POLITECNICO DI TORINO

Master Degree Course in Engineering and Management

Innovation Management and Entrepreneurship

Master Degree Thesis

Qualitative evaluation of Power BI and its application in an Aerospace Industry Cluster

Supervisors

Prof. Luca MASTROGIACOMO

Prof. Federico BARRAVECCHIA

Candidate

Chiara MARTUCCI

ACADEMIC YEAR 2023-2024

Abstract

The increasing influx of data in the contemporary business landscape underscores the crucial role of Data Visualization in enabling companies to extract meaningful insights and make informed decisions. As organisations struggle with ever-expanding data sets from diverse sources, effectively visualising and interpreting this information becomes essential. Data Visualization enhances understanding and facilitates the dissemination of complex insights within the organisation hierarchy.

This thesis aims to explore the key role of Data Visualization by comprehensively evaluating the various data visualisation tools available to enterprises. The primary objective was the identification of the most suitable tool to facilitate decision-making processes and improve operational efficiency. This study investigated two of the leading Data Visualization tools in the Business Intelligence market, Power BI and Tableau. Through a qualitative analysis, their respective features, functionality, and performance metrics were evaluated. Power BI stood out as one of the most effective data visualisation tools due to its extensive capabilities, user-friendly interface, and seamless integration due to its belonging to the Microsoft ecosystem.

The potential of the Power BI software was tested on the Umbria Aerospace Cluster, a consortium of 40 companies located in the Umbria region. Whereas the Cluster previously relied on conventional data analysis and monitoring methodologies, which often resulted in static and time-consuming processes, with the implementation of Power BI, the Cluster's database has experienced a significant increase in dynamism and operational efficiency. By harnessing Power BI's advanced visualisation features, the Umbria Aerospace Cluster gained invaluable insights into its operational landscape, identified optimisation opportunities, and steered decision-making processes grounded in data-driven intelligence.

In conclusion, this thesis has highlighted how the interactive dashboards provided by Power BI can enable stakeholders to intuitively explore data, unveil latent trends and decipher intricate patterns.

Contents

Chapter 1

Introduction

In today's dynamic business environment, companies are increasingly recognizing the importance of leveraging data to drive their growth and success. The growing amount of data available highlights the importance of Business Intelligence (BI), a powerful tool that empowers companies to uncover valuable insights from complex datasets. BI encompasses a range of processes and technologies aimed at collecting, analyzing, and visualizing data to facilitate strategic decision-making within companies. By transforming raw data into actionable insights, BI enables companies to optimize their operations, identify new opportunities, mitigate risks, and adapt to market shifts

Central to the BI ecosystem are Data Visualization tools, which serve as intuitive platforms for interpreting and communicating data effectively. Data Visualization not only aids in understanding the data but also simplifies the communication of intricate insights throughout the organization's structure. According to literature, among the available tools Power BI and Tableau emerge as prominent contenders, offering robust features and capabilities to create interactive dashboards and reports. In this thesis work, it was decided to focus on Power BI, examining its functionality and exploring how its integration can enrich the organizational environment. To emphasize the benefits of using Power BI in comparison with other

tools available on the market, it was decided to conduct a thorough comparative analysis with its main competitor, Tableau. In particular, differences in terms of structure and services offered were evaluated by emphasizing for both software strengths and weaknesses. Through this benchmark, the research endeavors to offer valuable insights into selecting the most suitable data visualization tool to meet the diverse needs of users and organizations.

Moreover, the selected tool, which is deemed most suitable based on factors like ease of implementation, compatibility with existing infrastructure, and advanced collaboration features, is employed for conducting an economic evaluation of an aerospace industry cluster called Umbria Aerospace Cluster. The aerospace sector, characterized by its emphasis on precision, efficiency and innovation, presents unique challenges such as stringent regulatory standards, intricate supply chains and rapidly changing market landscapes. In this demanding environment, the effective use of Business Intelligence and Data Visualization tools is of immense importance. Leveraging the insights derived from data analysis is crucial for aerospace companies to achieve operational excellence and make strategic decisions that drive growth. Through the integration of Power BI, it has been possible to revolutionize a once static report, which heavily relied on charts from different sources as Excel and secondary data. Previously, the creation and updates of this report were laborious and time-consuming. However, Power BI transformed this process, making it dynamic and interactive. It not only simplified report creation but also enabled real-time updates, overcoming the challenges of inefficiency and stillness associated with traditional methods.

Chapter 2

Business Intelligence for companies' strategy

In today's rapidly evolving business landscape, marked by the pervasive influence of digital technologies and the relentless pursuit of competitive advantage, for businesses it's crucial to pay utmost attention to the data they possess and to be able to extract meaningful insights from it.

Companies find themselves immersed in a sea of information streaming in from diverse sources such as customer interactions, sales transactions, social media platforms, IoT devices, and more. This deluge of data presents both opportunities and challenges for firms.

On one hand, the abundance of data offers immense potential for gaining valuable insights into customer behavior, market trends, operational efficiencies, and strategic opportunities. Analyzing this data can unveil hidden patterns, correlations, and predictive indicators that can inform critical business decisions and drive innovation. However, on the other hand, the sheer volume, velocity, and variety of data can overwhelm organizations, making it difficult to effectively manage, process, and derive actionable insights from it. Many companies struggle with data silos, disparate data formats, inconsistent data quality, and legacy systems that

hinder their ability to harness the full potential of their data assets.

Moreover, as data privacy regulations become increasingly stringent and cybersecurity threats proliferate, companies must also grapple with the challenge of ensuring data security, compliance, and integrity throughout the data lifecycle.

So, data is the most important asset a company possesses, but at the same time, deriving a competitive advantage from it is anything but simple. As companies straggling with huge volume and high complexity of data at their disposal, traditional methods of data processing and decision-making proved inadequate. Manual data collection and analysis were time-consuming, error-prone, and often failed to deliver timely insights to support strategic decision-making.

The concept of Big Data refers to data sets so large, complex, and fast that they become difficult to handle with traditional data processing tools. Doug Laney introduced the so called *Five V's*, a conceptual framework for understanding and managing the challenges associated with data management and analysis in business contexts affected by Big Data issue.

- Volume: it refers to the vast quantities of data generated and collected by organizations, stemming from diverse sources such as transactions, interactions, and sensor feeds. The exponential growth in data volume necessitates scalable infrastructure and storage solutions to accommodate the burgeoning data landscape. Effective management of data volume entails the deployment of distributed computing architectures and storage systems.

- Velocity: it denotes the speed at which data is generated, processed, and analyzed in real-time or near real-time [12]. This is crucial to enable organizations to derive actionable insights and respond promptly to dynamic market conditions. To address the challenges posed by data velocity, businesses must invest in high-performance computing resources, stream processing technologies, and event-driven architectures to enable timely data processing and

decision-making.

- Variety: it emphasizes the need to manage a wide range of data types and formats, including structured, unstructured, and semi-structured data. This heterogeneity poses challenges in terms of data integration, interoperability, and analysis. Advanced analytics techniques such as natural language processing (NLP), text mining, and machine learning algorithms are instrumental in extracting meaningful insights from disparate data sources.

- Value: it concentrates on the value that data can provide to the company when transformed into meaningful insights. The concrete result is achieved only when data are transformed into valuable information from which to derive knowledge in order to make targeted decisions that can be translated into oriented actions, activities and choices. To do this requires analytical tools.

- Veracity: it refers to the accuracy, reliability, and trustworthiness of data. Data governance frameworks, data quality assurance mechanisms, and metadata management practices play pivotal roles in mitigating risks related to data inaccuracies, inconsistencies, and biases.

In response, organizations began to explore innovative approaches to data management and analysis. The concept of Business Intelligence (BI) emerged as a comprehensive solution to the challenges posed by the abundance of data.

The term *Business Intelligence* was initially coined in the mid-1990s by the Gartner Group, a prominent global information technology firm. In September 1996, Gartner Group reports characterized Business Intelligence within forward-thinking enterprises as encompassing "information and applications accessible widely to employees, consultants, customers, suppliers, and the public." In the competitive landscape of contemporary markets, maintaining a competitive edge necessitates
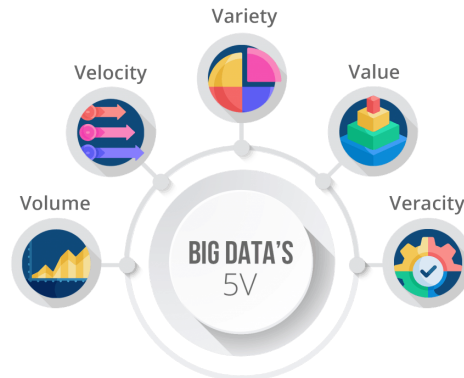
Figure 2.1: Model of Five V [4]

proactive measures. Effective decision-making in business contexts relies on more than just intuition; it demands access to accurate and up-to-date information [7].

The concept of Business Intelligence encompasses two distinct but interrelated definitions: one as a business process and the other as a technology-driven approach. In the literature, Business intelligence is quoted as the process of converting data into information, which is subsequently transformed into knowledge [6]. By integrating advanced analytics tools, data visualization techniques, and robust data governance frameworks, Business Intelligence enables organizations to harness the power of their data to drive informed decision-making and gain a competitive edge in the marketplace.

At its core, Business Intelligence involves a range of technologies and methodologies designed to collect, organize, analyze, and visualize data in a meaningful and actionable way. Moreover, Business Intelligence goes beyond mere data analysis to encompass the broader goal of fostering a data-driven culture within organizations. By promoting data literacy, encouraging collaboration across departments, and aligning data initiatives with strategic objectives, BI empowers organizations to leverage their data as a strategic asset and drive continuous improvement and innovation.

It is possible to represent the Business Information System in a graphical way through the Anthony's Pyramid, a framework for understanding the different levels of decision-making within an organization, the corresponding roles and responsibilities associated with each tier. The Anthony's Pyramid, introduced by Robert J. Anthony, is a conceptual model that elucidates how BI tools and techniques cater to decision-making needs across three primary organizational tiers: operational decisions, management decisions and strategic decisions.

Figure 2.2: Anthony's Pyramid *(own reworking)*

- At the base of the pyramid are operational decisions, which pertain to day-to-day activities and are typically made by frontline employees. These decisions are focused on executing tasks and processes to ensure the smooth functioning of the organization's operations. At the operational level, BI facilitates real-time data analysis and reporting to support day-to-day decisions.

- Moving up the pyramid, the middle tier encompasses management decisions, which are made by mid-level managers responsible for overseeing departments or functional areas within the organization. Within the management tier, BI empowers mid-level managers with comprehensive analytics and performance metrics to make informed decisions regarding resource allocation, process optimization, and performance improvement initiatives.

- At the top of the pyramid are strategic decisions, which are made by top-level executives and leaders of the organization. At the strategic level, BI supports

top-level executives in formulating long-term organizational strategies, identifying market opportunities and mitigating potential risks. Strategic BI capabilities encompass predictive analytics, scenario modeling and data-driven insights to guide strategic planning and business development initiatives.

There are four fundamental steps that organizations typically follow to harness the power of data effectively through Business Intelligence: data collection, data visualization, data analysis, and performance monitoring.

The first step involves data collection, wherein organizations gather relevant data from various internal and external sources, such as interviews or survey. The goal is to compile a comprehensive dataset that provides insights into different aspects of the business environment.

Once the data is collected, the next step is visualization. This involves transforming raw data into meaningful visual representations such as charts, graphs, dashboards, and reports. Visualization techniques enable stakeholders to gain a clear understanding of trends, patterns, and relationships within the data, making it easier to identify key insights.

Following visualization, organizations proceed to data analysis. This step involves applying statistical techniques and data mining algorithms to uncover actionable insights from the dataset in order to identify trends, correlations, and outliers that may impact business performance. This deeper understanding of the data enables organizations to make strategic decisions and optimize their operations.

Finally, the last step is monitoring. Continuous monitoring of Key Performance Indicators (KPIs) and metrics is essential to track the effectiveness of business strategies and initiatives. By monitoring data in real-time or on a regular basis, organizations can identify emerging trends, detect anomalies and respond promptly to changes in the business environment.

2.1 Business Intelligence Architecure

2.1.1 Data Collection and Storing

In the field of Business Intelligence, data collecting comprises a wide range of inputs required to provide relevant insights and facilitate informed decision-making inside organizations. These inputs, sourced from both internal and external sources, constitute the foundational elements upon which BI systems rely to derive actionable intelligence.

Internally, organizations gather data from various operational systems, including enterprise resource planning (ERP), customer relationship management (CRM), supply chain management (SCM), and financial management systems. Additionally, organizations may collect data from internal databases, spreadsheets and other structured repositories that house valuable information pertinent to their business processes.

CRM systems are designed to manage interactions and relationships with current and potential customers. These systems capture a wide range of data related to customer interactions, including contact details, purchase history, inquiries, complaints, feedback, and preferences. These systems help organizations track customer behavior across various touchpoints, such as website visits, emails, phone calls, and social media interactions. By consolidating and analyzing this data, organizations gain insights into customer preferences, buying patterns, and satisfaction levels. This information is invaluable for tailoring marketing strategies, improving customer service, and fostering long-term customer loyalty.

ERP systems, instead, integrate core business processes and functions, including finance, human resources, supply chain management, manufacturing, and inventory management, into a unified platform. ERP data encompasses a wide range of information, such as sales orders, purchase orders, invoices, inventory levels, production

13

schedules, employee payroll, and financial transactions. By centralizing and standardizing this data, ERP systems provide organizations with real-time visibility into their operations, enabling efficient resource allocation, inventory optimization, and cost control.

SCM systems capture a wealth of data related to the various stages of the supply chain, including procurement, production, inventory management, logistics, and distribution. This data includes information on supplier performance, order processing times, inventory levels, transportation costs, lead times, and demand forecasts. By integrating data from different sources within the supply chain, SCM systems provide organizations with a holistic view of their operations, enabling them to identify inefficiencies, bottlenecks, and opportunities for improvement.

Externally, organizations leverage data from a multitude of sources to enrich their BI initiatives. This includes market research data, industry reports, economic indicators, demographic data, social media feeds, and competitive intelligence. While internal data sources, such as transactional data from operational systems, offer valuable insights into organizational performance and customer interactions, external data sources complement this information by providing broader context and external perspectives, such as market trends, competitor activities and macroeconomic factors that may influence firm business performance.

One significant category of external data sources in BI is qualitative data obtained through interviews and surveys. Interviews and surveys allow organizations to directly engage with stakeholders, including customers, suppliers, industry experts, and other relevant parties, to gather firsthand insights, opinions, and feedback. For instance, conducting customer interviews or surveys can help organizations uncover unmet needs, preferences, and pain points, informing product development, marketing strategies, and customer service initiatives. Similarly, interviews with industry experts or surveys of market participants can provide valuable market intelligence, helping organizations identify emerging trends, competitive threats,

and strategic opportunities.

Moreover, external data sources extend beyond qualitative data to include quantitative data from market research reports, government databases, economic indicators, and social media analytics. Market research reports offer comprehensive analyses of industry trends, market size, competitive landscapes, and consumer behavior, providing organizations with valuable insights into market dynamics and competitive positioning.
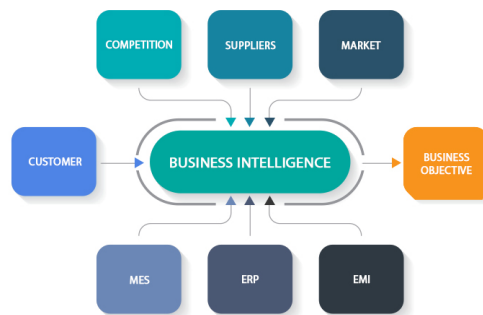

Figure 2.3: Data Inputs for Business Intelligence [5]

In the continuum of data management processes, the Data Warehouse (DW) emerges as a critical step subsequent to the data collection phase, assuming the role of a structured repository for the aggregation and organization of diverse datasets sourced from disparate origins. The data acquired must then undergo a process of extraction, cleaning and transformation before being loaded into the Data Warehouse, this process is referred to as Extract, Transform and Load (ETL).

During the extraction phase, data is retrieved from various sources, then the data undergoes transformation, where it is cleansed, standardized, and structured to meet the desired format and quality standards. This transformation phase often involves applying business rules, ensuring that the data aligns with organizational requirements and objectives. Finally, in the loading phase, the transformed data is loaded into the target destination, whether it's a data warehouse, data lake, or other storage platform. This loading process is typically optimized for efficiency

and scalability, ensuring that the data is readily accessible for further analysis.

The aim of ETL process is to provide complete, accurate, and consistent data to the DW, integrating with the Data Quality system. In fact, Data Quality is "a fundamental requirement for the entire information system" [21], precisely because if errors are present in the databases, inaccurate analyses are produced, which in turn result in incorrect decisions, thus worsening business performance. Data quality depends on three categories of quality factors, namely accuracy (conformity to actual values), completeness (presence of necessary data) and consistency (no contradictions).

In addition, the data warehouse preserves the history of the data, providing the evolution of all measured quantities over time, thus enabling the realization of temporal analyses. In addition to the data warehouse, there is another form of analysis-oriented database, namely Data Marts: these are thematic databases, concerning a single business process within the company, so they can be multiple, unlike the data warehouse which is instead unique.

