Airbnb Project

Diana Melendez, Ericka Houle and Steven Green

Agenda

Hypothesis

Data

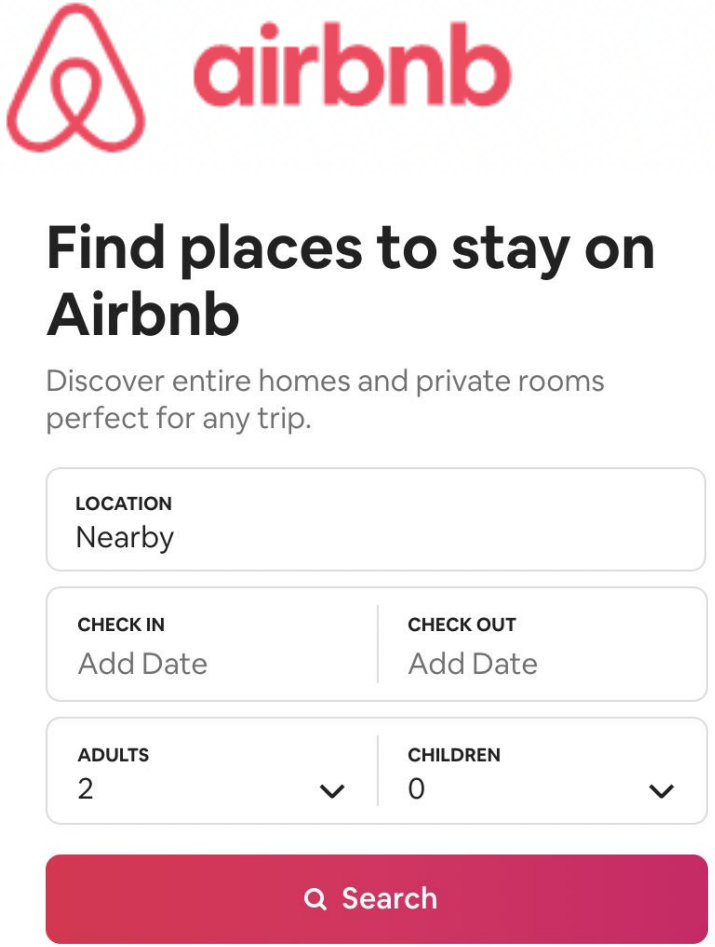
Inspiration

Analysis

Projection

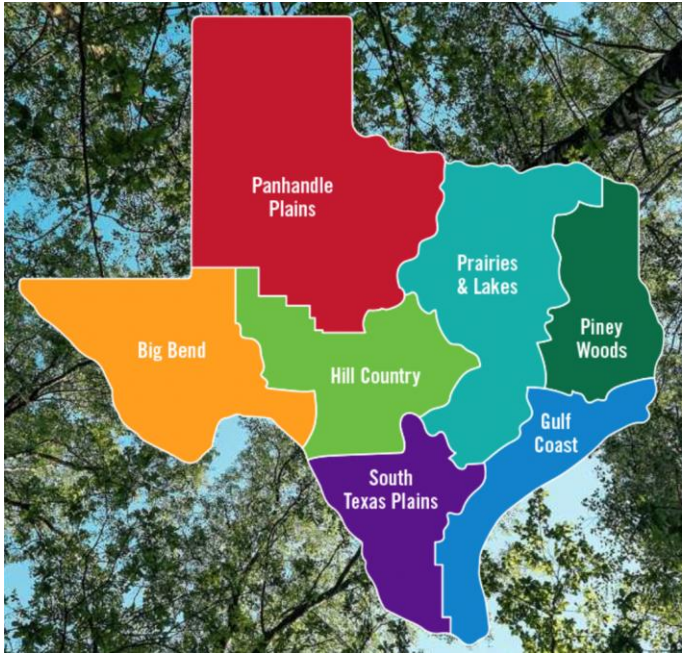
Limitations & Future Analysis

Q& A



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.



