UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Visión Electrónica
https://doi.org/10.14483/issn.2248-4728

VISIÓN ELECTRÓNICA

A RESEARCH VISION

Business Intelligence and Analytics (BI&A): data analysis and business analytics for the optimization and automation of the *Stanley Black and Decker* Sales Report

Inteligencia de Negocios y Analítica (BI&A): análisis de datos y analítica de negocios para la optimización y automatización del Informe de Ventas de Stanley Black and Decker

Sebastián Nuñez Mejia[1], Harold Vacca González[2], Ernesto Uribe Becerra[3]

INFORMACIÓN DEL ARTÍCULO

Historia del artículo:
Enviado: 12/05/2025
Recibido: 01/07/2025
Aceptado: 17/11/2025

Keywords:
Business Intelligence and Analytics (BI&A)
Tools
Sales report
Python
Power BI
Windows Task Scheduler
Power Automate
Automation
Optimization

Palabras clave:
Inteligencia de Negocios y Analítica (BI&A)
Herramientas
Informe de ventas
Python
Power BI
Programador de Tareas de Windows
Power Automate
Automatización
Optimización

ABSTRACT

Business Intelligence and Analytics (BI&A) encompasses a range of strategies and tools employed in corporate settings to enhance the collection, processing, and analysis of large datasets with speed and precision. This capability provides significant strategic value to organizations by improving operational visibility and understanding. Data science plays a crucial role in this context, enabling systematic data analysis and business analytics that directly inform organizational decision-making.

While the full adoption of BI&A requires considerable investment in software and hardware development, it is becoming essential in today's global landscape, which is characterized by increasingly complex and voluminous data. This is especially important for identifying critical tasks requiring high-quality execution and for optimizing and automating associated processes and procedures.

This paper addresses the challenge of generating sales reports at Stanley Black & Decker. After identifying key variables, such as time expenditure, data quality, and the multiplicity of archived information, we propose a workflow designed to optimize and automate report generation. Our innovative approach involves data processing and cleansing to automate data cleaning and ensure data quality; data visualization to create interactive and detailed reports; scheduled process automation to ensure regular and efficient task execution; and integration to automate workflows between applications and services, thus enhancing process coherence and efficiency.

The proposed solution sequentially and complementarily integrates Python scripts for extracting and transforming data from various sources, Power BI for analyzing and visualizing the transformed data, and Windows Task Scheduler along with Power Automate to configure and automatically schedule script execution for report updates.

1. MSc. in Big Data and Data Science, Universidad Internacional de Valencia, Spain. E-mail: snmsebastian@gmail.com ORCID: https://orcid.org/0009-0008-2248-046X

2. MSc. in Applied Mathematics, Universidad EAFIT, Medellín, Colombia. Professor and Director of the SciBas Research Group, affiliated with Universidad Distrital Francisco José de Caldas, Bogotá, Colombia. E-mail: hvacca@udistrital.edu.co ORCID: https://orcid.org/0000-0001-7017-0070

3. MSc. in Applied Linguistics for EFL, Universidad Distrital Francisco José de Caldas, Bogotá, Colombia. Professor and collaborator of the SciBas Research Group, affiliated with Universidad Distrital Francisco José de Caldas, Bogotá, Colombia. E-mail: euribeb@udistrital.edu.co ORCID: https://orcid.org/0009-0008-8018-0422

Cite this article as: S. Nuñez Mejia, H. Vacca González, E. Uribe Becerra, "Business Intelligence and Analytics (BI&A): data analysis and business analytics for the optimization and automation of the Stanley Black and Decker Sales Report", Visión Electrónica, vol. 19, no. 2, 2025.

The implementation of our solution resulted in a new sales report that achieved a 61% reduction in the time required and an 85% reduction in the number of files. Furthermore, automating the data cleaning process improved the reliability and quality of the information. Consequently, the integration of technological tools and BI&A led to greater process efficiency and enhanced the quality of information, which is crucial for decision-making at Stanley Black & Decker, both nationally and internationally.

RESUMEN

La Inteligencia de Negocios y Analítica (BI&A) abarca una variedad de estrategias y herramientas empleadas en entornos corporativos para mejorar la recopilación, el procesamiento y el análisis de grandes conjuntos de datos con rapidez y precisión. Esta capacidad proporciona un valor estratégico significativo a las organizaciones al mejorar la visibilidad y la comprensión operativas. La ciencia de datos juega un papel crucial en este contexto, permitiendo un análisis de datos sistemático y una analítica de negocios que informan directamente la toma de decisiones organizacionales.

Si bien la adopción total de BI&A requiere una inversión considerable en el desarrollo de software y hardware, se está volviendo esencial en el panorama global actual, que se caracteriza por datos cada vez más complejos y voluminosos. Esto es especialmente importante para identificar las tareas críticas que requieren una ejecución de alta calidad y para optimizar y automatizar los procesos y procedimientos asociados.

Este artículo aborda el desafío de generar informes de ventas en Stanley Black & Decker. Después de identificar variables clave, como el gasto de tiempo, la calidad de los datos y la multiplicidad de información archivada, proponemos un flujo de trabajo diseñado para optimizar y automatizar la generación de informes. Nuestro enfoque innovador involucra el procesamiento y la limpieza de datos para automatizar la limpieza de datos y garantizar su calidad; la visualización de datos para crear informes interactivos y detallados; la automatización de procesos programados para garantizar la ejecución regular y eficiente de las tareas; y la integración para automatizar los flujos de trabajo entre aplicaciones y servicios, mejorando así la coherencia y la eficiencia del proceso.

La solución propuesta integra de forma secuencial y complementaria scripts de Python para extraer y transformar datos de diversas fuentes, Power BI para analizar y visualizar los datos transformados, y el Programador de Tareas de Windows junto con Power Automate para configurar y programar automáticamente la ejecución de scripts para las actualizaciones de informes.

La implementación de nuestra solución resultó en un nuevo informe de ventas que logró una reducción del 61% en el tiempo requerido y una reducción del 85% en el número de archivos. Además, la automatización del proceso de limpieza de datos mejoró la confiabilidad y la calidad de la información. En consecuencia, la integración de herramientas tecnológicas y BI&A condujo a una mayor eficiencia de los procesos y mejoró la calidad de la información, lo cual es crucial para la toma de decisiones en Stanley Black & Decker, tanto a nivel nacional como internacional.

1 Introduction

In today's business environment, decision-making increasingly relies on analyzing vast datasets generated by organizational activities. However, this data often requires preprocessing to be effectively utilized. This need has led to the rise of Business Intelligence (BI), a concept introduced by Howard Dresner of Gartner in 1989. BI aims to integrate technologies that extract and process corporate data from management systems, presenting it in a way that enables users to derive meaningful insights and facilitate tasks crucial to achieving business goals. Following the establishment of this concept, Jay Liebowitz defined BI as a systematic process encompassing the collection, analysis, and management of internal and external information, as well as relevant knowledge, to improve decision-making within a company. Alberto Rozenfarb further proposed that BI is a discipline that integrates information from diverse sources, empowering analysts to explore a unified dataset according to their specific criteria.

This data is typically stored in a historical data repository, or Data Warehouse, originating from transactions generated in business management systems and other sources. Thus, BI fundamentally relies on analyzing an organization's accumulated data to extract detailed knowledge, known as business intelligence.

Within this framework, "assets" encompass various types of databases relevant to the company, including information on suppliers, customers, employees, and processes. BI processes are responsible for exploring and analyzing this data to identify trends in business behavior, enabling the strategic use of information to generate competitive advantages and support decision-making.

