Department of Information Technology
COLLEGE OF ENGINEERING
CENTRAL LUZON STATE UNIVERSITY
Science City of Muñoz, Nueva Ecija

AL1SE: DEEPSEEK-R1 AI INTEGRATED DATA VISUALIZATION SOFTWARE FOR THE DEPARTMENT OF STATISTICS OF CENTRAL LUZON STATE UNIVERSITY (CLSU)

A Capstone Project
Presented to the
Department of Information Technology

In Partial Fulfillment
of the Requirements for the Degree
BACHELOR OF SCIENCE IN INFORMATION TECHNOLOGY

By:

Bugtong, Franz R.
Juego, Camille D.
Pajaron, Jefferson R.

October 2025

DISCLAIMER

"This Capstone Project is submitted to the Department of Information Technology, College of Engineering, in partial fulfillment of the requirements for the degree, Bachelor of Science in Information Technology at the Central Luzon State University, Science City of Muñoz, Nueva Ecija. It is a product of the authors' own work except those that were indicated in the text. The project report or any portion thereof including the source code, or any section may be freely copied and distributed provided that the source is acknowledged."

APPROVAL SHEET

This capstone project proposal entitled **"ALISE: DEEPSEEK-R1 AI INTEGRATED DATA VISUALIZATION SOFTWARE FOR THE DEPARTMENT OF STATISTICS OF CENTRAL LUZON STATE UNIVERSITY (CLSU)"** prepared and submitted by **FRANZ REYES BUGTONG, CAMILLE DIEGO JUEGO, JEFFERSON RILLO PAJARON** in partial fulfillment of the requirements for the degree BACHELOR OF SCIENCE IN INFORMATION TECHNOLOGY, has been examined and is hereby endorsed.

DR. MARLON A. NAAGAS
Adviser

CAPSTONE 2 PROJECT ORAL PRESENTATION COMMITTEE

ASST. PROF MARYCRIS A. MANALANG
Chair
10-29-2025
Date Signed

MR. CENON CONRADO C. DIVINA
Member
10-16-2025
Date Signed

Accepted and approved in partial fulfillment of the requirements for the degree BACHELOR OF SCIENCE IN INFORMATION TECHNOLOGY, school year 2025-2026.

DR. ANJHLA C. TOLENTINO
Head, Department of Information Technology
10-30-2025
Date Signed

DR. ROY SEARCA JOSE P. DELA CRUZ
Acting Dean, College of Engineering
10-30-2025
Date Signed

CERTIFICATION OF MANUSCRIPT PROOFREADING

This is to certify that the undersigned has reviewed and edited the manuscript of FRANZ REYES BUGTONG, CAMILLE DIEGO JUEGO, JEFFERSON RILLO PAJARON, entitled, **"AL1SE: DEEPSEEK-R1 AI INTEGRATED DATA VISUALIZATION SOFTWARE FOR THE DEPARTMENT OF STATISTICS OF CENTRAL LUZON STATE UNIVERSITY (CLSU)."** This further certifies that the scope of editing is for proper English language, grammar, punctuation, spelling, and overall writing style.

Issued this 10th day of October 2025 at Central Luzon State University.

JENALYN B. PAGAY, PhD
Associate Professor V
Department of English and Humanities, CASS, CLSU

ACKNOWLEDGEMENT

The researchers would like to extend their heartfelt gratitude to the people and institutions who made the completion of this capstone project possible.

First and foremost, our deepest appreciation goes to our adviser, **Dr. Marlon A. Naagas**, for his constant support, encouragement, and guidance throughout the course of this project. His insights and expertise have greatly contributed to the successful development of our work.

We would also like to sincerely thank the **Department of Statistics**, especially **Asst. Prof. Mel Jean R. Estolano** and **Asst. Prof. Onneth O. Tejada**, for warmly welcoming us, generously giving their time whenever needed, and providing guidance in the development of our software. Their openness and support truly motivated us to push forward.

In addition, we would like to express our gratitude to our Capstone 1 coordinator and Chair Panel, **Asst. Prof. Marycris A. Manalang,** together with **Mr. Cenon Conrado C. Divina,** our Critic, for their valuable time, constructive feedback, and insightful recommendations that significantly improved the quality of this study.

We would also like to thank our **Capstone 2 coordinator, Mr. Ronnie Camilo F. Robles**, for guiding us throughout the process and patiently answering our inquiries. His assistance and dedication played an essential role in helping us successfully navigate this project.

To our parents, we are beyond grateful for the unwavering love, patience, and support they provided. Their encouragement kept us strong during the most challenging moments of this journey.

Lastly, we extend our gratitude to **ourselves as a team**, for persevering despite sleepless nights, weary minds, and exhausted bodies. Despite the challenges, we pushed through, complied with requirements, met deadlines, and gave our best until the end.

This project is not only a product of knowledge and skills but also of the collective effort, determination, and support of everyone mentioned above.

ABSTRACT

This capstone project developed AL1SE, a stand-alone, AI-integrated data visualization software tailored for the Department of Statistics. The system was designed to address the limitations of existing tools, such as limited visualization options, internet dependency, manual workarounds, and costly subscription fees. AL1SE was built using an iterative prototyping approach, enabling continuous refinement based on user feedback.

For evaluation, IT experts assessed the system using the ISO 25010 software quality model, which indicated that it performed acceptably. User testing was also conducted with faculty members and students through the original Technology Acceptance Model (TAM), and results likewise reflected acceptable performance in terms of usefulness and ease of use.

Overall, AL1SE provides faculty members with an offline, cost-free alternative that supports interactive and specialized visualizations, aligning with the university's commitment to accessible, high-quality education.

TABLE OF CONTENTS

LIST OF FIGURES

Figure No. **Title** **Page**

LIST OF TABLES

Table No. Title Page

CHAPTER I

INTRODUCTION

Data visualization tools play an important role in data analysis; they help in transforming raw data into more comprehensive visuals. With such notion, the faculty members of the Department of Statistics, College of Science are in need of extra tool for data analysis that will fill in the gap of existing alternatives that they are currently using. There are current freeware solutions available in the said department for data visualization but have limitations when it comes to specialized data visuals like 3D scatterplot, which requires manual programming and lot of external dependencies to implement. Moreover, subscription-based software solutions with advanced features are too expensive to utilize.

To address these limitations, the development of AL1SE, a data visualization software alternative with DeepSeek-R1 AI integration, is hereby proposed. This is a stand-alone tool to be designed specifically for the said department to empower the faculty members with features that are other free alternatives are lacking such as optional data cleaning, automated AI report based on summary statistic, AI plot recommendation based on uploaded data set, 3D and interactive data visuals, that do not require an internet connection, subscription fees, and manual programming or numerous of external dependencies to operate.

By ensuring cost free distribution, and offline accessibility, the project upholds Central Luzon State University's (CLSU) commitment to inclusive and high-quality education for all, aligning with Sustainable Development Goal 4 (SDG 4) to promote equal access to educational resources, and lifelong learning opportunities (Goal 4 | Department of Economic and Social Affairs, n.d.).

PROBLEM STATEMENT

General Problem

In the Department of Statistics, existing data visualization solutions are either limited to basic visualization features or require paid subscription fees, and stable internet connection to function, hence restricting their usability and accessibility. Moreover, specialized visuals require manual programming of R or Python, and other workarounds to implement, which limit faculty members to perform in-depth analysis, and rely on the existing alternatives they currently have.

Specific Problem

1. **Limited Visualization Options**– Existing freeware data visualization alternatives only offer basic, non-interactive graphs, which are hard to use in in-depth analysis with complex data sets.
2. **Internet Dependency** – Many available alternatives require stable internet connection to function, limiting their usability in offline environment.
3. **Costly Subscription Fees** – Existing data visualization software alternatives requires costly subscription fees, which are hard to utilize.
4. **Require Manual Programming** – To utilize specialized visuals in existing alternatives, faculty members must manually setup and write code in R or Python, which can be time consuming, and may require technical expertise.

OBJECTIVES

General Objectives

Generally, this project aims to develop AL1SE, an AI integrated, stand-alone data visualization software designed specifically for the Department of Statistics. This envision to provide a data visualization alternative for the faculty members that will be accessible anytime without the need for internet connection or paid subscription fees while having some of the data visualization specialized features like the other existing alternatives offer, without the intent to replace current tools, but to provide an additional option to enhance data visualization capabilities.

Specific Objectives

1. Provide a data visualization alternative for the Department of Statistics to support specialized and interactive visuals.
2. Develop AL1SE as a stand-alone software that can be accessed and used any time without requiring an internet connection.
3. Distribute AL1SE free of charge, ensuring its availability without financial constraints.
4. Eliminate the need for manual programming to access specialized and interactive data visualization by ensuring AL1SE is ready for use immediately after installation.

SCOPE AND LIMITATIONS

Scope

The proposed system, AL1SE, will focus on:

- **The Needs of the Department of Statistics:** The sole client of this project are the faculty members of the Department of Statistics. The system features will be tailored based on their specific workflow and requirements.
- **Interactive Data Visualization:** The system will focus on providing specialized and interactive data visuals for better analysis.
- **List of Visuals:** Scatterplot**,** Bar Chart, Line Graph, Pie Chart**,** Histogram, Box Plot**,** Violin Plot**,** Heatmap**,** Sankey Diagram**,** 3D Scatter Plot**,** 3D Cluster Plot**,** Contour Plot**,** Surface Plot**,** Scatter Matrix**,** Parallel Coordinates.
- **AI integration:** DeepSeek-R1 will be integrated for support, specifically for report generation and Ai plot recommendation.
- **Standalone and Offline:** AL1SE will be developed as a standalone software that will be used in an offline environment.
- **Specific Options in Data Cleaning:**
    - **Data Validation:** Check Null Values
    - **Data Cleaning:** Mode Imputation**,** Mean Imputation**,** Drop Null Values**,** Set Placeholder
    - **Formatting:** Text Formatting, Number Formatting, Date Formatting
- **Descriptive Statistics:**
    - **Measures of Central Tendency:** Mean, Median, Mode

- **Measures of Position/Percentiles:** Min, Max, Q1 (First Quartile/25th percentile), Q2 (Second Quartile/50th percentile - same as median), Q3 (Third Quartile/75th percentile)
- **Measures of Dispersion/Variability:** Range, Variance, Standard Deviation, IQR (Interquartile Range)
- **Measures of Shape:** Skewness (measures asymmetry of distribution), Kurtosis (measures tail heaviness/peakedness of distribution)
- **Basic Count:** Count (frequency/sample size)
- **Distribution Analysis:** Frequency Distribution (shows how often each value or range of values occurs in the dataset)

- **Correlation Analysis/Bivariate Statistics:**
  - Pearson Correlation (measures linear relationship strength)
  - Spearman Correlation (measures monotonic relationship strength, rank-based**)**
  -

Limitations

Despite AL1SE's capabilities, it will have the following limitations:

- **AI Reliability:** AI can make mistakes. It will be integrated for automation and support as a *co-pilot*, but analytical and statistical knowledge is still required.
- **Limited to Descriptive Analytics and Correlation Analysis:** The proposed system will be limited to descriptive analytics and correlation analysis specified on the scope.

- **Requires Statistics Knowledge:** The software will require statistical knowledge since some terms can only be understood if the user has prior knowledge in statistical terms.
- **Limited to Windows OS:** AL1SE will only support Windows OS deployment as per the Department of Statistics' specification.
- **Geospatial Plots:** This kind of plots will be out of our scope.
- **Processing Speed:** The DeepSeek/ AI features processing speed will depend on the computer hardware where AL1SE will be installed.

SIGNIFICANCE OF THE STUDY

This project addresses the gap of the existing data visualization alternatives:

- **Department of Statistics and Faculty Members**: This envisions to empower the faculty members of the Department of Statistics by providing them a standalone, offline, data visualization software with optional data cleaning, AI integration for report automation, which can be used without needing an internet connection or subscription fees, while also authorizing them to distribute the software to their students, and others under their discretion, for free.
- **Researchers**: Through this project, students of the Bachelor of Science and Information Technology can be provided with ideas for future studies related to data visualization, AI integration in a stand-alone tool, and software development.
- **Future IT Projects**: This could serve as a reference for IT students who are interested in creating similar tools like the proposed system, AL1SE.

DEFINITION OF TERMS

To ensure clarity and consistency, the following terms are defined:

1. **AL1SE:** The proposed data visualization software that will be developed in this study.
2. **DeepSeek-R1:** The AI reasoning model that will be integrated to AL1SE for report automation, and AI plot recommendations.
3. **Data Visualization**: The main feature of AL1SE, which is transforming data into graphical representation.
4. **Descriptive Analytics**: The primary type of analytics that we will integrate on AL1SE.
5. **Correlation Analysis:** Statistical method we will integrate on AL1SE, specifically Pearson and Spearman
6. **Freeware**: Refers to the free tools currently used by the Statistics Department.
7. **Subscription Based Software Solutions**: Refers to the paid version of the free tools Statistics Department are currently using.
8. **Specialized Data Visuals:** Refers to the interactive types of visualizations included in AL1SE, such as 3D scatter plot.
9. **Department of Statistics**: A department under the College of Science, Central Luzon State University (CLSU).
10. **Standalone Software**: How AL1SE will operate, without external dependencies, and without needing an internet connection.

CHAPTER II

REVIEW OF RELATED LITERATURE AND EXISTING ALTERNATIVES

INTRODUCTION

There are many data visualization tools available today in the industry. However, many lack features for accessible specialized data visualizations. Therefore, this study will use thematic organization to analyze existing studies and alternatives related to the proposed data visualization software, AL1SE, to address the gap, emphasize the importance of an accessible data visualization tool in data analysis, and to effectively tailor this proposed system with the involvement of faculty members of the Department of Statistics.

LOCAL AND FOREIGN LITERATURES

Importance of Interactivity in Visual Tools

Existing data visualization tools available in the Department of Statistics are limited to basic, non-interactive visuals, which can be challenging to use in understanding complex datasets. Steed (2025) highlighted the importance of interactive data visualization tools in data analysis that bring interactive and comprehensive results especially in complex datasets. Shin (2024), on the other hand, pointed out the persistent challenges in data visualization tools, that despite advancements of visualization tools, issues still persist when it comes to design and interpretation, since even experts often struggle to create effective visualizations. These issues highlight the relevance of integrating AI into data visualization tools for support and automation while also having interactive visual options.

Rationale for Choosing DeepSeek-R1 as AI Model

DeepSeek as the AI model was chosen primarily because it is completely free to

use, making it a practical and cost-effective AI model for educational use. As highlighted in the study by Ariyo Okaiyeto et al. (2025). DeepSeek R1 gained rapid adoption due to its open accessibility, being both free and powerful helps democratize AI by making advanced tools available even in resource limited environment. This aligns with the goal of developing AL1SE as an offline, AI-integrated data visualization software. Similarly, Wu (2025) emphasizes that DeepSeek promotes equal access to learning through AI-driven platforms, making it an ideal choice for tools like AL1SE. Its adaptability, open access, and proven performance in academic settings support our goal of providing a cost-efficient AI solution without relying on paid services or internet connectivity. A comparative study by Gao et al. (2025) compared DeepSeek against Claude, Gemini, GPT, and LLaMA across tasks like authorship and citation classification, found that DeepSeek outperformed GPT, Gemini, and LLaMA in most classification cases, while offering a more affordable and accessible alternative to Claude. The study also highlighted DeepSeek's similarity in output quality with high-end models like Claude and Gemini, supporting its effectiveness in generating accurate and readable AI-generated reports. These findings reinforce the developers decision to integrate DeepSeek-R1 in an offline AI-powered visualization tool, tailored for educational use. Moreover, even though DeepSeek is free, it still a competitive AI model that can go against with other existing paid AI models. (Hayder, 2025) highlighted DeepSeek-R1's advancements in AI, emphasizing its human-like reasoning, and work efficiency. Unlike traditional models, it avoids extra training, and is known for its adaptability and performance. DeepSeek-R1 is an ideal AI model to integrate into the proposed offline data visualization software to enhance our system, as stated in (*China Connects Everything to DeepSeek in Nationwide Plan - Asia Times*, n.d.). Yong (2025)

reported that DeepSeek-R1 is being integrated in various platforms including smart applications and major tech companies. The report states that DeepSeek-R1 surpassed ChatGPT in terms of popularity, becoming the No. 1 free app on the US download charts within days. While the article approach focuses on cloud-based integration, our proposed stand-alone data visualization software takes a different direction by integrating DeepSeek-R1 into an offline environment.

Stand-alone Data Visualization Software

Most existing freeware alternatives are cloud-based, and while some offline options exist, they often lack specialized and interactive visuals. Developing a standalone data visualization alternative that can be accessed offline would be a valuable solution for the Department of Statistics. (Poddar & Poddar, 2020) emphasized the advantages of using a stand-alone data visualization tool, and how beneficial it is to analyze and interpret data through interactive visualizations without requiring a constant internet connection, that this approach not only reduces data costs but also enhances accessibility. Despite the goodness of a stand-alone software, there is still a gap in making it less complicated, so that it can be used after installation, without needing manual setup or workaround. Caponpon et al. (2022) developed a market trend analysis for local RTW shops using data visualization. Their system, which relies on PHPMyAdmin in XAMPP, allows business owners to analyze sales trends but requires a local server for operation. This gap of needing a local server to operate will be addressed in this proposed software as AL1SE, as it does not need numerous of external dependencies to install.

EXISTING ALTERNATIVES

There are a lot of well-known alternatives for data analytics such as Power BI Free and Power BI pro, the most well-known data visualization and business intelligence tools developed by Microsoft (*Power BI - Data Visualization | Microsoft Power Platform*, n.d.), which provide interactive dashboards and robust analytics. Given that these are good alternatives, the free versions are only effective in basic charts as they lack specialized chart types out-of-the-box, and AI is not available, which limits their suitability for complex statistical visualizations. While the pro version offers advanced features, it requires a costly subscription fee to be utilize. Moreover, both versions of the software require manual programming of R or Python to be able to access the other data visuals same with Tableau public and Tableau desktop (*Business Intelligence and Analytics Software - Tableau US*, n.d.) which are widely used data visualization software known for its powerful and interactive visual analytic tools. Tableau public is the free version, it is cloud-based meaning users must upload their data online which requires internet connection to do, it stores data publicly which concerns privacy for sensitive data. Tableau desktop is the pro version, offering specialized analytics and interactive dashboards, along with offline support, making it powerful alternative. However, like any other pro version alternatives, it is not a practical option due to its expensive subscription fees, making it less accessible for people who prefer to use a free one. This alternative can be accessed offline but sharing and publishing need an internet connection, it also has specialized and interactive visuals, which still require manual setup. Statistical Package for the Social Sciences (SPSS) is another alternative tool that is widely used in academic and research, (*IBM SPSS Statistics*, n.d.)while it has point and click interface, it is not focused in data

visualization capabilities, as it lacks flexibility in customizing interactive visuals, and it also requires subscription fees to utilize. Additionally, Rapid Miner is a data science platform designed for advanced analytics (*Data Analytics and AI Platform | Altair RapidMiner*, n.d.), including text mining, and predictive analysis. It offers some visualizations that are out-of-the-box like AL1SE, and its strengths lie in automated data processing workflows rather than comprehensive, customizable data visuals. The free version of Rapid Miner is limited in usage, making it less suitable for unrestricted use. Additionally, there's a decision analytic software like Analytica (*Analytica - Visionary Modeling*, n.d.), which offers strong capabilities for quantitative analysis and decision making, but just like the other existing alternatives it lacks interactive data visualizations, and it needs paid subscription as free version comes with limitations. While costly subscription fees are gaps in most data visualization alternatives, it isn't in RAWGraphs, it is an open-source tool that provides customizable visualizations and offline functionality (*RAWGraphs*, n.d.), allowing users to create a variety of visual analytics. Despite its flexibility, it requires configuration and programming knowledge, making it challenging for users without technical expertise. Moreover, it lacks data visuals options, same with Jamovi. Jamovi is a user-friendly, open-source statistical software designed for data analysis and visualization (*Jamovi - Open Statistical Software for the Desktop and Cloud*, n.d.). Jamovi does not require internet connection but it lacks automation, has limited customization, and does not support interactive visualizations.

GAPS IN EXISTING RESEARCH

Previous studies and existing data visualization tools provided significant advancements in data visualizations; however, several limitations remain unaddressed,

highlighting the need for this study. Existing alternatives mentioned in this study such as Power BI and Tableau requires paid subscription or cloud-based storage, restricting accessibility for researchers with financial limitations and unreliable internet connection. Some existing alternatives only offer basic visuals, while specialized visuals usually need to be manually programed or configured. Open-source alternatives like RawGraphs and Jamovi, providing offline functionality but lacks specialized visuals.

While existing studies highlight AI's potential in data visualization, current tools do not fully integrate AI for automation in offline settings, most AI integration can be found in cloud-based application which requires internet connection to function. DeepSeek-R1 has been recognized for its specialized capabilities, but it is mostly used in a cloud-based applications, leaving a gap for stand-alone software. Additionally, AI integration typically comes with costly monthly payments to keep AI working, even then, an online AI integration with a paid subscription still comes with limited uses. Similarly, research on offline visualization tools acknowledge the benefits of using AI in an offline setting, yet existing solutions still lack summary statistics, and AI powered report generation.

This proposed study will address the mentioned limitations by developing AL1SE, a free, offline, stand-alone data visualization software with built in AI for automation. The proposed software will support the Department of Statistics in conducting their research and continue data exploration in a more efficient and accessible way.

Table 1 presents a summary of related research studies, highlighting their key findings, existing gaps or limitations, and how the proposed data visualization software, AL1SE, addresses these challenges.

Table 1. Gaps in Existing Research

Research Title	Key Findings	Gaps/Limitations	Current Project Solutions to these gaps
Success of DeepSeek and potential benefits of free access to AI for global-scale use	Underscored the practicality, hardware efficiency, and widely adoptability of DeepSeek as an AI Model	Lacks studies on integrating DeepSeek in an offline setting, and in data visualization tools.	DeepSeek will be utilized in an offline setting, eliminating financial barriers.
Interactive Data Visualization	Found the importance interactive data visualization for better analysis.	Lacks studies on the different interactive data visuals that can be used for better analysis.	The proposed data visualization tools will have interactive visuals like 3D heatmap.
Simulation, Representation, and Automation: Human-	Found that data visualization continues to face	Lacks studies on how AI can be integrated into offline data	AL1SE will have AI integration to enhance

Centered Artificial Intelligence for Augmenting Visualization Design	challenges in interpretation even among experts, how AI integration can enhance automation if implemented.	visualization tools.	automation despite of it being offline.
DeepSeek: Toward Global Education Empowerment for the Whole Society	Highlighted DeepSeek's role in democratizing AI access and supporting educational equity through open access models	Limited exploration of DeepSeek in offline, non-internet-dependent tools	AL1SE will integrate DeepSeek in a fully offline environment to ensure accessibility without internet dependence
A Comparison of DeepSeek and Other LLMs	Illustrated how DeepSeek outperformed GPT, Gemini, and LLaMA in most classification tasks; demonstrated	Few studies explore the application of DeepSeek in offline visualization or educational tools.	AL1SE will use DeepSeek to generate AI-powered reports and insights, offline.

	comparable output quality to Claude.		
China Connects Everything to DeepSeek in Nationwide Plan	Mentioned that China is pushing nationwide initiative to integrate DeepSeek in various application to improve efficiency.	Contradictory findings on whether AI integration improves accessibility and efficiency in offline environment.	DeepSeek will be used in an offline environment through integrating it to our proposed software.
Highlighting DeepSeek-R1: Architecture, Features and Future Implications	Showed that DeepSeek-R1 enhances AI driven automation with optimized architecture for diverse applications.	Limited studies on its offline deployment and real-world scalability in constrained environments.	AL1SE will test DeepSeek-R1 efficiency by integrating it in offline setting, where it doesn't rely on cloud resources.
Covid-19 Data Visualization and Data Analytics with a Smart Standalone	Demonstrated the benefits of stand-alone data visualization tools	The study did not explore AI integration for automation.	AL1SE will enhance stand-alone data visualization by

Mobile Application	in reducing costs and enhancing accessibility.		integrating AI for automation.
Market Trend Analysis On Local RTW Shop In Lucena City With Data Visualization	Developed a market trend analysis for local business using PHPMyAdmin and XAMPP.	Requires local server and manual workarounds to setup.	AL1SE eliminates the need for local server, and manual setup.
Interactive Data Visualization using Mondrian	Introduced a stand-alone java-based data visualization tool.	Requires java installation, external database setup, and manual configuration.	AL1SE eliminates external setup, as it can be use after installation.
Power BI Free	Provided interactive dashboards and analytics.	Requires internet connection, and specialized visuals need to be manually programmed.	AL1SE will operate offline, specialized visuals do not need to be programmed.
Power BI Pro	Offered collaborative analytics and	Requires paid subscription and internet connection,	AL1SE can be distributed freely without fees,

	specialized visualization features.	specialized visuals also require manual programming.	specialized visuals do not need to be manually programmed.
Tableau Public	Enabled users to create and share dashboards for free.	Requires cloud storages, and needs manual configuration is needed for specialized visuals.	AL1SE will not require internet connection or manual programming for specialized visuals.
Tableau Desktop	Supported offline functionality and specialized analytics	Expensive licensing fees, requires manual workload for specialized data visuals, and sharing needs internet.	AL1SE will be free, as well as AI support, and requires no internet connection for full functionality
SPSS	Emphasized how SPPS is widely used in academia, user	Limited interactive visualization, and requires subscription	AL1SE will offer interactive data visuals without

	friendly, point and click interface.	fees.	needing subscription fees to utilize.
Rapid Miner	Established the strength of automated workflows and advanced analytics	Limited flexibility in visual customization, free version has usage limits.	AL1SE will provide customizable visuals and can be use any time, unlimited.
Analytica	Emphasized the effective quantitative and decision analysis capabilities	Lacks interactive data visualization, and paid subscription is required.	AL1SE will have interactive visuals without needing subscription fees.
RAWGraphs	Mentioned that the RWAGraph is an open-source and allows customizable offline visualization	Requires programming knowledge, and lacks basic summary statistics	AL1SE will not require programming skill for it to be use.
Jamovi	Illustrated that Jamovi is an open-source statistical	Lacks interactive visuals	AL1SE will have interactive data visuals.

	software that can be used in an offline environment.		

CONCLUSION

The review of existing studies and alternatives reveal that despite the advancement in analytics tools, there are still gaps in providing accessible, offline capable, AI for automation of reports. Existing alternatives often require subscription, manual setup, and lacks interactive data visuals, limiting their accessibility and usability. Thus, this study aims to address these gaps by developing AL1SE, a standalone data visualization software with specialized visualizations, with client-specified descriptive analytics will serve as an extra tool for the Department of Statistics to continue their data exploration.

CHAPTER III

METHODOLOGY

RESEARCH DESIGN

This study will use developmental research design to guide the development of AL1SE, a proposed data visualization software alternative. This study will focus on understanding user needs, and refining system features through continuous feedback, and iterative development. During the requirements gathering phase, a semi-structured interview will be conducted to the Department of Statistics faculty members. This approach will help the researchers to identify limitations of existing visualization tools, and determine essential features for AL1SE.

In the system design phase, this study will use evolutionary prototyping model, enabling continuous refinements based on user feedback. Faculty members will interact with the early version of AL1SE, for them to provide insights that will shape iterations until all necessary features are incorporated to the system, and no further refinements are needed.

To ensure alignment with user expectations and software quality standard , IT expert testing will be conducted before deployment, following the ISO 25010 standard. User testing and evaluation will be performed after the deployment period using the original Technology Acceptance Model (TAM). The researchers will utilize Likert Scale for both the IT Expert testing and the User testing. In IT testing, an open-ended question will be added to get the IT Experts comments and suggestion for future development.

SYSTEM ARCHITECTURE/CONCEPTUAL FRAMEWORK

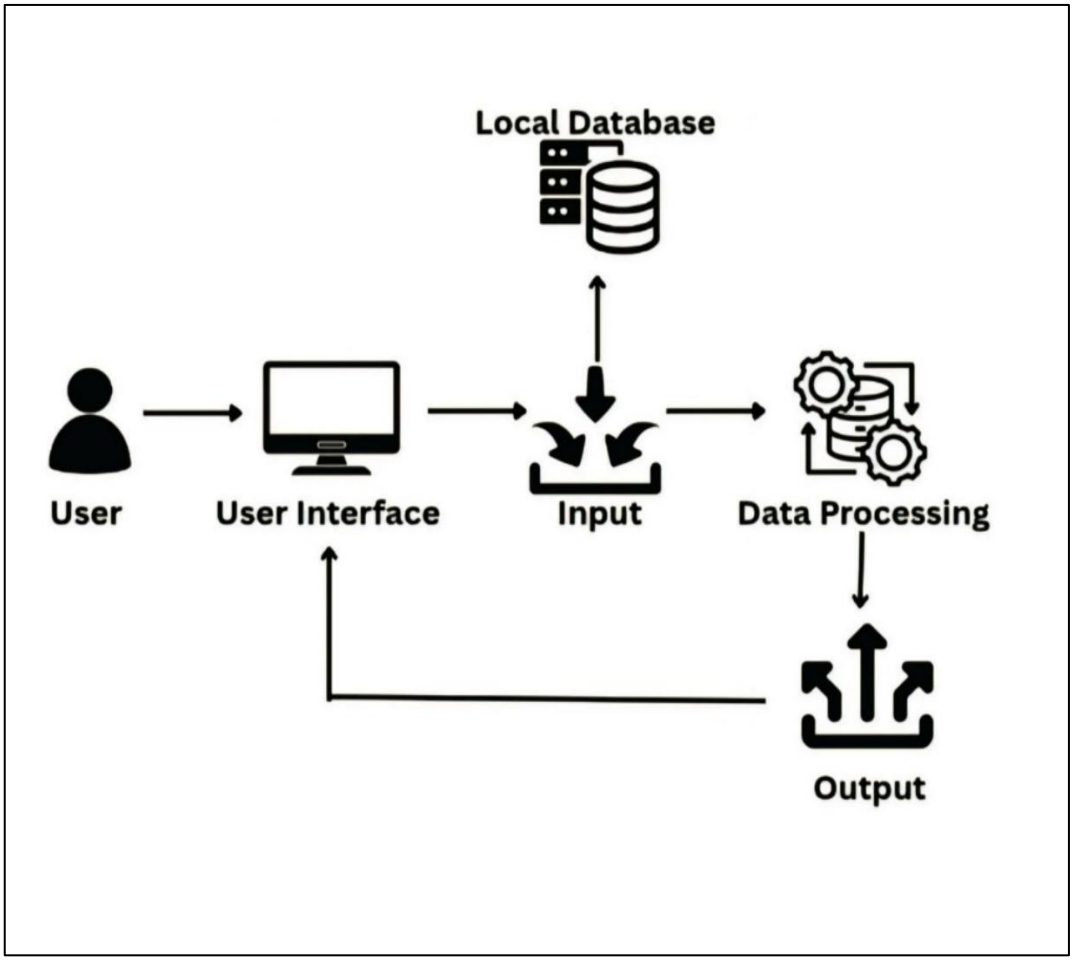


Figure 1. System Architecture of the Proposed Software

Figure 1 shows the system architecture of the proposed data visualization software, AL1SE. The user can access the system through a graphical user interface (GUI) on their computer, while data will be inputted through the interface, and will be stored in a local database. The system will process the data and will generate outputs like data visualization, which will be displayed back to the user, all in an offline, stand-alone environment.

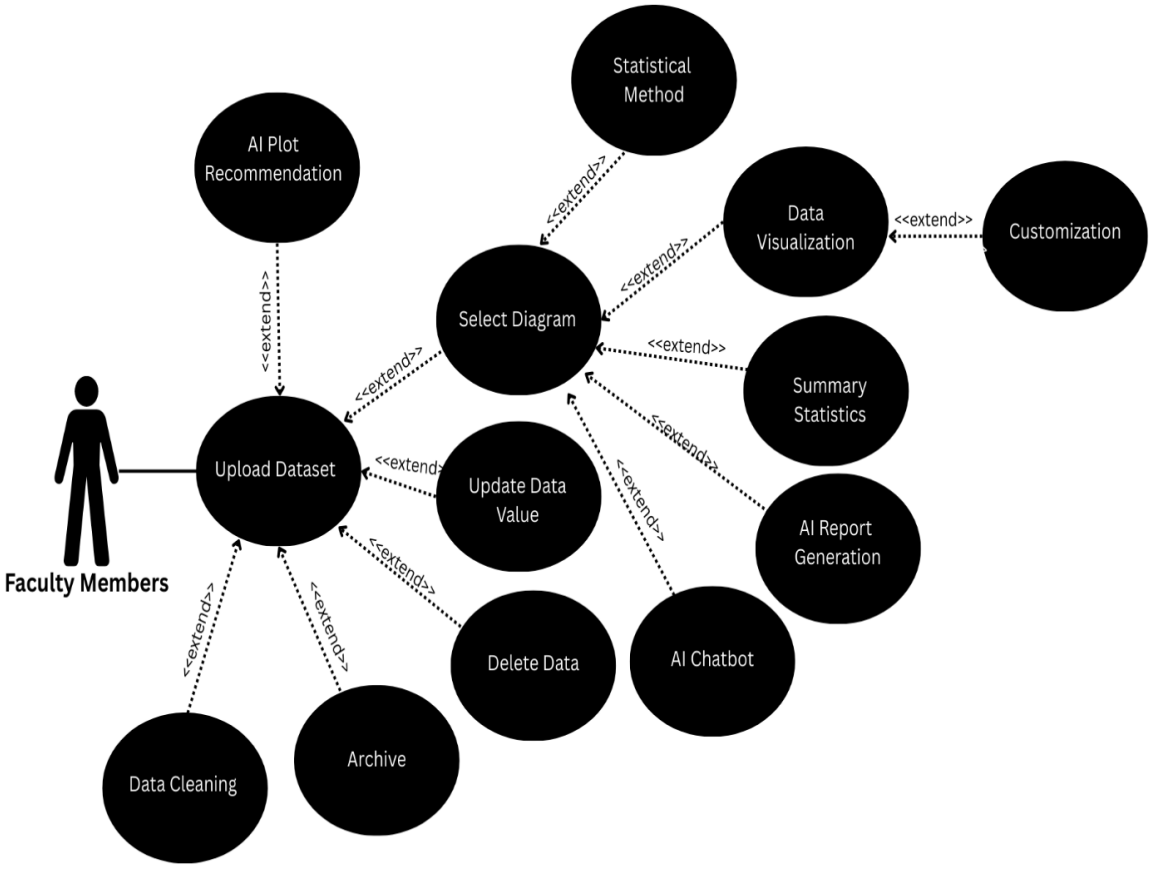


Figure 2.Use Case Diagram

Figure 2 illustrates the use case diagram of the system, which features a single actor representing the user. The faculty member can upload dataset to the system. There are optional actions such as selecting a diagram, updating data values, deleting data, and archiving files, extended from the upload process, depending on the user's preferences. Additionally, users have the option to automatically clean the uploaded data and view AI-generated plot recommendations based on the dataset.

Once a diagram is selected, the user is presented with four main options: data visualization, summary statistics, AI-generated report, and statistical treatment. Upon choosing a preferred visual format, users may further customize the graphs and charts if desired, or they can retain the default settings.

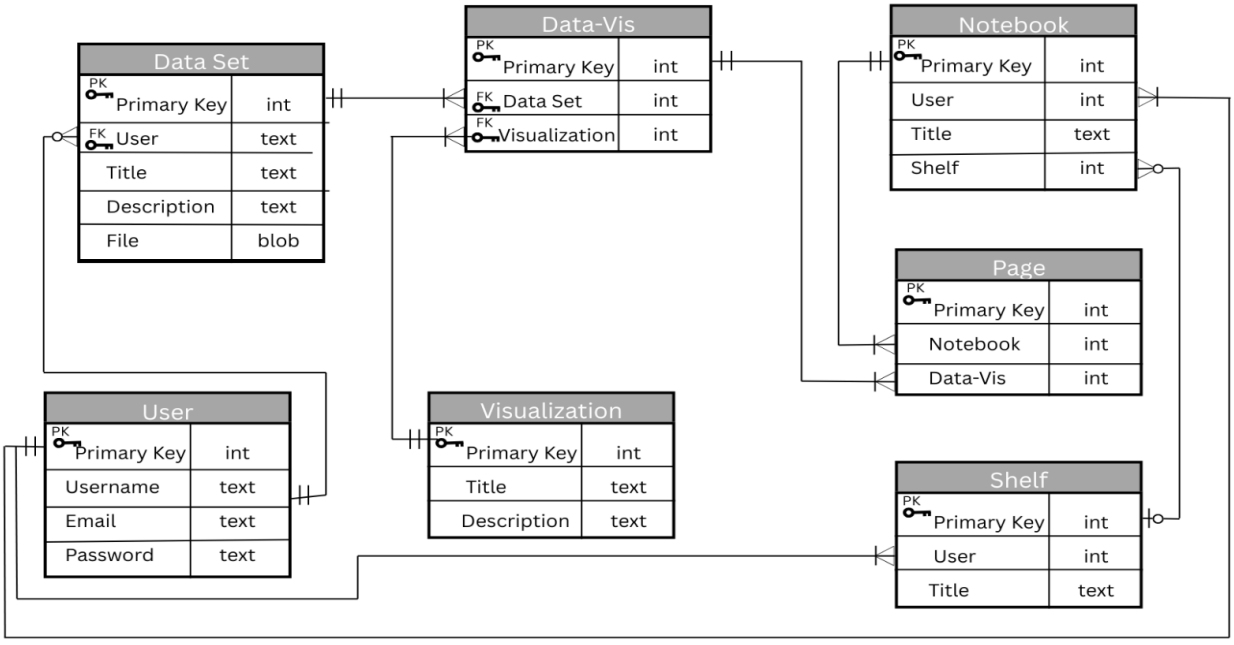


Figure 3. ERD Diagram

As shown in Figure 3, this database design enables users to upload, manage, and visualize datasets. Users can create and store multiple Data Sets, which include attributes like Title, Description, and associated File data. These datasets are stored in the Data Set table and linked to visualizations through the Data Vis table, which establishes a many-to-many relationship. This relationship allows each dataset to have multiple visualizations such as line charts, bar charts, and allows each visualization to be reused across different datasets. It also includes a Notebook entity that groups datasets and visualizations into pages for organization. Users can create Pages under their notebooks, with each page containing specific datasets and their corresponding visualizations. Furthermore, notebooks are grouped within a Shelf entity for higher-level categorization. This setup provides users with flexibility in organizing, viewing, and analyzing their data in various visual formats, while maintaining a structured hierarchy for better data management.

DEVELOPMENT TOOLS AND TECHNOLOGIES

The proposed system will be developed using Python and JavaScript as the primary programming languages to establish a strong foundation of the system. The backend will be built using the Django framework, while the frontend will utilize ReactJS to create an interactive user interface. SQlite will serve as the local database of our proposed system, as it is the default database for Django projects. Additionally, for overall development, Visual Studio Code (VS Code) will be used for its efficiency and well-organized coding environment. Additionally, AI integration will be implemented for automated report based on summary statistics using DeepSeek R1 7B Q6 model.

Table 2. Hardware Requirements

	Minimum	Recommended
Operating System	Windows 10/11	Windows 10/11
Processor	Intel Core i3 (10th Gen) / AMD Ryzen 3 3100	Intel Core i5 (10th Gen) / Ryzen 5 3600
Memory	8 GB Ram	16 GB Ram
GPU	Intel UHD 620 or equivalent	Nvidia GTX 1660 or better

Table 2 outlines the minimum and recommended hardware specifications for installing AL1SE to ensure efficient and smooth performance.

ALGORITHMS AND TECHNIQUES

The system will use DeepSeek R1 7B, a Large Language Model (LLM), to generate automated reports based on summary statistics data. The model we will use operates at 6-bit quantization (more bits, more accuracy), optimizing memory efficiency while maintaining accurate text generation. To ensure precise data handling of date fields, the system will integrate ChronoV2, a natural language date parser that standardizes various date formats extracted from unstructured text. Moreover, the Simple Statistics package will be used to compute necessary data for summary statistics feature, which will be the basis of the automated report.

IMPLEMENTATION PROCESS

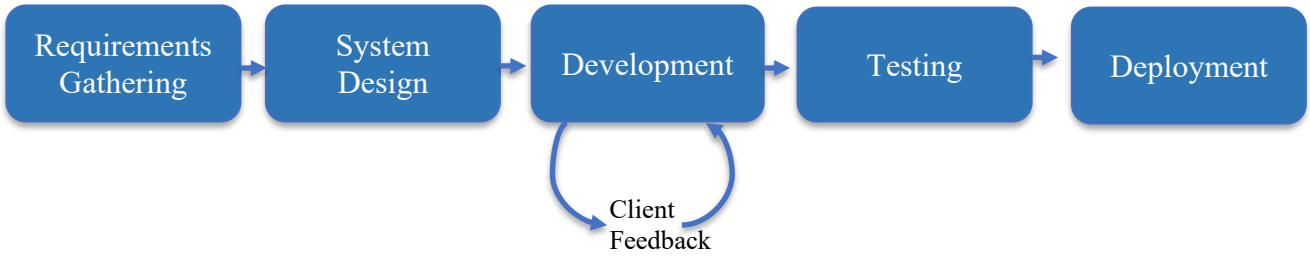


Figure 4. Agile Methodology

The development of AL1SE will follow the Agile methodology, ensuring flexibility, user-centered design, and adaptability to changes. For system design, the researchers will use prototype model, which aligns with Agile principles by emphasizing iterative development. Specifically, the researchers will use evolutionary prototyping which evolves through continuous build and refinement based on the Department of Statistics faculty members feedbacks. This approach allows interaction with early versions of the system, providing insights that will guide the refinements of the system. By integrating Agile with Prototype Model, the researchers ensure an adaptable process that prioritizes user.

3.1 Requirements Gathering

In the requirements gathering phase, the researchers will conduct a semi-structured interviews with open-ended questions to gain insights about the Department of Statistics' current data visualization tools. This phase aims to identify the existing limitations of the department's current data visualization tool and determine necessary features for AL1SE.

Sample Interview Questions:

1. Can you describe how often you use data visualization tools, and when do you use them?
2. What challenges have you encountered while using your current data visualization tools?
3. What specific features in a data visualization tool do you wish were available?
4. What types of data visualizations do you need or usually use?
5. How important is an accessible data visualization tool for you?

3.2 System Design

In this phase, the researcher will adopt the evolutionary prototyping model as described by Sommerville (2011), which emphasizes the continuous refinement of the system through iterative development. The deliverable of this approach is a working prototype that will incrementally improve through an iterative cycle of building and refinements. The process begins with the development of a basic system prototype, then the collection of feedback for every iteration. Each iteration will incorporate necessary adjustments and feature enhancements based on the gathered feedback. This cycle will continue to iterate until the system meets user requirements, and no further refinements are needed.

3.3 Development

This phase involves continuous refinement through iterative development based on client feedback. Since the researchers will use evolutionary prototyping, the development will start in the system design phase, and will continue throughout this phase. Unit testing will also be conducted concurrently, as developers will need to test each system feature individually whenever a new feature is added.

To manage iterations, the researchers will store different versions of the system on Google Drive, for easy access storage of the system different versions. This version control approach will help the researchers keep track of the system progress.

3.4 Testing

The developers will conduct unit testing every iteration, before presenting the working prototype. Since the approach is evolutionary prototyping, integration testing will not be conducted as a separate phase, as the interaction between features will already be evaluated throughout the iterative unit testing process, instead final developer testing will be conducted. IT expert testing will also be performed before the deployment, to evaluate the software's technical quality using the applicable ISO 25010 criterion. The evaluation will be carried out through surveys, combining the Likert Scale for quantitative assessment with open-ended questions for qualitative assessment.

3.5 Deployment

This is the phase where the researchers will deploy the final software to the Department of Statistics after the IT experts have completed their testing and no issues persist.

This phase includes:

- Software installation on designated computers.
- User Training and Documentation for faculty members.
- Initial Monitoring and support to address any deployment issues.

TESTING AND EVALUATION

After the deployment, AL1SE will undergo user testing, to ensure that all features are correctly aligned with the workflow of the Department of Statistics. The software will be evaluated using the original version of the Technology Acceptance Model (TAM) to measure user satisfaction and adoption. The evaluation will be carried out through a survey questionnaire, which will be a combination of Likert Scale for quantitative assessment, and open-ended questions for comments and suggestions.

CHAPTER IV

RESULTS AND DISCUSSION

This chapter presents the results of developing AL1SE, a stand-alone AI-integrated data visualization software for the Department of Statistics at Central Luzon State University (CLSU). The discussion is structured according to the development phases, following the methodology applied by the researchers. Each section highlights the implementation, testing, and evaluation processes carried out to ensure the project objectives are met accordingly.

