

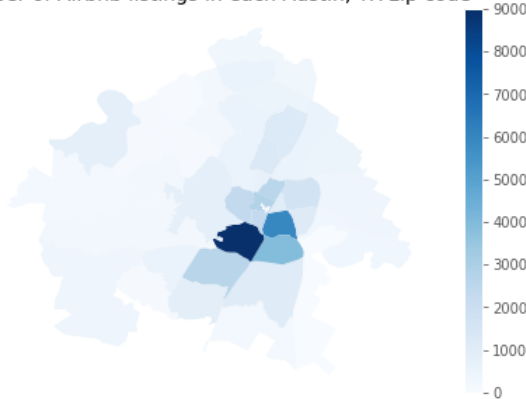
# Data Story: Exploratory Data Analysis

After sourcing and wrangling the dataset for my Airbnb Price Analysis project, I came up with a series of questions to help further explore this data set and get a better understanding of the feature relationships.

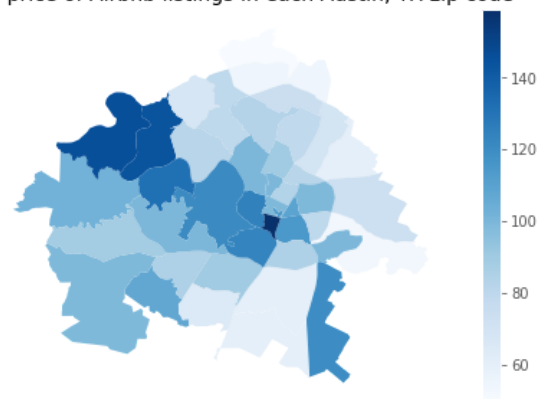
Which areas/neighborhoods in Austin, TX are the most expensive and have the most listings?

**Answer:** With some help from a similar project I found online, I was able to source some code and use the GeoPandas library to visually explore this question on an overlaid map of Austin, TX. I found that while the majority of the listings were located in downtown Austin, the median price point was more spread out. Downtown Austin was still one of the more expensive areas, however there were also some expensive areas to the west and northwest. I suspect that this is because Lake Travis is located in that direction, which is also a popular destination area for people traveling to the Austin area.

Number of Airbnb listings in each Austin, TX zip code



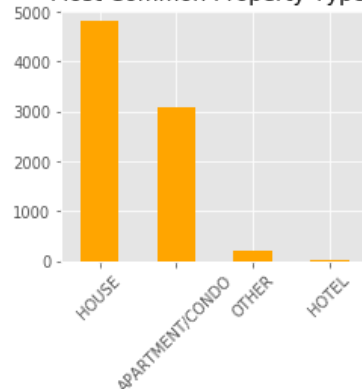
Median price of Airbnb listings in each Austin, TX zip code



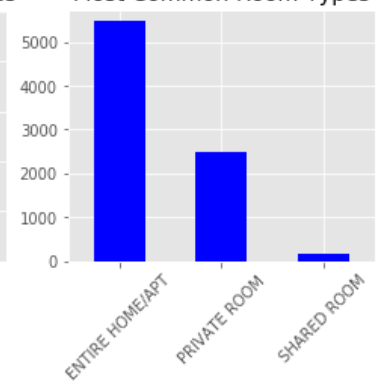
What are the most common listing types?

**Answer:** Houses are the most common property type for this data set, with the entire home/space being the most common room type.

