Tabela 1: Classificação da assimilação do *big data* pelas OAD e OID

Fonte: elaborado pelo autor.

No grupo das OAD, a maioria significativa (66,7%, n=20) reporta que consegue assimilar o BD e extrair valor dos dados. Um percentual menor (20%, n=6) afirma ter dificuldade na assimilação, mas ainda consegue gerar valor a partir dos dados, enquanto 13,3% (n=4) indicam que não conseguem assimilar o BD e tampouco extrair valor. Em contraste, no grupo das OID, mais de metade das empresas (53,3%, n=16) relatam dificuldades na assimilação dos dados, mas conseguem extrair algum valor. Aproximadamente 30% (n=9) das

OID afirmam que não conseguem assimilar nem gerar valor dos dados, e apenas 16,7% (n=5) conseguem assimilar o BD e gerar valor de maneira eficaz.

4.1.3 Maturidade técnica para lidar com *big data*

Sobre a questão “Como você classifica o nível de maturidade técnica e analítica da sua organização para lidar com *big data*?”. Os resultados obtidos são apresentados no gráfico 2.

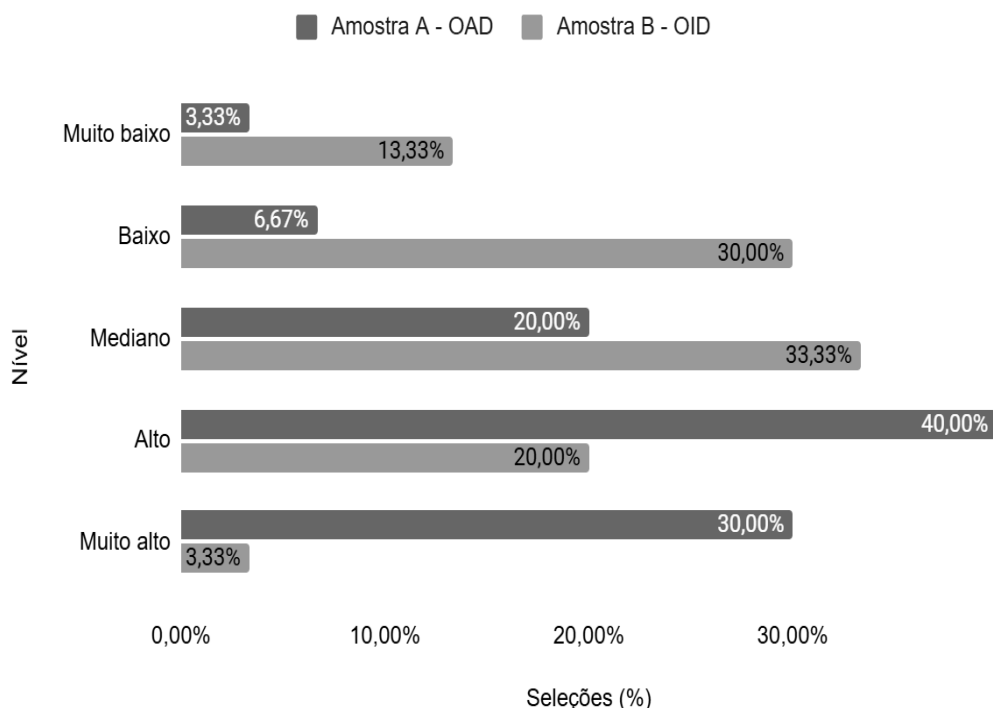


Gráfico 2: Comparativo do nível de maturidade para lidar com *big data*

Fonte: elaborado pelo autor.

A maioria no grupo das OAD, representando 70% do total, classifica sua maturidade como Alto (40%) ou Muito alto (30%). Esses valores indicam uma capacidade bem desenvolvida para gerenciar as etapas principais de arquitetura de dados. Uma fração de 10% (n=3) das OAD se encontra nas categorias de maturidade Muito baixo e Baixo. O cenário das OID revela uma incidência de processos com qualidade mediana (33%, n=10). Combinando as

categorias Muito baixo e Baixo, foi observado que 43,33% das OID reconhecem que suas capacidades estão aquém do ideal. 20% conseguem se posicionar como Alto.

4.1.4 Concordância com as premissas da cultura de dados

Nesta seção analisa-se os resultados da quarta pergunta: "Até que ponto você concorda com as afirmações a seguir sobre a cultura de dados na sua empresa?".

