Density of hotel accommodations in Madrid

This project focuses on analyzing the distribution and density of hotel accommodations in Madrid, providing valuable insights into the city's tourism landscape. By visualizing key metrics such as hotel density by postal code, the proportion of accommodation types, and identifying the most densely populated areas, this analysis aims to support urban planning, business strategies, and tourism development. The inclusion of interactive maps further enhances the accessibility and practical value of the results.

Tools and Libraries Used

The analysis was conducted using **R**, leveraging a variety of powerful libraries for data manipulation, visualization, and geospatial analysis:

- **Data manipulation and visualization:**
  - `ggplot2`, `tidyverse`, `dplyr`
- **Data import/export:**
  - `readxl`, `readr`, `openxlsx`
- **Geospatial analysis and mapping:**
  - `sf`, `tmap`, `leaflet`

Main Objective

The primary objective of this analysis is to explore and visualize the spatial distribution of hotel accommodations in Madrid. By identifying patterns and trends, the project provides actionable insights that can guide urban development policies, optimize tourism strategies, and inform investment decisions in the hospitality sector.

Data and Methodology

Data Sources and Processing

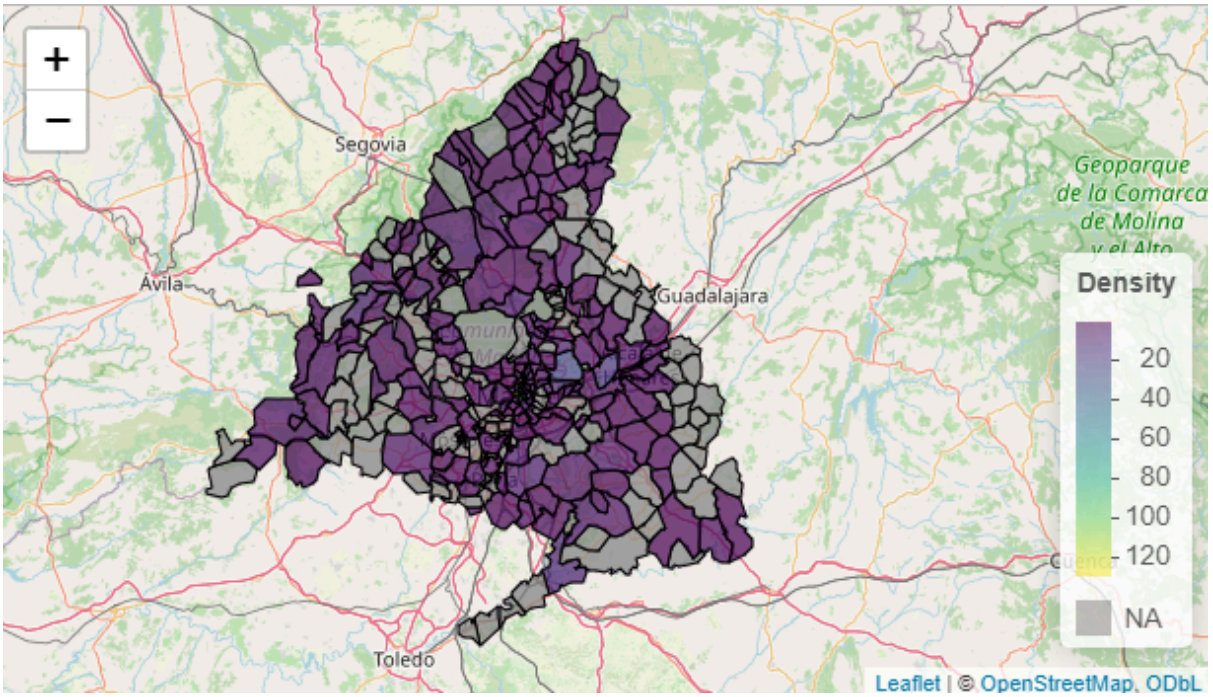
The dataset for this project was sourced from the Community of Madrid Open Data Platform, which provides detailed information on tourist accommodations in Madrid.

