

MALAWI UNIVERSITY OF SCIENCE AND TECHNOLOGY
NDATA SCHOOL OF CLIMATE AND EARTH SCIENCE
DEPARTMENT OF EARTH SCIENCES

DEVELOPMENT OF THE SURVEY SYSTEM

BRIGHT NKHOMA
GIS/020/20
THYOLO, MALAWI

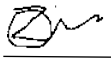
SUPERVISOR DR. CHISENGA

DISSERTATION SUBMITTED TO NDATA SCHOOL OF CLIMATE AND EARTHSCIENCES OF THE MALAWI UNIVERSITY OF SCIENCE AND TECHNOLOGY IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE BACHELOR OF SCIENCE DEGREE IN EARTH SCIENCES

DATE: 29TH APRIL, 2025

PROPOSAL APPROVED BY:

SUPERVISOR:

DR. C. CHISENGA Signature: 

ASSESSOR:
Title, Initials, Name Signature: _____

_____

Disclaimer

This document describes work undertaken as part of a programme of study at Ndata School of Climate and Earth Sciences of the Malawi University of Science and Technology. All views and opinions expressed therein remain the sole responsibility of the author(s), and do not necessarily represent those of the School.

COPYRIGHT STATEMENT

This copy of the proposal has been supplied on condition that anyone who consults it is understood to recognize that its copyright rests with its author and that no quotation from the dissertation and no information derived from it may be published without acknowledgement.

Acknowledgements

First and foremost, I would like to thank **God Almighty** for granting me strength, perseverance, and wisdom throughout the course of this project.

My sincere gratitude goes to my supervisor, **Dr. Chisenga**, for his invaluable guidance, insightful feedback, and constant encouragement. Your mentorship shaped this dissertation and helped me achieve a deeper understanding of both technical and scientific principles.

I also extend my heartfelt thanks to the **Malawi University of Science and Technology**, particularly the **Ndata School of Climate and Earth Sciences**, for providing me with the academic environment and resources necessary for the successful completion of this research.

Special thanks go to the **MUST Research Department** for their collaboration and support during the user testing phase. Your practical insights and constructive feedback greatly enriched the development of the MobileSurvey system.

I am grateful to my fellow students, colleagues, and friends for their motivation, assistance, and encouragement, especially during challenging stages of this project.

Lastly, I dedicate this work to my family for their unwavering support, patience, and prayers throughout this academic journey. Your belief in me has been my biggest source of inspiration.

Thank you all.

TABLE OF CONTENTS

CHAPTER 1 : Introduction 1
1.1 Background 1
1.2 Problem statement 2
1.3 Main and specific objectives 2
1.3.1 Main Objective 2
1.3.2 Specific Objectives 3
1.4 Research questions 3
1.5 Scientific Significance 3
1.5.1. Improved Scientific Understanding 3
1.5.2. Technical Contribution 5
1.5.3. Potential Use and Users 5
CHAPTER 2: Literature review 7
2.1 Currently used surveys 7
2.2. Challenges in Current Traditional Survey Systems 9
2.2.1 Missing Collaborative Features 9
2.2.2 Limited Data Interpretation and Visualization 9
2.2.3 Absence of Geo-Spatial Integration 9
2.2.4 Expensive to Collect Data 10
2.3 PROPOSED SYSTEM CHANGES 10
2.3.1 Mobile app 11
2.3.2 The website 12
2.4 USER REQUIREMENTS DEFINITION 14
2.4.1 Functional Requirements 14
2.4.2 Non Functional Requirements 14
CHAPTER 3: System Models and Architecture 16
3.1 system Overview 16
3.1.1 system architecture 16
CHAPTER 4: System requirements 18
4.1 Functional requirements 18
4.1.1 Authentication 18
4.1.2 Data analysis 18
4.1.3 Secure transactions 18
4.1.4 Survey creation and management 19
4.1.5 Data Collection and Responses 19
4.1.6 Collaboration and Access Control 19
4.1.7 Geo-Spatial Mapping and GIS Integration 20
CHAPTER 5: Methodology 21
5.1. Research Design 21
5.2 Threats to Internal and External Validity 21
5.2.1 Internal Validity 21
5.2.2 External validity 22
5.3 Variables and Conceptual Framework 24
5.3.1 Variables 24
5.3.2 Conceptual Framework 26
5.4 Target users 30

A. Researcher and Data Analysts 30
B. Field data Collectors 30
C. General Public and Business 30
D. Students 30
5.5 Core Features 31
A. Real-time Data Collection and Collaboration 31
B. Data Visualization and Geo-statistics: 31
C. generative AI for Statistical Insights 31
D. Secure Data Storage and Access 31
5.6 Materials 32
5.6.1 Next.js 32
5.6.2 FastAPI (Python Framework) 32
5.6.3 Express.js (Node.js Framework) 32
5.6.4 Google Firestore 33
5.6.5 Google Cloud Storage 33
5.6.6 Firebase Authentication 33
5.6.7 Mobile Devices (Smartphones/Tablets) 33
5.6.8 Laptops/Desktops and Smartphones 33
5.6.9 Internet Connectivity 33
5.7 Methods 34
5.7. 1 Problem Analysis 35
5.7.2 Requirements gathering 35
5.7.3 Research design 38
5.7.4 Implementation 47
5.3.5 User Testing 55
5.3.6 Data Collection 55
5.3.7 Data analysis 55
5.3.8 User Feedback 55
5.3.9 System Refinement 55
CHAPTER 6: Results and Discussion 56
6.1 Website 56
6.1.1 Questionnaire creation 56
6.1.2 Data Analysis (Dashboard) 58
6.1.3 Data analysis: Visualization 59
6.1.4 Data analysis: Hypothesis Testing (Inferential statistics) 59
6.1.5 Data analysis: Confidence interval (Inferential statistics) 60
6.1.6 Data analysis: Correlation/Regression (Inferential Statistics) 61
6.1.7 Data analysis: IDW (Geo-spatial analysis) 62
6.1.8 Data analysis : IDW on map view(Geo-spatial analysis) 63
6.1.9 Data analysis: Summary of responses (Descriptive Statistics) 63
6.1.10 Data analysis: Frequency distribution (Descriptive analysis) 64
6.1.11 Data analysis: Percentage Breakdown (Descriptive statistics) 64
6.1.12 Data Sharing and Commercialization 65
6.1.13 Collaboration 66
CHAPTER 7: Conclusions and Recommendations 67
7.1 Conclusions 67
7.1.1 Limitations 68
7.2 Recommendations for Future Work 68

8. Ethics and Professionalism 69
References 70

TABLE OF FIGURES

Figure 1. system architecture 17
Figure 2: Collaboration and team efficiency 27
Figure 3: system framework 28
Figure 4: System overview 29
Figure 5: Methods 34
Figure 6: Model diagram of a user Account 39
Figure 7: Survey management subsystem model 41
Figure 8: Data Collection Subsystem Model 42
Figure 9: Collaboration and Team Management Subsystem Model 43
Figure 10: Model of Data Analysis & Visualization Subsystem 44
Figure 11: Dataset and Market Place Model 46
Figure 12: Questionnaire creation 57
Figure 13: Data Analysis (Dashboard) 58
Figure 14: Data analysis: Visualization 59
Figure 15: Data analysis: Hypothesis Testing (Inferential statistics) 60
Figure 16: Data analysis: Confidence interval (Inferential statistics) 61
Figure 17: Data analysis: Correlation/Regression (Inferential Statistics) 62
Figure 18: Data analysis: IDW (Geo-spatial analysis) 62
Figure 19: Data analysis : IDW on map view(Geo-spatial analysis) 63
Figure 20: Data analysis: Summary of responses (Descriptive Statistics) 63
Figure 21: Data analysis: Frequency distribution (Descriptive analysis) 64
Figure 22: Data analysis: Percentage Breakdown (Descriptive statistics) 65
Figure 23: Data Sharing and Commercialization 65
Figure 24: Collaboration 66

Abstract

In the era of information, effective survey systems are essential for making sound decisions in all sectors of research, governance, and business. However, many traditional survey systems—especially in developing countries like Malawi—are plagued by inefficiencies in data acquisition, poor cooperation, inadequate integration of geo-spatial intelligence, and limited analytical capabilities (Chirwa, 2025). This dissertation introduces Mobile Survey, an innovative, integrated survey system developed to address these limitations.

The system comprises a mobile application and a website, facilitating real-time collaborative data collection, offline data synchronization, and advanced analysis utilizing statistical software and generative AI. It incorporates geo-spatial mapping functionality to enhance decision-making and includes a geo-portal component that enables users to share or monetize their data. Developed using cutting-edge technologies such as React Native, Next.js, FastAPI, and Firebase, Mobile Survey offers a secure, scalable, and user-friendly solution (Chirwa, 2025).

User testing indicated high usability and performance, though further refinement of the user interface and support for additional question types are necessary. This project makes both technical and scientific contributions to the field of survey systems by addressing gaps in collaboration, spatial integration, and AI-based analysis (Chirwa, 2025).

The findings of this study are applicable to researchers, government agencies, NGOs, and the general public, providing a robust foundation for efficient, intelligent, and inclusive data collection and analysis (Chirwa, 2025).

CHAPTER 1: Introduction

1.1 Background

Nowadays, many organizations and companies rely on survey data for better decision-making. This data can be raw data or geographic data. This data is analyzed before decisions are made. In Malawi, data is commonly collected using hard-copy forms; some even use systems like Google Forms and the Koby app. These systems are called survey systems. Survey systems are programs that are capable of collecting data and saving it in the database for future use. The collected data is then put into separate statistics software such as SPSS for analysis. The survey system that will be described here will have both data collection and data analysis capabilities. The data collection subsystem involves the following functionalities: creation of the survey form, assigning users to the project to collaborate in data collection, assigning rights or privileges to users on data management such as edit and delete, filling out the survey form, submitting the survey form, and displaying responses made by the user.

The system user interface was developed using Next.js. Next.js is a React framework that simplifies web development by providing features such as server-side rendering, static page generation, and optimized image handling (Next.js, n.d.). This system has two backend servers that are both used for data analysis. One server is written in Python using the FastAPI framework. The FastAPI framework is an open-source Python web framework used to develop web APIs with the hint types of Python 3.6 or higher. It enables data types to be validated even within JSON requests. It is based on standards such as JSON Schema, OAuth 2.0, and OpenAPI (FastAPI, n.d.). This server is mainly used for complex data calculations that are available in the Python environment, such as Pandas, NumPy, etc. The second server is an Express server using JavaScript and TypeScript languages. This server is used for retrieving, uploading, sorting, grouping, and deleting data from the database. The databases that were used are Google Cloud Storage (this is a cloud database owned by Google for storing

files) and Google Firestore (this is a NoSQL database that is used to store structured and unstructured documents in text format) (Google Cloud, n.d.; Firebase, n.d.). Users are authenticated using email and password, and this authentication is done using the Google Firebase Authentication API. This authentication system was used because it is secure and easy to implement (Firebase Authentication, n.d.).

1.2 Problem statement

Data collection in several low-income nations such as Malawi is still performed on paper forms using manual input methods that are highly unstructured and fragmented. This leads to data being unprocessed, analysis lagging, and little involvement of the community in how their data is used (Chirwa, 2025). Also, integrated systems for storing real-time data in the field, collaboration system tools, geospatial visualization, and monetization do not exist.

A fundamental challenge is the disconnect between those collecting data and those who benefit from the data, where valuable data gets aggregated in government or NGO databases with little feedback or utility for the communities contributing that data (Chirwa, 2025). Surprisingly, existing survey systems rarely offer location intelligence and spatial analysis—key tools for addressing and finding solutions to geographic challenges such as health outbreaks, land disputes, or resource allocation (Chirwa, 2025).

1.3 Main and specific objectives

1.3.1 Main Objective

Develop an integrated survey system using mobile and web technology, enhanced with geo-spatial and artificial intelligence, which enhances the data collection, data analysis, data interpretation, and data distribution capabilities of a community, researcher, and policy maker.

1.3.2 Specific Objectives

This project aims to:

1) Integrate collaborative features in the system to allow people to work as teams in data collection.
2) Implement analysis and geo-spatial mapping to enhance location-based decision making.
3) Implement a geo-portal where users will sell their survey data or share it for free.

1.4 Research questions

A. How can collaborative features be effectively implemented into a survey system to improve team efficiency in data collection and analysis?
B. How can analysis features be used to support location-based decision-making in survey systems?
C. What are the important features of a successful geo-portal for sharing or selling survey data?

1.5 Scientific Significance

This research has important scientific and practical value, as it fills many gaps in existing survey systems and contributes to the improvement of data collection, analysis, and decision-making processes. Below is the scientific significance of this research work:

1.5.1. Improved Scientific Understanding

1.5.1.1 Collaborative Data Collection

This research contributes to the understanding of how collaborative features can improve team efficiency in collecting and analyzing survey data. The integration of collaboration tools for real-time use has allowed the study to explore how teamwork can be optimized in data-driven environments, filling a knowledge gap in the design of collaborative survey systems (Zhang et al., 2020).

1.5.1.2 AI-Driven Insights

The integration of generative AI for explaining survey statistical results improves the understanding of how artificial intelligence can be used to interpret complex data. This research fills the gap between traditional statistical analysis and modern AI-driven insights, providing a framework for integrating these approaches effectively (Drosos et al., 2024).

1.5.1.3 Geo-Spatial Decision-Making

By incorporating geo-spatial mapping into survey systems, this research improves the understanding of how spatial data (location-based data) can be leveraged for decision-making. It contributes to the growing field of geo-spatial analysis by demonstrating its practical application in survey systems (Drosos et al., 2024).

1.5.1.4 Knowledge Gap Filling

Many of the existing survey systems do not have collaborative features, thereby making it difficult for teams to work efficiently. This research fills this gap by developing a system that supports real-time collaboration, task assignment, and role-based access, enabling teams to work seamlessly (SurveyMonkey, 2025). The following are the gaps this system tries to fill.

1.5.1.5 Limited Data Interpretation

Current data analysis tools often find it difficult to provide insightful explanations of statistical results. This research addresses this limitation by integrating generative AI to provide clear, actionable insights, making data analysis more accessible to non-experts (Drosos et al., 2024).

1.5.1.6 Absence of Geo-Spatial Features

The absence of geo-spatial capabilities in traditional survey systems makes it difficult to make decisions based on spatial data. This research fills this gap by

including geo-spatial mapping and analysis tools, thereby allowing users to make informed decisions based on geographic data (Beardsley, 2024).

1.5.2. Technical Contribution

1.5.2.1 Integrated Platform

This work offers a modern, integrated survey system that streamlines the process of data collection, geo-spatial mapping and analysis, AI-based assumptions, and real-time online collaboration all in a single platform. This strategy also yields an essential technical enhancement relative to existing tools (Drosos et al., 2024).

1.5.2.2 Generative AI Integration

Using generative AI to understand and communicate statistical results is a major addition to the survey field. With an innovative solution, this enhances the usability of data analysis tools by giving users intuitive insights (Drosos et al., 2024).

1.5.2.3 Geo-Portal Development

The launching of a geo-portal for sharing and selling survey data adds a brand-new technical capacity that allows users to charge or give away their data while ensuring the security and access of information. This feature has a lot of advantages and can be used by a wide range of users (Beardsley, 2024).

1.5.3. Potential Use and Users

1.5.3.1 Academic Researchers

The system can be utilized for the collection of survey data but also for a better analysis process of survey data so that researchers can efficiently focus on deriving insights rather than managing data collection processes (Drosos et al., 2024) .

1.5.3.2 Government Agencies

The geo-spatial features and AI-driven insights embedded in this system can be beneficial for government agencies working on urban planning, public health, and environmental monitoring (Beardsley, 2024).

1.5.3.3 Private Sector

The system can be used to help businesses conduct market research, gather feedback, and make strategic decisions.

1.5.3.4 Non-Governmental Organizations (NGOs)

The system can be used for field data collection in remote areas to create data-driven solutions for community-based projects.

1.5.3.4 General Public

Users, such as individuals or organizations, can share or sell survey data through the geo-portal, which will act as a marketplace for data.

CHAPTER 2: Literature review

The concepts of data collection and analysis are very important for decision-making in various sectors, such as academia, government, and business. However, traditional survey systems encounter challenges such as inefficiencies in data handling, lack of real-time collaboration, and limited geo-spatial integration (Smith & Johnson, 2020). Current improvements in mobile technology, cloud computing, and artificial intelligence introduce many opportunities to improve survey systems by enhancing accessibility, security, and analytical capabilities (Lee et al., 2021). This literature review explores existing research on survey systems, collaborative data collection, AI-driven insights, and geo-spatial integration with the purpose of identifying gaps and justifying the proposed system.

2.1 Currently used surveys

There are seven types of surveys that are commonly used to collect data. These types include paper surveys, mail surveys, telephone surveys, in-person surveys, panel surveys, focus group surveys (InMoment, 2024), and online surveys.

2.1.1 Paper surveys are an old-fashioned way to collect data using hard copies. This method is useful in areas where people have no access to smartphones or computers; it is also useful in areas where the internet is not accessible. Some people also find paper surveys flexible and easy to read since they are printed in large fonts (Snapsurveys, 2025).