Gráfico 3: Concordância das OAD com as premissas da cultura de dados

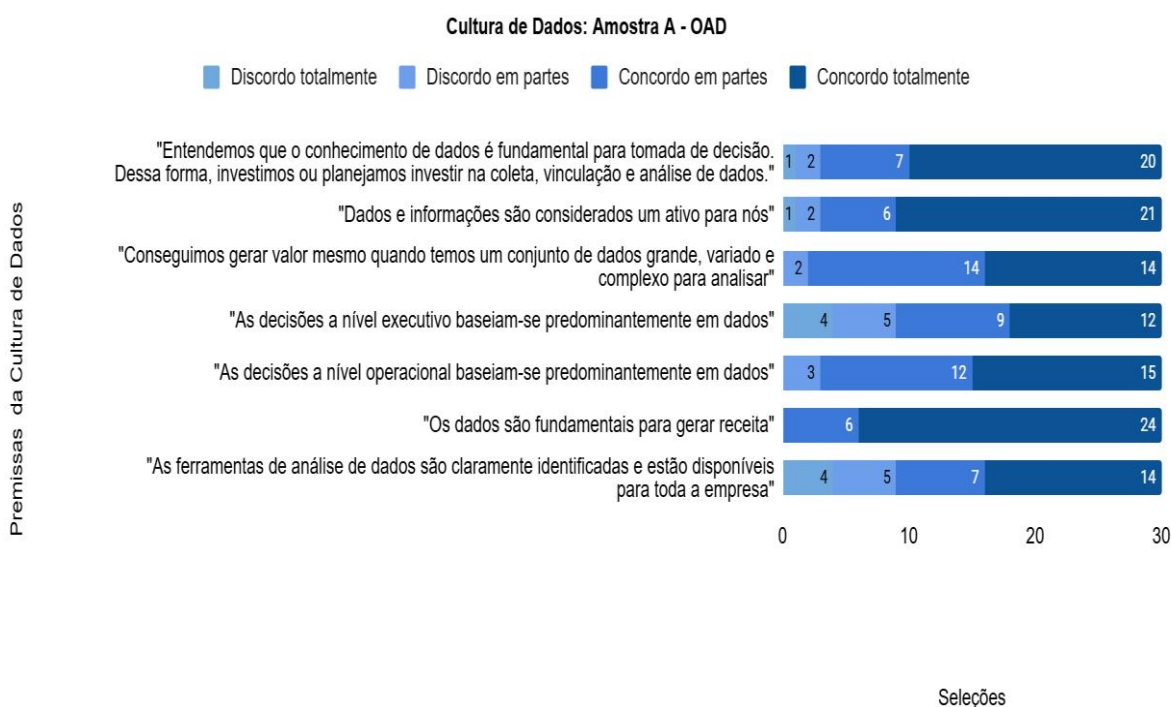


Gráfico 3: Concordância das OAD com as premissas da cultura de dados

Fonte: elaborado pelo autor.

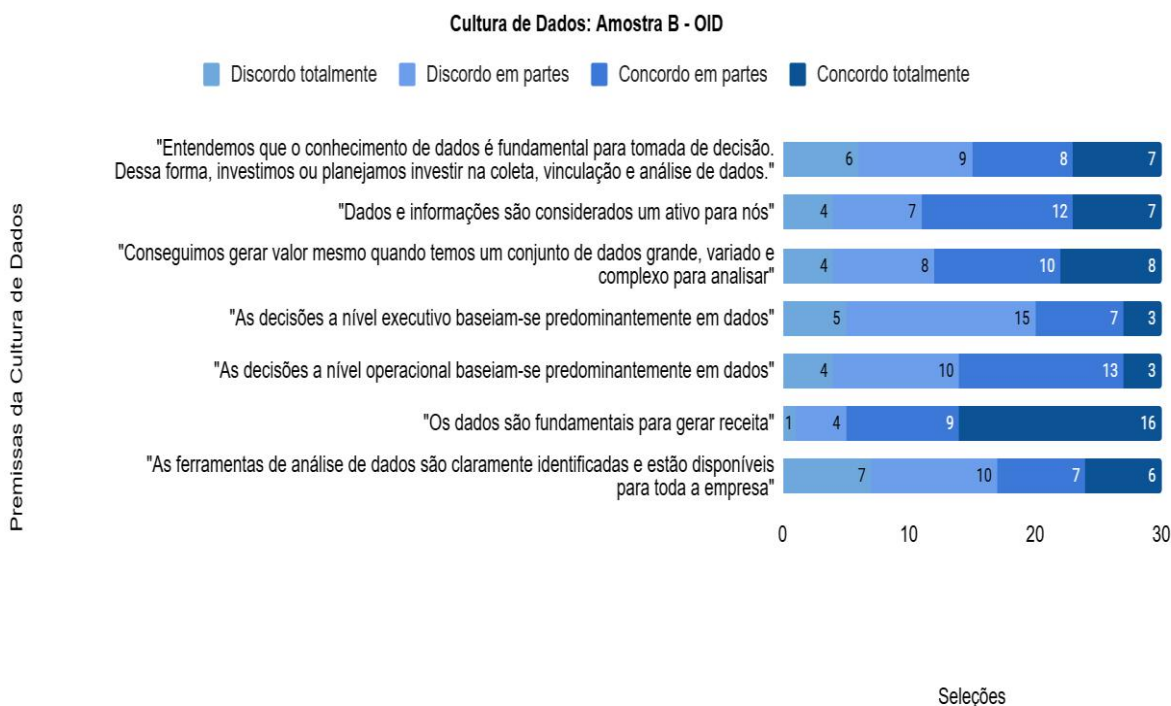


Gráfico 4: Concordância das OID com as premissas da cultura de dados

Fonte: elaborado pelo autor.

Os gráficos 3 e 4 revelam diferenças entre OAD e OID no uso de dados para a tomada de decisão. Na afirmação sobre a importância do conhecimento de dados e o investimento em análise, as OAD lideram com 20 respostas de "concordo totalmente" e 7 de "concordo em partes", enquanto as OID apresentam apenas 7 e 9, respectivamente. Na premissa de que "os dados são fundamentais para gerar receita", as OAD registram 24 concordâncias totais contra 16 das OID, indicando que estas ainda não exploram plenamente o potencial dos dados para aumentar a rentabilidade, possivelmente por falta de clareza sobre o retorno financeiro.

Quanto às ferramentas de análise de dados acessíveis a toda a empresa, as OAD têm 14 respostas de concordância total e 7 parciais, enquanto as OID mostram apenas 6 e 7, respectivamente, apontando desafios na democratização do acesso às informações. Em comum, tanto nas OAD quanto nas OID, as decisões operacionais tendem a ser mais embasadas em dados do que as decisões executivas.

