**Higher Nationals**

Internal verification of assessment decisions – BTEC (RQF)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **INTERNAL VERIFICATION – ASSESSMENT DECISIONS** | | | | | | |
| **Programme title** | BTEC Higher National Diploma in Computing | | | | | |
| **Assessor** | Mr. Lasitha Ranawaka | | **Internal Verifier** | |  | |
| **Unit(s)** | Unit 16: Computing Research Project (Pearson Set) | | | | | |
| **Assignment title** | Final Research Report – Big Data | | | | | |
| **Student’s name** | Mohamed Zaheer | | | | | |
| **List which assessment criteria the Assessor has awarded.** | **Pass** | | **Merit** | **Distinction** | | |
|  | |  |  | | |
| **INTERNAL VERIFIER CHECKLIST** | | | | | | |
| **Do the assessment criteria awarded match those shown in the assignment brief?** | | Y/N |  | | | |
| **Is the Pass/Merit/Distinction grade awarded justified by the assessor’s comments on the student work?** | | Y/N |  | | | |
| **Has the work been assessed accurately?** | | Y/N |  | | | |
| **Is the feedback to the student:**  Give details:   * Constructive? * Linked to relevant assessment criteria? * Identifying opportunities for improved performance? * Agreeing actions? | | Y/N Y/N  Y/N Y/N |  | | | |
| **Does the assessment decision need amending?** | | Y/N |  | | | |
| **Assessor signature** | |  | | | **Date** |  |
| **Internal Verifier signature** | |  | | | **Date** |  |
| **Programme Leader signature** (if required) | |  | | | **Date** |  |

|  |  |  |  |
| --- | --- | --- | --- |
| **Confirm action completed** | | | |
| **Remedial action taken**  Give details: |  | | |
| **Assessor signature** |  | **Date** |  |
| **Internal Verifier signature** |  | **Date** |  |
| **Programme Leader signature** (if required) |  | **Date** |  |

Higher Nationals - Summative Assignment Feedback Form

|  |  |  |  |
| --- | --- | --- | --- |
| **Student Name/ID** | Mohamed Zaheer / E171743 | | |
| **Unit Title** | Unit 16: Computing Research Project (Pearson Set) | | |
| **Assignment Number** |  | **Assessor** | Mr. Lasitha Ranawaka |
| **Submission Date** | 2024.06.26 | **Date Received 1st submission** |  |
| **Re-submission Date** |  | **Date Received 2nd submission** |  |
| **Assessor Feedback:**   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | **LO2 Conduct and analyse research relevant to a computing research project** | | | | | | | | **Pass, Merit & Distinction Descripts** | **P3** | **P4** | **M2** | **D1** |  |  | | **LO3 Communicate the outcomes of a research project to identified stakeholders** | | | | | | | | **Pass, Merit & Distinction Descripts** | **P5** | **M3** | **D2** |  |  |  | | **LO4 Reflect on the application of research methodologies and concepts** | | | | | | | | **Pass, Merit & Distinction Descripts** | **P6** | **P7** | **M4** | **D3** |  |  | | | | |
| **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.

**Assignment Feedback**

|  |  |  |  |
| --- | --- | --- | --- |
| **Formative Feedback: Assessor to Student** | | | |
| **Action Plan** | | | |
| **Summative feedback** | | | |
| **Feedback: Student to Assessor** | | | |
| **Assessor signature** | **Lasith.Ranawaka@esoft.lk** | **Date** |  |
| **Student signature** | **mhdzaheer2003@gmail.com** | **Date** |  |



**Pearson**

**Higher Nationals in**

**Computing**

Unit 16: Computing Research Project

(Pearson Set)

Final Research Report

**General Guidelines**

1. A Cover page or title page – You should always attach a title page to your assignment. Use previous page as your cover sheet and make sure all the details are accurately filled.
2. Attach this brief as the first section of your assignment.
3. All the assignments should be prepared using a word processing software.
4. All the assignments should be printed on A4 sized papers. Use single side printing.
5. Allow 1” for top, bottom, right margins and 1.25” for the left margin of each page.

**Word Processing Rules**

1. The font size should be **12** point and should be in the style of **Time New Roman**.
2. **Use 1.5 line spacing**. Left justify all paragraphs.
3. Ensure that all the headings are consistent in terms of the font size and font style.
4. Use **footer function in the word processor to insert Your Name, Subject, Assignment No, and Page Number on each pag**e. This is useful if individual sheets become detached for any reason.
5. Use word processing application spell check and grammar check function to help editing your assignment.

**Important Points:**

1. It is strictly prohibited to use textboxes to add texts in the assignments, except for the compulsory information. eg: Figures, tables of comparison etc. Adding text boxes in the body except for the before mentioned compulsory information will result in rejection of your work.
2. Carefully check the hand in date and the instructions given in the assignment. Late submissions will not be accepted.
3. Ensure that you give yourself enough time to complete the assignment by the due date.
4. Excuses of any nature will not be accepted for failure to hand in the work on time.
5. You must take responsibility for managing your own time effectively.
6. If you are unable to hand in your assignment on time and have valid reasons such as illness, you may apply (in writing) for an extension.
7. Failure to achieve at least PASS criteria will result in a REFERRAL grade.
8. Non-submission of work without valid reasons will lead to an automatic REFERRAL. You will then be asked to complete an alternative assignment.
9. If you use other people’s work or ideas 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.
10. If you are proven to be guilty of plagiarism or any academic misconduct, your grade could be reduced to A REFERRAL or at worst you could be expelled from the course

**Student Declaration**

I hereby, declare that I know what plagiarism entails, namely to use another’s work and to present it as my own without attributing the sources in the correct way. I further understand what it means to copy another’s work.

1. I know that plagiarism is a punishable offence because it constitutes theft.
2. I understand the plagiarism and copying policy of the Pearson UK.
3. I know what the consequences will be if I plagiaries or copy another’s work in any of the assignments for this program.
4. I declare therefore that all work presented by me for every aspects of my program, will be my own, and where I have made use of another’s work, I will attribute the source in the correct way.
5. I acknowledge that the attachment of this document signed or not, constitutes a binding agreement between myself and Pearson UK.
6. I understand that my assignment will not be considered as submitted if this document is not attached to the attached.

**mhdzaheer2003@gmail.com**

**Student’s Signature: Date:** 2024.06.26

**(*Provide E-mail ID*) (*Provide Submission Date*)**

Assignment Brief

|  |  |
| --- | --- |
| Student Name /ID Number | Mohamed Zaheer / E171743 |
| **Unit Number and Title** | Unit 16: Computing Research Project (Pearson Set) |
| Academic Year | 2024 |
| Unit Tutor |  |
| **Assignment Title** | **Final Research Project Report -Big Data** |
| Issue Date |  |
| Submission Date |  |
| IV Name & Date |  |
| **Submission Format:** | |
| * The submission is in the form of an individual written report. * This should be written in a concise, formal business style using single spacing and font size 12. * You are required to make use of headings, paragraphs and subsections as appropriate, and all work must be supported with research * referenced using the Harvard referencing system. * Please provide a referencing list using the Harvard referencing system. * The recommended word limit is minimum 4,500 words * Copy of the research proposal need to be attached to the final project report. | |

|  |
| --- |
| **Unit Learning Outcomes:** |
| **LO2**. Conduct and analyse research relevant to a computing research project  **LO3**. Communicate the outcomes of a research project to identified stakeholders  **LO4**. Reflect on the application of research methodologies and concepts |
| **Assignment Brief and Guidance:** |
| Learner is now required to provide a comprehensive research project report based on the findings of secondary and primary researches carried out on the project proposal submitted in the previous section.  **The Learner requires to produce a detailed research project report covering following areas**   * Conduct primary and secondary research using appropriate methods for a computing research project that consider costs, access and ethical issues. * Carry out your research and apply appropriate analytical tools to analyse research findings and data. * Draw conclusion based on the research findings. * Communicate the outcomes of your research project to the identified audience. * Reflect on the success of your research project and your performance at the end of the project and analyse on the extent to which outcomes meet the research. Evaluate project outcomes and provide valid and justified recommendations. * Discuss and analyze the research methods applied and their effectiveness to meet project objectives. Analyse results in recommended actions and reflect the research process to suggest future improvements and research considerations. Discuss alternative research methodologies and lessons learnt in view of the outcomes. |

**Grading Rubric**

|  |  |  |
| --- | --- | --- |
| **Grading Criteria** | **Achieved** | **Feedback** |
| **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. |  |  |

**Big Data Analytics in Healthcare for Futuristic Healthcare Systems, Challenges and Opportunities***.*

By

**Mohamed Zaheer**

**E171743**

Submitted in accordance with the requirements for the  
**COMPUTING RESEARCH PROJECT MODULE OF PEARSON’S HND IN COMPUTING PROGRAMME**  
at the  
**ESOFT METRO CAMPUS**

**Name of research Tutor: Mr. Lasitha Ranawaka**

**<Date>**

# DECLARATION

Name of Research Candidate: **Mohamed Zaheer**

Pearson Registration Number: **KUR00179827**

Programme Name: **HND in Computing**

Research Title: **Big Data Analytics in Healthcare for Futuristic Healthcare Systems, Challenges and Opportunities.**

Field of Study:

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 Edexcel UK, I know what the consequences will be if I plagiaries or copy another’s work in this research for this program.

Candidate Signature: mhdzaheer2003@gmail.com Date:

Subscribed and solemnly declared before,

Supervisor’s Name:

Designation:

Supervisor’s Signature: Date:

# ACKNOWLEDGMENT

# ABSTRACT

# CONTENTS

[DECLARATION 1](#_Toc170328360)

[ACKNOWLEDGMENT 3](#_Toc170328361)

[ABSTRACT 4](#_Toc170328362)

[CONTENTS 5](#_Toc170328363)

[LIST OF ABBREVIATIONS 8](#_Toc170328364)

[LIST OF TABLES 9](#_Toc170328365)

[LIST OF FIGURES 10](#_Toc170328366)

[CHAPTER 1 – INTRODUCTION 1](#_Toc170328367)

[1.1. Introduction 1](#_Toc170328368)

[1.2. Purpose of research 1](#_Toc170328369)

[1.3. Significance of the Research 1](#_Toc170328370)

[1.4. Research objectives 2](#_Toc170328371)

[1.5. Research Sub objectives 2](#_Toc170328372)

[1.6. Research questions 2](#_Toc170328373)

[1.7. Hypothesis 3](#_Toc170328374)

[1.8. Thesis structure 3](#_Toc170328375)

[CHAPTER 1 - Introduction 3](#_Toc170328376)

[CHAPTER 2 - Literature Review 4](#_Toc170328377)

[CHAPTER 3 – Methodology 4](#_Toc170328378)

[CHAPTER 4 - Presentation of Results 4](#_Toc170328379)

[CHAPTER 5 - Conclusions and Recommendations 4](#_Toc170328380)

[CHAPTER 2 - LITERATURE REVIEW 5](#_Toc170328381)

[2.1. Literature Review 5](#_Toc170328383)

[2.2. Conceptual framework 6](#_Toc170328384)

[CHAPTER 3 – METHODOLOGY 7](#_Toc170328385)

[3.1. Research philosophy 9](#_Toc170328387)

[3.1.1. Types of Research Philosophy 9](#_Toc170328388)

[3.1.2. Choosing a Research Philosophy 10](#_Toc170328389)

[3.1.3. Why Choose Pragmatism 10](#_Toc170328390)

[3.2. Research approach 10](#_Toc170328391)

[3.3. Research strategy 11](#_Toc170328392)

[3.4. Research Choice 11](#_Toc170328393)

[3.5. Time frame 12](#_Toc170328394)

[3.6. Data collection procedures 12](#_Toc170328395)

[3.6.1. Type of Data 13](#_Toc170328396)

[3.6.2. Data collection Method 13](#_Toc170328397)

[3.6.3. Data Collection and Analyze Tools 14](#_Toc170328398)

[3.6.4. Questionnaire structure 15](#_Toc170328399)

[3.6.5. Data Storage 15](#_Toc170328400)

