e-ISSN: 2236-0972
https://doi.org/10.5585/gep.v15i1.24673
Received: 11 June 2023
Approved: 14 Oct. 2023
Editor-in-chief: Luciano Ferreira da Silva
Scientific Editors: Flavio Santino Bizarrias e Renato Penha

Check for updates

PRACTICES AND BARRIERS FOR BIG DATA PROJECTS: A CASE STUDY ON A LARGE INSURANCE COMPANY

PRÁTICAS E BARREIRAS EM PROJETOS DE BIG DATA: UM ESTUDO DE CASO EM UMA GRANDE SEGURADORA

Marco Alexandre Terlizzi
Doutor em Administração
Escola de Administração de Empresas de São Paulo – FGV EAESP
São Paulo, SP – Brasil
marco.terlizzi@fgv.br

Felippe Eiji Tashiro de Oliveira
Mestre em Administração
Escola de Administração de Empresas de São Paulo – FGV EAESP
São Paulo, SP – Brasil
felippe.eiji@gmail.com

Eduardo de Rezende Francisco
Doutor em Administração
Escola de Administração de Empresas de São Paulo – FGV EAESP
São Paulo, SP – Brasil
eduardo.francisco@fgv.br

Abstract
The adoption of big data analytics is increasing in every major industry, which requires investments in new projects, technologies, architectures, and processes to allow the integration of big data platforms with legacy systems; however, many organizations have failed to incorporate it effectively into their decision-making processes and the benefits of big data analytics have not been adequately captured. This study aims to demonstrate the practices and barriers related to the implementation of a big data analytics project and suggest improvements to future projects. A case study was conducted on one of the largest insurance companies in Brazil with interviews with ten professionals engaged in the project (technicians, managers, and executives) and document analysis. The study advances the literature by adding two new findings that have not previously been identified: one new practice that was adopted to successfully implement a big data analytics project (implement automatic autoscaling alerts), as well as one barrier that prevents its proper adoption (complexity of access to multicloud data sources). The study also corroborated others practices and barriers that have previously been identified: four main practices (use specialized big data tools, integrate the platform with legacy systems, comply with privacy legislation, and ensure the documentation of technical architecture using business process modeling), and three barriers (high processing requirements of unstructured data analysis, failure to attend to business necessities at the right time, and project delays brought by bureaucratic interdepartmental processes). Finally, in our practical contribution, we propose an action plan to remove the main barriers that can impact the successful delivery of the project scope (integrate the new platform with the legacy). The project had excellent cases of payback right after the implementation, stimulating more innovation and advances.

Keywords: Project management. IT projects. Big data analytics platform. Insurance industry.

Resumo
A adoção do big data pelas organizações continua em expansão, exigindo investimentos em novos projetos, tecnologias, arquiteturas e processos que permitam a integração das novas plataformas de big data aos sistemas legados; entretanto, muitas organizações ainda não conseguiram integrar de forma eficaz o big data aos seus processos de tomada de decisão nem capturar de forma adequada seus benefícios. Este estudo tem como objetivo demonstrar as práticas e barreiras relacionadas à implementação de uma plataforma de big data e sugerir melhorias para projetos futuros. Realizamos um estudo de caso em uma das maiores seguradoras do Brasil por meio de análise documental e entrevistas com dez profissionais envolvidos no projeto (técnicos, gestores e executivos). O estudo expande a literatura atual com duas novas descobertas: uma nova prática que pode ser utilizada em uma plataforma de big data (alertas de escalonamento automático), bem como uma barreira que pode inibir sua adoção adequada (complexidade ao acessar fontes de dados *multicloud*). O estudo também corrobora práticas e barreiras identificadas anteriormente: quatro práticas (uso de ferramentas especializadas de big data, integração da nova plataforma aos sistemas legados, atendimento a legislação de privacidade, e uso de modelagem de processos na documentação técnica), e três barreiras (alto consumo de energia para processar dados não estruturados, não atendimento às necessidades do negócio no momento certo, e atraso no projeto causado por processos burocráticos interdepartamentais). Por fim, como contribuição prática, propomos um plano de ação para remover as principais barreiras que podem impactar o sucesso do escopo do projeto. O projeto gerou excelentes resultados pós-implantação, estimulando mais inovações e avanços.

Palavras-chave: Gerenciamento de projetos. Projetos de tecnologia. Plataforma de big data. Indústria de seguros.

How to cite the article
American Psychological Association (APA)

Terlizzi, M. A., Oliveira, F. E. T., & Francisco, E. R. (2024, Jan./Apr.). Practices and barriers for big data projects: A case study on a large insurance company. *Revista de Gestão e Projetos (GeP)*, *15*(1), 1-35. https://doi.org/10.5585/gep.v15i1.24673

1 Introduction

The insurance industry is a large worldwide investor in information and communication technology and responsible for 7% of the world's total related investments – US$ 200 billion in 2021 (Deloitte, 2021). Furthermore, one of the top priorities in such investments is legacy application modernization, which includes componentizing monolithic systems to support big data analytics with real-time data processing (Gartner, 2021). Employing big data analytics facilitates a deep understanding of the business through the sophisticated and rich environment that big data provides, enabling better decisions for generating value to the business (Gökalp et al., 2022).

Big data analytics is an essential area of study for practitioners and researchers (Chen et al., 2012; Silveira et al., 2015). It can be defined as managing, analyzing, and processing massive quantities of complex, real-time data to extract valuable business insights (Gupta & George, 2016). The adoption of big data analytics is increasing in every major industry (Mazzei & Noble, 2017); based on Statista forecasts, the big data market will grow to US$ 103 billion in worldwide revenue by 2027 (Statista, 2022). Especially in the insurance industry, the use of big data analytics has been increasing at a high rate in the last few years (Susep, 2022), demanding investments in new projects, technologies, architectures, and processes to allow the integration of big data platforms with legacy systems; such systems are responsible for storing organizations' core data, such as customer profiles, insurance claims, fraud history, and others (Kim et al., 2014).

Despite recent advances in big data analytics, there is substantial evidence showing that many organizations have failed to incorporate it effectively into their decision-making processes (Reggio & Astesiano, 2020; Tabesh et al., 2019) and project benefits have not been adequately captured (Marnewick & Marnewick, 2022). Benefits like the detection of malicious URLs especially with a huge volume of data (Li et al., 2020), predictive analytics for a competitive market, data mining techniques, velocity in analyzing real time data, data visualization capabilities for a clear monitoring of business, and specially the "what-if" scenario analysis that is essential for insurance companies that works constantly with risks. All the abovementioned benefits can positively impact analytic governance, strategic alignment, performance evaluation of the organization like analytic process, data management and provide more sophisticated analytic tools and techniques (Pour et al., 2023).

Furthermore, statistics show that only 20% of big data analytic platforms deliver business outcomes (Herschel et al., 2019). One of the top obstacles in this context is the uncertainty of a platform's technical capabilities (Kronz et al., 2021). Therefore, it is essential to understand what best practices are adopted to avoid barriers to successfully implementing a big data project. Theoreticians recognize that different versions of management are required under various circumstances, depending on an organization's country, sector, and size. Thus, it is crucial to expand this research field to include studies from different industries worldwide (Turner & Ledwith, 2018). To date, few studies have attempted to analyze how big data analytics is implemented in the insurance industry. Therefore, this study aims to address this gap and expand the research field by trying to answer the following question: What are the practices and barriers for implementing a big data analytics project in a large Brazilian insurance company? Accordingly, the aim of our article is to demonstrate the practices and barriers related to the implementation of a big data analytics project and suggest improvements to future projects. To address these aims and answer the research question, a case study was conducted on one of the largest insurance companies in Brazil with interviews and document analysis.

As a result, 5 main practices that were adopted to successfully implement a big data analytics project were identified (implement automatic autoscaling alerts, use specialized big data tools, integrate the platform with legacy systems, comply with privacy legislation, and ensure the documentation of technical architecture using business process modeling), as well as 4 barriers that prevent its proper adoption (complexity of access to multicloud data sources, high processing requirements of unstructured data analysis, failure to attend to business meetings at the right time, and project delays brought by bureaucratic interdepartmental processes); some of these have not previously been identified. Finally, an action plan to address these issues is presented.

This study proceeds by reviewing the related literature and then presenting methodology, results, and discussion sections. It finishes with conclusions and a discussion of the theoretical and practical implications of the findings. The relevant high-level interview questions are provided in the Appendix A.

2 Literature review

Advancements in information and communications technologies have resulted in the accumulation of a massive volume of data from various sources and in different formats, often referred to as big data (Alharthi et al., 2017; Mikalef et al., 2020). According to literature, big data has multiple dimensions that vary from the 3 Vs (volume, velocity, and variety) to the 7 Vs (volume, velocity, variety, veracity, value, visualization, and variability), as presented in Table 1 (Sivarajah et al., 2017; Mikalef et al., 2020).

Table 1.

Big Data Dimensions

Dimension	Characteristics
1) Volume	Refers to the enormous scale of data generated by various devices that requires innovative tools for the collection, storage, and analysis of such data.
2) Velocity	Refers to the rate (batch, near real-time, real-time, streams) at which the data are generated or updated, pointing to the real-time nature of big data.
3) Variety	Refers to the variation in types of data (structured, semistructured, unstructured) that can come in diverse and dissimilar forms from multiple sources, such as text, spreadsheets, audio, video, and sensors.
4) Veracity	Refers to the complex data structure that makes data imprecise, ambiguous, and inconsistent. For example, social media data can be biased and inaccurate.
5) Value	Refers to the knowledge/value extracted from vast amounts of structured and unstructured data without loss for end users.
6) Visualization	Refers to presenting data in a readable manner instinctively and effectively through different visual formats such as graphics.
7) Variability	Refers to constant changes in data; this means that the same data can be used in different processes for different purposes.

Source: Based on Sivarajah et al. (2017) and Mikalef et al. (2020).

