

Airbnb Rentals in Urban and Rural Counties of Texas

Diana Melendez, Ericka Houle and Steven Green

Introduction:

Airbnb has been a disrupter to the lodging industry and has changed the way people think about vacations and travel. Airbnb works on the premise of individuals putting their homes, rooms and spaces up for short term rentals. This allows ordinary people to cash-in on the opportunity to make a little extra cash. The question at hand is what is the best recipe to maximize profit for an entrepreneur?

Inspiration:

Initial research indicated that there has been a steady increase in both number of rentals and rates from 2015 through 2020 globally as indicated on opensoft's website. The question at hand is would Texas trend in the same direction as the global trend, and can we project how Airbnb would perform in the future.

Hypotheses:

Our goal of this assessment is to prove or disprove the following hypotheses:

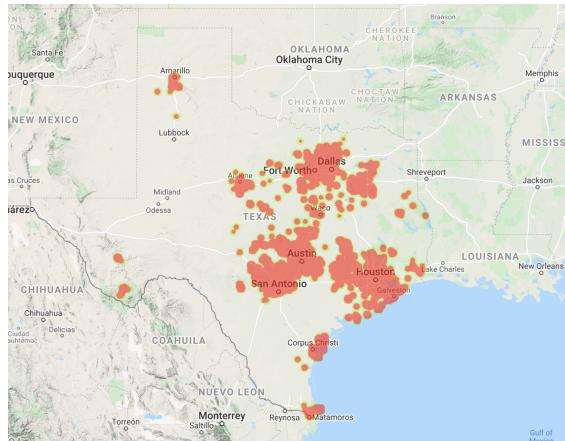
- There is less competition in rural areas due to higher rates.
- There are more rentals in areas with tourism attractions.
- Population and income affect Airbnb rates.

Sources of Data:

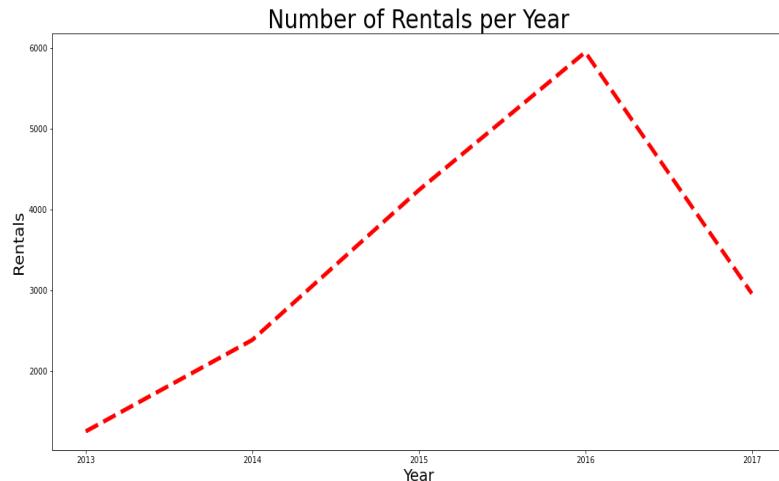
The data used for this analysis was acquired from Kaggle. We obtained two CSV files which included: Airbnb Property Data from Texas and US Census Demographic Data. This data included information such as county, room count, rate, population, household income and commute time. Combining this information with Google Maps API we were also able to create heatmaps and map markers for our data.

Analysis:

The following heatmap was created based on our Airbnb Counties csv that was narrowed down from the Airbnb Property Data from Texas. All other states were removed and we were left with the following information: Average Rate, Bedroom Count, City, Date of Listing, Latitude, Longitude, title and URL. The map shows all Texas Airbnb rentals from 2013 - 2017.



First, looking at the overall data we noticed that from 2013 through 2016 there had been a sharp increase in the number of rentals per year. This indicated that the popularity of Airbnb had been steadily increasing. In 2016 our diagram shows a sharp decrease in the number of rentals, this is due to only analyzing half of the year based on the data found. Later in the analysis we will dive further into whether we expect to see the same trend in 2017 as in the prior years.



