

DataWhiz : An adaptive role-based AI analytics platform for dynamic dashboard generation and automated decision support

Jeet J. Dodia *

Aniket Patel [†]

Department of Computer Science & Engineering

Institute of Computer Technology

Ganpat University

Mehsana 384012

Gujarat

India

Abstract

Organizations rely on data-driven decision-making, yet conventional business intelligence platforms demand substantial manual preprocessing and analytical expertise. This paper presents DataWhiz, an adaptive role-based AI analytics platform that automatically transforms structured datasets into intelligent dashboards. Through semantic understanding, automated preprocessing, KPI recommendation, dynamic visualization, predictive modeling, and AI-assisted insight generation, the framework generates context-aware reports tailored to specific professional roles. Evaluated across four public datasets from diverse domains, DataWhiz significantly minimizes manual analytical effort while delivering precise decision support, effectively bridging the accessibility gap for both technical and non-technical enterprise users.

Subject Classification: 68T07, 68T09, 62P25.

Keywords: Artificial intelligence, Business intelligence, Automated data analytics, Adaptive dashboard generation, Decision support systems, Predictive analytics, AutoML, Data visualization, Conversational analytics.

* ORCID-ID: 0009-0000-0253-1022

[†] ORCID-ID: 0000-0001-7616-2610

* E-mail: jeetdodia12@gmail.com (Corresponding Author)

[†] E-mail: ap02@ganpatuniversity.ac.in

1. Introduction

Organizations across finance, healthcare, marketing, education, manufacturing, and research generate massive volumes of structured data every day. Transforming these datasets into meaningful insights is essential for informed decision-making. However, the increasing complexity of data has made analysis challenging for many users, particularly those without expertise in statistics, programming, or business intelligence tools. As a result, organizations require intelligent analytics platforms that can automate data preparation, visualization, and interpretation while reducing manual effort.

Traditional business intelligence tools such as Microsoft Excel, Tableau, and Power BI provide powerful visualization capabilities but require users to manually clean datasets, identify key performance indicators (KPIs), select appropriate charts, configure dashboards, and interpret analytical results. These manual processes increase analysis time, require technical knowledge, and may lead to inconsistent reporting across different users and organizations. Although recent AI-based analytics solutions have introduced greater automation, most existing platforms still require significant user intervention [1]–[3] and do not automatically adapt analytical workflows according to dataset characteristics and user requirements.

To address these limitations, this paper proposes **DataWhiz**, an adaptive role-based AI analytics platform that automates the complete analytical workflow from dataset ingestion to intelligent decision support. The proposed framework automatically performs data preprocessing, duplicate removal, missing value handling, data type detection, semantic understanding of dataset attributes, KPI recommendation, visualization selection, predictive analytics, and AI-generated insight creation. Unlike conventional analytics systems, DataWhiz dynamically generates dashboards based on the uploaded dataset and provides conversational analytics that enables users to interact with their data using natural language queries.

The platform supports multiple professional roles, including Business Analyst, Financial Analyst, Marketing Analyst, Researcher, Healthcare Analyst, and Human Resource Manager. Based on the selected user role and detected dataset characteristics, the system recommends appropriate analytical workflows, visualizations, statistical summaries, forecasting models, and AI-generated reports. Furthermore, the platform is available in both web-based and desktop environments, allowing organizations to

choose between cloud-enabled collaboration and secure offline analysis depending on operational requirements.

Despite recent advances in Business Intelligence platforms, AutoML systems, and LLM-based decision support tools [1]–[8], existing solutions generally focus on isolated functionalities such as dashboard visualization, conversational querying, predictive analytics, or report generation. Few systems integrate automated preprocessing, semantic dataset understanding, adaptive KPI recommendation, visualization selection, predictive analytics, and role-based AI insight generation within a unified framework. Furthermore, most existing tools require substantial manual configuration and are not tailored to different professional users. To address these limitations, this work proposes DataWhiz, an adaptive role-based AI analytics platform that automates the complete analytical workflow from data ingestion to intelligent decision support.

