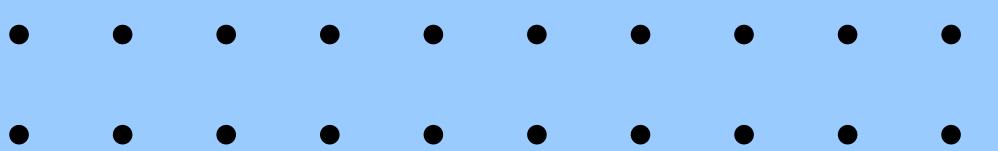




REAL ESTATE ANALYSIS

PRICE PREDICTIONS



REAL ESTATE MARKET 2022

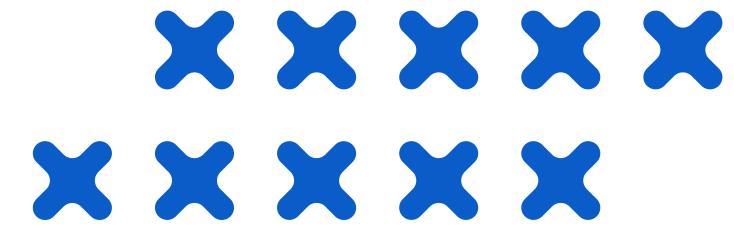
Median home

value is \$357,810

Median Household
Income: \$64,994

- Median Home Value (1-Year Forecast): +1.4%

- Homes Sold:
314,221
(-29.5% year
over year)



**BUYING OR SELLING A
HOME IS TOUGH!
SO MANY
CONSTANTLY
CHANGING FACTORS!**

**HOW TO MAKE THE
RIGHT DECISION?**

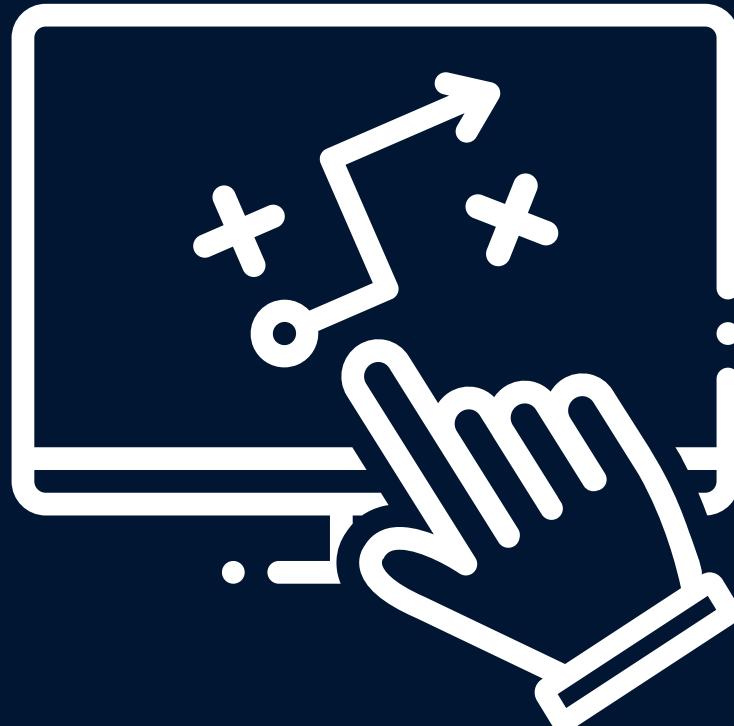


DATA

For the research, we used the following datasets:

- The King County, WA, real estate DataBase

We used the linear regression OLS method for modeling and finding parameters that can influence home prices the most.



