THE INTERNATIONAL JOURNAL OF SCIENCE & TECHNOLEDGE

Leveraging Data Analytics for Environmental Sustainability

Divyeshkumar Vaghani
Student, Department of Business Analytics, Gannon University, US

Abstract:
This research work is an assessment of how data analytics contributes to environmental sustainability. The objectives of the study were:

- *To determine how data analytics affects environmental sustainability,*
- *To find out barriers to data analytics and environmental sustainability, To discover opportunities for data analytics and environmental sustainability, and*
- *To ascertain how data analytics aligns with environmental conservation goals and values.*

This research adopts a quantitative questionnaire approach as the method of data collection. Questionnaires were administered to 50 persons; the study used descriptive and inferential statistics to analyse the data gathered, and regression analysis was used to generate inferences from the data, using Statistical Packages for Social Sciences (SPSS). The regression model shows that there is an overall statistical significance between the independent variables (effect of data analytics, barriers of data analytics, and opportunities of data analytics) and the dependent variable (environmental sustainability with a corresponding p-value of 0.038). Therefore, the study recommends that organisations and governments invest in robust data analytics infrastructure to effectively collect, process and analyse environmental data. Efforts should be made to address the barriers identified in the analysis, such as implementation challenges, data quality issues, and scalability concerns, and enhancing data literacy and building analytical skills among environmental professionals is crucial for maximising the potential of data analytics.

Keywords: *Analytical skills, barriers, big data, climate data model, data quality, environmental conservation, environmental data, opportunities*

1. Introduction

Data analytics is playing a revolutionary role in the manufacturing, retail, and healthcare sectors. Data analytics generally refers to the ability of society to use information in new ways to generate useful insights of substantial value (Mayer-Schonberger & Cukier, 2023). It has been termed by Mckinsey Global Institute (2021) as "the next cutting-edge technology for productivity, competition and innovation". Mayer-Schonberger and Cukier narrated the story of Google using search queries to visualise the spread of flu in the USA, pointing out one of the numerous ways in which data analytics is transforming healthcare. They also talked about the popular data analytics story from the retailer target, who can calculate customer pregnancies based on purchasing habits (Mayer-Schonberger & Cukier, 2023).

Although data analytics has been gaining popularity, it has not yet received a concrete definition (Dumbill, 2023). One of the most famous definitions includes the three "Vs" of data organisation: velocity, volume, and variety (Laney, 2021). Velocity means an increased speed of data formation and its use and interface, volume means the increased quantity of data to be organised, and variety refers to the numerous and often incompatible data layouts, structures, and semantics (Laney, 2021). However, over the years, more "Vs" have been added, which include visualisation, value, and veracity (Sowe, 2024).

McKinsey Global Institute (2021) defines data analytics as “the ability to capture, store, manage and analyse datasets whose size is beyond the ability of typical database software tools". MGI explains that data analytics arrangement changes not only over time due to technological advancements but also across sectors because of the variation in dataset software, size, and tools found in each sector (MGI, 2021). Simon (2023) asserts that there is no perfect definition of data analytics as the purposes of different companies shape the definition of data analytics.

Even though data analytics has been seen to have different definitions, a settlement has been reached on its potential to shape the decision-making process (Dumbill, 2023). Focus has been shifted to data analytics due to the rapid increase in the amount of information available; this information has been gathered through various media such as cell phones, computers, and information technology (IT) systems. MGI states that furthering data analytics has the potential to increase the productivity of different sectors by at least 0.5% a year through 2024; therefore, growing operating margins among individual companies by more than 60% are visible if they remain at the forefront of the data analytics movement.
MGI also forecasted the insightful potential importance of public sector applications of data analytics, indicating that the European Union could reduce costs for administrative activities by up to 20%, and annual productivity growth could be increased by 0.5% over the next ten years. Tambe (2023) discovered that a higher utilisation of data analytics technologies

could equate to 3% higher productivity than the average firm and vice versa for a lower utilisation of data analytics (Provost & Fawcett, 2023).

Environmental sustainability is largely not yet part of the popular lexicon of data analytics in action. This study intends to address this gap. However, the environmental sustainability effort, which generally creates and maintains the situation under which humans and their natural environment can exist in industrious harmony, allows for the fulfillment of the economic, social and other necessities of current and future generations needs a revolution of its own.

Climatic change continues to be at the front line of world-scale challenges. In its latest report, the Intergovernmental Panel on Climatic Change (IPCC) talked about the undeniable warming of the climatic system and noted the main role that human activities have played in promoting this change (IPCC, 2023). Furthermore, the condition of the ecosystem is declining, as indicated by the Millennium Ecosystem Assessment (2015), stating that human beings have severely altered the world's ecosystem, which has resulted in substantial and essentially irreversible loss in the variety of life on Earth.

Remarks that call for a strengthened effort to address concerns of environmental sustainability have been made in light of the new ways of data analytics and the insights that can be generated. The founder of the consulting firm SustainAbility & Volans, John Elkington, highlighted that data analytics and how we use data and numbers will play a pivotal role in sustainable outcomes in what he called a "Breakthrough Decade" (Elkington, 2024).

How, then, is data analytics utilised to strengthen environmental sustainability? What revolutionary data analytics applications are important to the environment as they have shown similar usefulness in the healthcare sector? How effective have the data analytics applications been to environmental sustainability as they have been to the retailer? This paper investigates the usefulness of data analytics in the realm of environmental sustainability across various sectors and companies. By gaining insights from sustainability sponsors (NGOs), service providers, policy specialists, compliers, and finally governments, I aim to provide a multi-perspective view of:

- How data analytics affects environmental sustainability;
- The barriers to data analytics and environmental sustainability;
- The opportunities for data analytics and environmental sustainability;
- How data analytics aligns with environmental conservation goals and values.

This paper is important because breadth on the topic is lacking, while other areas are gaining extensive coverage. Data analytics and environmental sustainability have been joined and featured with complexity. Graham (2023) states that because data can often be mapped, measured, and analysed, data analytics possesses the characteristics to produce profoundly new ways of knowing, acting, and being part of the world. Eventually, by developing an existing complexity, I look to create additional discussions on data analytics applications to environmental sustainability, adding more flesh to the topic and bringing it closer to parity with other fields.