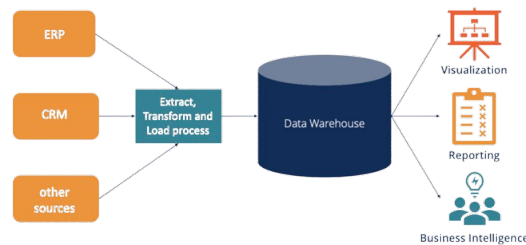Figure 2.4: From Data Collection to Data Visualization process

2.1.2 Data Visualization

Following the initial step of collecting data from different sources, the company is confronted with a plethora of data from which it has to derive maximum value for future endeavors. Data visualization draws upon principles of visual design and the theory of visual perception to effectively convey information within the domain

of Business Intelligence. By harnessing these principles, data visualizations aim to present complex data in a visually appealing and easily interpretable manner, enabling stakeholders to derive meaningful insights and make informed decisions. Visual design principles encompass various elements, including color, typography, layout and composition, which are strategically employed to create visually compelling graphics. These principles guide the arrangement and presentation of data elements to optimize clarity and comprehension.

Effective visualizations will allow a scientist both to understand their own data and to communicate their insights to others [29]. In the field of data visualization, different types of visualizations serve distinct purposes and are selected based on the nature of the data being represented.

Charts, such as bar charts, line charts, and pie charts, are commonly used to illustrate trends, comparisons and distributions in numerical data. Tables present data in a structured format, allowing for detailed examination and comparison of individual data points. Maps are utilized to visualize spatial data, revealing geographical relationships and patterns. Diagrams, including flowcharts and network diagrams, depict connections and processes, aiding in the comprehension of complex systems.
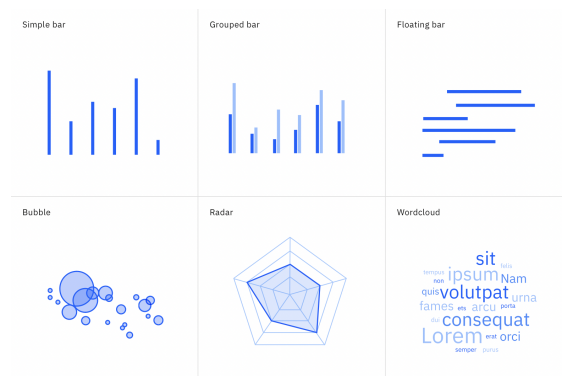

Figure 2.5: Simple bar, Grouped bar, Floating bar, Bubble, Radar, Wordcloud [13]

Data visualization holds paramount importance in the corporate landscape, serving as a conduit for the transmission of complex datasets into digestible visual formats. Its significance stems from its ability to clarify complex information, enabling stakeholders to discern key patterns, trends and relationships within datasets more efficiently than traditional textual or numerical representations alone. Within the business environment, data visualization serves multifaceted purposes.

By transforming convoluted data sets into visually understandable representations, it enables stakeholders to more intuitively grasp essential patterns, trends and relationships within the data. Through the use of intuitive graphical forms such as charts, graphs and maps, data visualization facilitates understanding and insight into business performance, market dynamics, consumer behavior and other relevant factors.

Key Performance Indicators (KPIs) are quantifiable metrics used to evaluate the success or performance of an organization, a department or a project in achieving predefined goals and objectives. These metrics serve as key metrics to measure progress and identify areas for improvement; KPIs are typically aligned with overall business objectives and are adjusted to reflect the organization's specific priorities. By visualizing KPIs, metrics, and other relevant data points through Data Visualization, organizations can identify areas of strength, weakness and opportunity, thereby enhancing the efficacy of their decision-making processes.

Moreover, data visualization serves as a powerful medium for communicating complex information to a diverse array of stakeholders, including executives, managers, employees, and external partners. By presenting data visually, organizations can convey key messages more effectively, fostering greater understanding, engagement, and alignment among stakeholders. Through the strategic utilization of visualizations, organizations can facilitate discussions, align stakeholders around common objectives, and drive collective action towards shared goals.

Dashboards are a core component of data visualization in Business Intelligence,

as they are comprehensive, interactive interfaces that consolidate and display key metrics, trends and insights derived from complex data sets in a visually intuitive way. In essence, dashboards act as a window into the underlying data, offering users a consolidated view of critical information relevant to their specific roles or objectives. By leveraging interactive elements such as filters, drill-down functionality and tooltips, dashboards enable users to dynamically explore data, uncover patterns, and gain actionable insights with minimal effort.

The design and development of effective dashboards require careful consideration of various factors, including user requirements, data sources, visualization techniques, and usability principles. A well-designed dashboard should strike a balance between aesthetic appeal and functional utility, ensuring that the visualizations are clear, informative, and aligned with the goals of the intended audience.

Moreover, dashboards play a crucial role in promoting data-driven decision-making across organizations by democratizing access to data and fostering a culture of transparency and accountability. By providing stakeholders with real-time access to relevant metrics and KPIs, dashboards empower them to monitor performance and track progress towards goals.
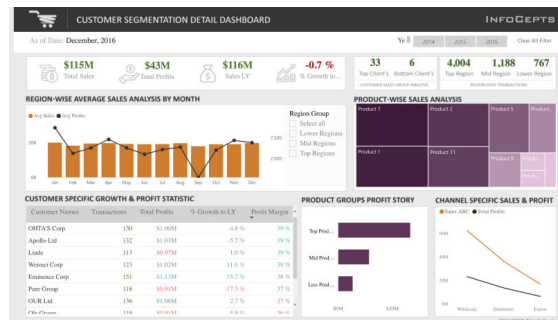

Figure 2.6: Dashboard exemplification [2]

As the demand for Data Visualization solutions continues to surge, a myriad of software and platforms have emerged to cater to diverse needs and preferences.

19

Whether it is creating interactive dashboards, generating insightful reports or crafting compelling visualizations, these software and platforms offer a diverse array of capabilities to suit various business requirements. The software available for developing a Business Intelligence project is diverse; below, there is a brief description of the three most popular platforms.

Tableau, founded in 2003, stands out for its user-friendly interface and robust visualization capabilities. It allows users to easily create interactive and visually appealing dashboards by simply dragging and dropping data fields. Tableau's patented technology, VizQL, enables users to explore and analyze data dynamically, facilitating quick insights and decision-making. Additionally, Tableau offers advanced analytics features, such as predictive analytics and statistical modeling, empowering users to uncover trends and patterns in their data. The platform supports a wide range of data sources and integrates seamlessly with various systems, making it a popular choice among organizations seeking powerful data visualization tools. [30]
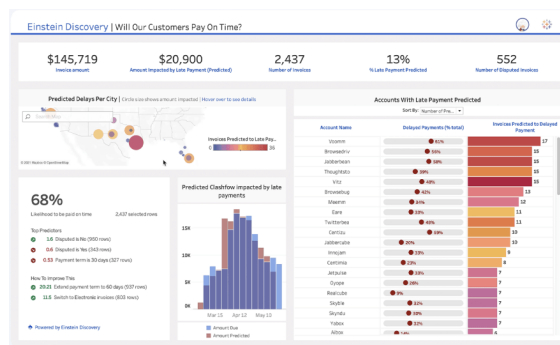
Figure 2.7: Tableau Dashboard [24]

Power BI, developed by Microsoft in 2005, is another leading data visualization software known for its integration capabilities and cloud-based services. Power BI offers a comprehensive suite of tools for data preparation, modeling, visualization and collaboration, all tightly integrated with the Microsoft ecosystem. Users can connect to multiple data sources, including Excel spreadsheets, SQL databases, and

cloud services, to create interactive reports and dashboards. Power BI's natural language processing feature enables users to ask questions about their data using plain language, making insights accessible to users at all skill levels. Moreover, Power BI's robust security and compliance features make it a preferred choice for organizations that prioritize data governance and regulatory compliance.
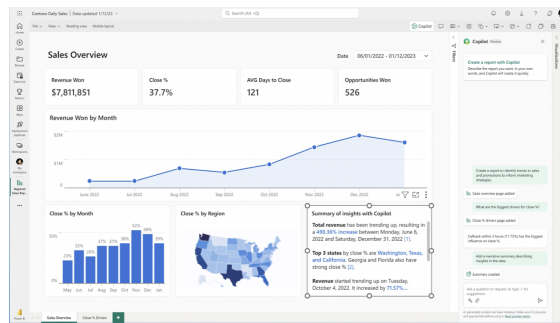

Figure 2.8: Power BI Dashboard [2]

Qlik, founded in 1993, is renowned for its associative data model and powerful data discovery capabilities. Unlike traditional BI tools, Qlik's associative engine allows users to explore data relationships intuitively, making it easier to uncover hidden connections. Qlik Sense, the company's flagship product, offers a self-service approach to data visualization, empowering users to create personalized dashboards and reports without relying on IT support. Qlik's platform also supports advanced analytics and augmented intelligence features. Furthermore, Qlik's open architecture and extensive API support facilitate seamless integration with existing systems and third-party applications, making it a flexible and scalable solution for organizations of all sizes.

When it comes to choosing a Business Intelligence software, market popularity is an important factor to consider. In terms of market share, both Tableau and Power BI have gained a significant position in the data visualization market and business intelligence space; although its market share may have been surpassed by Power BI and Tableau, Qlik remains a significant presence in the industry, especially
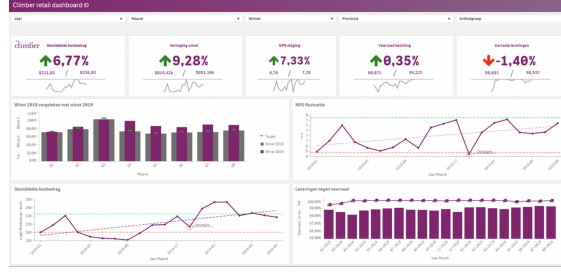
Figure 2.9: Qlik Dashboard [16]

in certain industries and regions. According to the latest analysis conducted by Statista in 2023, Power BI currently holds a 15.70% market share, positioning it as one of the leading solutions in the market. On the other hand, Tableau is not far behind, with a 14.90% market share, demonstrating its growing popularity among businesses. [16]

The specular characteristics of each software will be thoroughly discussed in Chapter 3.

2.1.3   Data Analysis

The conventional approach to Data Analysis involves manual examination and interpretation of data, which is prevalent across various industries. For example, in the healthcare sector, it is typical for experts to regularly examine present patterns and shifts in data, often on a quarterly basis. These specialists subsequently furnish a comprehensive report outlining their findings to the supporting healthcare entity. This report serves as the foundation for subsequent decision-making and strategizing in healthcare administration. However, this manual process is slow, costly and subjective, particularly as data volumes continue to escalate. With databases containing millions of records and numerous attributes, human analysis becomes impractical, necessitating automation to some extent. [8]

This increase underscores the pressing need for innovative computational theories and tools to aid in extracting valuable information from these massive datasets.

This emerging area of study, known as Knowledge Discovery in Databases (KDD), focuses on development of methodologies and technologies to assist humans in navigating and deriving knowledge from ever-growing volumes of data. The concept of Knowledge Discovery in Databases emerged from the first workshop in 1989, emphasizing that knowledge is the ultimate outcome of data-driven exploration. Its primary objective is to address the challenge of transforming extensive low-level data, which is often overwhelming and difficult to comprehend, into more manageable, abstract, or useful forms. This transformation may involve condensing the data into concise summaries, abstracting it into descriptive models of underlying processes or creating predictive models for forecasting future outcomes.

Throughout history, various terms have been used to describe the process of finding valuable patterns in data, including Data Mining, knowledge extraction, information discovery and data pattern processing. Data Mining, in particular, has been widely adopted by statisticians, data analysts and the Management Information Systems (MIS) communities, gaining traction in the field of databases as well. In essence, KDD encompasses the entire process of deriving useful knowledge from data, whereas Data Mining specifically refers to the step within this process where algorithms are applied to extract patterns. It is crucial to distinguish between the broader KDD process and the narrower Data Mining step.

Additional steps in the process, such as data preparation, selection, cleaning, integration of prior knowledge and interpretation of results, are vital to ensure that meaningful insights are obtained from the data. Blind application of Data Mining techniques, often criticized as data dredging, can lead to the discovery of spurious and irrelevant patterns, underscoring the importance of a comprehensive approach to knowledge discovery.

KDD is concerned with the holistic process of deriving knowledge from data, encompassing aspects such as data storage and access methods, scalability of algorithms to handle large datasets efficiently, interpretation and visualization of results

and the design of effective man-machine interactions. The process itself places particular emphasis on uncovering interpretable patterns that can be construed as valuable or insightful knowledge.
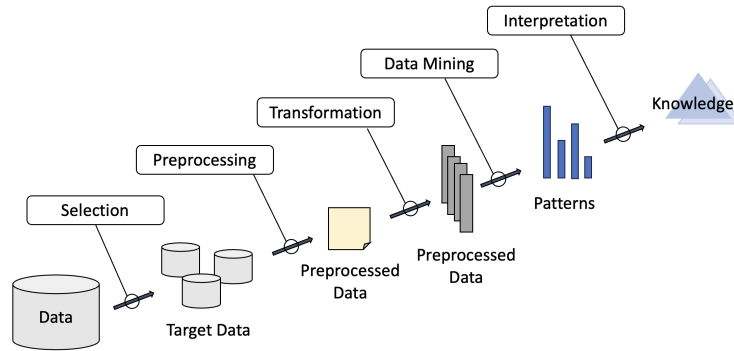

Figure 2.10: Knowledge Discovery in Databases (KDD) process

The Data Mining component of the KDD process often involves repeated iterative application of particular data-mining methods.

There are several software tools available for data analysis and data mining, each with its own features and advantages. These tools can be divided into several categories, including open-source software, commercial software and programming languages with specific libraries for data analysis. Some examples include Rapid-Miner, Weka, KNIME, Python with libraries such as *pandas* and *scikit-learn*, SAS and IBM SPSS. Each of these tools offers a range of capabilities for data analysis, data mining and predictive modeling, enabling users to explore and analyze their data effectively and efficiently. The choice of software depends on the specific needs of the project, the firm's skills and its own preferences.

2.1.4 Data Monitoring and Evaluations

In the intricate architecture of Business Intelligence tailored for enterprises, the culmination of data collection, visualization, and analysis unfolds into the stage of Data Monitoring and subsequent evaluations. This pivotal step encompasses a

dynamic process of continual surveillance and rigorous assessment, serving as the linchpin for informed decision-making and strategic refinement.

Data monitoring within this context encapsulates the vigilant observation of Key Performance Indicators, operational metrics and relevant data streams in real-time or through scheduled intervals. This active surveillance empowers organizations to promptly detect emerging trends, anomalies or shifts in patterns, providing invaluable insights into the ever-evolving landscape of their operations.

Complementing the vigilance of data monitoring, the evaluation component delves into a deeper analysis of collected data. It involves a methodical examination of performance metrics against predefined benchmarks, historical trends or targeted objectives. Through this evaluative lens, organizations can glean nuanced perspectives on the efficacy of their strategies, initiatives and operational processes.

The symbiotic interplay between data monitoring and evaluation forms the bedrock of a proactive and data-driven decision-making culture. By leveraging insights gleaned from ongoing surveillance and robust evaluation, businesses can not only navigate complexities with agility but also proactively identify opportunities for optimization and innovation.

Moreover, this iterative approach to data monitoring and evaluation fosters a culture of continuous improvement and adaptability within organizations. Armed with actionable insights, stakeholders can iteratively refine strategies, optimize resource allocation and align operations with evolving market dynamics.

In essence, data monitoring and evaluation serve as catalysts for organizational agility, resilience, and sustained growth in an ever-changing business landscape.

Chapter 3

Data Visualization Tools

In today's data-driven landscape, organizations rely heavily on Business Intelligence tools to extract meaningful insights from vast amounts of data. These tools empower decision-makers by providing visualizations, analytics, and reporting capabilities. Among the prominent players in this field are Microsoft Power BI and Tableau. The need for efficient BI solutions has grown exponentially due to the increasing complexity of business processes, data volumes, and the demand for real-time insights. Both Power BI and Tableau offer robust features to address these challenges, but they differ significantly in their approaches and functionalities.

This chapter aims to introduce the reader to the context of Data Visualization tools, focusing on the comparison between Power BI and Tableau. The study will explore their core features and architecture, outlining their limitations.

3.1   Power BI

Power BI is a cloud-based business analytics solution engineered by Microsoft for rapid and straightforward generation and dissemination of reports within a company. It comprises a suite of software services, applications, and connectors that synergize to transform disparate data sources into interactive, coherent, and readily

understandable insights. Power BI facilitates seamless connection to data sources, enables report and dashboard creation, and facilitates sharing with designated users. Its interactive dashboards facilitate effortless accessibility, contributing to timely decision-making and operational assistance.

Since its introduction in the market, Power BI has undergone significant development and expansion, evolving into a comprehensive platform for self-service and enterprise-level business intelligence. Over the years, Power BI has gained widespread adoption due to its user-friendly interface, powerful features, and seamless integration with other Microsoft products and services. As of 2024, it holds a prominent position in the market, competing with established BI tools such as Tableau and Qlik.

