iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL
FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: III Month of publication: March 2026

DOI: https://doi.org/10.22214/ijraset.2026.78665

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

# No-Code DB Analysis

Visweshkishan V[1], Santhosh Kumar N[2], Mrs. M. Sasikala[3]
*Department of Computer Science and Engineering, K.L.N. College of engineering*

***Abstract: This paper presents a no-code, AI-assisted dataset analysis system designed to simplify end-to-end data processing for tabular datasets. The system enables users to upload spreadsheet files (CSV, XLSX, XLS) through a web interface and automatically performs dataset profiling, cleaning, statistical analysis, visualization, and report generation. A FastAPI-based backend orchestrates the pipeline, while an external workflow automation tool integrates AI capabilities to recommend data cleaning and visualization strategies. The system combines deterministic data processing techniques with AI-driven planning to produce structured insights and downloadable PDF reports. The proposed solution demonstrates how lightweight web technologies and workflow automation can be integrated to create an accessible, automated data analysis pipeline without requiring programming expertise.***

***Keywords: No-Code Data Analysis, Dataset Profiling, Data Cleaning Automation, AI-Assisted Analytics, FastAPI, Data Visualization, Statistical Analysis, Automated Report Generation.***

## I. INTRODUCTION

With the rapid growth of data across industries, there is an increasing need for tools that enable efficient data analysis without requiring extensive programming knowledge. Traditional data analysis workflows involve multiple steps such as data cleaning, preprocessing, statistical analysis, and visualization, which often require expertise in tools like Python, R, or SQL.

To address this challenge, no-code and low-code platforms have emerged, enabling users to perform complex operations through simplified interfaces. However, many existing solutions either lack automation or require integration with multiple tools. This work introduces a unified system that automates the entire lifecycle of tabular dataset analysis. The proposed system allows users to upload datasets via a browser interface, after which the backend performs profiling, applies cleaning operations, conducts statistical analysis, generates visualizations, and produces a comprehensive PDF report. Additionally, the system integrates an AI-driven workflow to dynamically determine optimal cleaning and visualization strategies, enhancing adaptability across diverse datasets.

## II. LITERATURE REVIEW

[1] H. D. Mafukidze, A. Nechibvute, A. Yahya, I. A. Badruddin, S. Kamangar and M. Hussien, "Development of a Modularized Undergraduate Data Science and Big Data Curricular Using No-Code Software Development Tools," in IEEE Access, vol. 12, pp. 100939-100956, 2024, doi: 10.1109/ACCESS.2024.3429241.

Description of a modularized curriculum relying on no-code tools (NCTs) for non-computer science majors.Experimental study using a sample of 50 fourth-year students divided into control and experimental groups.Administration of post-survey questionnaires and assessment items to measure performance impacts.

[2] Popchev, Ivan & Orozova, Daniela. (2023). Algorithms for Machine Learning with Orange System. International Journal of Online and Biomedical Engineering (iJOE). 19. 109-123. 10.3991/ijoe.v19i04.36897.

Formulation of a multi-step algorithm for developing information flow using Orange's machine learning tools.Implementation of supervised learning (classification and regression) using Logistic Regression, Naïve Bayes, SVM, Decision Trees, and Neural Networks.Unsupervised learning application through clustering (k-means and hierarchical).Practical experiments conducted in the domain of "smart crop production".

[3] Dobesova, Zdena. (2024). Evaluation of Orange data mining software and examples for lecturing machine learning tasks in geoinformatics. Computer Applications in Engineering Education. 32. 10.1002/cae.22735.

Application of the "Physics of Notation" (PoN) theory by D. Moody to evaluate the visual vocabulary and effective cognition of Orange.Practical demonstration of machine learning tasks combining Orange with ArcGIS Pro.Use of density-based spatial clustering (DBSCAN) and neural networks followed by hierarchical clustering for spatial data.

[4] A. S. Abdelmagid and A. I. M. Qahmash, “Utilizing the Educational Data Mining Techniques ‘Orange Technology’ for Detecting Patterns and Predicting Academic Performance of University Students,” International Journal of Science and Learning, pp. 1415–1431, doi:10.18576/isl/120330.

Knowledge detection using the "K-Means" clustering algorithm to identify educational patterns.Predictive modeling using Linear Regression, Random Forest, KNN, Decision Tree, and SVM.Data source: Electronic courses (Blackboard LMS) for students at King Khalid University.

[5] M. A. Kuhail, S. Farooq, R. Hammad and M. Bahja, "Characterizing Visual Programming Approaches for End-User Developers: A Systematic Review," in IEEE Access, vol. 9, pp. 14181-14202, 2021, doi: 10.1109/ACCESS.2021.3051043.