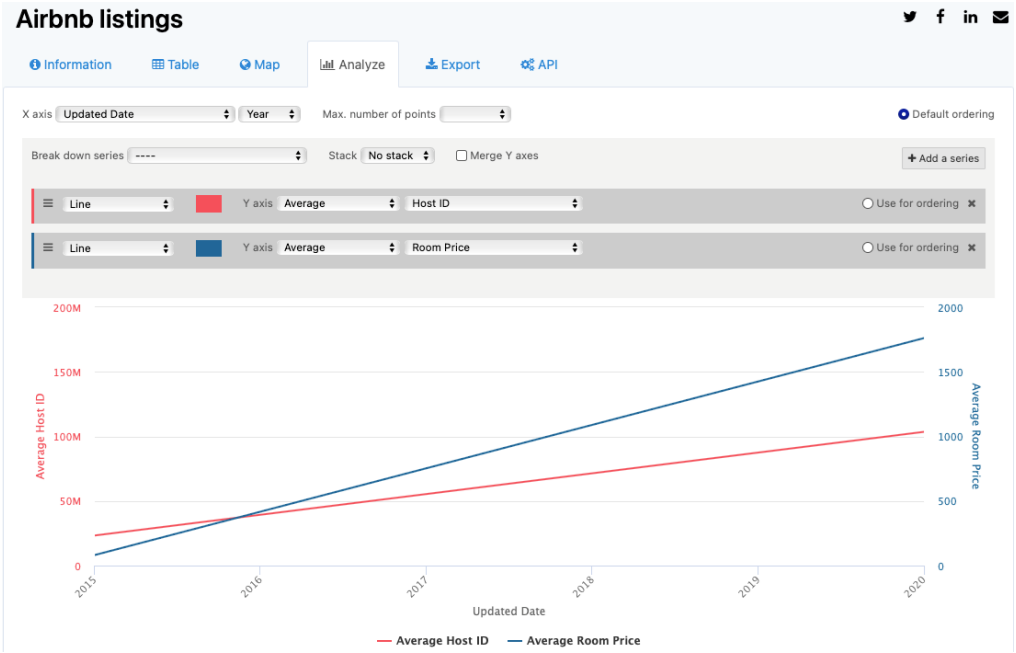
Datasets

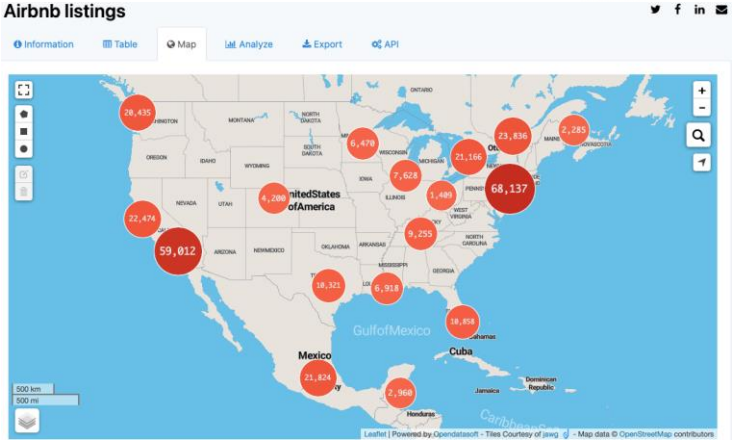
kaggle

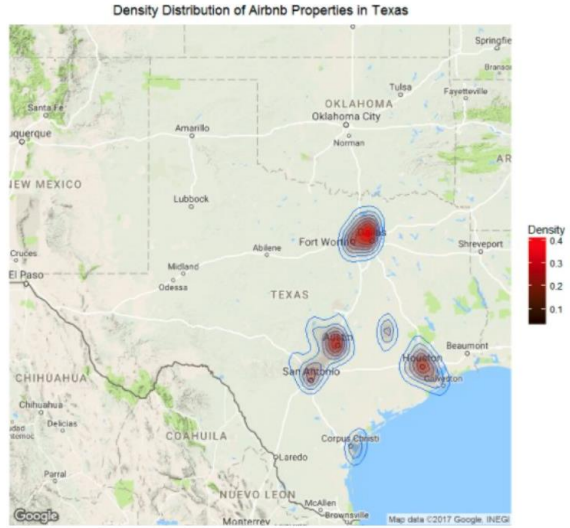
Inspiration

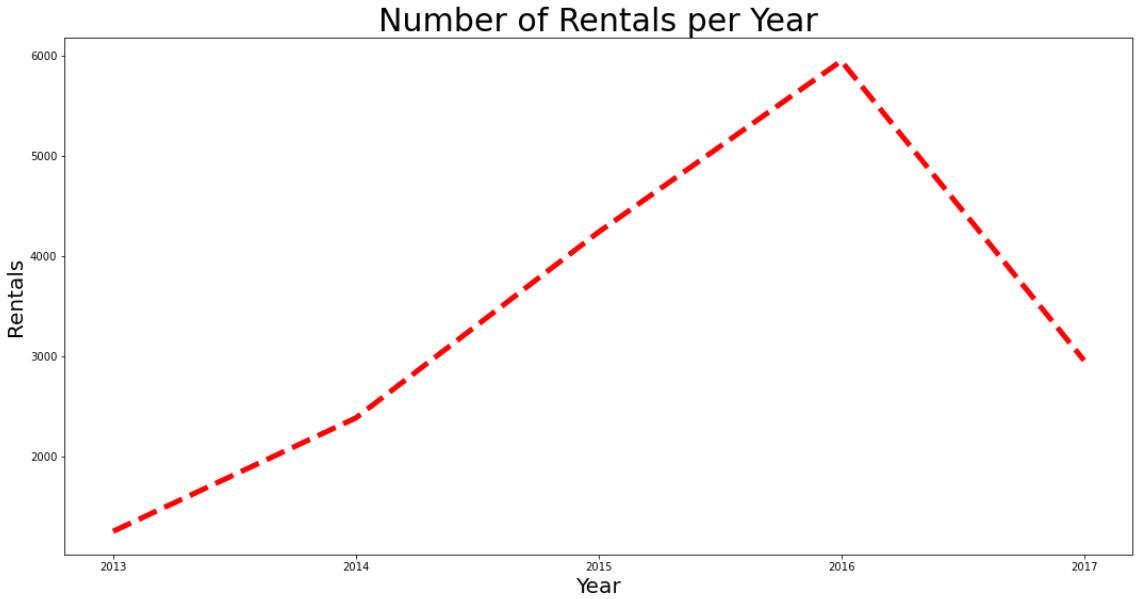
opendatasoft

KDnuggets™

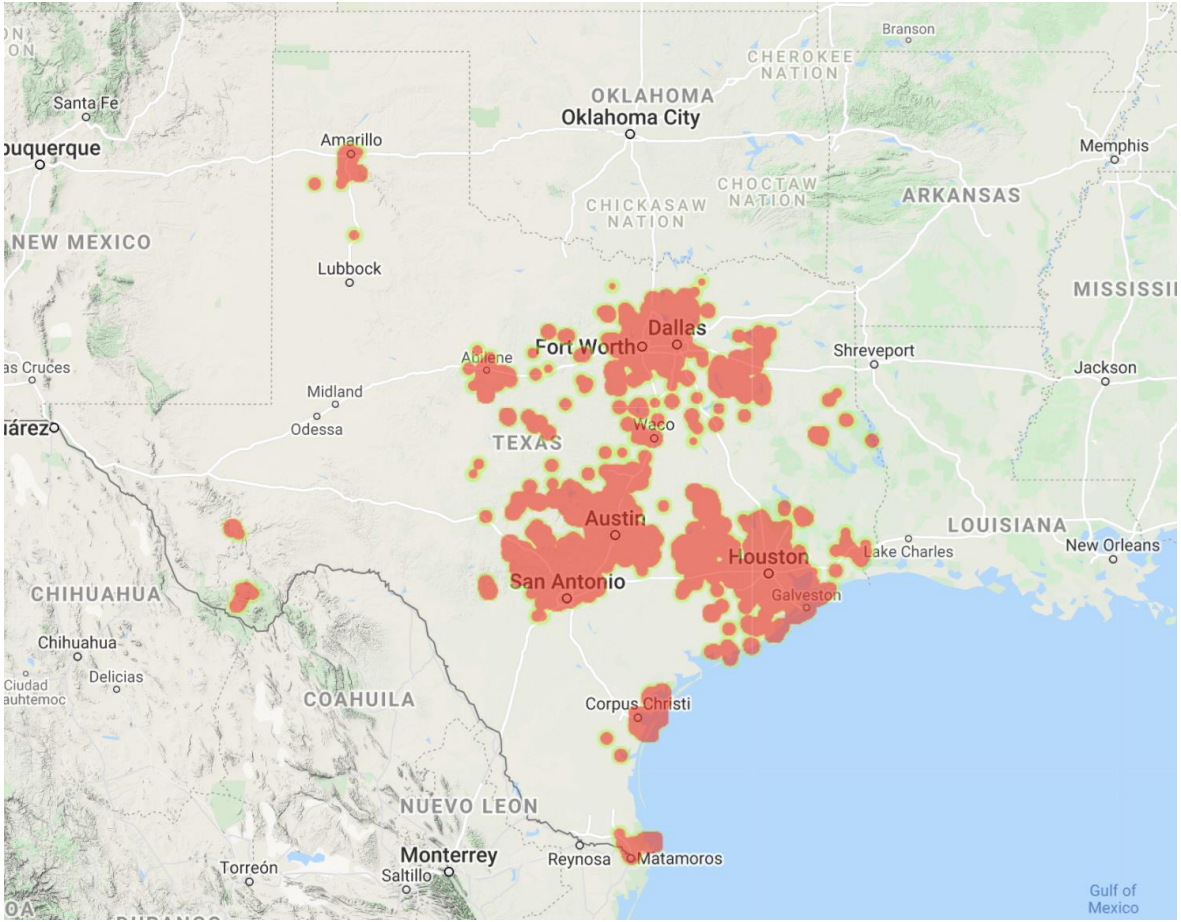







Number of Rentals per Year
Rentals
6000
5000
4000
3000
2000
2013
2014
2015
2016
2017
Year

