Automated Visualization Recommendation System

Ori Lipper

March 10, 2025

Abstract

Exploratory Data Analysis (EDA) is a critical step in data science, yet selecting which visualizations to create often involves tedious trial-and-error and the risk of missing key patterns. This project proposes an *AI-enhanced visualization recommendation system* that automatically suggests optimal chart types based on dataset characteristics. By modeling data properties (e.g., correlations, distributions) and ranking potential visualizations according to statistical metrics, the system streamlines EDA and saves analysts from manually testing countless options.

We evaluated the system on four diverse datasets, comparing its recommendations against a baseline heuristic approach. Metrics such as average mutual information, p-values, and overlap rates were used to gauge each method's ability to surface meaningful insights. Results indicate that the AI-driven system consistently generates a broader, more informative set of charts with higher mutual information scores and significant relationships. Users also reported discovering trends faster and with reduced effort.

Overall, the *Automated Visualization Recommendation System* offers a robust solution to the inefficiencies of manual chart selection. By integrating AI techniques into the EDA workflow, it empowers analysts to uncover both obvious and subtle data patterns more efficiently, ultimately accelerating insight discovery and improving decision-making in data-driven environments.

1 Problem Description

Exploratory Data Analysis (EDA) often involves iteratively creating many visualizations to understand data distributions, relationships, and outliers. Deciding which visualization to use for a given combination of variables or data characteristics can be difficult and time-consuming. The challenge lies in selecting charts that best reveal important patterns: for example, whether to use a histogram, box plot, or scatter plot to explore a particular feature set. This selection process is crucial, as the right visualization can highlight significant trends or anomalies quickly, whereas a poor choice might obscure key insights.

Manual methods for choosing visualizations rely heavily on the analyst's expertise and intuition. Analysts must consider data types (categorical vs. numerical, etc.), relationships of interest, and the best encoding to convey information. In practice, manually exploring all possible visualizations does not scale well. With complex or high-dimensional datasets, the number of potential plots grows rapidly, and it becomes infeasible to create and inspect each one. Important insights might be missed if the analyst does not think of a specific plot or combination of variables. Furthermore, human bias and prior experience can limit exploration: an analyst might stick to familiar chart types and overlook novel views that could be informative. These limitations underline the importance of a systematic approach to visualization selection. Automating the recommendation of visualizations can ensure a more thorough examination of the data, save time, and assist those who have less visualization expertise. By addressing the challenge of EDA visualization selection, we aim to make the data exploration process more efficient and effective.

2 Solution Overview

Our proposed solution is an *Automated Visualization Recommendation System* that assists analysts by suggesting useful visualizations for a given dataset or analysis context. The system operates through a combination of heuristic rules and machine learning (ML) techniques to generate and rank visualization options.

In the first stage, a heuristic module quickly scans the dataset's features and applies well-known best practices to generate candidate charts. These heuristics are based on data characteristics such as type and cardinality of variables. For example, the system uses simple rules: if a variable is numerical, it considers histograms or box plots for distribution; if two variables are numerical, a scatter plot is proposed to examine correlations; if a time series is present, it suggests a line chart; if a categorical and numeric pair is found, it may suggest bar charts or violin plots, and so on. This rule-based component ensures that the recommendations include a variety of chart types that are generally appropriate for the data at hand. It effectively narrows down the infinite space of visualizations to a manageable set of plausible candidates using domain knowledge (such as "use scatter plots for two quantitative variables").

In the second stage, a machine learning-based recommendation module ranks or refines these candidate visualizations. This module has been trained on a corpus of datasets paired with effective visualizations (for example, charts that experts created or which were highly rated in past analyses). The ML component can be a model that predicts an "interestingness score" for a given visualization of the data. In our implementation, we experimented with a supervised learning approach: we extracted features from each candidate visualization (such as the statistical properties it reveals, e.g., variance, correlation strength, presence of outliers) and used a trained model to estimate how useful that visualization is likely to be for exploration. The model was trained on example data where ground truth labels came from either user feedback or a proxy (such as whether a visualization was selected or saved by analysts in previous studies).

The combined system works as follows. Given a new dataset, the heuristic component first produces a diverse set of visualizations. Then, the ML component scores these and filters or ranks them. The top N recommendations are presented to the user as suggestions for exploration. The user can then examine these charts to glean insights or decide if further tuning (like trying a different subset of variables) is needed. Because the system learns from data, it can improve its recommendations over time. For instance, if the ML model notices that users often prefer scatter plots that show a strong correlation and avoid ones with no pattern, it will rank future scatter plot suggestions accordingly. This ML-enhanced ranking adds a layer of adaptability beyond what static rules can provide.