The study follows a Systematic Literature Review (SLR) process based on guidelines by Kitchenham and Charter.The researchers systematically analyzed 30 articles published between January 2010 and November 2020.Articles were evaluated against 12 dimensions, including VPL classification, interaction style, target users, domain, platform, and various empirical evaluation metrics.

## III. METHODOLOGY

The No-Code Dataset Analysis System was developed using the Iterative Prototyping Model, which emphasizes gradual enhancement through continuous testing and user feedback. This approach enabled progressive improvements in data processing accuracy, system efficiency, and user experience across multiple development cycles. Each iteration introduced refinements in dataset profiling, cleaning strategies, AI-assisted planning, and frontend-backend interaction, ensuring optimal performance and usability.

The system architecture is structured into four major components — Data Profiling, AI-Assisted Processing, Analysis Engine, and Web-based Interface — each playing a crucial role in delivering automated and accurate dataset analysis.

### *A. Data Profiling*

This phase focuses on transforming raw tabular datasets into structured metadata suitable for further processing. The system accepts datasets in CSV, XLSX, and XLS formats.

Once uploaded, the dataset undergoes profiling to extract essential information such as row count, column count, column types, missing values, duplicate records, and summary statistics.

Additionally, correlation matrices for numeric columns and unique value distributions for categorical columns are computed. This structured profiling enables the system to understand dataset characteristics and prepares it for intelligent processing. The profiling phase ensures that inconsistencies, redundancies, and data quality issues are identified early, improving the effectiveness of subsequent analysis steps.

### *B. AI-Assisted Processing*

The core of the system involves integrating an AI-driven workflow to automate decision-making in data cleaning and visualization planning.

The dataset profile is sent to an external workflow automation tool, where an AI model generates a structured plan consisting of cleaning strategies, relationship analysis requests, and chart suggestions.

The cleaning process includes operations such as handling missing values, removing duplicates, converting data types, normalizing numerical values, encoding categorical variables, and removing outliers. The system also includes fallback mechanisms to ensure continuity even if the AI workflow is unavailable.

This hybrid approach combines deterministic data processing with intelligent AI recommendations, significantly enhancing adaptability across diverse datasets.

### *C. Analysis Engine*

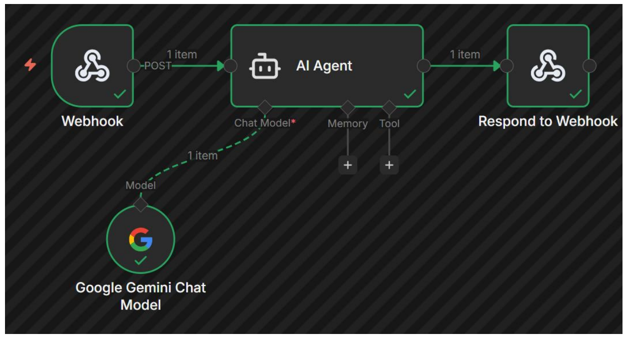
This module acts as the decision-making and computation core of the system. The cleaned dataset is analyzed using statistical techniques to uncover meaningful insights.

The engine performs correlation analysis using Pearson correlation to identify relationships between numeric variables. It also evaluates categorical influence through grouped aggregation, allowing the system to understand how categorical variables impact numeric outcomes.

Visualization plays a key role in this phase, where charts such as histograms, scatter plots, bar charts, box plots, and line graphs are automatically generated. These visualizations help users interpret complex data patterns easily.

The system also generates rule-based natural language insights, summarizing key findings such as strong correlations, trends, and anomalies. The analysis engine ensures accurate, efficient, and interpretable results without requiring manual intervention.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com