Business intelligence has two primary focuses: the technical and the business. The technical focus centers on the analysis and study of tools and technologies, while the business focus emphasizes the methodology for using information in business decision-making.

Numerous precedents exist, with experiences from the past decade being particularly noteworthy. For instance, research on Power BI and its utility in business management [3] demonstrated its potential as a business intelligence tool for SMEs, offering a useful and intuitive cost-effectiveness relationship for business decision-making through data analysis.

Other research focused on analyzing information generated using business intelligence and data science tools [4], involving these tools to analyze information in the manufacturing and quality sectors of PepsiCo. This resulted in reduced analysis time, centralized processes, and the incorporation of new office tools. A proposal for business intelligence using Microsoft Power BI as support for decision-making in the commercial area of ABC Company, a manufacturer of plastic products in Colombia [5], integrated business intelligence through Microsoft Power BI, enabling centralized information use and proactive decision-making through quality analysis, replacing difficult-to-interpret spreadsheets.

In the industrial application of business intelligence in a small company using Power BI [6], the tool was used to gather information from the company's activities, delivering high-impact benefits for business decision-making and fostering an organizational culture in data science to improve management. Similarly, in the implementation of business intelligence to improve the efficiency of decision-making in project management [7], BI was implemented in a telecommunications services organization, leading to reduced management errors, costs, and times in evaluated projects.

Furthermore, the implementation of a data mart as a business intelligence solution, under Ralph Kimball's methodology, aimed to optimize decision-making in the Finance Department of the General Comptroller's Office of the Republic [8]. This addressed limitations in accessing financial information, such as reliance on IT departments for data retrieval and manual data cleaning, which introduced errors and prolonged timelines. The implemented data mart automated data cleaning and consolidation through historical analysis, enabling the preparation of reports with reliable data.

In the educational field, decision support systems have been implemented to improve academic management, utilizing Ralph Kimball's methodology [9–14] for the development and implementation of data marts. These BI solutions include indicators for reducing time and activities, as well as interfaces that facilitate the visualization and interpretation of results, thereby improving academic decision-making.

In summary, delivering value to business users must be the primary objective of any BI system [15]. BI solutions help companies become more efficient by identifying specific areas that reveal new business opportunities. While companies often spend excessive time searching for information across multiple departmental data sources, BI centralizes information and makes it visually accessible through dashboards or reports, resulting in significant time savings [16]. As a result, organizations are aggressively leveraging their data assets by implementing and experimenting with data analysis techniques to drive business decisions and provide new functionalities. Today, it is difficult to find a successful company that does not utilize BI technology in its business, as highlighted in discussions of the benefits of combining Business Intelligence and BRMS [17].

The combination of Business Intelligence (BI) with Data Analytics is known as Business Intelligence & Analytics (BI&A). In the last decade, this approach has become indispensable in the business realm, where evidence-based decision-making and high-quality information are critical. However, its application has not been widespread in companies, as its implementation poses challenges, including managing new programs and advanced training for staff.

In this context, Stanley Black & Decker, a multinational company that markets electric and manual tools, as well as home improvement products in nearly 60 countries through physical and e-commerce channels, conducts periodic sales analyses as part of its positioning and productivity strategy. However, for its Latin American operations—encompassing Mexico, Argentina, Chile, and Uruguay—sales reports are required, and the retailer information and data are stored on the Datamind platform. The company has identified challenges related to the time spent managing these reports, the volume of information reported by each retailer, the manual cleaning and processing of data, and the number of files required for generating reports in Power BI.

Therefore, this paper presents a proposal to optimize and automate the process of obtaining weekly and monthly sales reports for Stanley Black & Decker in Mexico, Chile, Argentina, and Uruguay, from 2018 to the present. The aim is to reduce report generation time, increase data quality, and minimize the number of files used by leveraging tools such as Python, Power BI, Windows Task Scheduler, and Power Automate.

The article is structured as follows: first, a conceptual framework is established as a theoretical introduction and reference for the centralization and optimization strategy adopted in the research and development. Next, the materials and methods used for data processing, transformation, loading, and integration are described. Subsequently, the results are presented in terms of coding, storage, and timing. Then, a discussion of the results is conducted in relation to precedents from implementations in Latin America. Finally, the conclusions are presented, and future work is proposed based on web scraping techniques to add an extraction module to the automation.

2 Conceptual Framework

To provide context for this work, it's important to establish the conceptual boundaries. We recognize that process optimization, along with computational tools and their taxonomy, forms the core of implementations, along with defining associated processes like centralization, automation, coding, and effective execution tools.

Process optimization is defined as improving business operations to achieve greater efficiency, productivity, and profitability. This involves analyzing existing processes to identify and eliminate inefficiencies and bottlenecks. Common methods include Six Sigma, Lean Management, and Kaizen, which focus on improving process quality through continuous monitoring, analysis, and adjustment of business processes. Key benefits of process optimization include:

- **Increased efficiency:** Accelerating process management, which increases productivity.
- **Cost reduction:** Eliminating weak links and unnecessary steps in process execution.
- **Improved customer satisfaction:** Faster and more efficient processing of inquiries and orders increases customer satisfaction. Additionally, improving the quality of products and services positively impacts the company's image.
- **Recognition of competitive advantage:** Companies with efficient and profitable processes respond more quickly and flexibly to market changes.
- **Establishment of methodological cycles:** Such as PDCA (Plan, Do, Check, Act), extrapolated from management developments in the 1950s.

In summary, process optimization emphasizes the processing of information and documents. Companies collect and store large amounts of data in various formats and systems. However, automated data processing and intelligent routing (such as using workflows) is a process that requires time and is prone to errors, increasing the workload and business risk in a competitive environment, as presented in "Data Manipulation Process Optimization" [18].

However, effective data processing in company departments allows employees to quickly search for specific terms and locate relevant documents without having to read each one. By automating workflows, companies can operate more efficiently, thus increasing productivity. Methodologies like automated data mining can improve competitiveness by providing valuable insights regarding the organization's data.

In addition to the proliferation of data, data analysis refers to the process of examining, cleaning, transforming, and modeling data to discover useful information, draw conclusions, and support decision-making. This involves exploring and understanding datasets to obtain meaningful insights and uncover actionable patterns, trends, and relationships, as defined by ACM Computing Surveys [19].

Therefore, data analysis is a central component of Data Science, applied in a wide range of fields, from marketing and finance to scientific research and healthcare. This aims to facilitate informed decision-making and gain a competitive advantage in the business world, as outlined in "Data Science Essentials in Business Administration: A Multidisciplinary Perspective" [20].

2.1 Software

At this stage, it is important to examine data analysis tools, known as software and applications, that facilitate the processing, exploration, and interpretation of data. These tools allow analysts and professionals to process datasets effectively and efficiently, facilitating analysis, pattern recognition, report generation, and informed decision-making. A wide range of software and methods is covered, and the choice of tools will depend on the type of data and the goals of the analysis [20]. See Table 1.

2.2 Data Centralization

Data centralization is the process of consolidating information from various sources into a single location, typically in a database or a cloud data warehouse. Centralizing data makes it more accessible and secure, while also helping to establish a single source of quality and veracity that enhances decision-making [27].