The Power BI ecosystem encompasses multiple tools, including a Microsoft Windows desktop application called Power BI Desktop, an online Software as a Service (SaaS) known as Power BI Service, and Power BI Apps for mobile devices called Power BI Mobile.

- Power BI Desktop: Power BI Desktop is a desktop application that serves as the primary tool for creating data models, reports, and dashboards. It offers a robust set of data preparation and visualization capabilities, allowing users to import data from various sources and create interactive visualizations using a drag-and-drop interface.

- Power BI Service: The Power BI Service is a cloud-based platform where users can publish, share, and collaborate on Power BI reports and dashboards. It provides features such as data refresh, scheduled data updates, and role-based access control. Additionally, the Power BI Service offers integration with other Microsoft cloud services such as Azure and Office 365.

- Power BI Mobile: Power BI Mobile consists of native mobile apps available

for iOS, Android, and Windows devices. These apps allow users to access and interact with Power BI reports and dashboards on the go, enabling real-time decision-making and insights dissemination.

One of the key strengths of Power BI is its robust data modeling capabilities. Users can import data from a wide range of sources and then manipulate and shape it using Power Query, a powerful data transformation tool [1].

- Databases: Power BI supports connections to various databases such as SQL Server, MySQL, Oracle, and PostgreSQL;

- Cloud Services: It integrates seamlessly with popular cloud services like Microsoft Azure, Google Analytics, Salesforce, and SharePoint Online;

- Files: Power BI can connect to files stored locally or in cloud storage services such as OneDrive and SharePoint, including Excel spreadsheets, CSV files, and XML files;

- Streaming Data: Power BI supports real-time data streaming from sources like Azure Stream Analytics, enabling users to visualize and analyze data as it arrives.

By offering such diverse connectivity options, Power BI empowers users to access and analyze data from virtually any source, making it a versatile and powerful tool for business intelligence and analytics.
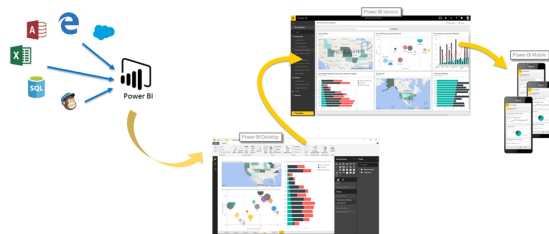

Figure 3.1: Power BI ecosystem [17]

3.1.1 Creation of Dashboard in Power BI Desktop

Power BI Desktop represents a powerful tool within the suite of Microsoft's business intelligence offerings, providing users with a comprehensive set of features for data analysis and visualization.

Power BI Desktop boasts an intuitive and user-friendly interface designed to streamline the process of data analysis and visualization. The interface is divided into several main areas, users can easily navigate between these areas to access various functionalities and tool. [28]

1. Ribbon Toolbar: This area is located at the top of the interface and contains a series of tabs, each providing access to various features and tools within Power BI Desktop. Common tabs include Home, Modeling, View, and Format.

2. Canvas Area: This area occupies the majority of the space in the Power BI Desktop interface and is where users can create and modify their reports and dashboards. It serves as a workspace where visual objects are created, combined, and reworked.

3. Data Pane: The Data Pane is typically located on the right side of the interface and provides a list of tables and fields present in the data model. Users can use this area to manage data sources, navigate between tables and fields, and create relationships between tables.

4. Visualizations Pane: This area provides a list of available visualizations. Users can drag and drop desired visualizations from the Visualizations Pane into the canvas area to create or modify their reports.

5. Filters Pane: The Filters Pane is a dedicated area for creating and managing filters applied to the data displayed within reports. Users can add filters to limit the data shown based on specific criteria, such as date ranges or specific categories.

6. Pages Pane: This area lists all the pages present in the current report. Users can create and manage multiple pages within a single report to organize their analysis in a more structured and cohesive manner.
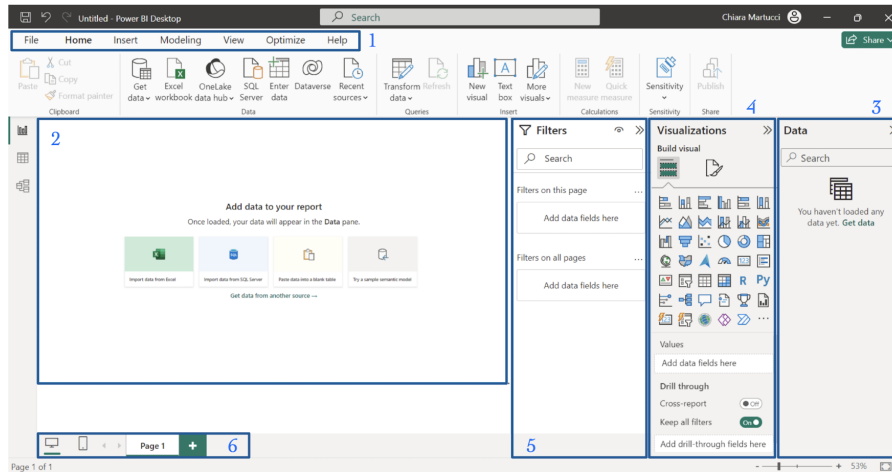

Figure 3.2: Power BI Desktop areas

3.1.2 Data Connection

In the introductory stage of creating a dashboard in Power BI, establishing connections to data is critical. The methods for connecting data sources in Power BI are data import, Direct Query and Live Connection.

The data import option in Power BI involves physically extracting rows of data from the selected source and storing them in memory within Power BI. This method is commonly used and offers several advantages, including improved query performance and efficient data compression. One noteworthy aspect of Power BI's data import mechanism is its use of xVelocity technology, it is an in-memory data storage and compression engine developed by Microsoft. xVelocity employs a columnar storage format, where data is stored in columns rather than rows. This format allows for highly efficient compression and retrieval of data, as it eliminates redundancy and maximizes data compression ratios. As a result, Power BI can handle

31

large volumes of data while minimizing memory usage and maximizing query performance.

Direct Query is a data connection method in Power BI that enables real-time querying of data directly from the underlying data source. With Direct Query, Power BI sends queries to the source database or system each time a user interacts with a report or visualization. This ensures that the data displayed in Power BI is always up-to-date, reflecting the latest information from the source. Direct Query is particularly advantageous when working with large or frequently changing datasets, as it eliminates the need to import and store data within Power BI. Instead, users can leverage the live connection to access real-time data without worrying about data freshness or storage constraints. However, it's important to note that Direct Query may introduce performance considerations, as queries are executed against the source system in real-time. Factors such as network latency, query complexity, and the performance of the source system can impact the responsiveness of reports and visualizations. [18]

Live Connection is another data connection method in Power BI that allows users to connect directly to an external data model or dataset, such as an Analysis Services cube or a tabular model. With Live Connection, Power BI serves as a visualization layer on top of the existing data model, enabling users to create reports and dashboards without importing data into Power BI. Live Connection is beneficial when organizations already have established data models in other reporting tools or platforms and want to leverage Power BI's visualization capabilities. By connecting directly to the existing data model, users can create interactive reports and visualizations without duplicating data or compromising data integrity. Still, it is essential to ensure that the external data model is properly maintained and updated, as any changes to the underlying data structure or calculations must be made in the original data model and not within Power BI. [18]

The difference between DirectQuery and Live Connection, is that with the latter, generally, there are no performance problems because it is supported only by the following types of data sources: SQL Server Analysis Services Tabular, SQL Server Analysis Services Multidimensional, Service Power BI. The reason why the performance does not suffer with these data sources is because they use the same xVelocity engine as Power BI or similar one.
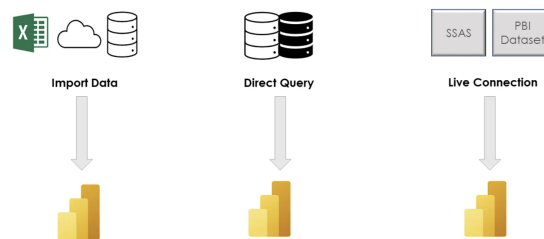

Figure 3.3: Data Connection methodologies in Power BI [19]

3.1.3 Data Preparation and Transformation

In any BI project, it is crucial that the data used be properly cleaned to obtain accurate results in reports and dashboards. Once the data connection has been established, whether through Live Connection, Direct Query, or data import, the subsequent phase involves Data Preparation. It emerges as arguably the most challenging, resource-intensive, and time-consuming aspect, as assessed by experts who estimate it consumes 60%-80% of the time and budget allocated to a standard analytics project. Numerous hurdles characterize this phase, encompassing fragmented and incomplete datasets, intricate system integration requirements, business data lacking structural uniformity, and a significant barrier posed by the demand for specialized skills. [20] Moreover, the inconsistency in the structure of business data adds another layer of complexity, requiring meticulous efforts to standardize and cleanse the data to ensure accuracy and reliability in analytics outcomes.

To address numerous challenges encountered in data preparation, Power BI

provides cutting-edge data preparation functionalities through Power Query, setting a new industry standard.

Power Query employs an intuitive, step-based interface based on a plethora of tools and features that allow users to easily apply the desired transformations, according to their requirements. Every action performed within the Power Query editor in Power BI automatically translates into a specialized scripting language called 'M'. Most of the transformations therefore, can be accomplished simply by interacting with the user interface of the Power Query editor but, for more complex business problems, there is a good chance that one will resort to editing M queries, written directly by the user.

Transformation operations sought by users may include removing duplicate rows, changing data types, splitting columns, creating calculated columns with custom formulas, adding indexes, and much more.
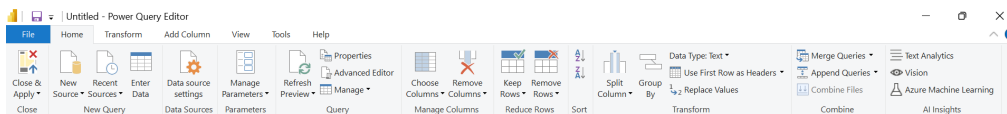
Figure 3.4: Power Query editor

Power Query empowers users to merge, aggregate, and append data from diverse sources seamlessly. This includes merging multiple tables based on shared columns, performing aggregation for creating summaries, and consolidating data from various origins into a unified table. Additionally, Power Query incorporates a query management panel facilitating users to oversee and control all queries within their data model. Through this panel, users can rename, modify, delete, or add queries as necessary, while also adjusting query settings like connection parameters and transformation steps. Notably, Power Query offers a unique feature wherein users can record and replay transformation steps executed on data. This feature enables the recording of a sequence of transformation operations for easy application to future datasets.

Figure 3.5: Transforming data with Power Query

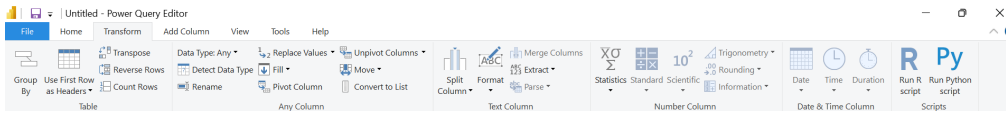## 3.1.4 Data Modeling

The transformational processes enacted within Power Query lay the groundwork for the subsequent construction of a robust data model within Power BI. Data modeling involves designing and structuring the underlying data model within Power BI to support analysis and visualization. In Power BI, the data model serves as the backbone of the analysis, defining relationships between different data tables and enabling users to perform cross-table analysis and calculations. Power BI uses a tabular data model, which consists of tables, columns, and relationships. Users can import data from multiple sources into Power BI and then define relationships between related tables based on common key fields. This allows users to create sophisticated reports and visualizations. In addition to establishing relationships, users can enhance their data model by creating calculated columns, measures, and hierarchies. Calculated columns enable users to derive new data based on existing columns, while measures allow users to perform aggregations and calculations fastly. Hierarchies provide a way to organize and drill down into data hierarchically, enhancing the user experience in reports and visualizations.

By mastering data preparation and data modeling techniques in Power BI, organizations can unlock the full potential of their data and drive informed decision-making. With clean and well-structured data at their fingertips, users can create compelling reports, interactive dashboards, and insightful visualizations that deliver actionable insights to stakeholders across the organization.
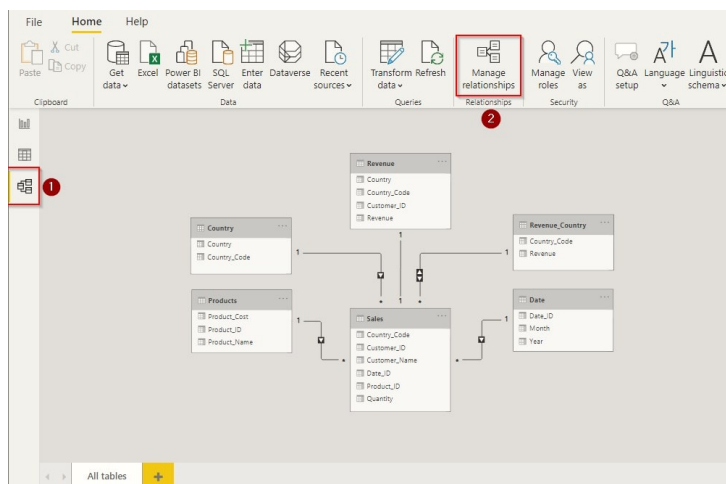
Figure 3.6: Data Model in Power BI [14]

3.1.5   Data Visualization

Within the software, users are provided with a myriad of visualization options, each tailored to specific data presentation needs. Whatever it is, the choice made significantly influences how data is visually represented and interpreted. The types of data visualizations in Power BI can be divided into several categories based on their nature and use.

- Tables and matrices present data in a structured tabular format, with rows and columns representing data dimensions. Tables offer a detailed view of individual data points, allowing users to examine precise values and relationships. Matrices extend this functionality by enabling data aggregation, subtotaling, and hierarchical organization, making them suitable for summarizing complex datasets.

- Bar and column visualizations are graphical representations of data where values are depicted using horizontal or vertical bars. They are commonly employed to compare data across different categories or to illustrate data distributions. These visualizations are versatile and can effectively convey both categorical and numerical data.

- Line visualizations present data points connected by lines, often used to visualize trends or patterns over time or other continuous dimensions. They provide insights into how data values change over a specified period, allowing for the identification of trends, seasonal patterns, or anomalies.

- Scatter plot visualizations plot individual data points on a two-dimensional chart, with each point representing the values of two variables. They are particularly useful for visualizing relationships between variables and identifying patterns or correlations. By examining the distribution of points, analysts can discern any underlying trends or Clusters within the data.

- Pie and donut visualizations display data as circular segments, with each segment representing a proportion of the whole. They are commonly used to illustrate the composition of a dataset, highlighting the relative sizes of different categories or subgroups. While pie charts show individual segments as slices of a circle, donut charts feature a central hole, allowing for additional data labels or annotations.

- Tree maps visualize hierarchical data structures by representing data categories as nested rectangles within a larger rectangle. Each rectangle's size and color represent quantitative values, allowing users to quickly grasp the relative size and importance of different data categories within the hierarchy.

- Map visualizations utilize geographical maps to represent spatial data, with data points plotted on specific locations or regions. They offer insights into geographic patterns, distributions, or trends, making them invaluable for analyzing location-based data. Map visualizations can display data using various markers, colors, or overlays, enhancing the understanding of spatial relationships within the dataset.

- Combo charts combine multiple chart types, such as lines, bars, and areas,

within a single visualization. This allows users to compare different datasets or data series on the same chart, providing a comprehensive view of multiple dimensions of the data within a single visual.

- Bubble chart visualizations depict data points as bubbles, with the size of each bubble representing a third variable. They are effective for visualizing relationships between three variables and can convey complex data patterns in a straightforward manner. By examining the position and size of bubbles, analysts can identify correlations or trends across multiple dimensions.

- Ribbon charts display data as a series of contiguous ribbons, where each ribbon represents a different category or data dimension. They are effective for visualizing data distributions and hierarchical relationships, providing insights into the relative proportions and distributions of data categories.

- Waterfall charts illustrate the cumulative effect of sequential positive and negative changes in a data series. They are commonly used for financial analysis and reporting to visualize the net effect of various income and expense components on a total value, facilitating easy identification of key contributors to overall changes.

- Gauge visualizations provide a concise representation of a single data value within a predefined range or threshold. They typically display data values as a dial, needle, or fill level within a gauge, allowing users to quickly assess performance against predetermined targets or benchmarks.

- Key Performance Indicator visuals display one or more key metrics or performance indicators as numerical values or icons. They are often used to monitor progress towards specific goals or objectives, providing at-a-glance insights into performance trends and deviations from target values.

Figure 3.7: Different Types of Data Visualization in Power BI [3]

Upon selecting a visualization type, users need to map the appropriate data fields to the visualization's axes, values, categories, and other relevant properties. This crucial step ensures that the visualization accurately reflects the insights derived from the dataset. Furthermore, Power BI offers extensive customization options for visualizations, enabling users to tailor the appearance and behavior of each visualization. From adjusting colors, fonts, labels, and scales to adding trendlines or reference lines, customization empowers users to craft visualizations that precisely convey their intended message.

