Revenue AI – Krystian Koniuszy 1st part of assignment. Solution and conclusions.

Intro

In the process of preparing this summary I have used provided data set for Barcelona Airbnb listings(http://data.insideairbnb.com/spain/catalonia/barcelona/2023-12-13/data/listings.csv.gz).

Downloaded csv file contained 18,321 rows of data, of witch 2265 have no data on `price` column rendering them useless for price analysis. Aforementioned rows where removed from data set. After initial data exploration an outlayer have been discovered in data containing price per unit 5 times higher that next highest value per listing with much more greater living standards, therefore this outlayer can be filtered out by tool made for this subject matter.

Conclusion is drawn around outlayered data point, that data may contain errors in data scrapping process which have been used by data set provider in order to collect the information or it may fall into human error category, we can not be sure basing solely on available data.

Outlier can be seen in 1st row of data in bellow screen shot of data.
Dashboard: https://reveniuaiairbnb-cu3vqzvaqbn6hsaincyblt.streamlit.app/

name	↓ price	neighbourhood	property_type	room_type	beds	host_name	listing_url
Guest suite in Barcelona · 1 bedroom · 1 bed · 1 shared bath	56,815	Barcelona, Catalunya, Spain	Private room in guest suite	Private room	1	Eric	https://www.airbnb.com/rooms/25927238
Rental unit in Barcelona · ★4.71 · 1 bedroom · 1 bed · 2 baths	11,000	None	Private room in rental unit	Private room	1	J	https://www.airbnb.com/rooms/37486088
Rental unit in Barcelona · ★4.87 · 3 bedrooms · 5 beds · 2 baths	9,999	Barcelona, Catalunya, Spain	Entire rental unit	Entire home/apt	5	Marta+Paul	https://www.airbnb.com/rooms/912316691776394062
Rental unit in Barcelona · ★New · 1 bedroom · 1 bed · 1 bath	9,999	Barcelona, Catalunya, Spain	Entire rental unit	Entire home/apt	1	Lofts	https://www.airbnb.com/rooms/867256504774715226
Hotel in Barcelona · 1 bedroom · 1 bed · 1 private bath	9,200	None	Room in hotel	Private room	1	ONA Hotels	https://www.airbnb.com/rooms/43479425
Hotel in Barcelona · 1 bedroom · 1 bed · 1 private bath	9,200	None	Room in hotel	Private room	1	ONA Hotels	https://www.airbnb.com/rooms/43479422
Hotel in Barcelona · 1 bedroom · 1 bed · 1 private bath	9,200	None	Room in hotel	Private room	1	ONA Hotels	https://www.airbnb.com/rooms/43479419
Hotel in Barcelona · 1 bedroom · 1 bed · 1 private bath	9,200	None	Room in hotel	Private room	1	ONA Hotels	https://www.airbnb.com/rooms/43479417
