Università di Catania

AirBnB Sales Analysis

Negar Laleh

University of Catania

Department of Business and Economics

Academy year 2023/2024

Description of Dataset

OccupancyRateByCity - Copia.xlsx

- **City**: The name of the city.
- **Occupancy Rate**: The rate of occupancy in the city

CitiesInSicily.xlsx

- **Città**: The name of the city.
- **Provincia**: The province where the city is located.

HouseInfo.xlsx

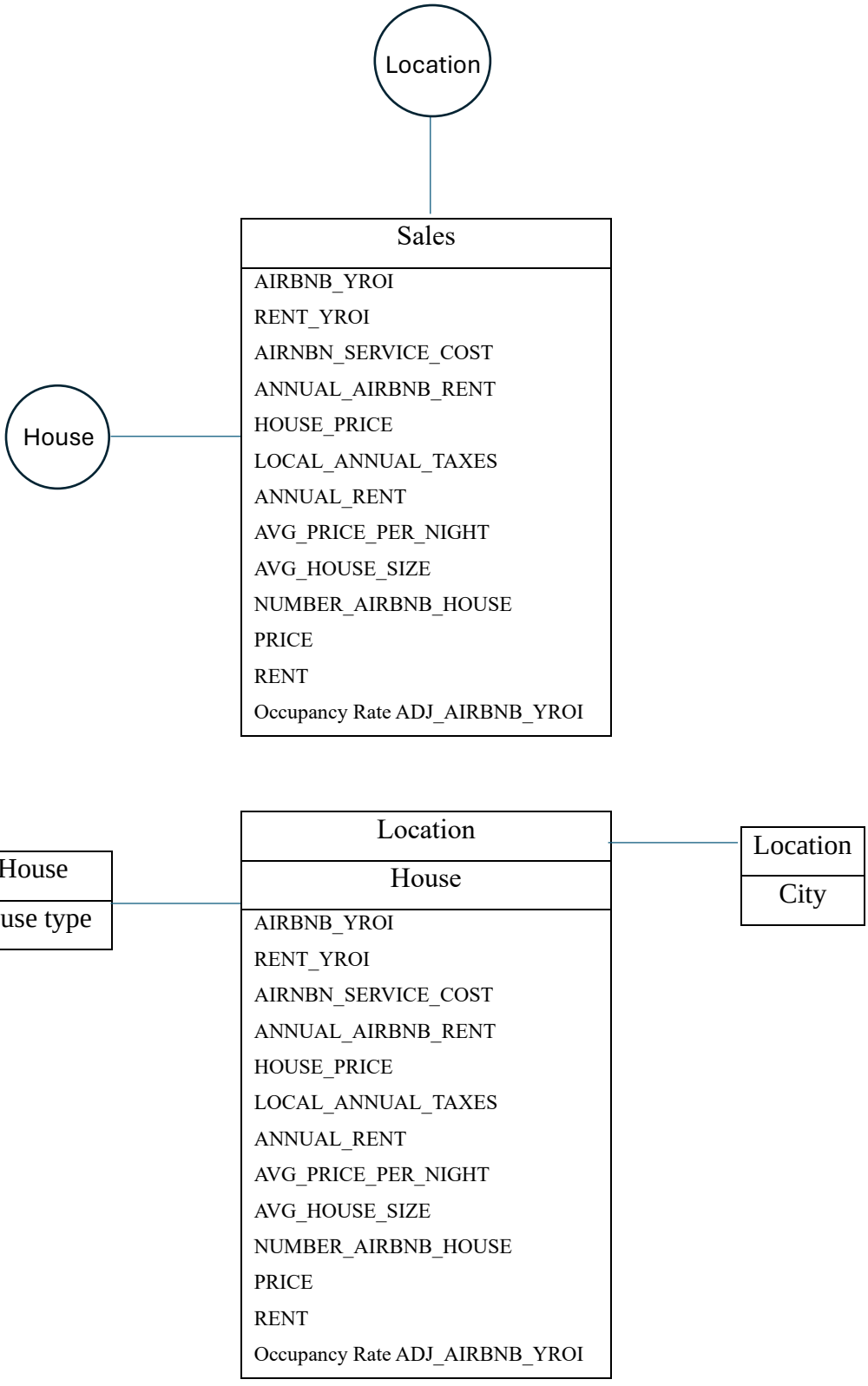
- **id**: An identifier for the house.
- **City**: The city where the house is located.
- **Bedrooms_Temp**: The number of bedrooms in the house.
- **Bathrooms**: The number of bathrooms in the house.