4.1 DATA GATHERING

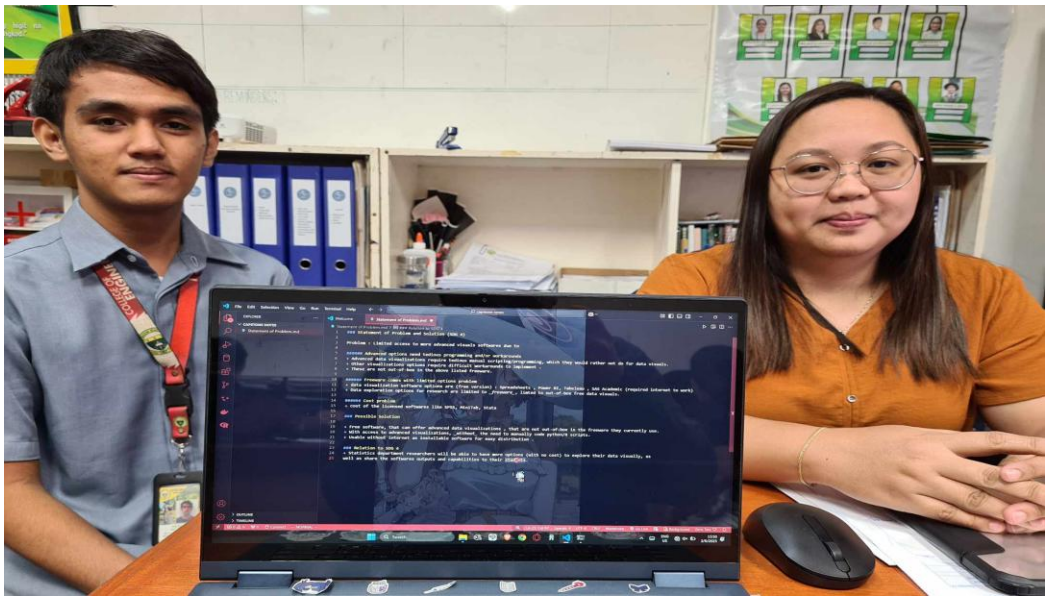


Figure 5. Interview

The researchers gathered data through the use of semi-structured interviews with the faculty members of the Department of Statistics. As shown in the figure above, one of the interviewees was the acting head of the department, Ma'am Onneth O. Tejada. This process ensured that the features of the developed software were aligned with the department's workflow and addressed its specific needs.

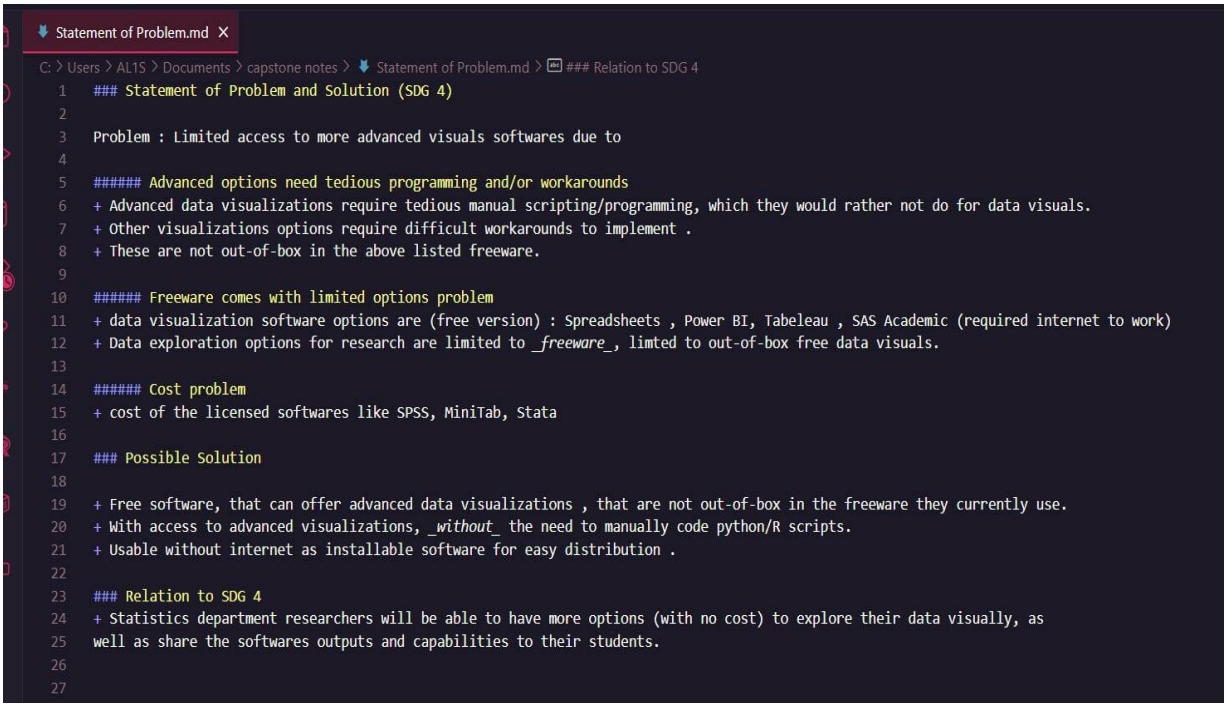


Figure 6. Summary of the Interview

As shown in Figure 6, the interviews revealed several challenges with the department's existing data visualization tools. Faculty members noted that available tools required manual scripting or programming in Python or R, which they preferred to avoid. Additionally, other visualization options required complex workarounds that were difficult to implement. The department was also constrained by the limited availability of freeware tools, which highlighted not just the cost-related issues of existing data visualization solutions, but also their accessibility, as most alternatives required a stable internet connection.

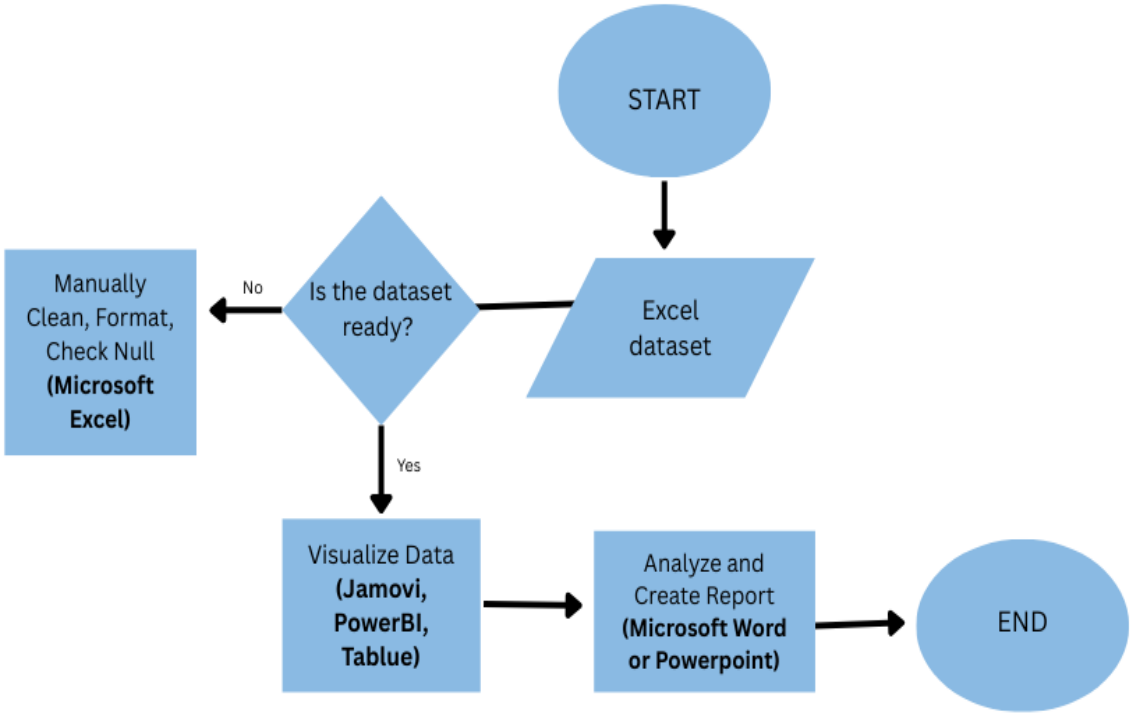


Figure 7. Department of Statistics Current Process

Figure 7 shows the current workflow of the faculty at the Department of Statistics. They manually clean, format, and edit datasets in Microsoft Excel before using different tools such as Jamovi, Power BI (free), and Tableau (public) for visualization. Once charts are produced, outputs are manually transferred into Microsoft Word or PowerPoint to prepare reports. Faculty members usually handle the analysis on their own, but advanced visuals often require coding in Python or R, which they prefer not to use. Incorporating additional visualizations would also require manually installing packages or extensions to their existing software. As a result, they are limited to the default visuals provided by these free tools..

Table 3. Comparative Scope of AL1SE vs Statistics Department Current Tool

Feature	AL1SE (Developed System)	Power BI (Free/Desktop)	Tableau (Public/Free)	Jamovi
Target Users	CLSU Statistics Department Faculty Members	Business/enterprise users	Business/enterprise users	Academic/statistics users
Deployment	Standalone, offline	Desktop app; cloud sharing needs paid plan	Public web publishing; private use requires paid	Desktop app, offline
Visual Types	Built-in advanced plots (scatter matrix, violin, 3D cluster, etc.)	Standard BI visuals; advanced require add-ons or Python/R	Broad visuals; advanced need Python/R or calc fields	Core statistical plots only
AI Support	Built-in AI (reports, plot suggestions)	No built-in AI in Desktop	No built-in AI in Public	None
Data Cleaning	Guided options (null check, imputation, formatting)	Power Query (manual setup)	Tableau Prep/calc fields	Basic GUI, R-based
Statistics	Full set (central tendency, dispersion, skew,	Basic summaries	Basic summaries	Rich statistical tests

	kurtosis, frequency)			
Extensibility	No coding needed; features pre-built	Supports Python/R scripts	Supports Python/R via TabPy/Rserve	Built on R, module-based
Best Use	Teaching, research.	Business dashboards	Public storytelling, enterprise dashboards	Academic statistics

Table 3 compares AL1SE with the Department of Statistics' existing tools such as Power BI, Tableau, and Jamovi. While these tools provide useful features, they also present limitations for the department's workflow. The free versions of Power BI and Tableau are free and offline, but it requires Python or R coding to generate specialized visuals, while Jamovi is mainly focused on statistical tests and offers limited visualization options.

AL1SE bridges these gaps by integrating advanced, ready-to-use statistical visuals, optional data cleaning, descriptive statistics, and correlation analysis (Pearson and Spearman) into one offline system. Unlike the general-purpose design of the existing tools, AL1SE is tailored specifically for the Department of Statistics. In addition, it features DeepSeek-R1 for AI-assisted reporting and plot recommendations, a capability not present in the alternatives.

Feasibility Study

This section presents the analysis of the project's viability, covering technical, economic, and operational aspects. It demonstrates that the proposed solution is a practical and well-considered approach to the problem.

Technical Feasibility

Table 4. Developers Laptop Specification

Component	Specification
Laptop Brand	Acer Aspire 7 R0S4
Operating System	Windows 11
Processor	AMD Ryzen 5 5500U
Memory	16 GB Ram
GPU	GTX 1650
Storage	512 NVMe SSD

AL1SE was developed using the team’s laptops, mainly an Acer Aspire 7 R0S4. This laptop had enough processing power and memory for programming, testing, and AI integration. Its GTX 1650 GPU and 16 GB RAM allowed smooth performance, while the NVMe SSD helped speed up application builds and overall system responsiveness.

Table 5. Statistics Department Computer Specification

Component	Specification
Model	Inspiron 24 S420 All-in-One
Operating System	Windows 11 Home Single Language (Version 24H2)
Processor	13th Gen Intel® Core™ i5-1335U (1.30 GHz)
Memory	16 GB RAM
Graphics	Intel® Iris® Xe Graphics (128 MB)
Storage	477 GB SSD
Display	24-inch Touchscreen Display

AL1SE was deployed on standard desktop computers in the Department of Statistics, such as the Inspiron 24 S420 All-in-One unit. This computer, equipped with a 13th Gen Intel Core i5 processor and 16 GB of RAM, provided sufficient performance for running the system smoothly. The built-in Intel Iris Xe Graphics ensured clear data visualizations, while the SSD storage supported fast data access and reliable system responsiveness. The 24-inch touchscreen display also made it easier for users to navigate the system interface efficiently.

Conclusion: Technically feasible

Economic Feasibility

AL1SE did not require significant financial resources for development or deployment. The project relied entirely on team effort, with the only expenses being personal costs incurred by the team. There were no licensing fees, subscriptions, or

hardware purchases required. The system's benefits included streamlined data visualization, reduced reliance on external software, and improved efficiency for the Department of Statistics, making the project cost-effective.

Conclusion: Economically feasible

Operational Feasibility

Table 6. Skills to Use the Software

Operational Requirement	Description
Statistical Knowledge	Users were expected to have a basic understanding of statistics to interpret visualizations and results effectively.
Computer Skills	Basic computer literacy was required, including navigating software, managing files, and using the user interface.
Analytical Thinking	Users also needed to analyze and make sense of the outputs generated by the system.

Table 6 presented the key operational requirements for using AL1SE. It outlined the essential skills and knowledge that faculty needed to operate the system and interpret its outputs effectively.

Conclusion: Operationally feasible

4.2 SYSTEM DESIGN

This section presents the system design of AL1SE, highlighting its flow and key components. It includes the system flowchart, conceptual framework (IPO), system architecture, use case diagram, context diagram, data flow diagram, and database design. As the project followed an evolutionary prototyping approach, the diagrams reflect the final intended system. To complement these models, screenshots of the user interface are also provided as visual references.

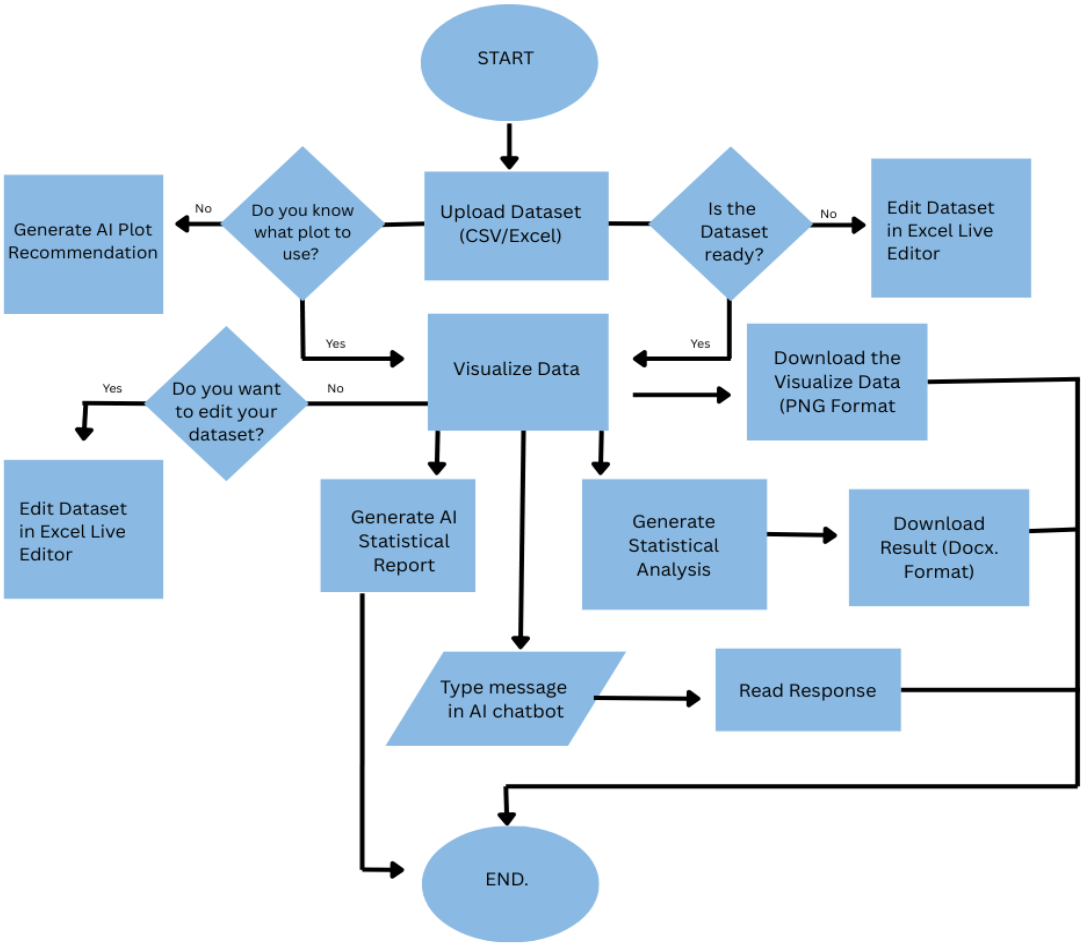


Figure 8. System Flowchart

This flowchart illustrates the end-to-end workflow of the system, beginning with dataset upload and progressing through data visualization, statistical analysis, and AI-

assisted interactions. It supports editing datasets via a live Excel interface, AI-driven plot recommendations, and the ability to download outputs. The user can also interact with an AI chatbot for additional insight.

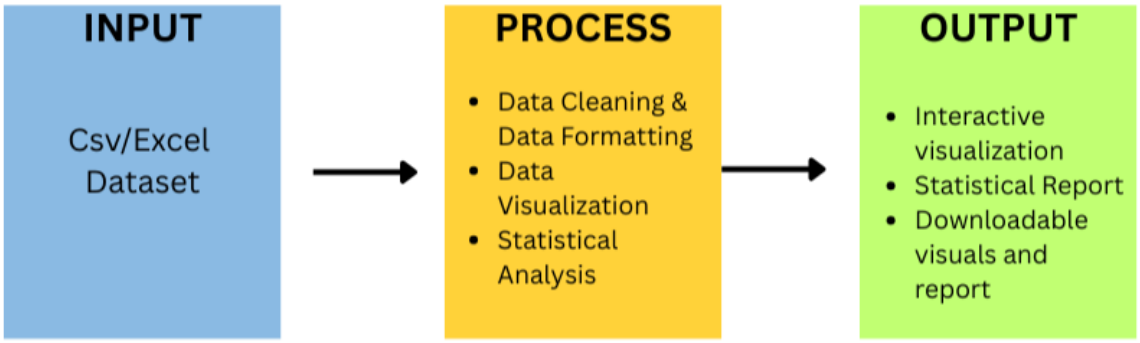


Figure 9. Conceptual Framework (IPO)

Figure 9 presents the conceptual framework, which outlines the high-level flow of the data analysis system. It begins with importing a dataset in CSV or Excel format (Input), followed by key processing steps including data cleaning or formatting, data visualization, and statistical analysis (Process). The system delivers final results in the form of interactive visualizations, statistical result, and downloadable reports (Output).

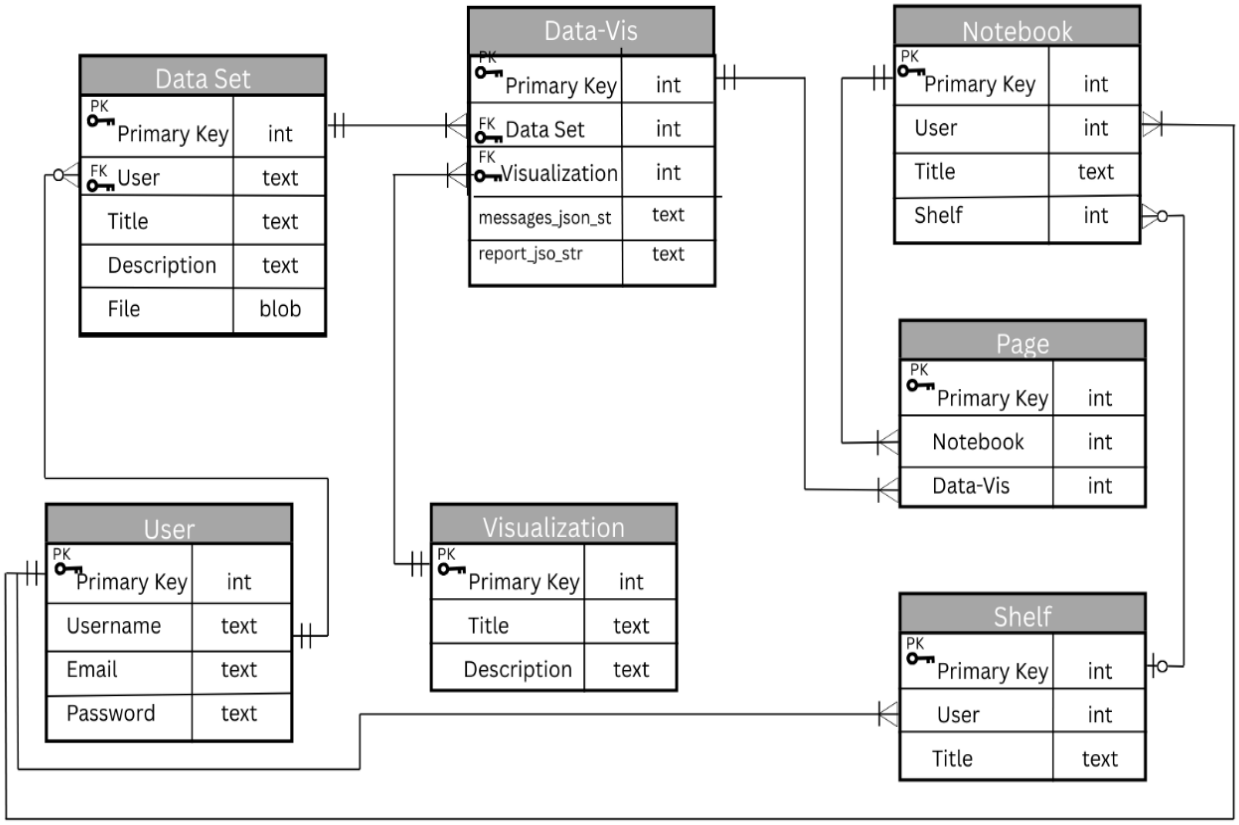


Figure 10. Final ERD Diagram

Figure 10 represents the finalized database structure of the developed data visualization software for the Department of Statistics. The system allows users to upload datasets, generate visualizations, and receive AI insights. The Data-Vis table acts as a bridge between datasets and visualizations while storing AI-generated outputs including chatbot messages (messages_json_str) and statistical reports (report_json_str). These outputs are stored as JSON strings, ensuring flexibility and easy integration with the AI components. Users can organize their work into notebooks and pages, with an additional layer of structure through customizable shelves.

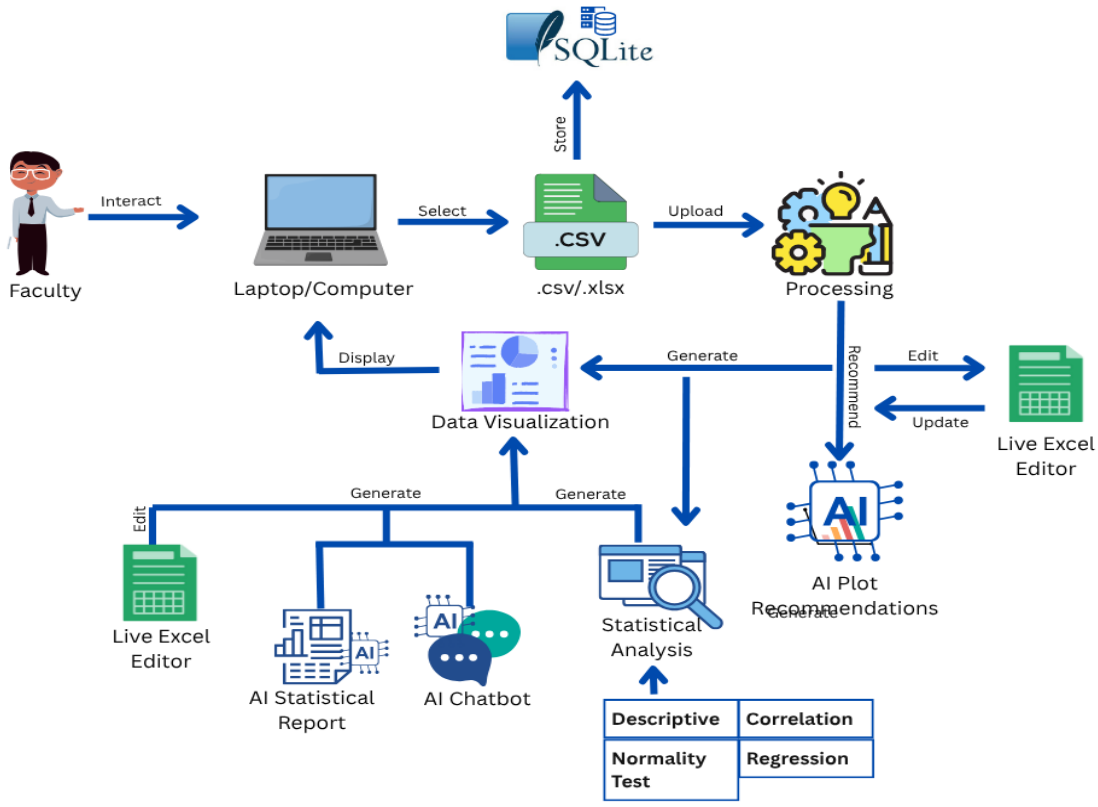


Figure 11. Final System Workflow Diagram

Figure 11 illustrates the system flow. It begins with uploading datasets in CSV or Excel format, which are stored in the local SQLite database. Users can then receive optional plot recommendations to guide their choice of visualizations. The system includes a spreadsheet editor for immediate data editing, AI-assisted statistical report generation, and an AI chatbot for user support. Finally, users can perform various analyses, including Descriptive and Correlation analyses.

AL1SE User Interface (UI)

This section presents some of the main pages of the system's user interface. The screenshots highlight the core functions of the software, including dataset management, data visualization, statistical analysis, and report generation. These examples provide an overview of the primary interfaces encountered when using the system. For reference, the complete set of system screenshots is provided in Appendix D.

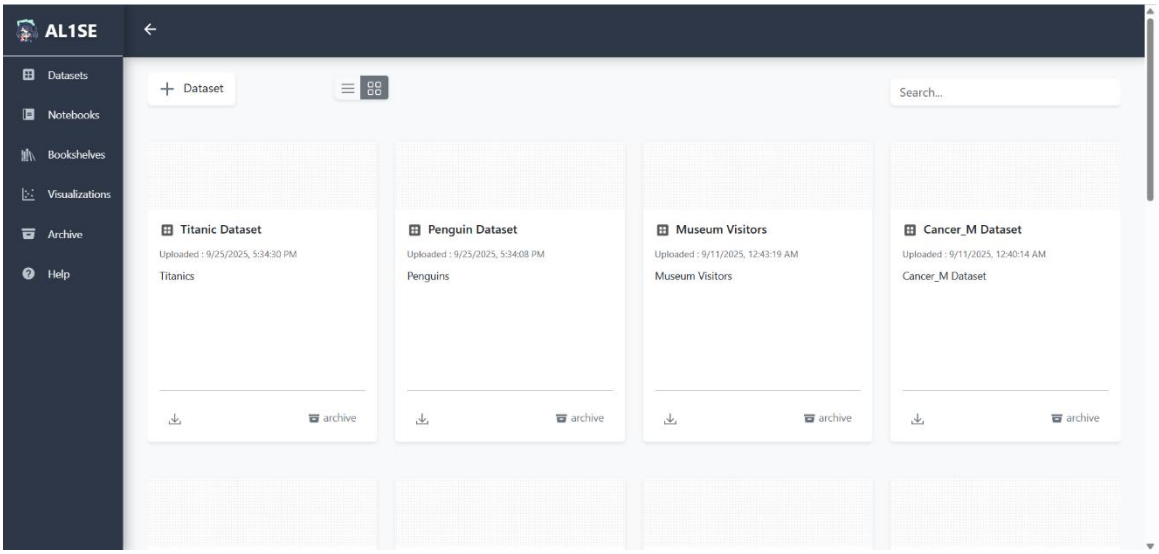


Figure 12. Dataset Page/Landing Page

The Dataset Page serves as the landing page of AL1SE, displaying all uploaded datasets. Users can quickly view dataset names, upload dates, and file types. Each card provides options to download or archive the dataset. The left sidebar allows navigation to other sections such as Notebooks, Bookshelves, Visualizations, Archive, and Help

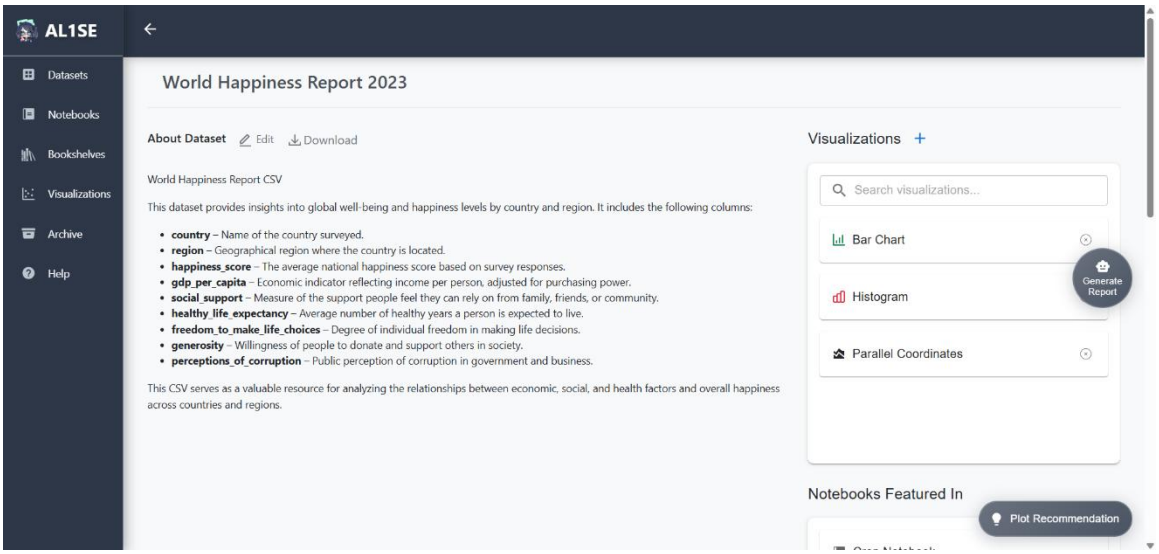


Figure 13. About Dataset Page

Figure 13 shows the page that appears after selecting a specific dataset. On this page, users can view details about the dataset, with options to edit or download it. Additionally, this is where users can add visualizations.

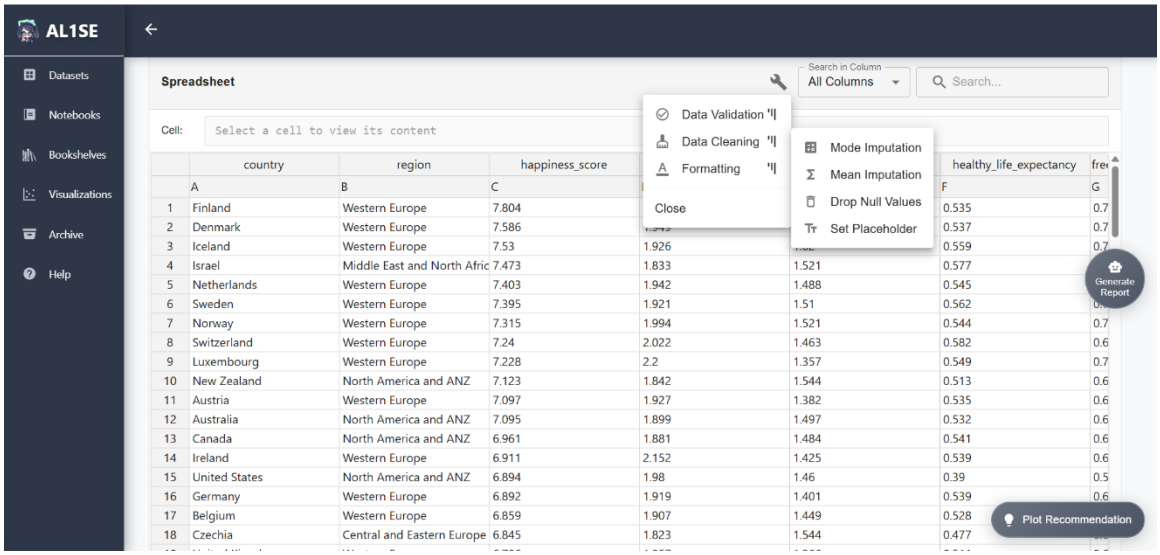


Figure 14. Live Excel Editor

Figure 14 shows the Live Excel Editor, where users can edit their dataset in real time. This feature supports essential functions such as data cleaning, data formatting, and data validation.

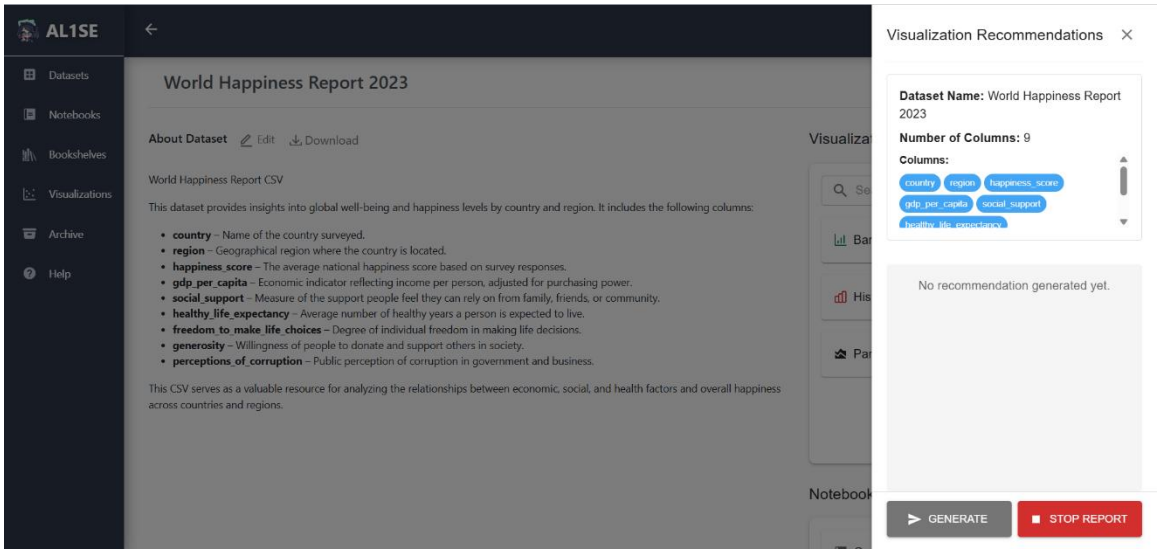


Figure 15. AI Plot Recommendation

Figure 15 shows the AI Plot Recommendation feature, which suggests appropriate chart types based on the dataset.

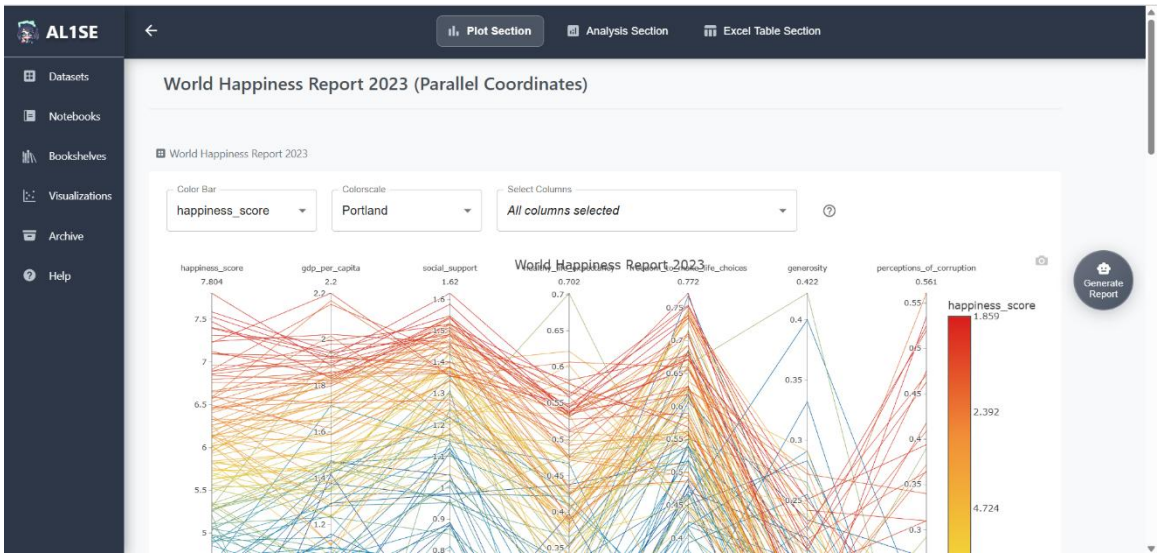


Figure 16. Sample Plot

Figure 16 shows a sample plot that allows users to zoom in and out, interact with the plot, and download the visualization in PNG format. By double-clicking on an item in the legend, users can also hide other columns to focus on visualizing a single column.

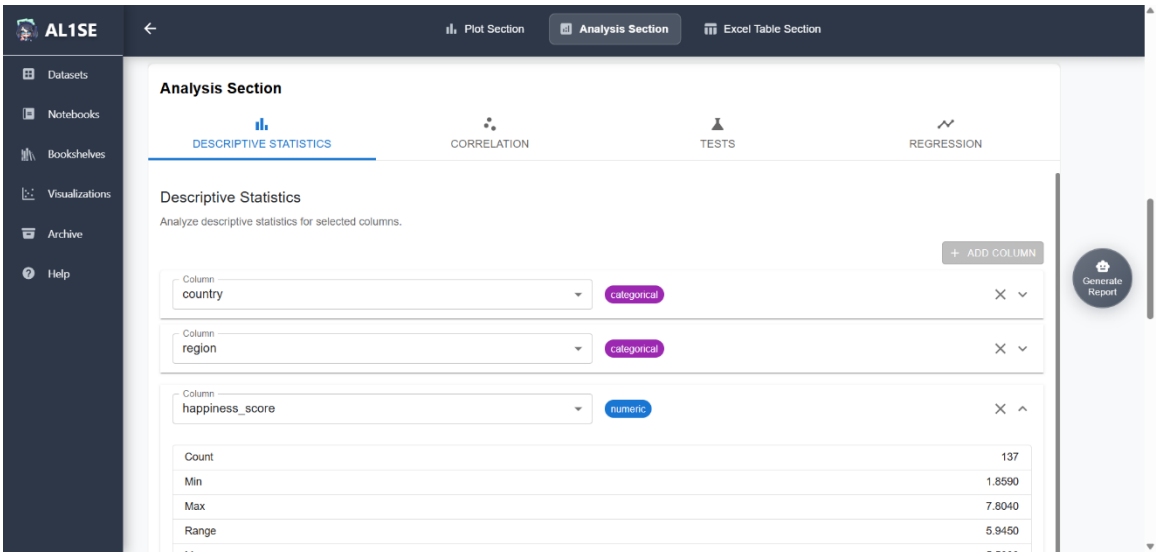


Figure 17. Analysis Tab

Figure 17 shows the Analysis tab, where users can perform different types of analysis, including descriptive statistics and correlation tests, particularly Pearson and Spearman correlations.

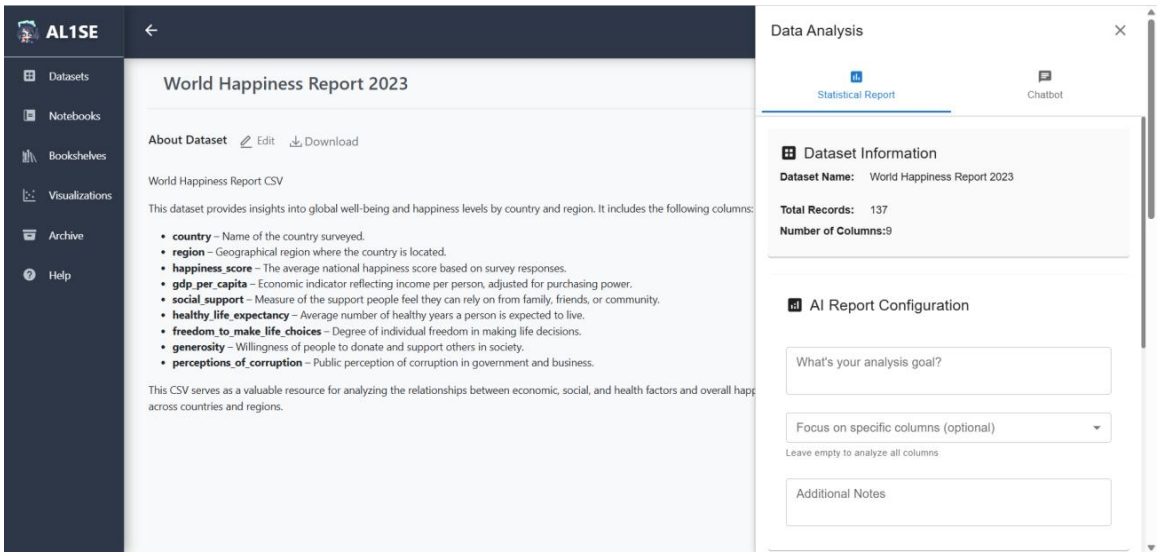


Figure 18. AI Statistical Report and Chatbot

Figure 18shows the AI statistical report and the AI Chatbot. These features assist users by generating answers and results based on the uploaded dataset, though they may still produce occasional errors or mistakes.

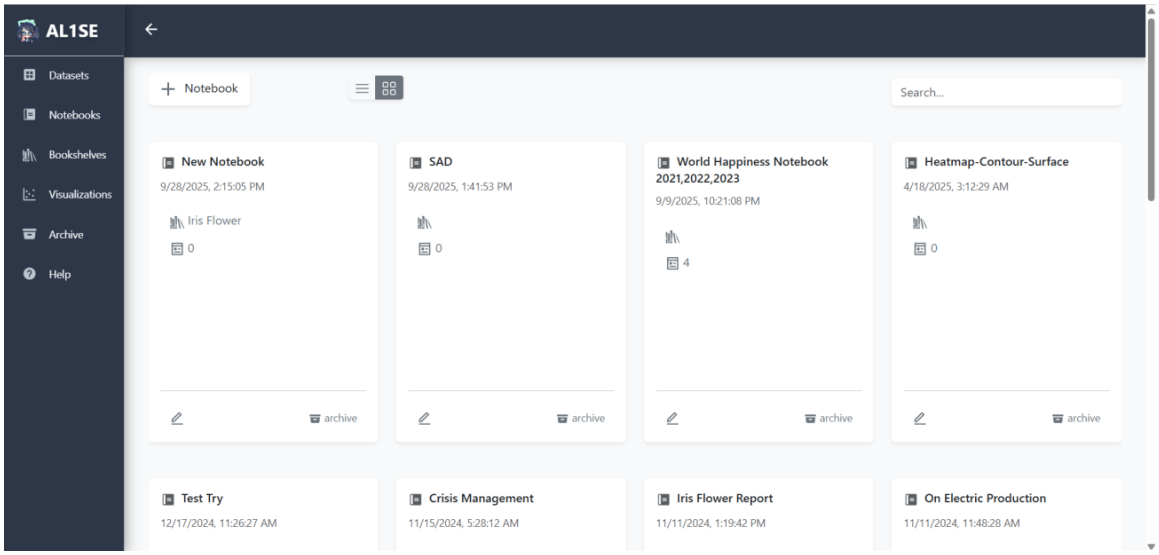


Figure 19. Notebook Page

Figure 19 shows the Notebook page, which consists of multiple visualizations, either the same datasets or different ones, organized within a single workspace.

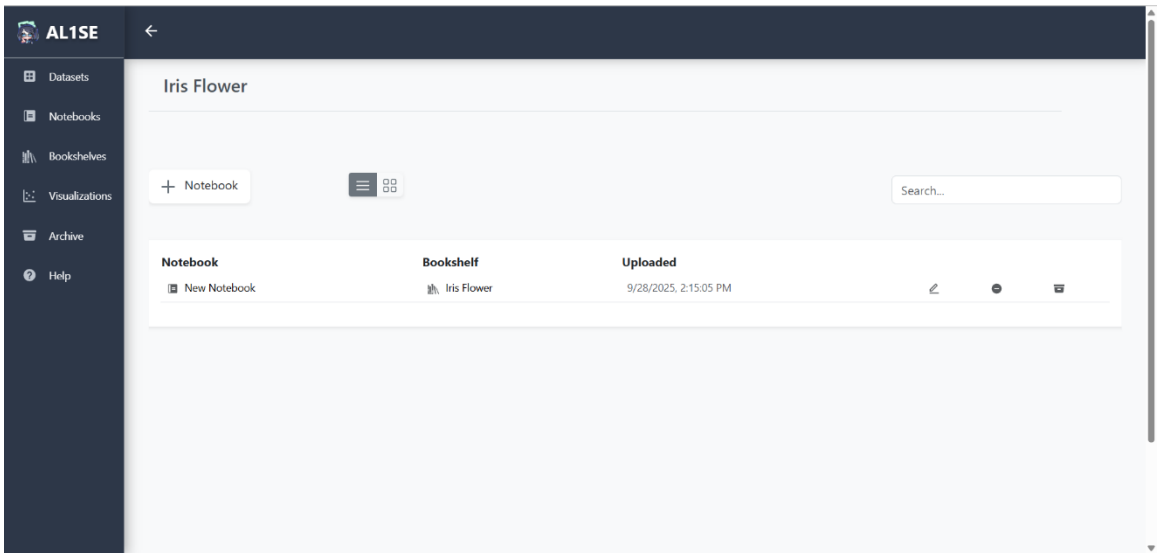


Figure 20. Bookshelves Page

Figure 20 shows the Bookshelves page, which consists of multiple notebooks organized within a single workspace.