2.1.2 Mail surveys are also old-fashioned, but they have been used for a long time. The reason why this method has been around for so long is that it feels authentic and provides access to an audience that may be less willing to respond online. Those demographics who are less likely to answer an online email survey are often familiar with mail surveys and will be willing to respond (Greenwald Research, 2022). In addition, mail surveys also have a wide geographic reach since most places are accessible by post (ScienceDirect Topics, n.d.).

2.1.3. Telephone surveys are also a common type of survey. This involves an interviewer calling a respondent and asking them questions. The interviewer will record the responses in a computer or charts based on what the respondent has answered. This system allows the interviewer to add a personal touch and also include follow-up questions, ask for response clarifications, and these capabilities are more important and are missing in other systems such as mail and paper surveys (SmartSurvey, 2025).

2.1.4. In-person surveys are a common method of data collection involving face-to-face interactions between interviewers and respondents. This approach allows interviewers to ask follow-up questions based on respondents' answers, facilitating deeper insights (ATLAS.ti, 2023). Additionally, in-person surveys are beneficial for collecting data from individuals with low literacy levels, as interviewers can explain questions in an understandable manner (Doyle, n.d.).

2.1.5. Panel surveys involve selecting a group of individuals as a sample, who are then asked to answer questions over a period of time. This longitudinal approach allows researchers to track changes and gain long-term insights from the target audience (SurveyMonkey, n.d.; QuestionPro, n.d.).

2.1.6. Focus group surveys involve a group of 8 to 10 participants who engage in discussions moderated by a professional to ensure the conversation remains on track and unbiased. This qualitative research method is instrumental in uncovering participants' attitudes and perceptions, providing valuable insights for market research (Scribbr, 2023; Wikipedia, 2025).

2.1.7. Online surveys involve people filling survey forms using their smartphones and computers. This system requires internet access to submit the survey response. The interviewer only creates the questions and adds them online, and then respondents are accessed to answer the questions, with the answers recorded and stored in a database. This approach is important because it is fast and can reach a lot of people at the same time; it can also reach people who are very far away, thereby providing insights from a wide range of people (SmartSurvey, 2025; Qualaroo, 2025).

2.2. Challenges in Current Traditional Survey Systems

2.2.1 Missing Collaborative Features

A lot of the existing survey systems do not support real-time collaboration, thereby making teamwork inefficient (Brown & Davis, 2019). Studies show that role-based access control and features for assigning tasks are important because they improve team coordination in data collection (Zhang et al., 2022). Cloud-based solutions, such as Google Forms and SurveyMonkey, offer limited collaborative functionalities, necessitating more advanced systems (Wilson, 2021).

2.2.2 Limited Data Interpretation and Visualization

The current survey tools mainly offer elementary statistical reports, which do not deliver substantial analytical findings (Miller, 2020). The implementation of interactive dashboards, together with dynamic charts through advanced visualization techniques, leads to better data understanding (Few, 2017). The implementation of AI for automated insights stands as an uninvestigated territory, according to Chen et al. (2023).

2.2.3 Absence of Geo-Spatial Integration

The absence of geo-spatial mapping in traditional survey systems restricts organizations from making location-based decisions (Garcia & Martinez, 2021). The integration of Geographic Information Systems in surveys using platforms such as ArcGIS Survey123 (ESRI, 2022) enables improved spatial analysis (Garcia & Martinez, 2021). The combination of statistical tools with seamless integration introduces a current challenge, according to Kumar et al. (2023).

2.2.4 Expensive to Collect Data

The systems that are used are very expensive. For example, paper surveys require papers to be printed and cost a lot, especially when the survey targets a lot of people.

2.3 PROPOSED SYSTEM CHANGES

The system we are developing aims to fill the gaps in the existing survey systems. Unlike paper and mail surveys, this survey system will not cost a lot of money, as there will be no need to print papers. This system is an online system with the capabilities of recording responses even in offline mode and then syncing the data when the internet becomes available. This system will be available on both mobile devices and computers. A separate mobile app is developed, and the main purpose of this app is to collect data. This is an advantage because there are a lot of people with smartphones compared to people with computers, and also mobile apps simplify things with just a click of the app icon, allowing users to access the system instead of going to the browser and typing in the site URI, which is considered a lot of work by some people.

This system also has a website whose main purpose is to analyze data collected from mobile devices. Even though the mobile app's main purpose is to collect data, this website also has a dedicated page for data collection, meaning that data collection can also be done using the website. The creation of survey forms is also done on the website. The questions are structured and strongly typed, thereby allowing only responses with predefined types; otherwise, they will be rejected. This is important for data quality and integrity.

The following are explanations of the functionalities of this system.

2.3.1 Mobile app

- Ability to authenticate the user. This app will allow the user to log into the system. This authentication will allow the user to be added to the survey project for collaboration
- Ability to respond to questions. This app is built with the capability to detect the type of response required by the question and provides the necessary way to respond to the question. This system has the following types of questions:

    1. Multiple choice single select. This question has a list of options where a user has to select only one.
    2. Multiple choice multiple select. This question has a list of options where the user can select one or more of the options.
    3. Short text. This question accepts text, and the text length is limited to a certain amount of words specified by the person creating the survey form.
    4. Long text. This question accepts text input of any length.
    5. Number. This question only accepts digits, which include integers, floats, and longs.
    6. Coordinates. This question accepts geographic coordinates. The app has a special input for this type of question, where it provides the user with the option of inputting the options manually or automatically detecting the current location
    7. Image. This question only allows image input. It opens the file explorer where the image will be uploaded to the system.
    8. Audio. This question expects audio input. It opens an audio recorder where the user will record their voice and upload it to the system.
    9. Date. This question accepts dates. It has a date picker from which the date will be selected.
    10. Time. This question accepts time. It has a time picker from which the time will be selected.

11. Likert scale. This is a scaling scale that includes the options that rate an object being studied. It allows the user to select only one option.

12. Yes/No. This only has yes and no options and only allows the user to select one.

- Ability to view responses. The surveyor will be able to check and review the responses that were collected.
- Ability to edit the responses. It is probable that mistakes can be made during data collection, and for clean data purposes, this data has to be corrected.
- Ability to have administrative power over other surveyors. By default, the person who created the survey form is an administrator, and this person has the rights to add and remove participants. Once the survey has been removed, he will not be able to fill out the survey form unless he is added again. The administrator will also be able to give administrative power to other participants.

2.3.2 The website

- Ability to create survey forms. It has well-structured user design that will allow the user to create structured and typed questions. The questions can also have follow up questions. The user can delete or edit questions if mistakes were made in question creation.
- Ability to answer questions. Just like the mobile app, this website is also able to to collect survey data.
- Ability to view and edit survey responses.
- Ability to have and share administrated power.
- Ability to analyze data. The collected survey data is accumulated in the cloud database. This data is the analyzed in the website. The following are the statistics that are done in this website:

    1. Response summaries. This provides the summaries of responses on each question. Each summary displays the following:

        A. The number of responses collected

B. The response that appears more frequent that others.

C. The number of unique responses

2. Percentage distribution. On each question that has categorical data as its responses, each response is assigned a percentage based on how many respondents selected it.

3. Confidential intervals. This measures the degree of uncertainty or certainty in a sampling method. They can take any number of probability limits, with the most common being a 95% confidence level (Hayes, A. (n.d.). What Is a Confidence Interval and How Do You Calculate It? Investopedia. Retrieved April 2, 2025, from https://www.investopedia.com/terms/c/confidenceinterval.asp).

4. Correlation and regression analysis. Correlation refers to how variable are lineary related to one another. Regression allows the researchers to predict or explain the variation in one variable based on another variable.

5. Visualization. Charts are used to display the data. These charts include:

A. Bar charts

B. Pie charts

C. Dougnut charts

D. Scatter plots

E. Rada charts

F. Polar Area charts

6. Geospatial analysis. This involves the following:

A. Mapping distribution of responses. This shows the location where the data was collected.

B. Heat map. Heat map depicts values for a main variable of interest across two axis variables as a grid of colored squares (Yi, M. (n.d.). A complete guide to heatmaps. Atlassian. Retrieved April 2, 2025, from https://www.atlassian.com/data/charts/heatmap-complete-guide).

C. Kriging. Kriging can be understood as a two-step process: first, the spatial covariance structure of the sampled points is determined by fitting a variogram; and second, weights derived

from this covariance structure are used to interpolate values for unsampled points or blocks across the spatial field (Columbia University Mailman School of Public Health. (n.d.). Kriging Interpolation. Retrieved April 2, 2025, from https://www.publichealth.columbia.edu/research/population-health-methods/kriging-interpolation).

- It stores and publishes data that other can download and use them in their projects. A user can choose whether to sell the data or just to share it for free. This feature also acts as a geoportal as it can store spatial data such as geo-json data. This feature will also encourage people to sell the data for profit at the same time proving valuable data that one an use in his or her business.

2.4 USER REQUIREMENTS DEFINITION

This survey system will provide the following services:

2.4.1 Functional Requirements

A. Authenticate the user
B. Create survey form
C. Submit survey responses to the cloud database
D. Save the data to the local storage when there is no internet access and sync the data when internet is available.
E. Manage responses. This includes editing and deleting.
F. Analyzes collected survey data.
G. Shares data for free or for a profit by selling it.

2.4.2 Non Functional Requirements

2.4.2.1 Userbility requirements

This system is built using React on both the mobile app and website. The mobile app user interface, which is more user-friendly and easier to

navigate, is developed using React Native, a React framework that simplifies mobile app development using JavaScript and TypeScript (Hidden Brains, 2024). The website user interface, which is more user-friendly and easier to navigate, was developed using Next.js, a React framework developed by Vercel that simplifies web development (Next.js, 2025).

2.4.2.2 Efficiency Requirements

This system requires the gadget to meet the following requirements in order to use this system.

2.4.2.3 Mobile device

i. The mobile device should have at least android version 8
ii. The device should be able to connect to the internet

2.4.2.4 Browser

i. The device should have a browser that will be used to access the website.

CHAPTER 3: System Models and Architecture

This survey system is a collection of many subsystems that both use the same database for data storage and authentication. The mobile app is a system on its own with subsystems such as authentication system, response answering subsystem, collaboration subsystem and local storage management subsystem for saving data when the user is offline. The website is also a system on it's with the following subsystems: authentication subsystem, data analysis subsystem, generative ai subsystem, geo portal and data sharing subsystem, collaboration subsystem, project participants management subsystem.

3.1 system Overview

3.1.1 system architecture

system architecture is the conceptual model that defines the structure, behavior, and views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system (Wikipedia contributors. (n.d.). Systems architecture. Wikipedia, The Free Encyclopedia. Retrieved April 8, 2025, from https://en.wikipedia.org/wiki/Systems_architecture).

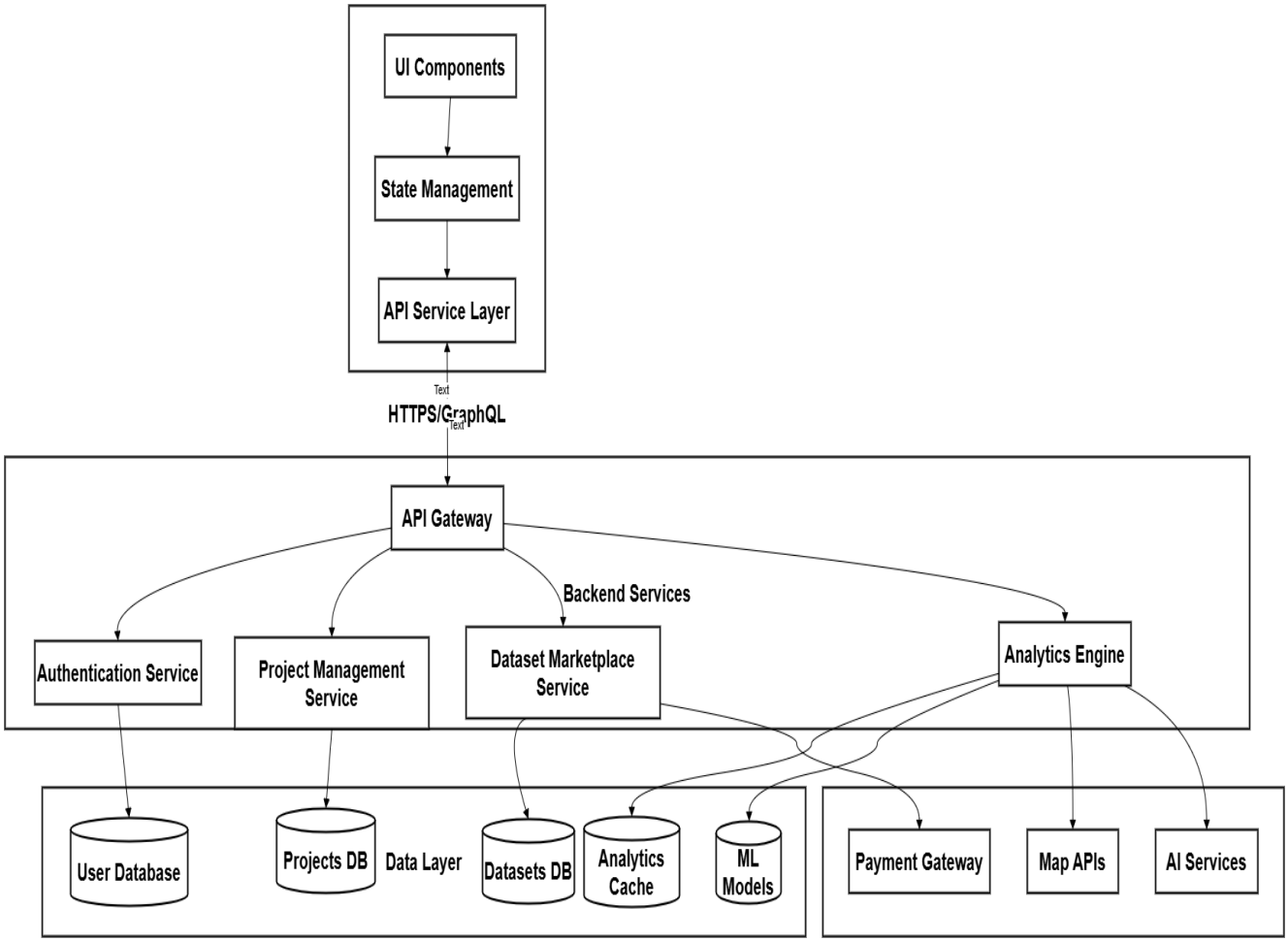


Figure 1. system architecture

CHAPTER 4: System requirements

4.1 Functional requirements

4.1.1 Authentication

The system is supposed to provide a strong authentication method. No one, even system programmers, should have access to the users' authentication credentials. This can be done using the encryption process. Encryption is the process of converting the original information into a collection of characters that requires a key to convert it back to the original text (Cloudflare, n.d.). This process will make users feel safe knowing that their data is secure (Eskuad, n.d.).

4.1.2 Data analysis

The system should be able to accurately calculate the commanded statistics method. The calculation should not take a lot of time. The calculation should not use a lot of the user's CPU; therefore, all calculations should be done on the server, and the results returned to the user, thereby allowing users with low-end computers to be able to analyze the data too (Vaia, n.d.). In visualization, the graphs or charts should be visible enough for better interpretation and analysis, and a wide range of charts should be available (Tableau, n.d.). The user should also be able to download these charts. The system should also be able to accurately present geographical data. The map used should be clear and big enough to show all data overlays clearly. The interpolation method used, kriging, should be able to provide an accurate variogram of the interpolated data and also be able to return the interpolation map so that the user can view and download the map (Esri, n.d.).

4.1.3 Secure transactions

Since users will be able to sell their datasets, the system should provide a secure way of handling users' money (Smith, 2022). The system should also keep records of all transactions, including payment details such as the buyer, the seller,

the dataset, and the time of transaction (Johnson & Lee, 2021). It should integrate with secure payment providers such as PayPal, Stripe, or mobile money (PayPal, 2023; Stripe, 2022). Additionally, the system should allow users to dispute transactions and request refunds if necessary (World Bank, 2021).

4.1.4 Survey creation and management

A. The system should be able allow the users to create survey forms. These survey forms should contain questions or fields that users will be able to answer.
B. The system should provide the user with different types of questions so that the user can choose one that fits his question.
C. The system should also allow the user to add follow up questions to the questions.
D. The system should also allow the user to edit the questions when the user has made a typo or any mistake in creating a survey from.
E. The system should also allow the user to delete the question.
F. The system should also allow the user to submit the survey form once finished creating it so that it should be ready for people to collect data.

4.1.5 Data Collection and Responses

A. The system should allow the users to submit their responses once they have finished answering it.
B. It should also have an auto save functionality whereby the data will be saved during internet loss.
C. The system should check and make sure it only allows the specified data types of the questions defined by the person who created the survey form to be submitted
D. User should also be able to edit and delete their responses even if they are already submitted

4.1.6 Collaboration and Access Control

A. Users should be able to edit or delete the responses that he or she made but should not have access to other people's responses.

B. The admin should be able to give data management rights to other users but these rights should not include deleting the project or other people's responses.

