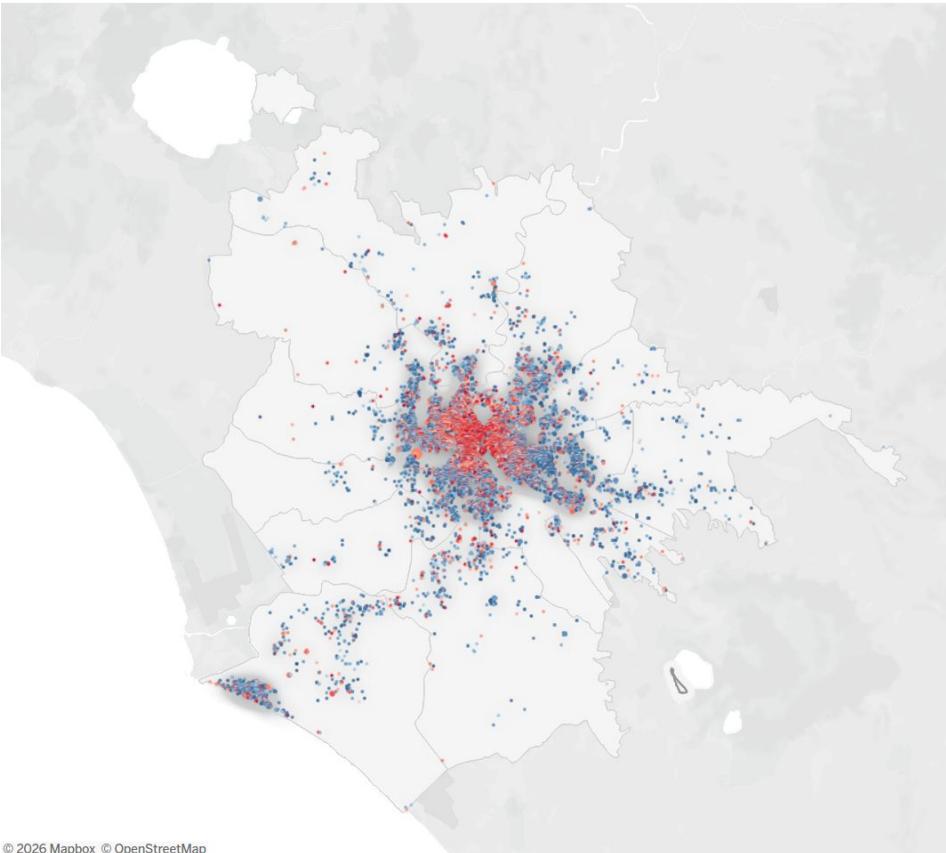


Data Visualization Pitch

Student Individual Assignment



The Roman Empire of AirBnB

Student name: Azriel Matthew Ortega
Student ID: 946814

Student e-mail address: a.ortega@campus.unimib.it

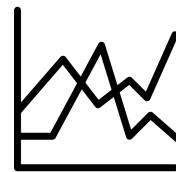
Research questions

How does the Airbnb landscape in Rome vary between its historic core and outer neighbourhoods, and how do "professional" hosts (multi-listing) differ from "casual" hosts in their strategy, pricing, and legal compliance?



Geographic Split

What are the pricing and popularity hotspots in Rome?



Market Dynamics

What is the dominant type of listing in each neighbourhood, and who is offering it?



Host Professionalization

How does host type correlate with availability and legal compliance?

About Data

Source: <https://insideairbnb.com/get-the-data/>

Inside Airbnb is a mission driven project that provides data and advocacy about Airbnb's impact on residential communities.

Primary issue with AirBnb data is that AirBnB data does not officially release raw data for public research. However, Inside Airbnb provides their data by web-scraping on the AirBnb website. Hence, the data is limited to when the data was retrieved from and is not a continuous live source.

Inside AirBnb provides numerous files to be used for "non-commercial, public analysis". The files `listings_summarize.csv`, `neighbourhoods.csv`, and `neighbourhoods.geojson` are used for this project.