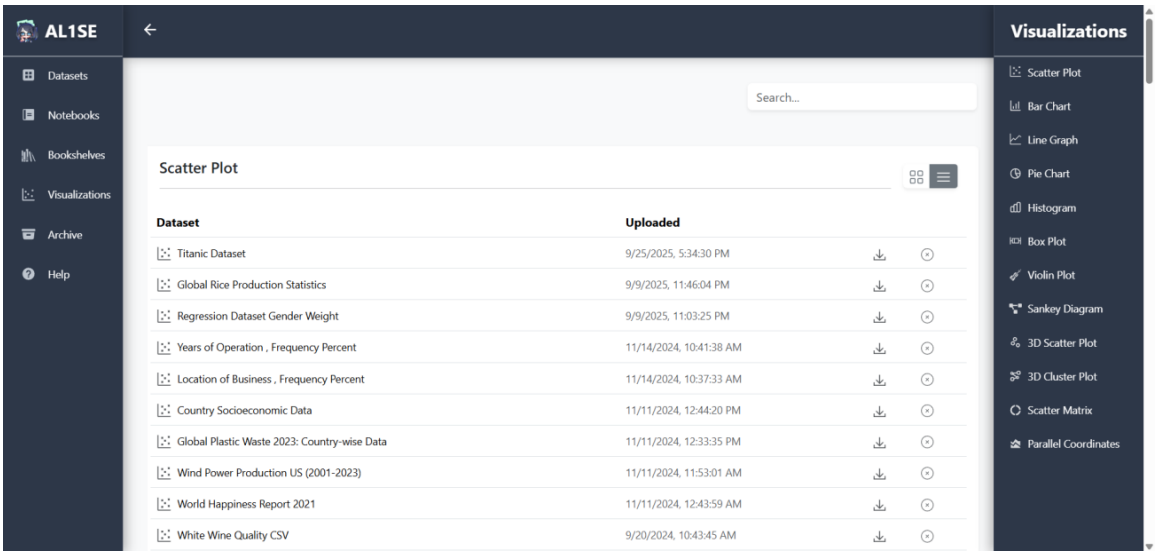


Figure 21. Visualization Page

Figure 21 shows the Visualizations page, where all created visualizations are displayed and categorized by their type.

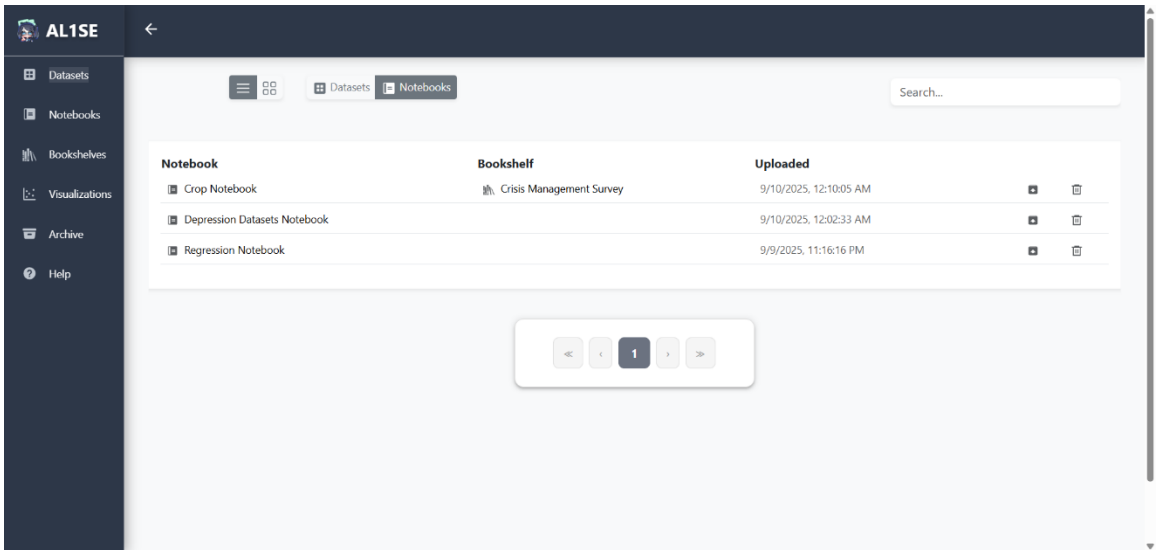


Figure 22. Archive Page

Figure 22 shows the Archive page, where users can view previously archived datasets and restore them by unarchiving.

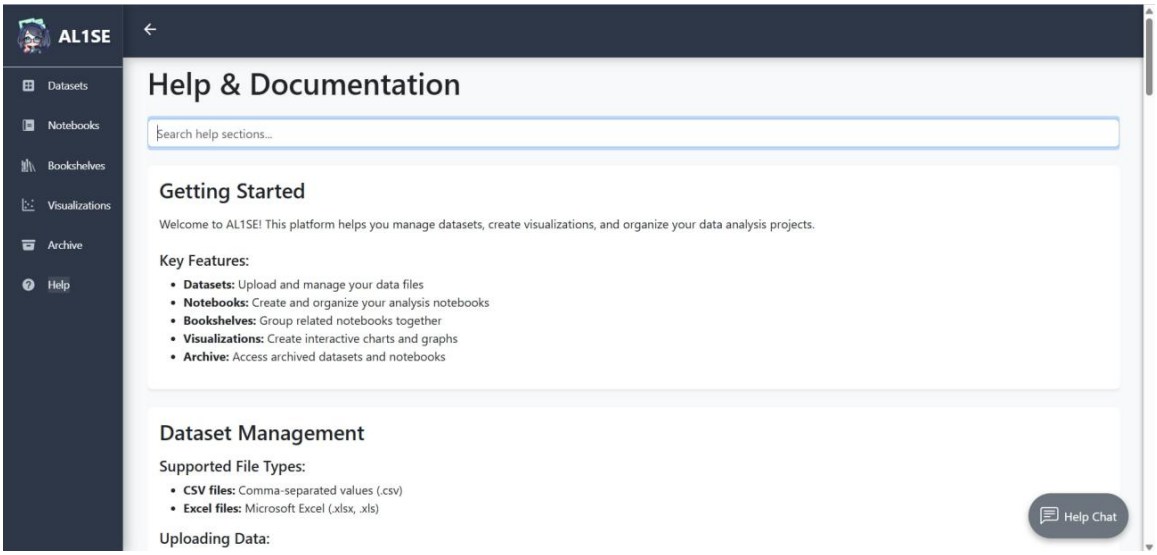


Figure 23. Help Page

Figure 23 shows the Help page, where users can access valuable information and guidance about the software.

4.3 DEVELOPMENT

To effectively design and develop AL1SE, a set of tools, frameworks, and programming languages was selected to ensure efficiency in building the software. These technologies served as the foundation for both backend and frontend development, database management, and AI integration.

Table 7. Technology Used for Development

Component	Technology / Tool Used	Version (If applicable)
Programming Language	Python (Backend)	3.11
	JavaScript	ES2024 (ECMAScript 2024)

Framework	Django (Backend)	5.2.6
	ReactJS (Frontend)	18.3.1
Runtime Environment	Node.js	20.18.0
Database	SQlite	3.45.3
Integrated Development Environment (IDE)	Visual Studio Code (VScode)	1.104.2
AI Model	DeepSeek	R1 1.5B model
AI Server	Ollama	0.5.7
Version Control	Google Drive	

Table 7 shows the technologies and tools used in developing AL1SE. The backend was built with Python and Django, while JavaScript and ReactJS handled the frontend interface. SQLite served as the database, and Visual Studio Code (VSCode) was used as the development environment. The system incorporates the DeepSeek R1 AI model, initially a larger version but later switched to the 1.5B model for faster responses. Google Drive was used for version control. Together, these tools provided an efficient and reliable environment for building an AI-integrated data visualization software for the Department of Statistics.

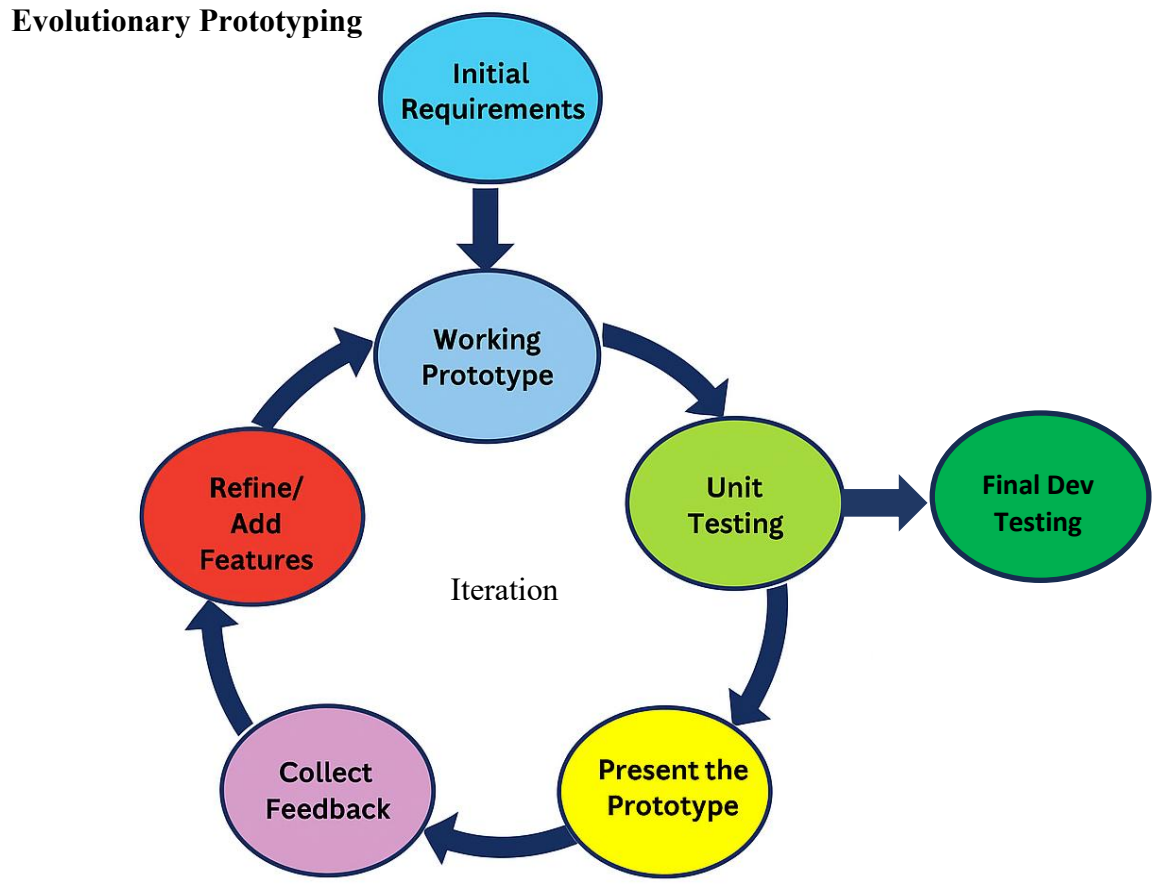


Figure 24. Evolutionary Prototyping

The researchers adopted evolutionary prototyping, beginning with data gathering to obtain the initial requirements of the system. After each round of unit testing, the Department of Statistics reviewed the working prototype, provided feedback, and suggested improvements to enhance its functionality. These inputs were incorporated into the following iteration to ensure that the software remained aligned with user needs. This cycle of testing, feedback, and refinement was repeated throughout development, with faculty members actively involved in every stage. The iterations were not limited to fixed schedules, as feature complexity varied and required flexibility. The process continued until the last iteration, after which the developers conducted final developer testing, completing the system development.

Initial Prototype Presentation

Figure 25. Initial Phase (Presentation of Working Prototype)

Figure 25 illustrates the presentation of the initial version of the developer's system prototype, which was presented to the Department of Statistics. During this presentation, the developers walk faculty members through the system's features, starting with importing CSV or Excel files to generate visuals. At this initial phase, the developers presented the initial design of the software and sample plots that the faculty can utilize, which consists of only 5 visualizations. After the prototype presentation, the developers asked the client what other visualizations they want to see in the software, as well as other user interface concern that can be modified and presented on the next prototype presentation.

Initial Phase Sample Output

Figures 26 and 27 below shows the landing page, and the sample plots during the first iteration, which were the initial working prototype we presented to the Department of Statistics.

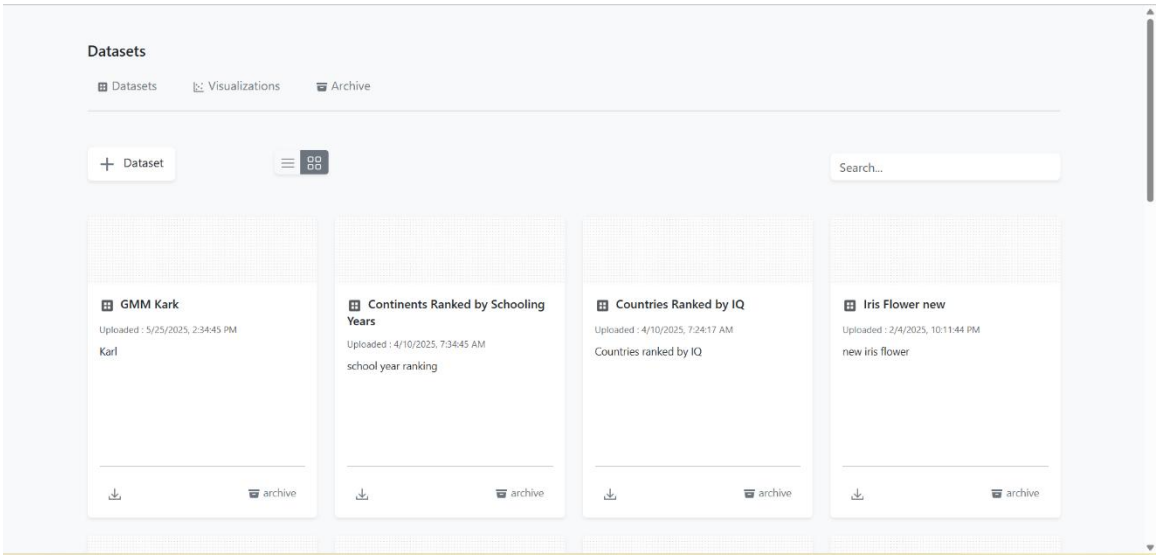


Figure 26. Landing Page/Datasets Page

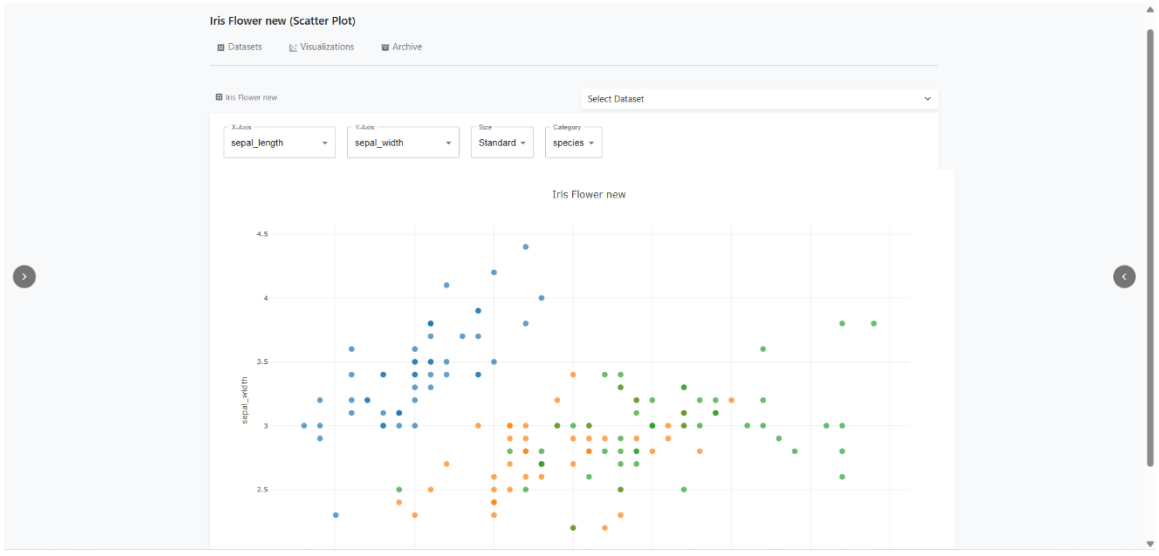


Figure 27. Scatterplot (sample plot developed)

First Iteration

Figure 28. First Iteration (Presentation of Working Prototype)

During the first iteration, as shown in Figure 28 above, we presented the new plots available in our data visualization software. The working prototype was developed based on the feedback gathered from the first iteration. We presented 15 data visualization options, including some 3D plots like the 3D scatterplot, which is interactive since it can be customized according to user needs, it can be rotated and zoomed for better analysis. To enhance user interaction, we also incorporated various visualization controls such as downloading plots as PNG, zooming in and out of the visualization area, panning, selecting data points using a box, autoscaling the visualization, and resetting the axes to default. Users can also filter, show, or isolate specific data categories for focused visualization.

First Iteration Sample Output

Figures 29 and 30 below show two of the fifteen plots developed and presented to the Department of Statistics during the first iteration.

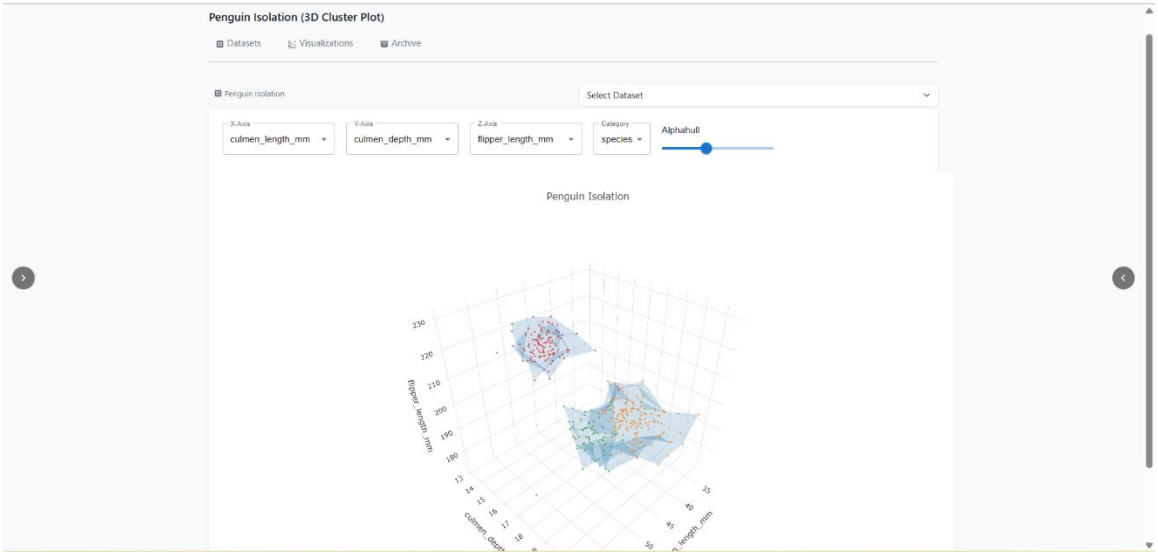


Figure 29. Data Visualization Page (3D Scatter Plot)

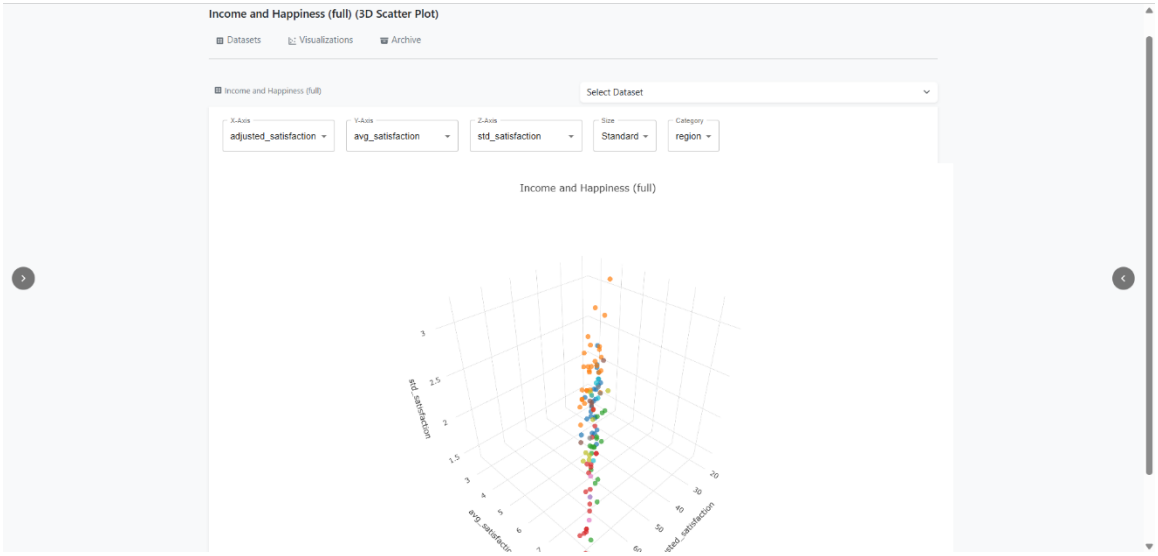


Figure 30. Data Visualization Page (3D Cluster Plot)

Second Iteration

Figure 31. Second Iteration (Presentation of Working Prototype)

Figure 31 shows our third iteration of the working prototype. At this stage, the spreadsheet editor was integrated, allowing users to update datasets directly within the software without relying on external tools. Descriptive statistics were enhanced to include both numerical and categorical summaries, with the option to remove specific items from the list. Correlation analysis was also introduced, enabling the calculation of both Pearson and Spearman correlations to help users better understand relationships within their datasets.

Second Iteration Sample Output

Figure 32 shows the developed Excel editor, which has different functionalities like checking of null values, formatting, and more. While Figure 33 shows the descriptive statistics tab.

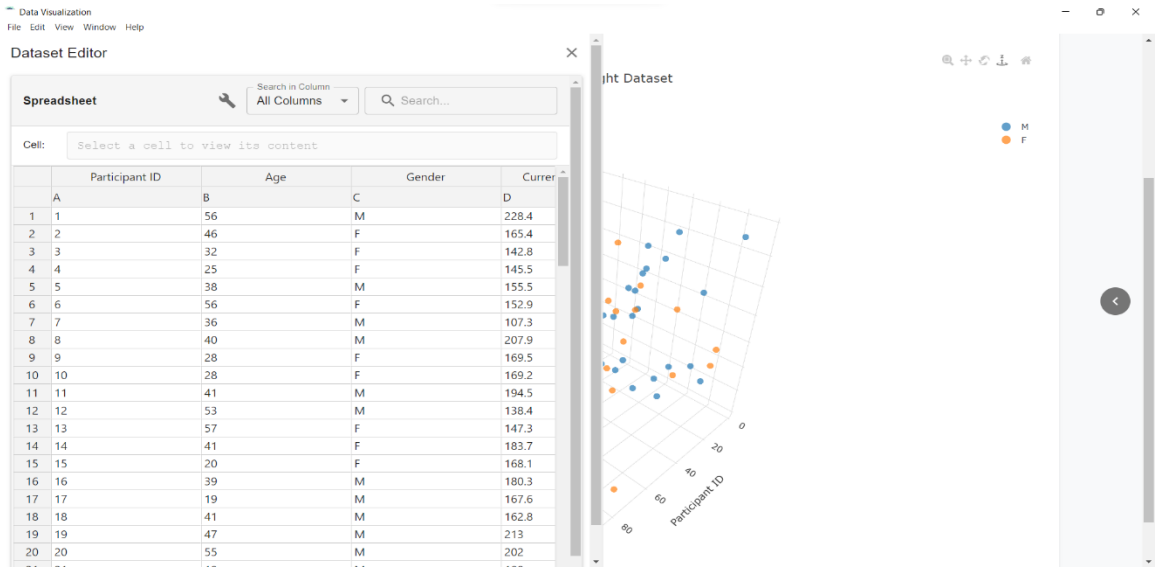


Figure 32. Data Visualization Page (Live Excel Editor Drawer)

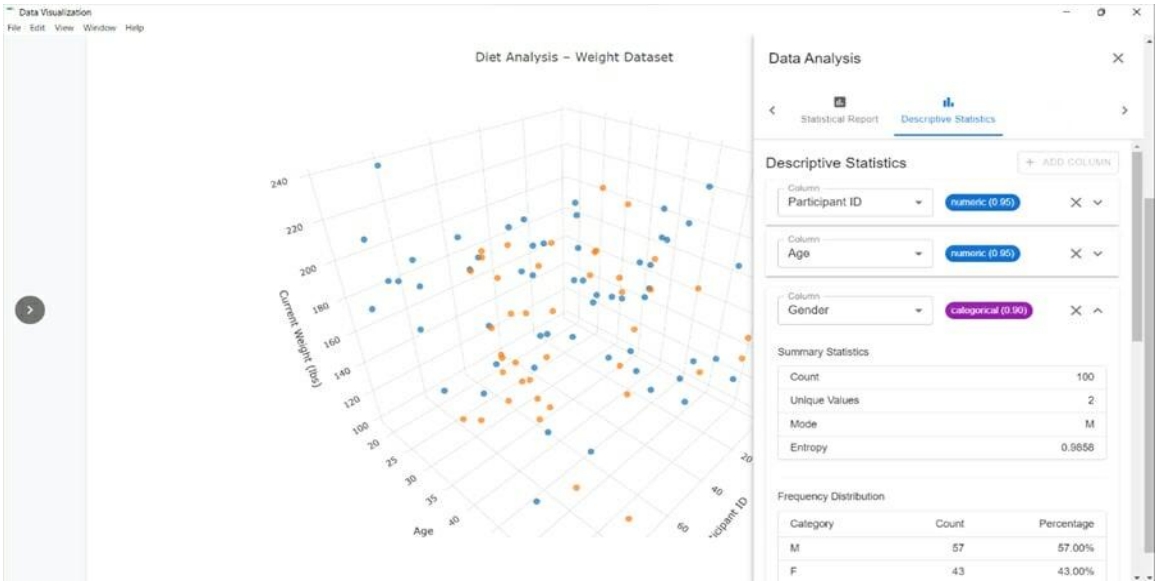


Figure 33. Data Visualization Page (Descriptive Statistics Tab)

Figure 34. Third Iteration (Presentation of Working Prototype Version)

Figure 34 shows our third iteration. At this point, the data visualization software, AL1SE, had an AI chatbot for user assistance. The chatbot was integrated to guide users, although statistical knowledge was still required since AI can make mistakes. In this iteration, the AI was enhanced with multiple functions: it recommended the suitable visualization type for the uploaded dataset, could stop the streaming of responses when prompted, and generated an automated statistical report, which was displayed in the Generate Report drawer. To further support users, the system also allowed the download of the generated report in .docx. format for offline use and documentation. Additionally, normality tests and simple linear regression were introduced, expanding the tool's statistical analysis capabilities beyond descriptive statistics.

Third Iteration Sample Output

Figure 35 presents the **normality test feature**, which evaluates whether the dataset follows a normal distribution. Figure 36, on the other hand, showcases the AI chatbot, which leverages the dataset and summary statistics to assist the user, though it may still make mistakes.

Figure 35. Normality Test Tab

Figure 36. Data Visualization Page (Chatbot Tab)

4.5 TESTING

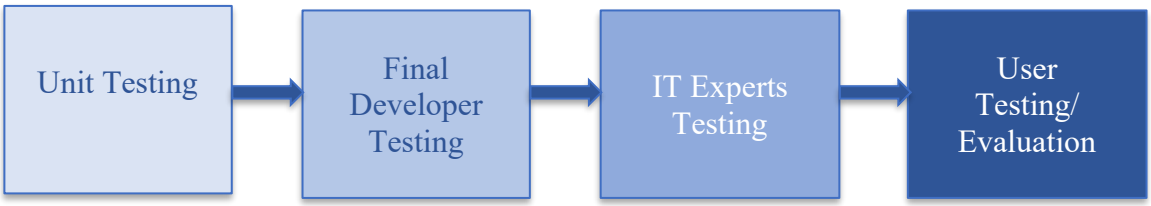


Figure 37. Testing Phases of AL1SE

The diagram presents the testing phase carried out for AL1SE. The process began with unit testing to evaluate individual features during each iteration. After the system was fully developed, final developer testing was conducted to validate the integration and overall functionality of all modules working together. Next, IT experts evaluated the system's technical quality based on ISO/IEC 25010 standards. Finally, after deployment, user evaluation was performed using the original Technology Acceptance Model (TAM) to assess the users' perceived usefulness and perceived ease of use of the system.

Developer Testing

The developer tested the software during development to ensure that all features worked together correctly. For unit testing, condensed test cases were created for each iteration to verify the features before presenting them to the Department of Statistics. The iteration cycled three times in varying spans of time, as no fixed timeframe was imposed for each feature. Given the evolutionary prototyping approach, integration testing was not conducted as a separate phase, since the interaction between features had already been verified during the final developer testing, which can be seen in Appendix D as it consisted of 138 test cases.

Table 8. Test Cases for the Initial Phase

TC ID	Test Scenario	Expected Result	Actual Result	Status
TC01	Upload CSV file	CSV uploads successfully	Uploaded successfully	Pass
TC02	Upload Excel file	Excel uploads successfully	Uploaded successfully	Pass
TC03	Upload Word file	System rejects Word file	System accepted Word file	Fail
TC04	Upload invalid format (image)	Error message displayed	Error shown	Pass
TC05	Display 5 visualizations (Histogram, Boxplot, etc.)	All render properly	Rendered correctly	Pass
TC06	Use Zoom tool on Histogram	Area zooms in/out	Worked correctly	Pass
TC07	Use Pan tool on Histogram	Chart pans smoothly	Worked correctly	Pass
TC08	Use Box Select on Histogram	Points selected in box	Worked correctly	Pass
TC09	Use Lasso Select on Histogram	Points selected via lasso	Worked correctly	Pass
TC10	Reset Axes on Histogram	Axes return to default	Worked correctly	Pass

In this phase, the focus was on testing the system's file upload functionality and basic visualization tools. Most features worked as expected, except for the handling of unsupported file formats (e.g., Word files), which was initially accepted by the system instead of being rejected.

Table 9. Test Cases for the First Iteration

TC ID	Test Scenario	Expected Result	Actual Result	Status
TC11	Display Scatterplot	Scatterplot renders correctly	Rendered correctly	Pass
TC12	Display Bar & Line Graphs	Both render correctly	Rendered correctly	Pass
TC13	Display Box & Violin Plots	Both render correctly	Rendered correctly	Pass
TC14	Display Heatmap	Heatmap renders correctly	Rendered correctly	Pass
TC15	Display Sankey Diagram	Sankey diagram renders correctly	Rendered correctly	Pass
TC16	Display 3D Scatter Plot	3D scatter plot renders correctly	Rendered correctly	Pass
TC17	Display 3D Cluster Plot	3D cluster plot renders correctly	Rendered correctly	Pass
TC18	Display Contour Plot	Contour plot renders correctly	Rendered correctly	Pass
TC19	Display Surface Plot	Surface plot renders correctly	Rendered correctly	Pass
TC20	Display Scatter Matrix & Parallel Coordinates	Both render correctly	Rendered correctly	Pass

In this iteration, the system was expanded from the initial 5 visualizations to a total of 15 by introducing additional chart types. The test cases were condensed into 10 scenarios for clarity, focusing on the proper rendering and functionality of the new visualizations. In addition, the issues identified in the initial phase, such as the acceptance of unsupported file formats, were now correctly handled.

Table 10. Test Cases for the Second Iteration

TC ID	Test Scenario	Expected Result	Actual Result	Status
TC21	Upload CSV dataset with 3,000 rows	Dataset uploads successfully	Dataset uploaded successfully	Pass
TC22	Check and count null values per column	Null values counted and displayed	Null values counted correctly	Pass
TC23	Mode imputation on null values	Null values replaced with the mode of the column	Null values replaced correctly	Pass
TC24	Apply uppercase formatting to a column	All values in the selected column converted to uppercase	Values converted to uppercase	Pass
TC25	Apply text formatting (Title Case) to a column	All values in the selected column converted to title case	Values converted to title case	Pass
TC26	Insert and delete rows	Row inserted above or below and deletable	Rows inserted and deleted successfully	Pass
TC27	Insert and delete columns	Column inserted left/right and deletable	Columns inserted and deleted successfully	Pass
TC28	Update dataset values	Edited cells update correctly in the dataset	Dataset values updated successfully	Pass

TC29	Generate descriptive statistics for numerical column	Statistics displayed correctly (mean, median, etc.)	Numerical descriptive statistics displayed correctly	Pass
TC30	Run correlation analysis (Pearson & Spearman)	Correlation matrices calculated and displayed	Correlation matrices displayed correctly	Pass

In the second iteration, the system was enhanced to include spreadsheet editing, descriptive statistics, and correlation analysis functionalities. The test cases were condensed into 10 key scenarios for clarity, covering core operations such as handling large datasets (up to 10,000 rows), managing null values, formatting text and numbers, inserting or deleting rows and columns, updating dataset values, and generating statistical analyses. All features tested in this phase worked as expected, demonstrating that the system could efficiently handle larger datasets and perform data analysis.

Table 11. Test Cases for the Third Iteration

Test Case ID	**Test Scenario**	**Expected Result**	**Actual Result**	**Status**
TC31	Open Generate Report Drawer	Drawer opens from the right side	Drawer opened correctly	Pass
TC32	Close Generate Report Drawer	Drawer closes when clicking “X”	Drawer closed successfully	Pass
TC33	Create AI Report (All Dataset)	AI generates a report based on the dataset	AI generated the report correctly	Pass
TC34	Create AI Report (Selected Columns)	AI generates report using selected columns	Report included only selected columns	Pass

TC35	Reset AI Report Form	Displays test results for data distribution	Form reset successfully	Pass
TC36	Stop AI Report while Loading	Stops AI process and retains user input	AI stopped and retained input	Pass
TC37	Regenerate AI Report	Refresh AI report using the same dataset	AI report refreshed correctly	Pass
TC38	Chatbot Interaction	Chatbot answers user queries about dataset	Chatbot responded correctly	Pass
TC39	Normality Test	Displays test results for data distribution	Results displayed correctly	Pass
TC40	Simple Linear Regression	Generates regression output	Regression results displayed correctly	Pass

In the third iteration, the focus shifted to AI-integrated functionalities, including the Generate Report drawer, Chatbot, plot recommendations, and downloading AI-generated reports. The test cases were condensed to 10 scenarios, covering the opening and closing of the report drawer, creating new reports, regenerating or stopping reports during processing, selecting specific columns for analysis, interacting with the chatbot, and downloading reports. In this phase, regression and normality tests were also incorporated to strengthen the statistical analysis features.

IT Expert Testing

One of the IT experts with eight years of experience was not impressed with the first version of our system in terms of the user interface (UI). The concern raised was that the extra spaces at the sides of the pages were not utilized for easier navigation. While some experts appreciated the simplicity of the design, we still considered this feedback and decided to optimize the white spaces to improve navigation and overall usability.

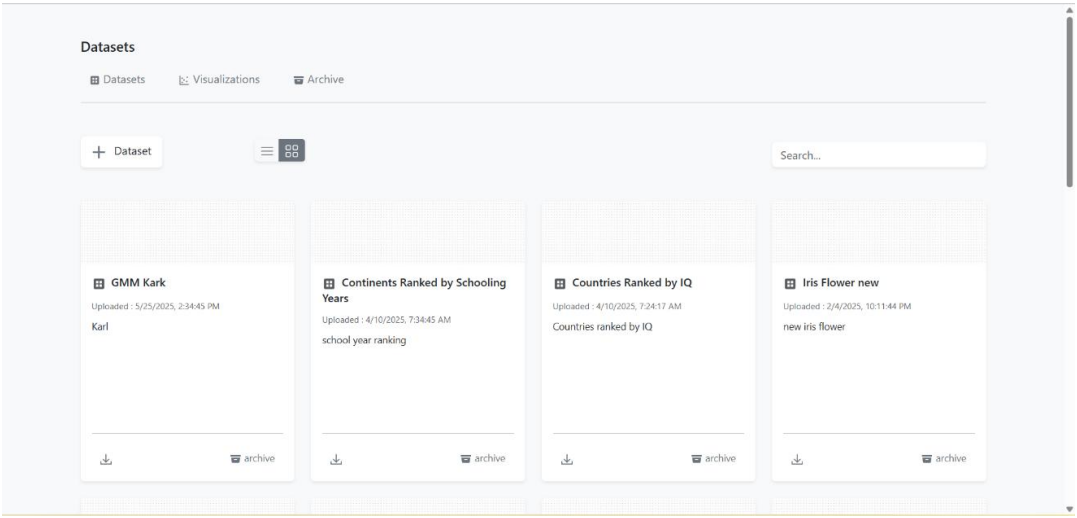


Figure 38. Before Applying Suggestion

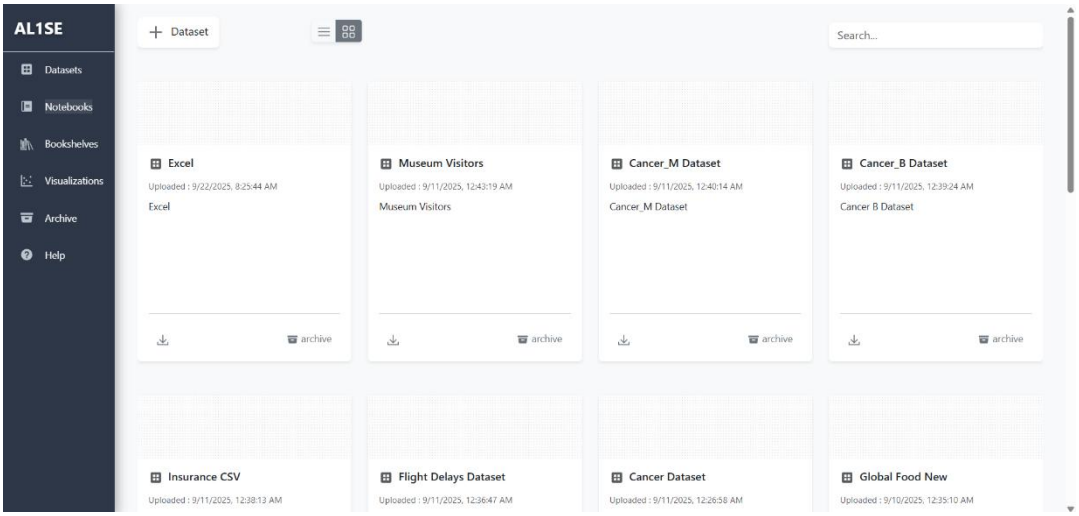


Figure 39. After Applying Suggestion

IT experts evaluated the software again before deployment to ensure its technical quality. The assessment was guided by the applicable ISO/IEC 25010 standards and conducted through a survey form administered via Google Forms. A 4-point Likert Scale was used for responses, where 1 = Strongly Disagree, 2 = Disagree, 3 = Agree, and 4 = Strongly Agree..

Table 12. Weighted Mean and Verbal Interpretation

Weighted Mean Range	Verbal Interpretation
1.00–1.74	Strongly Disagree
1.75–2.49	Disagree
2.50–3.24	Agree
3.25–4.00	Strongly Agree

.

Table 13. IT Experts Testing Result

Criterion	Statements	Average W.M	V.I
Functional Suitability	AL1SE provides the necessary functions for data visualization.	3.55	Strongly Agree
	The software generates accurate visualizations from CSV and Excel datasets.	3.18	Agree
	The features of AL1SE are well-suited for the tasks of data analysis and visualization.	3.45	Agree
TOTAL:	_	**3.39**	**Agree**

Performance Efficiency	AL1SE processes CSV and Excel files with minimal delay.	3.09	Agree
	The system maintains acceptable performance when handling datasets of different sizes.	3.27	Agree
TOTAL:	**_**	**3.18**	**Agree**
Compatibility	AL1SE correctly reads CSV and Excel files.	3.27	Agree
TOTAL:	**_**	**3.27**	**Agree**
Usability	AL1SE layout and interface make it easy to understand and use.	3.09	Agree
	The software can be navigated comfortably without requiring advanced technical knowledge.	2.64	Agree
	The system provides clear labels, prompts, and feedback that guide the user.	3.36	Agree
TOTAL:	**_**	**3.03**	**Agree**
Reliability	AL1SE is dependable for completing visualization tasks with minimal interruption.	3.27	Agree
	AL1SE provided consistent results when processing the same dataset multiple times.	3.18	Agree
	AL1SE recovers properly from file-related or operational errors.	3.27	Agree
TOTAL:	_	**3.24**	**Agree**
Portability	AL1SE runs decently on the installed Windows computer.	3.36	Agree
	The software requires minimal setup or configuration to operate on a supported Windows system.	3.27	Agree
TOTAL:	**_**	**3.32**	**Agree**
	Overall, I am satisfied with the performance of AL1SE.	3.55	Strongly Agree

Satisfaction	I would recommend AL1SE to others who need a data visualization tool.	3.55	Strongly Agree
TOTAL:		**3.55**	**Strongly Agree**

The evaluation of AL1SE by 11 IT Experts produced favorable results across the IS0/IEC 25010. The respondents had varying level of experience, with the highest recorded at eight years, further underscoring the credibility of their evaluation. For Functional Suitability, the overall weighted mean was 3.39 (Agree), indicating that the system ‘s feature were adequate for data visualization and analysis. Performance Efficiency obtained a mean of 3.18 (Agree), showing that the system processes with acceptable speed and consistency. Compatibility received a mean of 3.27 (Agree), confirming that AL1SE accurately reads standard file formats such as CSV and Excel.

For Usability , the overall mean was 3.03 (Agree), suggesting that the system is generally usable, though improvements may be made to better support user without prior statistical knowledge. Reliability was rated 3.34 (Agree), reflecting that AL1SE delivers consistent result with minimal errors. Portability obtained a mean of 3.32 (Agree), indicating that the software can be installed and run acceptably on supported Windows environment without complex setup requirements. Finally, Satisfaction achieved the highest rating, with an overall mean of 3.55 (Strongly Agree), showing that IT experts were highly satisfied with AL1SE and would recommend its use for data visualization tasks.

Overall, the results demonstrate that IT Experts considered AL1SE a functional, reliable, satisfactory tool for a cost-free data visualization option.

4.6 DEPLOYMENT

The first deployment on September 4, 2025 was unsuccessful due to installation issues. Some functions that worked in testing were either missing or not functioning on the designated computers. The developers postponed the deployment to the next day, treating the attempt as a trial run to uncover potential issues.

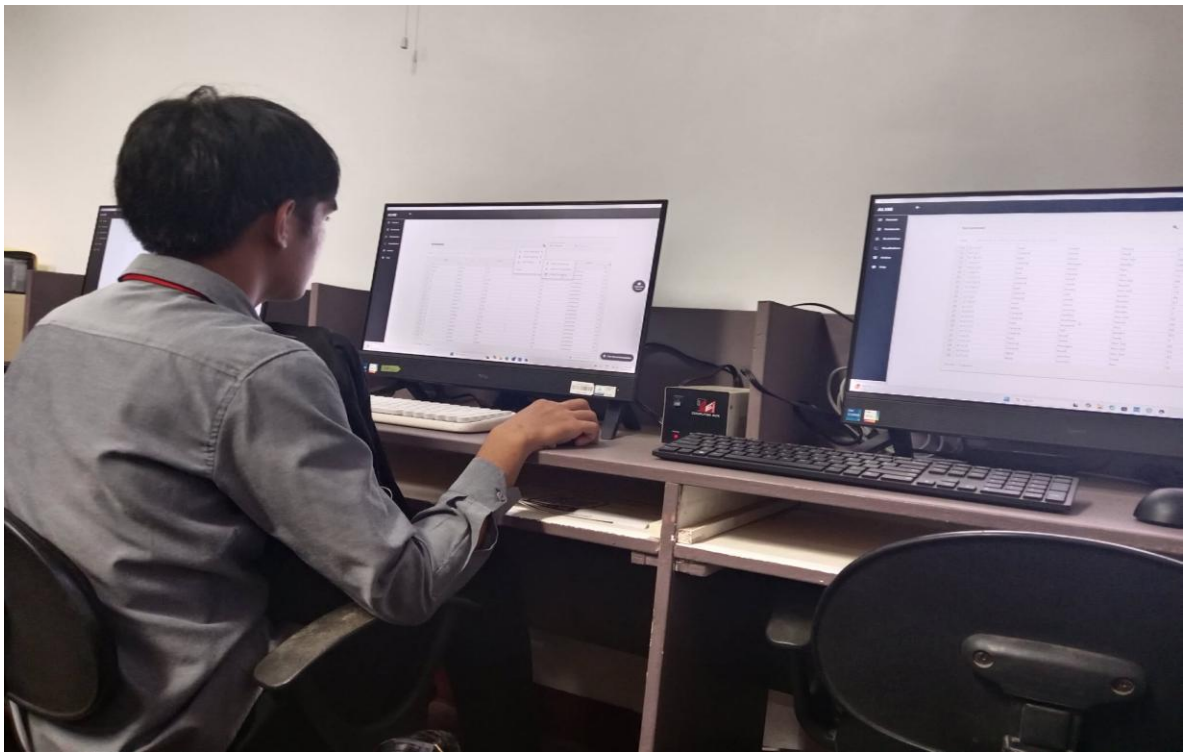


Figure 35. Deployment at the Department of Statistics

The software was deployed officially at the Statistics Department on September 5, 2025. The installable version of the data visualization software, AL1SE, was installed on four designated computers. The development team first ensured that the system was functioning properly on these machines before allowing the faculty to use it independently. Although the software could be evaluated instantly based on the original Technology Acceptance Model (TAM), we still deployed it for five days to give the users ample time to explore the features, identify potential issues, and provide more comprehensive feedback.

