OmniViz - A Coordinated Multi-View Dashboard for Validating Radviz

Giuseppe Fosci, Alessio Rago, Emilio Martino
Sapienza University of Rome
https://visualanalyticsprojectradviz.netlify.app/

Abstract

The analysis of complex tabular datasets poses a significant challenge, as no single visualization can capture every detail. In response, we propose a web-based dashboard that arranges multiple visualization techniques for a comprehensive exploration.

The centerpiece is Radviz, a powerful but often misinterpreted technique. This paper presents a visual analytics dashboard where Radviz is not used in isolation, but is contextualized and validated against a suite of traditional visualizations.

By placing these charts in a coordinated system, we empower users to critically examine the insights suggested by Radviz, understand the strengths and weaknesses of each view based on data type and volume, and build a more robust understanding of their data.

1 Introduction

Previous work has already highlighted the potential of Radviz for multivariate data exploration [2]. General proprieties of the technique and its suitability for identifying patterns in high-dimensional datasets have been well examined. Building on those findings, this paper extends the discussion by embedding Radviz within a standard visual analytics environment, a dashboard, allowing it to be accompanied by complementary views. The dashboard is designed for flexibility, placing all controls within a collapsible side panel. The user's journey begins with data loading, either from a list of pre-loaded examples or by uploading a custom file adhering to the format in Section 2.

Once loaded, the dashboard offers granular control over the analysis. Users can select which dimensions to include via checkboxes and adjust the number of data rows (tuples) with a slider, observing how all charts update in real-time.

2 Data Input and Format Requirements

The system accepts data in the Comma-Separated Values (CSV) format. To ensure successful parsing, the file must adhere to a specific structure.

2.1 Structural Rules

1. **Mandatory Header Row:** The first line must be a header containing column names.
2. **Identifier Column:** Exactly one column must contain unique, textual identifiers. The system automatically identifies this column.
3. **Numeric Attribute Columns:** All other columns must contain numeric data. Non-numeric or empty cells are interpreted as zero.

Name	Safety	Price	Horsepower
Toyota Camry	5	24000	203
Honda Civic	5	22000	158
Ford Mustang	4	27000	310

Table 1: Representation of data in a valid CSV file.

2.2 Data Classification for Radviz Analysis

The paper "To RadViz or not to RadViz" posits that the effectiveness of a Radviz plot is fundamentally dependent on the nature of the input data. To address the visualization's known issues, the authors propose a data classification scheme based on how normalization affects the integrity of the analysis. This system is founded on two essential concepts:

- **Real Domain:** The theoretical and complete range of possible values for a given attribute, defined by external constraints. For instance, the Real Domain for Italian university grades is invariably [18, 30].
- **Active Domain:** The actual range of values, from minimum to maximum, present for an attribute within the specific dataset under analysis. Standard min-max normalization typically operates on