AirBnBPrice.xlsx

- **id**: An identifier for the house.
- **Price**: price of the house size 1

conceptual model and a logical design

From BDA2324_05-DW-ConceptualModelingandDesign_Part1.pds, slide number 31-41 and BDA2324_06-DW-LogicalModelingAndDesign.pdf, slide 34


Location
Sales
AIRBNB_YROI
RENT_YROI
AIRNBN_SERVICE_COST
ANNUAL_AIRBNB_RENT
HOUSE_PRICE
LOCAL_ANNUAL_TAXES
ANNUAL_RENT
AVG_PRICE_PER_NIGHT
AVG_HOUSE_SIZE
NUMBER_AIRBNB_HOUSE
PRICE
RENT
Occupancy Rate ADJ_AIRBNB_YROI
House
House
House type
Location
House
AIRBNB_YROI
RENT_YROI
AIRNBN_SERVICE_COST
ANNUAL_AIRBNB_RENT
HOUSE_PRICE
LOCAL_ANNUAL_TAXES
ANNUAL_RENT
AVG_PRICE_PER_NIGHT
AVG_HOUSE_SIZE
NUMBER_AIRBNB_HOUSE
PRICE
RENT
Occupancy Rate ADJ_AIRBNB_YROI
Location
Location
City

Description of the OUTPUT File

Output.xlsx

- **City**: city the houses exist in
- **House Type**: the type of hose consisting of the numbers of the bathroom/s and bedroom/s
- **AirBnB_yroi**: annual Airbnb yield on investment (YROI) for a property
- **Rent_yroi:** annual rent yield on investment (YROI) for a property
- **Airbnb_servise_cost:** cost of the services provided by Airbnb
- **Annual_airbnb_rent:** annual rental income from an Airbnb property
- **House_price:** full price of the house
- **Local_annual_taxes:** the amount of the annual taxes
- **Annual_rent: the** price of the annual rent
- **Avg_price_per_night:** average price of the house per night
- **Avg_house_size:** average size of the house
- **Number_airbnb_houses:** number of the properties available to Airbnb
- **Occupancy rate:** occupancy rate of each city
- **ADJ_AIRBNB_YROI:** adjusted annual yield for an Airbnb property, considering additional costs and occupancy rate variations.