2. Literature Review

2.1. Data Analytics

Data analytics can be defined as the practice of processing vast amounts of unstructured, diverse, and complex data that cannot be processed with traditional analytical tools; it needs more enhanced technology to scrutinise and extract meaningful insights and make cognizant decisions (Appelbaum, 2016). Data analytics is a field that includes data, analytical technologies, and methods to get the needed insights from the data to improve business performance, deliver sustainable value, and provide competitive advantage (Wamba et al., 2017). In the past few years, data analytics has gained considerable interest in several domains. However, data in the context of environmental sustainability has a very limited legacy in the existing literature.

Data analytics is seen as the creation of new technologies and architectures structured to economically derive value from a large amount of data by making discovery and/or analysis (Mikalef, Pappas & Krogstie, 2018). Data analytics has transformed climatic science, proposing unprecedented prospects for advanced modeling and analysis. Hansen et al. (2020) noted the role of remote sensing machinery and satellite observations, showcasing their usefulness in enriching climate datasets. Rasp et al. (2018) studied machine learning methods, noting their potential in deciphering core climatic patterns from numerous datasets. The nuanced perspective provided by Stephens et al. (2022) grounds the significance of different data sources, including ground-based radars and citizen science initiatives.

Climatic change, a multifaceted global problem, needs improved technologies and methodologies for precise modeling and analysis. In recent years, integrating data analytics into climate change research has been seen as a revolutionary force, paving the way for new avenues of understanding and checkmating environmental sustainability. Data in the context of climatic studies means the many and diverse datasets obtained from different channels that deal with different areas of Earth's climatic system. These datasets are recognised by their velocity, variety, volume and complexity. Velocity pertains to the speed at which data is generated and processed, variety deals with the different types of data, which includes sensor readings, satellite observations, climate models, and socio-economic information, and volume refers to the massive number of data continuously generated. The complexity comes from the mutual interconnectedness within the climatic system. In climate change research, data analytics deals with information at different scales – from global climate models down to local and regional interpretations. The incorporation of these different data sources makes way for a more nuanced understanding of climate dynamics and their usefulness in the ecosystem, economies and societies.

There have been some studies on data and sustainability in the automobile industry (Bughin, Chui, & Mantika, 2020), whereas other studies either provide an anecdotal view or offer a limited depth of scrutiny. Some research addressed the usefulness of data analytics on environmental sustainability (Koo, Piratla & Matthews, 2015; Koseleva & Ropaite, 2017; Lokers et al., 2016). Other studies have investigated the relationship between data analytics and the financial performance of companies (Gunasekaran et al., 2017; Wamba et al., 2017). Only a few of these studies examine how data analytics can be leveraged to enhance environmental sustainability (Hazen et al., 2016; Jeble et al., 2018).

2.2. Environmental Sustainability

The history of environmental sustainability stems from the works of early visionaries. Plass (2016) conducted groundbreaking research that laid the foundation for understanding the greenhouse effect, while Manabe and Withrald (2015) introduced General Circulation Models (GCMs) that marked an important step in enhancing complex climate systems. Hansen et al. (1983) went further to advance these models by contributing to our changing comprehension of climate dynamics.

Sustainability is commonly referred to as commercial practices that meet the current needs without jeopardising the needs of upcoming generations (World Commission on Environment and Development, WCED, 2017). Sustainability has been utilised to balance and take responsibility for development and commercial activities (Hakovirta & Denuwara, 2020). It focuses on three main pillars: environmental, social and economic advancements (Govindan et al., 2024). These pillars have been considered widely as the foundation for developing many sustainability principles.

Environmental sustainability has attracted much attention recently due to growing concerns arising from climate change, global warming and scarcity of natural resources (Caniato et al., 2022; Fransoo, Gunther & Jammernegg, 2024). The urgency of dealing with climatic change and promoting environmental sustainability cannot be overemphasised. Global warming, more frequent and harsh weather conditions, and shifts in ecosystems pose reasonable threats to food security, biodiversity and human well-being. The resultant effects of inaction are dire, emphasising the need for instant and efficient measures to mitigate and adapt to these changes.

Environmental sustainability includes a complete approach to stabilising the needs of the present without affecting the capacity of the upcoming generations to meet their own needs. Environmental sustainability serves as a linchpin in this quest to provide the scientific bedrock upon which sustainable policies and practices can be articulated. The urgency of dealing with climate change is shown in the comprehensive assessments by the Intergovernmental Panel on Climate Change (IPCC). The Fifth Assessment Report (IPCC, 2024) shows scientific findings, pin-pointing the growing risks linked to global warming. Stern (2017) carried out research in environmental economics, which provides a distinct perspective on the economic need for sustainable practices and further states the need for significant global action.

2.3. Harnessing Data Analytics for Environmental Sustainability

There has never been a more pressing need for environmental sustainability in the contemporary world. Organisations are facing great pressure to implement eco-friendly methods as the repercussions of climatic change have become more glaring. Luckily, improvements in data analytics are providing organisations with effective technologies to boost sustainability initiatives.

Data analytics has emerged as a vital instrument for driving sustainability efforts. Reducing and regulating carbon footprints is one of the vital parts of environmental sustainability for companies. Quantifying and observing the greenhouse gas emissions caused by an organisation's operations is called carbon accounting. Data analytics is essential here because it helps organisations to correctly coordinate, evaluate and report their data on carbon emissions. Organisations can choose what aspect of their operation contributes more substantially to carbon emissions by using modernised analytical approaches. They can establish aggressive reduction objectives successfully and prioritise mitigation techniques. Furthermore, monitoring real-time emissions is made easier by data analytics, which helps organisations promptly modify their procedures and investments.

Figure 1: Data Analytics in Environmental Sustainability
Adopted from Sumanth (2020)

Data analytics supports adherence to voluntary reporting principles and legal duties regarding carbon emissions. By automating the assemblage and analysis of emissions data, organisations can ensure accuracy and openness in their operations and improve their reporting techniques. This can shape the organisation's environmental responsibility culture while also growing its reputation with investors.

Comprehensive analyses, also known as Life Cycle Assessments, are fabricated to examine a product or service's environmental effects at every stage of its life cycle, from the collection of raw materials to disposal. By analysing a great amount of data from different sources, data analytics help organisations perform life cycle assessments (LCAs) more specifically and efficiently. Organisations can look for new ways to minimise environmental effects at every step of a product's life cycle by using data analytics. For example, they can reduce energy usage, waste generation and emissions by augmenting the procurement of materials, manufacturing processes, packaging and transportation. Organisations can make more environmentally friendly goods and services that are demanded by consumers by integrating Life Cycle Assessment findings into their decision-making procedures.