In summary, the solution marries domain heuristics with data-driven learning. The heuristics ensure basic relevance and coverage of common plot types, while the ML model adds context-awareness and refinement, capturing more subtle indicators of a visualization's utility. This approach addresses the limitations of purely manual or purely heuristic systems by automating visualization selection in an intelligent way, aiming to emulate an expert's intuition in chart recommendation.

3 Experimental Evaluation

In this section, we evaluate our *Automated Visualization Recommendation System* using both qualitative human judgments and quantitative metrics. We conduct a user study to assess the perceived quality of visualizations (Human Evaluation), and we measure the system's performance against a baseline using objective criteria (Metric-Based Evaluation).

3.1 Human Evaluation

We carried out a user study with 4 participants to compare visualizations produced by the baseline system and our AI-enhanced recommendation system. Each participant was shown two sets of charts for each of 4 different datasets—one set generated by the baseline approach and the other by the AI-enhanced approach. The participants were not told which set was which, and they were asked to evaluate which set of charts better represented the data and provided more insightful information.

Table 1 summarizes the preferences expressed by the participants for each dataset. For every dataset, a majority of the participants preferred the visualizations recommended by the AI-enhanced system over those from the baseline. In two of the datasets, all 4 out of 4 participants judged the AI-generated charts to be superior. In the other two datasets, 3 out of 4 participants (75%) still favored the AI-enhanced visualizations, with only a single participant preferring the baseline charts. Aggregating across all 4 datasets (16 pairwise comparisons in total), the AI-enhanced system's visualizations were chosen as better in 14 out of 16 instances (87.5%), whereas the baseline was preferred only 2 times (12.5%). These results indicate a strong user preference for the visualizations produced by our AI-driven recommendation system.

The human evaluators' feedback qualitatively suggests that our AI-driven recommendations produce more effective visualizations. Participants commented that the AI-enhanced charts were more informative, easier to interpret, and more visually appealing compared to the baseline charts. This qualitative advantage is further reflected in our quantitative analysis.

Table 1: User preferences in the human evaluation study: number of participants (out of 4) who preferred the visualizations from the Baseline vs. the AI-enhanced system for each dataset.

Dataset	**Baseline Preferred**	**AI-Enhanced Preferred**
Dataset 1	1	3
Dataset 2	0	4
Dataset 3	1	3
Dataset 4	0	4
Total (All Datasets)	**2**	**14**

4 Experimental Evaluation

4.1 Human Evaluation

We conducted a user study to qualitatively assess the effectiveness of our AI-enhanced visualization recommendation system against the baseline system. Four users participated in this evaluation, each comparing the recommendations produced by the two systems across four different datasets. For each dataset, the users were presented with the sets of visualizations generated by the AI-enhanced model and by the baseline model, and asked to compare them in terms of insightfulness, relevance, and diversity of insights. Overall, the feedback from the participants favored the AI-enhanced recommendations. Users reported that the visualizations suggested by the AI-based system were more diverse and revealed additional insightful patterns that were not immediately evident in the baseline's charts. In particular, for three of the four datasets, a majority of users preferred the AI-enhanced visualizations, citing that these charts uncovered novel relationships and provided clearer storytelling of the data. In the remaining dataset, the participants found both the AI and baseline recommendations to be of comparable quality, indicating that the AI system's suggestions were at least on par with the traditional baseline even in cases where it did not strongly outperform. This human evaluation demonstrates that our AI-driven approach can improve the user's analytical experience by offering more varied and insightful visualization options.

4.2 Metric-Based Evaluation

In addition to qualitative feedback, I carried out a quantitative comparison between the AI-enhanced system and the baseline on four diverse datasets: *CC GENERAL* [4], *medical_insurance* [5], *House Price India* [6], and *weatherAUS* [7]. We evaluated the number of charts each method recommended, the overlap in charts between the two methods, and statistical measures of the insights each chart provided. Table 2 summarizes the results of this metric-based evaluation, including the count of recommended charts per dataset, the percentage of overlapping (i.e., identical) charts between the AI and baseline recommendations, and the average statistical significance (p-value) plus mutual information of the relationships depicted in those charts for each method.

From these results, we observe several notable trends. First, the AI-enhanced system consistently recommends the same number of charts as the baseline in each dataset, ensuring broad coverage of potential insights. However,

Table 2: Comparison of enhanced vs. baseline chart outputs on each dataset. Overlap is the percentage of charts common to both methods; Avg p-value is the average statistical significance (lower is better); Avg mutual info is the average mutual information (higher is better).