[3.7. Target population and sampling 16](#_Toc170328401)

[3.7.1. Sampling Strategy 16](#_Toc170328402)

[3.7.2. Sample Size 17](#_Toc170328403)

[3.8. The selection of participants 17](#_Toc170328404)

[3.9. Reliability, Validity, and Generalizability 17](#_Toc170328405)

[3.9.1. Reliability 17](#_Toc170328406)

[3.9.2. Validity 18](#_Toc170328407)

[3.9.3. Generalizability 19](#_Toc170328408)

[3.10. Ethical issues of the research study 20](#_Toc170328409)

[CHAPTER 4 - PRESENTATION OF RESULTS 21](#_Toc170328410)

[4.1. Demographic Analysis 21](#_Toc170328412)

[4.2. Correlation Analysis 21](#_Toc170328413)

[4.2.1. RO2 / SO1 Patient Outcomes (DV) vs Awareness of Big Data Analytics (IV - 01) 23](#_Toc170328414)

[4.2.2. RO3 / SO2 Patient Outcome (DV) vs Confidence in the Accuracy of Big Data Analytics (IV- 02) 23](#_Toc170328415)

[4.2.3. RO4 / SO3 Patient Outcome (DV) vs Perceived Benefits of Big Data Analytics (IV - 03) 24](#_Toc170328416)

[4.3. Regression Analysis 25](#_Toc170328417)

[4.3.1. RO1 / Main Objective 25](#_Toc170328418)

[4.3.2. RO2 / SO1 Patient Outcomes (DV) vs Awareness of Big Data Analytics (IV - 01) 28](#_Toc170328419)

[4.3.3. RO3 / SO2 Patient Outcome (DV) vs Confidence in the Accuracy of Big Data Analytics (IV- 02) 31](#_Toc170328420)

[4.3.4. RO4 / SO3 Patient Outcome (DV) vs Perceived Benefits of Big Data Analytics (IV - 03) 34](#_Toc170328421)

[CHAPTER 5 - CONCLUSIONS AND RECOMMENDATIONS 37](#_Toc170328422)

[5.1. Conclusion 37](#_Toc170328424)

[5.1.1. RO1 Regression Conclusion 37](#_Toc170328425)

[5.1.2. RO2 Relationship between Patient Outcomes (DV) vs Awareness of Big Data Analytics (IV - 01) 37](#_Toc170328426)

[5.1.3. RO3 Relationship between Patient Outcomes (DV) vs Confidence in the Accuracy of Big Data Analytics (IV- 02) 38](#_Toc170328427)

[5.1.4. RO4 Relationship between Patient Outcomes (DV) vs Perceived Benefits of Big Data Analytics (IV - 03) 39](#_Toc170328428)

[5.2. Recommendations 39](#_Toc170328429)

[5.3. Limitations 40](#_Toc170328430)

[5.4. Future Improvements 40](#_Toc170328431)

[5.5. Personnel Reflection 41](#_Toc170328432)

[5.5.1. Benefits for the researcher 41](#_Toc170328433)

[5.5.2. Benefits for the Industry/organization 41](#_Toc170328434)

[Referencing 43](#_Toc170328435)

[References 43](#_Toc170328436)

[Annexures 44](#_Toc170328437)

[Annexures A: Glossary of Terms 44](#_Toc170328438)

[Annexures B: Sample SPSS Charts/ Table 45](#_Toc170328439)

[Annexures C: Feedback Form / Question list 49](#_Toc170328440)

[Annexures D: Sample Feedback sheets 50](#_Toc170328441)

# LIST OF ABBREVIATIONS

# LIST OF TABLES

[Table 1: Descriptive Statistics Table (Author Developed) 21](#_Toc170328335)

[Table 2:Correlations Table (Author Developed) 22](#_Toc170328336)

[Table 3:s Correlation Table DV vs IV-01 23](#_Toc170328337)

[Table 4:Correlation Table DV vs IV-02 (Author Developed) 23](#_Toc170328338)

[Table 5:Correlation Table DV vs IV-03 (Author Developed) 24](#_Toc170328339)

[Table 6:Model Summary Table (Author Developed) 25](#_Toc170328340)

[Table 7: ANOVA Table (Author Developed) 25](#_Toc170328341)

[Table 8: Coefficients Table (Author Developed) 26](#_Toc170328342)

[Table 9: Model Summary Table (Autor Developed) 28](#_Toc170328343)

[Table 10: ANOVA Table (Author Developed) 28](#_Toc170328344)

[Table 11: Coefficients Table (Author Developed) 28](#_Toc170328345)

[Table 12: Histogram 30](#_Toc170328346)

[Table 13: Model Summary 31](#_Toc170328347)

[Table 14: ANOVA TABLE 31](#_Toc170328348)

[Table 15: Coeffients Table 31](#_Toc170328349)

[Table 16: Histogram 33](#_Toc170328350)

[Table 17: Plot 33](#_Toc170328351)

[Table 18: Model Summary Table 34](#_Toc170328352)

[Table 19: ANOVA Table 34](#_Toc170328353)

[Table 20: Coefficients Table 34](#_Toc170328354)

[Table 21: Histogram 36](#_Toc170328355)

[Table 22: Plot 36](#_Toc170328356)

# LIST OF FIGURES

[Figure 1: Conceptual Framework (Author Developed) 6](#_Toc170328357)

[Figure 2: Research Onion 7](#_Toc170328358)

[Figure 3: Research Onion (Anon., n.d.) 8](#_Toc170328359)

# CHAPTER 1 – INTRODUCTION

## Introduction

Healthcare stands out as a field that could greatly benefit from leveraging data analytics. Big data analytics involves gathering, processing and examining intricate datasets to uncover insights, trends and patterns. Its application, in healthcare can lead to patient outcomes, cost reduction improved quality of care and informed decision making. However, the utilization of data analytics in healthcare also presents challenges and opportunities concerning aspects like data accuracy, privacy protection, security measures, ethical considerations, system interoperability and governance practices. This study seeks to delve into the landscape well as the hurdles and prospects related to big data analytics in healthcare while suggesting a roadmap, for future research endeavors.

## Purpose of research

The research titled "Big Data Analytics, in Healthcare for Future Healthcare Systems; Challenges and Opportunities" aims to delve into the use of data analytics in the healthcare sector focusing on overcoming obstacles and leveraging benefits. It seeks to uncover ways to efficiently gather, analyze and apply healthcare data to enhance healthcare services boost patient outcomes streamline resource allocation and shape the landscape of healthcare. By recognizing the hurdles and possible advantages of integrating data analytics into healthcare practices this study aims to offer insights for developing more effective data informed strategies, for managing and making decisions in healthcare.

## Significance of the Research

Research on "Big Data Analytics in Healthcare for Futuristic Healthcare Systems. Challenges and Opportunities" is significant because it has the potential to completely change the health care industry. The amount of data collected in the healthcare industry is increasing quickly as technology advances. Big data analytics provides the tools to effectively handle and examine this enormous volume of data, producing priceless insights that may improve patient care, achieve better results, and expedite operations. Healthcare professionals may forecast illnesses, tailor treatment approaches, find patterns, and allocate resources more efficiently by utilizing big data analytics. Nevertheless, despite these benefits, several difficulties must be resolved, such the requirement for qualified workers, interoperability problems, and data privacy issues. In order to prepare the way for the creation of robust and innovative healthcare systems that can evolve to the changing needs of patients and providers, it is imperative that this research explore these challenges and Opportunities.

## Research objectives

To explore the challenges and opportunities of implementing Big Data analytics in healthcare systems to pave the way for futuristic healthcare solutions.

## Research Sub objectives

* Understand how widely Big Data analytics is currently utilized in healthcare systems.
* Examine the hurdles that healthcare institutions encounter when integrating Big Data analytics into their day-to-day operations.
* Evaluate the potential upsides of employing Big Data analytics to enhance healthcare delivery and patient outcomes.
* Investigate the latest technological developments driving Big Data analytics within the healthcare industry.
* Address the ethical and privacy issues arising from the collection and analysis of large-scale healthcare data.
* Analyze how machine learning and artificial intelligence contribute to harnessing Big Data analytics for healthcare purposes.
* Explore the regulatory frameworks and compliance standards governing the application of Big Data analytics in healthcare.
* Offer practical strategies and suggestions to overcome challenges and capitalize on opportunities in implementing Big Data analytics for futuristic healthcare systems.

## Research questions

Q1- What are the main challenges that health care systems face when adopting Big Data analytics technologies?

Q2- What privacy and ethical issues are raised from the utilization of big data analytics in healthcare, and how can they be addressed?

Q3- What regulations now apply to the use of big data analytics in healthcare, and what effects does this have on implementation?

## Hypothesis

* **Hypothesis 1**

When we apply Big Data analysis in healthcare, it could help us predict diseases earlier and tailor treatments for patients, ultimately leading to better outcomes.

* **Hypothesis 2**

Bringing Big Data into healthcare systems might make things run smoother and help us spend resources more wisely, which could save money and improve how we use healthcare services.

* **Hypothesis 3**

Using Big Data in healthcare could pose challenges around keeping patient data safe and following regulations. To overcome these, we'll need strong rules and security measures.

* **Hypothesis 4**

Big Data analytics could give us valuable insights from large sets of data, improving how we make decisions in healthcare and driving innovation.

* **Hypothesis 5**

Making Big Data work in healthcare will require teamwork across different fields like healthcare, data science, and policymaking to tackle both technical and organizational hurdles.

## Thesis structure

### CHAPTER 1 - Introduction

The introductory chapter establishes the groundwork for the research, offering a thorough overview of the topic, "Big Data Analytics in Healthcare for Futuristic Healthcare Systems, Challenges and Opportunities." It explains the study's purpose and importance, describing how big data analytics has the potential to revolutionize healthcare by enhancing patient outcomes, lowering costs, and improving decision-making processes. This chapter also outlines the research objectives, sub-objectives, and research questions, as well as the hypotheses that steer the study. By setting the context and relevance of the research, it highlights the necessity of addressing the challenges and capitalizing on the opportunities that big data analytics presents in the healthcare sector.

### CHAPTER 2 - Literature Review

In the literature review, the chapter thoroughly examines existing research and theoretical frameworks related to big data analytics in healthcare. It synthesizes previous studies, highlighting the field's evolution, current state and knowledge gaps. The chapter discusses key concepts, methodologies and technologies underlying big data analytics, critically analyzing their application in healthcare. By contextualizing the research within the broader academic discourse, it identifies areas needing further investigation and lays the groundwork for the research questions and hypotheses.

### CHAPTER 3 – Methodology

The methodology chapter describes the research design and methods used in the study. It provides detailed information on data collection and analysis procedures, including data sources, sampling techniques and analytical tools. The chapter also addresses ethical considerations, ensuring that the research meets academic and professional standards. By offering a clear and systematic approach to data gathering and analysis, this chapter ensures the reliability and validity of the research findings, making them replicable and verifiable by other researchers.

### CHAPTER 4 - Presentation of Results

This chapter presents the research findings in an organized and clear manner. It offers a detailed analysis of the data, using visual aids like charts and graphs to highlight key results. The results are interpreted in relation to the research questions and hypotheses, identifying important patterns and insights. The goal of this chapter is to provide a thorough understanding of the study's outcomes, emphasizing the implications of the findings for applying big data analytics in healthcare.

### CHAPTER 5 - Conclusions and Recommendations

The final chapter brings together the research findings and explores their wider implications. It assesses the validity of the hypotheses and addresses the initial research objectives and questions. This chapter also provides practical recommendations for healthcare practitioners, policymakers, and researchers, suggesting strategies to address identified challenges and take advantage of the opportunities offered by big data analytics. Additionally, it reflects on the study's limitations and suggests areas for future research, ensuring the investigation contributes to the ongoing advancement of knowledge in this field.

