Priscila Falcão dos Santos

SATURN:

Easy data analysis platform aimed at non-tech teams.

SÃO PAULO
2025

Priscila Falcão dos Santos

Saturn:

Easy data analysis platform aimed at non-tech teams.

Final Course Project submitted to the Institute of Technology and Leadership (INTELI), to obtain a bachelor's degree in Software Engineering.

Advisor: Prof. Afonso Brandão Lelis

SÃO PAULO
2025

Cataloging in Publication
Library and Documentation Service
Institute of Technology and Leadership (INTELLI)
Data entered by the author.
(Cataloging record with international cataloging data, according to NBR 14724. The record will be completed later, after approval and before the final version is deposited. The completion of the cataloging record is the responsibility of the institution's library.)

Sobrenome, Nome
Título do trabalho: subtítulo / Nome Sobrenome do autor; Nome e Sobrenome do orientador. – São Paulo, 2025.
nº de páginas : il.

Trabalho de Conclusão de Curso (Graduação) – Curso de [Ciência da Computação] [Engenharia de Software] [Engenharia de Hardware] [Sistema de Informação] / Instituto de Tecnologia e Liderança.
Bibliografia

1. [Assunto A]. 2. [Assunto B]. 3. [Assunto C].

CDD. 23. ed.

Acknowledgments

Special thanks to the advisor who accompanied the trajectory of this project, to Ana Paula, who guided me in joining Inteli and to Verônica who guided me to success so far.

Epigraph

“Research shows that companies that employ data-driven decision-making are more productive, have a higher market value, and deliver higher returns for their shareholders.” - Bang, Claus G, 2024. *Data-Driven Decision-Making for Business.*

Resumo

[Falcão dos Santos, Priscila. **Saturn: Easy data analysis platform aimed at non-tech teams**. 2025. 24 p. TCC (Graduação) – Curso Engenharia de Software, Instituto de Tecnologia e Liderança, São Paulo, 2025.]

O projeto Saturn é uma plataforma de análise de dados baseada em nuvem, concebida para enfrentar a dificuldade de empresas de pequeno porte em analisar e interpretar grandes volumes de dados, e a insustentabilidade de contratar analistas em massa devido ao alto custo. O objeto de estudo é o desenvolvimento de uma ferramenta acessível que democratize a análise de dados para áreas sem conhecimento técnico especializado, como negócios, administração e marketing. O objetivo principal é fornecer uma aplicação de análise de dados altamente personalizável que suporte a leitura e fusão de múltiplas fontes de dados, a criação facilitada de gráficos diversificados, e a exportação e a geração de reports. O método de desenvolvimento adotou uma arquitetura robusta e escalável, otimizada para baixo orçamento, baseada no modelo C4, utilizando Next.js, Node.js, Python (Flask), e Supabase. A solução processa dados usando um middleware em Python, realizando a análise e comunicando com o Supabase. Em conclusão, o projeto Saturn estabelece uma base tecnológica sólida para evoluir continuamente, visando o impacto de longo prazo na capacitação de pequenas empresas para a tomada de decisões estratégicas orientadas a dados.

Palavras-Chave: análise de dados; software; plataforma de dados; times não-tech.

ABSTRACT

[Falcão dos Santos, Priscila. **Saturn: Easy data analysis platform aimed at non-tech teams**. 2025. nº of pages. Final course project (Bachelor) – Course Software Engineering, Institute of Technology and Leadership, São Paulo, 2025.]

The Saturn project is a cloud-based data analytics platform designed to address the challenges faced by small enterprises in analyzing and interpreting large volumes of data, as well as the unsustainability of hiring analysts at scale due to high costs. The object of study is the development of an accessible tool that democratizes data analysis for domains without specialized technical knowledge, such as business, administration, and marketing. The primary objective is to provide a highly customizable data analytics application that supports the ingestion and integration of multiple data sources, the streamlined creation of diverse charts, and the export and generation of reports. The development methodology adopted a robust and scalable architecture optimized for low budgets, based on the C4 model, and leveraging Next.js, Node.js, Python (Flask), and Supabase. The solution processes data through a Python-based middleware, performing analysis and communicating with Supabase. In conclusion, the Saturn project establishes a solid technological foundation for continuous evolution, aiming for long-term impact in empowering small enterprises to make data-driven strategic decisions.

