data

MDPI

Article

GroupView: A Visual Framework for Exploring Group Membership Dynamics over Time

Mithilesh Kumar Singh and Klaus Mueller *

Department of Computer Science, Stony Brook University, Stony Brook, NY 11794, USA; mkssingh@cs.stonybrook.edu
* Correspondence: mueller@cs.stonybrook.edu

Abstract

Tracking group membership dynamics over time is a persistent challenge in visual analytics, particularly when dealing with complex, multidimensional datasets. Existing tools often struggle to visualize dynamic group transitions while preserving attribute relationships and maintaining consistent group definitions. We present GroupView, a visual framework designed to explore temporal data and group dynamics to address this. GroupView enables users to slice data into time-based segments and create dynamic groupings, facilitating the identification of trends and patterns that may otherwise remain hidden. Its features include automated grouping based on data similarities, combinatorial grouping for richer insights, and custom grouping for tailored analysis. A heuristic user study involving visualization experts provided feedback on usability and analytical value, highlighting the strengths of GroupView in intuitive exploration and insight discovery. These features position GroupView as a valuable tool for analysts and researchers working with evolving datasets, offering new avenues for uncovering trends and tracking group-level changes over time.

Keywords: interactive data visualization; data slicing; group membership; multidimensional data analysis; data exploration tool

check for updates

Academic Editor: Jamal Jokar Arsanjani

Received: 26 April 2025
Revised: 9 August 2025
Accepted: 11 August 2025
Published: 21 August 2025

Citation: Singh, M.K.; Mueller, K. GroupView: A Visual Framework for Exploring Group Membership Dynamics over Time. *Data* **2025**, *10*, 133. https://doi.org/10.3390/data10080133

1. Introduction

In domains such as socioeconomics, marketing, public health, and environmental monitoring, understanding how groups of individuals or entities evolve is critical to uncover meaningful insights and make informed decisions. For example, socioeconomic studies often track population movements between income brackets to assess the effectiveness of poverty reduction policies [1–3]. In marketing, businesses monitor shifts in consumer behavior, such as transitions between product categories or customer loyalty tiers, to refine their strategies [4,5]. Similarly, public health experts analyze changes in risk groups to improve interventions for diseases and other health outcomes [6–8], while environmental researchers examine species' transitions between threat levels due to climate change or pollution [9,10]. These use cases highlight the growing importance of tools that can effectively visualize group membership dynamics over time, capturing high-level trends and granular details about individual contributions [11,12].

Despite the importance of tracking group dynamics, visualizing these transitions presents significant challenges. Traditional tools, such as Sankey diagrams, Alluvial diagrams, and flow-based techniques, are widely used to depict transitions between categories, but they often abstract away critical attribute-level information and interactions [13–15].

This abstraction can obscure valuable insights, particularly in multidimensional datasets where individual attributes play a pivotal role in group behavior [16]. Existing approaches also often lack flexibility in defining and analyzing groups over time. Automated methods like k-means clustering focus on statistical groupings but struggle to capture dynamic changes in group definitions, making them less suitable for datasets where context plays a critical role [17–20]. While visualization tools such as C-Group [21] enable the exploration of dynamic group memberships in social networks, they tend to be tailored to specific dataset structures and do not generalize well to broader multidimensional data scenarios. These limitations highlight the need for more adaptive tools that not only track group membership over time but also preserve contextual relationships and attribute-level variations critical for robust analysis.

To address these challenges, we introduce GroupView, a visual framework designed to explore temporal data and group dynamics. GroupView enables users to slice datasets into time-based segments and create dynamic groupings, facilitating the identification of trends and patterns that may otherwise remain hidden. Unlike traditional tools, GroupView introduces three distinct grouping approaches: automated grouping for data-driven insights (e.g., k-means or k-modes clustering), combinatorial grouping for exploring feature interactions, and custom grouping for domain-specific analyses. By tracking old and new group memberships across these different grouping strategies, GroupView ensures consistent yet flexible group definitions. Leveraging principles from explainable clustering of k-means and k-medians [22], GroupView visually disaggregates each group to highlight the contribution of individual factors, making the grouping process transparent and intuitive for users. The framework tracks the initial group assignments, attracting new members to fall into the group and repelling those who drop out. This adaptability allows analysts to understand evolving trends and patterns better.

GroupView builds on existing visualization and grouping techniques, offering additional flexibility and functionality. For automatic grouping, GroupView leverages standard clustering algorithms, specifically, k-means for continuous data, k-modes for categorical data, and k-prototypes for mixed data types to generate robust data-driven groupings. In addition, the framework supports combinatorial and custom grouping options, allowing users to refine or redefine groupings based on domain-specific criteria and analytical goals. Importantly, GroupView explicitly visualizes the transitions and interactions between individual attribute categories or numerical brackets and group membership dynamics, bridging the gap between high-level patterns and detailed data-level insights. By adhering to the declarative principles of the Grammar of Graphics [23], our interface systematically maps the attributes of the data to visual elements such as color, size, and layout, ensuring complex multidimensional data are presented clearly and structurally. Furthermore, the framework supports dynamic time-axis-based data slicing, enabling exploratory analysis across multiple dimensions. This empowers users to track how group memberships evolve, uncovering meaningful changes driven by individual or group-level factors.

The remainder of this paper is structured as follows: In Section 2, we review related work on dynamic visualizations of group transitions, highlighting the gaps addressed by GroupView. Section 3 introduces the design and functionality of the GroupView framework, detailing its flexible grouping methods, time-slicing capabilities, and interactive visualization features. Section 4 presents various use cases to demonstrate the practical applications of GroupView in analyzing dynamic group transitions. Section 5 provides insights derived from these use cases, emphasizing the analytical depth and trends uncovered by the framework. Section 6 presents the heuristic evaluation results, illustrating the framework's usability and effectiveness through participant feedback. In Section 8, we discuss the implications of GroupView, its contributions, limitations, and potential for future advancements. Finally, Section 9 concludes the paper by summarizing key findings and outlining directions for future research.

2. Related Work

Effectively tracking dynamic group memberships and understanding the underlying factors driving transitions is a persistent challenge in visual analytics. Existing methods can be broadly categorized into flow-based visualization techniques, network-based approaches, and automated clustering methods. Each offers valuable insights but suffers from significant limitations when applied to evolving multidimensional datasets.

Flow-based visualization methods, such as Sankey diagrams [13], alluvial diagrams [14], and flow maps [24], are widely used to illustrate group transitions over time or across categories. Sankey and alluvial diagrams effectively convey high-level movement patterns but abstract individual contributions, making it difficult to see which attributes drive group dynamics. For example, Rosvall and Bergstrom [14] demonstrate the utility of alluvial diagrams for large networks but note limitations for fine-grained attribute-level changes. Guo [24] addresses scalability in flow maps, yet such approaches can struggle to preserve detailed group definitions over time. Complementing these critiques, a broader review [25] observes that many flow-based methods prioritize aesthetics over the clarity needed to explore multidimensional datasets.

While recent work has introduced enhancements to traditional flow-based visualizations, such as hybrid Sankey diagrams [26], improved alluvial diagrams [27], and hierarchical Sankey structures [28], these approaches still face limitations, such as the aggregation of flows that obscure attribute-level categories or numerical brackets, difficulty in clearly representing continuous and numeric attribute brackets within evolving groups, and limited interactivity for dynamically exploring subgroup changes over time.

Network-based visualization techniques, such as *Force-directed Graphs*, *Network Diagrams*, and *Hive Plots* [11], excel in revealing clusters and static group relationships but falter when applied to *dynamic group memberships*. For example, Shi et al. [21] introduced C-Group, which visualizes group dynamics in social networks, but it is tailored to specific datasets and lacks flexibility for generalizing to multidimensional scenarios. Although network-based methods can provide insight into static structures, they fail to account for temporal changes or interactions at the attribute category or numerical bracket level. As noted by recent works on multidimensional visualization techniques [29], a growing demand exists for approaches combining high-dimensional analysis with dynamic temporal data.

Automated clustering techniques, such as k-means [17], k-modes [19], and k-prototypes [30], are commonly used for statistical group identification. Although these methods effectively generate initial groupings, they lack flexibility for domain-specific analyses and dynamic datasets. MacQueen [17] and Huang [19] noted that clustering methods are inherently static and do not adapt group definitions as data change over time. Recent work on integrating machine learning and visualization [31] highlights the importance of leveraging contextual factors and clustering algorithms to refine group dynamics. On the other hand, Andrienko et al. [20] explored interactive clustering methods, emphasizing the role of user-driven refinement in handling large-scale trajectory data.

Efforts to address these limitations have led to methods incorporating temporal and multidimensional perspectives. Gotz and Sun [6] introduced an interactive system for analyzing patient cohorts and tracking group transitions over time to support clinical decision-making. Similarly, Shin et al. [7] applied clustering methods to visualize temporal patterns in patient data, providing insight into group dynamics in healthcare. Complementing these domain-specific applications, recent studies [32] underscore the need to balance aesthetics and clarity in the design of visual frameworks. Furthermore, work on semi-supervised learning for multidimensional visualization [33] demonstrates the potential to combine machine learning techniques with temporal analysis to uncover evolving patterns.

While flow-based, network-based, and automated clustering techniques each offer unique strengths, they also exhibit critical shortcomings. Flow-based methods excel at depicting high-level patterns but fail to capture individual contributions or maintain consistent group definitions. Network-based approaches effectively visualize static structures but lack the temporal flexibility needed for dynamic analyses. Automated clustering methods provide initial groupings but struggle to incorporate contextual factors or track group changes over time. Although recent multidimensional visualization and machine learning integration efforts offer promising directions [29], significant gaps remain in developing tools that dynamically track and adapt group memberships across diverse datasets.

3. GroupView: Design and Functionality

In this section, we detail the design and functionality of GroupView, a visual framework developed to support the dynamic analysis of group membership in complex datasets. GroupView is built around the principles of interactivity and flexibility, integrating traditional clustering techniques (e.g., k-means and k-modes) with dynamic time-axis-based slicing and version-based slicing. These features work together to maintain consistent yet adaptable group definitions as data evolves.

Our design emphasizes a user-centric workflow, guiding users from dataset loading and configuration to interactive exploration. Key components include (1) a configuration panel to set analysis parameters, (2) GroupView controls to enable data slicing and grouping, and (3) a Visual Explorer that incorporates outcome plots and clearly illustrates group composition through attribute-level categories or numerical brackets. Additional interactive features, such as group sorting and side-by-side comparisons, further support users in identifying high-level trends and detailed patterns in dynamic group behavior.

3.1. GroupView Configuration

The configuration phase in GroupView allows users to prepare their datasets for analysis through two main options: selecting an existing dataset or uploading a new dataset. Users can choose from a library of predefined use cases (e.g., poverty or diabetes datasets), as shown in Figure 1, or upload their datasets for more tailored analyses.

