

Dataset Description

Natalia Ramirez

2024-02-14

What is it?

The data was extracted from “**Airbnb Market Analysis & Real Estate Sales Data**,” with the following URL: [<https://www.kaggle.com/datasets/computingvictor/zillow-market-analysis-and-real-estate-sales-data>].

It is a collection of Airbnb properties listed in 2019 for the California Zipcodes 92314, 92315, 92284, and 92252. It encompasses metrics like generated revenue, occupancy, nightly rate, revenue, and availability. Moreover, it also dives into the commodities inside the real-state, such as the number of bathrooms or bedrooms.

What is the format of the data?

Each row represents a different property and all rows have a unique associated with them. The columns of the data set show characteristics of the real-state in quantitative format. One important thing to note is that some data is shown as integer while other is shown as float, thus conversion to standarize the data is needed. Finally, the float format uses ‘,’ instead of ‘.’. Therefore, data formatting is necessary for correct analysis.

Methods to capture the data

The data was collected from the public site Zillow which shows characteristics of properties in the market, moreover other data was extracted from the Airbnb site directly. The data was gathered monthly.

How large is the database?

market_analysis.csv has 26026 observations (properties listed) with 14 columns of metrics.

Quality of data

The data is almost completely clean and easy to plot without any pivoting function. It serves to analyze characteristics of properties that could be related with revenue growth and nightly rate.

Data shown

```
#Reading Airbnb Market_Analysis
data <- read.csv('/Users/nataliaramirez/Desktop/SoftwareDev_Project/market_analysis.csv', sep = ';')
head(data)
```

##	unified_id	month	zipcode	city	host_type	bedrooms	bathrooms	guests
## 1	10000347	2020-1	92315	Big Bear Lake	2-5 Units	3	2	10
## 2	10000347	2020-10	92315	Big Bear Lake	2-5 Units	3	2	10
## 3	10000347	2020-11	92315	Big Bear Lake	2-5 Units	3	2	10
## 4	10000347	2020-12	92315	Big Bear Lake	2-5 Units	3	2	10
## 5	10000347	2020-2	92315	Big Bear Lake	2-5 Units	3	2	10
## 6	10000347	2020-3	92315	Big Bear Lake	2-5 Units	3	2	10
##	revenue	openness	occupancy	nightly.rate	lead.time	length.stay		
## 1	899,9599915	31	0,064516129	449,9799957	52	3		
## 2	13950	31	1	450				
## 3	13500	30	1	450				
## 4	13950	31	1	450				
## 5	450	29	0,034482759	450	8	1		
## 6	0	31	0					