4.7 REVIEW/EVALUATION

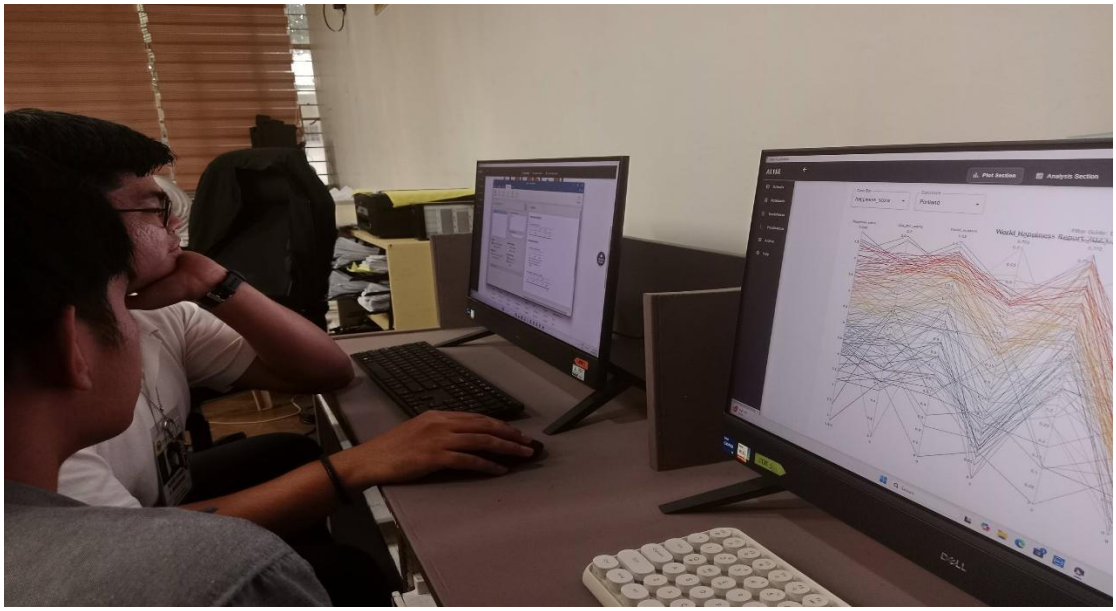


Figure 36. User Testing/Evaluation

After the five-day deployment, user testing was conducted through a physical survey form. Before the evaluation, the researchers demonstrated and explained the software features to ensure that participants were familiar with AL1SE. The users were then given hands-on experience with the software before feedback was gathered, to ensure they had the opportunity to explore the features of the developed data visualization software, AL1SE. The main respondents were faculty members of the Department of Statistics, as they are the primary intended users of the system. However, students were also included to provide additional feedback. A total of 15 responses were gathered: 5 from faculty members and 10 from students, representing 10 percent of the student population (100 students in total, according to the faculty). The survey utilized a Likert scale to measure two criteria based on the original Technology Acceptance Model (TAM): perceived usefulness and perceived ease of use.

Table 9 shows the evaluation results from the Statistics Department faculty, and additional student testers, highlighting positive feedback on AL1SE's usefulness and ease of use, with higher ratings for visualization availability, practicality, and interactive functions.

Table 14. User Testing/ Evaluation Result

User Type: 33.3% Faculty, 66.7% Student		
Experience with Data Visualization Tool: 20% Beginner, 80% Intermediate		
Perceived Usefulness (PU)	**Average W.M**	**V.I**
AL1SE helps me understand my datasets better.	3.04	Agree
The data visualizations I need are available in AL1SE.	3.96	Strongly Agree
AL1SE help me produce accurate and reliable visualizations.	3.04	Agree
I can complete data visualization tasks efficiently using AL1SE.	3.64	Strongly Agree
The cost-free access and no-coding requirement make AL1SE a practical tool for data visualization.	3.92	Strongly Agree
The AI-generated reports in AL1SE provide sufficient detail to help me prepare my own reports more efficiently.	2.96	Agree
The AI plot recommendations in AL1SE help me choose appropriate visualizations for my datasets.	3.68	Strongly Agree
The overall features of AL1SE are useful to perform analysis.	3.08	Agree
Perceived Ease of Use (PEOU)	**Average W.M**	**V.I**
AL1SE is easy to use without prior technical knowledge.	3.08	Agree

AL1SE allows me to customize visualizations according to my needs.	3.20	Agree
AL1SE operates consistently with minimal errors during my use.	3.28	Agree
I was able to learn how to use AL1SE quickly.	2.80	Agree
Interactive visualizations in AL1SE are easy to manipulate and explore.	3.52	Strongly Agree
Overall, using AL1SE feels simple and easy to use.	3.20	Agree

The survey results show that 10 out of 100 students (10%) participated in user testing and evaluation, while five of the six faculty members (83%) responded, ensuring almost complete representation of the faculty perspective. In terms of experience, 80% of the respondents identified as intermediate users, while 20% were beginners.

For Perceived Usefulness (PU), the overall weighted mean was 3.42 (Agree), indicating that respondents generally found AL1SE beneficial in understanding datasets, generating accurate visualizations, and completing tasks with acceptable performance. The highest ratings were given to the availability of visualizations (3.96) and the system's practicality due to cost-free access and no coding requirement (3.92).

For Perceived Ease of Use (PEOU), the overall weighted mean was 3.18 (Agree), showing that respondents considered AL1SE manageable. The highest score was for interactive visualizations being easy to manipulate (3.52), while the lowest was for learning to use AL1SE quickly (2.80). This suggests that although the system is manageable once explored, users may need more time and guidance to become familiar with its functions.

From the open-ended responses, users highlighted AL1SE's intuitive and user-friendly interface, AI integration, and high-quality visualizations as its strongest points. No major challenges were reported.

Overall, both PU and PEOU were rated Agree, reflecting that respondents held a positive perception of AL1SE's usefulness and ease of use, while also recognizing areas for improvement, particularly in guiding beginner users.

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

SUMMARY

Overall Summary

This study developed AL1SE, an AI-integrated data visualization software designed to address the limitations of the Department of Statistics' existing tools, which were costly, internet-dependent, and required programming knowledge. The proposed system enables users to manage datasets, generate visualizations, and access AI-assisted features such as plot recommendations and statistical reporting.

The system was developed using the evolutionary prototyping methodology, guided by insights gathered through interviews and consultations with the client. The development was carried out using Visual Studio Code, using Django for the backend, ReactJS for the frontend, and Python and JavaScript as programming languages, with DeepSeek integrated for AI features.

The study outcome is a functional data visualization software with AI integration, deployed at the Department of Statistics, which provides a cost-free platform for data visualization.

Key Results

The system was evaluated by IT experts based on the ISO/IEC 25010 software quality standards, with most criteria rated "Agree" and overall satisfaction rated "Strongly Agree." Additionally, the system was deployed for five days at the Department of Statistics to allow users to experience its features before they evaluated it.

Further user testing involving faculty members and students, guided by the Technology Acceptance Model (TAM), resulted in overall ratings of "Agree" for both Perceived Usefulness and Perceived Ease of Use. These results demonstrate that AL1SE is an effective, accessible, and well-accepted tool for data visualization by its intended users.

CONCLUSION

The first objective of this study was to provide a specialized and interactive data visualization alternative for the Department of Statistics, which was successfully achieved. AL1SE offers specialized visualization capabilities complemented by AI features and statistical analysis, effectively meeting the department's requirements as confirmed by user evaluation feedback.

The second objective was to develop a stand-alone, offline software, which was also fulfilled. AL1SE operates completely without internet access, ensuring reliability and usability regardless of connectivity, as demonstrated during deployment.

The third objective was to distribute AL1SE free of charge, which was met by making the software freely available to users, thereby removing financial barriers and encouraging broader adoption.

The fourth objective was to eliminate the need for manual programming, and it was accomplished through enabling users to generate visualizations without coding R or Python.

These findings support existing research on AI enhanced visualization tools while uniquely combining offline functionality with AI assistance. Overall, AL1SE successfully met its goals and contributed to making data visualization more accessible and user-friendly for the Department of Statistics.

RECOMMENDATION

Prioritized Recommendation 1: Online Access to External Datasets, which will enable access to online hosted datasets via link for visualization and analysis.

Prioritized Recommendation 2: Enhance the system's AI capabilities by transitioning from simple generative responses to **agentic AI**, which can proactively analyze datasets, suggest appropriate methods, and adapt to different levels of user expertise. This will improve usability and efficiency.

Future Research: Investigate the use of dedicated AI hardware to improve performance and scalability. Exploring this direction could ensure that AL1SE can handle more complex computations and advanced AI functions effectively.

REFERENCES

Analytica - visionary modeling. (n.d.). Retrieved May 26, 2025, from https://analytica.com/

Ariyo Okaiyeto, S., Bai, J., Wang, J., S Mujumdar, A., & Xiao, H. (2025). Success of DeepSeek and potential benefits of free access to AI for global-scale use. *International Journal of Agricultural and Biological Engineering*, *18*(1), 304–306. https://doi.org/10.25165/j.ijabe.20251801.9733

Business Intelligence and Analytics Software - Tableau US. (n.d.). Retrieved April 12, 2025, from https://www.tableau.com/

Caponpon, J. S., Larisma, J., & Ng Lungsod Ng Lucena, D. (2022). MARKET TREND ANALYSIS ON LOCAL RTW SHOP IN LUCENA CITY WITH DATA VISUALIZATION. *Dinsulan*, *1*(1), 1–1. https://ejournals.ph/article.php?id=22417&fbclid=IwY2xjawJNw-tleHRuA2FlbQIxMAABHffRfj_HavBW4_3GZLhY3q93vPUjlaqRge7eVq2z-PZJtSmVvQln54CGqw_aem_42DbaVF1m4r_Yx2ksiBtWA

China connects everything to DeepSeek in nationwide plan - Asia Times. (n.d.). Retrieved March 29, 2025, from https://asiatimes.com/2025/02/china-connects-everything-to-deepseek-in-nationwide-plan/?fbclid=IwY2xjawJNwX1leHRuA2FlbQIxMAABHQW8I_6BuSXYStFlr_FzjyqpGgkTCv1jdnlmVAU1XTnKelssjw1zDVeLkw_aem_rp5x7qAc73cWRT6nXBZmoA#

Data Analytics and AI Platform | Altair RapidMiner. (n.d.). Retrieved May 26, 2025, from https://altair.com/altair-rapidminer

Gao, T., Jin, J., Ke, Z. T., & Moryoussef, G. (2025). *A Comparison of DeepSeek and Other LLMs*. http://arxiv.org/abs/2502.03688

Goal 4 | Department of Economic and Social Affairs. (n.d.). Retrieved March 28, 2025, from https://sdgs.un.org/goals/goal4?fbclid=IwY2xjawJTpeFleHRuA2FlbQIxMAABHQf8MALjFJwo8qgHDS-YC2AQPoSanuFccRGQcI-dhclufPTUTPcReEcQlg_aem_TWreMdSF1QtMNc77uO598Q#progress_and_info

Hayder, W. (2025). Highlighting DeepSeek-R1: Architecture, Features and Future Implications. *International Journal of Computer Science and Mobile Computing*, *14*, 1–13. https://doi.org/10.47760/ijcsmc.2025.v14i02.001

IBM SPSS Statistics. (n.d.). Retrieved May 26, 2025, from https://www.ibm.com/products/spss-statistics

jamovi - open statistical software for the desktop and cloud. (n.d.). Retrieved March 29, 2025, from https://www.jamovi.org/

Poddar, A., & Poddar, M. (2020). Covid-19 Data Visualization and Data Analytics with a Smart Standalone Mobile Application. *2020 IEEE 17th India Council International Conference, INDICON 2020*. https://doi.org/10.1109/INDICON49873.2020.9342143

Power BI - Data Visualization | Microsoft Power Platform. (n.d.). Retrieved April 12, 2025, from https://www.microsoft.com/en-us/power-platform/products/power-bi

RAWGraphs. (n.d.). Retrieved March 29, 2025, from https://www.rawgraphs.io/

Shin, S. (2024). Simulation, Representation, and Automation: Human-Centered Artificial Intelligence for Augmenting Visualization Design. *ProQuest LLC*.

Sommerville, Ian. (2011). *Software engineering*. Pearson.

Steed, C. A. (2025). Interactive data visualization. *Data Analytics for Intelligent Transportation Systems*, 185–211. https://doi.org/10.1016/B978-0-443-13878-2.00002-3

Wu, F. (2025). DeepSeek: Toward Global Education Empowerment for the Whole Society. *Frontiers of Digital Education*, *2*(2), 26. https://doi.org/10.1007/s44366-025-0062-y

APPENDICES

Appendix A.

Sample Outputs

Column: culmen_length_mm | numeric (0.95)

Count	336
Min	32.1000
Max	59.6000
Range	27.5000
Mean	43.9815
Median	44.5000
Mode	41.1000
Variance	29.6909
Standard Deviation	5.4489
Skewness	0.0466
Kurtosis	-0.8781
IQR	9.1500
Q1 (25%)	39.4000
Q2 (50%)	44.5000
Q3 (75%)	48.5500

Descriptive Statistics Numerical

Spearman Correlation Results

Variable 1	Variable 2	Spearman's rho	P-value	Strength
culmen_length_mm	culmen_depth_mm	-0.21	0.002	Very Weak
culmen_length_mm	flipper_length_mm	0.682	0.002	Moderate
culmen_length_mm	body_mass_g	0.588	0.002	Moderate
culmen_depth_mm	flipper_length_mm	-0.522	0.002	Moderate
culmen_depth_mm	body_mass_g	-0.432	0.002	Weak
flipper_length_mm	body_mass_g	0.844	0.002	Strong

Spearman Correlation Analysis

Correlation ▼

Pearson's Correlations

			Pearson's r	p
culmen_length_mm	-	culmen_depth_mm	−0.232	< .001
culmen_length_mm	-	flipper_length_mm	0.654	< .001
culmen_length_mm	-	body_mass_g	0.592	< .001
culmen_depth_mm	-	flipper_length_mm	−0.582	< .001
culmen_depth_mm	-	body_mass_g	−0.472	< .001
flipper_length_mm	-	body_mass_g	0.872	< .001

JASP Pearson Correlation Analysis

Descriptive Statistics ▼

Descriptive Statistics

	culmen_length_mm
Valid	336
Missing	0
Mode	41.100[a]
Median	44.500
Mean	43.982
Std. Deviation	5.457
IQR	9.075
Variance	29.780
Skewness	0.047
Std. Error of Skewness	0.133
Kurtosis	−0.878
Std. Error of Kurtosis	0.265
Range	27.500
Minimum	32.100
Maximum	59.600
25th percentile	39.450
50th percentile	44.500
75th percentile	48.525

[a] The mode is computed assuming that variables are discreet.

JASP Descriptive Statistics

STATISTICAL ANALYSIS REPORT

Dataset: Normality Tests - sepal_length

Generated: 10/9/2025

Normality Tests

Test	Statistic	p-value / crit	Decision	Note
Shapiro–Wilk	0.9761	1.02e-2	Reject H0 (not normal)	—
Kolmogorov–Smirnov	0.0887	1.78e-1	Fail to reject H0 (normal)	Parameters estimated; p-value is approximate (Lilliefors).
Anderson–Darling (crit 5%)	0.8892	crit used	Reject H0 (not normal)	Compared against critical value 0.7670

Verified By: _____

Signature

Date: _____

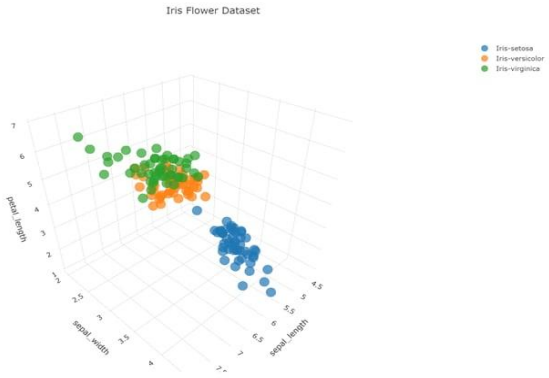
Sample Downloaded Report (Normality Test)

STATISTICAL ANALYSIS REPORT

Dataset: Iris Flower Dataset

Generated: 10/9/2025

Data Visualization



AI Analysis

Alright, so I'm trying to figure out how to analyze the Iris Flower Dataset. Let me start by understanding what the user is asking for. They mentioned that they want a comprehensive analysis of the dataset using a 3D scatter

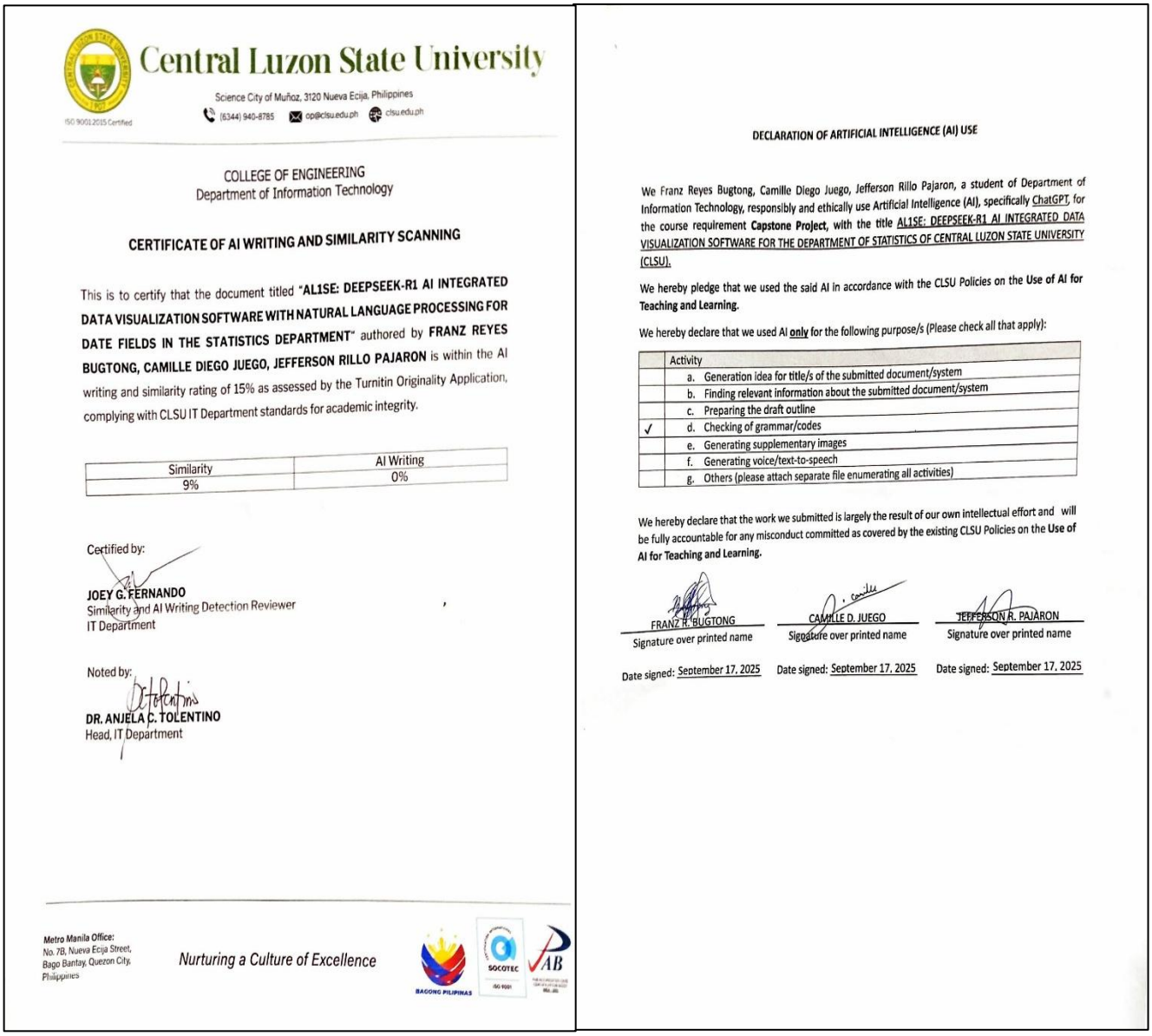
Sample Downloaded Report with Visual

Appendix B.

Deployment/ Demo Documentation

Appendix C.

Certificate of AI Writing and Similarity Scanning



Central Luzon State University

Science City of Muñoz, 3120 Nueva Ecija, Philippines
(6344) 940-8785 op@clsu.edu.ph clsu.edu.ph
ISO 9001:2015 Certified

COLLEGE OF ENGINEERING
Department of Information Technology

CERTIFICATE OF AI WRITING AND SIMILARITY SCANNING

This is to certify that the document titled "**AL1SE: DEEPSEEK-R1 AI INTEGRATED DATA VISUALIZATION SOFTWARE WITH NATURAL LANGUAGE PROCESSING FOR DATE FIELDS IN THE STATISTICS DEPARTMENT**" authored by **FRANZ REYES BUGTONG, CAMILLE DIEGO JUEGO, JEFFERSON RILLO PAJARON** is within the AI writing and similarity rating of 15% as assessed by the Turnitin Originality Application, complying with CLSU IT Department standards for academic integrity.

Similarity	AI Writing
9%	0%

Certified by:

JOEY G. FERNANDO
Similarity and AI Writing Detection Reviewer
IT Department

Noted by:

DR. ANJELA C. TOLENTINO
Head, IT Department

Metro Manila Office:
No. 7B, Nueva Ecija Street,
Bago Bantay, Quezon City,
Philippines

Nurturing a Culture of Excellence

BAGONG PILIPINAS SOCOTEC PAB

DECLARATION OF ARTIFICIAL INTELLIGENCE (AI) USE

We Franz Reyes Bugtong, Camille Diego Juego, Jefferson Rillo Pajaron, a student of Department of Information Technology, responsibly and ethically use Artificial Intelligence (AI), specifically ChatGPT, for the course requirement **Capstone Project**, with the title AL1SE: DEEPSEEK-R1 AI INTEGRATED DATA VISUALIZATION SOFTWARE FOR THE DEPARTMENT OF STATISTICS OF CENTRAL LUZON STATE UNIVERSITY (CLSU).

We hereby pledge that we used the said AI in accordance with the CLSU Policies on the **Use of AI for Teaching and Learning.**

We hereby declare that we used AI **only** for the following purpose/s (Please check all that apply):

	Activity
	a. Generation idea for title/s of the submitted document/system
	b. Finding relevant information about the submitted document/system
	c. Preparing the draft outline
✓	d. Checking of grammar/codes
	e. Generating supplementary images
	f. Generating voice/text-to-speech
	g. Others (please attach separate file enumerating all activities)

We hereby declare that the work we submitted is largely the result of our own intellectual effort and will be fully accountable for any misconduct committed as covered by the existing CLSU Policies on the **Use of AI for Teaching and Learning.**

FRANZ R. BUGTONG
Signature over printed name
Date signed: September 17, 2025

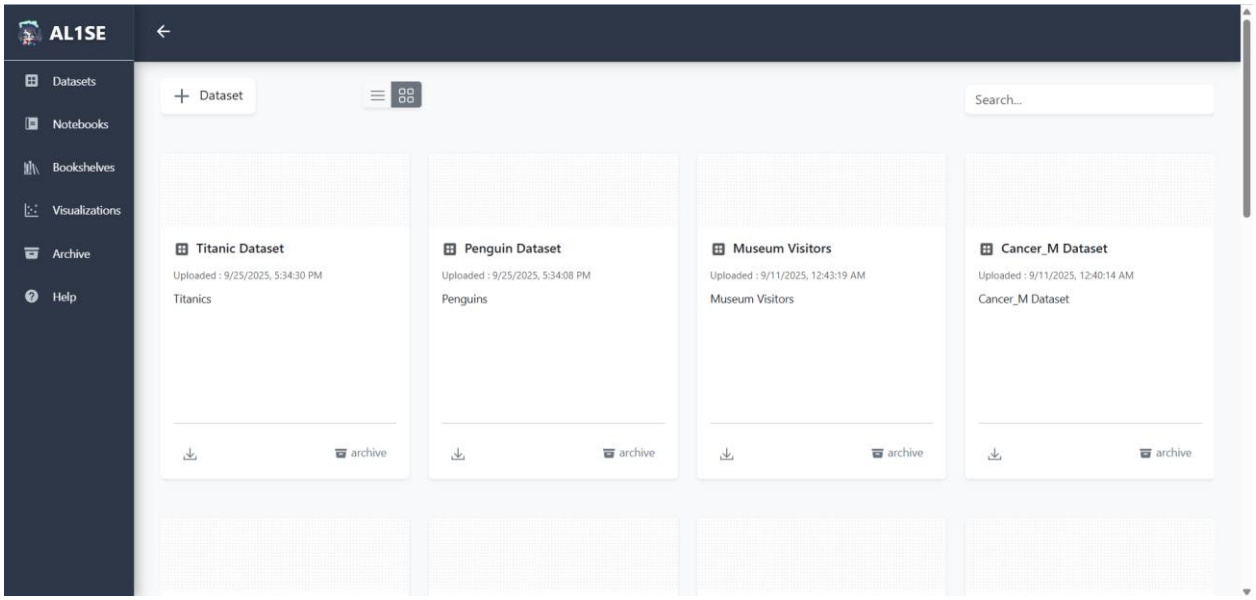
CAMILLE D. JUEGO
Signature over printed name
Date signed: September 17, 2025

JEFFERSON R. PAJARON
Signature over printed name
Date signed: September 17, 2025

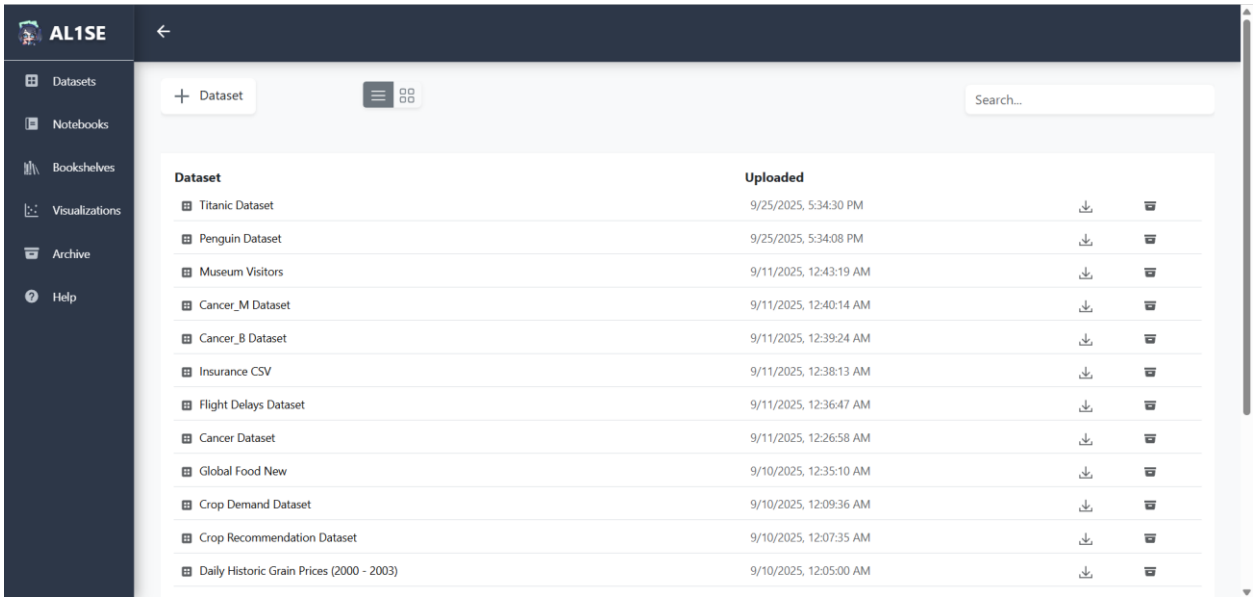
Appendix D.

AL1SE Complete User Interface (UI)

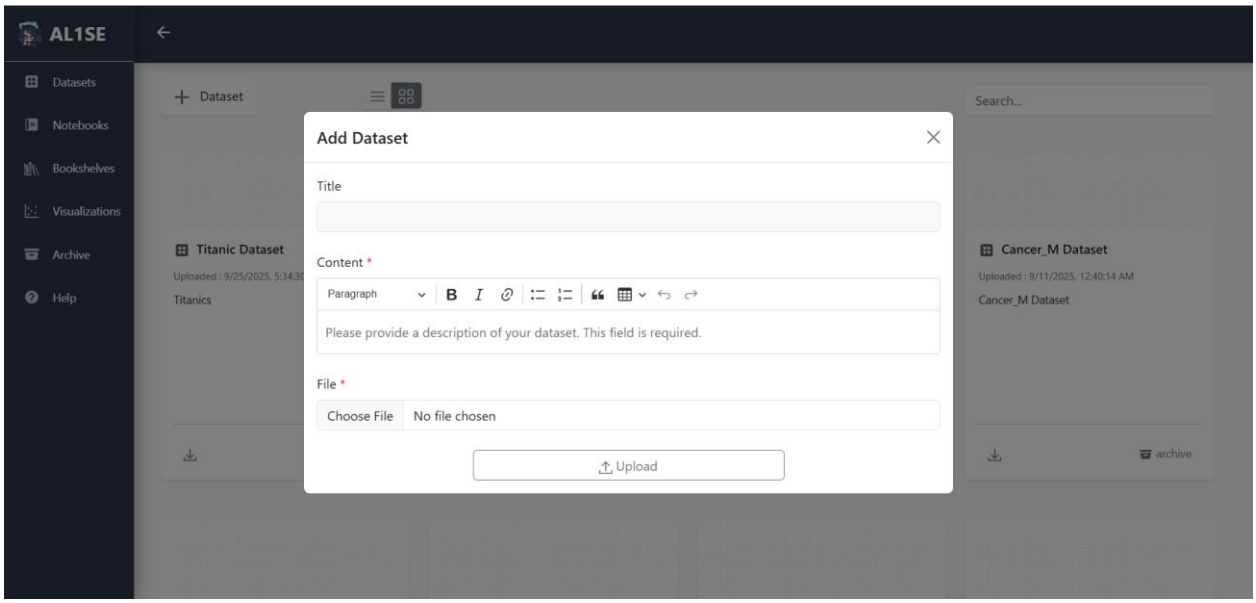
AL1SE Screenshots



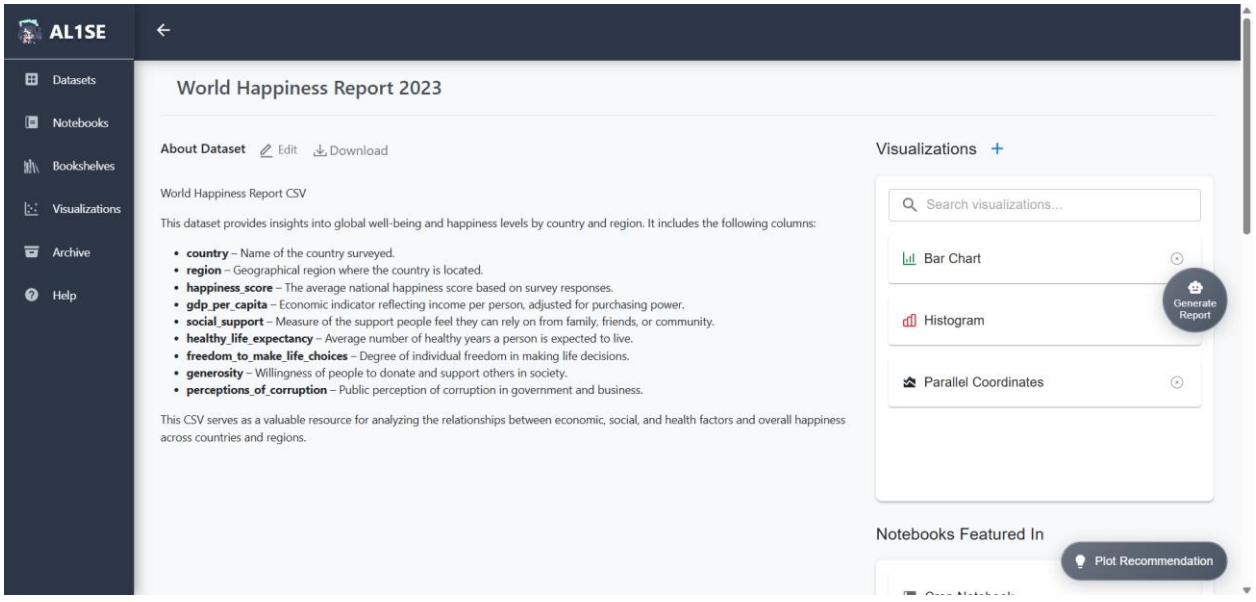
Datasets Page (Grid View)



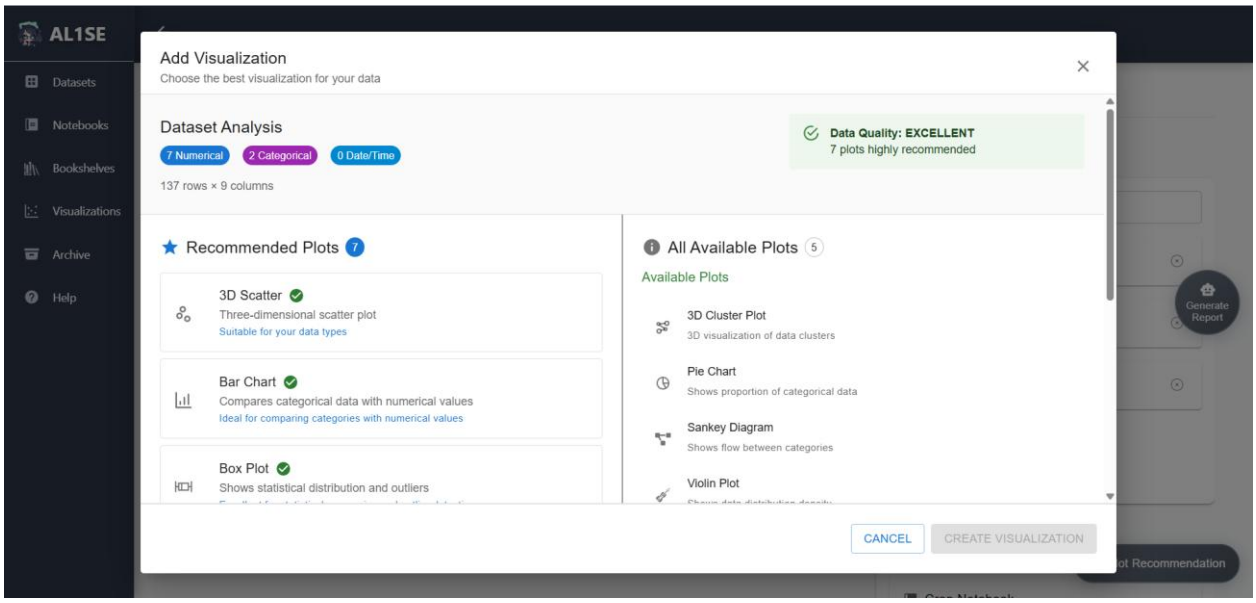
Datasets Page (Table View)



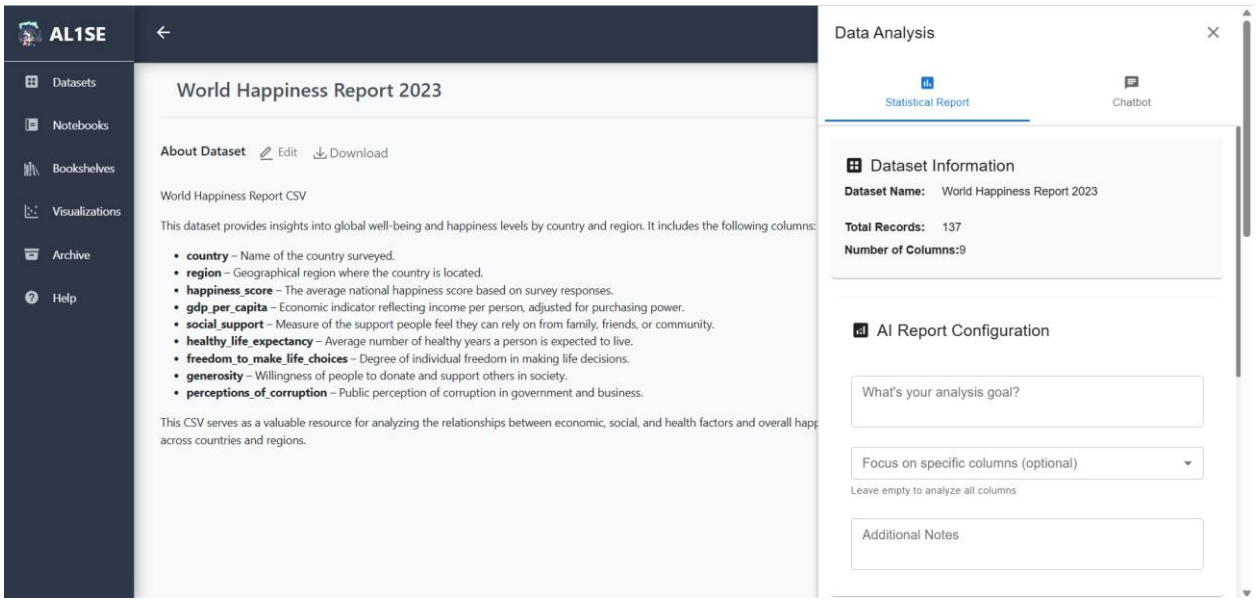
Add Dataset Modal



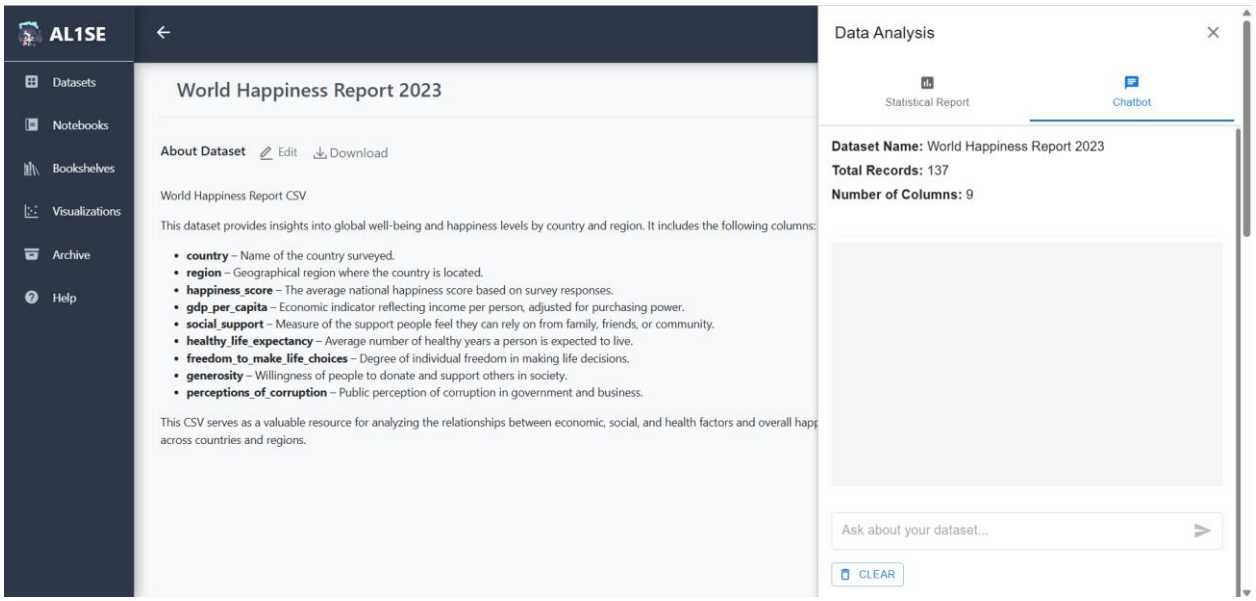
Dataset Page



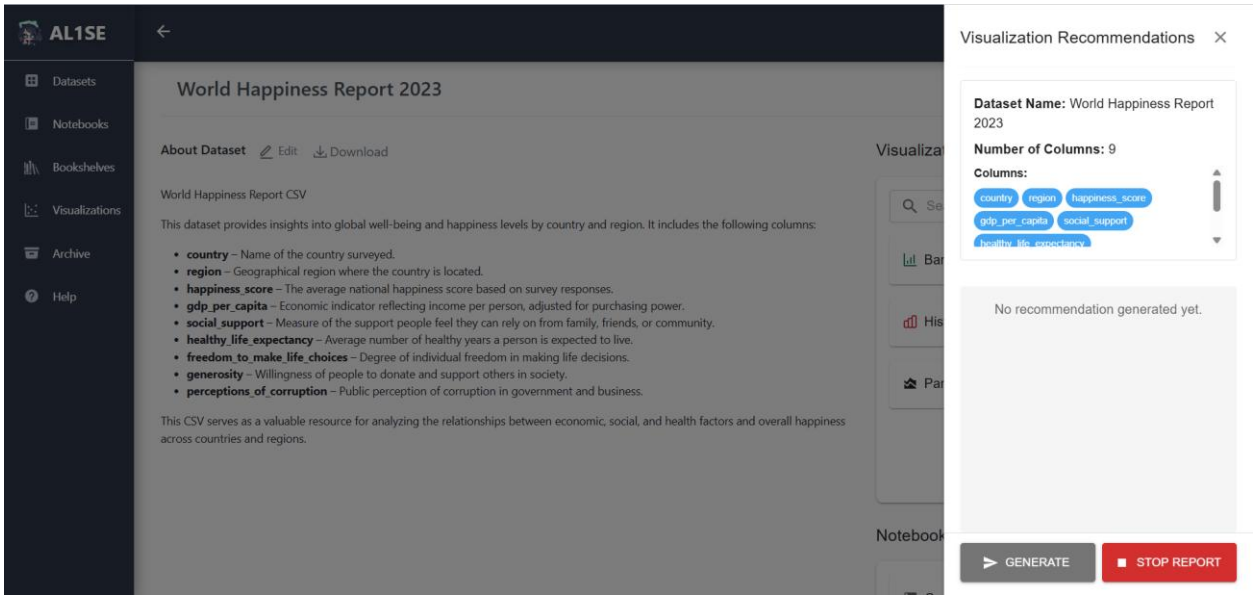
Add Visualization Modal



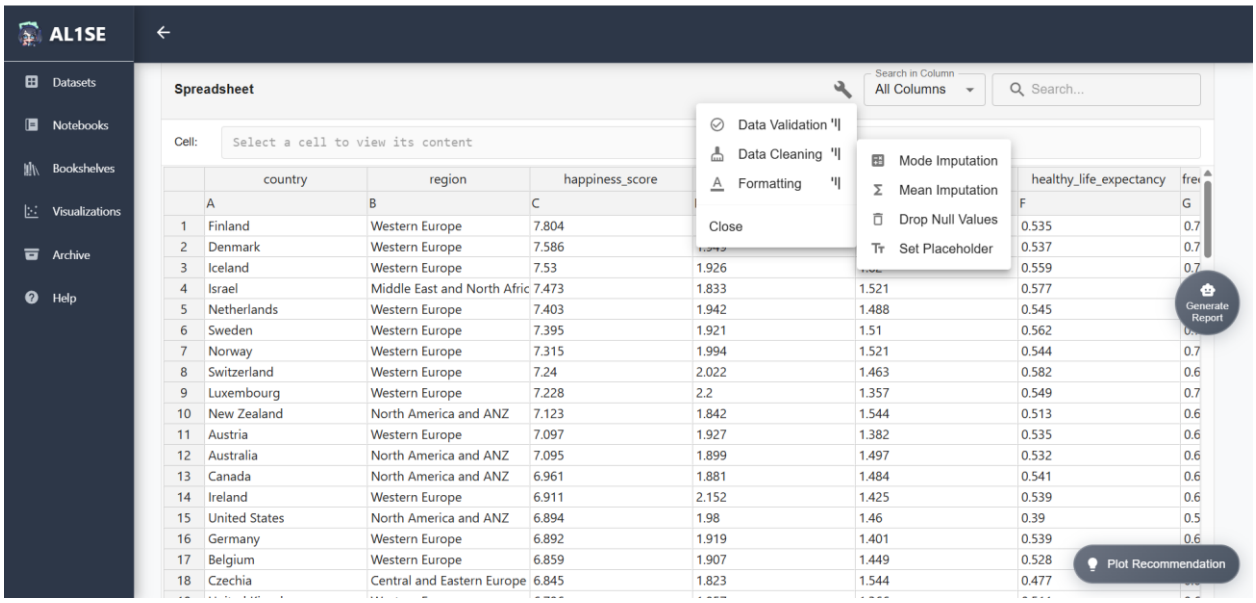
Dataset Page (Generate Report Drawer, Statistical Report Tab)