Once the desired visualizations have been created, integrating them into the dashboard is a simple process in Power BI. After selecting the visualization to be added, users can click on the "Pin visual" or "Pin to dashboard" option. This action pins the visualization to the dashboard, making it readily accessible for viewing and analysis. Then, users can resize, reposition, and rearrange visualizations within the canvas area to achieve the desired layout: Power BI offers extensive flexibility in this regard. This capability enables users to prioritize and emphasize specific data insights, ensuring clarity and coherence in the presentation of the dashboard.

In addition to data visualizations, Power BI allows for the incorporation of supplementary elements to enrich the dashboard's content and functionality. These elements may include images, explanatory text, interactive buttons, and text boxes. By integrating such elements, users can provide context, explanations, and additional resources to enhance the audience's understanding of the presented data.

3.1.6   Dashboard Sharing

Upon completing the dashboard creation process, users have the option to publish it to the Power BI Service for sharing with other users. This functionality enables the dissemination of insights and findings to a wider audience within an organization or beyond. The process of publishing and sharing a dashboard involves several key steps:

- Publishing the Dashboard: once the dashboard is ready, users can initiate the publishing process directly from Power BI Desktop or through the Power BI service. This action uploads the dashboard and associated visualizations to the Power BI cloud service, making it accessible from any device with an internet connection.

- Setting Permissions: after publishing the dashboard, users have the ability to set permissions and control access to it. This includes defining who can view, edit, or share the dashboard, as well as specifying any restrictions or limitations on its use. Power BI provides granular control over access permissions, allowing users to tailor access rights to specific individuals, groups, or roles within the organization.

- Sharing Options: Power BI offers multiple sharing options to suit different collaboration scenarios. Users can share the dashboard directly with specific individuals by providing their email addresses or usernames. Alternatively,

users can share the dashboard with groups or distribution lists, streamlining the sharing process for larger audiences. Additionally, Power BI supports sharing via secure links, enabling users to share dashboards with external stakeholders or partners while maintaining control over access.

- Collaborative Features: in addition to basic sharing functionality, Power BI provides collaborative features that enhance team collaboration and productivity. These include features such as commenting, annotation, and real-time collaboration, allowing users to engage in discussions, provide feedback, and collaborate on dashboard development and refinement.

- Monitoring and Tracking Usage: once the dashboard is shared, users can monitor its usage and track engagement metrics using built-in analytics and reporting features in Power BI. This includes tracking the number of views, interactions, and shares, as well as analyzing user engagement patterns and trends over time. By gaining insights into dashboard usage, users can optimize content, improve usability, and ensure the dashboard continues to meet the needs of its audience.

By leveraging the publishing and sharing capabilities of Power BI, users can effectively disseminate insights, foster collaboration, and drive data-driven decision-making across their organization.

3.2 Tableau

Tableau originated from a research project at Stanford University in 2003, where co-founders Chris Stolte, Pat Hanrahan, and Christian Chabot collaborated to enhance data analysis and accessibility through visualization. Their innovative technology, VizQL, revolutionized data representation by translating user actions into data queries, simplifying the process through an intuitive interface. Over the years,

Tableau has continued investing heavily in Research and Development, expanding its capabilities to incorporate machine learning, statistics, natural language processing, and smart data preparation to enhance analytical creativity. Beyond providing a comprehensive analytics platform, Tableau offers support and resources to foster a data-driven culture within organizations, enabling them to achieve valuable outcomes efficiently. Acquired by Salesforce in 2019, Tableau remains committed to empowering individuals and organizations to gain insights from their data. Today, Tableau is utilized worldwide across various sectors and industries, driving impactful change through data-driven decision-making [25].

The Tableau ecosystem comprises several core components that work together to facilitate the creation, sharing, and exploration of data visualizations and analytics. The main components of Tableau include:

- Tableau Desktop is the primary authoring tool used by analysts and data scientists to create interactive dashboards, reports, and visualizations. It provides a user-friendly interface for connecting to data sources, building visualizations using drag-and-drop functionality, and exploring insights through interactive filtering and drill-down capabilities.

- Tableau Server is a centralized platform for sharing, collaborating, and managing Tableau workbooks and dashboards within an organization. It allows users to publish their Tableau Desktop creations to a secure server environment, where they can be accessed and viewed by authorized users via web browsers or Tableau's mobile app.

- Tableau Online is a cloud-based deployment option for Tableau Server, offering similar functionality but hosted in the cloud by Tableau. It provides scalability, flexibility, and accessibility, allowing organizations to leverage Tableau's capabilities without the need for on-premises infrastructure.

Tableau offers versatile connectivity options to connect with a wide array of data sources. These include connections to:

- Databases: Tableau facilitates connections to both relational and non-relational databases such as MySQL, PostgreSQL, SQL Server, Oracle, among others;

- Cloud Services: Tableau seamlessly integrates with major cloud services like Amazon Redshift, Google BigQuery, and Microsoft Azure SQL Database;

- Files: Users can import data from various file formats including Excel, CSV, JSON, and PDF among others. Tableau allows for straightforward importation and integration of data from local files into visualization projects;

- Streaming Data: Tableau also supports connections to streaming data sources, enabling real-time analysis from sources like Apache Kafka and Amazon Kinesis.

The extensive data sources available within Tableau synergize seamlessly with its strategic collaboration with Salesforce, amplifying the scope and depth of analytical capabilities for informed decision-making within Salesforce's CRM environment. Salesforce's acquisition of Tableau underscores a commitment to bolstering its analytics capabilities, enabling seamless access to Tableau's robust data visualization tools within Salesforce's CRM environment. This integration empowers Salesforce users to harness Tableau's comprehensive data analysis functionalities directly within their CRM workflows, thereby facilitating deeper insights into customer data, sales performance, and business operations. Consequently, this strategic partnership not only expands Tableau's reach but also enriches Salesforce's CRM offerings with advanced analytics capabilities, thereby fostering a more comprehensive solution for organizations seeking to leverage data for informed decision-making and competitive advantage.

3.2.1 Creation of Dashboard in Tableau Desktop

Tableau Desktop is the keystone tool for designing and developing dashboards in Tableau. Its robust functionality and intuitive interface empower users to transform raw data into compelling visualizations and interactive dashboards. At its core, Tableau Desktop offers a comprehensive suite of functionalities tailored for data analysis and visualization. Users can connect to various data sources, as said, ranging from traditional databases to cloud services and files. Once connected, Tableau Desktop facilitates data exploration through its drag-and-drop interface, allowing users to effortlessly build visualizations without the need for complex coding or scripting.

Upon starting Tableau Desktop, users are greeted by an easy-to-use interface designed into several key work areas, each of which performs a specific function in the dashboard design workflow:

1. Data Source Pane: Located on the left-hand side of the interface, the Data Source Pane provides access to all connected data sources. Here, users can view the tables, fields, and relationships within the data source, as well as perform data preparation tasks such as filtering, grouping, and aggregation.

2. Canvas Area: Positioned in the center of the interface, the Canvas Area serves as the workspace where users build their dashboards. Here, users can drag and drop fields onto shelves to create visualizations, arrange and format elements, and design the overall layout of the dashboard.

3. Sheets and Dashboards Tabs: At the bottom of the interface, users can navigate between Sheets and Dashboards tabs. Sheets represent individual worksheets where users create and customize visualizations, while Dashboards serve as containers for organizing and presenting multiple visualizations together.

44

4. Toolbar: At the top of the interface, the Toolbar provides quick access to essential tools and functions. Users can use the Toolbar to perform actions such as connecting to data sources, saving workbooks, formatting visualizations, and sharing dashboards.

5. Data Pane: Positioned on the left-hand side of the Canvas Area, the Data Pane displays the fields and measures available in the current view. Users can use the Data Pane to add, remove, and rearrange fields, as well as apply filters and sorting to the data.
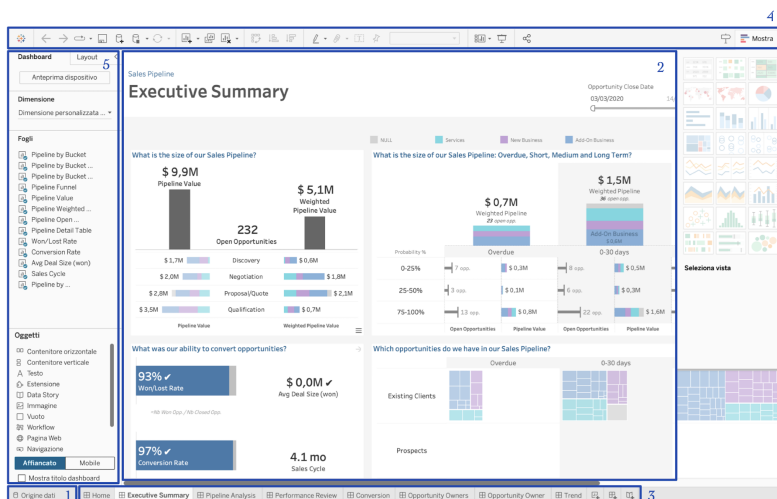
Figure 3.8: Tableau Desktop areas

3.2.2 Data Connection

In the process of building a dashboard in Tableau, the selection and connection of data sources serve as foundational steps, laying the groundwork for insightful data analysis and visualization.

Before embarking on dashboard creation, it is crucial to identify and select the relevant data sources that will drive the analysis. This entails a thorough understanding of the business requirements and the specific data needed to address

them effectively. Tableau's intuitive interface facilitates the selection process by offering seamless connectivity to various data repositories, including databases, cloud services, files, and streaming data sources. As already said, Tableau offers a rich array of data connectivity options, ensuring flexibility and compatibility with diverse data sources.

Once the data sources are identified and connected, Tableau offers various methods for importing and integrating data: users can opt for full data imports, where the entire dataset is brought into Tableau Desktop for analysis; alternatively, Tableau's data extraction capabilities allow for the optimization of performance and resource utilization by creating extracts of the data that can be stored locally or on Tableau Server. This latter approach is particularly beneficial for large datasets or scenarios requiring frequent data updates. Additionally, Tableau supports live connections to data sources, enabling users to analyze real-time data without the need for data replication.

3.2.3   Data Preparation and Transoformation

Tableau offers a comprehensive suite of advanced data preparation features aimed at enhancing the efficiency and effectiveness of data transformation and integration processes for analysis. These tools empower users to manipulate and refine datasets according to their specific analytical needs, ensuring the generation of accurate and actionable insights. One key feature is data blending, which enables users to seamlessly combine data from multiple sources within Tableau. By blending data, users can correlate information from disparate datasets without the need for complex data modeling or preprocessing steps. This capability is particularly valuable in scenarios where data resides in different systems or databases, allowing users to perform holistic analysis across various data sources. Moreover, Tableau's data modeling capabilities provide users with the flexibility to structure and organize data within Tableau Desktop. Users can define relationships between

different data sources, create calculated fields, and implement advanced analytical functions to derive meaningful insights from the data. This empowers users to tailor the data model to suit specific analytical requirements, ensuring the accuracy and relevance of analysis outcomes. In addition to data blending and modeling, Tableau offers robust calculation features that enable users to perform complex calculations and derive new insights from the data. Whether it's aggregating data, performing statistical analyses, or creating custom metrics, Tableau provides a wide range of built-in functions and formulas to support advanced data manipulation. Furthermore, Tableau's integration capabilities extend beyond its native environment, allowing seamless integration with other data management and analytics platforms. Through connectors and APIs, Tableau facilitates interoperability with various systems and tools, enabling organizations to leverage their existing data infrastructure and investments. This promotes data sharing and collaboration across different departments and teams, fostering a culture of data-driven decision-making and innovation within the organization.

In Tableau, the process of creating queries and leveraging VizQL for visualization is seamlessly integrated into the user interface. Tableau's interface provides users with a canvas where they can visually construct queries by dragging and dropping fields from the data source pane. This drag-and-drop functionality eliminates the need for users to write complex SQL queries manually, making data analysis more accessible to users with varying levels of technical expertise. For example, fields can be placed on the columns shelf to define the columns of the query result, on the rows shelf to define the rows, or on the filters shelf to filter the data based on specific criteria. As users interactively build their queries on the canvas, Tableau dynamically generates the corresponding SQL queries in the background. This process enables users to quickly iterate and explore different aspects of their data without having to write SQL code or navigate complex database structures

manually. In addition to simplifying the query creation process, Tableau's drag-and-drop interface also allows for real-time query execution. Users can interact with visualizations, apply filters, and drill down into data, with Tableau dynamically generating and executing queries to reflect their actions.

Behind the scenes, Tableau's visual query language, VizQL, powers the drag-and-drop functionality and visualization capabilities. VizQL translates user actions into visual representations of the underlying data, generating semantic queries based on interactions. This allows users to interactively explore and analyze data without needing to understand the intricacies of SQL or database structures. VizQL optimizes query execution by generating efficient queries tailored to the user's actions and selections. This optimization ensures that queries are executed quickly and accurately, even when working with large datasets or complex visualizations [26].

3.2.4 Data Visualization

Selecting the appropriate visualization types is paramount in the quest to effectively convey insights derived from data. Within the Tableau ecosystem, users are presented with an array of visualization options, each tailored to specific data analysis scenarios. The typologies are the same previously listed for Power BI, in the Section 3.1. Anyway, when embarking on the selection process of the right type of visualization tool, several key considerations come to the fore:

- Understanding Data Characteristics: Delving into the inherent attributes of the dataset is fundamental. Factors such as data structure, distribution, and relationships must be meticulously examined to discern the most fitting visualization types.

- Aligning with Analytical Goals: The crux of visualization selection lies in its alignment with analytical objectives and so the identification of the precise

insights and messages intended for communication to users. Whether the aim is to identify trends, compare values, or explore intricate relationships within the data, the chosen visualization type should seamlessly facilitate the fulfillment of these goals.

- Considering Audience Preferences: A nuanced understanding of the preferences and familiarity levels of the target audience is indispensable. Tailoring the dashboard to resonate with the audience's expectations and prior experiences with visualization types ensures maximum impact and comprehension.

- Exploration of Visualization Types: Within the Tableau repertoire, an exhaustive exploration of visualization types unveils a diverse array of options, each endowed with unique features and capabilities. From fundamental chart types like bar charts, line charts, and scatter plots to more advanced visualizations such as heatmaps, treemaps, and box plots, Tableau offers a versatile toolkit for visual storytelling.

Designing dashboards in Tableau is not merely a technical task; it is an art form that blends creativity with analytical precision to deliver compelling insights to stakeholders. Effective dashboard layout serves as the foundation for guiding users through data insights: it involves organizing visual elements in a structured and intuitive manner to facilitate easy comprehension and navigation. Key principles to consider include:

- Clarity: Ensuring the dashboard layout is clear and unambiguous, with a logical hierarchy of information that guides users through the data journey.

- Simplicity: Striving for simplicity by presenting only the most relevant information and eliminating unnecessary clutter that may distract users from key insights.

- User-friendliness: Designing the dashboard with the end user in mind, making it intuitive and easy to interact with, regardless of the user's level of expertise.

- Consideration of Space: Utilizing space efficiently to strike a balance between providing sufficient information and avoiding overcrowding, allowing each visual element to breathe and stand out.

Also color serves as a potent instrument in dashboard design, wielding a considerable impact on data comprehension and user engagement. Designers must ensure that chosen colors cater to the needs of all users, including those with visual impairments. Attention must be paid to factors such as color blindness and contrast ratios to maintain accessibility across diverse user groups. Thoughtful selection of colors is vital also to ensure readability and coherence, even when colors are combined. This approach mitigates visual clutter and fosters clarity, enabling users to navigate the dashboard seamlessly and glean insights effortlessly. Harnessing the psychological effects of color can evoke specific emotions and responses from users, enriching their engagement with the dashboard content. Strategic incorporation of colors tailored to the target audience enhances the immersive nature of the user experience.

Figure 3.9: Tableau Dashboard example [23]

3.2.5 Dashboard Sharing

Tableau stands out for its collaboration features which are designed to foster teamwork, facilitate knowledge sharing, and drive collective decision-making within organizations. By leveraging Tableau's robust collaboration capabilities, users can share, discuss, and iterate on insights, thereby maximizing the value derived from data-driven initiatives. One of the key aspects of Tableau's collaboration functionality is its ability to enable users to share dashboards with colleagues, stakeholders, and decision-makers. Once a dashboard is created, users can publish it to Tableau Server or Tableau Online, both of which serve as centralized platforms for hosting and sharing Tableau content. This ensures that authorized users can easily access and interact with the dashboard using web browsers or Tableau's mobile app, regardless of their location or device.

Furthermore, Tableau's real-time collaboration features empower users to work together on dashboards, fostering a collaborative environment where teams can collectively analyze data, identify insights, and make informed decisions. Users can simultaneously view and interact with the same dashboard, enabling real-time collaboration and eliminating the need for back-and-forth exchanges of static reports or screenshots. This promotes transparency, agility, and alignment across teams.

In addition to real-time collaboration, Tableau offers features that enable users to annotate visualizations and share insights within the platform. Users can add comments, annotations, and highlights directly onto visualizations, providing context, explanations, and observations to accompany the data. This enriches the narrative of the dashboard and facilitates deeper understanding and interpretation of the insights presented. Moreover, users can share their insights with specific individuals or groups, ensuring that relevant stakeholders are informed and engaged in the decision-making process.