Hotel in Barcelona · 1 bedroom · 1 bed · 1 private bath	9,200	None	Room in hotel	Private room	1	ONA Hotels	https://www.airbnb.com/rooms/43479394
Hotel in Barcelona · 1 bedroom · 2 beds · 1 private bath	9,100	Barcelona, Catalunya, Spain	Room in hotel	Private room	2	Ona Hotels	https://www.airbnb.com/rooms/43483589
Hotel in Barcelona · 1 bedroom · 2 beds · 1 private bath	9,100	Barcelona, Catalunya, Spain	Room in hotel	Private room	2	Ona Hotels	https://www.airbnb.com/rooms/43483560
Hotel in Barcelona · 1 bedroom · 1 bed · 1 private bath	9,100	Barcelona, Catalunya, Spain	Room in hotel	Private room	1	Ona Hotels	https://www.airbnb.com/rooms/43483536
Rental unit in Barcelona · ★4.55 · 1 bedroom · 1 bed · 1 shared bath	8,999	Barcelona, Catalonia, Spain	Private room in rental unit	Private room	1	Juan	https://www.airbnb.com/rooms/7105437
Hotel in Barcelona · 1 bedroom · 1 bed · 1 private bath	8,478	Barcelona, Catalunya, Spain	Room in hotel	Private room	1	Ona Hotels	https://www.airbnb.com/rooms/43483528
Hotel in Barcelona · ★4.83 · 1 bedroom · 1 bed · 1 private bath	6,011	Barcelona, Catalunya, Spain	Room in hotel	Private room	1	Acta Voraport	https://www.airbnb.com/rooms/42890622
Hotel in Barcelona · ★4.71 · 1 bedroom · 1 bed · 1 private bath	5,026	Barcelona, Catalunya, Spain	Room in hotel	Private room	1	Acta Voraport	https://www.airbnb.com/rooms/42893238
Rental unit in Barcelona · ★4.71 · 3 bedrooms · 6 beds · 1 bath	5,000	None	Entire rental unit	Entire home/apt	6	Diagonal	https://www.airbnb.com/rooms/600371086640483735
Rental unit in Barcelona · ★4.40 · 3 bedrooms · 6 beds · 1 bath	5,000	None	Entire rental unit	Entire home/apt	6	Diagonal	https://www.airbnb.com/rooms/49333514
Rental unit in Barcelona · ★3.86 · 3 bedrooms · 6 beds · 1 bath	5,000	None	Entire rental unit	Entire home/apt	6	Diagonal	https://www.airbnb.com/rooms/49333438
Rental unit in Barcelona · ★3.75 · 3 bedrooms · 6 beds · 1 bath	5,000	None	Entire rental unit	Entire home/apt	6	Diagonal	https://www.airbnb.com/rooms/49333175
Rental unit in Barcelona · ★4.50 · 3 bedrooms · 6 beds · 1 bath	5,000	None	Entire rental unit	Entire home/apt	6	Diagonal	https://www.airbnb.com/rooms/49332862
Rental unit in Barcelona · ★5.0 · 3 bedrooms · 6 beds · 2 baths	5,000	None	Entire rental unit	Entire home/apt	6	Diagonal	https://www.airbnb.com/rooms/48237481
Rental unit in Barcelona · 1 bedroom · 1 bath	4,500	None	Entire rental unit	Entire home/apt	None	Albert	https://www.airbnb.com/rooms/917928804879169822
Rental unit in Barcelona · 1 bedroom · 1 bed · 1 shared bath	4,000	None	Private room in rental unit	Private room	1	Jonatan	https://www.airbnb.com/rooms/44000385