Key words: data analysis; software; data platform; non-tech teams.

List of Illustrations

Summary

1 Introduction

The modern corporate landscape is defined by the critical need to manage and extract value from large volumes of data, with the ability to analyze and interpret such information being a decisive factor for success [Brynjolfsson et al., 2011]. The production of advanced analyses and executive reports for purposes such as performance monitoring and market analysis is essential for strategic decision-making and for maintaining competitiveness in an increasingly data-driven market [McAfee and Brynjolfsson, 2012]. However, small enterprises face a lack of specialized technical expertise and the cost barrier imposed by hiring analysts at scale, making the generation of insights unsustainable in the long term. Furthermore, the management of private and confidential data requires the implementation of robust security practices in compliance with regulations such as the LGPD, which is fundamental to user trust and business continuity. In light of this context, there is a clear demand for a tool capable of democratizing access to data analysis for areas without technical expertise.

1.1. **Partner Company Context:**

The Saturn project established a partnership with the company Sattwa, an organization in the food industrial sector specialized in food fortification and the production of macronutrient blends. This collaboration emerged from the recognition of the difficulties Sattwa faces in handling large volumes of data and in generating budgets and ongoing analyses—a common challenge among companies that rely on data-driven insights to support strategic decision-making.

The project's relevance to the partner is high, as the company does not have a dedicated technology team, and its financial department, in particular, lacks expertise in technology, statistics, and data management. This context reflects a broader market need for tools that democratize data analysis for non-technical professionals, such as those working in business and administration.

The partnership with Sattwa enabled the company to act as a stakeholder, providing continuous feedback and suggestions for improvement.

1.2. **Problem Definition (Corporate Pain Point):**

The central pain point driving the Saturn project lies in the inefficiency of data analysis processes within small and medium-sized enterprises, particularly those that handle large volumes of information and struggle to analyze and interpret them effectively. This problem is exacerbated by the absence of a dedicated technology team in many of these organizations and by the lack of specialized expertise in technology, statistics, and data management within critical departments such as finance, administration, and marketing. The traditional approach, in which data analysis is restricted to a small group of specialists, results in slow and ineffective decision-making processes and prevents the majority of employees from leveraging data to optimize their work. As a consequence of this misalignment, most employees report feeling overwhelmed or unhappy when interacting with data (74%) [Accenture, 2020]. This negative sentiment disrupts workflow efficiency, leading 36% of employees to seek alternative methods to complete tasks without using data and 14% to avoid the task entirely, which constitutes a tangible barrier to establishing a data-driven culture [Accenture, 2020].

The Saturn project aims to transform several baseline metrics that quantify this inefficiency. Operational cost metrics are directly impacted by the need to hire analysts at scale, a costly and unsustainable solution in the long term for small enterprises. In terms of productivity and engagement, the lack of data literacy in the market results in an average productivity loss of more than five working days (43 hours) per employee each year, due to procrastination and stress-related sick leave associated with data-related tasks. Only 21% of the global workforce reports full confidence in their data literacy skills, defined as the ability to read, understand, question, and work with information [Accenture, 2020]. This low level of confidence is reflected in decision-making practices, where nearly half of employees (48%) frequently choose to rely on intuition rather than data-driven insights, and only 37% of organizations are able to make decisions with high quality and speed [Accenture, 2020]. Furthermore, a significant volume of corporate data remains unused, with 60%

to 73% of all enterprise data never being analyzed. The system's objective, therefore, is to reverse these figures by providing an intuitive interface that encourages data usage by non-technical professionals, with fast response times, ideally within six seconds.

1.3. **Proposed Solution and Expected Contribution:**

The primary objective of Saturn's contribution is to create value by addressing the inefficiency caused by the lack of data literacy in small enterprises and the unsustainable need for hiring analysts at scale. Although the value delivered by the platform cannot be quantified in exact percentage terms prior to implementation, the solution aims to improve resource optimization and promote more accurate decision-making. It is estimated that, within three months of platform usage, clear results will be observable in terms of increased quality of strategic decisions and reduced wasted investments, transforming raw data into a strategic asset that provides a competitive advantage.

