



Milestone 2: Visualization Design Plan

VisualEstate Group: Élise Boyer, Charles Girardot, Gauthier Nielly

April 15, 2025

1 Project Overview and Objectives

Our project aims to visualize meaningful insights from real estate data across three major European cities: Berlin, London, and Madrid. The goal is to provide users, whether investors exploring unfamiliar markets or travelers searching for the ideal place to live, with an intuitive way to identify the most important features in their future home. In this second milestone, we present our Minimum Viable Product (MVP) along with more advanced ideas. We detail the key visualizations we plan to implement, as well as the tools and course concepts required to develop them. All the sketches of the visualizations are included at the end of the report.

2 Project Implementation Roadmap

2.1 Minimum Viable Product (MVP)

The core of our MVP is a website that provides clear and meaningful insights through basic exploration and visualization of our real estate datasets.

The visualizations will include:

- **A tree map** for each city, showing what types of information the dataset contains. This will be based on graph structures (related to Lecture 10 on trees and graphs).
- **A 2D choropleth map** to visualize the price range across each city, and **a density map** to show the spatial distribution of properties. The choropleth map will be build with GeoJSON, connecting to Lecture 7 on map-based visualizations. Link to D3.js tutorial or Plotly
- **Scatter plots** of surface area versus price, grouped by neighborhood, to help identify local pricing patterns. Other comparisons could be made, such as the price vs. the number of bedrooms, or by features such as the balcony, the garage, and others. The features could be selected in the scatter plot window.
- **A correlation heatmap** to highlight the key factors influencing property prices in each city.

The main visualization library used will be **D3.js** or **plotly** for interactive visualisations.

The site will consist of:

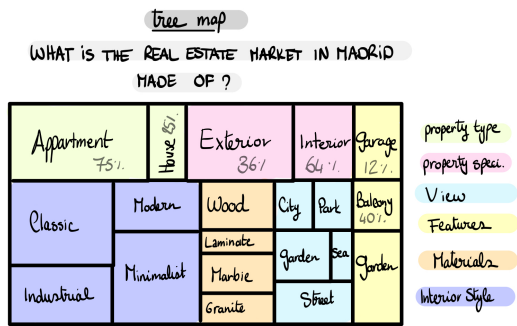
- A *homepage* providing general project information (e.g., which city is analyzed, how many data points, and what features are included).
- Separate pages for each city, presenting their respective visualizations and analyses.
- An *About* page detailing our methodology, data sources, limitations, and potential future improvements.

2.2 Additional Features and Enhancements

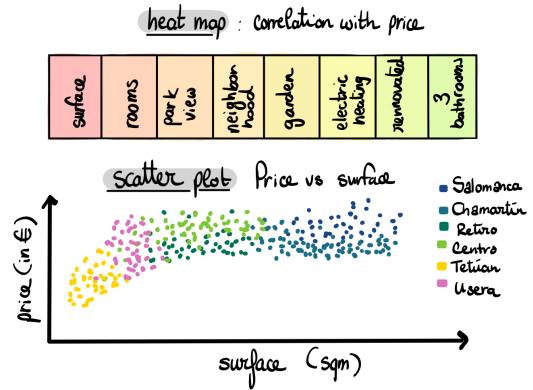
Even though these elements are not part of our MVP, we consider the site’s visual design and dynamic interactivity essential for usability and overall user engagement. While animations and polish are not strictly necessary to convey insights, they significantly enhance the user experience. To achieve this, we plan to use GSAP (GreenSock Animation Platform) for smooth and responsive animations throughout the website.