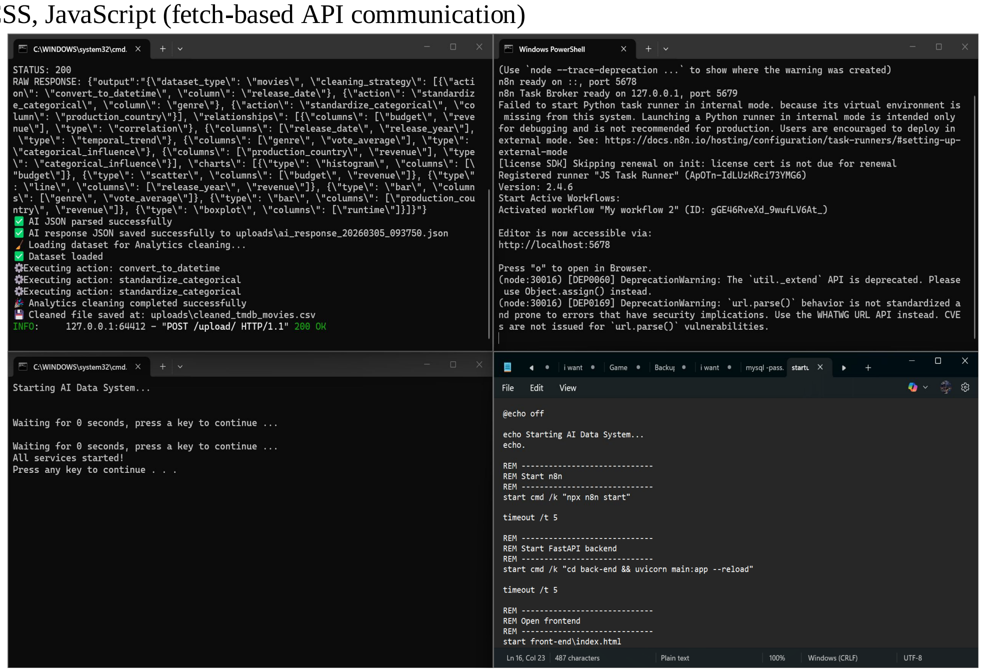

### D. *Web-based Interface*

The frontend of the system is developed using HTML, CSS, and JavaScript, providing a simple, clean, and user-friendly interface. Users can upload datasets through a drag-and-drop mechanism or file selector and initiate the analysis process with minimal effort. The interface communicates with the FastAPI backend via REST API calls, ensuring seamless data transfer and real-time processing. A loading indicator is displayed during processing, and upon completion, users are provided with a downloadable PDF report containing all analysis results.

The results are presented in a clear and structured format, enabling users to easily understand dataset insights. The modular design of the interface allows future enhancements such as dashboards, visualization panels, or integration with other systems.

### E. *Tools and Technologies*

Programming Languages: Python, HTML, CSS, JavaScript
Libraries: Pandas, NumPy, SciPy, Matplotlib, Seaborn, ReportLab
APIs Used: n8n Webhook (AI workflow integration)
Backend Framework: FastAPI
Frontend: HTML, CSS, JavaScript (fetch-based API communication)




## IV. MODELING AND ANALYSIS

The development of the No-Code Dataset Analysis System involved a structured process of system modeling and requirement analysis, ensuring that both functional and non-functional aspects were clearly defined prior to implementation. This phase was essential in transforming the conceptual idea of automated dataset analysis into a practical and efficient framework capable of handling diverse tabular data formats and generating meaningful insights.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

Through detailed analysis, the system was designed to process structured datasets from multiple sources, including CSV and Excel files, which often contain inconsistencies such as missing values, duplicate records, and mixed data types. The modeling phase focused on designing modular components for data profiling, AI-assisted processing, statistical analysis, and report generation — each optimized for efficiency, scalability, and reliability.

The overall system architecture ensures seamless integration between the FastAPI-based backend, AI workflow module, and web-based frontend, enabling a smooth flow from dataset upload to final report generation.

### *A. System Analysis*

Functional Requirements:

The functional requirements define the core operations that the system must perform to achieve its objectives.

The system accepts dataset files in multiple formats such as CSV, XLS, and XLSX through a web interface.

It performs dataset profiling, including detection of column types, missing values, duplicates, and summary statistics.

The system processes the dataset using AI-assisted planning to determine appropriate cleaning and visualization strategies.

It applies data cleaning operations such as handling missing values, removing duplicates, encoding categorical data, and normalizing numerical values.

The analysis engine performs statistical computations, including correlation analysis and categorical influence analysis.

The system automatically generates visualizations such as histograms, scatter plots, bar charts, and box plots.

It produces natural language insights summarizing key findings from the dataset.

A comprehensive PDF report is generated and made available for download via the web interface.

The system supports API-based communication between frontend and backend for seamless integration and processing.

Temporary files and intermediate outputs are managed efficiently to maintain system performance.

These functionalities make the system highly efficient, user-friendly, and suitable for applications such as business analytics, data exploration, and decision support systems.

Non-Functional Requirements:

The non-functional requirements define the quality and performance characteristics of the system.

High Processing Efficiency: Ensures quick dataset processing and report generation with minimal delay.

Scalability: Capable of handling datasets of varying sizes and supporting future expansion with additional modules.

Reliability: Provides consistent and accurate outputs even when processing complex or incomplete datasets.

Usability: Offers a simple and intuitive web interface for users with minimal technical expertise.

Maintainability: Modular architecture allows easy updates and improvements to individual components such as cleaning, analysis, or reporting modules.

Portability: Can be deployed across different environments, including local systems and cloud platforms.

Fault Tolerance: Includes fallback mechanisms to ensure continued operation even if the AI workflow component fails.

Interoperability: Supports integration with external tools and workflows through API-based communication.