Big data is seen as one of the most strategic resources of the 21st century; however, many organizations are implementing this new strategic resource (Alharthi et al., 2017). In the words of Gupta and George (2016, p. 5), "The intelligence gleaned from data will be of little use to an organization if its managers fail to foresee the potential of newly extracted insights". Therefore, implementing a successful big data project is not trivial; a team must address a list of barriers by adopting the best technical and managerial practices.

2.1 Big data analytics barriers

Big data offers companies gains in strategy and advantages in competition; however, there are a wide variety of associated challenges, such as a need for smart and big data sources, efficient access to real-time data, very tight project deadlines, and adequate networks for processing data (Araz et al., 2020; Terlizzi et al., 2016). Academics and practitioners have enumerated a list of barriers to fully realizing big data benefits in organizations. These barriers can be classified as managerial or technical in nature and are discussed in Table 2.

Table 2.

Barriers to Big Data Analytics

Barriers	Characteristics	Author(s)
1) Complexity of data (T)	Data are frequently stored in various formats, including unstructured databases and offline text files. Additionally, the volume of data is increasing progressively, making processing this data even more difficult.	(Johnson, 2012; Jones, 2013; Demchenko et al., 2013)
2) Legacy infrastructure not ready (T)	Legacy technologies were not designed to meet the growing requirements of big data analytics. The development of IT infrastructure for this purpose requires significant investments in software and hardware to support the analysis of millions of records in real time.	(turner et al., 2011; Tsai et al., 2023)
3) Ignorance of a technical guideline (T)	The benefits of a controlled process of configuration management are not always understood or realized by the project team. While traditional data sources benefit from time testing and documentation, big data sources are unstructured, and their use can often lead to mistakes. Usually, technical staff members do not use a guideline to define the limitations of these big data sources or workable solutions.	(Ali & Kidd, 2014; Lazer et al., 2014; Brave et al., 2022)
4) Lack of skills among staff (T)	Big data analytics professions will be among the top ten fastest-growing occupations in 2026. However, a lack of such skills may increase data entry errors that could result in the loss of valuable information and limitations on the value a business can gain from the data it captures.	(Hoffman & Podgurski, 2013; Rieley, 2022)
5) Lack of skills among managers (M)	Most big data investments fail due to a lack of technical understanding among managers and their inability to incorporate insights gathered from big data into organizational decisions.	(Collins, 2014; Ross et al., 2013; Mikalef et al. 2018)
6) Costly infrastructure and tools (M)	Developing or buying new data management systems involves significant capital investment in tools and infrastructure required for big data acquisition, storage, and processing.	(Sivarajah et al., 2017)
7) Privacy concerns (M)	Big data analytics often employ personal data collected for an entirely different purpose. Combining personal information with other data sources can entail numerous legal and ethical challenges, such as a potential leak of confidential information about individuals (e.g., medical records, financial situation, embarrassing behavior, and family relationships).	(Jones, 2013; Rijmenam, 2014; Price & Cohen, 2019; Krotov & Johnson, 2022)
8) Resistance to adopting big data (M)	Some organizations lack an overall understanding of how big data can improve their business operations; consequently, they see little value in these initiatives, which results in organizational resistance to big data.	(LaValle et al., 2011; McAfee et al., 2012; Villarejo-Ramos & Cabrera-Sánchez, 2019)
9) Lack of a data-driven culture (M)	Executives in many organizations rely heavily on their prior experiences or intuitive feelings instead of following evidence-based decision processes, as explained by the Upper Echelons	(Hambrick & Mason, 1984; Hambrick, 2007;

	Theory. If top managers do not value data-driven decision-making, their behavior will make implementing a big data platform difficult.	McAfee et al., 2012; Ross et al., 2013; Tabesh et al., 2019)
10) Different visions about big data and the associated goals	It is not trivial to disseminate a unified vision of organizational big data among all managers and staff. Unfortunately, many corporate decision-makers lack a shared understanding of what big data analytics is and what benefits it can provide for their business operations and outcomes.	(LaValle et al., 2011; Rogers et al., 2007; Terlizzi et al., 2017)

Note: (T) Technical, (M) Managerial
Source: Authors (2023).

2.2 Big data analytics practices

If there is a diversity of barriers to successfully implementing a big data project, it is not surprising that there is also a diversity of practices used to overcome these issues. Previous literature demonstrates that the big data practices adopted by organizations can differ according to the organization, sector of the economy, and country. These practices can be classified as managerial or technical in nature and are discussed in Table 3.

Table 3.

Big Data Analytics Practices

Practices	Characteristics	Author(s)
1) Use specialized tools and algorithms (T)	Specialized software tools and algorithms should be used to store, manage, and analyze complex data in an efficient, reliable, fast, and economical manner. Examples: MapReduce, Hadoop, MapR, Cloudera, and Hortonworks.	(Rijmenam, 2014; Francisco et al., 2019)
2) Integrate legacy with big data platforms (T)	Building a new and independent big data platform is the best option for new organizations. However, it is more feasible for organizations with legacy IT systems to integrate modern big data platforms into their legacy systems.	(Davenport & Dyché, 2013; Tsai et al., 2023)
3) Adopt a technical guideline (T)	Configuration management plays an integral role in maintaining the integrity of information throughout the project life cycle by controlling changes. Technical staff must analyze the pros and cons of using a big data source with the support of a methodology. For example, SMALL is a checklist based on five questions that should help big data users draw justifiable conclusions and prevent mistakes in interpretation.	(Whyte et al., 2016; Brave et al., 2022)
4) Collaborate with educational institutions (T)	Organizations should collaborate with educational institutions to align their curricula with the industry requirements for big data skills. In addition, industry and academic institutions should collaborate on providing practical training to address missing skills.	(Miller, 2014; Jin & Yao, 2022)
5) Encourage internal and external learning sessions (M)	Educate managers and staff with a working knowledge of data science. Motivate employees to gain understanding from external sources (e.g., degrees, certificates) and organize internal knowledge dissemination with frequent learning sessions between data scientists and managers.	(Dykes, 2017; Zettelmeyer, 2015)
6) Commit to long-term investment (M)	Implementing a big data platform is a long-term investment and depends on the commitment of top management. Therefore,	(Crittenden & Crittenden,

Practices	Characteristics	Author(s)
	managers should not expect to generate immediate returns, as this process takes time to pay off.	2008; Vidgen et al., 2017)
7) Comply with privacy legislation (M)	Organizations must understand and comply with privacy legislation and incorporate general best practices for handling sensitive customer data into their policies and operations. In addition, it is crucial to implement controls to protect the privacy of individuals whose data is collected (e.g., to anonymize data and store it in a secure and encrypted database).	(Alharthi et al., 2017; Krotov & Johnson, 2022; Terlizzi et al., 2019)
8) Define a clear big data strategy (M)	Successful cultural change can be achieved by having a clear organizational vision concerning big data, ensuring top management commitment, and accelerating its acceptance.	(Alharthi et al., 2017; Rogers et al., 2007)
9) Create the position of Chief Data Officer (M)	The coordination of data-driven decision-making processes at all levels of an organization is an essential responsibility of managers in addressing the challenges related to big data. This importance demands new executive positions, such as Chief Data Officer, to restructure the process and facilitate communication.	(Aiken & Gorman, 2013; Nie et al., 2019)
10) Establish a mutual understanding of big data goals (M)	Managers must communicate big data-related business goals to their teams. Building a multiskilled team comprising data scientists and engineers with technical and business knowledge is essential. This practice helps create a rich culture of open communication.	(Davenport & Patil, 2012; Mayhew et al., 2016; Mikalef et al., 2020)

Note: (T) Technical, (M) Managerial
Source: Authors (2023).

2.3 Big data analytics in insurance companies

The insurance industry has grown rapidly all over the world. Gross insurance premiums worldwide have more than doubled over the past twenty years, increasing from US$ 2.5 trillion in 2000 to US$ 5.2 trillion in 2020. Furthermore, insurance spending (the ratio of gross insurance premiums to gross domestic product), which is an indicator of the importance of the insurance industry within the domestic economy, also increased in the same period, rising from 8.6% in 2000 to 9.4% in 2020 (OECD, 2021). In Brazil, the insurance industry has also stood out, reaching an annual revenue of US$ 104 billion in 2021 and an insurance spending rate of 6.3% (CNSeg, 2022). Moreover, according to the Global Insurance Potential Index, Brazil ranks eighth in terms of potential for growth among the countries of the world and first within Latin America (Mapfre, 2022).

For many years, insurance has been a vital service to people who insure various assets, including homes, automobiles, businesses, and even their own lives (Bohnert et al., 2019). However, like other sectors, the insurance industry has been facing severe competition due to the advent of modern technologies (Moraes et al., 2017; Yang & Zhou, 2021). Therefore, to be more efficient and profitable, these companies are optimizing operational areas such as marketing, underwriting, claims, and reports through the aggressive adoption of big data analytics (Venkatesh, 2019).

Big data analytics has also added precision to actuarial work, leading to more accurate pricing and risk estimation. For example, new digital data sources reveal information about behaviors and lifestyle habits, allowing insurers to assess individual risks much more accurately. In addition, predictive analytics can help actuaries evaluate enormous amounts of data quickly and incorporate data that were previously either unavailable or considered irrelevant (Venkatesh, 2019).

In the insurance industry, a product's price must be set before its cost is known. This fact makes the process of rate making critical and significant. Indeed, insurance firms must predict how many claims will occur and the severity of those claims to set fair prices for their insurance products. For example, in the automobile insurance sector, claim prediction is the cornerstone of premium estimates. Furthermore, it is crucial to assign the correct insurance policy to each prospective policyholder in the insurance sector. A failure to foresee auto insurance claims with accuracy increases the cost of insurance policies for excellent clients and decreases the price of the product for bad clients (Hanafy & Ming, 2021).