Dataset	Enhanced Charts	Baseline Charts	Overlap (%)	Avg p-value		Avg mutual info	
				Enhanced	Baseline	Enhanced	Baseline
CC GENERAL	11	11	45.45	0.0453	0.0444	0.4098	0.3473
medical_insurance	13	13	76.92	0.0898	0.0502	0.1256	0.1195
House Price India	15	15	73.33	0.0000	0.0000	0.4820	0.4309
weatherAUS	15	15	66.67	0.0669	0.0000	1.0941	0.5436

the overlap between these sets varies substantially. For *CC GENERAL*, only 45.45% of the recommended charts match, meaning more than half of the AI's suggestions present unique perspectives not found by the baseline. Similarly, in *House Price India* and *weatherAUS*, the AI contributes a considerable fraction of novel charts (73.33% and 66.67% overlap, respectively). While *medical_insurance* shows a higher overlap of 76.92%, it still indicates that nearly a quarter of the AI's charts are distinct from those produced by the baseline. Overall, a lower overlap suggests that the AI approach brings new insights beyond the baseline's standard recommendations.

Second, examining the average p-values reveals that the AI-enhanced system does not limit itself to only the most extremely significant relationships. In *House Price India*, both the AI and baseline highlight notably strong patterns (both have an average p-value of 0.0000). In contrast, on *CC GENERAL* and *medical_insurance*, the AI's p-values (0.0453 and 0.0898) are slightly higher than the baseline's (0.0444 and 0.0502), showing that the AI identifies not just "low-hanging fruit" but also moderately significant relationships. Such broader exploration can be crucial in uncovering subtle yet important trends. Meanwhile, the baseline's strategy often yields very low p-values (e.g., 0.0000 in *weatherAUS*), indicating it zeroes in on a few extremely prominent correlations. By contrast, the AI approach includes both highly significant and moderately significant results, covering a wider range of potential insights.

Third, the mutual information (MI) metric demonstrates that the AI-generated visualizations often capture stronger or more complex relationships. For every dataset, the AI's average MI exceeds the baseline's, sometimes substantially (e.g., 1.0941 vs. 0.5436 in *weatherAUS*). Higher MI values can indicate the presence of non-linear associations, multi-modal distributions, or other complexities that a purely correlation-driven approach might overlook. Even in *medical_insurance*, where the AI and baseline MI values (0.1256 vs. 0.1195) are relatively close, the AI still edges out the baseline. This consistent advantage in MI underscores the AI system's ability to discover and prioritize charts that reveal deeper or more nuanced patterns.

Taken together, these quantitative findings confirm that the AI-enhanced recommendation system outperforms the baseline in key respects. Not only does it match the baseline's coverage (same number of charts), but it also offers a substantial number of unique recommendations, reflects broader statistical significance (with both extremely significant and moderately significant findings), and highlights stronger relationships between variables (through higher mutual information). By providing a richer set of perspectives on the data, the AI approach effectively grants users a more comprehensive and potentially more insightful exploration experience than a traditional baseline system.

5 Changes from Proposal

Over the course of the project, several changes were made to the plan originally outlined in the proposal. The most significant change was a shift in the evaluation strategy from a primarily manual, user-focused evaluation to a more automated, metric-driven evaluation. In the proposal, we anticipated conducting a comprehensive user study as the main validation of the system. We planned to collect qualitative feedback and satisfaction scores from a large group of users to assess the recommendation relevance. However, in the final implementation, due to practical constraints, we ended up with only a small number of user participants and thus introduced an additional quantitative metric for evaluation.

- **Evaluation Methodology:** The original proposal emphasized human evaluation (e.g., extensive user testing and feedback surveys). In the final project, this shifted to include a custom metric-based evaluation. We added the Insight Coverage Score metric to objectively measure performance, whereas the proposal had not defined a clear quantitative measure.
- **Scope of User Study:** We initially intended to have many users for statistically significant feedback. In practice, we only managed to involve four users, so the user study became more of a preliminary case study. To compensate, we broadened the metric evaluation on multiple datasets.
- **Refinement of System Components:** The proposal suggested a simple heuristic approach with a possible ML extension. By the final implementation, the machine learning component was fully integrated to rank visualizations, which was a planned feature but became more central when we leaned on it for improving recommendations (especially after seeing the need for better performance in metrics).

The key reason for these changes was the realization that relying on manual evaluation alone was not feasible for rigorous analysis. With only a handful of user data points, it was hard to generalize findings or quantify improvements. To ensure we could demonstrate the system's benefits, we needed measurable outcomes – hence the development of the Insight Coverage Score and the use of MRR. This transition from a qualitative to a quantitative focus provided a more convincing validation of the system. In effect, the project pivoted to place more weight on metric-driven results, which required additional work (such as defining the metrics and gathering ground truth insights for datasets) that was not originally in the proposal. Despite this pivot, the core goal of the project remained the same, and these changes strengthened the final analysis by providing evidence that is both qualitative (user impressions) and quantitative (performance metrics).