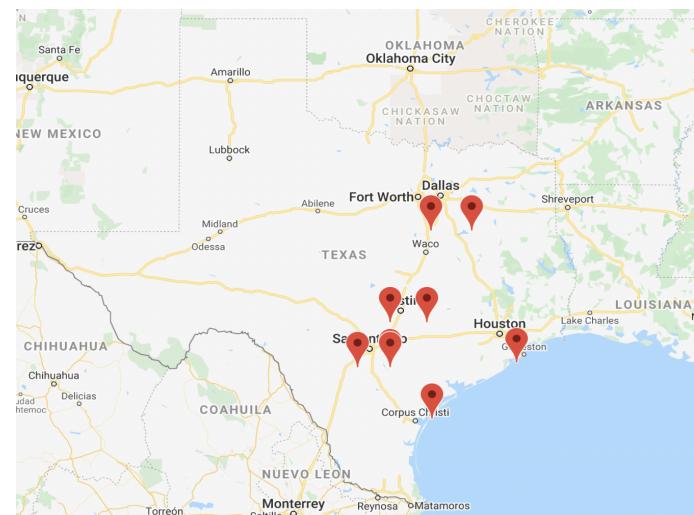
It was decided that we would look at the top five urban and rural counties as they comprised 52.5% of the overall data and provided a good representation of the overall data.

Urban counties consisted of the following:

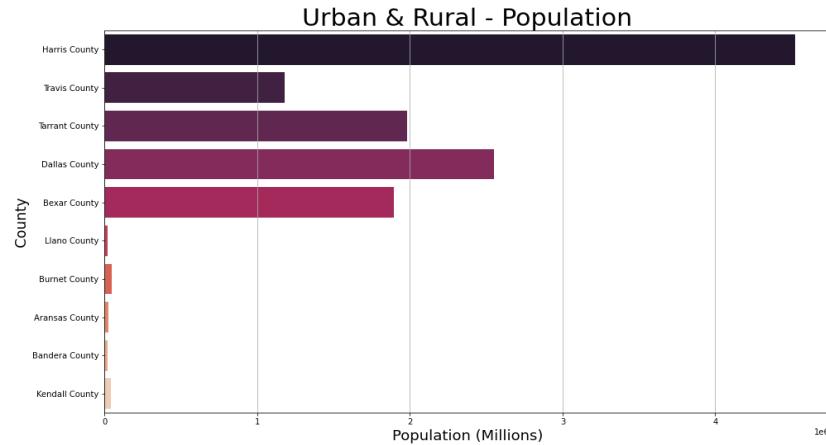
- Harris County
- Travis County
- Tarrant County
- Dallas County
- Bexar County

Rural counties consisted of the following:

