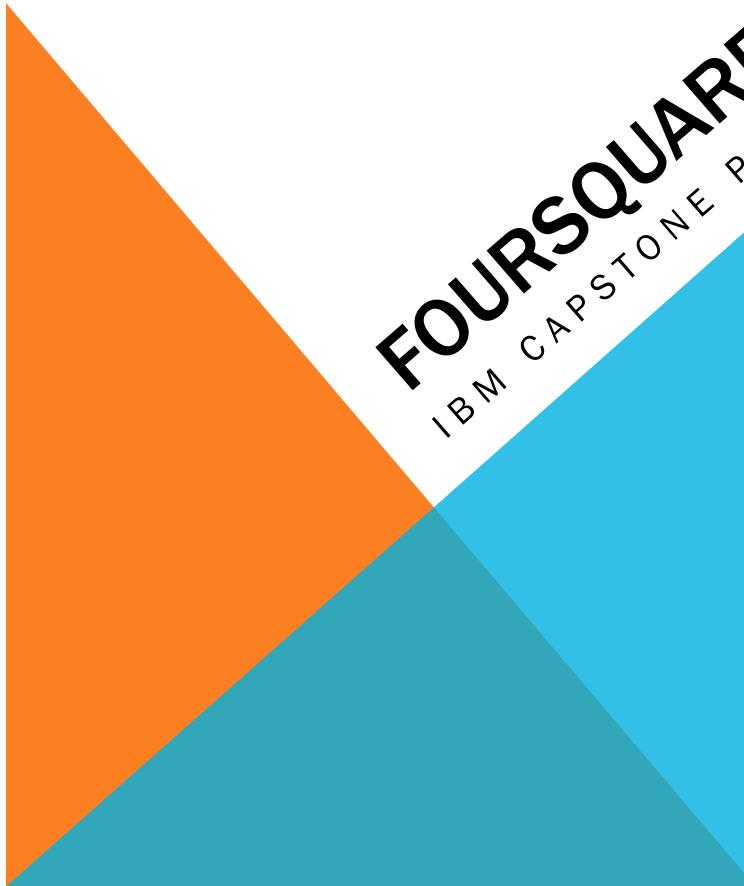


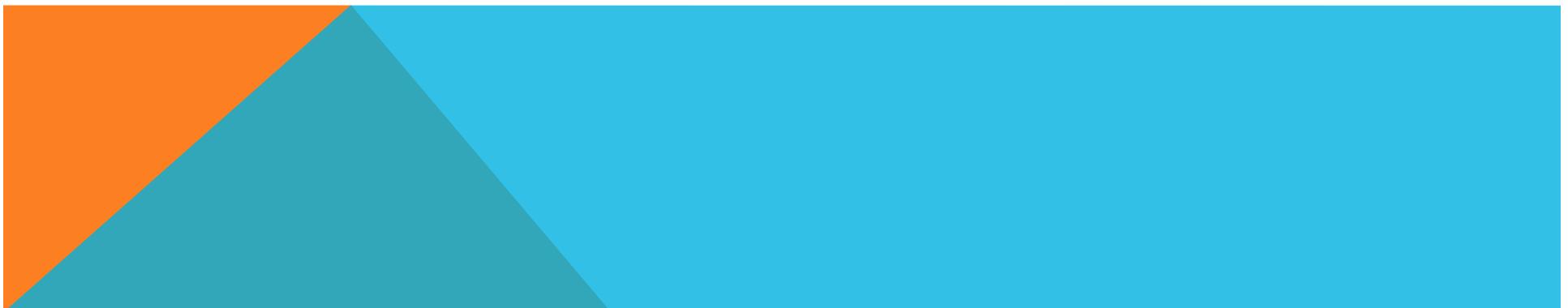
FOURSQUARE MEETS ZILLOW

IBM CAPSTONE PROJECT BY ADAM LAW



IBM CAPSTONE PROJECT

- Home appraisals are considered an art form.
- As there are many variables that can impact a home's market value, the appraisal (or estimation of a home's market value) is simply a 'feeling' at how much a home is worth, given by the appraiser.
- Because they are highly regulated, home appraisal professionals go through a rigorous training, continuing education and lengthy on-the-job experience before they become licensed appraiser. This is a must, as home appraisals are routinely relied on for mortgage financing by financial institutions.
- However, Zillow's introduction of the Zestimator has now offered reliable real estate pricing; challenging appraisals by moving real estate estimation from a 'feeling' to a science.
- Using Machine-Learning Regression and its ability to weight multiple factors, Zillow is able to accurately give the normal homeowner access to a free estimate of their home based on factors private to Zillow's proprietary regression.