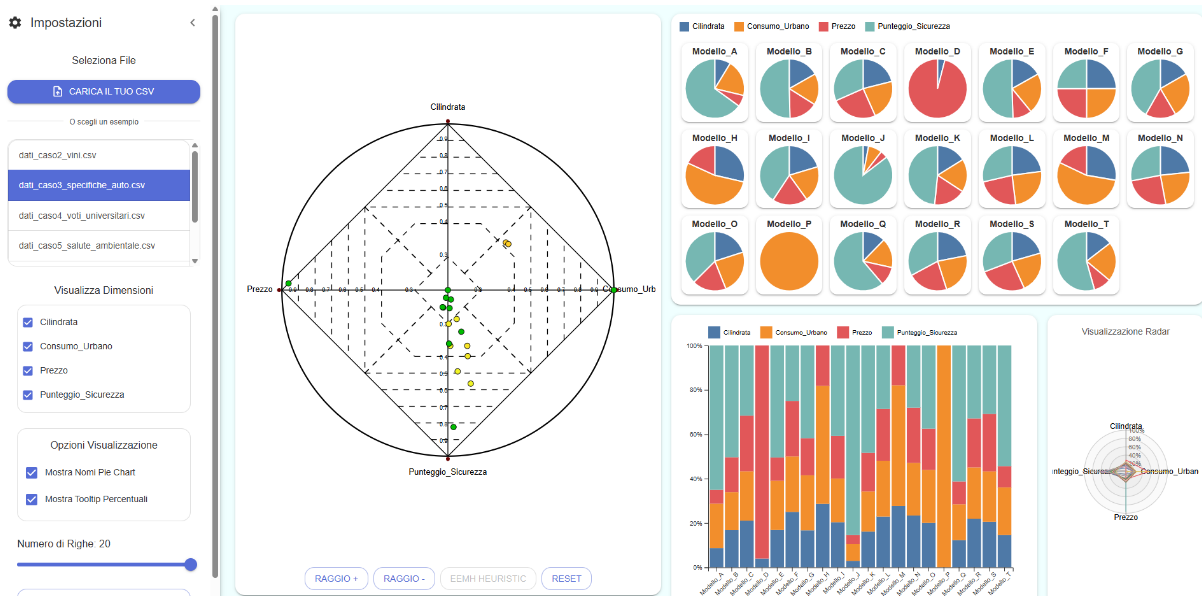

Figure 1: The OmniViz interface. On the left is the control panel. The main view contains the Radviz Chart (left), and the coordinated views: Pie Chart Grid (top right), Stacked Bar Chart and Radar Chart (bottom right).

this domain, which introduces instability as the domain can change when the dataset is updated.

Based on these definitions, the authors propose the following six-category classification to guide the use of Radviz.

- **Unconstrained Data:** This represents the most general and least suitable case for Radviz. The dataset contains attributes that are semantically unrelated and possess different, unaligned Active Domains (e.g., a wine dataset with attributes for acidity, alcohol content, and price). Applying a separate min-max normalization to each attribute renders their normalized values incomparable, leading to a visualization that is likely misleading.
- **Unconstrained Positive Data:** This is a specific sub-case of the previous category, where all attribute values are positive. Despite this constraint, the core problem remains: the attributes are still semantically unrelated with varying Active Domains. Consequently, it suffers from the same fundamental normalization issues and is equally ill-suited for reliable analysis in Radviz.
- **Data with Fixed Active Domains:** This category marks a significant improvement in stability. The user manually fixes the normalization range for each attribute, typically to its Real Domain (e.g., always normalizing grades within $[18, 30]$). This ensures that the visualization is **stable** and reproducible, as adding new data does not alter the positions of existing points. However, if these fixed domains differ across attributes, the challenge of comparing differently scaled values persists.
- **Data with Fixed and Common Active Domains:** This is a highly recommended scenario for using Radviz. It applies to datasets where all attributes are semantically related and share an identical, fixed normalization interval (e.g., a student's grades across multiple exams, all normalized within the ‘[18, 30]‘ range). The use of a common normalization formula, $p'_i = (p_i - \text{Min})/(\text{Max} - \text{Min})$, ensures that the **order of the original values is preserved** ($p'_i > p'_j \Rightarrow p_i > p_j$). This enables a meaningful and reliable analysis of attribute dominance.
- **Data with All Active Domains Fixed to [0, 1]:** In this excellent use case, all data values already exist within the ‘[0, 1]‘ interval. As a result, the min-max normalization step becomes an identity transformation ($p'_i = (p_i - 0)/(1 - 0) = p_i$) and can be omitted. The effects of normalization are completely removed, and the Radviz plot becomes a direct, undistorted representation of the original data.
- **Data that Represent Partitions:** This is the **ideal and intended use case** for Radviz. It is a significant sub-case of the previous category where not only do all values fall within ‘[0, 1]‘, but the values for each data point (row) also sum to 1. Ex-

amples include market share percentages, election results, or portfolio asset allocations. Since Radviz's primary strength is visualizing proportional relationships, it excels in this scenario, providing the most accurate and insightful analysis without any normalization-induced distortion.

3 System Architecture and Data Processing

OmniViz is built as a single-page application using React.js, with D3.js for visualization.

It should be noted that the Radviz chart is not a standard component within the D3.js library; its implementation was sourced from the A.W.A.R.E. (Advanced Visualizations and Visual Analytics Research) [3] working group of Sapienza University of Rome.

A central component, `App.jsx`, acts as an orchestrator, managing all application state. A key part of the architecture is an intelligent data processing pipeline that automatically classifies and conditionally normalizes the input data.

3.1 Numerical Data Classification

The pipeline's core logic is the `classifyNumerically` algorithm, which determines the intrinsic properties of the data to decide if normalization is needed.

It classifies datasets into one of three categories, in order of precedence: `partizionali` (if all rows sum to 1), `dominio_01` (if all values are in [0, 1]), or `generico` (all other cases).

