



**STATE UNIVERSITY
OF BANGLADESH**

Comparative Study of Data Analysis Tools and Technologies

M. Asif Hossain

A Thesis Proposal for the
Project Thesis Work

CSE-0400

Department of Computer Science and Engineering
4th December 2021

**Submitted To
Muhammad Masud Tarek**

Contents

1	Introduction	2
1.1	Research Topic	2
2	Advantages of My Research	3
3	Features of My Research	4
4	Problem Statements	5
5	Motivation of My Research	6
6	Proposed Methodology	7
6.1	Sources of Literature	7
6.2	Literature Types and Time Period	8
6.3	Research Process	8
7	Evaluation of Effectiveness of My Research	9
8	Importance of My Research	10
9	Scopes and Limitations of My Research	11
9.1	The scopes of this research	11
9.2	The limitations of this research	11
10	Estimated Budget and Time Duration	12
10.1	Estimate Budget	12
10.2	Time Duration	12

Chapter 1

Introduction

Data analysis is defined as a process of cleaning, transforming, and modeling data to discover useful information for business decision-making. The purpose of Data Analysis is to extract useful information from data and taking the decision based upon the data analysis.

Data analysis is important in business to understand problems facing an organisation, and to explore data in meaningful ways. Data in itself is merely facts and figures. Data analysis organises, interprets, structures and presents the data into useful information that provides context for the data.

In this paper, we will analyze, testing and find out the performance of the data analysis tools and technologies. Finally, we will get some answers of limitation, the effectiveness of the following technologies.

1.1 Research Topic

The final selected topic of this research is **“Comparative Study of Data Analysis Tools and Technologies.”**

Chapter 2

Advantages of My Research

This research will increase the Critical thinking of this author and this research will help to build up analytical skills through hands-on learning. And other advantages are:

- Defining academic, career, and personal interests.
- Expanding knowledge and understanding of a chosen field outside of the classroom.
- Developing one-on-one connections with distinguished faculty in their field.
- Building community with peers, faculty, and organizations on- and off campus.

Chapter 3

Features of My Research

The objective of the problem will be studied properly. The author will reach the conclusion with reference and the long-term studies. This research will be completed in the form of effective and analytic information.

Chapter 4

Problem Statements

- RQ1: In which cases which Quantitative and Qualitative data analysis tools and technologies preferable?
- RQ2: Are there any monopolies of data analysis tools or all most effective data analysis tools and technologies are open sources?
- RQ3: What are the sufficiency and deprivation of data analysis tools and technologies?

Chapter 5

Motivation of My Research

There are numerous tools and technologies for data analysis. All tools are not suitable for all work. Different tools have different roles in data analysis. Our study is to find out those differences, roles, effectiveness.

Chapter 6

Proposed Methodology

The research paper will be published after completing the three survey exercises and analyzing all of the data from those surveys by using various types of Data Analysis Tools and Technologies. Therefore we may be able to find out the acceptability, power, and limitations of the Data Analysis Tools and Technologies.

6.1 Sources of Literature

The research method for this study consisted of a search of available multifocal literature for the topics of the research. The following sources for literature and information search will be used:

1. IEEE Xplore Digital Library: To finding correct information, direct access to the IEEE Xplore interface will be used.
2. ACM Digital Library: Same as IEEE Xplore, to find out accurate information, direct access to the ACM Digital Library interface.
3. Google Scholar Search: If the two aforementioned failed to give results, Google Scholar Will be able to find unofficial links/scans/photographs of documents of interest.
4. The State University of Bangladesh Library and moodle.sub.edu.bd: SUB Library has many research papers and books which will help in this thesis and moodle.sub.edu.bd has all content for the all courses

such as PDF books, previous semesters works, course guidelines, and all courses videos.

5. Google Internet Search: Google is the fastest source to find some bitsof initial information on some specific topic.

6.2 Literature Types and Time Period

The main source of literature was from scientific publisher organizations such as IEEE and ACM. However, to broaden the sources of knowledge, information from practice (e.g. programming libraries, manufacturer's references) was included in the search process. Using multivocal literature.

Within the subject of the study, the time period studied consisted of the literature from the 1960s to the present year 2020.

6.3 Research Process

With a population of a minimum of a hundred population, this research will conduct with one real-world survey. Take advantage of the online survey by using the Google Form Two hundred population survey will also be conducted. For collecting big data, will be using Intel Extreme Tuning Utility log files. For calculating and analysis purposes the author will not use any dummy data in this research paper.

Chapter 7

Evaluation of Effectiveness of My Research

The main source of research will come from scientific publisher organizations such as IEEE and ACM. However, to increase the sources of knowledge, information from practice (e.g. programming libraries, manufacturer's references) will be included in the research paper. Data analysis different tools have different roles. We will find out by examining how those tools are effective from the different perspectives of the data analysis.

Chapter 8

Importance of My Research

Through this research, we will find answers to some of the significant questions. As well as In which cases Quantitative and Qualitative data analysis tools and technologies are preferable and the sufficiency and deprivation of data analysis tools and technologies. We will find out those answers by proper examination and study. The final result will help the data analytics and it will build a different narrative for the data analysis tools and technologies.

Chapter 9

Scopes and Limitations of My Research

9.1 The scopes of this research

- Understanding effective data analysis process.
- Results will help to develop more effective data analysis tools.
- Implement the output to the real-world data analysis fields.
- Start realizing the benefits of data analytics.

9.2 The limitations of this research

- Some output may bring complex and bias results.
- Data collected from different sources can vary in quality and format.
- Some tools may create issues with hardware and lab limitations may occur.

Chapter 10

Estimated Budget and Time Duration

10.1 Estimate Budget

- Survey Cost: Estimated Budget For completing the survey of 100 population estimate cost per head 50 BDT. Which is equivalent to 5,000 BDT.
- Conference Paper Publication: Cost may vary on the different platforms. So this cost is UNDEFINED.

10.2 Time Duration

Starting 1st June 2021

Ending 1st June 2022