The primary objectives of this research are as follows:

- To design a role-based AI analytics platform.
- To automate dashboard generation.
- To integrate AI-generated insights.
- To evaluate the platform on real-world datasets.

The major contributions of this work are summarized as follows:

1. Development of an adaptive AI-based analytics framework for automated data analysis.
2. Automatic semantic understanding of uploaded datasets for intelligent dashboard generation.
3. Dynamic KPI recommendation and visualization selection based on dataset characteristics.
4. Integration of automated data preprocessing, forecasting, SQL assistance, and AI-generated analytical insights within a unified platform.
5. Support for role-based analytical decision making and conversational analytics to improve accessibility for both technical and non-technical users.

The remainder of this paper is organized as follows. Section 2 presents the related work. Section 3 describes the proposed methodology and

system architecture. Section 4 discusses the adaptive dashboard generation framework. Section 5 presents the experimental evaluation and comparative analysis. Finally, Section 6 concludes the paper and outlines future research directions.

Related Work: Recent studies have explored LLM-based decision support systems for strategic decision-making, supply chain management, clinical decision support, financial analytics, and trustworthy AI-assisted decision making [1]–[6]. Machine learning-driven data analytics have also been applied to domain-specific applications, including healthcare monitoring, diagnostic support, and treatment analysis [7]. Similarly, artificial intelligence has been investigated for optimizing analytical and decision-making processes in digital marketing [8]. However, existing approaches generally focus on specific domains or individual analytical tasks. Unlike these systems, DataWhiz integrates automated preprocessing, semantic dataset understanding, KPI recommendation, predictive analytics, conversational analytics, and AI-generated insights within a unified role-based analytics platform.

2. Proposed Methodology

The proposed DataWhiz framework follows a fully automated analytics pipeline that transforms raw structured datasets into intelligent dashboards and decision-support reports with minimal user intervention. Instead of requiring users to manually clean data, configure visualizations, or select analytical models, the platform automatically identifies dataset characteristics and executes an adaptive analytical workflow. Figure 1 illustrates the overall architecture and workflow of the proposed DataWhiz framework.

The complete analytical workflow consists of eight sequential stages, namely data ingestion, preprocessing, semantic analysis, KPI identification, visualization recommendation, predictive analytics, AI-based insight generation, and dashboard export. Each stage operates independently while exchanging processed information with the subsequent stage, thereby ensuring modularity, scalability, and maintainability.

Step 1: Data Ingestion

Users upload structured datasets in CSV or Microsoft Excel format through either the web or desktop interface. After upload, the system validates the dataset by checking file integrity, supported format, duplicate column names, and schema consistency before initiating further analysis.

Step 2: Automated Data Preprocessing

The preprocessing engine automatically prepares the uploaded dataset by removing duplicate records, detecting missing values, identifying data types, converting date fields into appropriate formats, and standardizing categorical variables. Numerical columns are validated for invalid entries while inconsistent records are handled before analytical processing begins.

Step 3: Semantic Dataset Understanding

Rather than relying only on column names, the framework analyses the semantic meaning and statistical characteristics of dataset attributes to determine their analytical significance. Based on detected relationships, the platform identifies numerical, categorical, temporal, and geographical attributes required for dashboard generation.

Step 4: Intelligent KPI Recommendation

After semantic analysis, the system automatically generates relevant Key Performance Indicators (KPIs) according to dataset characteristics. Examples include total revenue, average sales, customer count, growth percentage, profit margin, employee statistics, and healthcare indicators depending on the uploaded dataset.

Step 5: Adaptive Visualization Recommendation

The visualization engine dynamically recommends suitable charts based on detected data types and analytical objectives. Time-series data are represented using line charts, categorical variables using bar or pie charts, geographical attributes using maps, and correlation analysis through heatmaps and scatter plots. This automatic visualization selection removes the need for manual dashboard configuration.

