Individual Assignment

Interactive Visualization of a Housing Dataset

Thomas Gantz[1]

[1]Université du Luxembourg, student number: 024746141f
e-mail: thomas.gantz.001@student.uni.lu
github.com/gantz-thomas-gantz/DataVis

Deadline: November 04, 2025

ABSTRACT

This work presents an interactive D3 + React visualization designed to help prospective homebuyers identify suitable properties. The interface features two linked views: a scatterplot mapping area against price for market overview, and a matrix organizing properties by bedroom and bathroom counts. In the matrix, circles represent individual properties with size encoding area and color encoding price per square meter. Coordinated interactions between views enable efficient property comparison and filtering while maintaining awareness of overall market patterns.

1. Design

The following section introduces the proposed task-driven visualization approach.

1.1. User Tasks and Requirements

The primary user is a prospective homebuyer seeking best value properties within specific constraints (budget, size, room count). This is a search task requiring efficient filtering and comparison and differs from exploratory or trend-discovery scenarios which aim for pattern recognition. The visualization supports three core tasks:

- **Price-value assessment:** Evaluate property value by comparing area against price, the most important decision criterion.
- **Room configuration filtering:** Filter by bedroom and bathroom counts, the second most important criterion.
- **Interactive refinement:** Iteratively refine searches by selecting regions of interest through linked interactions.

1.2. System Overview

The visualization consists of two coordinated views. The **scatterplot** (left) maps area (x-axis) against price (y-axis) for direct insight into price-value relationships. The **matrix** (right) organizes properties by bedroom (columns) and bathroom (rows) counts into discrete cells. Within cells, circles represent properties with size encoding area and color encoding price per square meter. Selections in either view filter and highlight the other, enabling rapid drill-down without context loss.

1.3. Scatterplot View

The scatterplot provides market overview by mapping each property's area against total price. Each point represents a single property. Users can brush (click and drag) to select rectangular regions, filtering the matrix to show only properties within the selected price-area range.

1.4. Matrix View

The matrix organizes properties into a grid where rows represent bathroom counts and columns represent bedroom counts. Each cell aggregates all properties matching that room configuration, making availability apparent.

Within each cell, circle size encodes area (larger circles = larger properties), helping identify properties with many rooms but limited space, or vice versa. Color encodes price per square meter to visually highlight cost-effective options.

1.5. Interaction Design

The visualization supports several interaction techniques:

- **Bidirectional filtering:** Rectangular brushing in the scatterplot filters the matrix by area-price range (coarse filtering). Clicking matrix cells filters the scatterplot for matching bedroom/bathroom configurations. Right-clicking in the matrix discards individual properties (fine filtering).
- **Hover details:** Hovering over circles displays detailed property information in the matrix and highlights the property's position in the scatterplot.
- **Navigation:** Back and forward buttons support iterative refinement of the property search.

2. Evaluation

This section evaluates the visualization based on (a) theoretical criteria established in the literature (Ghoniem 2025) and (b) a task-driven evaluation approach.

2.1. Design Principles

Visual encoding. The proposed visualization maps the data dimensions price, area, number of bedrooms, and number of bathrooms to the visual variable position. In addition, it maps price per square meter to color and area to size, which constitutes a natural visual metaphor, in the second matrix view.

All represented data are quantitative, either continuous or discrete. The use of position as the primary visual variable adheres to established standards of perceptual effectiveness. Size encoding is less precise, followed by color. However, assigning the more important data dimensions to the more effective visual variables is appropriate. Moreover, both area and color operate as pre-attentive visual features that enable rapid perception.

The encoding of price per square meter and area in the right matrix view remains open to discussion. Alternative visual variables such as length or angle could also be employed to represent these attributes.

Design process. The visualization design follows the established information visualization process. It is explicitly task-oriented, and the visualization was filtered to include only the most relevant criteria. The data, namely the housing properties, are coherently represented as the same geometric primitives in the form of points in both views. The difference between the two views lies in the specific visual attributes assigned to these primitives.

The chosen visualization is clear and effective. It conveys all relevant information contained in the dataset without introducing spurious visual patterns. The data can be readily perceived, allowing users to gain an overview first and to access further details on demand through zooming and filtering.

Multidimensional data representation. The dataset comprises multidimensional data with four or more dimensions. To address this, two parallel visualizations were implemented, complemented by small multiples in the right matrix view.

Standard alternatives such as parallel coordinates or dimensionality reduction techniques, for instance principal component analysis (PCA) or t-distributed stochastic neighbor embedding (t-SNE), were considered but deemed unsuitable as they were not aligned with the defined user task.

2.2. Task-Driven Evaluation

A user study is the most appropriate evaluation method for this task-driven design. The study would assess how effectively the visualization supports goal-directed analytical tasks, such as:

1. Identify the property with the lowest price within a specified area range.
2. Compare properties with similar prices but differing area and bedroom count.
3. Find the property with minimal area and maximal price satisfying specified bedroom/bathroom requirements and parking availability.

These goal-oriented tasks reflect realistic real-estate decision-making processes. The visualization should enable users to efficiently identify, compare, and filter properties through successive refinement steps.

A standard housing search website such as *SeLoger* would serve as a meaningful baseline, providing a familiar interface for comparison. Evaluating user performance across both systems would determine whether the proposed visualization offers advantages in efficiency and decision accuracy.

The anticipated workflow for the third task is illustrated below, beginning with an overview of all properties and proceeding through targeted filtering until the optimal choice is identified.

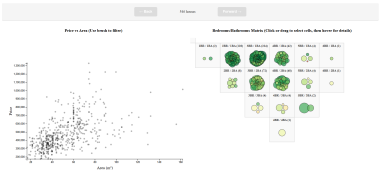

Fig. 1: Initial overview of the dataset.

Fig. 2: After narrowing down properties by area and budget.

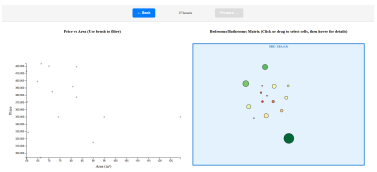

Fig. 3: After selecting the required number of rooms.

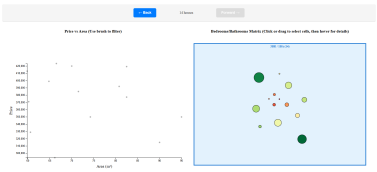

Fig. 4: After excluding one property without a parking spot.

3. Conclusion

3.1. Summary and Contributions

This work presents a task-driven visualization approach for real-estate property search. The main contributions are: (1) an analysis of user tasks and key data dimensions in property selection; (2) a coordinated multi-view design providing complementary visual perspectives beyond standard housing search websites; and (3) an interactive filtering framework supporting progressive refinement through constraint specification.

3.2. Limitations and Future Work

Several limitations need to be considered. The design has not yet been validated through user studies. Scalability to large datasets remains untested and may require aggregation or hierarchical filtering. In particular, the matrix view becomes cluttered with more than thirty data points.

Future work could explore alternative visual variables such as length or angle. A trend line in the scatter plot could help identify favorable properties offering larger area at lower price. Additionally, the use of green and red to indicate favorable and unfavorable options may present accessibility challenges for color-blind users.

References

Ghoniem, M. 2025, Lecture Notes on High Performance Data Analytics and Visualisation, Luxembourg Institute of Science and Technology.