Big data use has been changing how the insurance sector works. Insurers have vast amounts of data and are exploring diverse ways of utilizing these data (Arumugam & Bhargavi, 2019; Sood et al., 2022). Some previous literature has examined how wearables improve customer satisfaction and engagement (Seth & Gulati, 2022). Moreover, systematic literature reviews have found that social media is the most frequently cited modern technology (Zia & Kalia, 2022). Other work has found that the insurance sector has begun to develop big data solutions based on blockchain in the contexts of fraud prevention and risk assessment (Trivedi & Malik, 2022).

3 Research methodology

In Brazil, the implementation of big data projects in the insurance industry is recent, and, following a review of the literature, it was apparent that there is a lack of both qualitative and quantitative research in this area. As the subject under investigation is new and there was little existing research, the objective of this study emphasized the need for exploratory research. Therefore, the empirical research conducted in this study employed the case study method. Case studies provide researchers with an opportunity to understand the conditions that are present in a particular situation (Yin, 2017), and they are frequently used in IT studies (Sarker et al., 2012);

in particular, they are appropriate for investigating management aspects rather than technical aspects (Benbasat et al., 1987).

3.1 Case study design

The case study design is the logic that links the data to be collected with the conclusions, thus ensuring coherence in defining the research question; defining the unit and period of analysis; linking the data to the research question; and defining the criteria for interpreting the findings (Yin, 2017). It is recommended that data be collected from various sources. For example, interviews are one of the most important sources of facts, and documentation supports the use of evidence from other sources (Yin, 2017).

The insurance industry is one of the largest investors in technology; thus, it is expected to utilize superior IT management techniques (Berghout et al., 2011; Deloitte, 2021). To address the research question, a case study was conducted on a leading insurance company in Brazil, representing 20% of the industry's Net Income. Using the analogy of Siggelkow (2007): organizations that are ‘talking pigs’ can be used as sector benchmarks.

The studied organization has not authorized the disclosure of its name. It is one of the largest insurance companies in Brazil and the leader in the automobile, residential, and commercial segments. The company has an operational structure and distribution chain consisting of approximately 14.000 employees, 35.000 independent brokers, 12.000 service providers, and 120 branches across the country. It was founded in 1945; has more than fifty products in its portfolio, 11.3 million clients, possesses financial assets worth US$ 8.7 billion, and is considered one of the most valuable brands in Brazil (Interbrand, 2022).

In 2011, the company launched a Tier 3+ data center to support the increasing volume of services, call centers, and others that depend on digitalized platforms. However, currently, many databases are migrating to various cloud services, such as Microsoft Azure, Google Cloud Platform and Amazon Web Services. For this reason, the company also works with on-premises solutions such as multicloud data services. Orchestrating all these data flows is a relevant challenge to the company, as it must take trustworthy information to specific areas for analytics and monitoring in an organized and agile way without compromising the crucial services that are operating.

An organization is not necessarily required for the case but may be the specific context of the case (Martinsuo & Huemann, 2021). The unit of analysis used in this study is the IT project through which the big data platform has been implemented in in the focal company, especially the module that supports pricing strategy. There is no recommended number of interviews to be performed, but it is suggested to reach saturation and report the number of interviews performed in a study (Sarker et al., 2012). We reached saturation by collecting data through ten interviews with all professionals engaged in the project: technicians, managers, and executives. The interviews took place from October to December 2022, and during this period, internal documents were also collected from the company's Knowledge Management System, such as technical guides, policies, the business case, and project presentations. Hence, a high-level profile of these professionals is displayed in Table 4. The objective of the interviews was to investigate the main barriers to successfully implementing a big data analytics project and the main practices employed to do so.

Table 4.

High-Level Profile of The Professionals Interviewed

Professional	Years of practice	Responsible for
1) Director of pricing	14	Managing the different areas of pricing, appropriately positioning all the products of the company, and ensuring superior results according to the stipulated goals.
2) Manager of pricing	14	Managing the portfolio of twenty products in terms of pricing strategies and policies, management reports, market monitoring, developing new products, and risk analysis. Managed a team of twenty-eight high-performance analysts with different skills.
3) Coordinator of pricing	6	Managing pricing analysts technically to ensure that the right actions and decisions take place regarding each element of pricing. Additionally responsible for monitoring pricing actions and taking efficient and assertive action.
4) Director of product	2	Studying the market and opportunities for innovations, actions, and changes needed to keep products competitive with an unobstructed vision of the products he is responsible for.
5) Manager of product	9	Monitoring the rules and new opportunities related to the products he is responsible for. Is responsible for keeping in constant touch with clients to understand their needs and thus keep products as competitive as possible to maintain clients' satisfaction.
6) Coordinator of product	5	Maintaining thorough technical knowledge of the relevant rules and closely monitoring products' performance to keep products working correctly and change rules rapidly, especially in situations where the market changes, clients or brokers complain, or opportunities arise to be more competitive according to market monitoring.
7) Product owner	6	Integrating business necessities into the project, making sure it will be done according to the plan of each sprint, having a systemic view of all parts of the project, including the business needs, the technological implementation, the construction of dashboards for monitoring, analytics, and reporting on the project to the coordinators, manager, and director.
8) Scrum master	4	Organizing the activities inside the roadmap of the project such that they are finished within the period stipulated. When the developers need support from other areas, the scrum master helps with the best way to address situations that are sometimes complicated due to the company size.

Professional	Years of practice	Responsible for
9) Technology leader	9	Managing the developers in choosing the best technological alternatives and solving or avoiding problems that the developers sometimes cannot see. The technology lead must release the final project according to the necessities.
10) Developer	5	Implementing the elements of the project, such as the connections, data flows, and cloud settings, and making the variables for the development and analysis requested by the product owner available.

Note. Source: The authors.

During the interviews, the professionals presented some of the project's technical and management documents to facilitate understanding of the process. The non-confidential documents were shared with the authors by e-mail and certain relevant excerpts were reported in the study. Each interview was recorded using MS Teams and transcribed to MS Word as soon as possible afterward, as recommended by Miles and Huberman (1994). The interview transcriptions were read several times so that the researchers could become familiar with the data in greater detail (Eisenhardt, 1989). After that, using MS Excel, the main quotes extracted from the answers were coded individually by the authors and then grouped (Larsson, 1993). Finally, the grouped quotes were compared with the practices and barriers identified in previous literature.

4 Result analysis

This section discusses the motivations to integrate pricing strategy into the project, then it provides an understanding of the project development. It finishes by explaining in detail the implemented platform.

4.1 Project motivators

The pricing department of an insurance company is one of its most critical departments and is responsible for many factors that affect the final prices of products, such as risk modeling, commercial adjustments, reserves, and actuarial and other analytics studies intended to maintain products' health and ensure good results for the company (Arumugam & Bhargavi, 2019). Before the implementation of a big data platform in the focal company, the pricing department suffered in relation to managing data for analytical proposals and monitoring actions. Velocity is a crucial element of pricing strategy because changes in the environment must be quickly captured by big data analytics models to preserve and improve a company's results. The main big data challenges faced by this department are presented in Table 5.

Table 5.

Big Data Challenges of The Pricing Department

Challenge	Characteristics
Traditional architecture	The volume of data increases yearly, with new systems and data sources being created all over the company in a disorganized way. The fact that the company does not have a defined architecture or a data lake for centralizing all this information is a major issue for the pricing department, which depends on collecting this data to run its big data models in real time.
Restricted access to data	The pricing department depends on access to databases from all over the company; however, these databases are segregated according to insurance products (life, automobile, etc.) and owned by different departments. Thus, when the department does not have access to the database, it is necessary to request a timely extraction from the data owner, which is not ideal for analytics because such data may be static and obsolete. Moreover, queries and extractions depend on many intermediated platforms that can be overloaded, and it may take long to complete a request; this impacts business strategy, as the data may be too late for decision-making and actions.
Unavailability of data	The availability of some products' data in the final database is quite delayed due to volume and treatment processes. When such a product's data are available, certain important variables for risk analyses or monitoring are not accessible in the database, making decisions less supported.
Noncentralized processes	The data are separated according to the stage of the policy: quotation, subscription, transmission, issuing, incidents, renovation, and changes in contracts; therefore, the unification of the information from various stages is complex and requires many manual processes in the pricing department.
Low performance	Constructing databases takes too long and does not result in superior performance due to daily queries from all departments and a dependence on many intermediated servers that are attached to many intermediated processes.

Note. Source: The authors.

In the insurance industry, technology has improved, but value realization must be the focus (Terlizzi et al., 2017; Terlizzi et al., 2014). This industry must integrate systems and data while leveraging the cloud, focusing on increasing the centricity of the customer (Deloitte, 2023). To follow the market and meet company needs, the pricing department must act quickly and overcome its challenges. Establishing a real-time accurate pricing process is crucial to taking advantage of new market opportunities while mitigating the risk of losing money. Thus, the implementation of a big data platform is required to guarantee an efficient decision process and improve company competitiveness.

The company has been growing rapidly over the last few years; however, it is still based on a system architecture of silos, where more than twenty products are stored in decentralized databases with different data management maturities. On the one hand, some products are new and were constructed under a modernized system on the public cloud; on the other hand, the data of some old (legacy) products are stored in more than one environment.

The company faces a classic big data problem due to the enormous volume of data (Sivarajah et al., 2017), the velocity at which information is being produced all around the organization, the variety of different types of data and sources, the value of such data for analytics purposes, their veracity for precise decision-making, the directives of visualization according to business strategy, and all the variability in the data. This situation can be observed in the interviews when we question the professionals about the project motivators, as described in Table 6.

Table 6.