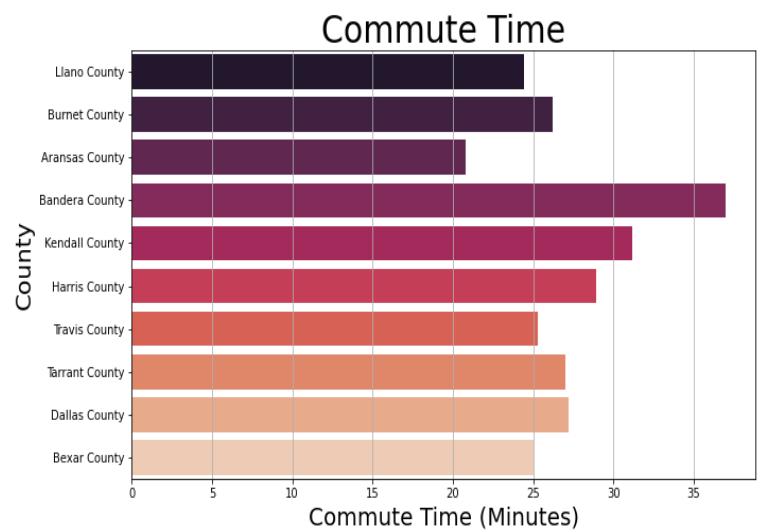
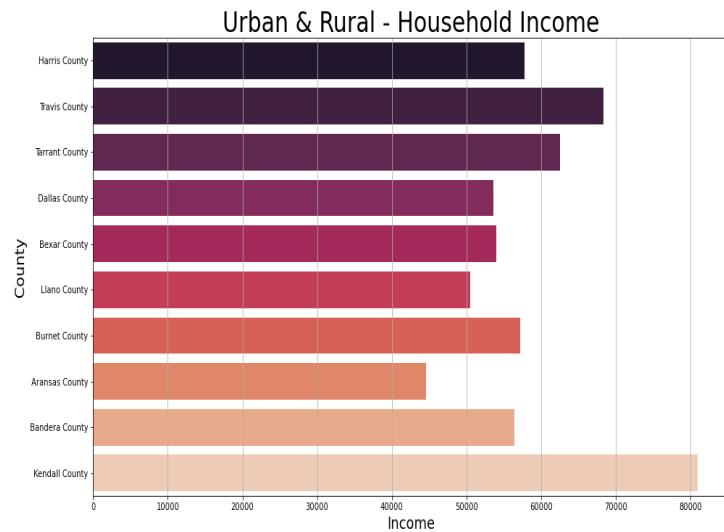
- Llano County (NW of Austin)
- Burnet County (NW of Austin)
- Aransas County (NE of Corpus Christi)
- Bandera County (NW of San Antonio)
- Kendall County (NW of San Antonio)



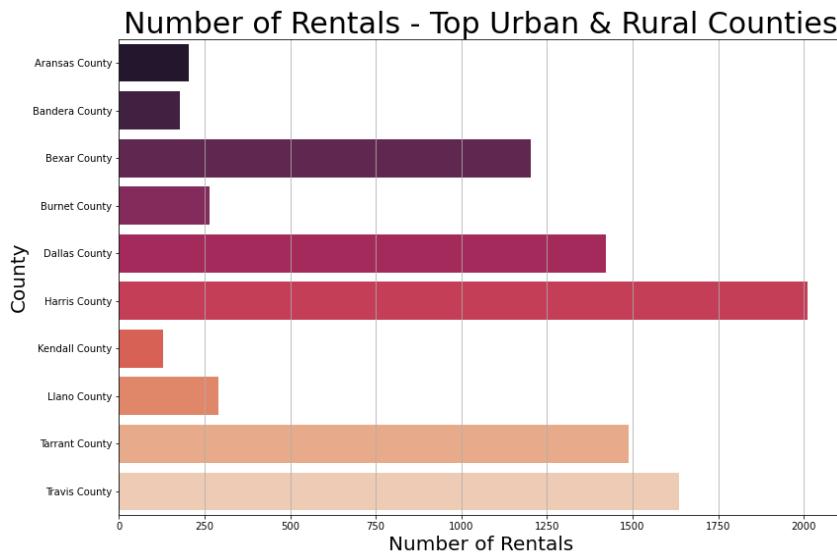
One question we had to address was what makes a rural county rural? This is decided by population. If a county has less than 50,000 people, we would consider it rural. There was a drastic difference between populations in urban and rural areas with the highest urban population of 1.9M people and the highest rural population of 45K.