### *B. Hardware and Software Requirements*

Hardware Requirements:

A multi-core processor (Intel i5 or above recommended)

Minimum 8 GB RAM (16 GB recommended for large datasets)

Optional GPU support for faster processing (not mandatory)

Sufficient storage for dataset uploads and generated reports

Software Requirements:

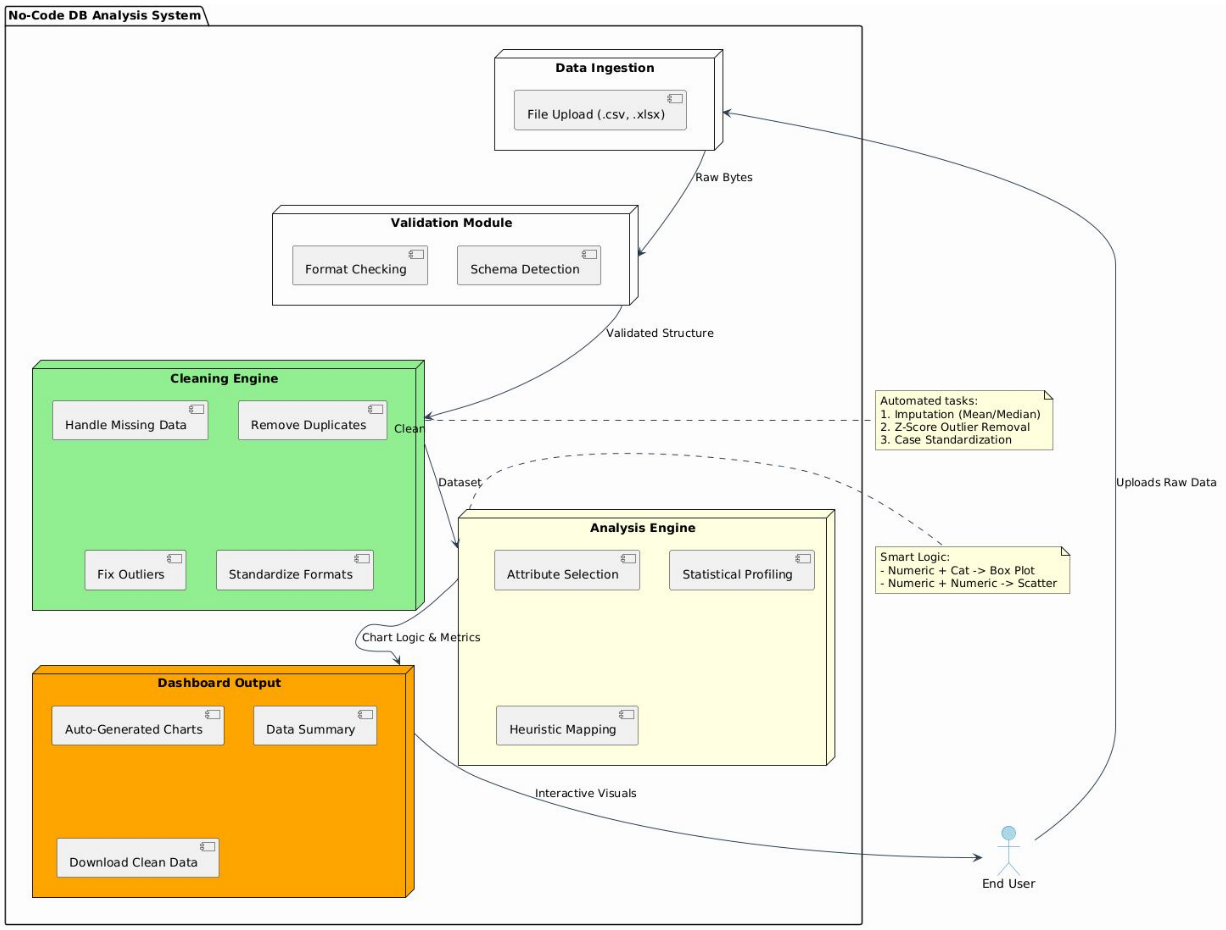
Programming Languages: Python, HTML, CSS, JavaScript

Libraries and Frameworks: Pandas, NumPy, SciPy, Matplotlib, Seaborn, ReportLab

Backend Framework: FastAPI

AI Integration: n8n workflow with AI model support

Frontend Technologies: HTML, CSS, JavaScript (Fetch API for communication)




## V. RESULTS AND DISCUSSION

The No-Code Dataset Analysis System was extensively tested on a variety of structured datasets in formats such as CSV, XLS, and XLSX, covering domains including business data, survey datasets, and general tabular records. The evaluation focused on assessing the system's ability to handle diverse data structures, detect inconsistencies, and generate accurate analytical insights with minimal user intervention. Prior to analysis, the datasets underwent a comprehensive profiling and preprocessing phase. This included identification of missing values, duplicate records, inconsistent data types, and outliers. Cleaning operations such as normalization, encoding, and statistical imputation were applied based on AI-assisted recommendations. These preprocessing steps significantly improved data quality, enabling more reliable statistical analysis and visualization.