Project Motivators

Motivator	Quotes
1) Analyze the high **volume** of data available in the company (Mikalef et al., 2020).	*"In a digital era, the more information we have, the better position in the market we can achieve and more assertive actions we take."* (Scrum master, #8) *"The project also opens the possibility of looking at the right product at the right moment due to the [large] number of products we work on."* (Manager of pricing, #2) The page in the documentation (Figure B1 – Appendix B) shows the use of the main dashboard that controls all the products in real-time, due to the substantial number of products. Each visualization controls the data during the day of all products in parallel. Automatic alerts notify anomalies and rules according to the business strategy. *"The project also opens the possibility of looking at the right product at the right moment due to the [large] number of products we work on." (Manager of pricing, #2)* The page in the documentation (Figure B1 – Appendix B) shows the use of the main dashboard that controls all the products in real-time, due to the enormous number of products. Each visualization controls the data during the day of all products in parallel. Automatic alerts notify anomalies and rules according to the business strategy.
2) Increase **velocity** of data and allow access to real-time data (Mikalef et al., 2020).	*"We have the intention to be as analytical as possible, as having high-velocity data is fundamental."* (Director of pricing, #1) *"Data use must be fast to have a much lower impact."* (Manager of pricing, #2) *"Some products suffer from a delay in information, especially due to the segregated data sources and different data maturities they have."* (Product owner, #7) According to the project's documentation, this Big Data architecture allows access to data in real-time with no need for software installed. Just accessing a Elasticsearch and Kibana according to Figure 7.
3) Explore the potential of the data **variety** that we have in the company (Mikalef et al., 2020).	*"The possibility to have more variables also opens many opportunities to have more sophisticated algorithms for business, especially in a department that works with pricing and risks."* (Product owner, #7) *"The project also opens the possibility of looking at the right product at the right moment due to the variety of products we work on."* (Manager of pricing, #2) The documentation shows the available variables in the search mechanism from different data sources in a unique environment for analytics and visualization purposes (Figure B2 – Appendix B).
4) Ensure data **veracity** to support better decisions (Sivarajah et al., 2017).	*"A trustworthy database where we have access to this information is precious, because it drives us and shows if we are going the right way or not." (Manager of pricing, #2)*

Motivator	Quotes
	"Precise information that you can use to drive the strategy and the business in a much more grounded and secure way." (Manager of product, #5) The automatic Monitoring of Elastic uses real-time alerts to guarantee a trustworthy database according to the quality rules implemented. This alert shows the status of the environment for detection and correction. This mechanism also shows alerts to the strategy of each product for faster decision-making and pricing action. Figure B3 (Appendix B) shows an example of the projects' documentation for the use and understanding types of alerts.
5) Obtain **value** from the vast amounts of data generated in the company (Alharthi et al., 2017).	"*When it comes to pricing, especially in insurance, selling is important when I realize by using the data that I'm acting wrong, more intensely or less intensely or that the environment is changing*." (Manager of pricing, #2)
6) **Visualize** information in an intelligent manner to facilitate taking actions (Alharthi et al., 2017).	*"Some products we don't sell more if the price is higher or lower, but for other ones, if you press a button, you going to change its behavior."* (Director of pricing, #1) *"We have improved in monitoring portfolios, business, performance, quality, the business vision, and its operation, control, and results."* (Coordinator of pricing, #6) The project's documentation explains each visualization possibility that the menu of creation in Kibana offers (Figure B4 – Appendix B). The documentation extracted explains possibilities to create visualizations in an easy form, like drag and drop editor, maps, and aggregation views until advanced visualizations like anomaly detections, machine learning, and others that use.
7) Analyze the **variability** of data from different angles (Sivarajah et al., 2017).	*"The first study we must do is to find out which products are more sensitive to moving the price. [...] We must build a matrix where we have pricing sensitivity vs. results."* (Director of pricing, #1) The Automatic Monitoring of the Elastic, as shown in Figure B3 (Appendix B), analyses through advanced algorithms, the factors impacting each product for a better and faster action. These algorithms read the data from the architecture in real-time by intelligent monitoring and suggest important points, giving different perspectives for business strategy actions.

Source: Authors (2023).

4.2 Dedicated team

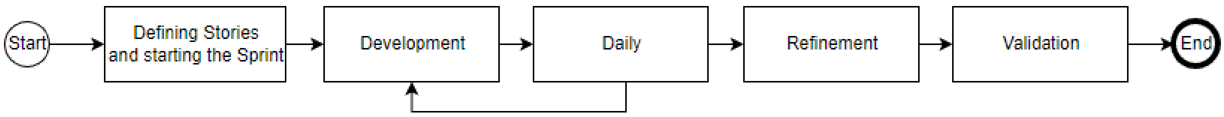
Implementing a big data platform in the studied organization has not been a trivial task. The traditional system architecture organized in silos, the involvement of many different departments and the enormous size of the company have brought great complexity since the beginning of the project. For this reason, first, the product owner elaborated an overview of the project, depicting the team budget, roadmap, and benefits. The product owner's formal power has a positive effect on ensuring project success (Sanchez, 2017).

The sponsor of the project, the director of pricing, approved the project, and the team organization started. A dedicated team was formed with seven members: the product owner, the scrum master, the technology lead, and four developers (responsibilities are described in Table 4). The creation of a dedicated team sponsored by the executive was a necessity and considered essential to project success (Sithambaram et al., 2021).

> *"Although it is a big, traditional company, we have fewer changes than a startup, in which the process is more agile, faster, and shorter."* (Manager of pricing, #2)
>
> *"Large backlogs and concurrence of projects is one of the main difficulties that we face while managing projects in the company. However, this project has superior performance because of the dedicated team, especially in terms of quickly implementing improvements in the platform."* (Director of pricing, #1)

Figure 1.

Project Development Routine



Source: Adapted from project documentation.

To develop this project, the team uses a flowing routine based on agile methodology (PMI, 2017), as demonstrated in Figure 1 and detailed in Table 7.

Table 7.

Project Development Routine

Process	Description
Defining stories and starting the sprint	The product owner defines the stories (the tasks to be developed during the sprint) for a two-week-sprint. With all the tasks mapped, the scrum master starts the sprint according to the business strategies described in the stories.
Development	According to the defined stories, the developers and technology lead start the development of the tasks that will be concluded at the end of the sprint. Usually, it is necessary to have support from the product owner in guiding the implementation regarding business rules and correctly integrating them into the technology to effectively provide value to the business.
Daily	With the objective of ensuring that everything is going correctly as defined, the team has a daily 15-minute meeting as a checkpoint to report their advances and the problems they are going through to the scrum master.
Refinement	After two weeks, they participate in a refinement meeting, where they discuss the effectiveness of the sprint. After this meeting, if all the tasks were completed and validated by the product owner, the sprint is ended. If not, they start a new sprint to finish the old tasks and any new tasks requested by the product owner. This is especially important to increase the experience and measure the productivity capacity of the team.
Validation	At the end of the sprint, the product owner checks the results with the coordinator and manager of pricing to validate or adjust the tasks. If the result is not ideal for them, a new story is defined to start in the next sprint.

Source: Adapted from project documentation (2023).

One of the most important aspects of the project is its focus on areas that are directly impacting the tasks due to the difficulty of finding an appropriate analyst to address them or drive the case to the right department, especially since this is a disruptive project and not familiar to the company.

> *"It is hard for the areas to understand the value and the changes that are necessary to make in the process and rules. We had so much difficulty explaining the project to the analysts in many departments, and sometimes we had to advance the topic to superiors; it was very onerous in the process."* (Product owner, #7)

> "*Our priority is not the priority of other departments that we have a dependency on in developing the project, especially in terms of capturing data for analytics*." (Scrum master, #8)

To guarantee an adequate flow of team tasks and efficient access to the several databases spread all over the company, the product owner has established a close partnership with four key technical departments (Table 8): (1) Information security, responsible for the security of traffic, routing, and access to data from the source to the endpoint; (2) IT transactional data, responsible for collecting and storing the real-time data received from clients; (3) cloud architecture, responsible for controlling the multicloud and on-premises routing between the data storage locations; and (4) data engineering, responsible for the treatment, accuracy, and quality of the data to be utilized.

One practice observed by the researchers is the extensive documentation made available using business process management notation, considered vital by the technical team.

"Business process management is gold, the register of the business; it is an intellectual register because if you detail the process, you know the areas of their involvement and risks. Some IT departments do not have a practice of documenting processes, and it makes the construction slower and more complex; on the other hand, we gain velocity." (Scrum master, #8)

Table 8.

Areas Involved in the Project

Department	Responsible for	Relevance to the project
Information security	The security of traffic, routing, and access to data.	Certifies all the steps of the project to make sure they will operate with company policy and helps speed up bureaucratic demands.
IT transactional data	Collecting and storing the real-time data received from clients (via e-commerce or brokers).	Indicates the sources of the several databases spread all over the company and ensures they are available for distribution.
Cloud architecture	Controlling the multicloud and on-premises routing between the data storage locations.	Supports the creation of a repository in the public cloud for storing and processing data in real-time. Transactional data

Department	Responsible for	Relevance to the project
		are stored with Amazon Web Services and processed on the Google Cloud Platform using a tool called "Big Query."
Data engineering	The treatment, accuracy, and quality of the data to be utilized, normalizing the data, and preventing the use of sensitive information.	Provides comprehension through the documentation they provide on the rules applied to the data for more effective validation. This department provides data via the Google Cloud Platform or in the local data center (on-premises) for products that are not on the cloud. The external data are from other services or other sources that are not created at the company; these data are used to enrich the transactional data.

