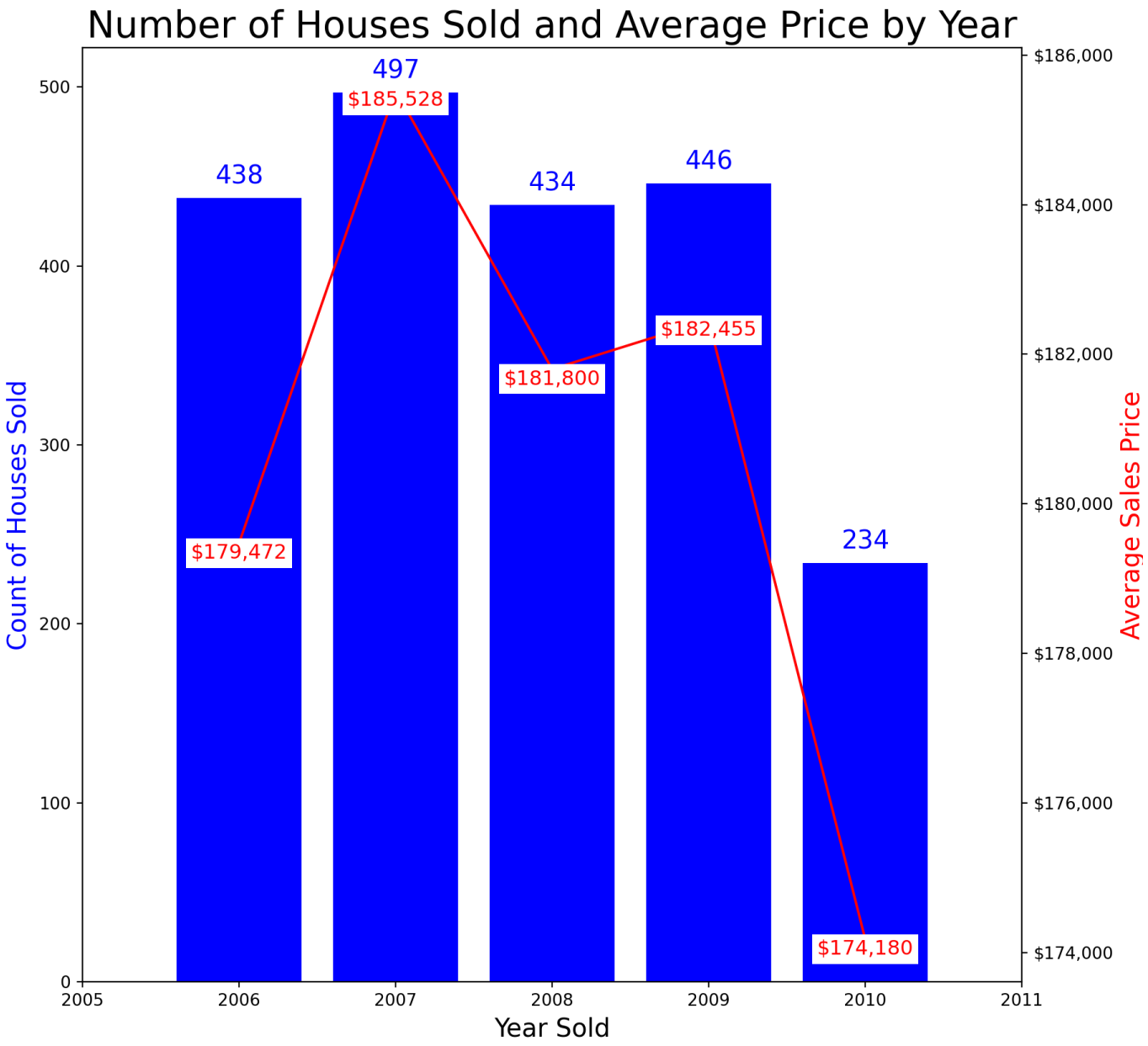
Predicting the Impact of the Great Recession on