4.1.7 Geo-Spatial Mapping and GIS Integration

A. The system should be able to collect location-based data such as coordinates.
B. The system should also be able to present the spatial data.

CHAPTER 5: Methodology

5.1. Research Design

This research employs a mixed-methods research design that combines both quantitative and qualitative approaches to fully assess the effectiveness and impact of the survey system (Creswell & Creswell, 2018). This means the design allows a solid evaluation of the investigation because it uses the strengths of both methodologies while minimizing their weaknesses (Johnson et al., 2020).

5.2 Threats to Internal and External Validity

5.2.1 Internal Validity

This section explains the degree to which the results are not affected by internal factors (Campbell & Stanley, 1963). The following are potential factors that can affect the results:

5.2.1.1 Selection bias.

This is the selection of samples based on their characters but not random sampling (Shadish et al., 2002). This includes selection of participants for data collection based on their skills. For example, on a survey project for collecting coordinates of places, people with GIS skills are likely to be selected based on experience on working with spatial data. This provides a sense of bias on data collection and the system should be able to collect any data regardless of the skills of the person collecting the data.

5.2.1.2 Instrumentation changes

During system development there can be frequent change of instrument changes (Kirk, 2013; Pressman, 2020). This often happens in consideration of system performance and user needs (Sommerville, 2016). There are a lot of tools for system development which includes text editors, system frameworks and system

dependencies (Martin, 2008). For example, if the system is using a certain dependency for data visualizations but this dependency takes a lot of user RAM and CPU it is not a good dependency because it can lead to poor user experience (Nielsen, 2012). Therefore, there is a need to use an alternative dependency that uses less of user RAM and CPU (McConnell, 2004).

Another factor that influences instrument changes is User Feedback (Krug, 2014). Users can find some system functionalities confusing and difficult to use and they tell the developers about the problems they faced during using the system (Norman, 2013). Therefore, the developers upgrade the system in order to meet user needs (Preece et al., 2015). This process is important but it slows down the project and making the project to miss the deadline the the process takes time (Brooks, 1995).

sParticipants' reactions during testing or data collection may be affected by external variables such as environmental conditions, time pressure or personal motivations (Shadish et al., 2002). For example, a user can provide very positive response because they know developers individually (Rosenthal, 1976), or they can run through a task due to fatigue or network problems (Dillman et al., 2014). These factors can distort the reliability of the collected data and make an incorrect conclusion about the efficiency or purpose of the system (Campbell & Stanley, 1963). To reduce this danger, where possible, tested in a controlled and neutral environment (Webb et al., 1966), and users should be encouraged to provide honest and fair feedback (Krug, 2014).

5.2.2 External validity

External validity refers to the extent that the results of the study can be generalized beyond the specific context in which research was conducted (Campbell & Stanley, 1963). In this study system project, it is important to ensure that the efficiency, purpose and performance of the system is not limited to a specific group of users, location or conditions (Shadish et al., 2002). Many factors can put external validity at risk:

5.2.2.1 Restricted sample

If the system is only tested with users of a particular geographical location, demographic group or professional background (e.g., only urban data collectors or IT professionals), the results cannot be generalized to a wider population (Bracht & Glass, 1968; Rothwell, 2005). For example, if only experienced smartphone users evaluate the app, the usability findings cannot reflect users with limited digital skills (Norman, 2013; Seffah et al., 2006). In order to ensure generalizability, studies must include participants from different fields, age groups, educational levels and technical backgrounds (Nielsen, 2012; Sauro & Lewis, 2016).

5.2.2.2. Specific technical environment

The performance and adaptability of the examination system may vary depending on device specifications, internet connectivity, and operating systems (Krug, 2014; Nielsen, 2012). If the system is only tested in areas with high-end smartphones or stable network connections, it may perform poorly in real-world conditions that include low-end devices or offline scenarios (Marsden et al., 2018; Sey et al., 2015). To ensure external validity, the system should be evaluated across different device specifications and connection conditions (Sauro & Lewis, 2016), particularly in areas with limited digital infrastructure (Wyche et al., 2016; Sambasivan et al., 2018).

5.2.2.3. Controlled Testing Conditions and Real-World Use

Testing the system in a controlled environment (e.g., a laboratory or classroom) cannot capture the complications of real-world usage, including distractions, environmental constraints, or unexpected user behaviors (Brewer, 2000; Dillman et al., 2014). These factors may significantly affect how the system performs during actual mass data collection (Shadish et al., 2002). External validity can be improved by conducting pilot studies or field tests under authentic conditions (Blandford et al., 2016; Kjeldskov & Graham, 2003).

5.2.2.4. Cultural and linguistic difference

The system may be designed in a language or cultural context familiar to developers but inaccessible to diverse user groups (Díaz et al., 2018; Winschiers-Theophilus et al., 2012). If interfaces, instructions, or question formats fail to incorporate local references or use culturally inappropriate language, this can significantly reduce system effectiveness and adoption in target communities (Dearden & Rizvi, 2008; Reinecke & Bernstein, 2011). Ensuring support for local languages and culturally appropriate designs is critical for widespread adoption and equitable access (Abdelnour-Nocera et al., 2017; Bidwell & Winschiers-Theophilus, 2015).

5.2.2.5. Temporary validity

Over time, both technical environments and user behaviors evolve (Rogers, 2016). A system that performs well today may become obsolete or less effective due to software updates, shifting user expectations, or changes in data collection practices (Kujala et al., 2018; Norman, 2013). Continuous monitoring, periodic updates, and responsive user feedback mechanisms are essential to maintain system relevance and effectiveness over time (Preece et al., 2015; Vredenburg et al., 2002).

5.3 Variables and Conceptual Framework

5.3.1 Variables

5.3.1.1 Independent and dependent variables

In this study, research focuses on evaluating the efficiency and purpose of an investigation system designed for data collection in different field environments. The variable included in the study is classified as follows:

5.3.1.1.1 Independent variables

These are factors that are manipulated or inspected by researchers to determine their impact on the dependent variables.

Variable	Operation details
User Experience Level	Categorized as beginner, intermediate, or advanced, based on the user's familiarity with mobile apps or digital tools.
Internet Connectivity	Availability of internet access during system usage (online vs. offline mode).
Device Specifications	Type of device used (low-end vs. high-end smartphones), based on RAM, CPU, and OS version.
Training Provided	Whether the user received training on how to use the system before data collection (Yes/No).
System Version	The version of the system used (initial vs. updated, with newer features or bug fixes).

5.3.1.1.2 Dependent Variables

These are the outcomes measured to evaluate the effectiveness and usability of the survey system.

Variable	Operational Definition
Data Accuracy	The degree to which collected data reflects true values, measured by

	comparing with verified records or expert validation.
Task Completion Time	The amount of time (in minutes or seconds) taken by a user to complete a standard survey form using the system.
User Satisfaction	Users' level of satisfaction, measured using a 5-point Likert scale questionnaire post-interaction.
System Usability Score (SUS)	A standardized usability scale (0–100) based on user responses to usability-focused statements.
Error Rate	Number of incorrect entries or failed submissions during a data collection session.

5.3.2 Conceptual Framework

The conceptual framework for this study demonstrates the relationship between system design elements (independent variables) and expected outcomes (dependent variables), facilitating assessment of the system's suitability for field data collection (Venkatesh et al., 2016). By integrating Technology Acceptance Model (TAM) constructs (Davis, 1989) with GIS-based decision-making frameworks (Malczewski, 2006), we developed a practical conceptual framework that bridges traditional survey methodologies with modern geospatial analysis capabilities (Goodchild, 2007; Couclelis, 2003). The framework contains 3 key elements:

A. Data Collection Module: Enabling geospatial real-time collaborative data collection.

B. Data Processing and AI analysis: The use of generative AI to improve data interpretation.

C. Geo-Portal for Data Utilization: Allowing users to share, sell and visualize their survey results.

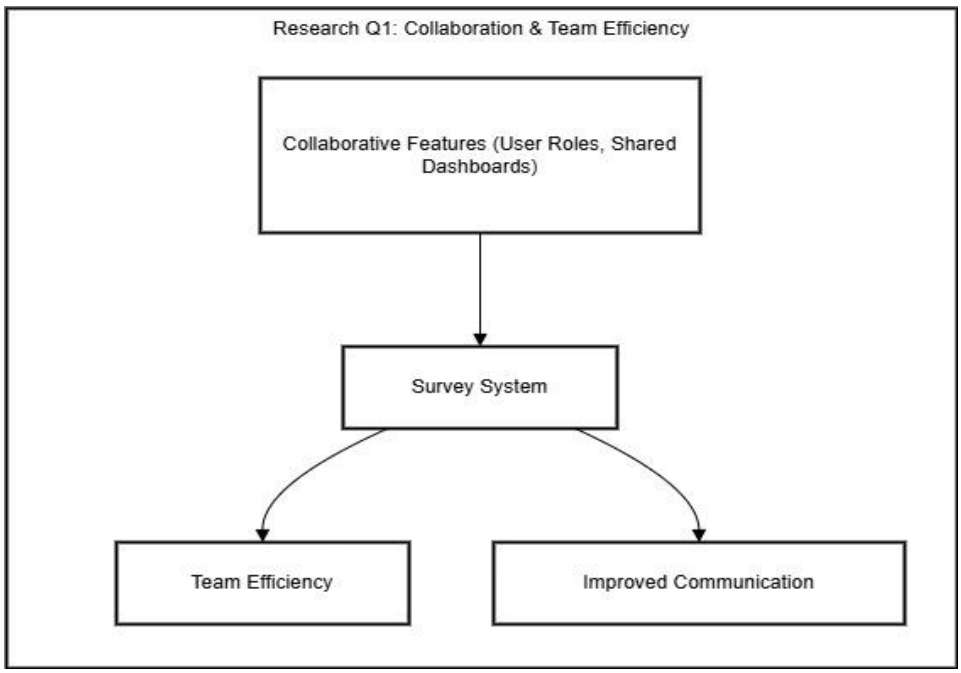


Figure 2: Collaboration and team efficiency

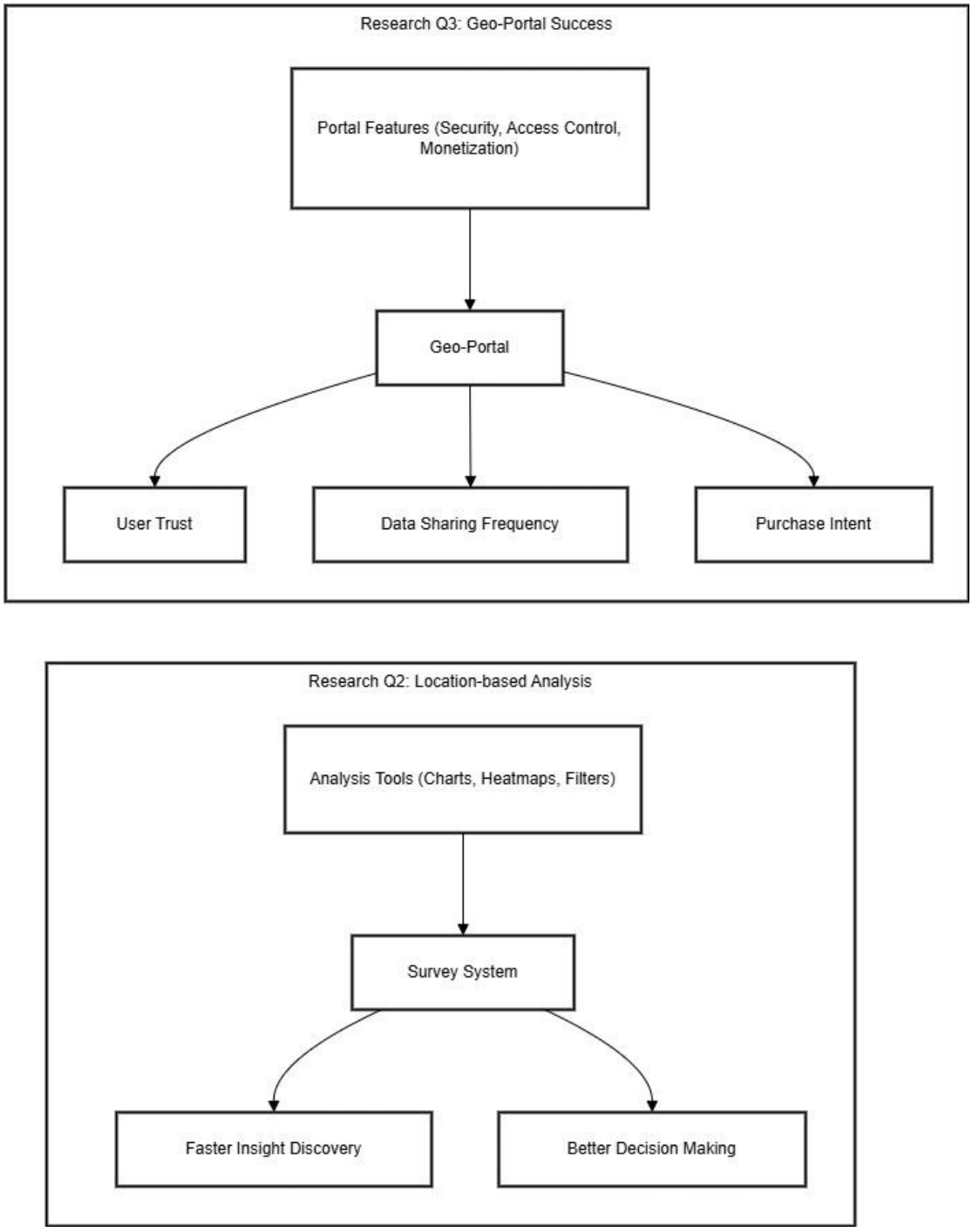


Figure 3: system framework

Such research design allows to estimate how effective the survey system is for data collection and analysis as well as usability and accessibility for different stakeholders. Design also allows for the identification of best practices and areas for improvement in order to develop better data-driven decision-making tools.

5.3.3 Study System and Design

This subsection describes the overarching design philosophy, system architecture, and methodological approach adopted for the software system.

5.3.3.1 System Overview

The name of the system is MobileSurvey. The purpose of this system is to simplify the process of data collection, analysis and data sharing. It also allows people to work as teams with the collaboration capabilities of the system.

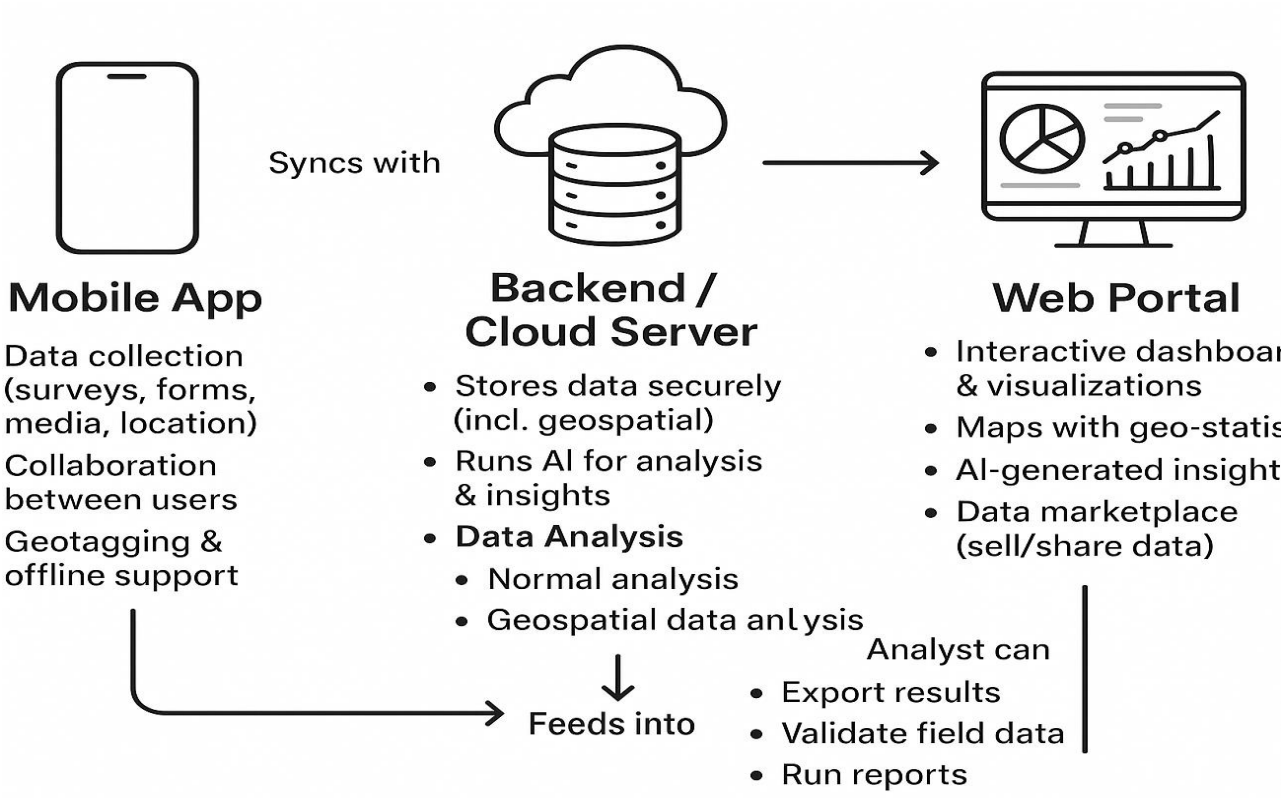


Figure 4: System overview

5.4 Target users

This system is designed to be used by a diverse range of users including:

A. Researcher and Data Analysts

These are people, organizations or companies who are seeking for data and data analysis in order to help them in decision making.