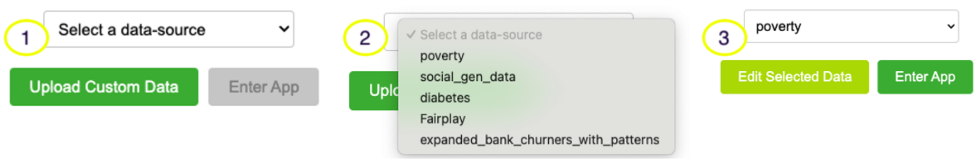

Figure 1. This figure demonstrates the three-step process for selecting a pre-existing use case dataset. (**Step 1**) Click the drop-down menu to view the available data sources. (**Step 2**) Choose a dataset, such as "poverty" or "diabetes," from the list. (**Step 3**) Click "Edit Selected Data" to reconfigure the dataset or click "Enter App" to proceed to the dashboard for analysis.

Users can edit metadata of pre-existing datasets to adjust parameters such as target variables or grouping features, aligning the dataset closely with their analytical objectives. Alternatively, users can upload new datasets by clicking the Upload Custom Data button. On the initial configuration screen (Figure 2), users upload datasets in CSV format, a widely accepted format for data visualization, and configure dataset fields for analysis, specifying roles for columns such as index, version, time, target, and identifying unused columns.

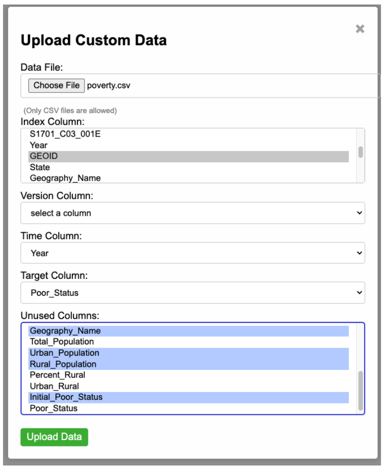

Figure 2. This figure demonstrates uploading and configuring a new dataset in GroupView. Users upload a CSV file and configure the dataset by selecting the appropriate columns for analysis. In this example, the Index Column is set to GEOID to identify each data entry uniquely, the Time Column is set to Year to track data over time, and the Target Column is set to Poor Status to represent the binary outcome (e.g., poor or not poor). The Version Column is left empty, meaning no version-based slicing will be applied. Additionally, several columns, such as Geography Name, Urban Population, Rural Population, and Percent Rural, are selected as Unused Columns, meaning they will be excluded from the analysis to simplify the dataset and focus only on the relevant variables.

Each configuration field plays a critical role in preparing the dataset:

- Index Column: Uniquely identifies each data entry, enabling GroupView to track changes over time or across versions. For example, the GEOID could serve as an index in a poverty tracking study, enabling users to track the poverty status over time. If left blank, GroupView automatically assigns a unique index.
- Version Column: Slices the dataset into segments based on different iterations or categories, such as regions, instances, variants, or periods. For example, in a customer retention study, the country column could serve as the version, allowing the analysis of each country separately. If no Version Column is selected, the dataset is analyzed as a single, unified version without segmentation.
- Time Column: Defines the progression axis, enabling users to analyze how groups evolve across sequential points, such as specific dates, years, quarters, iterations, or numeric indices (e.g., rounds or checkpoints). Depending on the dataset, the slider reflects granular time intervals or numeric progression steps, facilitating the analysis of trends, seasonal patterns, or progressive changes over the dataset's timeline.
- Target Column: Tracks binary outcomes (e.g., 0/1 for 'not poor'/'poor'), allowing for visualization of how such outcomes evolve within or across groups over time. The system currently supports binary targets to enable interpretable outcome transitions (e.g., 0→1, 1→0).
- Unused Columns: Allows for the exclusion of irrelevant columns, enabling users to focus exclusively on variables essential to their analytical goals.

Note on Target Variable: GroupView currently supports binary target variables to simplify and clarify outcome transition analysis. Binary outcomes allow us to define four intuitive transitions: 0→0, 0→1, 1→1, and 1→0. This is particularly useful for studying meaningful temporal changes, such as disease onset or job status. While multiclass targets are not directly supported in this version, users can analyze them by recoding into a binary format (e.g., one-vs-all) when appropriate.

In general, the configuration process allows users to tailor their datasets easily. Users specify key parameters, such as index, version, time, and target columns, by selecting a predefined use case or uploading a custom CSV file to structure their data appropriately. This streamlined setup lays the foundation for actively exploring dynamic group behaviors, enabling focused and actionable analyses of trends and patterns.

3.2. GroupView Interactive Dashboard

The GroupView Interactive Dashboard empowers users to explore and analyze complex datasets by seamlessly integrating two core components: GroupView Controls and GroupView Visual Explorer. Together, these components provide a dynamic and flexible environment for multidimensional data exploration, enabling users to track group dynamics and visualize evolving patterns.

3.2.1. GroupView Controls

The GroupView Controls offer interactive data segmentation and analysis tools, allowing users to define grouping strategies and track changes over time. As shown in Figure 3, Section (1), users can select data versions, apply grouping methods, aggregate outcomes, and adjust the time-axis control for detailed exploration.

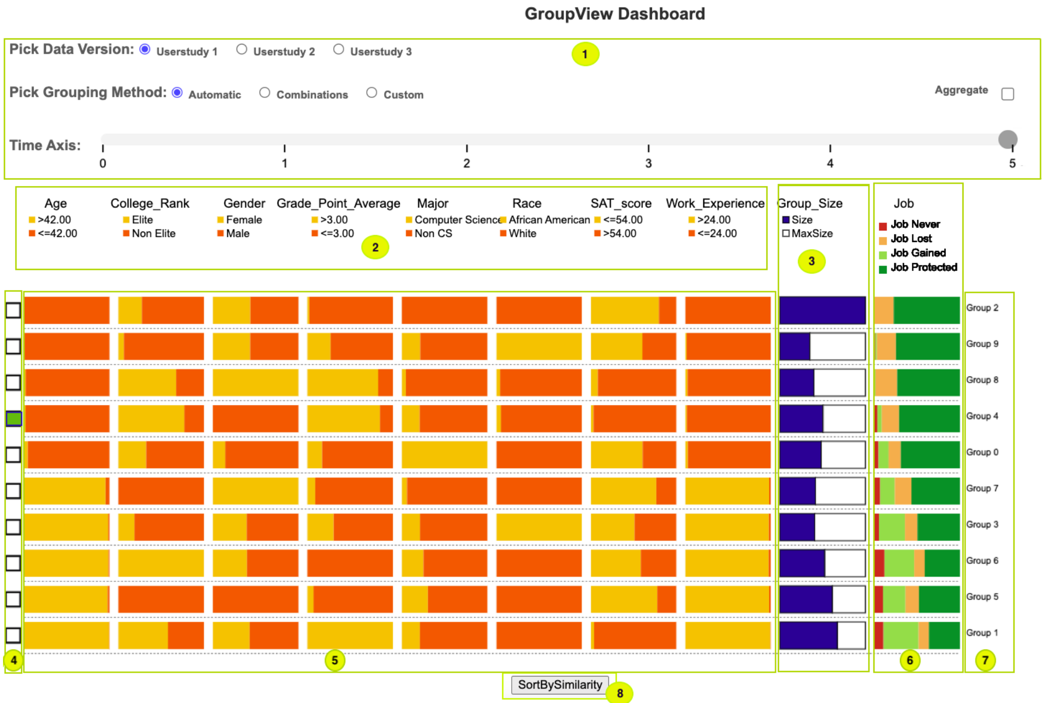

Figure 3. This figure illustrates the key features of the GroupView Interactive Dashboard using the FairPlay [34] dataset. The dashboard includes (1) GroupView Controls for selecting data versions, grouping methods, and Time Axis values; (2) Attribute Legends displaying feature categories (e.g., Age, Gender); (3) Group Size indicators; (4) a Group Selector for individual groups; (5) Feature Composition Rectangles showing each group's attribute distribution; (6) Target Outcome Legends and Plots visualizing outcomes such as Job Lost or Job Gained; (7) Group Labels; and (8) SortBySimilarity for group comparison. Outcomes reflect the five-round decision-making process from the FairPlay dataset, where round 0 represents the initial state and round 5 represents the final consensus; outcome changes (e.g., Job Lost) are measured relative to round 0. Groups are sorted by age, with the younger age groups on top.

The Pick Data Version control allows users to select a Version Column. Unique values from the Version Column, defined during the configuration phase, appear as radio buttons for easy selection. This functionality ensures clarity in the analysis by isolating distinct subsets of data, which is particularly useful for uncovering trends in the chosen data version.

The Pick Grouping Method control supports three options customized to different analytical goals:

- Automatic Grouping: This method involves clustering data points based on the similarity of features using algorithms such as k-means, k-modes, or k-prototypes, optimizing within-group homogeneity. k-means is used when the features are continuous, whereas k-modes is used for categorical data. In cases where the dataset comprises a mix of data types, a hybrid approach (e.g., k-prototypes) is used to handle both continuous and categorical features appropriately.
- Combination Grouping: This method automatically generates groupings by considering every possible combination of categories or numerical brackets across the selected features. For instance, if one feature has two categories and another has three, the algorithm will produce 2 × 3 = 6 distinct groups. This comprehensive approach enables a detailed exploration of how different attribute intersections influence group behavior, facilitating granular comparisons.
- Custom Grouping: This option empowers users to manually define groups based on specific analytical criteria tailored to their objectives. For example, a user might create groups such as Group 1, represented by gender: male, income bracket: $\leq 50K$, and Group 2, defined by gender: female, income bracket: $\geq 25K$. Users can focus their analysis on patterns most relevant to their research goals by explicitly specifying attribute categories or numerical brackets based on domain expertise.

Group transitions over time are tracked dynamically across all grouping methods. At the first time-axis value, groups are formed using one of the selected strategies: (*i*) automatic grouping, where group membership is determined by clustering algorithms applied to selected attributes; (*ii*) combination grouping, where each group corresponds to a unique combination of selected attribute values; and (*iii*) custom grouping, where users manually define groups using attribute filters.

These initial group definitions refer to the membership logic established at the beginning of the analysis. In combination and custom grouping, these definitions are explicit and interpretable. In automatic grouping, group definitions are derived from clustering assignments based on the initial centroids. At subsequent time-axis values, each member is re-evaluated by predicting its cluster assignment using the original clustering model. If the predicted cluster differs from the original, the member is considered to have transitioned. This approach avoids re-clustering and ensures consistency by using the initial clustering model as a fixed reference for group boundaries.

By maintaining consistent group definitions from the beginning, GroupView enables the analysis of how group memberships evolve over time, revealing patterns of stability, transition, and drift in dynamic datasets.

The Outcome Aggregation control allows users to summarize target outcomes across the selected time-axis values. Upon selecting the Aggregate checkbox, two precise aggregation methods become available:

- Initial Checkpoint: Compares outcomes at each selected time-axis value to the initial time-axis value, highlighting cumulative changes relative to the dataset's initial state.
- Previous Checkpoint: Compares outcomes between consecutive time-axis values, from the initial value up to the selected value, emphasizing incremental changes as the dataset progresses.

Users can select the Aggregate checkbox to examine trends up to the chosen time-axis value, suitable for long-term and short-term analyses. Combined with the interactive Time-Axis Control, this enables real-time exploration of changes through a slider.