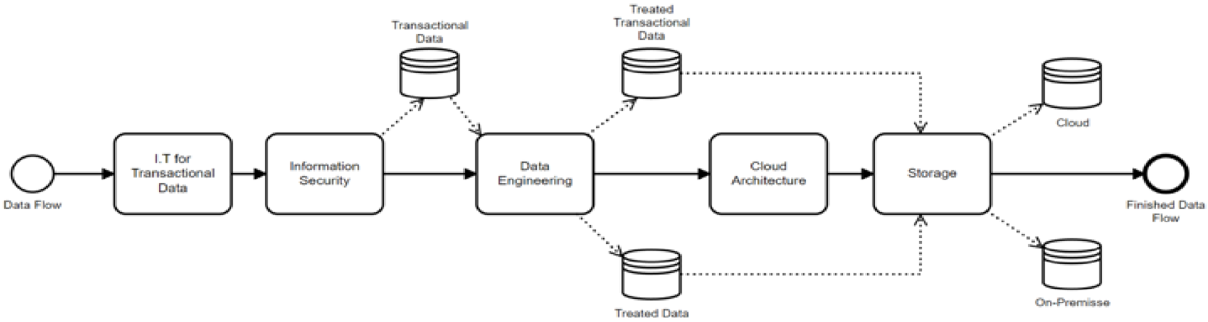
Source: Adapted from project documentation (2023).

In Figure 2, there is an original architecture of the macro–Business Process Management (BPM) of areas and Data Flow that helps so much the communication and development of the Big Data Project, since the Transactional data is mapped such as the areas that is responsible for the flow or even could be impacted. For this reason, this BPM makes it easier to understand the original architecture to develop the one presented in this research.

After the development, a new BPM was created as shown in Figure 2.

Figure 2.

Business Process Management of Involved Areas and Data Flow



Source: Project documentation

4.3 Final solution

The proposed solution should be capable of accessing data from multiple sources, efficient in processing structured and nonstructured data, and quick in attending to new business priorities.

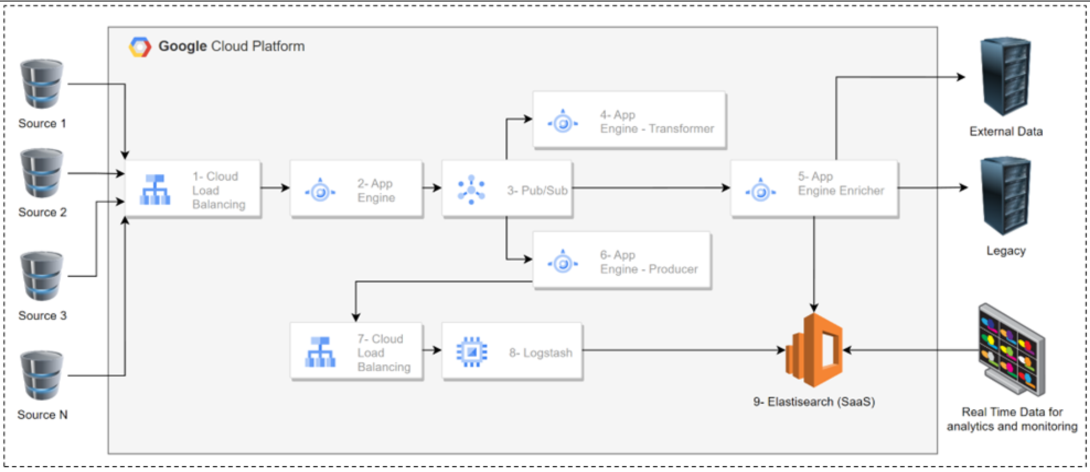
> *"Stop thinking old and start new things. Using data that we offer is more effective for the client. We must prioritize things that are a priority."* (Coordinator of pricing, #3)
>
> *"There are different systems spread throughout the company with much-dispersed information and difficulty in governance within each one of these systems."* (Manager of pricing, #2)
>
> *"The problem is preparing all the processes to become more digital, especially for the legacy we still have"* (Director of pricing, #1)
>
> "*The architecture in the company is a little bit conservative; some applications are two or three versions older*." (Developer, #10)

The proposed solution must follow three main guidelines to be in compliance with company policies: (1) no direct connection should be attached to the database, but a passive service should be used to receive the information that the transactional department is responsible for sending; (2) it is necessary to anonymize data according to the regulations of the company. The security department is responsible for checking policies and following the implementations of the project team; and (3) an agreement from data engineering is needed stating that data will be consumed before their stage.

In Figure 3, which shows the final real-time architecture, the connections starting from cloud load balancing can be observed. All the architecture was developed on the Google Cloud Platform. In general, this real-time data architecture provides autonomy in information control, with no intermediation of processes that can compromise efficiency in monitoring and accessing data for analytics.

Figure 3.

Final Real-Time Architecture



1- Cloud Load Balancing: distribute the user traffic across multiple instances.
2- App Engine: treat and enrich data.
3- Pub/Sub: communicate with the other components responsible for processing documents.
4- App Engine Transformer: normalize data using JSON to a better performance.
5- App Engine Enricher: read external data sources, such as legacy and external services, to enrich the real-time data with more variables to enhance their analytical potential.
5a- External Data: data created in the company or acquired from other companies.
5b- Legacy: historical data (client behavior, lifetime value, some products, and others).
6- App Engine Producer: generate a JSON file from the results of the processing.
7- Cloud Load Balancing 2: control the autoscaling resource due to the variability of data.
8- Logstash: collect, process, and store the data.
9- Elasticsearch: search mechanism and data analysis tool for monitoring the data flow.
10- Kibana: advanced visualization, alerts, and monitoring.

Source: Project technical documentation (2023).

This is an architecture of passive connections, where there are no manual queries and there is minimal impact on the corporate architecture. The objective is to have a high-performing architecture that is as noninvasive as possible. The main idea is to eliminate the line, reducing the heavy processing of each document with specific threads.

5 Discussion and conclusion

Based on the literature review and the discussion of the results of the case study, it is possible to answer the research question presented in this study: "What are the practices and barriers for implementing a big data analytics project in a large Brazilian insurance company?". Our research is important because it attempts to extend conceptions regarding the studied phenomenon of big data analytics projects in new directions. We advance the literature by adding two new findings that were not identified on previous literature: one new practice and one new barrier. Moreover, we corroborated others practices and barriers previously identified.

In this study, the implementation of a successful big data analytics project in an insurance company was observed, especially in relation to pricing. The new platform was constructed without affecting the original processes of the company. It was possible to identify the practices and barriers used by the project team. Concerning the practices observed in the studied organization related to the implementation of a big data analytics project, we found one new practice and corroborated four others identified by previous literature:

(1) Automatic autoscaling alert (technical). An automatic autoscaling and consumption alert was implemented. This is important because the finance department does not have specific information on the platform, just a financial overview of the costs. This component facilitates closer monitoring, cost alerts, and rapid adjustments to performance as needed and is essential for the autonomy of specific analytical studies of the architecture in relation to the breakeven point between cost and performance for products' sustainability, improving the platform's delivery value (this is a new finding not previously identified in the literature).

(2) The use of specialized big data tools (technical). Tools such as Logstash, Elasticsearch, and Kibana are essential for the scalability, support, and stability of the platform. As mentioned by Alharthi et al. (2017, p. 286): "*these massive data sets require new and innovative tools and approaches for capturing, storing, and analyzing data*". Specifically, the use of data visualization tools with visual elements such as charts, graphs, and maps provide an accessible way to understand trends, outliers, and patterns and thus simplify the decision-making process. This corroborates the results of previous studies (Alharthi et al., 2017; Rijmenam, 2014).

(3) Integrate with legacy systems (technical). Collecting data from legacy systems is the best way to guarantee data integrity. As mentioned by Davenport and Dyché (2013, p.20): "*at the majority of big companies, a coexistence strategy that combines the best of legacy data warehouse and analytics environments with the new power of big data solutions is the best of both worlds*". Legacy systems, or transaction processing systems, are the primary storage locations of data received from stakeholders, such as clients, suppliers, and employees. This corroborates the results of previous studies (Davenport & Dyché, 2013; Mikalef et al., 2020).

(4) Comply with privacy legislation (managerial). Acting in accordance with general data protection regulations, called "Lei Geral de Proteção de Dados" in Brazil, is not an option;

instead, it is an obligation and was established as a premise of the project. As mentioned by Krotov and Johnson (2022): "*If a potential legal or ethical issue surrounding big data is detected, this does not mean that the research project cannot be conducted. Numerous workarounds can be implemented to mitigate various ethical and legal problems*." Thus, all sensitive data must be anonymized before being stored in the big data analytics platform and are regularly assessed by the information security department. This corroborates the results of previous studies (Alharthi et al., 2017; Krotov & Johnson, 2022).

(5) Adequate use of technical documentation (technical). Traditional data sources from legacy systems have the benefit of being time-tested, and their vagaries are understood well and documented widely, but this was not the case for big data in its infancy, which led to mistakes (Lazer et al., 2014). However, this scenario evolved, and data sources from big data platforms are better documented (Brave et al., 2022). In our study, we observed an extensive use of standardized diagrams documenting the flow of activities and interactions between systems. The business process model and notation were adopted to facilitate understanding among the stakeholders involved in the project and avoid mistakes. This corroborates the results of a previous study (Brave et al., 2022; Whyte et al., 2016).

Regarding the barriers that can prevent the successful implementation of a big data analytics project, we found one new factor and corroborated three other factors identified by previous literature:

(1) Complexity of access to multicloud data sources (technical). Previous literature cites the difficulty of connecting with legacy systems because older architecture was not designed to support the high processing requirements of big data (Trelles et al., 2011). However, currently, a new barrier has emerged, that is, the complexity of integrating a big data platform simultaneously with multiple incompatible public cloud platforms from different suppliers (this is a new finding).

(2) High processing requirements of unstructured data analysis (technical). As commented by Demchenko et al. (2013, p. 49), "*volume is the most important and distinctive feature of big data which impose additional and specific requirements to all traditional technologies and tools currently used [...] demanding infrastructure components and management tools that allow fast infrastructures and services composition, adaptation and provisioning on demand for specific research projects and tasks*". In the same vein, we observed that in the studied company there are many noncentralized databases spread across

the studied company, especially those containing nonstructured data, which require elevated levels of processing power for absorption, enrichment, storage, analysis, and visualization. This corroborates the results previous studies (Johnson, 2012; Jones, 2013; Demchenko et al., 2013).

