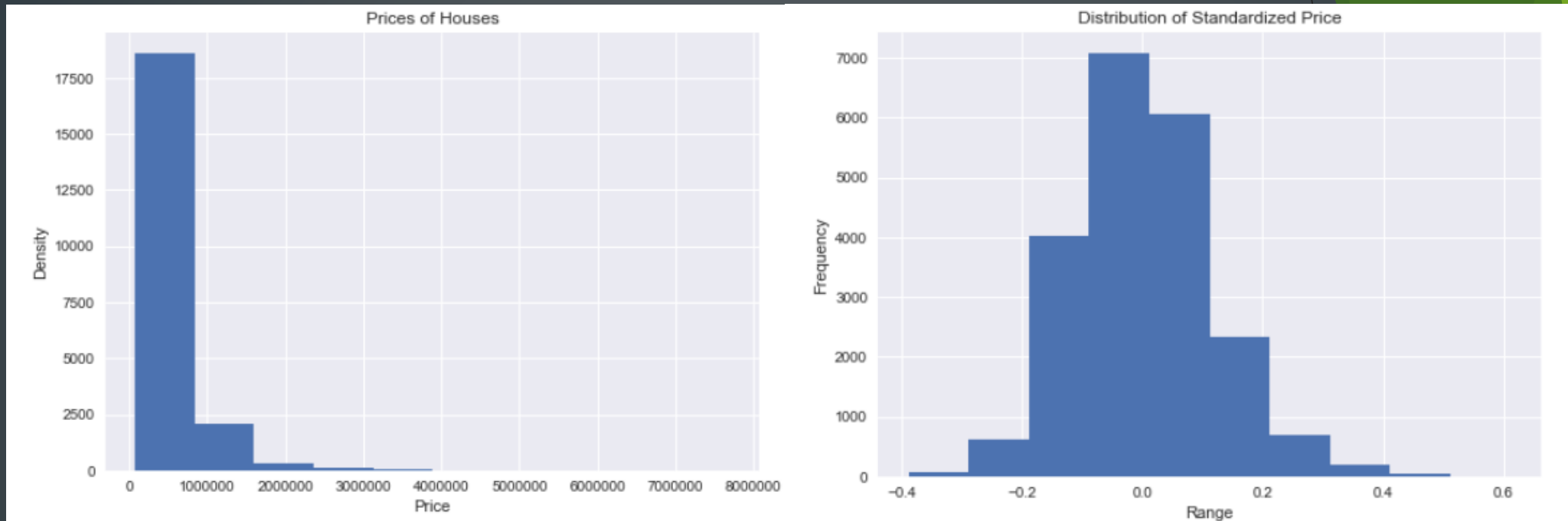


Housing Prices in King's County

What are the Relevant Variables?

- ▶ Price: Price of the house
- ▶ View: Score of the view that the house has
- ▶ Sqft_living: Square Feet of Living Room
- ▶ Yr_built: the year the home was built
- ▶ Condition: Condition of the house
- ▶ Waterfront: Whether the House was on the waterfront
- ▶ Floors: Number of floors on the house
- ▶ Zipcode: Zipcode of the house
- ▶ Sqft_lot15: Square Feet of the lot in 2015
- ▶ Yr_renovated: The year the home was renovated
- ▶ Long: Longitude of the house
- ▶ Lat: Latitude of the house
- ▶ Sqft_basement: Square Feet of the basement