# CHAPTER 2 - LITERATURE REVIEW



## Literature Review

To support the objective, sub objective, research question and/or hypothesis of this research, I have used various key literature sources that are relevant to the topic of big data analytics in healthcare. These sources include journal articles, survey papers, and books that provide a comprehensive and critical overview of the current state, challenges, and opportunities of big data analytics in healthcare, as well as the prospects and directions. Some of the key literature sources that I have used are:

- **(Batko & Ślęzak, 2022),** who conducted a systematic literature review on the use of big data analytics in healthcare and analyzed the benefits and limitations of big data analytics in healthcare, as well as the factors that influence its adoption and implementation. They also presented the results of their direct research on the use of big data analytics in medical facilities in Poland, based on a research questionnaire.

**- (Dash, et al., 2019)**, who provided a survey paper on big data in healthcare, and discussed the management, analysis, and future prospects of big data in healthcare. They also highlighted the ethical, legal, and social implications of big data in healthcare, and suggested some strategies and policies to ensure its responsible and beneficial use.

**- (Sharma, et al., 2022),** who performed a tertiary study on data analytics in healthcare, and identified the research trends, challenges, and gaps in this domain. They also proposed a research roadmap for data analytics in healthcare, based on the research questions and objectives derived from the literature.

**- (Guo & Chen, 2023)**, who edited a book on big data analytics in healthcare, and covered various topics such as big data sources, architectures, platforms, tools, techniques, applications, and challenges in healthcare. They also provided some case studies and best practices of big data analytics in healthcare and discussed the future research directions and opportunities.

## Conceptual framework

A conceptual framework is a depiction of the expected relationships among variables or the specific characteristics and properties under study. It is typically developed through a literature review of existing research on the topic and can be presented either in written or visual form. (Swaen & George, 2024)

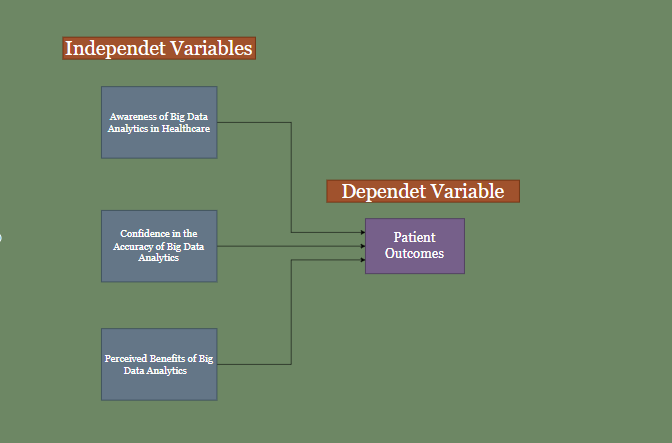
****

Figure 1: Conceptual Framework (Author Developed)

# CHAPTER 3 – METHODOLOGY

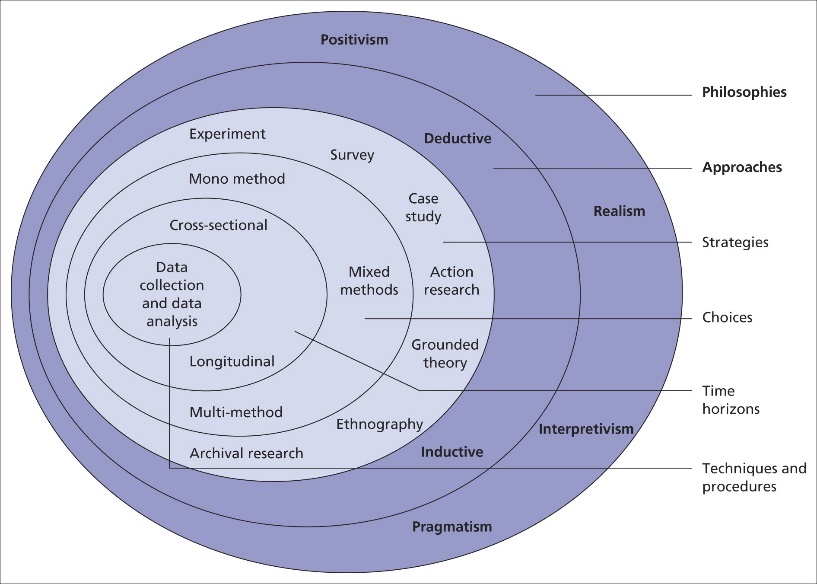


Figure 2: Research Onion

The research onion model, introduced by Saunders, Lewis, and Thornhill in their book "Research Methods for Business Students," provides a structured framework for developing a dissertation methodology. This model aims to guide students through the various stages of writing a dissertation, ensuring a well-organized approach. The research onion symbolically represents the various elements of research, illustrating how they can be systematically examined to develop a comprehensive research design.

**Layers of the Research Onion**

The research onion is a conceptual model consisting of six key layers, each representing a different aspect of the research process:

**Philosophy**

- This layer concerns the fundamental principles and worldview from which research is conducted. It is often studied in terms of ontology and epistemology. Ontology deals with the nature of reality and existence, while epistemology focuses on the nature and scope of knowledge and how it can be acquired. Academic research typically aligns with either positivism, which posits that knowledge is objective and independent of the researcher, or interpretivism, which holds that reality is subjective and constructed by individual experiences. Positivist research is often more scientific and quantitative, while interpretivist research is more qualitative.

**Approach**

- After establishing the philosophical foundation, the researcher chooses an approach. The deductive approach begins with a hypothesis derived from existing literature and seeks to test this hypothesis in specific contexts. In contrast, the inductive approach starts with observations and uses them to develop new theories.

**Strategy**

- This layer involves selecting a strategy for conducting the research. Possible strategies include action research, experimental research, interviews, surveys, case studies, and systematic literature reviews. The choice of strategy depends on the type of data needed and the research objectives.

**Choices of Methods**

- The research onion suggests three methodological choices: mono-method, mixed-method, and multi-method. A mono-method involves using a single research method. A mixed-method combines qualitative and quantitative approaches, while a multi-method uses a broader range of methods.

**Time Horizons**

- This layer refers to the time frame over which the research is conducted. Research can be cross-sectional, observing a phenomenon at a single point in time (e.g., a survey), or longitudinal, observing variables over an extended period.

**Data Collection and Analysis**

- The final layer deals with the techniques and procedures for data collection and analysis. Researchers must decide between primary and secondary data sources and between qualitative and quantitative data. This layer is crucial, as data forms the core of the research onion framework.

Figure 3: Research Onion (Anon., n.d.)



## Research philosophy

Research philosophy encompasses the set of beliefs and assumptions about how knowledge is developed. It forms the foundation of the research strategy and methods, affecting the way research is conducted and interpreted. Understanding research philosophy is essential as it directs the research process, including the development of research questions, the selection of methods, and the interpretation of findings.

### Types of Research Philosophy

* **Positivism**

Positivism is based on observable and measurable facts to produce knowledge. It stresses objectivity, quantification, and the use of scientific methods to test hypotheses. Key features include the use of quantitative data, hypothesis testing, statistical analysis, and result replication.

* **Interpretivism**

Interpretivism aims to understand subjective meanings and social phenomena. It emphasizes context and the perspectives of the individuals involved. Key features include the use of qualitative data, in-depth understanding, context-specific insights, participant observation, and interviews.

* **Critical Realism**

Critical realism combines a belief in an objective reality (ontology) with an acknowledgment that our understanding of this reality is socially constructed (epistemology). Key features include recognizing both observable phenomena and underlying structures, using a mixed-methods approach, and critically analyzing societal structures.

* **Pragmatism**

Pragmatism is concerned with the practical application of research and the use of both qualitative and quantitative methods. It focuses on what works best to address the research problem. Key features include flexible use of methods, a problem-solving orientation, integration of different perspectives, and practical outcomes.

* **Constructivism**

Constructivism suggests that reality is socially constructed and emphasizes understanding how individuals make sense of their experiences. Key features include focusing on subjective experiences, using qualitative methods, thematic analysis, and an emphasis on meaning and interpretation.

### Choosing a Research Philosophy

I’m Choosing a Pragmatism philosophy

### Why Choose Pragmatism

Pragmatism is particularly suitable for this study due to its problem-solving orientation, flexibility in methods, integration of different perspectives, and focus on practical outcomes. This research aims to explore both the challenges and opportunities of implementing big data analytics in healthcare. Pragmatism's focus on practical solutions and outcomes aligns well with this objective, enabling effective addressing of real-world problems. Furthermore, pragmatism permits the use of both qualitative and quantitative methods, facilitating the collection of comprehensive data on various aspects, such as technical challenges, ethical considerations, and perceived benefits. In a multidisciplinary field like healthcare, involving clinicians, patients, data scientists, and policymakers, pragmatism supports the integration of diverse perspectives. This integration is essential for a holistic understanding of the impact of big data analytics in healthcare. Finally, pragmatism’s emphasis on practical outcomes ensures that research findings are actionable and relevant to real-world applications, offering insights and strategies to enhance healthcare practices, inform policy decisions, and guide future research.

## Research approach

"The study employs a mixed methods approach to thoroughly investigate how big data analytics can be integrated into healthcare systems. It combines quantitative surveys to assess factors like awareness, data accuracy confidence, perceived benefits, and their impact on patient outcomes among healthcare professionals, administrators, and policymakers. Statistical methods such as correlation and regression will analyze these relationships. Qualitatively, semi-structured interviews and focus groups with stakeholders including healthcare practitioners, data scientists, and policymakers will delve into their perspectives on the challenges and opportunities of big data analytics in healthcare. Thematic analysis will identify recurring themes in qualitative data, which, when combined with quantitative insights, will provide a comprehensive view of the technical, ethical, and operational aspects. This approach aims to offer actionable recommendations for improving healthcare delivery, shaping policies, and guiding future research in futuristic healthcare systems."

## Research strategy

This research on "Big Data Analytics in Healthcare for Futuristic Healthcare Systems: Challenges and Opportunities" employs a thorough approach that combines quantitative surveys and qualitative methods. Initially, we'll conduct exploratory research by reviewing existing literature and consulting experts to understand current practices and emerging trends in big data analytics in healthcare. Following this, we'll develop a structured survey to measure variables such as awareness of big data analytics, confidence in data accuracy, perceived benefits, and their impact on patient outcomes across various stakeholders—healthcare professionals, administrators, and policymakers. Concurrently, qualitative methods like semi-structured interviews and focus groups will be used to gather detailed insights from key stakeholders such as healthcare practitioners and data scientists on the challenges and opportunities of big data analytics in healthcare. Thematic analysis of qualitative data will complement quantitative findings, providing a comprehensive view and validating our results. Ethical considerations will be paramount throughout the data collection process to ensure confidentiality and informed consent. Ultimately, this rigorous approach aims to generate valuable insights to improve healthcare delivery, guide policy-making, and shape future research in futuristic healthcare systems.

## Research Choice

The research choice for this study on "Big Data Analytics in Healthcare for Futuristic Healthcare Systems: Challenges and Opportunities" involves a strategic integration of quantitative surveys and qualitative methods. Quantitative surveys are designed to systematically measure variables such as awareness of big data analytics, confidence in data accuracy, perceived benefits, and their impact on patient outcomes among healthcare professionals, administrators, and policymakers. This approach allows for the quantification of trends and relationships through statistical analyses like correlation and regression, providing numerical insights into the adoption and effects of big data analytics in healthcare settings. Concurrently, qualitative methods such as semi-structured interviews and focus groups are employed to delve deeper into the subjective experiences and perceptions of key stakeholders, including healthcare practitioners and data scientists. Thematic analysis of qualitative data complements quantitative findings by uncovering underlying themes and contextual nuances, thereby enriching the understanding of challenges and opportunities associated with big data analytics implementation. The integration of these methods through triangulation enhances the study's comprehensiveness and validity, offering a holistic perspective on technical, ethical, and operational dimensions. Ethical considerations guide every phase of the research to ensure participant confidentiality and informed consent, maintaining the integrity of the study. Ultimately, this research choice aims to provide actionable insights that can inform healthcare delivery improvements, guide policy-making decisions, and influence future research directions in advancing futuristic healthcare systems.