HOW DID WE DO IT

1

Data cleaning

2

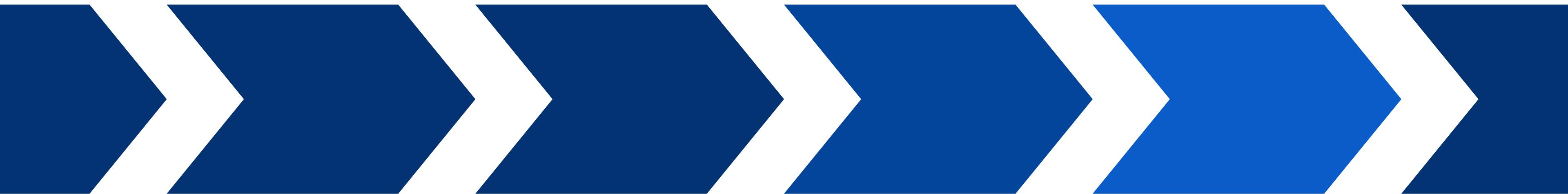
Data preparation

3

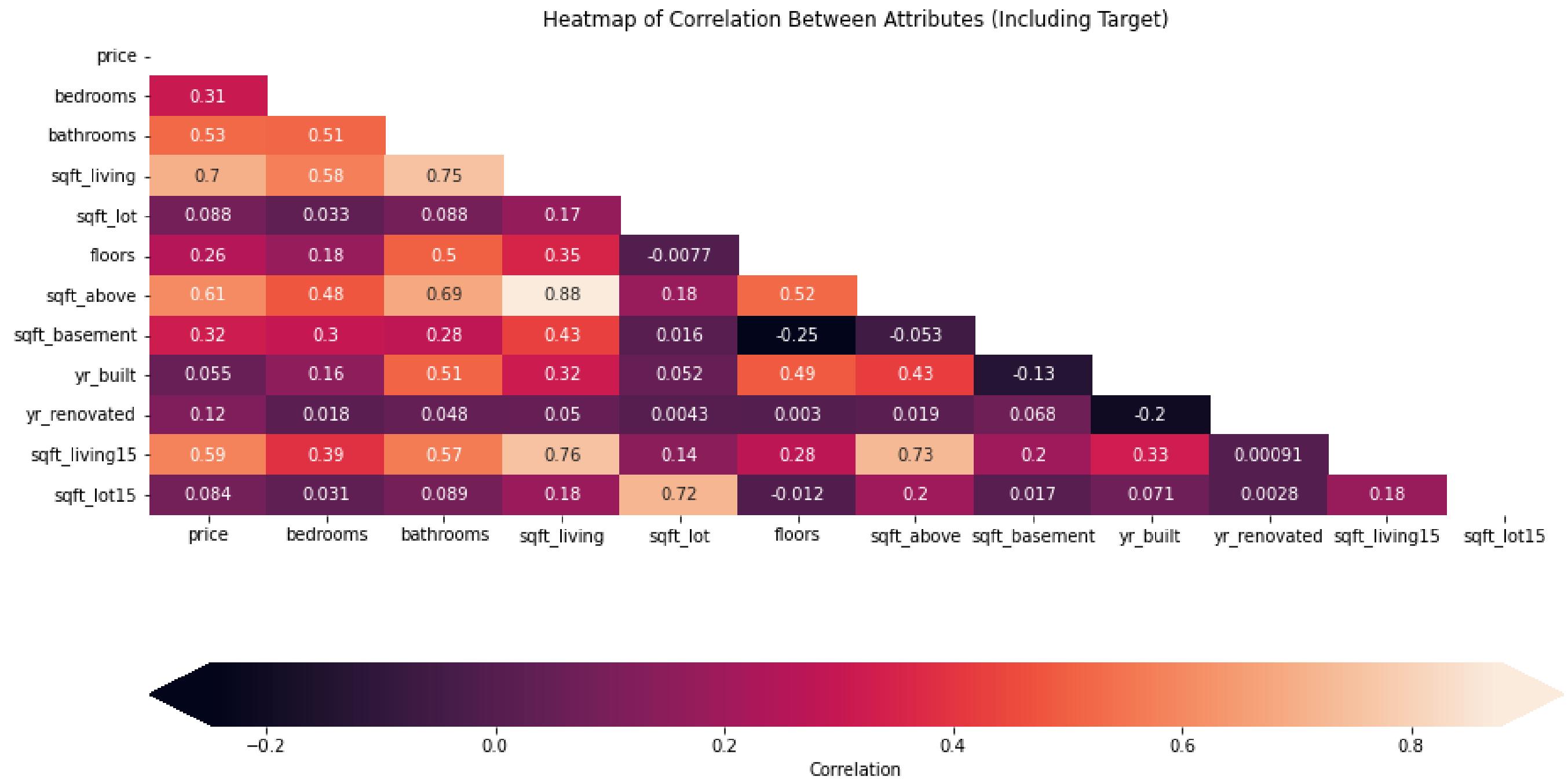
Model building
and verification

4

Further
recommendations



CORRELATIONS HEATMAP



CORRELATIONS

price

| | price |
|-------------------------------|----------|
| sqft_living | 0.702004 |
| sqft_above | 0.605481 |
| sqft_living15 | 0.586495 |
| bathrooms | 0.525029 |
| grade_11 Excellent | 0.356823 |
| sqft_basement | 0.323018 |
| bedrooms | 0.308454 |
| view_EXCELLENT | 0.307035 |



