Article

Data Storytelling and Decision-Making in Seaport Operations: A New Approach Based on Business Intelligence

Marco Gonçalves [1], Cátia Salgado [2], Amaro de Sousa [3] and Leonor Teixeira [1,4,*]

[1] Department of Economics, Management, Industrial Engineering and Tourism (DEGEIT), University of Aveiro, 3810-193 Aveiro, Portugal; marcofgoncalves@ua.pt

[2] Administração dos Portos de Sines e do Algarve (APS), 7521-953 Sines, Portugal; catia.salgado@apsinesalgarve.pt

[3] Institute of Telecommunications—Pole of Aveiro (IT), Department of Electronics, Telecommunications and Informatics (DETI), University of Aveiro, 3810-193 Aveiro, Portugal; asou@ua.pt

[4] Intelligent Systems Associate Laboratory (LASI), Institute of Electronics and Informatics Engineering of Aveiro (IEETA), University of Aveiro, 3810-193 Aveiro, Portugal

* Correspondence: lteixeira@ua.pt; Tel.: +351-234370361

Abstract: Seaports are experiencing several challenges due to the explosive growth of the maritime shipping business, which has led to the need for digitalized operations and more effective solutions. This article provides a comprehensive exploration of the process used to create a reliable business intelligence solution by analyzing the container delivery and pick-up services flow in one of Portugal's largest maritime container ports, using the CRISP-DM methodology. The solution, built with Microsoft Power BI®, provides the capability to identify and address data anomalies and present key performance indicators in visually dynamic dashboards. This solution empowers stakeholders to gain invaluable insights into the current and future operational status, thereby facilitating well-informed and adaptable decision-making, representing the main practical contributions. As a theoretical contribution, this study advances research by covering a gap in the literature and establishing the foundations for future business intelligence applications within the maritime industry, with a focus on addressing data dispersion challenges, enhancing logistics flow analysis, and reducing port congestion. The manuscript is structured into seven sections: introduction, literature review, port challenges, methodology, tool development, SWOT analysis, and conclusion.

Keywords: business intelligence; performance indicators; dashboard; CRISP-DM; seaports

1. Introduction

Maritime transport is a critical component of the global economy, facilitating the movement of 12 billion tons of goods in 2022 [1]. A significant portion of these goods is transported in containers, which offer a versatile and efficient means of moving products across various modes of transportation, such as ships, trucks, and trains. The demand for container transport is expected to grow, with a projected annual increase of 3% in container volume from 2024 to 2028 [2].

This growing reliance on maritime transportation comes with its own set of challenges, particularly in terms of port logistics. Efficient operations at container terminals are essential to maintain scheduled services and minimize disruptions [3]. Issues, such as peak-hour traffic and delays, require the development of systems capable of predicting and addressing these challenges [4]. While some ports have implemented truck scheduling systems to

reduce waiting times and emissions, there remains a pressing need for more comprehensive solutions to address the logistical complexities at terminal entrances [5].

In response to these challenges, the concept of smart ports has gained momentum, incorporating Industry 4.0 principles and digitalization to transform supply chain management. Smart port initiatives utilize advanced technologies to enable large-scale data analysis and process automation, boosting operational efficiency and resilience [6]. Issa Zadeh et al. [7] defines smart ports through three key sectors: intelligent logistics, intelligent infrastructure, and intelligent traffic. These sectors include technologies such as IoT, platforms, sensors, intelligent cargo handling, and energy management systems. When these components are effectively integrated, they optimize port operations, enabling them to meet growing demands with greater efficiency.

Despite recognizing the need for innovation, many ports still rely on manual solutions that lead to inefficiencies [6]. Modernizing ports through digitalization is essential, requiring the integration of systems and collaboration among stakeholders [8]. This transformation involves creating infrastructure for real-time data collection and processing, leveraging business intelligence tools to enhance planning and control of port operations [9]. Additionally, intelligent network technologies—such as IoT, information and communication technology (ICT), and smart energy systems—play a critical role in improving energy efficiency, sustainability, and the overall performance of seaports. These technologies enable secure communication, real-time data processing, and effective energy management, supporting the sustainable and efficient operation of smart ports [10].

International organizations, such as the United Nations Conference on Trade and Development (UNCTAD), advocate for increased investment in digitalization to enhance port efficiency and resilience. Furthermore, governments are urged to drive reforms in port infrastructure and operations to facilitate the adoption of digital solutions [11].

This research aims to drive the digital and ecological transition in transport, logistics, and port operations by introducing a decision support tool developed using Microsoft Power BI. While Power BI itself may not be a novel technology, the innovation in this research lies in how it is applied to integrate and analyze fragmented data from port operations, providing actionable insights that directly inform decision-making and improve operational efficiency. By addressing challenges, such as data dispersion, enhancing logistics flow analysis, and reducing port congestion, this investigation presents a compelling case for the widespread application of business intelligence in the maritime sector, emphasizing its positive impact on port management and operations. For this purpose, the application of the case took place in one of the largest container maritime ports in Portugal.

This paper is structured as follows: Section 2 reviews digital transformation in port logistics, focusing on data-driven decision-making and BI's impact on data analysis. Section 3 outlines the logistics processes at the terminal, highlighting container handling, service scheduling challenges, and the need for the Power BI tool. Section 4 describes the development methodology. Section 5 details the BI tool's design, covering requirements, data models, and dashboards. Section 6 evaluates the tool with a SWOT analysis. Section 7 summarizes the paper's contribution and future research directions.

2. Literature Review

2.1. Digital Transition in Port Logistics Operations

The history of seaports can be divided into five generations, each with significant technological advances and operational changes. Similarly, many forms of digitalization have been implemented in ports, amounting to three generations of digitization over the five generations of ports [1].

Before 1950, the first generation mainly involved manual port operations, such as cargo handling and paper-based unloading procedures. Over time, these procedures shifted to electronic formats through electronic data interchange (EDI), improving data exchange and processing [1,11]. The second generation, until 1980, transformed ports into centers of value-added services. This era saw increased raw material handling and the emergence of the first port communication systems (PCS), which improved coordination and efficiency in port operations [11,12]. Starting in 1980, the third generation saw the increasing importance of intermodal transport. Information technologies aimed at enhancing operational efficiency were introduced, with ports adopting terminal operating systems (TOS) to integrate data from various technologies and subsystems for more cohesive and effective management [11,13]. The fourth generation, from 1990, brought innovation through the integration of port companies into common administrations and the development of automated technologies, which are crucial for improving efficiency and automation in terminal operations [12,13]. The fifth generation, from 2010, introduced concepts such as "Smart Ports", "Industry 4.0", "Digitalization", and "Sustainability". Ports began using smart devices and mobile applications to optimize traffic and cargo flows. The integration of various control centers into a centralized system aims to allow for real-time data analysis, transforming how ports operate and adapt to modern needs [14].

Each generation brought important advancements, shaping modern ports and preparing them to meet the challenges and opportunities of the current global economy. The key aspects of each generation of port digitalization are summarized in Table 1, illustrating the continuous evolution and increasing complexity of port operations.

Table 1. Core elements of each port and digitalization generation, based on [12].

Port Generation	Digitalization Generation	Technology	Purpose	Contribution	Obstacle
First (Before 1960) Second (1960–1980) Third (1980–1990)	First (1950–1990)	EDI, PCS, TOS	- Transformation of all information exchanges between the various companies using the port from physical to electronic methods	- Better planning of process execution due to the availability of information	- Low adherence of the port community to the systems created, as companies do not abdicate from using physical documents
Fourth (1990–2010)	Second (1990–2010)	RFID, TAS, Laser	- Automation of terminal activities	- Increased efficiency and operational capacity	- Collected information is static - Low adoption of the truck appointment system (TAS)
Fifth (After 2010)	Third (since 2010)	Mobile Apps, Cloud, Smart devices	- Creation of intelligent processes that integrate port activity	- Better coordination and monitoring of port activity	- Inefficiencies in information flow - A certain reluctance among stakeholders to use disruptive processes

2.2. Decision-Making and the Potential of Data in Seaports

Decision-making is the cognitive process of reasoning used to select an action from among several alternatives, which is considered crucial in organizations because it has a decisive influence on their ability to adapt to change [15].

Ports generate a variety of data types, including logistical data, operational statistics, cargo and shipping information, as well as environmental and security data. Business intelligence tools play a crucial role in transforming this diverse and voluminous data into actionable insights. These tools employ processes such as extract-transform-load (ETL) and data virtualization to integrate data from disparate sources, ensuring that the information is accurate and accessible for analysis [16]. Advanced analytics and AI-driven insights, as part of BI, enable predictive capabilities, providing foresight into future trends and potential issues, which is essential for the dynamic environment of ports [17].

The potential of data in seaports is immense, as it supports the evolution of these key hubs in global trade networks into intelligent, efficient, and sustainable operations. Technologies such as the seaport data space (SDS) and big data architectures allow for secure data sharing and improved decision-making through key performance indicators (KPIs) displayed on dashboards. This integration can lead to decreased transaction costs and enhanced operational quality [18]. Additionally, integrating port community systems (PCS) is vital for the competitiveness of seaports, as it promotes coordinated communication among stakeholders and boosts operational efficiency [19].

Moreover, the role of data extends beyond the seaports themselves to their interconnections with inland freight distribution systems, such as dry ports. Research indicates that operations at dry ports can impact seaport competitiveness by improving performance, service variety, and capacity [20].

The increasing focus on data-driven decision-making in seaports is highlighted by [21], who discussed the role of digital twins in facilitating transparent and controlled seaport operations. Similarly, [22] emphasized the importance of advanced technologies and digitalization in optimizing seaport management. Furthermore, [20] underscored the impact of dry port operations on seaport competitiveness, suggesting that seamless integration can enhance overall performance and capacity.

Finally, [23] emphasized the importance of operational performance indicators and the optimization of logistics networks for enhancing seaport efficiency.

Decision-making must be efficient and effective, and data plays a key role in providing information to identify problems, analyze alternatives, and find solutions [24].

2.3. Business Intelligence—Benefits and Applications to Seaports

The concept of business intelligence (BI) is the discipline that integrates infrastructure to select, extract, process, and visualize data, enabling organizations to make data-driven decisions [25].