4.1.5 Desafios na geração de valor com dados

Os resultados da quinta pergunta: "Quais são os maiores desafios na geração de valor com os dados na sua empresa?" indicam os obstáculos enfrentados pelas empresas para extrair retorno dos dados. Nesse aspecto, é possível observar alguns pontos em comum entre as OAD e OID a partir do gráfico 5. Entre eles, um dos principais desafios identificados por ambos os grupos é a formação de equipes qualificadas para atuar com BD.

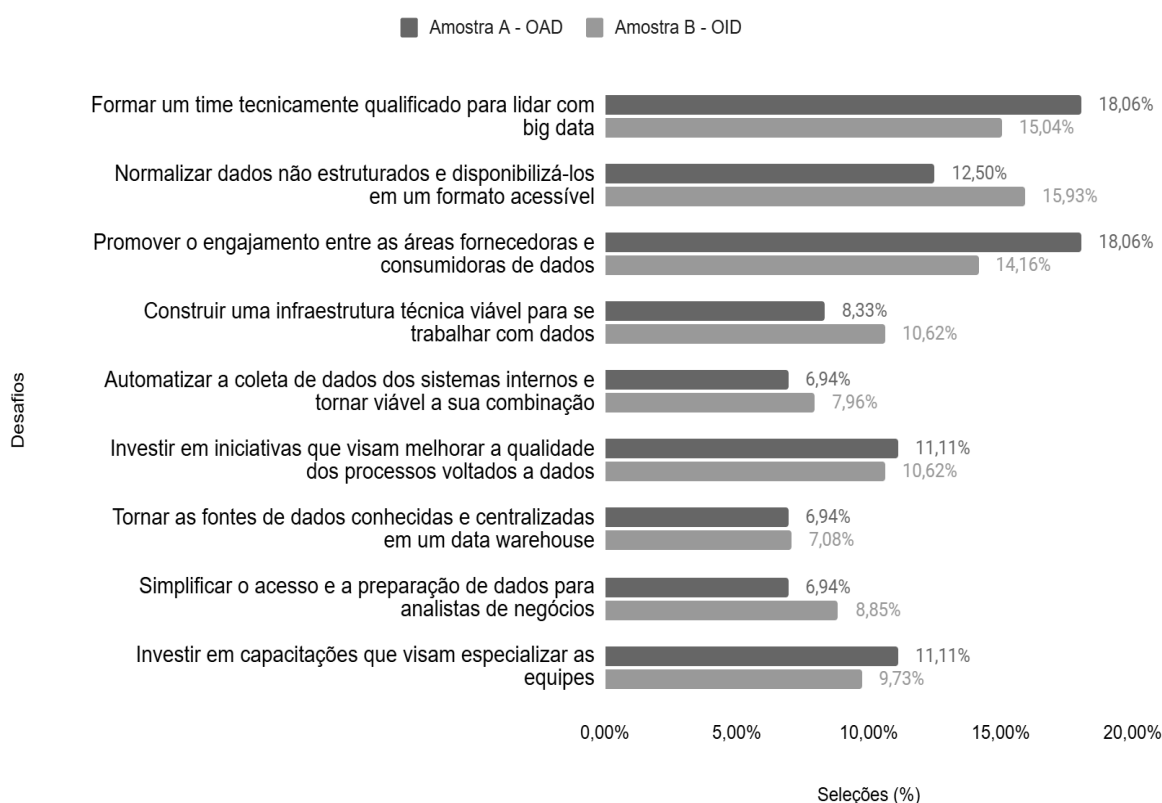


Gráfico 5: Comparativo dos desafios da geração de valor com dados

Fonte: elaborado pelo autor.

Destaca-se também a necessidade de promover o engajamento entre as áreas fornecedoras e consumidoras de dados como uma problemática comum entre os dois grupos, com 18,06% de seleção para as OAD e 14,16% para as OID.

4.1.6 Benefícios obtidos através do data-driven

Os resultados da sexta pergunta: "Quais benefícios sua empresa já obteve através do uso de dados?" indicam os retornos obtidos pelas empresas ao utilizarem dados como ferramenta estratégica.

