

Design Rationale: Coordinated Scatterplot and Parallel Coordinates for Housing Data

Prepared by Ludovic TEMGOUA ABANDA N.

November 3, 2025

Overview

This report describes the visual design of a React+D3 application that presents a housing dataset through two linked visualizations: an interactive scatterplot and a parallel coordinates plot. The goal is to support exploratory analysis tasks such as identifying relationships between numeric attributes (e.g. price, area), comparing multi-attribute patterns across homes, and selecting subsets to inspect simultaneously in both views.

The following sections explain the dataset characteristics, the mapping between visual encodings and user tasks, implementation choices that affect the visualization, and the strengths and compromises of the overall design.

1 Data and user tasks

The dataset consists of tabular records, each representing one housing unit. Attributes include numeric fields such as price, area, bedrooms, bathrooms, stories, and parking spaces, plus occasional categorical or ordinal attributes. The dataset is moderate in size, typically a few hundred records which is large enough to reveal structure yet small enough to render efficiently in the browser.

From these properties, the main user goals are:

1. **T1:** Compare pairwise relationships and trends (e.g. price vs. area).
2. **T2:** Detect multi-attribute patterns or clusters across many variables.
3. **T3:** Select and highlight subsets of interest across coordinated views.
4. **T4:** Inspect detailed values and identify outliers for individual records.

2 Visual encoding choices

2.1 Scatterplot

Encoding. The scatterplot maps two quantitative attributes to the X and Y axes (e.g. area on X and price on Y). Each record is drawn as a small circle of constant radius to minimize occlusion. Color and stroke are reserved for highlighting selected or hovered items. Interaction includes a two-dimensional brush for selecting multiple points and tooltips for showing precise values.

Justification. Position is the most accurate visual channel for quantitative data, making the scatterplot ideal for observing correlations, trends, and outliers (T1). The brush interaction

supports flexible range selection and coordinated highlighting across views (T3). By keeping the mark design simple, users focus on spatial patterns rather than graphical complexity.

2.2 Parallel coordinates

Encoding. Each record appears as a polyline crossing a set of vertical axes, one per attribute. The vertical position of the line on each axis encodes the numeric value of that attribute. Linear scales are used for continuous variables, while ordinal scales handle integer or categorical ones. Line opacity and color indicate selection state: unselected lines are faint, while selected ones are emphasized. Each axis supports brushing to filter by attribute, and hovering reveals full attribute values via tooltip.

Justification. Parallel coordinates are a standard method for examining multivariate numerical data. They directly support pattern discovery (T2) by allowing users to trace each record across attributes and spot consistent relationships (for instance, high price aligning with large area and more bedrooms). Coordinated selection with the scatterplot lets users link two-variable trends to their full multidimensional profiles (T3).

2.3 Coordinated interaction and identity

A key design decision is to maintain a stable data identity by assigning each record a unique index used across both visualizations. This ensures that a selection made in one view correctly highlights the same records in the other. The synchronization reinforces the sense that both views represent a single, consistent dataset rather than independent charts.

3 Implementation notes

To keep the visualization responsive and visually clear, several practical choices were made:

- CSS classes manage selection and hover states efficiently without excessive DOM updates.
- Axis tick strategies are adapted to attribute type: `area` uses wide intervals for readability, while small-integer fields (stories, parking) use step-1 ticks.
- Tooltips appear near the hovered element and provide immediate numeric detail without obscuring nearby marks.
- The app maintains lightweight state in React for cross-view coordination, while D3 handles scales and drawing for performance.

4 Design strengths and trade-offs

4.1 Scatterplot

Strengths. The scatterplot provides high perceptual accuracy for comparing two quantitative variables. Its simplicity makes it familiar and intuitive, encouraging exploration even by novice users.

Limitations. It only shows two dimensions at once, so understanding higher-dimensional structure requires linking with the parallel coordinates view. Overplotting can occur in dense areas, mitigated by small mark size and transparency.

4.2 Parallel coordinates

Strengths. Parallel coordinates excel at showing multi-attribute relationships and entire record profiles. Per-axis brushes enable detailed filtering, and coordination with the scatterplot provides both global and focused exploration paths.

Trade-offs. Visual clutter increases as more records are shown. Adjusting opacity helps, but axis order also affects perceived correlations. The current static order favors interpretability, while axis reordering could enhance flexibility in future iterations. Because parallel coordinates are less familiar to general audiences, subtle interaction cues (hover, highlighting, and tooltips) help users learn their logic quickly.

4.3 Color and highlighting

Color and stroke are used purely for selection and emphasis rather than encoding data variables. This avoids conflicting interpretations and ensures that highlighted items remain distinct across views. A limitation is that single-color highlighting restricts multi-group comparisons. Allowing several color-coded selection groups would enable side-by-side comparison of subsets—an avenue for future improvement.

5 Limitations and future work

Future extensions could include:

- Axis reordering and scaling options (linear/log/normalized) for advanced analysis.
- Adjustable opacity and tick density controls for user customization.
- Multi-color selection to compare several groups simultaneously.
- Export functions for saving selected subsets or images of the visualizations.

Conclusion

The coordinated scatterplot and parallel coordinates design effectively supports both pairwise and multi-dimensional exploration of housing data. The scatterplot offers precise quantitative comparison, while the parallel coordinates plot exposes broader attribute relationships. Through stable identity management and synchronized brushing, users can move fluidly between global and detailed perspectives.

The main trade-off is between visual simplicity and multi-dimensional expressiveness: although the combination introduces some visual clutter, careful use of opacity, color, and coordinated interaction maintains clarity. Overall, the design provides a robust, responsive environment for interactive data exploration within the browser.