BI has the potential to enhance decision-making in seaports by enabling the integration and analysis of diverse data sources. This can result in improved operational efficiency and better strategic planning [26]. Through data visualization and reporting tools, shipping data, cargo movements, and vessel tracking are presented in a clear and actionable manner [27]. Predictive analytics utilizes historical data to forecast cargo volumes, enabling ports to allocate resources effectively and prepare for future demands [28]. Monitoring operational performance through KPIs helps evaluate and improve the efficiency of port operations [29]. Furthermore, BI applications facilitate supply chain optimization by enhancing stakeholder collaboration through effective data sharing, ultimately leading to more streamlined and responsive supply chain management [30].

Several authors highlight the diverse benefits of BI in the maritime sector. For instance, the integration of BI with X-ray scanning technology has been proposed to improve maritime security by providing enhanced visibility in supply chain operations [31]. Additionally, the digitalization and automation of seaport infrastructures, supported by BI, contribute to the optimization of goods and people management, as well as facilitating a green energy transition [22]. Contradictions or interesting facts emerge when considering the varying stages of BI implementation across different seaports. While some ports, like those in Croatia, are in the developmental phase, focusing on simplifying business processes and stakeholder connectivity [32], others, such as those in China, have achieved notable advancements in automation and intelligence, but still face challenges in fully leveraging BI for business operations and decisions [33].

Digitalization is essential for the future of ports, as demonstrated by [7], who show that ports with smart energy infrastructures can significantly reduce carbon emissions and improve operational efficiency, while ports that do not adopt such solutions face inefficiencies and high environmental [10]. Moreover, the implementation of business intelligence (BI) enables ports to make quick decisions based on real-time data, optimizing logistics flows and reducing their environmental impact. In contrast, ports that do not adopt BI continue to operate with outdated data, resulting in operational inefficiencies and greater environmental impacts [10]. Therefore, digitalization is an urgent necessity to ensure the sustainability and competitiveness of ports in the future.

Ain et al. [34] highlight that BI systems offer a flexible technological solution for accessing data from multiple sources, enabling the accumulation, integration, and analysis of data to identify opportunities, strengths, and weaknesses. These systems support decision-making by facilitating advanced integration and management of structured and unstructured data, handling large volumes of data, empowering end users with enhanced processing capabilities to derive new insights, and providing solutions for ad hoc analysis and queries.

BI applications encompass data mining techniques, data visualization, and performance analysis, collectively converting raw data into valuable information. This transformation is achieved through the development of dashboards, which act as an interface between the system and users, visually summarizing performance indicators [26]. However, Kruglov et al. [35] argue that dashboard development and associated indicators should be customized to the end user, as one of the primary reasons for the rejection of this tool is the lack of information that users genuinely need.

Işık et al. [36] underscore the potential of BI systems to manage varying volumes of data generated at an increasing velocity due to business activities, combined with ongoing market changes, prompting companies to prioritize their implementation. This prioritization is motivated by their aim to differentiate themselves from competitors through more precise decision-making.

3. Initial Status and Challenges of the Port Under Study

In the port under study, containers are associated with either import or export flows. For exports, the supplier dispatches the container via a hauler, and the terminal operator manages its shipment and dispatch to the destination port. For imports, the container arrives via maritime transport, and the terminal operator coordinates unloading and delivery to a hauler for final customer transport.

The transportation process begins when a hauler receives a request from the exporter or importer. To enter the terminal, both the driver and the truck must be registered and authorized in the system. The hauler must then schedule a 1-h time window for logistical services (delivery, pickup, or both). For example, if a delivery is scheduled from 8:00 AM

to 9:00 AM, the driver must arrive within that timeframe, and the terminal operator must be prepared to execute the service.

Service scheduling can be done in advance using the JUL system or in person at the terminal's kiosk. Advance scheduling is preferred as it automatically validates the driver's entry authorization. In-person scheduling requires manual data entry into the GTOS system, which integrates with JUL.

There are two systems for recording data on container services: JUL, the port's global information system, which centralizes all activities, and GTOS, the terminal's operational system, which supports planning, execution, and control. JUL tracks appointments through various stages: "Pending" (awaiting validation), "Scheduled" (after validation), "Running" (service in progress), and "Realized" (completed). Other statuses include "Not performed" (service not started) and "Canceled" (due to hauler or terminal operator issues or expiry).

As container volumes increase, terminal operations face significant pressure, exposing inefficiencies. The main issue is the disorganization caused by a lack of demand forecasting and the terminal operator's inability to adapt resources effectively. The JUL system's inefficiency, due to low hauler adoption and unexpected cancellations, leads to truck queues and congestion at the port entrance.

Despite limited adherence to the scheduling platform, it remains useful for quantifying scheduled services. The port authority is focused on improving processes to address these inefficiencies and better manage the scheduling and execution of logistics services.

Technological advancements have led to the creation of isolated "data islands", with logistics data fragmented within JUL, resulting in integration challenges, particularly with GTOS. This lack of data consolidation leads to inefficient information flows, which hinders timely access to critical operational insights.

This work aims to develop a Power BI tool that extracts and transforms data from logistics services, providing insights that support decision-making and mitigate current inefficiencies. The BI system will aggregate data from various sources into a single repository, presenting relevant information in a user-friendly, graphical format to improve accessibility and understanding across the organization.

4. Development Methodology

The development of the BI tool has followed the cross-industry standard process for data mining methodology (Figure 1), comprising six main phases. This methodology guarantees a systematic and effective resolution of the challenges identified, although it is not able to fully integrate data from both systems. It is, therefore, a starting point for obtaining agile information to support decision-making.
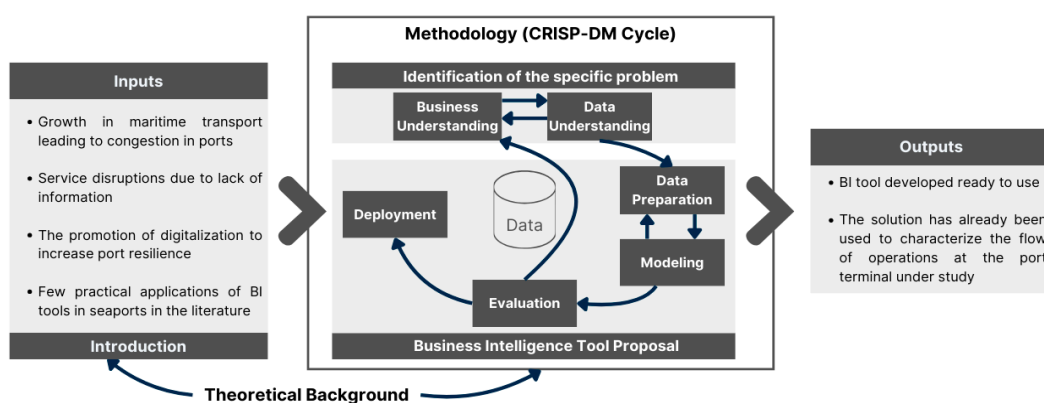

Figure 1. Diagram of the methodology adopted based on the CRISP-DM cycle.

As discussed in the previous chapters, the need for this BI solution arises from the investment in the digitalization of ports as a strategy to address the growth in maritime transport and mitigate the lack of agile information that compromises the effectiveness of port services. It is important to note that there are not many BI applications specifically targeted at this sector. While the CRISP-DM methodology is not typically considered agile, it was chosen for this study due to its structured approach, which provided a clear and systematic framework for addressing the challenges in port operations. The dynamic and complex nature of seaport environments requires a method that can both guide and adapt to evolving needs. Despite its more rigid structure, CRISP-DM was implemented to ensure a coherent process for data collection, transformation, and analysis. This approach enabled the identification and resolution of critical data fragmentation issues, contributing to better-informed decision-making.

During the early stages of the project, several meetings were held to understand how the port operates. The main goal was to identify the current issues and find opportunities to improve data analysis and decision-making procedures.

The second phase focused on identifying the key data sources for analysis. The main sources identified were the scheduling system "JUL", which handles appointments, and the operational system "GTOS", which manages operations at the terminal. A thorough evaluation of the data's quality was conducted to ensure completeness, accuracy, timeliness, consistency, and accessibility. Data integrity checks were performed to ensure consistent formatting, completeness, and the absence of null or duplicate records. Any discrepancies found were noted for future resolution in the source systems and the BI system.

In the third phase, the focus shifted to addressing inconsistencies in the data. This involved data cleansing and the application of data mining techniques such as classification and regression to create additional variables. A composite key was developed to cross-reference the data from the two systems, expanding analytical possibilities. Following this approach, final attributes and records were selected for loading into the BI system's data warehouse. KPIs, which are deemed necessary by stakeholders, were established to provide a comprehensive picture of scheduling and operational performance at the port.

Using the Kimball technique, which employs dimensional modeling to depict relationships between data tables, the fourth phase involved building the data warehouse. Snowflake and constellation schemas were used to organize the data effectively. Subsequently, dashboards featuring key performance indicators were created with a visually appealing and user-friendly interface to simplify data interpretation.

In the fifth phase, a comprehensive SWOT analysis was performed to assess the effectiveness of the proposed solution, with a particular focus on the dashboards, since they are the interface between the final solution and the end users. The evaluation aimed to ascertain the tool's ability to robustly and efficiently analyze the terminal's systems.

The sixth and final phase involved preparing the solution for implementation to analyze the operations of the terminal under study, which is currently under evaluation by the port authority. A final report that included a detailed account of every phase of the development process was also produced. The goal of this documentation is to demonstrate the solution's finished product and facilitate its duplication or environment adaptation, therefore guaranteeing the solution's ongoing improvement.

The process outputs involve delivering a functional business intelligence tool. This technology has been actively employed at the specific port being studied to carry out a sample analysis of the existing logistics operations and to assess its impact.

5. BI Tool: Features and Practical Outcomes

This section outlines the key findings related to the conceptualization and development of the business intelligence tool. The tool comprises three different modules, referred to in this paper as applications: one for data collected from the scheduling system (referred to as JUL), another for data collected from the operator's system coordinating terminal activities (referred to as GTOS), and a third for cross-referencing data from both systems (referred to as INT). Evaluating new parameters, such as the percentage of previously scheduled services and compliance with time windows, was only possible by cross-checking the data in the INT application.