(3) Failure to address business necessities at the right time (managerial). Rushing to reach business targets makes business units develop their own big data solutions that are not aligned with the corporate solution. In addition to increasing the inefficiency of duplicating tasks in the company and putting the ongoing project in check, this introduces a major risk, that is, the construction of multiple independent big data analytics platforms in the same organization. This corroborates the results of previous studies (Collins, 2014; Rieley, 2022; Ross et al., 2013; Mikalef et al., 2018). As mentioned by Mikalef et al. (2018), "*this problem is particularly critical for industry executives since [...] the lack of human skills and knowledge relating to big data and analytics is one of the main barriers in achieving business value*".

(4) Project delays brought by bureaucratic interdepartmental processes (managerial). "*Managers should stay committed to data-driven decision making and be persistent in providing structural support for big data initiatives*" (Tabesh et al., 2019, p. 353). However, as we observed, c onstructing a new big data analytics platform in a large company is not a trivial task; thus, this kind of IT project depends on several key people and departments (information security, data governance, IT architecture, IT infrastructure, etc.) for successful implementation. These for successful implementation. T these excessive dependences on divisions that have defined internal processes with established high service level agreements compromise the planned project schedule. This corroborates the results of previous studies (LaValle et al., 2011; McAfee et al., 2012; Tabesh et al., 2019).

Besides the findings, right after the conclusion of the project development, it had some successful value to the business that impacts positively the financial investments (payback) on the project, like (i) Detection and correction of null variable that impacts the pricing due to time-out problems, reducing from 18% to 3%, almost 3% more in revenue; (ii) fraud was reduced due to the intelligent algorithms for detection (iii) fast alerts in regions where pricing was not calibrated made the results increase due to fast actions.

5.1 Practical implications

The researchers expect that the findings of this study will complement those of previous research in project management and will be of interest to practitioners. Thus, this study has pointed out previous literature related to the adoption of a big data analytics platform in an organization, but simply approving the development of such a platform is not sufficient to ensure project success. If the level of adherence is low, it becomes necessary to identify and remove the barriers preventing the platform's adoption.

Exploring the causes of a problem can enrich the understanding of a given theory and allow readers to make more sense of complex organizational phenomena (Whetten, 1989). Therefore, the authors considered that the practical implications of this research would be incomplete if they only identified barriers and did not propose solutions for the cause of the problems. Thus, the study results were presented and discussed with the professionals who were previously interviewed. In this way, it was possible to validate the interpretation of the results with the support of the organization's specialists and to suggest an action plan to enhance the project's next steps.

Table 9 presents the study contribution to practitioners by proposing an action plan for the removal of the main barriers that make project implementation difficult. By the end of this study, the new proposed practices had been embraced by the team and were being evaluated and calibrated within the new project phase.

Table 9.

Action Plan

Barriers	Action plan	Comments
Complexity of access to multicloud data sources (T)	- Prevent overload; do not connect the big data platform directly to the transactional databases of the company. - Centralize all the necessary databases in a unique place (data lake). - Create a unique passive connection to integrate the big data platform into the multiple data sources: on-premises, multicloud, and mainframe.	*"The connection of different databases and their availability in unique and trustful place, facilities fast analysis, especially for always being updated."* (Product owner, #7) *"Preventing overload and using the passive connection was one of the most important gains to our clients, due to the easy and stable access to the database with no compromising the original architecture."* (Developer, #10)
High processing requirements of unstructured data analysis (T)	- Implement the platform on the public cloud to benefit from state-of-the-art technology. - Configure autoscaling to accelerate processing only when needed. - Use the search and analytics engine Elasticsearch to facilitate and optimize the processing of unstructured data.	*"The unstructured data in an engine that works perfectly with this type of technology, allied to autoscaling project in the cloud provides a high-performance environment independent of the number of simultaneous accesses."* (Developer, #10) *"Unstructured data with an appropriate engine was the key success for the project."* (Technology leader, #9)

Barriers	Action plan	Comments
Failure to address business necessities at the right time (M)	- Form a resolute big data analytics team that includes both business and IT professionals. - Explain clear objectives and key results to the team. - Use agile methodologies to ensure the correct backlog prioritization and constant delivery of solutions (minimum viable product). - Keep your technical team trained and up to date in analytics skills.	*"The alignment with all professionals impacted by the project was essential to the efficacy use of the platform, it shows that communication and sharing the benefits in big projects optimizes the use for a more assertive business necessity"* (Coordinator of product, #6) *"Agile methodologies accelerate the project, the solutions and escape from traditional flows, keeping the team in a constant learning."* (Product owner, #7)
Project delays brought by bureaucratic intradepartmental processes (M)	- Identify people and departments that are key to the project success. - Create partnerships with key people and departments endorsed by senior executives. - Make the key people and departments part of the project. If possible, get them involved in the objectives and key results.	*"My support to the project, with partnership with key people makes easier the development of the project, especially when they know the benefits that will be delivered."* (Director of pricing, #1) *"The dedicated team had a fundamental role in this project to notify all the problems during the development, for this reason, actions could be faster to overcome any problem that can affect the roadmap of the project"* (Manager of pricing, #2)

Source: The authors. (T) Technical, (M) Managerial.

5.2 Strengths and limitations

On the one hand, the strengths of this study included unrestricted access to one of the largest insurance companies in Brazil, which ensured data quality and enabled an in-depth analysis. In addition, the support and engagement of the organization's professionals who were interested in the study's results was especially important because the study was used to understand the organization's shortcomings and to develop an action plan for the project's next steps.

On the other hand, the study was limited by the fact that it employed a case study as its research approach. As a result, the findings are representative of only one company in the insurance industry at a particular point in time in a specific country, making it difficult to generalize the results to other industries. Practices may vary across organizations or across countries. If additional professionals in the organization had been interviewed, they may have had differing opinions, depending on their level of experience and the division of the organization in which they worked.

5.3 Further research

Future research, either quantitative or qualitative, should further examine a broader range of organizations in terms of size and industry sector by evaluating the following questions:

– How can the success of a big data analytics project be measured?

– Considering that a big data platform must evolve within an organization, what are the next deliverables that can potentialize its benefits to the business?

Finally, it is important to address this study's limitations by considering other stakeholders' perspectives within the organization.

References

Aiken, P., & Gorman, M. (2013). The Case for the Chief Data Officer: Recasting the C-Suite to Leverage Your Most Valuable Asset. Morgan Kaufmann Publishers.

Alharthi, A., Krotov, V., & Bowman, M. (2017). Addressing barriers to big data. Business Horizons, 60(3), 285–292. https://doi.org/10.1016/j.bushor.2017.01.002

Ali, U., & Kidd, C. (2014). Barriers to effective configuration management application in a project context: An empirical investigation. International Journal of Project Management, 32(3), 508–518. https://doi.org/10.1016/j.ijproman.2013.06.005

Araz, O. M., Choi, T.-M., Olson, D. L., & Salman, F. S. (2020). Role of Analytics for Operational Risk Management in the Era of Big Data. Decision Sciences, 51(6), 1320–1346. https://doi.org/10.1111/deci.12451

Arumugam, S., & Bhargavi, R. (2019). A survey on driving behavior analysis in usage-based insurance using big data. Journal of Big Data, 6(1), 86. https://doi.org/10.1186/s40537-019-0249-5

Benbasat, I., Goldstein, D. K., & Mead, M. (1987). The Case Research Strategy in Studies of Information Systems. MIS Quarterly, 11(3), 369–386. https://doi.org/10.2307/248684

Berghout, E., Nijland, M., & Powell, P. (2011). Management of lifecycle costs and benefits: Lessons from information systems practice. Computers in Industry, 62(7), 755–764. https://doi.org/10.1016/j.compind.2011.05.005

Bohnert, A., Fritzsche, A., & Gregor, S. (2019). Digital agendas in the insurance industry: The importance of comprehensive approaches†. The Geneva Papers on Risk and Insurance - Issues and Practice, 44(1), 1–19. https://doi.org/10.1057/s41288-018-0109-0

Brave, S. A., Butters, R. A., & Fogarty, M. (2022). The perils of working with big data, and a SMALL checklist you can use to recognize them. Business Horizons, 65(4), 481–492. https://doi.org/10.1016/j.bushor.2021.06.004

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. MIS Quarterly, 36(4), 1165. https://doi.org/10.2307/41703503

Collins, A. (2014). Big data skills to pay the big data bills. Technology Decisions. http://technologydecisions.com.au/content/it-management/article/big-data-skills-to-pay-the-big-data-bills-378997873

Crittenden, V. L., & Crittenden, W. F. (2008). Building a capable organization: The eight levers of strategy implementation. Business Horizons, 51(4), 301–309. https://doi.org/10.1016/j.bushor.2008.02.003

Davenport, T. H., & Dyché, J. (2013). Big Data in Big Companies. SAS Institute Inc. https://www.sas.com/en_in/whitepapers/bigdata-bigcompanies-106461.html

Davenport, T. H., & Patil, D. J. (2012, October 1). Data Scientist: The Sexiest Job of the 21st Century. Harvard Business Review. https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century

Deloitte. (2021). Pesquisa Febraban de Tecnologia Bancária. https://portal.febraban.org.br/pagina/3106/48/pt-br/pesquisa

Deloitte. (2023). Insurance Outlook. https://www2.deloitte.com/content/dam/insights/articles/us175368_cfs_fsi-outlook-insurance/DI_US175368_CFS_FSI-Outlook-Insurance.pdf

Demchenko, Y., Grosso, P., De Laat, C., & Membrey, P. (2013). Addressing big data issues in scientific data infrastructure. In 2013 International conference on collaboration technologies and systems (CTS) (pp. 48-55). IEEE.