Texas Airbnb Rentals



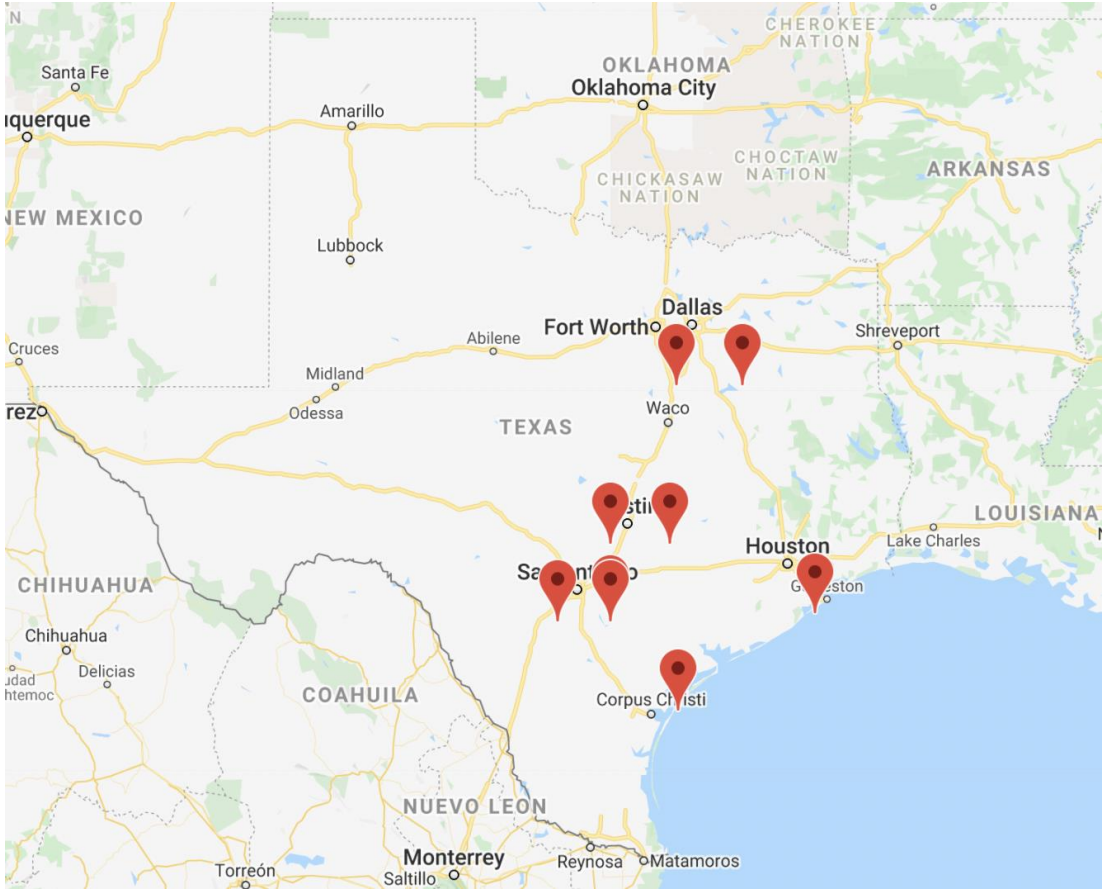
Top 5 Urban & Rural Rentals

Urban Counties

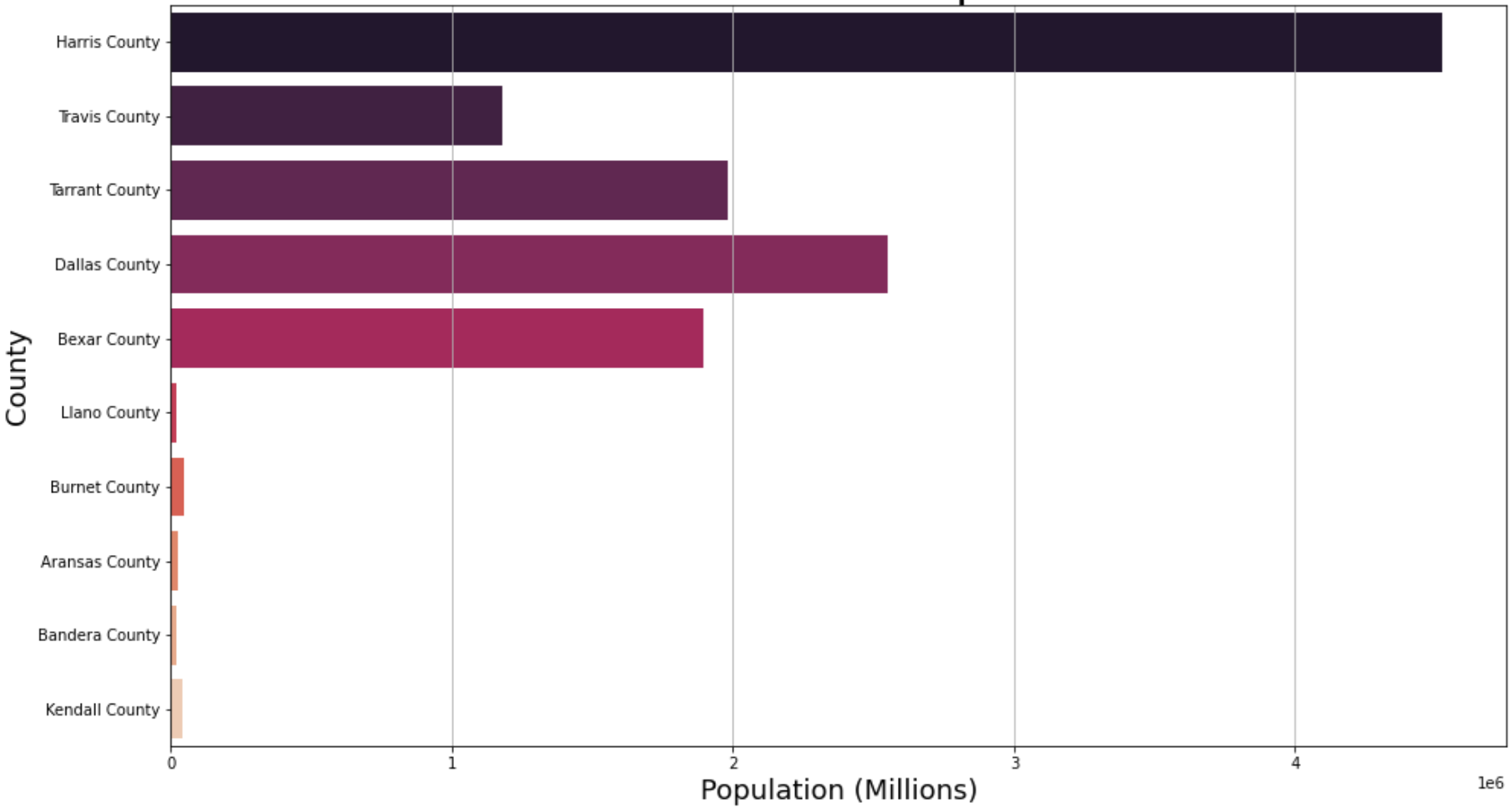
- Harris County
- Travis County
- Tarrant County
- Dallas County
- Bexar County