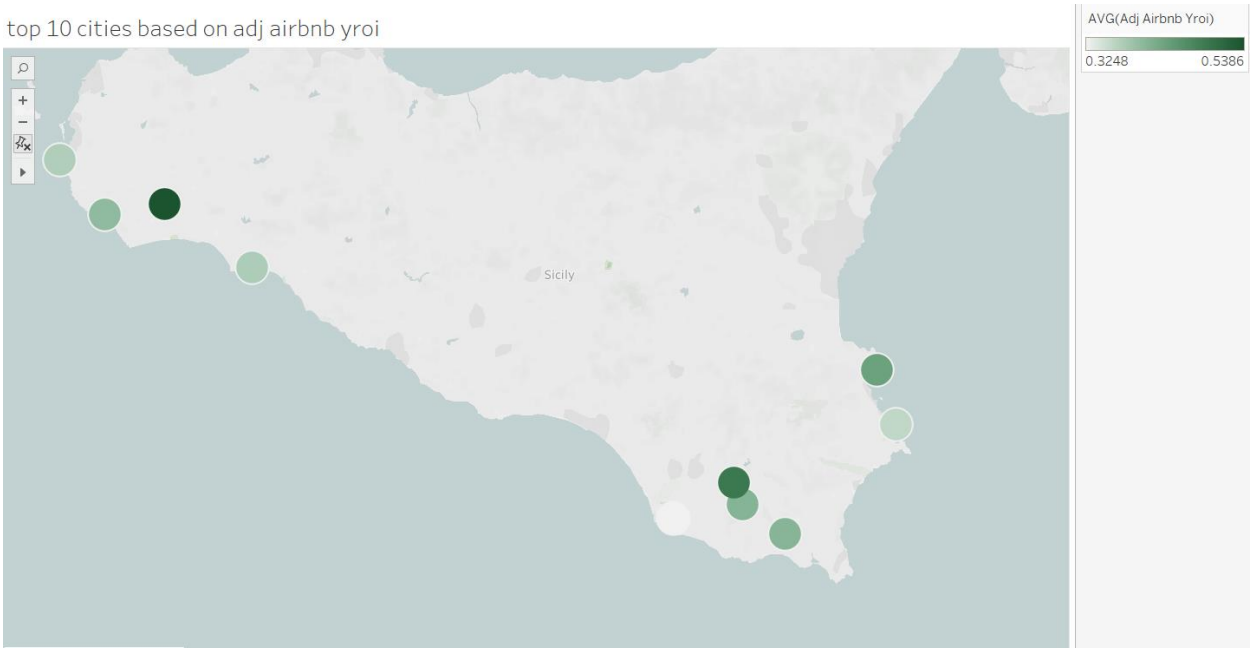
Description of the ETL Operations

In the initial step, the datasets "houseinfo.xlsx" and "AirbnbPrice.xlsx" were integrated based on property IDs. Subsequently, a data cleaning process was undertaken to rectify minor discrepancies and ensure data integrity before proceeding with the primary analysis. The variable representing house size was derived using a specified formula, while the classification of house type was established by combining the number of bedrooms and bathrooms. Following this, an aggregation method was employed to determine the count of properties available for Airbnb rental. A renaming procedure was executed before merging the current

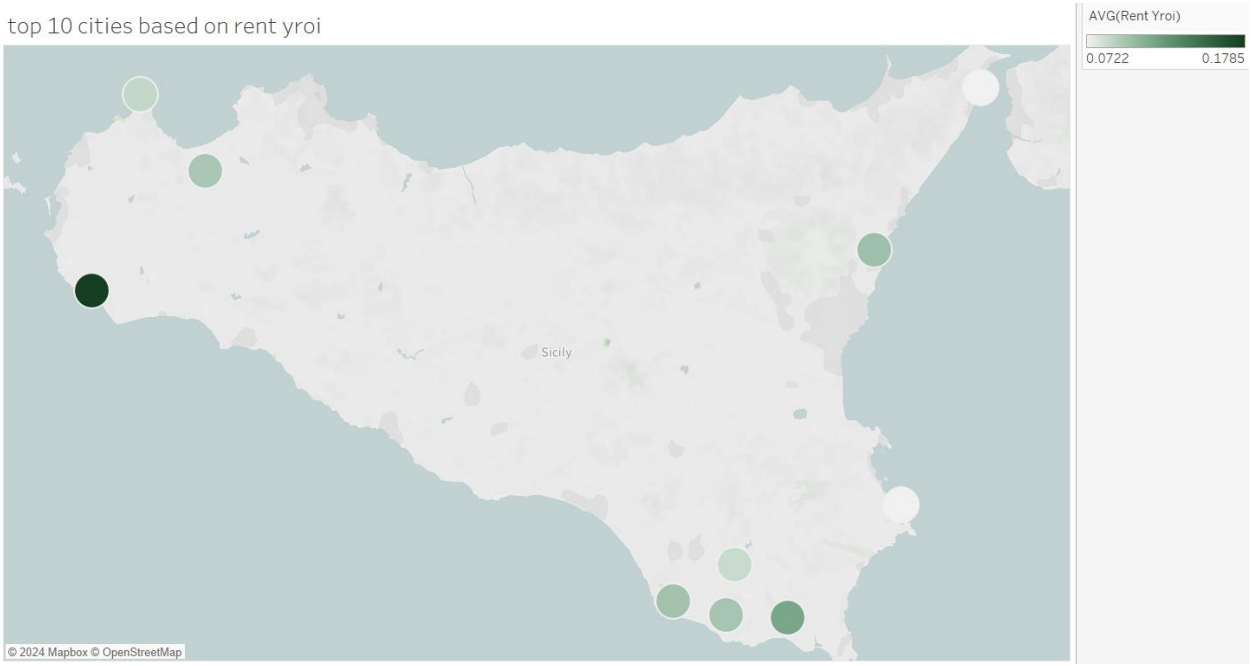
dataset with "CitiesInSicily_BUY_RENT.xlsx" based on their respective cities. A thorough assessment augmented this to rectify any inconsistencies in city names across both datasets. The dataset was improved after additional cleaning steps, which involved computing variables such as annual rent, house price, local annual taxes, annual Airbnb rent, Airbnb service cost, rent YROI, and Airbnb YROI. Upon joining with "OccupancyRateByCity.xlsx" based on city and subsequent optimization, the variable ADJ_AIRBNB_YROI was introduced, and the final output was extracted.

