Research Article

Big Data Analytics in Environmental Sustainability Supply Chain Management

Akanksha Sharma

University of Essex

Abstract

This paper examines the integration of Big Data Analytics (BDA) in Environmental Sustainability Supply Chain Management (ESSCM) within the aviation and automotive industries, given their data-intensive nature. Using a multidisciplinary approach, the research employs RBV framework to understand how these sectors balance environmental responsibility with BDA implementations. The methodology includes a comprehensive literature review in environmental sustainability and BDA, along with case studies focused on prominent automotive and aviation industries. Incorporating text analytics, the paper examines the extent to which BDA utilisation enhances environmental sustainability within complex supply chain systems. The study is motivated by the increasing significance of sustainable practices in the modern business landscape, coupled with the growing influence of BDA in shaping decision-making processes. Recognizing the need for a comprehensive understanding of how BDA can contribute to ESSCM within the aviation and automotive industries, this research seeks to bridge the gap in existing literature and offer valuable insights for companies navigating the intersection of technology, environmental responsibility, and supply chain dynamics. By providing a unique approach to advancing the discourse on technology and environmental responsibility, the paper underscores its potential to guide these industries in analysing key areas of performance, such as in technological innovation and green initiatives; and contributes to the existing literature by providing insights for further in-depth exploration.

Keywords: big data analytics (BDA); environmental sustainability; supply chain management (SCM); text mining; resource-based view (RBV); automotive industry; airline industry

Date of Submission: 01.09.2025 **Date of Acceptance:** 10.12.2025

Acknowledgements

I express my gratitude towards my supervisor, Dr. Debabrata Ghosh, whose advice, unwavering support, and important insights have played a pivotal role in facilitating this research endeavour. The enormous impact of their expertise, patience, and encouragement on the progress of my paper cannot be overstated.

I extend my sincere appreciation for their mentorship, constructive feedback, and dedication that made this work possible. I would like to express my appreciation to Dr. Fay Aljibory for her valuable support, motivation, and contributions that have enhanced the scholarly pursuit of this research endeavour.

Secondly, I express gratitude to my family for their unwavering support and understanding throughout my academic endeavours. Their support has been my pillar of strength.

Finally, I would like to express my gratitude to those individuals who have made direct or indirect contributions to this paper

1. Introduction

1.1 Background

To fully comprehend the significance of this study, a comprehensive understanding of its subject matter and potential benefits is essential for organizations and individuals seeking valuable insights. The exploration of Big Data Analytics (BDA) and its applications in environmental sustainability within Supply Chain Management (SCM) establishes the foundation for addressing critical concerns. BDA, categorized into five Vs (value, volume, variety, veracity, and velocity), contributes to practical ideas, sustainable development goals, enhanced eco-effective performance, and competitive advantages (Tao et al., 2018). The implementation of two fundamental BDA viewpoints enhances operational excellence: the collection of significant data from internal and external contexts (Frank et al., 2019) and the integration of big data for real-time adjustments. In recent years, sustainable supply chain management (SSCM) has incorporated environmental, social, and economic objectives within a company's supply chain operations. As organizations aspire to achieve sustainability outcomes across environmental, social, and economic dimensions (Carter and Rogers, 2008; Seuring and Muller, 2008), this study focuses on SSCM's environmental dimension in the influential automotive and aviation industries, with the aim of elucidating challenges and unveiling discoveries.

The environment faces severe damage from natural and human-induced causes in the 20th century. Environmental degradation, driven by industrialization, economic expansion, and resource pursuit, gives rise to numerous issues threatening Earth's ecosystems. The combustion of fossil fuels, particularly oil, coal, and gas, contributes to climate change and greenhouse gas (GHG) emissions, resulting in global warming and its associated consequences (Das and Sharma,

DOI: https://doi.org/10.5526/esj.457

2023). Furthermore, the transportation sector plays a substantial role in GHG emissions and air pollution; elevated emissions impact air quality, expedite climate change, and necessitate sustainable transportation alternatives (Parpieva et al., 2024). Therefore, this study focuses on two major industries, automotive and aviation, which are increasingly compelled to address environmental concerns and adopt sustainable practices. The discussion on these transportation industries encompasses an overview of the environmental causes and effects in each sector, emphasizing technological advancements through BDA. For this study, major factors contributing to environmental concerns and BDA implementation are examined using text analytics.

1.2 Research Rationale

The proficient application of big data analysis can augment supply chain sustainability by providing insights into the environmental impact of operations (Wang et al., 2020). Companies can identify areas for improvement and reduce their environmental footprint by analyzing data on energy use, carbon emissions, waste creation, and other environmental issues. This study connects significant industries, providing a comprehensive understanding of global transportation businesses by considering various aspects of environmental hazards and BDA uses. However, a limited number of scholarly papers explore the significance of BDA and its technologies for insightful conclusions. For this section of the study, secondary source data, derived from company databases and scholarly papers, constructs a conceptual framework and conclusion covering topics such as fuel usage, emissions, logistics, material sources, technological transformations, advancements, technical adaptations, and product lifecycle information. The study assesses BDA implementation using text mining and analytics to evaluate supply chain environmental sustainability. Organizations commit to minimizing their environmental impact and making informed decisions by collecting and analysing vast volumes of data. Integrating big data (BD) and analytics into the supply chain is crucial for firms to measure and assess their real-time environmental impact, promoting sustainability.

1.3 Research Aim, Objective, and Question

1.3.1 Research Aim

The primary aim of this paper is to conduct a thorough examination of the application of BDA in the context of ESSCM. The overarching objective of this study is to investigate how companies utilize BDA in the supply chains of the aerospace and automotive industries to enhance environmental sustainability. The ultimate focus is on understanding how these industries employ BDA to identify concerns related to environmental sustainability while integrating the Resource-Based View (RBV) framework comprehensively. While applying this theoretical framework, the study aims to provide a distinctive approach by introducing text analytics methods to construct case studies and derive insights through visualization techniques. These insights are intended to be valuable for both companies and scholarly researchers in further exploring operational aspects.

DOI: https://doi.org/10.5526/esj.457

1.3.2 Research Objective

This research is based on the below-mentioned objectives:

1. To examine the integration of BDA and ESSCM in general and in automotive and aerospace companies' performance in the supply chain and environmental sustainability.
2. To evaluate the industry's effectiveness in using big data analysis or its techniques for the sustainable development of nature in the supply chain.
3. To employ big data's text mining analysis to build company-based case studies for examining a company's database (reports).
4. To discuss the correlation between BDA and the environmental sustainability of the two industries through their relationship and time series analysis.

1.3.3 Research Question

As this paper focuses on BDA and ESSCM, considering the transportation industry, the main question that this study addresses is: **To what extent are companies implementing BDA and addressing environmental concerns?**

To enhance the study, this paper will also discuss:

RQ1 - To what extent is BDA being used and talked about globally and in the Scholarly world?

RQ2 - How is big data and BDA being helpful to improve the environmental sustainability of supply chains?

RQ3 - How could Text Analytics be useful to build case studies and evaluate results related to the company's initiative towards environmental concerns?

RQ4 - What challenges, limitations, and concerns are associated with the utilization of BDA and text analytics?

The research question involves two key industries – Automotive and Aviation, which broadens the scope of the study and enables a thorough examination of the use of BDA in various supply chain situations.

1.4 Methodology

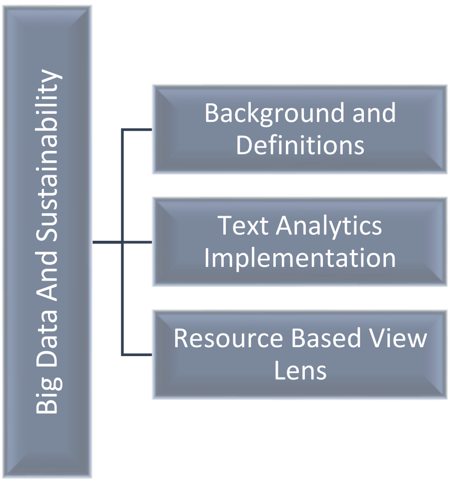
The utilisation of BDA has become significant in industries' efforts to reduce their carbon footprint and ensure environmental responsibility. Data analytics has the potential to transform SCM in the aviation and automotive industries, improving efficiency, environmental impact, and sustainability. BDA enables the processing and analysis of large data sets. The advancement of technology and the increasing complexity of supply chains have led to the accumulation of substantial data within businesses. This study gained a comprehensive understanding of BDA

DOI: https://doi.org/10.5526/esj.457

technology, while implementing text analytics, and derive insights through predictive analysis, time series analysis, correlation analysis, and data visualisation using R programming, python, and Tableau. These methods generated new opportunities and provide future researchers with a different perspective. The alignment of strategies between the objectives and research questions is crucial for understanding the role of BDA in promoting environmental sustainability in the automotive and aviation industries. This approach ensures a seamless progression in this study as they explore an intricate network of inquiries, goals, and practical implications. This scholarly endeavour concludes with a thorough and insightful investigation that improves intellectual engagement. The rest of the paper is organized as follows: Section 2 presents the Literature Review, Section 3 the Methodology of the study, Section 4 presents Finding and Analysis through numerical values and data visualisation, Section 5 is about Discussion based on section 4, and final Section 6 discusses Limitations of the paper.

2. Literature Review

BDA has gained considerable popularity due to its diverse applications, but a precise definition remains elusive. BDA refers to technologies and architectures designed to extract economic value from large amounts of data, achieved by rapidly capturing, analysing, and discovering data patterns (Mikalef et al., 2018). Companies employ innovative technological methods for enhanced information processing, enabling better decision-making and a competitive advantage (Hofman, 2018; Gunasekaran et al., 2017; Rialti et al., 2019). Industries like aviation and automotive face increasing pressure to address environmental concerns and embrace sustainable practices, prioritising environmental sustainability in their corporate strategies. Industrialists recognise that considering only cost, quality, and productivity is no longer sufficient for competitiveness (Sureeyatanapas and Yang, 2021). Therefore, this literature review investigates the intricate relationship between BDA and ESSCM within the aviation and automotive industries. It aims to clarify the implementation of BDA in ESSCM and critically evaluate whether BDA aligns with and promotes environmental sustainability in these industries through the use of resource-based view lenses while applying text analytics.



DOI: https://doi.org/10.5526/esj.457

Figure 1.1: Literature Review Model

Source: Developed by the researcher

2.1 Definition of BDA

In adherence to Moore's Law, wherein the number of transistors on microchips doubles roughly every two years (Tardi, 2020), the Big Data (BD) paradigm has continually evolved. The term BD, initially used by John R. Mashey, predates the Fourth Industrial Revolution (Bag et al., 2020). BD has expanded from the 3Vs to as many as 8Vs, leading to confusion. Doug Laney introduced the foundational 3Vs—Volume, Velocity, and Variety (Lanely, 2001). Later, “Veracity” and “Value” were added (Motau & Boursier, 2016; Singh & Marr, 2015). Some researchers extended it further, but this study considers five: volume, velocity, variety, veracity, and value (Dahdouh, 2020).

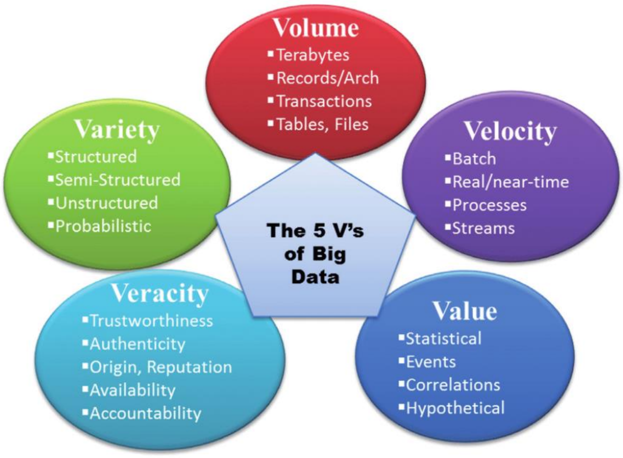


Figure 2.1.1: 5 Characteristics of Big Data

Source: (Dahdouh, 2020)

BD is better defined by its analytical processes rather than its size (Agaba, 2014). BDA involves gathering, arranging, and analysing extensive, continuously generated data to uncover insights for decision-making. However, organisations face challenges such as unreliable sources, data privacy issues, and the need for both technical expertise and human skills (Wang et al., 2019; Hazen, 2014). Advanced analytics methods such as data mining, predictive analytics, and text mining play a crucial role (Bank, 2022). Large amounts of unstructured data complicate machine learning, requiring lengthy data processing (Mergel, 2019; Muller, 2018; Nagar, 2021; Nasiri, 2020; Zhou, 2018).

2.2 ESSCM

Environmental sustainability entails the judicious utilisation of Earth's resources to meet current demands while safeguarding future generations (Golicic, 2013). SCM strategically coordinates diverse facets within a firm and across the supply chain to enhance performance (Mentzer et al.,

DOI: https://doi.org/10.5526/esj.457

2001). Environmental supply chain practices encompass eco-conscious endeavours throughout a product's lifecycle. Acquaye (2014) stresses the need for firms to adopt practices that are economically, environmentally, and socially sustainable (Schaltegger et al., 2008).

In sustainability, management strategies incorporate assessments of product life cycles and their impact on carbon reduction (Acquaye et al., 2011; Koh et al., 2013; Weber & Peters, 2009). Factors such as GHG emissions, CO2 levels, and waste management are critical. Ritchie's (2020) sector-specific analysis emphasises the complexity of global emissions and the need for holistic decarbonization approaches.

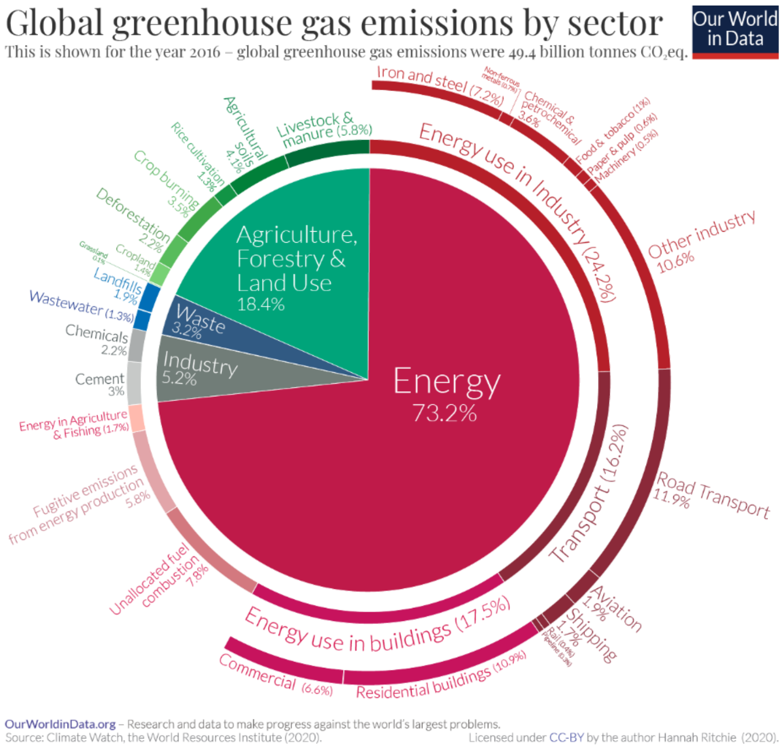


Figure 2.2.1: Global Greenhouse Gas Emissions by Sector Chart

Source: Ritchie (2020)

The WEF (2023) highlights that the US and EU account for 37% of cumulative global GHG emissions since 1850. Intra-organizational models focus on ecological raw materials, waste reduction, and energy conservation (Rao, 2002). Zhu, Crotty, and Sarkis (2008) compared ESSCM practices in the automotive industry in China and the UK. Mollenkopf et al. (2010) classified energy efficiency initiatives and examined adoption barriers.

2.3 Integrated Literature on Correlation of BDA and ESSCM:

The adoption of digital tools enhances business capabilities. BDA enables effective management of vast data for competitive advantage (Pettit et al., 2019). It strengthens SCM by improving production efficiency and reducing delays between manufacturers and users. Sensors, trackers,

DOI: https://doi.org/10.5526/esj.457

cloud computing, and warehouse digitization facilitate real-time SCM. Using descriptive, predictive, and prescriptive analytics, BDA improves purchasing and supply management (Min, 2010; Tumpa et al., 2019). BDA impacts economic, social, ethical, and legal aspects (Shubbak, 2013).

Economically, BDA supports sustainable firm performance and broader growth by enabling more informed, efficient decision-making, while also introducing a digital divide where unequal access to analytical capabilities deepens existing economic inequalities (Ertz et al., 2024; Cuquet and Fensel, 2018; Custers et al., 2017).

Socially, BDA can enhance public services and citizen engagement but simultaneously raises concerns about discrimination and privacy infringements, as data-driven systems may reproduce or intensify underlying biases (Cuquet et al., 2017; Custers et al., 2017).

Ethically, debates focus on privacy, autonomy, accountability, and the need for transparent data practices to uphold public trust and support responsible data governance (Custers et al., 2017; Guan & Zhou, 2017).

Legally, BDA challenges existing regulatory frameworks on data protection, intellectual property, accountability, and human rights, underlining the urgency of updating legal standards to protect individuals in an increasingly data-driven environment (Guan and Zhou, 2017; Cuquet et al., 2017; Custers et al., 2017).

Robust models are vital for scenario forecasting, influencing environmental, social, and economic supply chain performance (Jeble & Dubey, 2017; Rao, 2002). Combining diverse data types enhances SCM performance (Zhong et al., 2016; Fernando et al., 2018). BDA allows environmentally informed decisions, though research lacks clarity on how it directly affects green supply chain integration (Song et al., 2017).

2.4 BDA and ESSCM in Automotive Industry

Data shapes vehicle development, manufacturing, services, and connected technologies (Luckow, 2015). The industry's digitalization, driven by connected devices, amplifies demand for data analytics. Studies emphasise BDA's role in cost reduction, consumer behaviour, corporate environmental management, and barrier identification (Zhang et al., 2017; Beier, 2022).

BDA improves SCM through real-time data collection and predictive analysis (Giacosa et al., 2021), enhancing logistics, reducing waste, and improving energy efficiency (Chung, 2021). Big data warehouses support integration of diverse data sources (Silva, 2021). ML and BDA aid in optimising logistics and transportation (Zhu, 2021). BDA balances efficiency and environmental goals (Singh, 2022; Giacosa, 2022).

DOI: https://doi.org/10.5526/esj.457

Use cases include traffic prediction, autonomous driving, telematics data, and data lake architectures (Dremel, 2018). Structural challenges remain, but BDA contributes significantly to addressing environmental concerns in the automotive sector.

2.5 BDA and ESSCM in Airline Industry

The airline industry employs BDA to enhance ESSCM practices under environmental pressures (Savchenko, 2023). IoT applications improve operations, communication, and sustainability. Technologies like BD, additive manufacturing, and blockchain support greener SCM (Ramirez-Peña, 2020).

BDA improves fuel efficiency, reduces emissions, and streamlines maintenance, saving costs while supporting environmental responsibility (Aydin, 2019). Motivations include regulations, safety, efficiency, and sustainability (Burmester, 2018). It also drives innovation in biofuels and electric propulsion.

BD strengthens aviation safety and predictive models (Barak & Dahooei, 2018; Ni et al., 2019; Douglas, 2014). It also improves service levels (Cao et al., 2015; Kim & Shin, 2016; Park & Pan, 2018). However, barriers include lack of expertise, high costs, skills shortages, and privacy concerns (Hashem et al., 2015; Izzo, 2019; Mohamed, 2021). Despite challenges, technological advancement is crucial for addressing aviation's environmental concerns.

2.6 Text Mining and Text Analytics

This section employs text mining and analytics to extract insights from textual data such as social media, corporate documents, and news articles (Mohamed, 2021). Text mining extracts valuable insights from large datasets and is particularly relevant as much data exists in text form (Sánchez, 2008).