Supply chains are intertwined webs of organisations that manufacture and distribute products and services. Looking for environmental hazards and areas for advancement entails achieving openness across the supply chain. To enhance supply chain transparency, data analytics helps businesses by collecting and evaluating numerous volumes of data from suppliers, manufacturers, distributors and other stakeholders. With the use of advanced analytical methods like machine learning and data mining, organisations can locate possible environmental hotspots in their supply chains. This helps them cooperate with suppliers to enforce sustainable practices and promote advancements along the whole value chain. Data analytics helps organisations gain the trust of customers and stakeholders who value environmental sustainability by encouraging transparency and accountability.

The predictive analysis forecasts imminent behaviours and patterns by using machine learning approaches, statistical algorithms, and preceding data. Predictive analytics, when utilised in environmental sustainability, helps businesses foresee new opportunities and problems and proactively navigate through them via innovation. Organisations can find ways for sustainable innovation by investigating historical performance data and external factors like regulatory changes and market trends. Predictive analytics assists organisations in anticipating fluctuations in customer preferences towards environmentally pleasant equipment and also helps in making new products to satisfy varying market needs. Predictive models assist organisations in avoiding waste and reducing their environmental effects by enhancing resource allocation and energy management.

A vital part of environmental sustainability is resource coordination, which involves making efficient use of materials, water, energy and other natural resources. Precision resource management is aided by data analytics, giving organisations insights into how best to utilise resources and reduce waste. Organisations can find inefficiencies in the use of their resources and apply concentrated interventions to increase effectiveness by using data-driven optimisation techniques. Predictive maintenance models, for example, may notice equipment breakdowns before they occur and enhance maintenance plants to save energy usage and downtime. Likewise, industrial process enhancement using data analytics can lower material waste and energy usage without reducing product quality.

2.4. Barriers to Data Analytics and Environmental Sustainability

While the incorporation of data analytics holds immense benefits for advancing environmental sustainability research, it also presents several challenges. One key problem is the great volume of data generated, which needs improved computational skills for analysis. The speed of data generation is another problem in terms of real-time analysis, most especially when monitoring rapidly dynamic climate events.

The quality of data and interoperability are key issues, as datasets from different sources could have different resolutions, formats and accuracy levels. So, ensuring compatibility and consistency across different datasets is crucial for making reliable models and arriving at informed decisions.

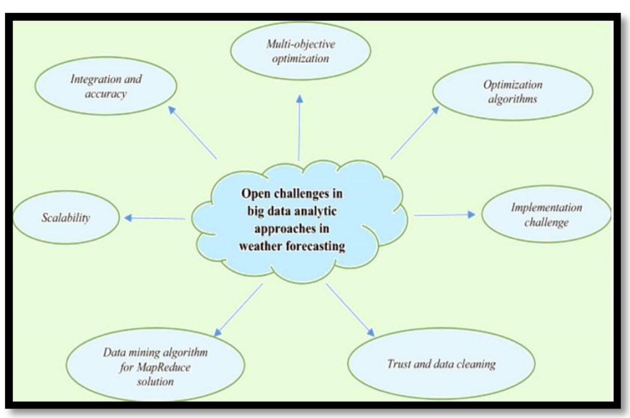

Figure 2: Data Analytics Approach in Weather Forecasting
Adopted from Sumanth (2020)

Environmental sustainability, propelled by the use of data analytics, has undoubtedly changed our understanding of environmental trends and possible impacts. However, the integration of heavy datasets into environmental sustainability research is not without its challenges and limitations. Here, three key limitations are assessed.

2.4.1. Data Quality and Integrity

This is one of the most disturbing challenges in leveraging data analytics for environmental sustainability; as vast amounts of different data are developed into models, inaccuracies and biases may arise. Incoherent data collection techniques, varying measurement standards, and data gaps can jeopardise the reliability of models, thereby leading to mistakes in predictions and policy recommendations. To address these issues, stringent data validation processes must be adopted, protocols must be standardised, and collaboration between stakeholders must be enhanced to get better quality and integrity for the datasets used.

2.4.2. Computational Challenges

The immense volume and intricacy of data in environmental sustainability research demand considerable computational power and resources. Environmental sustainability models often require sophisticated simulations, necessitating high-performance computing abilities. The storage and processing of large datasets demand improved technologies, posing financial and logistical challenges. To overcome this hurdle, investors must invest in groundbreaking technology and effective algorithms and collaborate together to optimise resources. To balance computational demands with environmental sustainability, it is necessary to ensure that the importance of climate modelling does not come at the cost of too much energy consumption.

2.4.3. Ethical Considerations in Environmental Sustainability

The ethical considerations of handling environmental sustainability data cannot be overemphasised. Data analytics has been gaining ground in the study of environmental sustainability, and ethical considerations surrounding data security and transparency become pertinent. Ensuring good data governance, securing sensitive information, and stimulating open access to non-sensitive data are vital elements in handling environmental sustainability data. Checking data openness for collaborative research and guarding individual privacy are delicate tasks that require continuous dialogue, robust policies and loyalty to ethical frameworks.

2.5 Opportunities for Data Analytics and Environmental Sustainability

The field of environmental sustainability continues to transform. Future prospects hold exciting opportunities driven by improvements in data analytics and new technologies. In this section, key areas that are promising to shape the future landscape of environmental sustainability through data analytics will be discussed.

2.5.1. Improvements in Data Analytics for Environmental Sustainability

The continuous modification of data analytics techniques is expected to meaningfully enhance environmental sustainability. Machine learning algorithms, for example, are becoming sophisticated in drawing patterns and insights from numerous datasets, leading to more precise climatic models. Predictive analytics, advanced statistical methods, and data mining play pivotal roles in deciphering complicated climatic relations, giving researchers unprecedented technologies for interpreting and predicting environmental changes. Collaborations between climatic researchers and data scientists are likely to grow, leading to innovative approaches to interpreting and analysing the ever-changing volume of climate-related data.

2.5.2. Prospective Technological Breakthroughs

Breakthroughs in technology on the horizon present promising prospects for changing environmental sustainability modelling. Quantum computing, for example, has the prospect of exponentially increasing computational power, making simulation and analysis more accurate. Moreover, improvements in satellite systems and sensor technologies can provide more real-time data, further changing climatic models. Incorporation of blockchain technology may improve the security, transparency, and traceability of climatic data, addressing key ethical concerns and encouraging trust among stakeholders. The interaction of these technological breakthroughs could drive climatic research into a new realm of efficiency and accuracy.