1.4. **Business Objectives:**

The primary objective of Saturn's contribution is to create value by addressing the inefficiency caused by the lack of data literacy in small enterprises and the unsustainable need for hiring analysts at scale. Although the value delivered by the platform cannot be quantified in exact percentage terms prior to implementation, the solution aims to improve resource optimization and promote more accurate decision-making. It is estimated that, within three months of platform usage, clear results will be observable in terms of increased quality of strategic decisions and reduced wasted investments, transforming raw data into a strategic asset that provides a competitive advantage.

1.5. **Structure of the thesis/dissertation:**

The structure of the thesis is organized into subsequent chapters that detail the complete journey of the Saturn project, from its strategic and technical justification to the assessment of its business impact. The Solution Development chapter establishes the business justification for democratizing data analysis, focusing on small enterprise domains without technical expertise (such as administration and marketing), and substantiates the choice of an architecture optimized to be low-cost yet scalable, describing the sprint-based development methodology. The Specification and Development section details the MVP, which is built upon a three-layer architecture documented using the C4 Model, and implements critical requirements such as data integration, the creation of customizable visualizations, and the execution of statistical and exploratory analyses via Python scripts. This section also addresses code quality through SonarQube. The Assessment of Impact and Business Contribution chapter defines corporate success metrics, aiming to increase decision accuracy and resource optimization, with expected results within three months. Finally, the Conclusion outlines the project's continuity through future modules already planned to focus on Security and LGPD compliance (Module 2), User Experience Evolution (Module 3), and Process Optimization (Module 4).

2 Solution Development

2.1 Applied Rationale

The Saturn project adopted a three-layer architecture with the objective of balancing quality, development speed, and the creation of a lightweight, low-budget application. The architecture was optimized for scalability, leveraging a modern and efficient technology stack.

2.1.1 **Business Area Rationale:**

In the context of the partner company Sattwa, the Saturn project focused specifically on the financial and management areas, which are fundamental to the industrial operation of food fortification but have historically had limited proximity to technology and lack technical expertise in statistics and data manipulation. For professionals in these domains, concepts such as budget monitoring, financial forecasting, and the analysis of financial statements are vital; however, reliance on manual processes or complex tools makes the interpretation of large volumes of data a significant challenge. The absence of data literacy in these areas leads to a reliance on intuition rather than objective insights, which can result in inefficiencies and misguided strategic decisions.

2.1.2 **Technological rationale for the solution:**

The frontend, developed using Next.js 15 with TypeScript 5 and styled with Tailwind CSS, is responsible for ensuring a lightweight application and delivering an intuitive and clean interface. This choice of interface technologies accelerates development and enhances usability, which is crucial for the non-technical target audience. For the Backend as a Service (BaaS), Supabase was selected, providing a PostgreSQL database, Supabase Auth for authentication, and Supabase Storage for file management. This approach minimizes infrastructure management overhead and aligns directly with the project's budgetary constraints.

Quality assurance and processing power reside in the combination of the middleware (Node.js/TypeScript) and the Models layer implemented in Python 3. This hybrid architecture is essential for both performance and the execution of complex analyses. The middleware orchestrates the execution of Python scripts that leverage established libraries such as pandas and numpy for data manipulation and statistical computation, as well as scikit-learn for advanced analyses such as Principal Component Analysis (PCA). To maintain code quality and robustness throughout development, the project employed SonarQube for static analysis, evaluating maintainability and reliability, and utilized Docker for containerization, ensuring portability and consistency across the development environment.

2.1.3 **Fundamentals of Management and Development Methods:**

The management and development of the Saturn project are grounded in the Scrum methodology, with the platform being built incrementally to ensure agility and alignment with business needs. The work cycle was structured into biweekly sprints, enabling technical progress and deliverables to be presented to the partner for validation at the end of each two-week period. This model of frequent communication was essential for feedback exchange, through which the partner evaluated progress, assisted in backlog prioritization, and suggested adjustments or new features to be implemented in subsequent stages.