REGRESSION MODEL

OLS Regression Results

| Dep. Variable: | price | R-squared: | 0.552 | | | |
|--------------------|------------------|---------------------|-------------|-------|-----------|-----------|
| Model: | OLS | Adj. R-squared: | 0.552 | | | |
| Method: | Least Squares | F-statistic: | 5187. | | | |
| Date: | Wed, 26 Oct 2022 | Prob (F-statistic): | 0.00 | | | |
| Time: | 21:24:45 | Log-Likelihood: | -2.9157e+05 | | | |
| No. Observations: | 21082 | AIC: | 5.831e+05 | | | |
| Df Residuals: | 21076 | BIC: | 5.832e+05 | | | |
| Df Model: | 5 | | | | | |
| Covariance Type: | nonrobust | | | | | |
| | coef | std err | t | P> t | [0.025 | 0.975] |
| const | 9.312e+04 | 6739.937 | 13.817 | 0.000 | 7.99e+04 | 1.06e+05 |
| sqft_living | 272.5097 | 3.132 | 87.007 | 0.000 | 266.371 | 278.649 |
| bathrooms | 1.069e+04 | 3393.441 | 3.149 | 0.002 | 4033.833 | 1.73e+04 |
| grade_11 Excellent | 3.119e+05 | 1.35e+04 | 23.154 | 0.000 | 2.86e+05 | 3.38e+05 |
| bedrooms | -4.638e+04 | 2282.776 | -20.316 | 0.000 | -5.09e+04 | -4.19e+04 |
| view_EXCELLENT | 5.586e+05 | 1.43e+04 | 39.111 | 0.000 | 5.31e+05 | 5.87e+05 |
| Omnibus: | 13217.471 | Durbin-Watson: | 1.991 | | | |
| Prob(Omnibus): | 0.000 | Jarque-Bera (JB): | 488165.457 | | | |
| Skew: | 2.459 | Prob(JB): | 0.00 | | | |
| Kurtosis: | 26.055 | Cond. No. | 1.92e+04 | | | |



CONCLUSIONS

The overall model was statistically significant ($R^2 = .552$, $p < .000$).

It was found that the Square Footage of the Living Area significantly predicted price ($\beta = 272.5097$, $p < .000$).

It was also found that:

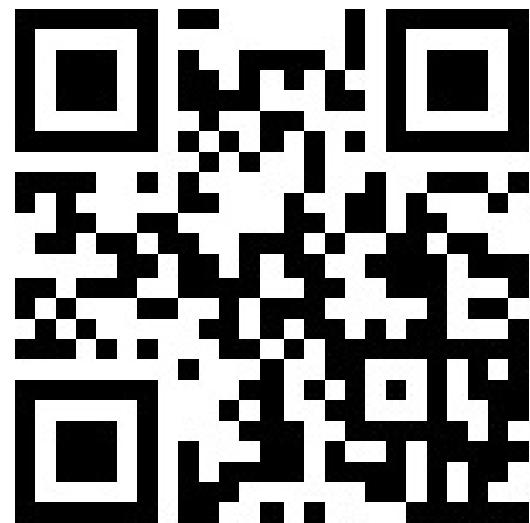
- Number of bathrooms significantly predicted price ($\beta = 1.069e+04$, $p < .000$)
- Number of bedrooms significantly predicted price ($\beta = -4.638e+04$, $p < .000$).



We suggest further investigation of how features such as 'Grade' and 'View' affect home prices.



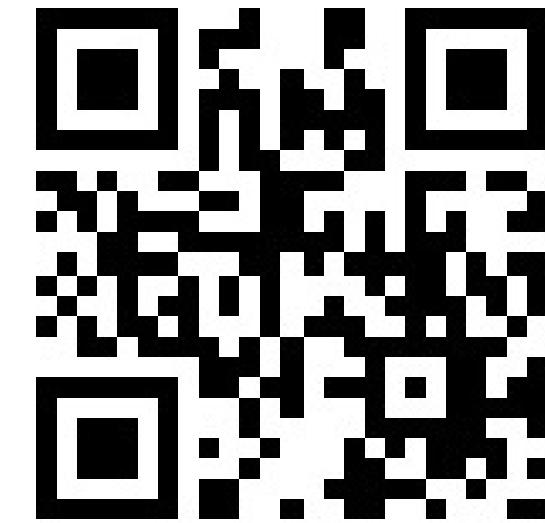
THE TEAM



LinkedIn



Andrei Mikhailov
DATA SCIENTIST



am@mikhailovandrei.com

Thank you.