Overall, Tableau's collaboration capabilities play a pivotal role in driving collaboration, fostering cross-functional communication, and accelerating the pace of

decision-making within organizations. By providing a collaborative environment where users can easily share, discuss, and iterate on insights, Tableau empowers teams to harness the full potential of their data and drive meaningful business outcomes [23].

3.3 Comparative Analysis: Power BI vs Tableau

In today's data-driven business landscape, organizations rely on robust Business Intelligence tools to gain actionable insights from their vast repositories of data. Two prominent contenders in the BI software market, Power BI and Tableau, have emerged as leaders in empowering businesses to extract value from their data assets. Developed by Microsoft and Tableau Software (now part of Salesforce), respectively, these platforms offer sophisticated features for data visualization, analysis, and reporting.

The Gartner Magic Quadrant report is an annual assessment conducted by Gartner, a research and consulting firm, providing a comprehensive overview of key players and products within a specific technological sector. Within the realm of Business Intelligence and Analytics software, the Gartner Magic Quadrant report evaluates and analyzes leading solution providers based on various criteria, including completeness of vision and ability to execute. Companies are categorized into four quadrants: Leaders, Challengers, Visionaries, and Niche Players.

For the 16th year in a row, Microsoft retains its leadership position in Gartner's Magic Quadrant 2023 for Analytics and Business Intelligence (ABI) due to the excellence of Power BI [10]. Notably, Power BI is closely trailed by Tableau in the Leaders quadrant. Consequently, this Thesis will proceed in analyzing and comparing these two software solutions.

In the previous section, an extensive examination was conducted of the features of Power BI and Tableau individually, highlighting their respective strengths and

Figure 1: Magic Quadrant for Analytics and Business Intelligence Platforms
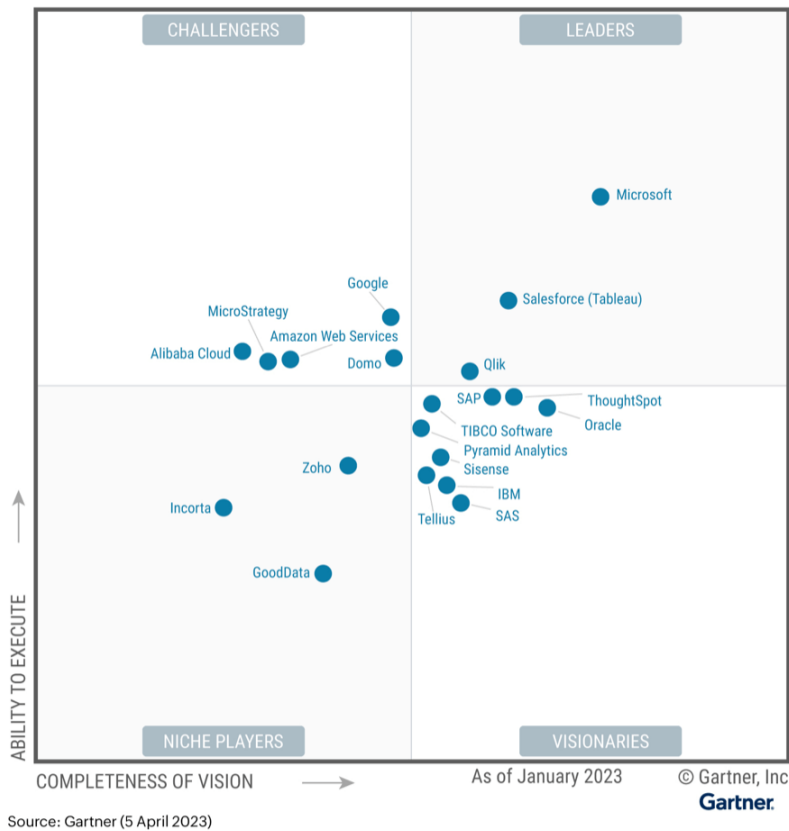

Figure 3.10: Gartner Magic Quadrant 2023 [22]

weaknesses. The focus now shifts to a qualitative comparative analysis between the two software solutions. Tableau and Power BI are recognized as industry leaders in data visualization solutions, each offering a unique set of capabilities and advantages. This phase aims to thoroughly compare the functionalities, performance, and usability of Power BI and Tableau, to identify significant differences and evaluate which tool is best suited to the specific needs of the organization.

Conducting a comparative analysis between two leading competitors is a strategic process aimed at gaining insights into their respective strengths, weaknesses, opportunities, and threats. The cornerstone is the definition of evaluation criteria, they enable a consistent and objective assessment of the two software available

acting as a guide to explore the intricacies of the two platforms and determine which one is best suited to the specific needs of the organization or project. The 9 evaluation criteria chosen for the scope of this Thesis are:

1. Ease of Deployment: This criterion examines how seamlessly each platform facilitates the setup and implementation process within an organization. It assesses factors such as the simplicity of installation procedures, compatibility with existing systems and infrastructure, and the availability of resources to guide users through the deployment process.

2. Data Visualization capabilities: This criterion assesses the range and sophistication of visualization options offered by each platform. It includes considerations such as the variety of chart types, interactivity features, and customization options available for creating visually appealing and insightful visualizations.

3. Data Connectivity and Integration: This criterion evaluates the ease of connecting to various data sources and the compatibility with different data formats. It also considers the integration capabilities with other tools and systems, such as databases, cloud services, and third-party applications.

4. Usability and User Experience: This criterion focuses on the intuitiveness of the user interface and the overall user experience provided by each platform. It considers factors such as ease of navigation, workflow efficiency, and the availability of user-friendly features for both novice and advanced users.

5. Performance and Scalability: This criterion assesses the performance metrics, such as processing speed, response time, and scalability of each platform. It evaluates their ability to handle large datasets and complex analyses efficiently, without compromising performance.

6. Integration with Artificial Intelligence (AI): This criterion evaluates the extent to which each platform integrates with AI capabilities, such as machine learning algorithms, natural language processing, and predictive analytics. It assesses the platform's ability to leverage AI-driven insights to enhance data analysis, automate repetitive tasks, and uncover hidden patterns or trends in the data.

7. Collaboration and Sharing: This criterion evaluates the collaboration features and sharing options available in each platform. It includes considerations such as real-time collaboration tools, version control, and accessibility of reports and dashboards for sharing with internal teams and external stakeholders.

8. Cost and Licensing: This criterion examines the pricing structure, licensing options, and total cost of ownership (TCO) for each platform. It includes considerations such as subscription fees, additional costs for advanced features, and scalability options to accommodate varying budget constraints.

9. Support and Community: This criterion considers the level of customer support provided by the vendor and the availability of resources such as documentation, tutorials, and community forums. It also evaluates the size and activity of the user community, which can provide valuable insights and support for users.

The comparative analysis was developed based on data gathered from surveys conducted with customers, feedback from users of each platform, and information from the official Gartner website [11]. Gartner considers various aspects and provides a numerical ranking of results obtained for each criterion of the two software solutions.

3.3.1 Ease of Deployment

The qualitative evaluation can now start from the first criteria. In terms of Ease of Deployment, both Power BI and Tableau present efficient processes, although they exhibit nuanced distinctions in their methodologies and integration functionalities. Implementing Power BI is often characterized by its user-friendly installation process and straightforward integration within an organization. With its strong integration with the Microsoft ecosystem, Power BI smoothly integrates with other Microsoft products and services, such as Azure, Excel, and SharePoint. This integration simplifies the setup and configuration process, as users can leverage existing Microsoft infrastructure and familiarity with Microsoft tools. Additionally, Power BI's cloud-based architecture allows for easy scalability and accessibility, enabling organizations to deploy and manage analytics solutions with minimal infrastructure requirements. Tableau offers flexible deployment options, including on-premises installations and cloud-based deployments on platforms like Tableau Online and Tableau Server. While Tableau's initial setup may require more technical expertise compared to some other BI tools, its deployment options, including on-premises, cloud, and hybrid deployments, cater to diverse organizational requirements. Tableau Server and Tableau Online provide centralized management capabilities, facilitating the deployment and distribution of analytics content across the organization. Although Tableau's deployment may involve additional configuration steps, particularly for on-premises installations, its flexibility and scalability make it a suitable choice for organizations seeking comprehensive analytics solutions tailored to their specific infrastructure and deployment preferences.

Overall, the two software taken in consideration offer relatively easy implementation processes, the choice between them may depend on factors such as existing infrastructure, organizational preferences, and technical expertise. In the Gartner report, both Power BI and Tableau received a score of 4.4 out of 5, highlighting the subjective nature of the company decision.
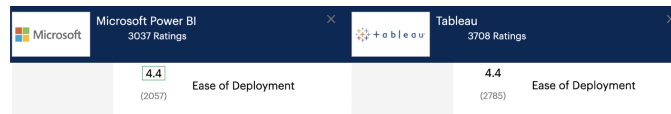
Figure 3.11: Ease of Deployment results [11]

3.3.2 Data Visualization capability

In terms of Data Visualization capabilities, Power BI offers a diverse range of visualization options, encompassing bar charts, line charts, pie charts, scatter plots, maps, and more. Its intuitive interface facilitates the creation of interactive and engaging data visualizations without requiring advanced technical skills. Moreover, Power BI provides advanced customization features, enabling users to modify colors, shapes, labels, and other elements to tailor visualizations to their specific needs. However, some more advanced visualizations may necessitate the use of extensions or custom packages, which may not be readily available in the default visualization library. Tableau is renowned for its extensive library of highly customizable and advanced data visualizations, including charts, maps, tables, scatter plots, and more. Its easy-to-use user interface and drag-and-drop functionality simplify the creation of complex and interactive visualizations. However, the learning curve for Tableau may be steeper for less experienced users, particularly those unfamiliar with data analysis, requiring a deeper understanding of expressions and calculation formulas. Both Power BI and Tableau provide a wide range of visualization capabilities, but Tableau is often considered more advanced and flexible in terms of customization and advanced features. However, Power BI excels in ease of use and native integration with other Microsoft solutions. For those seeking a platform that prioritizes user-friendliness and ease of learning, Power BI may present a favorable option. Conversely, individuals in need of a solution offering extensive customization capabilities and advanced analytical features might find Tableau to be more suitable.

In the Gartner report, both platforms received a commendable score of 4.6 out

of 5, reflecting a high level of satisfaction among users. This parity suggests that the decision between the two platforms is largely subjective and dependent on individual organizational needs and preferences. Therefore, organizations should carefully evaluate also other factors to determine the best fit for their specific requirements.

Figure 3.12: Data Visualization capability results [11]

3.3.3   Data Connectivity and Integration

When it comes to Data Connectivity and Integration, both Power BI and Tableau demonstrate robust capabilities, each with its own strengths and weaknesses. Power BI showcases seamless connectivity to a broad spectrum of data sources, spanning databases, cloud services, files, and APIs. Its integration with Microsoft ecosystem tools such as Azure, Excel, and SharePoint is highly efficient. Built-in connectors for popular data sources simplify data import and transformation processes. Moreover, Power BI supports real-time data streaming and automatic data refreshes, ensuring data insights remain current. However, Power BI may encounter limitations with complex data transformations or specific data sources. Additionally, achieving advanced integration may necessitate additional setup or custom development. Tableau offers extensive connectivity options, spanning databases, cloud platforms, web applications, and more. Its flexible architecture facilitates seamless integration with third-party tools and systems through APIs and extensions. Tableau's data preparation features enable users to clean, blend, and reshape data from multiple sources effortlessly. Furthermore, it supports live connections to data sources, enabling real-time access for analysis. Despite its robust connectivity options, Tableau users may face challenges in configuring and maintaining connections to certain data sources. Additionally, advanced integration features may

require additional technical expertise or support from IT professionals. Power BI prioritizes native integration with Microsoft ecosystem tools, making it particularly suitable for organizations heavily invested in Microsoft technologies. On the other hand, Tableau offers greater flexibility in integrating with a variety of third-party tools and systems, catering to diverse organizational environments. Considering the evaluation, both Power BI and Tableau excel in Data Connectivity and Integration. However, Power BI's seamless integration with the Microsoft ecosystem grants it an advantage for organizations deeply embedded in Microsoft technologies. Therefore, the ranking would favor Power BI in this aspect.

Indeed, Gartner reports a score of 4.3 for Power BI and 4.2 out of 5 for Tableau. Despite this, both scores reflect exceptionally high levels of satisfaction and performance for both platforms. The difference between the scores is minimal, indicating that both Power BI and Tableau excel in this criterion.
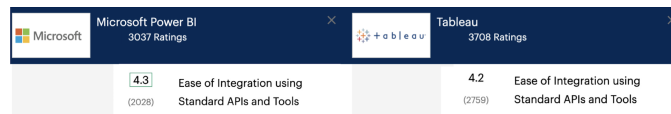

Figure 3.13: Data Connectivity and Integration results [11]

3.3.4 Usability and User Experience

When evaluating Usability and User Experience, Power BI boasts an intuitive user interface, making it accessible to users with varying levels of technical expertise [9]. Its seamless integration with the Microsoft ecosystem provides familiarity for users already accustomed to Microsoft products, enhancing ease of adoption. The platform offers a wide range of pre-built templates, visuals, and drag-and-drop functionalities, enabling users to create compelling dashboards and reports effortlessly. Tableau is reputable for its user-centric design and interactive capabilities, empowering users to explore data intuitively and derive meaningful insights. Its drag-and-drop interface simplifies the creation of complex visualizations, allowing users

to build interactive dashboards with ease. Tableau offers extensive customization options, enabling users to tailor visualizations to their specific requirements. Furthermore, Tableau's robust community support and extensive documentation contribute to a rich user experience, providing resources for users to enhance their skills and leverage advanced features effectively. While both Power BI and Tableau prioritize user experience, they differ in their approach and offerings. Power BI's integration with the Microsoft ecosystem offers familiarity and convenience for users already using Microsoft products, enhancing the adoption. On the other hand, Tableau's extensive customization options and interactive capabilities provide users with more flexibility in creating tailored visualizations and exploring data. Summarizing users' evaluations, Power BI may have an edge for organizations heavily invested in the Microsoft ecosystem, thanks to its seamless integration and familiarity. However, Tableau's extensive customization options and interactive capabilities make it a preferred choice for users seeking greater flexibility and control over their data visualization and analysis workflows.

In this category, Power BI achieved a score of 4 out of 5 which is higher than Tableau one (3.8) due to its seamless integration with the Microsoft ecosystem, intuitive interface, advanced collaboration features, and regular updates. However, also the evaluation of usability and user experience is subjective and influenced by individual preferences and specific use cases.

Figure 3.14: Usability and User Experience results [11]

3.3.5 Performance and Scalability

Power BI is known for its efficient Performance and Scalability, particularly in handling large datasets and complex analyses. Its integration with Azure services

provides robust infrastructure support, allowing for high-speed data processing and real-time analytics. Additionally, Power BI's cloud-based architecture enables seamless scalability, enabling organizations to expand their data capabilities as needed. However, some users have reported occasional performance issues when dealing with extremely large datasets or complex calculations. Tableau also offers strong performance and scalability capabilities, with its in-memory data engine facilitating fast data processing and visualization rendering. Tableau Server provides centralized management and scalability features, allowing organizations to distribute workloads across multiple nodes for enhanced performance. However, Tableau's on-premises deployments may require additional hardware resources to achieve optimal performance, and organizations may need to carefully manage server configurations to avoid performance bottlenecks.

In the criteria of Performance and Scalability, both Power BI and Tableau offer strong capabilities. Organizations should evaluate their specific performance requirements, data volumes, and deployment preferences to determine which platform best meets their needs. In the report made available by Gartner Power BI slightly exceeds the tableau rating with a ranking value of 2.7 and 3.6 respectively, highlighting a slight consumer preference for it.

Figure 3.15: Performance and Scalability results [11]

3.3.6 Integration with Artificial Intelligence

Both solutions showcase distinctive Integration with Artificial Intelligence. Power BI offers robust integration with (AI) capabilities through Azure Machine Learning. Users can leverage pre-built machine learning models or develop custom models directly within Power BI to analyze data, generate forecasts, and identify patterns

or anomalies. Power BI's integration with Azure services provides seamless access to AI functionalities, enabling users to leverage advanced analytics without leaving the Power BI environment. However, some users find the integration process complex, particularly when setting up and configuring Azure Machine Learning models within Power BI. Tableau provides integration with AI through the Tableau API and third-party platform integrations. Users can leverage external machine learning models to enrich data and integrate AI predictions directly into their analyses and visualizations. While Tableau's open API architecture allows for flexible integration with AI tools and services, users may encounter challenges in setting up and configuring external AI models within Tableau. Additionally, Tableau's approach to AI integration may require more technical expertise compared to Power BI's seamless integration with Azure Machine Learning. The major difference between Power BI and Tableau in terms of AI integration lies in their approach and ease of use. Power BI offers native integration with Azure Machine Learning, providing users with direct access to a wide range of AI capabilities within the Power BI environment. Tableau, on the other hand, relies on external integrations and third-party AI tools, giving users greater flexibility but potentially requiring additional setup and configuration.