2.5.3. Implications for Upcoming Environmental Sustainability Research

Upcoming environmental sustainability research could be shaped by data analytics, the integration of technological improvements, and interdisciplinary collaboration. Advanced models will offer stakeholders and policymakers more accurate information for creating targeted strategies to navigate and adapt to climatic change. Real-time observation and prediction abilities will enable quicker responses to harsh weather conditions and long-term climatic trends. Moreover, the availability of reliable, high-quality data will strengthen multinational cooperation and information-sharing, thereby leading to a worldwide approach to addressing climatic issues. As technologies become more intuitive, citizen science initiatives will also progress, engaging communities in the collection and analysis of data to contribute to a more detailed understanding of local and global climate changes.

2.6. Application of Data Analytics in Environmental Sustainability

Case studies show how data analytics has been applied to improve environmental sustainability through climate change analysis. For example, satellite information has been used to monitor deforestation patterns in the Amazon rainforest, initiating timely interventions to curb environmental degradation. Furthermore, the incorporation of economic and social data into environmental sustainability models has facilitated a more nuanced understanding of the interaction between factors affecting climate dynamics. These case studies show the flexibility of data analytics, giving room for more detailed predictions of climatic trends, early identification of natural hazards, and enhanced decision-making in environmental sustainability.

Examining the successful execution of sustainability practices, case studies show the positive outcomes stemming from decisions guided by climatic change models and data analytics in streamlining policies and practices that add to environmental sustainability. From eco-friendly agricultural practices to renewable energy initiatives, these case studies showcase the positive effects of incorporating data-driven insights and sustainable strategies.

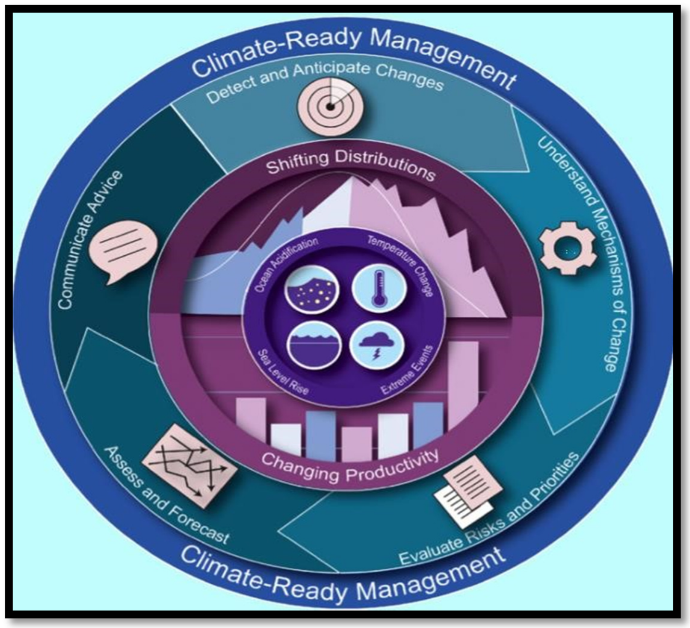

Figure 3: Robust Climate Change Modeling
Adopted from Sumanth (2020)

The real-world effect of using data analytics for environmental sustainability is seen in case studies that highlight positive influences. For example, a case study can illustrate how a community-based observation program powered by data analytics resulted in enhanced water resource management and enhanced resilience to climatic challenges.

These case studies point out the tangible benefits of harnessing sustainable practices guided by insights obtained from climate change modelling and data analytics. They show evidence of positive outcomes in the environment, community engagement, and the efficiency of data-driven decision-making in navigating the effects of climate change.

The investigation of case studies in environmental sustainability brings the empirical concepts discussed in earlier sections to life. These examples from the real world show how the incorporation of data analytics and climate change modelling informs environmental sustainability, leading to positive effects for the communities and environment. By studying successful applications, we derive valuable insights into the revolutionary potential of leveraging data analytics for environmental sustainability.

3. Methodology

This research adopts a quantitative questionnaire approach as the method of data collection. Considering the newness of both data analytics and environmental sustainability, quantitative questionnaires are the best approach to address such new, unexplained and intricate phenomena (Creswell & Poth, 2016). This is in line with Walsham (2016), who said that quantitative questionnaires are suitable for gaining insights into the diversity and difficulties of complex social phenomena. Moreover, this study wants to aid the understanding of how data analytics can promote environmental sustainability. Thus, quantitative questionnaires are used to answer these research questions (Merriam, 2015). Quantitative questionnaires provide an in-depth exploration of respondents' views and can grasp the meaning of their opinions and experiences on data analytics capabilities to promote environmental sustainability. Hence, this questionnaire technique offers a way to comprehend the subtleties of respondents' perceptions and experiences and, therefore, proceed to a detailed and complete understanding of these phenomena under examination (Daymon & Holloway, 2021).

Questionnaires were administered to 50 persons, including environmental sustainability experts (15), data scientists (15), and academic experts in environmental sustainability (10) and data analytics (10). The respondents were asked questions to provide their opinions, experience and perceptions regarding how data analytics affects environmental sustainability, the barriers of data analytics and environmental sustainability, opportunities for data analytics and environmental sustainability, and how data analytics aligns with environmental conservation goals and values.

The study used descriptive and inferential statistics to analyse the data gathered. This involved the use of simple percentages and table presentations. Furthermore, regression analysis was used to generate inferences from the data using Statistical Packages for Social Sciences (SPSS).

4. Results and Discussion

The data utilised for this study were collected via Google Forms sent through emails by 50 persons, which included environmental sustainability experts, data scientists, and academic experts in both environmental sustainability and data analytics. The instrument used for gathering data was the questionnaire. SPSS software was applied to analyse the data using reliability statistics, regression and correlation. The data was composed through an adopted questionnaire, and the Likert level was adopted for variables of effects of data analytics, barriers of data analytics, opportunities of data analytics, and environmental sustainability.