Housing Prices in Ames, Iowa

David Kanevsky, DSB 318

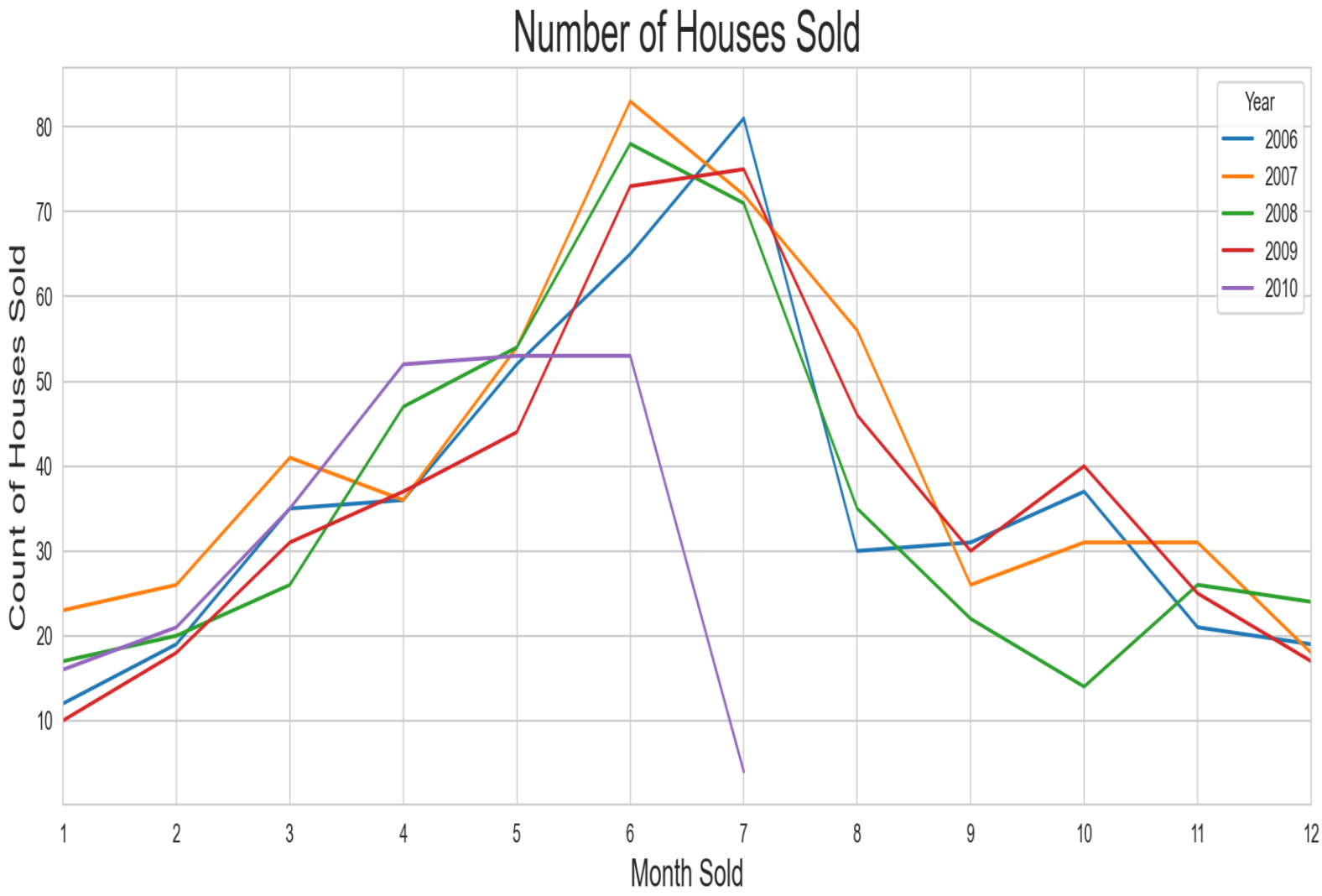