Figure 1: Dataset table from Price Analysis Dashboard.

Choice of data

In order to prepare this analysis a choice have been made to use full **listings.csv** file in order to drown the most insights into work.

As data frame originally consisted of 75 columns some like **'image_url'** where omitted as they did not bring any value into price analysis.

Some columns like **'neighborhood'** where almost the same just written in opposite order like: *„Barcelona, Spain", „Spain, Barcelona".*

Some other like rating where also deemed not necessary *(yes good reviews will correlate to higher prices and sometimes some high prices to bad reviews using The Snob Effect*) as to yield the most of it would be a good idea to create NLP ranking system to distinguish between bad and good opinions and this was not possible for this assignment in given period of time.

Columns for room_type, property_type, accommodates, beds, price, geo data where chosen as they all interconnected with each other and effect on each other as later shown in analysis.

Analysis

Mean costs fluctuation per room type driven by increase in number of beds.

While looking into cost of renting per room type for whole sample of data no surprises can be found as usual market machinations dictate average cost for one bed or for cost of accommodating one person. Clearly accommodations with lower living standards are far cheaper than their more luxurious counter offers, but by just increasing number of required beds leads to unusual development in average prices.

Figure 2: Whole data sample for mena costs for one bed and one person.

Just changing number of minium required beds to **4** leads to leveling mean price for one person, where the mean price per one bed gradually increases with increase in higher standards for room type as represented on whole data population.
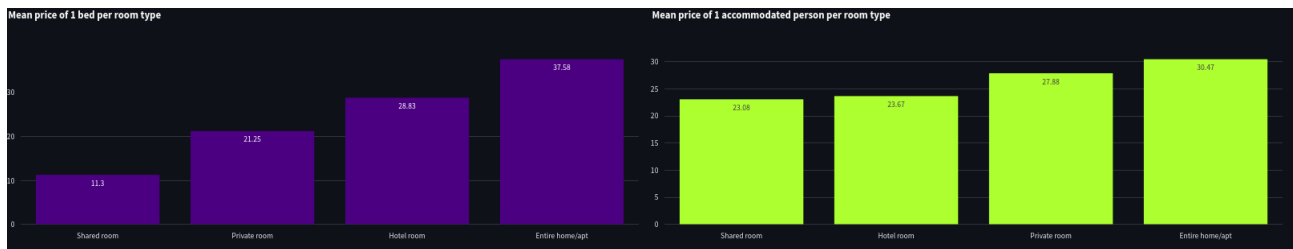
Figure 3: Mean price per one person leveling with number of minimum beds being 4.

Yet all it takes to change this dramatically is to increase minimum beds just by one.
Leaving us with configuration of **1810** units and more than double the gap between mean price for one bed for private room being **15.68$** and entire home apartment with **39.25$** per one bed and keeping this relation up to **10** minium beds with **29$** difference in favor of private room. This situation stays the same with increase of minium number of beds until we reach maximum available number of beds in all listings – being **30**, with just one available property.
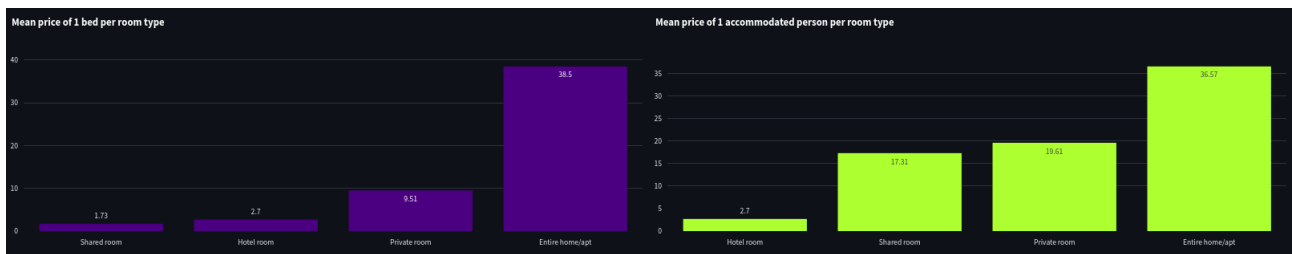

Figure 4: Cost advantage of private room over entire home/apt.

Mean costs falctuation per room type driven by incrise in nuber of accomodates.

No notable changes can be seen in data.

Mean costs fluctuations per property type driven by increase in number of accommodates.

For the most parts data remain the same and gradually increase with increasing prestige of rental property type. Only notable change happens in price when number of minimal accommodes rises to number of **9,** where all other remaining property types with capabilities to host ad least 9 people stay. In that scenario all prices expect for *private room in bed & breakfast* stay the same and even sometimes decrease while we increase number of minimal accommodes, where *private room in bed & breakfast* sores to **1790$** per bed make it **19.5 times** more expensive than closed cheapest alternative *entire chalet*.

Figure 5: Prices stabilisation and one property type huge increase in cost for one bed.

Mean costs fluctuations per property type driven by increase in number of minimum beds.

No notable changes can be seen in data.

Price population histogram.