The results are presented in the tables below:

S/N	How Does Data Analytics Affect Environmental Sustainability?	SA	A	UD	D	SD
1.	Reducing and regulating carbon footprints.	21(42%)	15(30%)	5(10%)	7(14%)	2(4%)
2.	Adherence to voluntary reporting principles and legal duties about carbon emissions	23(46%)	13(26%)	3(6%)	6(12%)	5(10%)
3.	Life Cycle Assessments	19(38%)	18(36%)	4(8%)	7(14%)	2(4%)
4.	Looking for environmental hazards and areas for advancement	17(34%)	22(44%)	3(6%)	5(10%)	3(6%)
5.	Predictive analysis forecasts	16(32%)	24 (48%)	1(2%)	6(12%)	3(6%)

Table 1: Responses on How Data Analytics Affects Environmental Sustainability
Source: Field Survey, 2024

Table 1 above shows responses regarding the effect of data analytics on environmental sustainability. Item 1 shows that 21 respondents, representing 2%, strongly agreed that data analytics helps reduce and regulate carbon footprints; 15 respondents, representing 30%, agreed; 5 respondents, representing 10%, were unsure; 7 respondents, representing 14%, disagreed, and 2 respondents, representing 4%, strongly disagreed.

Item 2 shows that 23 respondents, representing 4%, said that data analytics encourages adherence to voluntary reporting principles and legal duties about carbon emissions; 13 respondents, representing 26%, agreed; 3 respondents, representing 6%, were unsure; 6 respondents, representing 12%, disagreed, and 5 respondents, representing 10%, strongly disagreed.

Item 3 reveals that 19 respondents, representing 3%, strongly agreed that data analytics helps life cycle assessments, which in turn contributes to environmental sustainability; 18 respondents, representing 36%, agreed; 4 respondents, representing 8%, were unsure; 7 respondents, representing 14%, disagreed, and 2 respondents, representing 4%, strongly disagreed.

Item 4 shows that 17 respondents, representing 34%, strongly agreed that data analytics helps environmental sustainability by looking for environmental hazards and areas for advancement; 22 respondents, representing 44%, agreed; 3 respondents, representing 6%, were unsure; 5 respondents, representing 10%, disagreed, and 3 respondents, representing 6%, strongly disagreed.

Item 5 indicates that 16 respondents, representing 32%, said data analytics conducts predictive analysis forecasts, which help environmental sustainability; 24 respondents, representing 48%, agreed; 1 respondent, representing 2%, was unsure; 6 respondents, representing 12%, disagreed, and 3 respondents, representing 6%, strongly disagreed.

S/N	What Are the Barriers to Data Analytics and Environmental Sustainability?	SA	A	UD	D	SD
6.	Implementation Challenge	20(40%)	17(34%)	2(4%)	5(10%)	6(12%)
7.	Low speed of data production	16(32%)	23(46%)	3(6%)	4(8%)	4(8%)
8.	Low data quality and interoperability	24(48%)	15(30%)	5(10%)	3(6%)	6(12%)
9.	Large volume of data generated	22(44%)	18(36%)	3(6%)	3(6%)	4(%)
10.	Problems of scalability	18(36%)	25(50%)	1(2%)	1(2%)	5(10%)

Table 2: Responses on the Barriers to Data Analytics and Environmental Sustainability
Source: Field Survey, 2024

Table 2 shows the barriers to data analytics and environmental sustainability. The implementation challenge is the first item on the table, with 20 respondents, representing 40%, strongly agreeing; 17 respondents, representing 34%, agreeing; 2 respondents, representing 4%, being unsure; 4 respondents, representing 10%, disagreeing, and 6 respondents, representing 12%, strongly disagreeing.

Item 2 shows that 16 respondents, representing 32%, strongly agreed that the low speed of data production is a barrier to data analytics; 23 respondents, representing 46%, agreed; 3 respondents, representing 6%, were ambiguous and 4 respondents, representing 8%, disagreed and strongly disagreed, respectively.

Item 3 on the table shows that 24 respondents, representing 48%, strongly agreed that low data quality and interoperability are barriers to data analytics; 15 respondents, representing 30%, agreed; 5 respondents, representing 10%, were unsure; 3 respondents, representing 6%, disagreed, and 6 respondents, representing 12%, strongly disagreed.

Item 4 shows that 22 respondents, representing 44%, strongly agreed that the large volume of data generated is a hindrance to data analytics; 18 respondents, representing 36%, agreed; 3 respondents, representing 6%, were unsure and disagreed, and 4 respondents, representing 4%, strongly disagreed.

The final item on the table shows that 18 respondents, representing 36%, said that there are problems of scalability in data analytics; 25 respondents, representing 50%, agreed; a respondent, representing 2%, was unsure and disagreed respectively, and 5 respondents, representing 10%, strongly disagreed.

S/N	What Are the Opportunities for Data Analytics and Environmental Sustainability?	SA	A	UD	D	SD
11.	Potential to uncover insights and correlations within complex datasets	22(44%)	18(36%)	3(6%)	3(6%)	4(%)
12.	Gaining insights into the Earth's complex climate system	18(36%)	25(50%)	1(2%)	1(2%)	5(10%)
13.	Improvements in data analytics for environmental sustainability	20(40%)	17(34%)	2(4%)	5(10%)	6(12%)
14.	Successful implementation of sustainability practices	16(32%)	23(46%)	3(6%)	4(8%)	4(8%)
15.	Upcoming breakthroughs in technology	24(48%)	15(30%)	5(10%)	3(6%)	6(12%)

Table 3: Responses on the Opportunities for Data Analytics and Environmental Sustainability
Source: Field Survey, 2024

Table 3 shows the opportunities for data analytics and environmental sustainability. Item 1 shows that 22 respondents, representing 44%, strongly agreed that data analytics has the potential to uncover insights and correlations within complex datasets; 18 respondents, representing 36%, agreed; 3 respondents, representing 6%, were unsure and disagreed and 4 respondents, representing 4%, strongly disagreed.

Item 2 on the table shows that 18 respondents, representing 36%, said that data analytics helps us gain insights into the Earth's complex climate system; 25 respondents, representing 50%, agreed; a respondent, representing 2%, was unsure and disagreed, respectively; and 5 respondents, representing 10%, strongly disagreed.

Improvements in data analytics for environmental sustainability is the third item on the table, with 20 respondents, representing 40%, strongly agreeing; 17 respondents, representing 34%, agreeing; 2 respondents, representing 4%, being unsure; 4 respondents, representing 10%, disagreeing, and 6 respondents, representing 12%, strongly disagreeing.

Item 4 shows that 16 respondents, representing 32%, strongly agreed that successful implementation of sustainability practices is an opportunity for data analytics and environmental sustainability; 23 respondents, representing 46%, agreed; 3 respondents, representing 6%, were ambiguous; 4 respondents, representing 8%, disagreed and strongly disagreed, respectively.

