**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** | Research Proposal – 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** | 01 | **Assessor** | Mr. Lasitha Ranawaka |
| **Submission Date** | 23/03/2024 | **Date Received 1st submission** |  |
| **Re-submission Date** |  | **Date Received 2nd submission** |  |
| **Assessor Feedback:**   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | **LO1 Examine appropriate research methodologies and approaches as part of the research process** | | | | | | | | **Pass, Merit & Distinction Descripts** | **P1** | **P2** | **M1** | **D1** |  |  | |  | | | | | | | |  |  |  |  |  |  |  | |  | | | | | | | |  |  |  |  |  |  |  | |  | | | | | | | |  |  |  |  |  |  |  | | | | |
| **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)

Research Project Proposal

**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:23/03/2024**

**(*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 Proposal -Big Data** |
| Issue Date | 22/01/2024 |
| Submission Date | 23/03/2024 |
| IV Name & Date |  |
| **Submission Format:** | |
| **Research Project Proposal**   * 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. * Reference using the Harvard referencing system. * Please provide a referencing list using the Harvard referencing system. * The recommended word limit is minimum 2000 words. | |
| **Unit Learning Outcomes:** | |
| **LO1**. **Examine appropriate research methodologies and approaches as part of the research process**. | |
| **Assignment Brief and Guidance:** | |
| **Big Data**  Big data is a term that has become more and more common over the last decade. It was originally defined as data that is generated in incredibly large volumes, such as internet search queries, data from weather sensors or information posted on social media. Today big data has also come to represent large amounts of information generated from multiple sources that cannot be processed in a conventional way and that cannot be processed by humans without some form of computational intervention.  Big data can be stored in several ways: Structured, whereby the data is organised into some form of relational format, unstructured, where data is held as raw, unorganised data prior to turning into a structured form, or semi-structured where the data will have some key definitions or structural form but is still held in a format that does not conform to standard data storage models.  Many systems and organisations now generate massive quantities of big data on a daily basis, with some of this data being made publicly available to other systems for analysis and processing. The generation of such large amounts of data has necessitated the development of machine learning systems that can sift through the data to rapidly identify patterns, to answer questions or to solve problems. As these new systems continue to be developed and refined, a new discipline of data science analytics has evolved to help design, build and test these new machine learning and artificial intelligence systems.  Utilising Big Data requires a range of knowledge and skills across a broad spectrum of areas and consequently opens opportunities to organisations that were not previously accessible. The ability to store and process large quantities of data from multiple sources has meant that organisations and businesses are able to get a larger overall picture of the pattern of global trends in the data to allow them to make more accurate and up to date decisions. Such data can be used to identify potential business risks earlier and to make sure that costs are minimised without compromising on innovation.  However, the rapid application and use of Big Data has raised several concerns. The storage of such large amounts of data means that security concerns need to be addressed in case the data is compromised or altered in such a way to make the interpretation erroneous. In addition, the ethical issues of the storage of personal data from multiple sources have yet to be addressed, as well as any sustainability concerns in the energy requirements of large data warehouses and lakes.  The theme will enable students to explore some of the topics concerned with Big Data from the standpoint of a prospective computing professional or data scientist. It will provide the opportunity for students to investigate the applications, benefits and limitations of Big Data while exploring the responsibilities and solutions to the problems it is being used to solve.  **Choosing a research objective/question**  Students are to choose their own research topic for this unit. Strong research projects are those with clear, well focused and defined objectives. A central skill in selecting a research objective is the ability to select a suitable and focused research objective. One of the best ways to do this is to put it in the form of a question. Students should be encouraged by tutors to discuss a variety of topics related to the theme to generate ideas for a good research objective.  The range of topics discussed on Big Data, could cover the following areas:   * Storage models * Cyber security risks * Future developments and driving innovation. * Legal and ethical trade-offs   **Project Proposal should cover following areas.**   1. Definition of research problem or question. (This can be stated as a research question, objectives, or hypothesis) 2. Provide a literature review giving the background and conceptualisation of the proposed area of study. (This would provide existing knowledge and benchmarks by which the data can be judged) 3. Examine and critically evaluate research methodologies and research processes available. Select the most suitable methodologies and the process and justify your choice based on theoretical/philosophical frameworks. Demonstrate understanding of the pitfalls and limitations of the methods chosen and ethical issues that might arise. 4. Draw points (1–3, above) together into a research proposal by getting agreement with your tutor. | |