From a technological perspective, data centralization offers several advantages:

- **Decision-making:** It allows all team members to obtain a comprehensive and up-to-date view of the data they work with. For example, finance teams can gain a complete understanding of cash flow, revenues, and financial metrics. Having the most current information and an overall picture of all their data facilitates forecasting and enables more accurate and real-time strategic decision-making.
- **Efficiency:** It streamlines business operations by eliminating the manual collection of data from disparate sources. In the financial area, this could translate into a significant acceleration in monthly and quarterly reporting.
- **Integrity and compliance:** It maintains data consistency and minimizes errors and discrepancies in datasets.
- **Simplification of analysis and reporting:** It enables advanced analysis by having all relevant information in one place, making it easier to apply predictive models in machine learning.

- **Scalability and flexibility:** As an organization grows, it is essential to adapt to increasing data volumes and changing business needs. A centralized repository can be easily expanded or modified to incorporate new data sources and meet analytical requirements.
- **Data security:** It promotes better security measures and access control, allowing for improved management. Centralized security involves encryption standards and authentication mechanisms designed to protect confidential data from unauthorized access.
- **Data quality:** It facilitates data cleansing and standardization according to strict corporate standards, ensuring that teams and management have more reliable data for analysis.
- **Cost savings:** It reduces duplication of effort, as all data resides in one place. This decreases the need for redundant infrastructure and optimizes data processing workflows.

Table 1: Data Analysis Tools or Software and Applications to Facilitate the Processing, Exploration, and Interpretation of Data.

Tool	Functionality
Microsoft Excel	Spreadsheet software that enables data analysis and chart generation from tabular data. [21]
Microsoft Power BI	Data visualization and report creation tool that allows users to connect multiple data sources and create interactive dashboards. [6]
Tableau	Data visualization tool is known for its ability to create visually appealing visualizations and interactive dashboards. [22, 42].
D3.js and PlotlyI	Data visualization libraries that enable the creation of customized and dynamic charts and visualizations.
Python and R	Programming languages are widely used in data analysis and statistics, with extensive libraries and packages designed specifically for these purposes. [23]
SQL (Structured Query Language)	Programming language is used for managing and querying databases, essential for data analysis within relational database systems. [24]
RapidMiner and KNIME	Data mining tools are used to discover patterns and relationships within datasets.
Scikit-learn and TensorFlow	Machine learning frameworks that include libraries used to develop machine learning models and make predictions. [24].
Hadoop and Apache Spark	Big Data frameworks that enable the processing and analysis of large volumes of data.
QlikView, MicroStrategy, and Looker	Business Intelligence platforms that assist organizations in analyzing and visualizing business data. [25, 26].

Source: Own elaboration

2.2.1 Centralization and Tools like Dataflow in Power BI

Data visualization is critical for understanding and predicting trends. Microsoft Power BI stands out as a powerful tool for creating interactive reports and dashboards. Its capabilities and advantages include:

- **Connection to Multiple Sources:** Power BI can connect to a variety of data sources, from databases to web services.
- **Advanced Data Visualization:** It offers a wide range of visualization options to create easily interpretable reports.
- **Real-Time Updates:** It can connect to real-time data sources, keeping reports up to date.
- **Team Collaboration:** It facilitates team collaboration and online report sharing.
- **Integration:** It integrates efficiently with other Microsoft applications, such as Excel, and can link with other programs, such as R and Python in some cases [28, 29].

2.2.2 Automation of Updates

Business process automation refers to the use of technology to manage recurring tasks or processes within an organization. In this particular case, it allows companies to minimize costs, increase efficiency, and streamline processes that can be complex. Digital transformation is closely related to business automation, as it involves optimizing a business's processes through new technologies, as highlighted in "An intelligent future is calling" [30].

These processes also help to organize high-volume repetitive tasks that are easier to mechanize and lead to greater benefits within the organization. According to McKinsey Digital, teams spend nearly 20% of their time on tasks that could be automated, such as analyzing operational data and reviewing reports. For this reason, the same study notes that executives understand that task automation has become a trend and a necessity, and they are working to incorporate it, as brought up by "A conceptual model of automated financial reporting" [31].

In other words, when considering the automation of a process, it is important to take into account the repetitive tasks that depend on each company's improvement plan to make mission-critical and financial processes effective, as discussed in "On the relevance of reports" [32].

In terms of the data management tools that help process the obtained information, these can be divided into four major categories, as presented in Table 2 below:

Table 2: Data Management Analysis Tools.

Tool Category	Orientation
Cloud Data Management Tools	Cloud-based technologies that integrate various data sources through APIs, webhooks, or direct connections to databases.
ETL and Data Integration Tools	Technologies that enable the specification of data transformations, extraction of information from various sources, and continuous loading of data into a target database.
Master Data Management (MDM) Tools	Technologies that enable the visualization of complex master data across the organization.
Data Visualization and Analysis Tools	Technologies that facilitate the analysis and exploration of large datasets, as well as the generation of reports and dashboards to extract insights and drive critical decisions.

Source: Own elaboration

Finally, in this conceptual analysis, it is necessary to highlight Python as a fundamental pillar in data processing. As emphasized in sources [33–35], it is clear and simple syntax, along with its versatility for handling large volumes of data, makes it relevant due to its ability to perform analysis, design prediction models, generate graphs, and carry out data cleaning tasks. In this sense, Python presents itself as an interesting alternative for "Data Cleaning," thanks to its efficiency and speed in obtaining quality data.

3 Materials and Methodology

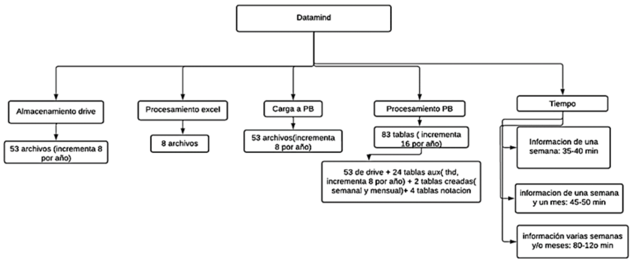
Stanley Black & Decker periodically analyzes its sales to understand the market and inform marketing decisions. These analyses involve collecting sales data reported by their customers (retailers) regarding the amount of money sold, products, and other indicators. Sales reports are generated on different platforms that vary by country, both in traditional channels and e-commerce.

In Latin America, which includes countries like Chile, Mexico, and Argentina (the latter also includes information from Uruguay), information tracking is done through the DataMind platform. Specifically, sales information from e-commerce platforms like Mercado Libre and Amazon is managed by the Data Manager assigned for this region.

Using this data, a sales report is prepared using Power BI Desktop, which allows for a clear visualization of sales behavior, both graphically and in tables. This report is updated weekly, and a global report is generated and updated at the end of each month. However, the information arriving at DataMind is raw and must be processed and cleaned for reporting purposes, as each country compiles the information differently or in different formats.

Initially, an analysis is conducted on how data processing is currently carried out. Then, according to the steps detailed in Figure 1, problems are identified (as shown in Figure 2) that hinder the organization, processing, transformation, and loading of data (ETL processes), as shown in Table 3.

Figure 1: Data process for generating the weekly and monthly sales report at Stanley Black & Decker.



Source: Own elaboration

Figure 2: Identification of issues associated with the generation of the weekly and monthly sales report at Stanley Black & Decker.