Item 5 on the table shows that 24 respondents, representing 48%, strongly agreed that an opportunity for data analytics and environmental sustainability is the upcoming breakthroughs in technology; 15 respondents, representing 30%, agreed; 5 respondents, representing 10%, were unsure; 3 respondents, representing 6%, disagreed, and 6 respondents, representing 12%, strongly disagreed.

S/N	How Does Data Analytics Align with Environmental Conservation Goals and Values?	SA	A	UD	D	SD
16.	Improvement in environmental sustainability through climate change analysis	17(34%)	22(44%)	3(6%)	5(10%)	3(6%)
17.	Initiating timely interventions to curb environmental degradation	16(32%)	24(48%)	1(2%)	6(12%)	3(6%)
18.	Eco-friendly agricultural practices and renewable energy initiatives	21(42%)	15(30%)	5(10%)	7(14%)	2(4%)
19.	Enhanced water resource management	23(46%)	13(26%)	3(6%)	6(12%)	5(10%)
20.	Enhanced resilience to climatic challenges	19(38%)	18(36%)	4(8%)	7(14%)	2(4%)

Table 4: Responses on How Data Analytics Aligns with Environmental Conservation Goals and Values
Source: Field Survey, 2024

Table 4 above shows responses on how data analytics align with environmental conservation goals and values. Item 1 shows that 17 respondents, representing 34%, strongly agreed that data analytics lead to improvements in environmental sustainability through climate change analysis; 22 respondents, representing 44%, agreed; 3 respondents, representing 6%, were unsure; 5 respondents, representing 10%, disagreed, and 3 respondents, representing 6%, strongly disagreed.

Item 2 indicates that 16 respondents, representing 32%, said data analytics helps in initiating timely interventions to curb environmental degradation; 24 respondents, representing 48%, agreed; 1 respondent, representing 2%, was unsure; 6 respondents, representing 12%, disagreed and 3 respondents, representing 6%, strongly disagreed.

Item 3 shows that 21 respondents, representing 2%, strongly agreed that data analytics had helped create eco-friendly agricultural practices and renewable energy initiatives; 15 respondents, representing 30%, agreed; 5 respondents, representing 10%, were unsure; 7 respondents, representing 14%, disagreed, and 2 respondents, representing 4%, strongly disagreed.

Item 4 shows that 23 respondents, representing 4%, said that data analytics encourages enhanced water resource management; 13 respondents, representing 26%, agreed; 3 respondents, representing, 6% were unsure; 6 respondents, representing 12%, disagreed, and 5 respondents, representing 10%, strongly disagreed.

Item 5 reveals that 19 respondents, representing 3%, strongly agreed that data analytics helps enhance resilience to climatic challenges; 18 respondents, representing 36%, agreed; 4 respondents, representing 8%, were unsure; 7 respondents, representing 14%, disagreed, and 2 respondents, representing 4%, strongly disagreed.

4.1. Data Analysis

Data analysis was done using regression analysis, utilising the Statistical Packages for Social Sciences (SPSS) software. The results generated are presented and interpreted below.

Variables Entered/Removed			
Model	Variables Entered	Variables Removed	Method
1	OPPORTUNITIES_OF_DATA_ANALYTICS, BARRIERS_OF_DATA_ANALYTICS, EFFECT_OF_DATA_Analytics[b]	.	Enter
a. Dependent Variable: ENVIRONMENTAL_SUSTAINABILITY			
b. All requested variables were entered.			

Table 5
Source: Author's Computation from SPSS, 2024

Model Summary										
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.953[a]	.907	.630	4.004	.907	3.268	3	1	.381	1.831
a. Predictors: (Constant), OPPORTUNITIES_OF_DATA_ANALYTICS, BARRIERS_OF_DATA_ANALYTICS, EFFECT_OF_DATA_ANALYTICS										
b. Dependent Variable: ENVIRONMENTAL_SUSTAINABILITY										

Table 6
Source: Author's Computation from SPSS, 2024

The model summary indicates that the independent variables (OPPORTUNITIES_OF_DATA_ANALYTICS, BARRIERS_OF_DATA_ANALYTICS, and EFFECT_OF_DATA_ANALYTICS) collectively explain 63% of the variance in the dependent variable (ENVIRONMENTAL_SUSTAINABILITY).

A value around 2 suggests no autocorrelation (the value here is 3.268), which is good.

ANOVA[a]						
Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	157.168	3	52.389	3.268	.038[b]
	Residual	16.032	1	16.032		
	Total	173.200	4			
a. Dependent Variable: ENVIRONMENTAL_SUSTAINABILITY						
b. Predictors: (Constant), OPPORTUNITIES_OF_DATA_ANALYTICS, BARRIERS_OF_DATA_ANALYTICS, EFFECT_OF_DATA_ANALYTICS						

Table 7

Source: Author's Computation from SPSS, 2024

The ANOVA table assesses the overall statistical significance of the regression model. Here, the F-statistic is 3.268 with a corresponding p-value of 0.038, indicating that the overall model is statistically significant at conventional levels (e.g., alpha = 0.05).

Coefficients								
Model		Unstandardised Coefficients		Standardised Coefficients	T	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	73.872	59.988		1.231	.043		
	EFFECT_OF_DATA_ANALYTICS	-2.709	1.931	-1.179	-1.403	.039	.131	7.627
	BARRIERS_OF_DATA_ANALYTICS	-1.042	1.334	-.530	-.781	.047	.201	4.974
	OPPORTUNITIES_OF_DATA_ANALYTICS	1.093	.367	1.388	2.978	.021	.426	2.345
a. Dependent Variable: ENVIRONMENTAL_SUSTAINABILITY								

Table 8

Source: Author's Computation from SPSS, 2024

The coefficients table displays the estimated coefficients for each independent variable. The "B" column represents unstandardised coefficients, while the "Beta" column represents standardised coefficients. The p-values in the "Sig." column indicate the statistical significance of each coefficient. All of the independent variables seem to be statistically significant predictors of environmental sustainability, as all p-values are below conventional thresholds (e.g., 0.05).