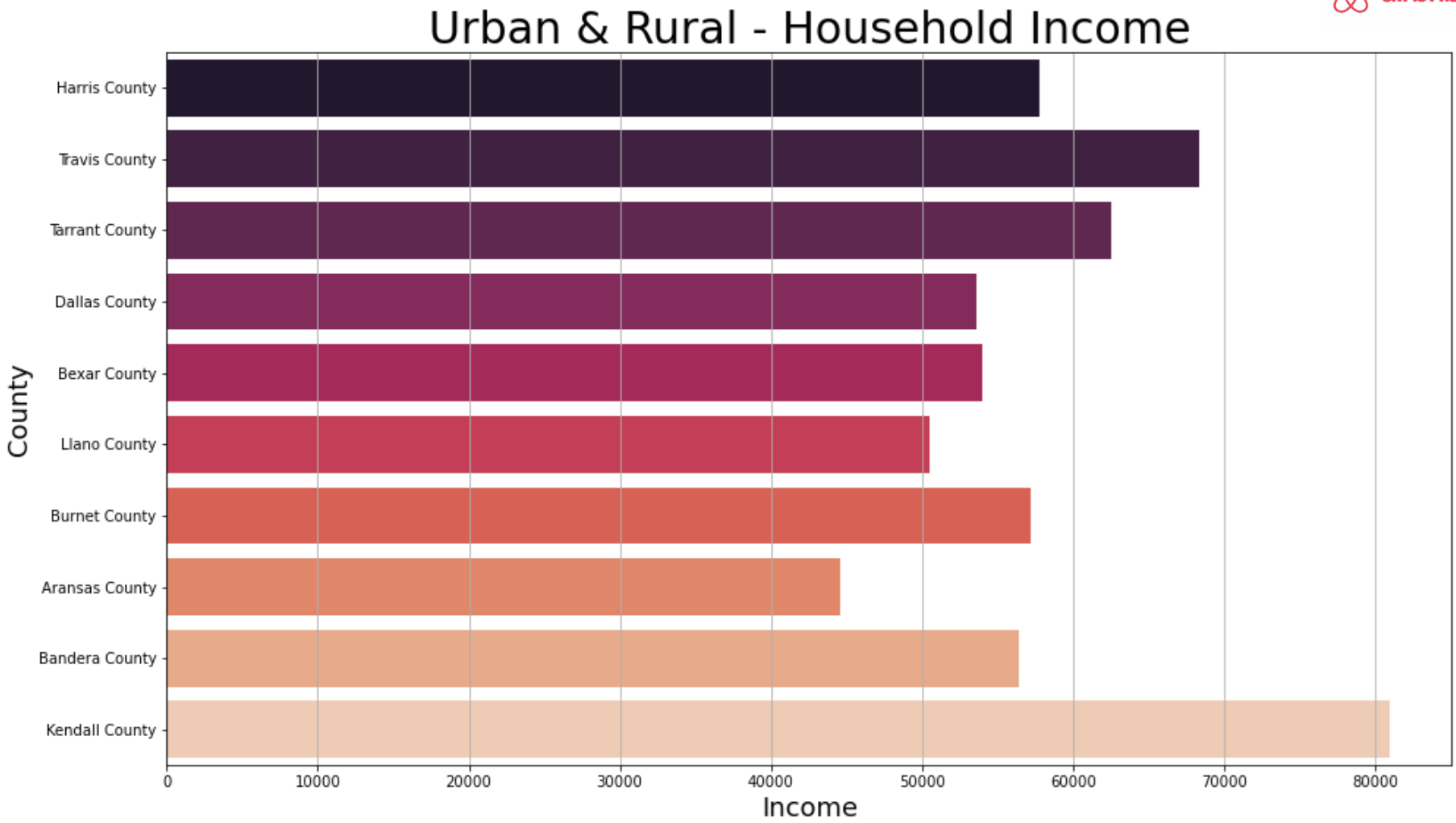
Rural Counties

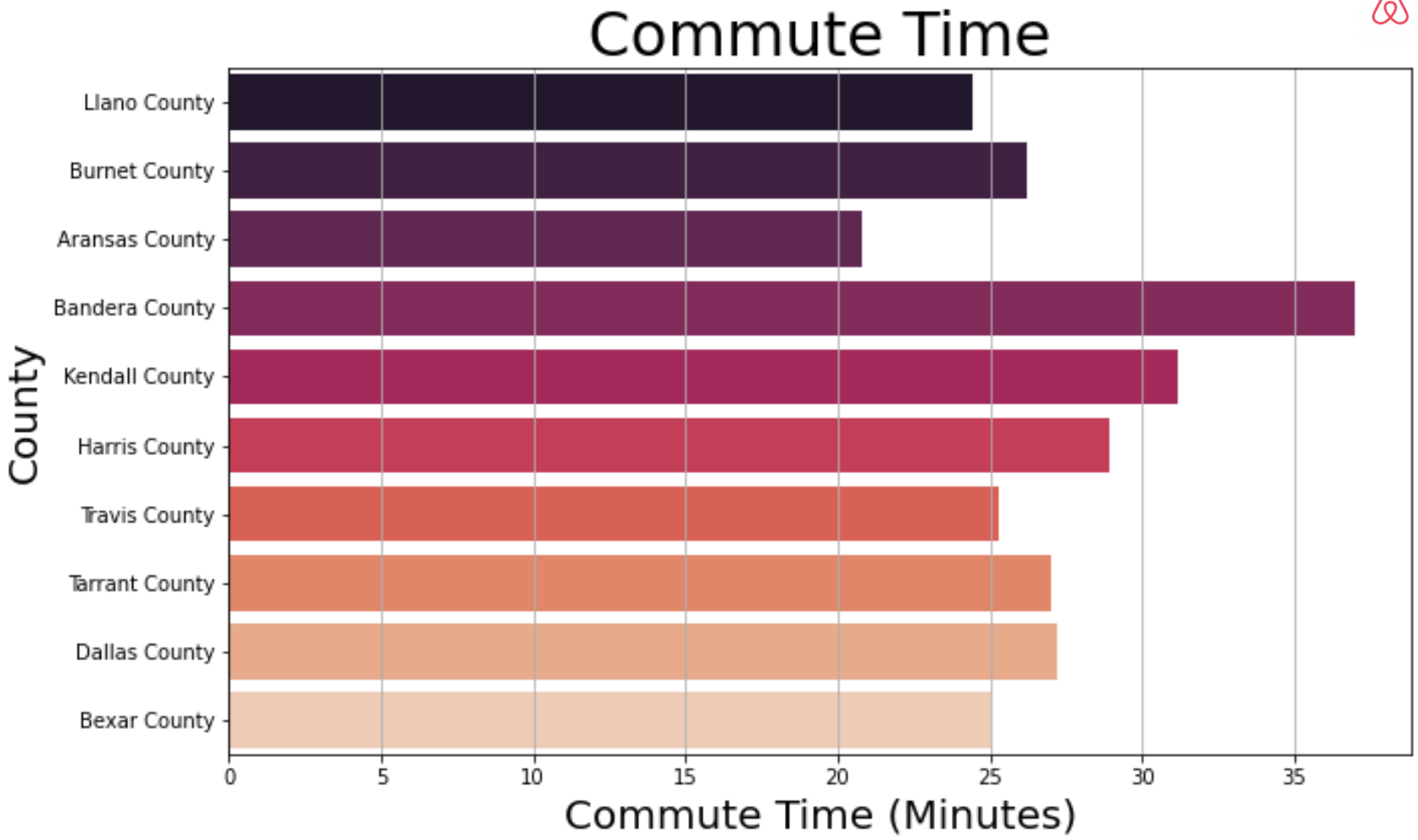
- Llano County
- Burnet County
- Aransas County
- Bandera County
- Kendall County