The Time-Axis Control (illustrated in Figure 3, Section (1)) is an interactive slider enabling users to navigate across the chosen time-axis values dynamically. Each value on this axis corresponds explicitly to the unique entries from the user-defined time column selected during the dataset configuration stage. These entries are ordered numerically, chronologically, or categorically (based on appearance order). Adjustments via the slider instantly update the group composition views and outcome plots, facilitating immediate exploration of group membership changes and outcome evolution. This capability is particularly effective for pinpointing trends, seasonality, and evolving patterns.

Together, the GroupView Controls provide a comprehensive, flexible framework for segmenting datasets, forming precise group definitions, aggregating outcomes, and tracking membership dynamics across specified time-axis values. They allow users to customize analyses according to domain-specific needs, whether investigating different data versions, exploring group behavior, or uncovering patterns across multiple temporal or categorical checkpoints. These controls form the foundation of GroupView's interactive and dynamic approach to multidimensional data exploration.

3.2.2. GroupView Visual Explorer

The GroupView Visual Explorer facilitates detailed exploration of group compositions and target outcomes across selected features. It combines intuitive visualizations with interactive tools, enabling users to analyze attribute-level dynamics and track how groups evolve over time. As shown in Figure 3, Sections (2) through (8), the Visual Explorer provides multiple features to support detailed analysis.

Attribute Legends and Group Sorting

Above each rectangular block on the dashboard, as shown in Figure 3, Sections (2), (3), and (6), labels (legends) represent different features or columns in the dataset. Each label is accompanied by color indicators corresponding to that feature's categories or numerical brackets. Categorical features display distinct categories, while continuous features are divided into user-specified or automatically generated numerical brackets (up to 7 by default) to maintain simplicity and clarity. This process, termed "feature discretization," facilitates grouping by creating visually uniform categories or brackets, simplifying comparisons and group formation.

Users can interact with the attribute legends to dynamically sort groups. Clicking on a legend attribute prioritizes the groups based on that attribute. For example, clicking on the Gender feature legend could sort groups by 'male' or 'female.' Clicking again reverses the order, helping users quickly compare attribute distributions. This sorting functionality supports intuitive pattern recognition, enabling users to identify correlations and trends effectively.

Group Composition and Target Outcome Visualization

The Group Composition feature in the Visual Explorer provides a detailed breakdown of how categories or numerical brackets are distributed within each group, as shown in Figure 3 (Sections (5), (3), and (6)). Each group is represented by rectangular blocks corresponding to specific features (e.g., gender, age, region) and their associated categories or numerical brackets. For example, a block for *Gender* might be filled 60% with the color representing "female" and 40% with the color representing "male," offering an intuitive overview of the group's gender distribution. Additionally, each group includes a dedicated rectangle indicating its relative size compared to other groups, allowing users to assess dataset distributions quickly; hovering over these rectangles reveals precise percentage values to support detailed analysis.

The Outcome Plot Visualization further enhances the exploration of group dynamics by illustrating how the target variable evolves over time. This visualization is available in two modes:

- Yearly Outcome Plot: Displays target outcomes at individual time points, enabling users to detect sudden changes or seasonal patterns.
- Aggregate Outcome Plot: Summarizes changes over time using two aggregation methods, Initial Checkpoint (which compares outcomes at each time point to the earliest state) and Previous Checkpoint (which compares outcomes between consecutive time points), thus highlighting both cumulative and incremental trends.

Each group's outcome is represented by a rectangular block divided into four color-coded categories, as shown in Figure 3, Section (6). The target values are binary (e.g., poor/not poor, employed/unemployed), and the categories are determined by comparing the initial target value (at the earliest time point) with the current value. The categories are defined as follows:

- Target Never (0,0): This category represents data points where the target has consistently remained 0 (e.g., a person who never had a job or always been rich).
- Target Gained(0,1): This indicates a transition from 0 to 1 (e.g., from not poor to poor or from not having a job to gaining employment).
- Target Protected (1,1): This category reflects stability where the target remains 1 (e.g., consistently poor or consistently employed).
- Target Lost (1,0): This represents a transition from 1 to 0 (e.g., from poor to not poor or from having a job to losing it).

When the Aggregate checkbox is selected, outcomes are calculated across all time points up to the current time value. Users can choose between Initial Checkpoint and Previous Checkpoint aggregation methods (Figure 4). These visualizations help identify patterns in group behavior and track shifts in target outcomes.

- Initial Checkpoint: Compares outcomes at each time point to the dataset's starting state (i.e., the earliest time value). For example, if the current time-axis value is 5, comparisons include time 0 to 1, time 0 to 2, and so on, up to time 5 (0–1, 0–2, . . . , 0–5). This approach provides a longitudinal view of cumulative changes over time. These outcomes are visualized in Figure 5.
- Previous Checkpoint: Compares outcomes between consecutive time points, capturing incremental changes (e.g., 0–1, 1–2, . . . , 4–5). This method identifies short-term trends or variations, such as month-over-month or year-over-year shifts, as shown in Section (B) of Figure 4. This approach is ideal for tracking short-term changes, such as month-over-month or year-over-year trends.

These aggregation modes allow users to examine changes in target outcomes from both a cumulative and incremental perspective, uncovering trends such as gradual improvements, abrupt transitions, or recurring patterns.

The defined outcome categories provide a comprehensive view of the target's progression, enabling users to identify shifts in group outcomes quickly. Customizable, color-coded proportions illustrate the percentage of members in each category, and interactive features, such as hovering to reveal precise values, support detailed analysis.

Overall, this dynamic visualization of target outcomes provides valuable insights into group behavior over time. By simultaneously depicting persistent states and transitions, the tool facilitates the detection of trends, anomalies, and disparities among groups. Moreover, the ability to customize labels, ensuring that outcome categories are presented in meaningful, context-appropriate terms, makes the visualization more accessible and understandable to experts and novices, thereby broadening its applicability across diverse domains such as poverty studies, healthcare analyses, and customer retention tracking.

Figure 4. Using the FairPlay [34] dataset, this figure highlights using the Aggregate checkbox and the Reference Point options in Section (A), where the user has selected the Previous Checkpoint radio button. This configuration visualizes the target outcomes vertically for each time step, showing outcomes for the selected time value of 5 and preceding time values (0 through 5), illustrated in Section (B). The figure also demonstrates multiple selected groups using the checkboxes, which activate the Compare button, as shown in Section (C). This functionality enables users to compare selected groups side by side, providing a detailed view of the evolution of group outcomes over time.

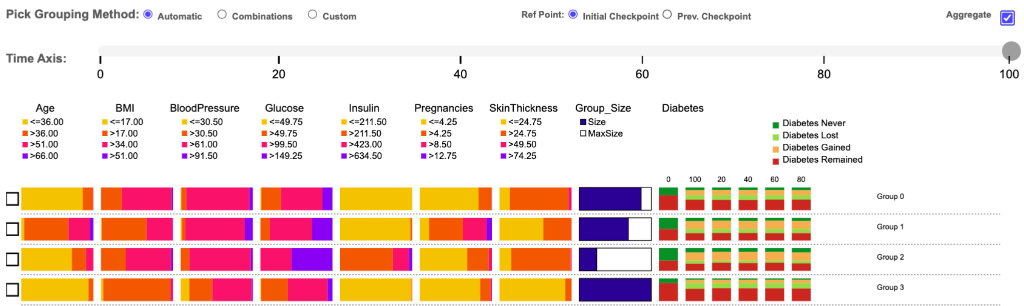

Figure 5. Diabetes dataset in GroupView with the Initial Checkpoint aggregation. Groups (Group 0–Group 3) were formed automatically from attributes such as age, BMI, and glucose. The time axis shows model training iterations (0, 20, 40, 60, 80, 100). Outcome categories are color-coded as shown in the legend; each iteration is compared to iteration 0 (ground truth), revealing performance drift and underfitting patterns.

Group Selection and Comparison

The Group Selection and Comparison feature combines dynamic time slicing with attribute-level analysis, enabling users to explore how group compositions and target outcomes evolve over time. Users can sort and compare groups interactively, gaining valuable insights into the relationships between groups.

Each group row includes a checkbox for selection (Figure 3, Section (4)). Selecting a single group activates the Sort by Similarity option, as shown in Figure 3, Section (8), which reorders the groups based on overlapping feature compositions. The selected group is moved to the top, while others are sorted in descending order of similarity using a Euclidean distance metric over the selected feature attributes. This functionality helps identify trends and anomalies between similar groups.

When multiple groups are selected, the Compare button appears, aligning the chosen groups side by side (Figure 4, Section (C)). This view simplifies the comparison of group characteristics and outcomes, allowing users to detect disparities and similarities easily.

Group labels (e.g., Group 0, Group 1) are assigned once at the first time-axis value, based on the initial group definitions formed through clustering, attribute combinations, or user-defined criteria. These labels are preserved across time to ensure consistency in tracking group membership and behavior. Although group ordering in the UI may change (e.g., after similarity-based sorting), the underlying label always refers to the original group identity, allowing for accurate longitudinal analysis.

The number of groups is fixed at initialization, either determined automatically (e.g., using a knee-point heuristic for clustering), generated from attribute combinations, or specified manually in the configuration. While group definitions remain unchanged, individual records may shift between existing groups over time as their attributes evolve. This design supports meaningful analysis of group membership transitions without altering the underlying group identities.

The GroupView Visual Explorer integrates intuitive visualizations, dynamic controls, and interactive features to provide a powerful tool for analyzing group dynamics and target outcomes. Its adaptability ensures its relevance across diverse domains, empowering users to uncover meaningful insights and make data-driven decisions with clarity.

3.3. Implementation Details

GroupView is implemented as a web application with a Python 3.x backend and a JavaScript/HTML/CSS frontend. The backend, built using the Flask web framework, manages data ingestion, preprocessing, and communication with the frontend through RESTful API endpoints.

The frontend uses HTML for structure, CSS for styling, and JavaScript for interactivity. Interactive visualizations are powered by D3.js, enabling dynamic rendering of group membership compositions and transitions over time.

Additional libraries, such as jQuery, support Document Object Model (DOM) manipulation and asynchronous data loading. Dataset metadata is stored in JSON configuration files, and data for analysis is provided in CSV format, either from preconfigured datasets or user uploads.

4. Use Cases

The power of GroupView lies in its versatility and adaptability across various contexts, making it an invaluable tool for exploring group dynamics and target outcomes. This section delves into four distinct use cases, highlighting the tool's ability to transform complex datasets into meaningful visual insights. Each use case was chosen for its unique challenges and opportunities, allowing us to demonstrate how GroupView can be applied to real-world scenarios, from tracking socioeconomic shifts to analyzing performance metrics. These examples showcase the flexibility of the tool, its ease of use, and the insights it can unlock, helping users make more informed decisions and uncover hidden trends.

4.1. FairPlay: Job Hiring Consensus

The FairPlay use case illustrates a collaborative decision-making scenario where multiple stakeholders negotiate job hiring outcomes across successive rounds. Each user study involves five stakeholders, each with distinct preferences for fairness and candidate selection criteria. Over these rounds, stakeholders refine their decisions to achieve a consensus on which candidates should be hired [34].