Collinearity Diagnostics							
Model	Dimension	Eigenvalue	Condition Index	Variance Proportions			
				(Constant)	EFFECT	BARRIERS	OPPORTUNITIES
1	1	3.823	1.000	.00	.00	.00	.00
	2	.156	4.949	.00	.00	.01	.36
	3	.021	13.626	.00	.06	.07	.31
	4	.001	82.260	1.00	.94	.92	.33
a. Dependent Variable: ENVIRONMENTAL_SUSTAINABILITY							

Table 9

Source: Author's Computation from SPSS, 2024

This table provides information on multicollinearity among the independent variables. The "Condition Index" assesses the severity of multicollinearity, with values above 30 indicating potential issues. Here, the values are relatively low, suggesting that multicollinearity might not be a significant problem in this model.

Residuals Statistics					
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	6.62	22.38	16.60	6.268	5
Residual	-3.186	1.747	.000	2.002	5
Std. Predicted Value	-1.593	.923	.000	1.000	5
Std. Residual	-.796	.436	.000	.500	5
a. Dependent Variable: ENVIRONMENTAL_SUSTAINABILITY					

Table 10
Source: Author's Computation from SPSS, 2024

This table shows statistics related to the residuals (the differences between predicted and observed values). The mean of residuals is close to zero, which indicates that, on average, the model's predictions are accurate.
The range of residuals is from -3.186 to 1.747, suggesting that the residuals are relatively small in magnitude.

4.2. Discussion of Finding

The following was discovered from the results shown in this section.

Table 1 shows that data analytics has several effects on environmental sustainability. These effects include reducing and regulating carbon footprints, encouraging adherence to voluntary reporting principles and legal duties about carbon emissions, helping life cycle assessments, helping environmental sustainability by looking for environmental hazards and areas for advancement, and data analytics conducting predictive analysis forecasts, which help environmental sustainability.

Table 2 reveals the barriers to data analytics and environmental sustainability. These barriers are seen to include implementation challenges, low speed of data production, low data quality and interoperability, large volume of data generated, which is hard to manage, and problems of scalability.

Table 3 indicates the opportunities for data analytics and environmental sustainability. These opportunities are seen to include the potential to uncover insights and correlations within complex datasets. Data analytics helps us gain insights into the Earth's complex climate system. There will be improvements in data analytics for environmental sustainability, and successful implementation of sustainability practices is an opportunity for data analytics and environmental sustainability. There might be future breakthroughs in technology.

Table 4 shows how data analytics align with environmental conservation goals and values. These include:

- Data analytics lead to improvements in environmental sustainability through climate change analysis,
- Data analytics help in initiating timely interventions to curb environmental degradation,
- Data analytics have helped the creation of eco-friendly agricultural practices and renewable energy initiatives,
- Data analytics encourage enhanced water resource management, and Data analytics help enhance resilience to climatic challenges.

The regression model shows that there is an overall statistical significance between the independent variables (effect of data analytics, barriers of data analytics, and opportunities of data analytics) and the dependent variable (environmental sustainability with a corresponding p-value of 0.038). All of the independent variables seem to be statistically significant predictors of environmental sustainability, as all p-values were below conventional thresholds of 0.05.

The regression model based on the provided data predicted that all the independent variables appear to be statistically significant.

5. Conclusion

The findings underscore the critical role of data analytics in advancing environmental sustainability efforts. Addressing barriers, capitalising on opportunities, and leveraging data analytics align closely with environmental conservation goals, highlighting the potential for data-driven approaches to contribute significantly to global sustainability endeavors.

Data analytics plays an important role in reducing and regulating carbon footprints, thereby encouraging adherence to reporting principles and legal duties, conducting life cycle assessments, identifying environmental hazards, and conducting predictive analysis. Despite the potential benefits of data analytics, there are several barriers to leveraging data analytics for environmental sustainability, which include implementation challenges, slow data production, low data quality and interoperability, managing large volumes of data, and scalability issues. There are significant opportunities for leveraging data analytics to improve environmental sustainability, which includes uncovering insights within complex datasets, gaining a deeper understanding of the Earth's climate system, improving data analytics techniques, aligning sustainability practices with data analytics, and potential technological breakthroughs. These opportunities highlight the potential for further advancements in leveraging data analytics for environmental conservation.

Data analytics aligns closely with environmental conservation goals and values by facilitating improvements in sustainability through climate change analysis, initiating timely interventions to curb environmental degradation, promoting eco-friendly agricultural practices and renewable energy initiatives, enhancing water resource management, and building resilience to climatic challenges.

The regression analysis demonstrates a statistically significant relationship between the independent variables (effects, barriers, and opportunities of data analytics) and the dependent variable (environmental sustainability), with a p-value of 0.038. This suggests that all independent variables are significant predictors of environmental sustainability, indicating the importance of considering various aspects of data analytics in driving environmental sustainability initiatives.

6. Recommendations

Based on the findings of the study, here are some recommendations:

- Organisations and governments should invest in robust data analytics infrastructure to effectively collect, process, and analyse environmental data.
- Efforts should be made to address the barriers identified in the analysis, such as implementation challenges, data quality issues, and scalability concerns.
- Enhancing data literacy and building analytical skills among environmental professionals is crucial for maximising the potential of data analytics.
- Collaboration among governments, businesses, research institutions, and non-profit organisations is essential for harnessing the full potential of data analytics for environmental sustainability.
- Continued investment in research and innovation is necessary to advance data analytics techniques and tools for environmental sustainability.
- Governments should integrate data analytics into environmental policy and planning processes to inform strategic decision-making and resource allocation.
- It is essential to prioritise ethical and responsible data use practices to ensure the privacy, security, and integrity of environmental data.
- Establishing monitoring and evaluation mechanisms is critical for assessing the impact of data analytics initiatives on environmental sustainability outcomes.