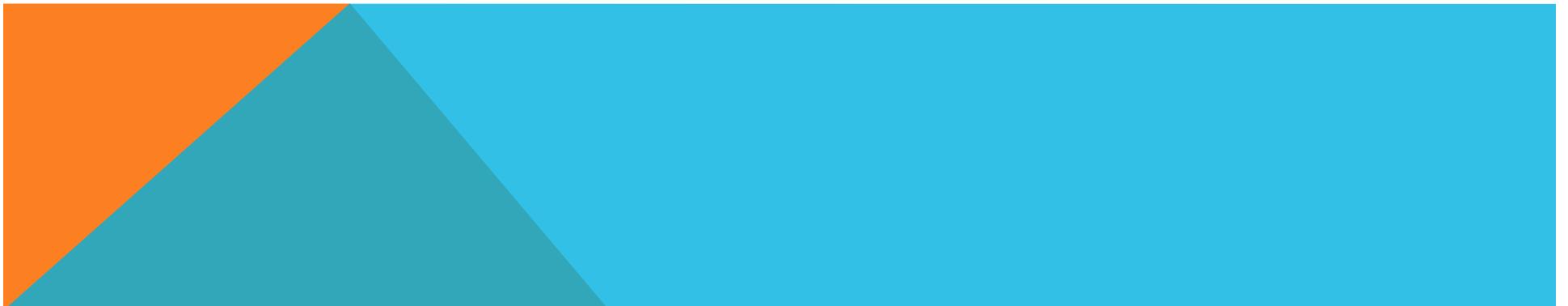
IBM CAPSTONE PROJECT

The impact of a home's value is based on many factors; square footage, number of bedrooms, number of bathrooms and utility features like garages or a washroom.

But most famously, the three most important factors for a home's value are 'Location! Location! Location!' Considered a Real Estate agent's mantra, this phrase is the belief that a home's location is the greatest variable of a home's value.

This phrase is the basis of our hypothesis:

Can the FourSquare api can be used to enhance Zestimator home prices by Location data analysis?



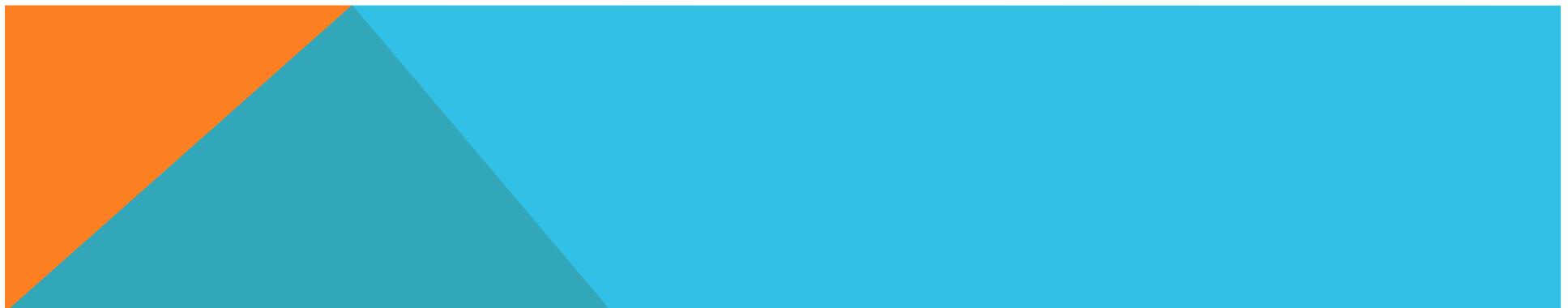
IBM CAPSTONE PROJECT

The Data

In order to test our Hypothesis, a baseline home regression estimator will need to be built using parsed real estate data (sqft, bedrooms, bathrooms, longitude, latitude, Zip) then an additional regression estimator will be created using the real estate data coupled with the location venue data from our FourSquare api. This data will include all available FourSquare data in a suitable radius from each home.