The dataset for this use case tracks the evolution of candidate hiring statuses throughout the negotiation process. GroupView's configuration is tailored to capture these dynamics: the Userstudy column differentiates between user studies, each involving five stakeholders; the Candidate Id column uniquely identifies individual candidates; and the Round column functions as the time axis, reflecting the progression of decisions over time. Stakeholders' preferences define custom grouping methods, where candidates are categorized by attributes such as gender (male/female) or college rank (elite/non-elite).

Using GroupView, the visualization provides insights into how candidate outcomes—hired or not—shift over time as stakeholders converge on a collective decision. It highlights the evolution of group preferences and the interplay of stakeholder priorities, revealing the factors influencing consensus. Detailed information about the data preparation and processing for this use case can be found in Appendix A.1.

This use case exemplifies GroupView's capability to navigate and analyze complex, multi-round decision-making processes, offering a clear and actionable view of how consensus emerges dynamically.

4.2. Tracking US County Poverty Status over Time

This use case focuses on analyzing the poverty status of US counties from 2012 to 2022 using a dataset that includes various socioeconomic and geographical characteristics. The goal is to track how poverty status evolves and explore the factors influencing these changes across different counties. For this analysis, we used the following features: GEOID, county name, year, poverty indicator, the total population, total labor force, the percentage of population below poverty, the percentage of males in poverty, the percentage of females in poverty, the unemployment rate, the median value of housing, and the demographic composition (the percentages of the White, Black, American Indian and Asian population). These features provide insights into economic disparities across different racial and geographic distributions.

For this study, GroupView was configured using two separate feature sets to analyze different aspects of poverty dynamics:

- The first analysis focused on understanding the racial and housing factors influencing poverty by examining demographic composition and housing values.
- The second analysis explored gender-based disparities by incorporating features related to male and female poverty percentages.

Both analyses shared the following core configurations:

- Index Column: GEOID, uniquely identifying each county across the dataset and year;
- Time Axis: Year, tracking changes in poverty status over time;
- Target Column: poverty indicator (Poor_Status), categorizing counties as either "Poor" or "Not Poor".

Analysis 1: Racial and Housing Factors

- Total Population;
- Median Housing Value;
- Percentage of Population Below Poverty;
- Demographic Composition: Percentages of White, Black, American Indian, and Asian populations.

Analysis 2: Gender-Based Disparities

- Total Population;
- Percentage of Population Below Poverty;
- Percentage of Males in Poverty;
- Percentage of Females in Poverty;

- Labor Force;
- Unemployment Rate.

The GEOID column serves as the Index Column, uniquely identifying each county in the dataset, while the year column acts as the Time Axis, tracking changes in poverty status over the years. The Poor_Status column is selected as the Target Column, representing whether a county is classified as "poor" or "not poor" in a given year. Columns not selected for analysis were designated as unused columns to streamline the evaluation process.

The dataset used in this use case was prepared by aggregating data from publicly available sources, primarily the *American Community Survey (ACS)* 1-Year and 5-Year estimates for the years 2012 to 2022. Supplemental socioeconomic and demographic datasets were integrated to enrich the analysis. Detailed metadata and preprocessing steps are outlined in Appendix A.5.

GroupView enables users to visualize how poverty status changes over years, revealing trends such as persistent poverty, counties that escape poverty, or those that fall into poverty over time. These insights are valuable for understanding the socioeconomic factors driving poverty dynamics and can support policy-making decisions and resource allocation.

4.3. Tracking the Progress of Training a Machine Learning Model

The third use case focuses on tracking the training progress of a machine learning model for predicting diabetes outcomes. The dataset comprises several health-related features, along with machine learning predictions, enabling the analysis of how accurately the model tracks diabetes progression over time. The data include columns such as Pregnancies, Glucose, BloodPressure, SkinThickness, Insulin, BMI, DiabetesPedigreeFunction, and Age, providing a comprehensive view of patients' health metrics. Additionally, the dataset includes a time index that tracks the model's prediction performance across multiple iterations.

GroupView was configured as follows in this use case: The Iteration Time column was set as the Time Axis, tracking changes in the model's predictions at regular intervals. The Patient ID (PID) served as the Index Column, uniquely identifying each patient in the dataset. The Diabetes column was used as the Target Column, representing whether the patient had been diagnosed with diabetes (1) or not (0) based on the machine learning model's predictions. The data were processed at 20-iteration intervals, capturing the model's performance in predicting the target outcome at different points in time.

Using GroupView (as shown in Figure 5), users can visualize how the machine learning model's accuracy evolves over time and how health metrics influence the model's predictions. By comparing the true diabetes status with predicted outcomes, GroupView helps identify where the model performs well and where it struggles, offering insights into potential areas for model improvement or further investigation. Detailed information about the data source and preparation can be found in Appendix A.10.

This use case demonstrates GroupView's ability to handle machine learning predictions and time-series data in healthcare applications. It showcases its potential to support informed decision-making in clinical and research contexts.

5. Insights

This section summarizes the key insights from each use case, highlighting how GroupView enabled meaningful exploration of data patterns, group dynamics, and target outcomes over time.

5.1. Consensus Building Patterns in FairPlay

The FairPlay use case provides a unique lens to analyze stakeholder decision-making dynamics during collaborative hiring processes. By leveraging GroupView's capability to track group compositions and target outcomes across multiple rounds, patterns highlighting key factors influencing hiring consensus emerge (see Figures 3 and 4). Figure 3 shows the Job Outcome at the final time-axis (round) value 5 compared to the initial time-axis (round) value 0, whereas Figure 4 displays the aggregate job outcomes across all time-axis (round) values relative to the previous round. Crucially, the group assignments remain the same in both figures, making it possible to cross-reference them to understand how job outcomes evolve until the final round.

5.1.1. Initial Observations (Round 0)

At first glance in the aggregate plot at Round 0, there are two distinct patterns in the distribution of jobs. Groups 0, 2, 4, 8, and 9 (collectively called G_{young}) show the highest proportion of Job Protected outcomes and the lowest proportion of Job Never, whereas the remaining Groups 1, 3, 5, 6, and 7 (G_{old}) feature a higher percentage of Job Never. This illustrates the baseline scenario before stakeholders started making changes using FairPlay.

- $\mathbf{G}_{\text{young}}$: This group is the union of Groups 0, 2, 4, 8, and 9. These groups are characterized by having younger candidates (Age $\leq$ 42) with lower work experience (Work_Experience $\leq$ 24). In the initial dataset, these groups held the majority of the jobs.
- $\mathbf{G}_{\text{old}}$: This group is the union of Groups 1, 3, 5, 6, and 7. They consist of older candidates (Age $>$ 42) with higher work experience (Work_Experience $>$ 24). Compared to G_{young}, these groups started with fewer jobs in the initial distribution.

The initial dataset clearly shows that a significantly larger share of job outcomes was allocated to the G_{young} groups compared to the G_{old} groups.

5.1.2. Round 1 Dynamics

Round 1 reveals a notable shift: most groups experience some form of Job Lost and Job Gained. A closer look shows that the highest Job Lost occurs in G_{young} groups, while the G_{old} groups garner the highest Job Gained. This is likely because G_{young} started with a disproportionate share of positions, and in an effort to distribute jobs more evenly, stakeholders appear to have reallocated some of those positions to underrepresented groups.

Within G_{young}, Groups 2, 8, and 9 stand out in that all members initially had jobs at Round 0 but then lost a majority of them by Round 1. Notably, varying indicators such as GPA or SAT score did not appear to shift this outcome, as both Groups 2 and 8 exhibit similar job-loss trends despite opposite ranges in GPA or SAT scores.

Looking closely within G_{young}, Groups 2, 8, and 9 are particularly interesting. All members in these groups had jobs at Round 0, but at Round 1, they experienced the highest rates of job loss. This shift suggests stakeholders were willing to sacrifice these jobs without a corresponding job gain or trade. Moreover, the contrasting GPA and SAT scores between Groups 2 and 8 did not translate into divergent outcomes; both groups ended up with similar overall outcomes. This observation suggests that these job indicator metrics did not significantly influence stakeholder decision-making despite their differing academic profiles.

Conversely, within G_{old}, Group 1 had the highest share of members without jobs at Round 0 but ended up with the highest job gains by Round 1. This correlates with the relatively higher GPA or SAT scores; it suggests a possible stakeholder strategy to ensure that groups that highlight certain merits, such as stronger academic indicators, receive more jobs.

Group 0 (in G_{young}) and Group 7 (in G_{old}) are also intriguing due to their similar group sizes and a near balance between **Job Lost** and **Job Gained**. Their exchange pattern hints that individuals may have moved between them or that these groups approached parity in some underlying attribute, such as SAT score distributions, leading to near-equivalent swapping of candidates.

5.1.3. Consensus and Final Outcome

Over multiple rounds, stakeholders continued refining these allocations in pursuit of a stable consensus. By Round 3, job statuses in each group began to "lock in," with negligible changes thereafter through the final round (Round 5). This indicates that stakeholders, without the benefit of GroupView's visualization, nonetheless converged on an arrangement they perceived as equitable based on their group goals and negotiations.

5.1.4. Concluding Observations

By Round 5, the visible convergence in job outcome distributions between the G_{young} and G_{old} groups reflects a decisive balancing effect. Although the G_{young} groups started with a higher share of jobs, iterative stakeholder negotiations gradually reconciled these initial discrepancies. This outcome reinforces the notion that stakeholder negotiations were not merely about shifting deficits or surpluses but rather about recognizing that each group brought valuable qualities to the table. In effect, the process effectively leveraged the diverse strengths, be it youth with lower work experience or the advanced academic and experiential credentials of older candidates, to arrive at a final, equitable distribution.

5.2. Poverty Trends in US Counties (2012–2022)

Exploration of the US county-level poverty data in GroupView uncovered several patterns that reveal the relationship between population characteristics and poverty outcomes over time, with a notable shift around the pivotal period of 2015–2017.

5.2.1. Persistent Poverty Trends

Certain county groups exhibited sustained poverty throughout the study period with little to no improvement. This is shown in Figure 6, where we re-ordered the groups by poverty percentage to highlight the disparities more clearly and make patterns across different demographics easier to identify.

- American Indian Communities (Group 2, top row in Figure 6): Counties with a majority American Indian population (>48%) remained in poverty from 2012 to 2020 with no significant changes. These counties also had consistently high poverty rates across the population, indicating entrenched structural disadvantages preventing economic recovery.
- Black-Majority Counties (Group 1 and Group 7; see Figure 7): Counties with moderate to high proportions of Black populations experienced some of the highest rates of sustained poverty. Group 1, despite showing slight improvement starting in 2016, had approximately 80% of its counties still classified as poor through 2020, indicating limited long-term progress. Similarly, Group 7 had the third-highest proportion of counties that remained in poverty throughout the period. The persistence of poverty in these counties suggests systemic barriers that continue to restrict economic mobility. Interesting in this respect is to compare these two groups with the large Group 0 (immediately below), which has a predominantly White population and has made greater gains in escaping poverty. In fact, comparing these three groups in order from top to bottom shows a clear shift from Black to White population, easily visualized by the shift in colors (purple and magenta) in the composition rectangles of these two

demographics. It is this visual representation that makes the correlation of wealth gains and demographic composition readily apparent.