Normalizing Data

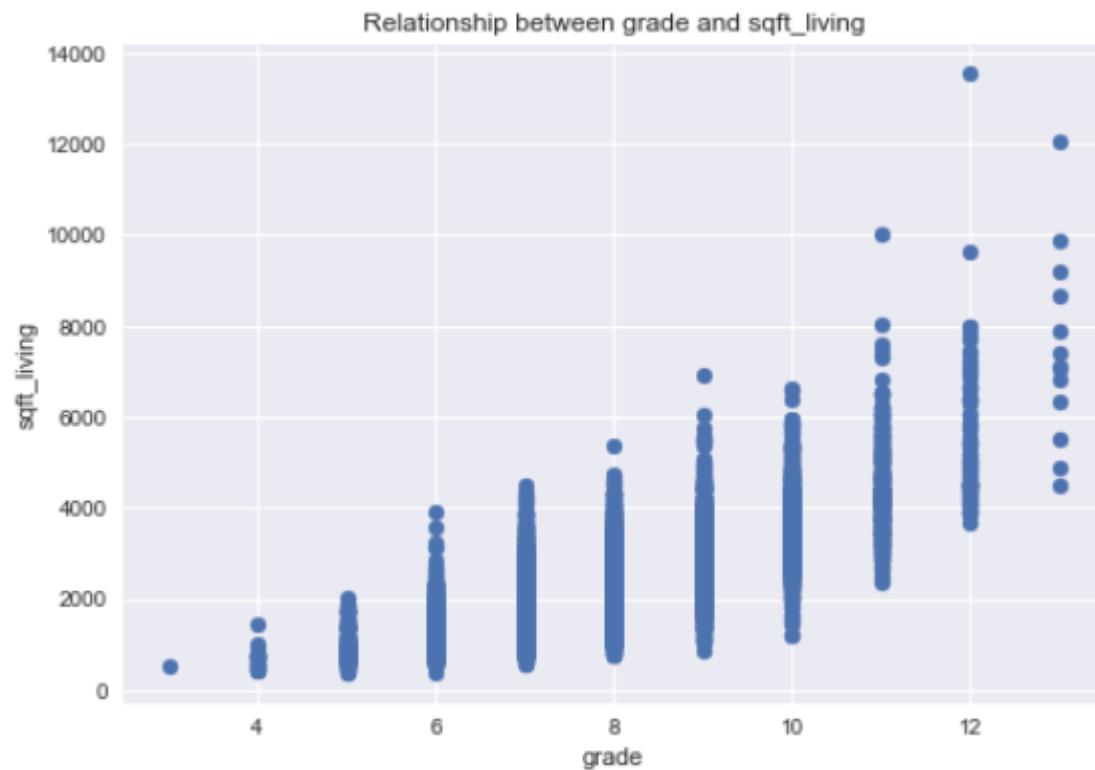


Here we can observe that the data is not normally distributed. Which is bad for modeling and regressions.

After Standardizing the data we get a much more preferable distribution.

Models are far more accurate after the data has been normalized and we can see results with greater accuracy.

Variable Selection



Our visualization shows us that as grade increases, sqft_living also goes up. This is bad because the impact of each individual effect gets hidden by the other one. The best option is generally to drop one of the variables and, generally drop the variable that drops the R2 by the least.

Interpreting Results/ Things to Consider When Buying a House.

- Latitude: Moving North seems to increase price
- View: A Higher view score increases Price
- Sqft_living: A bigger livingroom increases Price
- Yr_built: The later the house is built, the more it costs
- Condition: The condition of a house increases it's price
- Waterfront: A home on the waterfront will sell for more
- Floors: More Having More floors increases the price of the house
- Zipcode: As zipcode increases, prices decrease.
- Sqft_lot15: as the square feet in the lot increases, price increases.
- Yr_renovated: The later a home is renovated, the higher the price.
- Longitude: The further West a home is the less it costs.
- Sqft_basement: A larger basement is associated with a lower price.

OLS Regression Results

Dep. Variable:	price	R-squared:	0.708
Model:	OLS	Adj. R-squared:	0.708
Method:	Least Squares	F-statistic:	4265.
Date:	Mon, 29 Oct 2018	Prob (F-statistic):	0.00
Time:	11:37:35	Log-Likelihood:	28742.
No. Observations:	21082	AIC:	-5.746e+04
Df Residuals:	21069	BIC:	-5.735e+04
Df Model:	12		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
const	-0.0005	0.000	-1.070	0.285	-0.001	0.000
lat	0.2165	0.002	106.531	0.000	0.213	0.221
view	0.0882	0.003	34.259	0.000	0.083	0.093
sqft_living	1.0403	0.009	116.765	0.000	1.023	1.058
yr_built	-0.0374	0.002	-16.036	0.000	-0.042	-0.033
condition	0.0819	0.004	19.685	0.000	0.074	0.090
waterfront	0.0685	0.006	12.087	0.000	0.057	0.080
floors	0.0590	0.003	22.100	0.000	0.054	0.064
zipcode	-0.0399	0.002	-19.788	0.000	-0.044	-0.036
sqft_lot15	0.0376	0.014	2.609	0.009	0.009	0.066
yr_renovated	0.0250	0.002	10.089	0.000	0.020	0.030
long	-0.0631	0.005	-13.066	0.000	-0.073	-0.054
sqft_basement	-0.0641	0.006	-10.241	0.000	-0.076	-0.052

Recommendations:

- ▶ 1: Renovate your house!
 - ▶ Most of the variables in this dataset like the Zipcode, the Latitude, or the whether it is in front of water are immutable traits that you as a homeowner cannot change. However, renovating a house in the year 2000 has an estimated \$8000 impact on the price of the home!
- ▶ 2: Things you may consider when renovating:
 - ▶ Floors, the size of your living room in square feet tend to increase the price of your home as well. It may be worth to make these changes as they will increase the resale value of your homes.
- ▶ 3: Keep your home in the best possible condition:
 - ▶ The only other variable left that you are able to have an impact on is the amount of care you put into your home. Condition also correlates positively with price and each additional grade they have on condition increases the price of the home on average by \$23,070.