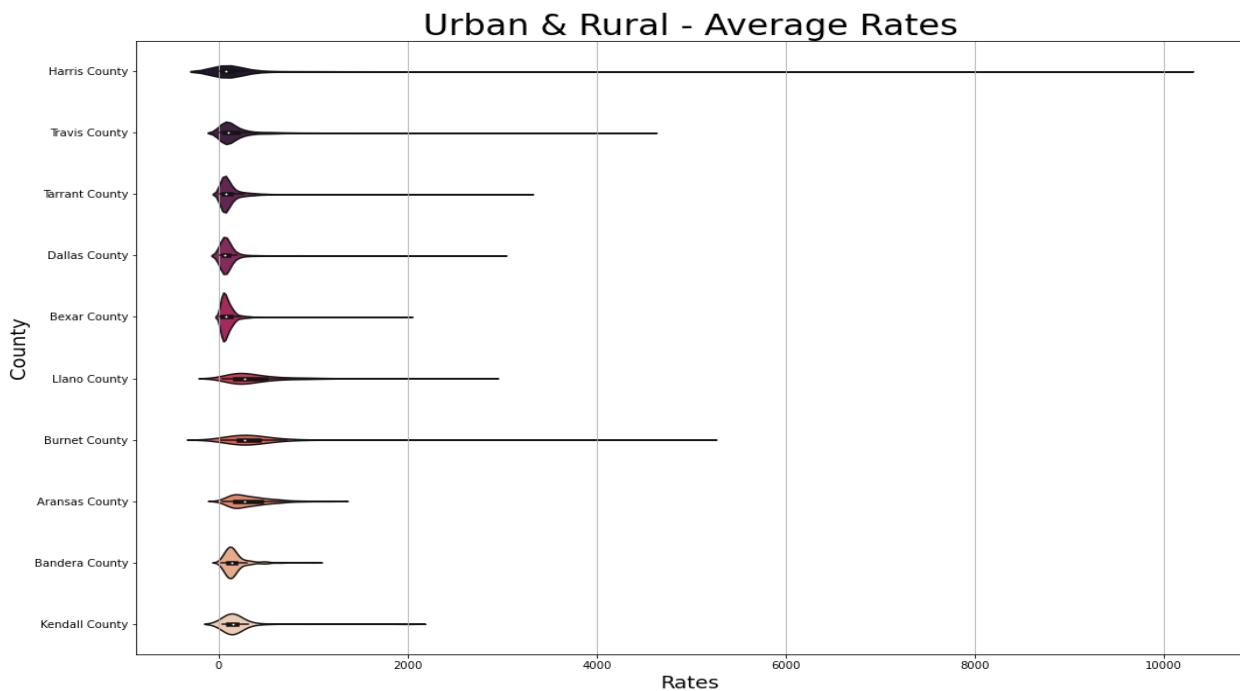
The data also indicated that the highest populated rural counties were around major metropolitan areas. Due to the proximity of the rural counties, we speculate that the salaries and commute times closely emulate those of urban and rural counties.

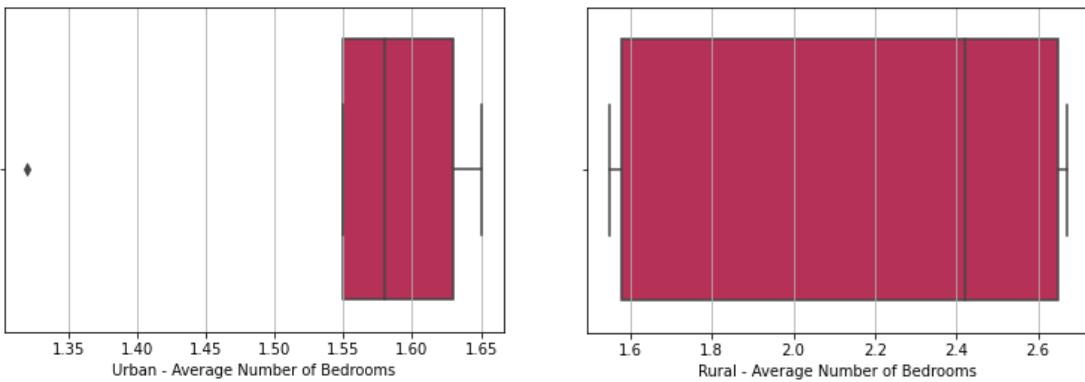


The number of rentals were slightly higher in urban counties than those of rural counties. The higher the population and housing density would allow for more opportunities to have rentals available.