The system leverages a hybrid approach combining deterministic statistical methods with AI-assisted planning. The AI workflow analyzes dataset metadata and suggests appropriate cleaning strategies, relationships to explore, and visualization types. This approach enhances adaptability across different datasets, allowing the system to dynamically adjust its processing pipeline without manual configuration. During experimentation, the system demonstrated strong performance in generating meaningful insights. Correlation analysis successfully identified relationships between numerical variables, while categorical influence analysis revealed patterns and trends within grouped data. Compared to traditional manual analysis workflows, the system significantly reduced the time and effort required to derive insights.

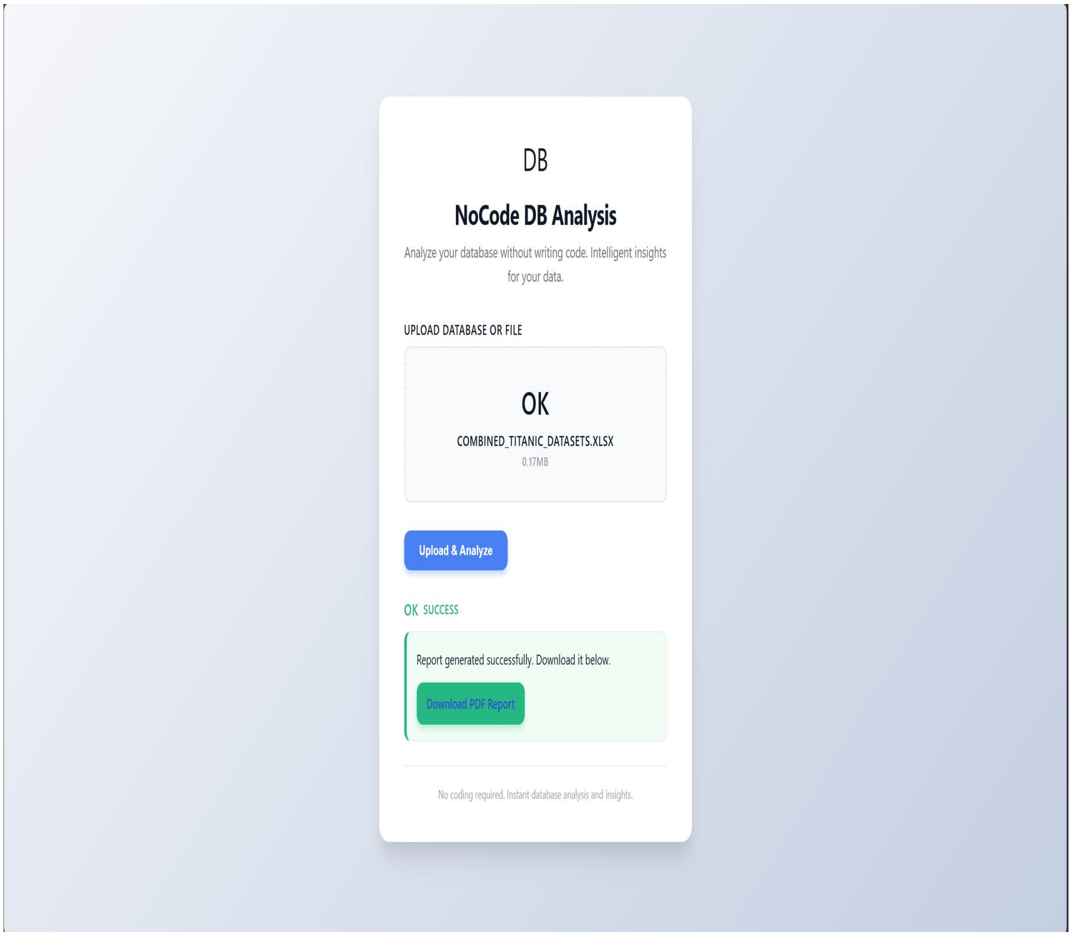
The automatic chart generation feature effectively visualized complex relationships using histograms, scatter plots, bar charts, box plots, and line graphs. These visual outputs improved interpretability and enabled users to quickly understand data distributions and trends without requiring technical expertise.

The FastAPI-based backend ensured efficient processing and smooth execution of the analysis pipeline, while the web-based frontend provided a simple and responsive user interface. Users were able to upload datasets, initiate analysis, and download comprehensive PDF reports with minimal latency. The seamless integration between frontend and backend contributed to a user-friendly and efficient experience.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

The system also demonstrated robustness through fallback mechanisms, ensuring that analysis could proceed even if the AI workflow component was unavailable. This increased reliability and made the system suitable for real-world usage scenarios.