**Useful links**

Useful resources for underlying principles, examples of articles and webinars on the theme:

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

**Grading Rubric**

|  |  |  |
| --- | --- | --- |
| **Grading Criteria** | **Achieved** | **Feedback** |
| **P1** Produce a research proposal that clearly defines a research question or hypothesis, supported by a literature review. |  |  |
| **P2** Examine appropriate research methods and approaches to primary and secondary research. |  |  |
| **M1** Analyse different research approaches and methodology and make justifications for the choice of methods selected based on philosophical/theoretical frameworks. |  |  |
| **D1** Critically evaluate research methodologies and processes in application to a computing research project to justify chosen research methods and analysis |  |  |

Research Proposal Form

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Student Name** | | Mohamed Zaheer | | | | | | | | |
| **Student number** | | E171743 | | | | **Date** | | 2024/03/22 | | |
| **Centre Name** | | Kurunegala | | | | | | | | |
| **Unit** | | Unit 16: Computing Research Project (Pearson Set) | | | | | | | | |
| **Tutor** | |  | | | | | | | | |
| **Proposed title** | | | | | | | | | | |
| **“Bigg Data Analytics in Healthcare for Futuristic Healthcare Systems”. Challenges and Opportunities.** | | | | | | | | | | |
|  | | | | | | | | | | |
| **Section One: Title, objective, responsibilities** | | | | | | | | | | |
| *Title or working title of research project (in the form of a question, objective or hypothesis): Research project objectives (e.g. what is the question you want to answer? What do you want to learn how to do? What do you want to find out?): Introduction, Objective, Sub Objective(s), Research Questions and/or Hypothesis* | | | | | | | | | | |
|  | | | | | | | | | | |
|  | | | | | | | | | | |
| **Section Two: Reasons for choosing this research project** | | | | | | | | | | |
| *Reasons for choosing the project (e.g. links to other subjects you are studying, personal interest, future plans, knowledge/skills you want to improve, why the topic is important): Motivation, Research gap* | | | | | | | | | | |
|  | | | | | | | | | | |
|  | | | | | | | | | | |
| **Section Three: Literature sources searched** | | | | | | | | | | |
| *Use of key literature sources to support your objective, Sub Objective, research question and/or hypothesis: Can include the Conceptual Framework* | | | | | | | | | | |
|  | | | | | | | | | | |
|  | | | | | | | | | | |
| **Section Four: Activities and timescales** | | | | | | | | | | |
| *Activities to be carried out during the research project (e.g. research, development, analysis of ideas, writing, data collection, numerical analysis, tutor meetings, production of final outcome, evaluation, writing the report) and How long this will take:* | | | | | | | | | | |
| **Milestone** | | | | | | | **Propose completion date** | | | |
|  | | | | | | |  | | | |
|  | | | | | | |  | | | |
|  | | | | | | |  | | | |
|  | | | | | | |  | | | |
|  | | | | | | | | | | |
| **Section Five: Research approach and methodologies** | | | | | | | | | | |
| *Type of research approach and methodologies you are likely to use, and reasons for your choice: What your areas of research will cover: Research Onion; Sample Strategy/Method; Sample Size* | | | | | | | | | | |
|  | | | | | | | | | | |
|  | | | | | | | | | | |
|  | | | | | | | | | | |
| **Comments and agreement from tutor** | | | | | | | | | | |
| Comments (optional): | | |  | | | | | | | |
| I confirm that the project is not work which has been or will be submitted for another qualification and is appropriate. | | | | | | | | | | |
| Agreed | Yes  No | | | Name |  | | | | Date |  |
|  | | | | | | | | | | |
| **Comments and agreement from project proposal checker (if applicable)** | | | | | | | | | | |
| Comments (optional): | | |  | | | | | | | |
| I confirm that the project is appropriate. | | | | | | | | | | |
| Agreed | Yes  No | | | Name |  | | | | Date |  |