This modular approach was chosen to allow separate analysis of the systems, which are not fully integrated and are operated by different entities. Developing the tool as a single application would compromise its integrity in the event of errors related to cross-sampling data and decrease its query performance.

5.1. Data Preparation

After importing the data from the systems, a quality study was undertaken to identify the non-conformities in the data according to five dimensions: completeness, accuracy, timeliness, consistency and accessibility. The solutions adopted for the non-conformities detected are presented in Table 2 according to their purpose: cleaning; exclusion.

To simplify and guide the interpretation of the analyses, variables were created from the initial attributes of the datasets.

The approach adopted was mainly based on constructing variables from attributes, mainly of the datetime type, using the regression technique and, where appropriate, classifying them into categories.

An example of this approach is the construction of the continuous variable "Appointment Advance" (Table 3), based on the difference between the start date of the time slot and the creation date of the appointment, where each value was classified using the variable "Creation Time". For example, when the "Appointment Advance" has a negative value, then the value shown by the "Creation Time" variable is 'Created during the time window'. This type of approach not only improves understanding of the values, but also makes it easier to compare them with other attributes, enriching the scope of the analysis.

Table 3 shows the variables constructed in the JUL dataset where "Delay End" and "End Type" follow the same approach. In order to simplify attributes that gave similar information, they were merged into one, as is the case with "Profile_Method", where the union of the "Profile" and "Method" attributes was considered relevant because drivers make exclusive use of the mobile application, while the carrier only uses the website, and null values always correspond to the integration. One of the crucial attributes for further analysis, the "Time Window", was created from the respective "Start" and "End" attributes. However, before joining these attributes, and to preserve the date provided by them, the "Service Date" variable was created.

Table 2. Main problems and solutions adopted in data preparation.

Dimension	Problem	Solution Adopted
Completeness	JUL data does not include in-person kiosk bookings, limiting the overall view of service flows.	Complementary analysis using GTOS data to characterize general service flows.
	GTOS data contains limited attributes to detail services and time slots.	Integration of JUL data and creation of a composite key to cross-reference information between the systems.
Accuracy	Problem 1: Incorrect bookings for booking and bill of lading (service type and container movement). (7283 occurrences in ~80,000 records)	Cleansing: Substitution of incorrect values in attributes, adjusting to 'Import' and 'Pick-up' as required.
	Problem 2: Delivery bookings associated with import movements. (1 occurrence in ~80,000 records)	Cleansing: Correction of the "Container Movement" attribute to "Export".
	Problem 3: Bookings marked as "Completed" with cancellation reasons. (706 occurrences in ~80,000 records)	Cleansing: Adjustment of the "Status" attribute to 'Cancelled' for applicable cases.
	Problem 4: Duplicate values in cancellation reasons. (1265 occurrences in ~80,000 records)	Cleansing: Replacement of 'Rejected by Terminal Operator' with 'Cancelled by Terminal Operator'.
	Problem 5: Times in UTC time zone (no adjustment for Portugal's summertime). (All records with time data)	Cleansing: Conversion of all timestamps to Portugal's time zone using a custom function.
Consistency	Problem 6: Missing booking IDs in multiple records.	Exclusion: Removal of these records
	Problem 7: Missing timestamps for booking evolution in 73% of records. (58,650 occurrences in ~80,000 records)	Exclusion: Removal of the booking timestamp attributes
	Problem 8: Missing end-of-time-slot dates in some records. (25 occurrences in ~80,000 records)	Exclusion: Removal of these records
	Problem 9: Delivery and pick-up bookings for the same container with inconsistent statuses. (9 occurrences in ~80,000 records)	Exclusion: Removal of these records
	Problem 10: Inconsistent formats in truck and container license plates.	Cleansing: Removal of non-alphanumeric characters from the respective attributes using a custom function
	Problem 11: Inconsistent formats in transport company names.	Cleansing: Manual mapping to align names between JUL and GTOS; development of an auxiliary function for standardization.
Timeliness	The period of the JUL dataset is between November 2021 and February 2024, and the GTOS dataset is between January 2022 and December 2023, so the data are up to date in time, where the ease of future updates is related to the use of the same SQL query and its accessibility.	
Accessibility	The process of extracting the data was carried out using an SQL query by specialized technicians from the respective systems. The query generated Excel files that served as the starting point for building the proposed BI tool. Therefore, it is not possible to assess the accessibility of the data, and it is important to note that future updates of the tool and the dataset will require the use of the same SQL query or, alternatively, adjustments based on the logic explained throughout the work.	
GTOS (specific)	Problem 12: Outliers in service duration.	Implementation of mechanisms to detect and exclude outliers, while allowing users the option to manually analyze them.

Table 3. Variables created in the JUL dataset.

Variable Name	Description	Type
Appointment Advance	Difference in minutes between the start of the scheduled time window and the creation time.	Double
Creation Time	Classification of Appointment Advance: 'Before Window', 'During Window', 'During Integration', or 'After Window'.	String
Profile_Method	Combination of "Profile" and "Method": 'Web-Transporter', 'App-Driver', 'Integration'.	String
Service Date	Date for which the service was scheduled.	Date
Time Window	Concatenation of the start and end times of the time window in a simplified format, e.g., '16:00–17:00'.	String
Delay End	Difference in minutes between the recorded completion timestamp and the end of the time window.	Double
End Type	Classification of how the scheduling was completed: 'Not Recorded', 'Cancelled', 'Auto Closed', 'Late Close', 'On-Time Close'.	String

Table 4 shows the variables created in the GTOS dataset. Initially, due to the lack of a unique identifier for each service, the "Service ID" was introduced. Since there was no attribute in this dataset that described the "Type of Service" according to the specification presented, this variable was constructed by classifying combinations of the attributes "Loads", "Unloads", and "Type of Journey". For example, a double trip corresponds to a delivery and lifting service, while if it is a single trip and there is a record of an unloaded container, it is a delivery, and the other possible case corresponds to the lifting of a container. To analyze the duration of the services and gauge the performance of the terminal operator in completing them, the variable "Service Duration" was constructed from the difference between the time the truck left and entered the terminal. Considering the lack of context about the time windows marked in this dataset, the variables "Entry Time Window" and "Exit Time Window" were created to group truck entries and exits into one-hour intervals throughout the day, as stipulated in the terminal operator's schedule. Before creating these variables and to preserve the date present in them, the "Service Date" was entered. Ideally, a service should be completed within the same time window in which the truck entered the terminal. To assess this compliance, the "Time Offset" variable was introduced, which represents the difference, in hours, between the truck's departure and entry time. Therefore, this metric quantifies the number of time windows that elapsed between the truck's departure and entry of the truck at the terminal.

Table 4. Variables created in the GTOS dataset.

Variable Name	Description	Type
Service ID	Unique identifier for each service.	Integer
Service Date	Datetime when the truck entered the terminal for the service.	Date
Service Type	Classification of the service: 'Delivery', 'Pickup', or 'Delivery and Pickup'.	String
Service Duration	Duration in minutes from the truck's terminal entry to its exit.	Double
Entry Time Window	1-h time slot during which the truck entered the terminal.	String
Exit Time Window	1-h time slot during which the truck exited the terminal.	String
Time Offset	Difference in hours between the time windows of truck exit and entry.	Integer

Mechanisms for identifying outliers were also added. Based on the interquartile range (IQR) technique, a method was developed to detect univariate outliers, shown in Figure 2 with the example of the variable "Average Service Duration", which corresponds to GTOS problem 12 identified in the data quality survey (Table 2, last row). The IQR method, intrinsic to the box-and-whisker plot, was chosen for this work as it is a simple, robust, and widely used technique from a general practical perspective, with no specific assumptions about the data distribution. Outliers identified by the method are excluded by default, but users have the option to analyze them further if required. However, this method can be applied in an equivalent way to other variables, depending on requirements.

```
1 Duration_Outlier =
2
3 var Q1 = PERCENTILEX.INC(ALL(fServices),fServices[Average Service Duration],.25)
4 var Q3 = PERCENTILEX.EXC(ALL(fServices),fServices[Average Service Duration],.75)
5
6 var InterQuartil_Range = Q3-Q1
7 var Upper_Fence = Q3 + InterQuartil_Range*1.5
8 var Lower_Fence = Q1 - InterQuartil_Range*1.5
9
10 return
11 SWITCH(TRUE(),
12     SELECTEDVALUE(fServices[Average Service Duration])>Upper_Fence, "Upper Outlier",
13     SELECTEDVALUE(fServices[Average Service Duration])<Lower_Fence, "Lower Outlier", "No")
```

Figure 2. Outlier identification method.

Table 5 summarizes the evolution of the attributes during data processing. Data preparation resulted in the exclusion of 5663 records in JUL and the standardization of the attributes in both datasets. In addition, seven variables were created in JUL and seven in GTOS, which will play a crucial role in the analyses. Finally, 16 attributes in JUL and 3 in GTOS that did not add value were excluded. Thus, 19 attributes were selected for the final JUL model and 11 attributes for the final GTOS model.

Table 5. Evolution of the number of attributes and records throughout data processing.

System	No. Initial Attributes	No. Initial Records	No. Final Attributes	No. Final Records
JUL	27	100.348	19	94.685
GTOS	7	111.984	11	111.984

5.2. Requirements and Indicators for the BI Tool

The requirements were raised with the stakeholders, considering the context of the seaport under study and the data available for collection from existing systems. The iterative process to define the final key performance indicators (KPIs) involved multiple stages. Initially, a comprehensive set of potential KPIs was identified, based on the raw data extracted from the operational systems. This preliminary list was evaluated in a series of meetings with stakeholders, where the relevance and applicability of each KPI were discussed in detail.

These discussions revealed that certain indicators were less critical or redundant and were subsequently excluded. Conversely, priority was given to the indicators that directly addressed identified monitoring needs, such as operational adjustments, resource optimization, and strategic decision-making. Ultimately, 14 KPIs were selected from the initial set, as they most effectively reflected the specific requirements of each user profile.