2.2 **Specification and Development :**

The specification and development of the system were conducted in alignment with the technical and organizational demands of the partner company, ensuring adherence to its practices, constraints, and strategic objectives. In this context, architectural decisions play a central role, as they directly impact the scalability, maintainability, security, and future evolution of the solution. Accordingly, the project was conceived based on a careful mapping of the partner's needs, ensuring that each system component, module, and interface reflected real business requirements. Furthermore, the selection of technologies, frameworks, and design patterns was not arbitrary but deliberately planned to address the specific characteristics of the proposal, balancing performance, cost, ease of integration, and support within the organization's technical ecosystem.

2.2.1 **Requirements and Specifications:**

Functional Requirements (FR):

FR1 – Data Import: The system must allow the import of CSV files.

FR2 – Creation of Customizable Charts: An intuitive interface for generating bar, pie, and scatter charts, allowing the customization of colors, legends, and axes.

FR3 – Statistical Analysis: The system must provide calculations for mean, weighted mean, median, mode, and standard deviation.

FR4 – Advanced Analyses and Insights: Execution of complex exploratory analyses, such as Principal Component Analysis (PCA) and correlation matrices, in addition to generating automatic insights via the pandas .describe() function.

Non-Functional Requirements (NFR):

NFR1 – Availability and Remote Access: Ensure high availability (96.9%) and secure access.

NFR2 – Usability: The interface must be simple, clean, and intuitive, enabling users without technical knowledge to perform analyses without difficulty.

NFR3 – Scalability and Low Cost: The architecture must be optimized for a low initial budget while remaining flexible to support growth in data volume and number of users.

NFR4 – Performance: The system must ensure fast response times, with a maximum limit of up to 6 seconds for analysis processing.

User Specifications and Use Cases

The platform is targeted at professionals in areas such as business, administration, finance, and marketing within small enterprises, who handle significant volumes of data but lack the technical expertise required to interpret them.

The main use cases include:

a) Data Project Management: The user creates organized projects to group multiple analyses around a specific topic.

b) Data Preparation: The user uploads CSV files to the storage bucket and uses system tools to clean or merge tables.
c) Visual and Statistical Exploration: Through components such as ChartAnalysis, the user selects axes and chart types to visualize trends or requests numerical analyses to validate financial or managerial hypotheses.
d) Dimensionality Reduction: For complex datasets, the user executes PCA to identify the most impactful variables prior to making a strategic decision.
e) Results Sharing: The user generates a final report and exports it for presentation in executive meetings, promoting evidence-based decision-making.

2.2.2 **Architecture and Technology:**

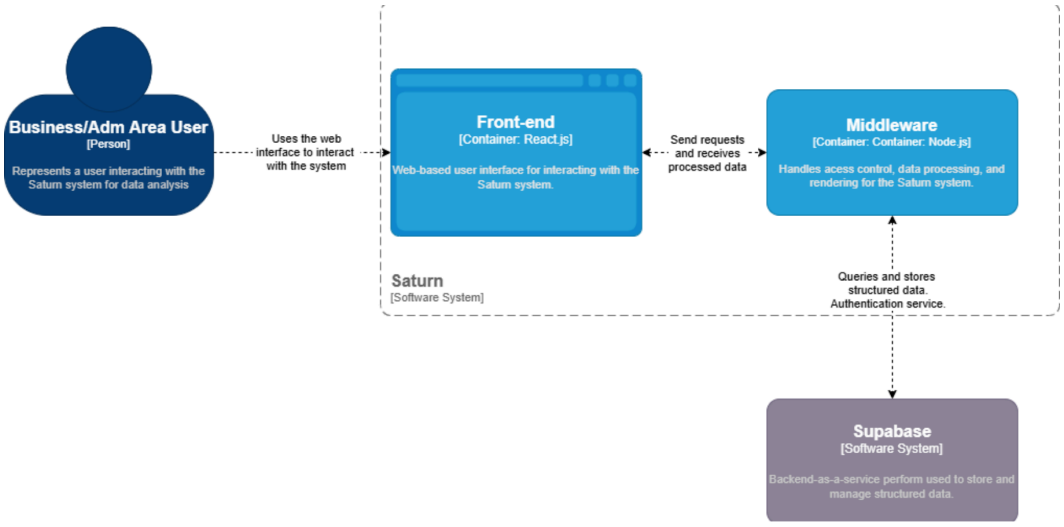
The Saturn project is based on a three-layer architecture (Presentation, Logic, and Processing), operating as a cloud-based platform to ensure remote access and low-cost scalability. The infrastructure is supported by Supabase, which provides the PostgreSQL database and file storage, enabling the application to remain lightweight and efficient for its target audience.