Dykes, B. (2017). Why Companies Must Close the Data Literacy Divide. Forbes. https://www.forbes.com/sites/brentdykes/2017/03/09/why-companies-must-close-the-data-literacy-divide/

Eisenhardt, K. M. (1989). Building Theories from Case Study Research. Academy of Management Review, 14(4), 532–550. https://doi.org/10.5465/AMR.1989.4308385

Francisco, E. D. R., Kugler, J. L., Kang, S. M., Silva, R., & Whigham, P. A. (2019). Beyond technology: Management challenges in the Big Data era. Revista de Administração de Empresas, 59, 375-378. https://doi.org/10.1590/S0034-759020190603

Gartner. (2021). 2022 CIO and Technology Executive Agenda: An Insurance Perspective. https://www.gartner.com/document/4008149

Gökalp, M. O., Gökalp, E., Kayabay, K., Koçyiğit, A., & Eren, P. E. (2022). The development of the data science capability maturity model: survey-based research. Online Information Review, 46(3), 547-567. https://doi.org/10.1108/OIR-10-2020-0469

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. Information & Management, 53(8), 1049–1064. https://doi.org/10.1016/j.im.2016.07.004

Hambrick, D. C., & Mason, P. A. (1984). Upper echelons: The organization as a reflection of its top managers. Academy of management review, 9(2), 193-206. https://doi.org/10.5465/amr.1984.4277628

Hambrick, D. C. (2007). Upper echelons theory: An update. Academy of management review, 32(2), 334-343. https://doi.org/10.5465/amr.2007.24345254

Hanafy, M., & Ming, R. (2021). Machine Learning Approaches for Auto Insurance Big Data. Risks, 9(2), Article 2. https://doi.org/10.3390/risks9020042

Herschel, G., Brethenoux, E., Idoine, C., Kronz, A., Hunter, E., & Horvath, M. (2019). Predicts 2019: Analytics and BI Strategy. https://www.gartner.com/document/3896764

Hoffman, S., & Podgurski, A. (2013). Big Bad Data: Law, Public Health, and Biomedical Databases. Journal of Law, Medicine & Ethics, 41(S1), 56–60. https://doi.org/10.1111/jlme.12040

Interbrand. (2022). Marcas Brasileiras Mais Valiosas 2021. https://learn.interbrand.com/hubfs/INTERBRAND/MBMV2021.pdf

Jin, M., & Yao, L. (2022). Influence Mechanism of Educational Leadership on Environmental Accounting Based on Big Data Algorithm. Journal of environmental and public health, 2022, 5690230. https://doi.org/10.1155/2022/5690230

Johnson, J. E. (2012). Big Data + Big Analytics = Big Opportunity. Financial Executives International. https://www.financialexecutives.org/FEI-Daily/March-2015/big-data-big-analytics-big-opportunity.aspx

Jones, M. (2013, April 1). Big Data Raises Big Questions. Government Technology. https://www.govtech.com/archive/Big-Data-Raises-Big-Questions.html

Kim, G.-H., Trimi, S., & Chung, J.-H. (2014). Big-data applications in the government sector. Communications of the ACM, 57(3), 78–85. https://doi.org/10.1145/2500873

Kronz, A., Jaffri, A., Tapadinhas, J., Sun, J., & Herschel, G. (2021). Predicts 2021: Analytics, BI and Data Science Solutions—Pervasive, Democratized and Composable. Gartner. https://www.gartner.com/document/code/735777

Krotov, V., & Johnson, L. (2022). Big web data: Challenges related to data, technology, legality, and ethics. Business Horizons. https://doi.org/10.1016/j.bushor.2022.10.001

Larsson, R. (1993). Case survey methodology: Quantitative analysis of patterns across case studies. Academy of management Journal, 36(6), 1515-1546. https://doi.org/10.5465/256820

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics, and the path from insights to value. MIT Sloan Management Review, 52(2), 21–32.

Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The Parable of Google Flu: Traps in Big Data Analysis. Science, 343(6176), 1203–1205. https://doi.org/10.1126/science.1248506

Li, T., Kou, G., & Peng, Y. (2020). Improving malicious URLs detection via feature engineering: Linear and nonlinear space transformation methods. Information Systems, 91, 101494. https://doi.org/10.1016/j.is.2020.101494

Marnewick, C., & Marnewick, A. L. (2022). Benefits realisation in an agile environment. International Journal of Project Management, 40(4), 454–465. https://doi.org/10.1016/j.ijproman.2022.04.005

Martinsuo, M., & Huemann, M. (2021). Designing case study research. International Journal of Project Management, 39(5), 417–421. https://doi.org/10.1016/j.ijproman.2021.06.007

Mayhew, H., Saleh, T., & Williams, S. (2016). Making data analytics work for you—Instead of the other way around | McKinsey. McKinsey Quarterly. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/making-data-analytics-work-for-you-instead-of-the-other-way-around

Mazzei, M. J., & Noble, D. (2017). Big data dreams: A framework for corporate strategy. Business Horizons, 60(3), 405–414. https://doi.org/10.1016/j.bushor.2017.01.010

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: The management revolution. Harvard Business Review, 90(10), 60–68.

Mikalef, P., Giannakos, M. N., Pappas, I. O., & Krogstie, J. (2018). The human side of big data: Understanding the skills of the data scientist in education and industry. In 2018 IEEE global engineering education conference (EDUCON) (pp. 503-512). IEEE. https://doi.org/10.1109/EDUCON.2018.8363273

Mikalef, P., Krogstie, J., Pappas, I. O., & Pavlou, P. (2020). Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities. Information & Management, 57(2), 103169. https://doi.org/10.1016/j.im.2019.05.004

Miles, M., & Huberman, M. (1994). Qualitative Data Analysis: An Expanded Sourcebook (Second Edition). Sage Publications.

Miller, S. (2014). Collaborative Approaches Needed to Close the Big Data Skills Gap. Journal of Organization Design, 3(1), 26. https://doi.org/10.7146/jod.9823

Nie, Y., Talburt, J., Dagtas, S., & Feng, T. (2019). The influence of chief data officer presence on firm performance: does firm size matter? Industrial Management & Data Systems, 119(3), 495-520. https://doi.org/10.1108/IMDS-03-2018-0101

PMI (2017). Agile Practice Guide. Project Management Institute.
Moraes, H. R. O. C., Cunha, M., Terlizzi, M. A. (2017). IT indicators and organizational performance: a study of the retail sector in Brazil. In CONF-IRM 2017. https://aisel.aisnet.org/confirm2017/16

Pour, M. J., Abbasi, F., & Sohrabi, B. (2023). Toward a Maturity Model for Big Data Analytics: A Roadmap for Complex Data Processing. International Journal of Information Technology & Decision Making, 22(01), 377-419. https://doi.org/10.1142/S0219622022500390

Price, W. N., & Cohen, I. G. (2019). Privacy in the age of medical big data. Nature medicine, 25(1), 37-43. https://doi.org/10.1038/s41591-018-0272-7

Reggio, G., & Astesiano, E. (2020). Big-Data/Analytics Projects Failure: A Literature Review. 2020 46th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), 246–255. https://doi.org/10.1109/SEAA51224.2020.00050

Rieley, M. (2022). Big data adds up to opportunities in math careers. U.S. Bureau of Labor Statistics. https://www.bls.gov/opub/btn/volume-7/big-data-adds-up.htm

Rijmenam, M. van. (2014). Think Bigger: Developing a Successful Big Data Strategy for Your Business (Illustrated edition). AMACOM.

Rogers, P., Meehan, P., & Tanner, S. (2007). Building a winning culture. Bain & Company. https://media.bain.com/Images/BB_Building_winning_culture.pdf

Ross, J. W., Beath, C. M., & Quaadgras, A. (2013). You May Not Need Big Data After All. Harvard Business Review. https://hbr.org/2013/12/you-may-not-need-big-data-after-all

Sanchez, O. P. (2017). Cost and time project management success factors for information systems development projects. International Journal of Project Management, 19. https://doi.org/10.1016/j.ijproman.2017.09.007

Sarker, S., Xiao, X., & Beaulieu, T. (2012). Toward an Anatomy of "Successful" Qualitative Research Manuscripts in IS: A Critical Review and Some Recommendations. In ICIS 2012 Proceedings. https://aisel.aisnet.org/icis2012/proceedings/ResearchMethods/12

Seth, P., & Gulati, K. (2022). Use of Wearable and Health Applications in Insurance Industry Using Internet of Things and Big Data. In Big Data: A Game Changer for Insurance Industry (pp. 1–13). Emerald Publishing Limited.

Siggelkow, N. (2007). Persuasion with Case Studies. Academy of Management Journal, 50(1), 20–24. https://doi.org/10.5465/AMJ.2007.24160882

Silveira, M., Marcolin, C. B., & Freitas, H. M. R. (2015). Uso Corporativo do Big Data: Uma Revisão de Literatura. Revista de Gestão e Projetos, 6(3), Article 3. https://doi.org/10.5585/gep.v6i3.369

Sithambaram, J., Nasir, M. H. N. B. M., & Ahmad, R. (2021). Issues and challenges impacting the successful management of agile-hybrid projects: A grounded theory approach. International Journal of Project Management, 39(5), 474–495. https://doi.org/10.1016/j.ijproman.2021.03.002

Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. Journal of Business Research, 70, 263–286. https://doi.org/10.1016/j.jbusres.2016.08.001

Sood, K., Dhanaraj, R. K., Balamurugan, B., Grima, S., & Maheshwari, R. U. (Eds.). (2022). Big Data: A Game Changer for Insurance Industry. Emerald Publishing Limited.