The relevant user profiles have been identified ("terminal operators", "port administrations", "haulers and drivers") with dedicated dashboards. Dashboards were implemented for haulers and their drivers to display the logistics operator's capacity availability for various time windows and days of the week. Terminal operators needed dashboards to showcase the most requested containers, periods of highest demand, and the haulers with the highest number of appointments, to adjust their resources and prevent service disruptions. For port administrations, the dashboards were required to present summary performance indicators such as the percentage of prior appointments and the time in advance they were conducted. In the latter case, the analysis provided by the dashboards is aimed to support the implementation of strategic port policies (for example, the implementation of mandatory prior scheduling).

A consistent requirement across all user profiles was the need for the tool to be user-friendly, with information presented clearly and concisely using appropriate graphs and diagrams. The selected indicators were also designed to be analyzed from three perspectives: daily trends, overall performance, and contribution to the total, offering flexibility in monitoring operations.

Out of 30 final attributes selected in the quality treatment (the sum of the fourth column values, Table 5), 14 KPIs were established and implemented using Power BI's data analysis expressions (DAX) language (Table 6). The indicators can be filtered by context, enabling diversified analysis without modifying their structure.

Table 6. KPIs defined and their description.

KPI	Description	Application
No. Appointments	Unique count of prior appointments in the JUL application.	
% Cancellation	Percentage of canceled prior appointments, calculated by dividing the number of canceled prior appointments by the total number of prior appointments.	
Average Appointment Advance	Average time (minutes) between appointment scheduling time and the start of the scheduled time window.	JUL
Containers per Appointment	Total number of containers handled by each prior appointment.	
No. Services	Unique count of services in the GTOS application.	
Average Duration of Service	Average time (minutes) taken per service.	
% Entry and Exit Same Time Window	Ratio between the number of trucks entering and leaving in the same time window and the total number of trucks that entered in the same time window.	
Availability Time Window	Difference between the operator's capacity and the number of services per time window divided by the total number of services.	GTOS
Containers Delivered and Picked Up	Total number of containers handled per service.	
% Scheduled Services	Percentage of services scheduled in advance obtained by dividing the number of previously scheduled services by the total number of services.	
No. Scheduled Services	Unique count of scheduled services in the INT application.	
% Compliance	Percentage of services completed within the scheduled time window obtained through the ratio between the number of services that comply the time window and the total number of services.	INT
Average Entry Delay	Average time (minutes) between the start of the time window and the time the truck entered the terminal	
Average Exit Delay	Average time (minutes) between the end of the time window and the time the truck exited the terminal	

The "No. Appointments" is the KPI for assessing the JUL's performance. By quantifying previous appointments, it is possible to monitor variations and identify trends that allow for adjustments in strategy during peak demand periods. Additionally, knowing the number of prior appointments is essential for measuring the adoption of this system by haulers and their drivers, which is crucial for quantifying the current problem.

The "% Cancellation" KPI is a measure of the proportion of previous appointments that have been canceled compared to the total number of prior appointments. This metric is significant as it highlights the effectiveness and reliability of JUL. A high cancellation rate can indicate various issues. Difficult-to-use systems might lead to scheduling errors, causing terminal operators to cancel. Additionally, a lack of flexibility could force haulers to cancel. Poorly adjusted operational capacity might also prevent terminal operators from meeting demand, leading to dissatisfaction among haulers. Therefore, maintaining a low cancellation rate is essential for ensuring smooth operations and high hauler satisfaction.

The "Average Appointment Advance" KPI gives the average time between the creation of the appointment and the start time window of the service. Monitoring these data is essential to avoid problems with operational capacity and flexibility. A shorter average time in advance may signal imminent service overload, leading to challenging resource management and delays. Conversely, a longer average time in advance enables more efficient operational planning, driving productivity gains.

The "Containers per Appointment" KPI shows the number of containers handled in each previous appointment, which helps in making more accurate adjustments in the allocation of resources. Also, the terminal operator's income depends on the movement of containers, so understanding the number of containers moved is crucial for managing operating costs and income effectively.

The "No. of Services" is a metric that captures the total count of services registered in GTOS. Unlike the "No. of Appointments", which records previously scheduled appointments, this indicator quantifies the actual provision of services at the port. It plays a crucial role in identifying potential expansion opportunities or the need to adjust service types. By analyzing this metric, terminal operators can identify areas with a growing demand for specific services, enabling them to make informed investment decisions to better meet customer needs and gain a larger market share.

The "Service Duration" KPI refers to the average time spent by the terminal operator on each service. This measure is valuable for identifying variations in the level of service provided to different haulers and/or at different times. Analyzing these data make it possible to plan the number of services that can be carried out within a specific timeframe, ensuring optimal and effective capacity utilization.

The "%Entry and Exit Same Time Window" KPI empowers terminal operators to evaluate the efficiency of services using only the data from its system (GTOS), and it is calculated as the ratio between the number of trucks entering and leaving in the same time window and the total number of trucks that entered in the same time window. A high value for this indicator implies effective coordination across different stages of the logistics process, which is essential for ensuring seamless operations. On the other hand, a low value may signal planning issues that result in delays, congestion, and dissatisfaction among haulers.

The "Availability Time Window" KPI assesses the unused capacity in each period through the difference between the operator's capacity and the number of services per time window divided by the total number of services. This indicator is essential for improving resource utilization and planning. By identifying periods of idle capacity, the terminal can reallocate resources, adjust the scheduling of operations to fill these periods, and reduce

costs associated with underutilization. At the same time, better management of available capacity makes it possible to increase operational efficiency.

The "Containers Delivered and Picked Up" KPI counts all the containers handled at the port. This metric provides a comprehensive count of the containers handled, unlike the "Containers per Appointment" indicator, which only counts containers from previous appointments. This helps in estimating the total operating costs and revenues related to container handling, which is crucial for ensuring the profitability of the business.

The "% Scheduled Services" is a KPI calculated by dividing the "No. of Appointments" by the "No. of Services", reflecting the percentage of services scheduled in advance at JUL. The optimal target for this metric is close to 100 percent. Lower values may indicate a lack of prior scheduling, while higher values could suggest missed appointments. Analyzing this metric provides insight into hauler adoption of the scheduling system and helps evaluate excessive demand or underutilization of resources.

The "No. of Scheduled Services" quantifies the total number of records in the third application (INT). Like the "No. of Services" and "No. of Appointments", this indicator is vital for assessing other indicators within this application.

The "%Compliance" KPI quantifies the percentage of services that are fulfilled within the scheduled time window, and is essential for assessing the haulers' commitment to comply with the agreed time window and the terminal operator's ability to provide the service as planned. High values for this indicator show that the services are fulfilled by those involved within the established schedule, improving operational efficiency and effectiveness. Therefore, this metric measures the punctuality of services, driving both internal efficiency and hauler satisfaction, which is essential for the growth and sustainability of the business.

The "Average Entry Delay" refers to the time difference between when the truck enters the terminal and the start time of the time window. The "Average Exit Delay" refers to the time difference between when the truck leaves the terminal and the end time of the time window. Monitoring and reducing these delays contribute to hauler satisfaction by providing a more predictable and reliable service.

The establishment of these requirements and the subsequent processing of the data turned raw data into actionable information, as the indicators established make it possible to measure the operational performance of seaports.

5.3. The BI Tool from the Perspective of Data Architecture and Technology

Regarding the architecture of the data model, the Kimball method was used due to its suitability for the context and requirements identified. This method was the most appropriate considering the data available, the lower complexity of the model, and the shorter development time. Furthermore, this design offers a straightforward solution with quick response times, increasing the tool's usability for end users.

In the applications for the separate systems (Figure 3), the data model was constructed in a uniform manner. It is important to note that only the date dimension ('dCalendar') was normalized, as an analysis of the fact tables ('fAppointments', left of Figure 3 'fServices', right of Figure 3) revealed that other dimensions would typically have just one or two attributes. This decision was based on prioritizing the performance of Power BI, which can effectively handle simpler data models, thus avoiding unnecessary complexity from adding dimensions with minimal attributes. The emphasis was placed on practicality and efficiency in data analysis, maintaining a straightforward and easily interpretable model.
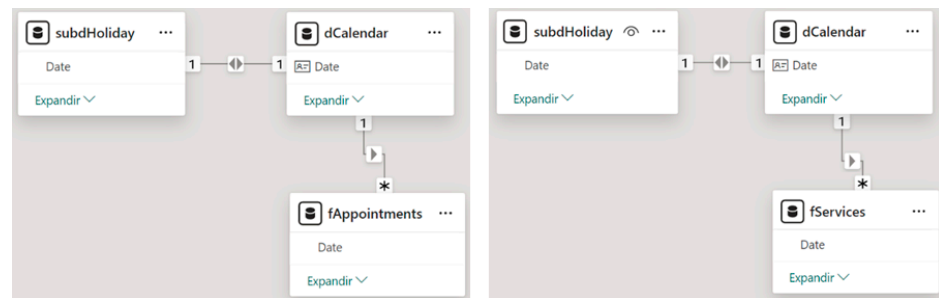
Figure 3. Data of the JUL application (**left**) and GTOS application (**right**).

In both models (Figure 3), the tables are linked using the "Date" attribute, which serves as the primary key of the 'dCalendar' table. This means that an appointment or service occurs exclusively on a date, which may correspond to a holiday at most, indicating a snowflake schema. The fact that the 'subdHoliday' table is a subdimension of the 'dCalendar' table further confirms this schema.

The third application (INT) included the 'dWindow', 'dTypeService', and 'dRoad-Hauler' dimensions. Each of these dimensions has a single attribute that acts as a primary key: "Window", "Type_Service", and "Name_RH" (Figure 4). Despite the initial assumption of avoiding the creation of dimensions with only one attribute, in this specific case, this approach was followed given the presence of multiple fact tables and the need to filter them simultaneously ('fAppointments' and 'fServices') to evaluate indicators such as "% Scheduled Services".
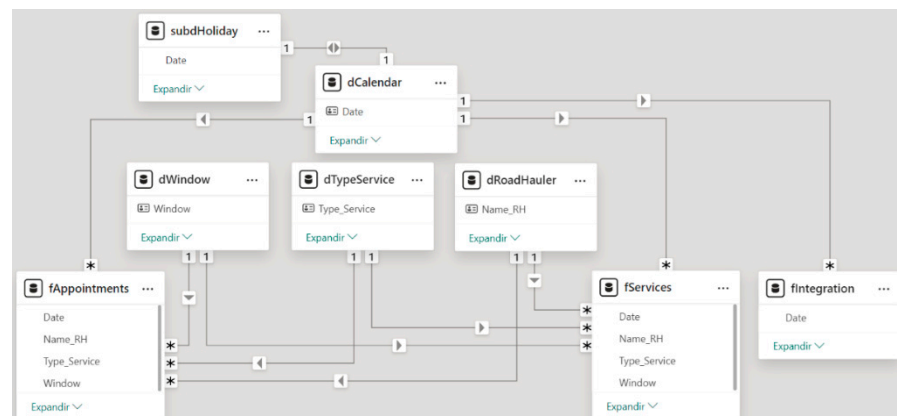
Figure 4. Data model of the INT application.