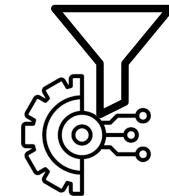
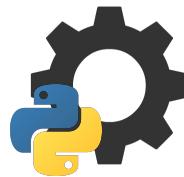
About Data

Source: <https://insideairbnb.com/get-the-data/>

The dataset has some quality and completeness concerns in terms of "Bookings" where Airbnb uses Reviews to estimate the number of bookings a listing have. Additionally, some guests stay without leaving reviews, so the "occupancy rate" for listings are likely to be understated. Moreover, zombie listings (listed in airbnb but haven't been booked) are included in the dataset. Price is often "base price" where other extra fees are not included. For host privacy, listing coordinates are anonymized (buffered within a 150m radius, hence spatial analysis reflects neighbourhood clusters rather than exact street addresses.

The Dataset requires cleaning before analysis such as price formatting, handling nulls, data parsing and outlier removals if needed.

Methodology



Data Collection

- Source: Publicly available Airbnb datasets
<https://insideairbnb.com/get-the-data/>
- Method: CSV download of listings and geographic GeoJSON files for Rome.

Processing & Cleaning

- Stack: Python using the Pandas library.
- Tasks: Removed listings with zero availability, handled missing price values, and filtered for the Rome metropolitan area.

Transformation

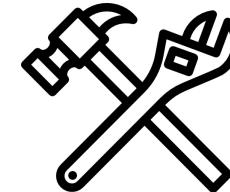
- Segmentation: Classified hosts into "Casual" (1 listing) vs. "Business" (2+ listings).
- Aggregation: Calculated median prices per neighborhood to avoid skewing from luxury outliers.
- Mapping: Joined listing coordinates with administrative boundary data for spatial analysis.

Methodology



Analytical and Visual Techniques

- Tools : Python and Tableau
- Descriptive Statistics: To define market composition.
- Comparative Analysis: Contrasting host counts vs. listing volume to identify market professionalization.
- Spatial Analysis: Using Choropleth map to identify geographic price correlations and tourist hotspots.