#### Research Ethics Approval Form

All students conducting research activity that involves human participants or the use of data collected from human participants are required to gain ethical approval before commencing their research. Please answer all relevant questions and note that your form may be returned if incomplete.

|  |  |  |  |
| --- | --- | --- | --- |
| **Section 1: Basic Details** | | | |
| **Project title:** | | **“Big Data Analytics in Healthcare for Futuristic Healthcare Systems”. Challenges and Opportunities.** | |
| **Student name:** | | M.W.M.M ZAHEER | |
| **Student ID number:** | | E171743 | |
| **Programme:** | | BTEC HND in Computing | |
| **School:** | | ESOFT Metro Campus | |
| **Intended research start date:** | | |  |
| **Intended research end date:** | | |  |
| **Section 2: Project Summary** | | | |
| *Please select all research methods that you plan to use as part of your project*   * Interviews: * Questionnaires: * Observations: * Use of Personal Records: * Data Analysis: * Action Research: * Focus Groups: * Other (please specify): | | | |
| **Section 3: Participants** | | | |
| *Please answer the following questions, giving full details where necessary.*  Will your research involve human participants?  Who are the participants? Tick all that apply:  Age 12-16  Young People aged 17–18  Adults  How will participants be recruited (identified and approached)?  Describe the processes you will use to inform participants about what you are doing:  **Studies involving questionnaires:**  Will participants be given the option of omitting questions they do not wish to answer?  Yes  No  If **“NO”** please explain why below and ensure that you cover any ethical issues arising from this.  **Studies involving observation:**  Confirm whether participants will be asked for their informed consent to be observed.  Yes  No  Will you debrief participants at the end of their participation (i.e. give them a brief explanation of the study)?  Yes  No  Will participants be given information about the findings of your study? (This could be a brief summary of your findings in general)  Yes  No | | | |
| **Section 4: Data Storage and Security** | | | |
| Confirm that all personal data will be stored and processed in compliance with the Data Protection Act (1998)  Yes  No  Who will have access to the data and personal information?  **During the research:**  Where will the data be stored?  Will mobile devices such as USB storage and laptops be used?  Yes  No  If **“YES”**, please provide further details:  **After the research:**  Where will the data be stored?  How long will the data and records be kept for and in what format?  Will data be kept for use by other researchers?  Yes  No  If **“YES”**, please provide further details: | | | |
| **Section 5: Ethical Issues** | | | |
| *Are there any particular features of your proposed work which may raise ethical concerns? If so, please outline how you will deal with these:* | | | |
| **Section 6: Declaration** | | | |
| I have read, understood and will abide by the institution’s Research and Ethics Policy:  Yes  No  I have discussed the ethical issues relating to my research with my Unit Tutor:  Yes  No  **I confirm that to the best of my knowledge:**  The above information is correct and that this is a full description of the ethics issues that may arise in the course of my research. | | | |
| Name: | M.W.M.M ZAHEER | | |
| Date: | 23/03/2024 | | |
| **Please submit your completed form to: ESOFT Learning Management System (ELMS)** | | | |

**THE RESEARCH PROPOSAL**

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

By

**M.W.M.M Zaheer**

E171743

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

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

# ACKNOWLEDGMENT

# EXECUTIVE SUMMARY

# CONTENTS

[ACKNOWLEDGMENT i](#_Toc162045876)

[EXECUTIVE SUMMARY ii](#_Toc162045877)

[CONTENTS iii](#_Toc162045878)

[INTRODUCTION 1](#_Toc162045879)

[1.1. Introduction 1](#_Toc162045880)

[1.2. Purpose of research 1](#_Toc162045881)

[1.3. Significance of the Research 1](#_Toc162045882)

[1.4. Research objectives 2](#_Toc162045883)

[1.5. Research Sub objectives. 2](#_Toc162045884)

[1.6. Research questions. 3](#_Toc162045885)

[1.7. Hypothesis 3](#_Toc162045886)

[LITERATURE REVIEW 4](#_Toc162045887)

[2.1. Literature Review 4](#_Toc162045889)