Thus, there are three fact tables in the model, each filtered by different dimensions, resulting in a constellation schema (Figure 4). The appointment ('fAppointments') and service ('fServices') are each associated with only one type of service, hauler, time window, and date. Additionally, the 'fIntegration' table contains records of where appointments from 'fAppointments' and services from 'fServices' were matched. This table is filtered only by "Date", following the approach used in the previous models (Figure 3).

This modeling approach ensured effective and efficient data analysis, while maintaining the simplicity and robustness of Power BI's performance.

5.4. The BI Tool from the Perspective of Interaction with Users

The dashboards serve as the core deliverable of the BI tool and act as the primary interface for engaging with end users. Customized dashboards were developed for each application to meet the specific needs of the identified user profiles. Additionally, each

application features a starting menu designed to enable smooth navigation through the range of dashboards, as illustrated in Figure 5.

Figure 5. Menu of the three applications (JUL top left, GTOS top right and INT bottom).

In terms of structure and presentation, there are shared views and elements among the different applications, making it easier to use the tool across various systems for parallel use and comparison of appointment volumes and services.

The applications share the following three views (Figure 5). The categories view is helpful for terminal operator analysis because it allows comparison of the system's primary indicators across all categories. The time view supports the terminal operator in spotting patterns and foreseeing potential issues by making it simpler to comprehend how indications change over time. The summary view (available only in JUL and GTOS analytics) provides a high-level overview that enables the port authority to efficiently assess the terminal's performance by summarizing the KPIs of the system.

Additionally, there are specific dashboards in each system, adapted to their particularities: in JUL, there is a dashboard that covers only cancellations; in GTOS, there is a dashboard that presents the availability of time windows; and in INT, a dashboard that assesses compliance with the time window.

Several dashboards were developed, each prominently displaying the identification of the source application and the specific object being analyzed, providing a clear and organized interface for users to make comparisons and evaluations. The five most relevant dashboards, out of a total of ten developed, will be presented next.

The "Categories View" dashboard aims to provide a detailed and versatile analysis of appointment data across different categories. This dashboard is illustrated when used in JUL analytics (Figure 6), and displays two indicators based on the selections made in the left menus. The top graph shows the "Average Appointment Advance", based on the category chosen in the top left selection menu (this top graph also provides a color gradient representation of the Number of Appointments). The bottom graph shows the "Number of Appointments", again based on the category chosen in the top left selection menu and on the selections made in the other two menus. The bottom graph aims to provide a comprehensive view of the distribution of the number of appointments across different values of the chosen category. These data can be further analyzed daily, as a total, or as a percentage of the total, giving users versatile options for assessing and understanding the appointment data.
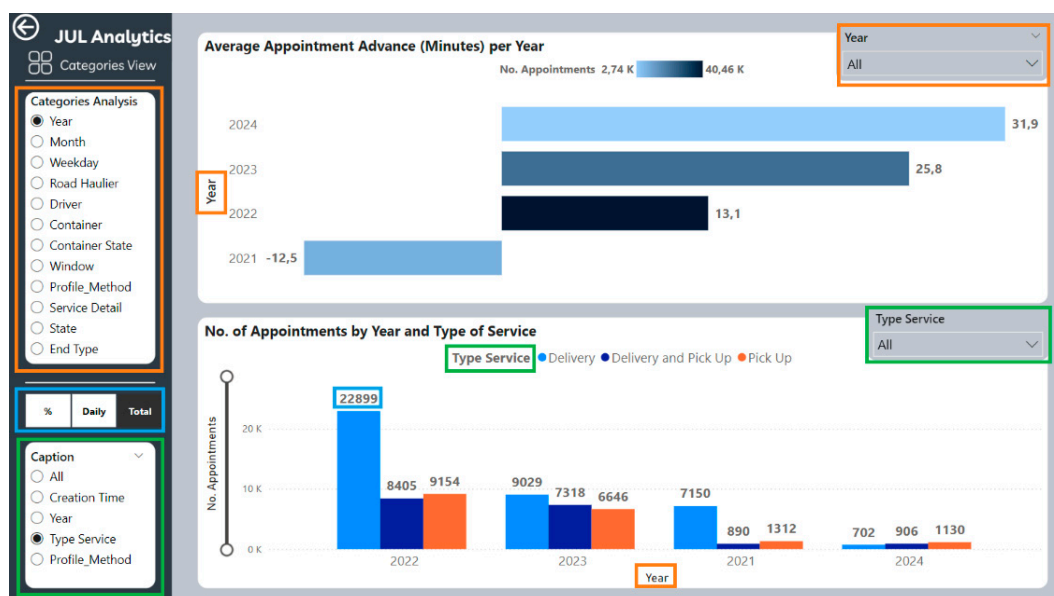
Figure 6. Categories view dashboard—JUL analytics.

This dashboard is important as it allows the profile of the terminal operator to explore the appointment data across various categories, uncovering unique trends and insights specific to each category. By doing so, it helps in making more informed and targeted decision-making processes, enabling users to make data-driven choices tailored to the specific trends and characteristics identified within each category.

In the example of Figure 6, the analysis category is the year (marked in orange). The upper graph shows that the time in advance of appointments has increased over the years, but 31.9 min is still a low value, leaving limited flexibility for the terminal operator. The year with the highest number of appointments is 2022, indicated by the darkest bar in the upper graph and shown in the lower graph. In the caption of the lower graph, the category is the type of service (marked in green), and the indicator is on a total basis (marked in blue). Delivery is the type of service with the highest number of appointments, although it dropped significantly from 2022 to 2023.

The only differences between this type of dashboard and the other applications are the indicators analyzed. In the GTOS application, there are the "Average Duration of Service", "% Entry and Exit Same Time Window", and "No. Services". In the INT application, there are the "% Scheduled Services", "No. Appointments", and "No. Services".

The "Summary View" dashboard aims to offer a comprehensive overview of the main performance indicators for the system being analyzed. This dashboard, when viewed in GTOS analytics (Figure 7), provides information on the daily number of services (outlined in blue). These services are categorized by type of service and grouped into containers delivered and picked up per day, allowing for a deep understanding of the various service activities. This level of detail is crucial for stakeholders, especially the port authority profile, as it enables them to closely monitor the system's performance and make well-informed decisions based on the most accurate and up-to-date data available.

Based on Figure 7's example, there were, on average, 204 services provided each day, with delivery and pick-up services accounting for the majority (109), delivery-only services coming in second with 51 services per day, and pick-up-only services coming in last with 44 services. This figure shows that 164 containers were delivered and 174 were collected daily on average. Furthermore, just 31% of the trucks departed within the same time window that they arrived, and each service took an average of 58 min.
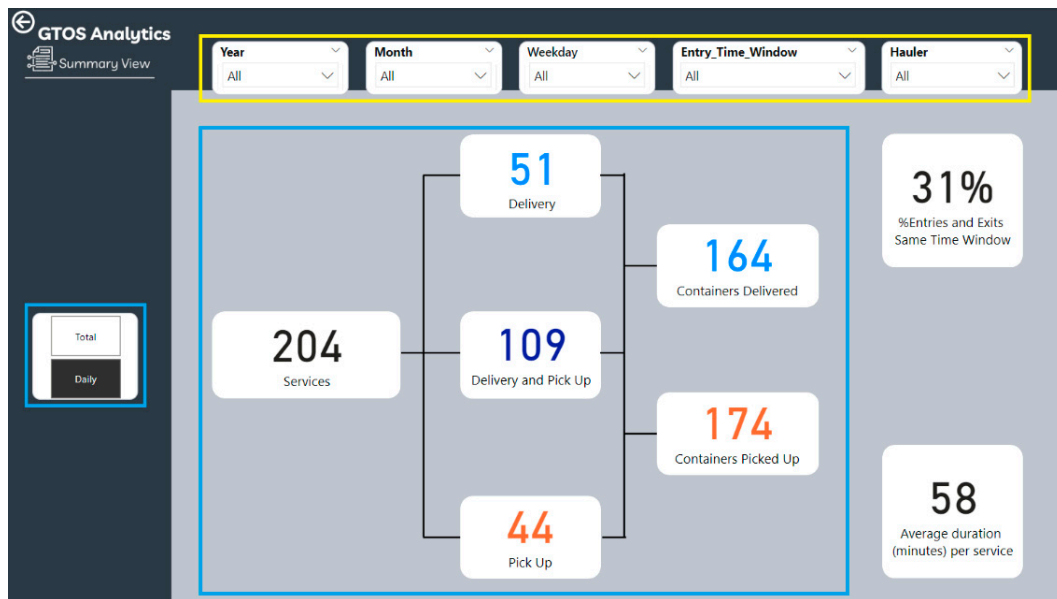
Figure 7. Summary view dashboard—GTOS analytics.

In the JUL application, this dashboard shows the "No. Appointments", " Containers per Appointment", "% Cancellation", and "Average Appointment Advance".