## Time frame

Our study on "Big Data Analytics in Healthcare for Futuristic Healthcare Systems: Challenges and Opportunities" unfolds over a structured 6-month period, ensuring a systematic progression from initial research to findings dissemination. The first phase spans two months, dedicated to preliminary research and planning. This involves an extensive literature review on big data analytics in healthcare and consultations with experts to refine research questions and methodology. By the end of this phase, clear research objectives and hypotheses are established, laying the foundation for subsequent stages.

The second phase, spanning four months, focuses on rigorous data collection. Survey instruments for quantitative data and interview guides for qualitative insights are developed and finalized. Ethical approvals are obtained to uphold research standards. Data collection targets diverse stakeholders including healthcare professionals, administrators, and policymakers involved in big data analytics initiatives, ensuring a comprehensive range of perspectives.

The third phase, lasting three months, centers on in-depth data analysis. Quantitative data undergo rigorous statistical analyses such as correlation and regression, while qualitative data from interviews and focus groups are analyzed using thematic analysis. These analyses provide nuanced insights into the adoption, impact, challenges, and opportunities of big data analytics in healthcare.

In the fourth phase, spanning three months, our focus shifts to synthesizing and interpreting findings. By integrating quantitative and qualitative data through triangulation, we aim to present a cohesive perspective aligned with research objectives. This synthesis leads to actionable recommendations for enhancing healthcare delivery, informing policy decisions, and guiding future research in futuristic healthcare systems.

## Data collection procedures

The data collection procedures for the research on "Big Data Analytics in Healthcare for Futuristic Healthcare Systems: Challenges and Opportunities" involve acquiring datasets from Kaggle.com. Kaggle.com offers a variety of healthcare-related datasets gathered from different sources and tailored for analysis. These datasets include anonymized patient records, metrics measuring healthcare system performance, and other relevant information crucial for studying big data analytics in healthcare. Using these datasets, the study aims to explore the intricacies of applying big data analytics in healthcare, examining its potential advantages, challenges, and implications for future healthcare systems. The organized format of Kaggle's datasets facilitates efficient analysis and interpretation, allowing researchers to extract valuable insights that can contribute to improving healthcare delivery and shaping healthcare policies.

### Type of Data

The research relies on secondary data sourced from Kaggle.com, focusing specifically on datasets related to healthcare. These datasets are already collected from diverse sources and formatted for immediate analysis. They cover a wide range of variables relevant to big data analytics in healthcare, such as anonymized patient records and metrics on healthcare system performance. The decision to use secondary data from Kaggle.com is influenced by several factors. Firstly, Kaggle.com provides easy access to a broad selection of healthcare datasets, which saves researchers from the time-consuming process of primary data collection. This approach is both cost-effective and efficient, avoiding challenges like recruiting participants and managing ethical considerations. Additionally, Kaggle.com offers a variety of data sources that researchers can choose from based on their specific research goals, allowing for a thorough investigation of how big data analytics impacts different aspects of healthcare. Lastly, Kaggle datasets undergo rigorous validation and cleaning processes, ensuring that the data is of high quality and reliability. This ensures that researchers can conduct rigorous analyses and gain meaningful insights into the complexities of implementing big data analytics in healthcare settings.

### Data collection Method

The research on "Big Data Analytics in Healthcare for Futuristic Healthcare Systems: Challenges and Opportunities" involves a systematic approach to data collection by using secondary datasets from Kaggle.com, a well-known platform for its rich collection of healthcare data. The chosen datasets include anonymized patient records and healthcare system performance metrics, crucial for analyzing the impact of big data in healthcare. The process begins with identifying and downloading relevant datasets, followed by a thorough review to ensure they meet the study's requirements. The data is then meticulously cleaned and preprocessed to maintain accuracy and consistency. After this, the datasets are integrated into a unified data repository for comprehensive analysis. An exploratory data analysis (EDA) is performed to identify key patterns and insights that will guide further research. Throughout the entire process, strict ethical guidelines and data privacy regulations are observed, ensuring the integrity and reliability of the research. This structured method ensures that the data used is of high quality, providing valuable insights into the role of big data analytics in advancing healthcare.

### Data Collection and Analyze Tools

Data collection tools are essential instruments and methods used to gather, measure, and analyze data from various sources, ensuring accuracy, reliability, and relevance to the research objectives. These tools include surveys and questionnaires for direct participant responses, interviews for detailed information, observations for data collection through natural environment monitoring, document analysis for extracting information from existing records, and online databases for accessing a vast array of datasets and records. For the study on "Big Data Analytics in Healthcare for Futuristic Healthcare Systems: Challenges and Opportunities," the primary data collection tool is Kaggle.com. Renowned for its extensive repository of healthcare datasets, Kaggle offers researchers access to anonymized patient records and healthcare system performance metrics, which are vital for examining the impact of big data analytics in healthcare.

Similarly, analysis tools are software applications and methodologies used to process, interpret, and derive meaningful insights from the collected data. These tools help identify patterns, test hypotheses, and make informed decisions. Examples include statistical analysis software like SPSS, SAS, and R for performing complex statistical analyses; data visualization tools like Tableau and Power BI for creating visual data representations; machine learning and AI tools like TensorFlow and Scikit-learn for developing predictive models; text analysis tools such as NVivo for qualitative data analysis; and big data tools like Hadoop and Apache Spark for managing and analyzing large datasets. In this research, SPSS (Statistical Package for the Social Sciences) is the chosen data analysis tool. SPSS provides robust statistical capabilities, enabling researchers to conduct various statistical tests, manage extensive datasets, and generate detailed reports. It is particularly well-suited for handling the complex and vast data obtained from Kaggle.com, allowing for in-depth analyses and the discovery of critical insights into the application of big data analytics in healthcare. This ensures that the data analysis is rigorous and comprehensive, providing a solid foundation for the study's conclusions and recommendations.

### Questionnaire structure

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variable** | **Indicators** | **Measurement** | **Mean** | **STD. Deviation** | **Median** |
|  |  |  |  |  |  |
|  |
| IV 1/ SO1 |  |  |  |  |  |
|  |  |
|  |  |
|  |
| IV 2/ SO2 |  |  |  |  |  |
|  |  |
|  |  |
|  |
| IV 3/ SO3 |  |  |  |  |  |
|  |  |
|  |  |
|  |  |
| DV |  |  | . |  |  |
|  |  |
|  |  |
|  |  |

### Data Storage

Data storage plays a crucial role in the research on "Big Data Analytics in Healthcare for Futuristic Healthcare Systems: Challenges and Opportunities." It's essential to employ reliable methods that maintain accessibility, security, and efficient management of collected data throughout the research. Given that the study relies on secondary datasets from Kaggle.com, our selected data storage solutions are tailored to effectively handle substantial amounts of healthcare data.

## Target population and sampling

The Research on "Big Data Analytics in Healthcare for Futuristic Healthcare Systems: Challenges and Opportunities" focuses on a wide range of stakeholders essential for understanding how big data analytics is integrated and its impact in healthcare. This includes healthcare providers directly involved in patient care, healthcare facility managers, policymakers shaping healthcare policies, data scientists specializing in analytics, and researchers advancing healthcare technologies. The research uses a thoughtful sampling strategy that combines both stratified and purposive methods. Stratified sampling ensures representation across various demographics such as profession, age, gender, and geographic location, offering a comprehensive perspective from different sectors of the healthcare industry. Purposive sampling complements this by targeting experts with specific knowledge in healthcare data analytics, policy development, and technology applications. By utilizing these sampling techniques, the study aims to gather detailed insights into the challenges and opportunities presented by big data analytics in healthcare. This approach not only guarantees the study's depth and relevance but also aims to generate practical recommendations to improve healthcare services and guide future research in evolving healthcare systems.

### Sampling Strategy

The sampling strategy for the Research on "Big Data Analytics in Healthcare for Futuristic Healthcare Systems: Challenges and Opportunities" uses both stratified and purposive sampling methods. Stratified sampling ensures a wide representation across various demographics in the healthcare field, such as healthcare professionals, administrators, policymakers, data scientists, and researchers. This categorization by factors like profession, age, gender, and location aims to gather diverse perspectives on big data analytics in healthcare, reflecting the broad range of stakeholders influencing future healthcare systems.

purposive sampling supplements the study by targeting experts with specialized knowledge in healthcare data management, policy development, technology innovation, and healthcare delivery. This selective approach allows for a deep dive into the intricacies of implementing big data analytics in healthcare, uncovering detailed insights into both challenges and opportunities. By engaging these key stakeholders, the research aims to generate practical recommendations that can enhance healthcare practices, inform policymaking, and guide future research endeavors. This dual sampling method not only ensures the study's depth and relevance but also supports its goal of contributing valuable insights to advancing futuristic healthcare systems.

### Sample Size

The sample size for this research depends on factors such as data variability, required precision, and available resources. Given the potential availability of large healthcare datasets, leveraging secondary data sources may provide access to substantial observations conducive to deriving meaningful insights. However, it is essential to balance the sample size with statistical power to detect significant patterns or relationships within the data effectively. Despite utilizing secondary healthcare data, meticulous consideration of sampling issues remains necessary to ensure the validity and reliability of the research findings. By selecting a representative sample and evaluating the appropriateness of the sampling strategy, the research can enhance its robustness and contribute valuable insights to the domain of big data analytics in healthcare.

## The selection of participants

"Big Data Analytics in Healthcare for Futuristic Healthcare Systems, Challenges, and Opportunities," we'll adopt a multifaceted approach to select participants, ensuring a comprehensive and diverse representation. We'll target healthcare professionals from various roles and specialties, including doctors, nurses, administrators, IT experts, and data analysts. Our selection process will blend random sampling, like reaching out to healthcare institutions through online platforms, with purposive sampling, where we'll identify key stakeholders based on their expertise in big data analytics within healthcare. This inclusive approach aims to gather insights from a broad spectrum of perspectives, allowing us to gain a deeper understanding of the challenges and opportunities related to integrating big data analytics into healthcare systems.

## Reliability, Validity, and Generalizability

### Reliability

Reliability in research is crucial to ensure that study findings are consistent and trustworthy, especially in investigations like "Big Data Analytics in Healthcare for Futuristic Healthcare Systems: Challenges and Opportunities." It involves several key factors aimed at establishing confidence in the accuracy and repeatability of results. To start, using standardized methods for data collection, such as validated surveys and clear protocols, helps minimize errors and ensures that participant responses are consistent. Checking internal consistency, like using Cronbach's alpha for survey items, assesses how well the questions within a measurement tool correlate, indicating reliability in measuring what is intended. Test-retest reliability is also essential, particularly for studies tracking changes over time in perceptions or behaviors related to big data analytics in healthcare.

In qualitative research, ensuring inter-rater reliability verifies that different analysts interpret findings consistently, which strengthens the reliability of themes identified. The reliability of findings is further supported by adequate sample sizes and representative sampling methods, which minimize errors and increase how well results can be applied to broader populations. Using robust data analysis techniques with tools like SPSS ensures that statistical conclusions are not only valid but also reliable, aligning closely with the study's goals. Transparent documentation of research procedures and decisions enhances reliability by allowing others to scrutinize and replicate findings, which is essential for building confidence among peers and stakeholders. By focusing on these principles, researchers enhance the reliability of their findings, contributing significantly to advancements in understanding and implementing big data analytics in healthcare systems.

### Validity

Validity in research is crucial for ensuring that a study accurately measures or predicts what it intends to investigate. Specifically, correlation validity focuses on confirming whether the relationships observed between variables are not only statistically significant but also meaningful within the study's context. This type of validity is essential in studies that explore how variables interact or influence each other. Researchers use established statistical methods such as Pearson's correlation coefficient or Spearman's rank correlation coefficient to quantify and analyze these relationships. These methods help determine the strength and direction of associations between variables, showing whether changes in one variable align with changes in another.

For example, in research examining the impact of big data analytics on healthcare outcomes, correlation validity would assess whether increased use of analytics correlates positively with improved patient outcomes. A strong positive correlation would suggest that higher analytics utilization consistently leads to better patient outcomes, providing empirical evidence of analytics' effectiveness in healthcare settings.