Source: Own elaboration

Below, five main tasks that require the implementation of codes in Python are identified:

a) ETL (Extract, Transform, Load) processes for the historical weekly and monthly information from DataMind. This information corresponds to the period between 2018 and 2023.

b) ETL processes for the current (2024) weekly and monthly information from DataMind.

c) ETL processes for historical information from the e-commerce sales channel (Mercado Libre and Amazon). This information also corresponds to the period between 2018 and 2023.

d) ETL processes for the current (2024) information from the e-commerce sales channel (Mercado Libre and Amazon), specifically for weekly information.

e) Sending an email indicating whether the ETL processes mentioned in points b and d were successfully completed. If the process is successful, Power Automate is executed to update the report in Power BI.

f) ETL process of the inventory information, with the goal of unifying the inventory across all countries.

Subsequently, for each of these tasks, an independent code is proposed, due to the specific guidelines derived from different data source origins. In particular, for point e, two codes are required to be implemented. The codes to be developed are presented in Table 4.

Table 3: Name assigned to the codes to be created for the automation and optimization process of the sales report for the company Stanley Black & Decker.

POINT	CODE NAME
a	ProcessETL_Datamind.py
b	Proces_Update_Datamind.py
c	ProcessETL_Meli_Amz_Historic.py
d	ProcessETL_Meli_Amz_Update.py
e	Datamind_Meli_Amz_Update.bat Datamind_Meli_Amz_Update.bat
f	ProcessETL_inventory.py

Source: Own elaboration

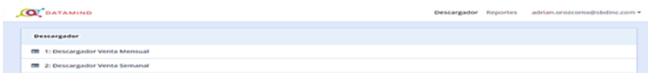
Table 4: Processing and description of the manual weekly and monthly sales report for the company Stanley Black & Decker, and the identified problems.

Process	Description	Problem
Storage	Sales information is located on the DataMind platform. For each of the 4 countries and each year (system initiated in 2018), there is a weekly and a monthly sales file. This results in an increase of 8 files per year in SharePoint, where the processed information is currently stored. Additionally, there are other files (calendars, special dates, etc.).	Currently, there are 52 files, increasing by 8 per year, which increases complexity and makes organization and ETL processes difficult.
Processing in Excel	The files from DataMind must be processed to complete missing information (SBU, brand, and sales channel). This process must be conducted for each of the 8 files of the current year. Additionally, the files are large, which slows down processing. Information formats (SBU, brand, sales channel) change, requiring specific notations (uppercase letters).	Processing must be done for each of the 8 files of the current year, complicated by file sizes and format variability. The update of information is not fully guaranteed.
File Loading to Power BI	Since 2019, information has been loaded into the same Power BI file, so each year the 8 new files necessary for the report must be added.	The substantial number of files slows down report processing and complicates their organization.
File Transformation in Power BI	A transformation process is applied to each file to create the report, ensuring that all files have a unified structure.	Problem 1 (Chile): In the retailer Mercado Libre, sales from París and Falabella are summed. Problem 2 (Mexico): Coppel does not report net sales, but rather units sold; also, The Home Depot (TDH) reports duplicated e-commerce sales. To solve some of the problems mentioned, auxiliary tables must be created each year.
Visualization	For each country, 7 pages are used to visualize the report, containing similar information structured by Country/Retailer cover, Weekly Sales, Monthly Sales, and corresponding charts.	The report presents redundancy, as the pages contain the same structure; the only difference is the source of information and the country from which they come. Therefore, this process could be automated.
Update	This process updates sales reports in Power BI weekly and monthly, starting in DataMind by downloading the files provided by clients, followed by cleaning them.	The times for this process vary depending on the weeks and months to be updated, averaging between 35 and 80 minutes.

Source: Own elaboration

Following this, we propose collecting raw sales data from the company for data processing and subsequent analysis. Sales information from traditional channels is obtained through the DataMind platform, a paid service used by Stanley Black & Decker. This platform allows various clients (retailers) to enter sales-related information on a weekly and monthly basis. Figure 3 illustrates the relationship between monetary value and the products sold. This data is downloaded for each of the following countries: Chile, Mexico, and Argentina, with the latter also including information from Uruguay.

Figure 3: DataMind platform used to obtain raw sales data for Stanley Black & Decker's traditional sales channel.



Source: Own elaboration

Information related to the Mercado Libre and Amazon e-commerce channels is obtained through the e-commerce Data Manager for the region. It's important to note that the process of obtaining this raw data remains manual throughout this project.

For data storage, the OneDrive cloud storage service is utilized, using the Stanley Black & Decker company email. A folder named "Dashboard LAG" is created, containing two main subfolders: "Data Flow" and "Dashboards."

The "Data Flow" folder includes four subfolders: "Datamind," "Master Products-Customers," "Shared Information for Projects," and "Emails."

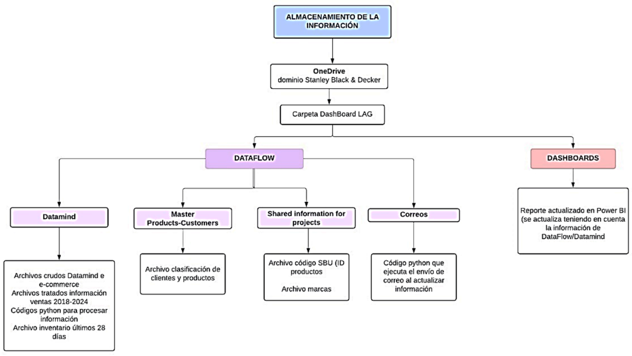
The "Datamind" folder contains the raw sales data files downloaded from DataMind (CSV files), as well as processed sales information from 2018 to the present, cleaned and formatted for consistency (CSV files). The scripts used to process this information are also stored here (ProcesoETL_Datamind.py, Proceso_Update_Datamind.py, ProcesoETL_Meli_Amz_Historico.py, and ProcesoETL_Meli_Amz_Update.py). Additionally, a unified inventory file for the mentioned countries, covering the last 28 days, is included.

The "Shared Information for Projects" and "Master Products-Customers" folders contain other important files. These folders hold information about SBUs (product identification), brand, calendar, and classifications of customers and products, respectively, all of which contribute to the descriptive analyses in the final report.

"The "Dashboards" folder will contain only the most recent Power BI report, updated with data from the files stored in the "Data Flow" folder.

The primary goal of this structure is to maintain a collection of clean, processed, and up-to-date files, centralizing all information required to generate the report, including both historical and processed data. The folder organization is detailed in Figure 4.

Figure 4: Folder organization for storage and centralization of information required for the sales report.



Source: Own elaboration

Data is extracted from the 'Datamind' folder, specifically from 'Data Flow', 'Master Products-Customers', and 'Shared Information for Projects', and loaded into 'Data Flow', Power BI's storage tool. This allows for cross-referencing of the contained information to conduct a descriptive analysis of the company's sales, as illustrated in Figure 5.

Figure 5: Folders in Dataflow used to create the report in Power BI.

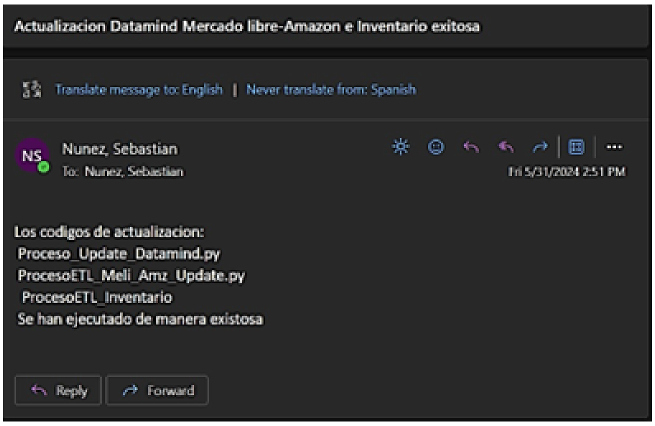
Source: Own elaboration

The Windows Task Scheduler is used to run a .bat file named 'Datamind_Meli_Amz_Update.bat' every Tuesday at 6:00 PM. This timing is chosen because the updated report is distributed every Wednesday.