The "Availability" dashboard (Figure 8) seeks to give terminal operators and haulers a clear and actionable picture of time window availability throughout the week. Using color coding, the dashboard effectively shows terminal utilization. Red indicates higher utilization, while green denotes lower utilization. This intuitive design offers valuable insights, and supports terminal operators in efficiently managing resource allocation. Moreover, the dashboard helps haulers make well-informed decisions about their movements, reducing unnecessary congestion at the terminal during peak times. By utilizing these data, the terminal operators and haulers profiles can optimize their processes and improve overall operational efficiency.
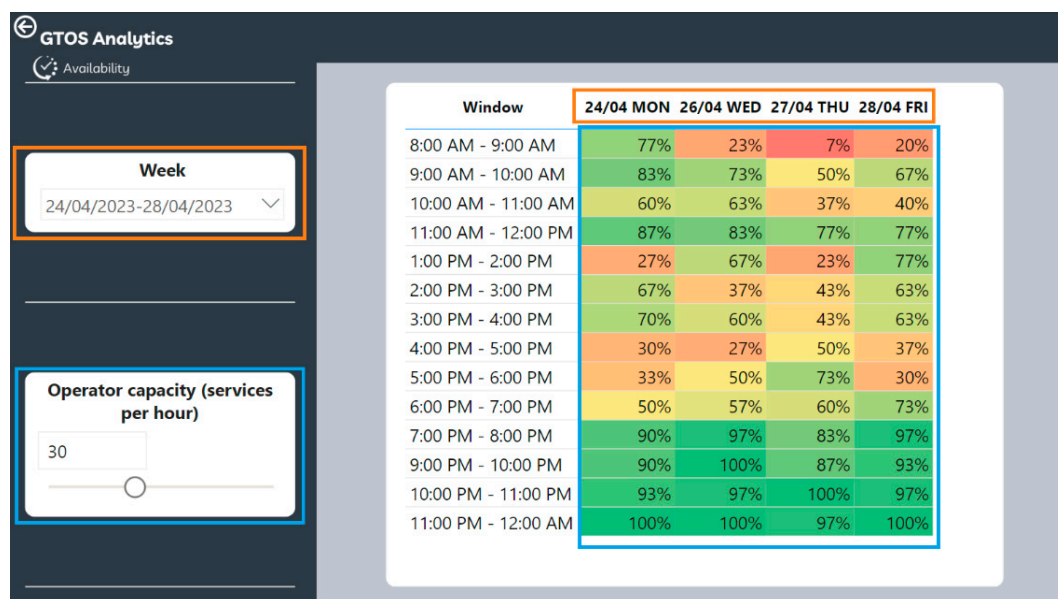
Figure 8. Availability dashboard—GTOS analytics.

In the example presented in Figure 8, the response capacity of the operator stands at 30 services per hour (displayed in blue). In this example, it becomes clear that the

time windows of highest demand during the analyzed week (highlighted in orange) were 8:00 AM–9:00 AM and 4:00 PM–5:00 PM (it is worth noting that the developed tool excludes public holidays from this analysis).

The "Time View" dashboard aims to provide a wide-ranging analysis of how system indicators change over time. It helps to identify patterns, trends, and seasonality in the data, enabling better insight into the temporal evolution of these indicators. The one present in INT analytics (Figure 9) boasts two primary graphs designed to provide an in-depth understanding of the data. The upper graph shows the indicator selected in the middle menu over the period chosen in the left menu, allowing users to gain valuable insights into performance patterns over time. Meanwhile, the lower graph offers a comparative view of the indicator's variation, allowing for the detailed observation of changes on a monthly or yearly basis. This comparison is crucial for the terminal operator's profile, as it enables the identification of seasonal patterns, detection of anomalies, and evaluation of the effectiveness of operational strategies across various time frames. Armed with these insights, this profile can proactively plan for future demands and challenges, as well as craft effective responses to the prevailing trends in the data.
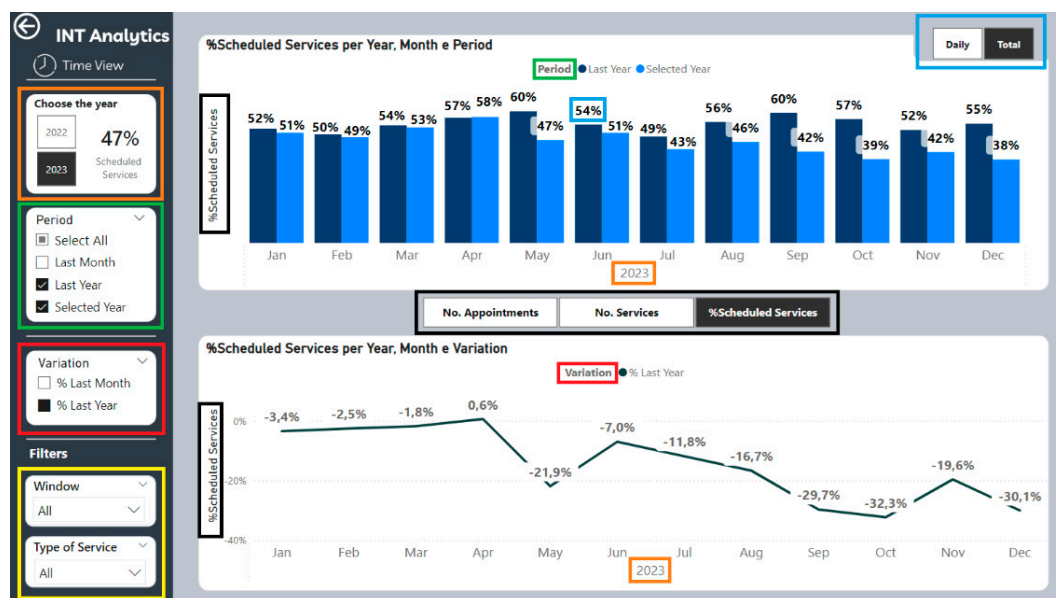

Figure 9. Time view dashboard—INT analytics.

In the example shown in Figure 9, the year 2023 was selected (highlighted in orange), and it was immediately apparent that the annual "%Scheduled Services" was 47%, after which the indicator "%Scheduled Services" was chosen (highlighted in black). Recall that the "%Scheduled Services" KPI indicates the percentage of services scheduled in advance at JUL. In the bar graph, the captions showing this indicator were selected, both for the year picked (2023, in light blue) and for the previous year (2022, in dark blue), and in the line graph, the variation compared to the previous year was selected (highlighted in red).

In this example, the upper graph shows that, in all months of 2023, the proportion of appointments dropped, except in April, when there was a slight increase (0.6%) compared to the same month in 2022. The graph below corroborates this trend by quantifying the decreases, which were significant in September, October, and December, reaching reductions of around 30%.

In the JUL application, this dashboard only presents the "No. of appointments", although there are more filter categories.

The purpose of the "Compliance" dashboard is to analyze the terminal operator's capacity to carry out services according to the schedule and to conduct a study of how effectively haulers and their drivers complied with the scheduled time windows when they had previously scheduled them. It is a sophisticated dashboard, available only in the INT application, which extracts data from various systems and allows the terminal operator profile to carry out an exhaustive analysis of activities. The dashboard consists of four graph components, each offering a unique and detailed analytical perspective. Users can choose specific indicators and categories to conduct a comprehensive examination. The graph in the top left visually displays the values of the time indicator across different elements within the selected category on the left menu. Meanwhile, the bottom left graph provides the "%Compliance" for these category elements, giving a detailed understanding of adherence levels. Additionally, the top right graph enables exploration of the relationship between category values and the number of scheduled services. Lastly, the bottom right graph visually depicts the distribution of services based on time window compliance, along with interactive filtering options that work seamlessly with selections made in the other graphs, allowing for a comprehensive and detailed analysis. This dashboard showcases the successful integration of various datasets, enabling innovative analysis of terminal data. While creating the "fIntegration" table was complex and had limitations (Figure 4), it was crucial for measuring previously inaccessible indicators, such as "%Compliance".

In Figure 10, the chosen category represents the type of service (highlighted in orange). It's evident that the delivery and pick-up service, despite being the most common (39.2%), shows the least compliance with the scheduled time window (11%), mainly due to departure delays averaging 26 min. Conversely, the booking pick-up service has the lowest compliance rate (1%) because, on average, drivers arrived at the terminal approximately two hours before the scheduled time window, completely missing the allocated period.
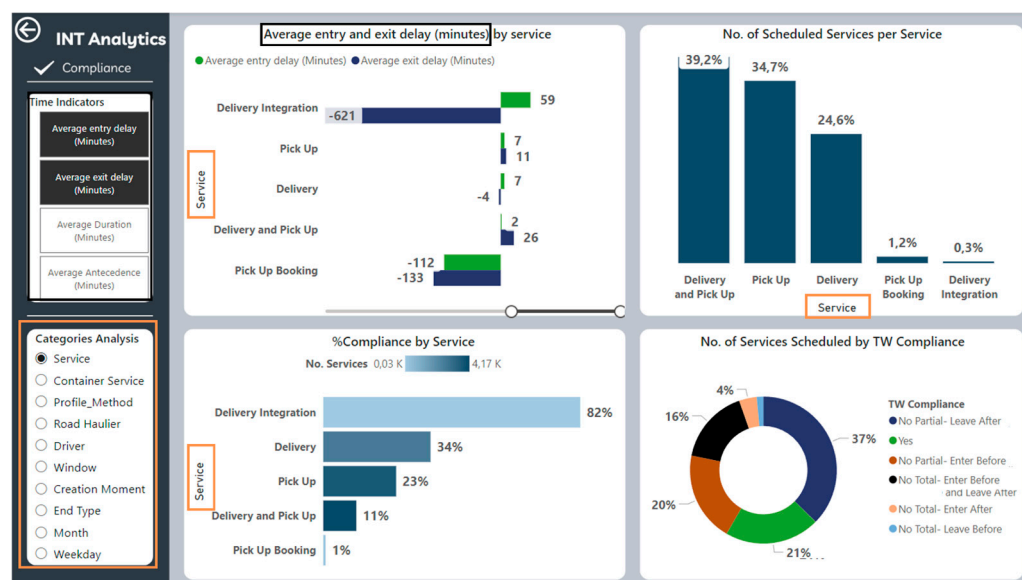

Figure 10. Compliance dashboard—INT analytics.

Overall, only 21% of the bookings adhered to their scheduled time window, indicating a clear need for improved punctuality and efficiency in fulfilling the appointment.

The presented dashboards showcase the tool's capacity to consolidate defined indicators in a straightforward and visually appealing format. This enables stakeholders to make better-informed decisions, encourages proactive analysis, and facilitates an agile response to emerging challenges and opportunities.