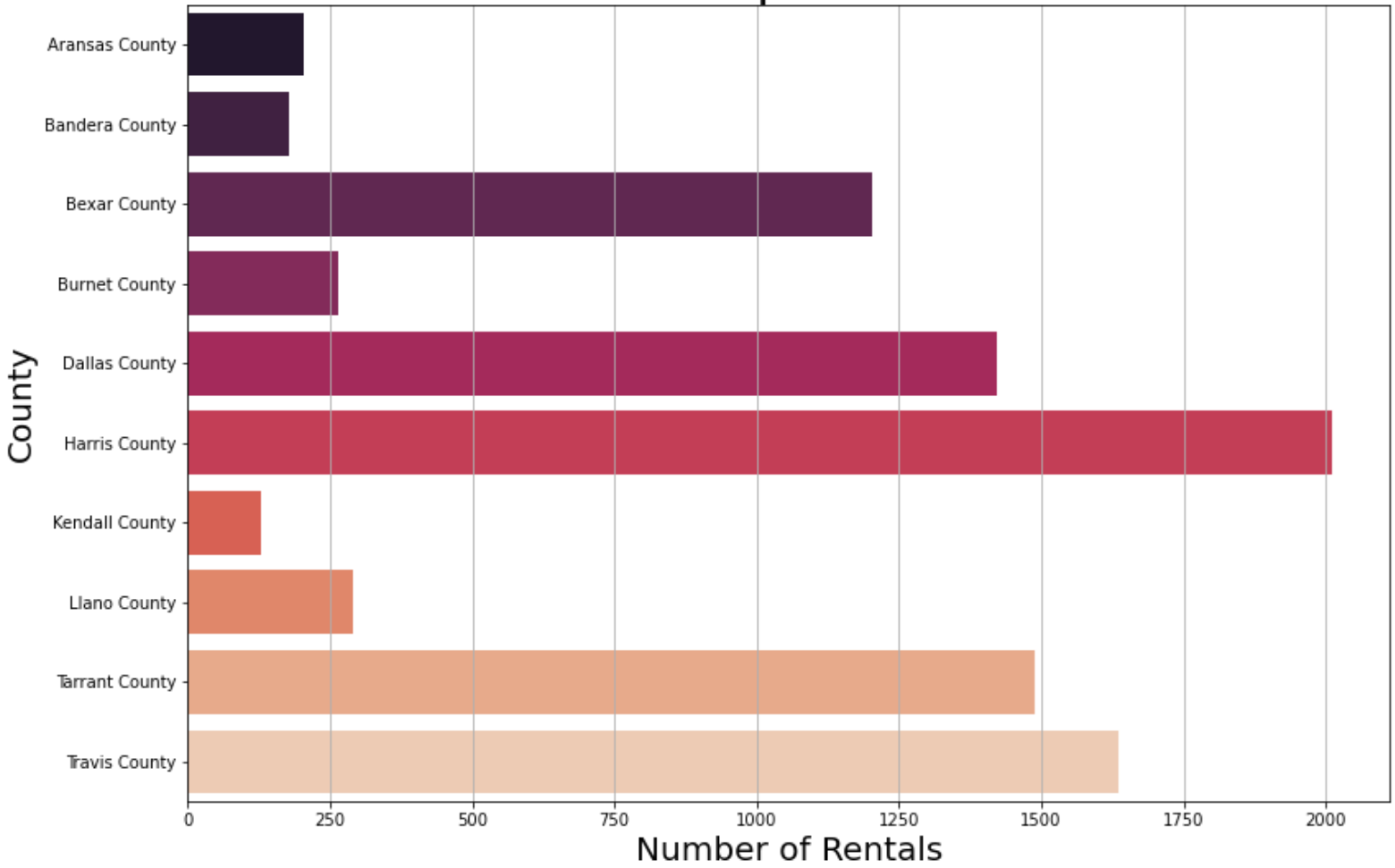
Urban & Rural - Population



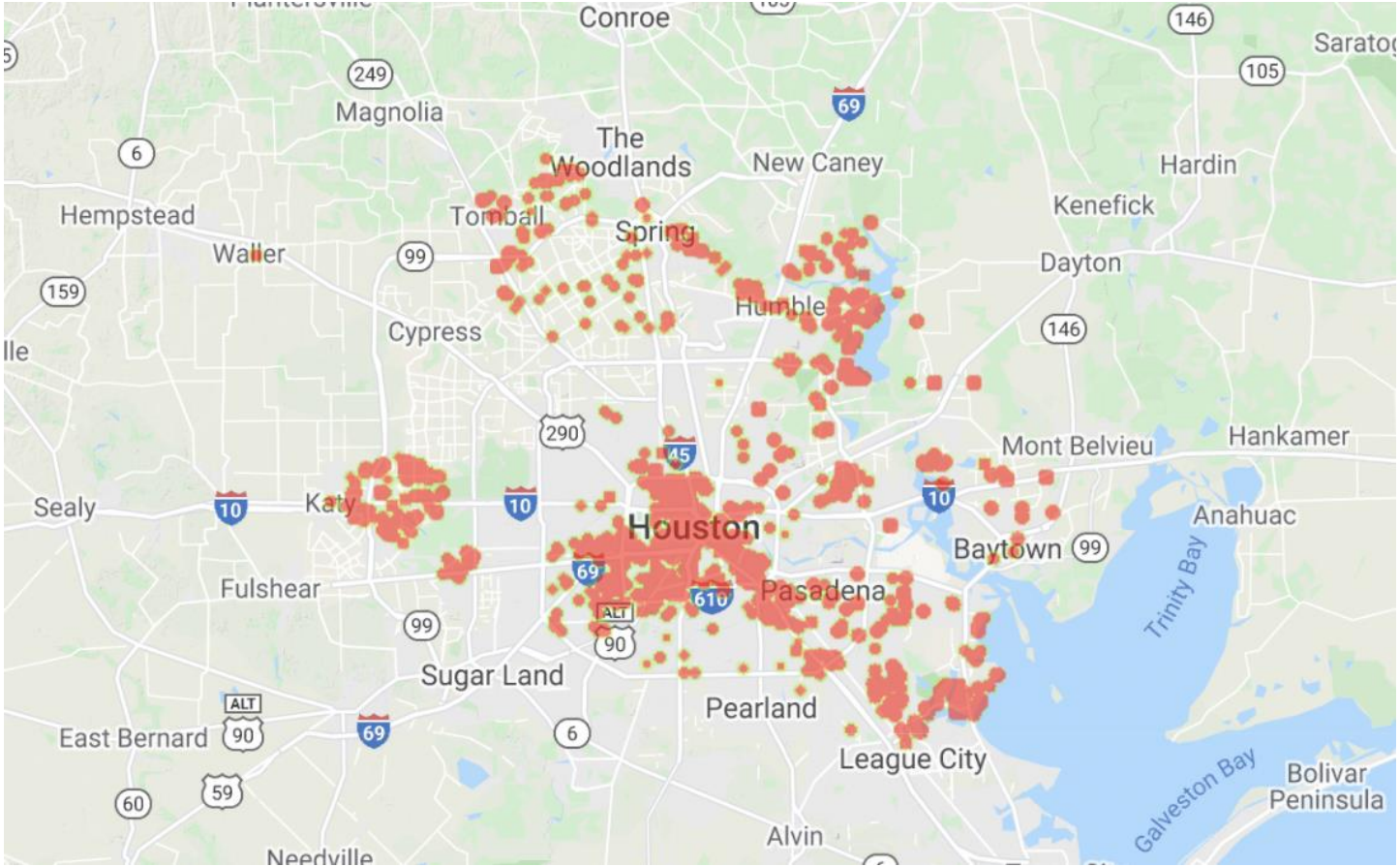

Urban & Rural - Household Income
Harris County
Travis County
Tarrant County
Dallas County
Bexar County
Llano County
Burnet County
Aransas County
Bandera County
Kendall County
County
0
10000
20000
30000
40000
50000
60000
70000
80000
Income


Commute Time
Llano County
Burnet County
Aransas County
Bandera County
Kendall County
Harris County
Travis County
Tarrant County
Dallas County
Bexar County
County
0
5
10
15
20
25
30
35
Commute Time (Minutes)