These three groups are interesting to compare since they share similar features in all other aspects, as can be easily observed by the similar composition rectangles of all other features. While the other groups below continue this trend in wealth gain, for them, the composition rectangles of other features are more diverse, indicating that these gains cannot be attributed to demographic shifts alone. Again, the visual representation of the composition rectangles makes this easy to recognize.

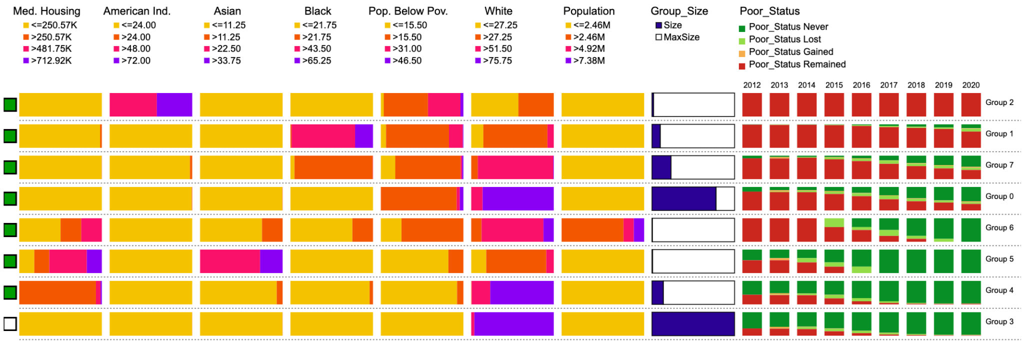

Figure 6. GroupView aggregate outcome plot from 2012 to 2020, showing year-over-year changes relative to its successive previous year. The plot visualizes group compositions across key features and their corresponding categories or numerical brackets: Median Housing, Percentage American Indian, Percentage Asian, Percentage Black or African American, Percentage Population Below Poverty, Percentage White, Total Population, and Group Size. The target outcome, Poor Status, is color-coded: Poor Status Never (green), Poor Status Lost (light green), Poor Status Gained (orange), and Poor Status Remained (red). Each stacked bar represents one year, illustrating temporal and regional disparities in poverty trends. Notably, the pivotal period of 2015–2017 marks significant shifts in poverty trajectories across racial and housing categories. This visualization highlights persistent poverty clusters, recovery trends, and demographic disparities, enabling the identification of high-risk counties and informing targeted policy interventions.

Figure 7. Zoomed excerpt of Figure 6 comparing Group 2, Group 1, and Group 7. The plot visualizes group compositions across seven key features (in left-to-right order): (1) Median Housing, (2) Percentage American Indian, (3) Percentage Asian, (4) Percentage Black or African American, (5) Percentage Population Below Poverty, (6) Percentage White, (7) Total Population, and (8) Group Size. The target outcome, Poor Status, is color-coded as follows: Poor Status Never (green), Poor Status Lost (light green), Poor Status Gained (orange), and Poor Status Remained (red).

5.2.2. Economic Recovery and Poverty Reduction

Some county groups demonstrated substantial improvements in poverty reduction over time.

- Asian-Dominant Counties (Group 5 and Group 6; see Figure 8): Counties with a moderate (>11%) to high (>33%) Asian population exhibited notable improvements. Group 5 transitioned from having nearly half its counties in poverty in 2012 to achieving total poverty elimination by 2016. Likewise, Group 6, with its substantial Asian population and a mix of other racial groups, moved out of poverty entirely by 2018, highlighting the association between Asian-majority regions and economic stability. In fact, there appears to be a correlation between the Asian dominance in a group and the speed of the wealth gain (compare the colors of the respective 'Asian' composition rectangles), albeit the groups themselves are very small (see the Group Size column).

Figure 8. Zoomed excerpt of Figure 6 comparing Group 5 and Group 6. The plot visualizes group compositions across seven key features (in left-to-right order): (1) Median Housing, (2) Percentage American Indian, (3) Percentage Asian, (4) Percentage Black or African American, (5) Percentage Population Below Poverty, (6) Percentage White, (7) Total Population, and (8) Group Size. The target outcome, Poor Status, is color-coded as follows: Poor Status Never (green), Poor Status Lost (light green), Poor Status Gained (orange), and Poor Status Remained (red).

- White-Dominant Counties (Group 0, Group 3, and Group 4; see Figure 9): White-majority counties exhibited two distinct recovery patterns based on median housing values and total population below poverty. Group 4, characterized by high median housing values, remained largely unaffected by poverty, reinforcing the link between housing wealth and economic stability. In contrast, Groups 3 and 0, with similar White-majority compositions but lower median housing values, experienced notable poverty reduction by 2020. However, Group 0 retained a higher share of persistent poverty, likely due to its significant proportion of the population's pre-existing historical poverty. This suggests that while housing value correlates with economic resilience, pre-existing poverty levels play a more decisive role in long-term economic mobility.

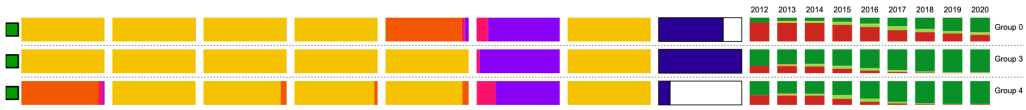

Figure 9. Zoomed excerpt of Figure 6 comparing Group 0, Group 3, and Group 4. The plot visualizes group compositions across seven key features (in left-to-right order): (1) Median Housing, (2) Percentage American Indian, (3) Percentage Asian, (4) Percentage Black or African American, (5) Percentage Population Below Poverty, (6) Percentage White, (7) Total Population, and (8) Group Size. The target outcome, Poor Status, is color-coded as follows: Poor Status Never (green), Poor Status Lost (light green), Poor Status Gained (orange), and Poor Status Remained (red).

5.2.3. Racial and Demographic Disparities in Poverty

Demographic composition played a significant role in poverty status and economic mobility.

- Black-Dominant Counties (Group 1): With a high proportion of Black population (>43%), this group experienced mixed trends. While there was a notable reduction in poverty between 2012 and 2015, about 20% of counties remained in poverty through 2020. This shift is partially attributed to a decline in population below poverty (>15% moving to <15%) and a decrease in White population percentages.
- White-Dominant Counties (Group 0, Group 3, and Group 6; see Figure 10): White-majority counties exhibited divergent poverty trends. Group 3, despite having a high White population and low median housing values, successfully reduced poverty by 2020, whereas Group 0 still retained approximately 20% of counties in poverty. Group 6, with a mixed racial composition, showed near-complete poverty elimination by 2019, aligning with urban county characteristics.

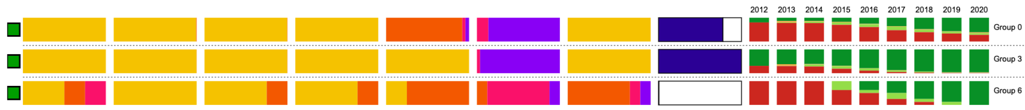

Figure 10. Zoomed excerpt of Figure 6 comparing Group 0, Group 3, and Group 4. The plot visualizes group compositions across seven key features (in left-to-right order): (1) Median Housing, (2) Percentage American Indian, (3) Percentage Asian, (4) Percentage Black or African American, (5) Percentage Population Below Poverty, (6) Percentage White, (7) Total Population, and (8) Group Size. The target outcome, Poor Status, is color-coded as follows: Poor Status Never (green), Poor Status Lost (light green), Poor Status Gained (orange), and Poor Status Remained (red).

5.2.4. Housing and Population Density Influence on Poverty

The relationship between housing values and the total population played a crucial role in shaping poverty trends.

- High-Median Housing as a Poverty Indicator (Groups 4, 5, and 6): Counties with moderate to high median housing values were nearly free from poverty by 2020, suggesting that housing wealth can serve as an indicator of economic stability.
- Low-Median Housing and Mixed Poverty Outcomes (Group 1 and Group 3): Counties with lower median housing values exhibited varied outcomes—some remained in poverty, while others transitioned out, indicating that housing alone is not a direct predictor of poverty without considering other socioeconomic factors.

5.2.5. Gender Disparities and Economic Mobility

Figure 11 highlights the impact of gender on poverty dynamics. Counties with higher combined male and female populations below the poverty line have experienced persistent economic hardship. In contrast, countries with lower overall gender-based poverty levels have shown faster transitions out of poverty. These trends underscore the importance of addressing gender disparities in achieving economic mobility through targeted employment and social policies.

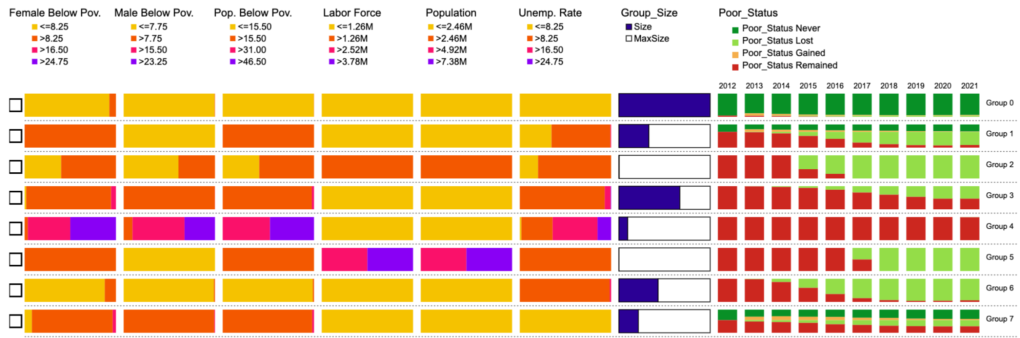

Figure 11. GroupView aggregate outcome plot from 2012 to 2021, visualizing gender disparities in poverty trends. Unlike the previous figure, which showed year-over-year changes, this plot compares each year's poverty status relative to the initial year (2012). The plot captures trends in Percentage Female Below Poverty, Percentage Male Below Poverty, Percentage Population Below Poverty, Labor Force, Total Population, and Unemployment Rate.

- Persistent Poverty in Counties with High Gender-Based Poverty:
  Counties in Groups 3, 4, and 7 exhibit the highest overall poverty rates. These are counties where poverty rates are high for both males and females. In these counties, the simultaneous elevation of poverty rates among both genders appears to intensify economic challenges. The most significant is Group 4, which showed no improvement over time, likely reflecting high unemployment rates, while Group 3 managed only moderate progress, with nearly half of its counties still in poverty by 2021. This

combined effect emphasizes that high poverty rates in both demographics reinforce each other, thereby impeding economic mobility.

- Faster Economic Recovery in Counties with Lower Gender-Based Poverty:
  In contrast, counties in Groups 0, 2, and 6 reported lower combined poverty levels, corresponding with substantial improvements by 2021. Additionally, Groups 1 and 5, despite maintaining moderately high female poverty rates, demonstrated rapid recovery, likely due to their lower male poverty levels. One reason could be that lower male poverty levels in these groups helped stabilize household income overall. In many households, men are still primary earners or contribute significantly to financial support. So, even when female poverty remains relatively high, the economic security provided by lower male poverty may facilitate quicker recovery.
  This overall pattern suggests that counties are better positioned for economic recovery when both male and female poverty rates are reduced or when a critical factor (such as male poverty) is particularly low. Nonetheless, even lower combined poverty may not fully translate into sustained progress in areas where high unemployment still persists (e.g., Groups 1 and 6).