The initial steps involved cleaning the data to ensure consistency and accuracy. This included organizing postal codes, removing unnecessary spaces, and filtering the dataset to

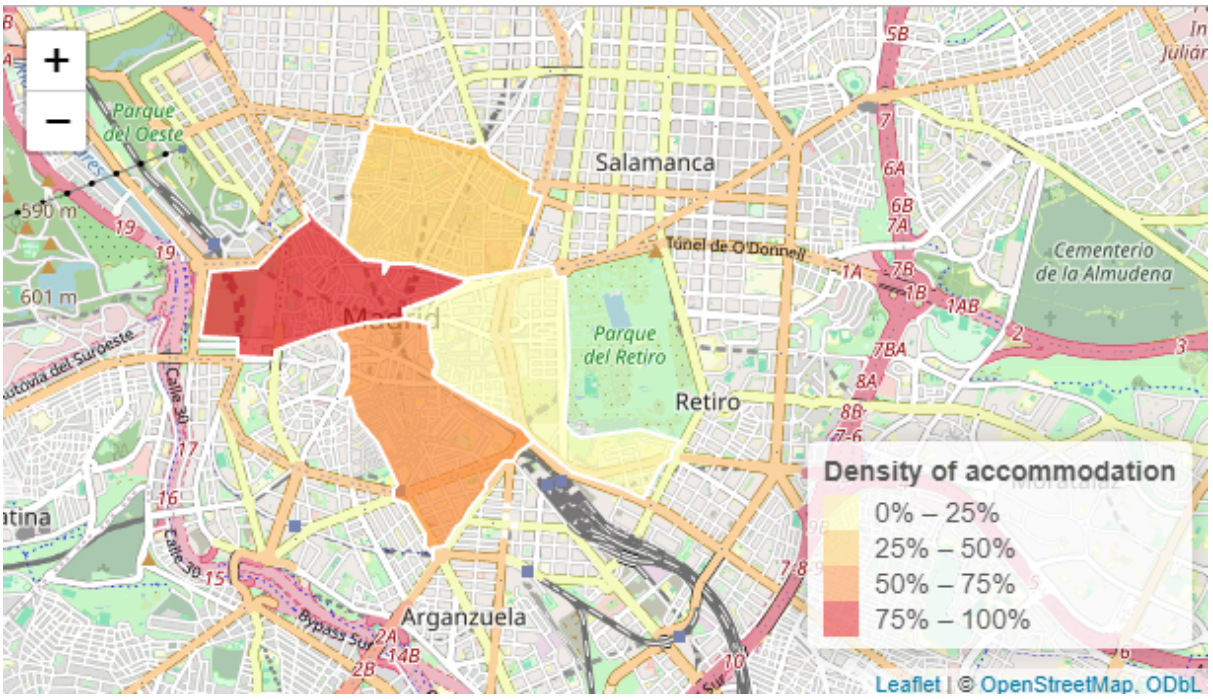
focus exclusively on relevant accommodation types: **hotels**, **hostels**, and **camping sites**. Hotels were further categorized into **hotels** and **apartment hotels** for more detailed analysis.

Techniques Used

1. **Data Cleaning:**
    - Ensured postal codes were properly formatted and removed extraneous characters.
    - Filtered the dataset to retain only key accommodation types (hotels, hostels, and camping sites).
2. **Data Manipulation and Analysis:**
    - Created an **interactive map** to visualize the hotel density across Madrid.
    - Filtered the data further to focus on the **five postal codes with the highest hotel density** for deeper analysis.
    - Generated a bar chart to highlight the **proportion of accommodation types**, providing insights into the most common categories.
3. **Geospatial Visualization:**
    - Developed a second **interactive map** that specifically highlights the **five densest areas**, offering a more focused view of high-density zones.



This map illustrates the distribution of hotel density across the Madrid region, depicted in shades of purple. The darker areas represent higher hotel density, while lighter areas indicate lower density. Gray regions are marked as "NA," likely denoting areas without relevant data or no hotels at all.