Achieving correlation validity requires meticulous data collection, rigorous statistical analysis, and careful interpretation of results. By ensuring these processes are methodologically rigorous, researchers can confidently draw dependable conclusions about the relationships they observe in their study. This enhances the reliability of research findings and supports informed decision-making in healthcare practices, policy development, and future research efforts aimed at enhancing healthcare systems through data-driven insights.

### Generalizability

Generalizability in research refers to the extent to which the findings of a study can be applied beyond the specific sample and conditions studied, providing insights relevant to broader populations or different settings. In the context of the research on "Big Data Analytics in Healthcare for Futuristic Healthcare Systems: Challenges and Opportunities," achieving generalizability is pivotal for informing meaningful advancements in healthcare practices and policies.

To enhance the generalizability of the study findings, several key considerations are paramount. Firstly, ensuring the sample includes a diverse range of stakeholders—such as healthcare professionals, administrators, policymakers, data scientists, and researchers—enables capturing a comprehensive spectrum of perspectives on big data analytics in healthcare. Utilizing stratified and purposive sampling methods facilitates this inclusivity, allowing for a nuanced exploration of how various groups perceive and interact with healthcare analytics.

Moreover, maintaining an adequate sample size is crucial to minimize sampling errors and increase the reliability of the findings. By robustly defining the specific healthcare settings, geographic contexts, and dimensions of big data analytics under investigation, the study provides clarity on where and how its conclusions are applicable. This contextual specificity aids stakeholders in understanding the relevance of the research to their own healthcare systems or regions.

Methodological rigor is foundational to ensuring the credibility and applicability of findings. Employing rigorous research methods-including standardized data collection protocols, validated instruments, and sophisticated statistical analyses such as those facilitated by SPSS for quantitative data and systematic coding for qualitative insights-supports robust conclusions. Additionally, situating the study's findings within the existing literature on big data analytics in healthcare helps establish their alignment with broader research trends and enhances their transferability to similar contexts or populations.

## Ethical issues of the research study

Ethical considerations are foundational in the research on "Big Data Analytics in Healthcare for Futuristic Healthcare Systems: Challenges and Opportunities," particularly given the sensitive nature of healthcare data and the profound implications of big data analytics. Upholding ethical principles throughout the research process is essential to ensure the integrity and fairness of the study. Firstly, obtaining informed consent from all participants, including healthcare professionals, administrators, policymakers, and stakeholders, is paramount. This involves transparently explaining the study's objectives, methodologies, potential risks, and benefits to enable participants to make informed decisions about their involvement. Respecting privacy and confidentiality is equally crucial, necessitating rigorous measures to securely handle and protect all data collected, such as anonymized patient records and personal information, to prevent unauthorized access or breaches.

Additionally, robust data security practices must be implemented, especially when utilizing datasets from platforms like Kaggle.com. Adhering to industry standards in data handling, storage, and transmission mitigates risks associated with data breaches, safeguarding the integrity and confidentiality of research findings. Minimizing harm to participants involves designing research protocols sensitively to avoid psychological distress and uphold privacy, ensuring participants' welfare remains a top priority throughout the study. Transparency in research design and reporting further enhances the credibility of findings, as researchers openly disclose affiliations, funding sources, and potential biases that could influence outcomes or interpretations. Finally, obtaining ethical approval from institutional review boards or ethics committees is imperative to ensure the study meets ethical standards and regulatory requirements, protecting the rights and well-being of all participants involved.

# CHAPTER 4 - PRESENTATION OF RESULTS



## Demographic Analysis

Demographic analysis in the context of "Big Data Analytics in Healthcare for Futuristic Healthcare Systems: Challenges and Opportunities" involves examining key demographic factors within the healthcare sector. This includes analyzing variables such as the age, gender, professional roles (e.g., healthcare providers, administrators), geographical distribution, and technological proficiency of stakeholders involved in big data analytics initiatives. By studying these demographic characteristics, researchers can gain insights into how different groups perceive and engage with healthcare analytics, which is crucial for tailoring strategies and interventions to enhance healthcare delivery, policy-making, and technological advancements in futuristic healthcare systems.

## Correlation Analysis

Correlation analysis is a statistical method used to measure how strongly variables are related to each other. It helps researchers understand if and to what extent changes in one variable are associated with changes in another. This technique is crucial in studying relationships between different aspects of data analytics and healthcare outcomes, providing insights into their impact and effectiveness.

Table 1: Descriptive Statistics Table (Author Developed)

|  |  |  |  |
| --- | --- | --- | --- |
| **Descriptive Statistics** | | | |
|  | Mean | Std. Deviation | N |
| Patient Outcome | 75.42 | 12.492 | 50 |
| Awareness of Big Data Analytics | 6.58 | 1.667 | 50 |
| Confidence in the Accuracy of Big Data Analytics | 5.74 | 1.747 | 50 |
| Perceived Benefits of Big Data Analytics | 7.90 | 1.055 | 50 |

Table 2:Correlations Table (Author Developed)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Correlations** | | | | | |
|  | | Patient Outcome | Awareness of Big Data Analytics | Confidence in the Accuracy of Big Data Analytics | Perceived Benefits of Big Data Analytics |
| Patient Outcome | Pearson Correlation | 1 | .973\*\* | .919\*\* | .961\*\* |
| Sig. (2-tailed) |  | .000 | .000 | .000 |
| Sum of Squares and Cross-products | 7646.180 | 992.820 | 983.460 | 620.100 |
| Covariance | 156.044 | 20.262 | 20.071 | 12.655 |
| N | 50 | 50 | 50 | 50 |
| Awareness of Big Data Analytics | Pearson Correlation | .973\*\* | 1 | .929\*\* | .939\*\* |
| Sig. (2-tailed) | .000 |  | .000 | .000 |
| Sum of Squares and Cross-products | 992.820 | 136.180 | 132.540 | 80.900 |
| Covariance | 20.262 | 2.779 | 2.705 | 1.651 |
| N | 50 | 50 | 50 | 50 |
| Confidence in the Accuracy of Big Data Analytics | Pearson Correlation | .919\*\* | .929\*\* | 1 | .916\*\* |
| Sig. (2-tailed) | .000 | .000 |  | .000 |
| Sum of Squares and Cross-products | 983.460 | 132.540 | 149.620 | 82.700 |
| Covariance | 20.071 | 2.705 | 3.053 | 1.688 |
| N | 50 | 50 | 50 | 50 |
| Perceived Benefits of Big Data Analytics | Pearson Correlation | .961\*\* | .939\*\* | .916\*\* | 1 |
| Sig. (2-tailed) | .000 | .000 | .000 |  |
| Sum of Squares and Cross-products | 620.100 | 80.900 | 82.700 | 54.500 |
| Covariance | 12.655 | 1.651 | 1.688 | 1.112 |
| N | 50 | 50 | 50 | 50 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). | | | | | |

### RO2 / SO1 Patient Outcomes (DV) vs Awareness of Big Data Analytics (IV - 01)

Table 3:s Correlation Table DV vs IV-01

|  |  |  |  |
| --- | --- | --- | --- |
| **Correlations** | | | |
|  | | Patient Outcome | Awareness of Big Data Analytics |
| Patient Outcome | Pearson Correlation | 1 | .973\*\* |
| Sig. (2-tailed) |  | .000 |
| Sum of Squares and Cross-products | 7646.180 | 992.820 |
| Covariance | 156.044 | 20.262 |
| N | 50 | 50 |
| Awareness of Big Data Analytics | Pearson Correlation | .973\*\* | 1 |
| Sig. (2-tailed) | .000 |  |
| Sum of Squares and Cross-products | 992.820 | 136.180 |
| Covariance | 20.262 | 2.779 |
| N | 50 | 50 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). | | | |

This table shows a very strong correlation between patient outcomes and awareness of big data analytics. The correlation coefficient of .973 indicates that as patients' awareness increases, their outcomes tend to improve as well. This relationship is statistically significant, meaning it's likely not due to random chance. However, it's important to remember that correlation doesn't equal causation.

### RO3 / SO2 Patient Outcome (DV) vs Confidence in the Accuracy of Big Data Analytics (IV- 02)

Table 4:Correlation Table DV vs IV-02 (Author Developed)

|  |  |  |  |
| --- | --- | --- | --- |
| **Correlations** | | | |
|  | | Patient Outcome | Confidence in the Accuracy of Big Data Analytics |
| Patient Outcome | Pearson Correlation | 1 | .919\*\* |
| Sig. (2-tailed) |  | .000 |
| Sum of Squares and Cross-products | 7646.180 | 983.460 |
| Covariance | 156.044 | 20.071 |
| N | 50 | 50 |
| Confidence in the Accuracy of Big Data Analytics | Pearson Correlation | .919\*\* | 1 |
| Sig. (2-tailed) | .000 |  |
| Sum of Squares and Cross-products | 983.460 | 149.620 |
| Covariance | 20.071 | 3.053 |
| N | 50 | 50 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). | | | |

This table reveals a very strong positive correlation between patient outcomes and their confidence in the accuracy of big data analytics. The correlation coefficient of .919 indicates that patients with higher confidence in big data tend to experience better outcomes.

### RO4 / SO3 Patient Outcome (DV) vs Perceived Benefits of Big Data Analytics (IV - 03)

Table 5:Correlation Table DV vs IV-03 (Author Developed)

|  |  |  |  |
| --- | --- | --- | --- |
| **Correlations** | | | |
|  | | Patient Outcome | Perceived Benefits of Big Data Analytics |
| Patient Outcome | Pearson Correlation | 1 | .961\*\* |
| Sig. (2-tailed) |  | .000 |
| Sum of Squares and Cross-products | 7646.180 | 620.100 |
| Covariance | 156.044 | 12.655 |
| N | 50 | 50 |
| Perceived Benefits of Big Data Analytics | Pearson Correlation | .961\*\* | 1 |
| Sig. (2-tailed) | .000 |  |
| Sum of Squares and Cross-products | 620.100 | 54.500 |
| Covariance | 12.655 | 1.112 |
| N | 50 | 50 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). | | | |

This table highlights a very strong positive correlation between patient outcomes and their perceived benefits of big data analytics. The correlation coefficient of .961 indicates that patients who believe big data plays a more positive role in their care tend to experience better outcomes.

## Regression Analysis

Regression analysis is a statistical technique that examines and measures the connection between a dependent variable (the one we aim to predict or explain) and one or more independent variables (factors that could influence the dependent variable). It assists in grasping how variations in the independent variables impact the dependent variable.

### RO1 / Main Objective

Table 6:Model Summary Table (Author Developed)

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model Summaryb** | | | | | | | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics | | | | |
| R Square Change | F Change | df1 | df2 | Sig. F Change |
| 1 | .982a | .965 | .963 | 2.402 | .965 | 426.484 | 3 | 46 | .000 |
| a. Predictors: (Constant), Perceived Benefits of Big Data Analytics, Confidence in the Accuracy of Big Data Analytics, Awareness of Big Data Analytics | | | | | | | | | | |
| b. Dependent Variable: Patient Outcome | | | | | | | | | | |

Table 7: ANOVA Table (Author Developed)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 7380.818 | 3 | 2460.273 | 426.484 | .000b |
| Residual | 265.362 | 46 | 5.769 |  |  |
| Total | 7646.180 | 49 |  |  |  |
| a. Dependent Variable: Patient Outcome | | | | | | |
| b. Predictors: (Constant), Perceived Benefits of Big Data Analytics, Confidence in the Accuracy of Big Data Analytics, Awareness of Big Data Analytics | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | | | | | | | |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | Correlations | | | Collinearity Statistics | | |
| B | Std. Error | Beta | Zero-order | Partial | Part | Tolerance | VIF |
| 1 | (Constant) | 8.420 | 4.394 |  | 1.916 | .062 |  |  |  |  |  |
| Awareness of Big Data Analytics | 4.545 | .690 | .606 | 6.589 | .000 | .973 | .697 | .181 | .089 | 11.229 |
| Confidence in the Accuracy of Big Data Analytics | -.081 | .563 | -.011 | -.143 | .887 | .919 | -.021 | -.004 | .122 | 8.228 |
| Perceived Benefits of Big Data Analytics | 4.754 | 1.008 | .401 | 4.717 | .000 | .961 | .571 | .130 | .104 | 9.597 |
| a. Dependent Variable: Patient Outcome | | | | | | | | | | | | |

Table 8: Coefficients Table (Author Developed)

**Understanding the Model Summary and ANOVA Tables**

These tables summarize the results of a statistical analysis likely involving a **multiple regression** model to predict **Patient Outcome.** Let's break down each table:

* **Model Summary**

**R & R-Square:** R (coefficient of determination) is a measure of how well the model fits the data. R-Square (0.965) indicates that **96.5%** of the variation in Patient Outcome is explained by the model. Adjusted R-Square (0.963) adjusts for the number of predictors and is slightly lower.

