International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 03 | March - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Insightify: A Web-Based AI-Powered Data Visualization Platform

Aditya Singh[1], Saisaketh Gudla[2], Kaushik Thumbare[3] , Priyanka Kharde[4], Puja Sarage[5]

[1]Department of Data Engineering, Universal College of Engineering
[2]Department of Data Engineering, Universal College of Engineering
[3]Department of Data Engineering, Universal College of Engineering
[4] Department of Data Engineering, Universal College of Engineering
[5] Assistant professor, Universal college of engineering

Abstract - Data visualization and analysis play a crucial role in decision-making across industries. However, many existing tools require technical expertise, making them inaccessible to non-technical users. Insightify is a web-based application designed to bridge this gap by providing an intuitive interface for CSV data analysis with AI-powered querying and insights. This paper presents the design, architecture, and implementation of Insightify, comparing it with existing solutions and highlighting its advantages.

***Key Words*:** Data Visualization, CSV Analysis, AI-Powered Insights, Web-based application, User-friendly experience.

1.INTRODUCTION

With the exponential growth of data, effective data visualization and analysis have become essential. Many businesses and individuals struggle to extract meaningful insights due to the complexity of traditional tools like Tableau and Power BI. While these tools offer extensive functionalities, they often require significant technical expertise, making them inaccessible to non-technical users. This gap has driven the demand for simpler, more intuitive visualization platforms that enable users to explore and analyze data without advanced programming knowledge.

One such attempt to enhance accessibility in data visualization was FeatureVista, a tool designed for software product line analysis [1]. FeatureVista improved the interpretability of feature dependencies and interactions, making complex software configurations more comprehensible. However, the study primarily focused on software engineering applications, leaving a gap in the development of general-purpose AI-driven visualization tools that cater to non-technical users across industries.

Insightify aims to bridge this gap by offering a user-friendly web application that enables users to upload CSV files, generate insightful visualizations, and receive AI-powered explanations. Unlike conventional platforms, Insightify integrates natural language querying, allowing users to interact with their data more intuitively. Additionally, the proposed system enhances accessibility by eliminating the need for advanced technical knowledge while ensuring scalability and security through cloud-based architecture.

Existing tools for data visualization often require coding knowledge, making them inaccessible to a broader audience. This limits the ability of businesses and individuals to derive actionable insights from their data efficiently. Our objectives are:

1. To develop an intuitive web-based application for CSV file visualization..
2. To integrate AI-powered data querying for natural language-based insights.
3. To ensure scalability and security through cloud-based architecture.

2. LITERATURE REVIEW

The study “FeatureVista: Interactive Feature Visualization for Software Product Lines” introduces a visualization tool for software product lines, aiding developers in understanding feature dependencies and interactions. Their study highlighted the tool’s ability to improve decision-making by providing an intuitive way to explore different configurations. However, the research lacked AI-powered automation for insights and did not address user accessibility for non-technical stakeholders[1].

The study “G6: A Web-Based Library for Graph Visualization” developed G6, a JavaScript-based library for scalable graph visualization. The study demonstrated how G6 enables dynamic, interactive graph representations directly in the browser. Despite its effectiveness in handling large-scale data, the research focused primarily on visualization without incorporating advanced analytical capabilities such as AI-driven insights or automated explanations[2].

The study ”Uncertainty-Oriented Ensemble Data Visualization”, explores uncertainty-aware visualization techniques using variable spatial spreading methods . Their approach improved data interpretation by incorporating uncertainty into ensemble visualizations. However, their methodology primarily catered to complex scientific datasets, limiting its usability for general business applications where non-expert users require simpler, more intuitive visualizations[3].

The study ”ConPlot: A Web-Based Application for Protein Contact Map Visualization” introduced ConPlot, a bioinformatics tool for protein structure visualization. The platform effectively integrated contact map analysis with other biological datasets, enhancing researchers' ability to study protein interactions. However, the study did not extend its capabilities to other data analysis domains, leaving a gap in the application of AI-powered insights for broader use cases[4].

The study ”Interactive Visualization of Terascale Data in the Browser” investigated the feasibility of using WebGL for real-time, browser-based visualization of terabyte-scale datasets. Their findings demonstrated that with appropriate optimization, even massive datasets could be interactively explored. However, their study focused primarily on technical scalability, neglecting the need for user-friendly interfaces that allow non-experts to extract meaningful insights from such large datasets[5].

The study “Visualization Tools for Collaborative Systems:

A Systematic Review" conducted a systematic review of visualization tools for collaborative systems, analyzing their applications in healthcare, education, and business. The study highlighted the importance of real-time updates, interactive visualizations, and user-friendly interfaces in facilitating teamwork and decision-making. However, it identified gaps such as the lack of AI-driven insights, limited accessibility for non-technical users, and restricted customization options[6].

3. METHODOLOGY

1. Data Flow & Processing:
    a. Data is imported using pandas.
    b. Filters applied via a web UI.
    c. AI-driven query responses generated by Gemini AI.
2. Visualization Techniques:
    a. Charts generated using Matplotlib & Streamlit.
    b. Automated insights with AI interpretation.
3. Security Measures:
    a. User authentication via Firebase/Auth0.
    b. Encrypted data storage for privacy.