In the Integration with Artificial Intelligence aspect, Power BI have a slight advantage over Tableau due to its seamless integration with Azure Machine Learning and direct access to AI functionalities within the Power BI environment. In fact, Power BI score is 4.6 out of 5 while the Tableau one is 4.2.
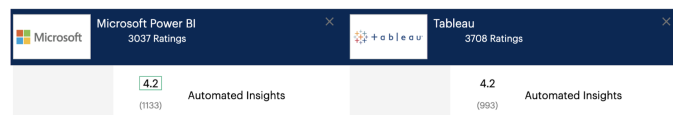

Figure 3.16: Integration with AI results [11]

3.3.7 Collaboration and Sharing capability

When assessing Collaboration and Sharing capabilities, both Power BI and Tableau offer distinct features and considerations. Power BI provides robust collaboration and sharing functionalities, facilitating seamless teamwork and knowledge exchange among users. Its integration with Microsoft Office 365 and SharePoint enables users to collaborate on reports and dashboards in real-time, leveraging familiar tools and workflows. Power BI's sharing options allow users to distribute reports securely within the organization or externally with stakeholders, ensuring data privacy and compliance. Additionally, Power BI's integration with Microsoft Teams enhances collaboration by enabling users to embed reports directly into team channels and chat conversations. However, some users find the sharing and collaboration features limited in the free version of Power BI, which may require upgrading to a paid plan for advanced functionality. Tableau offers comprehensive collaboration and sharing capabilities through its Tableau Server and Tableau Online platforms. Users can publish and share interactive dashboards securely, controlling access permissions and ensuring data governance. Tableau's subscription model allows organizations to scale collaboration features as needed, accommodating growing teams and increasing data sharing requirements. Additionally, Tableau's integration with Slack and other third-party collaboration tools enhances communication and collaboration among users. However, Tableau's collaboration features may require additional setup and configuration, particularly for on-premises deployments, and organizations may need to invest in server infrastructure to support large-scale collaboration.

To summarize, both Power BI and Tableau offer robust features for sharing and collaborating on reports and dashboards. Power BI may have a slight advantage for organizations already invested in the Microsoft ecosystem, thanks to its seamless integration with Office 365 and SharePoint. However, Tableau's centralized management capabilities and flexible scalability options make it a preferred choice for organizations requiring comprehensive collaboration features and governance

controls.

According to Gartner's assessment, Power BI received a ranking of 4.1 out of 5, while Tableau received a ranking of 4 out of 5 in this criterion. However, the specific requirements and preferences of an organization should ultimately dictate the choice between the two platforms.

Figure 3.17: Collaboration and Sharing results [11]

3.3.8 Cost and Licensing

In the evaluation of Cost and Licensing criteria, both Power BI and Tableau offer distinct pricing models and considerations. Power BI provides flexible pricing options tailored to the needs of different organizations. It offers both a free version with limited features and paid plans with advanced functionality and support. The paid plans are typically subscription-based, with pricing tiers based on user types and feature sets. Power BI also offers discounts for annual commitments and volume licensing, making it a cost-effective choice for organizations of all sizes. Additionally, Power BI's integration with the Microsoft ecosystem may provide added value for organizations already using Microsoft products and services. However, some users find the pricing structure complex, with additional costs for premium features and add-ons. Tableau offers several licensing options, including perpetual licenses, subscription-based models, and cloud-based deployments. Its pricing is typically based on factors such as the number of users, deployment type (on-premises or cloud), and feature requirements. Tableau's pricing can vary significantly depending on the deployment size and feature set, with larger organizations often requiring custom quotes for enterprise-scale deployments [27]. While Tableau may have a higher initial cost compared to Power BI, it offers comprehensive features and

scalability options that may justify the investment for organizations with complex data analytics needs. However, some users find Tableau's pricing prohibitive, particularly for smaller organizations or teams with limited budgets. In the Cost and Licensing criterion, both Power BI and Tableau offer competitive pricing options tailored to the needs of different organizations. Power BI may have a slight advantage for organizations seeking a cost-effective solution with flexible subscription options and integration with the Microsoft ecosystem. However, Tableau's comprehensive features and scalability options may justify the higher initial investment for organizations with complex data analytics needs.

Based on Gartner's evaluation, Power BI achieved a rating of 4.4 out of 5, whereas Tableau attained a rating of 4.2 out of 5 in this particular criterion.
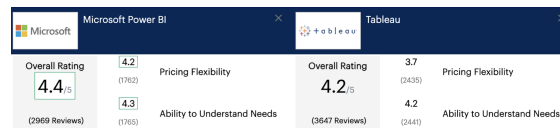

Figure 3.18: Cost and Licensing results [11]

3.3.9   Support and Community

When evaluating Support and Community, both Power BI and Tableau offer valuable resources for users, but there are differences in their support offerings and community engagement. Power BI provides extensive support resources through its online documentation, community forums, and user groups. Users can access a wealth of tutorials, articles, and troubleshooting guides to address common issues and learn new features. Additionally, Power BI offers official support channels, including email support, live chat, and phone support for users with paid subscriptions. The Power BI community is active and vibrant, with a large user base sharing insights, tips, and best practices through forums, social media groups, and local user groups. However, the official support options may be considered limited,

particularly for free users or those with basic subscriptions. Tableau offers compre-
hensive support resources, including online documentation, knowledge base articles,
and video tutorials, to help users get started and troubleshoot issues. Tableau's
official support includes email support, live chat, and phone support for users with
active maintenance or subscription agreements. Additionally, Tableau provides
dedicated support for enterprise customers, including access to a technical account
manager and priority response times. The Tableau community is known for its
active participation and engagement, with user groups, forums, and social media
channels providing opportunities for collaboration and knowledge sharing. In the
Support and Community criterion, both Power BI and Tableau offer robust support
resources and active user communities, making it challenging to determine a clear
winner. Power BI may have a slight advantage for users seeking accessible support
options and community engagement, while Tableau's targeted support offerings
may appeal to enterprise customers and users with specific technical requirements.

In the evaluation of Support and Community, Power BI earned a rating of 4.1
out of 5, while Tableau received a slightly lower rating of 4 out of 5. This suggests
that both platforms offer robust support resources and active user communities,
but Power BI may have a slight advantage in this aspect according to Gartner's
evaluation.

Figure 3.19: Support and Community results [11]

Delineating the comparison between Tableau and Power BI across nine distinct
criteria revealed a notable trend: Power BI excels over Tableau in six key aspects,
while both platforms stand on equal footing in the remaining three criteria. This
comprehensive analysis, based on user feedback and rigorous evaluation, unmistak-
ably underscores Power BI's superiority in meeting the diverse needs of users and

organizations.

The compelling evidence of Power BI's dominance leads to a consequential decision: to adopt Power BI as the primary software for the upcoming empirical analysis. This strategic selection not only underscores the platform's technical excellence but also underscores its alignment with the research objectives and requirements. By leveraging Power BI, the analytical framework is poised to benefit from robust capabilities and seamless functionality, facilitating a deeper and more impactful exploration of the research landscape.

Chapter 4

Application of Power BI

Business Intelligence is a cornerstone for companies, providing indispensable support for navigating the complexities of data management. Its overarching goal is to enable organizations not only to assess their performance, but also to decipher the underlying trends within their data sets. In addition, BI acts as a compass, guiding companies through the competitive landscape and providing in-depth comparisons with industry peers. This holistic approach not only sheds light on market positioning, but also identifies paths for growth and improvement. In essence, BI acts as a strategic ally, equipping companies with the tools and insights they need to thrive in an ever-changing marketplace.

In the context of the aerospace industry, it is common practice for companies to Cluster together to foster collaboration and innovation. These Clusters bring together a diverse array of stakeholders, including manufacturers, research institutions, academia, and government bodies, all united by a shared goal of advancing aerospace technologies and capabilities. By co-locating within these Clusters, companies benefit from proximity to key partners, suppliers, and talent pools, facilitating seamless collaboration and knowledge exchange. Moreover, Clustering encourages synergies among industry players, fostering a fertile environment for innovation and the rapid development of cutting-edge technologies.

Periodically, the *Politecnico di Torino* compiles a comprehensive report mapping the Umbria Aerospace Cluster. This report is generated from a large Excel file, which aggregates data from two primary sources: direct interviews with companies and the utilization of secondary data. The foundation of the first methodology adopted consists of conducting direct interviews with companies in the Cluster. Through these interviews, detailed and qualitative information is sought on the characteristics and internal dynamics of the companies. To this end, a survey instrument was developed specifically designed to investigate the characteristics of the cluster companies but also face-to-face interviews, providing a unique opportunity for direct observation of company operations, significantly enriching the understanding of internal dynamics. For some companies where physical visits were not feasible, remote interviews were conducted as an alternative. Collectively, the combination of face-to-face and remote interviews enabled a comprehensive and diversified insight into the Cluster's companies, laying a solid foundation for data analysis and processing. The second pillar of the methodology involves the utilization of secondary sources, including databases, company websites, and external references. These resources grant access to a wide range of information, with financial data from the AIDA Bureau Van Dijk proving particularly relevant for analysis. Overall, the integration of these two approaches - direct interviews and secondary sources - enabled a thorough and detailed mapping of the Umbria Aerospace Cluster's companies [15].

The combined approach adopted offers numerous advantages:

1. Comprehensive and balanced Data Collection: By utilizing both secondary data and direct interviews, a balanced dataset has been gathered, encompassing both quantitative and qualitative information. This has enabled a comprehensive and detailed view of the Cluster's companies.

2. Depth and contextualization of Information: Direct interviews have allowed

for the acquisition of contextualized information. Additionally, they have provided the opportunity to build a comprehensive framework of each company's operational, strategic, and cultural environment.

3. Flexibility: Thanks to the combination of in-person and remote interviews, it was possible to reach all the Cluster's companies, demonstrating flexibility in data collection methods.

These advantages highlight the robustness and adaptability of the chosen methodology, contributing to a thorough understanding of the Umbria Aerospace Cluster.

Figure 4.1: Methodology for mapping cluster companies

The report developed so far did not involve the use of any Business Intelligence software. However, following the comparative analysis conducted in Chapter 3 and the favorable outcome of Power BI, it was decided to proceed with the Cluster analysis using this software. This decision stems from the recognition of Power BI's effectiveness in analyzing and visualizing data, as demonstrated by its superior performance compared to its main competitor Tableau.

By leveraging the capabilities of Power BI, this thesis aims to improve understanding of the Umbria Aerospace Cluster, uncovering valuable insights and facilitating informed decision making.

4.1   Umbria Aerospace Cluster

The Umbria Aerospace Cluster was established in 2008 through the initiative of several leading companies, supported by Confindustria Umbria. Today, the association represents the umbrian regional industry in the aerospace, space, and defense sectors.

The Cluster currently comprises 40 member companies, as illustrated in Figure 4.2, generating an aggregate turnover exceeding 500 million euros and employing over 3000 workers. These companies possess expertise ranging from design to production, including proficiency in specialized processes and services. This diversification enables them to meet the demands of a constantly evolving market.

The member companies of the Umbria Aerospace Cluster are distinguished by their strong international orientation: over 75% engage in commercial relations with international clients, demonstrating a global presence and competitiveness. The Cluster actively promotes participation in major international events, fosters collaboration with national and international research institutes and entities, and invests in the training of its members.

Research and development play a central role for cluster companies, with investments reaching up to 20% of the average annual turnover. Most of the companies associated with the Umbria Aerospace Cluster are certified according to ISO quality standards and specific aerospace industry standards. This reflects the commitment and dedication of Cluster-affiliated companies to provide high-quality products and services, as well as their willingness to seize opportunities and attract clients requiring compliance with such certification requirements.

In general, the Cluster's mission is to promote growth and collaboration among member companies, with the aim of positioning Umbria as a leader in the aerospace, space, and defense sectors. This commitment is evidenced by the wide range of expertise, products, and services offered by member companies, which aspire to

make the Umbria Aerospace Cluster a reference point in the industry.

Cluster and its associated companies are in a continuous process of evolution and transformation. Therefore, the information contained in this report refers to the condition of the Cluster and its member companies as of July 2023. During the drafting of this thesis, some companies have joined the Umbria Aerospace Cluster. The companies *EN4 S.R.L., ERA Electronic System S.R.L., Meccanotecnica Umbra S.P.A., Sky Eye Systems S.R.L.,* and *Tifast Titanium S.R.L.* may not have been included in all the aggregated results presented in the document.

Figure 4.2: Umbria Aerospace cluster companies

The following analysis provides a quantitative overview of the entities constituting the Umbria Aerospace Cluster, aiming to offer a comprehensive insight into

the demographic and economic landscape of the Cluster. It encompasses an examination of diverse factors, including workforce size, company scale, and economic indicators. The objective is to gain a thorough understanding of the Cluster's composition and characteristics, shedding light on its role and impact within the aerospace industry.

4.1.1 Workforce and Enterprise Size

The employee count within the Umbria Aerospace Cluster has displayed notable fluctuations over the past five years, as illustrated in Figure 4.4. In 2018, the Cluster comprised 2.667 individuals in its workforce. Subsequently, there was a substantial surge in 2019, with the total number of employees reaching 2.969, marking an impressive 11% growth from the preceding year. Despite the global challenges posed by the pandemic, the Cluster managed to sustain its momentum in 2020, expanding its employee base to 3.132, representing a 5.5% increase compared to 2019. However, the trend took a downturn in 2021, with the number of employees slightly decreasing to 3.091. This decline continued in 2022, as the workforce diminished further to 3,062. These fluctuations likely reflect the lingering effects of the pandemic and may indicate a phase of consolidation following earlier periods of rapid growth.

It is essential to recognize that such variations in employment are common in specialized sectors like aerospace, influenced by various internal and external factors. Nonetheless, despite these fluctuations, the companies within the Umbria Aerospace Cluster have maintained a consistently high number of employees, underscoring the sector's significance to the regional economy. In the broader context of Umbria's labor market, the employees engaged through companies affiliated with the Cluster represent approximately 1.3% of the total workforce in the region. Looking ahead, growth projections for 2023-2024 and beyond are optimistic, with expectations of a return to pre-pandemic employment levels and gradual expansion

beyond that threshold.

Following the analysis conducted on the data showcased in the dashboard, attention is now shifted towards examining the dashboard itself. This dashboard, constructed using Power BI, operates autonomously, dynamically refreshing its content upon any modifications made to the source file, which, in this instance, is an Excel spreadsheet. This attribute endows the dashboard with a responsive essence, similar to a living entity that continually evolves to reflect the latest insights derived from data.

Within this dynamic Power BI dashboard, situated at the upper-left corner, are two visually prominent elements referred to as *cards* in Power BI. These card components serve a pivotal role in summarizing essential metrics or Key Performance Indicators relevant to the Umbria Aerospace Cluster. Specifically, they provide an immediate snapshot of vital statistics, including the total count of companies within the Cluster and the aggregate number of their establishments, encompassing both domestic and international locations. These cards offer a seamless and automated updating mechanism, ensuring that they remain synchronized with the underlying dataset. Consequently, as new companies join the Cluster and corresponding entries are appended to the Excel data source, these cards dynamically adjust to reflect the most current information available. This feature not only enhances the dashboard's usability but also underscores its capacity to deliver real-time insights, empowering users with up-to-date and accurate data at their fingertips.

Adjacent to these cards, positioned in the upper right corner, lies a dynamic element known as the Q&A box, embodying an innovative feature within Power BI. This interactive tool empowers users to engage directly with the dataset by formulating natural language queries, subsequently translated into actionable data inquiries yielding immediate responses. Notably, the Q&A box extends its utility by offering prompt suggestions, facilitating effortless exploration of pertinent insights. For instance, users can effortlessly inquire about metrics such as the number of

75

companies with international offices, thereby streamlining the analytical process and fostering a deeper understanding of the underlying data patterns.

Within the dashboard, positioned in the bottom-left corner, is a pie chart illustrating the count and relative percentage of each company's size. Specifically, to categorize the firms, the analysis relied on the segmentation outlined in Figure 4.3. This segmentation facilitated a comprehensive understanding of the distribution of company sizes within the Cluster, aiding in the assessment of its composition and dynamics.

Category	Employees		Turnover		Total Asset Value
Big company	≥250	*OR*	>50mln €	*AND*	>43mln €
Medium company	<250	*AND*	≤50mln €	OR	≤43mln €
Small company	<50	*AND*	≤10mln €	OR	≤10mln €
Micro company	<10	*AND*	≤2mln €	OR	≤2mln €

Figure 4.3: Categorization of Company Size within the Cluster

Completing the dashboard's layout, positioned in the lower right corner, is a dynamic bar chart depicting the workforce trends within the Cluster from 2018 to 2022, the latest year available in the AIDA secondary data source. This chart offers a visual representation of the Cluster's staffing trends over the specified period, facilitating the analysis of fluctuations and patterns in employment figures. This temporal analysis enables stakeholders to discern patterns and trends, informing strategic decision-making processes within the Cluster.

