Matt Fannin

Rental Listings Analysis

Summer Project 3: Rental Listing Analysis with 28 US Cities

Research Questions/Goals:

To maximize rental income, should a property owner rent traditionally (long term) or rent short term (Airbnb)?

What are the best cities to have rental property in?

What affects the amount of income a rental property can bring in?

Data Sources:

The data was collected from the Mashvisor API. I made 29 calls to the API, and created functions in order to convert it from a json object to a python one, so that I could then create a data frame. The code written to accomplish this, will be included in the Github repository. Once the data frame was made, I then exported it as a csv so that it could then be read into RStudio for analysis.

Set-Up

I first set my directory to the folder I was working in, so that I could read in the csv file I exported. The data was imported, and converted to a data frame object in R.

Analysis and Visualizations

I first wanted to analysis if it was more profitable to rent traditionally or rent through Airbnb. I used the summarize method to get the average rent of Airbnb, and traditional listings along with the average of return of investment (ROI). After doing this, the average rent per month from Airbnb was $529.31 dollars higher than traditional rental properties, in addition having a .65 higher ROI on average.

After doing this, I wanted to dive deeper into the data and wanted to attempt to see if there were certain cities that it was better to have a rental property in. To do this I made a smaller data set using the columns "city", "avg_property_price", and "avg_nightly_price." I then, created a new column seeing how much a property could make in year and sorting the cities by that new column "income_per_year." The cities that made the most were: San Francisco ($83,941), Nashville ($83,796), Boston ($78,006), Austin (71,135), and San Jose ($69,806). After finding this out I wanted to see how

long it would take to pay off the property, strictly using rental income. For the same five cities respectively: 20.2 years, 5.6 years, 13.26, 8.84, and 15.5 years. After doing this it seemed like good cities to have rental properties in are: Nashville, San Francisco, Boston, and Austin.

|	city	avg_property_price	avg_nightly_price	income_per_year	time_to_payoff
1	San Francisco	1695222.1	229.9755	83941.06	20.195387
2	Nashville	470686.9	229.5785	83796.15	5.617046
3	Boston	1034751.4	213.7158	78006.27	13.264978
4	Austin	629174.7	194.8927	71135.84	8.844694
5	San Jose	1081207.3	191.2506	69806.47	15.488640

In order to find out what affects the level of income a rental property can bring in, I was curious which cities are the biggest markets. To do this I took the values from the three columns: "investment_properties", "airbnb_listings", and "traditional_listings" and added them up to create a new column. I then arranged the resulting column in descending order to find the largest markets.

| city	investment_properties	airbnb_listings	traditional_listings	total
New York City	6252	5722	627	12601
Los Angeles	3883	4988	743	9614
Las Vegas	5291	991	33	6315
Dallas	3291	1299	97	4687
Philadelphia	2321	1859	263	4443
San Antonio	2758	1596	22	4376
Nashville	1217	2758	30	4005
San Diego	1388	2549	59	3996
Fort Worth	3132	587	64	3783
Washington D.C	1101	2549	97	3747

I wanted to see if if more expensive properties could be rented at a higher price point. So, I created a new column averaging the traditional rental income and the airbnb income, and created a scatterplot. Along with the scatterplot, I also plotted the linear regression line, created a linear regression line, and used the cor() method to find an r value equal to .948.

Property Value vs Rental Income
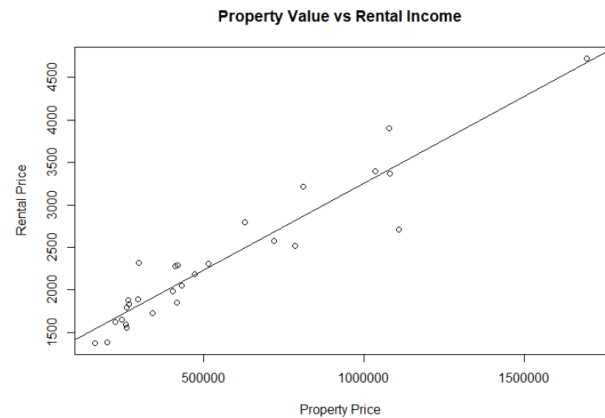
I now exlored if there were more occuapants In a rental property, if there would be more income. I followed a similar approach from above, creating a scatterplot along with plotting a linear regression line. In addition, I tested the predict() method. I used this by creating a data frame that contained a number of occupancies I wanted to predict a rental income for (50). Using the method, it predicted that with 50 occupants in a rental property it would bring in $1,972.94.

Rental Income vs Number of Guests
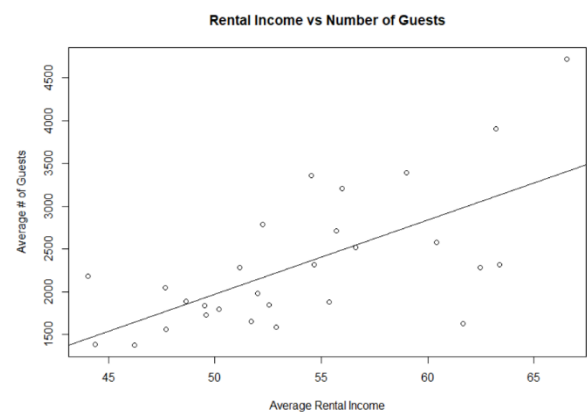
Lastly, I wanted to see if competition of rental properties had an affect on the price of rentals. I used the cor() method to get an r value equal to -.26 using the two columns "neighborhoods" and "avg_nightly_price", and made a scatterplot along with ploting the linear regression line. Then, I sorted the data frame by the two columns in descending order, to see if there was any common cities. The only common city that was in the top five of each was Nashville.

Competition in Rental Prices
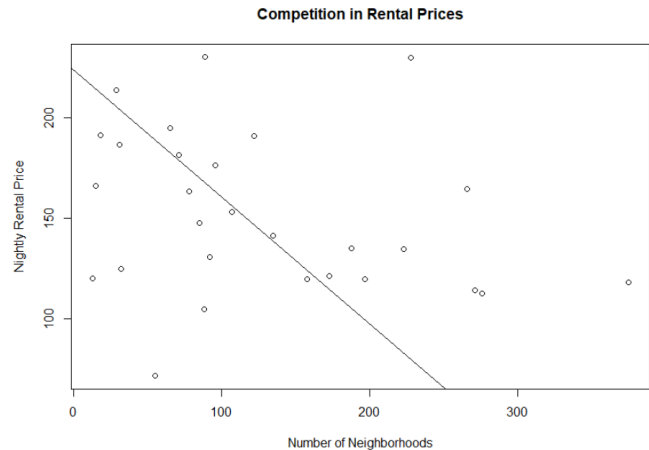
Results/Findings

- Renting through Airbnb seems to make more money (around $500) more per month, and have a better ROI. However, it seems more risky having more people in and out of the property compared to traditional renting. This could lead to more costs such as: cleaning or damage to the property itself.
- The best cities to rent in: San Francisco, Nashville, Boston, and Austin. They have the highest average nightly cost, and produce the most passive income per year. If I had to choose one that is the "best" it would be Nashville. This is because of the high rental rates, in addition to the property being valued fairly low compared to other big cities.
- Rental income is directly affected by how much the property costs, is inversely affected by the number of neighborhoods and competition (good for the consumer), and is directly correlated to the number of occupancies property has.

Skills Practiced/ Things Learned

- Piping
- Data Frame manipulation (select, mutate, arrange)
- Linear regression models and plotting them
- Predict()
- Barplots nad boxplot practice, along with scatterplots