For our housing market; we will choose the Orlando, Florida area.

The two estimators will be compared and if the estimator accuracy is increased by our FourSquare data, then we have proved our hypothesis.



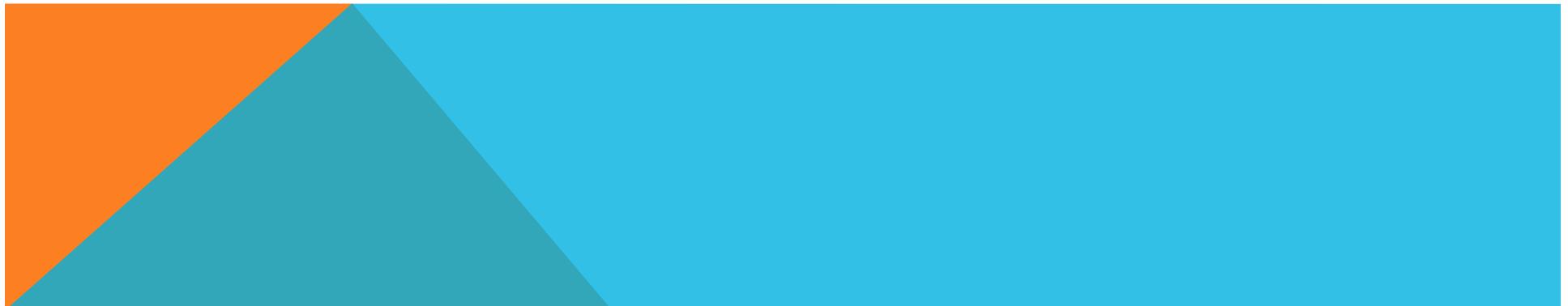
IBM CAPSTONE PROJECT

Baseline Estimator

The Base Random Forest Regressor will be trained on the basic parsed Home Features;
Bedrooms, Bathrooms, Square Feet, Latitude, Longitude and Zip

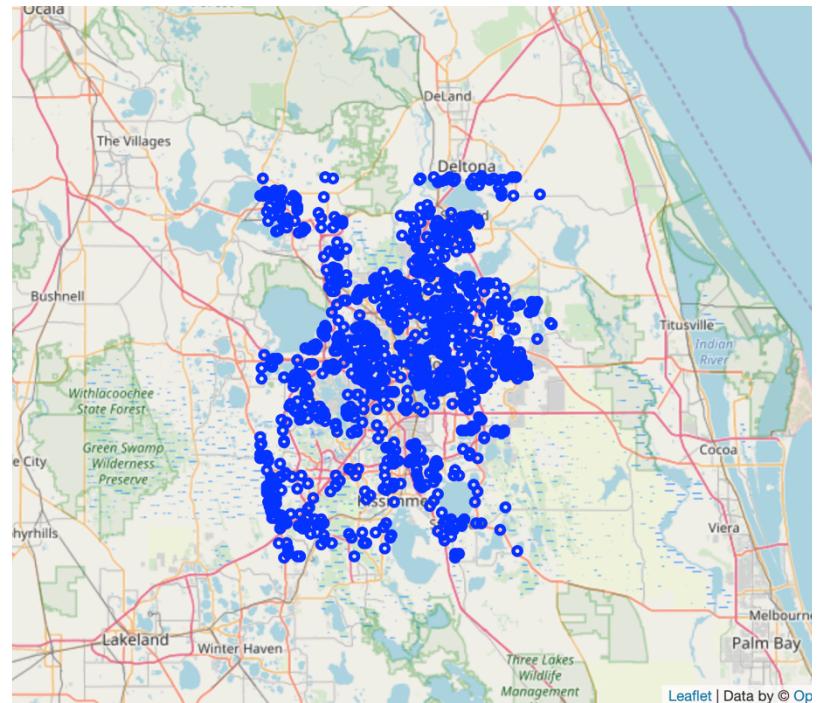
```
df.head()
```

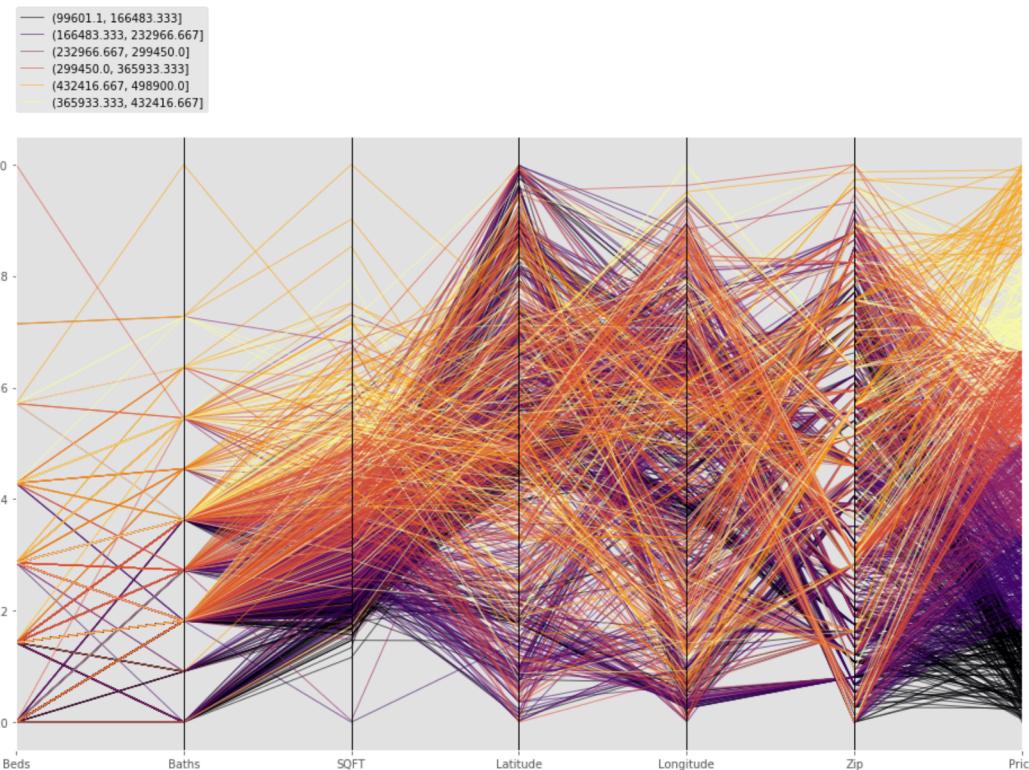
	Address	Beds	Baths	SQFT	Price	Latitude	Longitude	Zip
1	10080 Regent Park Dr # 2304, Orlando, FL 32825	2	2.5	1236	153000	28.541829	-81.243726	0
2	2550 N Alafaya Trl Apt 12107	2	2.0	1007	133500	28.575794	-81.211781	1
3	1802 Winter Green Blvd, Winter Park, FL 32792	3	2.0	1415	190000	28.615052	-81.314876	2
4	886 Commonwealth Ct, Casselberry, FL 32707	2	3.0	1295	167000	28.648725	-81.307811	3
5	510 Cranes Way Apt 201, Altamonte Springs, FL ...	2	2.0	1100	158000	28.670347	-81.378330	4



FOURSQUARE MEETS ZILLOW

Orlando Florida Area homes parsed from Zillow

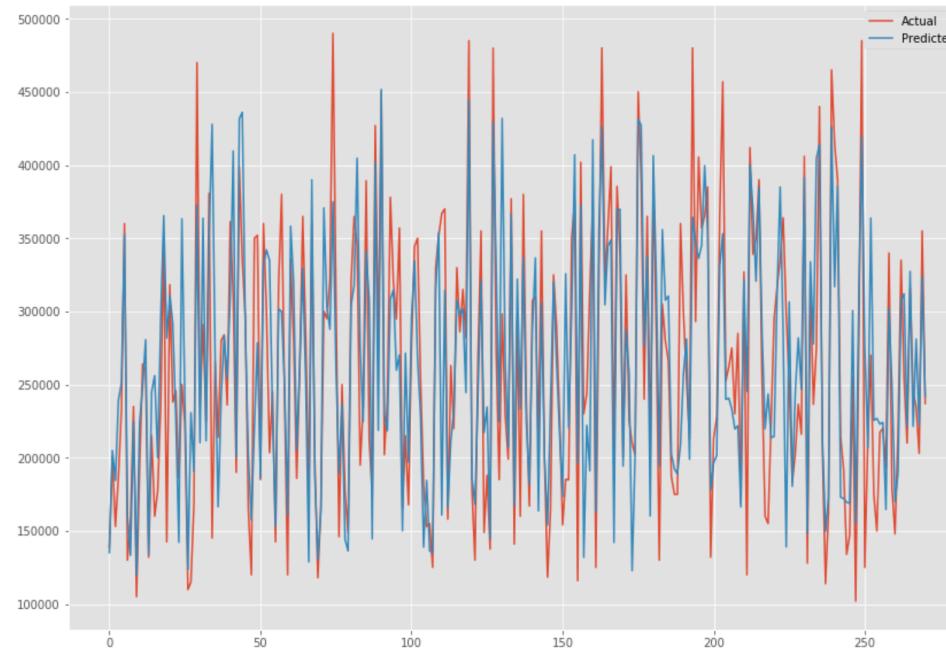




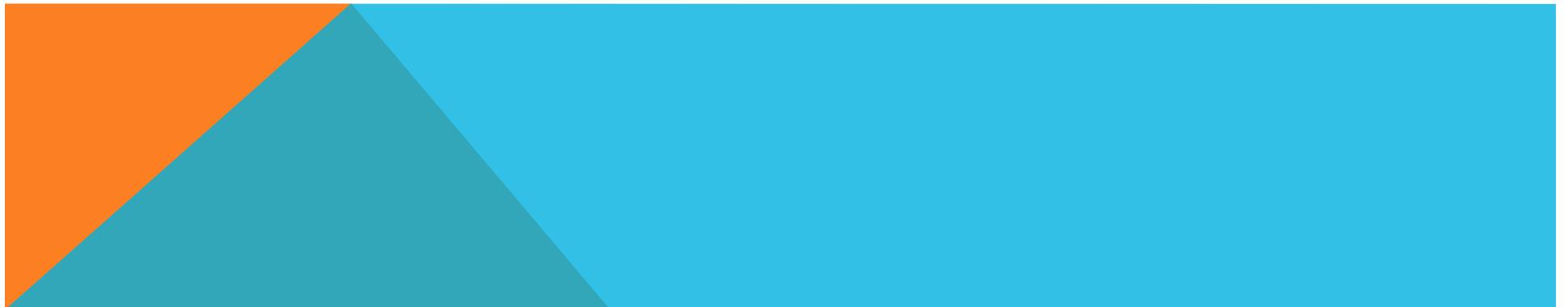
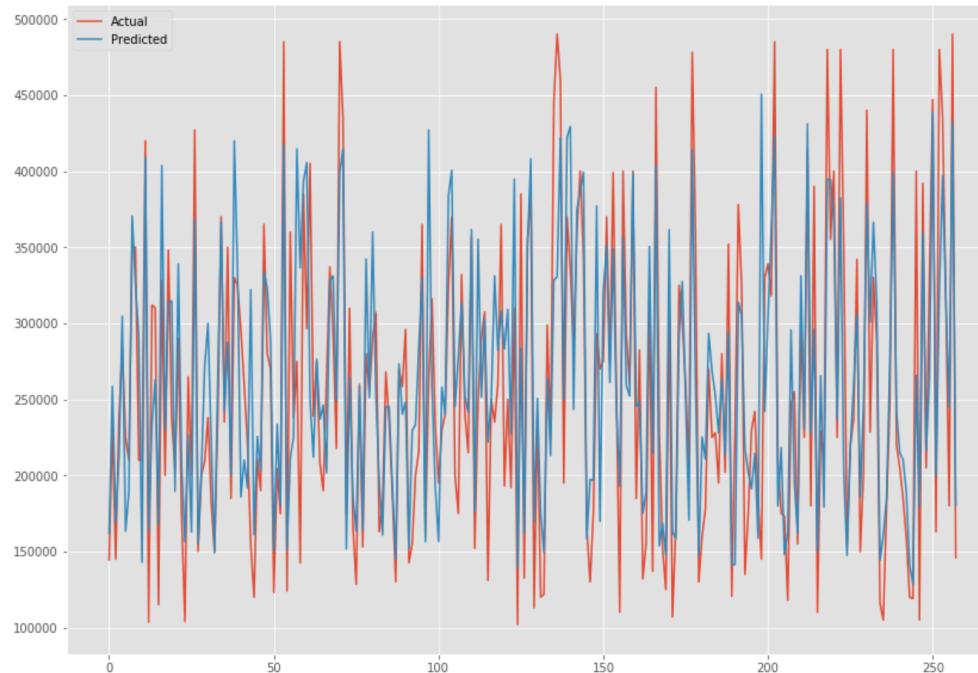
The Parallel Coordinates plot is helpful to understand associations in between each variable and the target price.



Our baseline estimator's performance ~ \$47,000 MAE



Now that we have the baseline estimator out of the way, we need to add our foursquare data to DataFrame. Our new accuracy is ~\$42,000 MAE which is an Improvement!

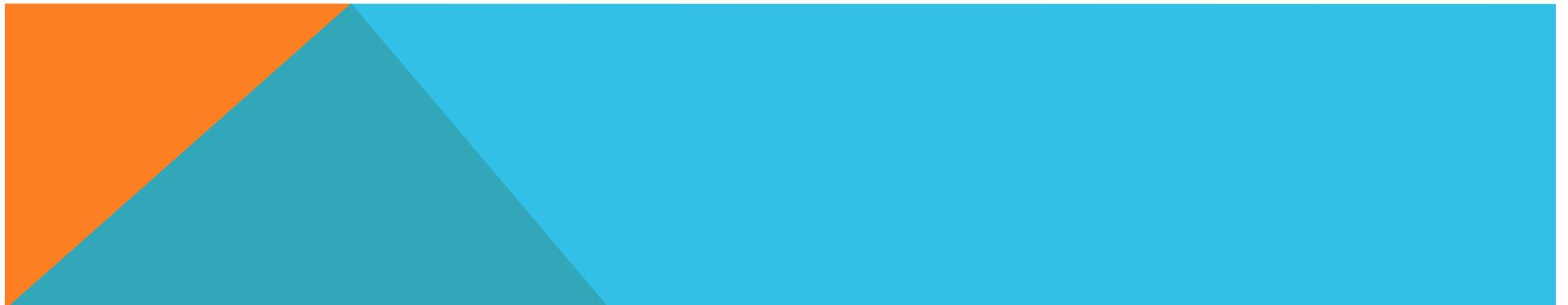


IBM CAPSTONE PROJECT

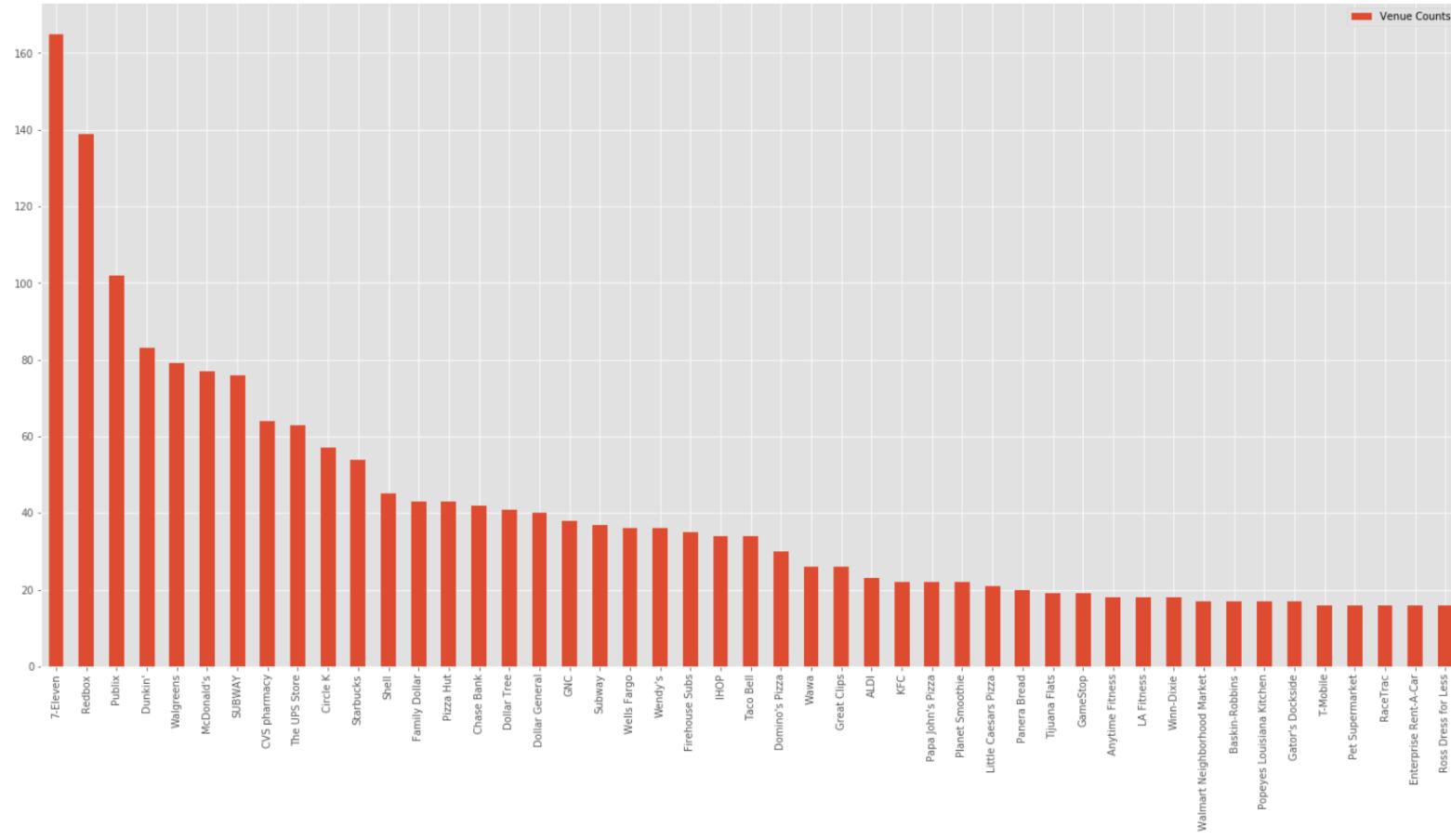
Results

```
In [50]: #This is where we compare our estimator accuracies
print('Final Score for Base Estimator:',initial_score , '// Final Score with Foursquare Data', Final_scores)
```

```
Final Score for Base Estimator: 47538.604896770536 // Final Score with Foursquare Data 42458.85154
927293
```



Top Venue Responses



With an increased accuracy of our Regressor, we can perform ‘everything else the same testing’ to extract the individual behavior of value added (or extracted) from our homes.