Overall, the results confirm that integrating AI-assisted planning with traditional data processing techniques can significantly enhance automation, efficiency, and usability in dataset analysis systems. The proposed solution effectively bridges the gap between manual data analysis and fully automated pipelines, providing a scalable and practical tool for data-driven decision-making.




## VI. CONCLUSION AND FUTURE WORKS

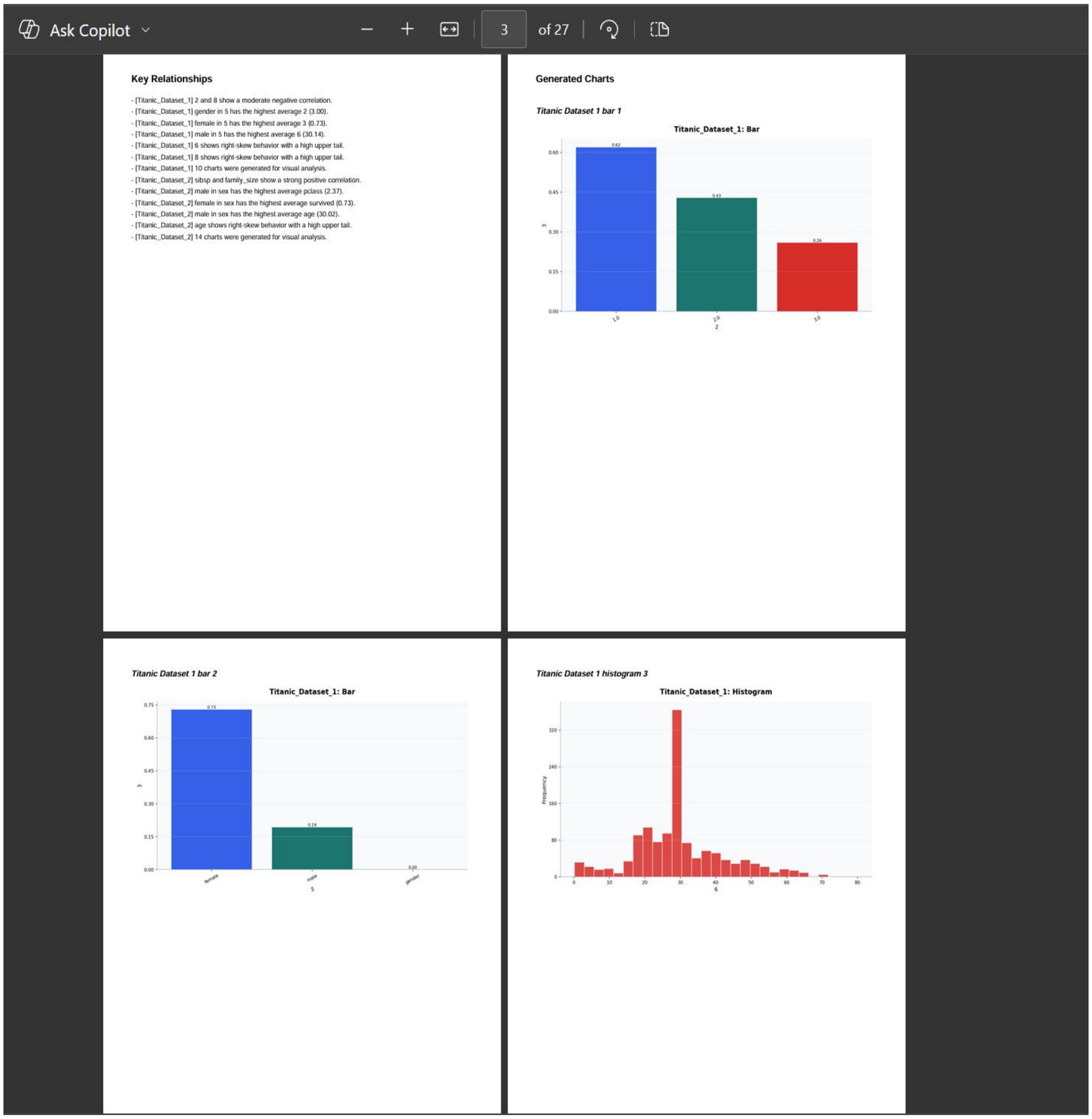
This project presents a robust AI-assisted No-Code Dataset Analysis System designed to automate the complete lifecycle of tabular data processing, from dataset upload to insight generation and report creation. By integrating intelligent workflow automation with traditional data analysis techniques, the system effectively simplifies complex analytical tasks and makes them accessible to users without programming expertise.