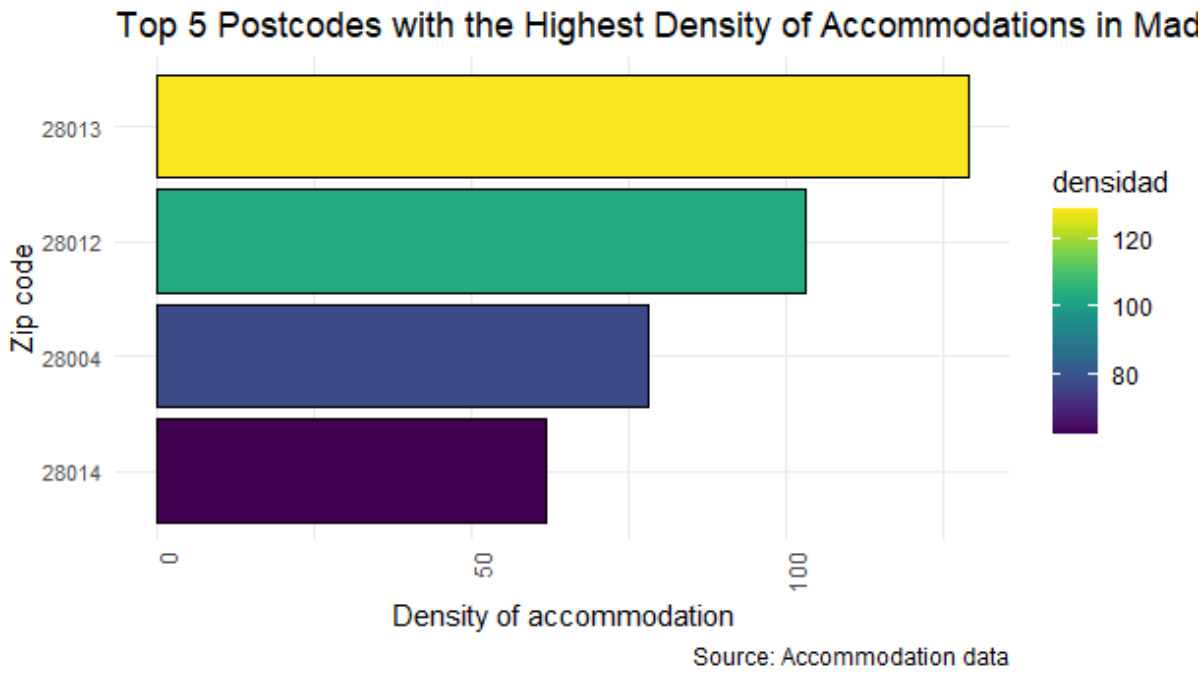


This map focuses on the central districts of Madrid, highlighting the density of accommodations within specific neighborhoods. The areas are color-coded based on density levels:

- **Red (75% - 100%)**: Highest density of accommodations, likely reflecting major tourist or business hubs.
- **Orange (50% - 75%)**: Moderate density, indicating areas with significant but not peak levels of accommodations.
- **Yellow (25% - 50%)**: Lower density, likely more residential or less tourist-oriented.
- **Light yellow (0% - 25%)**: Minimal accommodation presence, potentially quieter or less central areas.

Relevance:

1. **Tourism Hotspots**: The red zones likely correspond to areas like Gran Vía or Puerta del Sol, highlighting their importance for tourism and hospitality industries.
2. **Urban Planning**: This distribution informs urban developers and policymakers about overcrowding or the need for additional accommodations in underrepresented areas.
3. **Traveler Guidance**: The map helps tourists identify areas with a concentration of lodging options, aiding in travel planning and convenience.



This bar chart shows the top five postcodes in Madrid with the highest density of accommodations. Each bar represents a specific postal code, with colors indicating different density levels according to the color gradient provided.

Detailed Observations

1. **Diversity in Density**: While 28013 has the highest density, the differences between 28012 and 28014 are less pronounced, indicating a clustering of accommodations in areas central to Madrid's tourist and transport networks.
2. **Potential Link to Attractions**:
    - **28013** likely aligns with Sol and Gran Vía, renowned for entertainment, dining, and shopping.
    - **28012** could correspond to areas like Lavapiés or La Latina, known for cultural diversity, nightlife, and historical landmarks.
    - **28014** might represent the vicinity of Atocha (Madrid's main train station) and the Paseo del Prado, home to world-famous museums like the Prado Museum and Reina Sofía.
3. **Lower Rank Densities**: The postal codes in the lower part of the top five list still maintain high densities, showing that central Madrid is consistently dense in accommodations compared to the outskirts.
4. **Tourist Behavior Reflection**: These dense zones reveal patterns of tourist activity, as they correspond to areas rich in heritage, transportation connectivity, and vibrant nightlife.
5. **Infrastructure Load**: Higher density also implies these areas may face increased pressure on infrastructure, such as public transport, sanitation, and services tailored for tourists.