3.2 Adaptive Data Normalization

The final stage of the pipeline is conditional and is the key to ensuring fair visual comparisons. Normalization is applied **only** to datasets classified as `generico`.

For all other data types, this step is bypassed, and the original, unmodified data is used.

When required, a standard min-max normalization is applied to each dimension independently:

$$v' = \frac{v - \min(D)}{\max(D) - \min(D)}$$

This process is fundamental for making attributes with disparate scales comparable. For instance, in a dataset of cars, an attribute like 'Safety' (with a range of 3 to 5) would be visually insignificant compared to 'Price' (with a range of 10,000 to 80,000) if raw values were used.

Normalization converts both to a [0, 1] scale, ensuring that the top safety rating (5) and the highest price (80,000) both map to a value of 1, giving them equal visual weight in radial visualizations like Radviz and Radar Chart.

4 Visualization Components

4.1 Radviz Chart

The Radviz Chart is the centerpiece of the system, designed to provide a high-level overview of the entire dataset. It excels at displaying a large number of data points, making it an excellent tool for identifying global patterns, clusters, and outliers, thereby helping users generate initial hypotheses (e.g., "Why are these points grouped together?").

As the most scalable view in the dashboard, it is capable of rendering thousands of points to reveal the underlying structure of the data.

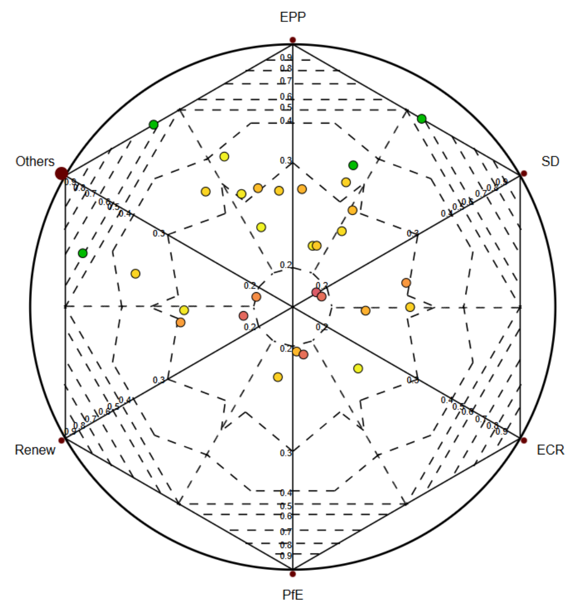

Figure 2: An example of the Radviz visualization. Data points (colored dots) are positioned inside a circular space based on the pull of dimensional anchors arranged on the circumference. Points closer to an anchor are more strongly influenced by that dimension.

Returning to the example of cars, suppose we notice a distinct cluster of data points positioned close to the dimensional anchor *Price* and simultaneously distant from the anchor *Urban Consumption*. This arrangement suggests an initial hypothesis: the points within this cluster represent luxury or sports vehicles, characterized by a high price and lower urban fuel efficiency. This would be more difficult to notice than, for example, a piechart and a stackedBarChart.

4.2 Radar Chart

The Radar Chart serves as the primary tool for detailed analysis and hypothesis validation. When a user selects one or more data points in another view, the Radar Chart offers a detailed profile of those items, plotting their values along explicit dimensional axes. This makes it the ideal component for confirming or refuting hypotheses about dimensional influence that were formed while observing the Radviz plot. However, it has the worst scalability of all the views and is best used for the detailed comparison of a very small set of items (fewer than 5).

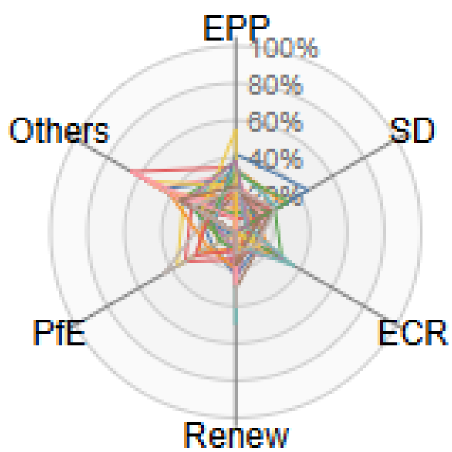

Figure 3: The Radar Chart displays selected data points as polygons. Each axis, radiating from the center, represents a dimension. The vertices of a polygon indicate the value of a data point for each dimension, making it ideal for profile analysis and direct comparison of a few items.