In addition to interactivity, we aim to incorporate more creative and original visualizations that present insights in a compelling and engaging way:

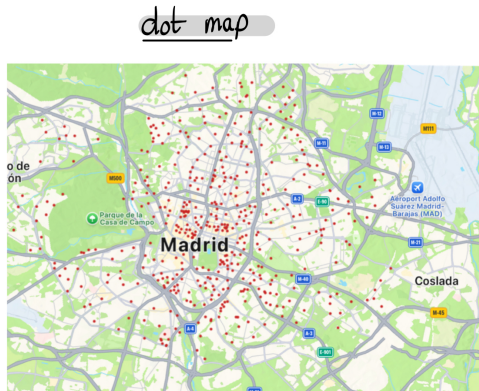
- **Feature-based Bar Plot:** As shown in Figure 1.d, users select desired features, and the bar chart displays the neighborhoods that best match, ranked by percentage.
- **Advanced 3D Maps:** Transforming standard 2D maps into an interactive 3D environment to instantly highlight more expensive neighborhoods, providing a clearer and more intuitive visualization of pricing data. Can be done using Mapbox GL or Deck GL.
- **Word Clouds:** As shown in Figure 1.c An intuitive way to summarize the most common features in each neighborhood, offering a qualitative overview of local trends.
- **Price Prediction Tool:** Allows users to input property details and receive an estimated price range based on our dataset—useful for buyers, sellers, or investors.
- **Addition of Neighborhood-Related Datasets:** This provides more detailed information about the surroundings of each property, enabling users to make decisions based not only on property features and price, but also on neighborhood amenities (e.g., bars, museums, parks, schools).



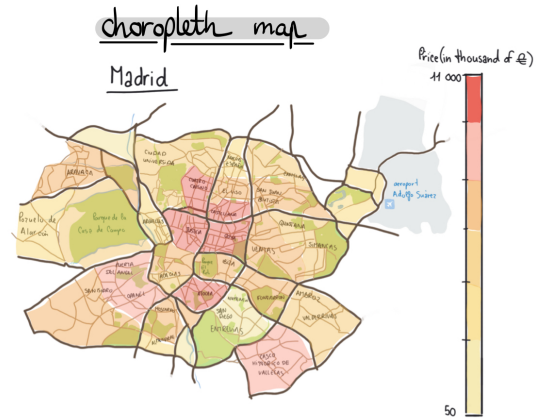
(a) **Treemap:** Depicts how the Madrid real estate market is composed by various property types and features. Each rectangle's size indicates the relative frequency or importance of that category.



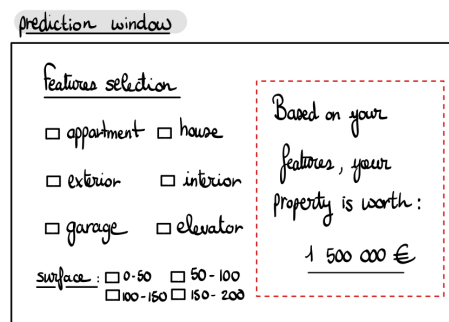
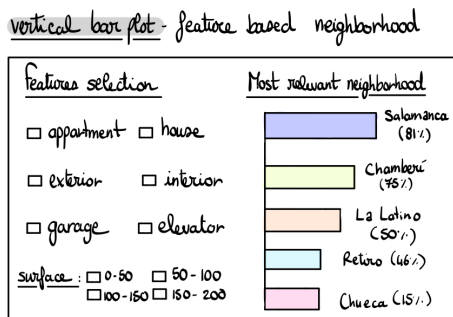
(b) **Heat Map + Scatter Plot:** Shows correlations between features (e.g., size, rooms) and price, alongside a scatter plot of price vs. surface area, color-coded by neighborhood.



(c) **Dot Map:** Displays property locations across Madrid. Each dot represents a listing, highlighting spatial distribution and density.



(d) **Choropleth Map:** Uses color shading to show average property prices by neighborhood, making the most and least expensive areas easy to spot.



(e) **Advanced Visualizations:** Depicts feature-based bar plots, a potential price-prediction interface, and other creative ways to explore amenities and neighborhood data for more immersive insights.

Figure 1: Overview of the primary visualizations for real estate data in Madrid. Each subfigure represents a unique technique to highlight data features, patterns, and spatial distributions.