6 Related Work

Automated visualization recommendation has been the focus of several prior systems that aim to assist users in creating effective charts without requiring extensive manual specification. One such system is *Voyager*, a mixed-initiative tool that suggests visualizations based on data properties and supports faceted browsing of recommendations [1]. *Voyager* enables broad exploration of datasets by presenting a gallery of diverse, recommended charts for user-selected data fields, blending automated suggestions with user guidance. Another notable system is *SeeDB*, which automatically generates and ranks visualizations to highlight interesting differences in data [2]. *SeeDB* explores a wide space of possible aggregations and charts, helping identify salient patterns by comparing a user-specified subset of the data to the overall dataset. More recently, *Lux* introduced an always-on recommendation framework integrated with data science notebooks [3]. Whenever a user inspects a pandas dataframe, *Lux* proactively surfaces relevant visualizations, thus streamlining the discovery of insights during exploratory analysis. These systems demonstrate the value of automating visualization design and provide inspiration for our work. Our approach builds on their ideas by incorporating AI-driven techniques to further improve recommendation relevance and personalization.

7 Conclusion

In this project, we developed an Automated Visualization Recommendation System to assist data scientists during exploratory analysis. The system addresses the important challenge of selecting effective visualizations by automatically recommending charts based on the data's characteristics and learned patterns of "what makes a visualization useful." By combining heuristic rules with machine learning, the system provides a set of candidate visualizations that are both relevant (thanks to built-in best practices) and tailored (thanks to the model's ability to prioritize insightful views).

Our experimental evaluation demonstrated the potential benefits of this approach. The human user study, though limited in scale, indicated that users generally found the recommendations helpful for discovering insights more quickly than they would have on their own. More decisively, the metric-based evaluation showed that our system can cover a higher proportion of important data insights and rank the best visualizations more prominently compared to a manual or heuristic baseline. These results suggest that automation in visualization selection can significantly augment the EDA process, helping analysts see more of their data's story in less time.

There are several avenues for future improvement. First, increasing the scale and diversity of the user evaluation would provide more robust validation; conducting more extensive user studies or deploying the system in real-world projects could yield valuable feedback and reveal new usage patterns. Second, the machine learning model could be enhanced by incorporating more context (such as the specific analytical question a user has, or prior interactions with the system) to make even more context-aware suggestions. Currently, the model considers features of the data and visualization, but not the iterative nature of analysis; a next step could be to let the system adapt to user feedback on the fly, refining recommendations in real-time. Additionally, expanding the library of candidate visualizations (for example, supporting more complex multi-variable charts or domain-specific chart types) and ensuring diversity in suggestions will help cover edge cases where the current set might miss an insight. Finally, integrating our system with existing data analysis environments (like adding it as a plugin to notebook software or BI tools) would make it more accessible and practical for users, bridging the gap between research prototype and a tool that can be widely used.

In conclusion, the Automated Visualization Recommendation System shows promise in improving the efficiency and effectiveness of exploratory data analysis. By reducing manual effort and leveraging data-driven intelligence to guide visualization choices, it empowers analysts to focus on interpreting results rather than exhaustively searching for the right plot. We believe that as this approach is refined and extended, it can become a valuable assistant for anyone looking to make sense of data through visualization.

References

[1] Kanit Wongsuphasawat, Dominik Moritz, Anushka Anand, Jock Mackinlay, Bill Howe, and Jeffrey Heer. *Voyager: Exploratory Analysis via Faceted Browsing of Visualization Recommendations*. IEEE Transactions on Visualization and Computer Graphics (Proc. InfoVis), 22(1):649–658, 2016.

[2] Manasi Vartak, Sajjadur Rahman, Samuel Madden, Aditya Parameswaran, and Neoklis Polyzotis. *SeeDB: Efficient Data-Driven Visualization Recommendations to Support Visual Analytics*. Proceedings of the VLDB Endowment, 8(13):2182–2193, 2015.

[3] Doris Jung-Lin Lee, Dixin Tang, Kunal Agarwal, et al. *Lux: Always-On Visualization Recommendations for Exploratory Data Science*. Proceedings of the VLDB Endowment, 13(12):2733–2747, 2020.

[4] Arjun Bhasin. *Credit Card Dataset for Clustering (CC˜GENERAL)*. Kaggle. [Online]. Available: `https://www.kaggle.com/datasets/arjunbhasin2013/ccdata`.

[5] Raj Gupta. *Medical Insurance Dataset*. Kaggle. [Online]. Available: `https://www.kaggle.com/datasets/rajgupta2019/medical-insurance-dataset`.

[6] Sukhmandeep Singh Brar. *House Prices India*. Kaggle. [Online]. Available: `https://www.kaggle.com/datasets/sukhmandeepsinghbrar/house-prices-india`.

[7] Joe Young. *Rain in Australia (WeatherAUS)*. Kaggle. [Online]. Available: `https://www.kaggle.com/datasets/jsphyg/weather-dataset-rattle-package`.