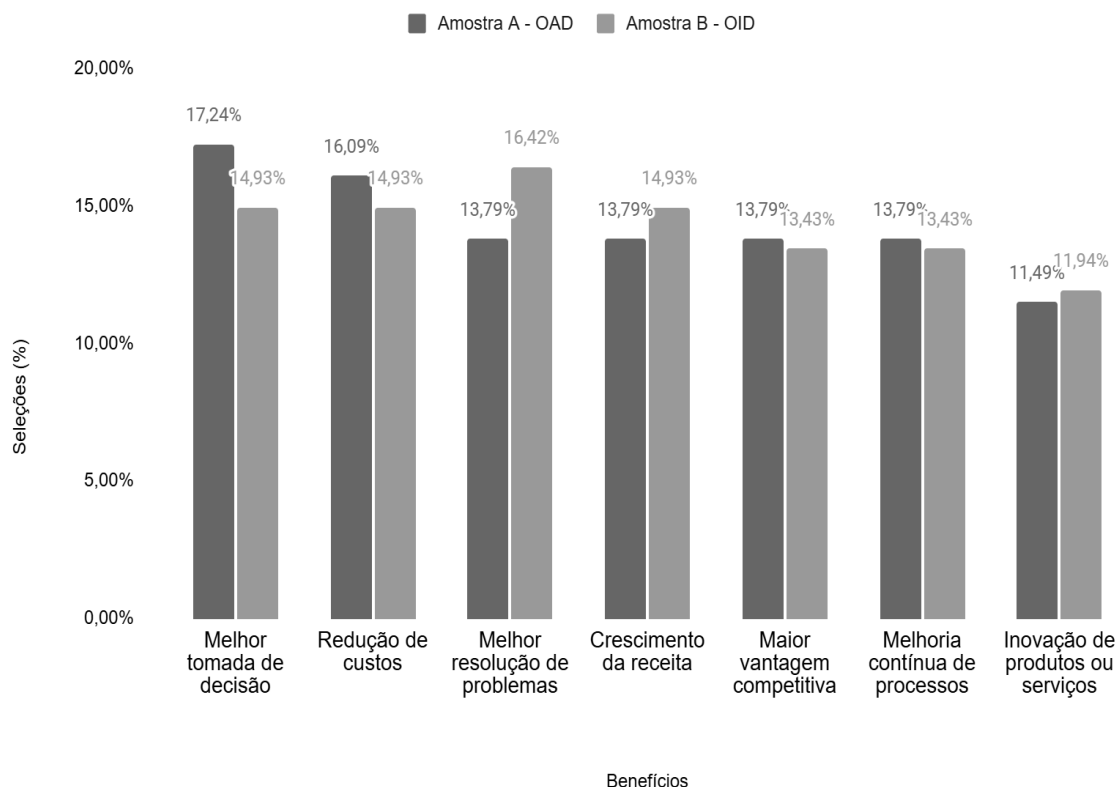


Gráfico 6: Comparativo dos benefícios obtidos com dados

Fonte: elaborado pelo autor.

Analisa-se, pelo gráfico 6, que as OID destacam o crescimento da receita como um dos principais benefícios, com 14,93%. Isso sugere que, independentemente do nível de maturidade em dados, as empresas já atingiram em algum momento um aumento de receita proporcionado pelo uso dos dados. Além disso, a melhor tomada de decisão é valorizada nas OAD, com 17,24%.

4.1.7 Cultura de Dados

O tópico 7 do questionário “Indique o nível de importância que sua empresa atualmente atribui às seguintes frentes de cultura de dados.” foi utilizado para levantar o nível de importância atribuído pelas organizações para cada frente da estrutura da cultura de dados.

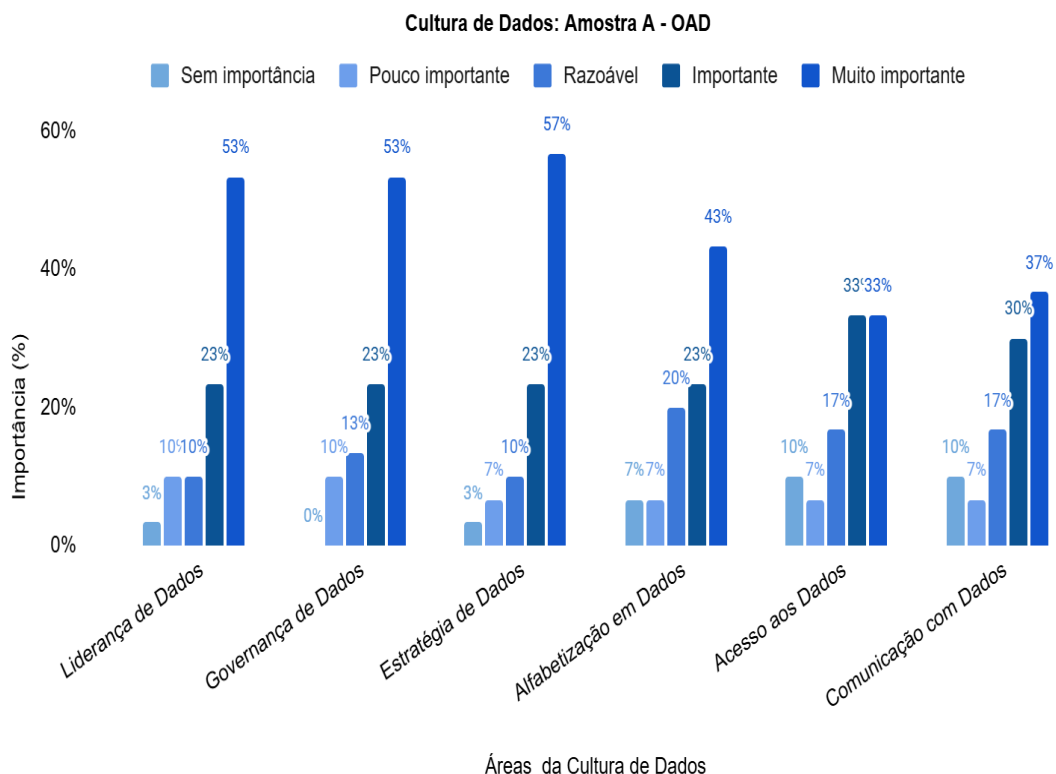


Gráfico 7: Importância atribuída pelas OAD aos pilares da cultura de dados

Fonte: elaborado pelo autor.

O gráfico 7 representa a percepção de relevância das OAD para cada pilar da cultura de dados. Verifica-se:

- Liderança de Dados: 53% das OAD consideram muito importante, e 23% importante, totalizando 76% de alta relevância.
- Governança de Dados: 53% veem como muito importante, e 23% como importante (76% no total).