A. System Overview

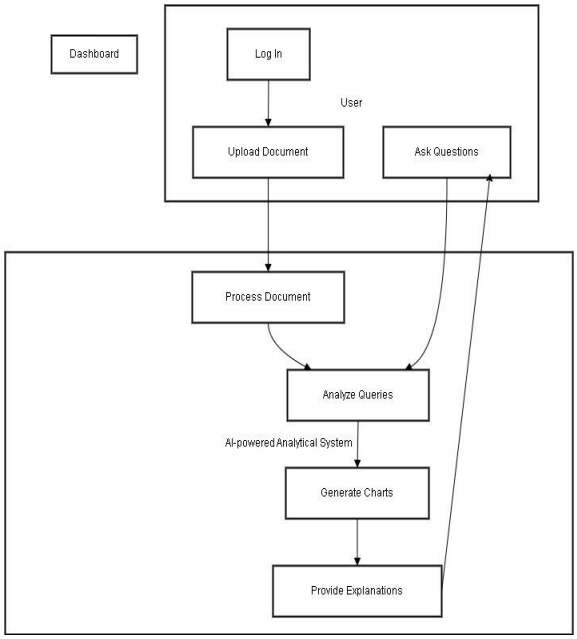

Fig -1: Flowchart of the insightify

Insightify enables users to upload CSV files, filter data, and generate visualizations (bar charts, line graphs, pie charts, etc.). AI-powered features allow users to query data in natural language, receiving intelligent insights. The platform ensures secure user authentication and is built with cloud scalability in mind.

B. System Architecture

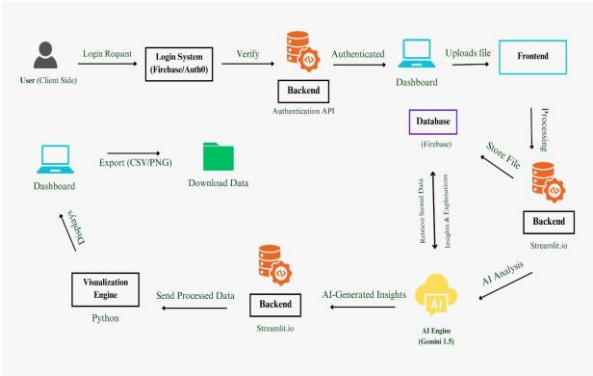

Fig - 2: System architecture of the insightify

1. Frontend: Built using Streamlit.io for interactive user experience.
2. Backend: Developed in Python, handling data processing, authentication, and API requests.
3. Database: Firebase stores user credentials and uploaded data.
4. AI Integration: Google Gemini AI powers data queries and automated explanations.
5. Data Processing: Uses pandas for filtering, cleaning, and visualization.
6. Security & Authentication: Implemented via Firebase/Auth0.

C. Key Functionalities

1. Data Import & Management: Users can drag-and-drop CSV files.
2. AI-Powered Querying: Users can ask questions and receive contextual explanations.
3. Real-time Visualization: Supports bar charts, line graphs, and pie charts.
4. Data Export: Users can save visualizations as PNG/JPG or export filtered data

4. RESULTS

Insightify was tested with multiple datasets of varying sizes. The system efficiently processed and visualized data, with real-time AI responses.