Through its amalgamation of interactive features and insightful visualizations, the Power BI dashboard serves as a powerful tool for exploring and understanding the demographic and economic landscape of the Umbria Aerospace Cluster.
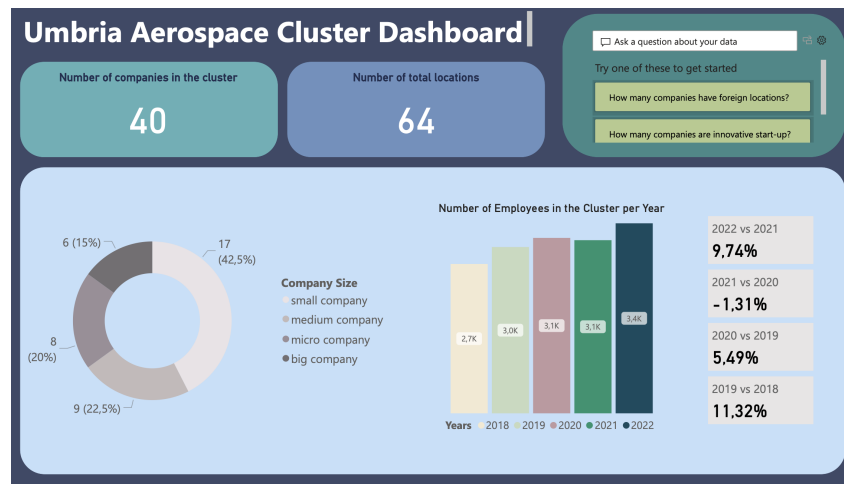
Figure 4.4: Workforce and Enterprise Size Dashboad on Power BI

4.1.2 Revenues from Sales and Services and Profits

The analysis of the Total Revenues from Sales and Services generated by the Cluster's companies over the past five years reveals a generally positive trend, as depicted in Figure 4.5. In 2018, the Cluster's Total Revenues stood at 496.318k€, emphasizing its significant contribution in the regional economy. This figure increased to 571.050k€ in 2019. However, the year 2020 witnessed a downturn, with revenues decreasing to 459.517k€ due to the economic challenges stemming from the global COVID-19 pandemic. Despite experiencing growth in 2021, reaching 492.701k€, it was slightly below expectations, possibly indicating the long-term effects of the pandemic. Nonetheless, by 2022, revenues rebounded to 537.867k€, showcasing the Cluster's resilience post-contraction. Forecasts for 2023 surpass previous record values, signaling optimism for the Cluster's future and the industry at large. Contributing to this optimism are factors such as post-pandemic economic recovery, investments in innovation and technology, and increasing demand for advanced aerospace solutions.

Examining the Average Revenues segmented by company size, provides deeper

77

insights into the Cluster's dynamics. These figures demonstrate significant fluctuations, reflecting the diverse composition of companies within the Cluster:

- Micro companies displayed volatility in average revenue over the analyzed period. Following a notable increase in 2019, micro enterprises experienced a significant decline in 2020, followed by a slight rebound in 2021 and a mild contraction in 2022, bringing average revenues back to 2020 levels.

- Small and medium-sized companies exhibited a growth trend from 2018 to 2019, followed by a contraction in 2020. Despite this contraction, these two categories remain crucial in the Cluster, contributing substantially to total revenues. The resurgence observed in 2021, and even more in 2022, indicates an economic recovery post-pandemic.

- Large companies showed a similar trend to small and medium ones, contributing to around 60% of the Total Cluster Revenue.

The analysis of the aggregate Profits of the Umbria Aerospace Cluster over the past five years reveals notable fluctuations. In 2018, Profits amounted to 33.094k€, underscoring the Cluster's effectiveness and profitability, which served as the primary driver of growth for the companies. The subsequent year, in 2019, witnessed a significant decline in profits, totaling 23.399k€. This trend continued in 2020, with profits further decreasing to 10.970k€, aligning with the downturn in Total Revenues attributed to the global economic impact of the COVID-19 pandemic. However, despite these challenges, the Cluster has shown resilience, maintaining significant profits under adverse conditions. A notable recovery in profits was observed in 2021, surging to 24.806k€. Yet, in 2022, profits experienced a significant contraction, dropping to 17.129k€. Future profit projections suggest optimism, anticipating a renewed growth trajectory similar to the post-pandemic period.

The analysis of average profits for each business size category provides a comprehensive overview of the financial performance of companies within the Cluster.

78

- In the case of micro companies, there was an increase in average profit from 2018 to 2019, followed by a decrease in 2020, which even reached negative values, indicating losses for micro enterprises. In the following years, a slight recovery is observed.

- Small and medium-sized compannies show similar trends, with variations in total contributions: they peaked in 2019 followed by a sharp decline in 2020; however, in 2021, the average profit returned to positive, followed by a significant increase in 2022, indicating a strong recovery capacity.

- Large companies showed a negative trend from 2018 to 2020, reaching the lowest point in this year. Subsequently, there was a significant recovery in 2021. However, in 2022, the average profit decreased again, although it did not reach the levels of the pandemic period.

Focusing now on the dashboard encompassing these datasets, it is possible to delineate two primary sections that have been crafted.

The initial section revolves around presenting the outcomes of the Revenue analysis concerning Sales and Services. Here, on the left side, there is an aggregated overview, consolidating data from all cluster companies, per year. Conversely, the right side entails a more detailed breakdown, segmenting these values based on the respective company sizes. This segmentation offers nuanced insights into Revenue distribution across various scales of enterprises within the Cluster.

Transitioning to the subsequent section, it mirrors a similar structure but focuses on Profit analysis. On the left side, the bar chart encapsulates this information, offering a graphical representation of profit trends over time. Meanwhile, the Clustered column chart on the right provides a comparative view, pitting the profits across different company sizes. This layout enables a comprehensive evaluation of profit dynamics within the Cluster, facilitating nuanced comparisons and insights into profitability trends based on company size categories.
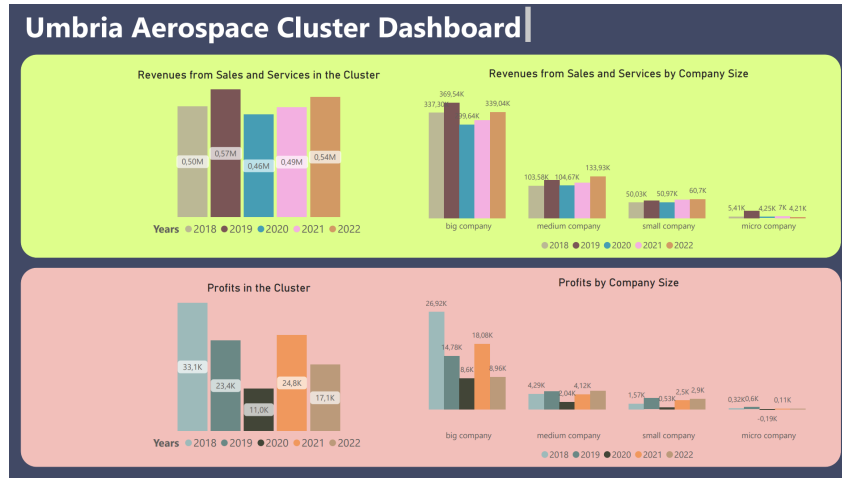
Figure 4.5: Revenues and Profits Dashboard on Power BI

4.1.3 Economic Indicators: ROS, ROE, ROA

Economic Indicators are fundamental tools used to assess the performance and financial stability of companies operating across different sectors. Among the multitude of financial indicators available, Return on Equity (ROE), Return on Assets (ROA), and Return on Sales (ROS) stand out as key metrics in assessing profitability and operational efficiency. Each indicator offers a unique perspective on the company's financial performance, shedding light on its ability to generate profits, utilize assets effectively, and manage sales revenue. Understanding these indicators and their implications is essential for stakeholders seeking to assess the overall financial strength and profitability of a company.

- Return on Sales: ROS is a metric that measures the profitability of the company's sales. It represents the ratio of net income to total sales revenue. ROS provides information on the percentage of profit generated from each unit of sale. A high ROS indicates that the company is generating a high profit relative to sales volume, while a low ROS may indicate low sales profitability.

- Return on Equity: ROE is a metric that measures the profitability of shareholders' investments in the company. It represents the ratio of the company's

net income to shareholders' equity. In other words, ROE indicates how profitable the company is in generating profit using the capital invested by shareholders. A high ROE is generally considered positive, indicating that the company is effectively using shareholders' capital to generate profits.

- Return on Assets: ROA is a metric that evaluates how efficiently the company uses its assets to generate profits. It represents the ratio of the company's net income to total assets. ROA provides insights into the return generated from all of the company's assets, regardless of the source of financing. A high ROA indicates that the company is efficient in utilizing its assets to generate profits.

The analysis of these indicators allows for providing an overview of the profitability of the Umbria Aerospace Cluster.

Return On Sales

ROS provides an overview of how different companies within the Cluster have been able to generate profits from their sales. Despite challenges, the ability of all categories of companies to generate a positive ROS demonstrates their resilience and ability to adapt to market conditions. Within the context of the Cluster, ROS performance exhibits different trends depending on the size of the company.

- Micro companies displayed a declining ROS from 2018 to 2020, reaching negative values, indicating losses generated from sales. However, there was a notable recovery in 2021, followed by a further increase in 2022, which bodes well for optimistic prospects in 2023.

- Small companies showed a positive ROS trend until 2019, followed by a sharp decline in 2020. Nevertheless, there was a remarkable recovery in 2021, with continued growth in 2022.

- Medium companies exhibited a trend similar to that of small enterprises, experiencing fluctuations in ROS over the years analyzed.

- Large companies, in contrast, experienced a substantial decline from 2018 to 2020, showing relatively stable trends from 2020 to 2022.

Return On Equity

Examining the performance of ROE within the Cluster reveals diverse dynamics depending on the size of the company.

- For micro companies, ROE exhibited a growing trend from 2018 to 2019, suggesting effective profit generation from invested capital. However, in 2020, it showed a negative ROE, indicating challenges in generating profits relative to the invested capital. This negative ROE might be attributed to factors such as increased expenses or decreased revenues during the COVID-19 pandemic. Nevertheless, there was a strong recovery in 2021 and 2022, indicating improved profitability and potential resilience to economic downturns.

- Small companies demonstrated a similar trend, with a less negative ROE in 2020 compared to micro companies. This suggests that small companies might have been better equipped to weather the challenges posed by the pandemic.

- Medium companies experienced a decline in ROE from 2018 to 20210, followed by a recovery in 2021 and 2022. Despite the initial decline, the subsequent increase in ROE indicates an improvement in profitability and efficiency in utilizing invested capital.

- Large companies witnessed a significant decline in ROE in 2019 compared to 2018 levels, touching nearly negative values. However, there was a trend of growth in 2020 and 2021, indicating potential improvements in profitability.

Nonetheless, there was a decline in 2022, suggesting ongoing challenges or fluctuations in performance.

Return On Assets

The analysis of the ROA trend provides further insights into the performance of companies across different size categories within the Cluster.

- Micro companies experienced a steady increase in ROA until 2019, reaching a peak in 2020, but then saw a slight downturn in 2021, maintaining stability in 2022. This decline may indicate challenges faced by this size-category companies in 2022, affecting their profitability from assets.

- Small companies exhibited an upward trend in ROA until 2019, followed by a contraction in 2020, with a modest recovery in 2021 and 2022. This suggests that while small companies encountered difficulties in 2020, they are gradually recovering.

- Medium companies displayed a ROA trend akin to small enterprises, although with a more pronounced recovery in 2022, signaling positive prospects for 2023.

- Large companies faced a declining ROA from 2018 to 2020, followed by a marginal uptick in 2021 and 2022. This pattern implies that large enterprises grappled with significant challenges in 2019 and 2020, impacting their profitability from assets.

Following the comprehensive analysis of the ROE, ROS, and ROA trends, the focus now shifts towards dissecting the composition of the Power BI dashboard illustrated in Figure 4.6. The dashboard comprises four distinct sections, each designed to provide insights into the financial performance of companies within the Cluster over time.

The first three sections prominently feature Clustered column charts, tailored to depict the trend of each financial indicator from 2018 to 2022. These charts are thoughtfully segmented based on the size classification of companies within the Cluster, facilitating a nuanced examination of how varying company sizes have navigated the financial landscape over the specified timeframe. In contrast, the fourth section of the dashboard shifts focus towards a summary view of the Cluster's financial performance in the latest available year, 2022. It has been done through the use of gauge visualization in Power BI. This snapshot serves as a crucial reference point, enabling to gauge the Cluster's performance in the most recent period and compare it against historical data to identify emerging trends or areas for improvement.
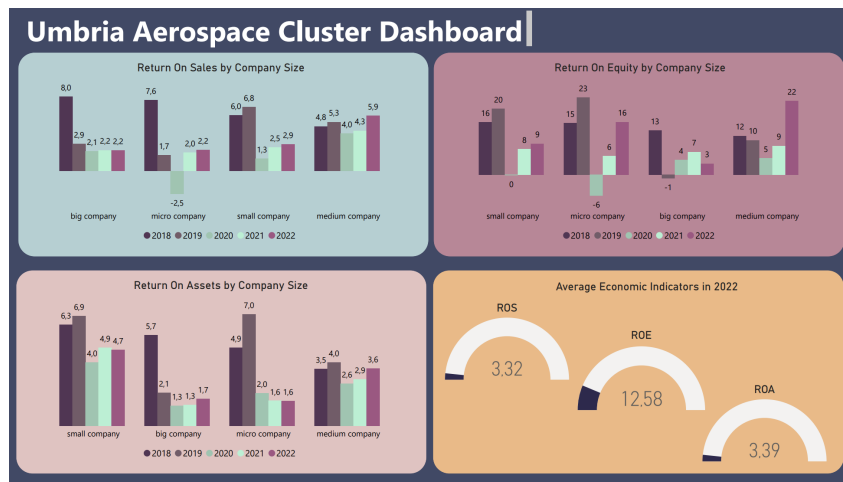
Figure 4.6: ROS, ROE and ROA Dashboard in Power BI

4.1.4 Impact on the Value Chain

Through comprehensive interviews conducted with each company within the Aerospace Cluster, it became feasible to quantify their involvement across different stages of the aerospace industry's value chain. Represented in Figure xx, a qualitative estimation of the impact, defined as the extent of significance and concentration that

companies dedicate to specific phases or segments of the value chain, has been provided.

The analysis highlights a substantial impact of cluster companies' activities particularly in the initial phases of the value chain, encompassing *Design and Development* as well as *Production of Parts and Components*. The pronounced engagement in the initial phases of the value chain suggests the ability to exert influence on the ultimate product or service in terms of innovation, quality, and cost-effectiveness. Such capability necessitates an in-depth understanding of aerospace technologies and requirements. Within the context of the Umbria Aerospace Cluster, the significant impact on the initial phases of the value chain underscores a concentration of specialized and technological expertise.

The *Testing* phase emerges as significantly impactful within the Cluster, with numerous specialized companies actively contributing and boasting specific expertise in this realm. This underscores the paramount importance placed on the quality and safety of aerospace products, critical factors in this industry. It is noteworthy that the majority of manufacturing companies within the Cluster possess advanced in-house testing facilities. These laboratories serve as integral components in the production process, ensuring product quality and adherence to technical specifications. This strong commitment to testing, encompassing both internal quality control and the development of specialized tools, underscores a deeply ingrained culture of quality within the Cluster.

Regarding the *System Integration* phase, cluster companies' impact appears moderately significant within the aerospace sector's value chain. Despite notable competencies in this domain within the Cluster, most companies seem to prioritize the production of individual components over their integration into finished products. This highlights a unique aspect of the Cluster, namely a prevalence of activities centered on component manufacturing rather than system integration. This observation suggests potential for further development of system integration

capabilities within the Cluster, aiming to establish a more integrated ecosystem capable of overseeing the entire aerospace production process.

cluster companies also demonstrate considerable activity in *Marketing and Sales* endeavors, reflecting the Cluster's effectiveness in positioning its products in the market and attracting customers. Many entities within the Cluster have established direct relationships with major players in the aerospace sector, qualifying them as Tier 1 suppliers. The active participation of cluster companies in international industry fairs is particularly notable, indicating a clear inclination to broaden their market reach, actively pursue new customers globally, and forge strategic relationships and partnerships. This proactive and market-driven approach by cluster companies underscores a strategic orientation toward customer needs and the delivery of high-quality solutions that align with customer expectations. Such endeavors contribute significantly to consolidating the Cluster's reputation and the associated companies in fostering enduring relationships with customers.