- Estratégia de Dados: 57% consideram muito importante, e 23% importante (80% no total).
- Alfabetização em Dados: 43% veem como muito importante, e 23% como importante (66% no total).
- Acesso aos Dados: 33.33% consideram muito importante, e outro 33.33% importante (66.66% no total).
- Comunicação com Dados: 37% veem como muito importante, e 30% como importante (67% no total).

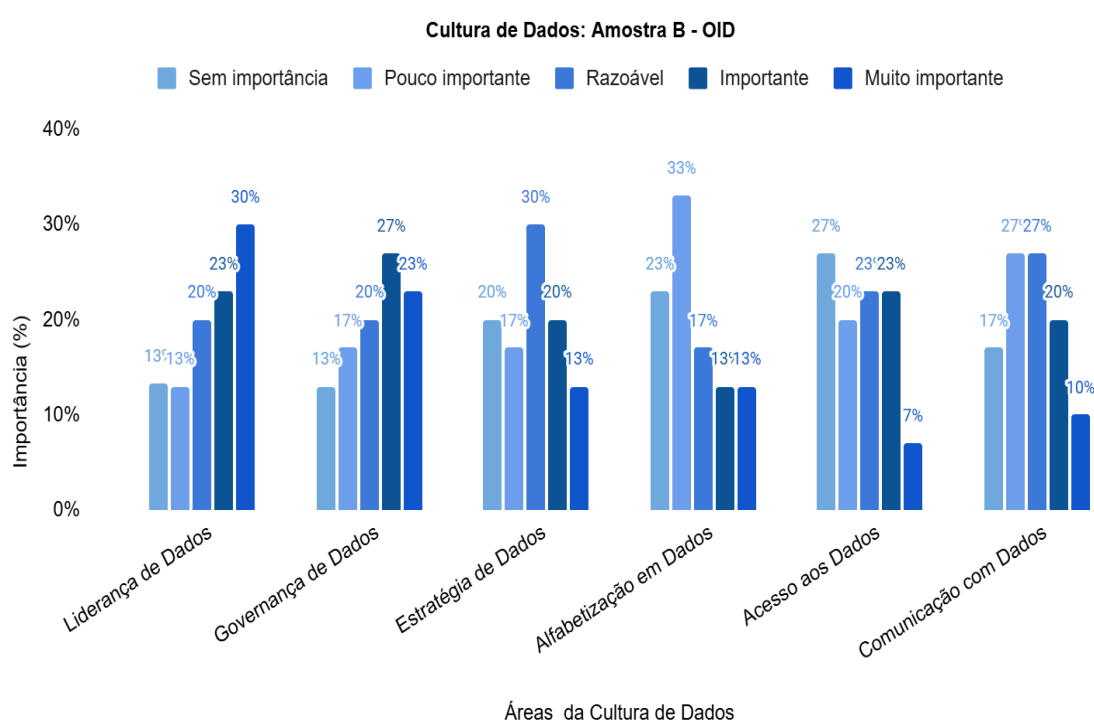


Gráfico 8: Importância atribuída pelas OID aos pilares da cultura de dados

Fonte: elaborado pelo autor.

O gráfico 8 representa a percepção das OID ao mesmo tópico. Verifica-se:

- Liderança de Dados: 30% consideram muito importante, e 20% importante, totalizando 50% de alta relevância.
- Governança de Dados: 23% veem como muito importante, e 27% como importante (50% no total).

- Estratégia de Dados: 13% consideram muito importante, e 20% importante (33% no total).
- Alfabetização em Dados: 13% veem como muito importante, e 17% como importante (30% no total).
- Acesso aos Dados: 23% consideram muito importante, e 7% importante (30% no total).
- Comunicação com Dados: 10% veem como muito importante, e 20% como importante (30% no total).

4.1.8 Obstáculos para implementação da cultura de dados

O tópico 8 do questionário “Quais são ou foram os principais obstáculos à implementação ou estabelecimento da cultura de dados na sua empresa?” aborda possíveis desafios enfrentados pelas empresas na adoção da cultura de dados.