Step 6: Predictive Analytics

The predictive analytics module applies machine learning algorithms for forecasting and trend analysis. Regression models, time-series forecasting, and feature importance analysis are automatically executed whenever the uploaded dataset contains sufficient predictive information. Generated prediction results are directly incorporated into the analytical dashboard.

Step 7: AI-Based Insight Generation

The platform employs a Large Language Model (LLM) [3], [6], [7] to automatically generate natural-language interpretations of analytical

results. Instead of requiring users to manually interpret charts and statistical outputs, the AI module summarizes significant trends, correlations, anomalies, and forecasting outcomes in an easily understandable form suitable for decision makers.

Step 8: Dashboard Generation and Export

Finally, all generated visualizations, KPIs, statistical summaries, forecasting outputs, and AI-generated insights are integrated into a unified interactive dashboard. Users may further explore the dashboard through conversational analytics, generate downloadable reports, or export analytical results for future use.

When a user selects a specific role, such as a Financial Analyst, the system dynamically adjusts the processing pipeline. For example, a Financial Analyst receives automated forecasting and trend decomposition graphs, whereas a Researcher gets correlation heatmaps, regression plots, and variable importance metrics.

The platform's modular design ensures that each stage from data cleaning to visualization is handled independently, allowing future expansion or replacement of components without reworking the entire system.

Algorithm 1. Adaptive Dashboard Generation

Input: Structured Dataset D

Output: Interactive Dashboard R

1. Read Dataset D
2. Validate Dataset Structure
3. Remove Duplicate Records
4. Detect Missing Values
5. Identify Data Types
6. Perform Semantic Analysis
7. Generate Relevant KPIs
8. Recommend Appropriate Charts
9. Execute Predictive Analysis (if applicable)
10. Generate AI-Based Insights
11. Build Interactive Dashboard
12. Export Reports

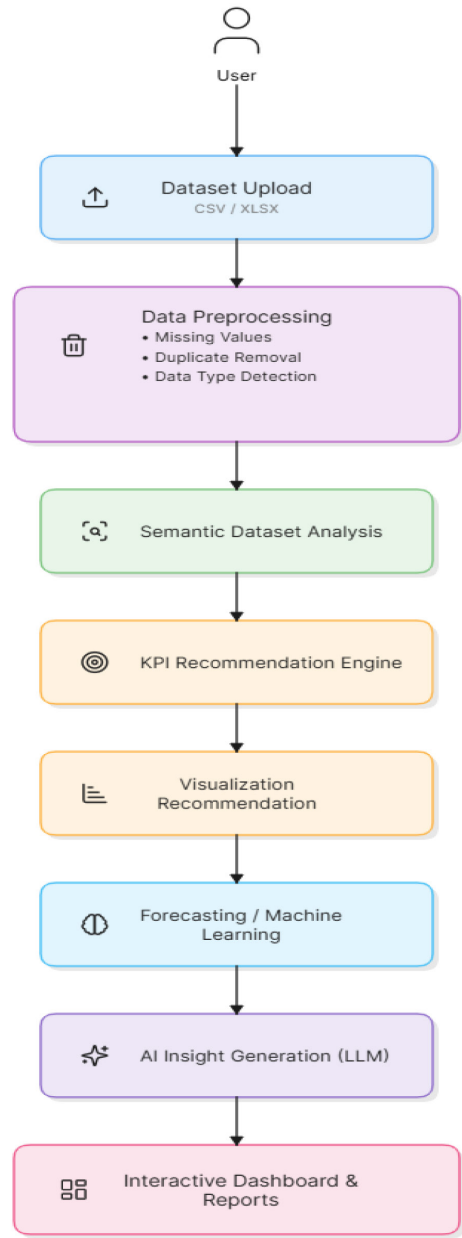


Figure 1
Overall architecture of the proposed DataWhiz framework.

3. Experimental Evaluation

The proposed DataWhiz framework was experimentally evaluated using multiple structured datasets from business, finance, healthcare, marketing, and human resource domains. The primary objective was to assess the effectiveness of the automated preprocessing engine, adaptive dashboard generation, visualization recommendation, predictive analytics, and AI-assisted insight generation modules under different analytical scenarios.