B. Field data Collectors

These are people who are experts on collecting and gathering data during survey.

C. General Public and Business

These are users who are seeking for geo spatial and statistical data that will help them in making good decisions in their business.

D. Students

This system can also be used by students who want to collect and analyze data for their academic or personal projects.

5.5 Core Features

The system includes the following functionalities:

A. Real-time Data Collection and Collaboration

This system allows researchers to collect data of the same project simultaneously. This also allows collaboration among field surveyors and data analysts.

B. Data Visualization and Geo-statistics:

A. Provide graphs and charts
B. Shows map to display response distribution.
C. Supports multiple projections such as 'WGS84 (Lat/Lon)', 'UTM Zone 36S', 'UTM Zone 33N', and 'Web Mercator'.
D. It is also capable for computing weighted geo data interpolation such as kriging.

C. generative AI for Statistical Insights

A. this is used to interpret and explain statistical results.
B. provides summary and insights on collected data.

Geo-Portal for Data Sharing and Monetization

a. Allows users to sell or share their collected data.
b. Ensures secure transactions and controlled access to datasets.

D. Secure Data Storage and Access

A. User login credentials are encrypted.
B. Provides role-based access control for different user levels.

5.6 Materials

The development and implementation of the proposed survey system required a combination of software tools, frameworks, and databases so that they should work together to achieve efficient data collection, data analysis, and data sharing. The materials that were used in this project include the following:

5.6.1 Next.js

This system utilizes Next.js (Vercel, 2023), a React framework designed to simplify web development through its support of JavaScript, TypeScript, and HTML (React Documentation, 2023). The framework enables both server-side rendering (SSR) and static site generation (SSG) (Zakas, 2022), architectural features that enhance our system's scalability and efficiency (Holmes & Marx, 2021). These capabilities align with modern web performance optimization practices (Google Web Fundamentals, 2022).

5.6.2 FastAPI (Python Framework)

This is the Python framework that was mainly used to perform complex and CPU-intensive calculations (Van Rossum & Drake, 2023). This framework was used together with libraries such as pandas (McKinney, 2022), GeoPandas (Jordahl et al., 2022), NumPy (Harris et al., 2020), etc.

5.6.3 Express.js (Node.js Framework)

This is another server that used TypeScript and JavaScript as its languages (Express.js Team, 2023). This server is mainly used for database data retrieval, sorting and grouping (Wilson, 2022).

5.6.4 Google Firestore

This is a cloud NoSQL database that is mainly used to store structured and unstructured documents (Google Cloud, 2023a). This is where we keep our survey data.

5.6.5 Google Cloud Storage

This is a file storage service that is used to store larger files such as GeoJSON data, images, or any binary files submitted through the survey (Google Cloud, 2023b).

5.6.6 Firebase Authentication

This is Google's authentication service that provides many functions of authentication. The function that we used is authentication with email and password (Firebase Documentation, 2023).

5.6.7 Mobile Devices (Smartphones/Tablets)

These were used to test the mobile version of the application during development and for data collection purposes in field trials (Android Developers, 2023; Apple Developer, 2023).

5.6.8 Laptops/Desktops and Smartphones

We needed gadgets to design, develop and test our system. So, the laptops were used for system design and implementation and mobile phones for mobile app data collection (Microsoft, 2023).

5.6.9 Internet Connectivity

We needed internet connection to upload and fetch data from cloud database. Submission of survey forms also require internet connection (Mozilla Developer Network, 2023).

5.7 Methods

The development of the survey system followed a structured process that ensured both the frontend and backend components were efficiently integrated and functional (Sommerville, 2016). The procedures carried out during the project are detailed below:

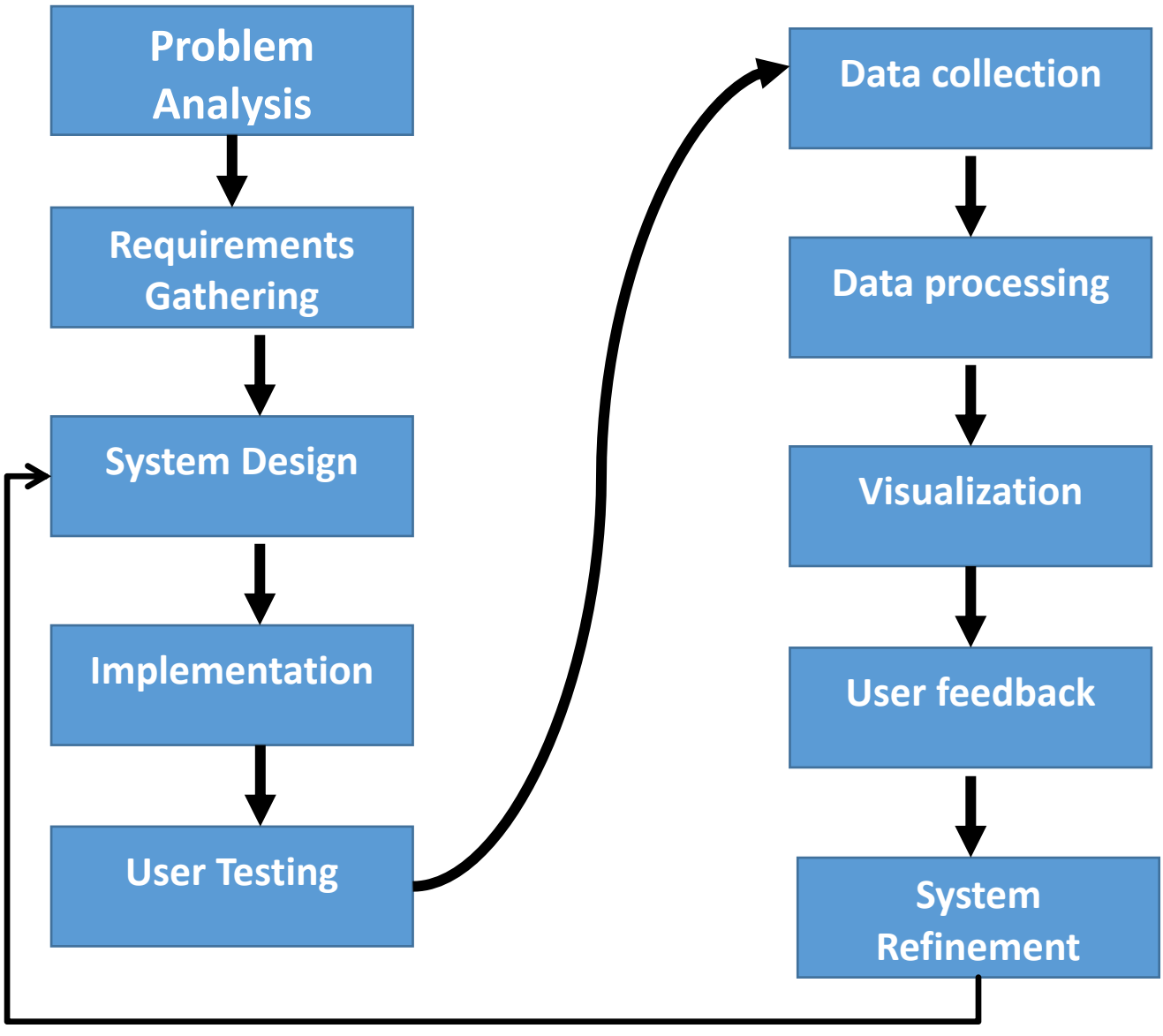


Figure 5: Methods

5.7. 1 Problem Analysis

This stage involves analyzing the limitations that the current survey systems have. It also involves understanding what the users are expecting from the system. We identified that the current survey systems have limitations in the following functionalities:

A. Collaboration features
B. Geo-spatial analysis
C. Data sharing

We did this by installing the current survey system in ou gadgets and started analyzing their functions and identifying their limitations.

5.7.2 Requirements gathering

This involves understanding requirements for a collaborative and efficient survey system. This also involved analysis of the requirements for the core features such as authentication, data syncing, offline data collection and data analysis and visualization.

5.7.2.1. Authentication

The system is supposed to provide a strong authentication method. No one even system programmers should have access to the user's authentication credentials. This can be done using encryption process. Encryption is the process of converting the original information into a collection of characters that requires a key to convert it back to the original text. This process will make users feel safe knowing that their data is secure.

5.7.2.2 Data analysis

The system should be able to accurately calculate the commanded statistics method. The calculation should not take a lot time. The calculation should not use a lot of the user's CPU therefore all calculations should be done on the server and the results returned to the user thereby allowing users with low end computers be able to analyze the data too. In visualization, the graphs or charts should be visible enough for better interpretation and analysis and a wide range of charts should be available. The user should also be able to download these charts. The system should also be able to accurately present geographical data. The map used should be clear and big enough to show all data overlays clearly. The interpolation method used, kriging, should be able to provide accurate variogram of the interpolated data and also be able to return the interpolation map so that the user can view and download the map.

5.7.2.3. Secure transactions

Since users will be able to sell their, the system should provide a secure way of handling user's money. The system should also keep records of all transactions. It should record the payment details including the buyer, the seller, the dataset and time of transaction. The system should integrate with secure payment providers such as PayPal, Stripe, or mobile money. The system should allow users to dispute transactions and request refunds if necessary.

5.7.2.4. Survey creation and management

A. The system should be able allow the users to create survey forms. These survey forms should contain questions or fields that users will be able to answer.
B. The system should provide the user with different types of questions so that the user can choose one that fits his question.
C. The system should also allow the user to add follow up questions to the questions.

D. The system should also allow the user to edit the questions when the user has made a typo or any mistake in creating a survey from.
E. The system should also allow the user to delete the question.
F. The system should also allow the user to to submit the survey form once finished creating it so that it should be ready for people to collect data.

5.7.2.5. Data Collection and Responses

A. The system should allow the users to submit their responses once they have finished answering it.
B. It should also have an auto save functionality whereby the data will be saved during internet loss.
C. The system should check and make sure it only allows the specified data types of the questions defined by the person who created the survey form to be submitted
D. User should also be able to edit and delete their responses even if they are already submitted

5.7.2.6. Collaboration and Access Control

A. Users should be able to edit or delete the responses that he or she made but should not have access to other people's responses.
B. The admin should be able to give data management rights to other users but these rights should not include deleting the project or other people's responses.

5.7.2.7. Geo-Spatial Mapping and GIS Integration

A. The system should be able to collect location-based data such as coordinates.
B. The system should also be able to present the spatial data.

5.7.3 Research design

This system was designed to be a combination of a number of subsystems whereby each subsystem is independent of other subsystems. The following are the designs of these subsystems

5.7.3 .1 User account management subsystem

The system only expects users to register and have an account. This account will be linked to the projects and datasets the user created. This will prevent people from accessing data that is not theirs. Having an account also helps to track, filter and update specific user data.

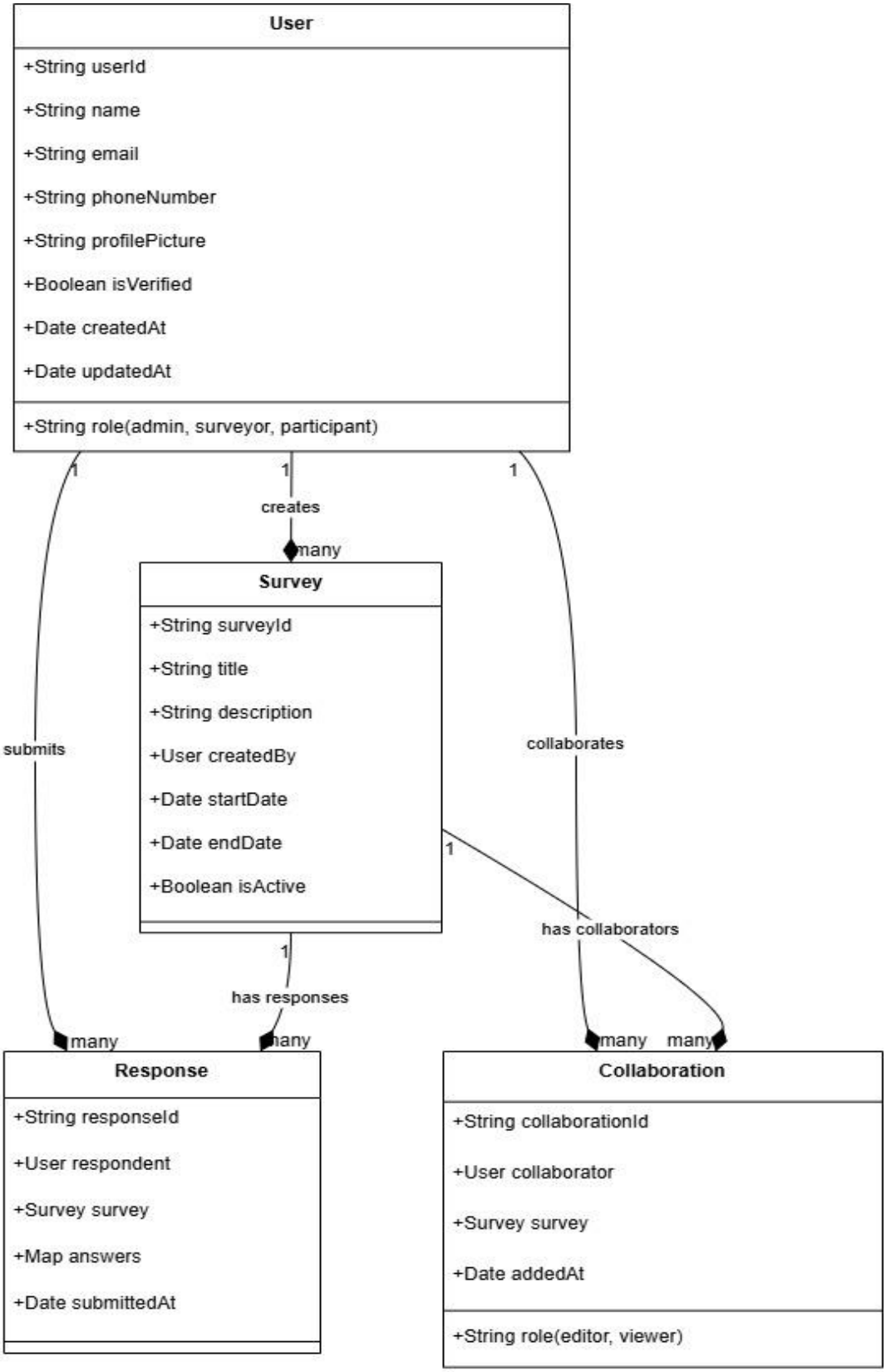


Figure 6: Model diagram of a user Account

5.7.3 .2 Survey Management Subsystem

This subsystem controls creation and management of surveys. This includes:

1. Creation of new survey form.
2. addition of questions
3. assigning type of question responses
4. Deleting questions
5. Editing questions
6. Submitting the survey form
7. Confirming the submission of survey form

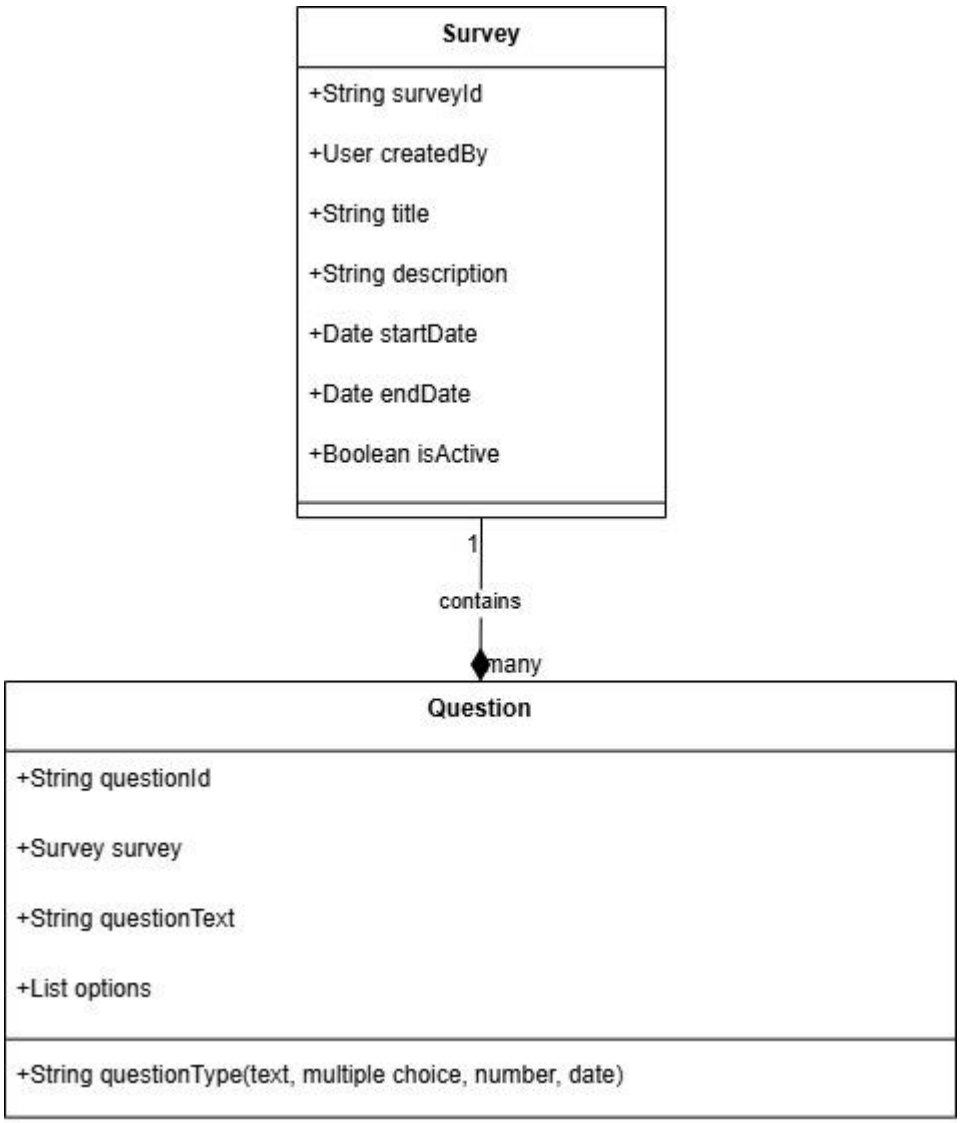


Figure 7: Survey management subsystem model

5.7.3.3 Data Collection Subsystem

This system uses approach of Object-Oriented Programming (OPP) whereby the response or the answer is an Object. This helps to easily track answers that corresponds to a specific question. In data collection subsystem, the user has to fill in answers and by typing text or clicking on the available options. The response is then saved in the answer object. Once the user has finished answering the questions, he can submit the answer object which is stored in the database.