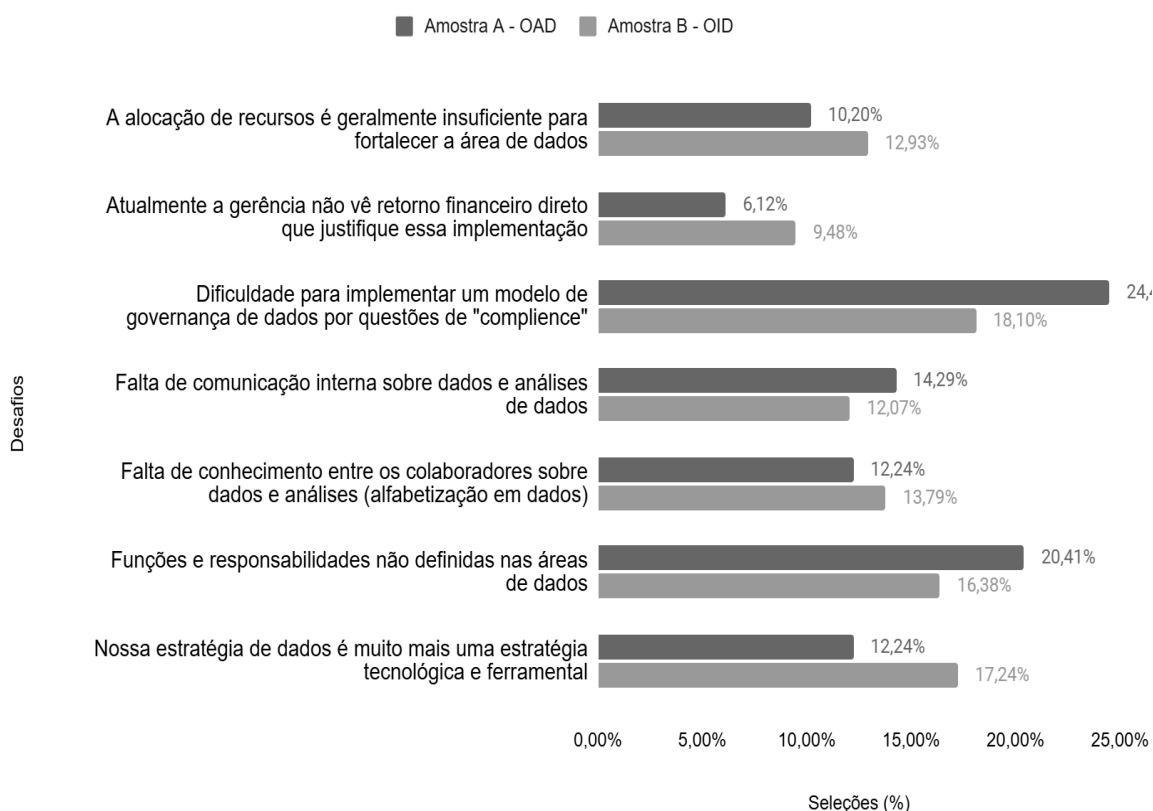


Gráfico 9: Comparativo dos desafios para implementar a cultura de dados

Fonte: elaborado pelo autor.

Através da análise comparativa do gráfico 9, ambos os grupos identificam dificuldades para implementação de um modelo de governança de dados por questões de "compliance" como o maior desafio, sendo mais acentuado nas OAD. Entretanto, as OAD dão maior ênfase à definição de funções e responsabilidades nas áreas de dados, enquanto as OID destacam questões tecnológicas e de alocação de recursos com maior frequência.

4.2 ANÁLISE INFERENCIAL DOS DADOS

4.2.1 Tipo de organização e o processo decisório

A fim de validar se empresas que apresentam uma cultura *data-driven* (representado pelo grupo das OAD) são mais eficientes no processo decisório baseado em dados, faz-se necessário investigar até que ponto as diferenças obtidas nos resultados da pesquisa se devem ao acaso ou indicam associação entre a cultura de dados e a eficácia na tomada de decisão. Nesse contexto, foram estabelecidas as seguintes hipóteses:

Hipótese nula (H_0): Não existe associação significativa entre o tipo de organização (OAD ou OID) e a forma como o processo de tomada de decisão é embasado. Ou seja, a forma como ocorre o processo decisório independe do tipo de organização.

Hipótese alternativa (H_1): Existe associação significativa entre o tipo de organização (OAD ou OID) e a forma como o processo de tomada de decisão é embasado. Ou seja, o *data-driven* influencia no processo decisório.

Para iniciar as análises no software R foi necessário importar a base de dados para o programa e então construir uma tabela de contingência dos resultados referentes à primeira pergunta do questionário. Os dados são demonstrados no gráfico 1 da seção 4.1.1. Nas análises, utilizou-se o R *commander* da biblioteca "Rcmdr", com auxílio da biblioteca "MASS" (Fox, 2005). Os resultados do teste Qui-quadrado e teste exato de fisher são demonstrados na tabela 2:

Teste Estatístico	Estatística de Teste	Graus de Liberdade (df)	p-valor
Teste Qui-Quadrado (Pearson)	$\chi^2 = 8.5342$	2	0.01402
Teste Exato de Fisher	Não aplicável	Não aplicável	0.01755

Tabela 2: R - Estatística de teste para a primeira associação

Fonte: elaborado pelo autor.

De acordo com os resultados apresentados, como o valor-p é menor que o nível de significância usual de 0.05, rejeita-se a hipótese nula. A aplicação da hipótese alternativa indica que existe uma associação estatisticamente significativa entre o tipo de organização (OAD ou OID) e a forma como as decisões são tomadas. O valor-p obtido para o teste exato de Fisher corrobora a conclusão do teste Qui-quadrado.

4.3.3 Nível de maturidade e capacidade de assimilação

A fim de validar se empresas que apresentam um nível maior de maturidade técnica e analítica relacionada ao BD possuem maior capacidade de assimilação do BD e geração de valor, foram realizados os testes acima para a segunda e terceira pergunta do questionário. Nesse contexto, a diferenciação entre OAD e OID não foi utilizada para que se tornasse possível o relacionamento dessas duas variáveis sob o formato de tabela de contingência. Assim, foram estabelecidas as seguintes hipóteses:

Hipótese nula (H_0): Não existe associação significativa entre o nível de maturidade técnica e analítica da empresa e a sua capacidade de assimilação do BD, bem como sua capacidade de geração de valor.

Hipótese alternativa (H_1): Existe associação significativa entre o nível de maturidade técnica e analítica da empresa e a sua capacidade de assimilação do BD, bem como sua capacidade de geração de valor.

Os resultados do teste Qui-quadrado e teste exato de fisher são demonstrados na tabela 3:

Teste Estatístico	Estatística de Teste	Graus de Liberdade	p-valor
Teste Qui-Quadrado (Pearson)	$X^2 = 48.999$	8	0.00000006361
Teste Exato de Fisher	Não aplicável	Não aplicável	0.0000001518

Tabela 3: R - Estatística de teste para a segunda associação

Fonte: elaborado pelo autor.