These insights illustrate GroupView's ability to uncover deeply embedded patterns in poverty dynamics by examining the interplay of demographic and economic factors over time. The findings reveal that the combined effect of gender-based poverty, along with factors like racial composition, median housing value, and employment conditions, plays a significant role in both poverty persistence and recovery.

5.3. Evolution of Diabetes Predictions in Machine Learning

The GroupView dashboard offers a dynamic perspective on the evolution of model predictions during a machine learning training process. In this setup, each time-axis value represents a training iteration checkpoint, starting with the ground truth at iteration 0 and progressing in increments of 20 up to 100. By using GroupView to visualize model outcomes across training stages, practitioners can track prediction performance across patient subgroups, identify error patterns, and detect when training progress stagnates or underfitting occurs. This subgroup-specific view provides insights beyond traditional aggregate performance metrics, enabling developers to pinpoint groups where the model struggles and adjust training data or model parameters accordingly.

5.3.1. Model Saturation and Underfitting

Visualizing the aggregate plot relative to the Initial Checkpoint (Figure 5) reveals a notable performance plateau after the first 20 training iterations. In Groups 0 through 3, the distribution of *Diabetes Gained* (orange, indicating false positives) and *Diabetes Lost* (light green, indicating false negatives) stabilizes early in the training process, with minimal improvements beyond iteration 20. This stagnation is further confirmed by the aggregate plot relative to the Previous Checkpoint (Figure 12), which shows no notable changes after iteration 20. Together, these views highlight underfitting, likely due to insufficient training data or limited model capacity, emphasizing the importance of early detection to prevent wasted training cycles.

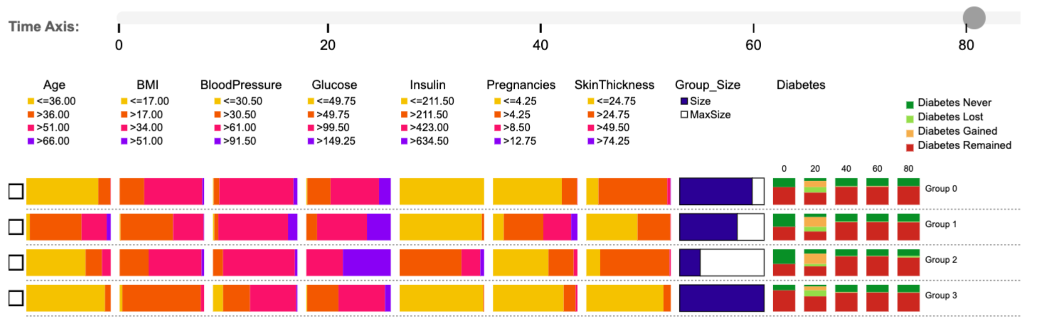

Figure 12. The diabetes dataset in GroupView with the Previous Checkpoint aggregate view, showing automatically clustered groups (Group 0 to Group 3) based on characteristics like Age, BMI, and Glucose. The time axis represents model training iterations (20, 40, 60, 80), with each checkpoint compared to its preceding checkpoint (e.g., iteration 40 is compared to 20). The target outcome, Diabetes Status, is color-coded: Diabetes Never (green), Diabetes Lost (light green), Diabetes Gained (orange), and Diabetes Remained (red). This view highlights whether model predictions change from one checkpoint to the next, revealing stagnation after iteration 20.

5.3.2. Group-Specific Prediction Patterns

While all groups exhibit the general pattern of underfitting, Group 2 shows relatively fewer false negatives (*Diabetes Lost*), suggesting better model performance for patients with its demographic composition, likely since this is the only group with high insulin. Conversely, Group 3 shows fewer false positives (*Diabetes Gained*), implying that the model better identifies non-diabetic cases for this subgroup. Somewhat similar in this behavior to Group 0, these groups are marked by younger populations, with relatively lower glucose levels, and lower pregnancy rates (which is correlated with age). Finally, Groups 0 and 1 experience higher rates of false positives and false negatives, indicating poorer generalization for patients within these groups.

These interpretations are consistent with the general medical understanding of diabetes risk factors. High insulin levels, as seen in Group 2, may reflect insulin resistance and thus improve the model's ability to detect diabetes. Conversely, Groups 0 and 3, characterized by lower glucose levels, fewer pregnancies, and younger age, are more likely to represent lower-risk populations, which could explain the lower false positive rates. It is important to note that while these associations are medically plausible, factors like data quality (e.g., missing insulin values) and diabetes type may influence the reliability of these patterns. This is discussed next.

5.3.3. Implications for Data Augmentation

The differences in subgroup prediction performance offer practical guidance for targeted data augmentation and model improvement:

- **Groups 0 and 1:** The higher rate of misclassifications suggests that additional training data are needed, particularly more samples representing both diabetic and non-diabetic patients from the demographic profiles captured by these groups.
- **Group 2:** Since this group shows fewer false negatives, it is likely better represented in the training data. However, adding more positive (diabetic) cases could enhance model sensitivity without compromising specificity.
- **Group 3:** This group's fewer false positives indicate that the model is adept at recognizing non-diabetic patients. Supplementing training data with more diabetic examples for this group could further balance the classification performance.

Combining insights from both the Initial Checkpoint (Figure 5) and the Previous Checkpoint (Figure 12) aggregate views, GroupView not only reveals underfitting and stagnation but also highlights the heterogeneity in model behavior across patient subgroups.

This underscores the limitations of relying solely on global performance metrics during training. Instead, subgroup-specific evaluation and data augmentation strategies—such as targeted data collection for underrepresented groups—can enhance the model's robustness and fairness, ensuring better generalization across diverse population segments.

6. Heuristic Evaluation Results

We conducted a heuristic evaluation, which provided valuable insights into the GroupView framework's usability, effectiveness, and areas for refinement. Five participants, including visualization experts and non-experts, assessed the system based on ease of use, visual clarity, and ability to support insight discovery during dynamic data exploration.

6.1. Summary of Findings

Participants evaluated GroupView using a Likert scale (1–5) and shared qualitative feedback. The findings are summarized across three key themes:

6.1.1. Ease of Use and Configuration

Participants found the configuration panel intuitive, assigning an average score of 5.0. They appreciated the step-by-step workflow and flexibility in selecting features, target variables, and time-axis values.

6.1.2. Visualization Clarity

Attribute-level categories, numerical bracket visualizations, and target outcome plots were positively received for their readability and interpretability. This aspect received an estimated average rating of 4.5.

6.1.3. Insight Discovery

While most users were able to identify actionable patterns, this area scored slightly lower (average 4.3), indicating the potential value of adding guidance or contextual prompts to aid less experienced users.

6.2. Qualitative Themes

Strengths: Participants praised the interactive visualizations and dynamic grouping features, particularly the ability to track changes in group membership over time. They also valued the temporal analysis capabilities for revealing trends and supporting exploratory insights.

Areas for Improvement: Suggestions included enhancing onboarding through detailed tooltips or walkthroughs and offering preset configuration templates to simplify the customization interface for non-experts.

6.3. Evaluation Against Design Objectives

- Usability: technical and non-technical participants consistently described the system as user-friendly.
- Insight facilitation: users reported discovering meaningful trends and relationships, supporting data-driven analysis.

7. Limitations and Future Work

The evaluation was conducted with a small sample size (five participants) as a formative usability study, aimed at gathering early feedback on the system's design and interactions. Such small-scale evaluations are common in HCI to efficiently surface the majority of usability issues, but they limit the generalizability of the findings. The absence of objective metrics (e.g., task accuracy or completion time) further constrains the depth of analysis. Additionally, the current outcome plots use stacked bar charts, which prioritize categorical clarity and space efficiency but may limit the perception of longitudinal trends compared to line charts.

Future work will expand the participant pool, incorporate objective task-based measures, and explore alternative visual encodings (e.g., line charts for trend analysis), alongside applying the system to a broader range of real-world datasets to enhance robustness and applicability.

8. Discussion

The insights from the use cases and the heuristic evaluation confirm GroupView's potential to address key challenges in dynamic data exploration while also highlighting areas for refinement. This discussion contextualizes these findings in terms of usability, comparative advantages, and opportunities for future development.

The results underscore GroupView's intuitive design—particularly for users familiar with visualization tools. Non-expert participants, however, reported a steeper learning curve, especially with the framework's more advanced features. To support broader adoption, future iterations could incorporate onboarding mechanisms, such as interactive tutorials or guided walkthroughs, to reduce barriers for less experienced users.

A defining strength of GroupView is its ability to integrate multidimensional grouping with temporal tracking—an area where many traditional tools fall short. In contrast to static visualizations, GroupView enables users to identify subtle patterns and explore changes in group behavior over time. Compared to widely used flow-based techniques such as Sankey diagrams, which excel at presenting high-level transitions but abstract away internal group structures, GroupView offers richer analytical depth. By preserving attribute-level detail and enabling flexible, dynamic group definitions, it supports more diagnostic and exploratory tasks, making it well-suited for data-driven decision-making in complex domains.

While the study relied on qualitative feedback from a small group of participants, this approach aligns with established best practices in usability research. As noted by Nielsen [35], a small number of evaluators—typically five to seven—can uncover the majority of usability issues. That said, future evaluations could complement this with quantitative task-based assessments (e.g., time-to-insight, accuracy) and a broader user sample to further validate effectiveness across domains and user expertise levels.

Looking forward, GroupView could benefit from integrating automated recommendations for grouping methods and offering preconfigured templates tailored to common analytical tasks. These enhancements would reduce cognitive load and further improve accessibility, particularly for non-expert users, without sacrificing the flexibility that makes GroupView a powerful tool for multidimensional, temporal data analysis.

Future work could also explore integrating automated recommendations for grouping methods or providing preconfigured templates tailored to common analytical scenarios. Such enhancements would reduce users' cognitive load and make the tool more accessible for non-expert audiences.

9. Conclusions

In conclusion, we have presented GroupView, a visual framework for exploring group dynamics over time. GroupView integrates automated clustering with flexible, user-defined grouping, dynamic time slicing, and version-based slicing, enabling users to track transitions in group membership and assess attribute-level contributions with clarity. One of the key contributions of our work is the outcome plotting functionality: the yearly outcome plot and the aggregate outcome plots (both relative to the initial year and the preceding year) allow users to monitor changes in outcomes over different time spans at a glance. Additionally, the use of attribute composition rectangles offers an intuitive visual explanation of group structure and dynamics, while the ability to sort and compare groups side by side further enhances analytical flexibility. Although GroupView does not radically depart from existing visualization tools, it provides practical improvements in interpretability and interactivity to help users better understand complex, evolving datasets. We acknowledge that further refinement and broader validation are necessary, and we view this work as a constructive step toward more accessible and transparent analysis of dynamic group behavior.