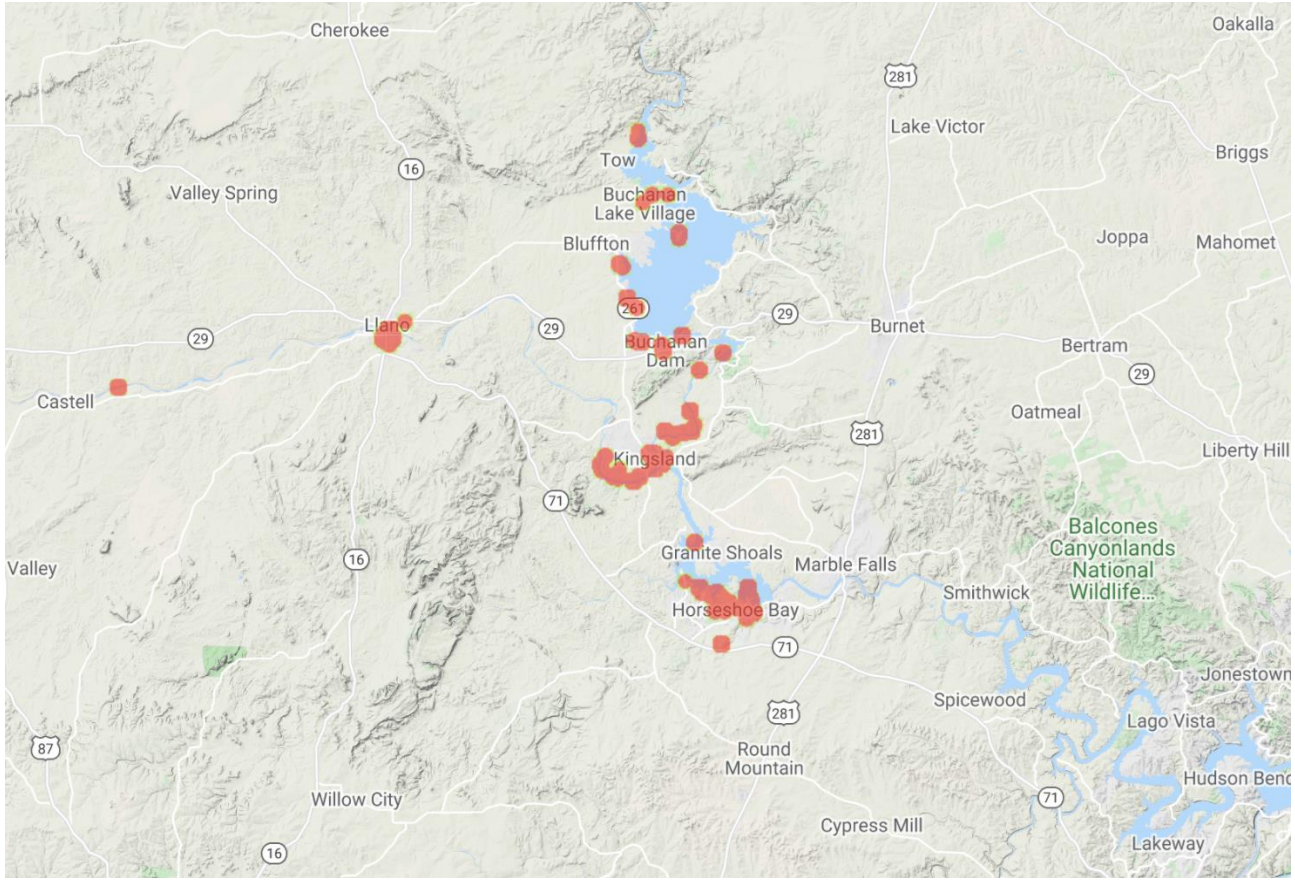
Number of Rentals - Top Urban & Rural Counties