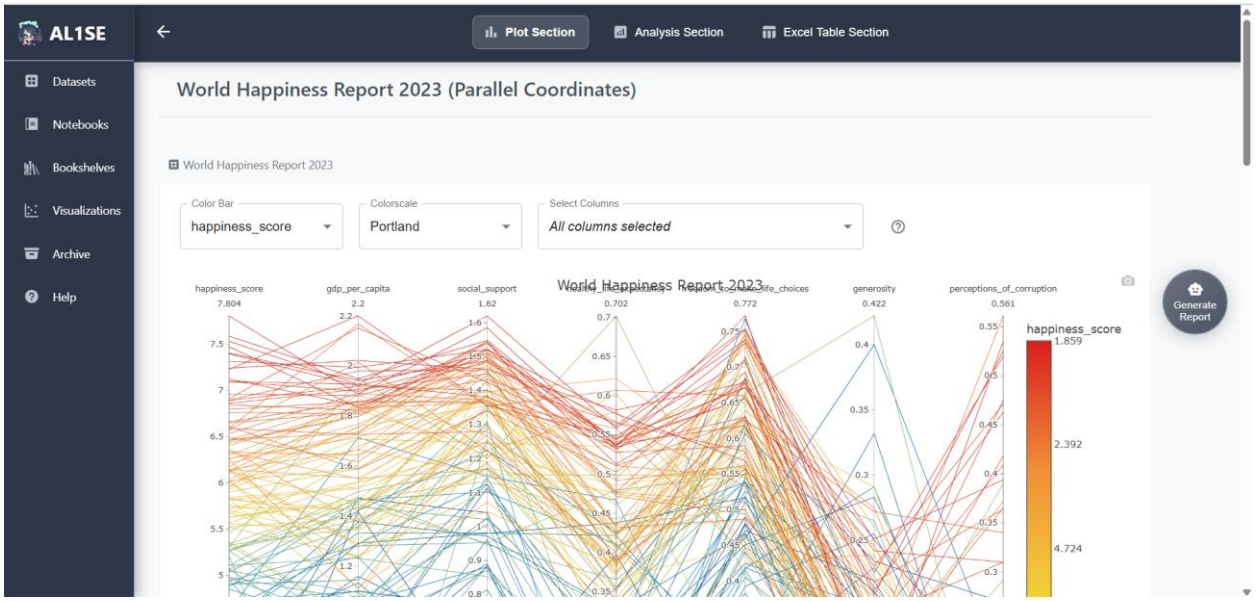
Dataset Page (Generate Report Drawer, Chatbot Tab)



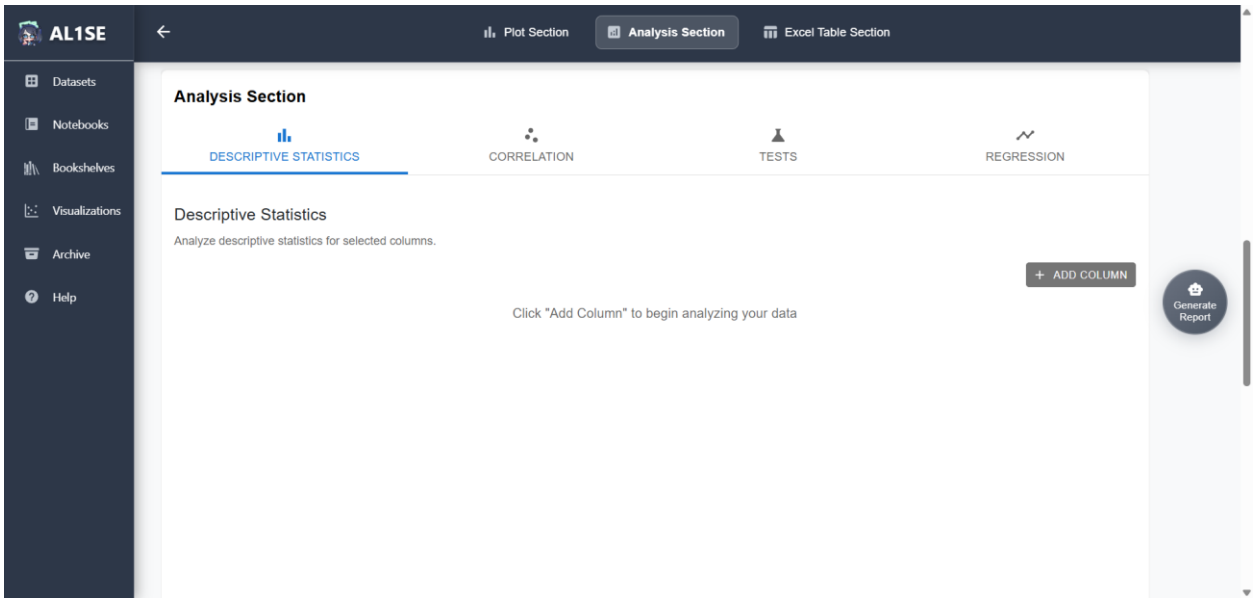
Plot Recommendation Drawer



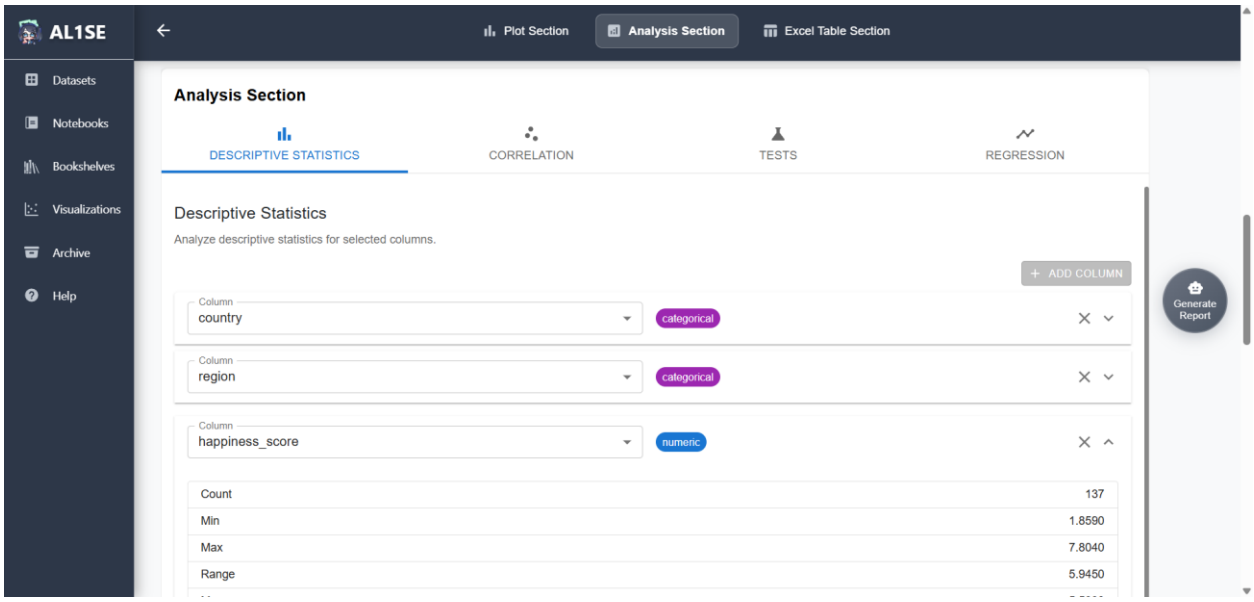
Dataset Page (Excel Table Component)



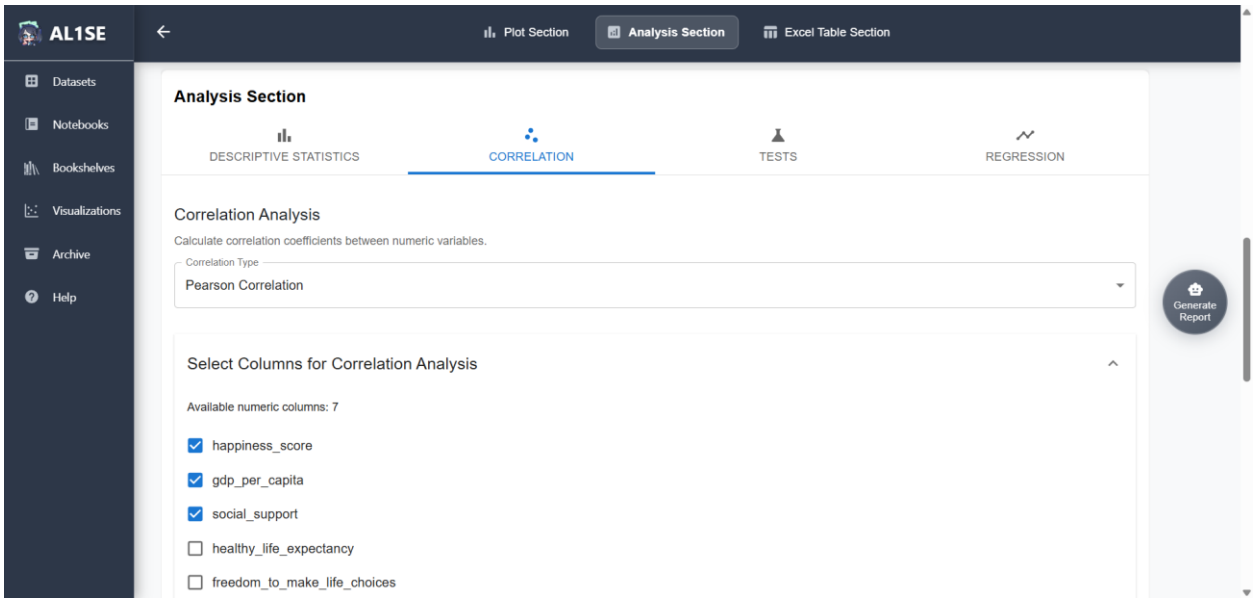
Datavis Page (Plot Section)



Datavis Page (Analysis Section)

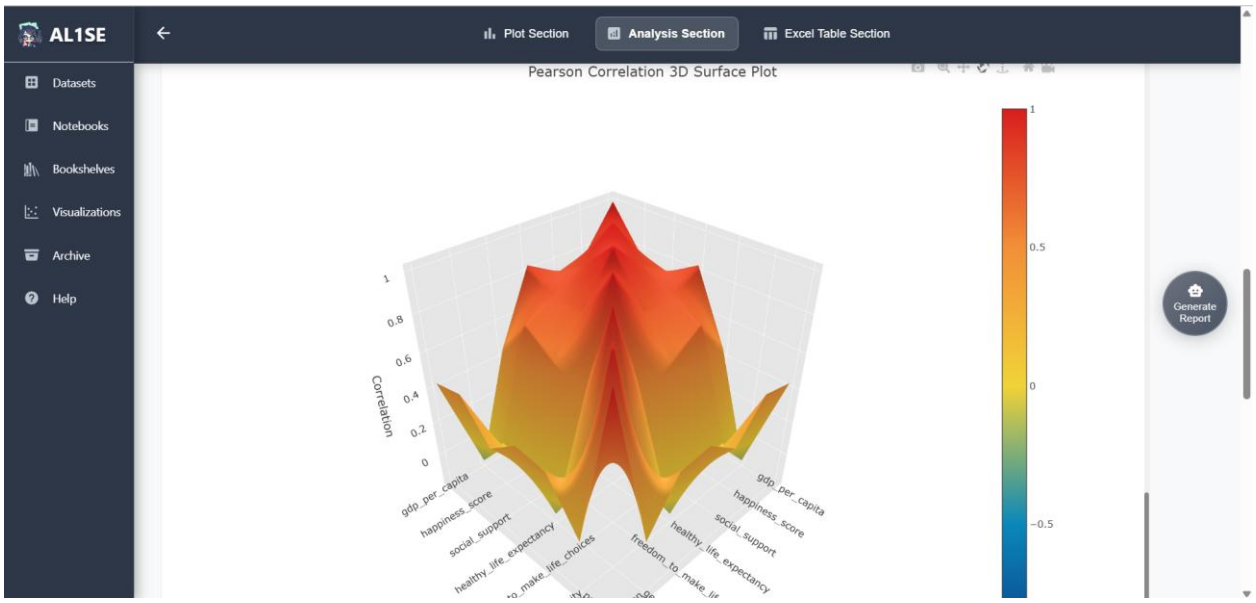


Datavis Page (Analysis Section, Descriptive Statistics Tab)

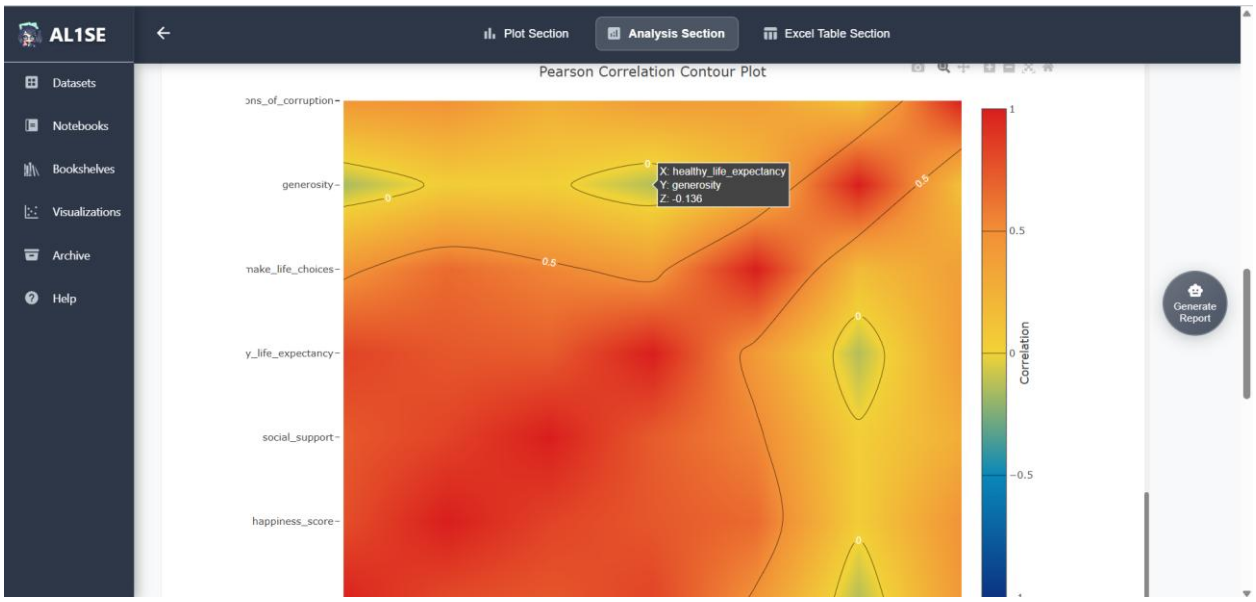


Datavis Page (Analysis Section, Correlation Tab)

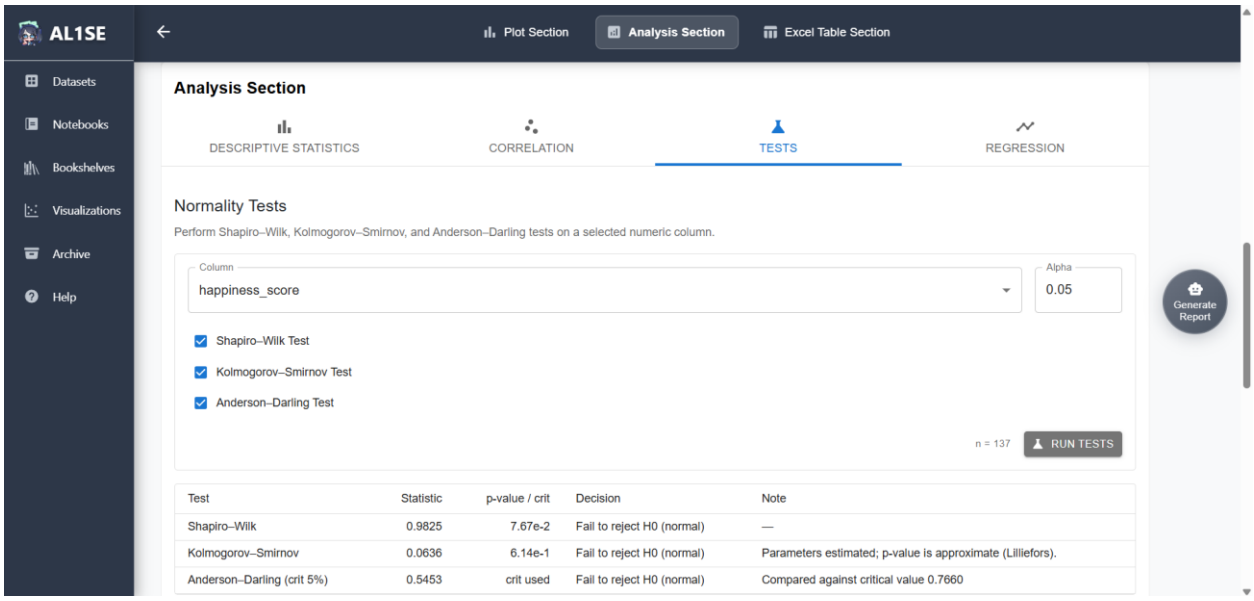
Datavis Page (Analysis Section, Correlation Tab, Heatmap Correlation)



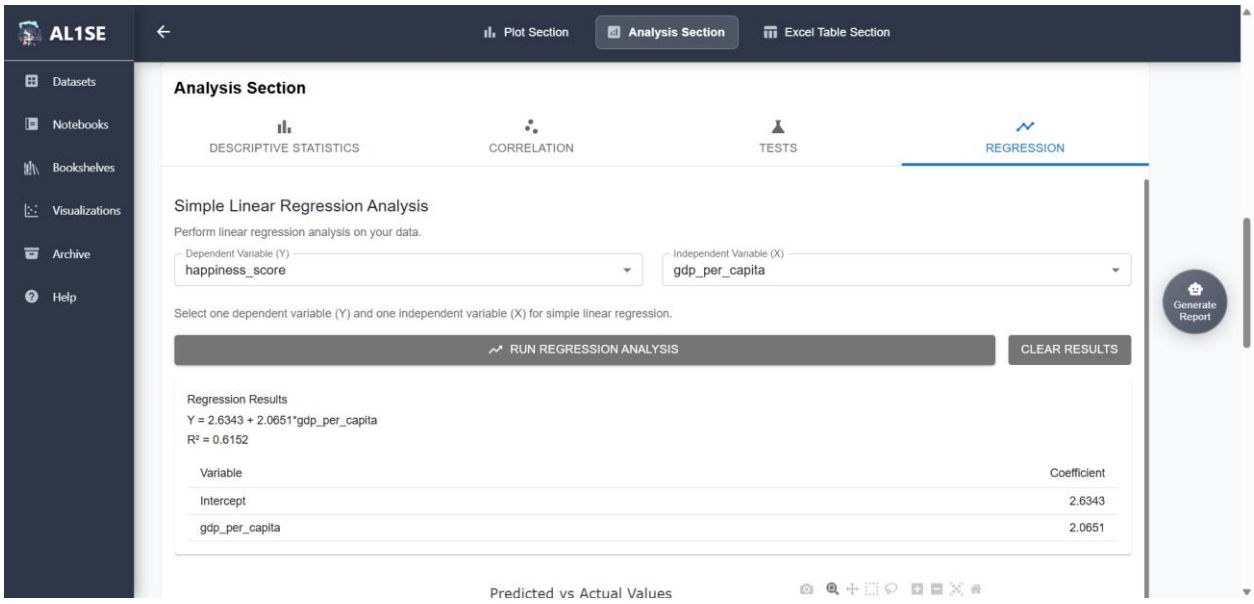
Datavis Page (Analysis Section, Correlation Tab, 3D Surface Correlation)



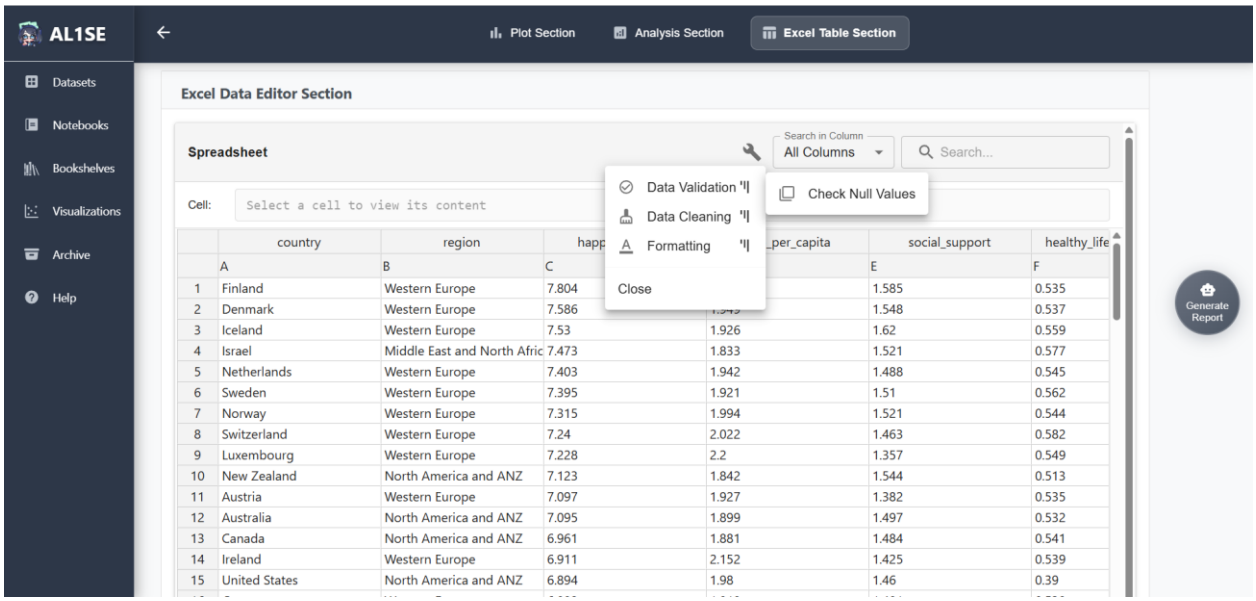
Datavis Page (Analysis Section, Correlation Tab, Contour Correlation)



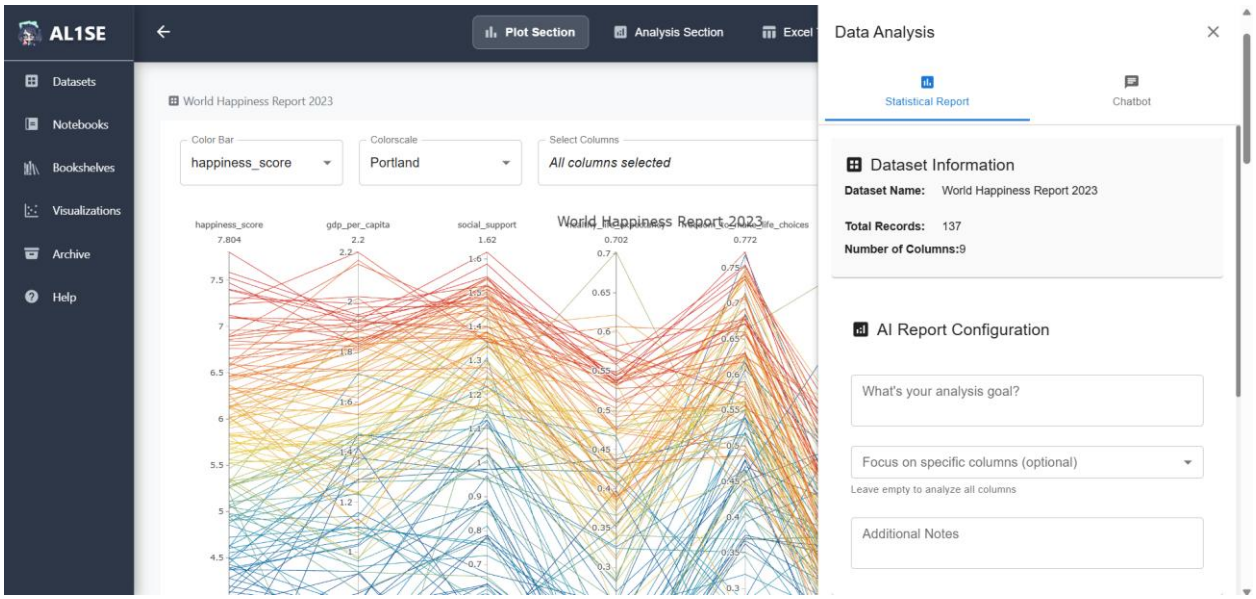
Datavis Page (Analysis Section, Normality Tests Tab)



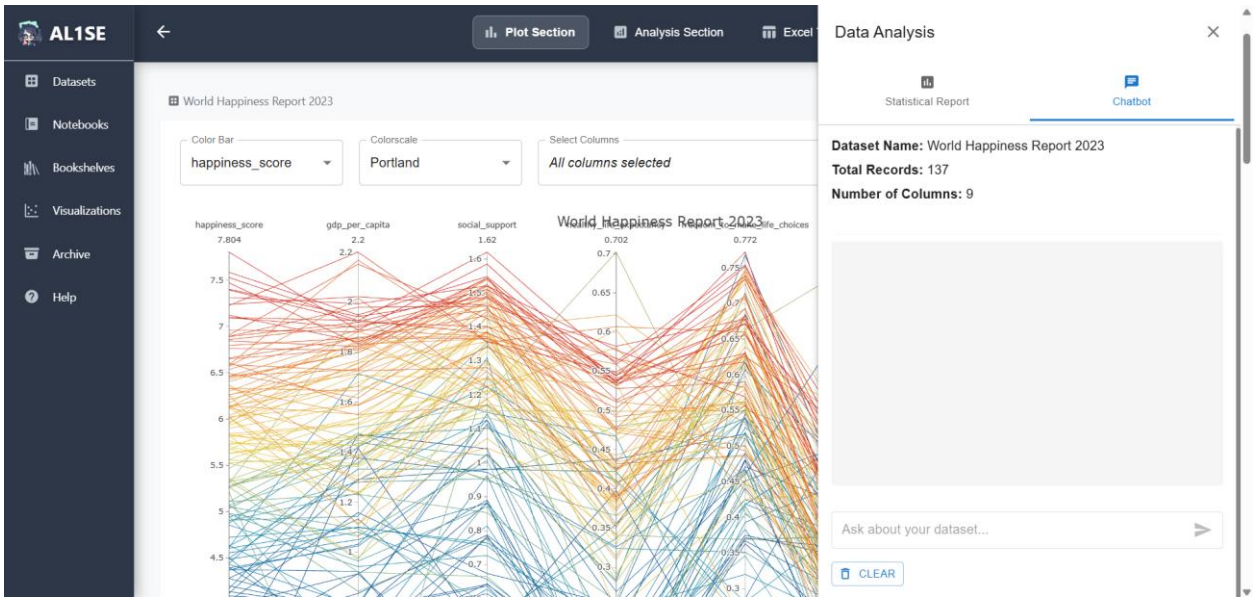
Datavis Page (Analysis Section, Regression Tab)



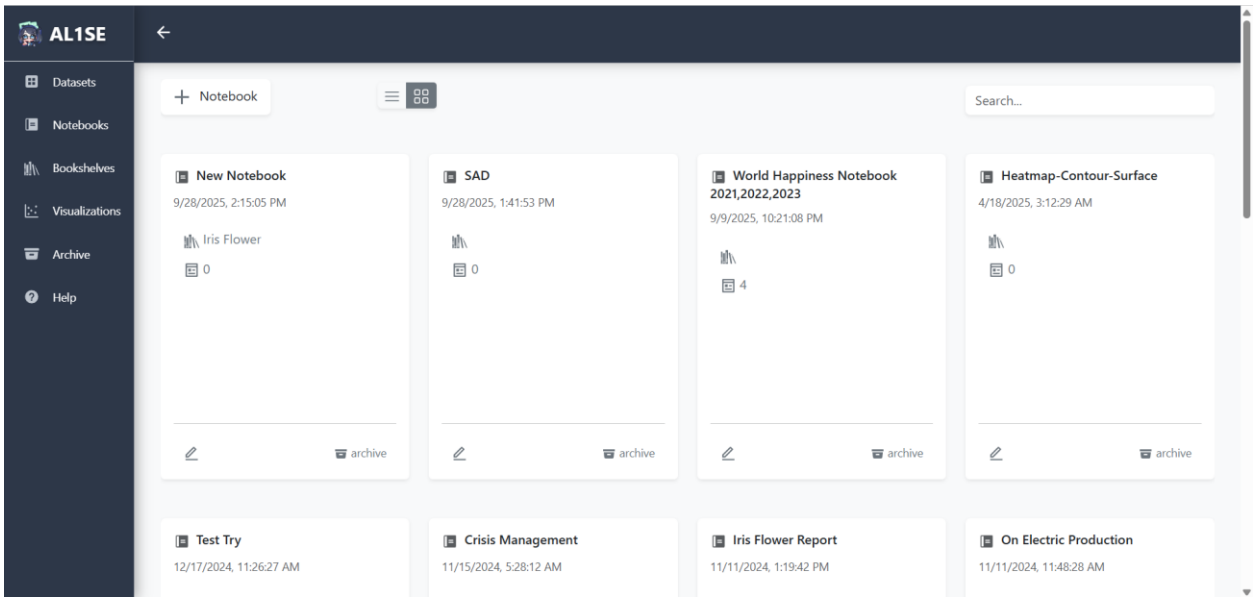
Datavis Page (Excel Table Section)



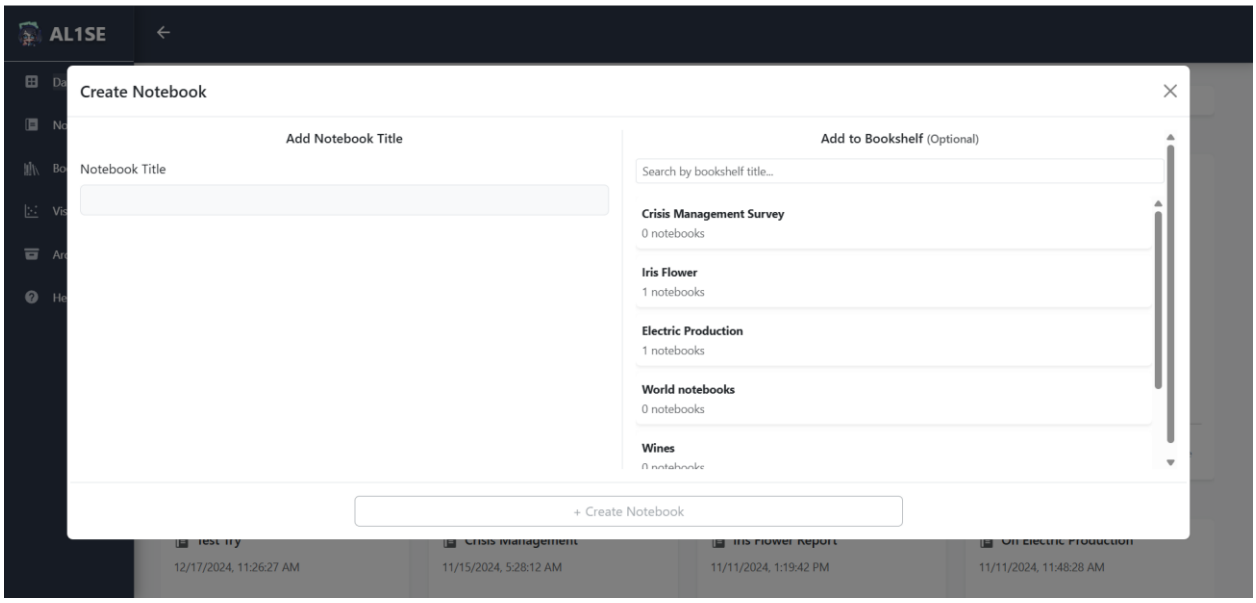
Datavis Page (Generate Report Drawer, Statistical Report Tab)



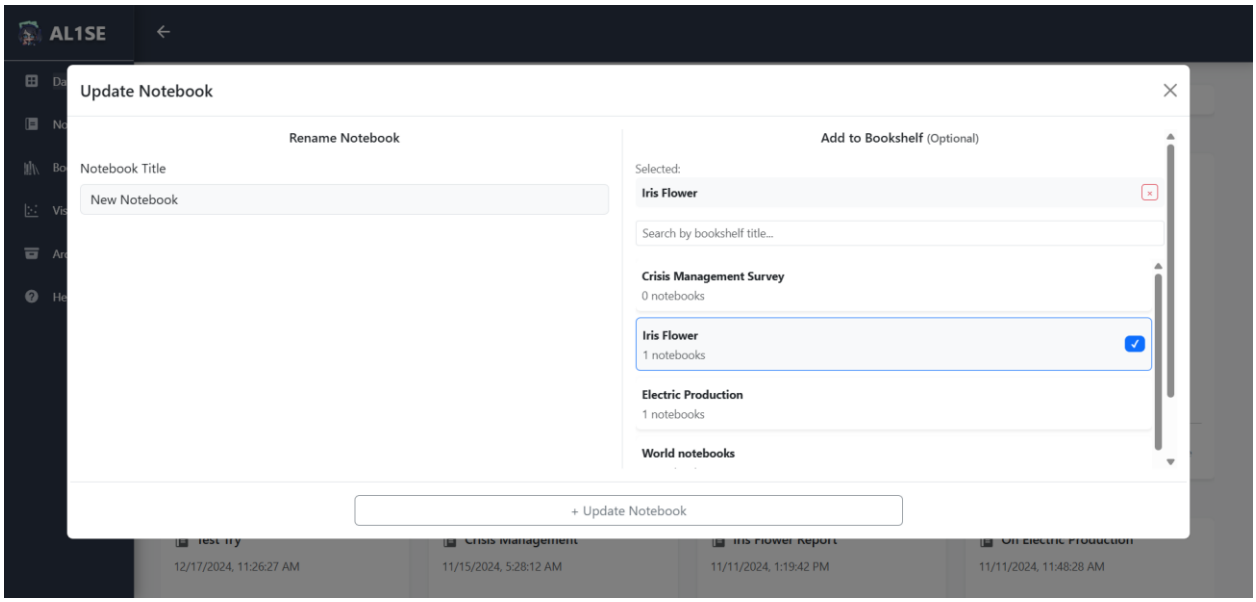
Datavis Page (Generate Report Drawer, Chatbot Tab)



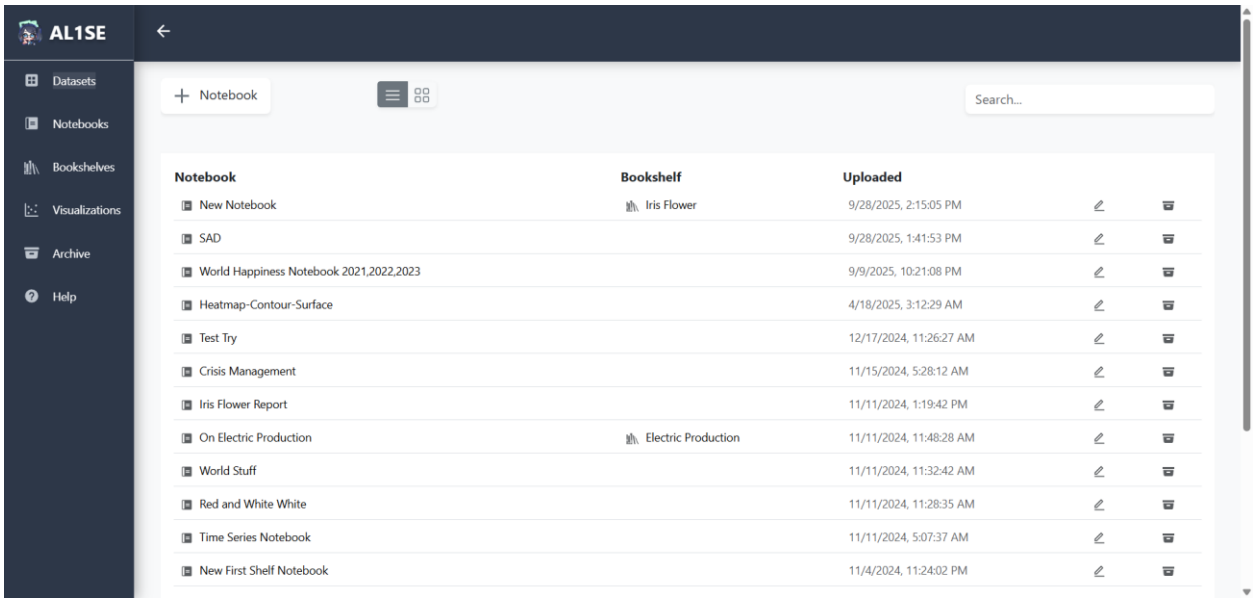
Notebooks Page (Grid View)



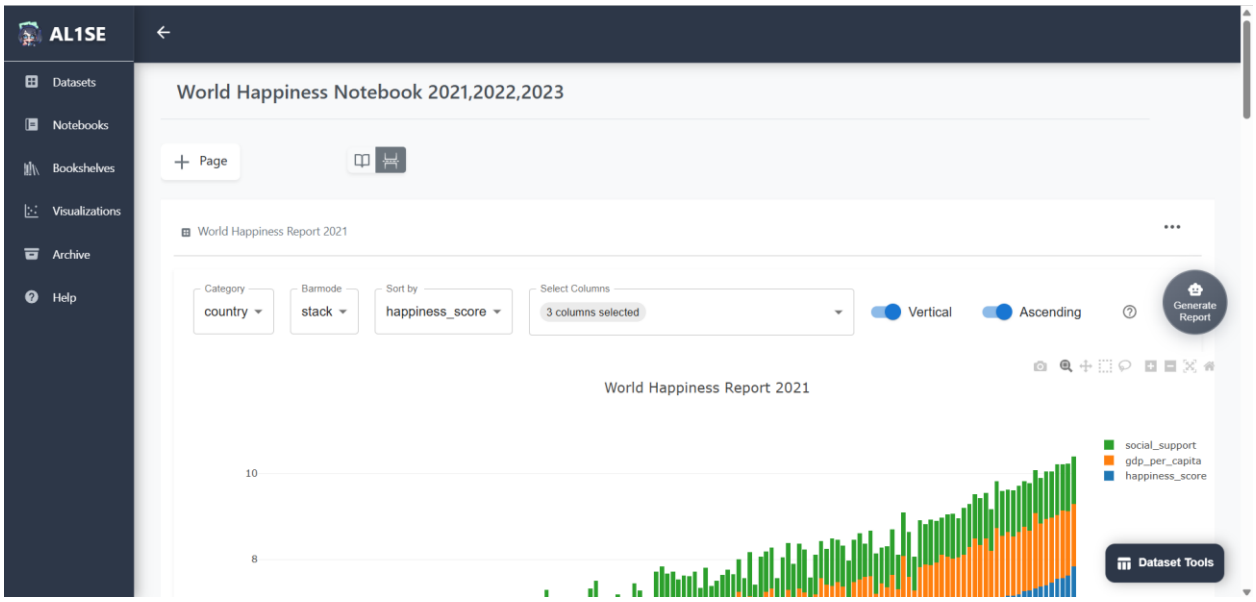
Create Notebook Modal



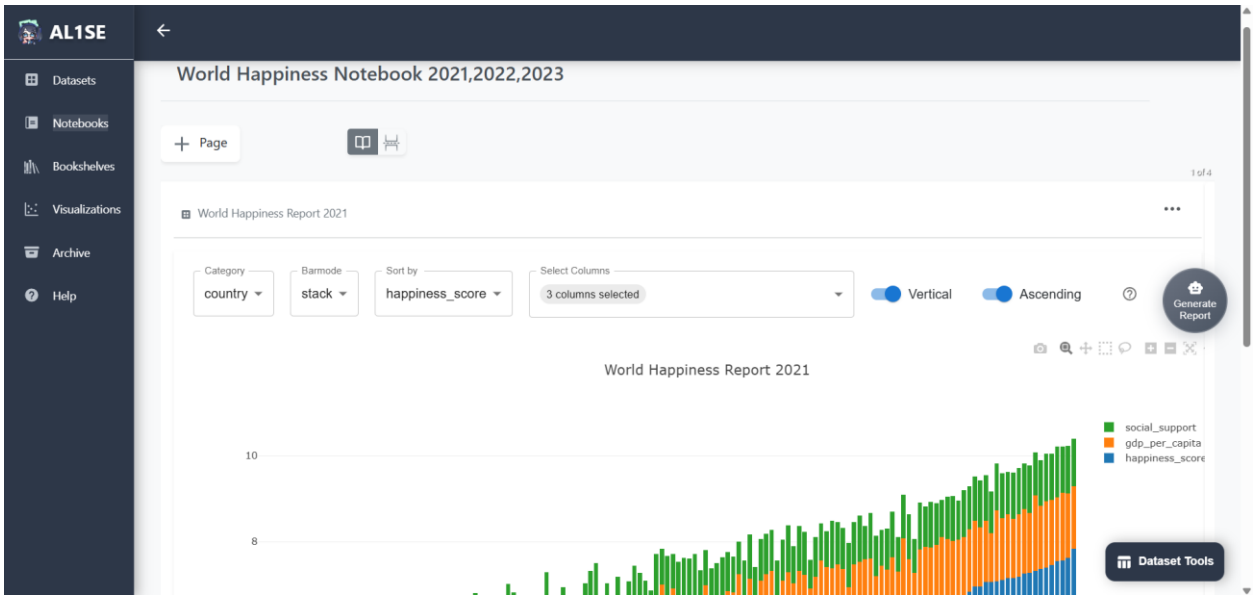
Update Notebook Modal



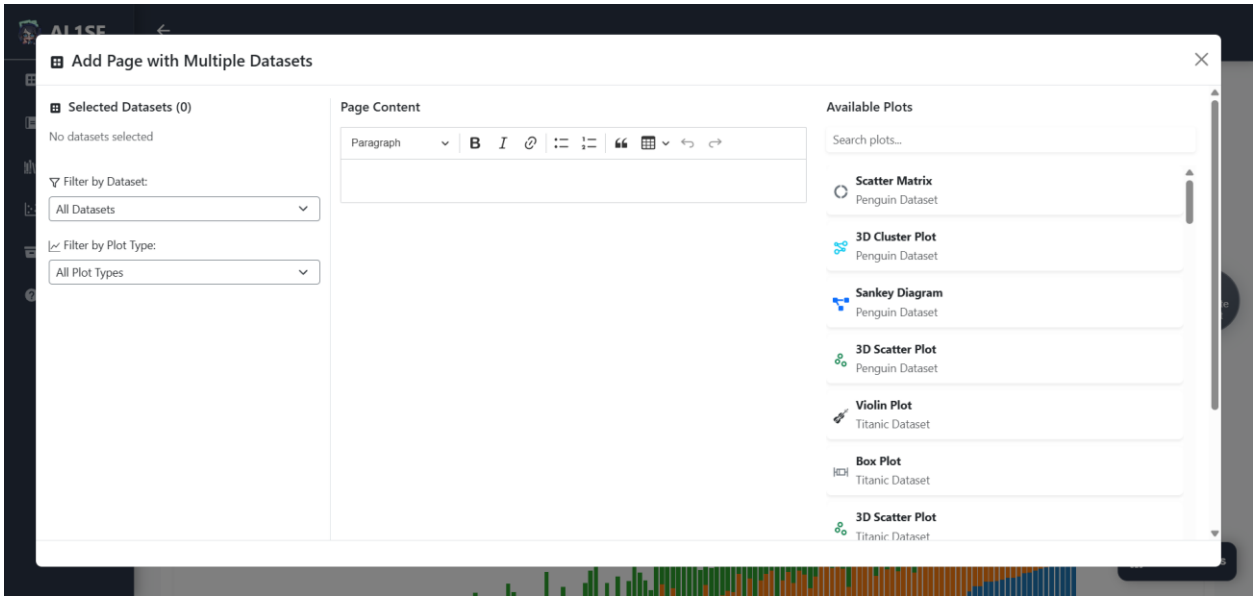
Notebooks Page (Table View)



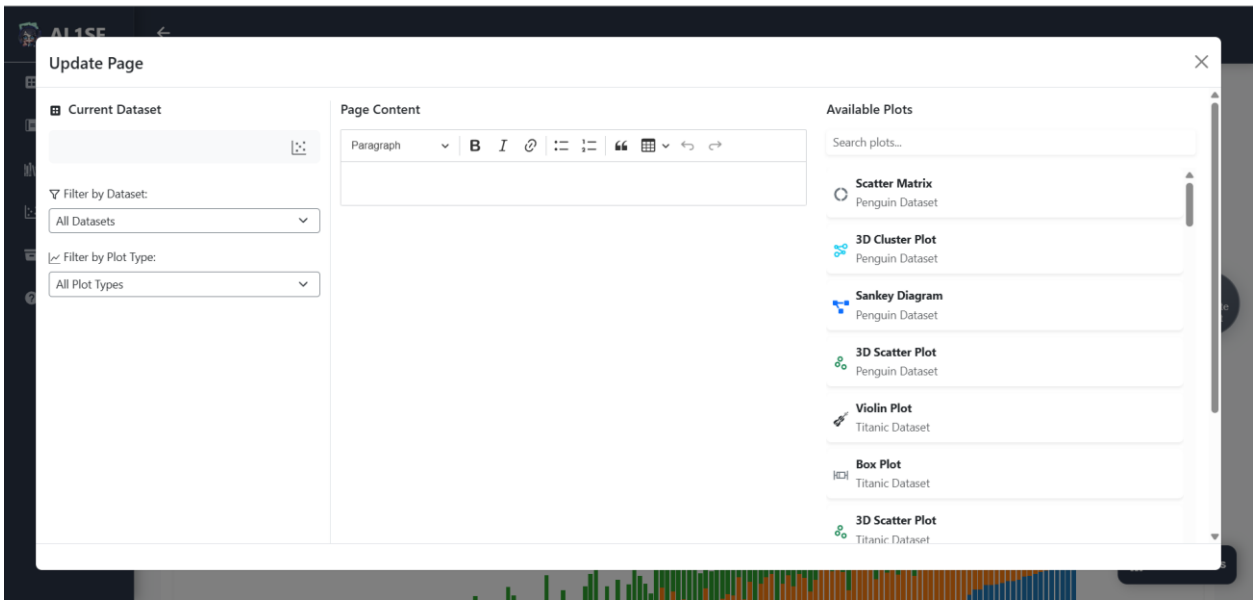
Notebook Page (Vertical Scrolling View)