Author Contributions: Conceptualization, M.K.S. and K.M.; methodology, M.K.S. and K.M.; software, M.K.S.; validation, M.K.S. and K.M.; formal analysis, M.K.S.; investigation, M.K.S.; resources, K.M.; data curation, M.K.S.; writing—original draft preparation, M.K.S.; writing—review and editing, K.M. and M.K.S.; visualization, M.K.S. and K.M.; supervision, K.M.; project administration, K.M.; funding acquisition, K.M. All authors have read and agreed to the published version of the manuscript.

Funding: Work reported in this publication was supported by SUNY System Administration under SUNY Research Seed Grant Award 23-01-RSG.

Informed Consent Statement: Not applicable, as the study did not involve sensitive personal data or identifiable information.

Data Availability Statement: The data supporting the reported results in this study are publicly available and have been processed as detailed in Appendix A: **FairPlay Data:** The dataset was sourced from publicly available user studies and preprocessed into a unified dataset. Appendix A.1 details the preprocessing and final dataset structure. **US County Poverty Data:** This dataset was compiled from publicly available American Community Survey (ACS) datasets. Detailed data preparation and column descriptions are provided in Appendix A.5. **Diabetes Data:** The dataset was obtained from publicly available resources and further extended to track machine learning model predictions. Details on the data source and preprocessing steps are outlined in Appendix A.10. The GroupView software is implemented as a Python (Flask) web application with a JavaScript/HTML/CSS frontend. The source code (version v1.0.0) is publicly available at https://github.com/mitsvision2/GroupView (accessed on 10 August 2025).

Acknowledgments: We thank Anthony J. Ripa and Tina Behzad for their constructive feedback, which informed the design and user experience of GroupView, particularly during its application in the FairPlay use case. ChatGPT (OpenAI; models: GPT-3.5 and GPT-4; accessed 2024–2025) was utilized to generate sections of this Work, including text, tables, graphs, code, data, citations, etc.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A

Appendix A.1. FairPlay Data

The dataset used for the FairPlay: Job Hiring Consensus use case was retrieved from a publicly available Github repository [36]. The data originally consisted of separate files for each user study, which were preprocessed and combined into a single dataset to enable unified analysis across all studies.

Below are the details of the data sources and preparation steps:

Appendix A.2. Data Sources

- Github Repository: The dataset was downloaded from a publicly available Github repository. Each file represented a user study conducted in the FairPlay framework, detailing stakeholder decisions and candidate outcomes.

- The data included attributes such as demographic information, academic background, and job outcomes for candidates, with multiple rounds of decision-making recorded for each user study.

Appendix A.3. Processing Steps

- Data Consolidation:
  - Combined individual files representing separate user studies into a single dataset for comprehensive analysis.
  - Appended a `Userstudy` column to identify the source of each record and a `Round` column to represent decision-making rounds as the time axis.
- Data Cleaning:
  - Standardized column names to ensure consistency across user studies.
  - Removed redundant or irrelevant columns not used in GroupView analyses.
- Feature Engineering:
  - Encoded categorical attributes such as `gender`, `race`, and `college rank` to align with GroupView's grouping methods.
  - Ensured compatibility of numerical columns like `Grade Point Average` and `SAT Score` for effective visualization.

Appendix A.4. Final Dataset

The final dataset consisted of approximately 40,000 rows and included the following key columns:

- `Gender, Race, Age, Work_Experience, Major, Grade_Point_Average, SAT_Score, College_Rank, Job, Userstudy, Round, id.`

This preprocessed dataset was prepared to enable dynamic visualizations and analyses using the GroupView framework. By consolidating and standardizing the data, the final dataset provides a robust foundation for exploring group dynamics, stakeholder decisions, and hiring outcomes across multiple studies and rounds.

Appendix A.5. US County Poverty Data

The dataset used to analyze poverty trends in US counties from 2012 to 2022 was compiled by integrating multiple sources from the *American Community Survey (ACS)* [37]. This dataset captures socioeconomic and demographic indicators at the county level, allowing for a detailed analysis of poverty dynamics.

Appendix A.6. Data Sources

The dataset was created using the following ACS sources:

- ACS 5-Year Estimates (2012–2022): Provides aggregated data covering all US counties, ensuring broader coverage, including smaller counties with less frequent updates.
- ACS Subject Tables: Supplemental datasets covering demographics, employment, housing, and poverty measures.

The datasets downloaded from ACS are as follows:

- B01003—Total Population: Contains the total estimated population per county (`B01003_001E`).
- B17001—Poverty Status by Sex and Age: Provides the count of individuals below the poverty level (B17001_002E) and breakdowns by gender (B17001_003E, B17001_017E).
- B23025—Employment Status: Includes the total labor force (B23025_002E) and total unemployed population (B23025_005E).

- B25077—Median Housing Value: Reports the median value of owner-occupied housing units (B25077_001E).
- B02001—Race Distribution: Contains demographic composition by race, including White (B02001_002E), Black or African American (B02001_003E), American Indian and Alaska Native (B02001_004E), and Asian (B02001_005E).

Appendix A.7. Data Processing and Preparation

To construct a comprehensive dataset, the following steps were undertaken:

- Data Extraction and Standardization: Each dataset for 2012–2022 was downloaded with a consistent selection of variables. The column names were mapped to meaningful descriptors using ACS documentation.
- Merging and Temporal Structuring: The datasets were combined by matching the counties using the GEOID column. A YEAR column was added to track yearly changes, forming a longitudinal dataset.
- Feature Engineering and Derived Metrics: Several derived metrics were computed to enhance the analysis:
  - Poverty indicators:
    - $\text{PERCENT\_POPULATION\_BELOW\_POVERTY} = \frac{\text{B17001\_002E}}{\text{B01003\_001E}} \times 100$
    - $\text{PERCENT\_MALE\_IN\_POVERTY} = \frac{\text{B17001\_003E}}{\text{B01003\_001E}} \times 100$
    - $\text{PERCENT\_FEMALE\_IN\_POVERTY} = \frac{\text{B17001\_017E}}{\text{B01003\_001E}} \times 100$
  - Economic indicators:
    - $\text{UNEMPLOYMENT\_RATE} = \frac{\text{B23025\_005E}}{\text{B23025\_002E}} \times 100$
  - Demographic composition:
    - $\text{PERCENT\_WHITE} = \frac{\text{B02001\_002E}}{\text{B01003\_001E}} \times 100$
    - $\text{PERCENT\_BLACK} = \frac{\text{B02001\_003E}}{\text{B01003\_001E}} \times 100$
    - $\text{PERCENT\_AMERICAN\_INDIAN} = \frac{\text{B02001\_004E}}{\text{B01003\_001E}} \times 100$
    - $\text{PERCENT\_ASIAN} = \frac{\text{B02001\_005E}}{\text{B01003\_001E}} \times 100$
- Defining Poverty Status: Counties were classified as "poor" or "not poor" according to the following thresholds:
  - A county is considered Poor if at least 20% of its population is below the poverty line ($\text{PERCENT\_POPULATION\_BELOW\_POVERTY} \geq 20$) or if unemployment is 8% or higher ($\text{UNEMPLOYMENT\_RATE} \geq 8$).
  - Otherwise, it is labeled as Not Poor.

Appendix A.8. Final Dataset

The final dataset includes the following columns:

- Identifiers: Year, GEOID, county name (Name)
- Economic and poverty indicators: Total population; percent population below poverty; percent male in poverty; percent female in poverty; unemployment rate; median housing value
- Demographics: percent White; percent Black; percent American Indian; percent Asian
- Poverty classification: poor status (binary indicator)

This dataset was explicitly curated to support the GroupView framework, enabling a dynamic exploration of poverty trends, identifying persistent poverty regions, and analyzing racial and economic disparities in US counties.

Appendix A.9. Summary of Data Sources

Table A1 provides an overview of the dataset, including the number of counties covered each year and the average poverty rate.

Table A1. Summary of data sources by year, including the number of unique counties and the average poverty rate across US counties.

Year	Unique Counties	Average Poverty Rate (%)
2012	3221	16.49
2013	3221	16.84
2014	3220	16.96
2015	3220	16.86
2016	3220	16.57
2017	3220	16.17
2018	3220	15.80
2019	3220	15.31
2020	3221	14.82
2021	3221	14.64
2022	3222	14.57

Appendix A.10. Diabetes Data

The dataset used for analyzing diabetes prediction was sourced from a publicly available Kaggle dataset: Pima Indians Diabetes Database [38]. The original dataset consists of 768 entries and nine columns capturing various health-related parameters for individuals. To extend its relevance to a model-training scenario, the dataset was utilized to train a two-layer neural network for diabetes classification, and the training progress was recorded.

Appendix A.11. Data Sources

Pima Indians Diabetes Database (Kaggle). Variables: Pregnancies (number of pregnancies), Glucose (plasma glucose concentration), BloodPressure (diastolic blood pressure, mm Hg), SkinThickness (triceps skinfold thickness, mm), Insulin (2-hour serum insulin, μU/mL), BMI (body mass index), DiabetesPedigreeFunction (family-history score), Age (years), and Outcome (binary diabetes diagnosis).

Appendix A.12. Processing Steps

- Loading the Dataset:
  - The dataset was loaded using pandas, with no missing values imputed, as the dataset was complete.
- Feature Scaling:
  - Continuous variables such as Glucose, BloodPressure, BMI, and others were normalized for efficient model training.
- **Splitting the Data:**
  - The data were split into training and test sets with an 80:20 ratio to evaluate model performance.
- **Augmenting for Model Training:**
  - Each training epoch's progress was recorded, tracking metrics like accuracy and loss over 100 iterations. This generated an additional time-indexed column, Iteration_Time.

Appendix A.13. Final Dataset

The final dataset, tailored for integration with the GroupView framework, contains the following key columns:

- Pregnancies, Glucose, BloodPressure, SkinThickness, Insulin, BMI, Age, Outcome, DiabetesPedigreeFunction, and Iteration_Time.

This prepared dataset was tailored explicitly for tracking model training performance and exploring training progress patterns using the GroupView framework.

Appendix B. Survey Results

The study measured participants' familiarity with data visualization tools and their experiences using the framework. A majority of the participants (75%) reported being very familiar with data visualization tools, while the remaining (25%) described themselves as somewhat familiar. The usability questions were categorized under two constructs: 'Ease of Use,' which ranged from 'Very Difficult' to 'Very Easy,' and 'Usefulness,' which ranged from 'Not Helpful' to 'Very Helpful.' Responses were analyzed on their original 1–5 Likert scale and grouped under their respective constructs for interpretability. These constructs are visualized using Likert-style charts to illustrate the distribution of responses and provide a clearer understanding of participant feedback.