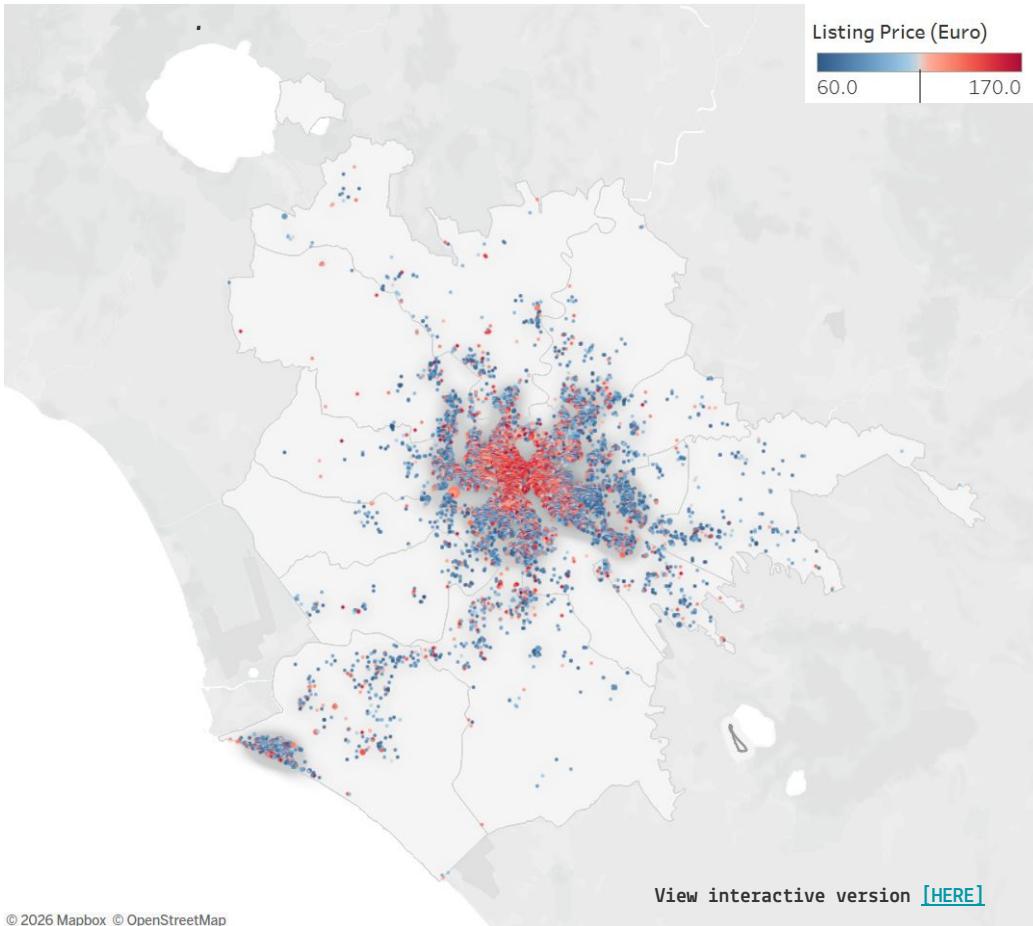


Tools and Reproducibility

- AI Integration : Utilized Gemini for code optimization of complex data visualizations and debugging.
- Reproducibility : All methodology process is documented in a **Jupyter Notebook**.
- Documentation: Shared via Github the source code and instructions. (View Documentation [HERE](#))

What are the pricing and popularity hotspots in Rome?

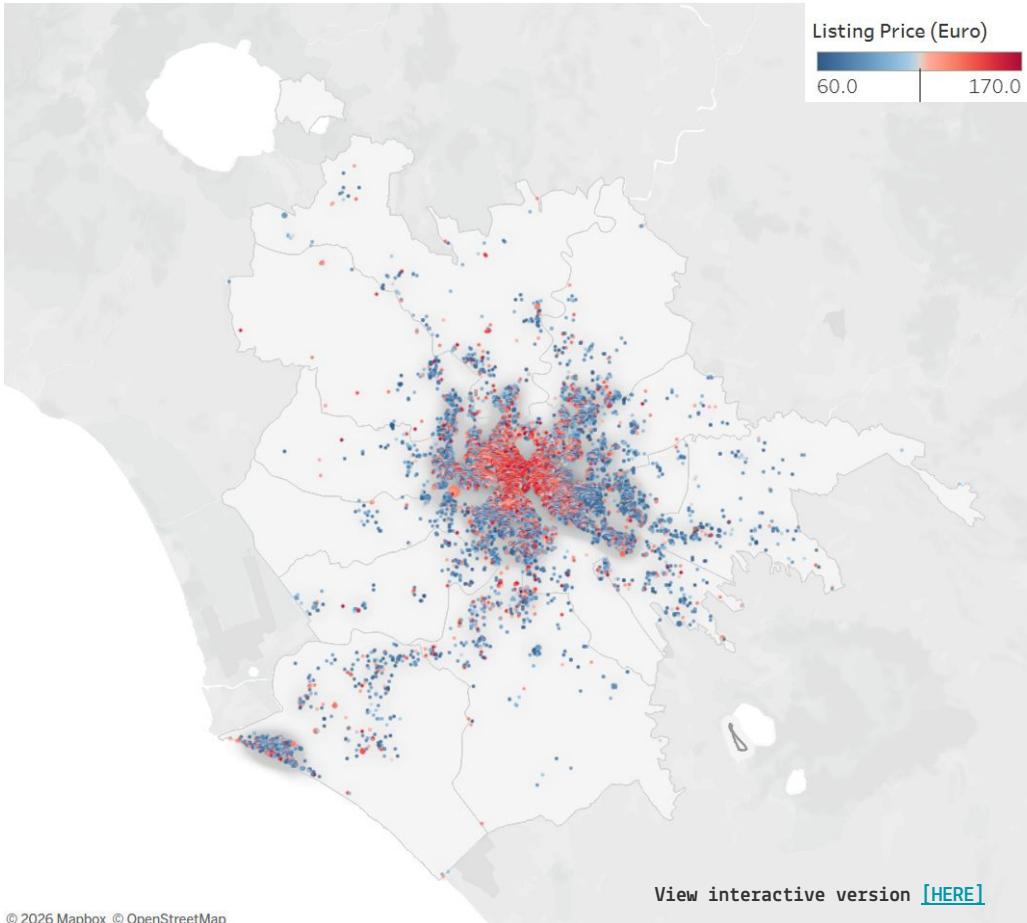
AirBnb Market Hotspots: Popularity vs Pricing



