Intergenerational Families in Kings Counties

Beau Morton

Overview

Client: National Association for Real Estate Brokers (NAREB) A black led organization that focuses on bringing together the minority professionals in the real estate industry to promote the meaningful exchange of ideas about our business and how best to serve their clientele.

Overview

How can home renovations increase the estimated value of homes for black families in King County, WA? (specifically focusing on intergenerational homes)

Data Preparation and Understanding

- 'Zipcode' was dropped as a variable
- Categorical variables were dropped, and dummy variables were created
- Regression models were created
- Homoscedasticity, normality, and multicollinearity tests were performed to confirm assumptions for linear regressions

Model 1

```
                           OLS Regression Results
==============================================================================
Dep. Variable:                  price   R-squared:                       0.493
Model:                            OLS   Adj. R-squared:                  0.493
Method:                 Least Squares   F-statistic:                     7003.
Date:                Sun, 17 Apr 2022   Prob (F-statistic):               0.00
Time:                        20:58:24   Log-Likelihood:             -3.0005e+05
No. Observations:               21597   AIC:                         6.001e+05
Df Residuals:                   21593   BIC:                         6.001e+05
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept    -3.815e+04   5252.893     -7.262      0.000   -4.84e+04   -2.79e+04
sqft_living    297.9354      4.483     66.458      0.000     289.148     306.723
sqft_above     -18.4232      4.479     -4.113      0.000     -27.202      -9.644
bathrooms    -3972.1769   3545.059     -1.120      0.263   -1.09e+04    2976.400
==============================================================================
Omnibus:                    14748.993   Durbin-Watson:                   1.982
Prob(Omnibus):                  0.000   Jarque-Bera (JB):           538245.600
Skew:                           2.806   Prob(JB):                         0.00
Kurtosis:                      26.804   Cond. No.                     9.59e+03
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 9.59e+03. This might indicate that there are
strong multicollinearity or other numerical problems.
```

Model 2

```
                           OLS Regression Results
==============================================================================
Dep. Variable:                  price   R-squared:                       0.507
Model:                            OLS   Adj. R-squared:                  0.507
Method:                 Least Squares   F-statistic:                     7398.
Date:                Sun, 17 Apr 2022   Prob (F-statistic):               0.00
Time:                        21:06:31   Log-Likelihood:             -2.9976e+05
No. Observations:               21597   AIC:                         5.995e+05
Df Residuals:                   21593   BIC:                         5.996e+05
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     7.704e+04   7684.108     10.026      0.000     6.2e+04    9.21e+04
bedrooms     -5.775e+04   2324.412    -24.846      0.000    -6.23e+04   -5.32e+04
sqft_living    313.8304      2.467    127.187      0.000     308.994     318.667
floors        3476.4322   3480.340      0.999      0.318   -3345.290    1.03e+04
==============================================================================
Omnibus:                    14415.709   Durbin-Watson:                   1.985
Prob(Omnibus):                  0.000   Jarque-Bera (JB):           491526.502
Skew:                           2.733   Prob(JB):                         0.00
Kurtosis:                      25.723   Cond. No.                     1.04e+04
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 1.04e+04. This might indicate that there are
strong multicollinearity or other numerical problems.
```