Summary Charts

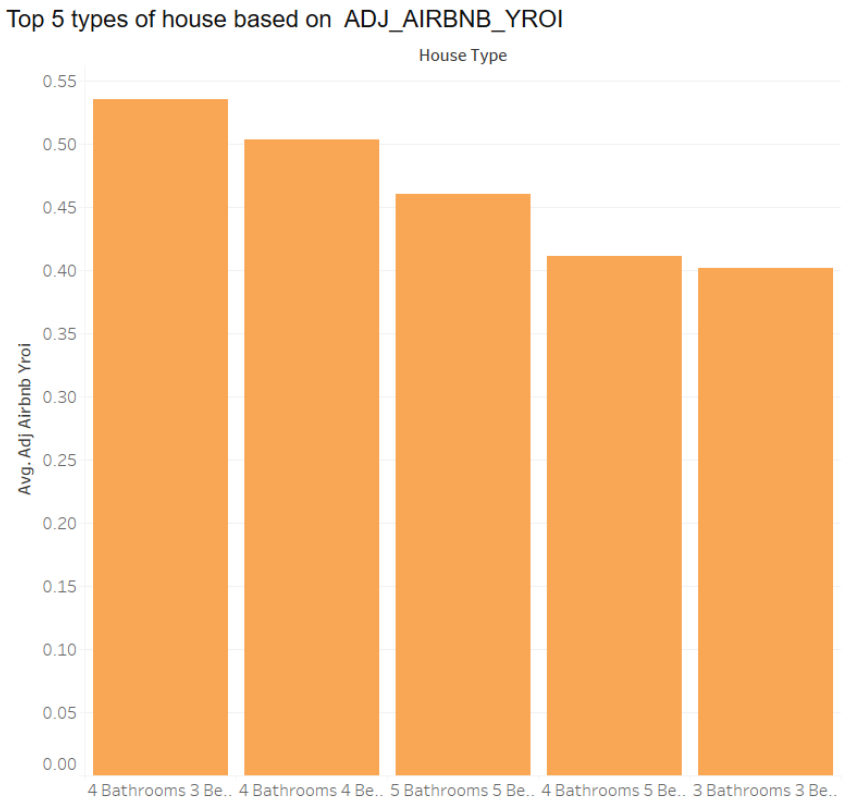
- Top 10 cities based on ADJ_AIRBNB_YROI