5.5. Analytical Insights from the Tool

The analysis conducted through the developed tool provides valuable insights into the logistics operations of the studied terminal.

Firstly, there is a gradual increase in the advance scheduling time over the years, although the average remains relatively low. Additionally, many appointments are created within the time window, making proactive planning by the terminal operator challenging.

When examining road hauler behavior, a significant observation is that several high-performing road haulers do not utilize prior scheduling, emphasizing the need to encourage the broader adoption of the application. Simplified scheduling, widely used in the early years of the analyzed data, has become less common over time. Drivers now predominantly create appointments via the mobile application, while road hauler managers demonstrate higher levels of responsibility, particularly in scheduling services with greater advance time and managing cancellations.

For the haulers in which the driver makes most of the appointments, they perform worse in terms of scheduling in advance, but comply with the time window more often. In this respect, the third hauler with the most appointments stands out, as it constantly scheduled the service after the start of the time window, and fulfilled 24% of the scheduled services. On the other hand, in the tenth carrier with the most services, in which it is the manager who coordinates the bookings, he created them, on average, two and a half hours before the opening of the time window, and only 7% of the services were completed within the time window.

Therefore, it can be concluded that scheduling the service further in advance does not ensure that the scheduled time window is met, because drivers prioritize arriving at the port, regardless of the scheduling. As a result, the duration of the services of the haulers who schedule in advance is compromised.

Between 2022 and 2023, the proportion of scheduled services decreased. The delivery service type remains the most frequently scheduled, but its significance has been declining since 2023, coinciding with an increase in the scheduling of combined delivery and pickup services. Notably, the scheduling trends do not reflect the actual demand proportions, as combined delivery and pickup services are the most sought after, the longest in duration, and the least likely to meet scheduled time windows. Encouraging advance scheduling for these critical services is essential for optimizing terminal operations.

Time window analysis reveals that the busiest windows are those with the shortest average appointment advance times. This creates capacity management challenges for the terminal and leads to an increased number of cancellations, often due to the expiration of validation dates or scheduling by drivers. The afternoon period (2:00 PM–6:00 PM) is the most in-demand, with peak truck arrivals observed between 4:00 PM and 5:00 PM. This specific time window has the lowest appointment advance time and the lowest scheduling proportion, with only a fraction of services fulfilled as planned.

Furthermore, services performed during this peak period experience the longest durations, contributing to delayed truck exits and further exacerbating operational pressures in subsequent time windows. Only after 6:00 PM does the percentage of trucks completing services within the same time window stabilize at the average level, highlighting the ongoing operational strain during peak hours.

In summary, the historical data analyzed underscores the need for more efficient scheduling and service management practices. Implementing these improvements will better align terminal operations with customer demands, enhance overall efficiency, and reduce operational bottlenecks during peak periods.

6. Evaluation: A SWOT-Analysis Overview

The business intelligence tool garnered positive feedback from key stakeholders, who validated the construction process and results, expressing satisfaction with the tool's capacity to quantify and analyze previously unmeasured aspects. However, to ensure a comprehensive assessment of the tool's robustness and quality, it was essential to consider several other factors.

As a result, additional evaluation criteria were considered to ensure a comprehensive assessment of the tool. This included usability, which examines the user experience and ease of use; functionality, which pertains to the range of analyses, information, and visualizations provided by the dashboards; flexibility, which evaluates the tool's capacity to adapt to future requirements; and effectiveness, which assesses whether the tool delivers valuable insights for decision-making.

Although feedback was collected from users during the testing phase, it is important to note that the tool was still in its early stages, and not all suggestions could be fully incorporated. The feedback provided valuable insights into potential improvements, but due to the developmental phase, only some of the identified issues were addressed in the current version. These insights will be considered for future updates of the tool.

To methodically structure this assessment, a SWOT analysis was conducted (see Table 7). This analysis facilitated the identification of the strengths and weaknesses that influence the tool's success, both internally and externally.

Table 7. SWOT analysis of the developed BI system.

	Positive Factors	Negative Factors
	Strengths	**Weaknesses**
Internal Factors	- The system measures new indicators (%Scheduled Services, %Completion) by integrating data from various sources and levels of detail into a unified dashboard. - It presents data in visually appealing graphics that enable quick interpretation. - It offers multiple functions, such as time analysis and categorization, and enables users to compare different categories in the graphs. - The structure and data model enable the creation of additional dashboards as needed. - Low effort and time required to perform future analyses, since the process is structured.	- Using the interface requires some training and understanding from the user, especially in the dashboards that allow dynamic analysis of various categories (Categories View). - The need for specific knowledge to configure the tool and deal with unforeseen events. - Although the mechanism for cross-referencing systems has achieved a representative and crucial sample given the circumstances, this is still limited, and it is necessary to obtain a total correspondence between appointments and services, which requires further integration of the systems involved.
	Opportunities	**Threats**
External Factors	- Develop the integration process between the scheduling system and the operating system, capitalizing on the advantages identified by the solution built. - Make the decision-making process more data-orientated using the tool. - Expanding the use of this type of solution to other port terminals.	- Data quality, which can undermine the credibility of the tool's analysis. - Resistance to adoption by users due to ignorance of the tool's benefits. - Power BI's ability to handle large data volumes may become a limitation. Increased data complexity and volume could require additional infrastructure to maintain performance and efficiency.

The SWOT analysis of the solution provides a comprehensive understanding of its transformative impact. The tool's strength lies in its ability to seamlessly integrate data from different sources and present it in an aesthetically appealing format, thereby facilitating

more comprehensive and detailed analyses. The transition from a manual process to an automated system has significantly reduced preparation time and greatly enhanced data quality. The flexibility to generate additional dashboards and perform dynamic analyses represents a strategic advantage. However, challenges such as the need for training and improved system integration must be overcome to fully maximize the tool's impact. Feedback from real users indicated that, while the interface is generally well-received, some users encounter difficulties with the complexity of dynamic analyses, suggesting a need for further training. Additionally, although integration with existing systems has yielded promising results, it remains incomplete, limiting the tool's full potential. To enhance broader adoption and improve efficiency, addressing these challenges through targeted training programs and system improvements will be essential.

To better understand this impact, it is relevant to compare the situation before and after the implementation of the BI solution. Before its introduction, data analysis was mainly manual and involved the laborious extraction and processing of information using Excel. This approach not only consumed significant time but also proved inefficient, leading to prolonged task completion times and shallow analyses due to the lack of system integration.

The implementation of the solution has brought a significant transformation. The automation and structuring of processes related to data extraction, processing, and visualization have led to a significant improvement in efficiency. The new tool has equipped users with interactive dashboards that provide a holistic perspective on indicators based on integrated and historical data, enabling real-time analyses. This transformation has not only reduced the time required for analysis preparation, but has also facilitated a shift towards more strategic and value-added activities. Consequently, the solution has not only addressed previous constraints, but has also unveiled new opportunities to address challenges and capitalize on emerging prospects.

The impact of the tool goes beyond operational efficiency. By identifying issues related to entry traffic, such as congestion, the solution enables the implementation of strategies that may include limiting the number of appointments per time window, requiring a 100% appointment rate, and optimizing truck arrival schedules. These measures contribute to reducing congestion and, consequently, the emissions of greenhouse gases. Additionally, the tool enables real-time analysis of operational efficiency, allowing the port to make data-driven decisions to minimize energy consumption and logistical waste.

Looking ahead, the tool could be enhanced with the inclusion of a module dedicated to environmental indicators. This expansion could provide deeper insights into energy intensity, carbon emissions, and the overall ecological footprint of the port's activities. However, integrating such indicators requires the development of real-time data collection systems that can track environmental parameters, such as truck energy consumption and pollutant emissions, during port operations. This future module would not only support sustainability, but also enable the port to meet the growing global demand for more transparent environmental practices.

7. Conclusions

The efficient collection and processing of data is crucial for success in any business. In today's fast-paced world, having access to high-quality information quickly is essential for business growth and continuity. BI tools address this need by consolidating and sharing information in a way that is accessible and tailored to different users' needs.

This article focused on the challenge of dispersing information within information systems, which impacts operations and causes congestion at a port terminal. Despite the extensive literature on the BI area, specific tools for the port sector are still under-researched. The main goal of this work was to develop a BI tool supporting the analysis of logistics

flows. Its application in a Portuguese seaport highlighted the value of digitalization and BI in enabling structured and rapid analyses, reducing information gaps, and facilitating data-driven decision-making.

Although the developed BI application met the needs of one of Portugal's largest container ports, the methodology applied and the solution presented offer a clear and robust pathway for adapting the CRISP-DM methodology to the operational context of different ports. The flexibility of CRISP-DM makes it inherently adaptable to various port environments, where the operational models and data sources may differ. For example, in a port with high cargo traffic and complex logistics, the data preparation phase would need to handle large, real-time datasets from multiple systems, such as container tracking, vehicle scheduling, and environmental monitoring. In contrast, smaller ports with simpler operations may focus more on specific KPIs related to cargo throughput and schedule adherence. This customization within the data preparation stage ensures that CRISP-DM can handle both high-volume and smaller-scale data with equal efficiency.

In the data modeling phase, for instance, the methodology can be adjusted to focus on different types of predictive models based on the port's operational goals. Larger ports may require predictive analytics for congestion forecasting or optimization of berth allocations, whereas smaller ports may focus on simpler dashboards that monitor operational efficiency in real time. Additionally, the deployment phase of the solution can be tailored to incorporate diverse port-specific software platforms (e.g., scheduling systems, sensor data for environmental monitoring, etc.) and integrate them into a unified dashboard to provide real-time insights. This adaptability ensures that the system can be seamlessly integrated into various IT infrastructures.

The evaluation phase of CRISP-DM also plays a critical role in ensuring that the solution remains relevant in diverse contexts. By continuously evaluating the impact of the BI tool in different port environments, the methodology allows for the continuous refinement and optimization of the solution. For example, feedback from port staff in a specific location could lead to adjustments in the way data are visualized, enhancing user engagement and decision-making. Thus, the CRISP-DM methodology provides a dynamic framework that can be fine-tuned according to each port's specific needs and operational complexities, ensuring broad applicability across various contexts.