The evaluation focused on five major performance criteria: (i) preprocessing accuracy, (ii) dashboard generation capability, (iii) visualization recommendation quality, (iv) predictive analytics performance, and (v) usability of AI-generated insights. Each uploaded dataset was processed through the complete analytical pipeline without requiring manual dashboard configuration or feature engineering.

To assess the robustness of the proposed framework, datasets containing missing values, duplicate records, categorical variables, numerical attributes, date columns, and multi-sheet Excel files were analyzed. The preprocessing engine successfully handled data cleaning, schema validation, missing value detection, duplicate removal, and automatic data type identification before initiating dashboard generation.

The adaptive visualization engine automatically selected appropriate charts based on dataset characteristics, including bar charts, line graphs, pie charts, scatter plots, correlation heatmaps, histograms, geographical maps, and KPI cards. Predictive models were generated whenever suitable numerical or temporal data were available, while AI-generated summaries explained statistical relationships, forecasting results, anomalies, and important business insights in natural language.

Experimental observations demonstrated that the proposed framework consistently generated complete analytical dashboards with minimal user intervention while maintaining reliable preprocessing, visualization, forecasting, and reporting performance across diverse datasets. The results indicate that the adaptive analytics framework significantly reduces manual analytical effort while improving accessibility for both technical and non-technical users.

Four publicly available structured datasets from different application domains were selected to evaluate the proposed framework, including business sales, finance, human resources, and healthcare datasets. These datasets contain different data types, categorical attributes, temporal

information, and varying record sizes, allowing assessment of the framework under diverse analytical conditions.

Table 1
Comparison of DataWhiz with Existing BI Platforms

Feature	Excel	Tableau	Power BI	DataWhiz
Automatic Data Cleaning	X	Partial	Partial	✓
Automatic KPI Generation	X	X	Partial	✓
Dashboard Generation	Manual	Manual	Manual	Automatic
AI Insight Generation	X	X	Limited	✓
Predictive Analytics	X	Partial	✓	✓
SQL Assistant	X	X	X	✓
Conversational Analytics	X	X	X	✓
Role-Based Analytics	X	X	X	✓
Report Generation	Manual	Manual	Partial	Automatic

Table 1 compares DataWhiz with widely used business intelligence platforms including Microsoft Excel, Tableau, and Power BI. While these platforms provide powerful visualization capabilities, they require users to manually perform data cleaning, KPI selection, dashboard configuration, and interpretation of analytical results. In contrast, DataWhiz integrates automated preprocessing, semantic dataset understanding, adaptive KPI recommendation, visualization selection, predictive analytics, conversational analytics, and AI-generated insight creation within a unified framework. These capabilities reduce manual analytical effort and improve accessibility for users with varying technical expertise.

The proposed DataWhiz framework was evaluated using four publicly available datasets from different application domains. The characteristics of the datasets, generated visualizations, AI insights, and execution status are summarized in Table 2.

Table 2
Dataset Evaluation Summary

Dataset	Domain	Records	Generated Charts	AI Insights	Status
Sales	Business	536,350	27	Yes	Successful
Finance	Finance	51,290	18	Yes	Successful
HR	Human Resources	40,000	11	Yes	Successful
Healthcare	Healthcare	55,500	15	Yes	Successful

Table 3
Runtime Performance Evaluation

Dataset	Records	Processing Time (s)	Data Integrity Score (%)	Duplicate Removal (%)	Dashboard Generation Time (s)
Sales	536,350	4.2	98.6	100	3.8
Finance	51,290	2.3	99.2	98	2.0
HR	40,000	1.9	97.8	97	1.8
Healthcare	55,500	2.5	99.1	100	2.2

Table 3 presents the runtime performance of the proposed framework across four datasets. DataWhiz consistently completed preprocessing and dashboard generation within a few seconds despite differences in dataset size and domain. The results demonstrate that the modular processing pipeline maintains efficient execution while automatically generating relevant visualizations and AI-assisted analytical summaries. These findings indicate that the proposed framework is suitable for rapid exploratory analytics with minimal manual intervention.