Statista. (2022). Big data market size revenue forecast worldwide from 2011 to 2027. https://www.statista.com/statistics/254266/global-big-data-market-forecast/

Susep. (2022). Painel de Inteligência do Mercado de Seguros. https://www2.susep.gov.br/safe/menuestatistica/pims.html

Tabesh, P., Mousavidin, E., & Hasani, S. (2019). Implementing big data strategies: A managerial perspective. Business Horizons, 62(3), 347–358. https://doi.org/10.1016/j.bushor.2019.02.001

Terlizzi, M. A., Brandimarte, L., Brown, S., & Sanchez, O. P. (2019). Privacy Concerns and Protection Motivation Theory in the Context of Mobile Banking. In the 27th European Conference on Information Systems (ECIS), Stockholm & Uppsala, Sweden, June 8-14, 2019. https://aisel.aisnet.org/ecis2019_rp/24

Terlizzi, M. A., Albertin, A. L., & de Moraes, H. R. O. C. (2017). IT benefits management in financial institutions: Practices and barriers. International Journal of Project Management, 35(5), 763–782. https://doi.org/10.1016/j.ijproman.2017.03.006

Terlizzi, M. A., de Souza Meirelles, F., & de Moraes, H. R. O. C. (2016). Barriers to the use of an IT Project Management Methodology in a large financial institution. International Journal of Project Management, 34(3), 467-479. https://doi.org/10.1016/j.ijproman.2015.12.005

Terlizzi, M. A., Bento, D. R., & Biancolino, C. A. (2014). Auditoria de Projetos no Banco Itaú. Revista Inovação, Projetos e Tecnologias, 2(1), 98-114. https://doi.org/10.5585/iptec.v2i1.17

Trelles, O., Prins, P., Snir, M., & Jansen, R. C. (2011). Big data, but are we ready? Nature Reviews Genetics, 12(3), Article 3. https://doi.org/10.1038/nrg2857-c1

Trivedi, S., & Malik, R. (2022). Blockchain Technology as an Emerging Technology in the Insurance Market. In Big Data: A Game Changer for Insurance Industry (pp. 81–100). Emerald Publishing Limited.

Turner, R., & Ledwith, A. (2018). Project Management in Small to Medium-Sized Enterprises: Fitting the Practices to the Needs of the Firm to Deliver Benefit. Journal of Small Business Management, 56(3), 475–493. https://doi.org/10.1111/jsbm.12265

Tsai, J. C.-A., Jiang, J. J., Klein, G., & Hung, S.-Y. (2023). Task Conflict Resolution in Designing Legacy Replacement Systems. Journal of Management Information Systems, 40(3), 1009-1034. https://doi.org/10.1080/07421222.2023.2229120

Venkatesh, S. (2019). Big Data—Can it make a big impact in the Insurance sector? Journal of the Insurance Institute of India, 6(4), 92–97.

Vidgen, R., Shaw, S., & Grant, D. B. (2017). Management challenges in creating value from business analytics. European Journal of Operational Research, 261(2), 626–639. https://doi.org/10.1016/j.ejor.2017.02.023

Villarejo-Ramos, Á. F., & Cabrera-Sánchez, J.-P. (2019). Factors affecting the adoption of Big Data analytics in companies. RAE - Revista de Administracao de Empresas, 59(6), 415-429. https://doi.org/10.1590/S0034-759020190607

Whetten, D. A. (1989). What constitutes a theoretical contribution?. Academy of management review, 14(4), 490-495. https://doi.org/10.5465/amr.1989.4308371

Whyte, J., Stasis, A., & Lindkvist, C. (2016). Managing change in the delivery of complex projects: Configuration management, asset information and 'big data.' International Journal of Project Management, 34(2), 339–351. https://doi.org/10.1016/j.ijproman.2015.02.006

Yang, W., & Zhou, J. (2021). Service Innovation of Insurance Data Based on Cloud Computing in the Era of Big Data. Complexity, 1–10. https://doi.org/10.1155/2021/2303129

Yin, R. K. (2017). Case Study Research and Applications: Design and Methods (6th ed.). Sage Publications, Inc.

Zettelmeyer, F. (2015). A Leader's Guide to Data Analytics. Kellogg Insight. https://insight.kellogg.northwestern.edu/article/a-leaders-guide-to-data-analytics

Zia, A., & Kalia, P. (2022). Emerging Technologies in Insurance Sector: Evidence from Scientific Literature. Big Data: A Game Changer for Insurance Industry, 43–63.

6 Appendices

6.1 Appendix A. Relevant high-level interview questions

Interview guide used in the semistructured interviews:

(1) What are the project motivators? Why is this platform relevant for the company? What are the expected benefits?
(2) What can you say about the main difficulties faced during the project execution?
(3) What can you say about the main practices observed during the project execution?
(4) Did the project meet your expectations? What could be done better?
(5) What can you say about digital transformation in the company and the potential problems?

6.2 Appendix B. Project Documentation

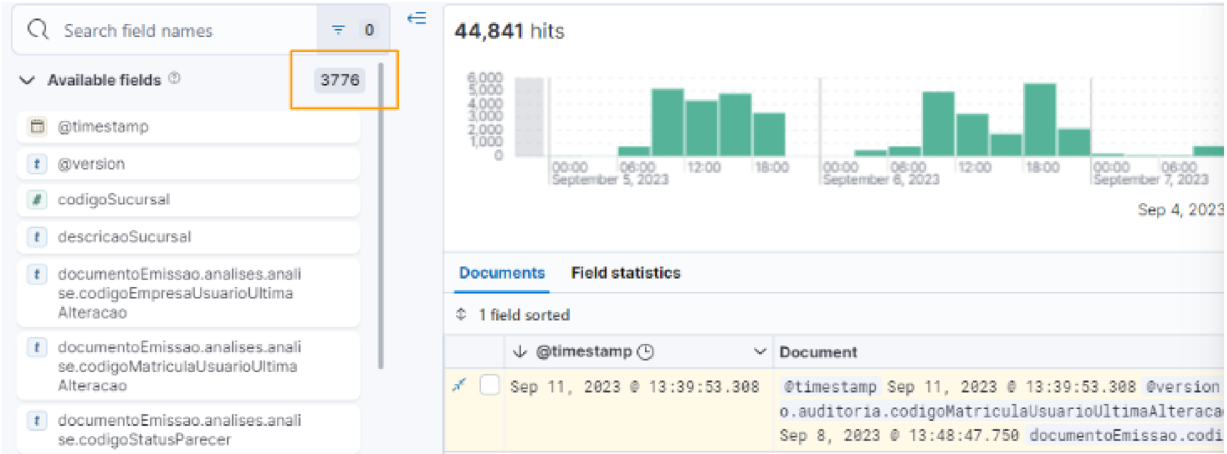
Figure B1.

Project Documentation – Dashboards For Controlling Products

Source: Project documentation.

Figure B2.

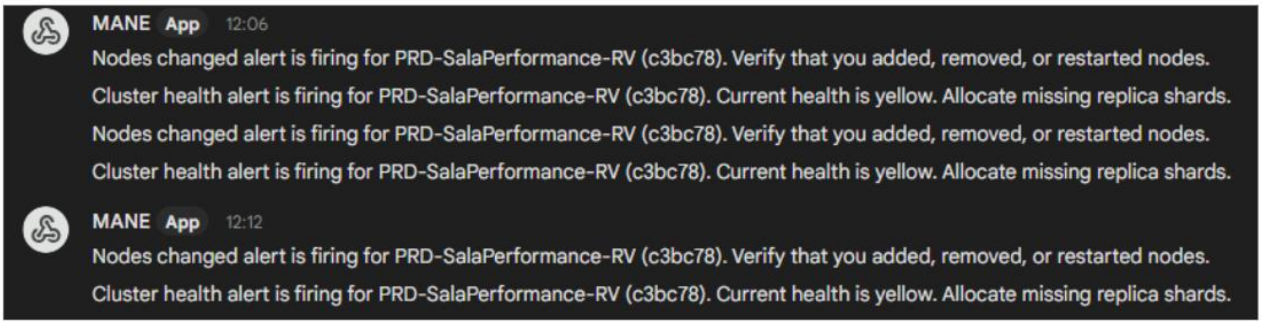
Project Documentation – Available Variables for Analytics



Source: Project documentation.

Figure B3.

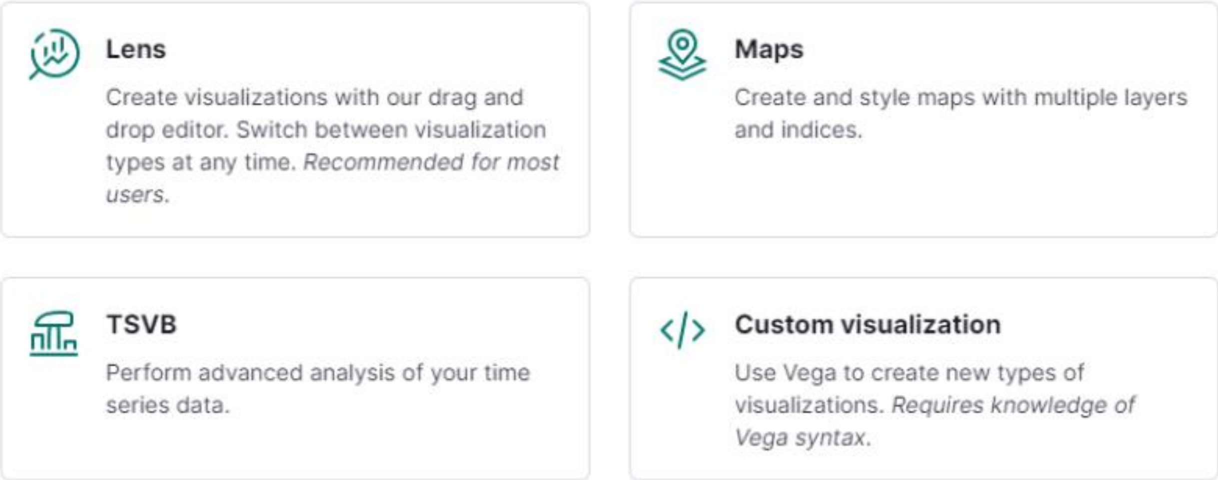
Project Documentation – Example of Automatic Monitoring of the Big Data Architecture



Source: Project documentation.

Figure B4.

Project documentation – Example of types of Visualization in Kibana



Source: Project documentation.