This .bat file executes the script 'Correo_Update_Datamind_Meli_A which, in turn, runs three additional scripts: 'Proceso_Update_Datamind.py', 'Proceso_Meli_Amz_Update.py', and 'ProcesoETL_inventario.py'. These scripts process the sales and inventory databases using raw data and then load the resulting files into the 'Dataflow/Datamind' folder.

Finally, the main script sends an email to confirm whether the processes have been executed successfully, as shown in Figure 6. If the execution is unsuccessful, the remaining tasks are not performed. The entire process is summarized in Figure 7.

Figure 6: Example of the email sent after the .bat script triggers the execution of all tasks.



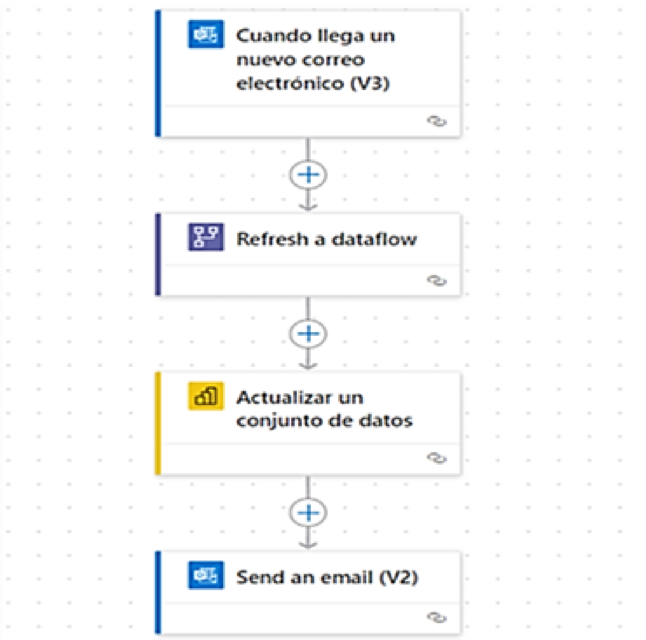
Source: Own elaboration

Figure 7: Process through the Windows Task Scheduler.

Source: Own elaboration

Once the previous phase is successfully completed, Power Automate updates the Data Flow with the refreshed files and automatically refreshes the Power BI Desktop report. If this process is completed correctly, a confirmation email is sent. The complete workflow is presented in Figure 8.

Figure 8: Workflow in Power Automate.



Source: Own elaboration

The result is an updated report presented in a clear and accessible format for decision-making in other areas of the company. Previously, the report's complexity stemmed from the inclusion of multiple countries, resulting in numerous pages. To address this, a country filter has been implemented, enabling users to view information specific to their selection. The report provides a descriptive analysis of the data.

To evaluate the new report and the optimization achieved through the new workflow, a metrics table (Table 5) will be created. These metrics will assess the time spent, the number of files used, and the number of tables required in Dataflow/Power BI to generate the report.

Table 5: Metrics Used to Evaluate Report Optimization.

MÉTRICAS QUE MIDEN LA OPTIMIZACIÓN DEL REPORTE			
MÉTRICA	**DESCRIPCIÓN**	**FÓRMULA**	**UNIDAD**
Tiempo de Actualización (Refresh Time)	Tiempo que tarda el proceso que conlleva actualizar el reporte	(Tiempo de actualización anterior−Tiempo de actualización actual)	minutos o segundos
Número de Archivos de Datos (Number of Data Files)	Cantidad total de archivos que se suben a Power BI o Dataflow antes y después de la optimización.	(Número de archivos anterior−Número de archivos actual)	enteros
Número de Tablas en el Modelo de Datos (Number of Tables in Data Model)	Cantidad total de tablas en el modelo de datos de Power BI.	(Número de tablas anterior−Número de ta blas actual)	enteros

Source: Own elaboration

4 Results

The objective of this project was to optimize Stanley Black & Decker's sales report by integrating tools such as Python, Power BI, Power Automate, and the Windows Task Scheduler. By leveraging these tools and centralizing information using OneDrive as cloud storage, a higher-quality report was produced in less time compared to the previous workflow. This implementation also reduced the occurrence of human errors.

This work establishes a foundation for the company to incorporate Business Intelligence and Analytics practices, which is crucial given the large volume of potentially valuable information within the data.

4.1 Development of Python Scripts

Python scripts were developed to optimize and automate the cleaning and standardization of the company's raw sales reports, as well as to perform other related tasks.

Four Python scripts were created to perform ETL processes on the raw sales files from both Datamind and e-commerce sources, covering historical data (2018-2023) and current data (2024).

In addition, two complementary scripts were developed: one to unify the inventory data for the last 28 days from the four countries (Chile, Mexico, Argentina, and Uruguay), and another to manage email notifications upon successful updates to the sales information (i.e., when the first four scripts are executed). This notification script triggers a series of tasks that culminate in the automatic update of the company's sales report in Power BI.

The developed scripts are listed in Table 6:

Table 6: Information on Python Scripts Created for Automation and Optimization of Stanley Black & Decker's Sales Report Process.

CÓDIGOS	INFORMACIÓN REQUERIDA	TIPO ARCHIVO	UBICACIÓN	DESCRIPCIÓN	RESULTADO
ProcesoETL_Datamind	Data Histórica Datamind 2018-2023 Copppel Fx rate	.py	DashboarLAG/Dataflow/Datamind	Código que procesa la información histórica que proviene de Datamind tanto semanal como mensual de todos los paises	df_datamind_historico.csv
Proceso_Update_Datamind	Data Actualización Datamind 2024 Copppel Fx rate	.py	DashboarLAG/Dataflow/Datamind	Código que procesa la información de actualizacion que proviene de Datamind tanto semanal como mensual de todos los paises	Datamind_actualizacion_historico.csv
ProcesoETL_Meli_Amz_historico	Data histórica Mercado libre-Amazon Gto Net	.py	DashboarLAG/Dataflow/Datamind	Código que procesa la informacion historica hasta 2023 de mercado libre y amazon	meli_amz_historico.csv
ProcesoETL_Meli_Amz_Update	Data actual Mercado libre-Amazon 2024 Gto Net	.py	DashboarLAG/Dataflow/Datamind	Código que procesa la información de 2024 de mercado libre y amazon	meli_amz_update.csv
ProcesoETL_inventario	Inventario de cada país	.py	DashboarLAG/Dataflow/Datamind/Inventario	Código que procesa la información de 2024 de inventario, unifica los inventarios de todos los países	df_datamind_inventario.csv
Correo_Update_Datamind_Meli_Am	Archivo	.py	DashboarLAG/Dataflow/Datamind/Correos	Código que ejecuta los códigos que actualizan la información de Datamind y meli_amz (ACTUAL) y envía correo de notificaciÓn	
Datamind_Meli_Amz_Update	Archivo	.bat	DashboarLAG/Dataflow/Datamind/Correos	Código que ejecuta de manera programada el código correo update datamind meli_amz.py	

Source: Own elaboration

It is important to note that each script requires specific source files for processing. Previously, both historical and current data files were required. In practice, the following files are essential, along with their corresponding functions:

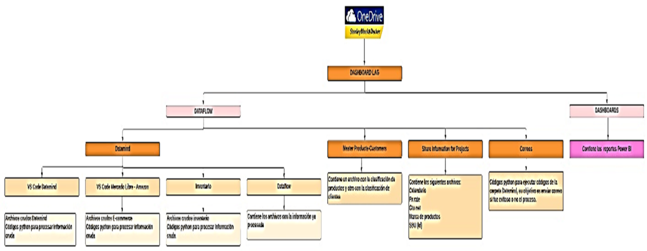
- **Coppel:** This file contains the unit price of the company's products. When cross-referenced with the Coppel report (a Mexican retailer that only provides units sold), the net sales in monetary terms can be calculated.
- **Fx rate:** This file includes the historical dollar exchange rate for each country for the relevant month. This is necessary because the company prefers to interpret reports with sales values converted from each country's local currency to US dollars.
- **Gto net:** This file allows for the calculation of the net selling price after applying discounts related to applicable economic activity taxes.