The system leverages a FastAPI-based backend for efficient processing and orchestration, combined with a lightweight web-based frontend for seamless user interaction. The integration of an AI-driven workflow enables dynamic decision-making for data cleaning, relationship analysis, and visualization planning, allowing the system to adapt to diverse datasets with minimal manual intervention. Additionally, the implementation of structured data profiling, automated preprocessing, statistical analysis, and visualization ensures accurate and meaningful insights.

The system demonstrates strong capability in handling real-world datasets with inconsistencies such as missing values, duplicates, and mixed data types. Compared to traditional manual analysis approaches, it significantly reduces processing time while improving usability and efficiency. Furthermore, the automated PDF report generation feature provides a comprehensive and portable summary of analytical results, making the system suitable for applications in business intelligence, data exploration, and decision support.

This work highlights the potential of combining AI-assisted planning with deterministic data processing to build scalable and user-friendly analytical platforms. It bridges the gap between complex data science workflows and accessible no-code solutions, enabling wider adoption of data-driven decision-making.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 14 Issue III Mar 2026- Available at www.ijraset.com*




*A. Future Enhancements*

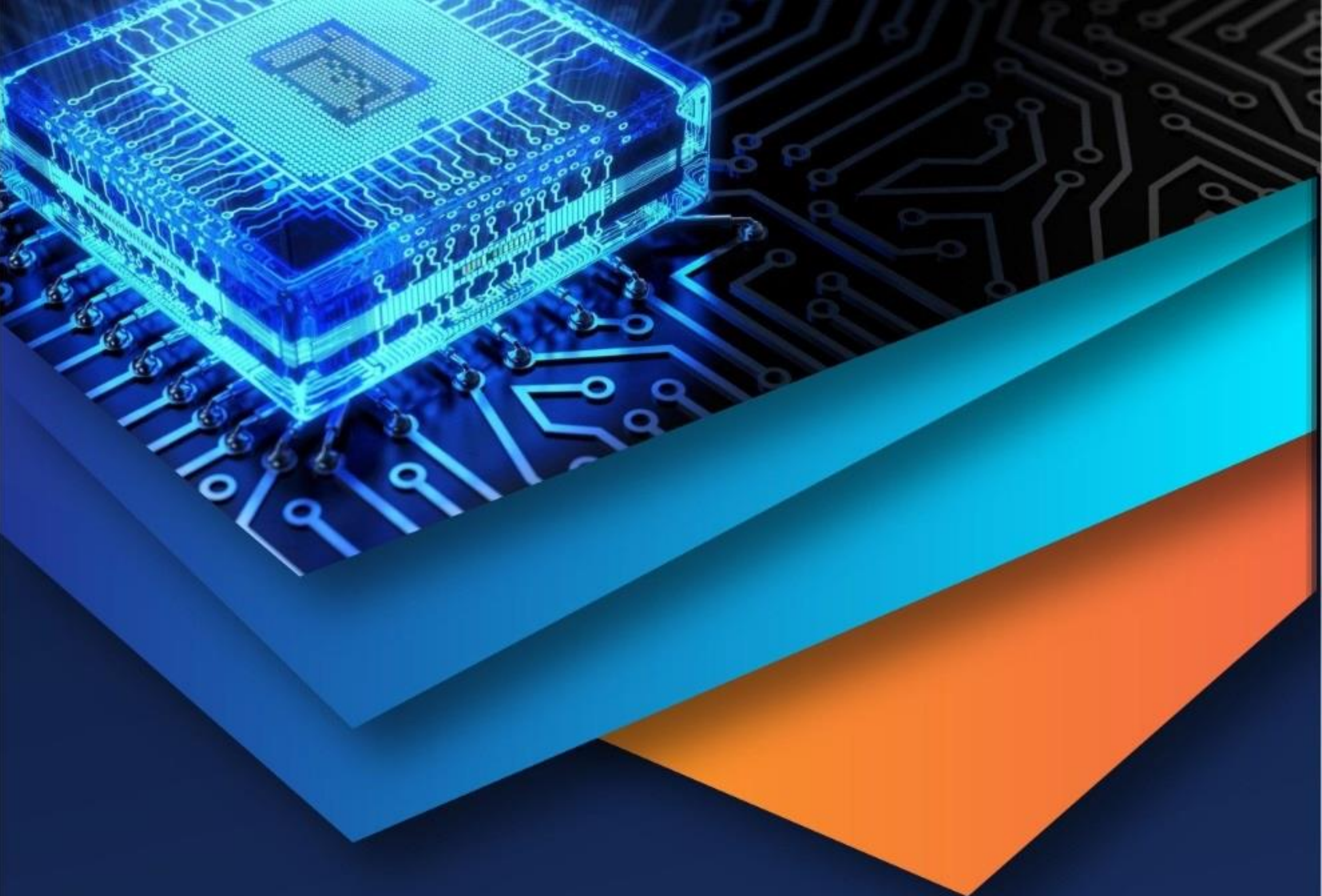
1) Advanced Data Visualization Dashboards: Develop interactive dashboards with dynamic charts and filters to provide deeper insights and real-time data exploration capabilities.
2) Cloud-Based Deployment and Scalability: Transition from a local setup to cloud-based infrastructure for improved scalability, accessibility, and performance in large-scale environments.
3) Automated Workflow Optimization: Enhance the AI workflow to dynamically optimize cleaning strategies and analysis pipelines based on dataset characteristics and past performance.
4) User Management and Authentication:: Implement multi-user support with authentication and role-based access control for secure and collaborative usage.
5) Integration with External Data Sources: Enable direct connections to databases, APIs, and data warehouses for real-time data ingestion and analysis.
6) Predictive Analytics and Machine Learning Integration: Extend the system to include predictive modeling, forecasting, and advanced analytics using machine learning techniques.
7) Enhanced User Interface and Experience: Improve the frontend with richer visual components, progress tracking, and interactive report previews.