Text mining is widely used in analysing literature on SCM (Mergel, 2019; Thomas, 2015; Park, 2018). Tao (2020) outlines steps such as cleaning, reduction, association analysis, classification, and clustering, supported by metadata for context. Visual representations, such as time series and correlation analysis, was an aid for this study (Tao et al., 2020).

DOI: https://doi.org/10.5526/esj.457

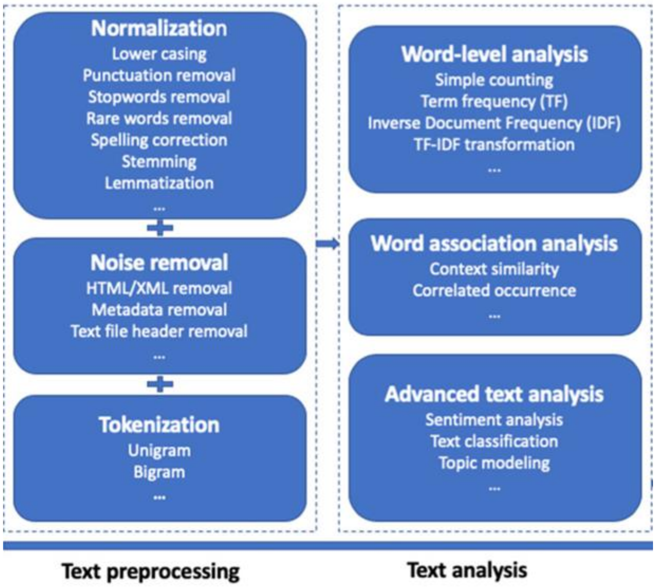


Figure 2.6.1: Text Analysis Framework

Source: Tao, Yang, and Feng (2020)

2.7 Theoretical Lens

Various theories support research on BDA and sustainability, including the Triple Bottom Line, panopticon, and Natural Resource-Based View (Seele, 2016; Mageto, 2021; Khatib, 2022). However, the Resource-Based View (RBV) is most relevant, explaining how BDA contributes to competitive advantage (Wamba et al., 2017). RBV is widely applied in related research (Dubey & Singh, 2019; Zhang, 2022; Vitari, 2020).

This study combines RBV with text analytics to assess how firms exploit environmental resources. RBV suggests competitive advantage can stem from deriving insights through text analytics to improve decision-making and innovation (Lubis, 2022).

DOI: https://doi.org/10.5526/esj.457

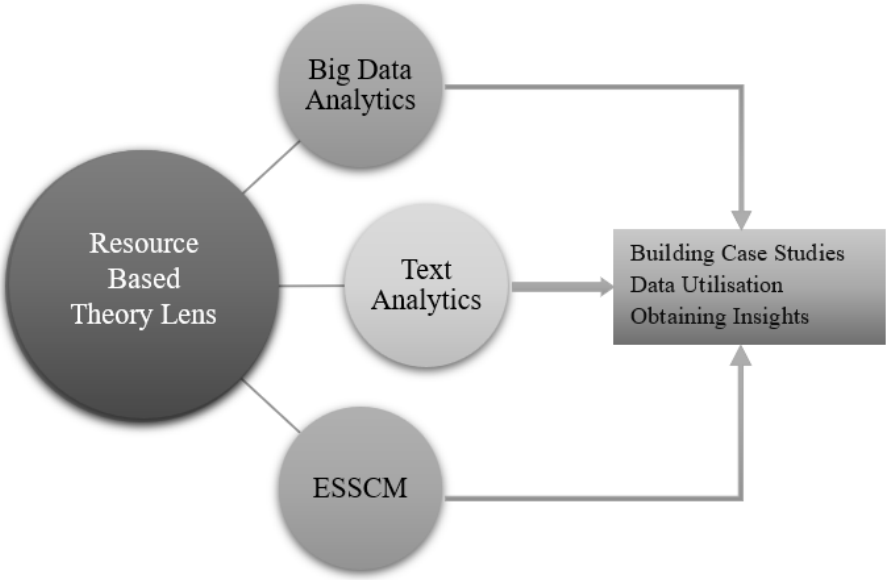


Figure 2.7.1: Conceptual Framework

Source: Developed by the researcher

2.8 Methodological Approaches of Previous Researchers and the Current Study

Studies employ a variety of methodologies to examine BDA and ESSCM. For example, Bag et al. (2020) applied a hybrid BWM and DEMATEL method to explore BDA-driven sustainable supply chains, while Dubey et al. (2017) examined big data and predictive analytics adoption through survey-based methods. Zhao et al. (2017) used empirical case studies to analyse BDA impacts in Chinese manufacturing. In another approach, Sharma and Panwar (2020) studied the Green IoT framework to reduce energy consumption and pollution in supply chains, while Yousefi and Tosarkani (2022) applied cognitive map modelling to evaluate risk management strategies.

Text mining has also been applied, such as Meyer et al. (2021) who investigated supply chain resilience during the pandemic using sentiment analysis. Machine learning and regression-based models have been used to forecast sustainability outcomes (Asha et al., 2022; Sathyan et al., 2020). DEA models and maturity frameworks have been adopted to measure airline supply chain efficiency (Rachman & Arviansysh, 2019; Hausladen & Schosser, 2020).

This section uniquely combines BDA, ESSCM, and text analytics within the RBV framework, positioning it as distinct from prior research.

Drawing from these arguments, the following hypotheses are proposed:

- **H0**: No significant correlation exists between the use of BDA tools and environmental sustainability in the automotive and airline industries.

DOI: https://doi.org/10.5526/esj.457

- **H1**: A significant positive correlation exists between the use of BDA tools and environmental sustainability in the automotive and airline industries.

2.9 Summary

This study identified the implications of BDA for environmentally sustainable supply chain management, highlighting features, trends, and challenges. BDA assists firms in responding to increasing data demands and competitive pressures, though challenges remain in data storage, complexity, adaptability, privacy, and return on investment (Goyal et al., 2020). Cultural barriers, skills shortages, and high costs further hinder adoption, yet analytics implementation is becoming increasingly inevitable.

The literature reflects diverse approaches across industries, with applications ranging from logistics optimization to carbon footprint reduction. While numerous studies exist on BDA in automotive industries, research on BDA and ESSCM in the airline industry is relatively limited, representing a critical gap. Thus, this section of the study investigates the roles of BDA and ESSCM in both aviation and automotive industries, supported by RBV and text analytics, contributing to a deeper understanding of digital sustainability transitions.

DOI: https://doi.org/10.5526/esj.457

3. Methodology

3.1 Research Philosophy, Approach, and Design

This study adopts a positivist research philosophy, which emphasises empirical observation and objective measurement. Positivism is appropriate because the research focuses on quantifiable data extracted from corporate disclosures rather than subjective interpretation (Phair & Warren, 2021).

A deductive approach was employed, moving from established theories such as the RBV towards hypothesis testing. The aim was not to generate new theory but to examine whether patterns in corporate reporting confirm or challenge existing claims about the role of BDA in sustainable supply chain management (Snieder & Larner, 2009).

The research design is quantitative and based on secondary data. Annual and sustainability reports (2019–2022) were analysed using text mining and keyword frequency methods. These documents were selected because they are systematically prepared, publicly available, and comparable across firms and industries.

The study applies a cross-sectional time horizon, focusing on a defined period of four years. While this does not capture long-term trajectories, it allows for comparison across firms and industries during a period of heightened environmental and digitalisation pressures.

Finally, the study's techniques include text analytics, keyword frequency analysis, correlation testing, and data visualisation. These methods enable transparent and replicable analysis of how BDA and environmental sustainability are represented in corporate disclosures, providing a basis for evaluating their alignment under the RBV framework.

3.2 Company Selection Process

The selection of companies in the airline and automotive sectors is a crucial aspect of the methodology employed in this paper. The following criteria were used to select companies for this study:

1. The company must be a global leader in the airline or automotive industry.
2. The company must have a strong commitment to sustainability.
3. The company must have implemented BDA in its operations.
4. The company must have publicly available annual and sustainability reports.

This reflects the rigorous approach of the study in this research. While selecting the companies, a list of the top 10 companies in both the automotive and airline industries was considered. From this list, Delta Airlines, United Airlines, Emirates, Tesla, Toyota, and Volkswagen were selected for further analysis. These companies were chosen due to various principal factors that

DOI: https://doi.org/10.5526/esj.457

collectively enhance the comprehensiveness and importance of this research, which are discussed in sections 3.2.1 and 3.2.2 for both industries.

3.2.1 Automotive Companies Selection

1. ***Technological Innovation:*** Innovation supports sustainable development through redesigning products, processes, and structures (OECD, 2009). Tesla, Toyota, and Volkswagen were selected for their varied approaches to innovation. Tesla ranked as the most valuable automotive brand in 2023, Toyota is the world's largest automaker by production volume, and Volkswagen is second (Statista, 2023).

2. ***Sustainability Initiatives:*** Each company pursues distinct sustainability strategies. Toyota prioritises hybrid technology and targets net carbon neutrality by 2050 (Toyota Motor Corporation, 2023). Tesla emphasises electric vehicles and aims to accelerate the transition to sustainable energy (Tesla, 2023). Volkswagen focuses on emission reduction and electric mobility, with a target of net-zero carbon by 2050 (Volkswagen Group, 2023).

3.2.2 Airline Companies Selection

1. ***Global Presence:*** Delta, United, and Emirates were chosen for their strong international operations. Delta is the largest airline by revenue, United ranks third, and Emirates leads in international passenger kilometres (Statista, 2023). Their inclusion provides a global view of BDA and ESSCM adoption.

2. ***Operational Diversity:*** Delta and United are legacy carriers with extensive domestic and international networks, while Emirates specialises in long-haul routes. Their diversity enables a comparative analysis of BDA across different operational models. Sustainability goals include net-zero emissions by 2050 for Delta and United, and a 50% reduction in emissions by 2050 for Emirates (Delta Airlines, 2023; United Airlines, 2023; Emirates, 2023).

3.3 Data collection

This study relies on secondary data from annual and sustainability reports (2019–2022) of Volkswagen, Tesla, Toyota, Delta, United, and Emirates. These reports provide information on sustainability initiatives, CSR activities, digitalisation, and growth patterns. Although reports do not explicitly detail BDA, they provide consistent insights into sustainability and digitalisation practices.

The collected data undergoes preprocessing and refinement to ensure accuracy and alignment with research objectives. The study applies below quantitative techniques:

1. **Text Analytics**: Automatic extraction of patterns and keyword frequencies from unstructured text (Anandarajan, Hill, and Nolan, 2018).

DOI: https://doi.org/10.5526/esj.457

2. **Correlation Analysis**: Measures the strength and direction of relationships between BDA and sustainability variables.

3. **Data Visualisation**: Word clouds and bar charts illustrate connections between BDA and sustainability keywords.

This approach facilitates the evaluation of technological expansion and environmental responsibility in the automotive and airline industries.

Table 3.3.1: Variables Used in Analysis

Variable Type	**Sub-Themes**	**Operationalisation**	**Source**
Independent Variables (BDA)	Data & Analytics, Technology & Software, AI & ML, Cybersecurity, Robotics, etc. (10 sub-themes)	Frequency counts of related keywords in company reports (2019–2022)	Annual & sustainability reports
Dependent Variables (ESSCM)	Sustainability & Conservation, Energy & Power, Supply Chain & Procurement, Waste & Recycling, etc. (13 sub-themes)	Frequency counts of related keywords in company reports (2019–2022)	Annual & sustainability reports
Control / Contextual	Industry Type (Automotive vs Airline)	Sector categorisation of companies	Statista / Company reports

Source: Developed by the researcher.

DOI: https://doi.org/10.5526/esj.457

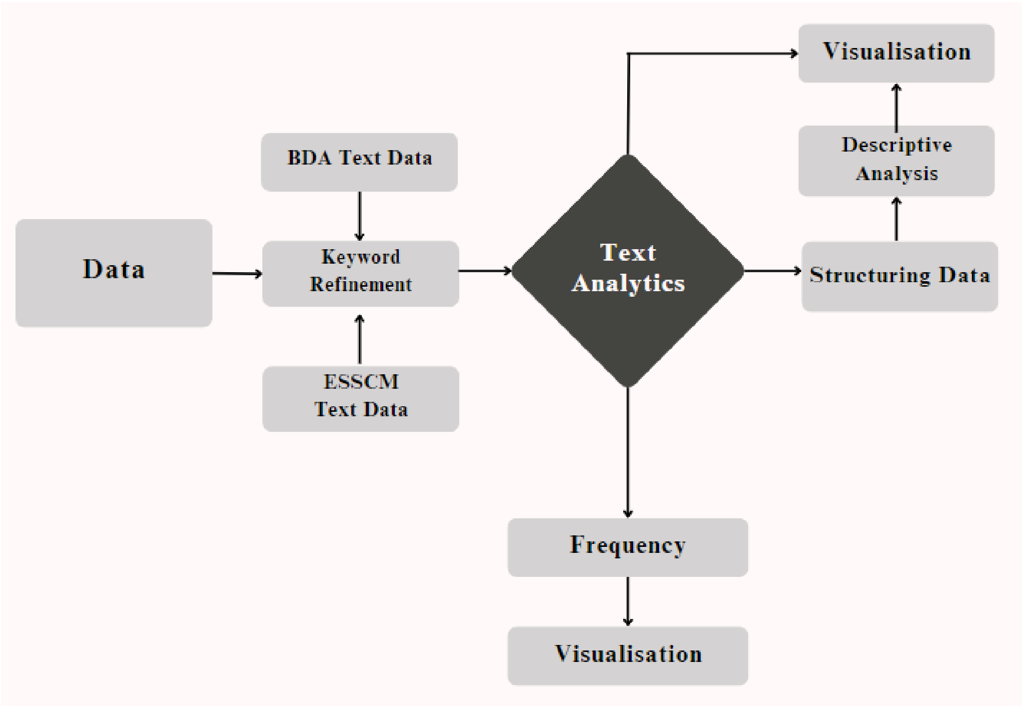


Figure 3.1: Process Flow Chart Framework

Source: Developed by the researcher.

4. Findings and Analysis

This section presents the findings and analysis of the study, examining the relationship between BDA and ESSCM across two industries: automotive (Volkswagen, Tesla, Toyota) and airlines (Delta, Emirates, United). The objective is to demonstrate how companies report on digitalisation and sustainability practices, and whether patterns of BDA adoption correlate positively with environmental sustainability efforts.

The section proceeds in a layered manner. First, the text mining procedure is outlined, describing how unstructured corporate reports were refined into structured datasets of keyword frequencies. Second, Wordclouds provide a visual introduction to reporting emphasis. Third, heatmaps quantify keyword frequencies by sub-theme, followed by radar charts to compare industry strengths across domains. Fourth, stacked bar charts show how domain contributions shift over time, and trend lines examine year-on-year dynamics for each company. Finally, scatterplots test correlations across industries, confirming or rejecting the research hypothesis. Company-specific results are available in the appendices, while this section focuses on industry-level insights.

This sequence is guided by the Findings and Analysis Framework (Figure 4.1.1), which divides the section into three stages: refinement, analysis, and findings. The framework demonstrates how the study moves systematically from data preparation, to analysis, and finally to interpretation.

DOI: https://doi.org/10.5526/esj.457

This layeed approach follows recommendations from Hastig and Sodhi (2020) and Tavana and Shaabani (2022), who emphasise the value of combining text mining with structured visualisation methods to analyse sustainability reporting.

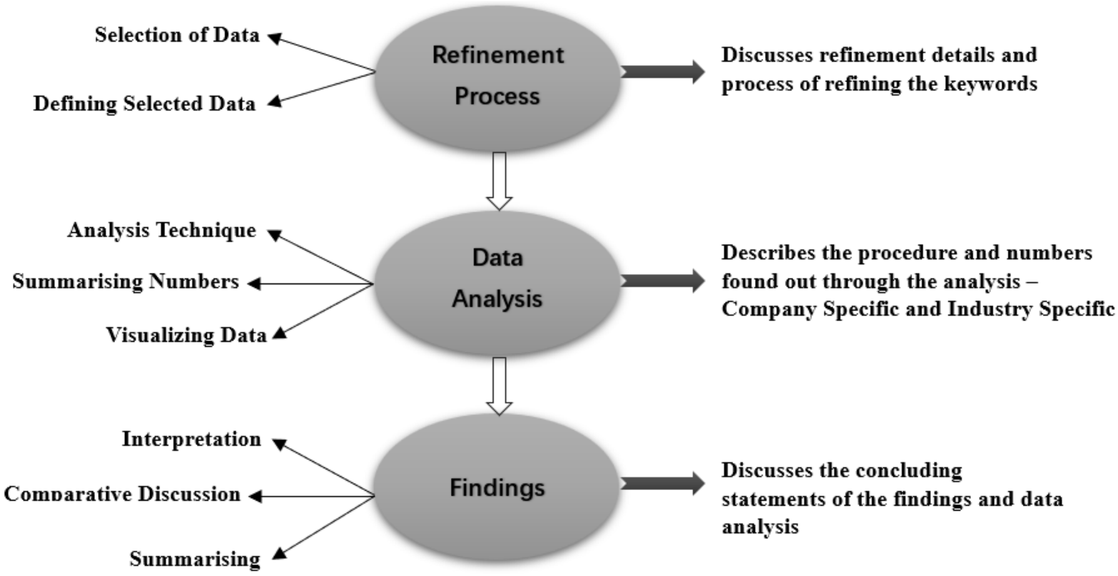


Figure 4.1: Findings and Analysis Framework

Source: Developed by the Researcher

4.1 Clarification on Frequency Counts

It must be noted that while annual and sustainability reports were the source of data, this study did not qualitatively interpret how each keyword was discussed. For example, when "sustainability" appeared, the analysis did not distinguish whether it referred to emissions, social responsibility, or renewable energy. Instead, the study relied strictly on frequency counts, transforming them into charts such as heatmaps and radar diagrams. This approach allowed statistical analysis while avoiding subjective interpretation.

4.2 Text Mining Procedure

Before presenting visualisations, it is essential to explain how the data was generated. The raw material comprised annual and sustainability reports from 2019–2022 for the six selected companies. These reports are extensive, often covering topics beyond technology and sustainability. The challenge was to extract only the keywords relevant to BDA and ESSCM sub-themes, ensuring comparability across firms and industries.

The text mining procedure followed three structured steps:

1. **Keyword Extraction and Counting**

DOI: https://doi.org/10.5526/esj.457

- Annual and sustainability reports were merged, converted from PDF into text, and uploaded into Google Spreadsheets.
- A sheet named *"Data"* contained the raw text, while a second sheet *"Output"* listed the defined keywords vertically.
- The COUNTIF formula was applied to calculate the number of times each keyword appeared. This approach is consistent with Manjunath and Viswanathan (2019), who identify spreadsheets as a low-cost but effective means for text-based frequency analysis, and with Tao et al. (2020), who note the role of text analytics in structuring unstructured corporate data."
- This method was chosen after testing alternatives: initial attempts with R Studio and libraries such as *pdftools* were less precise, while Google Spreadsheets offered accuracy and transparency. This approach has been recommended by Manjunath and Viswanathan (2019) as an efficient, low-cost solution for text-based frequency analysis.

2. **Data Preparation and Structuring**

- Raw keyword counts were unstructured and therefore reformatted in Microsoft Excel.
- Separate sheets were created for BDA and ESSCM, organised by sub-themes (10 for BDA and 13 for ESSCM).
- Company-specific totals were computed, followed by aggregation at the industry level.

3. **Final Dataset for Analysis**

- The final dataset allowed keyword frequencies to be examined at three levels: company-specific, sector-specific, and industry-level.
- This dataset provided the numerical foundation for all subsequent visualisations: Wordclouds, heatmaps, radar charts, stacked bar charts, trend lines, and correlation scatterplots.