The middleware, developed in Node.js with TypeScript, acts as the central orchestrator of the system and is responsible for request validation and the management of external script execution. One of the key advantages of this layer is that it is organized into specific modules and controllers, such as *choose_analysis.ts*, which centralizes the logic for selecting different categories of analysis. This modular structure is strategic, as it enables the rapid and straightforward integration of new features or types of analyses, such as predictive or filtered analyses, without requiring restructuring of the application's core.

The code organization reflects this emphasis on maintainability by employing a clear separation of directories and responsibilities. The source code is primarily divided between *src/frontend/*, which contains the Next.js interface, and *src/middleware/*, which holds the orchestration logic. At the data processing level (Models), Python scripts are organized into subdirectories categorized by analysis

type, such as */insight/*, */numeric/*, and */explainable*/, ensuring that each exploratory or statistical analysis component remains isolated and easy to maintain and extend.

Figure 1 - Project architecture.



Source: prepared by the author (2025).

2.2.3 **Metadata - technical specifications of the project:**

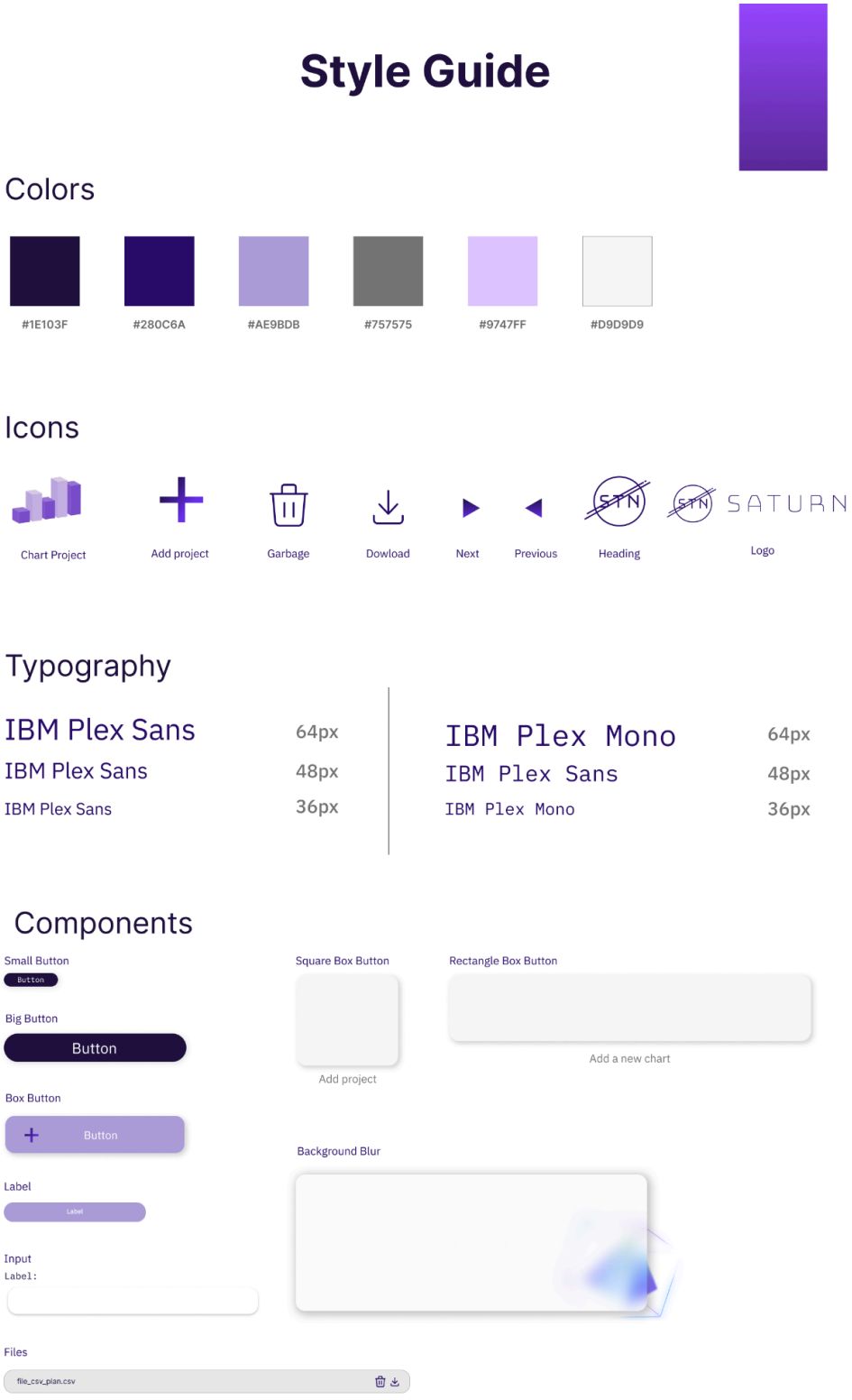
The metadata of the Saturn project is a JSONB object stored in the project table within the PostgreSQL database (managed via Supabase), serving as the central repository for all user configurations and results. Its structure is composed of three fundamental properties: name, which stores the user-defined identifier; the charts array, which details each created chart (specifying whether it is a bar, pie, or scatter chart, along with its axes and titles); and the items array, which stores the remaining analyses performed. This organization enables the platform to quickly retrieve all saved visualizations and the user's work history without the need to reprocess raw data on each access.