De acordo com os resultados apresentados, a estatística do teste Qui-quadrado pode não ser indicada para essa proposição, devido à presença de frequências esperadas menores que 1. Nesse caso, será utilizado somente o teste exato de Fisher, que apresenta valor de 0.0000001518. Como o valor-p é inferior a 0.05, rejeita-se a hipótese nula. O teste exato de Fisher indica que há uma associação significativa entre a maturidade técnica da organização e sua capacidade de lidar com o BD.

5 CONCLUSÕES

O trabalho objetivou analisar como as organizações lidam com o desafio de extrair valor do amplo volume de dados disponíveis e verificar se a cultura de dados é um fator de sucesso para uma empresa atingir ganhos e melhorar sua tomada de decisão. Os principais conceitos como BD, tomada de decisão e cultura de dados foram amplamente analisados na conceituação teórica do estudo.

A pesquisa demonstrou que o crescimento do BD pode ser um dos principais impulsionadores da transformação empresarial. Os dados representam ativos estratégicos que favorecem o crescimento da receita e o posicionamento competitivo, mas exigem das organizações infraestrutura tecnológica robusta e uma cultura *data-driven* para alcançarem resultados.

As Organizações Avançadas em Dados (OAD) mostraram maior eficiência em relação ao BD, melhor profundidade técnica e maiores chances de extrair benefícios, apresentando maior concordância com as premissas da cultura de dados e ênfase em liderança, governança e

estratégia. Os resultados inferenciais confirmaram que as OAD são mais eficazes que as OID na tomada de decisão, sendo o nível de maturidade técnica um fator determinante para assimilação e geração de valor a partir do BD.

O estudo reconhece limitações, como o uso predominante de variáveis qualitativas e amostra limitada a 60 organizações, o que restringe a generalização dos resultados. Sugere-se que estudos futuros incorporem métricas financeiras, ampliem a amostra e analisem detalhadamente as diferenças entre decisões operacionais e executivas no contexto da cultura de dados.

REFERÊNCIAS

AAKER, D.; KUMAR, V.; DAY, G. **Marketing research**. 5. ed. New York: John Wiley & Sons, 1995.

ALVAREZ, F. J.; SANCHEZ, F. S.; BELLIDO, R. Demographic and dwelling models by artificial intelligence: urban renewal opportunities in Spanish coast. **International Journal of Sustainable Development and Planning**, v. 13, n. 07, p. 941–953, 30 set. 2018.

AMAZON. **O que é estratégia de dados?** 2024. Disponível em: <https://aws.amazon.com/pt/what-is/data-strategy/>. Acesso em: 6 jun. 2024.

AKTER, S. et al. How to improve firm performance using big data analytics capability and business strategy alignment? **International Journal of Production Economics**, v. 182, p. 113–131, dez. 2016.

AWAN, U. et al. Big data analytics capability and decision-making: The role of data-driven insight on circular economy performance. **Technological Forecasting and Social Change**, v. 168, p. 120766, jul. 2021.

BANGE, C; GROSSER, T. **BARC Data Culture Survey 23: How to liberalize data access to empower data users**. 2023. Disponível em: <https://barc.com/research/barc-data-culture-survey-23/>. Acesso em: 15 mai. 2024.

BANSAL, S. K.; KAGEMANN, S. Integrating Big Data: A Semantic Extract-Transform-Load Framework. **Computer**, v. 48, n. 3, p. 42–50, mar. 2015.

BARC, **What is data communication?** 2024. Disponível em: <https://barc.com/de/glossar/data-analytics/data-communication/>. Acesso em: 28 Jun. 2024.

BERISHA, B.; MËZIU, E.; SHABANI, I. Big data analytics in Cloud computing: an overview. **Journal of Cloud Computing**, v. 11, n. 1, p. 24, 6 ago. 2022.

CHAUHAN, P.; SOOD, M. Big Data: Present and Future. **Computer**, v. 54, n. 4, p. 59–65, abr. 2021.

COX, M.; ELLSWORTH, D. **Application-controlled demand paging for out-of-core visualization**. Proceedings. Visualization '97 (Cat. No. 97CB36155). **Anais...IEEE**, [s.d.].

DEVORE, J. L. **Probabilidade e estatística para engenharia e ciências**. 1. ed. São Paulo: Cengage, 2019. 656 p. ISBN 978-85-221-2803-7.

DWIVEDI, Y. K. et al. Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. **International Journal of Information Management**, v. 57, p. 101994, abr. 2021.

FONSECA, J. J. S. **Metodologia da pesquisa científica**. Fortaleza: UEC, 2002.

FOX, J. The R Commander: A Basic Statistics Graphical User Interface to R. **Journal of Statistical Software**, v. 14, n. 9, p. 1–42, 2005. Disponível em: <https://www.jstatsoft.org/article/view/v014i09>. Acesso em: 16 nov. 2024.

GIL, A. C. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2007

GIL, A. C. **Como elaborar projetos de pesquisa**. 5. ed. São Paulo: Atlas, 2010. 184 p..

GOOGLE. **O que é governança de dados?** 2024. Disponível em: <https://cloud.google.com/learn/what-is-data-governance?> Acesso em: 11 jul. 2024.

GUALTIERI, M. **Hadoop Is Data's Darling For A Reason**. Forrester, 16 de janeiro de 2016. Disponível em: <https://www.forrester.com/blogs/hadoop-is-datas-darling-for-a-reason/>. Acesso em: 20 de mar. de 2024.

GUNTER, H. **Como elaborar um questionário**. (Série: Planejamento de Pesquisa nas Ciências Sociais, Nº 1). Brasília, DF: UnB, Laboratório de Psicologia Ambiental, 2003. Disponível em: <http://beco-do-bosque.net/XTextos/01Questionario.pdf>. Acesso em: 29 jun. 2013.

