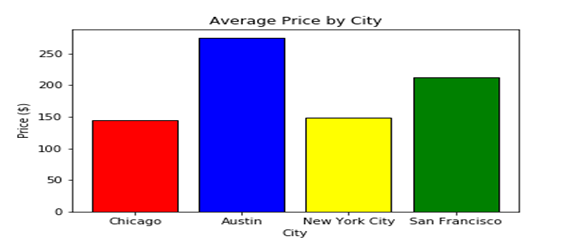
Steve Novis, Jeremy Coy, Godday Ogbo, Sanket Mishra 9/27/18

**Airbnb Data Analysis Write-Up**

Based on the analysis carried on the four cities - Austin, San Francisco, New York and Chicago - the following inferences can be drawn.

**Airbnb Prices per City**

Based on the information we had analyzed, we were able to visualize the differences between the four cities in a bar chart format. From what we found, Airbnbs in Austin are the most expensive on average while Chicago and New York are the least expensive.

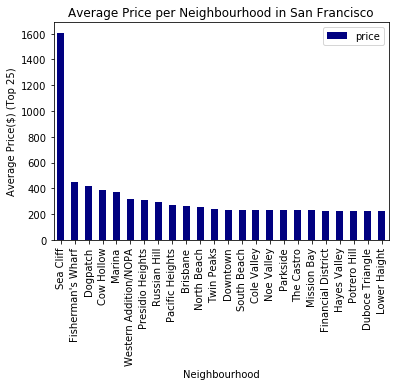
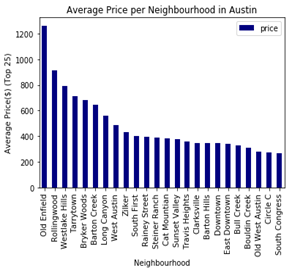


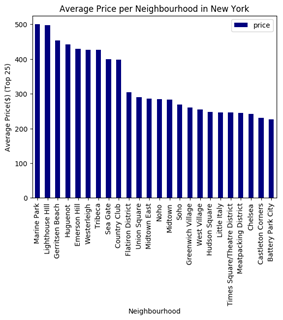
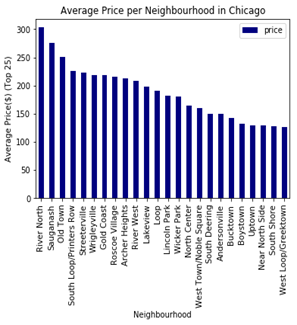
**Average Airbnb prices by neighborhood**

For visualization purposes, only the top twenty-five cities were shown in the charts. The bar charts below show the average price per neighborhood for Austin, San Francisco, New York, and Chicago. Based on what we found, we can see that some neighborhoods are more expensive than others. There can be varying reasons as to why this is the case, but there are notable cases in San Francisco and Austin.

Sea Cliff, San Francisco has an average price of about $1,600 per night, which is almost four times more expensive than the next. After some research we found that this is because Sea Cliff is home to prime real estate. It is at a place where there is a great view of both the Pacific and the Golden Gate Bridge, home to many famous celebrities and successful businessmen, and is situated close to a small public beach.

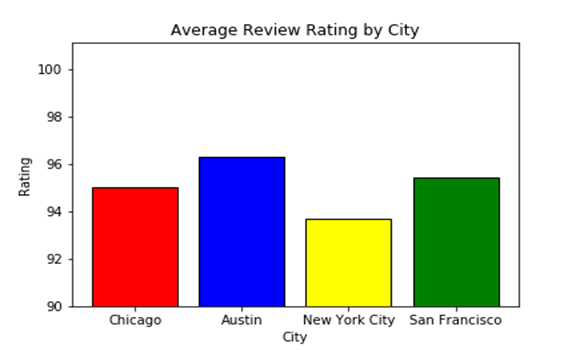
Old Enfield, Austin has an average price of $1,200 per night, a significant jump from the next most expensive neighborhood, Rollingwood. After some research, we found that this is because Old Enfield is where the University of Texas is and the standards of living in the neighborhood is above average. This will likely attract foreign businessmen, guests of the college, and people looking to visit students attending UT.

****

****

**Airbnb ratings by city**

The city with the highest rated Airbnb is Austin and followed by San Francisco, Chicago and New York City respectively. But while the chart may show big differences, the numbers are actually very close and differences in ratings between the cities are very minor.