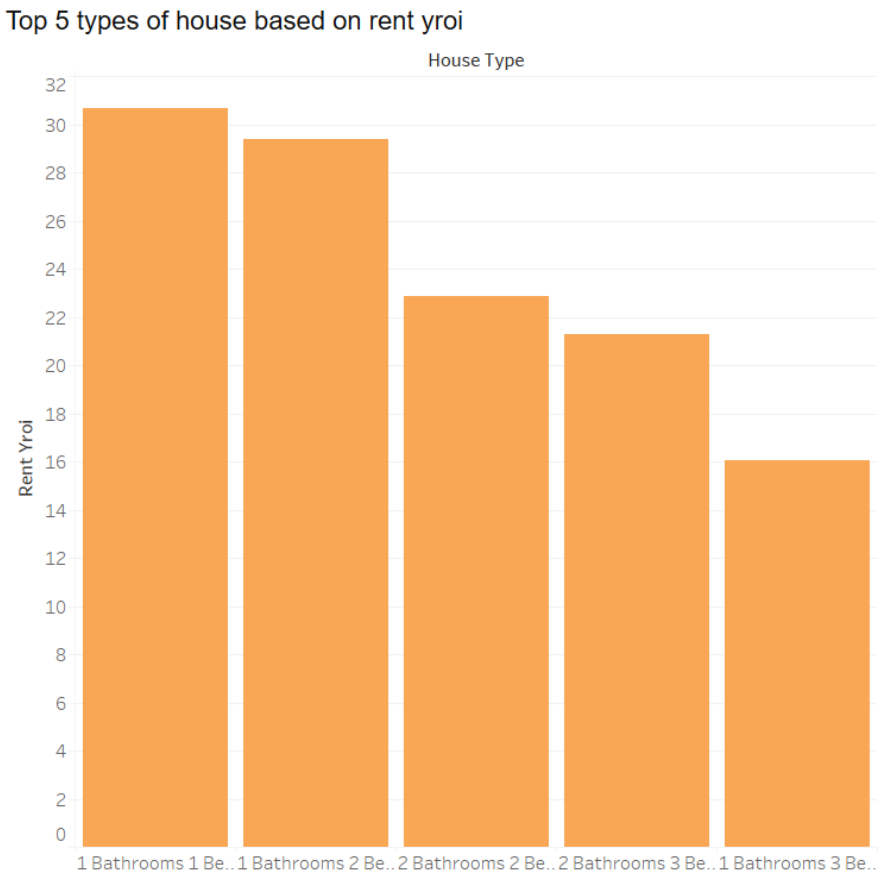
- Top 10 cities based on RENT_YROI



- Top 5 types of house based on ADJ_AIRBNB_YROI

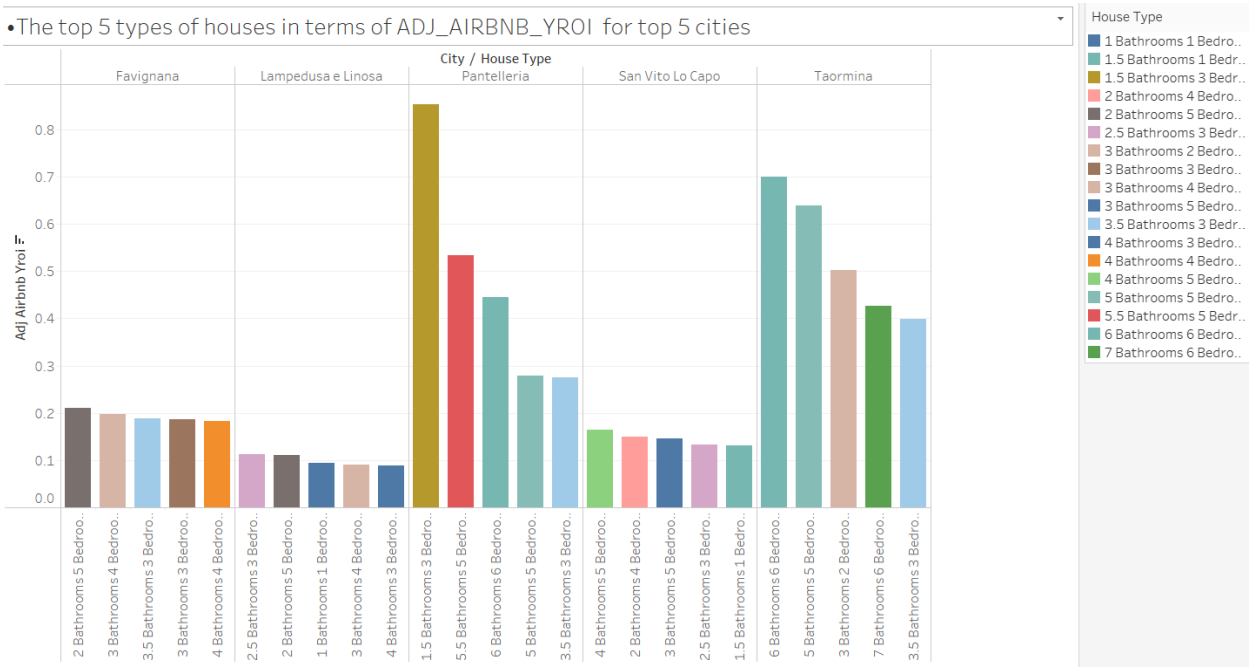


- Top 5 types of house based on RENT_YROI



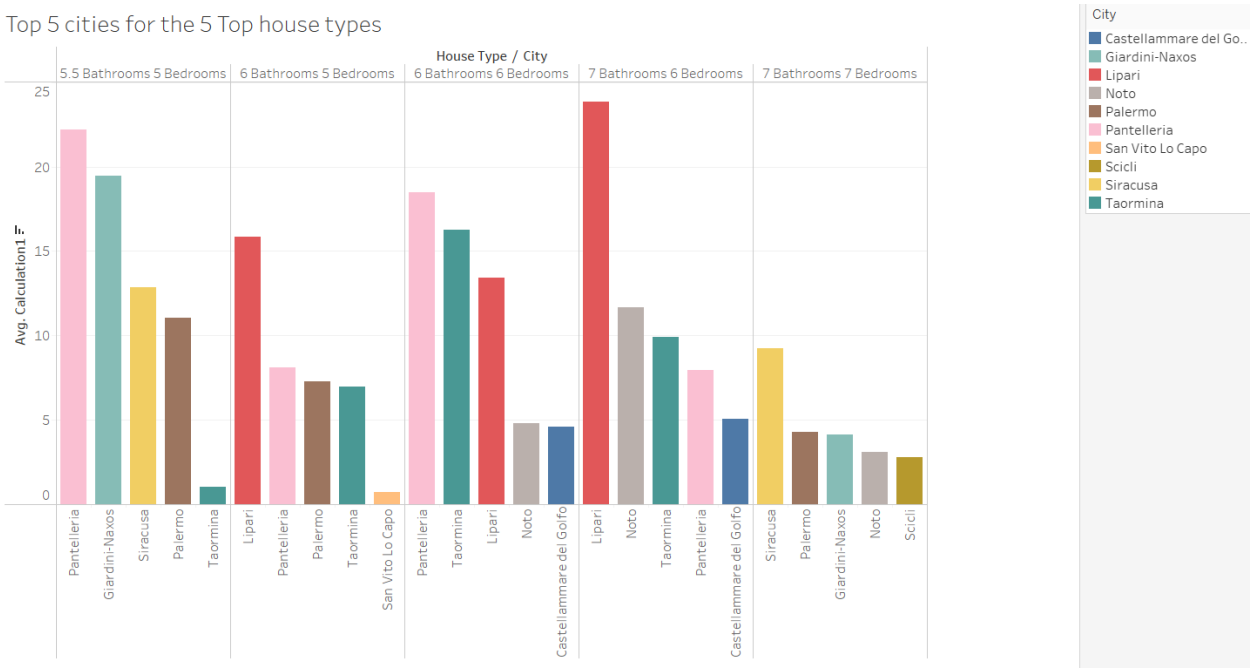
- **Additional Summary Charts**:
    - The top 5 types of houses in terms of ADJ_AIRBNB_YROI (consider only types with more than 20 samples) for the top 5 cities based on the Occupancy Rate (AVG, consider only cities with more than 300 houses).