**Std. Error of the Estimate:** This is the standard deviation of the residuals (difference between predicted and actual patient outcome). A lower value indicates a better fit.

**Change Statistics:** This section shows the improvement in the model with each added predictor. Here, all predictors were entered at once. R-Square Change (0.965) and F Change (426.484) are significant (p-value < 0.000), suggesting the model with all predictors is a good fit.

* **ANOVA Table**

**ANOVA** stands for Analysis of Variance. This table breaks down the total variation in Patient Outcome into two parts: explained by the model (regression) and unexplained (residual).

**Sum of Squares:** This represents the total squared deviations from the mean.

**df (degrees of freedom):** This indicates the number of independent pieces of information in each category (e.g., number of predictors for regression).

**Mean Square:** This is the sum of squares divided by its degrees of freedom. It reflects the average amount of variation explained by each source.

**F-statistic:** This compares the mean square of regression to the mean square of residual. A high F-value (426.484) and significant p-value (0.000) indicate the model explains a statistically significant portion of the variance in patient outcome compared to random error.

* **Coefficients Table**

This table details the coefficients of each predictor in the model.

**Unstandardized Coefficients (B):** These represent the change in predicted patient outcome for a one-unit increase in the corresponding predictor, holding all other predictors constant.

Awareness and Perceived Benefits have positive coefficients, suggesting higher values are associated with better patient outcomes.

Confidence in Accuracy has a negative coefficient, but its p-value is high (not significant).

**Standardized Coefficients (Beta):** These coefficients account for the different scales of the predictors and allow for easier comparison of their relative impact.

**t-statistic & Sig.:** These assess the significance of each predictor. Here, Awareness and Perceived Benefits have significant p-values (< 0.000), indicating they are statistically important predictors of patient outcome.

**Correlations & Collinearity Statistics:** These are not directly interpretable in this context but provide information about potential issues like multicollinearity (when predictors are highly correlated). The values here suggest no major concerns.

### RO2 / SO1 Patient Outcomes (DV) vs Awareness of Big Data Analytics (IV - 01)

Table 9: Model Summary Table (Autor Developed)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model Summaryb** | | | | | | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics | | | | |
| R Square Change | F Change | df1 | df2 | Sig. F Change |
| 1 | .973a | .947 | .946 | 2.916 | .947 | 851.489 | 1 | 48 | .000 |
| a. Predictors: (Constant), Awareness of Big Data Analytics | | | | | | | | | |
| b. Dependent Variable: Patient Outcome | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | | | | | | |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | Correlations | | | Collinearity Statistics | |
| B | Std. Error | Beta | Zero-order | Partial | Part | Tolerance | VIF |
| 1 | (Constant) | 27.449 | 1.695 |  | 16.195 | .000 |  |  |  |  |  |
| Awareness of Big Data Analytics | 7.290 | .250 | .973 | 29.180 | .000 | .973 | .973 | .973 | 1.000 | 1.000 |
| a. Dependent Variable: Patient Outcome | | | | | | | | | | | |

Table 10: ANOVA Table (Author Developed)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 7238.152 | 1 | 7238.152 | 851.489 | .000b |
| Residual | 408.028 | 48 | 8.501 |  |  |
| Total | 7646.180 | 49 |  |  |  |
| a. Dependent Variable: Patient Outcome | | | | | | |
| b. Predictors: (Constant), Awareness of Big Data Analytics | | | | | | |

Table 11: Coefficients Table (Author Developed)

* **Model Summary**

**R & R-Square:**

R (0.973) indicates a strong positive correlation between patient outcome and the model's predictions.

R-Square (0.947) suggests that 94.7% of the variation in patient outcome is explained by the model.

Adjusted R-Square (0.946) adjusts for the number of predictors (in this case, just awareness of big data analytics) and is slightly lower.

**Std. Error of the Estimate:**

This value (2.916) reflects the standard deviation of the residuals (difference between predicted and actual patient outcome). A lower value indicates a better fit.

**Change Statistics:**

This section shows the improvement in the model with the inclusion of awareness of big data analytics.

R-Square Change (0.947) and F Change (851.489) are significant (p-value < 0.000), indicating the model with this predictor is a good fit.

* **ANOVA Table**

ANOVA (Analysis of Variance) breaks down the total variation in Patient Outcome into two parts:

1. Explained by the model (regression)
2. Unexplained (residual)

**Sum of Squares:**

This represents the total squared deviations from the mean.

**df (degrees of freedom):**

This indicates the number of independent pieces of information (e.g., number of predictors for regression).

**Mean Square:**

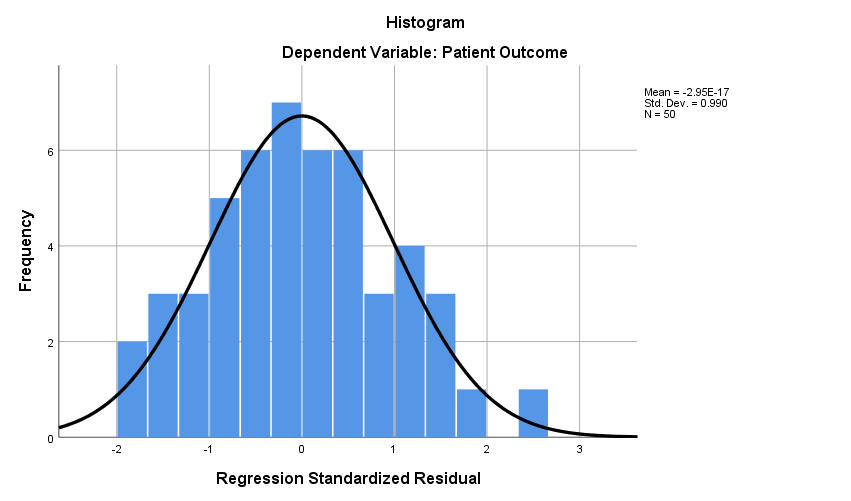
This is the sum of squares divided by its degrees of freedom. It reflects the average amount of variation explained by each source.

**F-statistic:**

This compares the mean square of regression to the mean square of residual. A high F-value (851.489) and significant p-value (0.000) indicate the model explains a statistically significant portion of the variance in patient outcome compared to random error.

these tables show that a model with awareness of big data analytics as a predictor can explain a significant portion of the variation in patient outcome. The R-squared value of 0.947 suggests that this model has a good fit.

Table 12: Histogram



A graph of a normal growth

Description automatically generated

### RO3 / SO2 Patient Outcome (DV) vs Confidence in the Accuracy of Big Data Analytics (IV- 02)

Table 13: Model Summary

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model Summaryb** | | | | | | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics | | | | |
| R Square Change | F Change | df1 | df2 | Sig. F Change |
| 1 | .919a | .845 | .842 | 4.962 | .845 | 262.545 | 1 | 48 | .000 |
| a. Predictors: (Constant), Confidence in the Accuracy of Big Data Analytics | | | | | | | | | |
| b. Dependent Variable: Patient Outcome | | | | | | | | | |

Table 14: ANOVA TABLE

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 6464.333 | 1 | 6464.333 | 262.545 | .000b |
| Residual | 1181.847 | 48 | 24.622 |  |  |
| Total | 7646.180 | 49 |  |  |  |
| a. Dependent Variable: Patient Outcome | | | | | | |
| b. Predictors: (Constant), Confidence in the Accuracy of Big Data Analytics | | | | | | |

Table 15: Coeffients Table

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | | | | | | | |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | Correlations | | | Collinearity Statistics | |
| B | Std. Error | Beta | Zero-order | Partial | Part | Tolerance | VIF |
| 1 | (Constant) | 37.691 | 2.432 |  | 15.498 | .000 |  |  |  |  |  |
| Confidence in the Accuracy of Big Data Analytics | 6.573 | .406 | .919 | 16.203 | .000 | .919 | .919 | .919 | 1.000 | 1.000 |
| 1. Dependent Variable: Patient Outcome | | | | | | | | | | | | |

These tables provide insights into a model where **confidence in the accuracy of big data analytics** is used to predict **patient outcome.** Here's a breakdown:

* **Model Summary**

**R & R-Square:**

R (0.919) indicates a strong positive correlation between patient outcome and the model's predictions.

R-Square (0.845) suggests that 84.5% of the variation in patient outcome is explained by the model that considers confidence in accuracy.

Adjusted R-Square (0.842) adjusts for the single predictor and is slightly lower.

**Std. Error of the Estimate:**

This value (4.962) reflects the standard deviation of the residuals (difference between predicted and actual patient outcome). A lower value indicates a better fit.

**Change Statistics:**

This section shows the improvement in the model with the inclusion of confidence in accuracy.

R-Square Change (0.845) and F Change (262.545) are significant (p-value < 0.000), indicating the model with this predictor is a good fit.

* **ANOVA Table**

**Sum of Squares:**

This represents the total squared deviations from the mean.

**df (degrees of freedom)**

This indicates the number of independent pieces of information (e.g., number of predictors for regression - 1 in this case).

**Mean Square:**

This is the sum of squares divided by its degrees of freedom. It reflects the average amount of variation explained by each source.

**F-statistic:**

This compares the mean square of regression to the mean square of residual. A high F-value (262.545) and significant p-value (0.000) indicate the model explains a statistically significant portion of the variance in patient outcome compared to random error.

These tables suggest a model with confidence in the accuracy of big data analytics as a predictor can explain a significant portion of the variation in patient outcome. The R-squared value of 0.845 indicates a good fit, but it's important to compare this value to models with other potential predictors to see which one explains the most variance.