Broader Implications

- **Spillover Demand**: Adjacent neighborhoods might experience spillover effects, increasing demand for accommodations as the top zones approach saturation.
- **Zoning and Regulation Needs**: High density might necessitate zoning laws or stricter regulations on short-term rental accommodations to ensure balanced urban growth.
- **Seasonal Trends**: The chart reflects overall density but could vary seasonally; zones like 28014 (near Atocha) might see fluctuations tied to rail traffic and events.

The analysis reveals that the highest accommodation density is concentrated in Madrid's central areas, specifically in postal codes 28013, 28012, and 28014. These zones correspond to prominent tourist and commercial hubs such as Gran Vía, Sol, and Atocha. Their high density reflects their importance as centers for cultural attractions, dining, and transportation connectivity, which make them highly appealing for both tourists and business

travelers. In contrast, peripheral areas exhibit significantly lower densities, highlighting a clear disparity between the city center and its outskirts.

These patterns also align closely with tourist behavior. Visitors tend to prioritize accommodations near landmarks, vibrant nightlife, and public transport, reinforcing the central areas' dominance. However, this concentration creates a potential challenge as high-density zones face increased pressure on urban infrastructure, such as transport networks, sanitation, and tourist services. This underscores the need for effective urban planning to address the strain caused by the large influx of visitors to these areas.

Practical Applications

From a practical perspective, this analysis provides valuable insights for managing tourism and urban growth. Authorities could use this information to implement policies that promote lesser-known neighborhoods or peripheral zones as alternatives to the oversaturated city center. This would help alleviate overcrowding and distribute tourist flows more evenly across the city.

Urban planning efforts could also benefit by focusing on improving infrastructure in high-density zones, such as public transportation and waste management systems, to enhance both visitor experience and residents' quality of life. Meanwhile, the data also presents opportunities for the private sector. Real estate developers and hotel operators can identify profitable areas for expansion in high-demand zones or consider investing in underutilized neighborhoods to diversify Madrid's accommodation offerings.

Additionally, sustainable tourism practices could emerge by encouraging tourists to explore and stay in lower-density areas. This would not only reduce environmental and social pressure on central neighborhoods but also support the economic development of less-visited parts of the city. Finally, local governments can use this analysis to regulate the growth of short-term rentals, balancing tourism needs with preserving the character and livability of Madrid's urban spaces.

Reflection and Improvements

While the current analysis offers valuable insights into the distribution of accommodations in Madrid, there are several areas where it could be improved or expanded to gain a more comprehensive understanding of the city's tourism and accommodation dynamics.

One potential improvement is incorporating **temporal data** into the analysis. Accommodation density might vary significantly throughout the year, particularly during high tourist seasons (such as summer or holidays). By including data on occupancy rates or average stays per area, we could understand how these densities fluctuate and better account for peak times when pressure on infrastructure is most intense. This would allow for a more nuanced view of seasonal tourism trends and help planners anticipate and manage high-demand periods more effectively.

Additionally, **socio-economic factors** could be explored to better understand the relationship between accommodation density and the broader socio-economic landscape of different neighborhoods. For example, how does accommodation density correlate with the

income levels, residential status, or demographic profile of an area? Analyzing this could provide a deeper understanding of the effects of tourism on local communities, such as rising rent prices or gentrification.

Another useful improvement would be the inclusion of **other types of accommodation** data, such as private rentals (Airbnb, for example), hostels, or boutique hotels. Currently, the analysis focuses on standard accommodations, but short-term rentals have become a significant part of the tourism landscape in many cities. Including these would offer a more complete picture of how accommodations are distributed throughout the city, especially in areas where traditional hotels might be less prevalent but private rentals are abundant.

Furthermore, expanding the analysis to include **transportation accessibility** would be valuable. By overlaying public transportation routes and distances from key tourist attractions, we could assess how accommodation density is affected by the ease of access to transportation hubs. This would provide a clearer idea of whether tourists are selecting accommodations based on proximity to transport or if central locations are still the most appealing due to their walkability.

Lastly, conducting a **comparative analysis with other major cities** could provide context for Madrid's accommodation density. How does Madrid compare to other European capitals like Barcelona, Paris, or Rome in terms of hotel distribution? This comparison could offer insights into broader trends in European tourism and urban development, as well as highlight best practices that Madrid might adopt.

By integrating these aspects into the analysis, we could develop more targeted recommendations for sustainable urban planning, tourism management, and policy development, while also supporting a more balanced and equitable distribution of tourism benefits across the city.