- This map examines Rome's Airbnb landscape by overlaying individual listings onto a median price choropleth. Using a color gradient from **blue (budget)** to **red (premium)**, it highlights the clear spatial relationship between high listing density and the city's most expensive administrative regions.
- **I Centro Storico** has the **greatest number of listings** as well has **the highest median price per listing**
- Majority of tourist landmarks are in Centro Storico such as the Pantheon, Piazza, Navona and Trevi Fountain
- Property owners find that having listings in this area have extremely high demand, hence the amount of listings and price are all high (#1 in both)
- **VII San Giovanni/Cinecittà** is 2nd most popular as it provides an easy access to the tourist landmarks via the Metro
- The zone is also cheaper ranking **#5** in the pricing, ideal for travelers who want access to the city center, but don't have the luxury to rent at Centro Storico
- The zone is also the most populous municipality in Rome, so **Housing stock is packed more in this area, providing more listings for hosts**
(<https://www.citypopulation.de/en/italy/romacity/>)

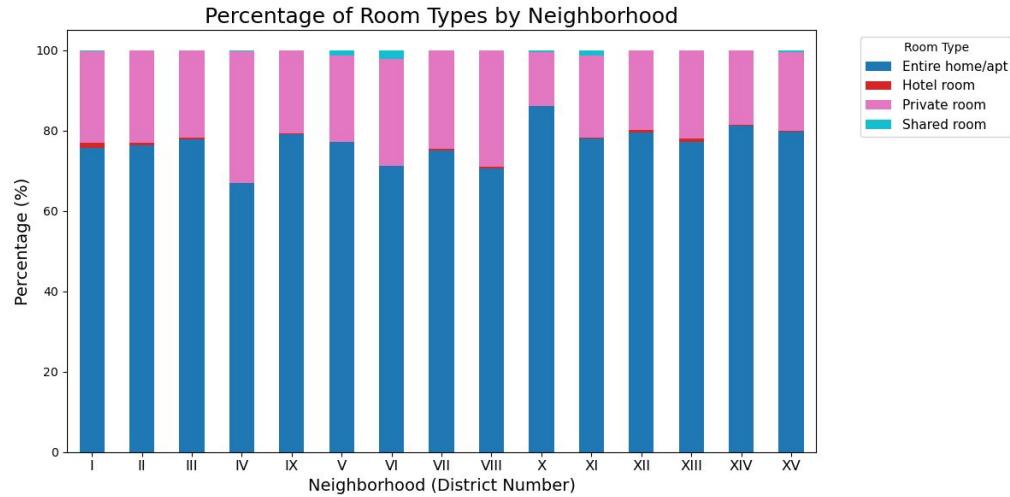
What are the pricing and popularity hotspots in Rome?