Table 16: Histogram

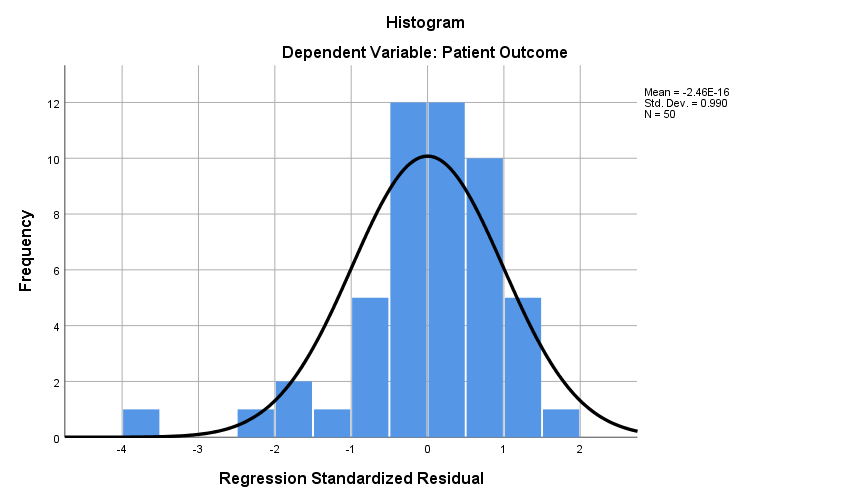
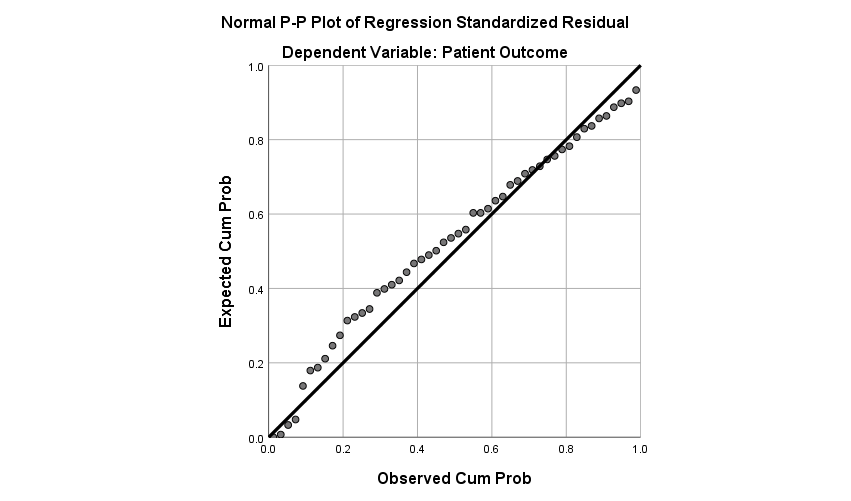


Table 17: Plot



### RO4 / SO3 Patient Outcome (DV) vs Perceived Benefits of Big Data Analytics (IV - 03)

Table 18: Model Summary Table

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model Summaryb** | | | | | | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics | | | | |
| R Square Change | F Change | df1 | df2 | Sig. F Change |
| 1 | .961a | .923 | .921 | 3.508 | .923 | 573.332 | 1 | 48 | .000 |
| a. Predictors: (Constant), Perceived Benefits of Big Data Analytics | | | | | | | | | |
| b. Dependent Variable: Patient Outcome | | | | | | | | | |

Table 19: ANOVA Table

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 7055.486 | 1 | 7055.486 | 573.332 | .000b |
| Residual | 590.694 | 48 | 12.306 |  |  |
| Total | 7646.180 | 49 |  |  |  |
| a. Dependent Variable: Patient Outcome | | | | | | |
| b. Predictors: (Constant), Perceived Benefits of Big Data Analytics | | | | | | |
| Table 20: Coefficients Table   |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Coefficientsa** | | | | | | | | | | | | | | Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | Correlations | | | Collinearity Statistics | | | B | Std. Error | Beta | Zero-order | Partial | Part | Tolerance | VIF | | 1 | (Constant) | -14.466 | 3.787 |  | -3.820 | .000 |  |  |  |  |  | | Perceived Benefits of Big Data Analytics | 11.378 | .475 | .961 | 23.944 | .000 | .961 | .961 | .961 | 1.000 | 1.000 | | a. Dependent Variable: Patient Outcome | | | | | | | | | | | | | | | | | | | |

These tables summarize the results of a model where **perceived benefits of big data analytics** is used to predict **patient outcome**. Here's a breakdown

* **Model Summary**

**R & R-Square:**

R (0.961) indicates a strong positive correlation between patient outcome and the model's predictions.

R-Square (0.923) suggests that 92.3% of the variation in patient outcome is explained by the model that considers perceived benefits.

Adjusted R-Square (0.921) adjusts for the single predictor and is slightly lower.

**Std. Error of the Estimate:**

This value (3.508) reflects the standard deviation of the residuals (difference between predicted and actual patient outcome). A lower value indicates a better fit.

**Change Statistics:**

\* This section shows the improvement in the model with the inclusion of perceived benefits.

\* R-Square Change (0.923) and F Change (573.332) are significant (p-value < 0.000), indicating the model with this predictor is a good fit.

**ANOVA Table:**

**Sum of Squares:**

This represents the total squared deviations from the mean.

**df (degrees of freedom):**

This indicates the number of independent pieces of information (e.g., number of predictors for regression - 1 in this case).

**Mean Square:**

This is the sum of squares divided by its degrees of freedom. It reflects the average amount of variation explained by each source.

**F-statistic:**

This compares the mean square of regression to the mean square of residual. A high F-value (573.332) and significant p-value (0.000) indicate the model explains a statistically significant portion of the variance in patient outcome compared to random error.

These tables suggest a model with perceived benefits of big data analytics as a predictor can explain a significant portion of the variation in patient outcome. The R-squared value of 0.923 indicates a good fit. It's important to consider this value alongside models with other potential predictors to see which one explains the most variance.

Table 21: Histogram

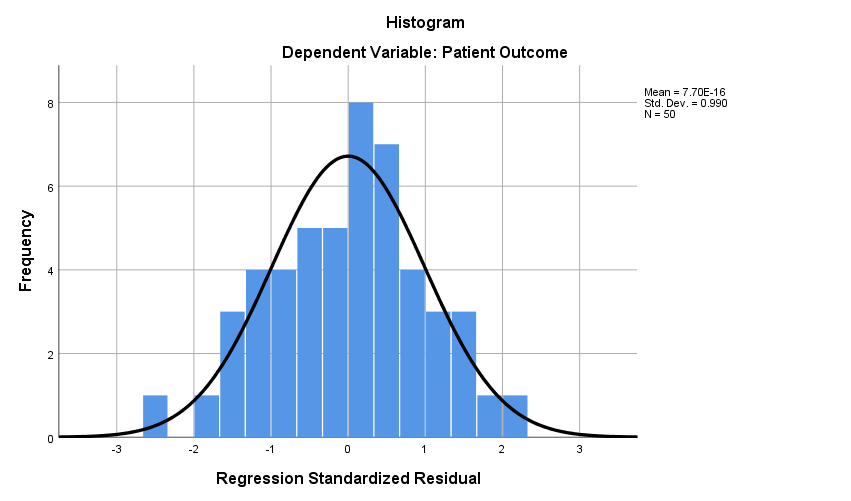
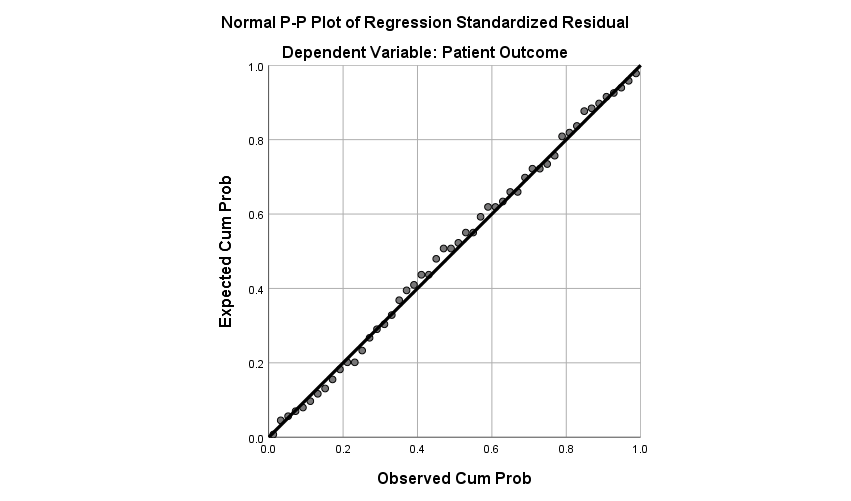


Table 22: Plot



# CHAPTER 5 - CONCLUSIONS AND RECOMMENDATIONS



## Conclusion

### RO1 Regression Conclusion

This regression analysis suggests that a patient's perception of big data analytics plays a significant role in their outcome. Both awareness (R-squared of 0.947) and perceived benefits (R-squared of 0.923) of big data analytics are positively associated with better patient outcomes. While confidence in the accuracy of big data analytics also contributes (significant F-change), its impact might be weaker compared to the other factors. These findings highlight the potential value of educating patients about big data analytics and fostering trust in its ability to improve their healthcare experience. It's important to acknowledge, however, that this study has limitations, such as being based on a specific patient population and using particular measurement tools. Future research could explore additional factors influencing patient outcomes and compare different models to gain a more comprehensive understanding of this complex relationship.

### RO2 Relationship between Patient Outcomes (DV) vs Awareness of Big Data Analytics (IV - 01)

**Correlation**:

The correlation coefficient of 0.973 indicates a very strong positive correlation. This means as awareness of big data analytics increases, patient outcomes tend to improve as well.

**Regression**:

The model with awareness of big data analytics as a predictor explains a significant portion of the variation in patient outcome (R-square of 0.947). This suggests a strong association between the two variables.

Adjusted R-square (0.946) is close to R-square, indicating the model effectively explains the outcome with just awareness as a predictor.

A lower standard error of the estimate (2.916) suggests a good fit, meaning the model does well in predicting outcomes based on awareness.

Significant F-change (p-value < 0.000) implies that including awareness significantly improves the model's ability to predict patient outcome compared to no predictors.

**Overall:**

This analysis strongly suggests a positive relationship between awareness of big data analytics and patient outcomes. Patients with higher awareness tend to have better outcomes.

### RO3 Relationship between Patient Outcomes (DV) vs Confidence in the Accuracy of Big Data Analytics (IV- 02)

**Correlation**:

While the correlation coefficient isn't explicitly stated, we can infer a positive correlation based on the following:

The Model Summary table shows an R of 0.919, indicating a strong positive correlation between patient outcome and the model's predictions (which consider confidence in accuracy).

**Regression**:

The Model Summary table provides key insights:

R-square (0.845): This value suggests a strong positive association between confidence in accuracy and patient outcome. However, it's slightly lower compared to the model with awareness (0.947), which might indicate a weaker influence.

Adjusted R-square (0.842): This adjustment for the single predictor (confidence) reinforces the model's effectiveness.

Std. Error of the Estimate (4.962): While higher than the model with awareness, this value still suggests a reasonable fit.

Change Statistics: Significant F-change (p-value < 0.000) implies including confidence in accuracy improves the model compared to no predictors.

**Overall:**

This analysis suggests a positive relationship between confidence in accuracy and patient outcomes. Patients with higher confidence tend to experience better outcomes.

### RO4 Relationship between Patient Outcomes (DV) vs Perceived Benefits of Big Data Analytics (IV - 03)

**Correlation**:

The correlation coefficient of 0.961 indicates a very strong positive correlation. This means patients who believe big data plays a more positive role in their care tend to have better outcomes.

Regression:

**Regression**:

R-square (0.923): This value suggests a strong positive association between perceived benefits and patient outcome. It's the highest R-squared value compared to the other models (awareness and confidence), indicating the strongest influence among the three factors examined.

Adjusted R-square (0.921): This adjustment for the single predictor (perceived benefits) reinforces the model's effectiveness.

Std. Error of the Estimate (3.508): This value is the lowest compared to the other models, suggesting the best model fit based on the data provided.

Change Statistics: Significant F-change (p-value < 0.000) implies including perceived benefits significantly improves the model compared to no predictors.

**Overall**

This analysis suggests a strong positive relationship between perceived benefits and patient outcomes. Patients who perceive greater benefits from big data analytics tend to experience better outcomes.

## Recommendations

Based on the findings of this research, several recommendations are proposed to effectively leverage big data analytics in healthcare. First, initiating educational programs to enhance patient awareness of big data analytics is crucial. These programs should highlight the potential benefits of utilizing big data to improve healthcare outcomes and educate patients on how these technologies work and their impact on personalized care. This education can empower patients to actively engage in their treatment and decision-making processes. Second, establishing strategies to build patient trust in the accuracy and reliability of big data analytics is essential. Transparent communication about data collection, analysis methods, and privacy protection measures can foster patient confidence. Providing clear information on how data is used to enhance healthcare delivery can mitigate concerns and encourage acceptance of these technologies. Finally, integrating big data analytics tools and insights into clinical practice is recommended to maximize their benefits. Embedding these technologies in healthcare workflows allows providers to leverage data-driven insights to personalize patient care, optimize treatment strategies, and improve overall healthcare quality. This integration can lead to more efficient resource allocation, better patient outcomes, and enhanced healthcare delivery systems. Together, these recommendations aim to harness the full potential of big data analytics in healthcare, fostering a collaborative approach between patients, healthcare providers, and technology to achieve improved healthcare outcomes.