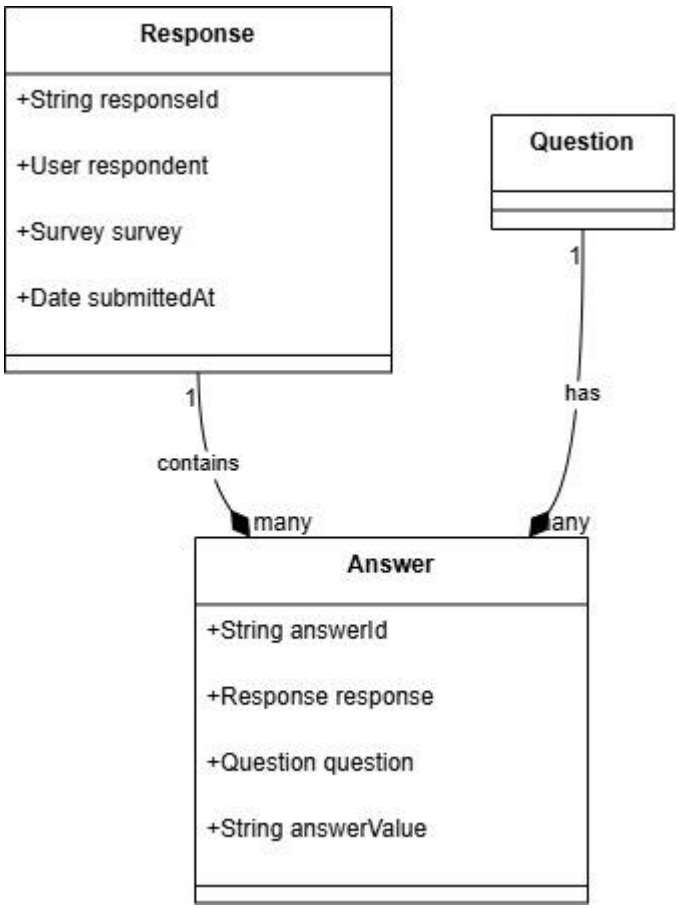


Figure 8: Data Collection Subsystem Model

5.7.3 .4 Collaboration and Team Management Subsystem

This system will allow people to work in teams. This subsystem allows user join and contribute to the project which is not theirs and also allowing them to work as a team.

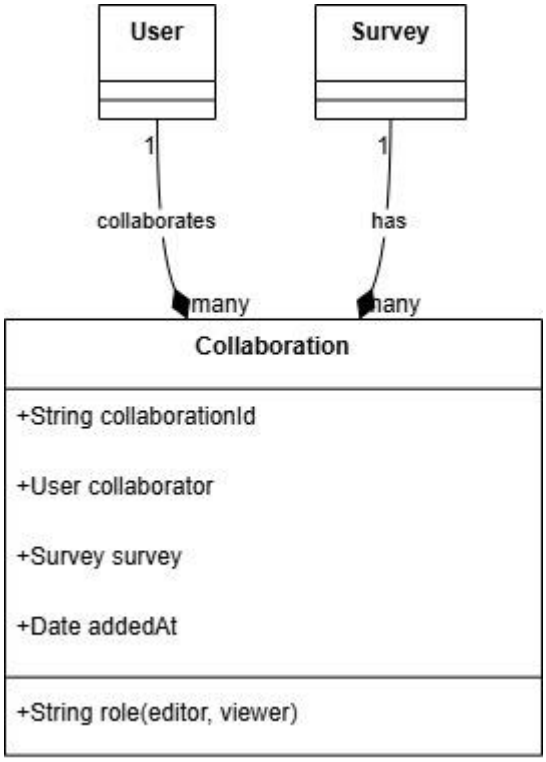


Figure 9: Collaboration and Team Management Subsystem Model

5.7.3 .5 Data Analysis & Visualization Subsystem

This part involves performing statistics calculations on survey data. This part fetches data from the database and perform statistical operation selected by the user. It has also a part of geo statistics where spatial data will also be analyzed.

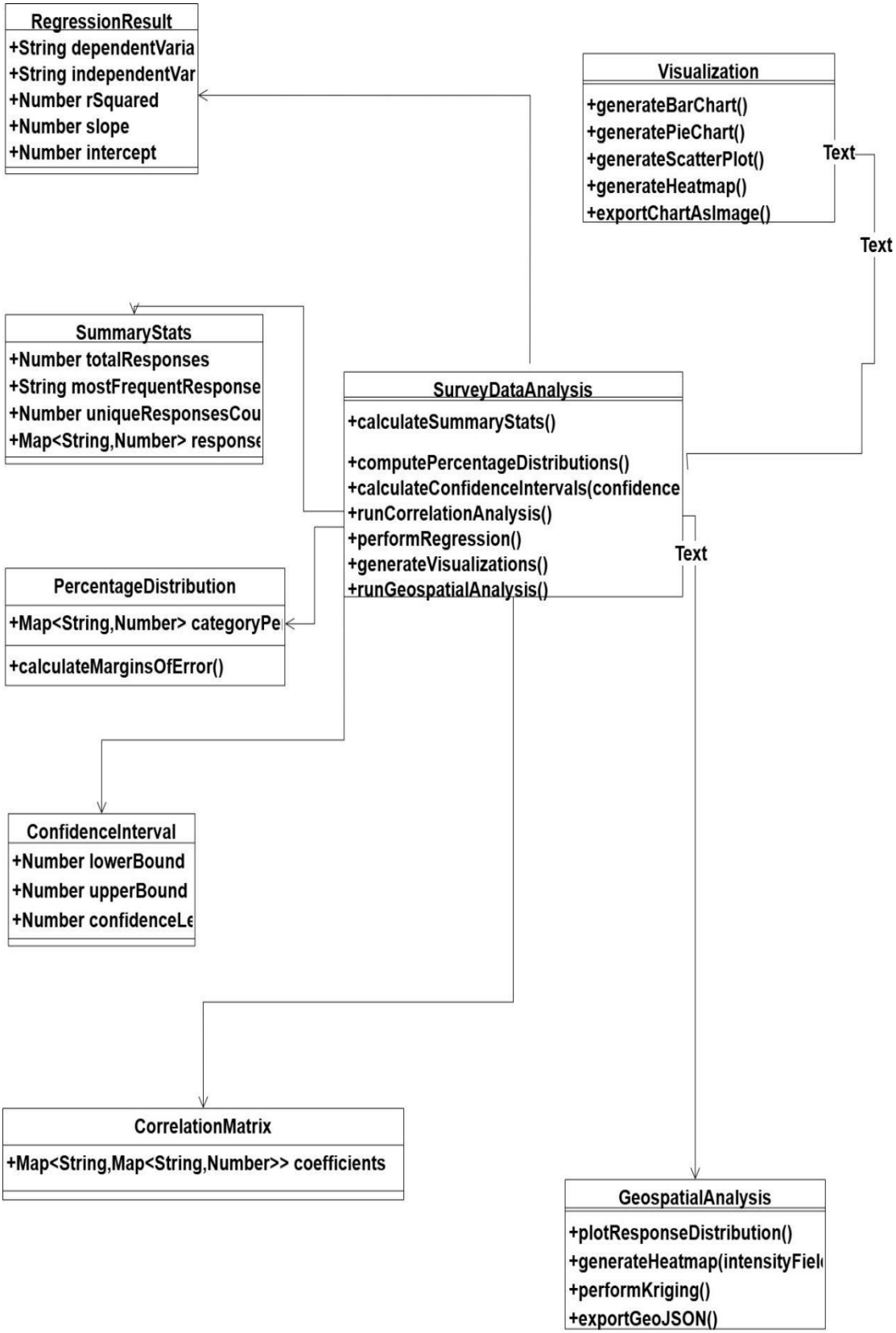


Figure 10: Model of Data Analysis & Visualization Subsystem

5.7.3 .6 Datasets and Market Place subsystem

User will be able share their survey data to the public. This system has a page for displaying datasets shared by users. A user can also upload an external dataset and share it with others. The user can choose whether the share the data for free or for a profit by selling it.

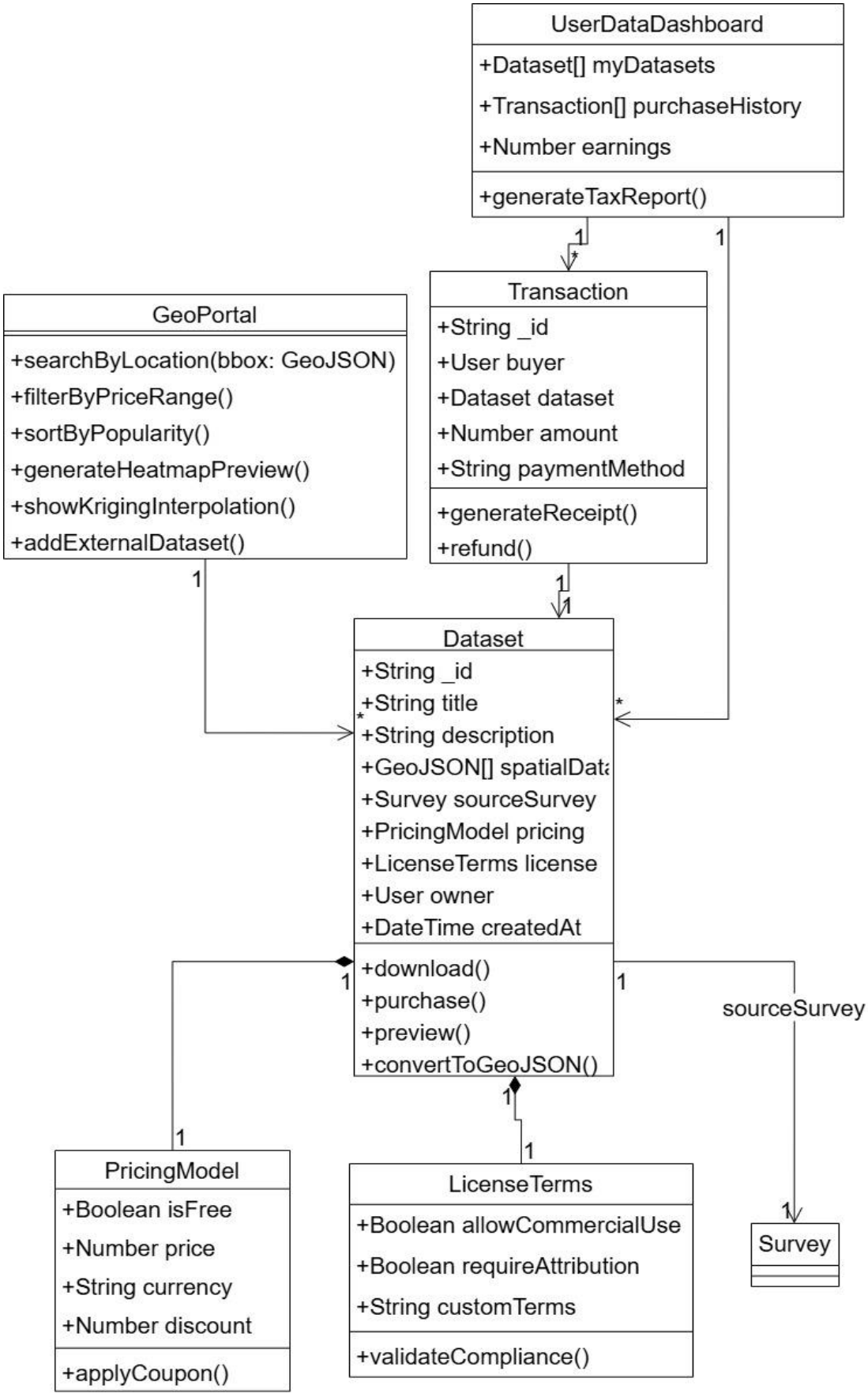


Figure 11: Dataset and Market Place Model

5.7.4 Implementation

After designing our system, we had to implement it. This involves coding the system. We had to code the user interface using NEXT js and our backend using python, JavaScript and Typescript.

5.7.4.1 Front End Development

This involves developing the user interface and managing client-side state (React Documentation, 2023). On the mobile app, the user interface was built using Expo-React-Native (Expo Team, 2023), a React framework that simplifies mobile app development. On the website, the user interface was built using Next.js (Vercel, 2023), a React framework for web development. The interface was constructed using React components - reusable UI elements like buttons and text inputs (Wieruch, 2022).

The mobile app contains these pages: Hero section, Home page, Project Preview page, Project Answer page, Document View page, Document Edit Page, and Project Collaboration page.

The website includes these additional pages: Project Edit Page, DataSet Page, Data Sharing Page, DataSet Preview Page, and Analysis Page.