This process ensured that the results are replicable and transparent, transforming unstructured text into structured frequency tables.

Refer to the table for the number of keywords considered.

Table 4.2.1: Numbers

DOI: https://doi.org/10.5526/esj.457

Number of Themes	2
Total Number of Sub-themes	23
Number of BDA Sub-themes	10
Number of ESSCM Sub-themes	13
Number of BDA Keywords	63
Number of ESSCM Keywords	63

Table 4.2.2: BDA Underlying Keywords

Keywords

Theme: BDA

Sub-theme:	**Underlying Keywords:**
Artificial Intelligence and Machine Learning	ML (with spaces behind and ahead), Algorithm, Analytical, Analytics, Artificial, Machine, Machine Learning
Data and Analytics	Analysis, Application, Big Data, Data, Data Management, Data Mining, Database, Data-Driven, dataset, dataSets, Predict, Prediction, Predictive, Processing, Information, Mining, Mathematic, Maths
Technology and Software	AWS (with spaces behind and ahead), Blockchain, Cloud, Computer, Digital, Electronic, Technology, Tool, Virtual, Software, Java, Program, Programming
Cybersecurity and Security	Cyber, Cybersecurity, Security, Securities, Surveillance
Business and Strategy	Business Intelligence, Strategy, Innovation
Ethics and Privacy	Ethical, Privacy
Robotics and Automation	Robot, Robotics
Technical and Technique	Technical, Techniques, Method

DOI: https://doi.org/10.5526/esj.457

Integration and Connectivity	Integration
Visualization and Data Volume	Visualisation, Visual, Volume

Source: Developed by the researcher.

Table 4.2.3: ESSCM Underlying Keywords

Keywords

Theme: ESSCM

Sub-theme:	**Underlying Keywords:**
Environmental Concerns	Carbon, CO2, Combustion, Emission, Emissions, Greenhouse, Pollution
Sustainability and Conservation	Conservation, Eco, Eco Friendly, Eco-Friendly, Environmental, Sustainable, Renewable, Renewal, Sustainability
Energy and Power	Electric, Energy, Power plant, Renewable, Power
Nature and Environment	Earth, Natural, Nature, Planet, Plant
Transport and Logistics	Transport, Logistics
Social and Corporate Responsibility	Corporate, Social, EPA, ESG
Supply Chain and Procurement	Procurement, Supplier, Suppliers, Supply, Supply Chain, SCM, SSCM
Waste and Recycling	Recycle, Waste, Wastage
Harm and Hazard	Harm, Harmful, Hazard
Resource Management	Material, Resources
Agriculture and Farming	Farmer, Farmers
Gas and Nitrogen	Gas, Nitrogen
Production and Process	Production, Process

Source: Developed by the researcher.

The finalised frequency data is as below:

DOI: https://doi.org/10.5526/esj.457

	Frequency		
Environmental Sustainability Supply Chain	Delta 2019-2022	Emirates 2019-2022	United 2019-2022
Environmental Concerns	6859	1212	1934
Carbon	1127	108	252
CO2	871	215	241
Combustion	89	0	4
Emission	2346	439	722
Emissions	2126	415	671
Greenhouse	186	34	37
Pollution	41	0	0
Pollut	73	1	7
Sustainability and Conservation	4746	1535	1967
Conservation	41	44	34
Eco	0	1	0
EcoFriendly	0	1	0
Eco-Friendly	3	6	5
Environmental	1676	343	426
Sustainable	615	403	796
Renewable	333	69	125
Renewal	74	5	14
Sustainability	2004	663	567
Energy and Power	2368	451	585
Electric	844	143	124
Energy	1179	237	336
Power plant	12	2	0
Renewable	333	69	125
Nature and Environment	1224	423	510
Earth	14	2	0
Natural	168	51	73
Nature	265	75	234
Planet	528	24	3
Plant	249	271	200

	Frequency		
Big Data Analytics	Delta 2019-2022	Emirates 2019-2022	United 2019-2022
Artificial Intelligence and Machine Learning	184	54	49
ML	0	0	0
Algorithm	24	1	0
Analytical	11	6	15
Analytics	19	14	16
Artificial	28	7	0
Machine	94	24	18
Machine Learning	8	2	0
Data and Analytics	3465	2531	3857
Analysis	262	133	259
Application	232	46	67
Big Data	8	0	0
Data	1062	242	1147
Data Management	3	0	1
Data Mining	0	0	0
Database	49	4	18
Data-Driven	10	2	0
dataSet	0	0	0
dataSets	0	0	0
Predict	97	13	117
Prediction	3	0	4
Predictive	9	1	0
Processing	101	20	28
Information	1504	2028	2136
Mining	125	41	80
Mathematic	0	0	0
Maths	0	1	0
Technology and Software	3500	1328	2070
AWS	4	0	0
Blockchain	12	3	5

Keywords Table | Raw | Volkswagen | Tesla | Delta | United | Emirates | Toyota | ESSCM | BDA

Figure 4.1: Frequency Numbers by Sub-themes

4.3 Wordclouds: Initial Keyword Impressions

Wordclouds provide a first visual impression of which keywords dominate reporting.

- **Automotive Wordcloud:** The largest terms are *environment, emission, sustainability, energy,* and *data*. This indicates that car manufacturers integrate sustainability with digitalisation.

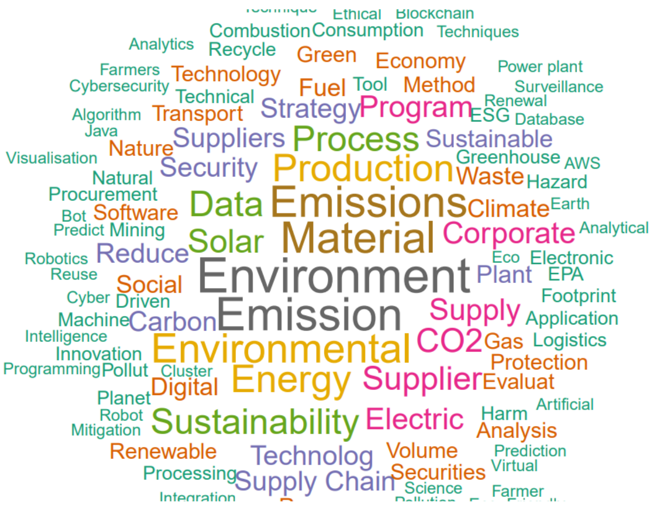


Figure 4.2: Wordcloud of Automotive Industry – BDA and ESSCM

DOI: https://doi.org/10.5526/esj.457

- **Airline Wordcloud:** The largest terms are *fuel, corporate, sustainability,* with relatively smaller BDA-related terms. This shows that airlines prioritise environmental compliance and reporting, with less emphasis on advanced analytics.

Figure 4.3: Wordcloud of Airline – BDA and ESSCM

Interpretation of Wordclouds: Together, these Wordclouds highlight an important point: both industries recognise the importance of environmental and sustainability-related themes as central to their corporate strategies. However, BDA appears more embedded in the automotive industry than in airlines, albeit still secondary to the environmental agenda. In contrast, airlines present sustainability as a compliance-driven and risk-avoidance priority, where BDA plays a less visible role. These findings reflect broader industry narratives; as noted by Luckow (2015) and Giacosa et al. (2021), automotive firms increasingly embed data analytics into environmental reporting, while airlines focus more on regulatory and compliance-related sustainability disclosures (Aydin, 2019)

4.4 Heatmaps: Frequency Distributions

Heatmaps quantify how often keywords were mentioned across sub-themes and years.

- Automotive BDA: Emphasis is strongest in *Data & Analytics* and *Technology & Software*. Toyota peaked at 1,257 mentions in 2020, Volkswagen at 1,117 in 2021. *AI & ML* remained consistently low (<100 annually). This pattern reflects wider post-COVID digitalisation trends in automotive supply chains (Chung, 2021; Silva, 2021).

DOI: https://doi.org/10.5526/esj.457

- Automotive ESSCM: *Sustainability & Conservation* dominated (Toyota reported 2,031 mentions in 2019), with *Supply Chain & Procurement* increasing sharply (VW: 444 in 2019 → 986 in 2022).
- Airline BDA: Much lower intensity; Delta and United averaged ~800 mentions in *Data & Analytics*, Emirates fewer (~600).
- Airline ESSCM: Strongest emphasis on *Sustainability* (~1,200 per year for Delta). Emirates grew in *Energy & Power* from 56 (2019) to 188 (2022). This stability reflects regulatory and operational pressures in aviation that enforce sustained sustainability practices (Ramirez-Peña et al., 2020; Savchenko, 2023)

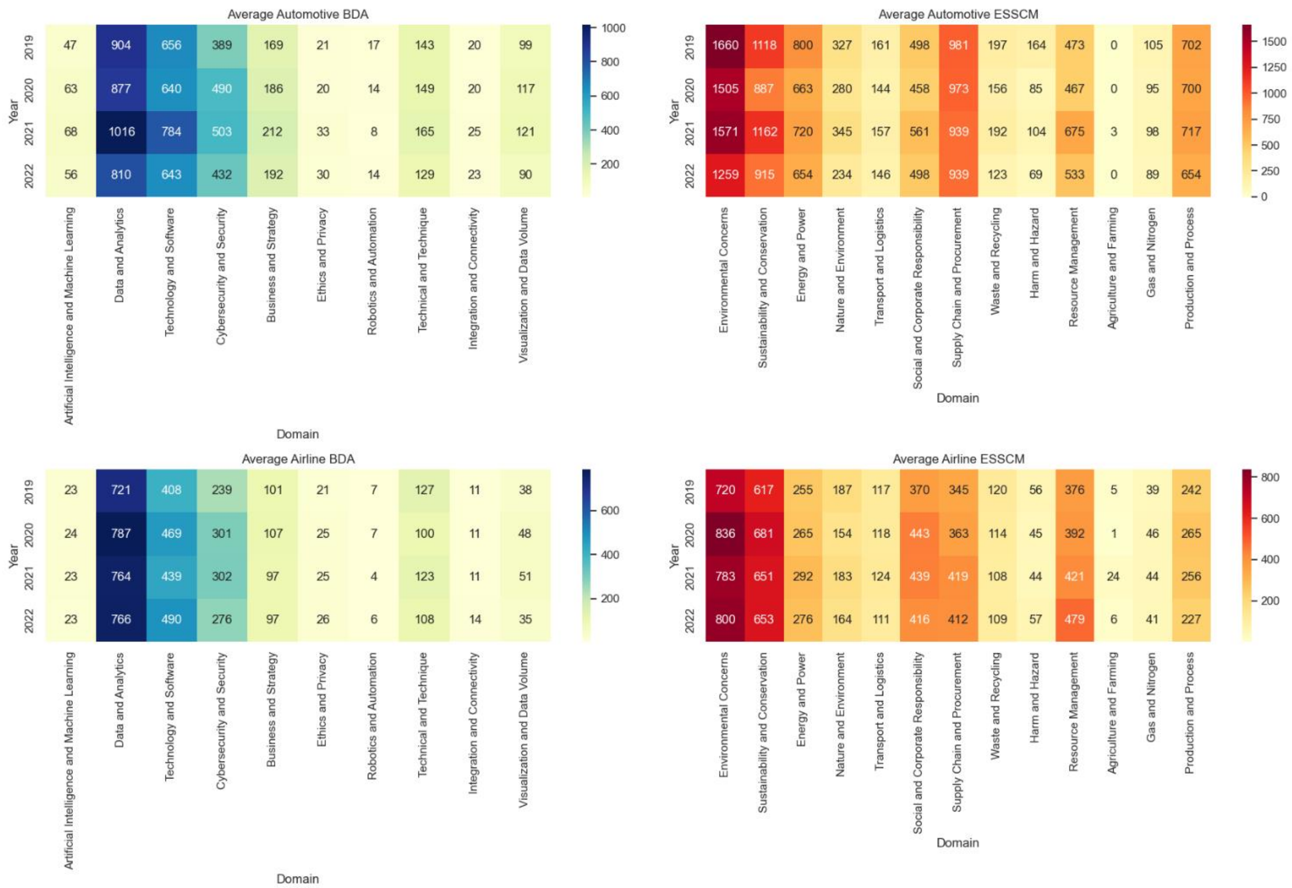


Figure 4.4: Four-panel heatmap (Auto BDA, Auto ESSCM, Airline BDA, Airline ESSCM)

These results confirm that the automotive sector is BDA-heavy, while airlines are ESSCM-heavy.

4.5 Radar Charts: Cross-Industry Profiles

Radar charts enable cross-industry comparison by mapping domain strengths.

DOI: https://doi.org/10.5526/esj.457

- **BDA Radar:** Automotive firms extend furthest in *Data & Analytics* and *Technology & Software*, while airlines lag significantly.

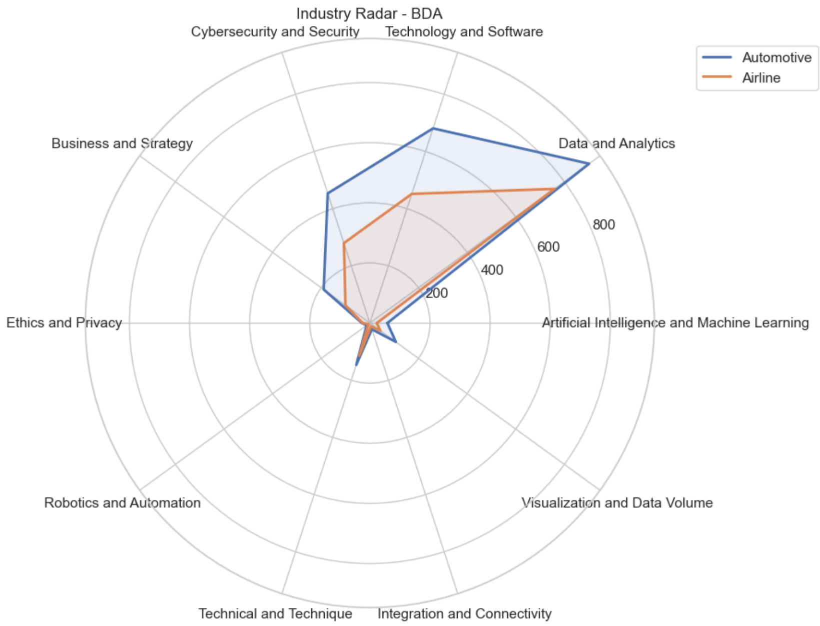


Figure 4.5: Industry Radar – BDA

- **ESSCM Radar:** Airlines extend further in *Energy & Power* and *Sustainability & Conservation*, while automotive firms contribute more in *Supply Chain & Procurement*.

DOI: https://doi.org/10.5526/esj.457

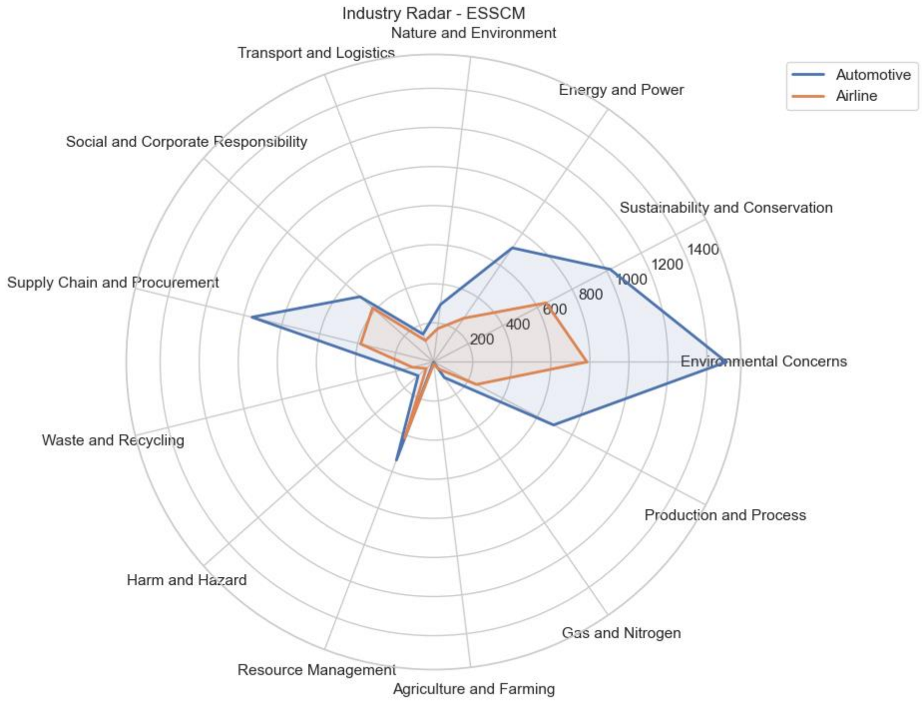


Figure 4.6: Industry Radar – ESSCM

Interpretation: These comparisons confirm that the two industries complement each other: one prioritises digital innovation, the other regulatory sustainability. This divergence aligns with Savchenko (2023), who highlights that aviation sustainability practices are largely driven by regulation, while automotive firms pursue innovation cycles enabled by BDA (Zhang et al., 2017)

4.6 Stacked Bar Charts: Contributions Over Time

Stacked bar charts show how domain contributions evolved between 2019 and 2022.

- **BDA Contribution:** Automotive shows sharp post-2020 growth in *Data & Analytics* and *Technology*. Airlines remain stable and comparatively low.

DOI: https://doi.org/10.5526/esj.457

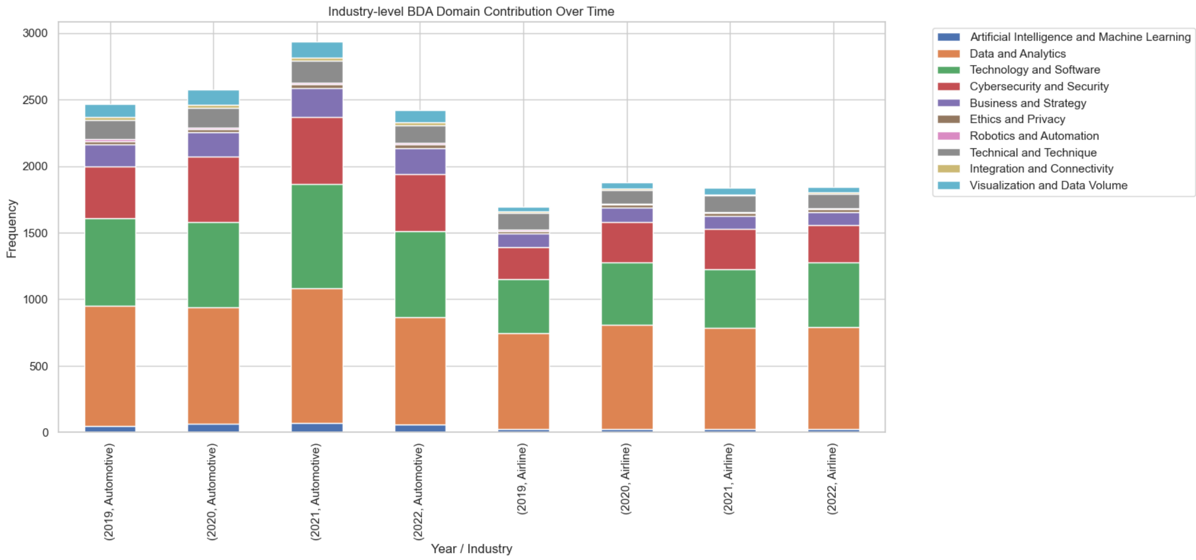


Figure 4.6: Industry-level BDA Domain Contribution Over Time

- **ESSCM Contribution:** Airlines consistently prioritise *Sustainability*. Automotive fluctuates: Toyota declines, Volkswagen's *Supply Chain* more than doubles.

