

Highlights

Decided to split into 10 models - one for each city

Linear model seemed appropriate and coefficients matched expected relationships with predictions

Found many variables work best as binary variables. For example, square footage of a basement does not matter as much as if there is a basement or not.

Review Progress

During this spring the following were completed:

- Theme 1:
 - Epic 1:
 - Data cleaning
 - Model building
 - Model testing
- These 2:
 - Epic 1:
 - Change attributes
 - Run new prediction for attribute

Demo/Analysis

Progress: Jupyter Notebook user demo.

- User inputs city and housing information
- Returns predicted house price
- Returns attribute to change price the most

```
Select city from list: Seattle(1), Renton(2), Bellevue(3), Kent(4), Redmond(5), Kirkland(6), Auburn(7), Sammamish(8),  
Federal Way(9), Issaquah(10)  
Please input the option number: 3
```

```
-----  
You chose Bellevue!
```

```
Please input the number of bedrooms: 4
```

```
Please input the number of bathrooms: 3
```

```
Please input the lot square footage: 2500
```

```
Please input the number of floors (1+): 2
```

```
Would you like a waterfront view? (y/n)y
```

```
Please input what condition the home should be in (1-5):4
```

```
Would you like a basement? (y/n)n
```

```
What year should the house be built in? (1900-2015)1960
```

```
Would you like a renovated house? (y/n)n
```

```
Your expected home price is: 363112.657640817
```

```
To decrease price most, if you changed bedrooms to 3 then your home price would be: 283427.08486101904
```

Lessons Learned

Consistent naming across the 10 different models and elements they use is important

RDS can store my coefficients as a table to make live predictions for all the models.

Recommendations

During the next spring, the following will be completed:

- Model scripts created
- Documentation
- Model unit testing