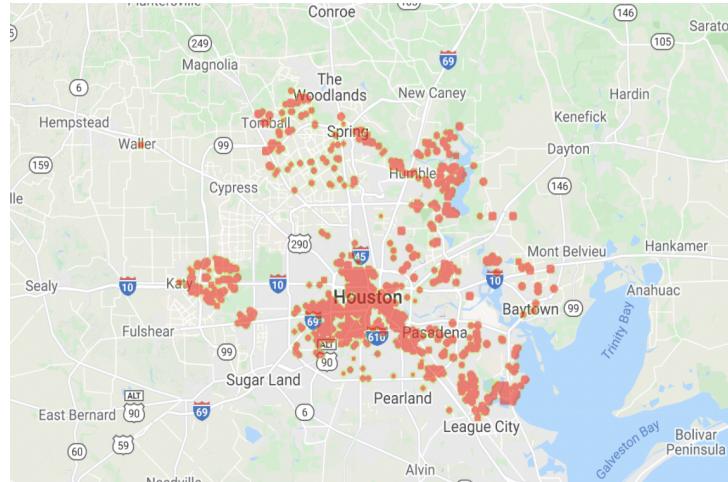


The average rental rate chart below indicates that the average rate in rural areas is slightly more than those in urban areas. When looking into the causes of this we found that there are more rentals in urban areas creating more competition and driving the rates down. When looking at the size of the rental we took the average number of bedrooms per county. It was noticed that there was an increase in the number of studio apartments in urban areas, which contributed to a lower average bedroom size. The combination of these factors resulted in lower rates.

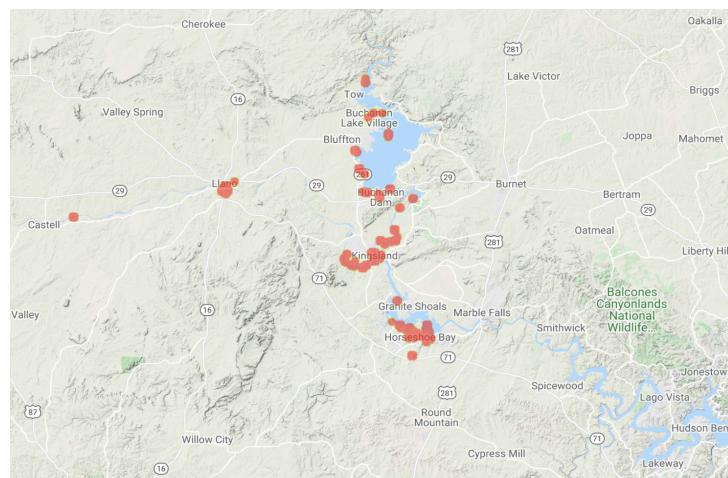




The Top number of rentals for an Urban County is Harris County which is the Houston area. There is a density of rentals with clusters in densely populated areas and near bodies of water. Background information about Houston is as follows: it is the location of 2 major airports, the 4th largest city in the nation, the world capital of the international energy industry, and Houston is a busy metropolitan area with many attractions.