The scripts "Proceso_ETL_Datamind.py" and "ProcesoETL_Meli_Amz_historico.py" are executed only once, as the historical data in the respective reports remains static.

4.2 Information Storage

Following the creation of the scripts, the OneDrive folder structure was organized as described in the Materials and Methods section to manage and store all files required for report preparation. The overall structure is shown in Figure 9.

Figure 9: Final Organization of OneDrive / DashboardLAG Folder.



Source: Own elaboration

The primary objective of centralizing all information is to provide various departments with access and to facilitate the connection of these files with Power BI Dataflow for report preparation. The folder that provides this direct connection is located at DashboardLAG/Dataflow/Datamind/Dataflow, as it contains all updated and processed files ready for use.

4.3 Composition and Generation of the Sales Report

As detailed in Table 7, the previous report required approximately 52 files and 24 tables in Power BI. Furthermore, an estimated annual growth of 8 files and 8 tables was projected. This resulted in a total requirement of 80 files and tables, making the report complex to update and maintain. Troubleshooting failures was also challenging due to this complexity.

Table 7: Number of files used to obtain the report according to the old version workflow.

ARCHIVOS NECESARIOS FLUJO DE TRABAJO VERSIÓN ANTIGUA				
UBICACIÓN	DRIVE	CANTIDAD	INFORMACIÓN COMPLEMENTARIA	NOTA
Carpeta Datamind Sell Out	Archivos que almacenan la informacion descargada de datamind desde 2019-2024	48	12 archivos por pais (2018-2024) 6 años: 6 archivos semanales - 6 mensuales	Por año incrementan en 8 archivos 2 por país(1 sem - 1men)
Carpeta Datamind Sell Out	Archivo que almacena la informacion de inventario(de todos los paises)	1		
Carpeta Datamind Sell Out	Archivo que almacena la informacion de precios de unidades vendidad por coppel	1		
Carpeta Datamind Sell Out	Archivo con un calendario	1		Es un calendario adaptado a Datamind
Carpeta Datamind Sell Out	Archivo que contiene información sobre la clasificación de los productos	1		
	TOTAL ARCHIVOS	52		
TABLAS AUXILIARES NECESARIAS EN POWER BI FLUJO DE TRABAJO VERSIÓN ANTIGUA				
UBICACIÓN	TABLAS AUXILIARES POWER BI	CANTIDAD		NOTA
Power BI	Tablas para solucionar problema de The Home Deppot a partir del 2022	24	8 archivos por año 4 para información semanal - 4 para información mensual	por cada año incrementan 8 tablas 4 semanales - 4 mensuales
Power BI	Tabla que almancena la notación de sbu	1		
Power BI	Tabla que almancena la notación de canal de venta	1		
Power BI	Tabla que unifica la información semanal	1		
Power BI	Tabla que unifica la información mensual	1		
	TOTAL TABLAS	28		

Source: Own elaboration

In contrast, the current implementation has significantly reduced these requirements to 8 files and one auxiliary table, totaling just 9. This represents a substantial improvement in process efficiency. The organization is now clear, with automatically processed files that centralize the required information, as shown in Table 8.

Table 8: Comparison of Files Used in the Traditional Workflow Versus the Optimized Sales Report.

ARCHIVOS NECESARIOS FLUJO DE TRABAJO NUEVA VERSIÓN		
UBICACIÓN	DATAFLOW	CANTIDAD
Dataflow datamind	Archivo que contiene la información histórica semanal y mensual de Datamind 2018-2023	1
Dataflow datamind	Archivo que contiene la información histórica semanal y mensual de mercado libre - amazon 2023	1
Dataflow datamind	Archivo que contiene la información semanal y mensual actual de Datamind 2024	1
Dataflow datamind	Archivo que contiene la información semanal y mensual actual de mercado libre - amazon 2024	1
Dataflow Master products - Customers	Archivo que contiene información sobre la clasificación de los productos	1
Dataflow Shared Information Proyect	Archivo que contiene notación de los sbu	1
Dataflow Shared Information Proyect	Archivo que contiene notación de las marcas	1
	Archivo que contiene calendario	1
	TOTAL ARCHIVOS	8
TABLAS AUXILIARES NECESARIAS FLUJO DE TRABAJO NUEVA VERSIÓN		
Power BI	Tabla que unifica la informacion historica y actual de datamind mercado libre y amazon	1
	TOTAL TABLAS	1

Source: Own elaboration

As shown in Table 9, the new workflow generates a report that is 44 pages shorter than the previous version, representing an 85% reduction in length.

Table 9: Comparative metric on the number of files used for the two reports.

Número de Archivos de Datos (Number of Data Files)	
Número archivos reporte anterior	52
Número archivos reporte actual	8
TOTAL ARCHIVOS	**44**

Source: Own elaboration

Table 10: Comparative metric in the number of tables used for the two reports.

Número de Tablas en el Modelo de Datos (Number of Tables in Data Model)	
Número tablas reporte anterior	80
Número tablas reporte actual	8
TOTAL TABLAS	**72**

Source: Own elaboration

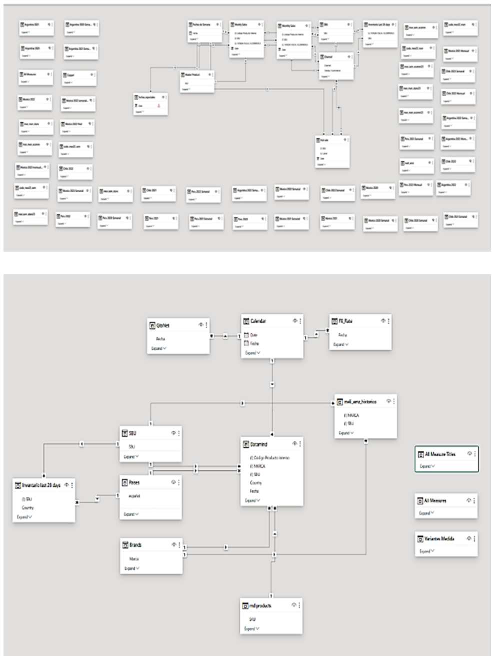
This, in turn, can be visualized with the logical diagrams that Power BI provides of the relationships between the tables used to generate the report. In Figure 11a and 11b, it can be observed that in the report of the previous version there is a greater number of tables compared to the current report.

