

# King County Pricing Analysis

Blake Elieff

October 23 2022



# Summary

Analysis of the king county dataset and building of a model that predicts house pricing given different features



# Outline

- **Business Problem**
- **Data & Methods**
- **Results**
- **Conclusions**

# Business Problem

- A model is required that is able to predict the sale price of housing in the King County area and answer the following questions:
  - How much does an extra bathroom add to price?
  - How much does an extra sqft add to the price?
  - How much does grade effect the price?

# Data & Methods

- 1 dataset (kc\_house\_data.csv)
- Used logarithmic scaling, Train split, one hot encoding

# Model

- In the end, this is the function that I was able to produce to work with my model
- For testing, I will begin with the median of each feature

```
def predict(sqft_living,bathrooms,bedrooms,grade, condition):  
    ...  
    sqft_living = The square foot of living area of the property  
    bathrooms = number of bathrooms  
    bedrooms = number of bedrooms  
    grade = numbered grade of the property  
    condition = numbered condition of the property'''  
  
    sqft_living_log = np.log(sqft_living)  
  
    bathrooms = bathrooms  
    bedrooms = bedrooms  
  
    if grade > 10:  
        high_grade = 1  
    else: high_grade = 0  
  
    if condition > 3:  
        good_condition = 1  
    else: good_condition = 0  
  
    X = [[sqft_living_log,bathrooms,bedrooms,high_grade, good_condition]]  
  
    price = np.exp(model.predict(X))  
    print('$',round(float(price),2))
```

# Results

- For a house that keeps the rest of the features the same, if we increase the bathroom number by 1 we could potentially have a price increase of \$35112

```
predict(sqft_living=1920,  
        bathrooms=2.25,  
        bedrooms=3,  
        grade=6,  
        condition=2)
```

\$ 465770.79

```
predict(sqft_living=1920,  
        bathrooms=3.25,  
        bedrooms=3,  
        grade=6,  
        condition=2)
```

\$ 500883.46

# Results

- Per square foot of living area added we add approximately \$194.7 to the sale price

```
predict(sqft_living=1920,  
        bathrooms=2.25,  
        bedrooms=3,  
        grade=6,  
        condition=2)
```

\$ 465770.79

```
predict(sqft_living=1921,  
        bathrooms=2.25,  
        bedrooms=3,  
        grade=6,  
        condition=2)
```

\$ 465965.49



# Results

- The difference between the house being high and low grade is approximately \$276017

```
predict(sqft_living=1920,  
        bathrooms=2.25,  
        bedrooms=3,  
        grade=6,  
        condition=2)
```

\$ 465770.79

```
predict(sqft_living=1920,  
        bathrooms=2.25,  
        bedrooms=3,  
        grade=11,  
        condition=2)
```

\$ 741788.67

# Conclusions

- This model is not able to predict the house prices for the King County however is useful to calculate change in price based off changes in feature values
- Adding a bathroom will add approx. \$35112 to the sale price
- Adding an extra sqft adds approx. \$195 to the sale price
- Changing from low to high grade adds approx. \$276,000 to the sale price



# Thank You!

Email: [blakeelieff@hotmail.com](mailto:blakeelieff@hotmail.com)

GitHub: @elieff