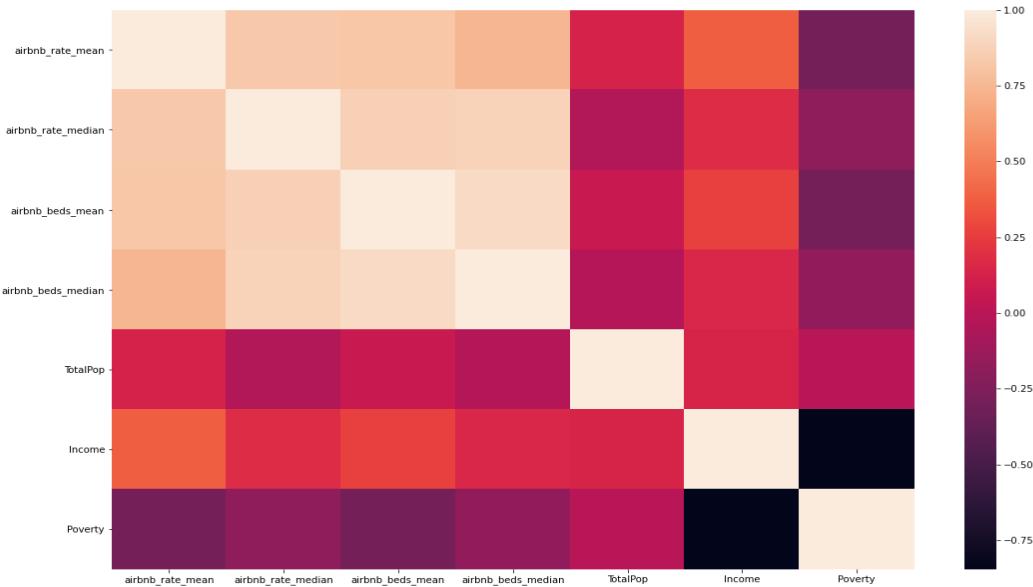
The Top number of rentals for a Rural County is Llano County which is northwest of Austin. Llano county is known as the Deer capital of Texas. There is a density of rentals with clusters near bodies of water and Kingsland, Texas. Kingsland is good for boating, fishing, swimming, golf, and picnics. Also located in Llano County is Buchanan Dam and the Colorado River. Rentals here could be lake house properties.



Correlations:

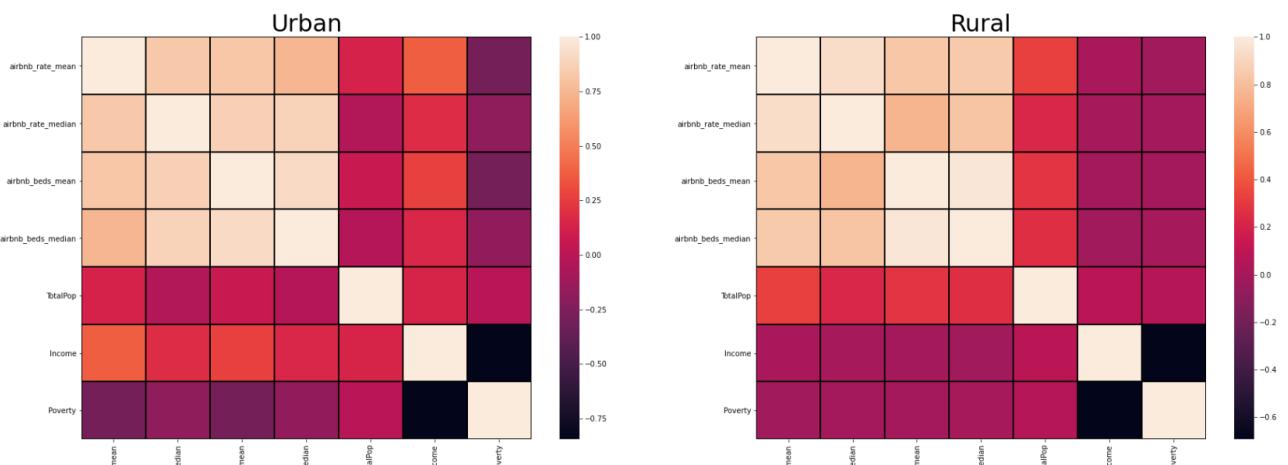
When we wanted to analyze the data to see if we could find correlations. The first thing we wanted to look at is the correlations of variables in the urban county tables. Looking at this heatmap we can see the correlations of income and rate. We also see the correlations of income with the average number of beds. This means we can show a relationship between income and rates and beds.