4.4 Report on Generation Times

The automation and optimization of the sales report have resulted in a significant reduction in report generation time. Data indicates that the new workflow has improved performance, reducing generation time by 60.18%. The previous version required approximately 65 minutes to generate the report, whereas the new workflow reduces this time to approximately 26 minutes (25.9), resulting in a time savings of 39.1 minutes, as shown in Table 11.

Notably, the only step with consistent time requirements is the download of raw data from the Datamind platform. As this platform requires permissions and cannot be linked for automation (e.g., via a Python script), this step must continue to be performed manually.

Figure 10: Logical diagrams in Power BI that show the relationships between the tables that feed the report obtained from the old workflow (a) and the current workflow (b).



a)

b)

Source: Own elaboration

Table 11: Metric of the amount of time spent between reports (a), Comparison of the amount of time taken to obtain the report through the traditional workflow (b) and through the proposed report optimization (c). The time unit is seconds.

a)

Tiempo de Actualización (Refresh Time)	
Tiempo actualización anterior	3905
Tiempo actualización actual	1555
TOTAL TIEMPO EN SEGUNDOS	**2350**
TOTAL TIEMPO EN MINUTOS	**39,10**

b)

FLUJO DE TRABAJO OBTENCIÓN REPORTE VERSIÓN ANTIGUA			
TAREA	TIEMPO	TOTAL ARCHIVOS	TOTAL TIEMPO
Descargar un archivo con la información semanal de una o varias semanas o mes de DATAMIND, por cada país (México, Chile, Argentina y Uruguay)	160	8	1280
Procesar cada archivo descargado y agregar la información a SharePoint (Almacenamiento en la nube)	280	8	2240
Actualizar dataset Power BI	385	1	385
Tiempo total que toma actualizar la información semanal y mensual de los 4 países			3905
TIEMPO EN MINUTOS			65,08

c)

FLUJO DE TRABAJO OBTENCIÓN REPORTE NUEVA VERSIÓN	
TAREA	TOTAL TIEMPO
Descargar los archivos con la información semanal de una o varias semanas o mes de DATAMIND, por cada país (México, Chile, Argentina y Uruguay)	160
Procesar la información descargada mediante Python	90
Procesar información Mercado libre y Amazon	20
Actualizar el almacenamiento masivo Dataflow Datamind	900
Actualizar dataset Power BI	385
Tiempo total que toma actualizar la información semanal y mensual de los 4 países	**1555**
TIEMPO EN MINUTOS	**25,9**

Source: Own elaboration

The Previous Workflow and the Current Improvement The previous version of the workflow for the report required manual data cleaning, involving processing each file individually. This meant managing the 8 existing information files: two per country (1 for weekly sales and 1 for monthly sales). In contrast, the current workflow uses the Python scripts Datamind_Meli_Amz_Update.bat and Correo_Update_Datamind_Meli_Amz.py to automatically execute 3 scripts for the ETL processes of the current information (Proceso_Update_Datamind.py, ProcesoETL_Meli_Amz_Update.py, and ProcesoETL_inventario.py). This significantly reduces processing time, as the processed data is obtained in approximately 90 seconds (1.5 minutes).

Additionally, a new step has been included in the current workflow: the incorporation of e-commerce information for analysis.

5 Results Discussion

An automated and optimized report has been successfully developed. It is considered automated because the Python scripts perform ETL (Extract, Transform, Load) processes automatically, generating clean and treated .csv files that provide complete, high-quality information. This addresses the previously identified difficulties presented by the raw data reported by retailers.

It is optimized because the process has seen a significant reduction in time (61%), number of files (85%), and number of tables needed (91%) for report preparation. Furthermore, the information has been completely centralized, making it accessible and clear for various departments within Stanley Black & Decker. Thanks to the integration of different technologies and Business Intelligence (BI) tools, a much better-optimized report was achieved compared to the previous version (see Annex 1).

Previous studies, such as the Optimization of Reporting and Administrative Reporting Processes in the Digital Sales Area of IBM Colombia, which aimed to optimize report preparation in IBM Colombia's digital sales area using VBA in Excel, found a significant reduction in creation time (by more than 60%), from 6 minutes to 1.2 minutes. These results align with the percentages found in the present study, leading to the conclusion that optimizing and automating processes provide valuable improvements in terms of efficiency and agility. Additionally, automation helps to reduce the occurrence of human errors, as discussed in the Evolution of Business Organizations [37].

On the other hand, studies such as the one presented in Modeling Impact of Human Errors on Data Unavailability and Data Loss of Storage Systems suggest that human errors can significantly influence data loss, potentially leading to misinterpretations of reality. This is similarly addressed in the study mentioned in Pace, Preventing Human Error [39].

Implementing data centralization and processing clearly improves the quality of information and, consequently, the interpretations derived from it. This aligns with the findings [40] that expert decisions vary depending on data quality, highlighting the importance of establishing data standards. For example, [41] highlights that the effectiveness of a decision is directly proportional to data quality.

Finally, this type of project is particularly relevant today. Stanley Black & Decker will have a better understanding of its sales, enabling more informed marketing decisions. Additionally, this would encourage the company to explore the use of Business Intelligence and Analytics (BI&A), strengthening its position in the market.

6 Conclusions and Further Studies

This project demonstrates the significant benefits of integrating tools such as Python, Power BI, Power Automate, and the Windows Task Scheduler for report optimization and automation. Significant reductions in time, the number of files, and the number of tables required for generating sales reports were observed, along with a notable improvement in data quality.

Automated data cleaning was achieved using Python scripts, ensuring better data quality and avoiding human errors that could occur with manual processes.

The sales report for Stanley Black & Decker was optimized and automated for both its traditional sales channel and e-commerce. This optimization resulted in a 61% decrease in report update time, an 85% reduction in the number of files used, and a 91% reduction in the number of tables required in Power BI Dataflow. The decreased update time reflects a significant improvement in process efficiency, enabling faster data access for users. Furthermore, the reduction in the number of files and tables indicates a consolidation and simplification of data sources, which likely reduces the complexity of managing information and minimizes potential integration errors.

Information was centralized through the creation of a OneDrive folder named Dashboard LAG. Within this folder, subfolders were established to effectively organize the information, facilitating access for any department within the company and providing a clear understanding of the files used.

These results not only demonstrate technical optimization but also imply a potential improvement in real-time, data-driven decision-making, thanks to the increased speed and simplicity of the update process.

In the future, it would be important to consider collecting data from platforms like Datamind using web scraping techniques to automatically extract large volumes of data from sources such as websites, as well as incorporating other applications that could offer business advantages.

Additionally, implementing similar projects in other areas and processes of multinational companies, such as updating customer information and data unrelated to Datamind, could prove highly competitive.

In this context, conducting predictive analytics to extract valuable information from data, such as exploring machine learning techniques, can provide insights not attainable through descriptive analyses. This constitutes a development area for a new generation of Data Science solutions, a cross-disciplinary field relevant to traditional technological and engineering solutions.

References

[1] R. F. Matamoros Z., "Implantación en una empresa de un sistema business intelligence saas/on demand a través de la plataforma litebi," Ph.D. dissertation, Universitat Politècnica de València, 2010.

[2] A. Rozenfarb, *Impacto de la business intelligence en el proceso de toma de decisiones*. México D.F: Instituto Politécnico Nacional, 2008.

[3] E. Iniesta, "Power bi y su utilidad en la gestión empresarial," M.S. thesis, Universidad de Valladolid, Valladolid, 2022.