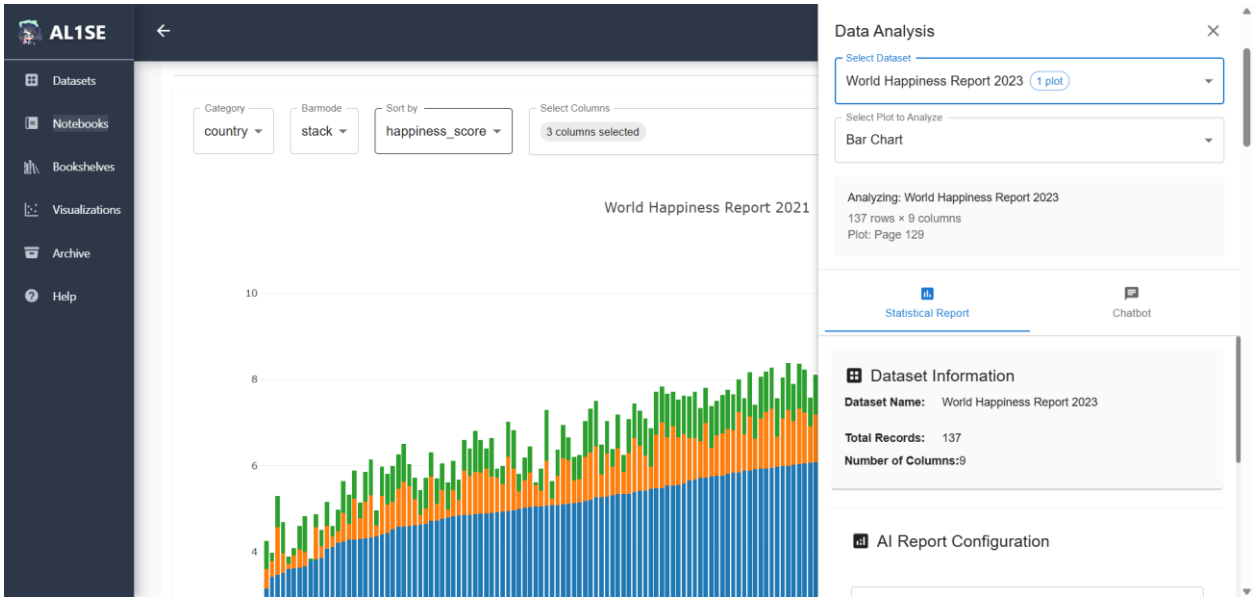
Notebook Page (Horizontal Scrolling View)



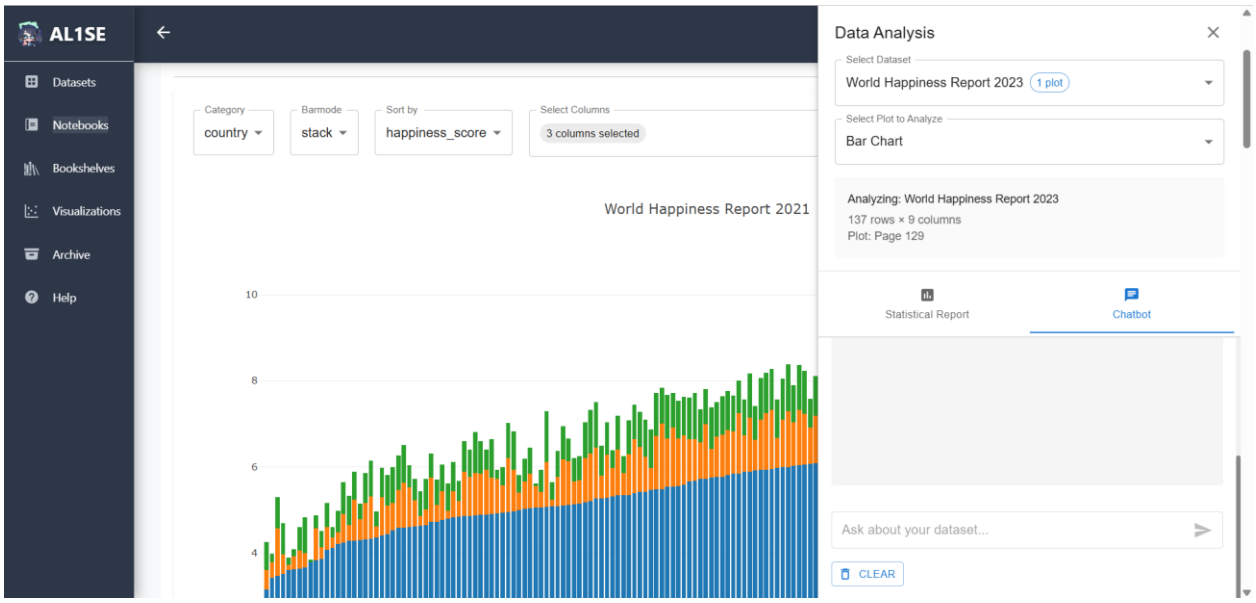
Add Page Modal



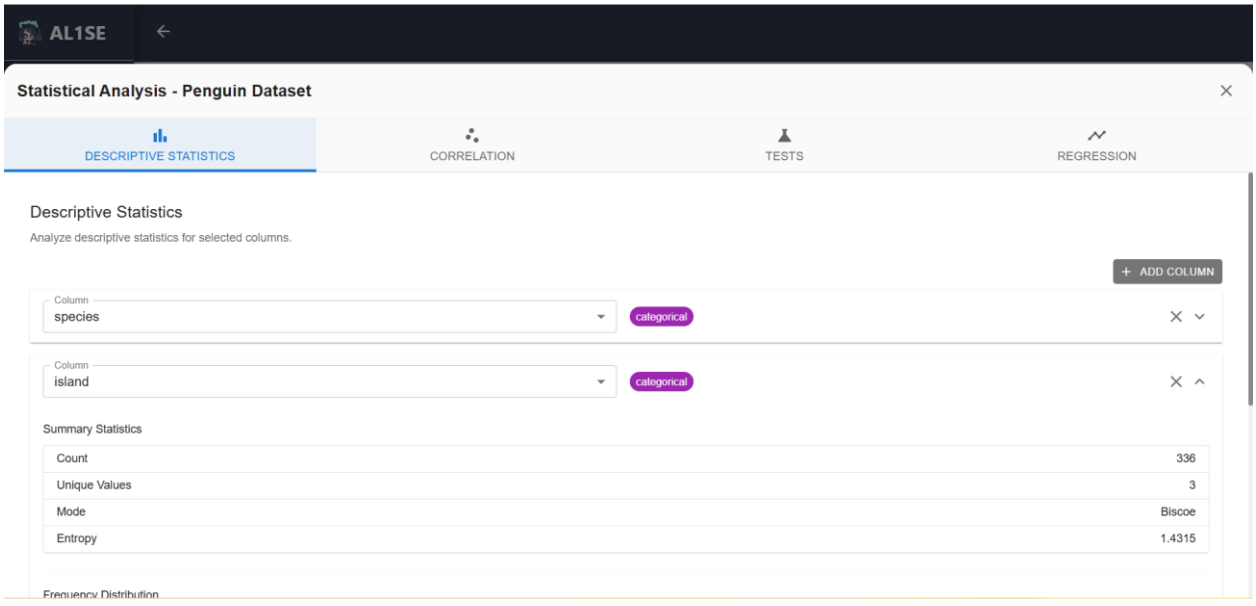
Update Page Modal



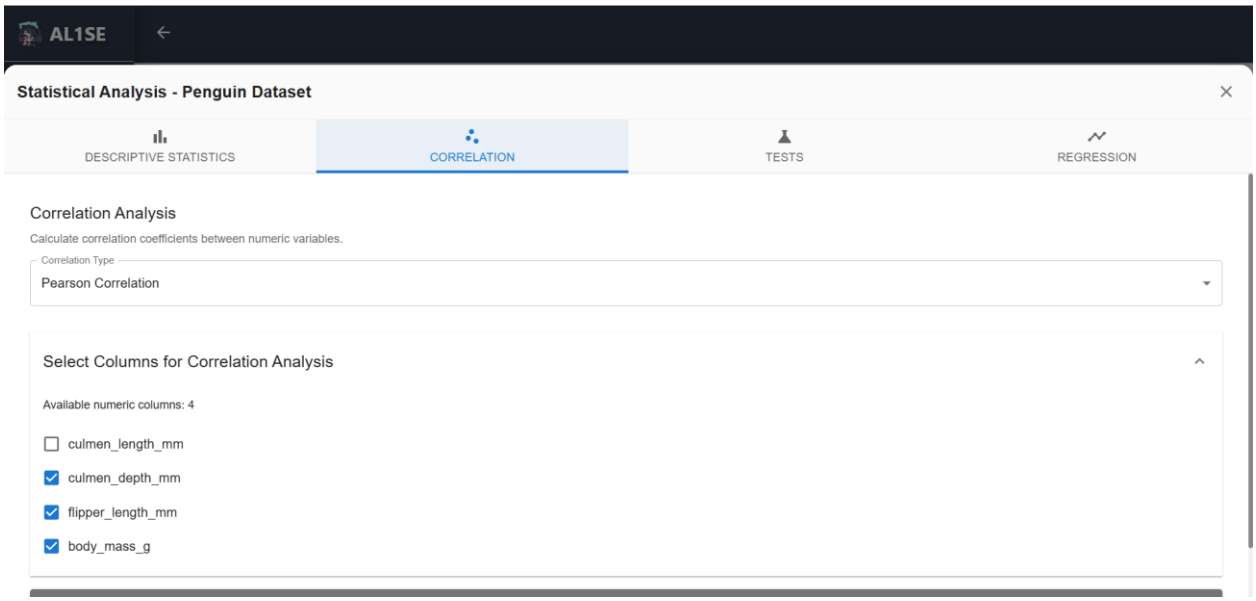
Notebook Page (Generate Report Drawer, Statistical Report Tab)



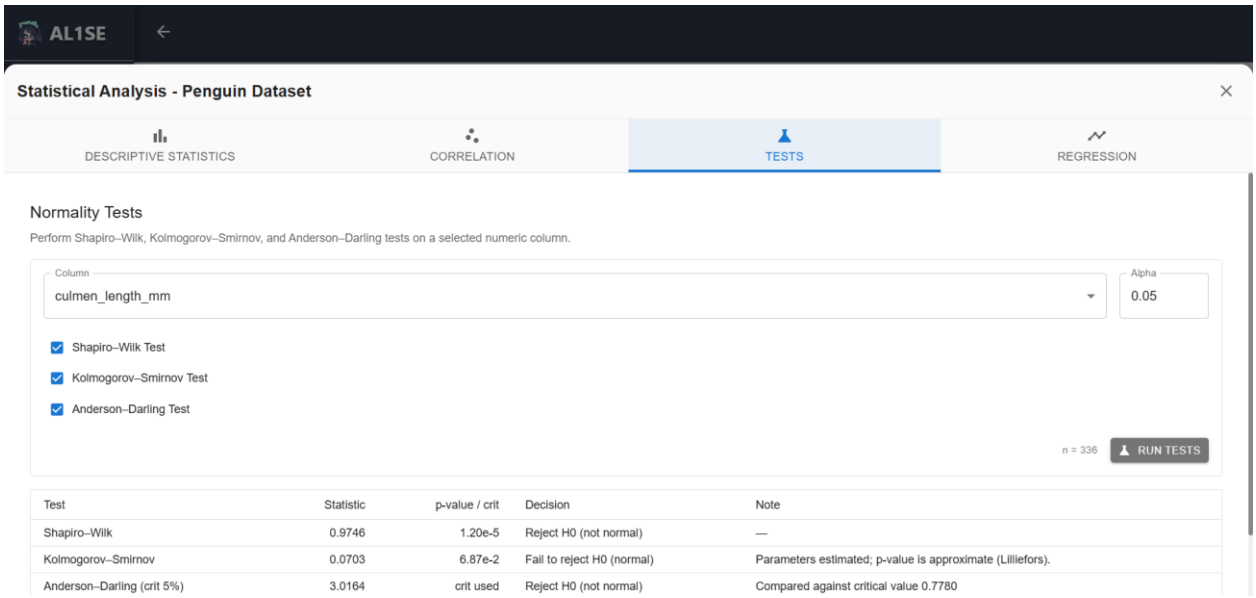
Notebook Page (Generate Report Drawer, Chatbot Tab)



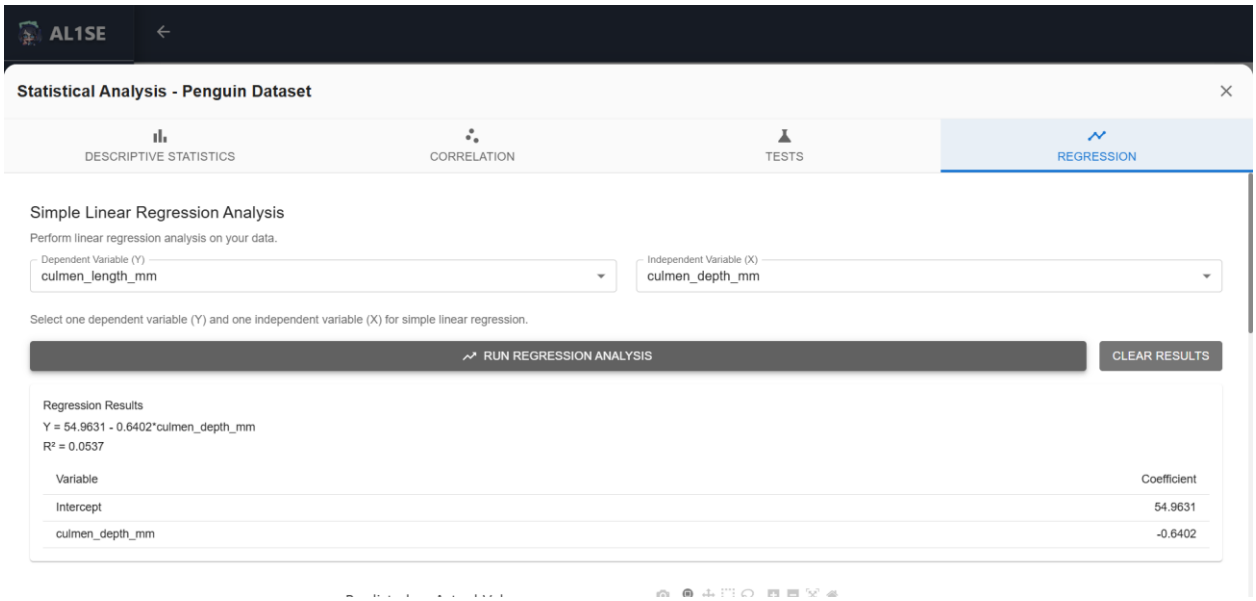
Notebook Page (Statistical Analysis, Descriptive Statistics Tab)



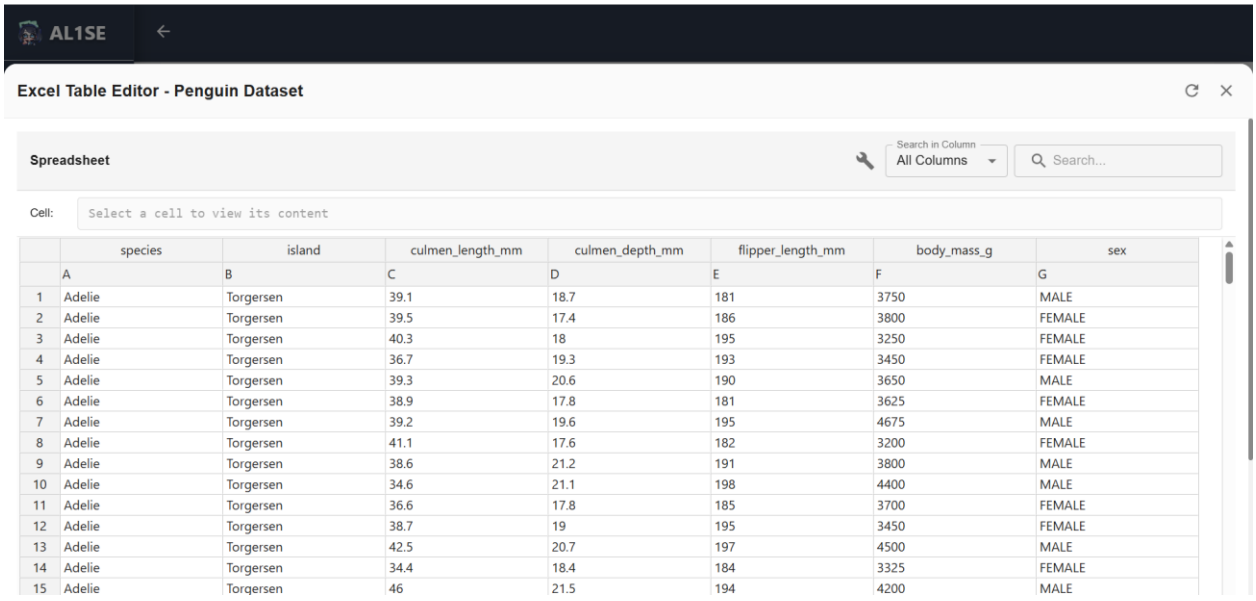
Notebook Page (Statistical Analysis, Correlation Tab)



Notebook Page (Statistical Analysis, Normality Tests Tab)

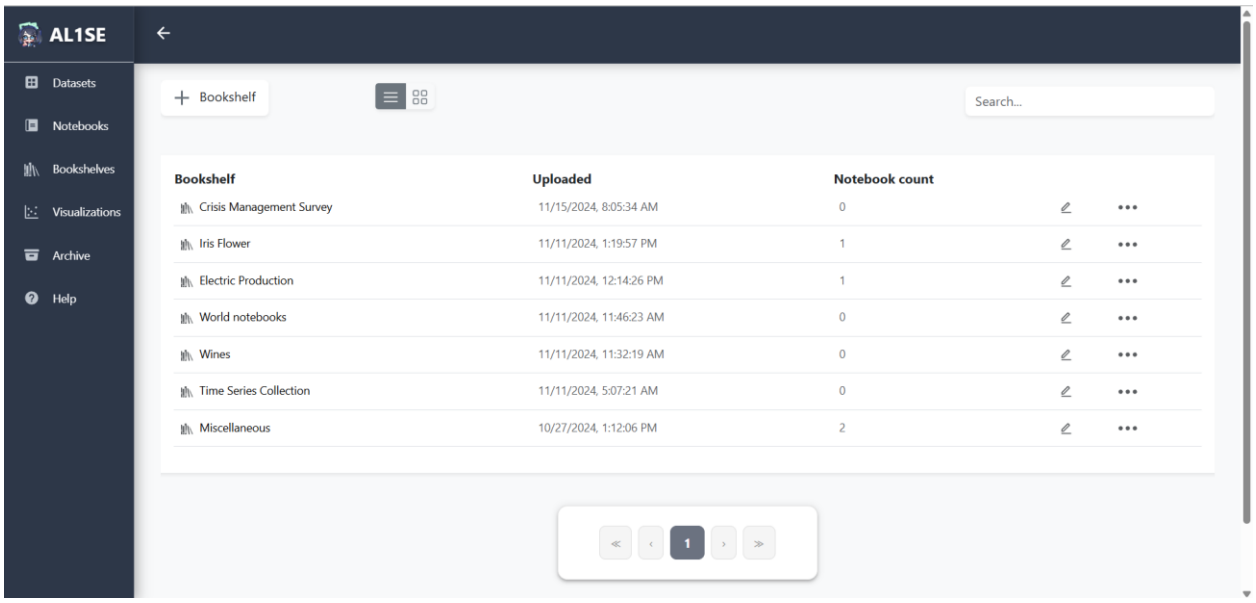


Notebook Page (Statistical Analysis, Regression Tab)

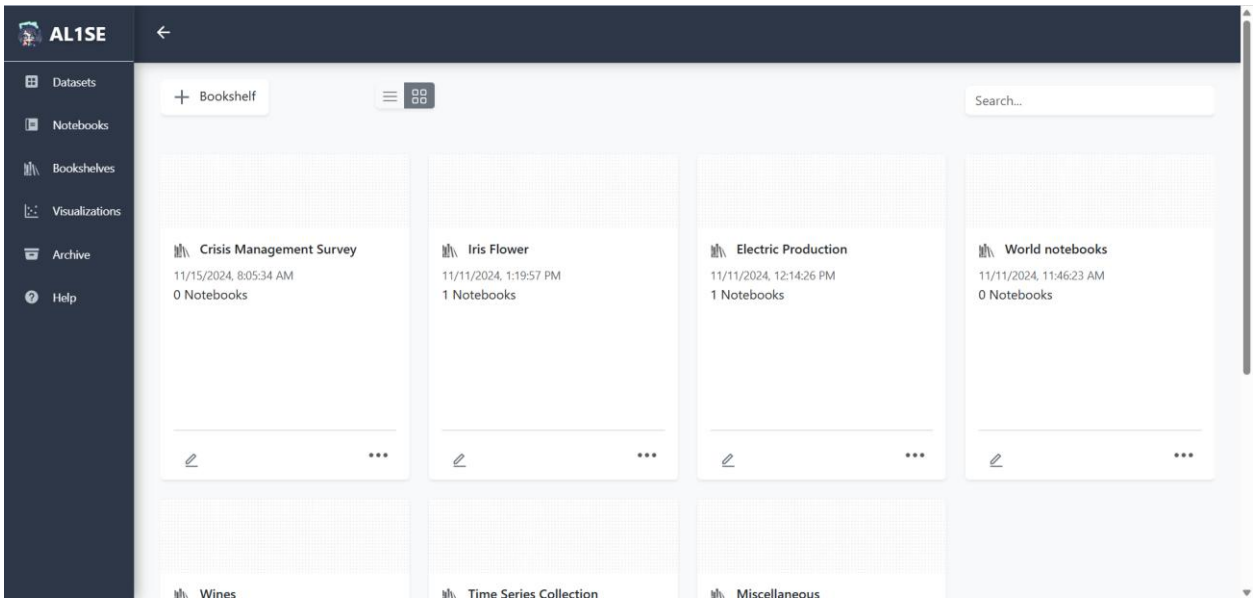


	species	island	culmen_length_mm	culmen_depth_mm	flipper_length_mm	body_mass_g	sex
	A	B	C	D	E	F	G
1	Adelie	Torgersen	39.1	18.7	181	3750	MALE
2	Adelie	Torgersen	39.5	17.4	186	3800	FEMALE
3	Adelie	Torgersen	40.3	18	195	3250	FEMALE
4	Adelie	Torgersen	36.7	19.3	193	3450	FEMALE
5	Adelie	Torgersen	39.3	20.6	190	3650	MALE
6	Adelie	Torgersen	38.9	17.8	181	3625	FEMALE
7	Adelie	Torgersen	39.2	19.6	195	4675	MALE
8	Adelie	Torgersen	41.1	17.6	182	3200	FEMALE
9	Adelie	Torgersen	38.6	21.2	191	3800	MALE
10	Adelie	Torgersen	34.6	21.1	198	4400	MALE
11	Adelie	Torgersen	36.6	17.8	185	3700	FEMALE
12	Adelie	Torgersen	38.7	19	195	3450	FEMALE
13	Adelie	Torgersen	42.5	20.7	197	4500	MALE
14	Adelie	Torgersen	34.4	18.4	184	3325	FEMALE
15	Adelie	Torgersen	46	21.5	194	4200	MALE

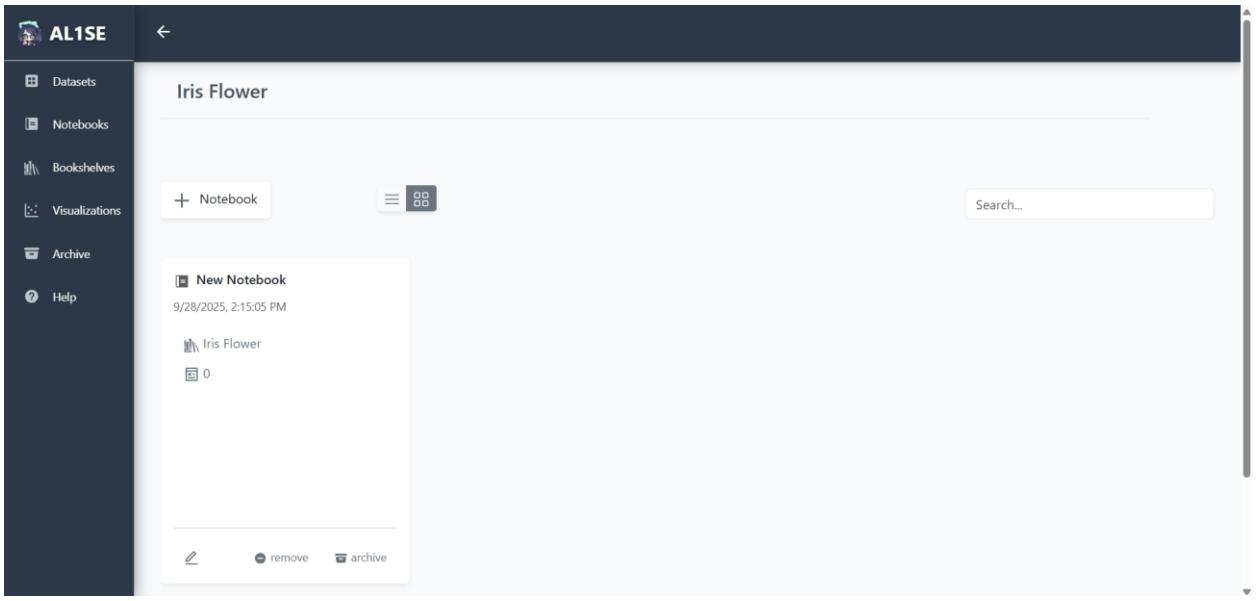
Notebook Page (Excel Table Editor)



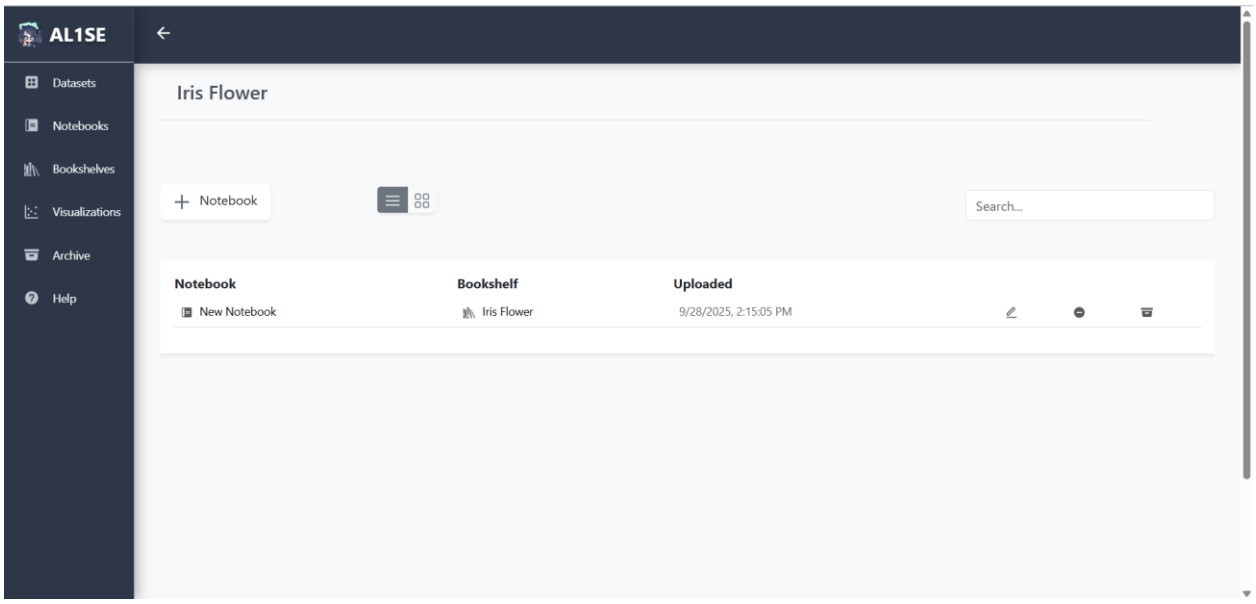
Bookshelves Page (Table View)



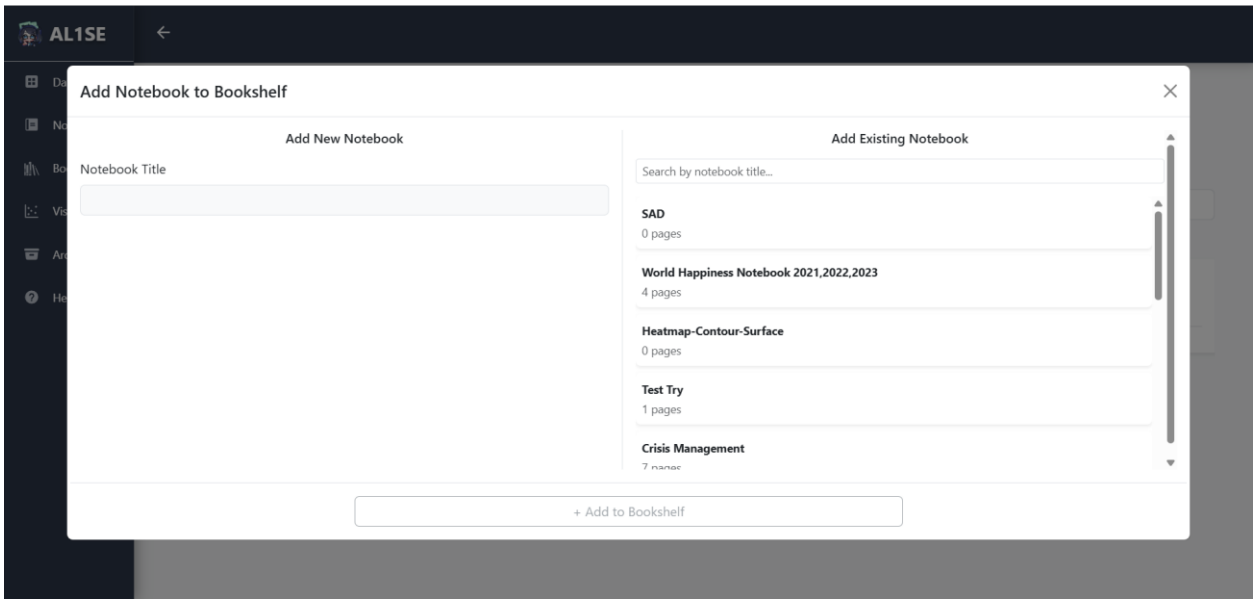
Bookshelves Page (Grid View)



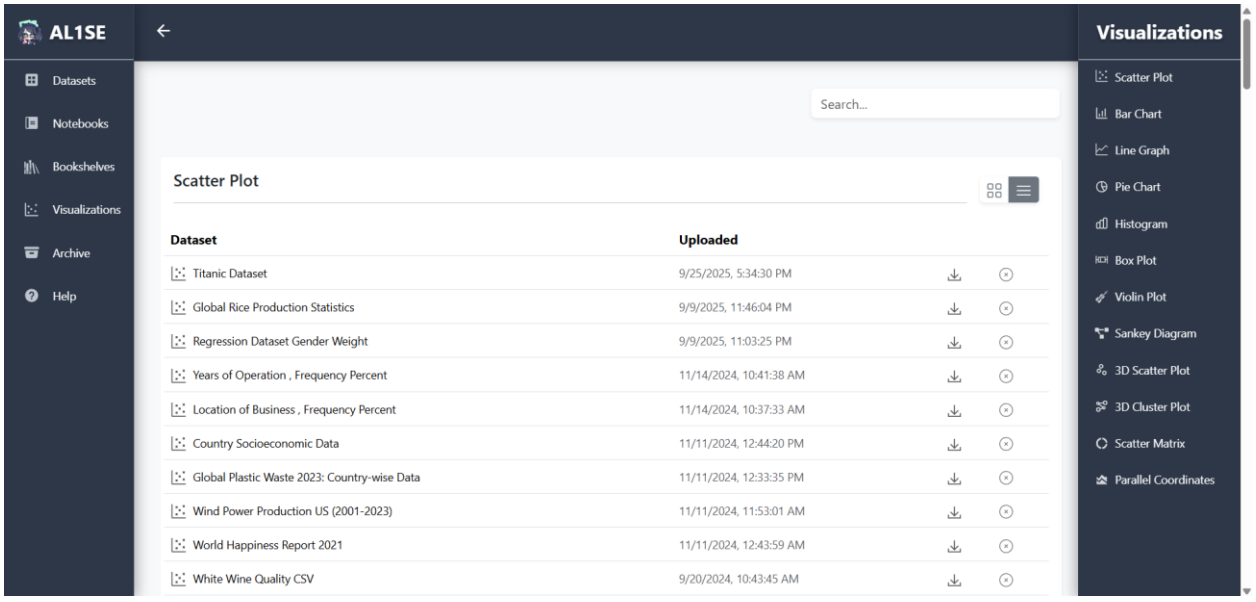
Bookshelf Page (Grid View)



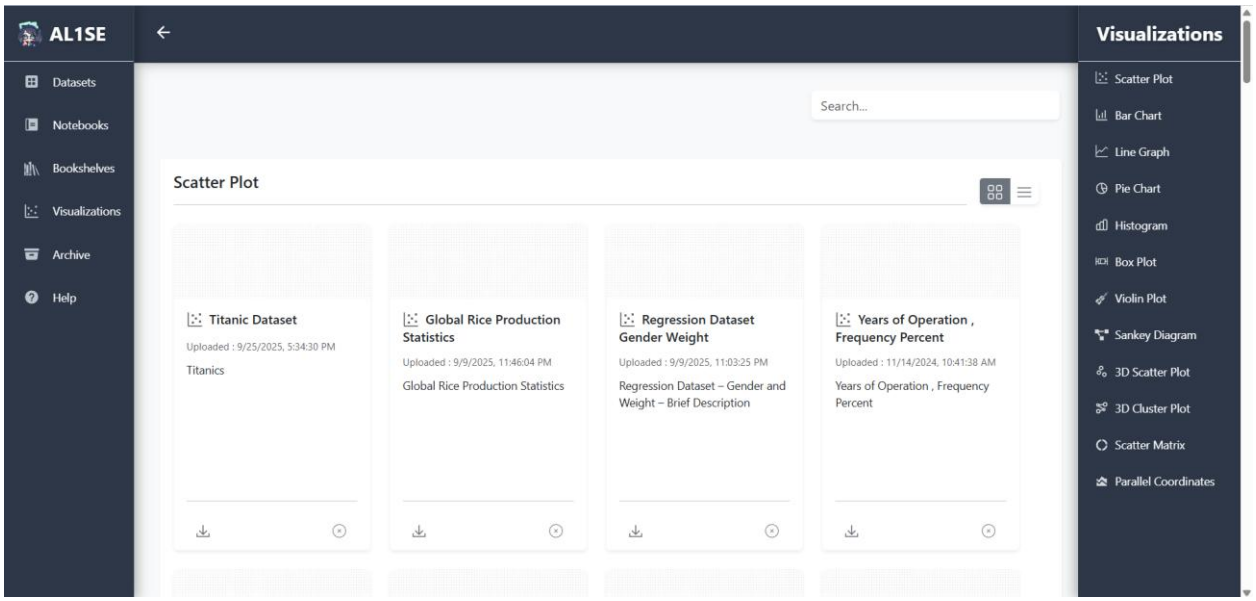
Bookshelf Page (Table View)



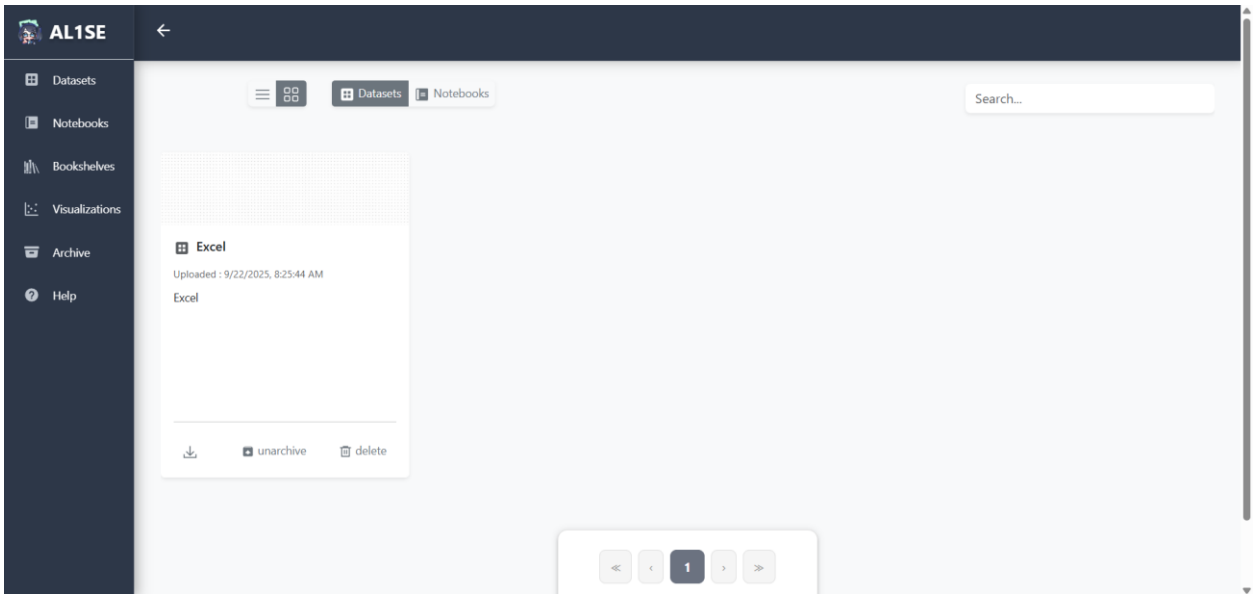
Add Notebook to Bookshelf Modal



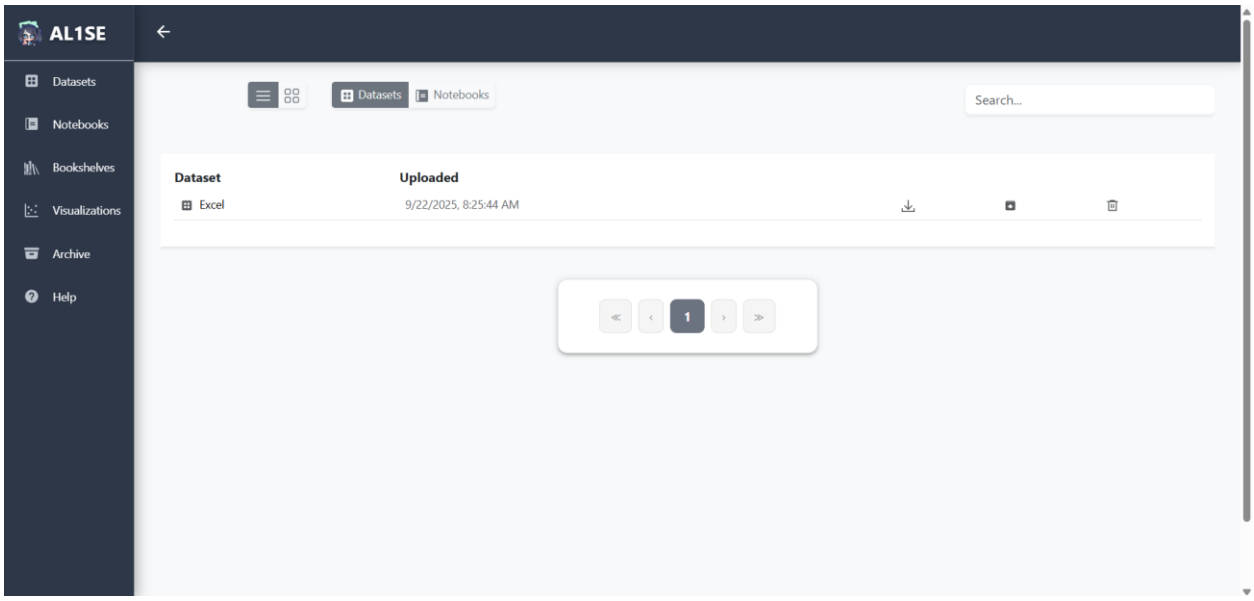
Visualizations Page (Table View)



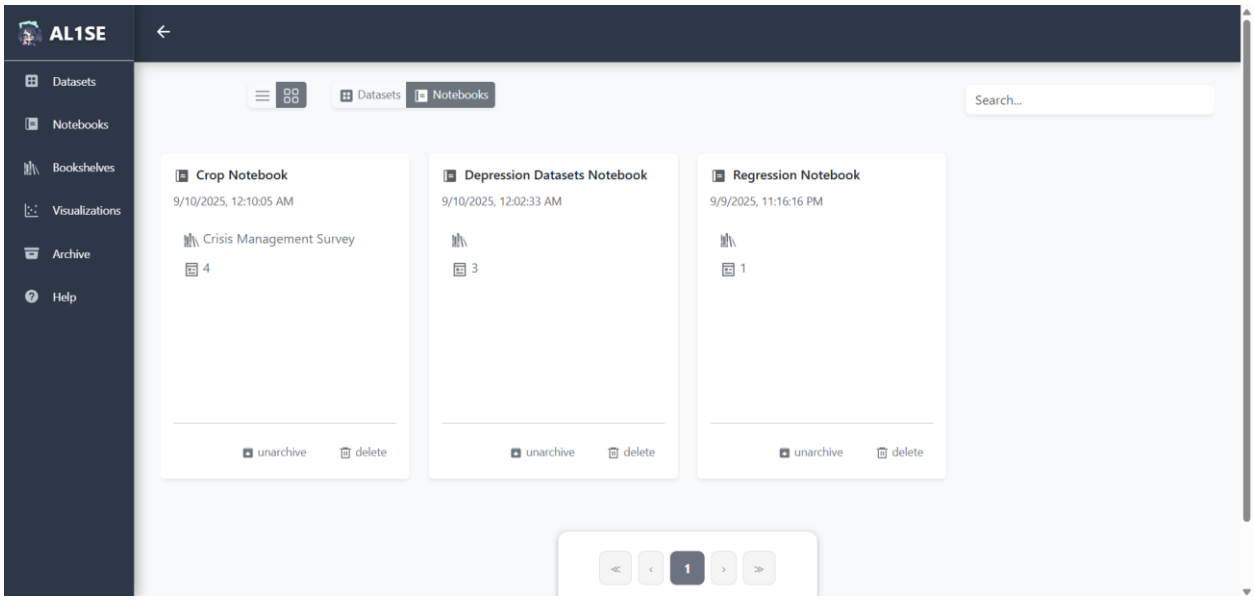
Visualizations Page (Grid View)



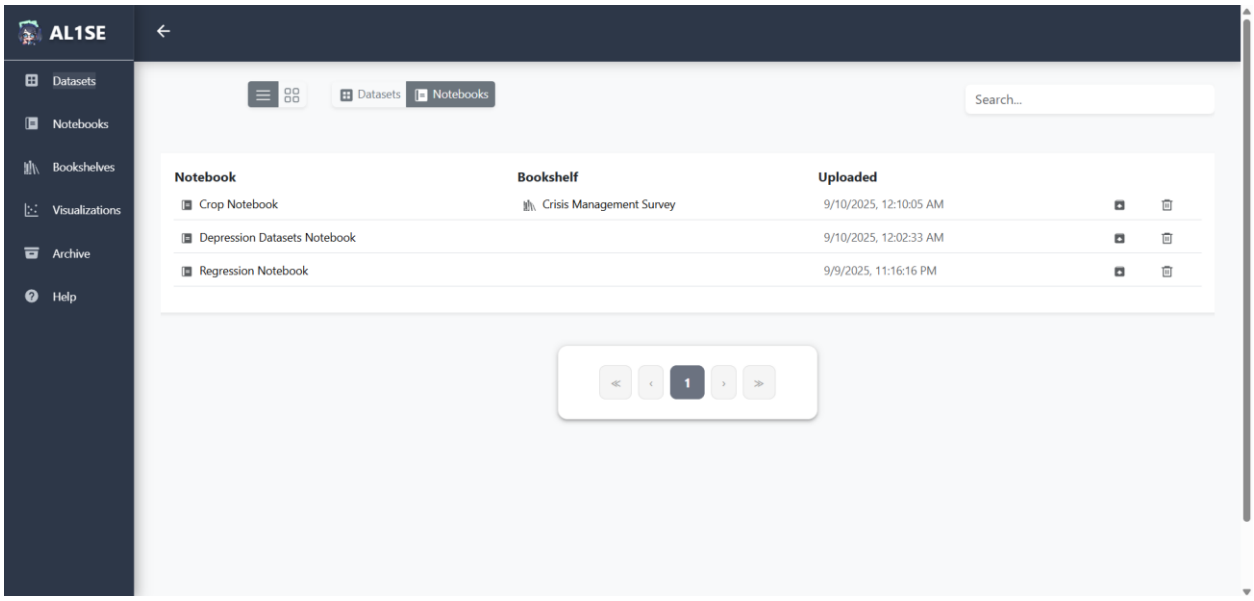
Archive Page (Datasets Tab, Grid View)