4. System Architecture and Implementation

The proposed DataWhiz framework was implemented using a modular client-server architecture. The frontend interface was developed using React for the web platform and Electron for desktop deployment, providing a consistent user experience across online and offline environments. The backend services were implemented using FastAPI (Python), which manages dataset processing, machine learning execution,

API communication, and report generation. Experiments were conducted on a Windows 11 system equipped with an Intel Core i5 processor, 16 GB RAM, Python 3.11, FastAPI, React, Electron, Pandas, NumPy, Plotly, Scikit-learn, Prophet, and Ollama LLM.

The analytical engine integrates Pandas, NumPy, Scikit-learn, Prophet, Plotly, and Matplotlib to perform automated preprocessing, statistical analysis, predictive modeling, and interactive visualization. AI-generated analytical summaries are produced using a Large Language Model (LLM), enabling natural-language interpretation of statistical outputs and forecasting results. The modular architecture allows each analytical component to operate independently, thereby improving scalability, maintainability, and future extensibility of the framework.

The implementation architecture of the developed DataWhiz platform, including the major interface components and workflow, is illustrated in Figure 2.



Figure 2
Architecture of the DataWhiz Framework.

5. AI-Based Insight Generation

The AI-based insight generation module is one of the primary contributions of the proposed DataWhiz framework. After completing data preprocessing, statistical analysis, visualization generation, and predictive modeling, the analytical outputs are automatically summarized

using a Large Language Model (LLM). Instead of requiring users to manually interpret graphs, statistical measures, and forecasting results, the AI module generates context-aware natural language explanations describing trends, correlations, anomalies, feature importance, and predictive outcomes.

The generated summaries are adapted according to the analytical context of the uploaded dataset. For example, business datasets emphasize revenue trends, customer behavior, and sales performance, whereas healthcare datasets focus on disease distribution, patient statistics, and treatment outcomes. Financial datasets generate explanations related to forecasting accuracy, profitability, and investment trends. This adaptive reporting mechanism enables both technical and non-technical users to quickly understand analytical results without requiring advanced statistical expertise.

In addition to automated reporting, the framework supports conversational analytics through an integrated chatbot interface. Users may submit natural-language questions related to the uploaded dataset, request additional visualizations, perform SQL-assisted analysis, and obtain AI-generated explanations without manually writing analytical queries. This functionality significantly improves accessibility and user interaction while reducing the complexity of exploratory data analysis.

6. Discussions

The experimental results demonstrate that the proposed DataWhiz framework effectively automates multiple stages of the analytical workflow, including data preprocessing, dashboard generation, visualization recommendation, predictive analytics, and AI-assisted interpretation. Unlike traditional business intelligence platforms that require extensive manual dashboard configuration, the proposed framework dynamically adapts analytical outputs according to dataset characteristics and user requirements.

The integration of semantic dataset understanding, adaptive visualization selection, forecasting, and conversational analytics enables the framework to reduce manual analytical effort while improving dashboard consistency and interpretability. These findings are consistent with recent AI-driven analytics research reported in healthcare and intelligent marketing applications [7], [8]. The modular architecture further enhances scalability by allowing independent enhancement of

preprocessing, visualization, forecasting, and AI modules without affecting the remaining analytical pipeline.

Although the proposed framework demonstrates promising analytical capabilities, several limitations remain. The current implementation primarily supports structured tabular datasets and relies on external large language models for AI-generated insights.

Compared with conventional dashboard development platforms, the proposed framework automates preprocessing, semantic analysis, KPI recommendation, visualization selection, forecasting, and AI-assisted report generation within a single pipeline. This integrated workflow significantly reduces manual configuration while maintaining dashboard quality across different application domains. Future research may extend the framework to support unstructured datasets, multimodal analytics, federated learning, and domain-specific fine-tuned language models for enterprise-scale decision support.