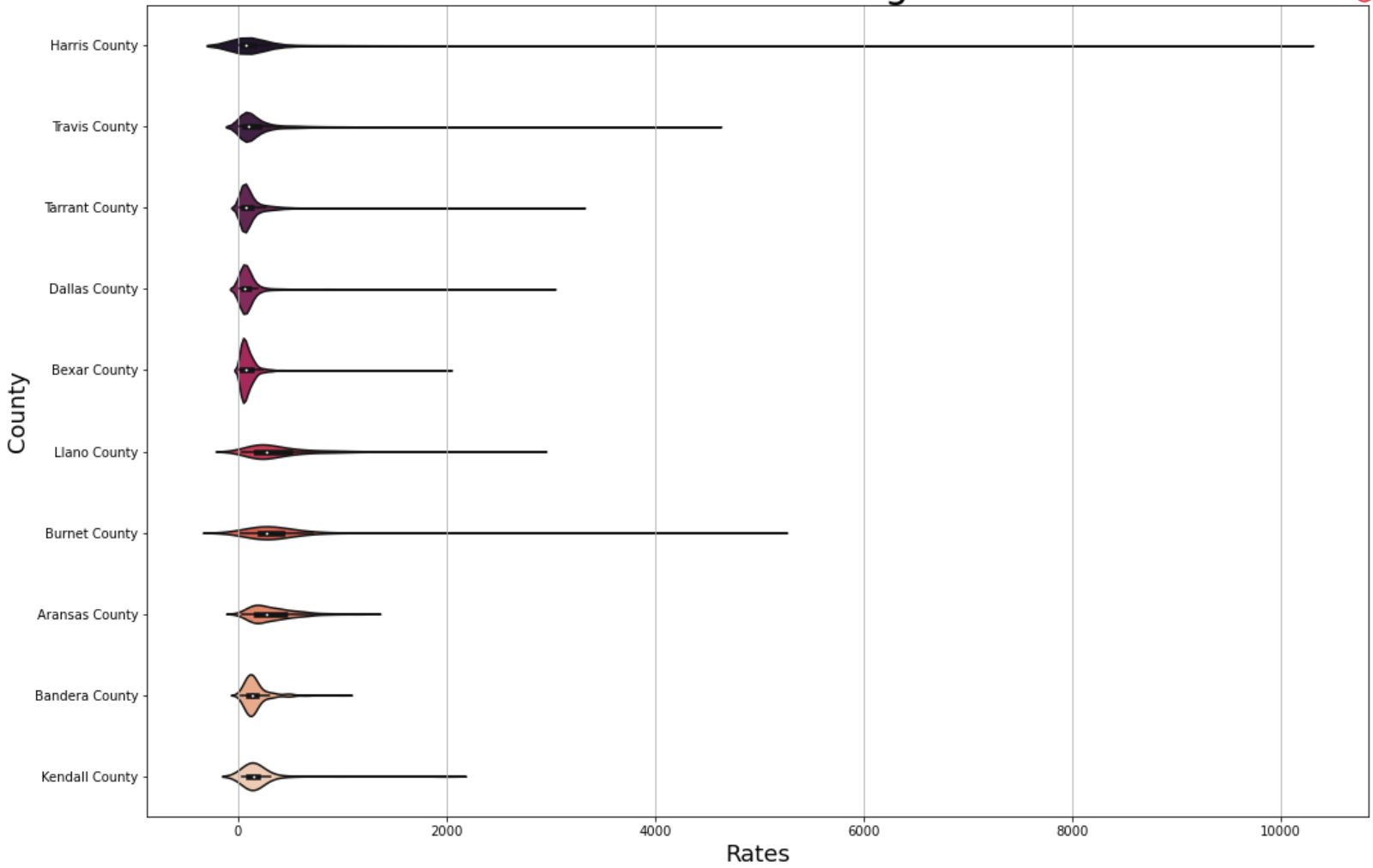
Harris County - Urban



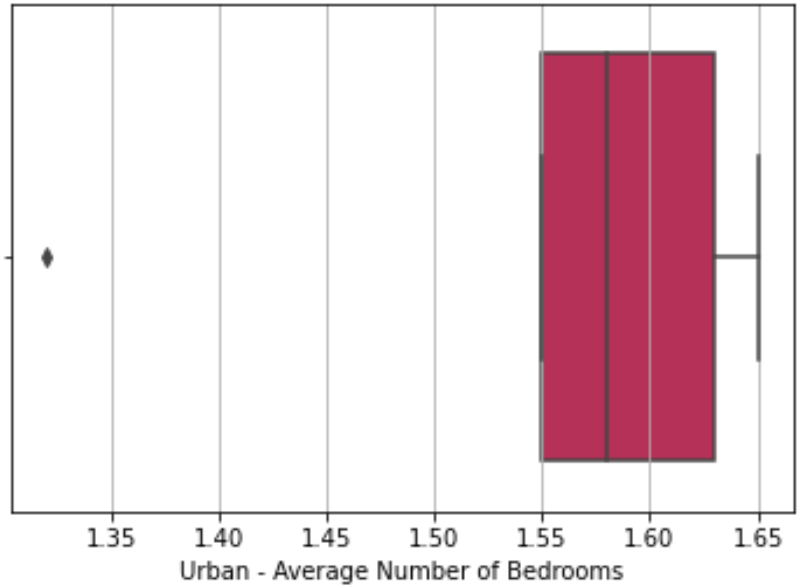
Llano County - Rural

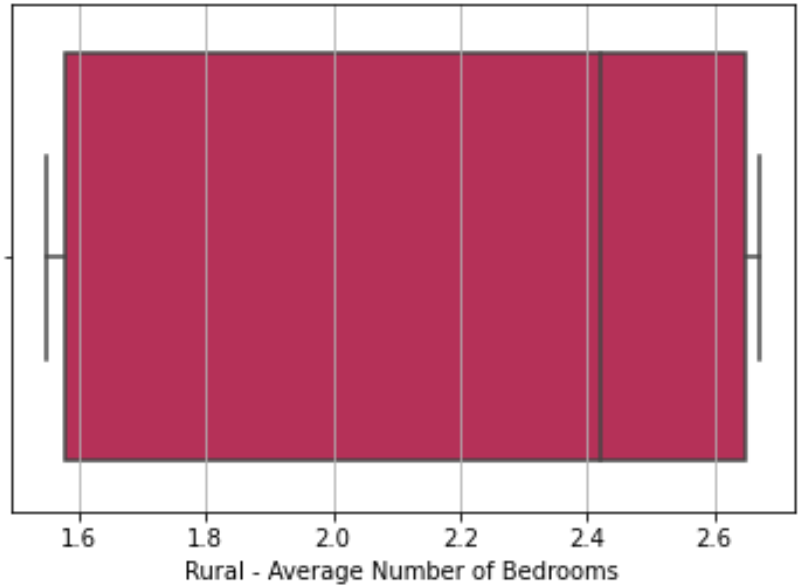


Urban & Rural - Average Rates

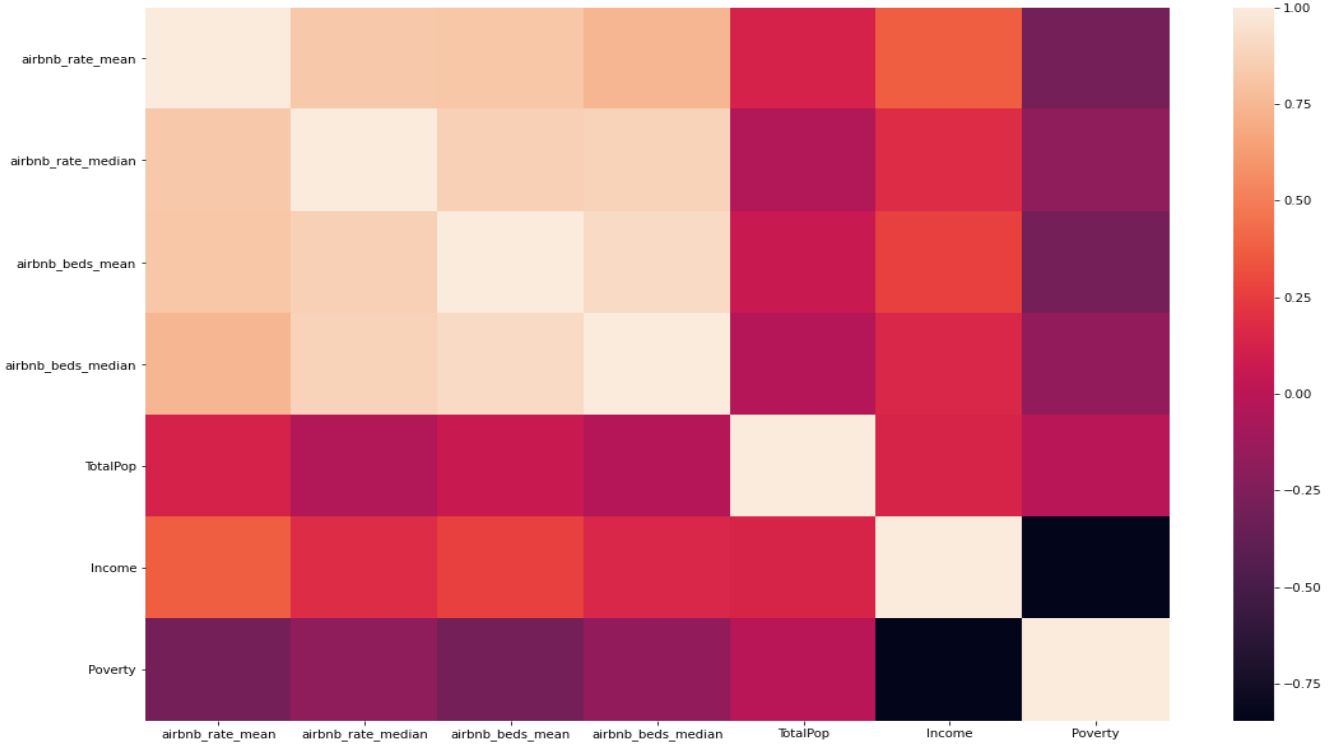


Average Number of Bedrooms

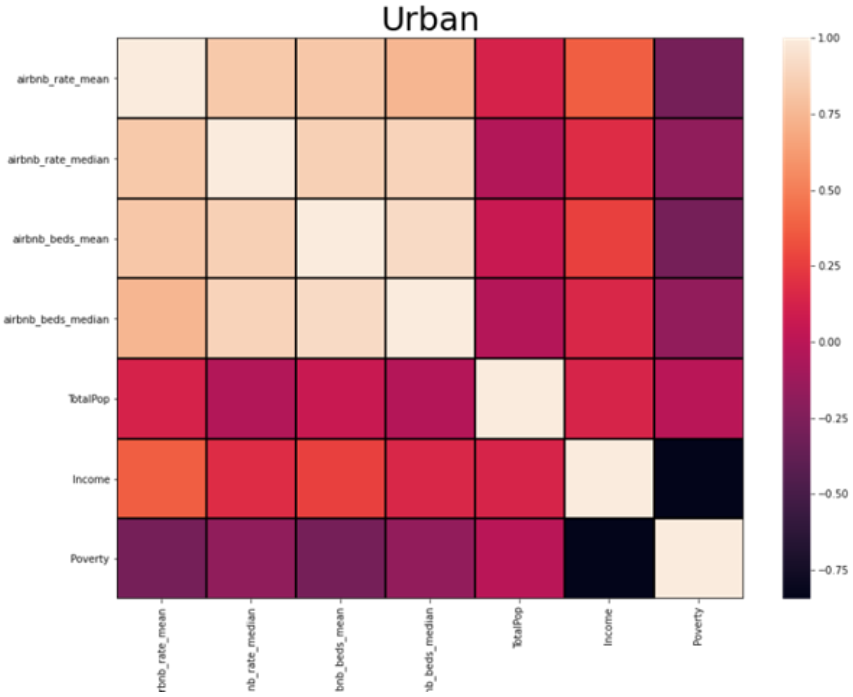




Correlation Heatmap - Urban



Correlation Heatmaps Urban vs Rural



OLS Regression Results-Urban

OLS Regression Results

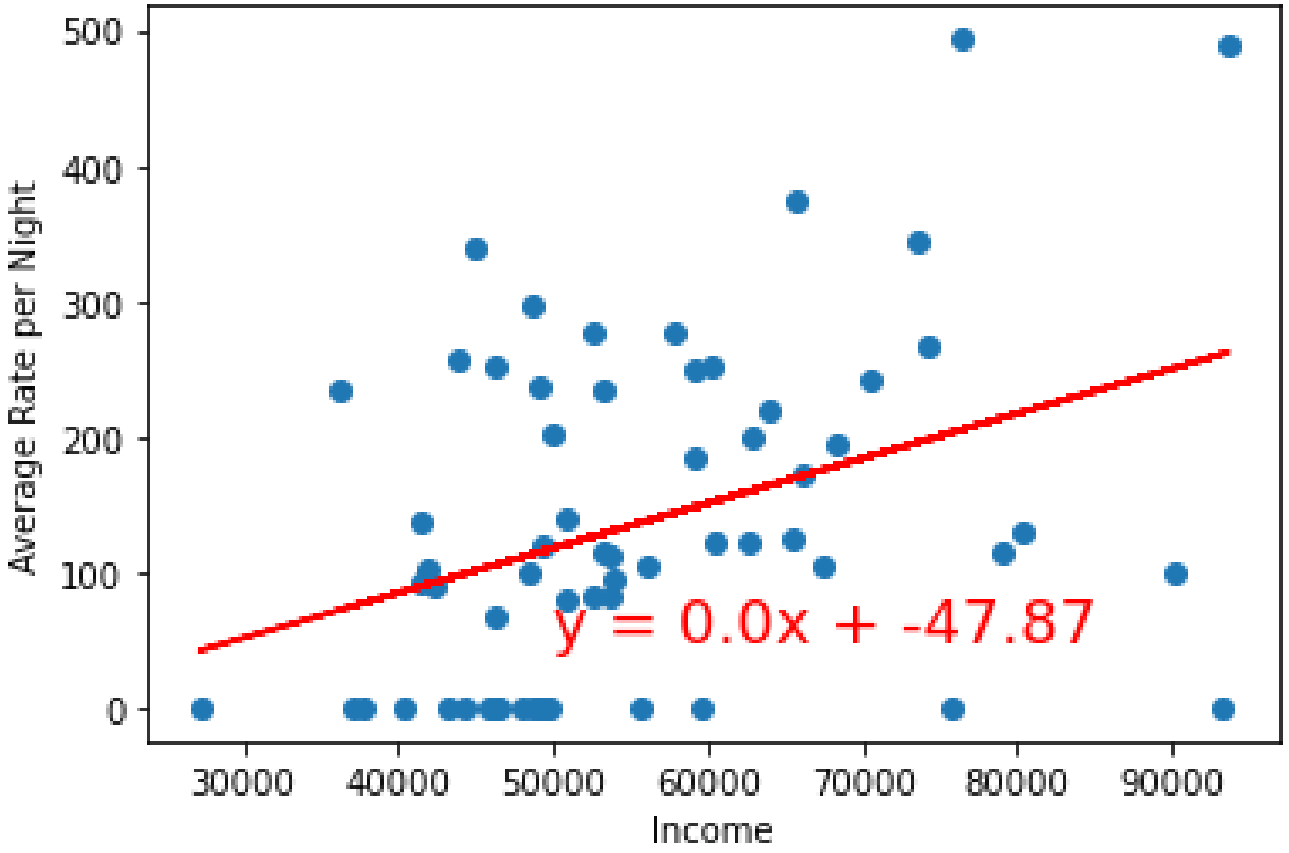
Dep. Variable:	airbnb_rate_mean	R-squared:	0.140
Model:	OLS	Adj. R-squared:	0.126
Method:	Least Squares	F-statistic:	9.907
Date:	Wed, 04 Aug 2021	Prob (F-statistic):	0.00255
Time:	14:01:18	Log-Likelihood:	-387.89
No. Observations:	63	AIC:	779.8
Df Residuals:	61	BIC:	784.1
Df Model:	1		
Covariance Type:	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-Urban



OLS Regression Results-Rural

OLS Regression Results

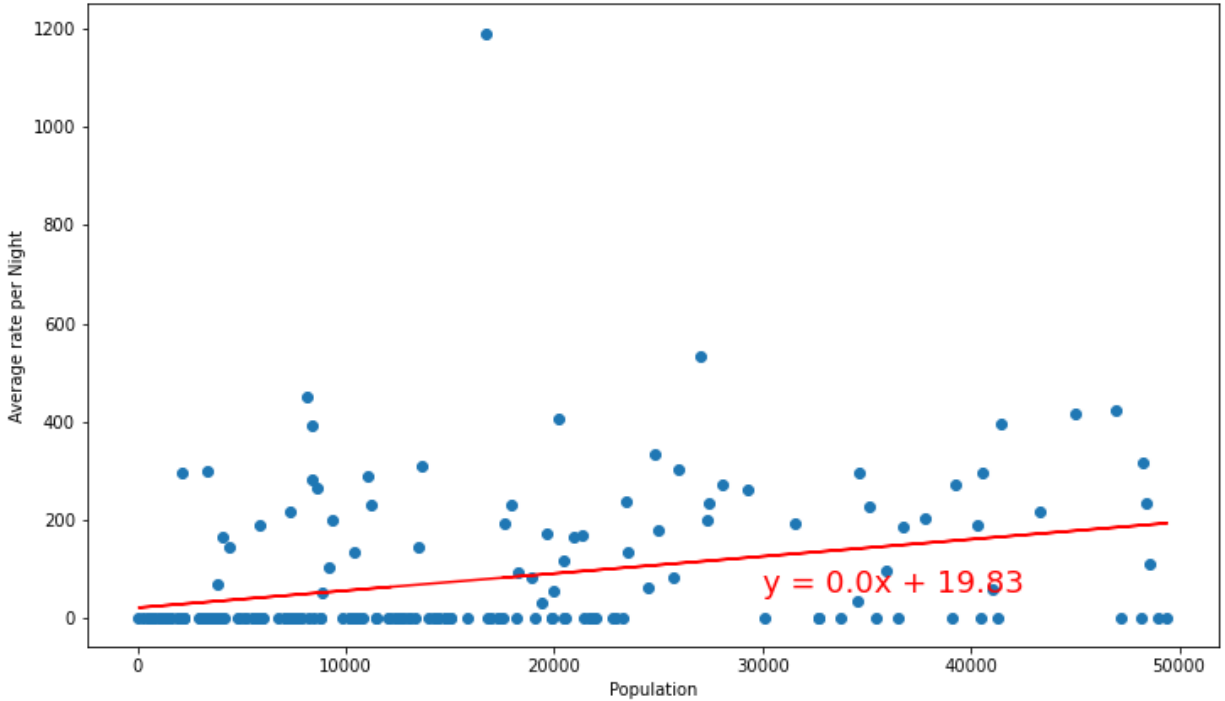
Dep. Variable:	airbnb_rate_mean	R-squared:	0.104
Model:	OLS	Adj. R-squared:	0.099
Method:	Least Squares	F-statistic:	21.85
Date:	Wed, 04 Aug 2021	Prob (F-statistic):	5.59e-06
Time:	16:38:15	Log-Likelihood:	-1208.6
No. Observations:	191	AIC:	2421.
Df Residuals:	189	BIC:	2428.
Df Model:	1		
Covariance Type:	nonrobust		

	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

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

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.

Regression Chart-Rural



Conclusion

We recommend operating an airbnb rental in a rural area outside a major metropolitan center that has outdoor destinations to maximize profit.

- **There will be higher rates in rural areas due to less competition.**
  - Rates are slightly higher in rural areas. We also found that the rental size is larger which could contribute to the higher rates.

There are more rentals in areas with tourist attractions.

- (i.e., Houston Theme Parks, Museums and business travel)

Population and income affect Airbnb rates.

- 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 & Future Work

- Data: Not current
  - Airbnb data is from 2013-2017 (only 6 months of 2017 data available)
  - Census data is from 2017
  - Airbnb data when compiled only shows data from 106 of 254 counties in Texas.

- Future work:
  - Look at the effects of rentals during COVID
  - How long will it take for the market to recover?
  - Seasonality of Airbnb 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=CjwKCAjw0qOIBhBhEiwAyvVcfzmOWPSjpt_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

Q & A Time!