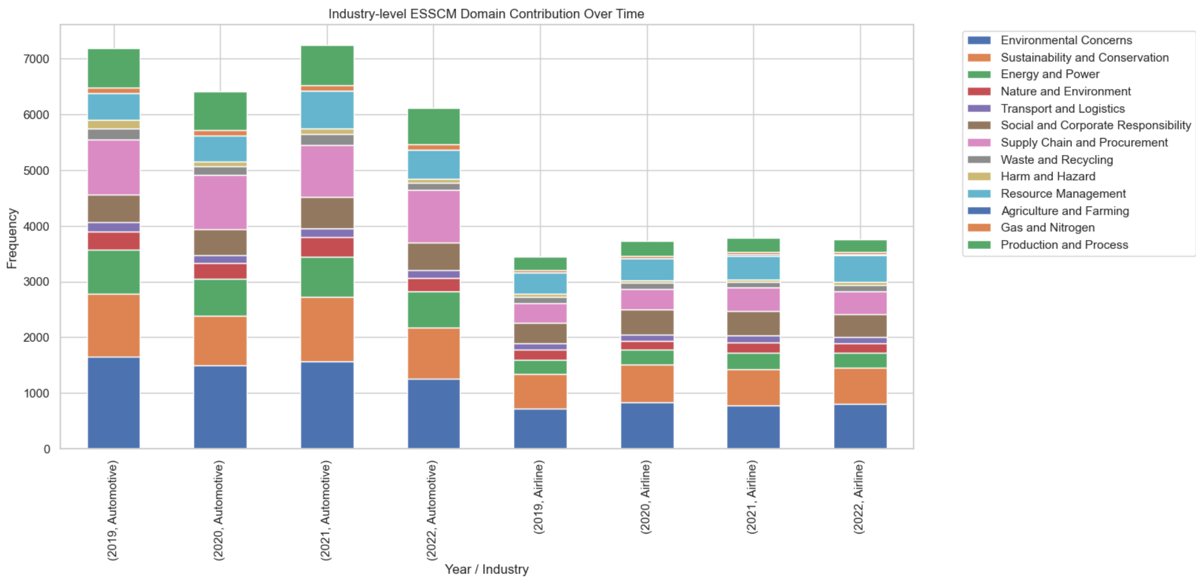


Figure 4.7: Industry-level ESSCM Domain Contribution Over Time

Synthesis of Visual Analysis

The numerical results highlight three central insights:

DOI: https://doi.org/10.5526/esj.457

1. **Magnitude Differences** → Automotive companies often exceed 1,000 mentions annually in BDA, while airlines remain in the 600–800 range.
2. **ESSCM Focus** → Airlines report 1,000+ sustainability mentions annually, making them stronger in ESSCM despite lower BDA intensity.
3. **Temporal Patterns** → Automotive investments surged after 2020 (e.g., Volkswagen's *Supply Chain* doubled from 444 → 986), while airlines grew more modestly but steadily.

These contrasting trajectories are consistent with Zhu et al. (2008), who observed that sustainability and supply chain priorities differ across industries depending on regulatory context and market dynamics.

4.7 Trend Analysis: Yearly Dynamics

Line charts provide a clearer view of yearly movement.

- Airlines: Delta shows stability in both BDA (~3,000–3,400) and ESSCM (~6,000–7,000). Emirates grows incrementally, while United surges dramatically in 2022 (~3,500 BDA; ~5,000 ESSCM). This surge reflects wider aviation pressures, including ICAO's CORSIA framework and the EU Emissions Trading System, which enforce sustainability commitments (Burmester, 2018; Ritchie, 2020).

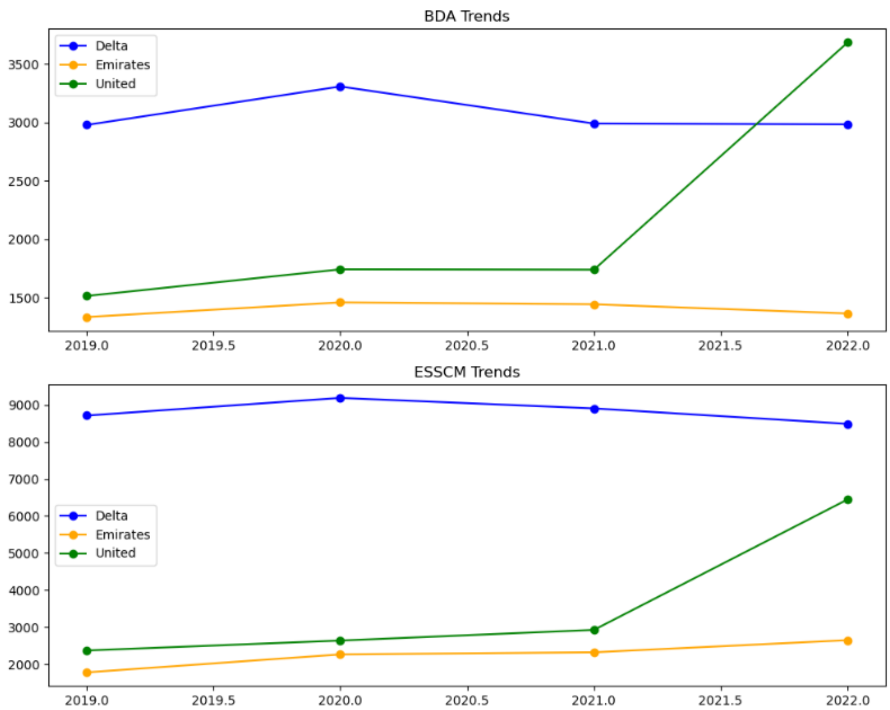


Figure 4.8: Airline BDA/ESSCM Trends

DOI: https://doi.org/10.5526/esj.457

- Automotive: Volkswagen peaks in BDA (~3,600 in 2021) and ESSCM (~9,000 in 2021), then declines. Toyota remains steady in BDA but declines in ESSCM (12,000 → 8,000). Tesla is stable in BDA but declines in ESSCM (9,500 → 7,500). This reflects broader sectoral transitions where firms shift emphasis from traditional sustainability to EV and technology innovation (Beier, 2022; Singh, 2022).

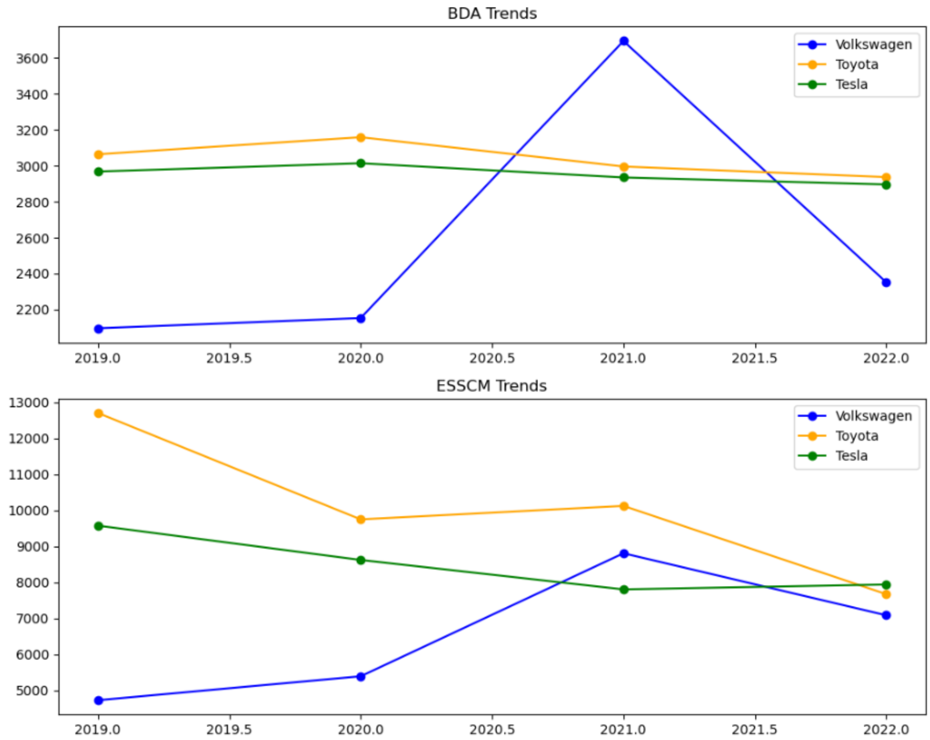


Figure 4.9: Automotive BDA/ESSCM Trends

This confirms airlines as steady and compliance-driven, while automotive firms are volatile and innovation-driven.

4.8 Industry-Level Scatterplot Findings

Scatterplots were created to compare average values of automotive vs airline practices across both BDA and ESSC domains.

- **BDA Comparison**

The scatterplot demonstrated a very strong positive correlation (r ≈ 0.97) between the two industries in BDA adoption. This indicates that increases in BDA-related activities within the automotive industry are closely mirrored by corresponding increases in the airline industry.

DOI: https://doi.org/10.5526/esj.457

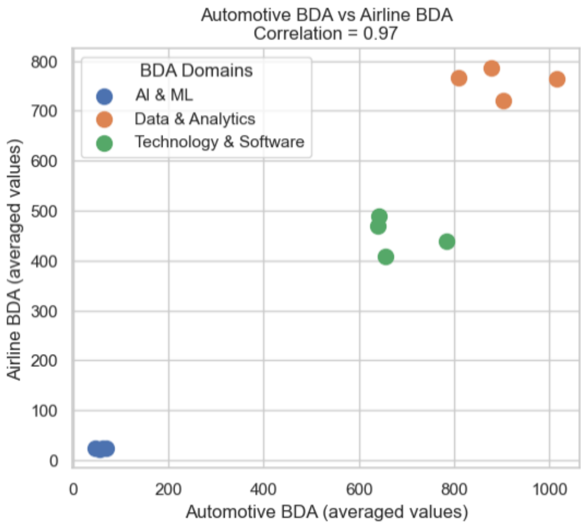


Figure 4.10: Automotive BDA vs Airline BDA

- For example, in *Data & Analytics*, Toyota and Volkswagen consistently reported over 1,000 references annually, while Delta and United averaged ~800 mentions.
- *Technology & Software* also displayed a clear parallel, with automotive firms frequently exceeding 600–1,100 mentions, while airlines such as Emirates and United remained closer to 300–500 mentions.
- *AI & ML* showed the lowest overall values, but trends were still aligned: automotive firms averaged 50–70 mentions per year, while airlines reported fewer than 50 annually.

Interpretation: The near-perfect correlation indicates that even though the automotive industry operates at a higher scale of BDA integration, the pattern of adoption is highly synchronised with the airline sector.

- **ESSC Comparison**

For ESSC practices, the scatterplot revealed a moderate positive correlation (r ≈ 0.69). This suggests some alignment between the industries, but with greater variability.

DOI: https://doi.org/10.5526/esj.457

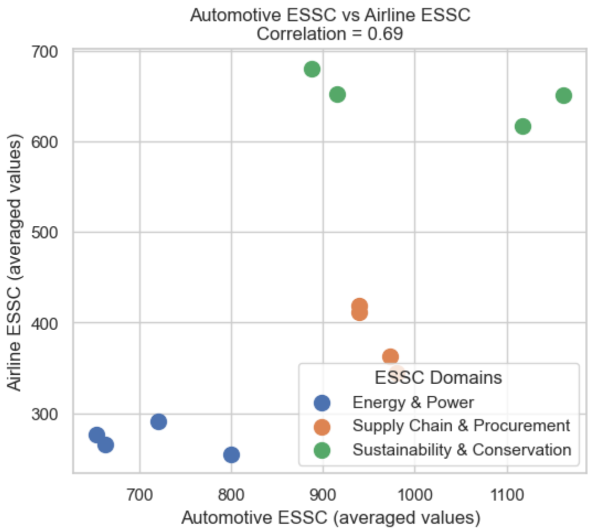


Figure 4.11: Automotive ESSCM vs Airline ESSCM

- In *Sustainability and Conservation*, Toyota reported 2,031 mentions in 2019 but dropped to 1,074 in 2022, showing volatility. By contrast, Delta reported a steadier ~1,200 mentions annually.
- *Energy & Power* was a central focus for Tesla (1,185 mentions in 2019; 836 in 2022), while Delta maintained consistent values around 600 mentions annually.
- *Supply Chain and Procurement* grew more sharply in automotive, with Volkswagen rising from 444 in 2019 to 986 in 2022, whereas airlines like Delta ranged between 804–936 mentions.

The weaker correlation highlights that while both industries are committed to sustainability, their trajectories differ: automotive firms fluctuate with strategic shifts and innovation cycles, whereas airlines show steadier, compliance-driven emphasis. These strong positive correlations support the argument that BDA acts as an enabling resource for sustainability, consistent with findings in Dubey and Singh (2019) and Wamba et al. (2017).

4.9 Conclusion

This section has demonstrated how automotive and airline industries report BDA and ESSCM practices. The text mining procedure converted unstructured corporate reports into structured keyword datasets, forming the foundation for all analysis. Wordclouds revealed industry narratives, heatmaps quantified frequency intensity, radar charts compared profiles, stacked bar charts showed contribution dynamics, line charts traced yearly trajectories, and scatterplots confirmed correlations.

DOI: https://doi.org/10.5526/esj.457

The findings confirm distinct sectoral orientations: automotive companies are BDA-heavy, volatile, and innovation-driven, while airlines are ESSCM-heavy, stable, and compliance-oriented. These results support the RBV perspective that digital resources such as BDA can provide firms with a valuable capability that enhances environmental sustainability outcomes (Wamba et al., 2017; Dubey & Singh, 2019).

In the next section, we will discuss in detail how the findings align with the Hypothesis and answers research questions.

5. Discussion

This section interprets the findings presented in section 4 in relation to the research questions, the hypothesis, and the Resource-Based View (RBV). Whereas the findings section presented descriptive evidence, here the focus is on explaining what these results mean, how they align with or diverge from the literature, and how they contribute to theory and practice.

The discussion is structured around the four research questions (RQ1–RQ4) and the central hypothesis. Each subsection integrates the results with existing scholarship and RBV, before moving on to contributions, implications, and methodological limitations.

5.1 Hypothesis Testing

The hypothesis testing confirmed expectations:

H0 (no significant correlation) was rejected.

H1 (a significant positive correlation exists between BDA and ESSCM) was supported.

A near-perfect BDA correlation ($r = 0.97$) and moderate ESSCM correlation ($r = 0.69$) demonstrate a systematic association between digitalisation and sustainability practices. While prior studies (Wamba et al., 2017; Dubey & Singh, 2019) relied on surveys and qualitative case studies, this research contributes by using text mining of corporate reports to provide quantitative confirmation. This methodological extension strengthens the evidence base for the BDA–sustainability link, addressing calls for more empirical validation in the literature.

5.2 Research Question Discussion

RQ1 – Extent of BDA Usage

Findings reveal that automotive firms embed BDA strategically in domains such as Data & Analytics and Technology, while airlines adopt analytics mainly for compliance. This partly aligns with Luckow (2015) and Giacosa (2021), but challenges their framing of digitalisation as a universal enabler. This paper's results highlight sectoral divergence: despite similar regulatory pressures, airlines lag in embedding BDA as a VRIN resource. This extension underscores that

DOI: https://doi.org/10.5526/esj.457

BDA adoption is not uniform but shaped by industry-specific drivers—innovation cycles in automotive versus compliance in aviation.

From an RBV perspective, this suggests that automotive firms are beginning to treat BDA as a *valuable* and *rare* resource that underpins innovation and competitive advantage. Airlines, however, have yet to exploit its full potential, limiting their ability to leverage analytics as a VRIN (Valuable, Rare, Inimitable, Non-substitutable) resource.

Answer: BDA is widely referenced across both industries, but automotive firms embed it more intensively and strategically, while airlines primarily adopt it for operational compliance.

RQ2 – How BDA Supports ESSCM

The results show strong alignment between Data & Analytics and Sustainability & Conservation, suggesting that higher BDA intensity translates into greater sustainability emphasis. This confirms Min (2010), Zhong et al. (2016), and Fernando et al. (2018), but crucially addresses the ambiguity raised by Song et al. (2017), who argued that the mechanisms were unclear. By offering statistical confirmation rather than qualitative inference, this study extends prior work and demonstrates how BDA enables measurable sustainability outcomes, not just efficiency gains.

From the RBV lens, this integration of BDA into sustainability practices demonstrates its role as a resource that is *valuable* and *difficult to imitate*. Firms such as Tesla, which showed particularly strong AI–Sustainability correlations, exemplify how proprietary analytical capabilities can generate sustainability outcomes that competitors may struggle to replicate.

Answer: BDA supports ESSCM by enhancing operational efficiency, reducing environmental impact, and strengthening sustainability reporting. Automotive firms lead in this integration, but airlines also show positive alignment, particularly in energy and emissions management.

RQ3 – Usefulness of Text Analytics

The text mining approach demonstrated that text analytics can successfully transform unstructured corporate reports into structured, comparable datasets. By applying keyword frequencies, the study was able to identify patterns in reporting emphasis and statistically test correlations across industries.

This supports the claims of Tao et al. (2020) and Mergel (2019), who argued that text mining provides scalable, replicable insights from large textual datasets. It also builds on Meyer et al. (2021), who applied sentiment analysis to supply chain resilience reporting. Unlike qualitative case studies, text analytics provides consistency and comparability across firms and industries.

Theoretically, text analytics itself can be conceptualised as a **dynamic capability** under RBV. It equips firms (and researchers) with the ability to continuously restructure and interpret information, extracting value from otherwise unmanageable volumes of text.

DOI: https://doi.org/10.5526/esj.457

Answer: Text analytics is an effective tool for building case studies and evaluating corporate initiatives, offering replicable and scalable insights into BDA–ESSCM integration.

RQ4 – Challenges, Limitations, and Concerns

Challenges remain. Airlines' weaker BDA adoption reflects structural barriers such as cost and legacy systems (Hashem et al., 2015; Mohamed, 2021), while volatility in Toyota's ESSCM reporting suggests symbolic disclosure rather than substantive strategy, echoing Shubbak's (2013) critique of 'greenwashing.' The keyword-frequency approach further underlines that emphasis may not equal action. From an RBV perspective, this shows that while BDA is valuable and rare, its uneven distribution across firms creates winners and laggards—Tesla and Toyota gain competitive edge, whereas Emirates and United risk falling behind. Leaders such as Tesla and Toyota exploit it effectively, while laggards like Emirates and United are disadvantaged by weaker capabilities.

Answer: Key challenges include cost, expertise, and reporting consistency. Methodologically, text mining captures emphasis but not context, necessitating complementary qualitative approaches.

5.3 Theoretical Integration

The findings align with RBV but also extend it. Traditionally, RBV applied to tangible or human resources; here, results demonstrate that digital capabilities such as BDA and text analytics also function as VRIN resources. Automotive firms leverage analytics as a rare and inimitable capability, while airlines illustrate the risks of underutilisation. Moreover, text mining itself can be conceptualised as a dynamic capability, enabling firms to continuously restructure and interpret sustainability information. This theoretical extension shows that RBV remains highly relevant in the digital era but must now incorporate intangible data-driven capabilities. Wordclouds and correlations revealed strategic differences: automotive companies frame BDA as transformative, while airlines frame sustainability as compliance.

This extends existing literature by showing that RBV applies not only to tangible assets but also to digital knowledge capabilities (Seele, 2016; Mageto, 2021).

5.4 Contributions to Literature

This section contributes in three ways. First, it confirms prior claims that BDA enhances supply chain efficiency and sustainability (Giacosa et al., 2021; Chung, 2021). Second, it extends existing research by applying text mining to corporate disclosures, providing quantitative confirmation of BDA–ESSCM alignment, where most prior studies relied on surveys or case studies. Third, it addresses a major gap by offering one of the first comparative analyses of BDA–ESSCM integration in airlines and automotive firms, demonstrating synchronisation in BDA adoption ($r \approx 0.97$) but divergence in sustainability emphasis ($r \approx 0.69$). This dual contribution—methodological and sectoral—positions the study as a novel step in bridging digitalisation and environmental sustainability research.