5.7.4.2 Backend Development

Our system employs two servers for backend functionality. The first is a FastAPI server (FastAPI Team, 2023), a Python framework used for complex statistical and geospatial analysis (McGibbney, 2021). This server handles the following functionalities:

5.7.4.2.1 IDW

This is a Geo-spatial interpolation method. This function is used to predict the value of a location where its data was not collected using it neighboring data. The following is the formula for this function:

$Z(x_0) = (\Sigma\ (1 / d(x_0, x_i) \hat{}\ p) * Z(x_i)) / \Sigma\ (1 / d(x_0, x_i) \hat{}\ p)$

Where:

- $Z(x_0)$: interpolated value at location x_0
- x_i: known data points
- $Z(x_i)$: value at data point x_i
- $d(x_0, x_i)$: distance between x_0 and x_i
- p: power parameter (commonly 2)
- N: number of known data points

Pseudocode Overview

```
Function getIDW(x, y, power = 2, arrays...)
   // arrays = [xValues[], yValues[], zValues[]]

   Combine xValues, yValues, zValues into a single list of points:
      coordsWeight = zip (xValues, yValues, zValues)

   For each point in coordsWeight:
      topSum += Z(xi) / distance (x, y, xi, yi) ^ power
      bottomSum += 1 / distance (x, y, xi, yi) ^ power

   Return interpolated value:
      result = topSum / bottomSum
      Return result
```

Function Breakdown in Pseudocode

Function zip(arrays):

Return a new list where each element is a tuple of the ith items from each input array

Function getDistancePower(x, y, x1, y1, power):

Return ((x - x1) ^ 2 + (y - y1) ^ 2) ^ (power/2)

Function getIDW(x, y, power, arrays):

coordsWeight = zip(xValues, yValues, zValues)

Initialize sumTop = 0

Initialize sumBottom = 0

For each [x_i, y_i, z_i] in coordsWeight:

distance = getDistancePower(x, y, x_i, y_i, power)

weight = 1 / distance

sumTop += z_i * weight

sumBottom += weight

Return sumTop / sumBottom

5.7.4.2.2 Kriging

This is also a Geo-spatial interpolation method whereby it produces a map showing the extent and distribution of interpolated values. The following is the pseudocode for this function:

FUNCTION init_kriging(data, pathList)

TRY

LOAD boundary from GeoJSON file

PARSE data into coordinates (x, y) and corresponding values

```
FILTER points:
    FOR each point in coordinates
        IF point is inside boundary
            ADD point to filtered coordinates
            ADD value to filtered values

IF no valid points remain
    RETURN empty list

NORMALIZE values using MinMaxScaler

INITIALIZE result buffers list

CREATE variogram model with filtered coordinates and normalized values
PLOT variogram and save to buffer
APPEND variogram image buffer to results

DEFINE grid:
    CALCULATE boundary bounds (minx, miny, maxx, maxy)
    CREATE a mesh grid over the boundary
    FLATTEN grid into list of coordinates

MASK:
    CHECK which grid points lie within boundary
    STORE mask as Boolean array

PERFORM ordinary kriging interpolation:
    INITIALIZE OrdinaryKriging with variogram
    COMPUTE estimated values for grid coordinates
    SET values outside boundary to NaN
    RESHAPE estimated values to grid shape
```

```
        PLOT kriging result:
            PLOT interpolated surface using contourf
            PLOT original sample points with scatter
            ADD colorbar, labels, title

            CALCULATE map width in meters using geodesic distance

            DETERMINE appropriate scale bar:
                SET label in meters or kilometers
                COMPUTE position and draw scale bar rectangle
                ADD text label above scale bar

            ADD north arrow to map
            ENABLE grid on plot

        ADJUST layout and save plot to buffer
        APPEND result image buffer to results

        DEFINE upload path using pathList
        UPLOAD all image buffers to Firebase
        RETURN download URLs

    EXCEPT any error
        PRINT error message
        RETURN empty list
```

5.7.4.2.3 Summary of Responses

This function provides the general statistical of the data. The following is the pseudocode for this function:

```
Function SummaryOfResponses(data):
    count = length of data
    mean = sum of all numbers in data / count
    min = smallest number in data
    max = largest number in data
    variance = sum of (each number - mean) ^ 2 / count
    stdDev = square root of variance
    Return {count, mean, min, max, stdDev}
```

5.7.4.2.4 Frequency Distribution

This function shows how many times a unique value appears in the data. The following is the pseudocode for this function:

```
Function FrequencyDistributions(data):
    freq = empty dictionary
    For each value in data:
        If value is in freq:
            freq[value] += 1
        Else:
            freq[value] = 1
    Return list of {value, count} pairs from freq
```

5.7.4.2.5 Percent Breakdown

This function displays the number of times a unique value appears in the data but expressed as percentage. The following is the pseudocode for this function:

```
Function PercentageBreakdown(data):
    freqList = FrequencyDistributions(data)
    total = sum of all counts in freqList
```

```
    For each item in freqList:
        percentage = item.count / total * 100
        Create new object with value and percentage
    Return list of {value, percentage}
```

5.7.4.2.6 Confidential Interval

This function calculates the range of values we expect of result to fall into if repeat the experiment again with a certain level of confidence expressed as a percentage. The following is the pseudocode for this function:

```
Function ConfidenceIntervals(data):
    mean = average of data
    variance = sum of (each number - mean) ^ 2 / length of data
    stdErr = square root of variance / sqrt(length of data)
    CI = [mean - 1.96 * stdErr, mean + 1.96 * stdErr]
    Return {mean, '95% CI': CI}
```

5.7.4.2.7 Hypothesis Testing (T-test)

Hypothesis testing is a systematic procedure for deciding whether the results of a research study support a particular theory which applies to a population. The following is the pseudocode for this function:

```
Function TTest(x, y):
    meanX = average of x
    meanY = average of y
    varX = variance of x
    varY = variance of y
    t = (meanX - meanY) / sqrt(varX/length(x) + varY/length(y))
    df = length(x) + length(y) - 2
    pValue = 2 * (1 - tDistributionCDF(abs(t), df))
    Return {tValue: t, degreesOfFreedom: df, pValue}
```

5.7.4.2.8 Correlation and Regression

Correlation determines the strength and direction of a linear relationship between two variables, while regression aims to predict or explain the variation in one variable based on another. The following is the pseudocode for this function:

```
Function CorrelationRegression(x, y):
    n = length of x
    sumX = sum of x
    sumY = sum of y
    sumXY = sum of x[i] * y[i]
    sumX2 = sum of x[i] ^ 2
    sumY2 = sum of y[i] ^ 2

    r = (n * sumXY - sumX * sumY) / sqrt((n * sumX2 - sumX ^ 2) * (n * sumY2 - sumY ^ 2))

    slope = (sumXY - sumX * sumY / n) / (sumX2 - sumX ^ 2 / n)
    intercept = (sumY - slope * sumX) / n

   Return {correlation: r, regression: {slope, intercept}}
```

The second server that we used is the express server. This is a node server. This server is used for database management which includes data retrieving, updating, deleting, sorting and grouping.

5.3.5 User Testing

After developing our application, we had to test to see how it will work on the real. This system was tested by MUST research department on their two surveys. The first one was name "FDS FOR CONTROL GROUP" AND "FGDS FOR TREATMENT GROUP". These were agriculture-based surveys. The application was installed in tablets so that they should be used. People to collect the data were selected the MUST research department.

5.3.6 Data Collection

When the app was installed in the tablets, people went to the field to collect the data using the app. The survey was a three-day survey. The collected data was saved in our database. When people finished collecting the data, we had to export the data so that is could be analyzed.

5.3.7 Data analysis

After collecting data, the data was analyzed. We used the system to generate visualizations such as graphs and charts.

5.3.8 User Feedback

In this stage we received the feedback from the people who were using the app. These feedbacks were both positive and negative feedback. The positive feedback was that the app was easy to navigate and to use, data collection was easy and straight forward. The negative feedback included that the user interface needed to improved, missing of some question data types such as roaster table.

5.3.9 System Refinement

Here, we gathered all the user feedback and redesigned the system to meet user requirements including improving user interface and also adding missing functionalities. This resulted in a new improved version of the system.

CHAPTER 6: Results and Discussion

In this research, we have managed to develop a survey system that is is capable of data collection, analysis, collaboration and data sharing. The mobile app is able to collect data and sync this data to the cloud database. The websites are able to:

1. create questionnaires
1. fill and submit responses and also sync data to the database
2. analyze the data
3. Share and commercialize the survey data

The following are the function of this system

6.1 Website

6.1.1 Questionnaire creation

This page is used to create new survey form or a questionnaire. This page three main sections which include title input section, question input method and a submission button. The title input field is where the survey title in written. The question input field is where the question is written. There is a field next to question field that allows the user to specify the question type. The question types include:

I. Multiple Choice (Multiple Select),
II. Number,
III. Short Text,
IV. Long Text,
V. Yes/No,
VI. coordinates,
VII. Rating Scale,
VIII. Likert Scale,
IX. Ranking,

X. Matrix (Single Select per Row),

XI. Matrix (Multiple Select per Row),

XII. Date Picker,

XIII. Time Picker,

XIV. File Upload,

XV. Numeric Input,

XVI. Slider,

XVII. Signature,

XVIII. Image Choice,

XIX. Audio Recording

This system allows the question to have follow up questions too. The generated questions can be edited and also deleted. After finishing adding the questions, then the finish button can be clicked to create the questionnaire the added question.

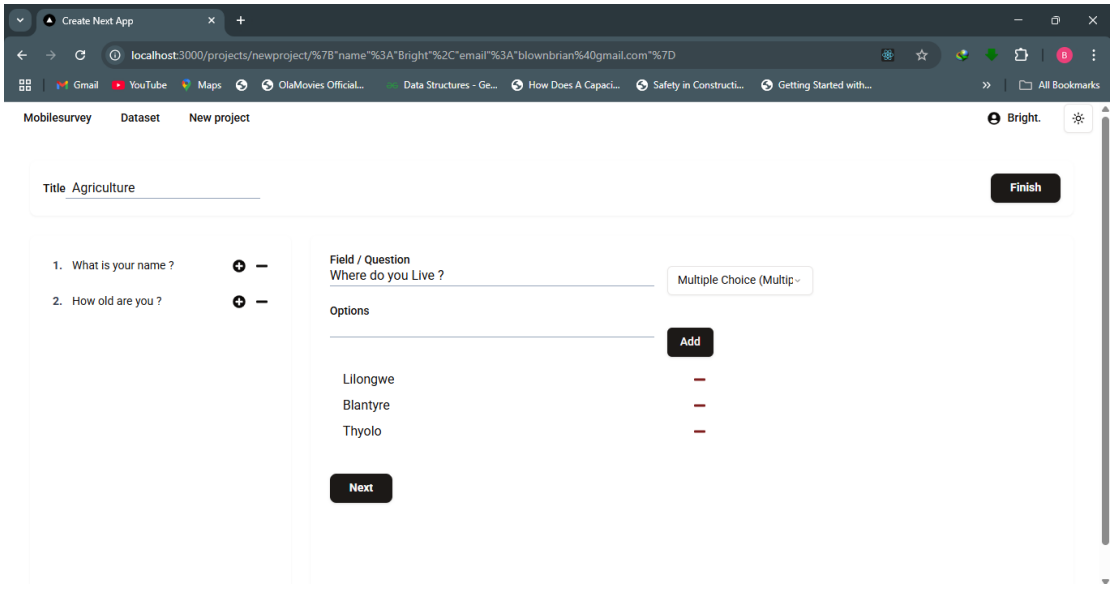


Figure 12: Questionnaire creation

6.1.2 Data Analysis (Dashboard)

This page displays data in a table format where each field is represented by a column. A field in the case of the questionnaire is a question where the answers are cells below the column. When we click the column or raw head, the descriptive statistics are shown on the page. This page user interface was designed to be similar with excel spread so that users can find this easy to use.

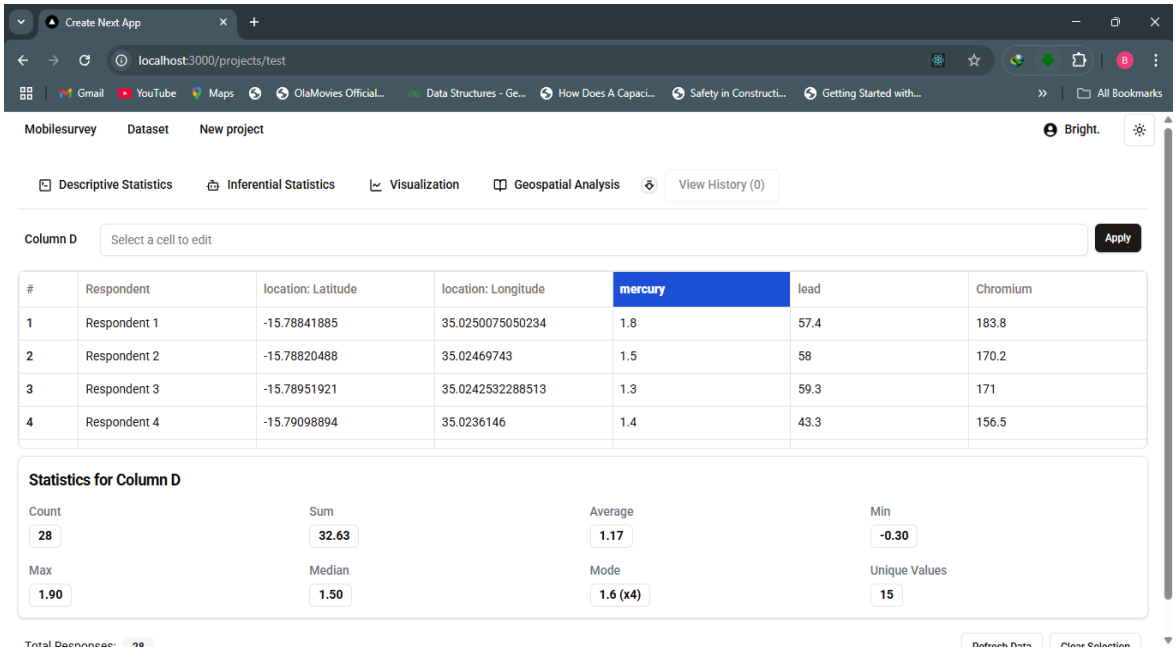


Figure 13: Data Analysis (Dashboard)

6.1.3 Data analysis: Visualization

This page show charts of the data. Firstly, the user has to select the columns he or she wants to visualize. This visualization shows graphs such as bar graph, line graph and pie chart. It is also possible to display to graphs on the same axis.

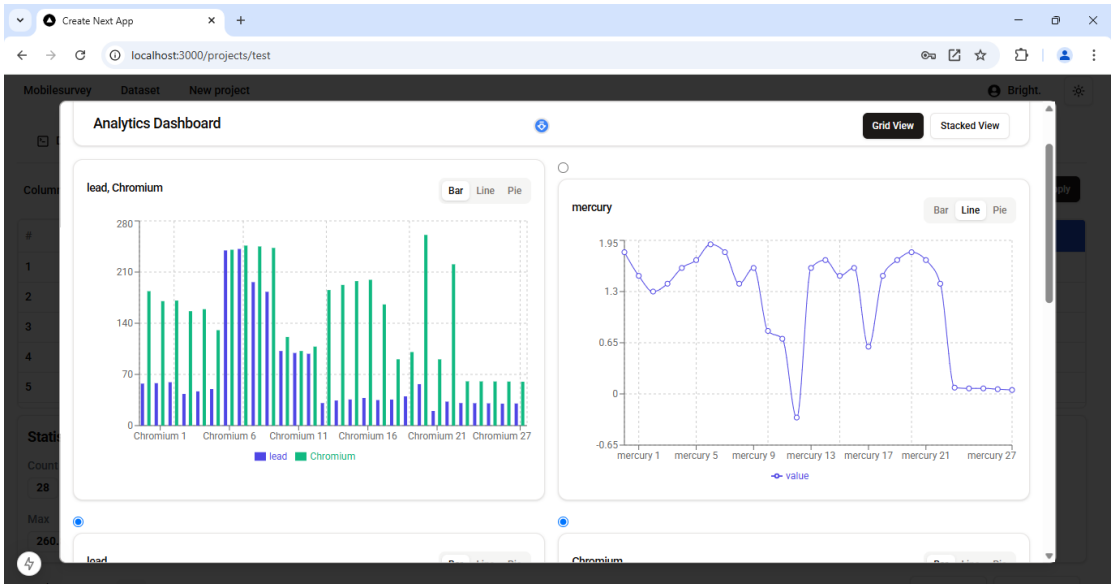


Figure 14: Data analysis: Visualization

6.1.4 Data analysis: Hypothesis Testing (Inferential statistics)

This page calculates hypothesis testing of the selected data variables. Hypothesis testing is a systematic procedure for deciding whether the results of a research study support a particular theory which applies to a population. A t test is a statistical test that is used to compare the means of two groups. It is often used in hypothesis testing to determine whether a process or treatment actually has an effect on the population of interest, or whether two groups are different from one another (Bevans, R. 2023).

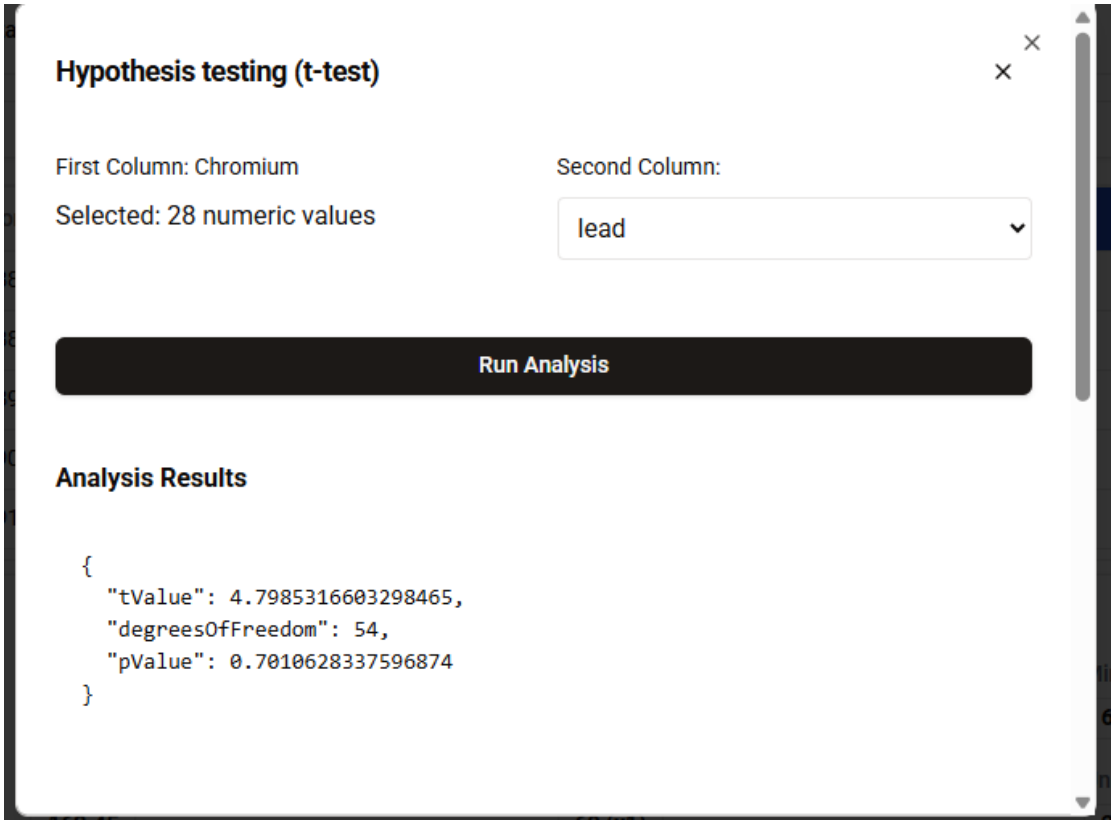


Figure 15: Data analysis: Hypothesis Testing (Inferential statistics)

6.1.5 Data analysis: Confidence interval (Inferential statistics)

This page calculates confidence interval of the data. The mean of your estimate plus and minus the variation in that estimate is known as a confidence interval. inside a certain degree of confidence, this is the range of values you anticipate your estimate will fall inside if you retake your test. In statistics, probability is often referred to as confidence (Bevans, R. 2023).