The dashboard illustrated in Figure 4.7 has been developed interactively using Power BI. The key element of the dashboard is the Radar Chart. It is noteworthy that the Radar Chart is not among Power BI's default visualization options but can be easily added by downloading it from the "Get more visuals" feature. This flexibility empowers users to customize their dashboards to suit their specific needs and leverage a wide range of available data visualization tools. The Radar Chart is a data visualization tool representing information in a spider web-like diagram, with different categories positioned along radial axes and values depicted as points connected by lines. In this specific context, the Radar Chart illustrates the phases of the aerospace industry value chain where cluster companies are concentrated, arranged clockwise from the most impacted to the least impacted. Adjacent to the Radar Chart, there is a filter pane allowing users to select a specific category of company size within the Cluster. This enables interactive exploration of how companies' concentration in certain value chain phases varies based on their size.
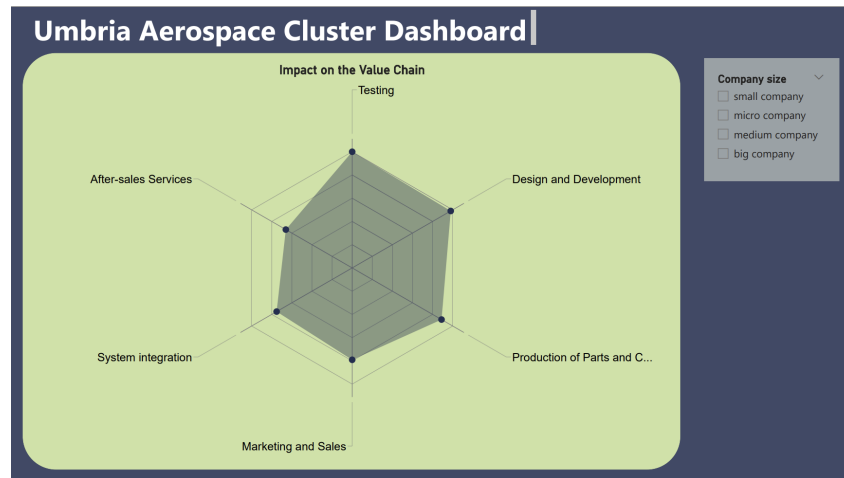
Figure 4.7: Impact on the Value Chain Dashboard in Power BI

4.1.5 Business Skills: Specialization Areas, Engineering Skills, Production Skills

In the context of assessing the skills of companies within the Cluster, a methodology based on the Aerospace Industry Competency Model has been applied. This model is internationally recognized and focuses on the specific skills needed in the aerospace sector. Using this model has allowed for a thorough and comprehensive analysis of the Cluster's competencies, providing detailed insights into the capabilities of the participating companies. The analysis primarily focuses on the "Industry-Wide Technical Competencies," which are unique to the aerospace industry and offer valuable insights into the technical and operational capabilities of the companies. These competencies are categorized into three main dimensions:

- Specialization areas represent the sectors in which cluster companies demonstrate high competence and consolidated experience, particularly distinguishing themselves for the quality and innovation of the products and services offered.

- Design competencies pertain to the companies' ability to design innovative and high-quality solutions in line with the most advanced standards of the

87

aerospace sector.

- Production competencies refer to the companies' capacity to implement technologies and production models typical of the aerospace sector, ensuring efficiency and quality.

Specialization Areas

The breakdown of expertise areas within the Umbria Aerospace Cluster presents a clear picture of its strengths in the industry, as is possibl to see in Figure 4.8.

The most significant field is that concerning *Industry Standards and Certifications*, reflecting the Cluster's emphasis on compliance with industry regulations. This emphasis is crucial for maintaining high standards and ensuring alignment with regulatory requirements. Next in importance is *Quality Assurance* indicating the Cluster's commitment to producing top-quality products and services, which is often seen as a key factor in setting it apart in the aerospace sector. The importance of the aeronautical sector highlights the significant involvement of many cluster companies in this particular segment of the industry. *Aeronautics* emerges as one of the Cluster's primary areas of specialization, demonstrating a wide range of skills possessed by participating companies in this field, particularly in the design, production, and maintenance of aircraft components and systems. While *Avionics* and *Training* areas show limited significance, some companies offer specialized expertise in these fields. Lastly, there is a noticeable gap in competencies related to *Propulsion*, *Aerodynamics*, and *Thermodynamics*.

Overall, these findings depict the Umbria Aerospace Cluster as possessing a diverse range of specialized skills, with particular emphasis on "*Standards and Certifications*, *Quality Assurance*, and *Aeronautics*.
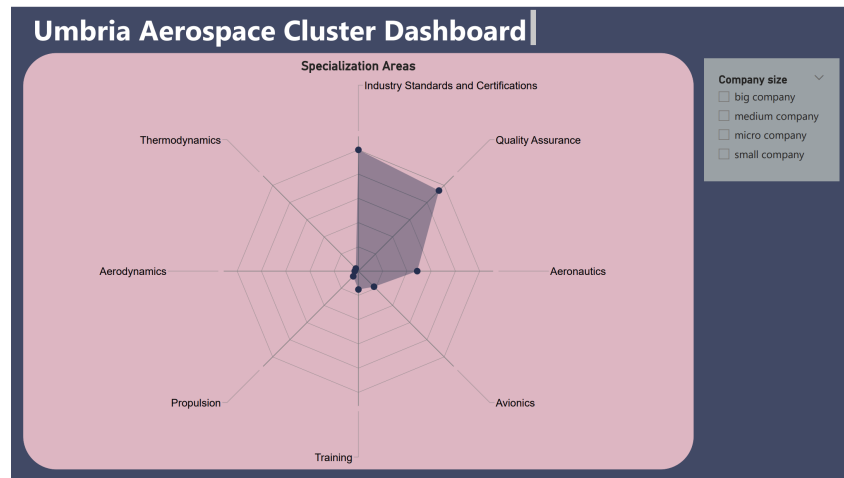
Figure 4.8: Specialization Areas Dashboard in Power BI

Engineering Skills

The Umbria Aerospace Cluster showcases a diverse array of technical proficiencies, particularly in fields such as *Mechanical Engineering, Electronical Engineering, Software Engineering, Systems Engineering, Production Process Engineering* as depicted in Figure 4.9. This blend of capabilities enables the Cluster to tackle a wide spectrum of technical hurdles and devise various innovative solutions tailored to the aerospace sector's needs. With the aerospace industry continuously evolving, the Cluster stands well-prepared to capitalize on its existing expertise and cultivate fresh competencies to address forthcoming challenges.

Mechanical engineering assumes a fundamental role within the Cluster, with firms possessing these skills making substantial contributions to the ecosystem by crafting and industrializing components and products indispensable to the aerospace domain. Equally crucial are the competencies in *Electronical Engineering, Software Engineering, Systems Engineering, Production Process Engineering* within the cluster. These disciplines intersect to furnish comprehensive, integrated solutions. *Systems Engineering* entails overseeing the design and management of intricate systems, ensuring seamless integration of all components. Meanwhile, *Electronic* and

89

Software Engineering focus on devising circuits, systems, and software essential across various aerospace domains, including flight control, navigation, and communication systems. *Production Process Engineering* forms the bedrock of the cluster, with numerous companies, particularly those involved in metal component processing and production, exhibiting advanced proficiency in this realm. These entities prioritize enhancing process efficiency, minimizing costs, boosting productivity, and upholding stringent quality standards.

Electrical and *Structural Engineering* also play vital roles within the Cluster. The first expertise aids in conceptualizing and developing aircraft electrical systems, encompassing power distribution and control mechanisms. In contrast, the second one concentrates on designing and analyzing aircraft structures to ensure structural integrity and safety.

It is noteworthy that certain areas like *Chemical Engineering, Optical Engineering*, and *Launch System Engineering* currently do not constitute primary sectors for cluster activities. However, this is not indicative of a lack of expertise but rather reflects a strategic concentration of capabilities on sectors deemed more pivotal for the cluster's operations. The absence of a pronounced presence in these specific domains aligns with the strategic direction embraced by cluster entities.
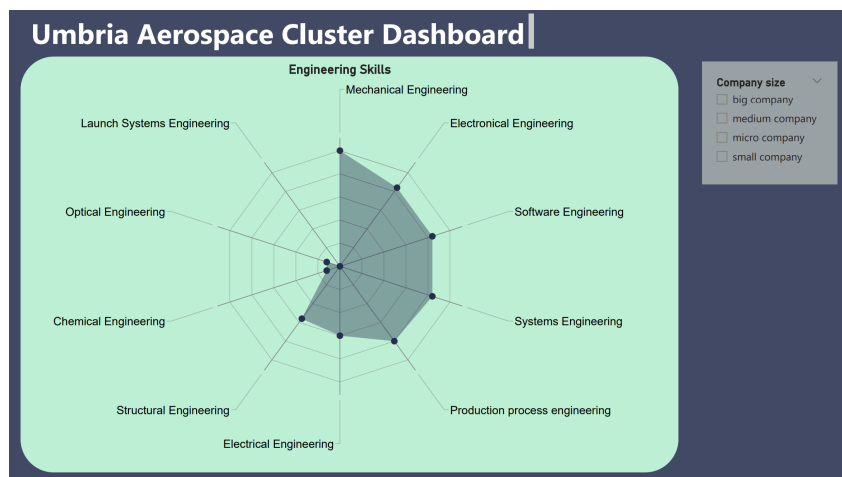
Figure 4.9: Engineering Skills Dashboard in Power BI

Production Skills

Through interviews conducted with companies within the Umbria Aerospace Cluster, a comprehensive and dynamically evolving landscape of technical and productive skills is revealed in Figure 4.10. In particular, the production of aeronautical components encompasses a wide range of procedures, ranging from component assembly to high-speed mechanical processing, from composite material production to electronic product manufacturing, as well as joining, welding, and assembly phases, non-destructive testing, and coating and special processes.

High-speed machining, crucial for the aerospace cluster, represents a technological strength thanks to the use of state-of-the-art machinery, including latest-generation CNC machines. This technology allows working on components of various sizes with extremely high precision. It is noteworthy that despite a substantial presence of companies engaged in mechanical component production, there is not significant overlap in activities and competitive dynamics remain contained. Instead, instances of synergy among these companies suggest a preference for collaborative and complementary engagements over direct competition. This collaborative spirit facilitates partnerships, knowledge exchange, technology sharing, support network establishment, and specialized workload distribution. Furthermore, it encourages diversification of solutions, enhances resilience against market fluctuations, and delivers a more comprehensive service to clients. The substantial deployment of CNC machines within the Cluster not only signifies the companies' commendable technological prowess but also underscores their adeptness at keeping abreast of technological advancements. These machines necessitate specialized skills for operation and maintenance, underscoring the Cluster's commitment to personnel training and the cultivation of advanced competencies. Another notable cluster strength lies in its capability to conduct *Non-destructive Testing*, a critical competency ensuring the quality and safety of aeronautical components. This encompasses a range of techniques aimed at identifying defects or malfunctions

91

without causing harm to the component itself.

A significant aspect lies in the expertise in *Electronics Material Processing.* This proficiency reflects the growing importance of electronic and digital technologies in the aerospace sector, spanning applications from flight control systems to communication systems. Several companies within the cluster specialize in designing, producing, and testing electronic products. They offer a diverse range of tailored solutions, maintaining strict standards of quality and reliability. Importantly, these companies' skills extend beyond aerospace, finding application in other industrial sectors such as automotive, consumer electronics, and industrial automation.

Another crucial aspect of the cluster's expertise lies in managing *Joining, Eelding, and Assembly Processes.* These techniques, essential for integrating various components into more complex structures, demand high levels of skill and precision. Cluster companies are proficient in executing a range of joining processes, employing both traditional methods and cutting-edge techniques.

Within the Cluster, the *Aerospace Component Assembly Procedures* is a fundamental expertise, crucial for transforming individual components into complete aerospace systems. These processes involve a series of highly specialized assembly techniques, requiring precision and detailed knowledge of the materials and components used. Some cluster companies are proficient in effectively managing the complexity of these processes.

Although not the focal point of the cluster's activities, skills related to *Special Processes* and *Coating* are still present. This diversity of expertise allows the Cluster to offer finished products, ready for assembly, which do not require further external processing. This expands the cluster's offerings, providing added value to customers and contributing to a more efficient and integrated aerospace value chain.

Despite these highly developed skills, there are areas where the cluster shows less presence, such as *Composites Manufacturing.* This could represent an opportunity for future development and expansion, given the increasing importance of composite
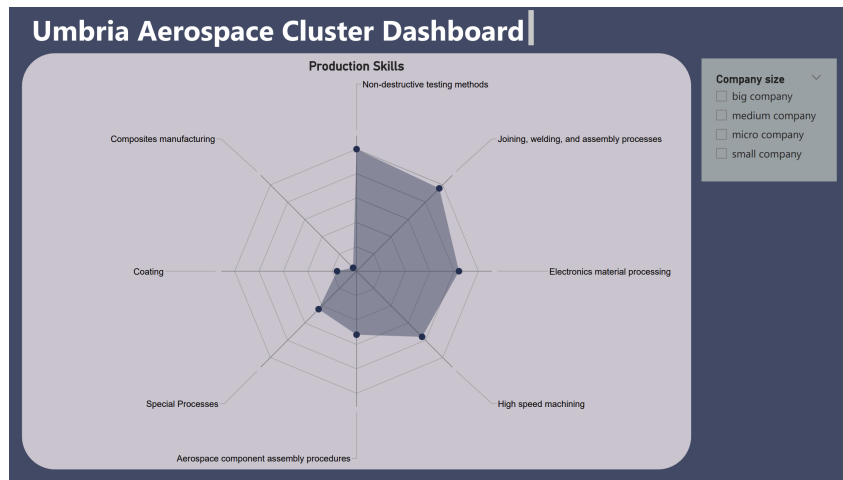
materials in the aerospace sector.

Figure 4.10: Production Skills Dashboard in Power BI

Chapter 5

Conclusions and Future Directions

The foregoing discussion has attempted to emphasize the idea that the integration of Business Intelligence into organizational frameworks represents a fundamental step in the modern view of organizational strategy. However, despite the associated costs, such as the one linked to the integration into existing company infrastructures, the benefits have proven to outweigh them. The ability to extract actionable insights from vast datasets and present it in interactive visual dashboards enables companies to make informed decisions and maintain competitiveness in dynamic modern industries.

Among all the software options for data visualization considered, Power BI has stood out for its exceptional capabilities and versatility. Its intuitive interface, robust functionality and tight integration with the Microsoft ecosystem make it a preferred choice for organizations worldwide. In the specific context of this research analysis, Power BI proved efficient in streamlining and automating the economic analysis of the Umbria Aerospace Cluster, making each company's self-assessment and comparison with competitor fast and effective through the messages conveyed by the dashboards.

Peering into the future, the integration of Artificial Intelligence could drive the

evolution of data visualization tools, representing a promising avenue for advancement. One potential future direction for data visualization tools is the incorporation of AI-driven algorithms for data analysis and interpretation. These algorithms could help identify patterns, trends, and anomalies in data more efficiently and accurately than traditional methods. Furthermore, AI can enhance the user experience by interacting with data using conversational commands, making data exploration more accessible to a broader audience.

The arguments above prove that Power BI's dynamic and interactive dashboards offer a superior alternative to traditional slide presentations powered by spreadsheets, providing the audience with real-time insights and a deeper understanding of complex datasets. For this reason, it is not ruled out that in the future there may be a partial or complete replacement of previous methods, making room for automated and efficient software like Power BI.

Bibliography

[1] Rhythm Bhavsar. Data visualization: Unveiling insights and enhancing data science. page 1, 2023.

[2] Power BI. Power bi dashboard. page website, 2024.

[3] Innovature BPO. Microsoft power bi: Architecture and features of this powerful data visualization tool. page 1, 2023.

[4] Beauty Chakraborty. What are 5v of big data. *Smart Edge*, 2022.

[5] Crossroad Elf. Business intelligence. 2022.

[6] Matteo Golfarelli et al. Beyond data warehousing: What is next in business intelligence? page 1, 2004.

[7] Mohammad Shariat et al. Conceptualizing business intelligence architecture. *School of Business and Industry Florida University*, page 1.

[8] Usama Fayyad et al. From data mining to knowledge discovery in databases. page 37, 1996.

[9] Enterprise DNA Experts. What is power query: A step-by-step guide for data analysis. page 1, 2023.

[10] Gartner. Magic quadrant. page 1, 2023.

[11] Gartner. Microsoft power bi vs tableau. page 1, 2024.

[12] Hero Group. Key characteristics of big data: Volume, variety, velocity, and more. page 1, 2023.

[13] IBM. Comparisons' graphs. page website, 2022.

[14] Krishan Kumar. Introduction to data modeling in power bi. page 1, 2023.

[15] D.A. Maisano E. Raguseo F. Franceschini L. Mastrogiacomo, F. Barravecchia. Mappatura e analisi dell'umbria aerospace cluster. page 7, 2023.

[16] Lorelei. Tableau vs power bi. 2023.

[17] Microsoft. Learn. page 1, 2024.

[18] Learn Microsoft. Live connection and directquery comparison. page 1, 2024.

[19] Tolulope Olowookere. Connection types in power bi. page 1, 2023.

[20] Adi Regev. Introducing: Power bi data prep with dataflows. page 1, 2018.

[21] Alessandro Rezzani. Business intelligence. page 26, 2017.

[22] Data Semantics. Power bi vs tableau – what is the difference and which one should you choose? page 1, 2023.

[23] Tableau. Explore the tableau exchange. page 1, 2024.

[24] Tableau. Tableau dashboard. 2024.

[25] Tableau. What is tableau? page 1, 2024.

[26] Tableau. What is vizql? page 1, 2024.

[27] Unknown. 5 software for manufacturing process: Streamlining solutions. page 1, 2024.

[28] PBI Visuals. Understand power bi desktop. page 1, 2024.

[29] Michael L. Waskom. statistical data visualization. page 1, 2021.

[30] Hatice Özşahan. 21 best data management software tools in 2024. page 1, 2024.

List of Figures

99