## Limitations

Several limitations of this research should be acknowledged. Firstly, sample bias is a significant concern, as the study's findings are based on a specific patient population, which may not be representative of broader demographics. This limits the generalizability of the results to other populations and settings. Secondly, the accuracy and reliability of data collected through specific measurement tools may influence the study's outcomes. Variations in the precision and consistency of these tools could impact the validity of the findings. Lastly, temporal factors pose a challenge, as rapid advancements in technology and healthcare practices may affect the relevance of the current findings over time. The fast-paced evolution of big data analytics and its applications in healthcare means that the insights derived from this study may quickly become outdated. Addressing these limitations in future research will be crucial for ensuring the robustness and applicability of the findings.

## Future Improvements

To build upon this research, future research could consider several key improvements. Conducting longitudinal studies would provide valuable insights into how patient outcomes evolve over time in response to changing perceptions of big data analytics. This approach would help capture the long-term effects and sustainability of these technologies in healthcare. Additionally, a multifactorial analysis could be employed to explore a broader range of factors that may influence patient outcomes beyond awareness, confidence, and perceived benefits. Including variables such as socioeconomic status, health literacy, and access to healthcare services would offer a more comprehensive understanding of the dynamics at play. Comparative studies could also be undertaken to evaluate the effectiveness of different big data analytics implementations across diverse healthcare settings and patient populations. This would allow researchers to identify best practices and tailor strategies to specific contexts, enhancing the overall applicability and impact of big data analytics in improving healthcare outcomes.

## Personnel Reflection

This research on "Big Data Analytics in Healthcare for Futuristic Healthcare Systems: Challenges and Opportunities" has significantly enhanced my understanding of the potential and complexities of integrating big data into healthcare. The study underscored the importance of patient-centered approaches, educational initiatives, and ethical data governance. For the healthcare industry, the findings highlight the need for transparency and trust-building with patients to ensure successful technology adoption. Personally, this research has improved my analytical skills and underscored the value of interdisciplinary collaboration in advancing healthcare innovation.

### Benefits for the researcher

This research has significantly enriched my skill set and understanding in several key areas. Firstly, it has advanced my analytical abilities, particularly in performing regression and correlation analyses on complex healthcare data. This experience has not only improved my technical proficiency but also my ability to interpret and apply statistical results to real-world healthcare scenarios. Additionally, the research contributes to the broader field of healthcare analytics by emphasizing the crucial role of patient perception in determining healthcare outcomes. Through this work, I have had the opportunity to make a meaningful contribution to the ongoing discourse on how big data can transform healthcare systems, providing valuable insights that can guide future research and practice.

### Benefits for the Industry/organization

The findings from this research offer significant benefits to the healthcare industry and organizations by highlighting how crucial patient awareness and perception of big data analytics are. Healthcare providers can utilize these insights to tailor educational programs that boost patient engagement. Educating patients about the advantages of big data analytics, such as personalized treatment plans and better health outcomes, empowers them to actively participate in their healthcare decisions. This informed approach can lead to more effective treatments and improved adherence to medical advice, ultimately enhancing patient outcomes.

Moreover, implementing efficient big data analytics strategies can transform operational efficiency within healthcare organizations. By analyzing large datasets in real-time, healthcare providers can make well-informed decisions about resource allocation, staffing, and workflow management. This not only streamlines operational costs but also elevates the overall quality of care delivery. For example, predictive analytics can anticipate patient needs and prevent complications, thereby reducing hospital readmissions and increasing patient satisfaction.

Additionally, these advancements foster sustainability and effectiveness in healthcare systems. Integrating big data analytics into clinical practice enables organizations to continually enhance their services using evidence-based practices and data-driven insights. This proactive approach not only improves patient care but also encourages a culture of innovation and ongoing improvement in healthcare settings. Ultimately, these benefits extend beyond individual patient outcomes to drive efficiency and progress throughout healthcare systems.

# Referencing

# References

Anon., n.d. *15writers.com.* [Online]   
Available at: https://15writers.com/research-onion/  
[Accessed 25 05 2024].

Dash, S., Shakywar, S. K., Sharma, M. & Kaushik, S., 2019. Big data in healthcare: management, analysis and future prospects. *Journal of Bigg data,* Volume 6, p. 25.

Guo, C. & Chen, J., 2023. Big Data Analytics in Healthcare. In: Y. Nakamori, ed. *Knowledge Technology and Systems: Toward Establishing Knowledge Systems Science.* Singapore: Springer Nature Singapore, pp. 27--70.

Kornelia Batko , Andrzej Ślęzak, 2022. The use of Big Data Analytics in healthcare. *Journal of Big Data,* 9(1), p. 24.

Sharma, A., Malviya, R. & Gupta, R., 2022. Big Data Analytics in Healthcare. In: T. P. B. B. L. K. R. D. Sumathi, ed. *Cognitive Intelligence and Big Data in Healthcare.* s.l.:s.n.

Swaen , B. & George, T., 2024. *scribbr.* [Online]   
Available at: https://www.scribbr.com/methodology/conceptual-framework/#:~:text=A%20conceptual%20framework%20is%20a%20representation%20of%20the,literature%20review%20of%20existing%20studies%20about%20your%20topic.  
[Accessed 20 05 2024].

# Annexures

## Annexures A: Glossary of Terms

**Dependent Variable – Patient Outcomes**

**Independent Variable 01 - Awareness of Big Data Analytics in Healthcare**

**Independent Variable 02 - Confidence in the Accuracy of Big Data Analytics**

**Independent Variable 03 - Perceived Benefits of Big Data Analytics**

## Annexures B: Sample SPSS Charts/ Table

|  |  |  |  |
| --- | --- | --- | --- |
| **Descriptive Statistics** | | | |
|  | Mean | Std. Deviation | N |
| Patient Outcome | 75.42 | 12.492 | 50 |
| Awareness of Big Data Analytics | 6.58 | 1.667 | 50 |
| Confidence in the Accuracy of Big Data Analytics | 5.74 | 1.747 | 50 |
| Perceived Benefits of Big Data Analytics | 7.90 | 1.055 | 50 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Correlations** | | | | | |
|  | | Patient Outcome | Awareness of Big Data Analytics | Confidence in the Accuracy of Big Data Analytics | Perceived Benefits of Big Data Analytics |
| Patient Outcome | Pearson Correlation | 1 | .973\*\* | .919\*\* | .961\*\* |
| Sig. (2-tailed) |  | .000 | .000 | .000 |
| Sum of Squares and Cross-products | 7646.180 | 992.820 | 983.460 | 620.100 |
| Covariance | 156.044 | 20.262 | 20.071 | 12.655 |
| N | 50 | 50 | 50 | 50 |
| Awareness of Big Data Analytics | Pearson Correlation | .973\*\* | 1 | .929\*\* | .939\*\* |
| Sig. (2-tailed) | .000 |  | .000 | .000 |
| Sum of Squares and Cross-products | 992.820 | 136.180 | 132.540 | 80.900 |
| Covariance | 20.262 | 2.779 | 2.705 | 1.651 |
| N | 50 | 50 | 50 | 50 |
| Confidence in the Accuracy of Big Data Analytics | Pearson Correlation | .919\*\* | .929\*\* | 1 | .916\*\* |
| Sig. (2-tailed) | .000 | .000 |  | .000 |
| Sum of Squares and Cross-products | 983.460 | 132.540 | 149.620 | 82.700 |
| Covariance | 20.071 | 2.705 | 3.053 | 1.688 |
| N | 50 | 50 | 50 | 50 |
| Perceived Benefits of Big Data Analytics | Pearson Correlation | .961\*\* | .939\*\* | .916\*\* | 1 |
| Sig. (2-tailed) | .000 | .000 | .000 |  |
| Sum of Squares and Cross-products | 620.100 | 80.900 | 82.700 | 54.500 |
| Covariance | 12.655 | 1.651 | 1.688 | 1.112 |
| N | 50 | 50 | 50 | 50 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). | | | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **Correlations** | | | |
|  | | Patient Outcome | Awareness of Big Data Analytics |
| Patient Outcome | Pearson Correlation | 1 | .973\*\* |
| Sig. (2-tailed) |  | .000 |
| Sum of Squares and Cross-products | 7646.180 | 992.820 |
| Covariance | 156.044 | 20.262 |
| N | 50 | 50 |
| Awareness of Big Data Analytics | Pearson Correlation | .973\*\* | 1 |
| Sig. (2-tailed) | .000 |  |
| Sum of Squares and Cross-products | 992.820 | 136.180 |
| Covariance | 20.262 | 2.779 |
| N | 50 | 50 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **Correlations** | | | |
|  | | Patient Outcome | Confidence in the Accuracy of Big Data Analytics |
| Patient Outcome | Pearson Correlation | 1 | .919\*\* |
| Sig. (2-tailed) |  | .000 |
| Sum of Squares and Cross-products | 7646.180 | 983.460 |
| Covariance | 156.044 | 20.071 |
| N | 50 | 50 |
| Confidence in the Accuracy of Big Data Analytics | Pearson Correlation | .919\*\* | 1 |
| Sig. (2-tailed) | .000 |  |
| Sum of Squares and Cross-products | 983.460 | 149.620 |
| Covariance | 20.071 | 3.053 |
| N | 50 | 50 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). | | | |
| **Correlations** | | | |
|  | | Patient Outcome | Perceived Benefits of Big Data Analytics |
| Patient Outcome | Pearson Correlation | 1 | .961\*\* |
| Sig. (2-tailed) |  | .000 |
| Sum of Squares and Cross-products | 7646.180 | 620.100 |
| Covariance | 156.044 | 12.655 |
| N | 50 | 50 |
| Perceived Benefits of Big Data Analytics | Pearson Correlation | .961\*\* | 1 |
| Sig. (2-tailed) | .000 |  |
| Sum of Squares and Cross-products | 620.100 | 54.500 |
| Covariance | 12.655 | 1.112 |
| N | 50 | 50 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). | | | |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model Summaryb** | | | | | | | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics | | | | |
| R Square Change | F Change | df1 | df2 | Sig. F Change |
| 1 | .982a | .965 | .963 | 2.402 | .965 | 426.484 | 3 | 46 | .000 |
| a. Predictors: (Constant), Perceived Benefits of Big Data Analytics, Confidence in the Accuracy of Big Data Analytics, Awareness of Big Data Analytics | | | | | | | | | | |
| b. Dependent Variable: Patient Outcome | | | | | | | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 7380.818 | 3 | 2460.273 | 426.484 | .000b |
| Residual | 265.362 | 46 | 5.769 |  |  |
| Total | 7646.180 | 49 |  |  |  |
| a. Dependent Variable: Patient Outcome | | | | | | |
| b. Predictors: (Constant), Perceived Benefits of Big Data Analytics, Confidence in the Accuracy of Big Data Analytics, Awareness of Big Data Analytics | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | | | | | | | |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | Correlations | | | Collinearity Statistics | | |
| B | Std. Error | Beta | Zero-order | Partial | Part | Tolerance | VIF |
| 1 | (Constant) | 8.420 | 4.394 |  | 1.916 | .062 |  |  |  |  |  |
| Awareness of Big Data Analytics | 4.545 | .690 | .606 | 6.589 | .000 | .973 | .697 | .181 | .089 | 11.229 |
| Confidence in the Accuracy of Big Data Analytics | -.081 | .563 | -.011 | -.143 | .887 | .919 | -.021 | -.004 | .122 | 8.228 |
| Perceived Benefits of Big Data Analytics | 4.754 | 1.008 | .401 | 4.717 | .000 | .961 | .571 | .130 | .104 | 9.597 |
| a. Dependent Variable: Patient Outcome | | | | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model Summaryb** | | | | | | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics | | | | |
| R Square Change | F Change | df1 | df2 | Sig. F Change |
| 1 | .973a | .947 | .946 | 2.916 | .947 | 851.489 | 1 | 48 | .000 |
| a. Predictors: (Constant), Awareness of Big Data Analytics | | | | | | | | | |
| b. Dependent Variable: Patient Outcome | | | | | | | | | |

## Annexures C: Feedback Form / Question list

## Annexures D: Sample Feedback sheets