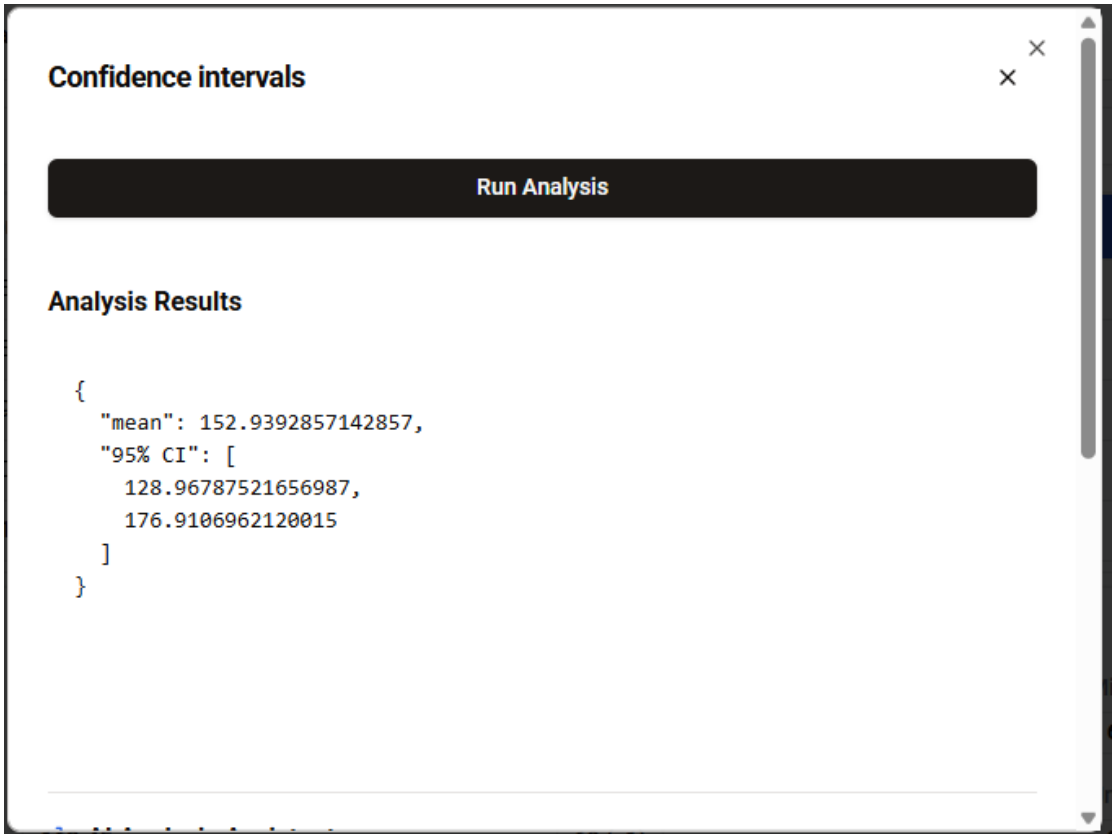


Figure 16: Data analysis: Confidence interval (Inferential statistics)

6.1.6 Data analysis: Correlation/Regression (Inferential Statistics)

This page calculates correlation and regression. In daily life, the term "correlation" is used to indicate some kind of link. We may say that we have observed a connection between wheeziness episodes and foggy days. Nonetheless, correlation is the statistical term used to indicate a relationship between two quantitative variables. Additionally, we assume that the relationship is linear, meaning that for every unit rise or reduction in one variable, the other variable grows or decreases by a constant amount. Regression analysis, which determines the optimal straight line to summarize the correlation, is another method that is frequently employed in these situations (Brown RA, 1093).

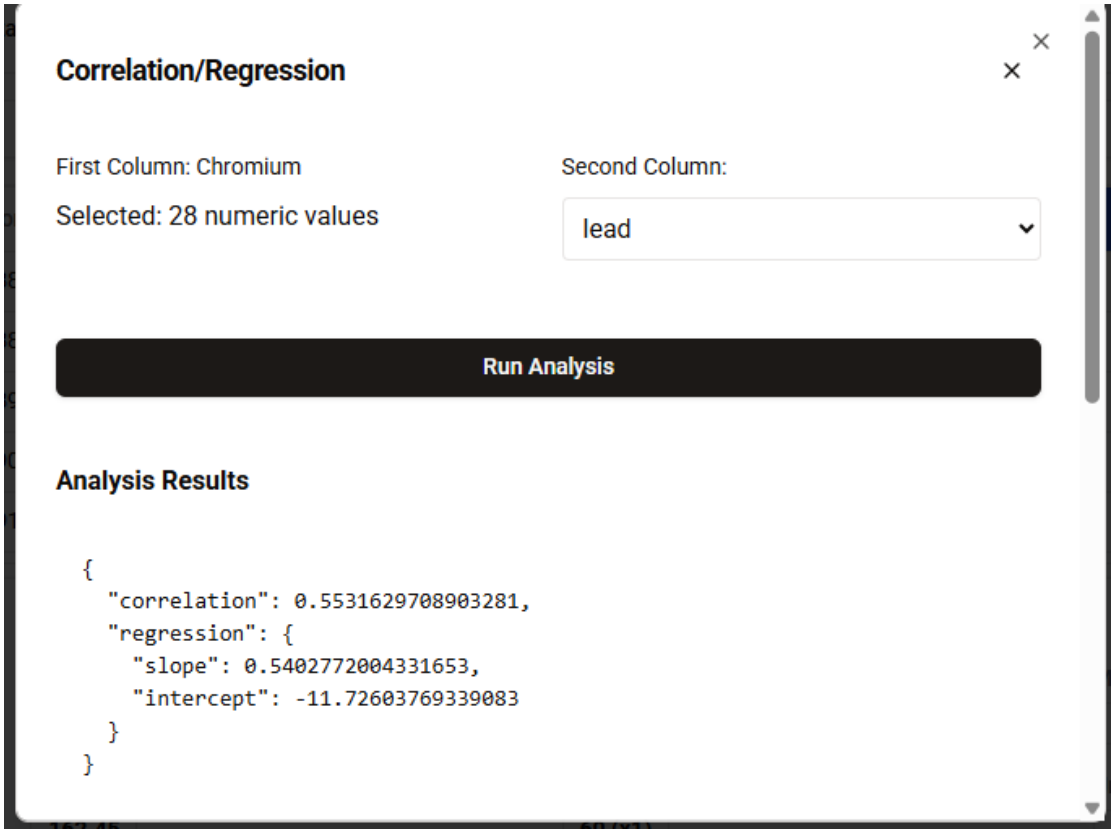


Figure 17: Data analysis: Correlation/Regression (Inferential Statistics)

6.1.7 Data analysis: IDW (Geo-spatial analysis)

This page calculates IDW of Geo-referenced data. Cell values are determined by inverse distance weighted (IDW) interpolation, which uses a linearly weighted combination of a collection of sample points. The inverse distance determines the weight. A location-dependent variable's surface should be the one being interpolated (Esri. (n.d.)).



Figure 18: Data analysis: IDW (Geo-spatial analysis)

6.1.8 Data analysis : IDW on map view(Geo-spatial analysis)

This system uses the map coordinates to calculate IDW interpolated values.

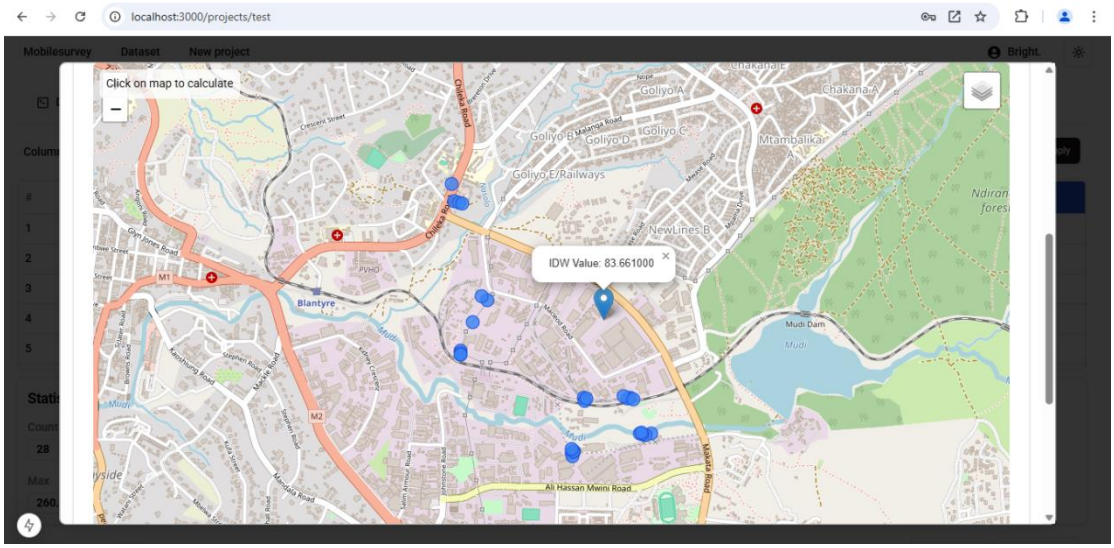


Figure 19: Data analysis : IDW on map view(Geo-spatial analysis)

6.1.9 Data analysis: Summary of responses (Descriptive Statistics)

This page calculates descriptive statistics of the selected column. These include count, average, min, max, mean and standard deviation.

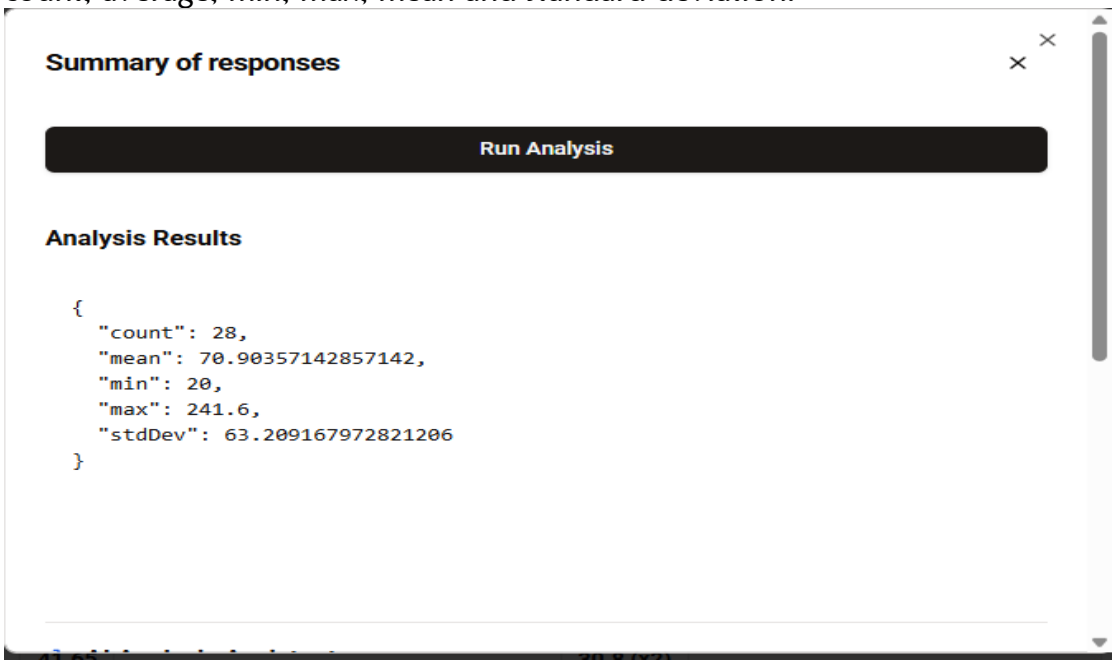


Figure 20: Data analysis: Summary of responses (Descriptive Statistics)

6.1.10 Data analysis: Frequency distribution (Descriptive analysis)

This page calculates frequency distribution. The number of observations for every potential value of a variable is described by a frequency distribution (Turney, S. 2023).



Figure 21: Data analysis: Frequency distribution (Descriptive analysis)

6.1.11 Data analysis: Percentage Breakdown (Descriptive statistics)

This page calculates percentage distribution of unique values of a certain field. This is displayed as a percentage.

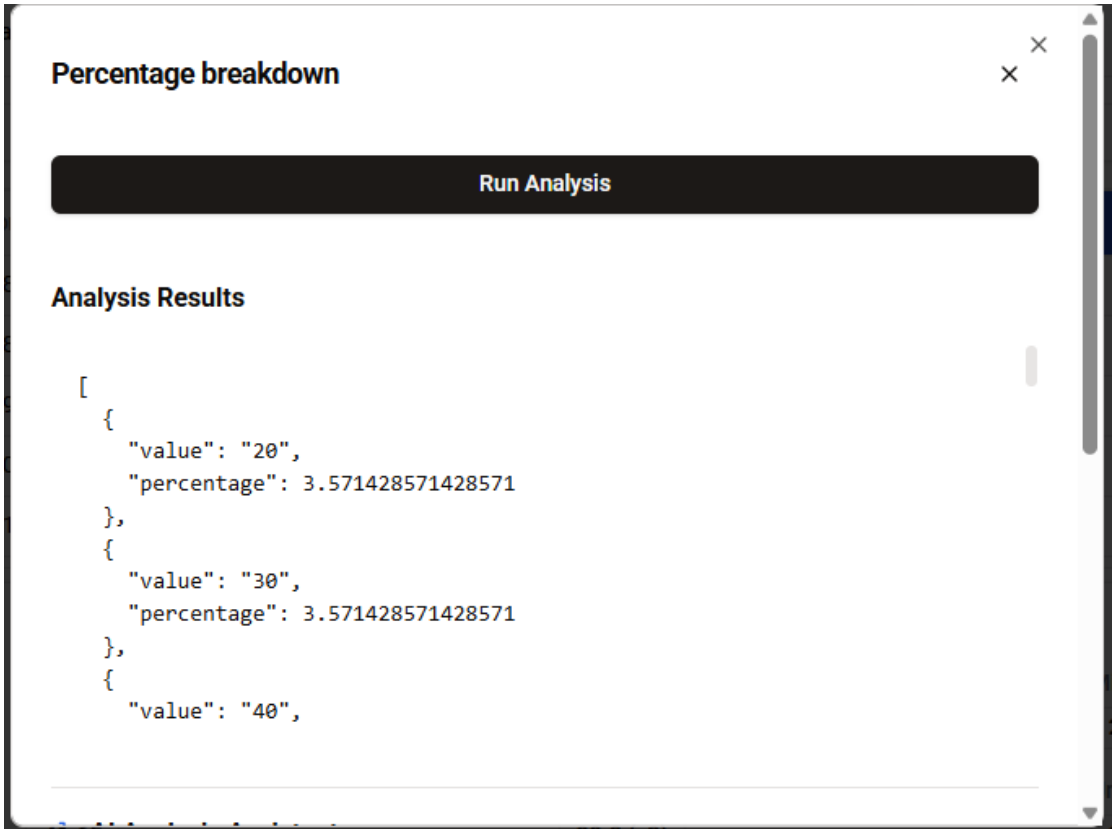


Figure 22: Data analysis: Percentage Breakdown (Descriptive statistics)

6.1.12 Data Sharing and Commercialization

This page is used for data sharing and commercialization. The survey data can be shared with others for free or at a price. This place also allows a user to share external datasets such as shapefiles, work sheets etc.