## VII. ACKNOWLEDGMENT

We extend our sincere gratitude to Mrs.M.SASIKALA for their invaluable guidance throughout this research. We also thank K.L.N. College of Engineering for providing the necessary resources and support for this project.

## REFERENCES

[1] H. D. Mafukidze, A. Nechibvute, A. Yahya, I. A. Badruddin, S. Kamangar and M. Hussien, "Development of a Modularized Undergraduate Data Science and Big Data Curricular Using No-Code Software Development Tools," in IEEE Access, vol. 12, pp. 100939-100956, 2024, doi: 10.1109/ACCESS.2024.3429241.
[2] Popchev, Ivan & Orozova, Daniela. (2023). Algorithms for Machine Learning with Orange System. International Journal of Online and Biomedical Engineering (iJOE). 19. 109-123. 10.3991/ijoe.v19i04.36897.
[3] Dobesova, Zdena. (2024). Evaluation of Orange data mining software and examples for lecturing machine learning tasks in geoinformatics. Computer Applications in Engineering Education. 32. 10.1002/cae.22735.
[4] S. Abdelmagid and A. I. M. Qahmash, "Utilizing the Educational Data Mining Techniques 'Orange Technology' for Detecting Patterns and Predicting Academic Performance of University Students," International Journal of Science and Learning, pp. 1415–1431, doi:10.18576/isl/120330.
[5] M. A. Kuhail, S. Farooq, R. Hammad and M. Bahja, "Characterizing Visual Programming Approaches for End-User Developers: A Systematic Review," in IEEE Access, vol. 9, pp. 14181-14202, 2021, doi: 10.1109/ACCESS.2021.3051043
[6] P. Abichandani, C. Sivakumar, and M. McGrath,"Internet-of-Things Curriculum, Pedagogy, and Assessment for STEM Education: A Review of Literature,"IEEE Access, vol. 10, pp. 38351–38369, 2022.
[7] G. V. Helden et al.,"The Use of Digital Peer Assessment in Higher Education: An Umbrella Review,"IEEE Access, vol. 11, pp. 23456–23478, 2023.
[8] M. J. Ramzan et al.,"A Conceptual Model to Support Learning Pathways for Data Science Education,"IEEE Access, vol. 9, pp. 132000–132015, 2021.
[9] S. K. Sharma and X. Wang,"Toward Massive Machine Type Communications in Ultra-Dense Cellular IoT Networks: Current Issues and Machine Learning-Assisted Solutions,"IEEE Communications Surveys & Tutorials, 2022.
[10] A. Sarker,"Machine Learning: Algorithms, Real-World Applications and Research Directions,"IEEE Access, vol. 9, pp. 123456–123478, 2021.
[11] M. Chen, Y. Hao, K. Hwang, L. Wang, and L. Wang,"Disease Prediction by Machine Learning Over Big Data From Healthcare Communities,"IEEE Access, 2021.
[12] Z. Zhou et al.,"Deep Learning-Based Data Analytics for Smart Cities,"IEEE Network, 2022.
[13] X. Sun and J. Liu,"Big Data Analytics in Education: A Review,"IEEE Access, 2023.
[14] M. J. Ramzan et al.,"A Conceptual Model to Support Learning Pathways for Data Science Education,"IEEE Access, vol. 9, pp. 132000–132015, 2021.
[15] R. K. Sharma et al.,"Explainable AI for Machine Learning: A Survey,"IEEE Access, 2022.
[16] A. Sarker,"Machine Learning: Algorithms, Real-World Applications and Research Directions,"IEEE Access, vol. 9, pp. 123456–123478, 2021




10.22214/IJRASET

INDEX COPERNICUS 45.98

ISRA JIF IMPACT FACTOR: 7.129

SJIFactor TOGETHER WE REACH THE GOAL IMPACT FACTOR: 7.429

ISI


# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY