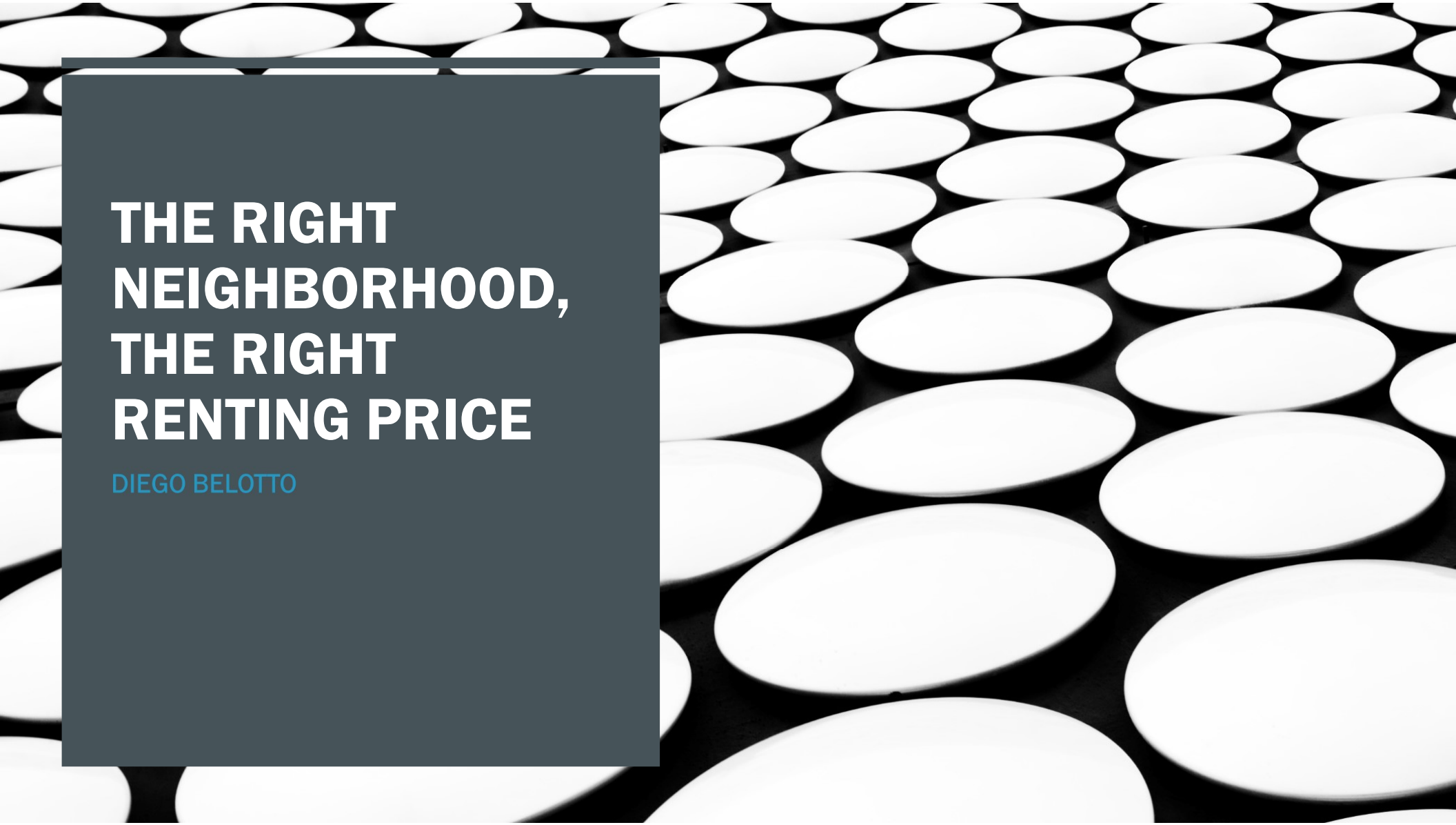


THE RIGHT NEIGHBORHOOD, THE RIGHT RENTING PRICE

DIEGO BELOTTO

THE VALUE OF PICKING THE STARTING NEIGHBORHOOD AND KNOWING THE FAIR RENTING PRICE

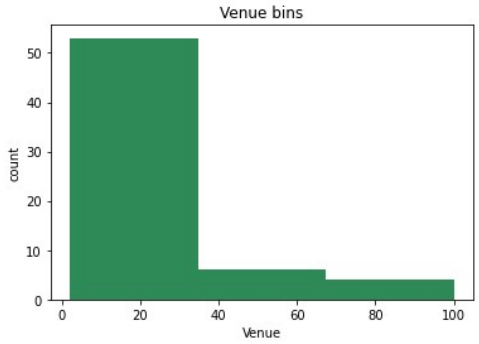
- Asunción is the capital of Paraguay and at the same time is the city with the largest quantity of venues and diversity of activities of the country
- For a young Paraguayan who wants to be independent or an adventurous tourist it seems as the ideal place to rent a house or an apartment.
- There is a vast quantity of factors that need to be weighted to make a renting decision
- We are going to try to help our audience to get a better understanding of the city using basic data analytics techniques and give tools to them to determine what neighborhoods would be suitable to start hunting.

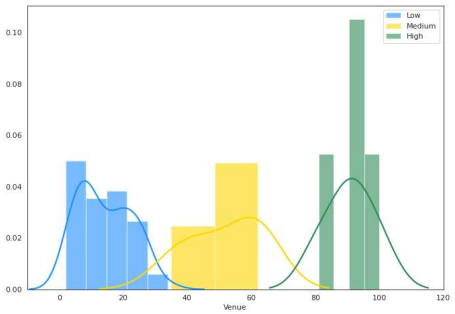
DATA ACQUISITION AND CLEANING

- Asuncion neighborhood data scrapped from http://geohidroinformatica.itaipu.gov.py/geoserver/wfs?typename=cih%3Aparaguay_2012_barrrios_localidades_wgs84&outputFormat=csv&version=1.0.0&request=GetFeature&service=WFS
- Venues data for each neighborhood using de Foursquare API https://api.foursquare.com/v2/venues/
- Renting data scrapped from https://www.infocasas.com.py/alquiler/inmuebles/asuncion/
- 68 neighborhoods analyzed
- 900 Renting observations scrapped; model developed with 285 observation and 284 features

EXPLORATORY ANALYSIS

- Distribution of venues per neighborhood and most common venue

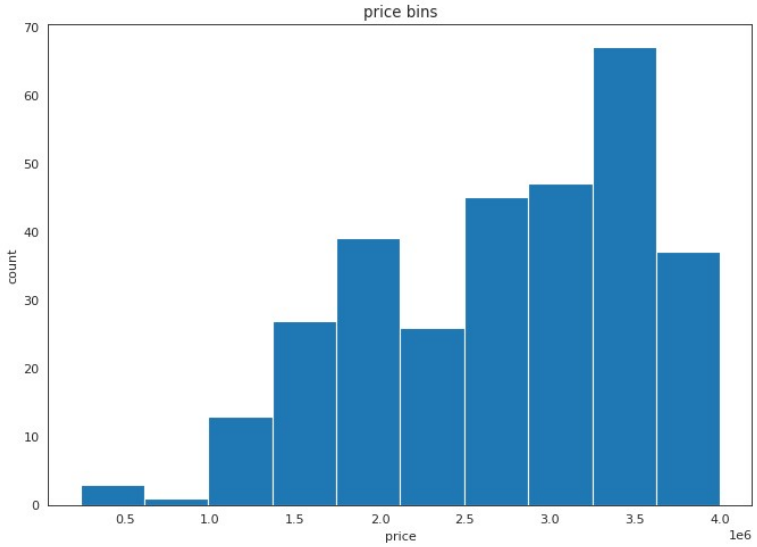




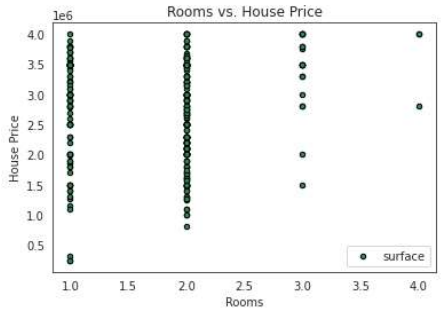
	1st Most Common Venue
Pizza Place	8
Fast Food Restaurant	6
Burger Joint	6
Bar	5
Ice Cream Shop	5
Restaurant	4
Athletics & Sports	2
Hotel	2
Plaza	2
Brewery	2
Health & Beauty Service	2
Park	2

RENTING PRICE ANALYSIS

- Renting prices of our data tends to be in the highest price
- Average price does not reflect very well the relationship between the price and the quantity of rooms

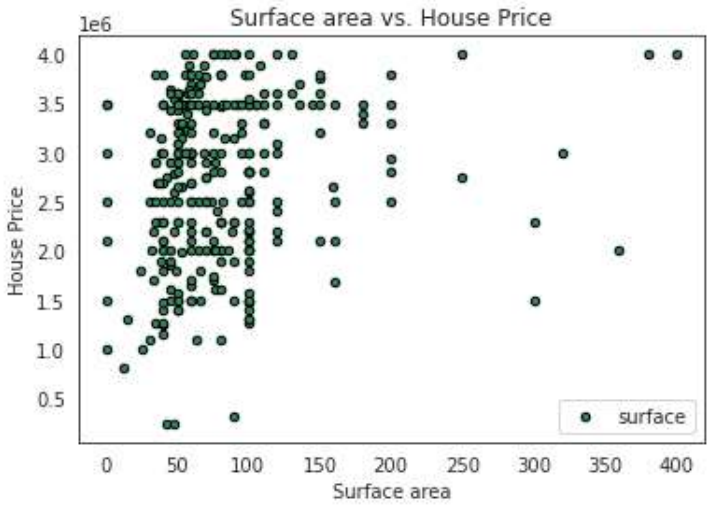


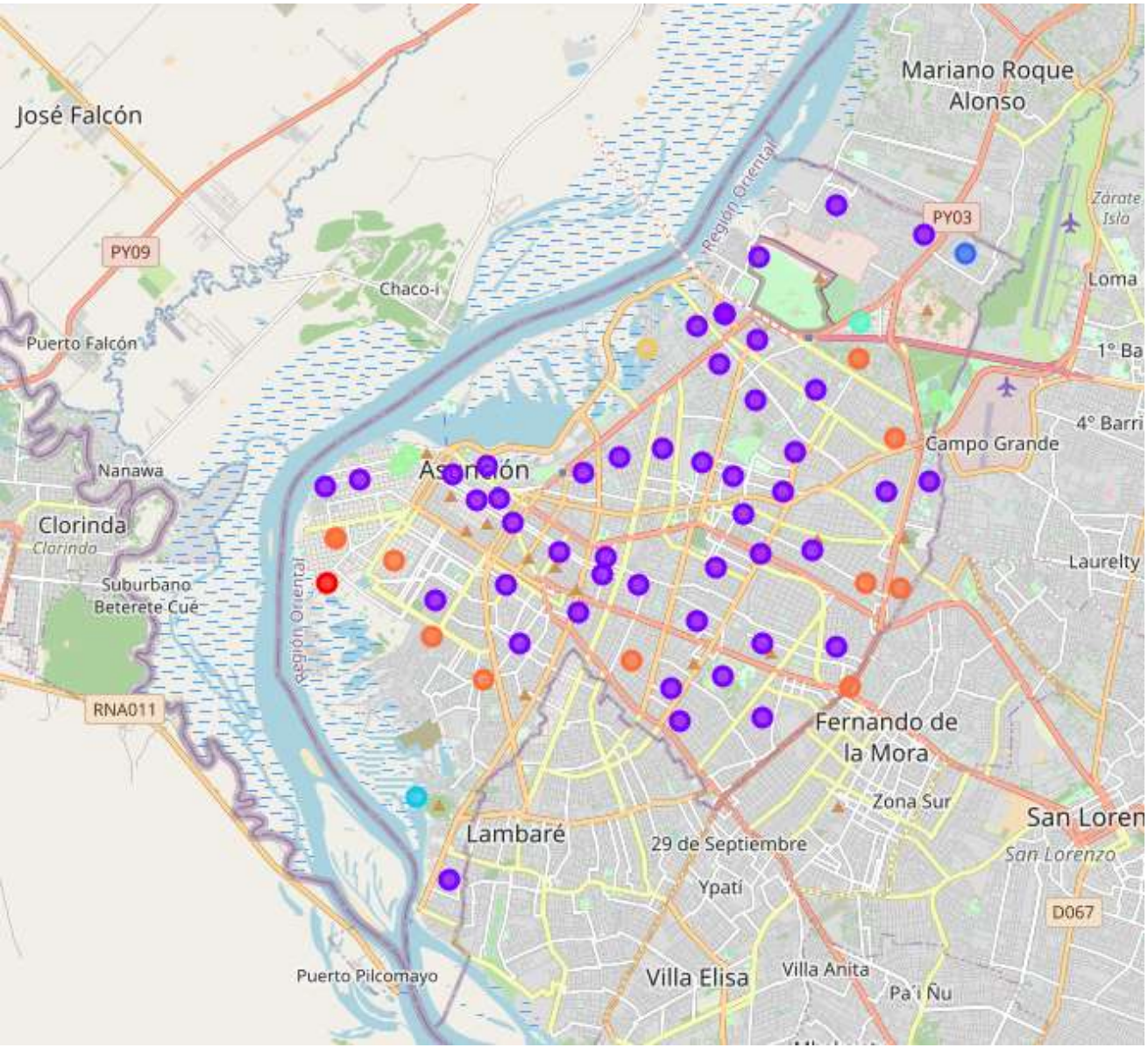
	tipoPropiedad	dormitorios	Precio_Gs
0	Casas	1.000000	2,671,428.571429
1	Casas	2.000000	2,540,000.000000
2	Casas	3.000000	3,360,000.000000
3	Casas	4.000000	3,600,000.000000
4	Departamentos	1.000000	2,771,423.837209
5	Departamentos	2.000000	2,631,527.777778
6	Departamentos	3.000000	3,383,333.333333



RENTING PRICE ANALYSIS

- Relationship between surface area and price is similar to the relationship of price with number of rooms
- There is no variable with a high correlation with price