consider the cities with 300 or more houses. Then, select the top 5 cities based on occupancy rate. For these cities, consider house types with more than 20 samples and choose the top 5 house types based on ADJ_AIRBNB_YROI. As a result, we get the following chart:



- The top 5 cities in terms of Occupancy Rate (consider only cities with more than 300 houses) for the top 5 types of houses based on the ADJ_AIRBNB_YROI/RENT_YROI.

By choosing the top 5 house types besed on ADJ_AIRBNB_YROI/RENT_YROI, and choosing the 5 top cities based on the occupancy rate we obtain the following chart:



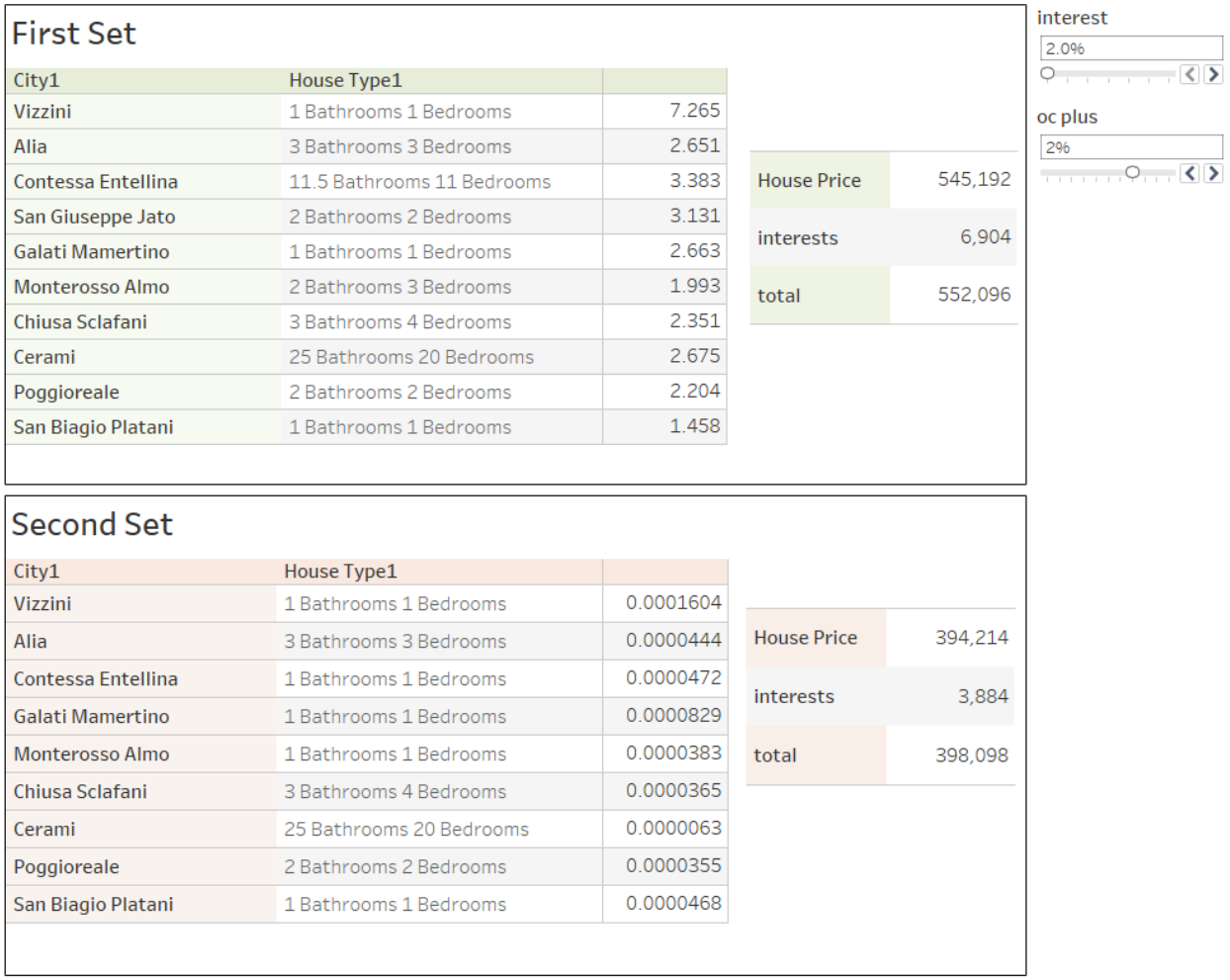
As it can be seen the cities are the same for house types