4.3 Stacked Bar Chart

This chart provides a compositional overview of all visible items, allowing for comparisons of how different attributes contribute to each data point. The visual encoding of the chart is adaptive and depends on the data type:

- For `partizionali` and `dominio_01` data, the segments in each bar represent the original, unmodified values, thus preserving their absolute meaning and proportions.
- For `generico` data, the segments represent the normalized values. This reveals the internal composition of a data point after its attributes have been brought to a common scale.

The Stacked Bar Chart scales moderately well and is effective for comparing the composition of up to 50-100 items.

4.4 Pie Chart Grid

This view employs a "small multiples" approach, dedicating an individual pie chart to each data item to show its proportional breakdown. By placing these small charts in a grid, it allows for a direct comparison of the internal distribution of attributes across different items. While effective for analyzing small, selected subsets of data (fewer than 15 tuples), this view suffers from poor visual scalability, as the charts become too small and cluttered to be legible with a larger number of items.

5 Discussion and Evaluation

5.1 Strengths of the Coordinated Approach

The primary strength of OmniViz is its ability to build user confidence through a tight loop of hypothesis and validation. A user can form a hypothesis in Radviz, immediately click on the relevant points, and see this hypothesis confirmed or refuted in the Radar Chart's explicit axes.

5.2 Analytical Workflow Synergy

The complementary features and scalability of each view enable a powerful analytical workflow: a user can start with a large-scale overview in **Radviz**, reduce the data scope to make the **Stacked Bar Chart** readable, and finally select a few items to validate hypotheses with precision in the **Radar Chart**.

6 Insights from Coordinated Views

A significant contribution of the OmniViz dashboard lies in its ability to facilitate a seamless analytical workflow, guiding the user from a high-level, exploratory overview to a detailed, validated understanding of the data.

This synergy between the different visualization components allows for the discovery of nuanced insights that would be difficult to discern from a single chart alone. The coordinated system is designed to foster a cycle of hypothesis generation and iterative refinement, leading to more robust and trustworthy conclusions. The following examples, derived from real-world datasets, illustrate this process.

6.1 Trade-offs in Automotive Data (Dataset 3)

When analyzing a dataset of vehicles, the Radviz chart might immediately reveal a distinct cluster of data points positioned between the dimensional anchors for *Safety* and *Urban Consumption*. This spatial arrangement suggests a direct relationship: cars in this group tend to be safer but also have higher fuel consumption, representing a clear design trade-off. The position of the cluster, relatively distant from the *Cilindrata* anchor, also implies that engine size is a less significant factor for this particular group of vehicles.

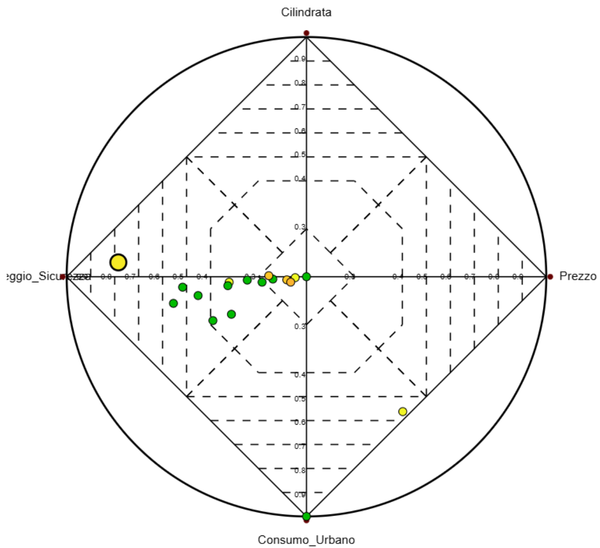

Figure 4: There is a cluster between the safety score and the urban consumption anchors, directly showing that the engine size is not much relevant and that safe cars often have a high consumption, representing a clear trade-off

During this same exploration, a single outlier might appear between the *PunteggioSicurezza* and *Cilindrata* anchors. Its unique position could suggest an ideal vehicle: one with a high safety score, powerful engine, and, by its distance from other anchors, potentially a low price and low consumption.

However, if the data point is colored yellow (a color that might be used to flag unreliable data), this visual cue provides critical context. The user can immediately infer that while this point seems interesting, it should be treated with caution and may not be a valid data point for analysis.