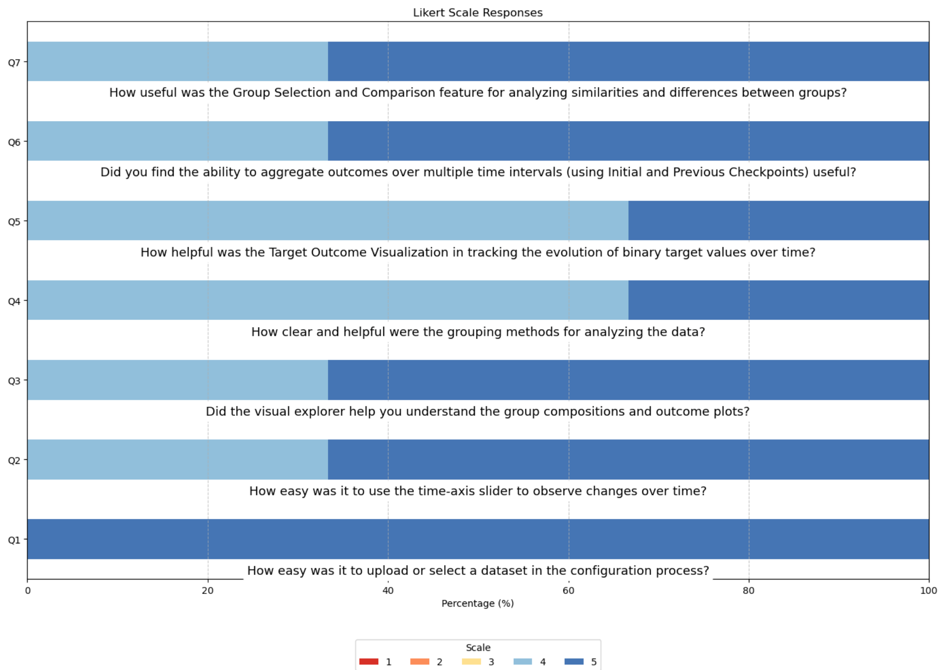

Figure A1. Likert scale responses for selected questions.

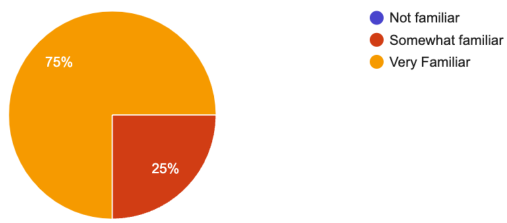

Figure A2. Familiarity with data visualization tools.

References

1. Chetty, R.; Friedman, J.N.; Hendren, N.; Jones, M.R.; Porter, S.R. *The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility*; Technical report; National Bureau of Economic Research: Cambridge, MA, USA, 2018. [CrossRef]
2. Alkire, S.; Foster, J. Counting and multidimensional poverty measurement. *J. Public Econ.* **2011**, *95*, 476–487. [CrossRef]
3. Narayan, A.; Van der Weide, R.; Cojocaru, A.; Lakner, C.; Redaelli, S.; Mahler, D.G.; Ramasubbaiah, R.G.N.; Thewissen, S. *Fair Progress?: Economic Mobility Across Generations Around the World*; World Bank Publications: Washington, DC, USA, 2018.
4. Gupta, S.; Hanssens, D.; Hardie, B.; Kahn, W.; Kumar, V.; Lin, N.; Ravishanker, N.; Sriram, S. Modeling customer lifetime value. *J. Serv. Res.* **2006**, *9*, 139–155. [CrossRef]
5. Weiss, L. Egocentric processing: The advantages of person-related features in consumers' product decisions. *J. Consum. Res.* **2022**, *49*, 288–311. [CrossRef]
6. Zhang, Z.; Gotz, D.; Perer, A. Interactive visual patient cohort analysis. In Proceedings of the IEEE VisWeek Workshop on Visual Analytics in Health Care, Seattle, WA, USA, 17 October 2012.
7. Shin, G.; McLean, S.; Hu, J.; Gotz, D. Visualizing temporal patterns by clustering patients. In Proceedings of the Visual Analytics in Healthcare, Washington, DC, USA, 15 November 2014.
8. Romm, K.F.; Metzger, A.; Gentzler, A.L.; Turiano, N.A. Transitions in risk-behavior profiles among first-year college students. *J. Am. Coll. Health* **2022**, *70*, 2210–2219. [CrossRef]
9. Pereira, H.M.; Leadley, P.W.; Proença, V.; Alkemade, R.; Scharlemann, J.P.; Fernandez-Manjarrés, J.F.; Araújo, M.B.; Balvanera, P.; Biggs, R.; Cheung, W.W.; et al. Scenarios for global biodiversity in the 21st century. *Science* **2010**, *330*, 1496–1501. [CrossRef]
10. Reynolds, M.D.; Sullivan, B.L.; Hallstein, E.; Matsumoto, S.; Kelling, S.; Merrifield, M.; Fink, D.; Johnston, A.; Hochachka, W.M.; Bruns, N.E.; et al. Dynamic conservation for migratory species. *Sci. Adv.* **2017**, *3*, e1700707. [CrossRef]
11. Vehlow, C.; Beck, F.; Weiskopf, D. Visualizing Group Structures in Graphs: A Survey. *Computer Graphics Forum* **2017**, *36*, 201–225. [CrossRef]
12. Hadlak, S.; Tominski, C.; Schulz, H.J.; Schumann, H. Visualization of attributed hierarchical structures in a spatiotemporal context. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 1497–1513. [CrossRef]
13. Riehmann, P.; Hanfler, M.; Froehlich, B. Interactive sankey diagrams. In Proceedings of the IEEE Symposium on Information Visualization, Minneapolis, MN, USA, 23–25 October 2005; INFOVIS 2005; IEEE: Piscataway, NJ, USA, 2005; pp. 233–240. [CrossRef]
14. Rosvall, M.; Bergstrom, C.T. Mapping change in large networks. *PLoS ONE* **2010**, *5*, e8694. [CrossRef]
15. Xu, Z.; Zhou, B.; Yang, Z.; Yuan, X.; Zhang, Y.; Lu, Q. NeatSankey: Sankey diagrams with improved readability based on node positioning and edge bundling. *Comput. Graph.* **2023**, *113*, 10–20. [CrossRef]
16. Bach, B.; Pietriga, E.; Fekete, J.D. Graphdiaries: Animated transitions andtemporal navigation for dynamic networks. *IEEE Trans. Vis. Comput. Graph.* **2013**, *20*, 740–754. [CrossRef]
17. MacQueen, J. Some methods for classification and analysis of multivariate observations. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA, 21 June–18 July 1965 and 27 December 1965–7 January 1966; Volume 1, pp. 281–297.
18. Jain, A.K. Data clustering: 50 years beyond K-means. *Pattern Recognit. Lett.* **2010**, *31*, 651–666. [CrossRef]
19. Huang, Z. Extensions to the k-means algorithm for clustering large data sets with categorical values. *Data Min. Knowl. Discov.* **1998**, *2*, 283–304. [CrossRef]
20. Andrienko, G.; Andrienko, N.; Rinzivillo, S.; Nanni, M.; Pedreschi, D.; Giannotti, F. Interactive visual clustering of large collections of trajectories. In Proceedings of the 2009 IEEE Symposium on Visual Analytics Science and Technology, Atlantic City, NJ, USA, 12–13 October 2009; IEEE: Piscataway, NJ, USA, 2009; pp. 3–10. [CrossRef]
21. Kang, H.; Getoor, L.; Singh, L. Visual analysis of dynamic group membership in temporal social networks. *ACM SIGKDD Explor. Newsl.* **2007**, *9*, 13–21. [CrossRef]
22. Moshkovitz, M.; Dasgupta, S.; Rashtchian, C.; Frost, N. Explainable k-means and k-medians clustering. In Proceedings of the 37th International Conference on Machine Learning (ICML 2020), Online, 13–18 July 2020; Volume 119, pp. 7055–7065. Available online: http://proceedings.mlr.press/v119/moshkovitz20a.html (accessed on 15 August 2025).
23. Wilkinson, L. *The Grammar of Graphics*, 2nd ed.; Springer: New York, NY, USA, 2005. [CrossRef]
24. Guo, D. Flow mapping and multivariate visualization of large spatial interaction data. *IEEE Trans. Vis. Comput. Graph.* **2009**, *15*, 1041–1048. [CrossRef]
25. Jiang, T.; Hou, Y.; Yang, J. Literature Review on the Development of Visualization Studies (2012–2022). *Eng. Proc.* **2023**, *38*, 89. [CrossRef]
26. Lupton, R.C.; Allwood, J.M. Hybrid Sankey diagrams: Visual analysis of multidimensional data for understanding resource use. *Resour. Conserv. Recycl.* **2017**, *124*, 141–151. [CrossRef]

27. Poddar, M.; Sohns, J.T.; Beck, F. Not Just Alluvial: Towards a More Comprehensive Visual Analysis of Data Partition Sequences. In Proceedings of the 29th International Symposium on Vision, Modeling, and Visualization, Munich, Germany, 10–13 September 2024; The Eurographics Association, 2024.
28. Porter, W.P.; Murphy, C.P.; Williams, D.R.; O'Handley, B.J.; Wang, C. Hierarchical Sankey Diagram: Design and Evaluation. In Proceedings of the Advances in Visual Computing: 16th International Symposium, ISVC 2021, Virtual Event, 4–6 October 2021; Proceedings, Part II; Springer: Cham, Switzerland, 2021; pp. 386–397.
29. Ortigossa, E.S.; Dias, F.F.; Nascimento, D.C.d. Getting over high-dimensionality: How multidimensional projection methods can assist data science. *Appl. Sci.* **2022**, *12*, 6799. [CrossRef]
30. Huang, Z. Clustering large data sets with mixed numeric and categorical values. In Proceedings of the 1st Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), Singapore, 23–24 February 1997; pp. 21–34.
31. Taylor, L.; Gupta, V.; Jung, K. Leveraging Visualization and Machine Learning Techniques in Education: A Case Study of K-12 State Assessment Data. *Multimodal Technol. Interact.* **2024**, *8*, 28. [CrossRef]
32. Quispel, A.; Maes, A.; Schilperoord, J. Aesthetics and Clarity in Information Visualization: The Designer's Perspective. *Arts* **2018**, *7*, 72. [CrossRef]
33. Ziraki, N.; Bosaghzadeh, A.; Dornaika, F. Semi-supervised Learning for Multi-View Data Classification and Visualization. *Information* **2024**, *15*, 421. [CrossRef]
34. Behzad, T.; Singh, M.K.; Ripa, A.J.; Mueller, K. FairPlay: A Collaborative Approach to Mitigate Bias in Datasets for Improved AI Fairness. *Proc. ACM Hum.-Comput. Interact.* **2025**, *9*, CSCW084. [CrossRef]
35. Nielsen, J.; Landauer, T.K. A mathematical model of the finding of usability problems. In Proceedings of the INTERACT'93 and CHI'93 Conference on Human Factors in Computing Systems, Amsterdam, The Netherlands, 24–29 April 1993; pp. 206–213.
36. Behzad, T.; Singh, M.K.; Ripa, A.J.; Mueller, K. FairPlay Dataset: Consensus-Driven Data Debiasing for Hiring Decisions. 2024. Available online: https://github.com/mitsvision2/FairPlay-User-Study-Data (accessed on 18 December 2024).
37. United States Census Bureau. American Community Survey (ACS). 2024. Available online: https://www.census.gov/programs-surveys/acs (accessed on 22 December 2024).
38. National Institute of Diabetes and Digestive and Kidney Diseases. Pima Indians Diabetes Database. 2020. Available online: https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database (accessed on 18 December 2024).