[2.2. Conceptual framework 4](#_Toc162045890)

[METHODOLOGY 5](#_Toc162045891)

[3.1. Research philosophy. 5](#_Toc162045893)

[3.2. Research approach 5](#_Toc162045894)

[3.3. Research strategy 6](#_Toc162045895)

[3.4. Research Choice 6](#_Toc162045896)

[3.5. Time frame 6](#_Toc162045897)

[3.6. Data collection procedures 8](#_Toc162045898)

[3.6.1. Type of Data 8](#_Toc162045908)

[3.6.2. Data Collection Method 10](#_Toc162045909)

[3.6.3. Data Collection and Analyze Tools 10](#_Toc162045910)

[3.7. Sampling 11](#_Toc162045911)

[3.7.1. Sampling Strategy 11](#_Toc162045913)

[3.7.2. Sample Size 11](#_Toc162045914)

[3.8. The selection of participants 12](#_Toc162045915)

[References 14](#_Toc162045916)

# 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.

# LITERATURE REVIEW



## Literature Review

To bolster the objectives, sub-objectives, research inquiries, and/or hypotheses of this study, several influential literature sources pertinent to big data analytics in healthcare have been incorporated. These sources encompass academic articles, review papers, and scholarly books, providing a thorough and critical analysis of the present landscape, challenges, and potentials of big data analytics in healthcare, along with projections for the future. Key literature sources utilized include.

- (Batko & Ślęzak, 2022) conducted a systematic review of literature on big data analytics in healthcare, analyzing its advantages and constraints, along with the factors influencing its adoption and execution. They also shared findings from their firsthand research on big data analytics utilization in medical facilities in Poland, based on a research questionnaire.

- (Dash, et al., 2019) presented a survey paper on big data in healthcare, discussing its administration, analysis, and future prospects. They also highlighted ethical, legal, and societal implications, proposing strategies and policies for responsible and beneficial utilization.

- (Sharma, et al., 2022) performed a tertiary study on data analytics in healthcare, identifying research patterns, challenges, and gaps, and suggesting a research roadmap based on derived research inquiries and objectives from the literature.

- (Guo & Chen, 2023) edited a comprehensive book on big data analytics in healthcare, covering topics like sources, architectures, platforms, tools, techniques, applications, and challenges. They also provided case studies and best practices, along with discussions on future research paths and opportunities.

## Conceptual framework

The conceptual framework for future research on big data analytics in healthcare revolves around four main pillars: data, analytics, outcomes, and challenges. The data component encompasses various healthcare data types, focusing on aspects like quality, privacy, and interoperability. Analytics involves methods and tools for processing and analyzing data, considering factors affecting their adoption. Outcomes explore the benefits, limitations, and ethical considerations of applying big data analytics in healthcare. Challenges encompass obstacles and opportunities, including technical, organizational and ethical concerns. The framework emphasizes the interconnected nature of these components, such as the data-analytics-outcomes (DAO) cycle for transforming data into insights and feedback loops for continuous improvement. This framework provides a structured approach for future research, aiding in the identification of research questions, objectives and methodologies for investigating big data analytics in healthcare.

# METHODOLOGY



## Research philosophy.

The research philosophy chosen for this study plays a critical role in shaping the approach taken to understand big data analytics in healthcare for futuristic healthcare systems. By adopting a positivist perspective, the focus lies on conducting objective analyses of data to uncover underlying patterns and relationships within healthcare datasets. Alternatively, embracing a pragmatist approach provides flexibility, allowing for the integration of both quantitative and qualitative methods to comprehensively address the research objectives. This philosophical stance ensures a balanced exploration of the complexities surrounding big data analytics in healthcare.

## Research approach

A mixed-methods approach is deemed most appropriate for investigating "Big Data Analytics in Healthcare for Futuristic Healthcare Systems." This methodological choice seamlessly integrates quantitative data analysis, such as statistical assessments of healthcare datasets, with qualitative data collection methods, such as interviews or case studies with healthcare professionals. Through this comprehensive approach, a thorough exploration of the challenges and opportunities associated with implementing big data analytics in healthcare can be achieved, offering valuable insights into the subject matter.

