

Anthony Palumbo - Principal Data Scientist at Boston Realty

Introduction

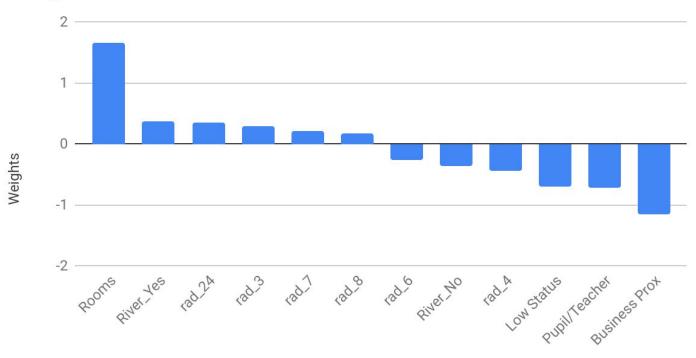
- Looking at median value of owner-occupied homes in \$1000s.
- 2 Main Categories-
 - Size
 - Location

-Area quality

How to use this model:

-Predict value of new houses

Weights of Predictors for House Price



Rooms: i.e: Size of the House

-For every extra room, the value of the house goes up \$1,668

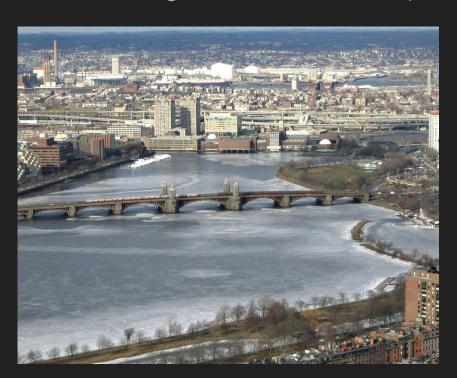


Location Location



The Importance of the River

-Added \$728 as opposed to not being on the Charles River(2*River_yes)



Distance to Businesses

- -Closer you are to business center, the more expensive the housing
- -Highest "negative" predictor in the data



Highway Proximity

- -Had a significant effect on price
- -Need to find codes



Quality of the Area



Student Teacher Ratio

- As the amount of students for every teacher went up, housing prices went down
- Approx. looses \$707.90



Status

- -As lower status increases, price decreases
- -\$698/% of lower status



Unimportant Factors

Unimportant factors = Predictors had little effect on house price

- -Zone, **crime rate**, non-retail business acres, nitrogen oxide concentration, age of the home, tax
- -Black is engineered feature for which the

Conclusion

- -Biggest 2 categories/ Things to Look for
 - -House size
 - -House location
- -Accuracy of Data: Able to explain 57.5% of the variation

QUESTIONS?

Technical Notes

Rooms	1.668
River_Yes	0.364
rad_24	0.3527
rad_3	0.2845
rad_7	0.2145
rad_8	0.179157
rad_6	-0.271
River_No	-0.364
rad_4	-0.4303
Low Status	-0.6982
Pupil/Teacher	-0.7079
Business Prox	-1.1579

The Model

```
Target_value = 3.754e+01 + -5.53e-02*w1 + 5.269e-02*w2 + -3.996e-03*w3 + -3.907e-02*w4 + 1.668e+00*w5 + -2.016e-02*w6 + -1.157e+00*w7 + -4.460e-03*w8 -7.079e-01*w9 + 1.004e-02*w10 + -6.98230657e-01*w11 + -3.6401e-01*w12 + 3.640e-01*13 + -1.798e-01w14 + -1.181e-01*w15 + 2.845e-01*w16 + -4.303e-01*w17 + -3.068e-02*w18 -2.719e-01w19 + 2.145e-01*w20 + 1.791e-01*w21 + 3.52777177e-01*w22
```