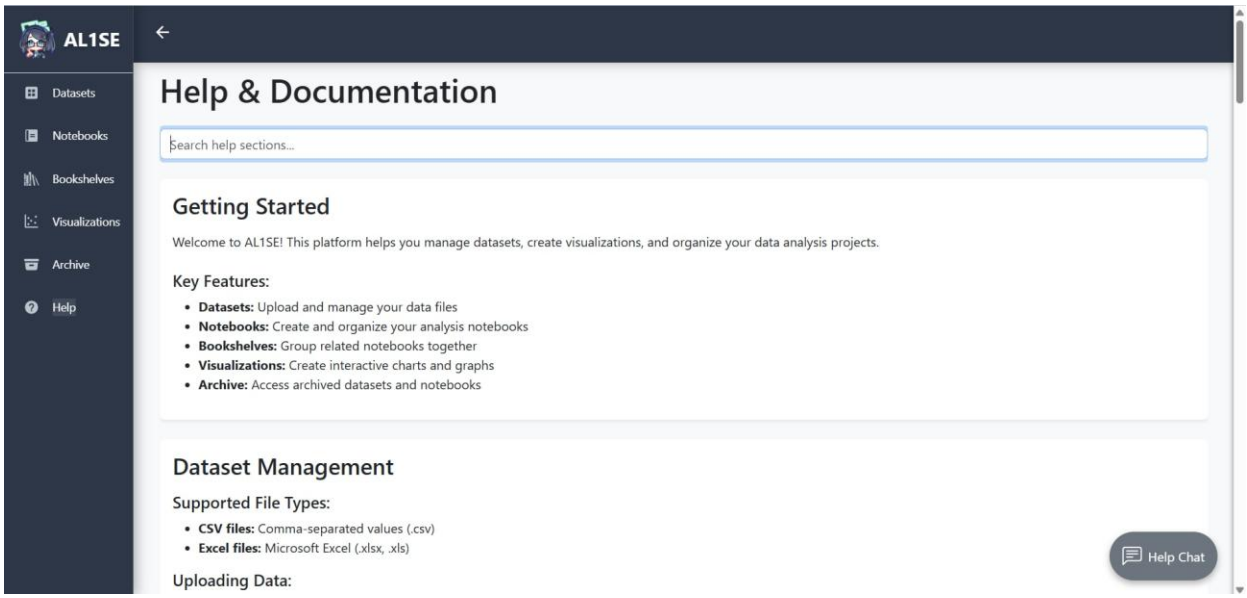
Archive Page (Datasets Tab, Table View)



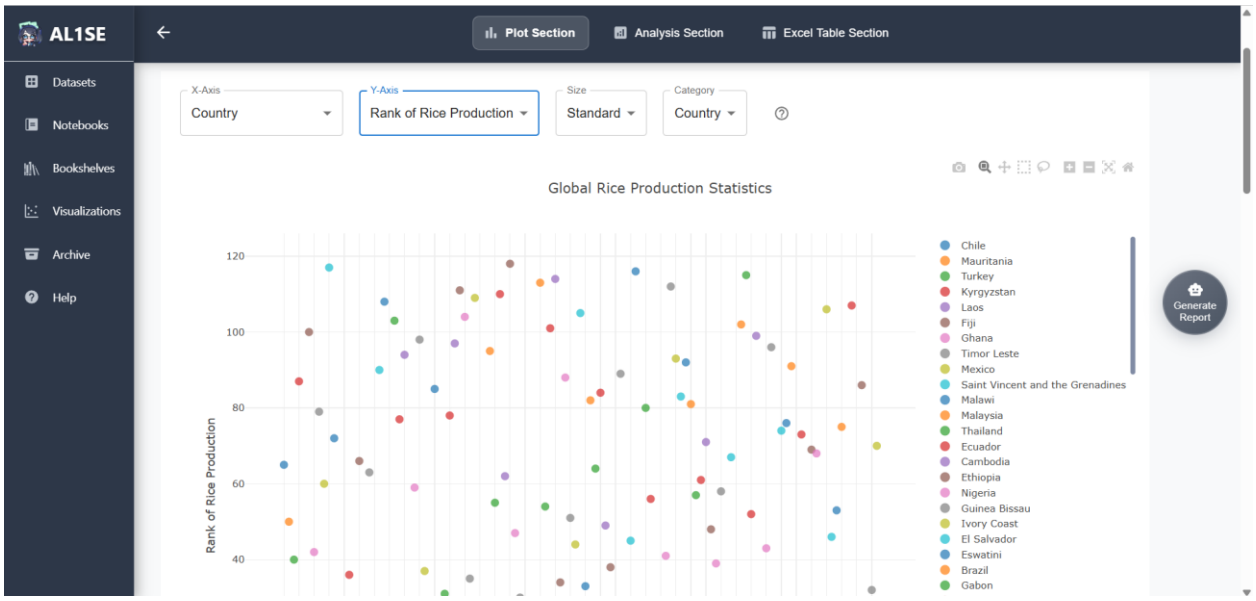
Archive Page (Notebooks Tab, Grid View)



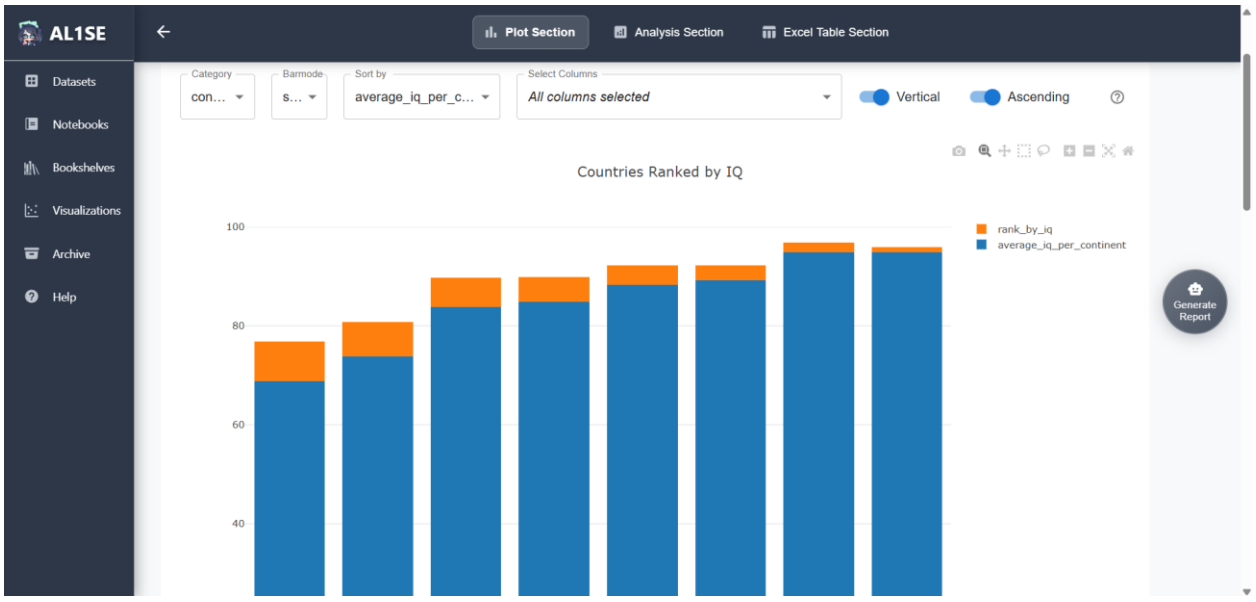
Archive Page (Notebooks Tab, Table View)



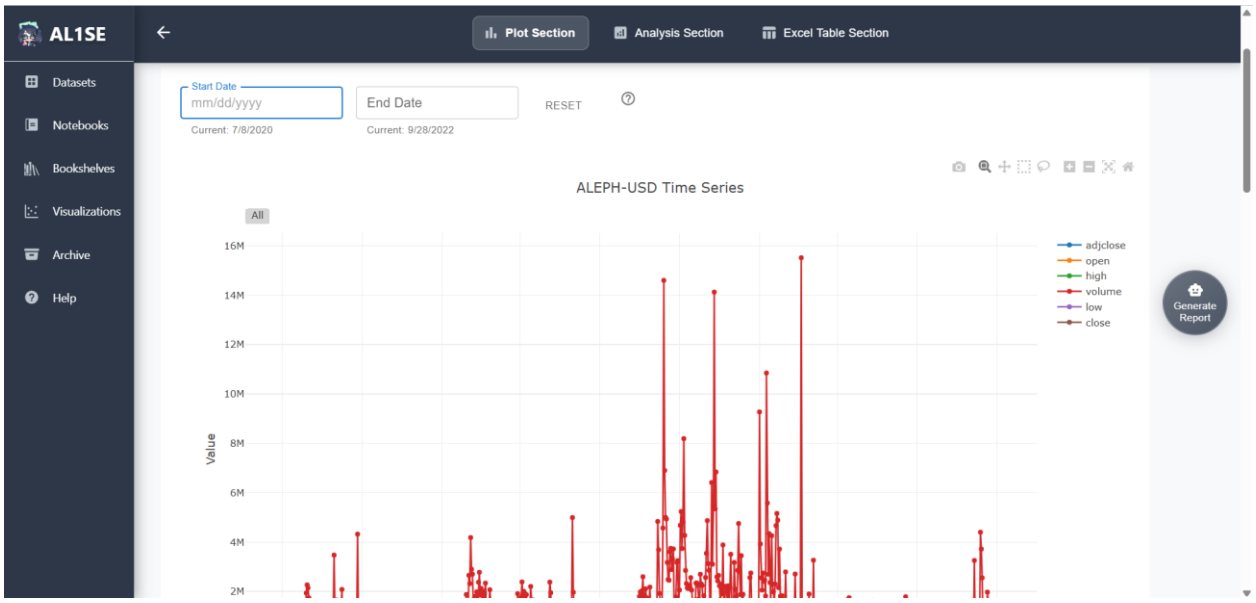
Help Page



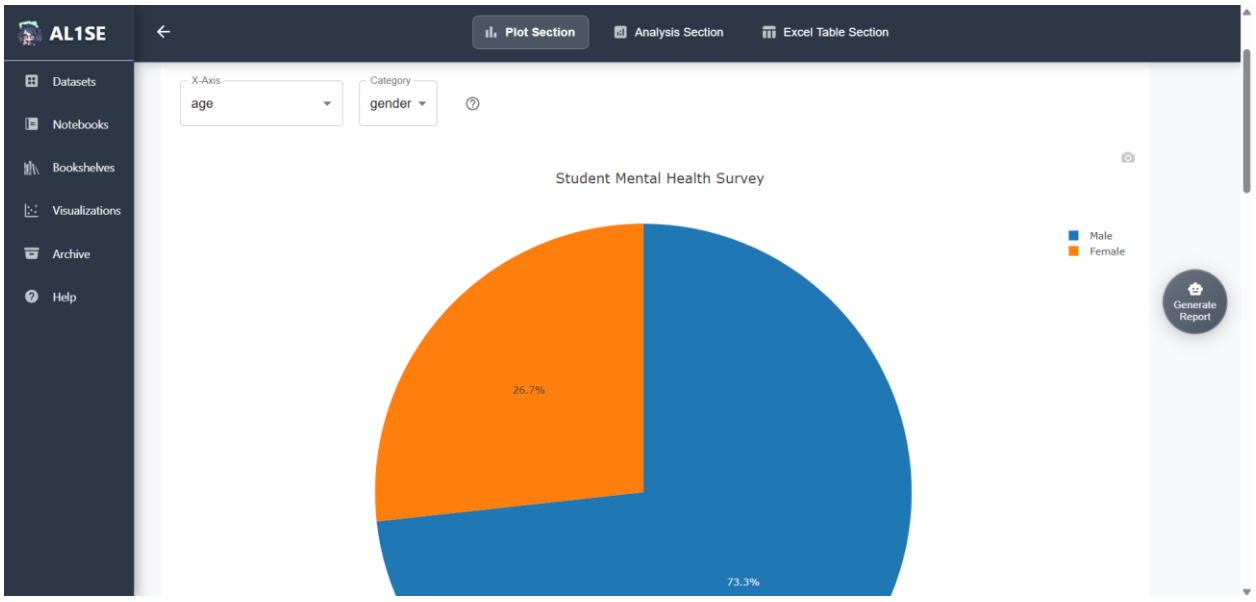
Scatterplot



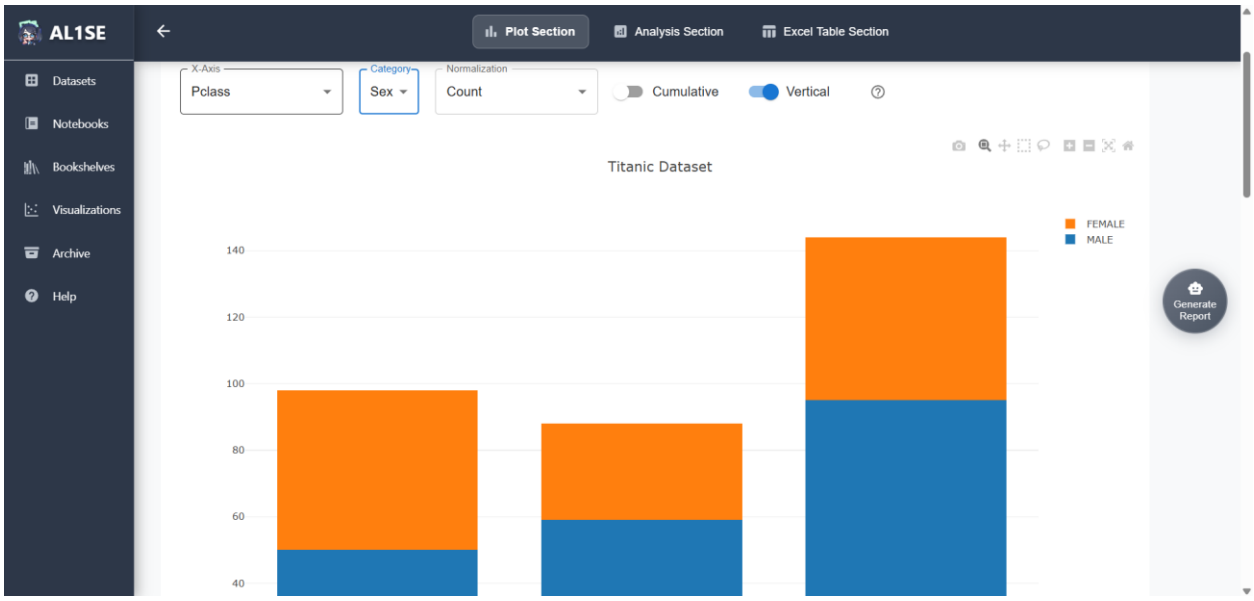
Bar Chart



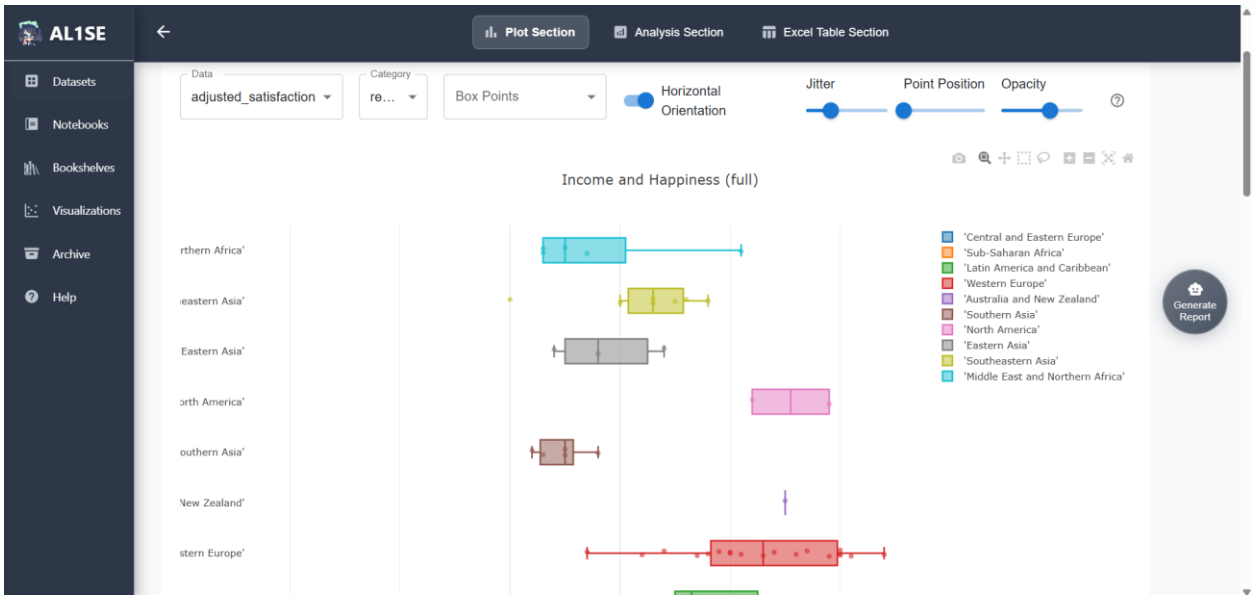
Line Graph



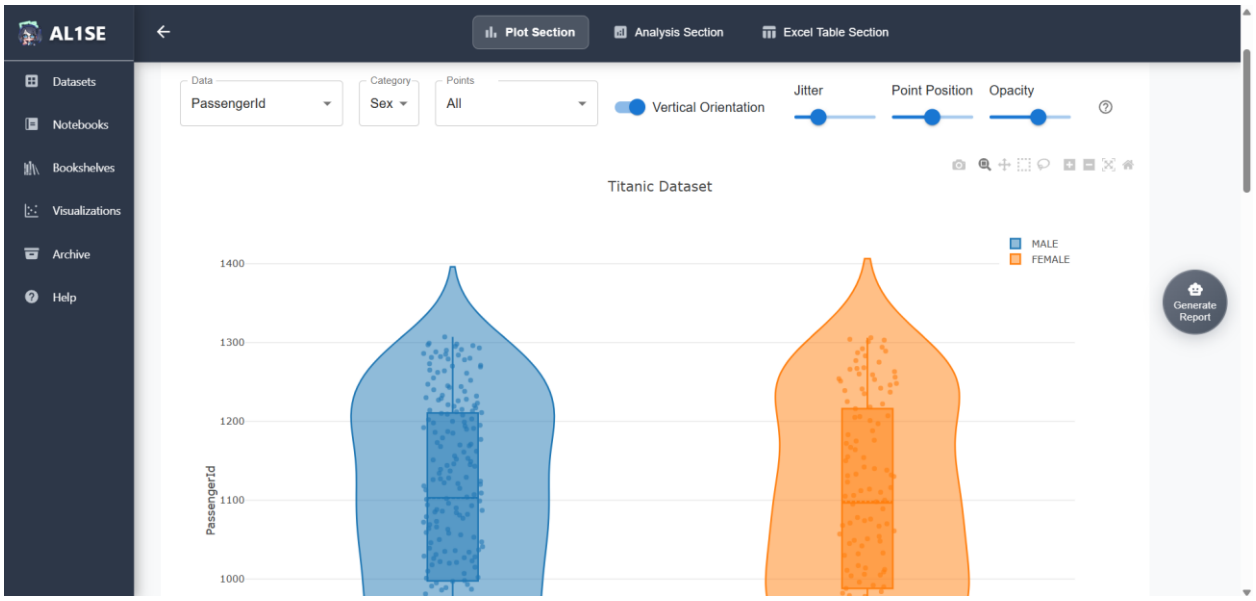
Pie Chart



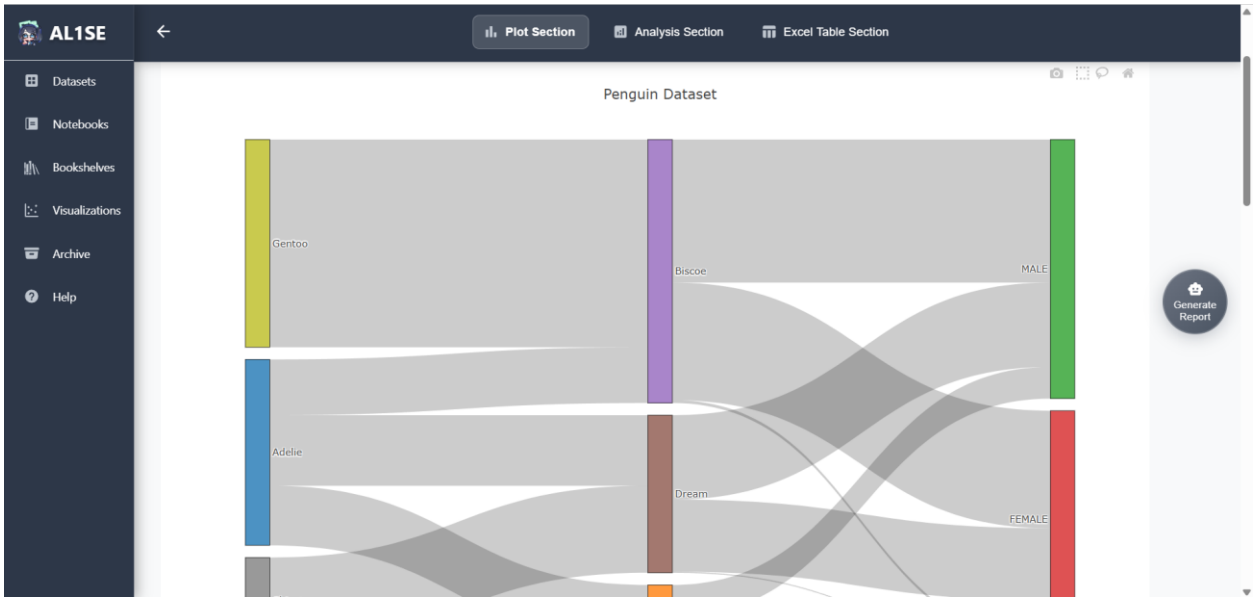
Histogram



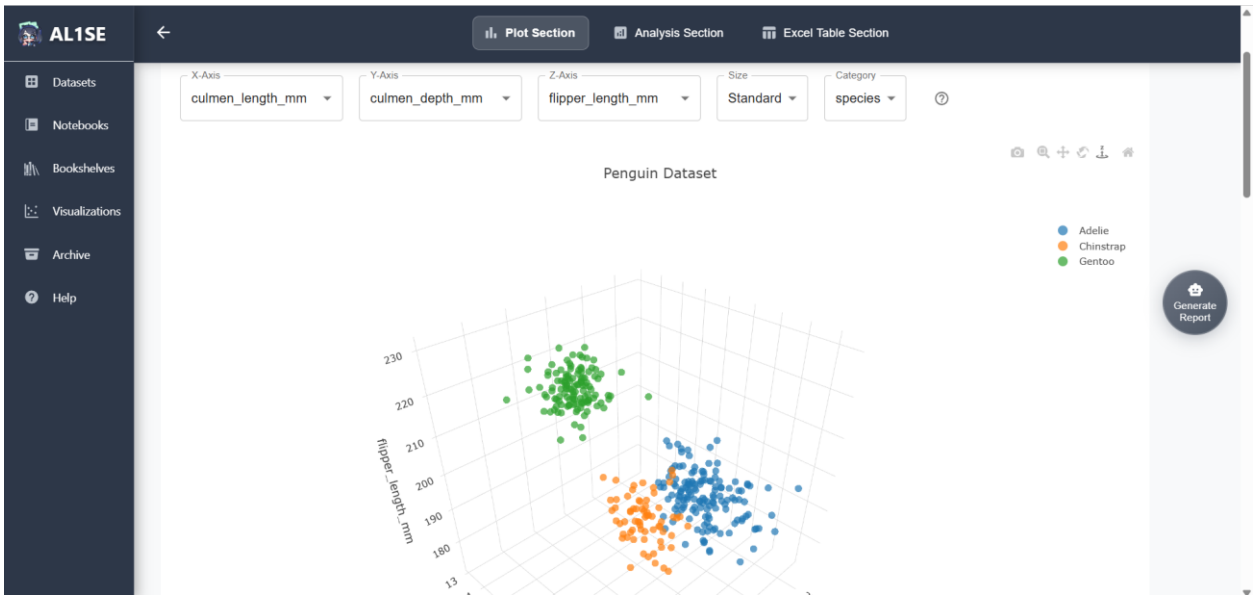
Box Plot



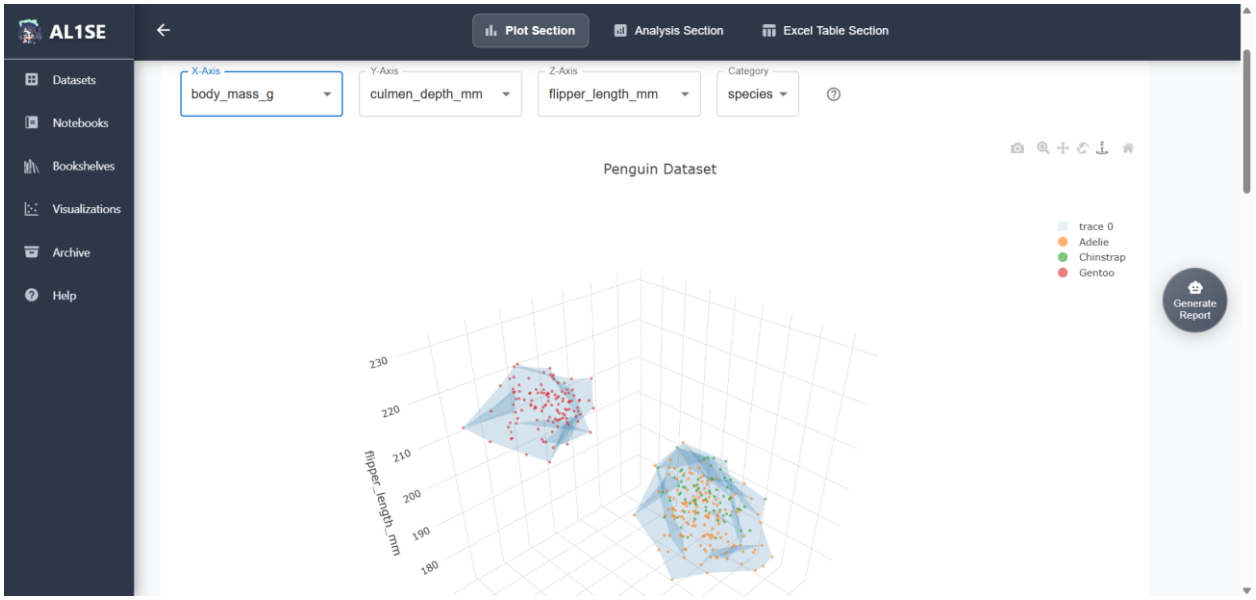
Violin Plot



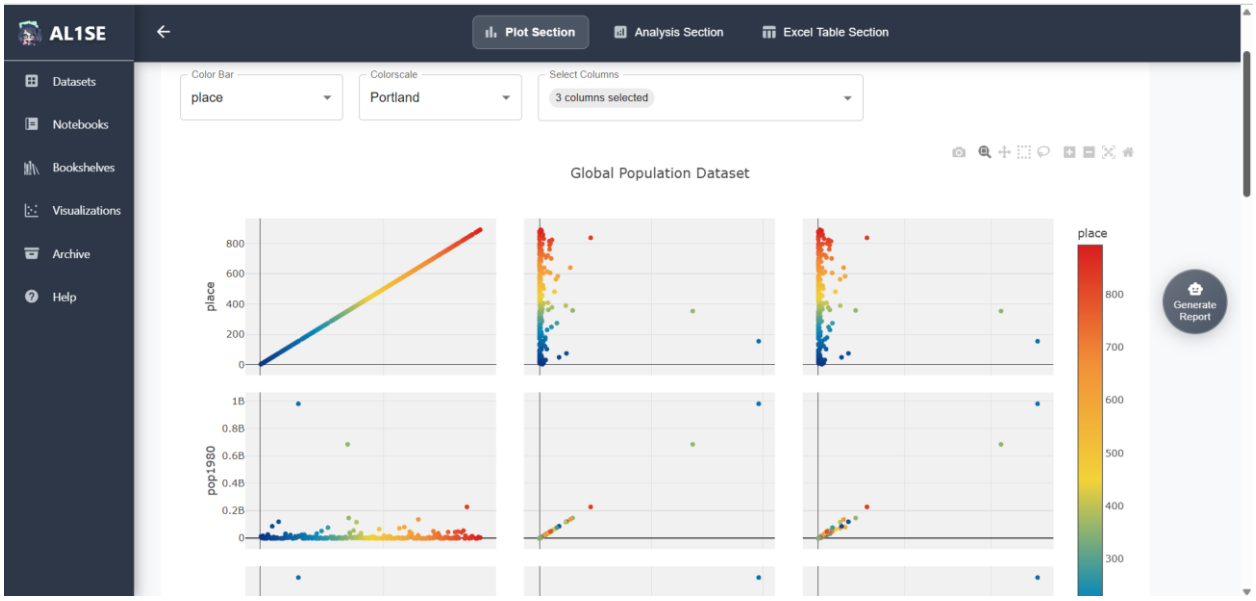
Sankey Diagram



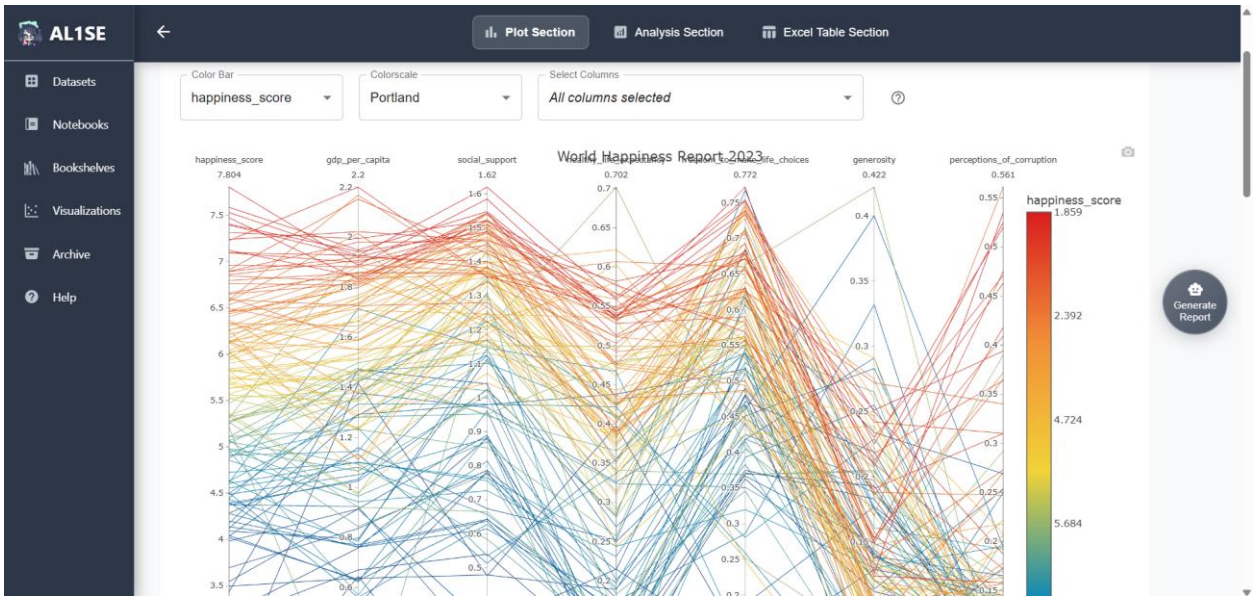
3D Scatterplot



3D Cluster plot



Scatter Matrix



Parallel Coordinates

Appendix E.

Final Developer Testing (Test Cases)

Landing Page/Data Visualization Page:

Test Case ID	Test Scenario	Test Steps	Test Data	Expected Results	Actual Result	Status
TC01	Download Functionality	1. Click the download button of “Penguin CSV” dataset	Uploaded dataset “Penguin CSV”	It will download the dataset to its original format.	It downloaded the dataset to its original format.	Pass
TC02	Archive Functionality – Archiving	1. Click the archive button of “Penguin CSV”.2. Click “Yes, archive it”	Uploaded dataset “Penguin CSV”	A prompt will say the dataset is archived and it will archive the dataset.	A prompt popped up saying that the dataset is archived, and the dataset is shown in the Archive Page.	Pass
TC03	Uploading Dataset with CSV file	1. Click the Add Dataset.2. Add a Title “Penguin CSV”.3. Add description “Penguins”.4. Upload “penguins.csv”.5. Click Upload.	Title: “Penguin CSV”, Description: “Penguins”, File: “penguins.csv”	It should prompt saying the dataset is successfully uploaded.	It prompted saying the dataset is successfully uploaded.	Pass
TC04	Uploading Dataset with Excel file	1. Click the Add Dataset.2. Add a Title “Penguin CSV”.3. Add description “Penguins”.4. Upload	Title: “Penguin CSV”, Description: “Penguins”, File: “penguins.xlsx”	It should prompt saying the dataset is successfully uploaded.	It prompted saying the dataset is successfully uploaded.	Pass

		“penguins.xlsx”.5. Click Upload.				
TC05	Selecting a Dataset	1. Click “Penguin CSV” anywhere inside the card.	Uploaded “Penguin CSV” dataset	It should redirect to the “Penguin CSV” dataset Dataset Page.	It redirected to the “Penguin CSV” dataset Dataset Page.	Pass
TC6	Dataset Page Edit Button – Editing Title, Description & File	1.Click the Edit Button. 2.Type a new Title “Iris CSV Dataset” 3.Edit the description “Iris flowers are Amazing” 4.Upload “Iris-Csv” 4.Click update button 5.Press Ok	Uploaded “Penguin CSV” dataset Insert a Title “Iris CSV Dataset” & Insert a Description “Iris flowers are Amazing” and upload the “Iris CSV” dataset	It should prompt saying “Success Dataset Detail updated successfully” the name will change, the description will change and the file will change.	It prompted saying “Success Dataset Detail updated successfully” the name changed, the description changed and the file changed.	Pass
TC7	Dataset Page Download Button	1.Click the Download button.	Uploaded “Penguin CSV” dataset	It will automatically download the “Penguin CSV” dataset.	It automatically downloaded the “Penguin CSV” dataset.	Pass
TC8	Dataset Page Add Visualization	1.Click the add Visualization button. 2.Click 3D Scatter 3.Click Create Visualization	Uploaded “Penguin CSV” dataset	It will create a selection of visualization type that has been applied to	It created a selection of visualization type that has been applied to	Pass

				that dataset.	that dataset.	
TC9	Dataset Page Search Visualizations	1.Click the search visualization search bar. 2.Typed "Scatterplot"	Uploaded "Penguin CSV" dataset	It should filter visualization type namely "scatterplot"	It filtered visualization type namely "scatterplot"	Pass
TC10	Dataset Page Remove Visualization type	1.Click the "X" button of the right side of a entry which is 3D Scatter plot. 2.Click Yes in removing the visualization type. 3.Click Ok	Uploaded "Penguin CSV" dataset	It should remove the visualization type to the list of visualization types of the dataset.	It should removed the visualization type to the list of visualization types of the dataset.	Pass
TC11	Section Change Analysis Section	1.Press Analysis Section button at the header	Uploaded "Penguin CSV" dataset in 3D scatterplot visualization	It will scroll down to the Analysis Section.	It scrolled down to the Analysis Section.	Pass
TC12	Section Change Excel Table Section	1.Press Excel Table Section button at the header	Uploaded "Penguin CSV" dataset in 3D scatterplot visualization	It will scroll down to the Excel Table Section.	It scrolled down to the Excel Table Section.	Pass
TC13	Section Change Plot Section	1.Press Excel Table Section button at the header 2.Press Plot Section	Uploaded "Penguin CSV" dataset in 3D scatterplot visualization	It will scroll up to the Plot Section.	It scrolled up to the Plot Section.	Pass
TC14	Scatterplot Change X & Y Values	1.Press the dropdown for X-axis and select "culmen_length_mm"	Uploaded "Penguin CSV" dataset in	It will display the X & Y values in a scatterplot	It displayed the X & Y values in a scatterplot	Pass

		2.Press the dropdown for Y-axis and select "culmen_depth_mm"	Scatterplot visualization	visualization.	visualization.	
TC15	Barchart change category, barmode, and sort by values.	1.Select "species" in category. 2.Select "stack" in barmode. 3.Select "body_mass_g" in sort by	Uploaded "Penguin CSV" dataset in Barchart visualization	It will display the category, and sort by values in a barchart visualization.	It displayed the category, and sort by values in a barchart visualization.	Pass
TC16	Line Graph change Start Date and End Date	1.Click on the start date and set 9/21/2018 2.Click on the end date and set 9/28/2022	Uploaded "AOG-USD Time Series" dataset in Line Graph Visualization	It will display the values inputted in the start and end dates in a line graph visualization.	It displayed the values inputted in the start and end dates in a line graph visualization.	Pass
TC17	Pie Chart change X-axis and category	1.Select "culmen_length_mm" in X-axis. 2.Select "sex" in category	Uploaded "Penguin CSV" dataset in Pie Chart visualization	It will display the X-axis, and category values in a Pie Chart visualization.	It displayed the X-axis, and category values in a Pie Chart visualization.	Pass
TC18	Histogram change, Category, Normalization	1.Select X-Axis value to "culmen_length_mm". 2.Select category value to "species". 3.Select Normalization to Count	Uploaded "Penguin CSV" dataset in Histogram visualization	It will display the X-axis, and category values in a Histogram visualization	It displayed the X-axis, and category values in a Histogram visualization	Pass

TC19	Boxplot change Data, Category and Box points	1.Select Data to “culmen_length_mm”. 2.Select category to “species”. 3.Select box points to “All”.	Uploaded “Penguin CSV” dataset in Boxplot visualization	It will display the Data, and category values in a Boxplot visualization	It displayed the Data, and category values in a Boxplot visualization	Pass
TC20	Violinplot change Data, Category and Box points	1.Select Data to “culmen_length_mm”. 2.Select category to “species”. 3.Select box points to “All”.	Uploaded “Penguin CSV” dataset in Violinplot visualization	It will display the Data, and category values in a Violinplot visualization	It displayed the Data, and category values in a Violinplot visualization	Pass
TC21	Sankey Diagram Data Manipulation	1.Drag the horizontal bars and change the postion	Uploaded “cars sankey” dataset in Sankey Diagram visualization	It will change the position of the horizontal bars.	It changes the position of the horizontal bars.	Pass
TC22	3D Scatterplot Change X, Y, & Z Values	1.Press the dropdown for X-axis and select “culmen_length_mm” 2.Press the dropdown for Y-axis and select “culmen_depth_mm” 3.Press the dropdown for Z-Axis and select “flipper_length_mm”	Uploaded “Penguin CSV” dataset in 3D Scatter plot visualization	It will display the X , Y & Z values in a 3D Scatterplot visualization.	It displayed the X & Y values in a 3D Scatterplot visualization.	Pass
TC23	Parallel Coordinates change color bar & colorscale	1.Click the dropdown of change color to “culmen_length_mm”.	Uploaded “Penguin CSV” dataset in Parallel	It will display the change color & colorscale	It displayed the change color & colorscale	Pass

		2.Click the dropdown of colorscale to “Viridis”.	Coordinates visualization	in the Parallel Coordinates visualization.	in Parallel Coordinates visualization.	
TC24	Parallel Coordinates add data filters	1.Press “culmen_depth_mm” line. 2.Drag it to 58-60	Uploaded “Penguin CSV” dataset in Parallel Coordinates visualization	It will filter the lines that will pass through.	It filtered the lines that will pass through.	Pass
TC25	Parallel Coordinates remove data filters	1.Press “culmen_depth_mm” line. 2.Double click the violet line.	Uploaded “Penguin CSV” dataset in Parallel Coordinates visualization	It will remove the filters.	It removed the filters.	Pass
TC26	3D Clusterplot Change X, Y, & Z Values	1.Press the dropdown for X-axis and select “culmen_length_mm” 2.Press the dropdown for Y-axis and select “culmen_depth_mm” 3.Press the dropdown for Z-Axis and select “flipper_length_mm”	Uploaded “Penguin CSV” dataset in 3D Scatter plot visualization	It will display the X , Y & Z values in a 3D Clusterplot visualization.	It displayed the X & Y values in a 3D Clusterplot visualization.	Pass
TC27	Scatter matrix change color bar & colorscale	1.Click the dropdown of change color to “culmen_length_mm”. 2.Click the dropdown of colorscale to “Viridis”.	Uploaded “Penguin CSV” dataset in Scatter Matrix visualization	It will display the change color & colorscale in Scatter Matrix visualization.	It displayed the change color & colorscale in Scatter Matrix visualization.	Pass

TC28	Edit Functionality	1.Click the Edit button of “Crop Notebook” Notebook. 2.Change the name to “Crop Notebooks” 3.Change the Bookshelf to “Crop Booshelf”	Created Notebook “Crop Notebook”	It will update the notebook.	It updated the notebook.	Pass
TC29	Archive Functionality Archiving	1.Click the archive button of “Crop Notebook”. 2.Click “Yes, archive it”	Created Notebook “Crop Notebook”	A prompt will say the Notebook is archive and it will archive the Notebook.	A prompt pop up saying that the Notebook is archived and the Notebook is archived shown in the Archive Page	Pass
TC30	Creating Notebook with all forms answered	1.Click the Add Notebook . 2.Add a Title “Crop Notebook” 3.Select a bookshelf “Crop Bookshelf” 4. Click Create Notebook	Type “Crop Notebook” in Title, “Crop Bookshelf” in Bookshelf	It should prompt saying the Notebook is successfully created.	It prompted saying the Notebook is successfully created.	Pass
TC31	Selecting a Notebook	1.Click “Crop Notebook ”. anywhere inside the card.	Created Notebook “Crop Notebook”	It should redirect to the “Crop Notebook ” Notebook Page	It redirected to the “Crop Notebook ” Notebook Page	Pass

Notebook Page:

Test Case ID	Test Scenario	Test Steps	Test Data	Expected Results	Actual Result	Status
TC32	Notebook Add Page	1. Click the Add Page Button.2. Select a new page to add "Penguin Dataset" in Scatterplot visualization.3. Click Create button.	Uploaded "Penguin CSV" dataset	It should prompt saying "Success Pages created successfully" and add the page to the notebook.	It prompted saying "Success Pages created successfully" and added the page to the notebook.	Pass
TC33	Notebook Add Page – Filter dataset & plot type	1. Click the Add Page Button.2. Select the dropdown of Filter by Dataset: "Penguin Dataset".3. Filter by Plot: Scatterplot.	Uploaded "Penguin CSV" dataset	It should show the filtered plots.	It showed the filtered plots.	Pass
TC34	Notebook Add Page – Creating Contents	1. Click the Add Page Button.2. Select a new page to add "Penguin Dataset" in Scatterplot visualization.3. Type content "Penguins".4. Click Create.	Uploaded "Penguin CSV" dataset	It should show the content on the notebook.	It showed the content on the notebook.	Pass
TC35	Notebook Page Edit Notebook	1. Click the Edit Page Button.2. Select a dataset to change the page into	Uploaded "Crop Analysis" dataset	It should update the page.	It updated the page.	Pass

		“Crop Analysis”.				
TC36	Notebook Page Delete Notebook	1. Click the Delete Page Button.2. Click “Yes, Delete it”.	Uploaded “Crop Analysis” dataset	It should delete the page.	It deleted the page.	Pass
TC37	Notebook Page Open Dataset Tools	1. Click the Open Dataset Tools Button at the lower right side.2. Select a Dataset “Penguin CSV”.	Uploaded “Penguin CSV” dataset	It should set the dataset for Excel and Analysis Tabs.	It set the dataset for Excel and Analysis Tabs.	Pass

Bookshelves Page:

Test Case ID	**Test Scenario**	**Test Steps**	**Test Data**	**Expected Results**	**Actual Result**	**Status**
TC38	Edit Functionality	1. Click the Edit button of “Iris Bookshelf” Bookshelf.2. Change the name to “Iris Bookshelf 2”.	Created Bookshelf “Iris Bookshelf”	It should update the Bookshelf.	It updated the Bookshelf.	Pass
TC39	Creating Bookshelf with all forms answered	1. Click the Add Notebook.2. Add a Title “Iris Bookshelf”.3. Click Upload.	Title: “Iris Bookshelf”	It should prompt saying the Bookshelf is successfully created.	It prompted saying the Bookshelf is successfully created.	Pass
TC40	Selecting a Bookshelf	1. Click “Iris Bookshelf” anywhere inside its row.	Created Bookshelf “Iris Bookshelf”	It should redirect to the “Iris Bookshelf” Bookshelf Page.	It redirected to the “Iris Bookshelf” Bookshelf Page.	Pass

TC41	Create Notebook to Bookshelf	1.Click the New Button. 2.Select Add Notebook 3.Add Notebook "Crop Notebook" 4.Click Add to Bookshelf.	Created Bookshelf "Iris Bookshelf	It should prompt saying "Success Pages created successfully " and adds the page to the Bookshelf .	It prompted saying "Success Pages created successfully " and adds the page to the Bookshelf .	Pass
TC42	Add Existing Notebook to Bookshelf	1.Click the New Button. 2.Select Add Notebook 3.Find Notebook "Crop Notebook" 4.Select "Crop Notebook" 5.Click Add to Bookshelf	Created Bookshelf "Iris Bookshelf & "Crop Notebook "	It should prompt saying "Success Pages created successfully " and adds the notebook to the Bookshelf .	It prompted saying "Success Pages created successfully " and adds the notebook to the Bookshelf .	Pass
TC43	Bookshelf Page Edit Notebook	1.Click the Edit Button. 2.Change the title to "Iris Notebook" Bookshelf	Created "Iris Notebook " Notebook.	It should update the page.	It should updated the page.	Pass
TC44	Bookshelf Page Delete Notebook	1.Click the Delete Page Button. 2.Click "Yes, Delete it".	Created "Iris Notebook " Notebook.	It should delete the page.	It should delete the page.	Pass