Correlation Heatmap - Urban



After recognizing that we wanted to look at the correlations between urban and rural areas and how those correlations compare and differ. It was in this data we confirmed that Airbnb rates are based on different conditions in urban and rural areas. Urban areas have the highest correlation when it comes to income, while rural areas have the highest correlation on population. We wanted to drill down and investigate this further.

Correlation Heatmaps Urban vs Rural



Urban Correlations with OLS Regression

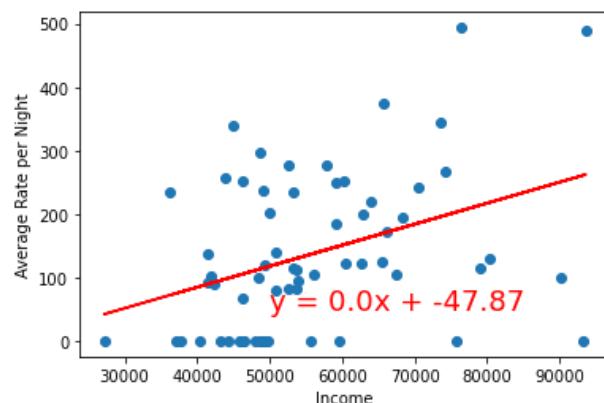
The data is shown to have a slight positive correlation between income and rates as indicated by the r-squared value. An r-value below 0.3 indicates no correlation between two variables, our data shows an r-squared value of 0.14 which results in an r-value of 0.374. This output and resulting graph indicates there is a small positive correlation between income and rates which trends upwards with income as rates increase.

OLS Regression Results						
Dep. Variable:	airbnb_rate_mean	R-squared:	0.140			
Model:	OLS	Adj. R-squared:	0.126			
Method:	Least Squares	F-statistic:	9.907			
Date:	Wed, 04 Aug 2021	Prob (F-statistic):	0.00255			
Time:	14:01:18	Log-Likelihood:	-387.89			
No. Observations:	63	AIC:	779.8			
Df Residuals:	61	BIC:	784.1			
Df Model:	1					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	-47.8696	60.496	-0.791	0.432	-168.839	73.100
Income	0.0033	0.001	3.147	0.003	0.001	0.005
Omnibus:	2.085	Durbin-Watson:	1.911			
Prob(Omnibus):	0.352	Jarque-Bera (JB):	2.035			
Skew:	0.413	Prob(JB):	0.362			
Kurtosis:	2.695	Cond. No.	2.38e+05			

Notes:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 2.38e+05. This might indicate that there are strong multicollinearity or other numerical problems.

Regression Chart:



OLS Regression Results-Rural

OLS Regression Results						
Dep. Variable:	airbnb_rate_mean	R-squared:	0.104			
Model:	OLS	Adj. R-squared:	0.099			
Method:	Least Squares	F-statistic:	21.85			
Date:	Wed, 04 Aug 2021	Prob (F-statistic):	5.59e-06			
Time:	16:38:15	Log-Likelihood:	-1208.6			
No. Observations:	191	AIC:	2421.			
Df Residuals:	189	BIC:	2428.			
Df Model:	1					
Covariance Type:	nonrobust					
<hr/>						
	coef	std err	t	P> t	[0.025	0.975]
const	19.8261	15.174	1.307	0.193	-10.107	49.759
TotalPop	0.0035	0.001	4.674	0.000	0.002	0.005
<hr/>						
Omnibus:	183.991	Durbin-Watson:	2.071			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	4608.124			
Skew:	3.584	Prob(JB):	0.00			
Kurtosis:	25.971	Cond. No.	3.13e+04			
<hr/>						