DOI: https://doi.org/10.5526/esj.457

5.5 Practical Implications

The findings suggest clear actions for practice. Automotive firms should embed BDA directly into sustainability strategies, moving beyond efficiency to innovation-led environmental outcomes. Airlines must shift from compliance-driven reporting to operational integration of analytics to avoid lagging behind competitors. Regulators should standardise reporting metrics across industries to reduce symbolic disclosure and ensure comparability. Managers across both sectors must prioritise investment in analytics skills and infrastructure to fully exploit BDA as a VRIN resource. Importantly, risks accompany inaction: firms that underutilise BDA risk reputational greenwashing, while overreliance on symbolic disclosure undermines competitiveness. These implications stress that analytics is not only a technical tool but a strategic enabler of sustainable supply chain transformation.

5.6 Conclusion

This section has interpreted the results in light of RQ1–RQ4, the hypothesis, and RBV. The findings confirm that BDA adoption is positively correlated with ESSCM, but with distinct orientations: automotive firms lead in analytics integration, while airlines focus on sustainability compliance. The section also outlined theoretical and practical contributions, acknowledged methodological boundaries, and positioned the research within RBV.

The next section concludes the thesis, providing an overall synthesis, extended limitations, and directions for future research.

6. Conclusion

This section concludes the study by synthesising the key findings, reflecting on the study's limitations, and outlining directions for future research. Whereas section 5 provided detailed discussion and interpretation, this section takes a broader view, drawing the study together and positioning it within the wider research landscape.

6.1 Overall Conclusion

This research set out to investigate the relationship between Big Data Analytics (BDA) and Environmental Sustainability in Supply Chain Management (ESSCM) within the automotive and airline industries. The study confirmed that BDA adoption and sustainability practices are positively aligned, though the form and intensity of this alignment differ across sectors.

The findings show that automotive firms are more focused on integrating analytics into innovation and operational processes, whereas airlines emphasise sustainability practices largely in response to compliance and regulatory pressures. Despite these differences in orientation, both

DOI: https://doi.org/10.5526/esj.457

industries demonstrate that digital capabilities and sustainability goals are increasingly interconnected rather than separate domains.

The study also demonstrated the value of text mining as a methodological approach for transforming unstructured corporate reports into structured, comparable insights. This methodological contribution highlights how emerging analytical tools can advance both academic research and managerial practice by enabling systematic monitoring of sustainability emphasis.

From a theoretical standpoint, the results reinforce the Resource-Based View (RBV), confirming that BDA can be understood as a VRIN (Valuable, Rare, Inimitable, Non-substitutable) resource. Furthermore, text analytics is positioned as a dynamic capability, enabling firms to adapt and respond to evolving sustainability demands through better use of data.

Overall, the thesis provides strong evidence that BDA is not only a technological enabler but also a strategic driver of environmental sustainability. The integration of analytics into sustainability reporting and practice enhances both compliance and competitiveness, marking a shift towards a more data-driven approach to sustainable supply chain management.

6.2 Limitations

This study's scope was deliberately defined to balance comparability with depth. The following boundaries should be recognised:

1. **Data Source:** The analysis relied on annual and sustainability reports, which ensured consistency across firms but excluded informal or internal practices not disclosed publicly. Corporate reporting may also reflect reputational motives.

2. **Keyword Frequency Approach:** The use of keyword counts prioritised replicability and comparability across firms. While this allowed statistical testing, it did not capture the context or meaning of terms, such as whether "sustainability" referred to emissions, CSR, or energy management.

3. **Timeframe:** The period of 2019–2022 was chosen to capture recent trends, including the effects of COVID-19. While timely, it represents a snapshot rather than a long-term trajectory.

4. **Sample Focus:** Six leading firms were purposively selected to reflect industry leaders whose strategies often set wider benchmarks. This provides depth but limits generalisability to smaller companies or other geographies.

5. **Quantitative Emphasis:** The study's reliance on frequency counts and correlations ensured objectivity but excluded qualitative insights such as managerial intent or cultural framing.

These boundaries reflect deliberate methodological choices rather than weaknesses, but they highlight areas where future research could extend the work.

DOI: https://doi.org/10.5526/esj.457

6.3 Future Research Directions

Building on these boundaries, several avenues for future investigation are recommended:

1. **Expanding Data Sources:** Future studies could integrate other disclosures such as press releases, investor briefings, or stakeholder reports, alongside annual filings, to capture richer perspectives. Social media or news datasets may also offer insights into external perceptions of BDA–ESSCM practices.

2. **Beyond Keyword Counts:** More advanced methods such as natural language processing (NLP), topic modelling, or sentiment analysis could capture not just the frequency but also the framing and tone of sustainability discourse.

3. **Longer Time Horizons:** Extending the timeframe would enable the study of long-term sustainability transitions, including pre-COVID baselines and future trajectories. Forecasting models could test whether BDA adoption predicts subsequent sustainability performance.

4. **Broader Industry Coverage:** Extending analysis to other resource-intensive industries such as shipping, logistics, or energy would test whether the observed BDA–ESSCM correlation is sector-specific or more universal.

5. **Causal Linkages:** Future work could go beyond correlation to examine causality between BDA investment and measurable environmental outcomes such as carbon reduction or supply chain resilience.

6. **Theoretical Diversification:** While RBV provided a strong lens, combining it with Institutional Theory or Stakeholder Theory could reveal how external pressures and societal expectations interact with internal capabilities.

7. **Mixed-Methods Approaches:** Combining text mining with interviews or surveys of managers would enrich the findings, allowing both statistical comparability and qualitative interpretation of organisational intent.

6.4 Final Remarks

This paper demonstrated that Big Data Analytics is a critical enabler of sustainable supply chain management. By applying text mining to corporate reports, it showed that analytics and sustainability are not isolated but positively correlated, albeit with different emphases across industries. The findings contribute to literature, theory, and practice by demonstrating BDA as a strategic VRIN resource, highlighting text analytics as a dynamic capability, and offering evidence for regulators and practitioners.

While the study focused on two industries and a defined timeframe, it provides a foundation for future research into digital sustainability transitions. As global pressures on environmental

DOI: https://doi.org/10.5526/esj.457

performance intensify, the integration of BDA into ESSCM will be increasingly central to firms' ability to compete, comply, and innovate.

References:

Abbasi, M. and Nilsson, F., 2012. Themes and challenges in making supply chains environmentally sustainable. *Supply Chain Management: An International Journal*, 17(5), pp.517-530.

Acquaye, A., Genovese, A., Barrett, J. and Lenny Koh, S.C., 2014. Benchmarking carbon emissions performance in supply chains. *Supply Chain Management: An International Journal*, 19(3), pp.306-321.

Acquaye, A.A., Duffy, A.P. and Basu, B., 2011. Stochastic hybrid embodied CO2-eq analysis: An application to the Irish apartment building sector. *Energy and Buildings*, 43(6), pp.1295-1303.

Anandarajan, M., Hill, C. and Nolan, T., 2019. *Practical text analytics: Maximizing the Value of Text Data*. Springer.

Aydın, U., Karadayı, M.A., Ülengin, F. and Ülengin, K.B., 2019. Enhanced performance assessment of airlines with integrated balanced scorecard, network-based superefficiency DEA and PCA methods. *Multiple Criteria Decision Making*, 25, pp.225-247.

Bag, S., Wood, L.C., Xu, L., Dhamija, P. and Kayikci, Y., 2020. Big data analytics as an operational excellence approach to enhance sustainable supply chain performance. *Resources, Conservation and Recycling*, 153, p.104559.

Barak, S. and Dahooei, J.H., 2018. A novel hybrid fuzzy DEA-Fuzzy MADM method for airlines safety evaluation. *Journal of Air Transport Management*, 73, pp.134-149.

Beier, G., Kiefer, J. and Knopf, J., 2022. Potentials of big data for corporate environmental management: A case study from the German automotive industry. *Journal of Industrial Ecology*, 26(1), pp.336-349.

Cao, Q., Lv, J. and Zhang, J., 2015. Productivity efficiency analysis of the airlines in China after deregulation. *Journal of Air Transport Management*, 42, pp.135-140.

Carter, C.R. and Rogers, D.S., 2008. A framework of sustainable supply chain management: moving toward new theory. *International Journal of Physical Distribution & Logistics Management*, 38(5), pp.360-387.

Cuquet, M. and Fensel, A. (2018). The societal impact of big data: A research roadmap for Europe. *Technology in Society*, 54, pp.74–86.

DOI: https://doi.org/10.5526/esj.457

Custers, B., La Fors, K., Jozwiak, M., Esther, K., Bachlechner, D., Friedewald, M. and Aguzzi, S. (2017). *Lists of Ethical, Legal, Societal and Economic Issues of Big Data Technologies*. [online] papers.ssrn.com.

Ertz, M., Imen Latrous, Ahlem Dakhlaoui and Sun, S. (2024). The impact of Big Data Analytics on firm sustainable performance. *Corporate Social Responsibility and Environmental Management*.

Guan, P. and Zhou, W. (2017). Business Analytics Generated Data Brokerage: Law, Ethical and Social Issues. *Communications in Computer and Information Science*, volume 759(CCIS), pp.167–175.

Chung, S.H., 2021. Applications of smart technologies in logistics and transport: A review. *Transportation Research Part E: Logistics and Transportation Review*, 153, p.102455.

Dahdouh, K., Dakkak, A., Oughdir, L. and Ibriz, A., 2020. Improving online education using big data technologies. *The Role of Technology in Education*, p.15.

Douglas, C.C., 2014. An open framework for dynamic big-data-driven application systems (DBDDAS) development. *Procedia Computer Science*, 29, pp.1246-1255.

Dubey, R. and Singh, T., 2019. Big data analytics and environmental sustainability: Resource-based view insights. *Journal of Cleaner Production*, 231, pp.262-273.

Das, A. and Sharma, A. (2023). Climate change and the energy sector. *Elsevier eBooks*, (9780323908757), pp.1–6.

Parpieva, U., Najmutdinova, G., Khaydarov, I., Zakirova, G. and Karimova, N. (2024). Issues of environmental changes and transportation. *BIO Web of Conferences*, [online] 141(04015), p.04015.

Fernando, Y., Jabbour, C.J.C. and Wah, W.X., 2018. Pursuing green growth in technology firms through the connection between big data analytics and business model innovation. *Journal of Cleaner Production*, 188, pp.463-472.

Giacosa, E., Culasso, F. and Crocco, E., 2021. How big data analytics is changing the automotive industry. *Technological Forecasting and Social Change*, 162, p.120382.

Giacosa, E., Culasso, F. and Crocco, E., 2022. Customer agility in the modern automotive sector: how lead management shapes agile digital companies. *Technological Forecasting and Social Change*, 175, p.121362.

DOI: https://doi.org/10.5526/esj.457

Golicic, S.L. and Smith, C.D., 2013. A meta-analysis of environmentally sustainable supply chain management practices and firm performance. *Journal of Supply Chain Management*, 49(2), pp.78-95.

Hashem, I.A.T., Yaqoob, I., Anuar, N.B., Mokhtar, S., Gani, A. and Ullah Khan, S., 2015. The rise of "big data" on cloud computing: Review and open research issues. *Information Systems*, 47, pp.98-115.

Hazen, B.T., Skipper, J.B., Ezell, J.D. and Boone, C.A., 2016. Big data and predictive analytics for supply chain sustainability: A theory-driven research agenda. *Computers & Industrial Engineering*, 101, pp.592-598.

Laney, D., 2001. META Delta. Application Delivery Strategies, 949(4), pp.1-4.

Luckow, A., Kennedy, K., Manhardt, F., Djerekarov, E., Vorster, B. and Apon, A., 2015. Automotive big data: Applications, workloads, and infrastructures. *2015 IEEE International Conference on Big Data (Big Data)*, pp.1201-1210.

Mageto, J., 2021. Natural resource-based view of the firm and environmental sustainability. *Business Strategy and the Environment*, 30(5), pp.2425-2439.

Mentzer, J.T., DeWitt, W., Keebler, J.S., Min, S., Nix, N.W., Smith, C.D. and Zacharia, Z.G., 2001. Defining supply chain management. *Journal of Business Logistics*, 22(2), pp.1-25.

Mergel, I., Edelmann, N. and Haug, N., 2019. Defining digital transformation: Results from expert interviews. *Government Information Quarterly*, 36(4), p.101385.

Meyer, A., Walter, W. and Seuring, S., 2021. The impact of the coronavirus pandemic on supply chains and their sustainability: a text mining approach. *Frontiers in Sustainability*, 2, p.1.

Min, H., 2010. Artificial intelligence in supply chain management: theory and applications. *International Journal of Logistics: Research and Applications*, 13(1), pp.13-39.

Mohamed, H.A.E.S. and Al-Azab, M.R., 2021. Big Data Analytics in airlines: Opportunities and challenges. *Journal of Association of Arab Universities for Tourism and Hospitality*, 21(4), pp.77-112.

Motau, B. and Boursier, C., 2016. Big Data: The five Vs. *Journal of Information Technology Research*, 9(3), pp.56-65.

OECD, 2009. Sustainable manufacturing and eco-innovation: Framework, practices, and measurement. Paris: OECD Publishing.

Phair, C. and Warren, K., 2021. Positivism in social science research. *Research Methods Review*, 8(2), pp.112-119.

DOI: https://doi.org/10.5526/esj.457

Ramirez-Peña, M., Mayuet, P.F., Vazquez-Martinez, J.M. and Batista, M., 2020. Sustainability in the aerospace, naval, and automotive supply chain 4.0: A descriptive review. *Materials*, 13(24), p.5625.

Rao, P., 2002. Greening the supply chain: a new initiative in South East Asia. *International Journal of Operations & Production Management*, 22(6), pp.632-655.

Ritchie, H., 2020. Sector by sector: where do global greenhouse gas emissions come from? *Our World in Data*. Available at: https://ourworldindata.org/ghg-emissions-by-sector.

Savchenko, L., Ovdiienko, O., Boichenko, M. and Mostenska, T., 2023. Assessment of the relationship between CSR environmental goals implementation and financial/operational indicators of airlines. *Scientific Journal of Silesian University of Technology. Series Transport*, 119.

Schaltegger, S., Bennett, M., Burritt, R.L. and Jasch, C., 2008. Environmental management accounting (EMA) as a support for cleaner production. Springer Netherlands.

Seele, P., 2016. Digitization as a driver of environmental sustainability: The panopticon as a model for managing transparency. *Business Strategy and the Environment*, 25(7), pp.413-425.

Seuring, S. and Müller, M., 2008. From a literature review to a conceptual framework for sustainable supply chain management. *Journal of Cleaner Production*, 16(15), pp.1699-1710.

Shubbak, M., 2013. Self-driving cars: Legal, social, and ethical aspects. SSRN. Available at: https://ssrn.com/abstract=2931847.

Silva, N., Barros, J., Santos, M.Y., Costa, C., Cortez, P., Carvalho, M.S. and Goncalves, J.N., 2021. Advancing logistics 4.0 with the implementation of a big data warehouse: A demonstration case for the automotive industry. *Electronics*, 10(18), p.2221.

Snieder, R. and Larner, K., 2009. *The art of being a scientist: A guide for graduate students and their mentors*. Cambridge University Press.

Song, M., Cen, L., Zheng, Z., Fisher, R., Liang, X., Wang, Y. and Huisingh, D., 2017. How would big data support societal development and environmental sustainability? *Journal of Cleaner Production*, 142, pp.489-500.

Tao, D., Yang, P. and Feng, H., 2020. Utilization of text mining as a big data analysis tool for food science and nutrition. *Comprehensive Reviews in Food Science and Food Safety*, 19(2), pp.875-894.

Tavana, M., Shaabani, A., Raeesi Vanani, I. and Kumar Gangadhari, R., 2022. A review of digital transformation on supply chain process management using text mining. *Processes*, 10(5), p.842.

Toyota Motor Corporation, 2023. *Sustainability Report 2023*. Available at: https://global.toyota.

DOI: https://doi.org/10.5526/esj.457

Volkswagen Group, 2023. *Sustainability Report 2023*. Available at: https://www.volkswagen-group.com/.

Wamba, S.F., Gunasekaran, A., Akter, S., Ren, S.J., Dubey, R. and Childe, S.J., 2017. Big data analytics and firm performance: Effects on profitability, revenue, and cost. *Journal of Business Research*, 70, pp.356-365.

Wang, L., Ding, J., Pan, L., Cao, D., Jiang, H. and Ding, X., 2019. Artificial intelligence facilitates drug design in the big data era. *Chemometrics and Intelligent Laboratory Systems*, 194, p.103850.

World Economic Forum, 2023. Global per capita emissions explained. Available at: https://www.weforum.org/agenda/2023/05/global-per-capita-emissions-explained-charts/.

Zhang, Y., Ren, S., Liu, Y. and Si, S., 2017. A big data analytics architecture for cleaner manufacturing and maintenance processes of complex products. *Journal of Cleaner Production*, 142, pp.626-641.

Zhong, R.Y., Xu, C., Chen, C. and Huang, G.Q., 2016. Big data analytics for physical internet-based intelligent manufacturing shop floors. *International Journal of Production Research*, 55(9), pp.2610-2621.

Zhu, Q., Crotty, J. and Sarkis, J., 2008. A cross-country empirical comparison of environmental supply chain management practices in the automotive industry. *Asian Business & Management*, 7(4), pp.467-488.