7. References

i. Appelbaum, D. (2016). Securing Big Data provenance for auditors: The Big Data provenance black box as reliable evidence. *Journal of Emerging Technologies in Accounting, 13*(1), 17–36.
ii. Bughin, J., Chui, M., & Manyika, J. (2020). Clouds, big data, and smart assets: Ten tech-enabled business trends to watch. *McKinsey Quarterly, 56*(1), 75–86.
iii. Caniato, F., Caridi, M., Crippa, L., & Moretto, A. (2022). Environmental sustainability in fashion supply chains: An exploratory case-based research. *International Journal of Production Economics, 135*(2), 659–670.
iv. Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.
v. Daymon, C., & Holloway, I. (2021). *Qualitative Research Methods in Public Relations and Marketing Communications*. Taylor & Francis.
vi. Dumbill, E. & Mohin, S. (2023). "Opportunities at the Intersection of Health and Data". Big Data. Vol. 1, No. 3. Mary Ann Liebert, Inc.
vii. Dumbill, E. (2023). "Making Sense of Big Data". Big Data. Vol. 1, No. 1. Mary Ann Liebert, Inc.
viii. Elkington, J. (2024). Get Ready for the Breakthrough Decade… Volans. Retrieved from: http://volans.com/2024/05/get-ready-for-the-breakthrough-decade/
ix. Fransoo, J. C., Günther, H., & Jammernegg, W. (2024). Environmental sustainability in supply chains. *Flexible Services and Manufacturing Journal*, *26*,1–4.
x. Govindan, K. A., Susana G. C. H., & Cruz-Machado, V. (2024). Impact of supply chain management practices on sustainability. *Journal of Cleaner Production, 85*, 212–225.
xi. Govindan, K. S., Stefan, Z. Q., & Azevedo, S. G. (2016). Accelerating the transition towards sustainability dynamics into supply chain relationship management and governance structures. *Journal of Cleaner Production, 112*(3), 1813–1823.
xii. Graham, M. & Shelton, T. (2023). "Geography and the future of big data, big data and the future of geography". Dialogues in Human Geography. 255–261. Sage.
xiii. Gunasekaran, A. P., Thanos, D. R., Wamba, S. F., Childe, S. J., Hazen, B., & Akter, S. (2017). Big data and predictive analytics for supply chain and organisational performance. *Journal of Business Research, 70*, 308–317.
xiv. Hakovirta, M., & Denuwara, N. (2020). How COVID-19 redefines the concept of sustainability. *Sustainability, 12*(9),1–4.
xv. Hansen, J. (2020). "GLOBAL SURFACE TEMPERATURE CHANGE." Reviews of Geophysics, vol. 48, no. 4, December 2020. DOI: https://doi.org/10.1029/2020RG000345
xvi. Hazen, B. T., Skipper, J. B., Ezell, J. D., & Boone, C. A. (2016). Big data and predictive analytics for supply chain sustainability: A theory-driven research agenda. *Computers & Industrial Engineering, 101*, 592–598.
xvii. IPCC (2023) "Climate Change 2023: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovern mental Panel on Climate Change" [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.
xviii. IPCC. Climate Change 2024: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Core Writing Team, RK Pachauri and L.A. Meyer (eds.), IPCC, Geneva, Switzerland, 2024, pp. 151.

xix. Jeble, S., Dubey, R., Childe, S. J., Papadopoulos, T., Roubaud, D., & Prakash, A. (2018). Impact of big data and predictive analytics capability on supply chain sustainability. *The International Journal of Logistics Management, 29*(2), 513–538.

xx. Koo, D., Piratla, K., & Matthews, C. J. (2015). Towards sustainable water supply: schematic development of big data collection using internet of things (IoT). *Procedia Engineering, 118*, 489–497.

xxi. Koseleva, N., & Ropaite, G. (2017). Big data in building energy efficiency: understanding of big data and main challenges. *Procedia Engineering, 172*, 544–549.

xxii. Laney, D. (2021). "3D Data Management: Controlling Data Volume, Velocity, and Variety". Meta Group Inc. Retrieved from http://blogs.gartner.com/douglaney/files/2022/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-andVariety.pdf.

xxiii. Lokers, R., Knapen, R., Janssen, S., Randen, Y., & Jansen, J. (2016). Analysis of Big Data technologies for use in agro-environmental science. *Environmental Modelling & Software, 84*, 494–504.

xxiv. Manabe, Syukuro (2015). "The Effects of Doubling the CO2 Concentration on the Climate of a General Circulation Model." Journal of the Atmospheric Sciences, vol. 32, no. 1, 2015, pp. 3–15.

xxv. Manyika, J. et al. (2021). "Big Data: The next frontier for innovation, competition, and productivity". McKinsey Global Institute.

xxvi. Mayer-Schönberger, V., & Cukier, K. (2023). "Big Data: A Revolution that will transform how we Live, Work and Think". John Murray. London, UK.

xxvii. Merriam, S.B. (2015). *Qualitative Research: A Guide to Design and Implementation*. John Wiley & Sons.

xxviii. Mikalef P, Pappas I.O., Krogstie J. (2018). Big data analytics capabilities: A systematic literature review and research agenda. Information Systems and e-Business Management, 16(3); 547–578. 10.1007/s10257-017-0362-y

xxix. Millennium Ecosystem Assessment (2015). "Ecosystems and Human Well-Being: Synthesis". Island Press, Washington DC. Retrieved from: http://www.unep.org/maweb/documents/document.356.aspx.pdf.

xxx. Plass, Gilbert N (2016). "The Carbon Dioxide Theory of Climatic Change." Tellus, vol. 8, no. 2, May 2016, pp. 140–154.

xxxi. Provost, F. & Fawcett, T. (2023). "Data Science for Business: What you need to know about data mining and data-analytic thinking". O'Reilly Media.

xxxii. Rasp, Stephan (2018). "Deep learning to represent subgrid processes in climate models." EARTH, ATMOSPHERIC, AND PLANETARY SCIENCES, vol. 115, no. 39, 2018, pp. 9684–9689. DOI: https://doi.org/10.1073/pnas.1810286115

xxxiii. Simon, P. (2023). "Too Big to Ignore: The Business Case for Big Data". John Wiley & Sons, Inc. Hoboken, New Jersey.

xxxiv. Sowe, S. K. & Zettsu, K. (2024). "Curating Big Data Made Simple: Perspectives from Scientific Communities." Big Data. Vol. 2, No. 1. 23–33. Mary Ann Liebert, Inc.

xxxv. Stephens, G. (2020). "An update on Earth's energy balance in light of the latest global observations." Nature Geosci, vol. 5, 2022, pp. 691–696. https://doi.org/10.1038/ngeo1580

xxxvi. Stern, N (2017). The Economics of Climate Change. The Stern Review, Cambridge University Press, 2017.

xxxvii. Sumanth T. (2020). Climate Change Modeling and Analysis: Leveraging Big Data for Environmental Sustainability. *International Journal of Computer Engineering and Technology.*

xxxviii. Walsham, G. (2016). Doing interpretive research. *European Journal of Information Systems 15*(3), 320–330.

xxxix. Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research, 70*, 356–365.

xl. World Commission on Environment and Development (WCED). (2017). *Report of the World Commission on Environment and Development (2017): Our common future.*