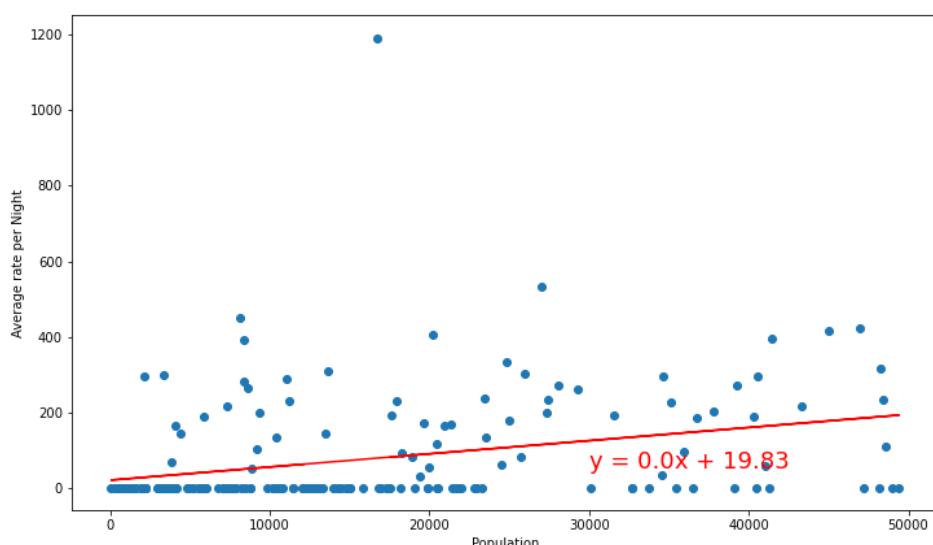
Notes:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 3.13e+04. This might indicate that there are strong multicollinearity or other numerical problems.

Like the urban regression results, the rural regression results also show a slight positive increase. The r-squared is .104 which means the r-value is .322. This still means that there is a positive correlation between population and rates.

Regression Chart-Rural

The following Regression Chart for Rural counties shows a positive slope in the trendline. This means that in rural areas, when the population increases the average rate per night also increases.



Conclusion:

For our hypotheses we can now conclude the following: we recommend operating an airbnb rental in a rural area outside a major metropolitan center that has outdoor destinations to maximize profit.

From the correlation data, we can see that rural areas that have a larger total population have higher rates than urban areas. Moreover, urban area rates are affected by income rather than population. This would mean that it is best to target rural counties with a total population of 40,000 to 50,000 people.

Limitations:

There were a few limitations that we discovered in our data sets. The census data was limited to 2017, which was out of date. Our airbnb data contained data from 2013 through only the first two quarters of 2017, which was not ideal when analyzing the data. We also noticed that our data did not include all the counties in Texas (only 106 out of 254 counties). This was easily identified in our heat layer map data when we observed a lack of Airbnb rentals in both Midland and El Paso.

Future Work:

Through this analysis we were able to come up with solid recommendations for our entrepreneurs, but it also brought up some additional questions and additional analysis that we would like to explore further. Some topics that we would like to explore are as follows:

- How has the pandemic affected the airbnb market?
- How long will it take to recover from the pandemic?
- What is the best time of year to list a property (i.e. season and holidays)?
- Redefine what is considered rural counties, and see if there are similar trends the farther you get away from metropolitan areas.
- Study the impact of local laws concerning short term rentals.

References:**Datasets:**

<https://www.kaggle.com/PromptCloudHQ/airbnb-property-data-from-texas>

<https://www.kaggle.com/muonneutrino/us-census-demographic-data>

Inspiration:

<https://www.kdnuggets.com/2017/08/insights-data-mining-airbnb.html>

<https://public.opendatasoft.com/explore/dataset/air-bnb-listings/map/>

Graphics:

https://www.traveltexas.com/things-to-do/outdoor-adventure/?gclid=CjwKCAjw0qOlBhBhEiwAyyVcfzmOWPSjpt_oiR-jt5UaiZhYC_qyne15B3Tc1uD5HuzdE6kB0FIRtBoCUPAQAvD_BwE

<https://www.airbnb.com>

Definition of Urban vs Rural

<https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural/2010-urban-rural.html>