7. Conclusion and Future Work

DataWhiz demonstrates how adaptive AI-driven analytics can simplify the complete data analysis workflow, from data preprocessing to intelligent dashboard generation and AI-assisted insight creation. The proposed framework reduces manual effort while improving analytical accuracy, visualization quality, and decision support for users with different professional roles. Experimental evaluation indicates that the platform consistently produces reliable dashboards and meaningful insights across multiple business domains. Future work will focus on integrating multimodal datasets, real-time streaming analytics, federated learning, explainable AI (XAI), and domain-specific fine-tuned language models. Additional enhancements will include collaborative multi-user environments, cloud-native deployment, and advanced AutoML optimization, making DataWhiz suitable for enterprise-scale intelligent analytics.

Acknowledgement

The author sincerely thanks Dr. Aniket Patel, ICT Ganpat University, for his continuous guidance, valuable suggestions, and support throughout the development and evaluation of the proposed DataWhiz framework. The author also acknowledges the use of open-source software, publicly available datasets, and research resources that supported the implementation and evaluation of this work.

References

- [1] M. Alkayyal, S. Malberg, and G. Groh, "An LLM-based decision support system for strategic decision-making," in *Machine Learning and Knowledge Discovery in Databases*. Applied Data Science Track and Demo Track—ECML PKDD 2025, Lecture Notes in Computer Science, vol. 16022. Cham, Switzerland: Springer, pp. 460–464 (2025), doi: 10.1007/978-3-032-06129-4_31.
- [2] K. K. L. Wong, L. Deng, Y. Liu, A. Li, and J. Tu, "Design and implementation of a local LLM-driven decision support system for AI-enhanced SCM," *J. Eng. Design*, vol. 37, no. 6 (2026), doi: 10.1080/09544828.2025.2610791.
- [3] C. Lawless, J. Schoeffler, L. Le, K. Rowan, S. Sen, C. St. Hill, J. Suh, and B. Sarrafzadeh, "'I want it that way': Enabling interactive decision support using LLM and constraint programming," *ACM Trans. Interact. Intell. Syst.*, vol. 14, no. 3, Art. no. 22 (2024), doi: 10.1145/3685053.
- [4] C. Abdellaoui, S. Ferré, C. Boussadi, C. Burgun, and A. Delamarre, "Using large language models to automate the comparison and integration of evolving clinical practice guidelines into clinical decision support systems," *npj Digit. Med.*, vol. 9 (2026), doi: 10.1038/s41746-026-01726-4.
- [5] P. N. Patil and P. S. Patil, "A LLM-based agent workflow for financial decision support using financial databases," *Saudi J. Econ. Manage. Res.*, vol. 10, no. 6, pp. 215–220 (2025), doi: 10.36348/sjemr.2025.v10i06.001.
- [6] A. Jain, A. Mulay, D. Verma, A. Pandey, P. Ramu, and A. Garimella, "DECISIVE: Guiding user decisions with optimal preference elicitation from unstructured documents," in *Proc. 64th Annu. Meeting Assoc. Comput. Linguistics (ACL)*, vol. 1, pp. 31774–31786 (2026), doi: 10.18653/v1/2026.acl-long.1465.
- [7] S. M. F. Tauhid, S. Tanweer, Md. T. Nafis, M. A. Ahad, and S. M. F. Malik, "Machine learning-driven data analytics for improved diagnostic accuracy, treatment efficacy, and real-time monitoring in smart health care," *J. Inf. Optim. Sci.*, vol. 46, no. 7, pp. 2291–2317 (2025), doi: 10.47974/JIOS-2031.
- [8] A. Chauhan, S. Kumar, A. Kumar Karn, and S. Sharma, "Role of artificial intelligence in optimizing digital marketing—an empirical study," *J. Inf. Optim. Sci.*, vol. 46, no. 2, pp. 439–451 (2025), doi: 10.47974/JIOS-1926.

Received July, 2026