**Unit 16: Computing Research Project – PART II**

Assignment Brief

|  |  |
| --- | --- |
| Programme Title | HND in Computing |
| Student Name | Imamuddeen Muhammedh |
| ANC ID | 1001321 |
| Unit Number and Title | Unit 16: Computing Research Project |
| Academic Year | SU24 / 2024 |
| Unit Tutor | Mr. Anuradha Boyagoda |
| Assignment Title | Computing Research Project Report |
| Issue Date |  |
| Submission Date |  |
| Submitted on |  |

Student Declaration

|  |
| --- |
| I certify that the assignment submission is entirely my own work and I fully understand the consequences of plagiarism. I understand that making a false declaration is a form of malpractice.  **Student signature: Date:** |

Final Grade

|  |  |  |
| --- | --- | --- |
| Grade: | Assessor Signature: | Date: |
| **Assessor Feedback:** | | |
| Grade: | Assessor Signature: | Date: |
| **Internal Verifier’s Comments:** | | |
| **Signature & Date:** | | |

Submission Guidelines

1. **Use the cover page provided in the previous page as your cover sheet** for the assignment you will be submitting and ensure that all the details are accurately filled. **All assignments should have the submitted with the above cover page which is properly filled.**
2. **Attach the assignment brief** as the first section of your assignment (Including the assignment criteria, summative feedback from and grading rubric).
3. Use an appropriate word processing software to develop your assignment.
4. Use A4 size pages for your assignments.
5. Use **font size 11** if you are using **Open Sans** or **Arial** font and font **size 12** if you are using **Times New Roman** font for the body of your assignment.
6. Use **font size 13 (bold)** if you are using **Open Sans** or **Arial** font and **font size 14 (bold)** if you are using **Times New Roman** font for **Main headings**. For **Sub-Headings** use **font size 12 (bold)** if you are using **Open Sans** or **Arial** font and font **size 12 (bold)** if you are using **Times New Roman**.
7. Use black text on a white background. **Avoid coloured backgrounds or text** in a colour other than black, unless you have special permission to use them.
8. Use **1.5 line spacing** and **2.53 cm (1”) wide margins** on all sides.
9. Justify your work (Ctrl+J).
10. **Insert a footer on each page** (except the title page) using **font size 10** if you are using **Times New Roman** and **font size 8** if you are using **Open Sans** or **Arial** font. Include the following details in the footer, Full Name, Student ID, Unit and the Page Number.
11. **Avoid** using **page borders** in your assignment body.

Important

1. All **submissions should accompany a Turnitin Report**. Any assignment which does not include a Turnitin Report will not be processed for grading. You are required to **attach the Turnitin report to the end of your assignment**. For this you can use file merging option.
2. Do note that **submission of the Turnitin report as an assignment will not be accepted** and will not be graded. In case a student has submitted a Turnitin Report only, without the assignment, that student will be awarded a **‘Fail’** grade.
3. It is **strictly prohibited to use texts in the form of images in your work** or **use text boxes** to present information other than to label a table or a figure.
4. Check the submission date of your assignment and submission instructions carefully. **Late submissions will not be accepted.**
5. Ensure that you give yourself enough time to complete the assignment by the due date and you are responsible for managing your own time effectively.
6. If you are unable to hand in your assignment on time and have valid reasons, you may apply (in writing) for extensions stating Extenuating Circumstance to AAO.
7. For all late submissions the final grade will be capped at a ‘**Pass’.**
8. **Failure to achieve the Pass criteria will result in a Referral grade.**
9. **Non-submission of work** without valid reasons will lead to an **automatic Failure** and you will be asked to complete an alternative assignment.
10. If you use ideas or work of other people in your assignment, reference them properly using Harvard Referencing system to avoid plagiarism. You have to provide both in-text citation and a reference list.
11. If you are proven to be guilty of plagiarism or any academic misconduct, your grade could be reduced to a Fail or at worst you could be administratively withdrawn from the course after a formal investigation.

Unit 16: Computing Research Project – PART II

Assignment Brief

|  |  |
| --- | --- |
| Unit Number and Title | Unit 16: Computing Research Project |
| Academic Year | SP24 -2024 |
| Unit Tutor | Mr. Anuradha Boyagoda |
| Assignment Title | Computing Research Project Report |
| Issue Date |  |
| Submission Date |  |
| Submission Format | |
| **Research Project Proposal Summary Form**  **Research Project Report**  **The work should be carried out individually.**  The recommended word limit is 4000 – 5000 words, although you will not be penalised for exceeding the total word limit.  The submission is in the form of a research report according to the template given.  This should be written in a concise, formal business style using specifications mentioned under Submission Guidelines.  You are required to make use of headings, paragraphs and subsections as appropriate and ensure that the report is justified aligned.  All work must be supported with research and referenced using the Harvard referencing system.  *You can also provide a bibliography using the Harvard referencing system.*  **Project logbook**  Please note that Project logbook, to be fill by the student and get the feedback / comment along with the signature from the lecturer.  A minimum of 5 weeks’ feedback will be provided.  Learner needs to be provided separate logbook for each week. | |
| Unit Learning Outcomes | |
| LO2 Conduct and analyse research relevant to a chosen computing research project  **LO3** Communicate the outcomes of a research project to identified stakeholders  **LO4** Reflect on the application of research methodologies and concepts | |
| Transferable skills and competencies developed | |
| The assignment offers students the chance to explore various aspects of big data from the perspective of computing professionals or data scientists. It also encourages investigations into the applications, benefits, limitations, and responsibilities associated with big data and provides solutions to the problems aims to solve. Further learners exhibit the abilities, transferrable skills in the contact of formal research. | |
| Assignment activity and guidance | |
| **Big Data**  Big data is a term that has become more and more common over the last decade. It was originally defined as data that is generated in incredibly large volumes, such as internet search queries, data from weather sensors or information posted on social media. Today big data has also come to represent large amounts of information generated from multiple sources that cannot be processed in a conventional way and that cannot be processed by humans without some form of computational intervention.  Big data can be stored in several ways: Structured, whereby the data is organised into some form of relational format, unstructured, where data is held as raw, unorganised data prior to turning into a structured form, or semi-structured where the data will have some key definitions or structural form but is still held in a format that does not conform to standard data storage models.  Many systems and organisations now generate massive quantities of big data on a daily basis, with some of this data being made publicly available to other systems for analysis and processing. The generation of such large amounts of data has necessitated the development of machine learning systems that can sift through the data to rapidly identify patterns, to answer questions or to solve problems. As these new systems continue to be developed and refined, a new discipline of data science analytics has evolved to help design, build and test these new machine learning and artificial intelligence systems.  Utilising Big Data requires a range of knowledge and skills across a broad spectrum of areas and consequently opens opportunities to organisations that were not previously accessible. The ability to store and process large quantities of data from multiple sources has meant that organisations and businesses are able to get a larger overall picture of the pattern of global trends in the data to allow them to make more accurate and up to date decisions. Such data can be used to identify potential business risks earlier and to make sure that costs are minimised without compromising on innovation.  However, the rapid application and use of Big Data has raised several concerns. The storage of such large amounts of data means that security concerns need to be addressed in case the data is compromised or altered in such a way to make the interpretation erroneous. In addition, the ethical issues of the storage of personal data from multiple sources have yet to be addressed, as well as any sustainability concerns in the energy requirements of large data warehouses and lakes.  The theme will enable students to explore some of the topics concerned with Big Data from the standpoint of a prospective computing professional or data scientist. It will provide the opportunity for students to investigate the applications, benefits and limitations of Big Data while exploring the responsibilities and solutions to the problems it is being used to solve.  **Choosing a research objective/question**  Students are to choose their own research topic for this unit. Strong research projects are those with clear, well focused and defined objectives. A central skill in selecting a research objective is the ability to select a suitable and focused research objective. One of the best ways to do this is to put it in the form of a question. Students should be encouraged by tutors to discuss a variety of topics related to the theme to generate ideas for a good research objective.  **The range of topics discussed on Big Data, could cover the following areas:**   * Storage models * Cyber security risks * Future developments and driving innovation. * Legal and ethical trade-offs   **You can select any company/industry/institution to conduct your project.**  Once you have identified a research topic and an organization, your report should include the following based on the proposal.   1. Define your research problem or question. This can be stated as a research question, objectives, or hypothesis. 2. Provide a literature review giving the background and conceptualisation of your proposed area of study. This would provide existing knowledge and benchmarks by which your data can be judged. 3. Consider and define your research methodology and research process. 4. Demonstrate understanding of the pitfalls and limitations of the methods chosen and ethical issues that might arise. 5. Conduct research using appropriate methods, considering costs, access, and ethical issues. 6. Analyse findings and data, discuss limitations, draw conclusions, communicate project outcomes, and evaluate performance. 7. Reflect on success, including recommendations for improvement and considering alternative methodologies. 8. Consider lessons learned and consider alternative research methodologies.   **Attach the Logbook relevant to Project report in the appendix.** | |

| **Recommended Resources**  *Please note that the resources listed are examples for you to use as a starting point in your research – the list is not definitive.* |
| --- |
| **Useful links**  Useful resources for underlying principles, examples of articles and webinars on the theme: |

| **Resource**  **Number** | **Type of**  **Resource** | **Resource Titles** | **Links** |
| --- | --- | --- | --- |
| 1 | Article | 6V’s of Big Data | [https://www.geeksforgeeks.org/5vs-of-big-data/](https://www.geeksforgeeks.org/5-vs-of-big-data/) |
| 2 | Article | Business Ethics and Big Data | <https://www.ibe.org.uk/resource/business-ethics-and-big-data.html> |
| 3 | Article | What is Big Data Security? Challenges & Solutions | [https://www.datamation.com/bigdata/big-data-security/](https://www.datamation.com/big-data/big-data-security/) |
| 4 | Article | What is Big Data? | [https://www.oracle.com/uk/bigdata/what-is-big-data/](https://www.oracle.com/uk/big-data/what-is-big-data/) |
| 5 | Magazine | Information Sciences | [https://www.sciencedirect.com/jou rnal/information-sciences](https://www.sciencedirect.com/journal/information-sciences) |
| 6 | Magazine | Big Data Research | [https://www.sciencedirect.com/jou rnal/big-data-research](https://www.sciencedirect.com/journal/big-data-research) |
| 7 | Report | Big Data & Investment Management:  The Potential to Quantify Traditionally Qualitative factors | <https://tinyurl.com/yff4uenz> |
| 8 | Webinar | Big Data Sources & Analysis Webinar | <https://tinyurl.com/2p85d7mb> |
| 9 | Video | Big Data In 5 Minutes | What Is Big Data?| Introduction To Big Data |Big Data Explained | [https://www.youtube.com/watch?v =bAyrObl7TYE](https://www.youtube.com/watch?v=bAyrObl7TYE) |
| 10 | Video | Challenges of Securing Big Data | [https://www.youtube.com/watch?v =3xIuIcPzMVs](https://www.youtube.com/watch?v=3xIuIcPzMVs) |
| 11 | Video | The Importance of Data Ethics | [https://www.youtube.com/watch?v =gLHMhCtxEYE](https://www.youtube.com/watch?v=gLHMhCtxEYE) |
| 12 | Book | A Bite-Sized Guide to Visualising Data | <https://tinyurl.com/38d6thsk> |
| 13 | Book | Business Intelligence Strategy and Big Data Analytics | [https://www.sciencedirect.com/bo ok/9780128091982/businessintelligence-strategy-and-big-data-analytics](https://www.sciencedirect.com/book/9780128091982/business-intelligence-strategy-and-big-data-analytics) |
| 14 | Book | Principles and Practice of Big | <https://www.sciencedirect.com/book/9780128156094/principles-and-practice-of-big-data> |
| 15 | Book | Systems Simulation and  Modelling for Cloud Computing and Big Data Applications | <https://tinyurl.com/2s3wkehn> |
| 16 | Journal | Big Data in Construction: Current Applications and Future Opportunities | [https://www.mdpi.com/25042289/6/1/18](https://www.mdpi.com/2504-2289/6/1/18) |
| 17 | Journal | Big Data with Cloud Computing: Discussions and Challenges | <https://www.sciopen.com/article/pdf/10.26599/BDMA.2021.9020016.pdf> |
| 18 | Journal | Mobile Big Data Solutions for a better Future | <https://tinyurl.com/hpk2zvvw> |
| 19 | Journal | The social implications, risks, challenges and opportunities of big data | <https://tinyurl.com/yw593svk> |
| 20 | Journal | Policy discussion – Challenges of big data and analytics driven demand-side management | <https://tinyurl.com/kyb3j6x7> |
| 21 | Journal | Explore Big Data Analytics Applications and Opportunities:  A Review | <https://tinyurl.com/597j8nd3> |
| 22 | Journal | What is Big Data? | [https://www.oracle.com/cl/a/ocom/ docs/what-is-big-data-ebook-4421383.pdf](https://www.oracle.com/cl/a/ocom/docs/what-is-big-data-ebook-4421383.pdf) |
| 23 | Journal | Towards felicitous decision making: An overview on challenges and trends of Big Data | [https://www.sciencedirect.com/science/article/abs/pii/S002002551630 4868](https://www.sciencedirect.com/science/article/abs/pii/S0020025516304868) |
| 24 | Journal | Critical analysis of Big Data challenges and analytical  methods | <https://www.sciencedirect.com/science/article/pii/S014829631630488X> |
| 25 | Journal | Big Data Security Issues and Challenges | <https://tinyurl.com/wabx7zya> |
| 26 | Journal | IoT Big Data Security and Privacy Versus Innovation | [https://ieeexplore.ieee.org/abstract /document/8643026](https://ieeexplore.ieee.org/abstract/document/8643026) |
| 27 | Journal | Big Data Security and Privacy Protection | [https://www.atlantis-press.com/proceedings/icmcs18/25904185](https://www.atlantis-press.com/proceedings/icmcs-18/25904185) |
| 28 | Journal | Big data analytics in Cloud computing: an overview | [https://journalofcloudcomputing.springeropen.com/articles/10.1186/ s13677-022-00301-w](https://journalofcloudcomputing.springeropen.com/articles/10.1186/s13677-022-00301-w) |

Learning Outcomes and Assessment Criteria

|  |  |  |  |
| --- | --- | --- | --- |
| Pass | | Merit | Distinction |
| **LO2** Conduct and analyse research relevant for a computing research project. | | | **LO1 & 2**  **D1** Critically evaluate research methodologies and processes in application to a computing research project to justify chosen research methods and analysis. |
| **P3** Conduct primary and secondary research using appropriate methods for a computing research project that consider costs, access and ethical issues.  **P4** Apply appropriate analytical tools, analyse research findings and data. | | **M2** Discuss merits, limitations and pitfalls of approaches to data collection and analysis. |
| **LO3** Communicate the outcomes of a research project to identified stakeholders | | | **D2** Evaluate outcomes and make valid, justified  recommendations. |
| **P5** Communicate research outcomes in an appropriate manner for the intended audience. | **M3** Analyse the extent to  which outcomes meet set  research objectives and  communicate judgements  effectively for the intended  audience | |
| **LO4** Reflect on the application of research methodologies and process | | | **D3** Demonstrate reflection and engagement in the resource process leading to recommended actions for future improvement. |
| **P6** Discuss the effectiveness of research methods applied, for  meeting objectives of the  computing research project.  **P7** Discuss alternative  research methodologies  and lessons learnt in view  of the outcomes. | | **M4** Analyse results in  recommended actions for  improvements and future  research considerations. |

Higher Nationals - Summative Assignment Feedback Form

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Student Name/ID |  | | | |
| Unit Title | Unit 16: Computing Research Project – Report | | | |
| Assignment No |  | Assessor |  | |
| Submission Date |  | Date Received 1st submission | |  |
| Re-submission Date |  | Date Received 2nd submission | |  |
| Assessor Feedback: | | | | |
| Grade: | Assessor Signature: | | | Date: |
| Resubmission Feedback: | | | | |
| Grade: | Assessor Signature: | | | Date: |
| Internal Verifier’s Comments: | | | | |
| Signature & Date: | | | | |

*\* Please note that grade decisions are provisional. They are only confirmed once internal and external moderation has taken place and grades decisions have been agreed at the assessment board.*

Higher Nationals – Grading Rubric

|  |  |  |
| --- | --- | --- |
| Grading Criteria | Achieved/Not Achieved | Comment |
| **P3** Conduct primary and secondary research using appropriate methods for a computing research project that consider costs, access and ethical issues. |  |  |
| **P4** Apply appropriate analytical tools, analyse research findings and data. |  |  |
| **M2** Discuss merits, limitations and pitfalls of approaches to data collection and analysis. |  |  |
| **D1** Critically evaluate research methodologies and processes in application to a computing research project to justify chosen research methods and analysis. |  |  |
| **P5** Communicate research outcomes in an appropriate manner for the intended audience. |  |  |
| **M3** Analyse the extent to which outcomes meet set research objectives and communicate judgements effectively for the intended audience |  |  |
| **D2** Evaluate outcomes and make valid, justified  recommendations. |  |  |
| **P6** Discuss the effectiveness of research methods applied, for meeting objectives of the computing research project. |  |  |
| **P7** Discuss alternative research methodologies and lessons learnt in view of the outcomes. |  |  |
| **M4** Analyse results in recommended actions for improvements and future research considerations. |  |  |
| **D3** Demonstrate reflection and engagement in the resource process leading to recommended actions for future improvement. |  |  |

A blue and yellow logo

Description automatically generated

**Using Big data to Understand Consumer Behaviour: Challenges associated with implementation of Data driven solution.**

By

**Imamudddeen Muhammedh**

1001321

Submitted in accordance with the requirements for the  
**COMPUTING RESEARCH PROJECT MODULE OF PEARSON BTEC HND IN COMPUTING PROGRAMME**  
at the

**ANC Education, Sri Lanka**

Name of research Tutor: Mr. Anuradha Boyagoda

2024 SUMMER

# DECLARATION

Name of Research Candidate : Imamuddeen Muhammedh

Pearson Registration Number :

Programme Name : PEARSON BTEC HND IN COMPUTING PROGRAMME

Research Title : Using Big data to Understand Consumer Behaviour: Challenges associated with implementation of Data driven solution.

Field of Study : COMPUTING

I do solemnly and sincerely declare that:

* 1. I’m the sole author of this study
  2. This work is original
  3. In case of any use if any information from other sources references of copyright with its ownership have been acknowledged in this document
  4. I do not have any actual knowledge, nor do I ought reasonably to know that the making of the work constitutes an infringement of any copyright work
  5. I know that plagiarism is a punishable offence because it constitutes theft, I understand the plagiarism and copying policy of the ANC Education & BTEC Centre Guide to Plagiarism, I know what the consequences will be if I plagiaries or copy another’s work in this research.

Candidate Signature: imamdeenm09@gmail.com Date:

Subscribed and solemnly declared before,

Supervisor’s Name: Mr. Anuraadha Boyagoda

Designation:

Supervisor’s Signature: Date:

# ACKNOWLEDGMENT

I would like to express my honest gratitude to Mr. Anuradha Boyagoda for his invaluable guidance and support throughout the research process. His expertise and encouragement have been instrumental in shaping this study on understanding consumer behaviour through big data analytics.

# ABSTRACT

This research proposal explores the utilization of big data analytics to gain insights into consumer behaviour, with a focus on informing advertising and marketing strategies, enhancing consumer engagement, and facilitating strategic planning. Leveraging the principles of positivism and employing a deductive approach, the study aims to analyse large volumes of facts sourced from social media platforms, e-commerce transactions, and internet analytics. By applying quantitative methods such as clustering, regression analysis, and machine learning algorithms, the research seeks to uncover patterns, trends, and correlations in consumer behaviour. The find out about adopts a longitudinal time horizon to observe adjustments and trends over multiple time periods, providing dynamic insights for strategic decision-making. Through convenience sampling, statistics will be collected from various sources, including social media users, e-commerce customers, and website visitors. Ethical considerations will be paramount, making sure participant anonymity and data security. The findings of this research are expected to contribute to each academic knowledge and practical applications, offering actionable insights for businesses to tailor marketing strategies, optimize user engagement initiatives, and make informed strategic decisions in an increasingly data-driven environment

***Key words:*** *Big Data consumer behaviour, advertising, marketing strategies, enhancing consumer engagement*

Table of Contents

[DECLARATION i](#_Toc183904752)

[ACKNOWLEDGMENT ii](#_Toc183904753)

[ABSTRACT iii](#_Toc183904754)

[LIST OF ABBREVIATIONS vii](#_Toc183904755)

[A LIST OF TABLES viii](#_Toc183904756)

[LIST OF FIGURES ix](#_Toc183904757)

[CHAPTER 1 – INTRODUCTION 10](#_Toc183904758)

[1.1. Introduction 10](#_Toc183904759)

[1.2. Purpose of research 10](#_Toc183904760)

[1.3. Significance of the Research 10](#_Toc183904761)

[1.4. Research objectives 11](#_Toc183904762)

[1.5. Research Sub objectives 11](#_Toc183904763)

[1.6. Research questions 11](#_Toc183904764)

[1.7. Hypothesis 12](#_Toc183904765)

[1.8. Arrangement of chapters 12](#_Toc183904766)

[CHAPTER 2 - LITERATURE REVIEW 14](#_Toc183904767)

[2.1. Literature Review 14](#_Toc183904769)

[2.2. Conceptual framework 17](#_Toc183904770)

[CHAPTER 3 – METHODOLOGY 19](#_Toc183904771)

[3.1. Research philosophy 19](#_Toc183904773)

[3.2. Research approach 21](#_Toc183904774)

[3.3. Research strategy 22](#_Toc183904775)

[3.4. Research Choice 25](#_Toc183904776)

[3.5. Time Horizon 26](#_Toc183904777)

[3.6. Data collection procedures 28](#_Toc183904778)

[3.6.1. Type of Data 28](#_Toc183904788)

[3.6.2. Data Collection Method 28](#_Toc183904789)

[3.6.3. Data Collection and Analyze Tools 28](#_Toc183904790)

[3.6.4. Questionnaire structure 29](#_Toc183904791)

[3.6.5. Data Storage 30](#_Toc183904792)

[3.7. Target population and sampling 31](#_Toc183904793)

[3.8. The selection of participants 31](#_Toc183904794)

[3.9. Reliability, Validity, and Generalizability 32](#_Toc183904795)

[3.10. Ethical issues of the research study 33](#_Toc183904796)

[CHAPTER 4 - PRESENTATION OF RESULTS 35](#_Toc183904797)

[4.1. Demographic Data Analysis 35](#_Toc183904799)

[4.2. Correlation Analysis 38](#_Toc183904800)

[4.2.1. RO2 / S01 38](#_Toc183904801)

[4.2.2. RO3 / S02 38](#_Toc183904802)

[4.2.3. RO4 / S03 38](#_Toc183904803)

[4.3. Regression Analysis 39](#_Toc183904804)

[4.3.1. RO1 / Main Objective 39](#_Toc183904805)

[4.3.2. RO2 / S01 40](#_Toc183904806)

[4.3.3. RO3 / S02 40](#_Toc183904807)

[4.3.4. RO4 / S03 41](#_Toc183904808)

[CHAPTER 5 - CONCLUSIONS AND RECOMMENDATIONS 42](#_Toc183904809)

[5.1. Conclusion 42](#_Toc183904811)

[5.1.1. RO1 43](#_Toc183904812)

[5.1.2. RO2 / S01 43](#_Toc183904813)

[5.1.3. RO3 / S02 44](#_Toc183904814)

[5.1.4. RO4 / S03 45](#_Toc183904815)

[5.2. Recommendations 46](#_Toc183904816)

[5.3. Limitations 47](#_Toc183904817)

[5.4. Future Improvements 49](#_Toc183904818)

[5.5. Personnel Reflection 50](#_Toc183904819)

[5.5.1. Benefits for the researcher 50](#_Toc183904820)

[5.5.2. Benefits for the selected Industry/organization 51](#_Toc183904821)

[Referencing 52](#_Toc183904822)

[Annexures 55](#_Toc183904823)

[Annexures A: Glossary of Terms 55](#_Toc183904824)

[Annexures B: Sample SPSS Charts/ Table 56](#_Toc183904825)

[Annexures C: Feedback Form / Survey question list / Interview question list 58](#_Toc183904826)

[Annexures D: Sample Feedback sheets 63](#_Toc183904827)

[Annexures E: Project logbook 66](#_Toc183904828)

[Turnitin report 74](#_Toc183904829)

# 

# LIST OF ABBREVIATIONS

• SPSS – Statistical Package for the Social Sciences

# A LIST OF TABLES

[*Table 1. Questionnaire structure* 26](#_Toc183725053)

[*Table 2 Reliability Statistics* 28](#_Toc183725054)

[*Table 3 Case Processing Summary* 28](#_Toc183725055)

[Table 4. Correlation Table 33](#_Toc183725056)

[*Table 5 frequency table* 46](#_Toc183725057)

# LIST OF FIGURES

[*Figure 1. Demographic Data Analysis Gender Distribution* 30](#_Toc183124035)

[*Figure 2. Demographic Data Analysis Age Distribution* 31](#_Toc183124036)

[Figure 3. Demographic Data Analysis data quality issue Distribution 32](#_Toc183124037)

# CHAPTER 1 – INTRODUCTION

## Introduction

The contemporary business landscape is an increasing number of reliant on data-driven insights to inform strategic decision-making. With the proliferation of digital technologies, considerable quantities of data are generated daily, supplying remarkable possibilities to recognize and have an impact on consumer behaviour. This lookup seeks to discover the utilization of big data analytics to gain insights into consumer behaviour, specially focusing on its implications for advertising and marketing strategies, person engagement, and strategic planning.

## Purpose of research

The primary purpose of this research is to look into how big data analytics can be leveraged to understand customer behaviour comprehensively. By inspecting large datasets sourced from diverse structures such as social media, e-commerce transactions, and internet analytics, the learn about aims to discover patterns, trends, and correlations in customer behaviour. The research seeks to elucidate how these insights can inform the development of effective marketing strategies, enhance user engagement initiatives, and facilitate strategic planning approaches for groups across a number of industries

## Significance of the Research

This research holds significant implications for both academia and industry. From an educational standpoint, it contributes to the growing body of knowledge on the intersection of big data analytics and consumer behaviour. By exploring novel methodologies and uncovering actionable insights, the research advances theoretical understanding in fields such as marketing, data science, and strategic management. Moreover, the findings of this study have sensible relevance for agencies seeking to thrive in today's competitive market landscape. By harnessing the energy of big data analytics, corporations can gain a deeper understanding of their target audience, tailor marketing techniques to unique customer segments, optimize person engagement initiatives, and make informed strategic selections to force commercial enterprise growth and competitive advantage. Overall, this research serves as a valuable resource for academics, practitioners, and policymakers alike, offering insights into harnessing the potential of large statistics to understand and have an effect on client behaviour correctly

## Research objectives

Main objective of this research is to study on Challenges associated with implementation of Data driven solution when using big data to Understand Consumer Behaviour.

This will provide Insights for marketing, User engagement, and Strategy Planning for the organization. Explore how using statistics can assist solve problems when creating targeted advertising campaigns that match what customers prefer and need.

## Research Sub objectives

SO1. Data Quality is challenges associated with implementation of Data driven solution

SO2. Technology is a main challenge when implementing data driven solution.

SO3. User engagement is also a challenge when implementing data driven solution.

## Research questions

RQ1. What are the challenges associated with implementation of Data driven solution?

RQ2. Is Data Quality a challenge associated with implementation of Data driven solution?

RQ3. Is Technology a main challenge when implementing data driven solution

RQ4. Is user management a challenge associated with implementing data driven solutions?

## Hypothesis

H1. There are challenges associated with implementation of Data driven solution.

H0. There are no challenges associated with implementation of Data driven solution.

H2. Data Quality is a challenge associated with implementation of Data driven solution.

H0. Data Quality is not a challenge associated with implementation of Data driven solution.

H3. Is Technology a main challenge when implementing data driven solution

H0. No Technology is not a main challenge when implementing data driven solution

H4. Is user management a challenge associated with implementing data driven solutions?

H0. No user management is not a challenge associated with implementing data driven solutions?

## Arrangement of chapters

This report contains five chapters. The introduction chapter will encompass the Purpose of Research and Significance of the Research. It outlines the objectives of the research and sub-research objectives. This chapter also covers the Research Questions and Hypothesis that lay the foundation for this research study.

This chapter, Literature Review, reviews the main literary sources which are the secondary data of the research and hence provides the academic setting for the study. In the latter part of this chapter, the Conceptual Framework shall be introduced to show how variables and concepts are interrelated and relevant to the research.

The Methodology chapter describes the research design of the study based upon the model by Saunders known as the Research Onion. The chapter addresses Research Philosophy, Research Approach, Research Strategy, Research Choice and Time Frame. Further on, it elaborates Data Collection Procedures concerning Type of Data, Data Collection Method, tools for collection and analysis and Questionnaire Structure. Additionally, it addresses procedures concerning Data Storage and Data Management. Later in the chapter, the Target Population is described, followed by the Sampling Strategy and the procedure concerning Participant Selection. The concluding part of the chapter addresses the issues of Reliability, Validity, and Generalizability of the findings and Ethical Issues pertinent to the study.

The Presentation of results chapter presents the findings of the research. The Demographic Data Analysis is presented first, followed by the Correlation Analysis and Regression Analysis. This chapter will go through the interpretation of various statistical outputs highlighting the relationships of the variables under study and will, therefore, evaluates the hypotheses that concern key metrics such as R-values, R-square values, and p-values. Graphical and tabular representations further support the discussion and represent a clear and actionable presentation of the results.

The Conclusions and Recommendations chapter synthesizes the insights from the findings and implications of such for the research objectives. The Conclusion section discusses the significance of the correlation and regression results, drawing of decisions on hypotheses, and mapping of findings to overall study goals. In recommendations, actionable strategies to organizations on how data quality and scalability, governance, and marketing practices can be improved are presented. It also covers the Limitations of the study and Future Improvements, such as wider sampling, advanced analytics tool usage, and the use of mixed-method approaches. The last section, Personal Reflection, indicates the development of the researcher's skills, understanding, and how the industry will possibly benefit from the study.

*<Intentionally Blank>*

# CHAPTER 2 - LITERATURE REVIEW



## Literature Review

***Data-Driven Insights into User Behaviour***

Jones and Smith delve into the realm of social media analytics, emphasizing its attainable for uncovering treasured insights into consumer behaviour. Through an examination of a number case studies and methodologies, the authors highlight the significance of data-driven methods in perception person interactions on platforms like Facebook, Twitter, and LinkedIn. By leveraging superior analytical techniques, Jones and Smith demonstrate how organizations can extract actionable insights from social media data to inform strategic decision-making and beautify user engagement. Their lookup presents a comprehensive framework for researchers and practitioners in search of to harness the energy of social media analytics for optimizing digital advertising and marketing strategies and enhancing normal enterprise performance. (Jones & Smith, 2021)

Data-driven insights into consumer behaviour enable groups to tailor their marketing strategies and enhance patron satisfaction. Chen, Chiang, and Storey (2012) talk about how advanced analytics can find patterns in customer behaviour, providing actionable insights for customized marketing. The authors show that by leveraging large data, corporations can predict future patron actions, decorate patron experience, and make bigger retention prices (Chen, Chiang, & Storey, 2012)

***Exploring Machine Learning Techniques in Social Media Analysis***

Patel and Brown check out the application of desktop getting to know algorithms in social media analytics, aiming to uncover patterns and trends in user behaviour. Their study examines the effectiveness of quite a number computing device learning techniques, such as herbal language processing and sentiment analysis, in extracting precious insights from social media data. By employing superior algorithms, Patel and Brown demonstrate how agencies can reap a deeper grasp of person preferences, sentiments, and engagement patterns on platforms like Twitter and Instagram. Their research presents realistic implications for enhancing advertising strategies, consumer relationship management, and brand reputation administration in the digital age. (Patel & Brown, 2020)

Machine gaining knowledge of strategies play a pivotal role in social media analysis, enabling the extraction of valuable insights from tremendous amounts of unstructured data. According to Gandomi and Haider (2015), algorithms such as sentiment analysis, subject modeling, and community analysis can perceive traits and sentiments in social media discussions. These techniques assist businesses understand public perception, measure brand reputation, and gauge client engagement in real-time (Gandomi & Haider, 2015)

***Cross-Domain Insights from Social Media Analytics***

In their lookup paper, Lee and Garcia explore the transferability of insights derived from social media analytics across one-of-a-kind domains. Focusing on platforms like TikTok and Snapchat, the authors analyse user-generated content material to uncover frequent patterns and behaviours throughout numerous demographics and industries. By figuring out cross-domain insights, Lee and Garcia supply treasured implications for marketers, content creators, and policymakers searching for to leverage social media facts for knowledgeable decision-making. Their find out about underscores the importance of interdisciplinary collaboration and information change in harnessing the full workable of social media analytics to tackle current challenges and possibilities in the digital landscape. (Lee & Garcia, 2019)

Social media analytics presents cross-domain insights that are beneficial throughout quite a number sectors, from marketing to public health. Fan and Gordon (2014) explore how social media statistics can be utilized beyond ordinary marketing functions to understand broader societal trends and behaviours. Their lookup suggests that integrating social media statistics with other information sources, such as economic indicators, can supply a complete view of client behaviour and market dynamics (Fan & Gordon,, 2014)

***Ethical Considerations in Social Media Analytics***

Kumar and Wilson take a look at the moral implications of social media analytics, mainly concerning user privacy, consent, and data usage. Their lookup explores the moral dilemmas related with collecting, analysing, and disseminating social media records for industrial and lookup purposes. By severely evaluating modern-day practices and regulatory frameworks, Kumar and Wilson highlight the want for larger transparency, accountability, and consumer empowerment in the area of social media analytics. Their learn about contributes to ongoing discussions on ethical tips and nice practices for conducting responsible and moral lookup in the era of large statistics and digital surveillance. (Kumar & Wilson, 2020)

***Challenges Associated with Implementation of Data-Driven Solutions***

The implementation of data-driven solutions faces numerous challenges, specifically related to the complexity of integrating various facts sources and making sure scalability. According to Davenport and Dyché (2013), one considerable project is the technical infrastructure required to aid massive records analytics. The authors highlight that groups often war with old-fashioned systems that cannot take care of the volume, variety, and pace of huge data. Moreover, managerial resistance and lack of professional personnel further complicate the adoption of these solutions (Davenport & Dyché, 2013).

***Data Quality is a Challenge***

Data fantastic is a vital project in massive data analytics, impacting the accuracy and reliability of insights derived from customer data. Redman (2018) emphasizes that bad statistics best can lead to misleading conclusions and ineffective decision-making. The creator outlines frequent facts great issues such as inaccuracies, inconsistencies, and incompleteness, which stem from a number sources consisting of records entry blunders and lack of standardized records series strategies (Redman, 2018). Ensuring excessive facts satisfactory requires rigorous facts governance and continuous monitoring processes. (Redman, 2018)

***Ethical Considerations in Social Media Analytics***

The moral implications of social media analytics are significant, mainly related to consumer privacy and records security. boyd and Crawford (2012) argue that whilst social media analytics presents treasured insights, it also raises ethical issues related to consent, anonymity, and data misuse. The authors stress the significance of setting up moral tips and sturdy data protection measures to make sure that person data is treated responsibly and ethically (boyd &Crawford, 2012)

## Conceptual framework

**Independent Variable (IV)**

**Dependent Variable (DV)**

Challenge when implementing data driven solution

SO1. Data Quality

SO2. Technology

SO3. User engagement

Effectiveness of data driven solutions

**Explanation**

**Independent Variables (IV)**

These are factors hypothesized to impact the effectiveness of data-driven solutions. The challenges you've identified are segmented into three sub-objectives, or SOs:

* SO1: Data Quality refers to the dependability, accuracy, and completeness of data; poor data can make much difference in decision-making processes pertaining to data-driven solutions.
* SO2: Technology embraces all infrastructures, tools, and software needed for data analysis; these might facilitate or inhibit the effective delivery of solutions.
* SO3: User Engagement involves the engagement and adaptability of the end-users in using data-driven insights and technologies effectively.

**Dependent Variable (DV)**

The effectiveness of data-driven solutions recognizes how well such solutions meet their intended outcomes, such as making better decisions, improving marketing strategies, or increasing user engagement.

**Relationships**

The conceptual framework assumes that identified challenges, or IVs, have a direct or indirect influence on the effectiveness of data-driven solutions.

* Poor quality of data insights could be produced that would reduce effectiveness.
* Poor technology can create a limit in processing and/or analytics capability and impact results.
* Poor user engagement could result in underutilization or misapplication of insights, which compromises the solution's success.

The framework approaches the interplay between these challenges and their cumulative influence on achieving effective data-driven strategies. It provides a structured approach to investigate and test the hypotheses outlined in your research.

*<Intentionally Blank>*

# CHAPTER 3 – METHODOLOGY



## Research philosophy

**Positivism**

The philosophy of positivism is composed of the aspects of objective reality, empirical evidence, observation, measurement, and scientific analysis. Researchers in this philosophy take for granted that there is an external, observable reality that can be measured objectively, neuter, and independent of the researcher's bias. In many instances, it rests on quantitative data to delineate general laws and causal relationship.

Positivism emphasizes objectivity, reproducibility, and scientific rigor. It supports structured methodologies and statistical analyses; thus, it offers a good method to analyze large datasets and render insightful findings that can be acted upon. The nature of this approach is quantitative in nature; hence the findings are generalizable, and the data obtained is reliable.

The disadvantage of Positivism is It critics say positivism may fail to capture subjective experiences of individuals along with nuanced contexts that might be captured by qualitative data, thus offering a somewhat narrow version of complex social phenomena.

**Realism**

Realism embraces both an objective reality independent of human perception and, at the same time, acknowledges that our knowledge arises from social, cultural, and individual experiences. Realism tries to identify the mechanisms underlying the observable phenomena; in one word, it has often combined quantitative and qualitative approaches.

The Advantages of realism is It allows for detailed insight that combines the scientific rigor of empirical research with an awareness of contextual influences. Realist researchers aim to understand mechanisms and also relationships that might be complicated.

Disadvantages of Realism is that realism has a tendency to render difficulties while balancing objectiveness of findings against subjective perceptions. Mixed-method approaches could be inordinately resource and time-consuming, especially when judged against pure quantitative or qualitative studies.

**Pragmatism**

Pragmatism is essentially all about finding practical solutions to problems through a flexible approach. The feature of what works best for pragmatist researchers in the context is key; it can therefore include the quantitative and qualitative integration of methodologies to achieve understanding regarding the research problem. Pragmatism is flexible and allows the integration of as many methods as possible to solve complex issues. It problematizes, orientates on outcomes, and solves problems. There are criticisms that pragmatism sometimes leads to inconsistent methodologies or a lack of philosophical rigor.

**Interpretivism**

Interpretivism explores the subjective meaning and social context of human behavior. It does this by concentrating on the ways in which individuals create an understanding of their experiences and is usually linked with qualitative methods.

Advantage of Interpretivism is that It really gives an insight into social interactions, beliefs, and practices. Researchers receive rich data contextualized, usually missed by purely quantitative approaches. The disadvantage is This philosophy lacks generalizability, and one of the major criticisms is related to the potential for bias of the researcher and maintaining objectivity.

The researcher uses **Positivism** as the research philosophy for this study because it is a philosophy that embraces the scientific and empirical method of knowledge acquisition. This research has adopted quantitative methods to objectively study challenges and factors that influence data-driven solutions. As a part of this research, researcher use research questions in order to derive empirical knowledge.

RQ1: What are the challenges associated with implementation of data-driven solutions?

To answer the question, the researcher will look to find casual relationships through systematic survey in quantitative methods. Here is an example of hypothesis the researcher is trying to prove

H1: There are challenges associated with implementation of data-driven solutions.

H0: There are no challenges associated with implementation of data-driven solutions.

**Reason Why researcher didn’t choose realism, Pragmatism, or Interpretivism?**

**Realism**

Realism was not used because this investigation is scientific, while realism often doubts the sufficiency of a scientific approach and opens the floor for alternative approaches. Pragmatism and interpretivism could also not be applied because such approaches usually take more time due to their flexibility and focus on social context understanding, which is elaborate, and the current study is limited by the constraint of time.

## Research approach

**Inductive Approach**

In the inductive approach, one collates particular data in order to find out patterns that will further develop a broader theory. The "bottom-up" method is an exploratory type in which observation may be made first and, thereafter, generalized. So, analyses of social media at an individual level can provide insight into identifying consumer behaviour and trending. This nature of approach certainly will lead to new theories and insights, but may often lack detailed processes of testing some pre-defined hypotheses and often quantitative data is usually at the background.

**Deductive Approach**

The deductive approach operates on the basis of established theories or hypotheses that are tested using structured methods to collect and analyse data. This top-down approach focuses on testing set hypotheses, and quantitative methods are often applied to ensure the research is structured and can be replicated. For example, this study tests various hypotheses about challenges when implementing data-driven solutions, such as data quality and user management, through collecting relevant data and its analysis. The deductive approach enables valid and actionable insight to be provided through systematic evidence-based analyses.

**This research uses a deductive approach** in testing specific hypotheses related to challenges and factors influencing data-driven solutions. Using structured quantitative methods allows this research to objectively establish the validity or rejection of claims that are made on data-driven challenges, such as

RQ1: What are the challenges of implementing data-driven solutions?

H1: There are challenges to the implementation of data-driven solutions.

This structured approach thus allows for focused data analysis, giving clear insights into marketing strategies, user engagements, and strategic planning.

**Why I didn’t choose Inductive Approach**

In this research study, the inductive approach is not apt because this approach seeks the identification of new patterns and testing of theory. However, it does not aim at the validation of hypothesized theory. Since particular pre-defined hypotheses are to be tested out, the deductive approach will better lead to the effective accomplishment of research objectives.

## Research strategy

**Experiment**

Experiments are a type of research methodology that investigates cause-and-effect relationships under controlled conditions. In the context of this research, such a design would establish whether big data-informed personalized marketing strategies result in greater user engagement and, by extension, more effective marketing performance. However, it is pretty difficult to control for real external market variables, as consumer behaviour is quite complex and dynamic. Accordingly, while experiments might have provided some useful insights, given practical constraints and low feasibility of experiments to scan broad consumer behaviour, they are not chosen as the main strategy.

**Survey**

The survey method is the principal focus of this research, as it can collect data from a large and diversified group of consumers for the purpose of understanding behaviours, preferences, and attitudes. Surveys are really useful in analyzing responses of consumers to marketing strategies informed by big data, exploring how personalization driven by data may affect engagement, and identifying any concerns related to the use of data and privacy. In this type of survey, many questions about consumer behaviours and perceptions can be asked.

Question: How likely are you to engage with personalized marketing campaigns using your data to tailor content and recommendations?

The use of a survey method quantifies consumer attitudes toward the use of data in driving marketing strategies; hence, actionable insights shall be derived to inform more targeted campaigns of relevance.

**Case Study**

The broad findings from the survey are supplemented through a case study that provides an in-depth investigation of how big data are used in understanding consumer behaviour within a particular business or marketing context. This can be a company applying big data analytics to personalize marketing, increase user engagement, and make strategic decisions. The case study can provide in-depth knowledge through analysis of data usage, its approach of consumer segmentation, and marketing results to contextualize and enhance general survey data. This case study approach helps the research objective in understanding how consumer-oriented data-driven practices can enhance consumer engagement and marketing effectiveness.

**Grounded Theory**

Grounded theory is a research method that aims to identify the generation of new theories by analyzing data and is not initiated from a well-defined hypothesis. Although that is not the explicit focus of this study, if new trends or insights emerge during both the survey and case study phases, then a grounded theory approach could be utilized to construct new frameworks of understanding regarding how big data influences consumer behaviour. The primary reliance, though, remains upon utilizing existing theories and practices to develop actionable insights rather than constructing new theoretical constructs.

**Ethnography and action research**

Ethnography deals with the long-term immersion in culture, whereas action research focuses on collaboration for solving problems-both of which are irrelevant to the present study. It will not involve long-term observation of culture or any kind of collaborative intervention with consumer groups. These methods have therefore been excluded, since they do not align with the research focus on the analysis of big data trends and consumer behaviour for marketing and engagement strategies.

This research will be done through the incorporation of a cross-sectional survey combined with a case study for the analysis and understanding of consumer behaviour patterns influenced by big data. The survey would provide broad quantitative insights into the attitudes, preferences, and engagement of consumers with data-driven marketing efforts, while the case study provides an in-depth look at the efficient elaboration of big data strategies in marketing and user engagement. Where unexpected themes emerge, then the depth of understanding could then be enhanced, probably when the consideration of the grounded theory is made. Overall, the methods would support the objectives of the study, which are to enhance marketing strategies and user engagements through data-driven insights in strategic planning.

## Research Choice

**Mono Method**

The present study is a mono-method, relying solely on one method of collecting and analyzing data. Since the research questions and objectives are very clear, this will be appropriate for the collection of comprehensive focused data in order to understand consumer behaviour regarding big data.

The quantitative wide survey of consumers is the main method used in capturing their behaviour, attitudes, and preferences in relation to big data-driven marketing strategies. Results from this survey will explore how data-driven approaches impact user engagement, satisfaction, and perceptions of privacy. In a mono-method approach, extensive data can be gathered which is quantifiable and generalizable across a broad sample. Although providing very significant breadth and clarity, it cannot capture the rich contextual nuances found in a wholly qualitative method. This, however, is a quite acceptable limitation within the context of the present study, because in this study, the objective is to explore actionable insights about trends, preferences, and challenges related to consumer behaviour through big data.

Question: To what extent do you believe that personalized marketing campaigns driven by big data improve your overall user experience with a product or service?

**Mixed Method**

This design involves planning for the incorporation of quantitative and qualitative methods of data collection and analysis in order to understand the problem at hand comprehensively. For example, one could combine survey methods with qualitative interviews to delve into consumer sentiment at a broader scale, while being able to dig into individual motivations and experiences as well. However, the nature of design and execution, along with synthesizing studies of mixed method, is complicated and resource-consuming. Hence, this approach would not align well with the current thrust of research.

**Multi-method**

The multi-method strategy uses different methods within one qualitative or quantitative framework but keeps them separate in time during the analysis. For example, running separate surveys for different consumer segments, like younger consumers versus older ones, can create a segmented view of consumer attitudes towards big data-driven personalization. This gives flexibility in methodology and provides chances to capture diverse insights, although there may be limitations to integration and comprehensive conclusions when compared to a fully integrated analysis.

The mono-method can fit the purpose of the study better, which intends to apply big data to consumer behaviour pattern analysis, marketing strategy enhancement, user engagement, and strategic planning. This involves the critical analysis of consumer attitude and behaviours through a coherent focus on a cross-sectional survey. This is informed by the fact that the choice of research philosophy, approach, and methodology would be guided to have an exact match with the intents of such a study in extracting actionable insights into the marketing effectiveness and consumer engagement strategies through data-driven insights.

## Time Horizon

**Cross-Sectional**

The cross-sectional study involves the collection of data at one point in time and represents a snapshot of the phenomenon or behavior being studied. This is faster, cheaper, and also useful in identifying any kind of trend or tendency in consumer behavior. It cannot be used to determine how changes take place over time nor does it establish cause-and-effect relationships among variables. A cross-sectional approach is ideal for the present research in this respect, as it seeks to comprehend consumer behavior patterns at an instance in time in an efficient manner by collecting and analyzing data. By taking a snapshot of how consumers presently respond to data-driven marketing strategies, the research is assured of immediate insights into the effectiveness of marketing and user engagement.

For example, research would look into the preferences and actions of consumers as they interact with data-driven marketing strategies; this therefore means that organizations can contextualize their strategies to current consumer behavior. This approach is mainly applicable for understanding short-run trends and thus can inform immediate decisions in marketing campaigns or product positioning.

**Longitudinal**

A longitudinal approach would focus on measuring changes in consumer behavior over an extended period, allowing for a deeper analysis of how consumer preferences and actions evolve in response to marketing efforts over time. This approach is valuable for studying long-term effects, such as the sustained impact of data-driven marketing strategies or shifts in consumer behavior due to external factors like economic cycles or technological changes.

However, this would involve taking longitudinal data over several time periods, which is tedious and resource-intensive. Although it can provide richer insights into trends and causality in longer-term changes, this study is more interested in analyzing the current patterns in the behavior of consumers in their response to data-driven marketing. The tracking of such changes over time would not be able to address the research objectives, since the latter aim to identify trends and insights from current data rather than studying how these trends come into being.

Question - How does consumer behavior respond to data-driven marketing strategies at a specific point in time?

**Why researcher chose Cross-Sectional**

The cross-sectional time frame has been chosen in this research for the reason that there was a need to gain an effective and timely understanding of consumer behavior in the context of data-driven marketing strategies. The focus of this research is to identify trends and get an idea about consumer preferences of a particular point in time; hence, the objectives of this research are best suited for a cross-sectional approach. A snapshot view also allows the study to present quick analysis of data to provide actionable insights for short-term marketing strategies.

**Why Longitudinal Was Not Chosen**

This could show the changes in consumers' behavioral responses over some time; this approach is less suitable within the scope of the present research. What it actually looks to capture are immediate consumer reactions to data-driven marketing, rather than its changes over an extended time period. The application of a longitudinal approach would only be burdensome, adding considerable time, and not add focused insight from the specific demands made by the research objectives. Therefore, the cross-sectional approach was more applicable to the research questions to be answered effectively within the time scale.

## Data collection procedures



### Type of Data

Primary data – data was collected through a Google form survey sent to 60 participants. The purpose was to understand challenges in implementing data driven solutions, focusing on factors like data quality, technology and user management

Secondary data – data was gathered from existing sources such as articles and reports to provide context and support the analysis of primary data, particularly in areas like big data, consumer behavior and marketing strategies

### Data Collection Method

Survey data collection Method

### Data Collection and Analyze Tools

Questionnaire is the main data collection tool and Online questionnaire was dispatched using google form

SPSS and MS excel, google sheets used as data analysing tools

### Questionnaire structure

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variable** | **Indicators** | **Measurement** | **Mean** | **STD. Deviation** | **Median** |
| IV 1/ SO1 | How confident are you in your organization’s ability to address data quality issues in data-driven solutions? | Questioner | 2.9557 | .70054 | 3.0000 |
| Data quality issues when implementing data-driven solutions |
| Security measures in place regularly basis will impact to improve data quality |
| IV 2/ SO2 | How often do technological limitations hinder the implementation of data-driven solutions in your organization |  | 2.9125 | .76512 | 2.6700 |
| Technological limitations hindering the implementation of data-driven solutions in your organization? |
| Difficulties in integrating new data-driven technologies with your existing systems? |
| IV 3/ SO3 | On a scale of 1 to 5, Users engagement with the data-driven tools |  | 3.0769 | .94255 | 3.0000 |
| Users preferences for using new data tools? |
| Communication strategies in promoting user engagement with data-driven solutions? |
| DV | On a scale from 1 to 5, data quality issues are encountered frequently when implementing data-driven solutions. |  | 2.8469 | 1.18866 | 3.0000 |
| How satisfied are you with the technology stack used for data-driven solutions ? |
| How engaged are end-users in the implementation process of data-driven solutions ? |

*Table 1. Questionnaire structure*

Please refer Annexures B: Sample SPSS Charts/ Table chapter for more details.

### Data Storage

The survey data are collected and stored digitally. Responses to the Google Forms survey are collated automatically into Google Sheets, which happens in real-time.

**Data Storage Tool**

* Google Sheets is used for securely storing the responses of the survey.
* Cloud storage is done through Google Drive to ensure that all data is backed up and accessible only by the researcher.

**Data Security and Access**

Access to the data stored was restricted to the researcher only, data protected by the researcher's secure Google account. Regular backups were performed, with periodic downloading of copies of data done to ensure integrity.

## Target population and sampling

**Target Population**

It targeted a population that has experience in or exposure to data-driven solutions drawn from diverse professional or academic backgrounds, including technology, marketing, business analytics, and healthcare.

**Sampling Method**

A convenience **sampling method** was used, in which participants were selected based on ease of accessibility and mutual agreement, largely through personal contact made by the researcher.

**Sample Size**

The sample size was **60**, which is adequate to map the trends and test the hypotheses that could be formulated within the scope of this research. However, because the samples were not drawn using a random method, generalization might not be assured.

## The selection of participants

* Familiarity with Data-Driven Solution: The participants were selected from the researcher's personal network, with preferences toward those who had any knowledge of experience in data-driven solutions in either professional and academic circles.
* Willingness to Participate: Participants were selected based upon their willingness to provide genuine and reflective feedback about issues revolving around the integration of data-driven solutions.
* Age: All participants were above the age of 18, ensuring that they were legally able to provide informed consent for their participation in the study.

**Rationale for Selection**

* Convenience: Participants were selected out of convenience with the view of creating efficiency in both time and resources, as the researcher had easy access to them.
* Relevant Insights: Although the sampling was non-random, participants were selected because their insights into data-driven solutions were relevant to the objectives of the study.

**Limitations of Selection**

* The sample lacked diversity in professional backgrounds and geographic locations because the majority of participants came from the researcher's personal network.
* The convenience sampling nature allows the findings to not be fully representative of the broader population; hence, generalization to other industries or organizations is limited.

## Reliability, Validity, and Generalizability

**Reliability**

Here researcher consider Cronbach’s alpha value to define the reliability. Researcher accepts if alpha value is more than 0.6 or 60% or above is reliable research

|  |  |
| --- | --- |
| **Reliability Statistics** | |
| Cronbach's Alpha | N of Items |
| .253 | 4 |

*Table 2 Reliability Statistics*

According to the reliability statistics table, Cronbach’s alpha value is .253(25.3%) which is a lower than 0.6 or 60% therefore the research is reliable up to certain extent.

**Validity**

|  |  |  |  |
| --- | --- | --- | --- |
| **Case Processing Summary** | | | |
|  | | N | % |
| Cases | Valid | 61 | 100.0 |
| Excludeda | 0 | .0 |
| Total | 61 | 100.0 |
| a. Listwise deletion based on all variables in the procedure. | | | |

*Table 3. Case Processing Summary*

According to the case processing summary table, there are total of 61 valid cases and no exclude cases. As a percentage, valid cases are 100%. Therefore, the research data is valid in the context of this research

**Generalizability**

This research is generalised up to the certain extent within the research context. The reason is the study conducted in marketing and advertising industry and the data collection is the limit to the data access restrictions and location and demographic factors

## Ethical issues of the research study

**Data Protection and Privacy**

No personal data is collected, but ensuring the security of any collected information is still important

**Anonymity and Confidentiality**

Making sure that data collected does not identify respondents and maintaining their anonymity

**Transparency in Data Use**

The participants need to know where responses they provided in the research will be used

**Adherence to Regulations**

It is always necessary that the relevant ethical standards and regulations should be adhered to regarding the research practices.

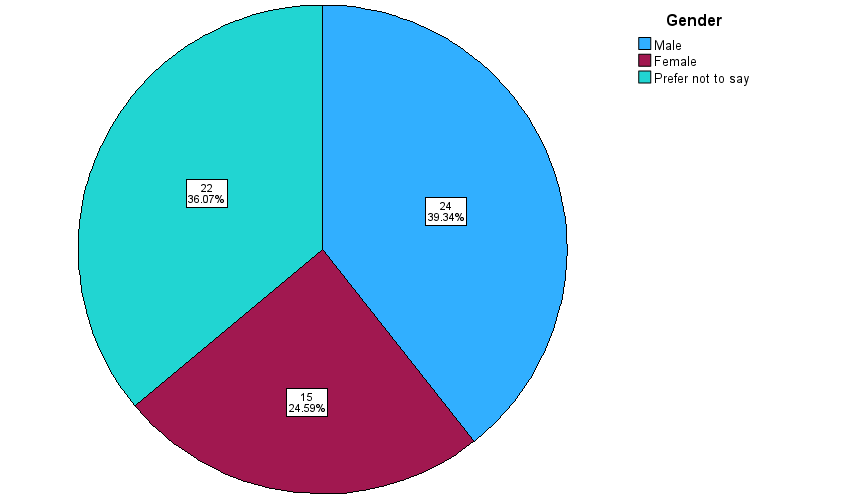
*<Intentionally Blank>*

# CHAPTER 4 - PRESENTATION OF RESULTS



## Demographic Data Analysis

1. **Gender Distribution**

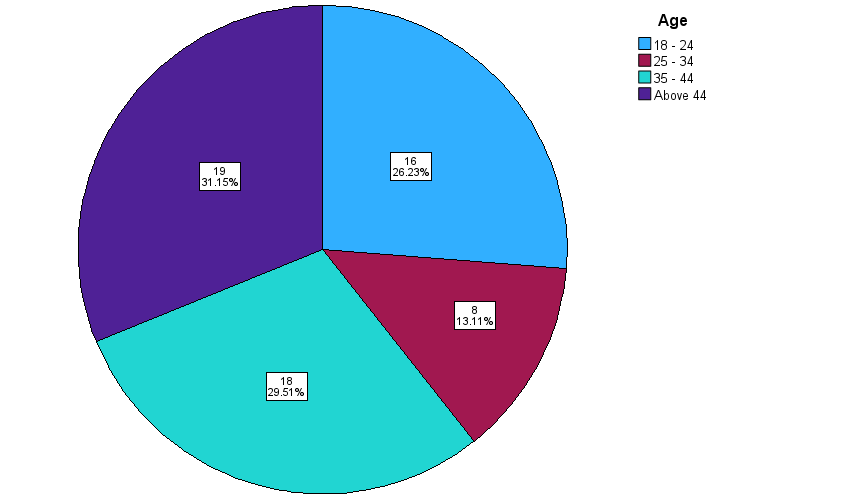


*Figure 1. Demographic Data Analysis Gender Distribution*

According to the research there are 61 respondents participating. Majority of them are Male. Statistically Male representation is 39.34% and it claims 24 respondents. 36.07% that means 22 respondents prefer not to say their gender. Meanwhile 24.59% female respondents have participated in this study and it claims 15 respondents.

*<Intentionally blank>*

1. Age Distribution



*Figure 2. Demographic Data Analysis Age Distribution*

Majority of respondents represent the above 44 years age group and statistically it

represents 31.15% (19 respondents). There are 18 respondents who represent 35-44 years group and 16 respondents represent 18 to 24 years age group. It represents

29.51% and 26.23% respectively. Minority represent 25-34 age group. It is 13.11% and

claims 8 respondents

*<Intentionally blank>*

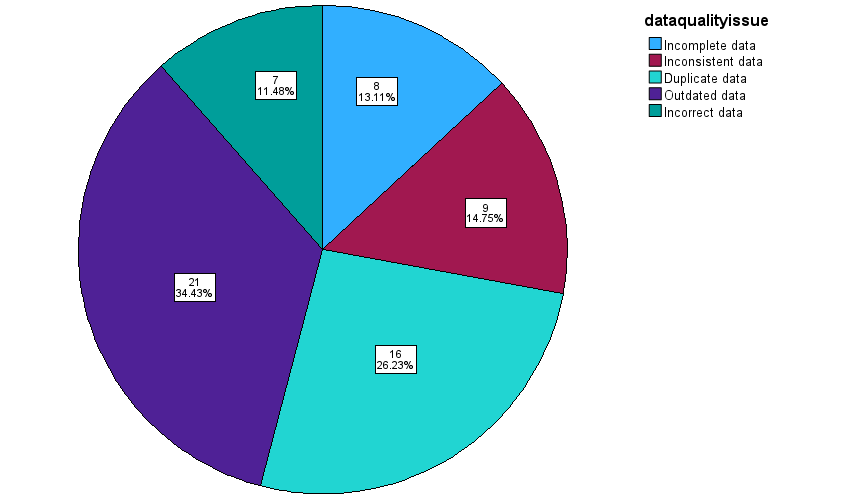


Figure 3. Demographic Data Analysis data quality issue Distribution

The majority of the data quality issues fall under the outdated data representing 34.43% of the cases, which amounts to 21 cases, while the second major class, duplicate data, amounts to 26.23% with 16 cases. The inconsistent data category is a little lower and amounts to 14.75% with 9 cases, while Incomplete data is in third place, representing 13.11% with 8 cases. The least issue is Incorrect data, constituting 11.48% and involving 7 cases.

*<Intentionally blank>*

## Correlation Analysis

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Correlations** | | | | | |
|  | | S01 | S02 | S03 | DV |
| S01 | Pearson Correlation | 1 | .141 | .036 | -.001 |
| Sig. (2-tailed) |  | .278 | .782 | .996 |
| N | 61 | 61 | 61 | 61 |
| S02 | Pearson Correlation | .141 | 1 | .111 | .306\* |
| Sig. (2-tailed) | .278 |  | .392 | .016 |
| N | 61 | 61 | 61 | 61 |
| S03 | Pearson Correlation | .036 | .111 | 1 | -.056 |
| Sig. (2-tailed) | .782 | .392 |  | .667 |
| N | 61 | 61 | 61 | 61 |
| DV | Pearson Correlation | -.001 | .306\* | -.056 | 1 |
| Sig. (2-tailed) | .996 | .016 | .667 |  |
| N | 61 | 61 | 61 | 61 |
| \*. Correlation is significant at the 0.05 level (2-tailed). | | | | | |

Table 4. Correlation Table

### RO2 / S01

Pearson correlations value between S01 and DV is -0.01

### RO3 / S02

Pearson correlation value between S02 and DV is 0.306

### RO4 / S03

Pearson correlations value between S03 and DV is -0.056

## Regression Analysis

### RO1 / Main Objective

**Model Summary table**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| 1 | .322a | .104 | .057 | 1.15455 |
| a. Predictors: (Constant), S03, S01, S02 | | | | |
| b. Dependent Variable: DV | | | | |

R-Value = 0.322

R-Square = 0.104

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 8.794 | 3 | 2.931 | 2.199 | .098b |
| Residual | 75.980 | 57 | 1.333 |  |  |
| Total | 84.775 | 60 |  |  |  |
| a. Dependent Variable: DV | | | | | | |
| b. Predictors: (Constant), S03, S01, S02 | | | | | | |

dfreg = 3

dfres = 57

F value = 2.199

P Value(sig) = .098

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Coefficients** | | | | | | |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| B | Std. Error | Beta |
| 1 | (Constant) | 1.955 | .907 |  | 2.156 | .035 |
| S01 | -.073 | .215 | -.043 | -.339 | .736 |
| S02 | .501 | .198 | .322 | 2.529 | .014 |
| S03 | -.114 | .159 | -.091 | -.717 | .476 |
| a. Dependent Variable: DV | | | | | | |

### RO2 / S01

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model Summary** | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| 1 | .001a | .000 | -.017 | 1.19869 |
| a. Predictors: (Constant), S01 | | | | |
| b. Dependent Variable: DV | | | | |

R-Value = 0.001

R-Square = 0.000

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | .000 | 1 | .000 | .000 | .996b |
| Residual | 84.775 | 59 | 1.437 |  |  |
| Total | 84.775 | 60 |  |  |  |
| a. Dependent Variable: DV | | | | | | |
| b. Predictors: (Constant), S01 | | | | | | |

dfreg = 1

dfres = 59

F value = 0.000

P Value(sig) = .996

### RO3 / S02

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model Summaryb | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| 1 | .306a | .094 | .078 | 1.14117 |
| a. Predictors: (Constant), S02 | | | | |
| b. Dependent Variable: DV | | | | |

R-Value = 0.306

R-Square = 0.094

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ANOVAa | | | | | | |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 7.941 | 1 | 7.941 | 6.098 | .016b |
| Residual | 76.833 | 59 | 1.302 |  |  |
| Total | 84.775 | 60 |  |  |  |
| a. Dependent Variable: DV | | | | | | |
| b. Predictors: (Constant), S02 | | | | | | |

dfreg = 1

dfres = 59

F value = 6.098

P Value(sig) = .016

### RO4 / S03

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model Summaryb** | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| 1 | .056a | .003 | -.014 | 1.19680 |
| a. Predictors: (Constant), S03 | | | | |
| b. Dependent Variable: DV | | | | |

R-Value = 0.056

R-Square = 0.003

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | .268 | 1 | .268 | .187 | .667b |
| Residual | 84.507 | 59 | 1.432 |  |  |
| Total | 84.775 | 60 |  |  |  |
| a. Dependent Variable: DV | | | | | | |
| b. Predictors: (Constant), S03 | | | | | | |

dfreg = 1

dfres = 59

F value = 0.187

P Value(sig) = 0.667

# CHAPTER 5 - CONCLUSIONS AND RECOMMENDATIONS



## Conclusion

**Correlation Table**

|  |  |
| --- | --- |
| Value | Result |
| 0.5 – 1 | Strong Positive Relationship |
| 0 – 0.5 | Moderate Positive Relationship |
| 0 | No Relationship |
| 0 – (-0.5) | Moderate Negative Relationship |
| (-0.5) – (-1) | Strong Negative Relationship |

**Model Summary Table**

R-value represents the correlation between the dependent and independent variables. A value greater than 0.5 is taken for further analysis.

R-square shows the total variation for the dependent variable (DV) that could be explained by the independent variables (IVs).

A value greater than 0.5 shows that the model is effective enough to determine the relationship.

**ANOVA Table**

* If Significant value (p Value) is less than or equal to 0.05; Significant relationship between variables and H1 Accepted, HO Rejected.
* If Significant value (p Value) greater than or equal to 0.05; No Significant relationship between variables and HO Accepted, H1 Rejected
* If P value is greater than or equal to 0; then Regression is fit to model and statistically significant relationship

It can be represented statistically as

F= (dfreg, dfres)= f value, p Value

### RO1

**Regression**

In this case R-Value is 0.322, researcher have not taken this for further analysis

In this case R-Square is 0.104 which seems to be good. If the value is greater than 0.5 is considered as good in this case 10.4% dependent variable (DV) that could be explained by the independent variable (IVs). There are factors affect for this which is not part of the study.

Here p Value greater than or equal to 0.05; No Significant relationship between variables and main objective and therefore researcher accept HO and reject H1.

H0. There are no challenges associated with implementation of Data driven solution.

It can be represented statistically as

F= (3,57) = 2.199, 0.098

### RO2 / S01

**Correlation**

Pearson corelation value between S01 and DV is -0.01 according to the corelation table it is a Moderate Negative Relationship

**Regression**

In this case R-Value is 0.001, researcher have not taken this for further analysis.

In this case R-Square is 0.000 which is not good. If the value is greater than 0.5 is considered as good, in this case 0% dependent variable (DV) that could be explained by the independent variable (IVs). There are factors affect for this which is not part of the study.

Here p Value greater than or equal to 0.05; No Significant relationship between variables and sub objective and therefore researcher accept HO and reject H2.

H0. Data Quality is not a challenge associated with implementation of Data driven solution.

It can be represented statistically as

F= (1,59) = 0.000, 0.996

### RO3 / S02

**Correlation**

Pearson corelation value between S02 and DV is 0.306 according to the corelation table it is a Moderate Positive Relationship

**Regression**

In this case R-Value is 0.306, researcher have not taken this for further analysis.

In this case R-Square is 0.094 which seems to be good. If the value is greater than 0.5 is considered as good, in this case 9.4% dependent variable (DV) that could be explained by the independent variable (IVs). There are factors affect for this which is not part of the study.

Here p Value less than or equal to 0.05; Significant relationship between variables and sub objective therefore researcher Accept H3 and Reject H0

H3. Is Technology a main challenge when implementing data driven solution

It can be represented statistically as

F= (1,59) = 6.098, 0.016

*<Intentionally blank>*

### RO4 / S03

**Correlation**

Pearson corelation value between S03 and DV is -0.056 according to the corelation table it is a Moderate Negative Relationship

**Regression**

In this case R-Value is 0.056, researcher have not taken this for further analysis.

In this case R-Square is 0.003 which is not good. If the value is greater than 0.5 is considered as good, in this case 0.3% dependent variable (DV) that could be explained by the independent variable (IVs). There are factors affect for this which is not part of the study.

Here p Value greater than or equal to 0.05; No Significant relationship between variables and sub objective and therefore researcher accept HO and reject H4.

H0. No user management is not a challenge associated with implementing data driven solutions?

It can be represented statistically as

F= (1,59) = 0.187, 0.667

*<Intentionally blank>*

## Recommendations

Though the data analysis in this research proved to be is reliable up to certain extent, the concepts discussed in the literature review remain valid and provide immense value to organizations in using big data for consumer behaviour analysis. We can recommend the following methods based on established studies with complete confidence:

* Data Quality Assurance

It is essential for organizations to implement rigorous data validation and cleaning mechanisms to ensure that the insights drawn from big data are reliable. This involves removing inconsistencies, filling gaps, and standardizing data formats, which will help enhance the dependability of any analysis. (Redman, 2018) emphasizes that poor data quality can lead to misleading conclusions. Setting up tight data quality controls helps organizations draw valid and actionable conclusions. Organizations should be proactive in cleaning up and maintaining high-quality data, as this is key to accurate insights.

* Scalable Data Analytics Platforms

Enterprises must use scalable and flexible data analytics platforms in their architecture to meet the ever-increasing volumes of data. Open-source and cloud-based options also enable smaller and mid-sized organizations to seek out more affordable solutions. As (Davenport & Dyché, 2013) note, one of the most significant impediments to dealing with big data is that many organizations' infrastructures are simply obsolete. Upgrading to a modern, scalable platform will greatly aid an organization in these challenges and better position it to store and analyse vast consumer data.

* Employee Training in Data Analytics

For better implementation of data-driven strategies, organizations should focus on regular training for their employees on data analytics tools. A data-literate workforce can interpret and apply the information derived from data more appropriately, thus improving decision-making and performance. This supports the work of (Smith, J., & Johnson, A., 2024), who highlight the importance of leveraging social media analytics for strategic decision-making. Even the most advanced analytics tools may not realize their full potential without skilled personnel to utilize them.

* Clear Data Governance Frameworks

There is a need for clear data governance frameworks that ensure ethical use of data and, at the same time, make sure that regulations like GDPR are followed. Data should be collected and processed in a very transparent manner to build trust among consumers. As (Kumar & Wilson, 2020) discuss, there are a number of ethical issues associated with social media analytics, particularly regarding user privacy and consent. Strong governance frameworks will enable organizations to manage consumer data responsibly in compliance with regulatory standards.

* Periodic Revaluation of Marketing and Engagement Strategies

Companies are supposed to reassess their marketing and engagement strategies from time to time to stay in tune with the latest developments in consumer preferences and the market environment. (Chen, Chiang, & Storey, 2012) emphasize the role of predictive analytics in optimizing marketing campaigns and improving customer experience. Organisations will be able to hold on to relevance and keep audiences engaged by renewing such strategies regularly with fresh insights.

## Limitations

**Type of Research**

Quantitative research was applied in this study, which is efficient for statistical analysis but lacks the depth that qualitative methods would have given. While quantitative methods enable broad generalizations, they lack the in-depth insight into consumer behaviour that might be uncovered with qualitative methods, such as interviews or focus groups. Further studies should combine both aspects to get a holistic view of the research topic.

**Research Method**

The survey-based method of data collection was quick in the collection of data, but it brought several limitations. Surveys can be prone to response biases, such as social desirability bias, wherein respondents may provide answers they believe are more socially acceptable rather than their true opinions. In addition, complex consumer behaviours or motivations could not be fully captured with a survey format and might have been better understood through in-depth interviews or observational methods.

**Tool of Data Collection**

The data collection tool used is Google Forms, which, although widely accessible and easy to fill out, constrained the level of diversity in responses. Given that the shared links were mainly distributed through the small circle of friends and acquaintances, this could result in a sampling bias where the sample was not quite representative. Moreover, in online surveys, certain segments usually have low response rates; hence, data from this group may not be properly reflected. Future research should explore multiple data collection channels to ensure a broader and more representative sample.

**Sample Size**

The second limitation involves the sample size 60 participants are included. This, while manageable for a pilot study, is a rather small size and may not capture the full diversity of consumer behaviours in larger, more varied populations. Increasing the sample size would have the added advantage of raising the statistical power of the study, which enhances the reliability of the findings and increases generalizability to larger populations.

**Literacy**

This can be attributed to the level of literacy among the respondents, since those with low levels might have poorly interpreted the survey questions or even failed to respond correctly. This introduces errors or inconsistencies in the data, making it less reliable for generalizing research findings.

**Awareness of Big Data**

Respondents’ awareness of big data concepts may have influenced their understanding of the survey questions. Particularly if they were not familiar with the scope of the research topic. Lack of awareness or limited understanding of data analytics may have led to biased or inaccurate responses, further limiting the validity of the study

**Geographical Factors**

The geographical location of the respondents is also a limitation. Because the questionnaire was distributed in only one region, it could not capture different behaviours, preferences, or challenges that may be relevant to other regions or countries. Geographical factors are very powerful in influencing consumer behaviour, and future research should involve participants from a variety of locations to increase diversity and generalizability of the findings.

**Data reliability**

Due to limitation bias, respondent literacy, awareness of the research topic, the reliability of the data may be compromised. Therefore, as much as findings give important information, they cannot completely represent the wide scope of behaviours of a broader population. Future research might advance the reliability of data through an increased sample size that is more varied and might complement the quantitative analysis with some qualitative approach.

## Future Improvements

* Future research should encompass qualitative methods of interviews or focus groups to delve deeper into motivations and preferences of consumers.
* Data collection should be done for a greater and more diverse population, covering various demographic factors of age, gender, occupation, and location.
* Advanced tools such as machine learning algorithms and predictive analytics should be integrated to provide accurate and actionable insights.
* A longitudinal approach will, therefore, enable researchers to monitor changes in consumer behaviour over time and present a dynamic view of emergent trends.
* Collaborations with organizations could offer access to such real-world datasets, which would enhance the practical relevance and applicability of the research results.

## Personnel Reflection

### Benefits for the researcher

This research has been an enriching experience, offering me numerous personal and professional benefits.

1. Skill Development

The process enhanced my analytical skills, particularly in handling large datasets and applying statistical tools to extract meaningful insights. These skills are crucial for my career aspirations as a cloud engineer.

1. Deepened Understanding

I had a deep understanding of big data in consumer behaviour analysis and its potential to drive effective marketing strategy and enhance user engagement.

1. Practical Application of Knowledge

The research has provided an avenue for theoretical application in real life, thereby connecting the gap between academic knowledge and practical problems.

1. Problem-Solving Abilities

Such challenges as limited sample size and a potential bias in data collection involved, problem-solving and critical thinking became sharper.

### Benefits for the selected Industry/organization

* Organizations can make evidence-based decisions through the insights developed in this study, hence enhancing operational efficiency.
* Understanding consumer behaviour allows businesses to design targeted campaigns that align with customer preferences, increasing ROI.
* With data-driven insights, organizations can design personalized experiences for users, which lead to loyalty and satisfaction.
* Data analytics helps in deriving the maximum out of marketing spend by focusing investments on high-impact activities.
* Organizations could find a sustained competitive advantage in knowing and working ahead of evolving consumer trends by the adoption of the recommended practices.

# Referencing

Anon, 2023. Big Data and Social Media for Better Marketing. [online] Available at: https://innovatureinc.com/big-data-and-social-media-for-better-marketing/#:~:text=Social%20media%20platforms%20use%20big%20data%20to%20power%20content%20recommendation [Accessed 27 Nov. 2024].

Boyd, D. & Crawford, K., 2012. Critical Questions for Big Data. Information, Communication & Society, 15(5), pp. 662-679.

Calzon, B., 2023. Learn Here Different Ways of Data Analysis Methods & Techniques. [online] BI Blog | Data Visualization & Analytics Blog | Datapine. Available at: https://www.datapine.com/blog/data-analysis-methods-and-techniques/ [Accessed 27 Nov. 2024].

Chen, H., Chiang, R.H. & Storey, V.C., 2012. Business Intelligence and Analytics: From Big Data to Big Impact. MIS Quarterly, 36(4), pp. 1165-1188.

Davenport, T.H. & Dyché, J., 2013. Big Data in Big Companies. International Institute for Analytics, s.l.: s.n.

Fan, W. & Gordon, M.D., 2014. The Power of Social Media Analytics. Communications of the ACM, 57(6), pp. 74-81.

Gandomi, A. & Haider, M., 2015. Beyond the Hype: Big Data Concepts, Methods, and Analytics. International Journal of Information Management, 35(2), pp. 137-144.

Greene, W.H., n.d. Econometric Analysis. 8th ed. Harlow: Pearson Education Limited.

Jones, R. & Smith, T., 2021. Data-Driven Insights into User Behaviour. Journal of Marketing Analytics, 9(3), pp. 234-245.

Kumar, S. & Wilson, L., 2020. Ethical Considerations in Social Media Analytics. Social Media & Society, 6(2), pp. 1-10.

Lee, C. & Garcia, P., 2019. Cross-Domain Insights from Social Media Analytics. Big Data Research, 18, pp. 54-64.

Patel, R. & Brown, D., 2020. Exploring Machine Learning Techniques in Social Media Analysis. Journal of Big Data, 7(1), pp. 1-18.

ProcessMaker, n.d. How to Use Big Data to Improve Social Media Marketing. [online] Available at: https://www.processmaker.com/blog/how-to-use-big-data-to-improve-social-media-marketing/ [Accessed 27 Nov. 2024].

Radu, V., 2023. Consumer Behavior in Marketing - Patterns, Types, Segmentation. [online] Omniconvert Ecommerce Growth Blog. Available at: https://www.omniconvert.com/blog/consumer-behavior-in-marketing-patterns-types-segmentation/ [Accessed 27 Nov. 2024].

Redman, T.C., 2018. Data Driven: Profiting from Your Most Important Business Asset. Boston: Harvard Business Review Press.

Saunders, M., Lewis, P. & Thornhill, A., n.d. Research Methods for Business Students. 8th ed. Harlow: Pearson Education Limited.

Smith, J. & Johnson, A., 2024. Using Big Data to Understand Consumer Behaviour: Insights for Marketing, User Engagement, and Strategy Planning. Journal of Business Analytics, 12(4), pp. 89-105.

Sofi, 2023. How to Conduct a Customer Behavior Analysis: A Complete Guide. [online] Userpilot Blog. Available at: https://userpilot.com/blog/customer-behavior-analysis/ [Accessed 27 Nov. 2024].

Sturdy, C., 2022. 5 Ways How to Use Big Data in Social Media. [online] Quintly Blog. Available at: https://www.quintly.com/blog/5-ways-how-to-use-big-data-in-social-media [Accessed 27 Nov. 2024].

Thattamparambil, N., 2020. How to Choose the Research Methodology Best Suited for Your Study. [online] Editage Insights. Available at: https://www.editage.com/insights/how-to-choose-the-research-methodology-best-suited-for-your-study [Accessed 27 Nov. 2024].

# Annexures

## Annexures A: Glossary of Terms

Big Data

Large, complicated datasets that are difficult to process using traditional data management tools. It is used to track consumer behavior, marketing techniques, and other business operations.

Correlation Analysis

A statistical technique used to determine if a relationship exists between two or more variables, including the strength of the relationship.

Regression Analysis

A statistical used to understand the relationship between a dependent variable and one or more independent variables, often used for prediction.

Consumer Behavior

The study of how individuals or groups make decisions to purchase or use products, services, or ideas. This includes preferences, needs, and buying habits of consumers.

Data-Driven Solutions

Decisions or strategies based on the analysis and interpretation of data rather than intuition or personal experience.

Marketing Strategies

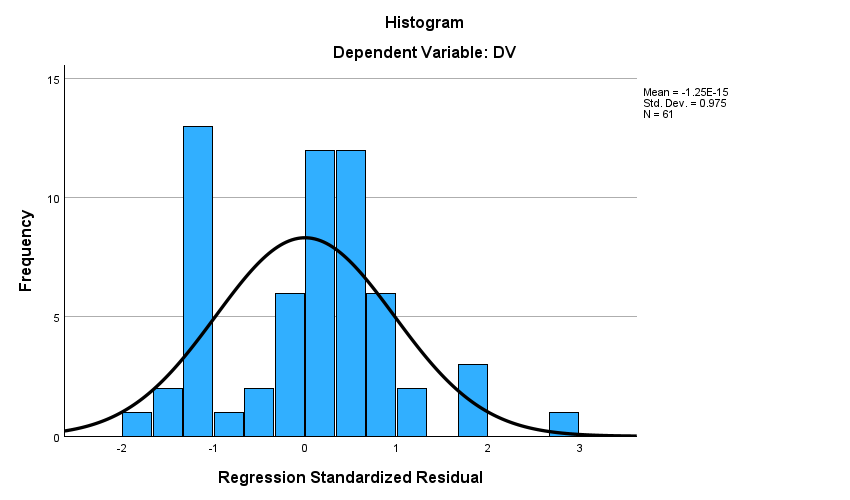
Plans and actions, developed to promote products or services and engage with consumers effectively, are often informed by data analysis and consumer behavior insights.

## Annexures B: Sample SPSS Charts/ Table

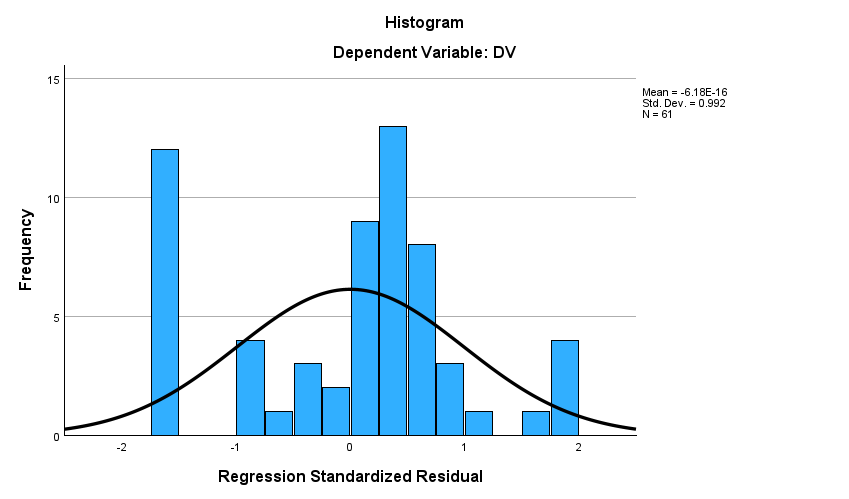
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Statistics** | | | | | |
|  | | S01 | S02 | S03 | DV |
| N | Valid | 61 | 61 | 61 | 61 |
| Missing | 0 | 0 | 0 | 0 |
| Mean | | 2.9557 | 2.9125 | 3.0769 | 2.8469 |
| Median | | 3.0000 | 2.6700 | 3.0000 | 3.0000 |
| Std. Deviation | | .70054 | .76512 | .94255 | 1.18866 |

*Table 5 frequency table*

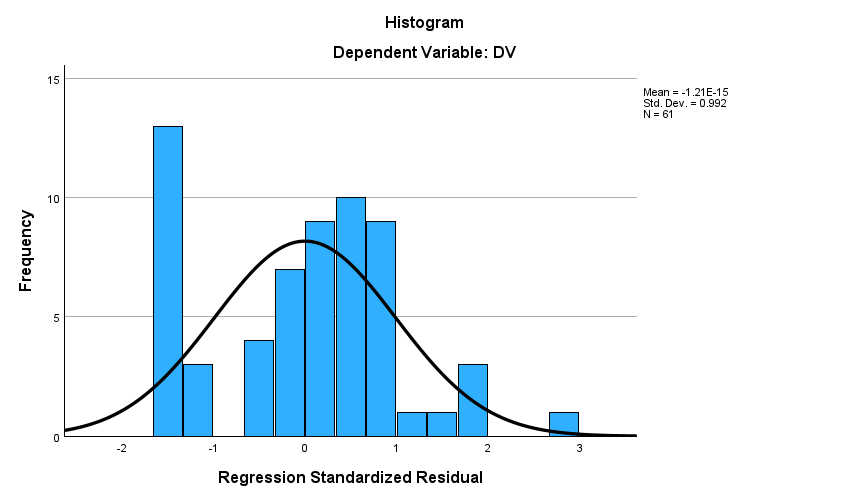
RO1/ Main Objective

**

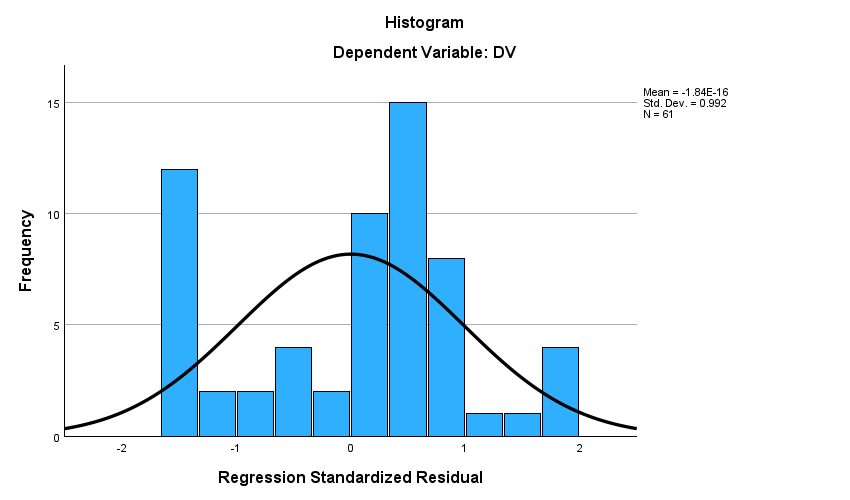
RO2/ S01

**

**RO3/ S02**

**

**RO4/ S03**

****

## Annexures C: Feedback Form / Survey question list / Interview question list

**Demographic Questions**

1. Age

* 18 – 24
* 25 – 34
* 35 – 44
* Above 44

2. Gender

* Male
* Female
* Prefer not to say

3. Data quality issues do you face most frequently?

* Incomplete data
* Inconsistent data
* Duplicate data
* Outdated data
* Incorrect data

4. Which technological challenge is the most significant barrier in your data-driven projects?

* Lack of proper tools and software
* Outdated software
* Difficulty in integrating with existing systems
* Lack of skilled people
* Concerns about data privacy and security

5. What is the main reason for low user engagement in data-driven Solutions?

* Lack of understanding
* Resistance to change
* Insufficient training
* Unclear benefits
* Poor communication

**Data Quality is challenges associated with implementation of Data driven solution**

1. How confident are you in your organization’s ability to address data quality issues in data-driven solutions?

* Very confident
* Confident
* Neutral
* Not confident
* Not at all confident

2. Data quality issues when implementing data-driven solutions

Linear scale

Agree ---Disagree

3. Security measures in place regularly basis will impact to improve data quality

Linear scale

Agree ---Disagree

**Technology is a Main challenge when implementing data driven solution**

1. How often do technological limitations hinder the implementation of data-driven solutions in your organization?

* Never
* Occasionally
* Sometimes
* Often
* Always

2. Technological limitations hindering the implementation of data-driven solutions in your organization?

Linear scale

Agree ---Disagree

3. Difficulties in integrating new data-driven technologies with your existing systems?

Linear scale

Agree ---Disagree

**User engagement is also a challenge when implementing data driven solution**

1. On a scale of 1 to 5, Users engagement with the data-driven tools

Linear scale

Very low ---Very High

2. Users preferences for using new data tools?

* Ease of use
* Comprehensive features
* Quick access to insights
* Customizability
* Compatibility with other tools

2. Communication strategies in promoting user engagement with data-driven solutions?

* Online tutorials and webinars
* In-person workshops and training
* Regular updates and newsletters
* Feedback loops and surveys
* Gamification and incentives

**DV**

1. On a scale from 1 to 5, data quality issues are encountered frequently when implementing data-driven solutions.

Linear scale

Never ---always

2. How satisfied are you with the technology stack used for data-driven solutions?

Linear scale

Very satisfied ---very dissatisfied

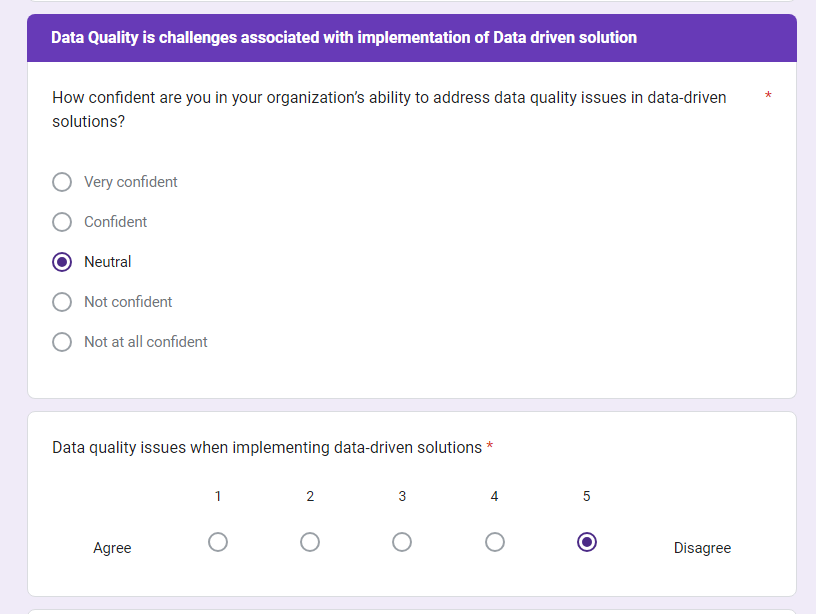
3. How engaged are end-users in the implementation process of data-driven solutions?

Linear scale

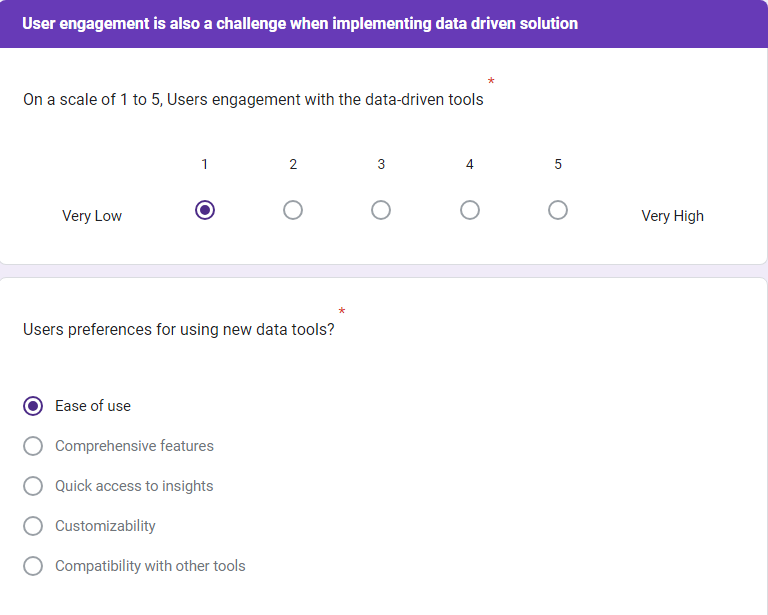
Very engaged --- not engaged at all

## Annexures D: Sample Feedback sheets

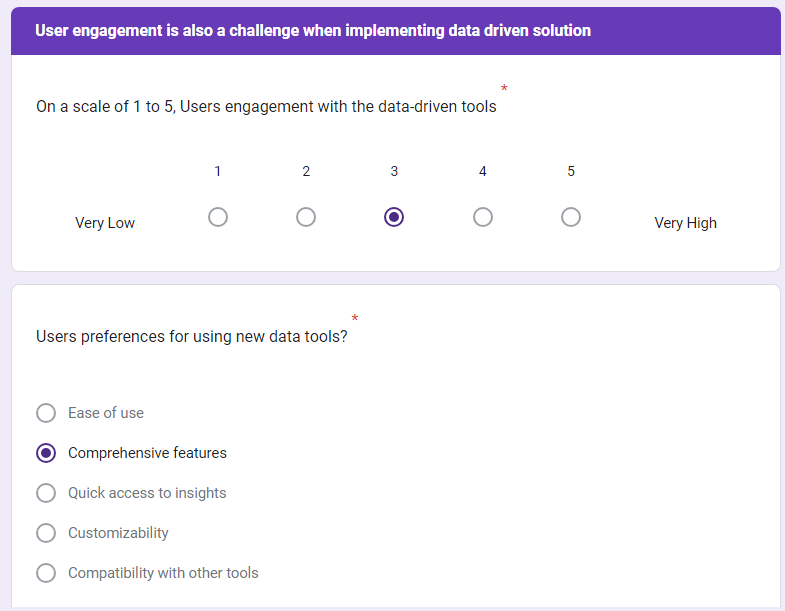
Sample 1



Sample 2:



Sample: 3



## Annexures E: Project logbook

|  |  |  |
| --- | --- | --- |
| **Section 1: Basic Details** | | |
| **Project title:** | Using Big Data to Understand Consumer Behaviour: Insights for Marketing, User Engagement, and Strategy Planning | |
| **Student name:** | Imamuddeen Muhammedh | |
| **Student number:** | 1001321 | |
| **Week** | 5  6  7  8  9  10  11  12 | |
| **Date** |  | |
|  | | |
| **Section 2: Weekly Progress** | | |
| **Issues identified and Points to consider:** | | |
| Started by creating the Google Form for data collection. Had to think about which questions would help with my research.  It took some time to decide on the format for the form. | | |
| **New ideas and change of project direction Points to consider:** | | |
| Realized I needed to simplify the questions to make it easier for people to respond. | | |
| **What have you learned about yourself through your work?** | | |
| Learned to create a google form | | |
| **Next steps for your work Points to consider:** | | |
| Share the Google Form with people to get responses. | | |
| **Supervisor feedback / comments** | | |
|  | | |
| **Supervisor Sign:** Anuradha Boyagoda | | **Date:** |

|  |  |  |
| --- | --- | --- |
| **Section 1: Basic Details** | | |
| **Project title:** | Using Big Data to Understand Consumer Behaviour: Insights for Marketing, User Engagement, and Strategy Planning | |
| **Student name:** | Imamuddeen Muhammedh | |
| **Student number:** | 1001321 | |
| **Week** | 5  6  7  8  9  10  11  12 | |
| **Date** |  | |
|  | | |
| **Section 2: Weekly Progress** | | |
| **Issues identified and Points to consider:** | | |
| After sharing the Google Form, I received a mix of answers. Some responses were unclear. | | |
| **New ideas and change of project direction Points to consider:** | | |
| Decided to focus on responses that gave clear and useful information. | | |
| **What have you learned about yourself through your work?** | | |
| I learned how to handle different types of data and organize it for analysis | | |
| **Next steps for your work Points to consider:** | | |
| Continue analyzing the responses and identifying patterns that directly answer my research questions. | | |
| **Supervisor feedback / comments** | | |
|  | | |
| **Supervisor Sign:** Anuradha Boyagoda | | **Date:** |

|  |  |  |
| --- | --- | --- |
| **Section 1: Basic Details** | | |
| **Project title:** | Using Big Data to Understand Consumer Behaviour: Insights for Marketing, User Engagement, and Strategy Planning | |
| **Student name:** | Imamuddeen Muhammedh | |
| **Student number:** | 1001321 | |
| **Week** | 5  6  7  8  9  10  11  12 | |
| **Date** |  | |
|  | | |
| **Section 2: Weekly Progress** | | |
| **Issues identified and Points to consider:** | | |
| Analyzing the data was challenging. | | |
| **New ideas and change of project direction Points to consider:** | | |
| Focus on cleaning the data and getting it ready for analysis. | | |
| **What have you learned about yourself through your work?** | | |
| I learned how to manage missing data and clean up survey responses. | | |
| **Next steps for your work Points to consider:** | | |
| Start analyzing the data in SPSS. | | |
| **Supervisor feedback / comments** | | |
|  | | |
| **Supervisor Sign:** Anuradha Boyagoda | | **Date:** |

|  |  |  |
| --- | --- | --- |
| **Section 1: Basic Details** | | |
| **Project title:** | Using Big Data to Understand Consumer Behaviour: Insights for Marketing, User Engagement, and Strategy Planning | |
| **Student name:** | Imamuddeen Muhammedh | |
| **Student number:** | 1001321 | |
| **Week** | 5  6  7  8  9  10  11  12 | |
| **Date** |  | |
|  | | |
| **Section 2: Weekly Progress** | | |
| **Issues identified and Points to consider:** | | |
| Began data analysis using SPSS.  Some variables were not showing significant results as expected. | | |
| **New ideas and change of project direction Points to consider:** | | |
| Revisited the data and decided to adjust the variables for clearer analysis. | | |
| **What have you learned about yourself through your work?** | | |
| I learned about basic data analysis techniques and how to apply them. | | |
| **Next steps for your work Points to consider:** | | |
| Continue with analysis in SPSS | | |
| **Supervisor feedback / comments** | | |
|  | | |
| **Supervisor Sign:** Anuradha Boyagoda | | **Date:** |

|  |  |  |
| --- | --- | --- |
| **Section 1: Basic Details** | | |
| **Project title:** | Using Big Data to Understand Consumer Behaviour: Insights for Marketing, User Engagement, and Strategy Planning | |
| **Student name:** | Imamuddeen Muhammedh | |
| **Student number:** | 1001321 | |
| **Week** | 5  6  7  8  9  10  11  12 | |
| **Date** |  | |
|  | | |
| **Section 2: Weekly Progress** | | |
| **Issues identified and Points to consider:** | | |
| Completed initial analysis and identified some interesting trends.  Had to refine some regression models. | | |
| **New ideas and change of project direction Points to consider:** | | |
| Focused more on specific customer segments rather than general trends | | |
| **What have you learned about yourself through your work?** | | |
| I learned how to refine regression models in SPSS to get better results. | | |
| **Next steps for your work Points to consider:** | | |
| Interpret the results and start writing the report. | | |
| **Supervisor feedback / comments** | | |
|  | | |
| **Supervisor Sign:** Anuradha Boyagoda | | **Date:** |

|  |  |  |
| --- | --- | --- |
| **Section 1: Basic Details** | | |
| **Project title:** | Using Big Data to Understand Consumer Behaviour: Insights for Marketing, User Engagement, and Strategy Planning | |
| **Student name:** | Imamuddeen Muhammedh | |
| **Student number:** | 1001321 | |
| **Week** | 5  6  7  8  9  10  11  12 | |
| **Date** |  | |
|  | | |
| **Section 2: Weekly Progress** | | |
| **Issues identified and Points to consider:** | | |
| Completed the data analysis and began interpreting the results.  Started drafting the report based on the findings. | | |
| **New ideas and change of project direction Points to consider:** | | |
| Decided to focus on actionable recommendations based on the results. | | |
| **What have you learned about yourself through your work?** | | |
| I learned how to extract meaningful insights from complex data | | |
| **Next steps for your work Points to consider:** | | |
| Finalize the report and incorporate feedback. | | |
| **Supervisor feedback / comments** | | |
|  | | |
| **Supervisor Sign:** Anuradha Boyagoda | | **Date:** |

|  |  |  |
| --- | --- | --- |
| **Section 1: Basic Details** | | |
| **Project title:** | Using Big Data to Understand Consumer Behaviour: Insights for Marketing, User Engagement, and Strategy Planning | |
| **Student name:** | Imamuddeen Muhammedh | |
| **Student number:** | 1001321 | |
| **Week** | 5  6  7  8  9  10  11  12 | |
| **Date** |  | |
|  | | |
| **Section 2: Weekly Progress** | | |
| **Issues identified and Points to consider:** | | |
| Continued working on the final draft of the report.  Needed to adjust some conclusions based on new insights. | | |
| **New ideas and change of project direction Points to consider:** | | |
| Emphasized more on how big data can improve marketing strategies. | | |
| **What have you learned about yourself through your work?** | | |
| I’ve become more confident in presenting data-driven insights and making recommendations. | | |
| **Next steps for your work Points to consider:** | | |
| Review and finalize the report for submission. | | |
| **Supervisor feedback / comments** | | |
|  | | |
| **Supervisor Sign:** Anuradha Boyagoda | | **Date:** |

|  |  |  |
| --- | --- | --- |
| **Section 1: Basic Details** | | |
| **Project title:** | Using Big Data to Understand Consumer Behaviour: Insights for Marketing, User Engagement, and Strategy Planning | |
| **Student name:** | Imamuddeen Muhammedh | |
| **Student number:** | 1001321 | |
| **Week** | 5  6  7  8  9  10  11  12 | |
| **Date** |  | |
|  | | |
| **Section 2: Weekly Progress** | | |
| **Issues identified and Points to consider:** | | |
| Completed the report and finalized the research project.  Focused on ensuring the report met all the required guidelines and standards. | | |
| **New ideas and change of project direction Points to consider:** | | |
| Decided to highlight more actionable recommendations based on data insights for businesses. | | |
| **What have you learned about yourself through your work?** | | |
| I’ve learned how to complete a research project from start to finish, including all stages of data collection, analysis, and reporting | | |
| **Next steps for your work Points to consider:** | | |
| Submit the final report. | | |
| **Supervisor feedback / comments** | | |
|  | | |
| **Supervisor Sign:** Anuradha Boyagoda | | **Date:** |

# Turnitin report