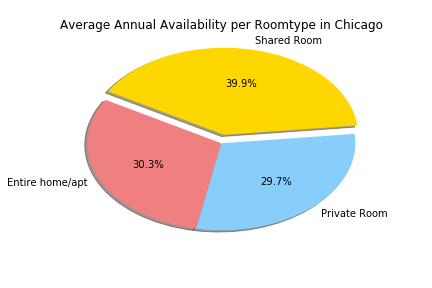
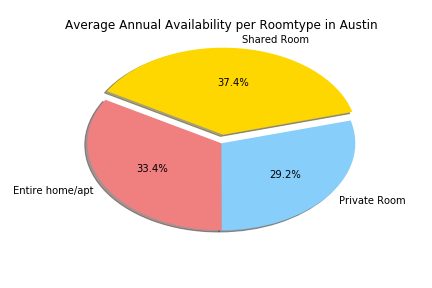


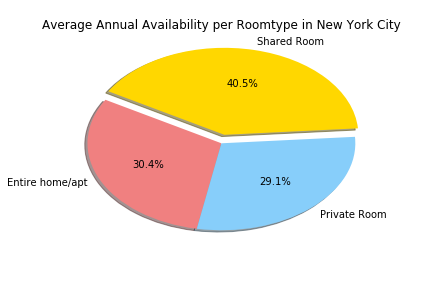
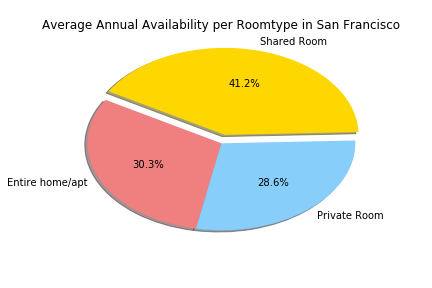
**Average availability by room type for each city**

The pie charts below show percentages of the average availability for each city divided based on room type. Note that availability is the number of times that the Airbnb was vacant and ready in the past year. The way the pie chart is divided tells us which types of rooms were more available - and thus vacant - in the past year.

The three room types listed are private rooms, shared rooms, and entire homes/apartments. Private rooms are singular rented rooms, shared rooms are multiple rooms for rent that are likely used in a shared living space, and entire homes/apartments are renting out the entire living space of a home or apartment.

Based on the data we found, shared rooms are the most commonly vacant across all cities cities, while entire homes and private rooms have mostly similar vacancy amounts. Thus, we can say that the most frequently occupied room types are private rooms, followed by entire homes, and shared rooms are the least frequently occupied room type.

**Conclusion**

From our analysis of Airbnb data, we were able to find many surprising trends between the four cities. We found that Austin has the highest review rating and average price per city (with the highest neighborhood being $12,000 a night). We were surprised to see that Austin was so expensive and highly rated compared to the other cities. This is likely because Austin is not only a nice vacation destination for the nature-lover, but it is also home to the University of Texas. We inferred that the combination of these two likely drove up prices for Airbnbs.

We learned that Sea Cliff, San Francisco’s average price is almost four times as expensive as the next neighborhood. We concluded that this was caused by the prime real estate value of the property and the area. We were able to recognize that this price was highly affected by the high-income neighborhoods, travelling destinations, and the view.

Lastly, we found that the most occupied rooms are private rooms followed by entire homes with shared rooms being the most available. We were able to make several reasonings about each city compared to each other and within their neighborhoods. There is much room to expand this analysis with better data or more time. Some topics to further pursue (within their respective areas) are: historical Airbnb pricing vs hotel pricing, historical Airbnb availability vs hotel pricing, and historical Airbnb ratings vs hotel ratings.

**Limitations**

There were several limitations that we encountered throughout the completion of the analysis. In the beginning, we had planned to compare Airbnb prices to nearby hotels and see if an increase in Airbnb availability affected hotels. After looking for applicable datasets, we found a detailed dataset that had been compiled on Airbnb. One limitation on this Airbnb data set was that there was no historical pricing data, which meant we could not track trend changes over time. Additionally, there was minimal data about hotel pricing that we could access. Since we could not look at Airbnb pricing in relation to hotels over time, we redefined the analysis. We had to narrow our scope down into just Airbnb trends and expanded that scope by looking at various cities and comparing and contrasting them with each other.

As we worked out the analysis of this dataset, we had a few inherent limitations within the data set itself. One example is the average rating reviews we calculated. There is nothing that stops users from not submitting reviews, which means that these reviews could be skewed if residents did not submit feedback. This could mean that the reviews themselves could be inaccurate. Another limitation that we recognized was that the average price per city or neighborhood could be influenced by external conditions that we did not recognize. This means there could be pricing trends or behaviors that we are not aware of.