AirBnb Market Hotspots: Popularity vs Pricing

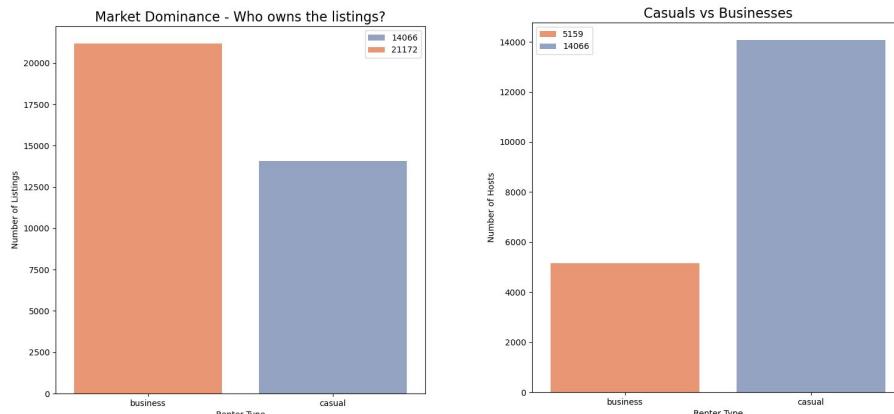


- **XIII Aurelia** and **II Parioli/Nomentano** are high-end alternatives to the center, offering slightly lower prices than Centro Storico but with as well a decent amount of Airbnb listings. This neighbourhood is slightly more accessible to the center of rome, and usually holds more residential "high-end" experience.
- **V Prenestino/Centocelle** neighbourhood is ranked 6th in listing count, but 2nd lowest in price, making a primary choice for **budget travelers**
- Proximity to student-heavy Pigneto district and a generally lower cost of living makes this an attractive neighbourhood for young travelers and long stay guests who prioritize value over proximity
- **VI Roma Delle Torri** shows the lowest density and lowest prices. The lack of tourism infrastructure and distance from the center results in minimal Airbnb activity, catering to non-tourist stays.

What is the dominant type of listing in each neighbourhood, and who is offering it?

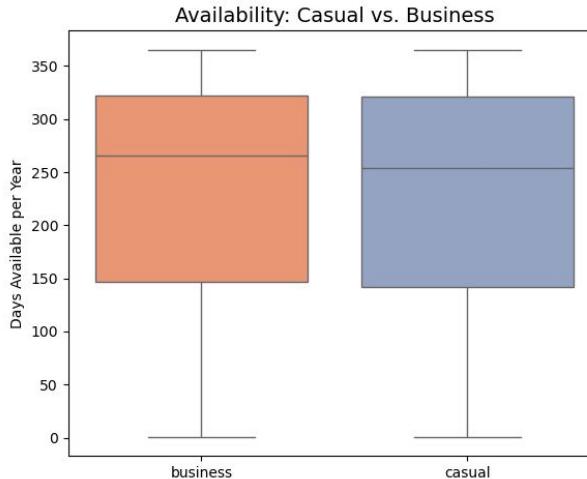


- This stacked bar chart illustrates the distribution of the four Airbnb room types across different areas of Rome, normalized to 100% to show relative proportions.
- As shown in the chart, **"Entire apartment/home"** is dominant in every neighbourhood. This indicates that the Rome market has moved from the original "home-sharing" concept to a more professionalized landscape where full units are now the primary offering.
- This also shows that **shared** or **private** rooms represent a diminishing minority of the inventory



- This visualization compares the total number of listings (left) against the total number of unique hosts (right), categorized into **"Casual"** (1 listing) and **"Business"** (2+ listings).
- The contrast between the two charts is striking. While the **majority of the hosts** are **casual hosts**, the **majority of the "inventory" or listings** is controlled by a smaller group of **business hosts**.
- This highlights a significant concentration of market power, where **business hosts** manage the bulk of the city's available listings

How does host type correlate with availability and legal compliance?



- The 1st plot illustrates the availability of listings per host type ([Casual vs. Business](#))
- **Business hosts** show **concentrated high-availability**, reflecting professionalized inventory management and a commitment to year-round market presence.
- **Casual hosts** show a lower median and greater statistical dispersion, suggesting more intermittent hosting patterns, **likely driven by personal use of the property, or non-commercial objectives**



- The 2nd plot explains the Legal Compliance of the owners, where if a certain listing has a license provided to AirBnB.
- **Business hosts** are significantly **more likely to display a valid license** than **casual hosts**.
- **Business hosts** operate with **higher transparency** because they are more 'visible' to tax authorities and face higher institutional risk. **Casual hosts** may struggle with the administrative complexity of Rome's bureaucracy or perceive their smaller scale as a shield against enforcement.
- **NOTE:** The "License" field is self-reported by the host. While a "Licensed" status indicates that a permit was provided, the dataset does not verify the current validity of that specific permit with the Comune di Roma

LICENCE

Slides are shared with the following license:

The Roman Empire of AirBnB by Azriel Matthew Ortega,
is licensed under
CC BY-NC 4.0

Base Map Data: [© OpenStreetMap contributors](#). Data is available under the Open Database License (ODbL).