## Research strategy

The research strategy employed in this study encompasses both deductive and inductive reasoning methodologies. Deductive reasoning involves the systematic testing of hypotheses derived from existing theories or frameworks related to big data analytics in healthcare. In contrast, inductive reasoning entails the collection of data and the identification of patterns to develop new theories or insights about the topic. By combining these deductive and inductive approaches, a robust exploration of the research questions is ensured, allowing for a nuanced understanding of the complexities inherent in big data analytics in healthcare.

## Research Choice

The chosen research methodology leans towards a descriptive or exploratory study. A descriptive study offers a detailed overview of the current state of big data analytics in healthcare, shedding light on the challenges and opportunities it presents within futuristic healthcare systems. Conversely, an exploratory study seeks to identify new perspectives or areas of inquiry within the field, potentially leading to the development of innovative solutions or approaches. This deliberate research choice enables a comprehensive examination of big data analytics in healthcare, contributing to the advancement of knowledge in this rapidly evolving domain.

## Time frame

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Task & Time | Jan | | | | Feb | | | | Mar | | | | Apr | | | |
| W1 | W2 | W3 | W4 | W1 | W2 | W3 | W4 | W1 | W2 | W3 | W4 | W1 | W2 | W3 | W4 |
| Finalize Research Proposal |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Literature Review |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Define Research Questions |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Develop Data Collection Tools |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Contact Potential Participants |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Begin Participant Requirements |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Data Collection |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Collected quantitative data |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Continue data collection |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Analize qualitative data |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Combine qualitative & quantitative |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Write the research report |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Review & finalize the report |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Submit the final report |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

## Data collection procedures

The process of collecting data involves various methods and techniques used to gather relevant information for the study. In this research, we utilize structured surveys, interviews with experts in the healthcare field, and the analysis of healthcare data. These approaches allow us to obtain comprehensive datasets that offer insights into the practical challenges and opportunities of applying big data analytics in shaping the future of healthcare systems.



### Type of Data

For this research I got the dataset from [Kaggle.com](https://www.kaggle.com/datasets/prasad22/healthcare-dataset)

**Column Names Explanati**ons

* **Name**

The "Name" column in the dataset likely contains the names of the patients. It's of string type, allowing for the storage of textual data, and serves as a key identifier for individuals within healthcare systems.

* **Age**

The "Age" column contains the patient's age in years, typically stored as an integer. It's crucial for understanding demographics, assessing health risks, and tailoring treatments.

* **Gender**

The "Gender" column indicates the gender of each patient, typically categorized as 'Male', 'Female', or 'Other'. This data is essential for tailoring healthcare interventions to specific gender-related needs and understanding disparities in healthcare outcomes.

* **Blood** **Type**

The "Blood Type" column contains the blood type of each patient, usually denoted as 'A', 'B', 'AB', or 'O'. This information is crucial for medical procedures like blood transfusions and helps anticipate potential complications.

* **Medical Condition**

The "Medical Condition" column describes the diagnosed health condition of each patient, aiding in treatment planning and monitoring.

* **Date of Admission**

The "Date of Admission" column records the date when patients were admitted to the hospital. Storing this information as datetime data allows for precise tracking of admissions, optimizing resource allocation and facilitating timely interventions.

* **Doctor**

The "Doctor" column contains the name or identifier of the attending physician responsible for the patient's care. This information facilitates communication among healthcare providers and ensures continuity of care during the patient's hospital stay.

* **Hospital**

The "Hospital" column contains the name or identifier of the healthcare facility where each patient was admitted. This information is vital for tracking patient encounters and coordinating care across different healthcare settings.

* **Insurance provider**

The "Insurance Provider" column contains the name or identifier of the patient's insurance company. This information is crucial for managing billing, reimbursement, and tailoring treatment plans based on coverage.

* **Billing Amount**

The "Billing Amount" column contains the total cost billed for each patient's treatment. This numerical data, stored as either float or integer types, quantifies expenses and facilitates financial management for healthcare institutions and patients**.**

* **Room Number**

The "Room Number" column specifies the room where each patient was accommodated during their hospital stay. It can be represented as a string or integer, depending on the room numbering system. This information is crucial for logistical organization and efficient patient management within the healthcare facility.