KISH, L. **Survey sampling**. New York: John Wiley & Sons, Inc., 1965.

MCDANIEL JR., C. D.; GATES, R. **Fundamentos de pesquisa de marketing**. 4. ed. Rio de Janeiro: LTC, 2005. 312 p.

MALHOTRA, N. K. **Pesquisa de marketing: uma orientação aplicada**. 6. ed. Porto Alegre: Bookman, 2012. 735 p.

MATTAR, F. N. **Pesquisa de marketing: metodologia, planejamento, execução e análise**. 2. ed., v. 2. São Paulo: Atlas, 1994.

MATTAR, F. **Pesquisa de marketing**. São Paulo: Atlas, 1996.

MATTAR, F. N. **Pesquisa de marketing**. 3. ed. compacta. São Paulo: Atlas, 2001. 275 p.

MICROSOFT. **Relational vs. NoSQL data**. 2022. Disponível em: <https://learn.microsoft.com/en-us/dotnet/architecture/cloud-native/relational-vs-nosql-data>. Acesso em: 6 jun. 2024.

MINTZBERG, H.; RAISINGHANI, D.; THEORET, A. The structure of “unstructured” decision processes. **Administrative science quarterly**, v. 21, n. 2, p. 246, 1976.

MIRARAB, A.; MIRTAHERI, S. L.; ASGHARI, S. A. Value creation with big data analytics for enterprises: a survey. **TELKOMNIKA (Telecommunication Computing Electronics and Control)**, v. 17, n. 6, p. 2790, 1 dez. 2019.

MISRA, N. N. et al. IoT, Big Data, and Artificial Intelligence in Agriculture and Food Industry. **IEEE Internet of Things Journal**, v. 9, n. 9, p. 6305–6324, 1 maio 2022.

MONTGOMERY, D. C.; RUNGER, G. C. **Applied statistics and probability for engineers**. 7. ed. Nashville, TN, USA: John Wiley & Sons, 2018.

MYLER, L. **Better data quality equals higher marketing ROI**. Disponível em: <https://www.forbes.com/sites/larrymyler/2017/07/11/better-data-quality-equals-higher-marketing-roi>. Acesso em: 13 abr. 2024.

NETFLIX. **Research areas analytics: Driving insights from data**. 2024. Disponível em: <https://br.investing.com/academy/statistics/faturamento-netflix/>. Acesso em: 7 jul. 2024.

NUTT, P. C. Types of Organizational Decision Processes. **Administrative Science Quarterly**, v. 29, n. 3, p. 414, set. 1984.

OUSSOUS, A. et al. Big Data technologies: A survey. **Journal of King Saud University - Computer and Information Sciences**, v. 30, n. 4, p. 431–448, out. 2018.

PARADZA, D.; DARAMOLA, O. Business Intelligence and Business Value in Organisations: A Systematic Literature Review. **Sustainability**, v. 13, n. 20, p. 11382, 15 out. 2021.

PUCPR. **Do dado à decisão: o que é e como construir uma cultura data-driven**. 2024. Disponível em: <https://posdigital.pucpr.br/blog/cultura-data-driven>. Acesso em: 6 jul. 2024.

REHMAN, M. H.-U.; CHANG, V.; BATOOL, A.; WAH, T. Y. Big data reduction framework for value creation in sustainable enterprises. **International Journal of Information Management**, v. 36, n. 6, p. 917–928, 2016.

RIALTI, R. et al. Big data and dynamic capabilities: a bibliometric analysis and systematic literature review. **Management Decision**, v. 57, n. 8, p. 2052–2068, 12 set. 2019.

SANTOS, A. R. **Metodologia científica: a construção do conhecimento**. Rio de Janeiro: DP&A, 1999.

SAMARA, B. S.; BARROS, J. C. **Pesquisa de marketing: conceitos e metodologia**. 4. ed. São Paulo: Pearson, 2007. 273 p.

SHARDA, R.; DELEN, D.; TURBAN, E. **Business Intelligence e Análise de Dados para Gestão do Negócio**. 4^a ed. Porto Alegre: Bookman, 2019. 584 p.

SHIMIZU, T. **Decisão nas organizações: introdução aos problemas de decisão encontrados nas organizações e nos sistemas de apoio à decisão**. São Paulo: Atlas, 2001. 313 p.

SIMON, H. A. **The new science of management decision**. Old Tappan, NJ, USA: Prentice Hall, 1977.

TABLEAU. **A Guide To Data Driven Decision Making: What It Is, Its Importance, & How To Implement It**. 2024. Disponível em: <https://www.tableau.com/learn/articles/data-driven-decision-making>. Acesso em: 8 jul. 2024.

TABLEAU. **Data culture drives data and AI success**. 2024. Disponível em: <https://www.tableau.com/why-tableau/data-culture>. Acesso em: 15 jul. 2024.

TERA. **Liderança de Dados: definindo a expertise que sua organização precisa**. 2024. Disponível em: <https://medium.com/somos-tera/lideran%C3%A7a-de-dados-definindo-a-expertise-que-sua-organiza%C3%A7%C3%A3o-precisa-b1dcce17f59>. Acesso em: 8 jul. 2024.

DAVENPORT, T; BEAN, R. **Data and AI Leadership Executive Survey 2022**. Boston: [s.n.].

WANG, J. et al. Big Data Service Architecture: A Survey. **School of Computer & Communication Engineering**, v. 21, 2020.