[4] F. Rolando and J. C. Sanchez S., "Análisis de la información generada utilizando herramientas de inteligencia de negocio y ciencia de datos," Ph.D. dissertation, Universidad Nacional de Mar del Plata. Facultad de Ingeniería, Argentina, 2022.

[5] J. L. Parra C., M. A. Rincón R., and D. P. Romero S., "Propuesta de inteligencia de negocios mediante la herramienta microsoft power bi, como soporte para la toma de decisiones del área comercial de la empresa abc manufacturera de productos plásticos," Tech. Rep., 2019.

[6] F. J. Rivera R., "Industrial aplicación de business intelligence en una pequeña empresa mediante el uso de power bi," M.S. thesis, Universidad de Valladolid Escuela de ingenierias industrials, Valladolid, 2018.

[7] M. E. Carhuaricra I. and J. I. Gonzales C., "Implementación de business intelligence para mejorar la eficiencia de la toma de decisiones en la gestión de proyectos," Tech. Rep., 2017.

[8] A. Rojas Z., "Implementación de un data mart como solución de inteligencia de negocios, bajo la metodologı́a de ralph kimball para optimizar la toma de decisiones en el departamento de finanzas de la contralorı́a general de la república," Tech. Rep., 2014.

[9] R. Kimball and J. Caserta, *The Data Warehouse ETL Toolkit*. Indianapolis: Wiley, 2004.

[10] R. Kimball and R. Merz, *The Data Webhouse Toolkit: Building the WebEnabled Data Warehouse*. Wiley, 2000.

[11] R. Kimball and M. Ross, *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*, Second Edition. New York: Wiley, 2002.

[12] R. Kimball and M. Ross, *The Kimball Group Reader; Relentlessly Practical Tools for Data Warehousing and Business Intelligence*. Indianapolis: Wiley, 2010.

[13] R. Kimball et al., *The Data Warehouse Lifecycle Toolkit*, 2nd Edition. New York: Wiley, 2008.

[14] R. Kimball et al., *The Data Warehouse Lifecycle Toolkit*. New York: Wiley, 1998.

[15] P. Muryjas and M. Wawer, "Business intelligence as a support in human resources strategies realization in contemporary organizations," *Aktualni problemy ekonomiky*, no. 2, pp. 183–190, 2014.

[16] P. Křı́ž, "Business intelligence řešenı́ pro společnost 1188," M.S. thesis, Brno University of Technology, 2015. [Online]. Available: https://dspace.vutbr.cz/bitstream/handle/11012/39466/Jan-Kriz-Masterthesis.pdf?sequence=1

[17] A. Hassanifar, *How business intelligence and brms bring benefits*, 2015.

[18] G. Mensa F., "Optimización del proceso de manipulación de datos," M.S. thesis, Universitat Politècnica de Catalunya, 2022.

[19] L. Cao, "Data science," *ACM Computing Surveys*, vol. 50, no. 3, pp. 1–42, 2017. DOI: 10.1145/3076253

[20] M. Afsharian, "Data science essentials in business administration: A multidisciplinary perspective," *Decision Analytics Journal*, p. 100 442, 2024.

[21] S. W. Palocsay, I. S. Markham, and S. E. Markham, "Utilizing and teaching data tools in excel for exploratory analysis," *Journal of Business Research*, vol. 63, no. 2, pp. 191–206, 2010.

[22] S. A. Murphy, "Data visualization and rapid analytics: Applying tableau desktop to support library decision-making," *Journal of Web Librarianship*, vol. 7, no. 4, pp. 465–476, 2013. DOI: 10.1080/19322909.2013.825148

[23] S. A. Fahad and A. E. Yahya, "Big data visualization: Allotting by r and python with gui tools," in *2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE)*, IEEE, 2018, pp. 1–8.

[24] I. Setiawan, "Perbedaan data engineer, data scientist dan data analyst," *Widya Accarya*, vol. 12, no. 2, pp. 306–309, 2021.

[25] R. J. Podeschi, "Experiential learning using qlikview business intelligence software," *Information Systems Education Journal*, vol. 13, no. 4, p. 71, 2015.

[26] A. Ibarra, *Business intelligence y la información como un activo estratégico – nebi*, Recuperado en marzo 2019, 2014. [Online]. Available: http://www.nebi.co/bi/business-intelligence-y-lainformacion-como-un-activo-estrategico/

[27] Astera, *Data ware house concepts: Kimball vs. inmon approach*, Acessado en jun. 2021, 2020. [Online]. Available: https://www.astera.com/type/blog/data-warehouse-concepts/

[28] R. Wade, *Advanced Analytics in Power BI with R and Python*. 2020. DOI: 10.1007/978-1-4842-5829-3

[29] V. Naneva and K. Stefanova, "Optimization techniques for business intelligence reports," in *AIP Conference Proceedings*, vol. 2333, AIP Publishing, 2021.

[30] Zendesk, *An intelligent future is calling—here's a look back at zendesk relate 2023*, 2023. [Online]. Available: https://www.zendesk.es/blog/look-back-zendesk-relate-2023/

[31] I. A. Baldursson, "A conceptual model of automated financial reporting," Ph.D. dissertation, 2022.

[32] M. Schulz, P. Winter, and S. K. T. Choi, "On the relevance of reports—integrating an automated archiving component into a business intelligence system," *International Journal of Information Management*, vol. 35, no. 6, pp. 662–671, 2015.

[33] R. González D., *Python para todos*, En línea, 2011. [Online]. Available: https://drive.google.com/file/d/0B5VsS3feMSSgNTU4OGJkZDktMjgzNy00NGNlLWFkYj view?resourcekey=0-bAgfYp82Tah5u0Xaj7pGgQ

[34] D. Dasari and P. S. Varma, "Employing various data cleaning techniques to achieve better data quality using python," in *2022 6th International Conference on Electronics, Communication and Aerospace Technology*, IEEE, 2022, pp. 1379–1383.

[35] A. Rawat, "A review on python programming," *International Journal of Research in Engineering, Science and Management*, vol. 3, no. 12, pp. 8–11, 2020.

[36] G. A. Arias S., "Optimización de los procesos de elaboración de reportes e informes administrativos en el área de ventas digitales de ibm colombia," Tech. Rep., 2019.

[37] L. Antwiadjei, "Evolution of business organizations: An analysis of robotic process automation," *Eduzone: International Peer Reviewed/Refereed Multidisciplinary Journal*, vol. 10, no. 2, pp. 101–105, 2021.

[38] M. Kishani and H. Asadi, "Modeling impact of human errors on the data unavailability and data loss of storage systems," *IEEE Transactions on Reliability*, vol. 67, no. 3, pp. 1111–1127, 2018.

[39] K. A. Barchard and L. A. Pace, "Preventing human error: The impact of data entry methods on data accuracy and statistical results," *Computers in Human Behavior*, vol. 27, no. 5, pp. 1834–1839, 2011.

[40] I. N. Chengalur-Smith, D. P. Ballou, and H. L. Pazer, "The impact of data quality information on decision making: An exploratory analysis," *IEEE transactions on knowledge and data engineering*, vol. 11, no. 6, pp. 853–864, 1999.

[41] C. Batini, C. Cappiello, C. Francalanci, and A. Maurino, "Methodologies for data quality assessment and improvement," *ACM Computing Surveys*, vol. 41, no. 3, pp. 1–52, 2009. DOI: 10.1145/1541880.1541883

[42] Tableau Software, *Todos los derechos reservados*, En línea, 2023. [Online]. Available: https://www.tableau.com/es-es/learn/tutorials/ondemand/forecast-9