* **Admission Type**

The "Admission Type" column categorizes how patients were admitted to the hospital, such as 'Emergency', 'Elective', or 'Urgent'. Stored as strings, this information aids in prioritizing care, allocating resources, and optimizing operational processes within the healthcare facility.

* **Discharge Date**

The "Discharge Date" column records the date when patients were discharged from the hospital. Storing this information as datetime data allows for accurate tracking of hospital stays, facilitating continuity of care and resource planning.

* **Medication**

The "Medication" column contains information about the medications prescribed to patients during their hospital stay. Stored as strings, this data is vital for documenting treatment plans and ensuring medication safety and adherence.

* **Test Results**

The "Test Results" column contains the outcomes of medical tests conducted on patients during their hospitalization. Stored as strings, this data is crucial for diagnosis, treatment monitoring, and communication among healthcare professionals.

### Data Collection Method

The selected sources for data collection include Kaggle.com, Google Dataset Search, and Google Scholar. Kaggle.com provides access to a diverse range of datasets, including those relevant to healthcare, facilitating the retrieval of pertinent data for analysis. Additionally, Google Dataset Search offers a comprehensive platform for discovering datasets sourced from research institutions and government agencies, further enriching the data collection process. Google Scholar serves as a valuable resource for accessing academic papers, research articles, and scholarly publications related to big data analytics in healthcare, aiding in understanding the current state of research and identifying relevant studies and datasets.

### Data Collection and Analyze Tools

For data collection and analysis, the chosen tools are Python and Google Colab. Python, equipped with libraries such as pandas, NumPy, and scikit-learn, provides versatility and efficiency in data collection, preprocessing, analysis, and machine learning tasks. Google Collab offers a convenient cloud-based environment for running Python code, allowing researchers to collaborate, share, and execute data analysis workflows seamlessly. These tools collectively empower researchers to effectively collect and analyze data for their study on big data analytics in healthcare for futuristic healthcare systems, addressing challenges and exploring opportunities in the field.

## Sampling

"Big Data Analytics in Healthcare for Futuristic Healthcare Systems. Challenges and Opportunities," sampling remains pivotal in ensuring the reliability and validity of the study's findings. Given the focus on healthcare data and analytics, the sampling process aims to select a subset of existing datasets that accurately represents the population of interest.



### Sampling Strategy

The sampling strategy involves assessing available healthcare datasets to ensure they adequately capture the relevant population characteristics aligned with the research objectives. While researchers may not directly control the sampling process for existing datasets, they can evaluate the representativeness and suitability of the data for addressing the research questions. Depending on the nature of the healthcare data, different sampling approaches may be utilized. For instance, if the datasets encompass a broad spectrum of healthcare information, a random sampling method may be appropriate to select records without introducing bias. Alternatively, if the datasets are segmented into various healthcare domains or patient demographics, a stratified sampling approach ensures adequate representation of each subgroup.

### 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," the process of selecting participants entails identifying pertinent entities and data points within healthcare datasets that correspond to the study's objectives. Since the research relies predominantly on secondary data sources, the term "participants" refers to entities within these datasets rather than individual human subjects.

The participants in this research encompass a range of entities such as healthcare organizations, patient records, medical procedures, diagnostic codes, treatment outcomes, and other healthcare-related data elements. These selections are made in accordance with the research goals, which seek to investigate the challenges and prospects of leveraging big data analytics in shaping the future of healthcare systems.

When choosing participants from healthcare datasets, it's crucial to ensure their relevance to the research inquiries and the broader study scope. For example, patient records containing demographic details, medical histories, and treatment procedures may be chosen to analyze patterns in healthcare delivery and patient results. the participant selection process should reflect the diverse and intricate nature of healthcare systems, considering variables like the types of healthcare facilities, specialized areas, geographic locations, and the demographics of the patient population served.

By meticulously selecting participants aligned with the research objectives and offering valuable insights into the challenges and opportunities of big data analytics in healthcare, the study can produce meaningful outcomes and contribute to advancements in healthcare analytics and the design of futuristic healthcare systems.

# References

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.