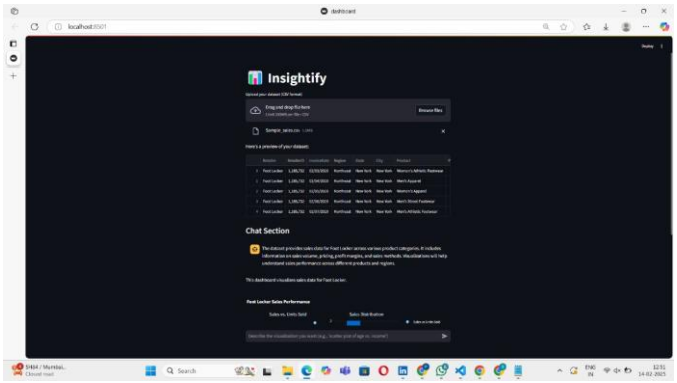

Fig -3: Home page of the insightify

Fig - 4: Pie chart generated for the data given to the insightify

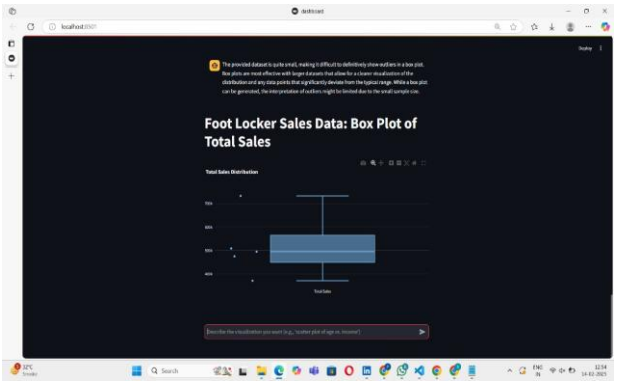

Fig. - 5: Box plot generated for the data given to the insightify

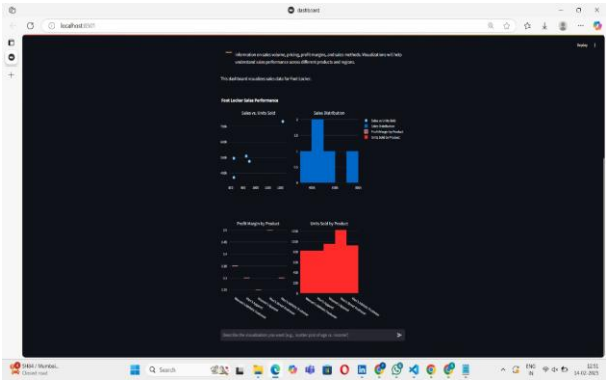

Fig - 6: Comparison chart generated by insightify

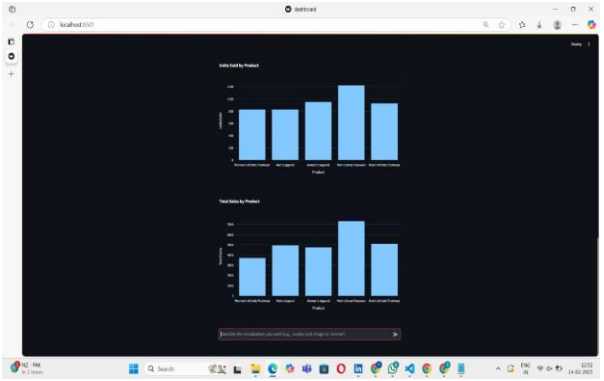

Fig. 7 Bar chart generated for the data given to the insightify

Table - 1:
Comparison of the existing models

Features	Tableau	Google Data Studio	Insightify
Ease of use	Medium	High	Very High
AI Integration	No	No	Yes
Web based	Yes	Yes	Yes
Free Access	No	Yes	Yes

The results indicate that Insightify outperforms existing tools in accessibility and AI integration, making it a superior choice for non-technical users.

5. CONCLUSION

Insightify successfully bridges the gap between data complexity and user accessibility, providing an intuitive AI-powered data visualization platform. Future enhancements include support for time-series analysis, advanced filtering, and additional visualization types. The proposed system sets a new standard for data democratization, ensuring that insights are accessible to all users, regardless of technical expertise.

ACKNOWLEDGEMENT

We take this opportunity to express our deep sense of gratitude to our project guide and project co-ordinator Mrs. Puja Sarage for her continuous guidance and encouragement throughout the project. We would also like to thank Mr. Sandesh Patil, the Head of Department of Data Engineering, for his encouragement and support. We would also like to thank our principal Dr. J.B. Patil and the management of Universal College of Engineering, Vasai, Mumbai for providing us all the facilities and the work friendly environment.

REFERENCES

1. A. Bergel, R. Ghzouli, T. Berger, and M. R. V. Chaudron, "FeatureVista," *Proceedings of the 25th ACM International Systems and Software Product Line Conference - Volume A*, pp. 196–201, Sep. 2021, doi: https://doi/10.1145/3461001.3471154.

2. Y. Wang *et al.*, "G6: A web-based library for graph visualization," *Visual Informatics*, vol. 5, no. 4, pp. 49–55, Dec. 2021, doi: https://doi.org/10.1016/j.visinf.2021.12.003.

3. M. Zhang, L. Chen, Q. Li, X. Yuan, and J. Yong, "Uncertainty-Oriented Ensemble Data Visualization and Exploration using Variable Spatial Spreading," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 2, pp. 1808–1818, Feb. 2021, doi: https://doi.org/10.1109/tvcg.2020.3030377.

4. F. Sánchez Rodríguez *et al.*, "ConPlot: web-based application for the visualization of protein contact maps integrated with other data," *Bioinformatics*, vol. 37, no. 17, pp. 2763–2765, Jan. 2021, doi: https://doi.org/10.1093/bioinformatics/btab049.

5. W. Usher and V. Pascucci, “Interactive Visualization of Terascale Data in the Browser: Fact or Fiction?,” *arXiv.org*, 2020. https://arxiv.org/abs/2009.03254.

6. M. Vásquez-Bermúdez, C. Sanz, M. A. Zangara, and J. Hidalgo, “Visualization Tools for Collaborative Systems: A Systematic Review,” *Communications in Computer and Information Science*, pp. 107–122, 2021, doi: https://doi.org/10.1007/978-3-030-88262-4_8.

7. D. R. Brademan *et al.*, “Argonaut: A Web Platform for Collaborative Multi-omic Data Visualization and Exploration,” *Patterns*, vol. 1, no. 7, p. 100122, Oct. 2020, doi: https://doi.org/10.1016/j.patter.2020.100122.

8. Google, "Google Gemini AI Documentation," *Google Developers*, 2024. [Online]. Available: https://developers.google.com/gemini-ai.