DOI: https://doi.org/10.5526/esj.457

Appendices

Appendix A: Raw Data Structure:

1. **Keyword Counts:**

Keywords_COUNT

=COUNTIF(Data!$A$1:$FD$57452, "*" & A2 & "*")

KeyWords	Count
ML	0
Algorithm	0
Analysis	29
Analytical	4
Analytics	7
Application	30
Artificial	6
AWS	0
Big Data	4
Blockchain	1
Bot	2
Business Intelligence	0
Cloud	11
Cluster	4
Computer	7
Cyber	21
Cybersecurit	5
Data	106
Data Management	1
Data Mining	0
Database	3
Data-Driven	1
dataSet	0
dataSets	0
Digital	123
Driven	23
Electronic	21
Ethical	7

DOI: https://doi.org/10.5526/esj.457

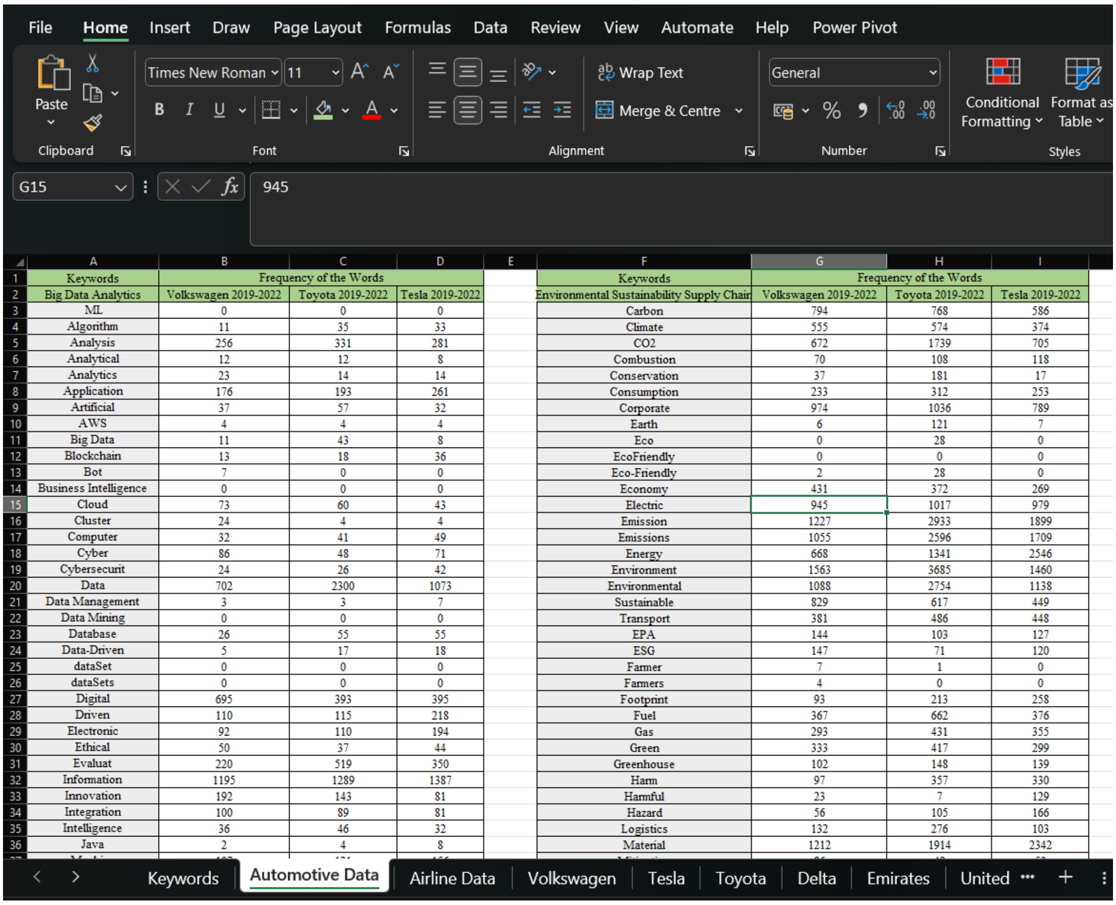
2. Data Sheet obtained in Step 2 – Unstructured Data

Keywords	Annual Report and Sustainability Report - Year (2019-2022)			
Big Data Analytics	2019	2020	2021	2022
ML	0	0	0	0
Algorithm	5	5	1	0
Analysis	50	44	100	62
Analytical	1	2	8	1
Analytics	4	5	9	5
Application	55	44	44	33
Artificial	10	9	9	9
AWS	77	1	0	60
Big Data	3	4	4	0
Blockchain	7	3	3	0
Bot	2	1	2	2
Business Intelligence	0	0	0	0
Cloud	29	23	11	10
Cluster	4	3	10	7
Computer	4	9	13	6
Cyber	16	21	29	20
Cybersecurit	2	7	10	5
Data	131	136	294	141
Data Management	0	1	1	1
Data Mining	0	0	0	0
Database	7	7	6	6
Data-Driven	0	1	1	3
dataSet	0	0	0	0
dataSets	0	0	0	0
Digital	180	184	195	136
Driven	21	23	36	30
Electronic	14	23	44	11
Ethical	14	9	16	11
Evaluat	48	45	95	32
Information	202	209	571	213
Innovation	29	41	66	56
Integration	19	21	37	23
Intelligence	10	9	8	9
Java	1	1	0	0

Keywords	Annual Report and Sustainability Report - Year (2019-2022)			
Environmental Sustainability Supply Chain	2019	2020	2021	2022
Carbon	101	125	226	342
Climate	93	108	167	187
CO2	162	157	175	178
Combustion	13	14	22	21
Conservation	4	5	17	11
Consumption	39	40	118	36
Corporate	153	212	351	258
Earth	3	1	1	1
Eco	0	0	0	0
EcoFriendly	0	0	0	0
Eco-Friendly	0	0	2	0
Economy	41	80	112	198
Electric	225	214	257	249
Emission	250	254	444	279
Emissions	217	217	391	230
Energy	112	127	230	199
Environment	321	314	492	436
Environmental	202	216	345	325
Sustainable	129	125	374	201
Transport	71	94	120	96
EPA	266	268	22	238
ESG	1	13	346	87
Farmer	0	0	6	1
Farmers	0	0	4	0
Footprint	20	30	29	14
Fuel	64	58	173	72
Gas	69	52	91	81
Green	40	55	142	96
Greenhouse	20	14	42	26
Harm	13	14	39	31
Harmful	2	2	17	2
Hazard	11	13	26	6
Logistics	26	29	36	41
Material	170	209	586	247

DOI: https://doi.org/10.5526/esj.457

3. Automotive Industry-wise Data Structure:

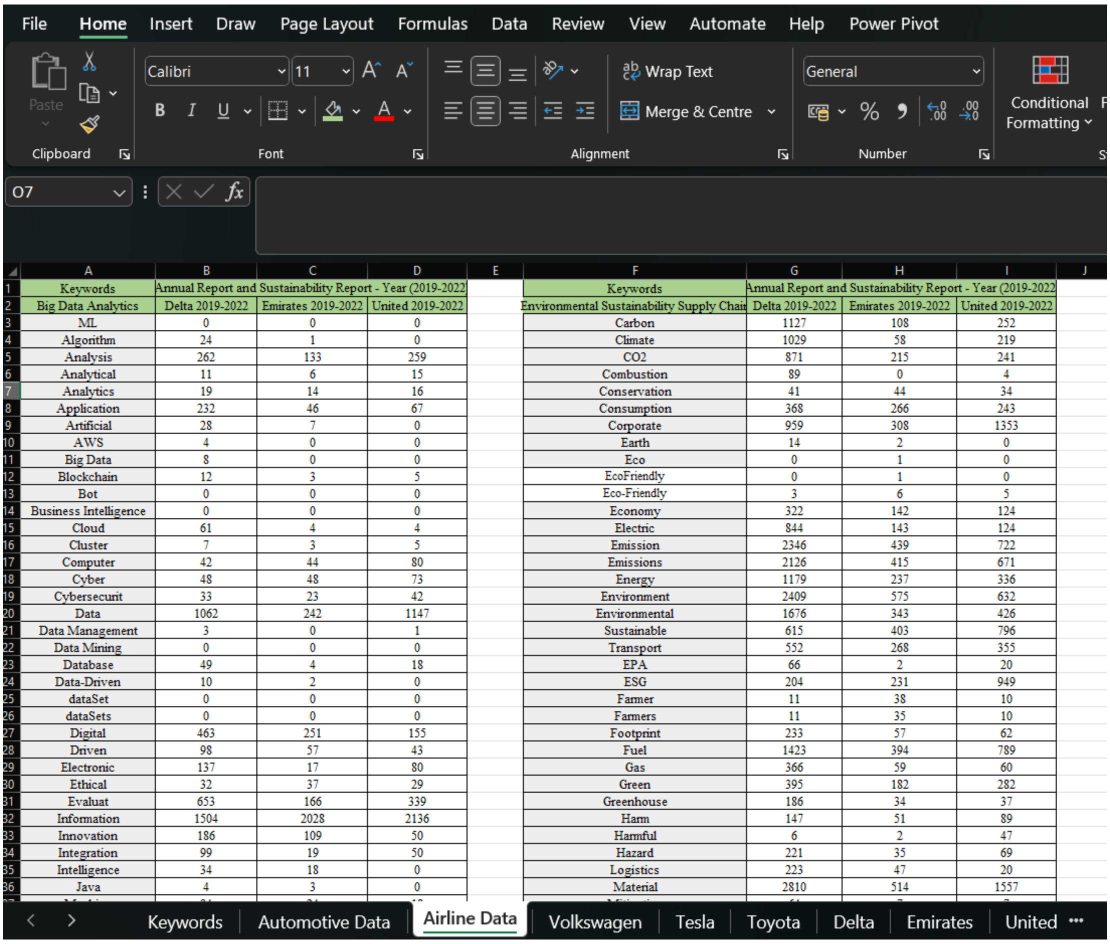


Keywords	Frequency of the Words		
Big Data Analytics	Volkswagen 2019-2022	Toyota 2019-2022	Tesla 2019-2022
ML	0	0	0
Algorithm	11	35	33
Analysis	256	331	281
Analytical	12	12	8
Analytics	23	14	14
Application	176	193	261
Artificial	37	57	32
AWS	4	4	4
Big Data	11	43	8
Blockchain	13	18	36
Bot	7	0	0
Business Intelligence	0	0	0
Cloud	73	60	43
Cluster	24	4	4
Computer	32	41	49
Cyber	86	48	71
Cybersecurit	24	26	42
Data	702	2300	1073
Data Management	3	3	7
Data Mining	0	0	0
Database	26	55	55
Data-Driven	5	17	18
dataSet	0	0	0
dataSets	0	0	0
Digital	695	393	395
Driven	110	115	218
Electronic	92	110	194
Ethical	50	37	44
Evaluat	220	519	350
Information	1195	1289	1387
Innovation	192	143	81
Integration	100	89	81
Intelligence	36	46	32
Java	2	4	8

Keywords	Frequency of the Words		
Environmental Sustainability Supply Chain	Volkswagen 2019-2022	Toyota 2019-2022	Tesla 2019-2022
Carbon	794	768	586
Climate	555	574	374
CO2	672	1739	705
Combustion	70	108	118
Conservation	37	181	17
Consumption	233	312	253
Corporate	974	1036	789
Earth	6	121	7
Eco	0	28	0
EcoFriendly	0	0	0
Eco-Friendly	2	28	0
Economy	431	372	269
Electric	945	1017	979
Emission	1227	2933	1899
Emissions	1055	2596	1709
Energy	668	1341	2546
Environment	1563	3685	1460
Environmental	1088	2754	1138
Sustainable	829	617	449
Transport	381	486	448
EPA	144	103	127
ESG	147	71	120
Farmer	7	1	0
Farmers	4	0	0
Footprint	93	213	258
Fuel	367	662	376
Gas	293	431	355
Green	333	417	299
Greenhouse	102	148	139
Harm	97	357	330
Harmful	23	7	129
Hazard	56	105	166
Logistics	132	276	103
Material	1212	1914	2342

DOI: https://doi.org/10.5526/esj.457

4. **Airline Industry-wise Data Structure:**



Keywords	Annual Report and Sustainability Report - Year (2019-2022)		
Big Data Analytics	Delta 2019-2022	Emirates 2019-2022	United 2019-2022
ML	0	0	0
Algorithm	24	1	0
Analysis	262	133	259
Analytical	11	6	15
Analytics	19	14	16
Application	232	46	67
Artificial	28	7	0
AWS	4	0	0
Big Data	8	0	0
Blockchain	12	3	5
Bot	0	0	0
Business Intelligence	0	0	0
Cloud	61	4	4
Cluster	7	3	5
Computer	42	44	80
Cyber	48	48	73
Cybersecurit	33	23	42
Data	1062	242	1147
Data Management	3	0	1
Data Mining	0	0	0
Database	49	4	18
Data-Driven	10	2	0
dataSet	0	0	0
dataSets	0	0	0
Digital	463	251	155
Driven	98	57	43
Electronic	137	17	80
Ethical	32	37	29
Evaluat	653	166	339
Information	1504	2028	2136
Innovation	186	109	50
Integration	99	19	50
Intelligence	34	18	0
Java	4	3	0

Keywords	Annual Report and Sustainability Report - Year (2019-2022)		
Environmental Sustainability Supply Chain	Delta 2019-2022	Emirates 2019-2022	United 2019-2022
Carbon	1127	108	252
Climate	1029	58	219
CO2	871	215	241
Combustion	89	0	4
Conservation	41	44	34
Consumption	368	266	243
Corporate	959	308	1353
Earth	14	2	0
Eco	0	1	0
EcoFriendly	0	1	0
Eco-Friendly	3	6	5
Economy	322	142	124
Electric	844	143	124
Emission	2346	439	722
Emissions	2126	415	671
Energy	1179	237	336
Environment	2409	575	632
Environmental	1676	343	426
Sustainable	615	403	796
Transport	552	268	355
EPA	66	2	20
ESG	204	231	949
Farmer	11	38	10
Farmers	11	35	10
Footprint	233	57	62
Fuel	1423	394	789
Gas	366	59	60
Green	395	182	282
Greenhouse	186	34	37
Harm	147	51	89
Harmful	6	2	47
Hazard	221	35	69
Logistics	223	47	20
Material	2810	514	1557

5. **Data Filtering**

=VLOOKUP($A10,'B:\ RESOURCES\[Considered Keywords.xlsx]Tesla'!$A$1:$E$66,COLUMN(D9),0)

=SUM(B12:B29)

Appendix B: Company-Level Visualisations

1. **WordCloud: Airline**

install.packages("wordcloud")

library(wordcloud)

data <- data.frame(

Term = c(

"ML", "Algorithm", "Analysis", "Analytical", "Analytics",

DOI: https://doi.org/10.5526/esj.457

"Application", "Artificial", "AWS", "Big Data", "Blockchain",

"Bot", "Business Intelligence", "Cloud", "Cluster", "Computer",

"Cyber", "Cybersecurity", "Data", "Data Management", "Data Mining",

"Database", "Data-Driven", "dataSet", "dataSets", "Digital",

"Driven", "Electronic", "Ethical", "Evaluat", "Information",

"Innovation", "Integration", "Intelligence", "Java", "Machine",

"Machine Learning", "Mathematic", "Maths", "Method", "Mining",

"Predict", "Prediction", "Predictive", "Privacy", "Processing",

"Program", "Programming", "Robot", "Robotics", "Science", "Securities",

"Security", "Software", "Strategy", "Surveillance", "Technical",

"Technique", "Techniques", "Technolog", "Technology", "Tool",

"Virtual", "Visualisation", "Volume", "Carbon", "Climate", "CO2",

"Combustion", "Conservation", "Consumption", "Corporate", "Earth",

"Eco", "EcoFriendly", "Eco-Friendly", "Economy", "Electric", "Emission",

"Emissions", "Energy", "Environment", "Environmental", "Sustainable",

"Transport", "EPA", "ESG", "Farmer", "Farmers", "Footprint", "Fuel",

"Gas", "Green", "Greenhouse", "Harm", "Harmful", "Hazard", "Logistics",

"Material", "Mitigation", "Natural", "Nature", "Nitrogen", "Planet",

"Plant", "Pollut", "Pollution", "Power plant", "Process", "Procurement",

"Production", "Protection", "Recycle", "Reduce", "Renewable", "Renewal",

"Resources", "Reuse", "SCM", "Social", "Solar", "SSCM", "Supplier",

"Suppliers", "Supply", "Supply Chain", "Wastage", "Sustainability", "Waste"

),

Freq = c(

DOI: https://doi.org/10.5526/esj.457

```
0, 25, 654, 32, 49, 345, 35, 4, 8, 20, 0, 0, 69, 15, 166, 169, 98, 2451, 4, 0, 71, 12, 0, 0, 869, 198, 234,
98, 1158, 5668, 345, 168, 52, 7, 136, 10, 0, 0, 1053, 246, 227, 7, 10, 243, 149, 3653, 22, 16, 9, 153,
915, 1110, 580, 1363, 1, 308, 139, 125, 1372, 802, 360, 112, 0, 414,

1487, 1306, 1327, 93, 119, 877, 2620, 16, 1, 1, 14, 588, 1111, 3507, 3212, 1752, 3616, 2445, 1814,
1175, 88, 1384, 59, 56, 352, 2606, 485, 859, 257, 287, 55, 325, 290, 4881, 78, 292, 574, 44, 555, 720,
81, 41, 14, 2034, 355, 1042, 615, 220, 1356, 527, 93, 462, 101, 4, 1423, 68, 0, 1561, 1190, 1061, 683,
0, 3234, 1207

)

)

set.seed(123)

wordcloud(words = data$Term, freq = data$Freq, scale = c(3, 0.5), colors = brewer.pal(3, "Dark2"))


filtered_data <- data[data$Term != "Information", ]


set.seed(123)

par(plt = c(0.1, 1, 0.1, 1))

wordcloud(

words = filtered_data$Term,

freq = filtered_data$Freq,

scale = c(1.5, 0.5),

colors = brewer.pal(8, "Dark2"),

min.freq = 1,  # Minimum frequency for inclusion

random.order = FALSE,  # Disable random word order

rot.per = 0,  # Disable word rotation

max.words = 1500  # Adjust the maximum number of words displayed

)
```


DOI: https://doi.org/10.5526/esj.457

2. WordCloud Automotive:

```
library(tm)

library(wordcloud)

data <- data.frame(

Term = c(

"ML", "Algorithm", "Analysis", "Analytical", "Analytics",

"Application", "Artificial", "AWS", "Big Data", "Blockchain",

"Bot", "Business Intelligence", "Cloud", "Cluster", "Computer",

"Cyber", "Cybersecurity", "Data", "Data Management", "Data Mining",

"Database", "Data-Driven", "dataSet", "dataSets", "Digital",

"Driven", "Electronic", "Ethical", "Evaluat", "Innovation",

"Integration", "Intelligence", "Java", "Machine",

"Machine Learning", "Mathematic", "Maths", "Method", "Mining",

"Predict", "Prediction", "Predictive", "Privacy", "Processing",

"Program", "Programming", "Robot", "Robotics", "Science", "Securities",

"Security", "Software", "Strategy", "Surveillance", "Technical",

"Technique", "Techniques", "Technolog", "Technology", "Tool",

"Virtual", "Visualisation", "Volume", "Carbon", "Climate", "CO2",

"Combustion", "Conservation", "Consumption", "Corporate", "Earth",

"Eco", "EcoFriendly", "Eco-Friendly", "Economy", "Electric", "Emission",

"Emissions", "Energy", "Environment", "Environmental", "Sustainable",

"Transport", "EPA", "ESG", "Farmer", "Farmers", "Footprint", "Fuel",

"Gas", "Green", "Greenhouse", "Harm", "Harmful", "Hazard", "Logistics",

"Material", "Mitigation", "Natural", "Nature", "Nitrogen", "Planet",

"Plant", "Pollut", "Pollution", "Power plant", "Process", "Procurement",
```


DOI: https://doi.org/10.5526/esj.457

```
"Production", "Protection", "Recycle", "Reduce", "Renewable", "Renewal",

"Resources", "Reuse", "SCM", "Social", "Solar", "SSCM", "Supplier",

"Suppliers", "Supply", "Supply Chain", "Wastage", "Sustainability", "Waste"

),

Freq = c(

0, 79, 868, 32, 51, 630, 126, 12, 62, 67, 7, 0, 176, 32, 122, 205, 92, 4075, 13, 0, 136, 40, 0, 0, 1483,
443, 396, 131, 1089, 416, 270, 114, 14, 384, 28, 0, 0, 1245, 655, 166, 4, 13, 168, 289, 2762, 47, 130,
56, 157, 896, 2453, 1304, 2167, 6, 623, 102, 88, 2349, 998, 534, 167, 1, 1286,

2148, 1503, 3116, 296, 235, 798, 2799, 134, 28, 0, 30, 1072, 2941, 6059, 5360, 4555, 6708, 4980,
1895, 1315, 374, 338, 8, 4, 564, 1405, 1079, 1049, 389, 784, 159, 327, 511, 5468, 188, 474, 874, 83,
302, 1774, 382, 232, 68, 4115, 580, 4206, 1252, 520, 2088, 947, 159, 975, 211, 0, 1582, 3429, 0,
3320, 2460, 3037, 2097, 0, 3973, 1486

)

)

set.seed(123)

par(plt = c(0.1, 1, 0.1, 1))

wordcloud(

words = data$Term,

freq = data$Freq,

scale = c(1.5, 0.5),  # Increase the scaling for larger words

colors = brewer.pal(20, "Dark2"),

min.freq = 1,

random.order = FALSE,

rot.per = 0,

max.words = 500  # You can adjust this value to display more words

)
```


DOI: https://doi.org/10.5526/esj.457

3. Trend Analysis Airline-

```
import pandas as pd

import matplotlib.pyplot as plt


# Create a DataFrame from your new data

data = {

    'Name': ['Delta', 'Delta', 'Delta', 'Delta', 'Emirates', 'Emirates', 'Emirates', 'Emirates', 'United',
'United', 'United', 'United'],

    'Year': [2019, 2020, 2021, 2022, 2019, 2020, 2021, 2022, 2019, 2020, 2021, 2022],

    'BDA_Count': [2980, 3309, 2992, 2985, 1336, 1461, 1445, 1366, 1516, 1744, 1741, 3685],

    'ESSCM_Count': [8712, 9187, 8904, 8488, 1782, 2267, 2323, 2650, 2372, 2640, 2928, 6447]

}


df = pd.DataFrame(data)


# Plot the trends for each company and category

fig, axes = plt.subplots(nrows=2, ncols=1, figsize=(10, 8))


for company, color in zip(df['Name'].unique(), ['blue', 'orange', 'green']):

    company_data = df[df['Name'] == company]

    axes[0].plot(company_data['Year'], company_data['BDA_Count'], label=company, marker='o',
linestyle='-', color=color)

    axes[1].plot(company_data['Year'], company_data['ESSCM_Count'], label=company,
marker='o', linestyle='-', color=color)


axes[0].set_title('BDA Trends')

axes[1].set_title('ESSCM Trends')
```


DOI: https://doi.org/10.5526/esj.457

```
axes[0].legend()

axes[1].legend()

plt.tight_layout()

plt.show()
```

4. Trend Analysis Automotive-

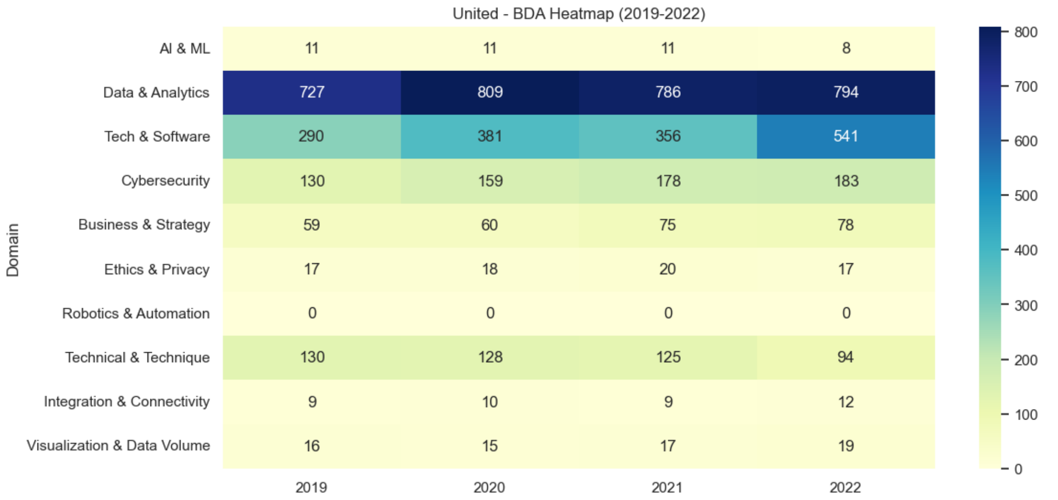
```
import pandas as pd

import matplotlib.pyplot as plt


# Create a DataFrame from your data

data = {

    'Company': ['Volkswagen', 'Volkswagen', 'Volkswagen', 'Volkswagen', 'Toyota', 'Toyota', 
'Toyota', 'Toyota', 'Tesla', 'Tesla', 'Tesla', 'Tesla'],

    'Year': [2019, 2020, 2021, 2022, 2019, 2020, 2021, 2022, 2019, 2020, 2021, 2022],

    'BDA_Count': [2096, 2153, 3694, 2352, 3064, 3159, 2996, 2937, 2968, 3014, 2935, 2896],

    'SupplyChain_Count': [4727, 5392, 8807, 7086, 12702, 9748, 10122, 7672, 9574, 8621, 7804, 
7940]

}


df = pd.DataFrame(data)


# Plot the trends for each company and category

fig, axes = plt.subplots(nrows=2, ncols=1, figsize=(10, 8))


for company, color in zip(df['Company'].unique(), ['blue', 'orange', 'green']):

    company_data = df[df['Company'] == company]
```

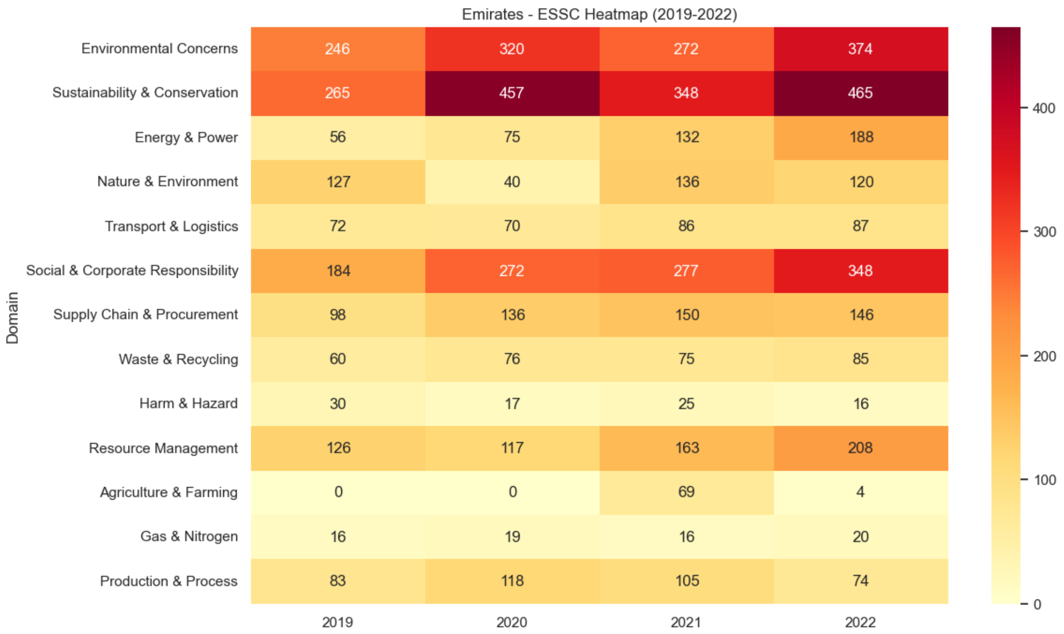

DOI: https://doi.org/10.5526/esj.457

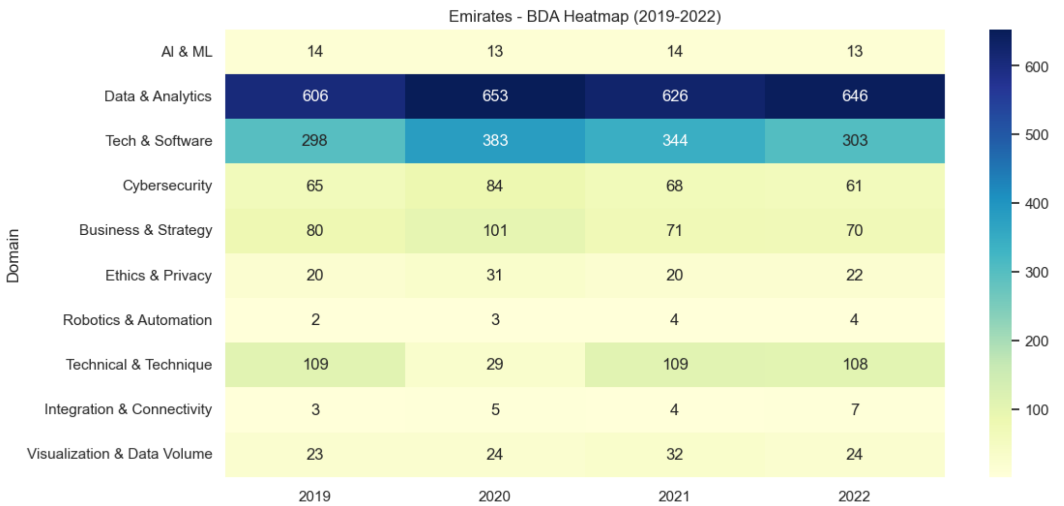
```
    axes[0].plot(company_data['Year'], company_data['BDA_Count'], label=company, marker='o', linestyle='-', color=color)

    axes[1].plot(company_data['Year'], company_data['SupplyChain_Count'], label=company, marker='o', linestyle='-', color=color)


axes[0].set_title('BDA Trends')

axes[1].set_title('ESSCM Trends')

axes[0].legend()

axes[1].legend()

plt.tight_layout()

plt.show()
```