Within the items array, the metadata records the results of Python scripts dedicated to numerical analyses (mean, standard deviation), insights (comprehensive statistical descriptions via the .describe() method), and explanatory analyses such as the correlation matrix and PCA. Each entry in this array not only

contains the processed values but also stores the path of the original file in Supabase Storage used for that specific analysis, ensuring traceability of the data source. The system is designed to be incremental and extensible, allowing the middleware to append new analysis objects to the JSONB structure as the user progresses in their exploration, thereby facilitating the future integration of additional functionalities.

2.2.4 **UX Definitions**

Figure 2 - Style Guide.



Source: prepared by the author (2025).

2.2.5 **Testing and Technical Evaluation:**

The testing strategy of the Saturn project was based on a hybrid approach, combining continuous code inspection with incremental validations alongside the user. Automated technical evaluation was conducted using SonarQube, which performed static analyses to identify security vulnerabilities, reliability issues, and code redundancies. Complementarily, the Vitest framework with the jsdom library was employed for the implementation of unit and integration tests, focusing on the robustness of middleware and frontend functionalities. At the corporate level, User Acceptance Testing (UAT) was conducted organically through biweekly meetings with business partners, allowing stakeholders to validate each new MVP feature and provide critical feedback to ensure that the solution addressed the real challenges of the financial and management areas.

The evaluation results demonstrate that the solution meets the required quality standards, achieving an A grade in Security and Maintainability in the final SonarQube reports. Although the initial test coverage was identified as an area for improvement, the development cycle prioritized the implementation of automated tests to enhance system reliability in the final stages. In terms of performance, the platform was validated to meet the non-functional requirement of a response time of up to six seconds, ensuring agility in the processing of statistical and exploratory analyses. Finally, the MVP validation confirmed the integrity of the five pages and 21 components developed, ensuring that critical workflows, such as Supabase authentication and automated insights generation, operate according to the established technical specifications.

2.3 Assessment of Impact and Contribution to the Business

The assessment of impact and business contribution considers how a facilitated data analytics platform can strengthen organizational strategy by transforming dispersed data into actionable information for different management levels. In this context, the proposed solution aims not only to measure the return on investment of time and resources but also to demonstrate gains in decision-making

agility, opportunity identification, and risk mitigation. By democratizing access to technical analysis and management tools, the platform enables departments with limited technological expertise to utilize indicators, reports, and dashboards intuitively, thereby expanding the intelligent use of data throughout the organization.