Just by glancing at our histogram we can clearly see that with full non filtered data we are working with right sided asymmetric distribution, clearly caused by demand in customers wanting to rent cheeper properties.
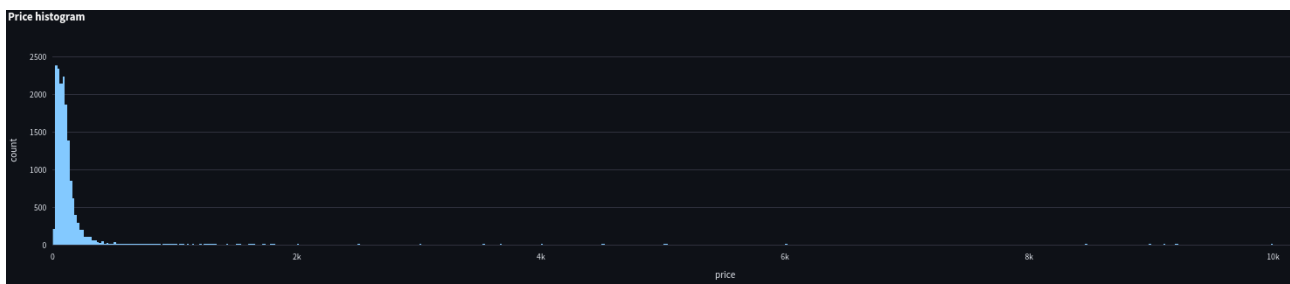

Figure 6: Price population histogram.

Geographical analysis

Thanks to longitude and latitude being present in **listings.csv** data set we are able to check on aerial view of listing items in Barcelona. As long establish saying in real estate say: *"3 the most important things in real estate is: location, location, location",* same can be sad about Airbnb listing properties in Barcelona. Highest density of properties is in city center as it is the most attractive for tourist to rent their rooms with quick and short distance to the monuments and entertainment centers. The highest costing properties are located beyond busy city center giving their future renters peace, quietness, space and luxury.
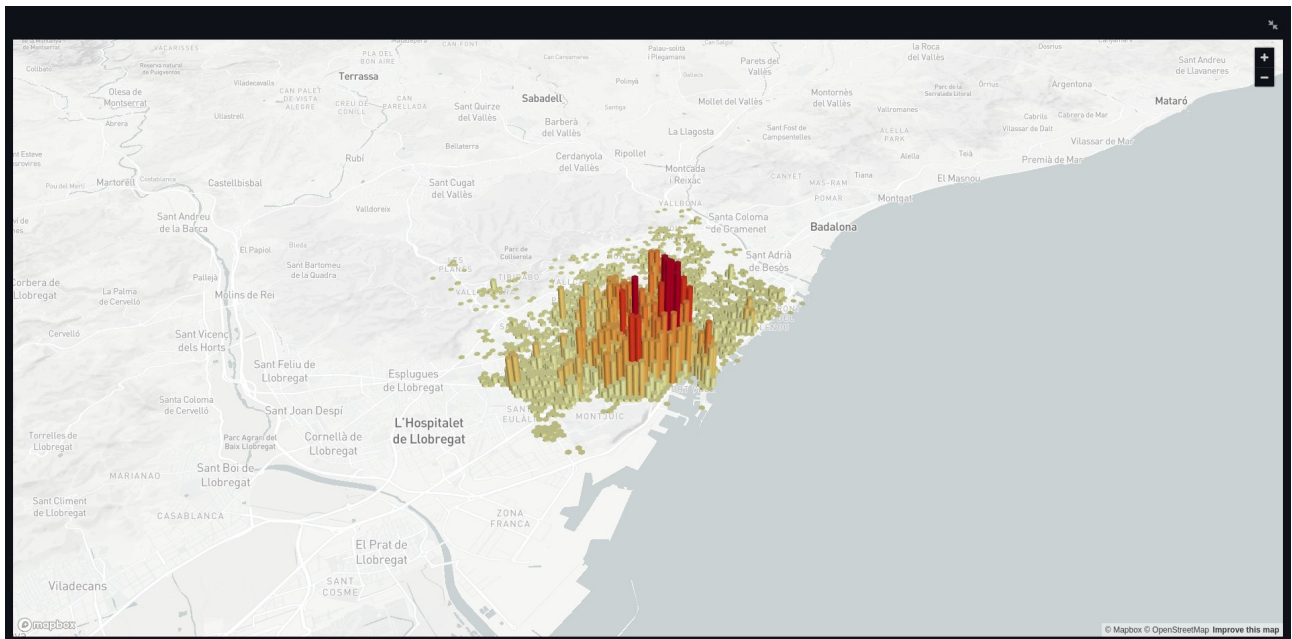
Figure 7: Geograhical placement, density of listings in relation to price.