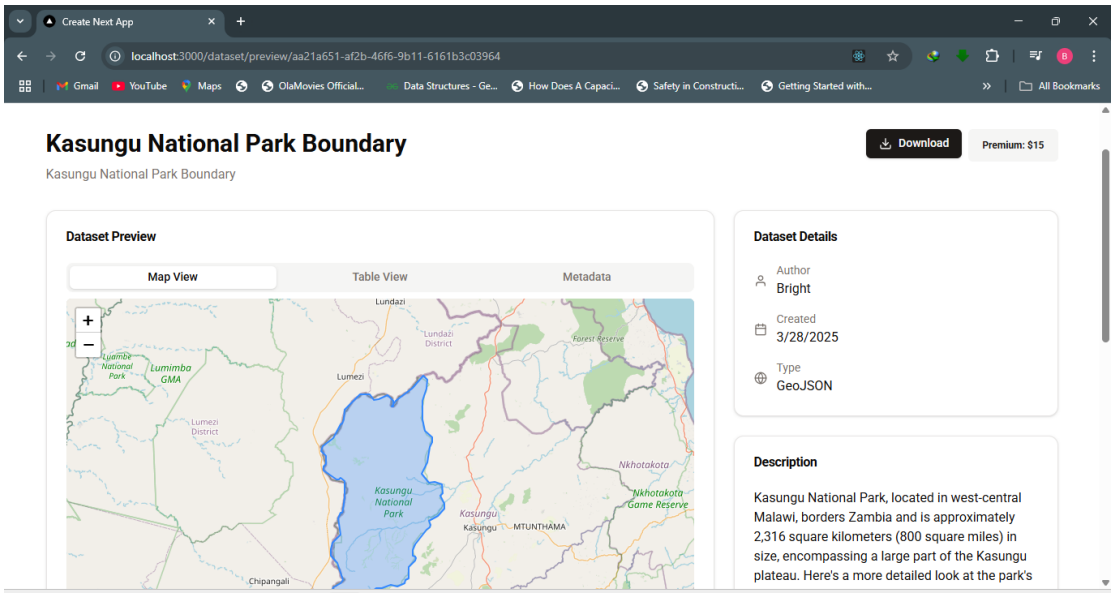


Figure 23: Data Sharing and Commercialization

6.1.13 Collaboration

This page provides collaboration functionalities. This is where the project admin adds users to participate and collaborate in the survey project. The project admin can also give admin privileges to other members.

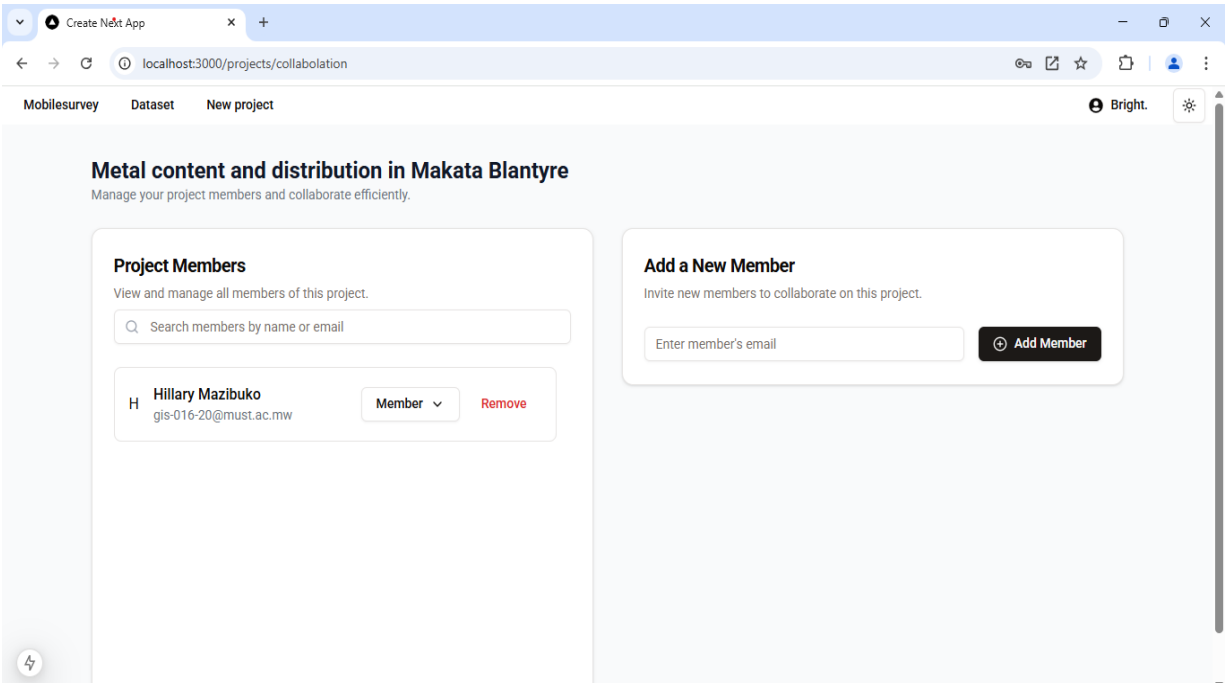


Figure 24: Collaboration

CHAPTER 7: Conclusions and Recommendations

7.1 Conclusions

The main purpose of this project was to develop a survey system that solves the challenges that exist in the current survey systems. This system introduces integrated, collaborative, AI-powered, and Geo-spatially enabled data collection and analysis platform. This project has successfully achieved its main and specific objectives and it has demonstrated promising results both in terms of technical innovation and practical impact.

Key findings of the project include:

1) Collaboration: This survey system allows people to work as teams. It also allows the project admin to grant privileges to other users over the survey project.

2) AI and Statistical Insights: By including the generative AI in traditional statistics, the system does not only analyze the data but also it explains statistical results, making data interpretation accessible to non-experts.

3) Geo-Spatial Analysis: The inclusion of geospatial analysis including mapping, kriging, and other techniques empowers the users to visualize and to make location-based decisions using the data.

4) Data Sharing and Monetization: The geo-portal allows the users to either sell or freely share their survey datasets thereby contributing to the growing trend of open data and data monetization.

5) Robust Architecture: The use of the modern tools such as React Native, Next.js, FastAPI, and Firebase ensures good performance, reliability, and scalability of both the mobile and web components of this system.

6) These outcomes demonstrate that the system can also enhance the data driven decision making in the academic research, government, NGOs, and the private sectors.

7.1.1 Limitations

Despite the system's successful development and deployment, some limitations were observed:

1) The user interface needed further refinement in order to enhance usability for the users with limited digital literacy.

2) Some of the advanced question types like the roster tables were not included in the initial version.

3) The testing sample size was limited and largely homogeneous thereby affecting the generalizing of the results.

4) Offline capabilities, while present, could be further optimized for areas with extremely limited connectivity.

7.2 Recommendations for Future Work

To build on this work, the following recommendations are proposed:

1. User Interface Improvements: The future iterations should include the design improvements guided by usability testing involving diverse users.

2. Expanded Question Types: by adding the advanced features like roster tables and matrix-type questions will make the system more comprehensive.
3. Localization and Multilingual Support: by introducing the support for the local languages will enhance the usability in diverse regions.

4. Broader Testing: The future research should involve a more diverse user base from the different regions, professions, and technical backgrounds in order to ensure the system's robustness and universality.

5. Machine Learning-Based Predictive Analytics: Incorporating machine learning models can further improve the insights by identifying the trends and predicting future responses or behaviors.

8. Ethics and Professionalism

In carrying out this research, it is of primary importance to abide by the highest standards of the ethical and the professional conduct. Plagiarism is one the example of ethical and professional misconduct in the scientific research. Plagiarism of any type and format is strictly prohibited. For clarity, plagiarism is defined as an attempt to pass off someone else's work as your own.

References

Abdelnour-Nocera, J., Dunckley, L., & Sharp, H. (2017). An approach to the evaluation of usefulness as a social construct using technological frames. International Journal of Human–Computer Interaction, 33(3), 221-237. https://doi.org/10.1080/10447318.2016.1232908

Android Developers. (2023). Testing your app. https://developer.android.com/studio/test

Apple Developer. (2023). Xcode testing workflow. https://developer.apple.com/testflight/

Bidwell, N. J., & Winschiers-Theophilus, H. (2015). At the intersection of indigenous and traditional knowledge and technology design. Informing Science Press.

Blandford, A., Furniss, D., & Makri, S. (2016). Qualitative HCI research: Going behind the scenes. Morgan & Claypool.

Bracht, G. H., & Glass, G. V. (1968). The external validity of experiments. American Educational Research Journal, 5(4), 437–474. https://doi.org/10.3102/00028312005004437

Brewer, M. B. (2000). Research design and issues of validity. In H. T. Reis & C. M. Judd (Eds.), Handbook of research methods in social and personality psychology (pp. 3-16). Cambridge University Press.

Brooks, F. P. (1995). The mythical man-month: Essays on software engineering. Addison-Wesley.

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. Houghton Mifflin.

Cook, T. D., & Campbell, D. T. (1979). Quasi-experimentation: Design & analysis issues for field settings. Rand McNally.

Couclelis, H. (2003). The certainty of uncertainty: GIS and the limits of geographic knowledge. Transactions in GIS, 7(2), 165-175. https://doi.org/10.1111/1467-9671.00142

Creswell, J. W. (2014). Research design: Qualitative, quantitative and mixed methods approaches (4th ed.). Sage.

Creswell, J. W., & Creswell, J. D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.). Sage.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319-340. https://doi.org/10.2307/249008

Dearden, A., & Rizvi, H. (2008). Participatory IT design and participatory development: A comparative review. Proceedings of PDC 2008.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). Internet, phone, mail, and mixed-mode surveys: The tailored design method (4th ed.). Wiley.

Díaz, J., Sales, A., Arellano, B., & Clemmensen, T. (2018). Cultural user experience with personas. Behaviour & Information Technology, 37(10-11), 1025-1036. https://doi.org/10.1080/0144929X.2018.1505950

Expo Team. (2023). Expo React Native Documentation. https://docs.expo.dev/

Express.js Team. (2023). Express.js documentation. https://expressjs.com/

FastAPI Team. (2023). FastAPI Documentation. https://fastapi.tiangolo.com/

Firebase Documentation. (2023). Email/password authentication. https://firebase.google.com/docs/auth

Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. GeoJournal, 69(4), 211-221. https://doi.org/10.1007/s10708-007-9111-y

Google Cloud. (2023a). Cloud Firestore documentation. https://cloud.google.com/firestore

Google Cloud. (2023b). Cloud Storage documentation. https://cloud.google.com/storage

Google Web Fundamentals. (2022). Rendering on the Web. https://web.dev/rendering-on-the-web/

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., ... Oliphant, T. E. (2020). Array programming with NumPy. Nature, 585(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Holmes, S., & Marx, M. (2021). Next.js in Action. Manning Publications.

Jordahl, K., Van den Bossche, J., Fleischmann, M., McBride, J., Wasserman, J., Gerard, J., Tratner, J., Perry, M., Badaracco, A. G., Farmer, C., Hjelle, G. A., Snow, A. D., Cochran, M., Gillies, S., Culbertson, L., Bartos, M.,

Eubank, N., Maxalbert, Albert, Bilogur, A., ... Ren, C. (2022). GeoPandas: Python tools for geographic data. https://geopandas.org

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2020). Toward a definition of mixed methods research. Journal of Mixed Methods Research, 1(2), 112-133. https://doi.org/10.1177/1558689806298224

Kirk, R. E. (2013). Experimental design: Procedures for the behavioral sciences (4th ed.). Sage.

Kjeldskov, J., & Graham, C. (2003). A review of mobile HCI research methods. In Proceedings of the 5th International Mobile HCI Conference (pp. 317-335). Springer. https://doi.org/10.1007/978-3-540-45233-1_23

Krug, S. (2014). Don't make me think, revisited: A common sense approach to web usability (3rd ed.). New Riders.

Kujala, S., Roto, V., Väänänen-Vainio-Mattila, K., Karapanos, E., & Sinnelä, A. (2018). UX Curve: A method for evaluating long-term user experience. Interacting with Computers, 23(5), 473-483. https://doi.org/10.1016/j.intcom.2011.06.005

Malczewski, J. (2006). GIS-based multicriteria decision analysis: A survey of the literature. International Journal of Geographical Information Science, 20(7), 703-726. https://doi.org/10.1080/13658810600661508

Marsden, N., Bangor, A., & Kortum, P. (2018). The impact of device type on UX measurement. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 62(1), 1244-1248. https://doi.org/10.1177/1541931218621283

Martin, R. C. (2008). Clean code: A handbook of agile software craftsmanship. Prentice Hall.

McConnell, S. (2004). Code complete: A practical handbook of software construction (2nd ed.). Microsoft Press.

McGibbney, L. J. (2021). Geospatial Python. Manning Publications.

McKinney, W. (2022). Python for Data Analysis (3rd ed.). O'Reilly Media.

Microsoft. (2023). Windows development environment setup. https://learn.microsoft.com/en-us/windows/dev-environment/

Mozilla Developer Network. (2023). Web APIs. https://developer.mozilla.org/en-US/docs/Web/API

Nielsen, J. (2012). Usability engineering. Morgan Kaufmann.

Norman, D. (2013). The design of everyday things (Revised ed.). Basic Books.

Polit, D. F., & Beck, C. T. (2017). Nursing research: Generating and assessing evidence for nursing practice (10th ed.). Wolters Kluwer.

Pressman, R. S. (2020). Software engineering: A practitioner's approach (9th ed.). McGraw-Hill.

Preece, J., Sharp, H., & Rogers, Y. (2015). Interaction design: Beyond human-computer interaction (4th ed.). Wiley.

Python Software Foundation. (2023). Python Language Reference. https://docs.python.org/3/reference/

React Documentation. (2023). React Components. https://react.dev/learn

Reinecke, K., & Bernstein, A. (2011). Improving performance, perceived usability, and aesthetics with culturally adaptive user interfaces. *ACM Transactions on Computer-Human Interaction, 18*(2), 1-29. https://doi.org/10.1145/1970378.1970382

Rogers, Y. (2016). HCI theory: Classical, modern, and contemporary. *Synthesis Lectures on Human-Centered Informatics, 5*(2), 1-129. https://doi.org/10.2200/S00736ED1V01Y201603HCI034

Rosenthal, R. (1976). Experimenter effects in behavioral research (Enlarged ed.). Irvington.

Rothwell, P. M. (2005). External validity of randomised controlled trials: "To whom do the results of this trial apply?". The Lancet, 365(9453), 82-93. https://doi.org/10.1016/S0140-6736(04)17670-8

Sambasivan, N., Batool, A., Ahmed, N., Cutrell, E., & Toyama, K. (2018). "They don't leave us alone anywhere we go": Gender and digital abuse in South Asia. CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. https://doi.org/10.1145/3173574.3173920

Sauro, J., & Lewis, J. R. (2016). Quantifying the user experience: Practical statistics for user research (2nd ed.). Morgan Kaufmann.

Seffah, A., Donyaee, M., Kline, R. B., & Padda, H. K. (2006). Usability measurement and metrics: A consolidated model. Software Quality Journal, 14, 159-178. https://doi.org/10.1007/s11219-006-7600-8

Sey, A., Coward, C., Bar, F., Sciadas, G., Koepke, L., & Rothschild, C. (2015). Connecting people for development: Why public access ICTs matter. Technology & Social Change Group, University of Washington.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin.

Sommerville, I. (2016). Software engineering (10th ed.). Pearson.

Trochim, W. M. K. (2020). Research methods: The essential knowledge base (2nd ed.). Cengage.

Van Rossum, G., & Drake, F. L. (2023). Python 3.12.0 documentation. Python Software Foundation. https://docs.python.org/3/

Vercel. (2023). Next.js Documentation. https://nextjs.org/docs

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2016). Unified theory of acceptance and use of technology: A synthesis and the road ahead. Journal of the Association for Information Systems, 17(5), 328-376. https://doi.org/10.17705/1jais.00428

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., ... SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. Nature Methods, 17, 261–272. https://doi.org/10.1038/s41592-019-0686-2

Vredenburg, K., Mao, J.-Y., Smith, P. W., & Carey, T. (2002). A survey of user-centered design practice. Proceedings of CHI 2002, 471-478. https://doi.org/10.1145/503376.503460

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). Unobtrusive measures: Nonreactive research in the social sciences. Rand McNally.

Wieruch, R. (2022). The Road to React. Leanpub.

Wilson, E. (2022). Node.js design patterns (3rd ed.). Packt Publishing.

Winschiers-Theophilus, H., Chivuno-Kuria, S., Kapuire, G. K., Bidwell, N. J., & Blake, E. (2012). Being participated: A community approach. Proceedings of PDC 2012.

Wyche, S. P., Schoenebeck, S. Y., & Forte, A. (2016). "Facebook is a luxury": An exploratory study of social media use in rural Kenya. *Proceedings of the ACM on Human-Computer Interaction, 1*(CSCW), 1-22. https://doi.org/10.1145/3134695

Zakas, N. (2022). Understanding Next.js: The React Framework for Production. Smashing Magazine.

Bevans, R. (2023, June 22). An Introduction to t Tests | Definitions, Formula and Examples. Scribbr. Retrieved April 28, 2025, from https://www.scribbr.com/statistics/t-test/

Bevans, R. (2023, June 22). Understanding Confidence Intervals | Easy Examples & Formulas. Scribbr. Retrieved April 28, 2025, from https://www.scribbr.com/statistics/confidence-interval/

Brown RA, Swanson-Beck J. Medical Statistics on Personal Computers , 2nd edn. London: BMJ Publishing Group, 1993

Esri. (n.d.). How IDW works. ArcGIS Pro Documentation. Retrieved April 28, 2025, from https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/how-idw-works.htm

Turney, S. (2023, June 21). Frequency Distribution | Tables, Types & Examples. Scribbr. Retrieved April 28, 2025, from https://www.scribbr.com/statistics/frequency-distributions/