Visualizations Page:

Test Case ID	**Test Scenario**	**Test Steps**	**Test Data**	**Expected Results**	**Actual Result**	**Status**

TC45	Delete Functionality	1. Click the delete button of "Iris flower dataset".2. Click "Yes, Delete it".	Uploaded dataset "Iris flower dataset"	A prompt will say the dataset is deleted and it will remove the dataset.	A prompt popped up saying the dataset is deleted, and it was removed.	Pass
TC46	Right Sidebar Scrolling	1. Click Violin Plot in the right sidebar.	Loaded CSV/Excel Data	It will scroll down automatically to the Violin Plot section.	It scrolled down automatically to the Violin Plot section.	Pass

Archive Page:

Test Case ID	Test Scenario	Test Steps	Test Data	Expected Results	Actual Result	Status
TC47	Unarchive Functionality	1. Click the unarchive button of "Penguin CSV".2. Click "Yes, unarchive it".	Archived dataset "Penguin CSV"	A prompt will say the dataset is unarchived and it will reappear in the Datasets Page.	A prompt popped up saying the dataset is unarchived, and it was shown in the Datasets Page.	Pass
TC48	Delete Functionality	1. Click the delete button of "Penguin CSV".2. Click "Yes, Delete it".	Archived dataset "Penguin CSV"	A prompt will say the dataset is deleted and it will be removed.	A prompt said the dataset is deleted, and it was removed.	Pass

Sidebar:

Test Case ID	Test Scenario	Test Steps	Test Data	Expected Results	Actual Result	Status
TC49	Notebook Page Button	1. Click the Notebooks Page button.	–	It will redirect to the Notebooks Page.	It redirected to the Notebooks Page.	Pass
TC50	Bookshelves Page Button	1. Click the Bookshelves Page button.	–	It will redirect to the Bookshelves Page.	It redirected to the Bookshelves Page.	Pass
TC51	Visualizations Page Button	1. Click the Visualizations Page button.	–	It will redirect to the Visualizations Page.	It redirected to the Visualizations Page.	Pass
TC52	Archive Page Button	1. Click the Archive Page button.	–	It will redirect to the Archive Page.	It redirected to the Archive Page.	Pass
TC53	Datasets Page Button	1. Click the Archive Page button.2. Click the Datasets Page button.	–	It will redirect to the Datasets Page.	It redirected to the Datasets Page.	Pass
TC54	Help Page Button	1. Click the Help Page button.	–	It will redirect to the Help Page.	It redirected to the Help Page.	Pass

Generate Report Component:

Test Case ID	Test Scenario	Test Steps	Test Data	Expected Results	Actual Result	Status
TC55	Open Generate Report Button	1. Click the Generate Report Tab Button on the middle right side.	Uploaded “Penguin CSV” dataset	It should open the drawer from the right side.	It opened the drawer from the right side.	Pass

TC56	Creating a Report for Statistical Report Tab	1. Click the Generate Report Tab Button.2. Type "Tell me more about this dataset" in *What's your analysis goal?*3. Click Generate Report button.	Uploaded "Penguin CSV" dataset & typed analysis goal.	It should load, then the AI will generate a report about the dataset.	It loaded, then the AI generated a report about the dataset.	Pass
TC57	Reset Form for Generate Report	1. Click the Generate Report Tab Button.2. Type analysis goal.3. Click Reset Form.	Uploaded "Penguin CSV" dataset & typed analysis goal.	It should return to default.	It returned to default.	Pass
TC58	Create New Report while AI is loading	1. Click the Generate Report Tab Button.2. Type analysis goal.3. Click Generate Report button.4. Click	Uploaded "Penguin CSV" dataset & typed analysis goal.	It should go back to the AI report configuration but retain the values typed by the user.	It returned to the AI report configuration and retained the typed values.	Pass

		Create New button.				
TC59	Regenerate Report while AI is loading	1. Click the Generate Report Tab Button.2. Type analysis goal.3. Click Regenerate Report button.	Uploaded “Penguin CSV” dataset & typed analysis goal.	It should refresh the AI report.	It refreshed the AI report.	Pass
TC60	Stop Report while AI is loading	1. Click the Generate Report Tab Button.2. Type analysis goal.3. Click Stop Report button.	Uploaded “Penguin CSV” dataset & typed analysis goal.	It should go back to the AI report configuration but retain the values typed by the user.	It returned to the AI report configuration and retained the typed values.	Pass
TC61	Stop Report while AI is streaming	1. Click the Generate Report Tab Button.2. Type analysis goal.3. Click Stop Report button.	Uploaded “Penguin CSV” dataset & typed analysis goal.	It should stop the AI report streaming.	It stopped the AI report streaming.	Pass
TC62	Create New Report	1. Click the Generate	Uploaded “Penguin	It should go back to the AI report	It returned to the AI report configuration	Pass

	while AI is streaming	Report Tab Button.2. Type analysis goal.3. Click Generate Report button.4. Click Create New button.	CSV" dataset & typed analysis goal.	configuration but retain the values typed by the user.	and retained the typed values.	
TC63	Regenerate Report while AI is streaming	1. Click the Generate Report Tab Button.2. Type analysis goal.3. Click Regenerate Report button.	Uploaded "Penguin CSV" dataset & typed analysis goal.	It should refresh the AI report.	It refreshed the AI report.	Pass
TC64	Chatbot Tab Chat	1. Click the Generate Report Tab Button.2. Click Chatbot Tab.3. Type "What is the dataset all about".4. Press Enter.	Uploaded "Penguin CSV" dataset & chatbot query.	The AI should respond with instructions or information.	The AI responded with instructions/information.	Pass

TC65	Chatbot Tab Stop Button	1. Click the Generate Report Tab Button.2. Click Chatbot Tab.3. Type chatbot query.4. Press Enter.5. Click Stop.	Uploaded “Penguin CSV” dataset & chatbot query.	The AI should stop streaming the messages.	The AI stopped streaming the messages.	Pass
TC66	Chatbot Tab Clear Button	1. Click the Generate Report Tab Button.2. Click Chatbot Tab.3. Type chatbot query.4. Press Enter.5. Click Clear.	Uploaded “Penguin CSV” dataset & chatbot query.	The AI chat messages should be removed.	The AI chat messages were removed.	Pass
TC67	Select Focus Dataset	1. Select a dataset from the dropdown to focus into the notebook “Tester”. 2. Select a plot to analyze “Violin Plot”.	Uploaded “Tester” dataset in Notebook.	It should load the dataset to the Generate Report drawer.	It loaded the dataset to the Generate Report drawer.	Pass

TC68	Download Report	1. Click the Generate Report Tab Button.2. Type analysis goal.3. Click Generate Report button.	Uploaded “Penguin CSV” dataset & typed analysis goal.	It should download the report in DOC file format.	It downloaded the report in DOC file format.	Pass

Excel Component:

Test Case ID	Test Scenario	Test Steps	Test Data	Expected Results	Actual Result	Status
TC69	Check Null Values per column	1. Click the Gear Icon 2. Click Data Validation 3. Check Null Values 4. Select in dropdown Count Null Values per Column 5. Select Column Country 6. Apply	Uploaded “Penguin CSV” dataset	It will show the Null values per column.	It shows the Null values per column.	Pass
TC70	Check Count Total Null Values	1. Click the Gear Icon 2. Click Data Validation 3. Check Null Values 4. Select in dropdown Count Total Null Values 5. Apply	Uploaded “Penguin CSV” dataset	It will show the Count Total Null Values.	It shows the Count Total Null Values.	Pass
TC71	Check Percentage of Null Values per column	1. Click the Gear Icon 2. Click Data Validation 3. Check Null Values 4. Select in dropdown Percentage of Null Values per	Uploaded “Penguin CSV” dataset	It will show the Percentage of Null Values per column.	It shows the Percentage of Null Values per column.	Pass

		Column 5. Select Column Country 6. Apply				
TC72	Check Percentage of Null Values in Dataset	1. Click the Gear Icon 2. Click Data Validation 3. Check Null Values 4. Select in dropdown Percentage of Null Values in Dataset 5. Apply	Uploaded “Penguin CSV” dataset	It will show the Percentage of Null Values in Dataset.	It shows the Percentage of Null Values in Dataset.	Pass
TC73	Mode Imputation	1. Click the Gear Icon 2. Click Data Cleaning 3. Click Mode Imputation 4. Click Country 5. Apply	Uploaded “Penguin CSV” dataset	It will impute the null values with the mode.	It imputes the null values with the mode.	Pass
TC74	Mean Imputation	1. Click the Gear Icon 2. Click Data Cleaning 3. Click Mean Imputation 4. Click Country 5. Apply	Uploaded “Penguin CSV” dataset	It will impute the null values with the mean.	It imputes the null values with the mean.	Pass
TC75	Drop All Null Values	1. Click the Gear Icon 2. Click Data Cleaning 3. Click Drop Null Values 4. Select in dropdown Drop all null values 5. Apply	Uploaded “Penguin CSV” dataset	It will drop all null values in the dataset.	It dropped all null values in the dataset.	Pass
TC76	Drop Null Values (single column)	1. Click the Gear Icon 2. Click Data Cleaning 3. Click Drop Null Values 4. Select in dropdown Null values single column 5. Select Country 6. Apply	Uploaded “Penguin CSV” dataset	It will drop null values from the selected column.	It dropped null values from the selected column.	Pass

TC77	Set Placeholder to All Null Values	1. Click the Gear Icon 2. Click Data Cleaning 3. Click Set Placeholder 4. Select in dropdown All null values 5. Add placeholder “Missing” 6. Apply	Uploaded “Penguin CSV” dataset & typed “Missing”	It will replace all null values with “Missing”.	It replaced all null values with “Missing”.	Pass
TC78	Set Placeholder to Single Column Null Values	1. Click the Gear Icon 2. Click Data Cleaning 3. Click Set Placeholder 4. Select in dropdown Single column 5. Select Country 6. Add placeholder “Missing” 7. Apply	Uploaded “Penguin CSV” dataset & typed “Missing”	It will replace null values of the Country column with “Missing”.	It replaced null values of the Country column with “Missing”.	Pass
TC79	Uppercase All Columns	1. Click the Gear Icon 2. Click Formatting 3. Click Text Formatting 4. Select Format Type UPPERCASE 5. Select All columns 6. Apply	Uploaded “Penguin CSV” dataset	It will convert all text values to UPPERCASE.	It converted all text values to UPPERCASE.	Pass
TC80	Uppercase Single Column	1. Click the Gear Icon 2. Click Formatting 3. Click Text Formatting 4. Select Format Type UPPERCASE 5. Select Column species 6. Apply	Uploaded “Penguin CSV” dataset	It will convert the species column to UPPERCASE.	It converted the species column to UPPERCASE.	Pass

TC81	Lowercase All Columns	1. Click the Gear Icon 2. Click Formatting 3. Click Text Formatting 4. Select Format Type lowercase 5. Select All columns 6. Apply	Uploaded “Penguin CSV” dataset	It will convert all text values to lowercase.	It converted all text values to lowercase.	Pass
TC82	Lowercase Single Column	1. Click the Gear Icon 2. Click Formatting 3. Click Text Formatting 4. Select Format Type lowercase 5. Select Column species 6. Apply	Uploaded “Penguin CSV” dataset	It will convert the species column to lowercase.	It converted the species column to lowercase.	Pass
TC83	Title Case All Columns	1. Click the Gear Icon 2. Click Formatting 3. Click Text Formatting 4. Select Format Type Title Case 5. Select All columns 6. Apply	Uploaded “Penguin CSV” dataset	It will convert all text values to Title Case.	It converted all text values to Title Case.	Pass
TC84	Title Case Single Column	1. Click the Gear Icon 2. Click Formatting 3. Click Text Formatting 4. Select Format Type Title Case 5. Select Column species 6. Apply	Uploaded “Penguin CSV” dataset	It will convert the species column to Title Case.	It converted the species column to Title Case.	Pass
TC85	Remove Spaces All Columns	1. Click the Gear Icon 2. Click Formatting 3. Click Text Formatting 4. Select Format Type Remove	Uploaded “Penguin CSV” dataset	It will remove spaces from all values.	It removed spaces from all values.	Pass

		Spaces 5. Select All columns 6. Apply				
TC86	Remove Spaces Single Column	1. Click the Gear Icon 2. Click Formatting 3. Click Text Formatting 4. Select Format Type Remove Spaces 5. Select Column species 6. Apply	Uploaded “Penguin CSV” dataset	It will remove spaces from the species column.	It removed spaces from the species column.	Pass
TC87	Convert Spelled Number All Columns	1. Click the Gear Icon 2. Click Formatting 3. Click Number Formatting 4. Select Format Type Convert Spelled Number 5. Select All columns 6. Apply	Uploaded “Penguin CSV” dataset	It will convert spelled numbers to digits in all columns.	It converted spelled numbers to digits in all columns.	Pass
TC88	Convert Spelled Number Single Column	1. Click the Gear Icon 2. Click Formatting 3. Click Number Formatting 4. Select Format Type Convert Spelled Number 5. Select Column species 6. Apply	Uploaded “Penguin CSV” dataset	It will convert spelled numbers in the species column.	It converted spelled numbers in the species column.	Pass
TC89	Limit Decimal Points All Columns	1. Click the Gear Icon 2. Click Formatting 3. Click Number Formatting 4. Select Format Type Limit Decimal Points 5. Select “2” in Decimal Places dropdown 6.	Uploaded “Penguin CSV” dataset & selected 2 decimal points	It will limit decimal points to 2 for all numeric values.	It limited decimal points to 2 for all numeric values.	Pass

		Select All columns 7. Apply				
TC90	Limit Decimal Points Single Column	1. Click the Gear Icon 2. Click Formatting 3. Click Number Formatting 4. Select Format Type Limit Decimal Points 5. Select “2” in Decimal Places dropdown 6. Select Column body_mass_g 7. Apply	Uploaded “Penguin CSV” dataset & selected 2 decimal points	It will limit decimal points to 2 in the body_mass_g column.	It limited decimal points to 2 in the body_mass_g column.	Pass
TC91	Change Date Format All Columns	1. Click the Gear Icon 2. Click Formatting 3. Click Date Formatting 4. Select Format Type ISO 5. Select All columns 6. Apply	Uploaded “AOG-USD Time Series” dataset	It will convert all date formats to ISO format.	It converted all date formats to ISO format.	Pass
TC92	Change Date Format Single Column	1. Click the Gear Icon 2. Click Formatting 3. Click Date Formatting 4. Select Format Type **ISO** 5. Select Column **volume** 6. Apply	Uploaded “AOG-USD Time Series” dataset	It will convert the **volume** column date format to ISO.	It converted the **volume** column date format to ISO.	Pass
TC93	Search Filter by Column	1. Click the dropdown beside the search bar 2. Select **sex** 3. Type “Male” in the search bar 4. Press Enter	Uploaded “Penguin CSV” dataset & searched “Male”	It will filter rows where the **sex** column = Male.	It filtered rows where the **sex** column = Male.	Pass

TC94	Search Filter All Columns	1. Click the search bar 2. Type “Male” 3. Press Enter	Uploaded “Penguin CSV” dataset & searched “Male”	It will filter rows with “Male” across all columns.	It filtered rows with “Male” across all columns.	Pass
TC95	Insert Row Above	1. Right-click cell (species, row 1) 2. Click Insert Row Above	Uploaded “Penguin CSV” dataset	It will insert a new row above.	It inserted a new row above.	Pass
TC96	Insert Row Below	1. Right-click cell (species, row 1) 2. Click Insert Row Below	Uploaded “Penguin CSV” dataset	It will insert a new row below.	It inserted a new row below.	Pass
TC97	Delete Row	1. Right-click cell (species, row 1) 2. Click Delete Row	Uploaded “Penguin CSV” dataset	It will delete the row.	It deleted the row.	Pass
TC98	Insert Column Left	1. Right-click cell (species, row 1) 2. Click Insert Column Left 3. Type “Column1” 4. Press OK	Uploaded “Penguin CSV” dataset	It will insert a new column to the left.	It inserted a new column to the left.	Pass
TC99	Insert Column Right	1. Right-click cell (species, row 1) 2. Click Insert Column Right 3. Type “Column2” 4. Press OK	Uploaded “Penguin CSV” dataset	It will insert a new column to the right.	It inserted a new column to the right.	Pass
TC100	Delete Column	1. Right-click cell (species, row 1) 2. Click Delete Column	Uploaded “Penguin CSV” dataset	It will delete the column.	It deleted the column.	Pass
TC101	Excel Formula	1. Click the first row of culmen_length_	Uploaded “Pengui	It will compute the formula and	It computed the formula and	Pass

		mm 2. Double-click the cell 3. Type formula =SUM(D1+D2)	n CSV" dataset	display result in the formula bar and cell.	displayed result in the formula bar and cell.	
TC102	Update Dataset Values	1. Click the first row of column **species** 2. Double-click the cell 3. Rename "Adelie" to "Adelies" 4. Click Update Dataset	Uploaded "Penguin CSV" dataset	It will prompt that dataset is updated successfully and values updated.	It prompted success and updated dataset values.	Pass

Search Component:

Test Case ID	Test Scenario	Test Steps	Test Data	Expected Results	Actual Result	Status
TC103	Search Functionality with match	1. Click the search bar.2. Type "Iris".	Uploaded "Iris CSV" dataset	It will display datasets that start with "Iris".	It displayed datasets that start with "Iris".	Pass
TC104	Search Functionality without match	1. Click the search bar.2. Type "xyz".	Uploaded dataset without the name "xyz"	It will not display any datasets.	It did not display any datasets.	Pass

Pagination Component:

Test Case ID	Test Scenario	Test Steps	Test Data	Expected Results	Actual Result	Status
TC105	Page Pagination	1. Scroll down to the bottom of the page.2. Click "2" in the pagination.	Uploaded 17 datasets	It should show the 2nd page of datasets uploaded.	It showed the 2nd page of datasets uploaded.	Pass
TC106	Page Pagination Right Side Button	1. Scroll down to the bottom of the page.2.	Uploaded 17 datasets	It should show the next page	It showed the next page of	Pass

		Click “>” in the pagination.		of datasets uploaded.	datasets uploaded.	
TC107	Page Pagination Double Right Side Button	1. Scroll down to the bottom of the page.2. Click “>>” in the pagination.	Uploaded 33 datasets	It should show the last page of datasets uploaded.	It showed the last page of datasets uploaded.	Pass
TC108	Page Pagination Left Side Button	1. Scroll down to the bottom of the page.2. Click “2” in pagination.3. Click “<” in the pagination.	Uploaded 17 datasets	It should show the previous page of datasets uploaded.	It showed the previous page of datasets uploaded.	Pass
TC109	Page Pagination Double Left Side Button	1. Scroll down to the bottom of the page.2. Click the last page in pagination.3. Click “<<” in the pagination.	Uploaded 33 datasets	It should show the first page of datasets uploaded.	It showed the first page of datasets uploaded.	Pass

UI Change Component:

Test Case ID	Test Scenario	Test Steps	Test Data	Expected Results	Actual Result	Status
TC110	Changing to Table View	1. Click the Table View Button	Currently in Grid View & Loaded CSV/Excel Data	It will display the datasets in a table format.	It displayed the datasets in table format.	Pass
TC111	Changing to Grid View	1. Click the Grid View Button	Currently in Table View & Loaded CSV/Excel Data	It will display the datasets in a card format.	It displayed the datasets in a card format.	Pass
TC112	Loading Screen	1. Click Notebooks Page	Loaded Notebooks	It will show a statistical loading screen.	It showed a statistical loading screen.	Pass

Plot Recommendation Component

Test Case ID	Test Scenario	Test Steps	Test Data	Expected Results	Actual Result	Status
TC113	Dataset Page Back Button	1. Click the upper left side button “<-“.	Uploaded “Penguin CSV” dataset	It should redirect to the Datasets Page.	It redirected to the Datasets Page.	Pass
TC114	Open Plot Recommendation Drawer	1. Press Chat Recommendation Button at the lower right side	Uploaded “Penguin CSV” dataset	It will open the drawer from the right side.	It opened the drawer from the right side.	Pass
TC115	Generate Plot Recommendation	1. Press Chat Recommendation Button at the lower right side2. Press Generate Button	Uploaded “Penguin CSV” dataset	The AI shall create a prompt suggesting the best visualization type for the dataset.	The AI created a prompt suggesting the best visualization type for the dataset.	Pass

Plot Tools:

Test Case ID	Test Scenario	Test Steps	Test Data	Expected Results	Actual Result	Status
TC116	Download Plot as PNG	1. Press the camera button at the upper right side.	Uploaded “Penguin CSV” dataset in Scatterplot visualization	It will download the plot as PNG.	It downloaded the plot as PNG.	Pass
TC117	Zoom	1. Press the Zoom tool.2. Drag an area of the plot.	Uploaded “Penguin CSV” dataset in Scatterpl	It will zoom into the selected area of the plot.	It zoomed into the selected area of the plot.	Pass

			ot visualization			
TC118	Pan	1. Press the Pan tool.2. Drag the plot area to pan.	Uploaded “Penguin CSV” dataset in Scatterplot visualization	It will pan the visualization.	It panned the visualization.	Pass
TC119	Box Select	1. Press the Box Select tool.2. Drag the mouse to select an area of the plot.	Uploaded “Penguin CSV” dataset in Scatterplot visualization	It will select the points within the drawn box.	It selected the points within the drawn box.	Pass
TC120	Lasso Select	1. Press the Lasso Select tool.2. Drag the mouse in a lasso shape to select points.	Uploaded “Penguin CSV” dataset in Scatterplot visualization	It will select points using lasso.	It selected points using lasso.	Pass
TC121	Zoom In	1. Press the Zoom In button.	Uploaded “Penguin CSV” dataset in Scatterplot visualization	It will zoom in on the visualization.	It zoomed in on the visualization.	Pass

TC122	Zoom Out	1. Press the Zoom Out button.	Uploaded “Penguin CSV” dataset in Scatterplot visualization	It will zoom out of the visualization.	It zoomed out of the visualization.	Pass
TC123	Autoscale	1. Press the Autoscale button.	Uploaded “Penguin CSV” dataset in Scatterplot visualization	It will autoscale the visualization to fit the data.	It autoscaled the visualization to fit the data.	Pass
TC124	Reset Axes	1. Press the Zoom Out button.2. Press the Reset Axes button.	Uploaded “Penguin CSV” dataset in Scatterplot visualization	It will reset the axes to default.	It reset the axes to default.	Pass
TC125	Unselect a Legend	1. Click “Adelie” in the legend.	Uploaded “Penguin CSV” dataset in Scatterplot visualization	It will filter out the “Adelie” data from the plot.	It filtered out the “Adelie” data from the plot.	Pass
TC126	Isolate Trace	1. Double-click “Chinstrap” in the legend.	Uploaded “Penguin CSV” dataset	It will isolate the “Chinstrap” data in the plot.	It isolated the “Chinstrap” data in the plot.	Pass

			in Scatterplot visualization			
TC127	Orbital Rotation	1. Click the Orbital Rotation tool.2. Drag the visualization to rotate in orbital motion.	Uploaded “Penguin CSV” dataset in 3D Scatterplot visualization	It will rotate the visualization in orbital motion.	It rotated the visualization in orbital motion.	Pass
TC128	Turntable Rotation	1. Click the Turntable Rotation tool.2. Drag the visualization to rotate in turntable motion.	Uploaded “Penguin CSV” dataset in 3D Scatterplot visualization	It will rotate the visualization in turntable motion.	It rotated the visualization in turntable motion.	Pass
TC129	Select Column	1. Press the Select Column dropdown.2. Select “sex” and “culmen_depth_mm”.	Uploaded “Penguin CSV” dataset in Scatter Matrix visualization	It will show only the “sex” and “culmen_depth_mm” columns.	It showed only the “sex” and “culmen_depth_mm” columns.	Pass

Analysis Tabs Component:

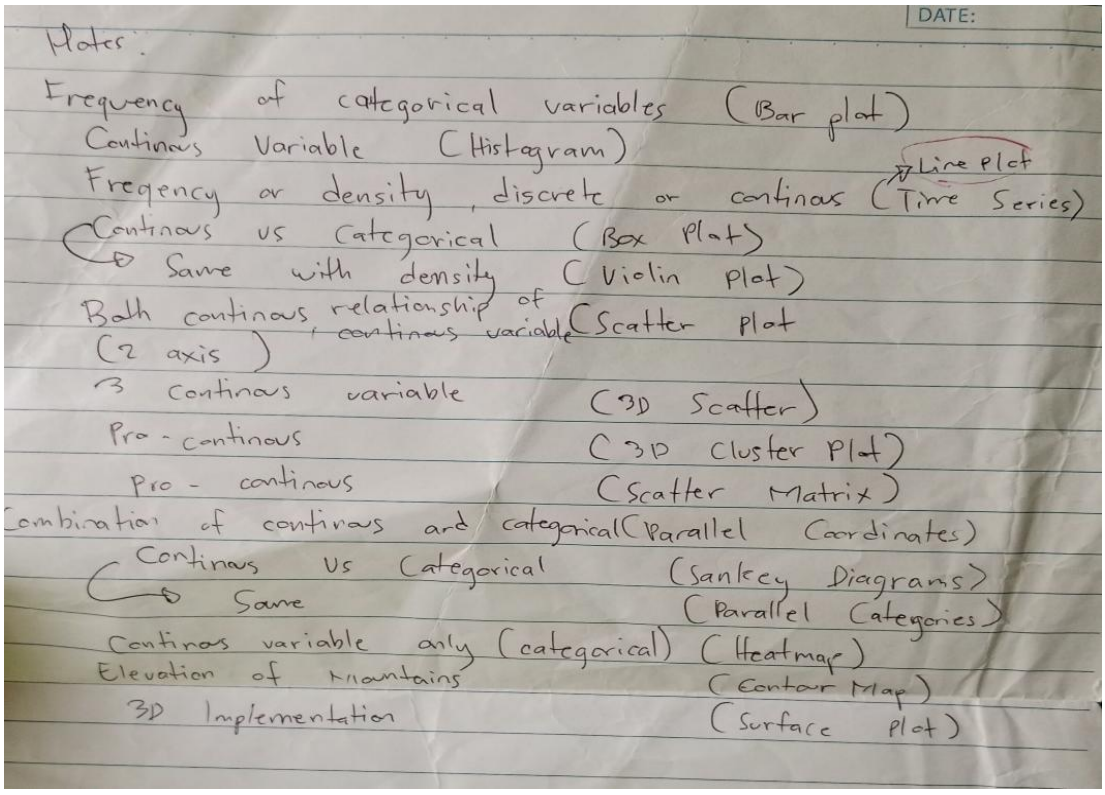
Test Case ID	**Test Scenario**	**Test Steps**	**Test Data**	**Expected Results**	**Actual Result**	**Status**
TC130	Descriptive Statistics (Numerical)	1. Click Descriptive Statistics tab.2. Click Add Column → “culmen_length_m	Uploaded “Penguin CSV” dataset	It will show the numerical descriptive	It showed the numerical descriptive	Pass

		m".3. Click the body of the accordion to show results.		statistics of the dataset.	statistics of the dataset.	
TC131	Descriptive Statistics (Categorical)	1. Click Descriptive Statistics tab.2. Click Add Column → "sex".3. Click the body of the accordion to show results.	Uploaded "Penguin CSV" dataset	It will show the categorical descriptive statistics of the dataset.	It showed the categorical descriptive statistics of the dataset.	Pass
TC132	Pearson Correlation	1. Click Correlation tab.2. Select Pearson Correlation from dropdown.3. Expand accordion.4. Tick All Columns.5. Run.	Uploaded "Penguin CSV" dataset	It will calculate the Pearson Correlation.	It calculated the Pearson Correlation.	Pass
TC133	Spearman Correlation	1. Click Correlation tab.2. Select Spearman Correlation from dropdown.3. Expand accordion.4. Tick All Columns.5. Run.	Uploaded "Penguin CSV" dataset	It will calculate the Spearman Correlation.	It calculated the Spearman Correlation.	Pass
TC134	Visualize Correlation (Heatmap)	1. Click Correlation tab.2. Select Spearman Correlation.3. Expand accordion.4. Tick All Columns.5. Run Analysis.6. Tick Show Correlation Heatmap.	Uploaded "Penguin CSV" dataset	It will visualize the correlation in a heatmap.	It visualized the correlation in a heatmap.	Pass
TC135	Visualize Correlation (3D Surface Plot)	1. Click Correlation tab.2. Select Spearman Correlation.3. Expand accordion.4. Tick All Columns.5. Run	Uploaded "Penguin CSV" dataset	It will visualize the correlation in a 3D surface plot.	It visualized the correlation in a 3D surface plot.	Pass

		Analysis.6. Tick Show 3D Surface Plot.				
TC136	Visualize Correlation (Contour Plot)	1. Click Correlation tab.2. Select Spearman Correlation.3. Expand accordion.4. Tick All Columns.5. Run Analysis.6. Tick Show Contour Plot.	Uploaded "Penguin CSV" dataset	It will visualize the correlation in a contour plot.	It visualized the correlation in a contour plot.	Pass
TC137	Normality Test	1. Click Tests tab.2. Select "culmen_length_mm" from dropdown.3. Tick All Tests.4. Click Run Test.	Uploaded "Penguin CSV" dataset	It will calculate Shapiro–Wilk, Kolmogorov–Smirnov, and Anderson–Darling tests on the selected numeric column.	It calculated Shapiro–Wilk, Kolmogorov–Smirnov, and Anderson–Darling tests on the selected numeric column.	Pass
TC138	Simple Linear Regression	1. Click Regression tab.2. Select Simple Linear Regression.3. Set Dependent Variable = "culmen_length_mm".4. Set Independent Variable = "culmen_depth_mm".5. Run Analysis.	Uploaded "Penguin CSV" dataset	It will calculate the Simple Linear Regression of the dataset.	It calculated the Simple Linear Regression of the dataset.	Pass

Appendix F.

List of Plots Gathered for Development



Appendix G.

User Evaluation Form

DATA PRIVACY NOTICE

We respect your privacy. This survey is part of a research study on **AL1SE: Deepseek-R1 AI Integrated Data Visualization Software For The Department Of Statistics Of Central Luzon State University (CLSU).**

By answering, you are helping us evaluate the software's usability, effectiveness, and overall user experience using the Original Technology Acceptance Model (TAM).

- **What we'll collect:** Your answers about usability, perceived usefulness, ease of use, and some basic information like your department/program (name is optional).
- **How we'll use it:** Your responses will be used only for academic research. Results will be combined and reported anonymously, ensuring no individual can be identified.
- **How we'll protect it**: Your answers will be securely stored in a secured location, accessible only to the research team. Data will be retained only for the duration of the capstone project and will then be properly disposed of in compliance with the **Data Privacy Act of 2012 (RA 10173).**
- **Your rights:** You may skip any question, withdraw anytime, or request for your data to be removed.

If you have any questions, please contact:

Researchers/Developers Team
Camille D. Juego – juego.camille@clsu2.edu.ph
Jefferson R. Pajaron – pajaron.jefferson@clsu2.edu.ph
Franz R. Bugtong – bugtong.franz@clsu2.edu.ph

CONSENT

By putting a check in the box below, you confirm that:

☐ I have read and understood the Data Privacy Notice, and I freely give my consent to participate in this survey.

USER EVALUATION FORM

I. Demographics

FIELD	RESPONSE
Name (Optional):	
Department/Program:	
Experience with Data Visualization Tools:	☐ None ☐ Beginner ☐ Intermediate ☐ Advanced
User Type:	☐ Faculty ☐ Student

II. User Evaluation (TAM Original)
Instructions: Please rate the following statements based on your experience with AL1SE.

1 = Strongly Disagree 2 = Disagree **3 = Agree** **4 = Strongly Agree**

Statement	Rating (1-4)
Perceived Usefulness (PU)	
AL1SE helps me understand my datasets better.	

The data visualizations I need are available in AL1SE.	
AL1SE helps me produce accurate and reliable visualizations.	
I can complete data visualization tasks efficiently using AL1SE.	
The cost-free access and no-coding requirement make AL1SE a practical tool for data visualization.	
The AI-generated reports in AL1SE provide sufficient detail to help me prepare my own reports more efficiently.	
The AI plot recommendations in AL1SE help me choose appropriate visualizations for my datasets.	
The overall features of AL1SE are useful to perform better analysis.	
Perceived Ease of Use (PEOU)	
AL1SE is easy to use without prior technical knowledge.	
AL1SE allows me to customize visualizations according to my needs.	
AL1SE operates consistently with minimal errors during my use.	
I was able to learn how to use AL1SE quickly.	
Interactive visualizations in AL1SE are easy to manipulate and explore.	
Overall, using AL1SE feels simple and easy to use.	

III. Comments and Suggestions

1. What did you like most about AL1SE?

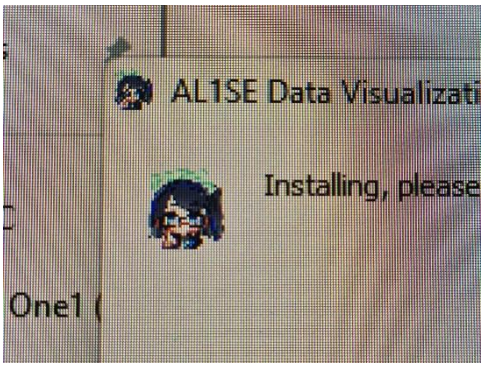
2. Were there any difficulties or challenges you faced while using AL1SE?

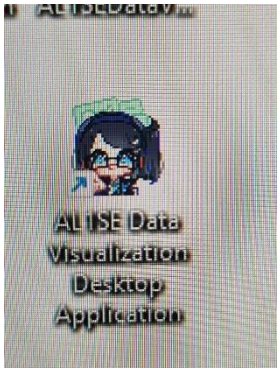
3. Are there any other features you would like to see added to AL1SE in future enhancements?

Appendix H.

AL1SE Logo

Artist: Chimanju





Appendix I.

AL1SE Demonstration Video

Appendix J.

System Output Verification Certificate

VERIFICATION CERTIFICATE

This is to certify that, as a statistician at the **Department of Statistics, Central Luzon State University (CLSU)**, I have reviewed and verified the outputs of the software developed for Capstone 2, entitled **"AL1SE: Data Visualization Software for the Department of Statistics at Central Luzon State University (CLSU)"**, created by Franz R. Bugtong, Camille D. Juego, and Jefferson R. Pajaron.

The core features of the software, which are not AI-based, are accurate and correct. Moreover, I have reviewed the AI-integrated features, and they performed acceptably as additional tools to provide guidance.

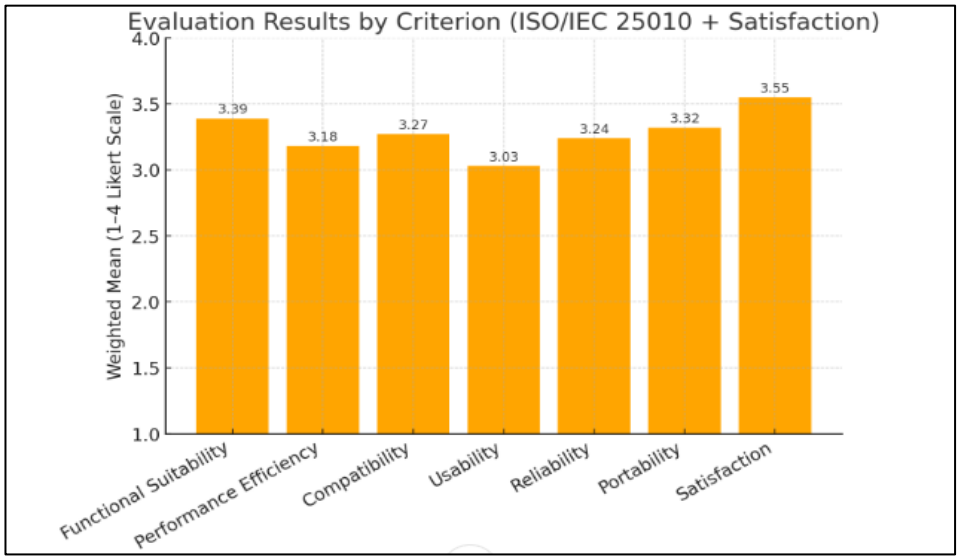
Verified by:

MSt JEAN R ESTOLANO 10/09/2025

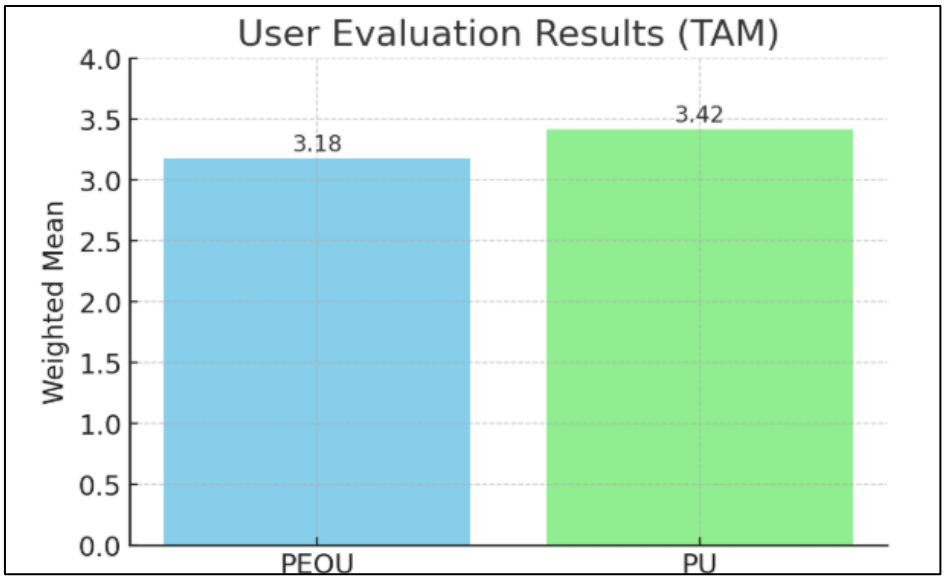
SIGNATURE OVER PRINTED NAME

Appendix K.

Testing Evaluation Results (Graph)



IT Experts Evaluation Result



User Evaluation Result

Appendix L.

Curriculum Vitae Per Member

Franz R. Bugtong

+639673923878 | LinkedIn | GitHub | bugtong.franz@clsu2.edu.ph | franzbugtong0411@gmail.com

Central Luzon State University, Science City Of Munoz, Nueva Ecija

EDUCATION

- **Central Luzon State University** — Aug 2022 - Present
  Bachelor of Science in Information Technology, Major in Systems Development — Science City of Munoz, Philippines
  - Relevant Coursework: Full Stack Development (Django and ReactJS), Interactive Data Visualization, REST API development, Artificial Intelligence, Database Management, JWT Token Authentication, RSA and AES Cryptography

PROJECTS

- **Data Visualization Software for the Department of Statistics (CLSU)** — Aug 2024 - Nov 2024
  Tools: Python (Django REST Framework), JavaScript (ReactJS, D3.JS, NodeJS, Axios etc.),
  - Led a team of eight in a Scrum project, developing software for a client in Central Luzon State University.
  - Utilized Django (Python) and ReactJS (JavaScript), allowing for full stack development.
  - Utilized D3.JS, ReactJS and npm packages for generating a wide range of interactive Data Visualizations (scatter plots, line plots, bar plots) for several dataset formats.

- **AI Integrated Data Visualization Software for the Department of Statistics (CLSU)** — Dec 2024 - Present
  Tools: Python (Django REST Framework), JavaScript (ReactJS, Plotly.JS, NodeJS, Axios etc.), AI (DeepSeek-R1, Llama 3.2, Ollama)
  - Led a team of three in a Capstone/Thesis project in Central Luzon State University.
  - Utilized Django (Python) and ReactJS (JavaScript) to facilitate backend and frontend development.
  - Utilized Plotly.JS, ReactJS, Plotly.JS and npm packages for generating a wide range of interactive Data Visualizations (3D Scatter plots, Parallel Coordinates etc.) and data analytics.
  - Integrated Ollama in Django backend to facilitate AI integration for report automation and chat bot assistance.
  - Performs Data Visualization, Data Analytics, Automated Report Generation, and AI chat bot assistance for given dataset.

- **End-to-End Encrypted Messaging with Hybrid RSA-AES encryption** — Mar 2025 - Mar 2025
  Tools: Python (Django, RSA, AES encryption packages), JavaScript (ReactJS, Material UI etc.)
  - Led a team of three in a cryptography project in Central Luzon State University.
  - Utilized Django, RSA, and AES (Python) packages for achieving end-to-end encryption.
  - Utilized Django REST Framework to facilitate backend development.
  - Utilized ReactJS, JWT Token authentication, and Material UI to facilitate frontend development.
  - Performed End-to-End encryption, ensuring password-protected messages for every user.

SKILLS

- **Programming Languages:** Python, JavaScript
- **Frameworks and Libraries:** Django, ReactJS, Plotly.JS
- **AI:** Ollama (Python), DeepSeek-R1, Llama 3.2
- **Technologies:** Vim, Git

HONORS AND AWARDS

- **College Scholar** — Dec 2024
  Awarding Institutions: Central Luzon State University
  - Achieved a GWA of under 1.75 for the 1st semester of third year college.

Camille D. Juego

+639515872908|juego.camille@clsu2.edu.ph | camillejuego2@gmail@gmail.com
cj72003 (Camille Juego) |
Central Luzon State University, Science City of Munoz, Nueva Ecija

OBJECTIVE

Seeking to contribute expertise in web development using PHP and Python (Django, CodeIgniter, MVC), UI/UX design with Figma and FlutterFlow, WordPress, project management, research, and technical writing.

EXPERIENCE

- **CLSU Department of Statistics** — *Aug 2024 - present*
  Project Manager — Munoz, Philippines
  - Documented all project meetings and prepared detailed minutes (MoMs).
  - Tracked team tasks and reminded members of deadlines and deliverables.
  - Coordinated and organized all project reports and documentation.
  - Led the writing of the project's research paper/thesis.
- **COMELEC Munoz** — *May 2025*
  DESOTechnicalSupport Staff — Munoz,Philippines
  - Handled all technical issues, including network troubleshooting and system errors.
  - Prepared and managed reports and documentation related to technical operations.

EDUCATION

Central Luzon State University — *2022 - present*
BachelorofScienceinInformationTechnology — Munoz, Philippines
* GPA: 1.5 ◦ 3rd year, 2nd semester

Munoz, National Highschool - ANNEX — *2020*
Senior Highschool — Rizal, Philippines
* Average: 91

Munoz, National Highschool - ANNEX — *2018*
Junior Highschool — Munoz, Philippines
* Average: 88

PROJECTS

AI Integrated Data Visualization Software for the Department of Statistics (CLSU) — *Dec 2024 - present*
Tools: Python (Django REST Framework), JavaScript (ReactJS, D3.JS, NodeJS, Axios etc.),
* Led the writing of the project's research paper/thesis.

Payroll Management System (Tito Solutions Company) — *August 2024 - November 2024*
Tools: Python, Django, HTML, CSS, Bootstrap)
* Managed all project documentation, including SRS, user manuals, and technical reports.
* Developed backend processes and designed frontend components for admin and employee side.

ADDITIONALINFORMATION

Languages: English, Filipino
Interests: Designing, Writing, Traveling

Jefferson R. Pajaron

+63-963-533-5269 | pajaron.jefferson@clsu2.edu.ph | jeffersonpajaron@gmail.com

§ Calezarb (Jefferson Pajaron) |

Central Luzon State University, Science City Of Munoz, Nueva Ecija

OBJECTIVE

Seeking to contribute expertise in web development with React.js, PHP, and Django, along with cross-platform app building in Electron.

EXPERIENCE

• CLSU Department of Statistics

Aug 2024 - present

Programmer — Munoz, Philippines

- Developed a data visualization software.
- Designed the User Interface of the software.
- Implemented Automated AI Reports, Excel functionalities, and Statistical Analysis for the Software
- Created a desktop distributable version of the software.

EDUCATION

Central Luzon State University — 2022 - present
Bachelor of Science in Information Technology — Munoz, Philippines
*GPA: 1.4 ◦3rd year 2nd semester

Rizal National Highschool — 2020
Senior Highschool — Rizal, Philippines
*Average: 91

New Era University — 2018
Junior Highschool — Quezon City, Philippines
*Average: 87

PROJECTS

AI Integrated Data Visualization Software for the Department of Statistics (CLSU) — Dec 2024 - present
Tools: Python (Django REST Framework), JavaScript (ReactJS, D3.JS, NodeJS, Axios etc.),
*Developed Excel functionality for the system.
*Utilizedscipy,stats-models,pandas,scikit-learnforstatisticalanalysisofthesystem.
*Implemented automated AI reports generation and a chatbot using ollama.
*Utilized electronjs in creating a distributable software for the system

SoundWave a Music Streaming Webpage built using PHP — Aug 2025 - Aug 2025
Tools: PHP, Ajax, Jquery.),
*Developed the CRUD functionality of the system.
*Implemented real-time data synchronization with the database.
*Designed the User Interface of the webpage.
*Built responsive and asynchronous components using jQuery and Ajax

ADDITIONAL INFORMATION

Languages: English, Filipino
Interests: Programming, Reading