2.3.1 **Defining Corporate Success Metrics:**

To monitor the success of the Saturn project from a business perspective, the KPIs should focus on:

a) Adoption by Non-Technical Users: Measure the percentage of professionals in business, administration, and marketing who actively use the platform. Success is defined by the system's ability to enable users without technological expertise to independently extract valuable insights.
b) Reduction of Operational Costs (ROI): Compare the platform's cost to the high expense of hiring specialized data analysis professionals, which represents one of the main pain points identified. This KPI focuses on the savings generated for small enterprises that can now perform advanced analyses without increasing payroll costs.
c) Efficiency in Generating Insights: Measure the average time a user takes to transform raw data into an exportable executive report (PDF/PNG). The goal is to drastically reduce the time spent on manual data cleaning and chart creation, thereby increasing team productivity.
d) Frequency of Advanced Analysis Use: Monitor how often functionalities such as PCA (Principal Component Analysis) and the Correlation Matrix are executed. A high usage volume indicates that the platform is effectively elevating the company's analytical maturity, shifting decision-making from intuition to evidence-based insights.

2.3.2 **Results and Impact Analysis:**

It was not possible to measure implementation results during the system development period. Implementation at the company will take place in the future (2 months after delivery of the current work).

2.3.3 **Critical Success Factors and Lessons Learned:**

The choice of a cloud-based architecture (Supabase) and a modular middleware enabled the development of a scalable, low-cost solution specifically designed to democratize access to data analysis for users in business and marketing areas. On the other hand, the process faced challenges such as the concentration of production workload on a single individual, which delayed the implementation of more complex platform features.

As a key lesson learned, the project demonstrated that the potential positive impact on company outcomes, including reduced costs for specialists and increased agility in decision-making, depends on overcoming users' lack of data fluency, requiring an intuitive interface that minimizes reliance on constant technical support.

3 Conclusion

The Saturn project has established itself as a strategic solution for democratizing data analysis in small enterprises, specifically serving professionals in non-technical areas such as management, finance, and marketing. Through a three-layer architecture (Frontend, Middleware, and Models) and the integrated use of technologies such as Next.js, Python, and Supabase, the project demonstrated that it is possible to transform raw and "messy" data into valuable assets through statistical analyses and customizable visualizations. The application of the Scrum methodology in biweekly sprints was essential to ensure incremental and transparent development, allowing the platform to evolve in continuous alignment with feedback from the business partner while maintaining low operational costs.

Beyond the successful development of functionalities, the project validated its technical robustness by achieving an A grade in security and maintainability via SonarQube, providing a solid foundation for software scalability. The business impact is reflected in empowering users to perform complex tasks, such as dimensionality reduction (PCA), without the need for dedicated technical teams, thereby reducing costs and minimizing reliance on intuition in strategic decision-making. For the next phases, the roadmap includes strengthening data security and compliance with the LGPD, as well as enhancing the user experience, ensuring that Saturn remains a reliable, ethical, and competitive tool in the data-driven market.

References

[Brynjolfsson, E.; Hitt, L. M.; & Kim, H. H. (2011). **Strength in Numbers: How does data-driven decision making affect firm performance?** Proceedings of ICIS. Available at: Accessed on: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1819486. Dec. 16, 2025.]

[McAfee, A., Brynjolfsson, E. (2012). **Big data: The management revolution. Harvard Business Review.** Vol. 106. No. 5. pp. 133-139. Available at: https://www.jstor.org/ stable/43861002. Accessed on: Dec. 16, 2025.]

[ACCENTURE. **The human impact of data literacy: a leader's guide to democratizing data, boosting productivity and empowering the workforce**. Report 2020. Available at: https://www.accenture.com/content/dam/accenture/final/a-com-migration/r3-3/pdf/pdf-118/accenture-the-human-impact-data-literacy.pdf. Accessed on: Dec. 16, 2025.]

[OECD (2019). **Data Analytics in SMEs: Trends and Policies**. OECD Publishing. Available at: https://www.oecd.org/content/dam/oecd/en/publications/reports/2019/06/data-analytics-in-smes_1535d46b/1de6c6a7-en.pdf. Accessed on: Dec. 16, 2025.]

[Data Analytics Adoption in Small and Medium Enterprises. **International Journal of Trend in Scientific Research and Development (IJTSRD)**. Available at: https://www.ijtsrd.com/papers/ijtsrd97280.pdf. Accessed on: Dec. 16, 2025.]

[The Importance of Data Analytics for Small and Medium Enterprises. NWDLTD Blog. Available at: https://nwdltd.com/blog/f/the-importance-of-data-analytics-for-small-and-medium-enterprises. Accessed on: Dec. 16, 2025.]