Model 3

```
                        OLS Regression Results
==============================================================================
Dep. Variable:                  price   R-squared:                       0.509
Model:                            OLS   Adj. R-squared:                  0.509
Method:                 Least Squares   F-statistic:                     7451.
Date:                Sun, 17 Apr 2022   Prob (F-statistic):               0.00
Time:                        22:15:14   Log-Likelihood:            -2.9972e+05
No. Observations:               21597   AIC:                         5.994e+05
Df Residuals:                   21593   BIC:                         5.995e+05
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     8.424e+04   6637.291     12.692      0.000    7.12e+04    9.72e+04
bedrooms     -5.959e+04   2327.075    -25.605      0.000   -6.41e+04   -5.5e+04
sqft_living    318.6333      2.383    133.727      0.000     313.963     323.304
sqft_lot        -0.3847      0.043     -8.922      0.000      -0.469      -0.300
==============================================================================
Omnibus:                    14308.034   Durbin-Watson:                   1.985
Prob(Omnibus):                  0.000   Jarque-Bera (JB):           477721.465
Skew:                           2.710   Prob(JB):                         0.00
Kurtosis:                      25.394   Cond. No.                     1.73e+05
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 1.73e+05. This might indicate that there are
strong multicollinearity or other numerical problems.
```
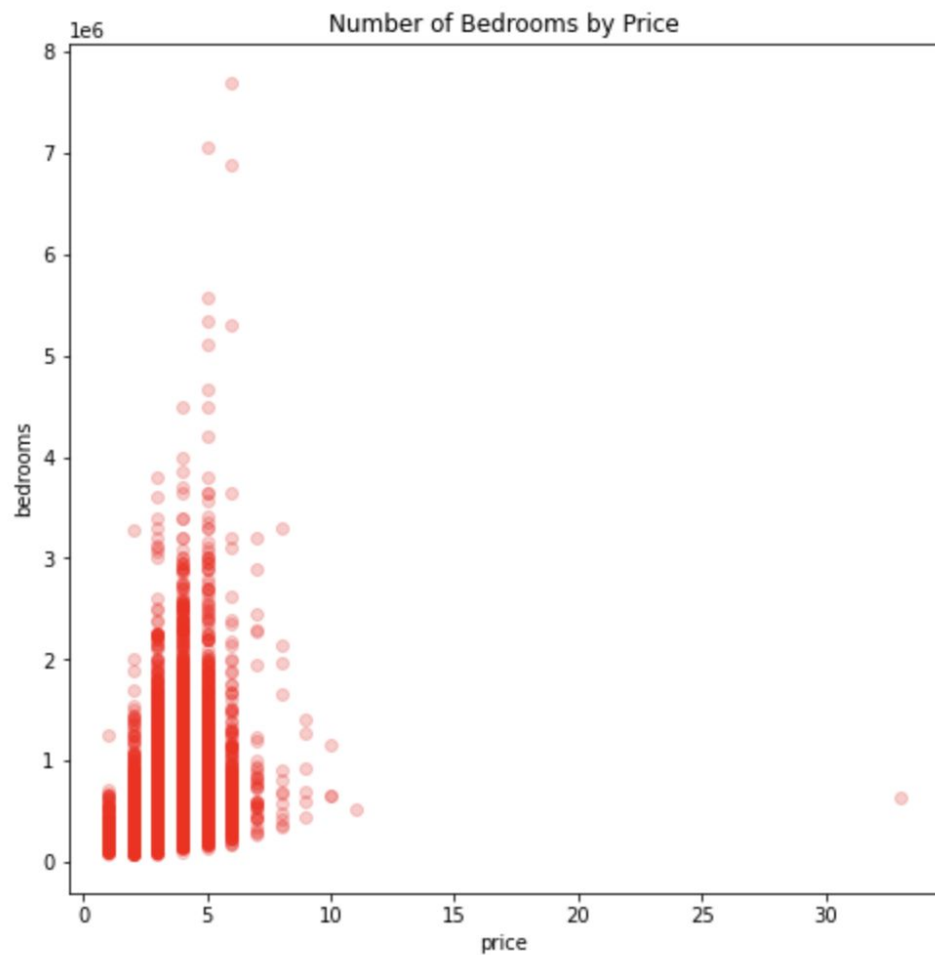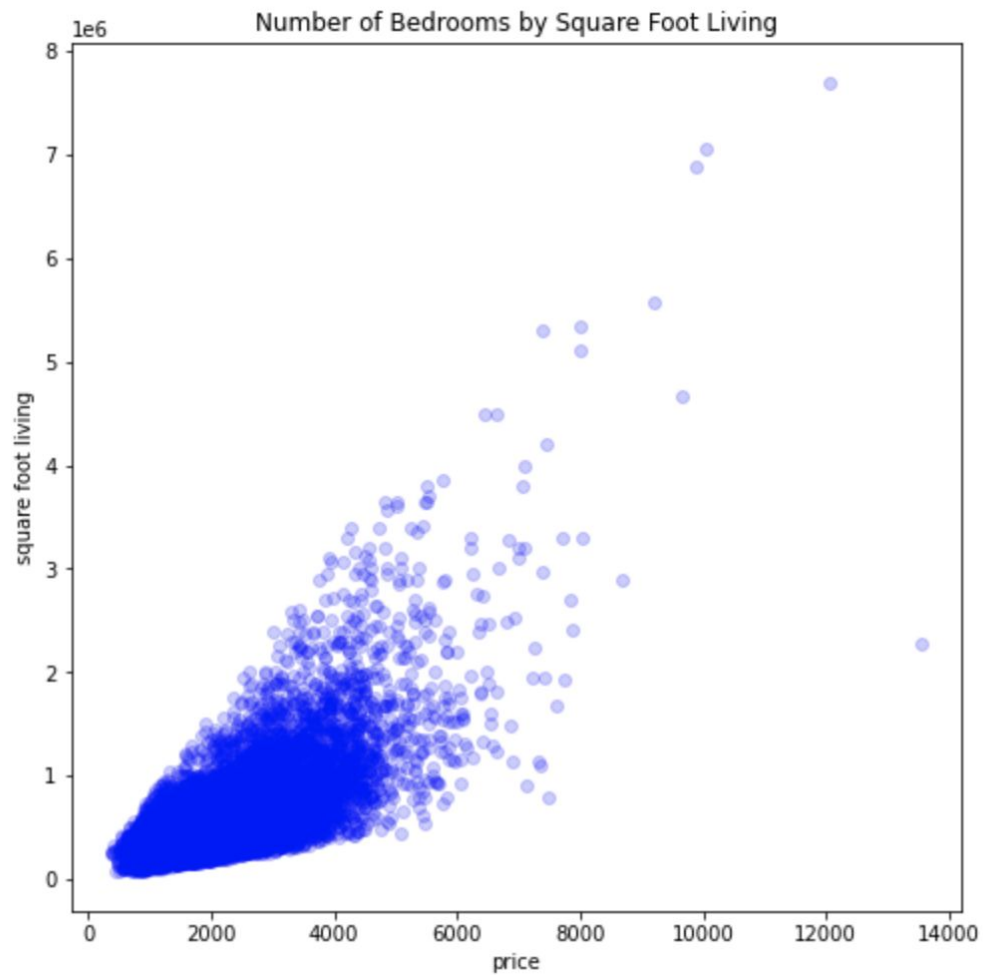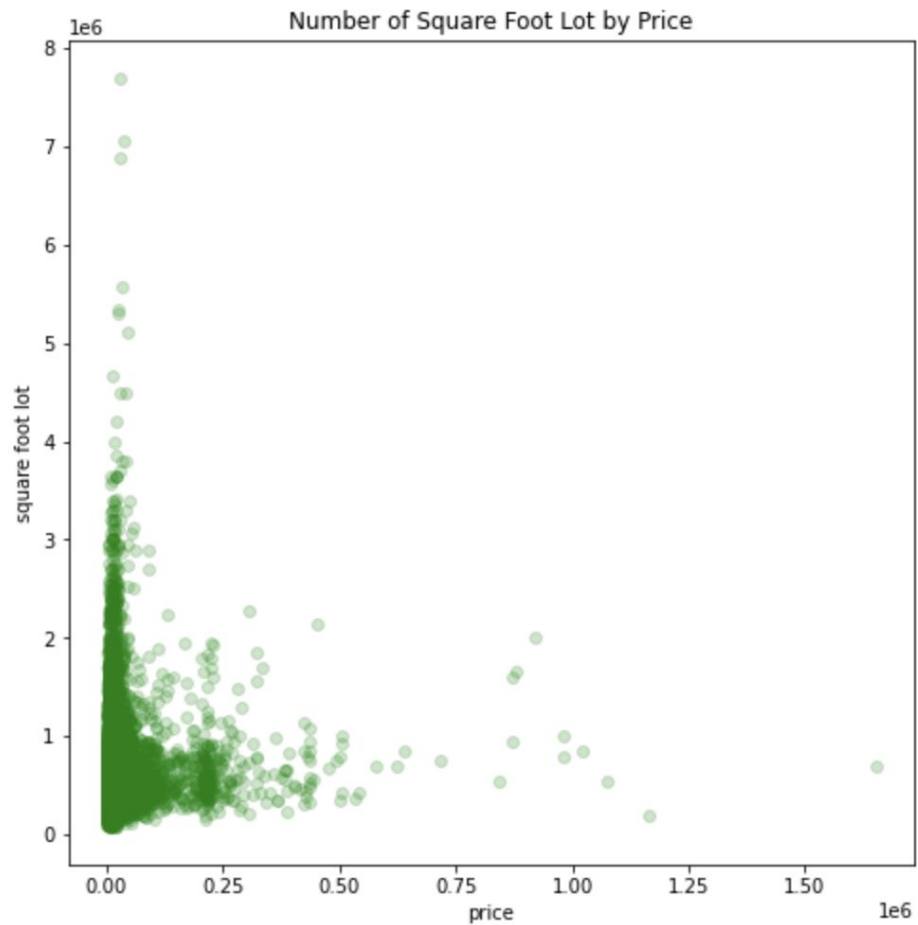
Number of Bedrooms by Price

Number of Bedrooms by Square Foot Living

Number of Square Foot Lot by Price

Conclusion

The third model has showed square foot lot space, square foot living spaces, and bedrooms spaces were good variables to further analyse.

However, when we look at these variables, we can also see that there are many outliers that need to be dropped to get a clearer understanding of normality before moving forward.

Next Steps

In the future, we as analysts should sharpen our focus on our normality testing, by dropping the outliers in this data.

I recommend continuing to explore square foot living space, square foot lot space, and bedrooms space as variables. They have the lowest random chance, they also offer the most opportunity for renovation in the future. I would also recommend, web scraping for public health data about what black families need as they grow intergenerational families.