Most Common Property Types

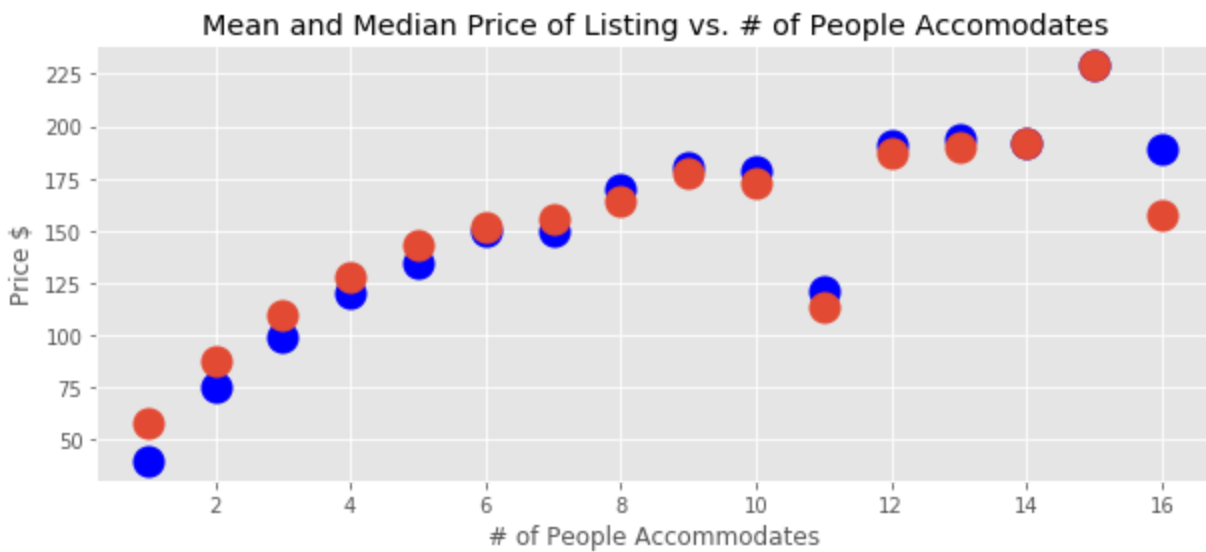


Most Common Room Types



How do Airbnb prices correlate with the number of people a listing accommodates?

**Answer:** It appears that on average, price has a positive relationship with the # of people that an Airbnb listing accommodates, which is a logical pretense of this data set. Also, based on the data in this scatterplot, it seems that this relationship is logarithmic, meaning that the impact of increasing the listing price by accommodating more people subsequently diminishes. I found that the Pearson correlation coefficient for these two features was 0.45, which isn't strong, but still a positive correlation.



**\*\*Mean is in blue, Median is in red**

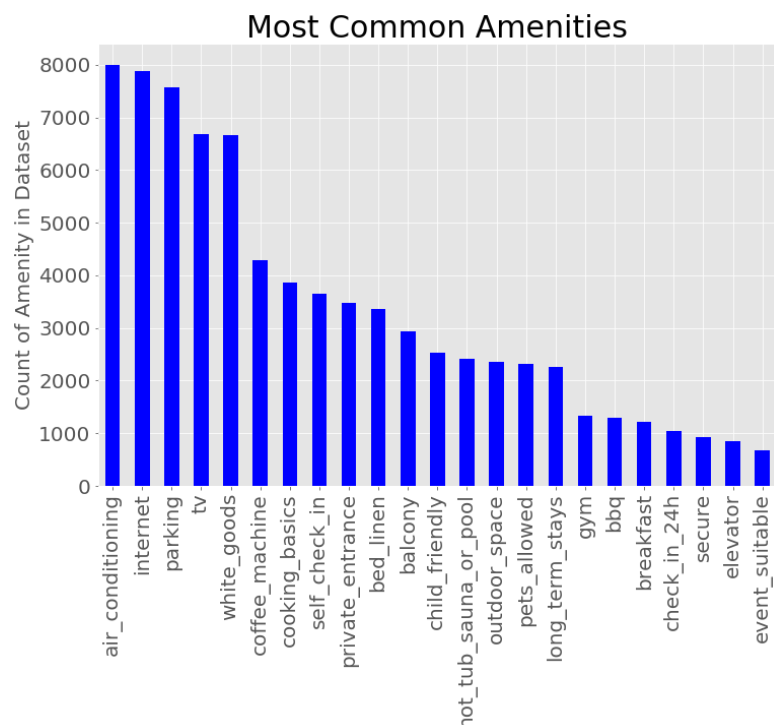
What is the value of being a superhost? Is it worth getting verified?

**Answer:** Neither being a superhost or verified seems to have any impact on median listing prices



What are the most common amenities, and which amenities are likely to increase?

**Answer:** The Most common amenities include A/C, internet, parking, TV and white goods (linens etc.) However it appears that these aren't necessarily the most important amenities to increase listing price. Parking, for example is the 3rd most common amenity, however the data shows a negative impact on the median listing price when it is included in a listing. This could just be that since most of the more expensive listings are located downtown, they simply do not have parking spaces. Also, the majority of listings are not child friendly, however the listings that are child friendly have a significantly higher median price.



## Moving Forward

Based on my initial EDA of the data, I suspect that the largest impacts on listing prices will come from a handful of certain amenities, the geographic proximity to downtown Austin and the number of people that a listing accommodates.