As a theoretical contribution, this study fills a significant gap in the literature regarding BI tools for supporting port operations. While BI methodologies have been widely applied in other logistics sectors, this work represents a pioneering adaptation of CRISP-DM for maritime ports. The methodology is extended to address challenges specific to the port environment, such as data integration from diverse, often siloed systems, and the need for real-time analysis to support decision-making. Moreover, the flexibility of CRISP-DM to adapt to different operational scales and contexts extends its applicability not just to large ports, but to smaller and mid-sized terminals as well.

Developing the solution presented several challenges with limitations. Data collection and processing were challenging due to non-conformities, which were addressed through automated processes. However, the tool still faces challenges in autonomously handling unforeseen non-conformities. Additionally, the absence of certain attributes in the systems analyzed limited the sample, and the dashboard structure prioritized comprehensive analyses over usability, potentially impacting user adherence and necessitating additional training. Feedback from real users highlighted those certain aspects, such as the complexity of dynamic analyses, may hinder user adoption without adequate training. Additionally, as the solution scales, Power BI's ability to handle larger datasets may necessitate additional infrastructure to maintain optimal performance and efficiency.

Future work should take a proactive approach to continuous data analysis to ensure the BI tool remains relevant and efficient. Simplifying dashboard analysis categories can enhance the user experience, while the development of new indicators and dashboards must keep pace with operational evolution. The integration of real-time data can provide more accurate and up-to-date information. Furthermore, extending the solution to other ports and logistics sectors could broaden its impact and utility, potentially serving as a model for other port terminals. The methodology presented here offers a clear roadmap for adapting BI solutions to diverse port contexts, showcasing CRISP-DM's flexibility and potential for application across various operational environments. The solution's generalizability is, thus, not only feasible, but also beneficial for enhancing operational efficiency, resilience, and decision-making in port logistics globally.

References

1. Belmoukari, B.; Audy, J.F.; Forget, P. Smart port: A systematic literature review. *Eur. Transp. Res. Rev.* **2023**, *15*, 4. [CrossRef]
2. Song, D. A literature review, container shipping supply chain: Planning problems and research opportunities. *Logistics* **2021**, *5*, 41. [CrossRef]
3. Lange, A.K.; Nellen, N.; Jahn, C. Truck appointment systems: How can they be improved and what are their limits? In Proceedings of the International Conference of Logistics, Hamburg, Germany, 21–23 September 2022. [CrossRef]
4. Abu Aisha, T.; Ouhimmou, M.; Paquet, M. Optimization of container terminal layouts in the seaport—Case of port of Montreal. *Sustainability* **2020**, *12*, 1165. [CrossRef]
5. Li, N.; Chen, G.; Govindan, K.; Jin, Z. Disruption management for truck appointment system at a container terminal: A green initiative. *Transp. Res. Part D Transp. Environ.* **2018**, *61*, 261–273. [CrossRef]
6. Heikkilä, M.; Saarni, J.; Saurama, A. Innovation in smart ports: Future directions of digitalization in container ports. *J. Mar. Sci. Eng.* **2022**, *10*, 1925. [CrossRef]
7. Issa Zadeh, S.B.; Esteban Perez, M.D.; López-Gutiérrez, J.-S.; Fernández-Sánchez, G. Optimizing Smart Energy Infrastructure in Smart Ports: A Systematic Scoping Review of Carbon Footprint Reduction. *J. Mar. Sci. Eng.* **2023**, *11*, 1921. [CrossRef]
8. Molavi, A.; Lim, G.J.; Race, B. A framework for building a smart port and smart port index. *Int. J. Sustain. Transp.* **2020**, *14*, 686–700. [CrossRef]
9. González-Cancelas, N.; Molina Serrano, B.; Soler-Flores, F.; Camarero-Orive, A. Using the SWOT methodology to know the scope of the digitalization of the spanish ports. *Logistics* **2020**, *4*, 20. [CrossRef]
10. Issa Zadeh, S.B.; López Gutiérrez, J.S.; Esteban, M.D.; Fernández-Sánchez, G.; Garay-Rondero, C.L. Scope of the Literature on Efforts to Reduce the Carbon Footprint of Seaports. *Sustainability* **2023**, *15*, 8558. [CrossRef]
11. Beresford, A.K.C.; Gardner, B.M.; Pettit, S.J.; Naniopoulos, A.; Wooldridge, C.F. The UNCTAD and WORKPORT models of port development: Evolution or revolution? *Marit. Policy Manag.* **2004**, *31*, 93–107. [CrossRef]
12. Heilig, L.; Schwarze, S.; Voß, S. An analysis of digital transformation in the history and future of modern ports. In Proceedings of the Fifty International Conference on System Sciences, Honolulu, HI, USA, 4–7 January 2017. [CrossRef]

13. Heilig, L.; Stahlbock, R.; Voß, S. From Digitalization to Data-Driven Decision Making in Container Terminals. In *Handbook of Terminal Planning*, 1st ed.; Böse, J., Ed.; Springer: Suderburg, Germany, 2020; pp. 125–154. [CrossRef]

14. Min, H. Developing a smart port architecture and essential elements in the era of Industry 4.0. *Marit. Econ. Logist.* **2022**, *24*, 189–207. [CrossRef]

15. Morelli, M.; Casagrande, M.; Forte, G. Decision making: A Theoretical review. *Integr. Psychol. Behav. Sci.* **2022**, *56*, 609–629. [CrossRef] [PubMed]

16. Muniswamaiah, M.; Tappert, C.; Agerwala, T. Data Virtualization for Decision Making in Big Data. *Int. J. Softw. Eng. Appl.* **2019**, *10*, 45–53. [CrossRef]

17. Eboigbe, E.; Nnabugwu, O.; Farayola, O.; Daraojimba, C.; Olatoye, F. Business Intelligence transformation through AI and Data Analysis. *Eng. Sci. Technol. J.* **2023**, *4*, 285–307. [CrossRef]

18. Sarabia-Jacome, D.; Palau, C.E.; Esteve, M.; Boronat, F. Seaport Data Space for Improving Logistic Maritime Operations. *IEEE Access* **2020**, *8*, 4372–4382. [CrossRef]

19. Caldeirinha, V.; Nabais, J.L.; Pinto, C. Port Community Systems: Accelerating the Transition of Seaports toward the Physical Internet—The Portuguese Case. *J. Mar. Sci. Eng.* **2022**, *10*, 152. [CrossRef]

20. Jeevan, J.; Chen, S.-L.; Cahoon, S. The impact of dry port operations on container seaports competitiveness. *Marit. Policy Manag.* **2018**, *46*, 4–23. [CrossRef]

21. Neugebauer, J.; Heilig, L.; Voß, S. Digital Twins in the Context of Seaports and Terminal Facilities. *Flex. Serv. Manuf. J.* **2024**, *36*, 821–917. [CrossRef]

22. Clemente, D.; Taveira-Pinto, F.; Cabral, T.; Rosa-Santos, P. Blue Seaports: The Smart, Sustainable and Electrified Ports of the Future. *Smart Cities* **2023**, *6*, 1560–1588. [CrossRef]

23. Wei, H.; Sheng, Z. Dry Ports-Seaports Sustainable Logistics Network Optimization: Considering the Environment Constraints and the Concession Cooperation Relationships. *Pol. Marit. Res.* **2017**, *24*, 143–151. [CrossRef]

24. Graf, C.; Six, M. The effect of information on the quality of decisions. *Cent. Eur. J. Oper. Res.* **2014**, *22*, 647–662. [CrossRef]

25. Solanki, V. Evolution of Business Intelligence Tools. *Int. J. Res. Appl. Sci. Eng. Technol.* **2023**, *11*, 1149–1151. [CrossRef]

26. Gurcan, F.; Ayaz, A.; Menekse Dalveren, G.G.; Derawi, M. Business intelligence strategies, best practices, and latest trends: Analysis of scientometric data from 2003 to 2023 using machine learning. *Sustainability* **2023**, *15*, 9854. [CrossRef]

27. Gorman, M.F.; Clarke, J.P.; de Koster, R.; Hewitt, M.; Roy, D.; Zhang, M. Emerging practices and research issues for big data analytics in freight transportation. *Marit. Econ. Logist.* **2023**, *25*, 28–60. [CrossRef]

28. Sugrue, D.; Adriaens, P. A data fusion approach to predict shipping efficiency for bulk carriers. *Transp. Res. Part E Logist. Transp. Rev.* **2021**, *149*, 102326. [CrossRef]

29. Duru, O.; Galvao, C.B.; Mileski, J.; Robles, L.T.; Gharehgozli, A. Developing a comprehensive approach to port performance assessment. *Asian J. Shipp. Logist.* **2020**, *36*, 169–180. [CrossRef]

30. Jafari, T.; Zarei, A.; Azar, A.; Moghaddam, A. The impact of business intelligence on supply chain performance with emphasis on integration and agility–a mixed research approach. *Int. J. Product. Perform. Manag.* **2023**, *72*, 1445–1478. [CrossRef]

31. Lim, Y.K.; Park, J.W.; Min, H. Supply chain analytics for enhancing the maritime security. *Int. J. Logist. Syst. Manag.* **2017**, *28*, 164. [CrossRef]

32. Jovic, M.; Kavran, N.; Aksentijevic, S.; Tijan, E. The Transition of Croatian Seaports into Smart Ports. In Proceedings of the 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics, Opatija, Croatia, 20–24 May 2019. [CrossRef]

33. Lei, H.; Yifei, H.; Yi, G. The research of business intelligence system based on data mining. In Proceedings of the International Conference on Logistics, Informatics and Service Sciences, Barcelona, Spain, 27–29 July 2015. [CrossRef]

34. Ain, N.; Vaia, G.; DeLone, W.; Waheed, M. Two decades of research on business intelligence system adoption, utilization and success—A systematic literature review. *Decis. Support Syst.* **2019**, *125*, 113113. [CrossRef]

35. Kruglov, A.; Strugar, D.; Succi, G. Tailored performance dashboards—An evaluation of the state of the art. *PeerJ Comput. Sci.* **2021**, *7*, e625. [CrossRef] [PubMed]

36. Işık, Ö.; Jones, M.C.; Sidorova, A. Business intelligence success: The roles of BI capabilities and decision environments. *Inf. Manag.* **2013**, *50*, 13–23. [CrossRef]