DOI: https://doi.org/10.5526/esj.457

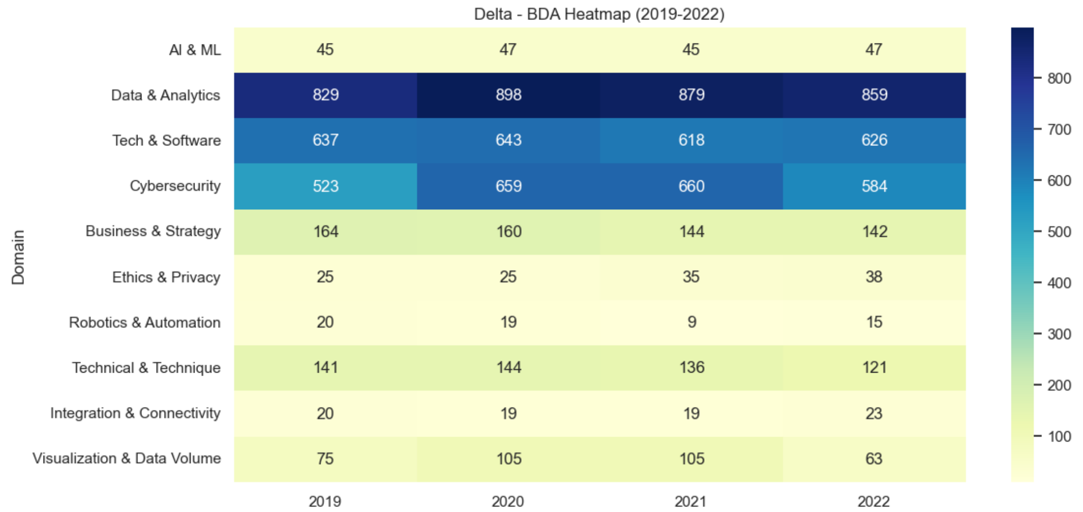
5. Company-wise Heatmaps –



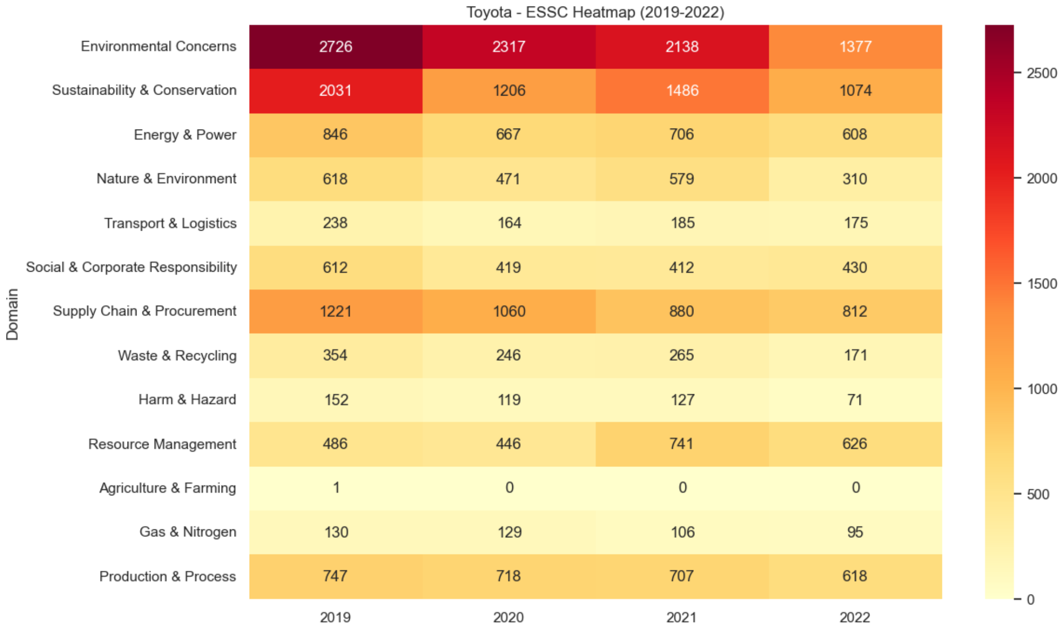


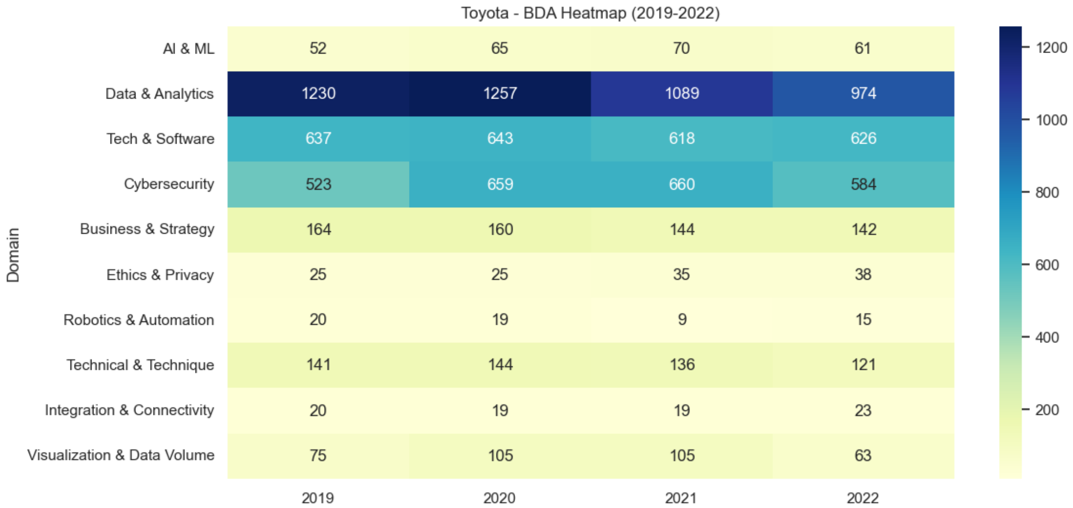


DOI: https://doi.org/10.5526/esj.457

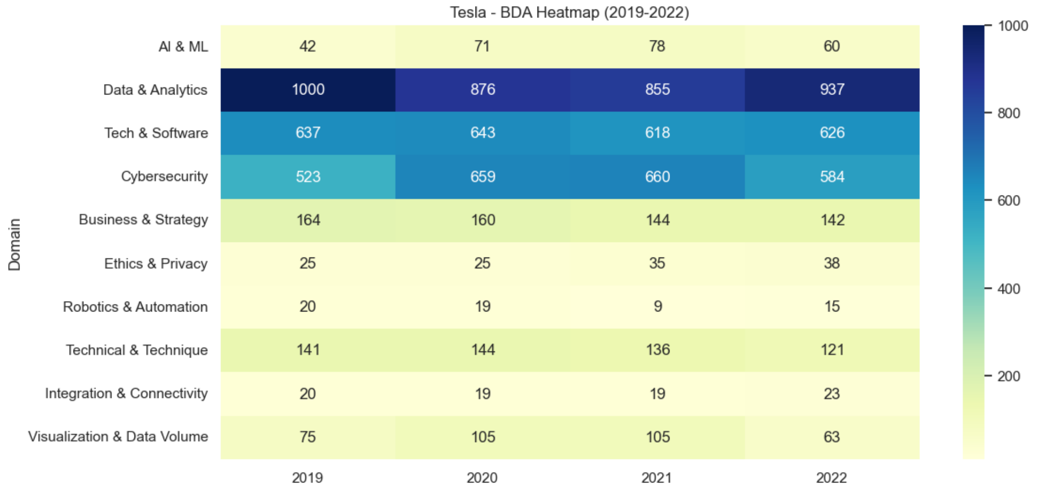


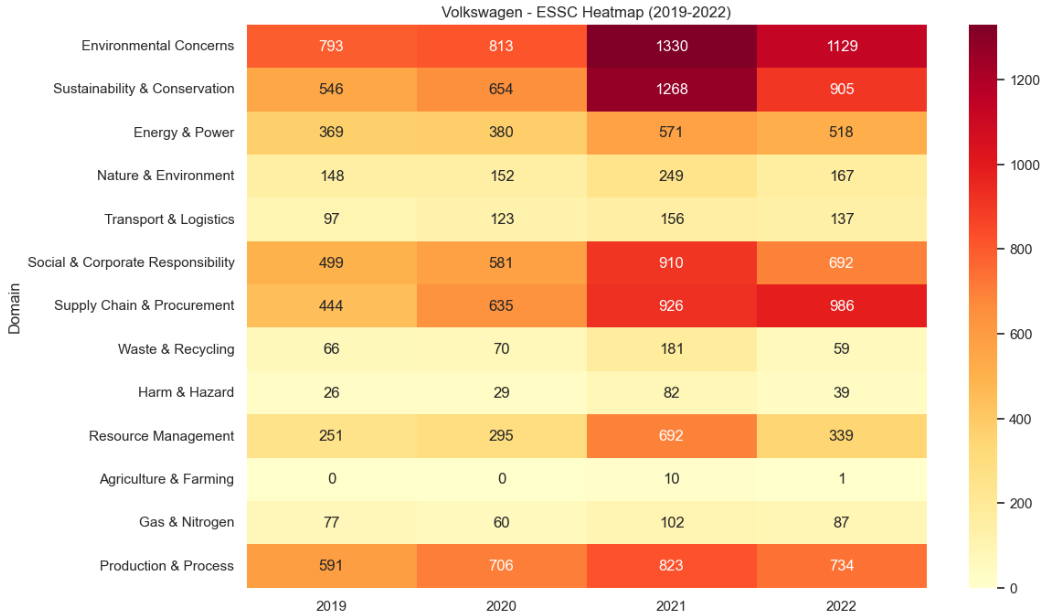
DOI: https://doi.org/10.5526/esj.457



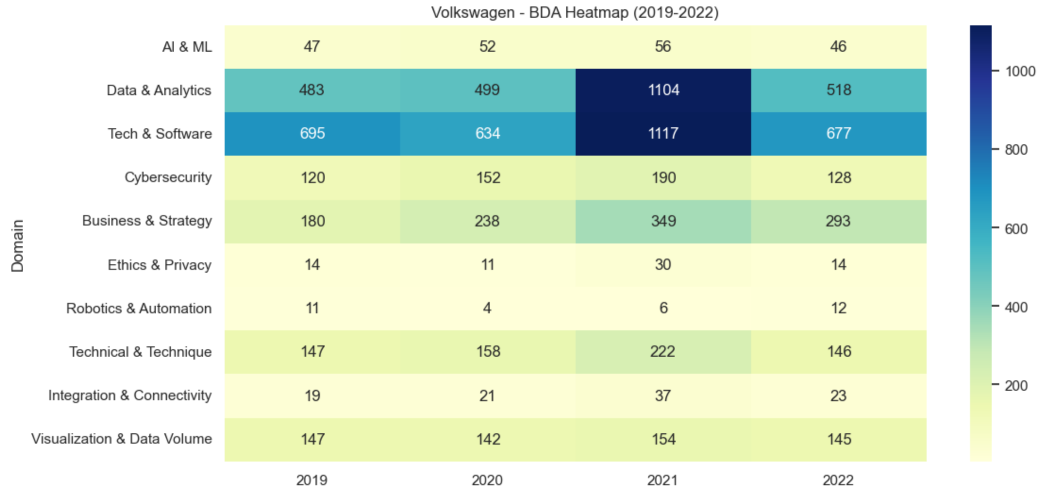


DOI: https://doi.org/10.5526/esj.457



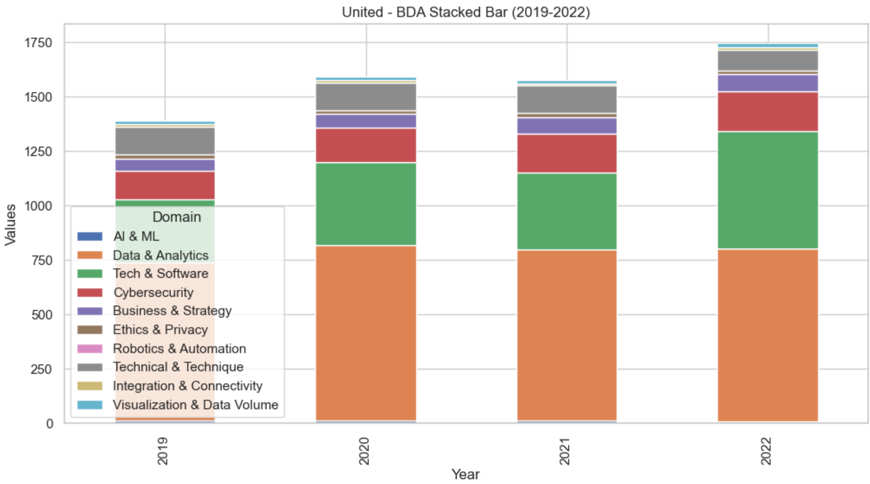


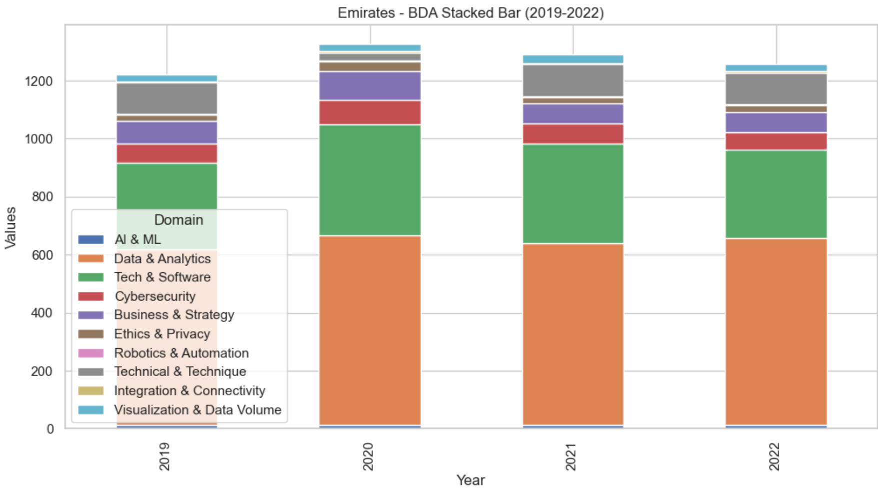
DOI: https://doi.org/10.5526/esj.457

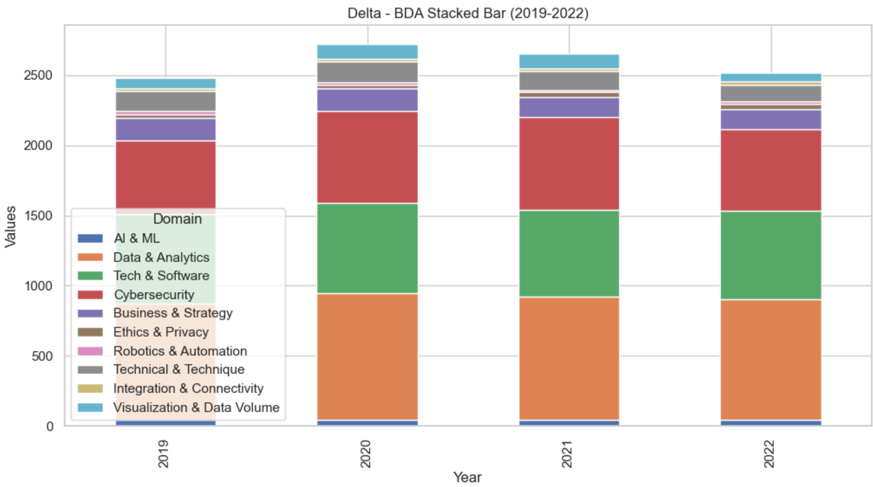


DOI: https://doi.org/10.5526/esj.457

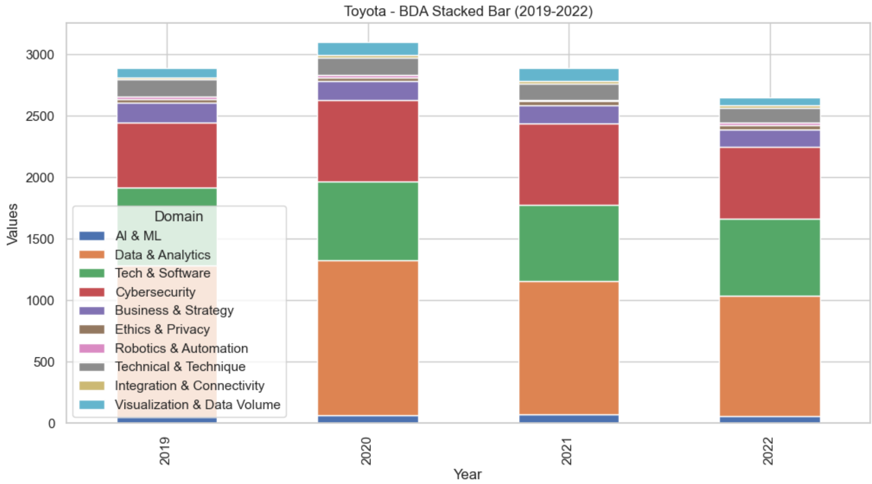
6. Company-wise Stacked charts –

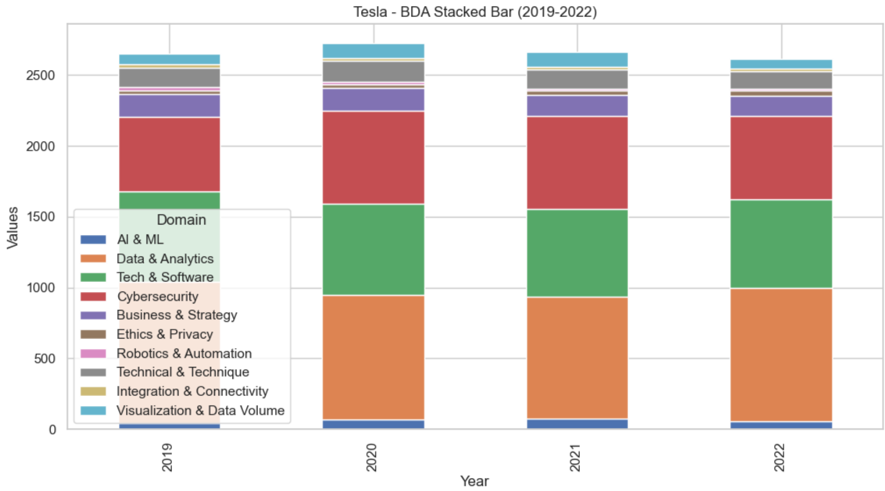


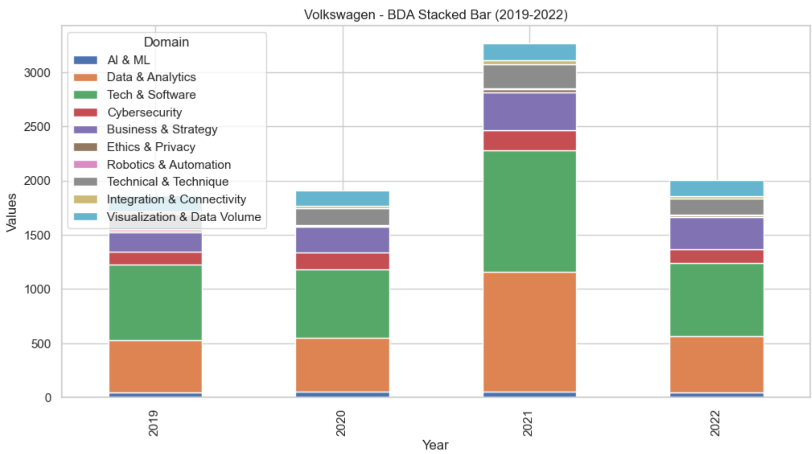




DOI: https://doi.org/10.5526/esj.457

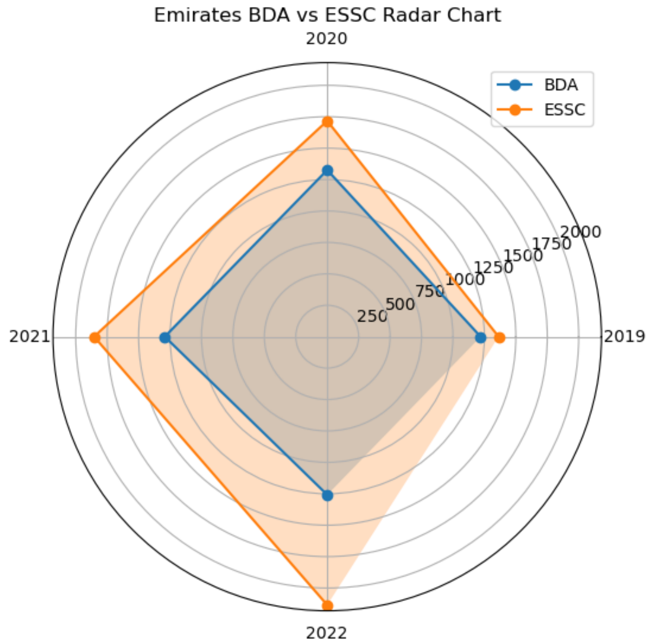


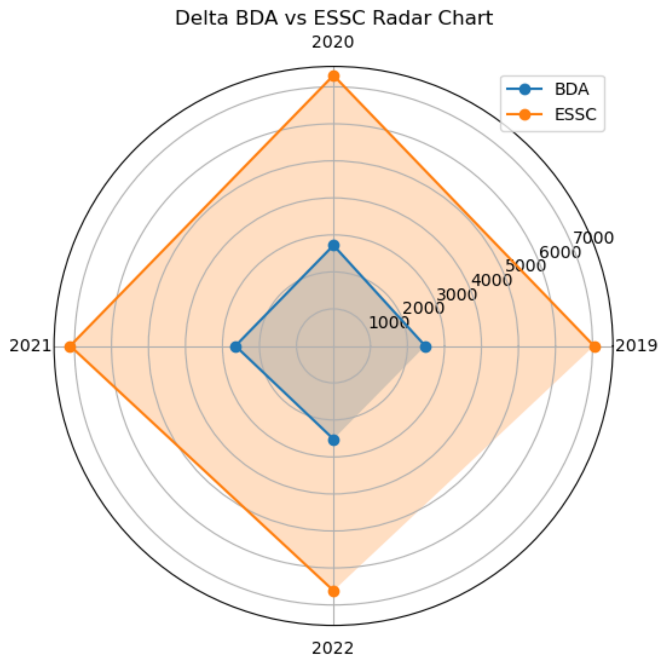




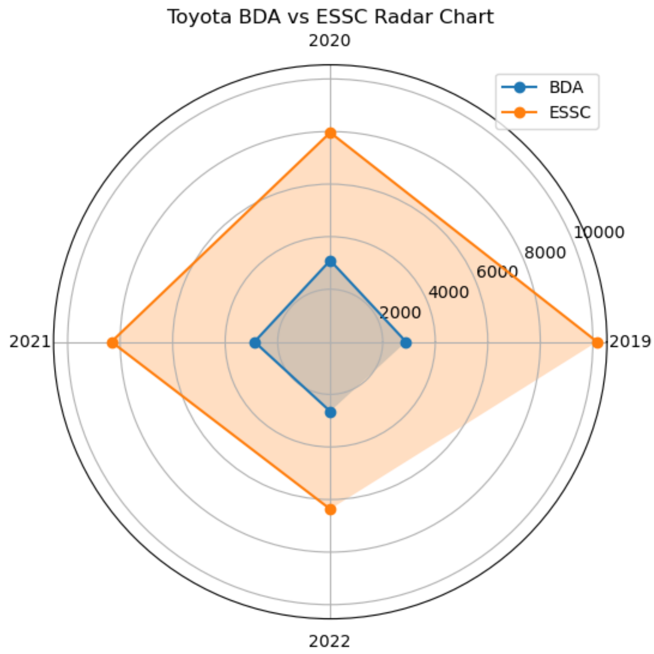
DOI: https://doi.org/10.5526/esj.457

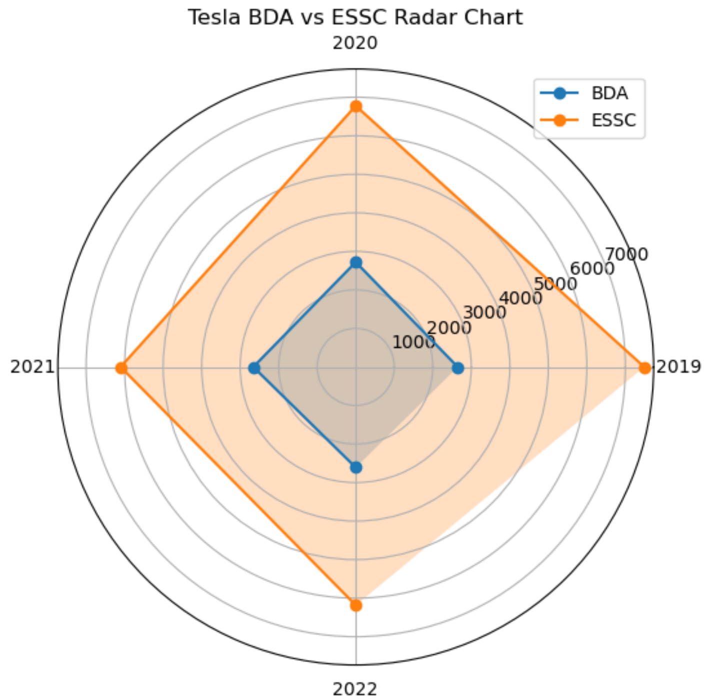
7. Company-wise Individual Radar for Airline and Automotive Industry:



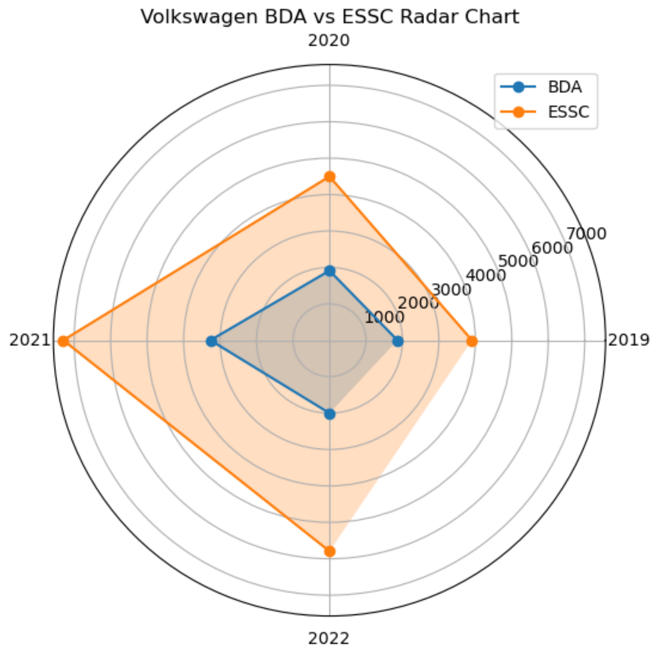


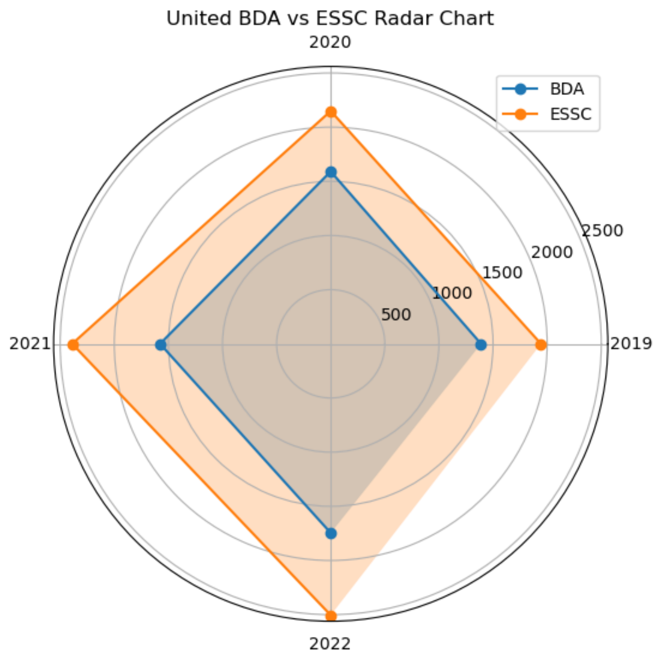
DOI: https://doi.org/10.5526/esj.457





DOI: https://doi.org/10.5526/esj.457





DOI: https://doi.org/10.5526/esj.457

Appendix C: Data Sources – Company Reports (2019–2022)

1. Automotive Firms

- Volkswagen Annual Reports (2019–2022):
  https://annualreport2019.volkswagenag.com/
  https://annualreport2020.volkswagenag.com/
  https://www.volkswagen-newsroom.com/en/publications/corporate/annual-report-2021-835
  https://annualreport2022.volkswagenag.com/

- Volkswagen Sustainability Reports (2019–2022):
  https://www.volkswagen-group.com/en/publications/corporate/sustainability-report-2019-1668
  https://www.volkswagen-group.com/en/publications/corporate/sustainability-report-2020-1667
  https://www.volkswagen-group.com/en/publications/corporate/sustainability-report-2021-1947
  https://www.volkswagen-group.com/en/publications/more/group-sustainability-report-2022-1644

- Toyota Annual Reports (2019–2022):
  https://www.toyota-industries.com/investors/items/2020_annual_financial_report_E.pdf
  https://www.toyota-industries.com/investors/items/2021_annual_financial_report_E.pdf
  https://www.toyota-industries.com/investors/item/2022_annual_financial_report_E.pdf
  (Toyota 2019 annual in sustainability doc link)

- Toyota Sustainability Reports (2019–2022):
  https://global.toyota/pages/global_toyota/sustainability/report/sdb/sdb19_en.pdf
  https://global.toyota/pages/global_toyota/sustainability/report/er/er20_en.pdf
  https://global.toyota/pages/global_toyota/sustainability/report/sdb/sdb21_en.pdf
  https://global.toyota/pages/global_toyota/sustainability/report/sdb/sdb22_en.pdf

- Tesla Annual Reports (2019–2022):
  https://ir.tesla.com/financial-information/sec-filings

- Tesla Sustainability (Impact) Reports (2019–2022):
  https://www.tesla.com/ns_videos/2020-tesla-impact-report.pdf
  https://www.tesla.com/ns_videos/2021-tesla-impact-report.pdf
  https://www.tesla.com/ns_videos/2022-tesla-impact-report-highlights.pdf

DOI: https://doi.org/10.5526/esj.457

1. Airline Firms

- Delta Annual Reports (2019–2022): https://www.annualreports.com/Company/delta-air-lines-inc
- Delta Sustainability (ESG) Reports (2019–2022): https://news.delta.com/delta-releases-2022-esg-report
- United Airlines Annual Reports (2019–2022): https://ir.united.com/financials/sec-filings
- United Sustainability Reports: https://crreport.united.com/documents/United-Corporate-Responsibility-summary.pdf
- Emirates Annual Reports (2019–2022): https://c.ekstatic.net/ecl/documents/annual-report/2022-2023.pdf
- Emirates Sustainability/ESG Reports (2019–2022): https://www.emiratesnbd.com/-/media/enbd/files/investor-relations/common-pdf/esg_report_2022.pdf

Copyright Statement

DOI: https://doi.org/10.5526/esj.457