MODELING

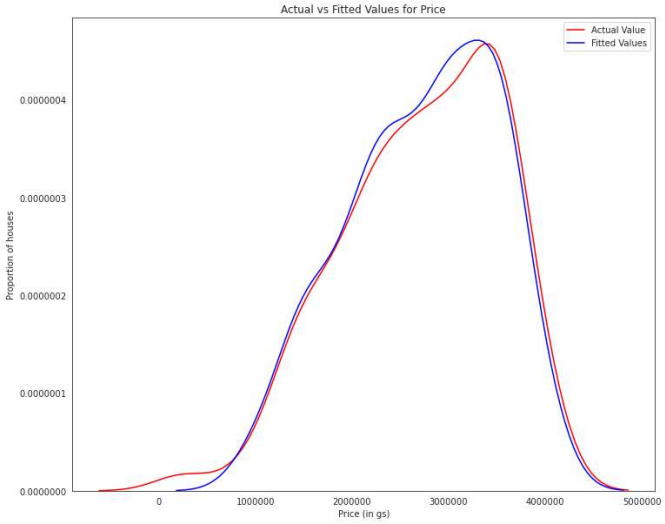
- Classification
- Kmeans

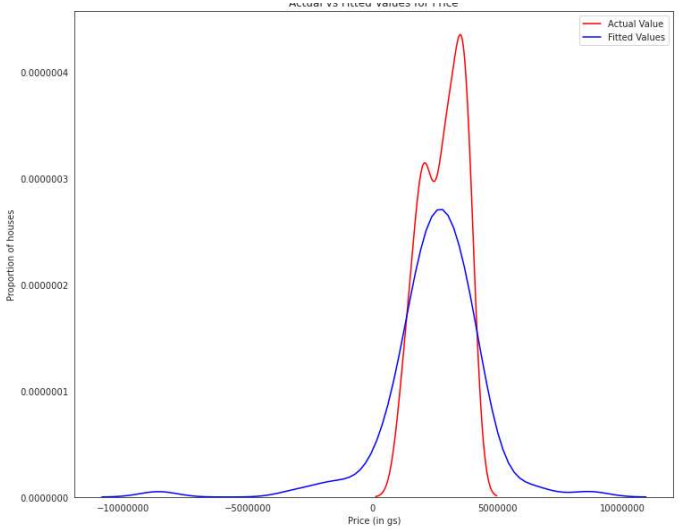
- The cluster with the most neighborhoods extends across the entire city

- Most neighborhood near the limit of the city are in other clusters, probably more affected by the near cities

MODELING

- Regression
- Linear
- Train Data R2 0.88
- Test Data R2 -5.60

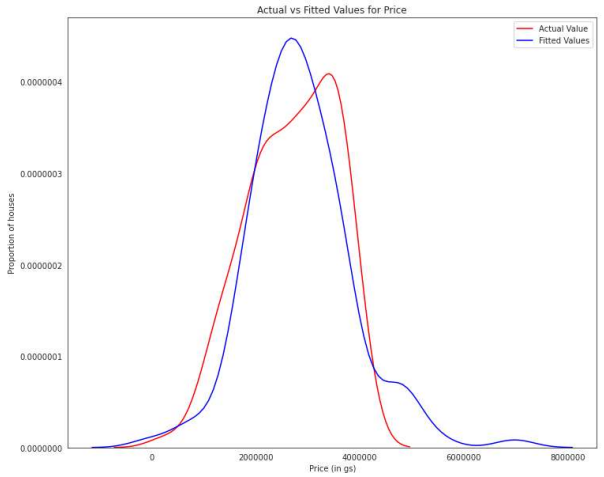


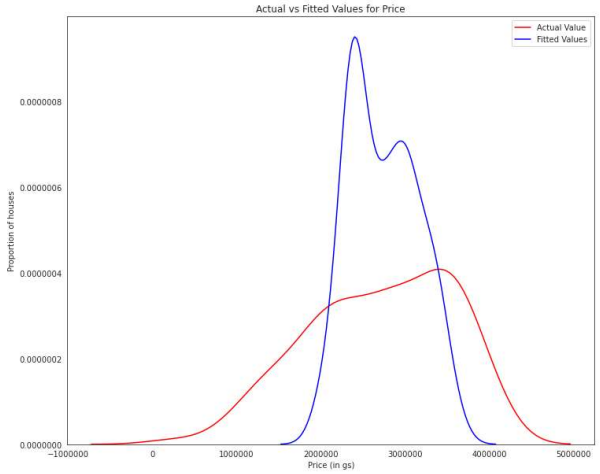


MODELING

- Ridge
- Train R2 Score 0.892
- Test R2 Score -0.455

- KNN
- Train R2 Score 0.40
- Test R2 Score 0.03





CONCLUSION AND FUTURE DIRECTION

KMeans algorithm classified each neighborhood in 8 clusters, getting that most of these neighborhoods were grouped in one cluster, giving the impression that the city is homogenous

The performance of the three were very poor at predicting a price using the test data, but with high scoring with the test data. Meaning a high bias.

There are other variables with more relevance than the venues near each neighborhood or even the size or the quantity of rooms of a property.

Future research could be improved adding data such as:

the transportation access, security level, ratings of the venues (having quantity doesn't mean quality), point of interest

As well to extend the universe of observation data to prevent bias when training the model.