6.2 Performance Patterns in Student Grades (Dataset 4)

In a dataset of student evaluations, Radviz can offer a clear, holistic view of grade distribution. In one scenario, the points may appear widely dispersed, indicating that there are no well-defined clusters of students with similar performance profiles across all subjects.

The values are distributed evenly, suggesting a diverse range of academic strengths and weaknesses. Despite this general dispersion, an analyst might notice a subtle gravitational pull of the majority of data points towards the *Programming* anchor. At first glance, this suggests that, on average, students performed better in this subject compared to others. This kind of high-level trend, which captures the overall "center of mass" of the dataset, is an insight that Radviz is uniquely suited to provide.

Such information is not as easily or immediately understood by observing the Pie Chart Grid or the Stacked Bar Chart, which are more effective for comparing the specific grade compositions of individual students rather than identifying aggregate patterns across the entire cohort.

6.3 Evaluating Attribute Dominance in Wine Data

The dashboard also excels at helping users critically assess the data itself. When visualizing a wine dataset, Radviz might show that nearly all data points are heavily pulled towards the *Density* anchor, indicating its predominance over other attributes like *Acidity* or *Alcohol*.

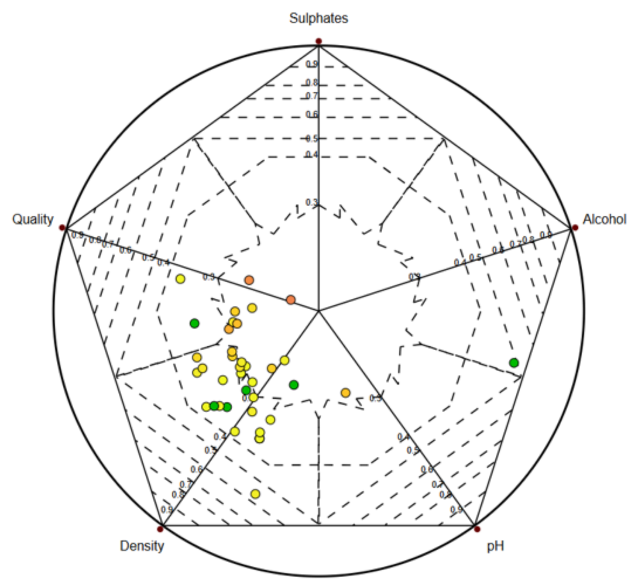

Figure 5: he Radviz visualization of the wine dataset clearly shows the predominance of the *Density* attribute. Nearly all data points cluster near its dimensional anchor, overwhelming the influence of other factors

However, by selecting these points and examining them in the other views, a crucial insight emerges: this high density value is not correlated with the wine's *Quality*. In fact, it is a standard, nearly constant value across almost all samples. In this case, the coordinated system demonstrates a limitation of Radviz when faced with a non-discriminatory attribute. The visualization correctly shows the attribute's high numerical value, but it is the context from other charts—like the Pie Chart Grid and Stacked Bar Chart, which clearly show density as a consistently large slice or bar segment in every single item—that reveals this attribute is not useful for analysis.

This confirms that for this specific question, Radviz does not add significant value beyond what simpler, well-known charts can provide, thereby validating the visualization choice itself.

7 Radviz Optimization with EEMH

To enhance Radviz, we integrate the **Effectiveness Error Minimization Heuristic (EEMH)**[1].

Since Radviz's layout is sensitive to anchor ordering, this algorithm computationally seeks an optimal circular arrangement of the DAs to minimize visual clutter. When applied, both the Radviz anchors and the Radar Chart axes perform a synchronized animation to their new positions.

8 Conclusion

OmniViz is an effective demonstration of how a coordinated, multi-view dashboard can be used not only to analyze data, but to critically evaluate the visualizations themselves. By combining a high-scalability overview technique (Radviz) with low-scalability detail-on-demand views, the system allows users to form hypotheses at a large scale and validate them with precision at a small scale.

References

[1] M. Angelini et al. "Effectiveness Error: Measuring and Improving RadViz Visual Effectiveness". In: *IEEE Transactions on visualization and computer graphics* (2021).

[2] Mauro Ficorella et al. "To RadViz or not to RadViz". In: *Eurographics Conference on Visualization (EuroVis)* (2023).

[3] La Sapienza University. *A.W.A.R.E.* Accessed: 2025-09-08. URL: http://aware.diag.uniroma1.it/.