For the first two summary charts, maps were used to examine if distance affected the results. Bar charts were employed to clearly display differences, using different colors to distinguish between cities and house types. Specifically, this approach helped us identify that some house types or cities are available in multiple cities or house types chosen in the additional summary charts.

STEP 4 Core:

For the first set, I selected the top 10 cities based on their ADJ_AIRBNB_YROI. From each of these cities, I then chose the highest-scoring house based on the ADJ_AIRBNB_YROI again. I used the same criteria for the second set to select the top cities. However, I chose the top houses with the highest Composite Investment Score Formula.

$$\text{Composite Score} = 0.4\left(\frac{rent_{yroi}}{house\ price}\right) + 0.4\left(\frac{AirBnB_{yroi}}{house\ price}\right) + 0.2\left(\frac{1}{house\ price}\right)$$



First Set

City1	House Type1	
Vizzini	1 Bathrooms 1 Bedrooms	7.265
Alia	3 Bathrooms 3 Bedrooms	2.651
Contessa Entellina	11.5 Bathrooms 11 Bedrooms	3.383
San Giuseppe Jato	2 Bathrooms 2 Bedrooms	3.131
Galati Mamertino	1 Bathrooms 1 Bedrooms	2.663
Monterosso Almo	2 Bathrooms 3 Bedrooms	1.993
Chiusa Sclafani	3 Bathrooms 4 Bedrooms	2.351
Cerami	25 Bathrooms 20 Bedrooms	2.675
Poggioreale	2 Bathrooms 2 Bedrooms	2.204
San Biagio Platani	1 Bathrooms 1 Bedrooms	1.458

House Price	545,192
interests	6,904
total	552,096

interest
2.0%

oc plus
2%

Second Set

City1	House Type1	
Vizzini	1 Bathrooms 1 Bedrooms	0.0001604
Alia	3 Bathrooms 3 Bedrooms	0.0000444
Contessa Entellina	1 Bathrooms 1 Bedrooms	0.0000472
Galati Mamertino	1 Bathrooms 1 Bedrooms	0.0000829
Monterosso Almo	1 Bathrooms 1 Bedrooms	0.0000383
Chiusa Sclafani	3 Bathrooms 4 Bedrooms	0.0000365
Cerami	25 Bathrooms 20 Bedrooms	0.0000063
Poggioreale	2 Bathrooms 2 Bedrooms	0.0000355
San Biagio Platani	1 Bathrooms 1 Bedrooms	0.0000468

House Price	394,214
interests	3,884
total	398,098

Subsequently, a separate table presents the total house prices, the interest to be paid, and the total amount of money required.

My first choice was the second set. The first set was selected as an alternative to aid in the "what if" analysis, considering the interest rate. However, the inclusion of this parameter did not alter my decision, and the second set is still the optimal set of houses to purchase.