2024-04-19

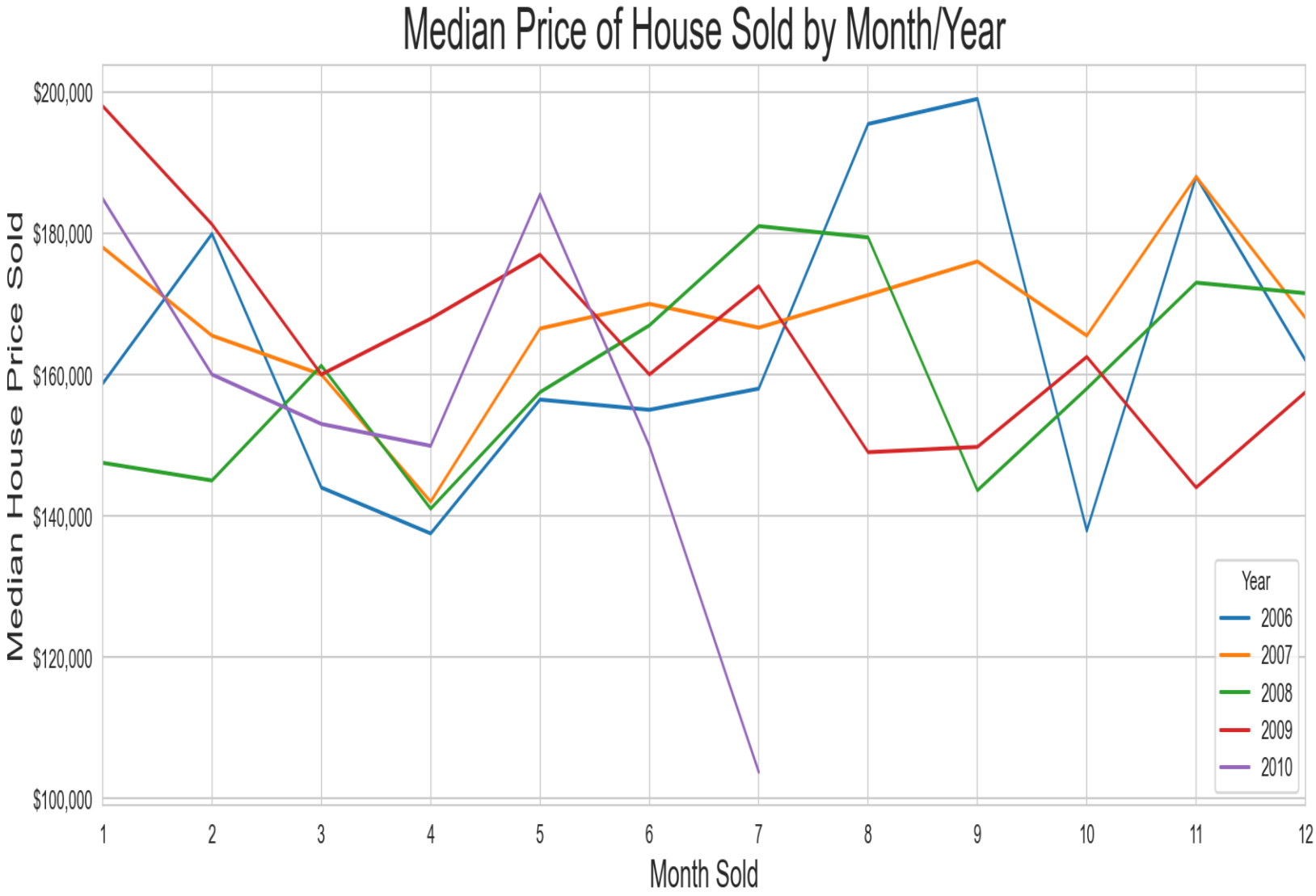
Initial and Cleaned Data

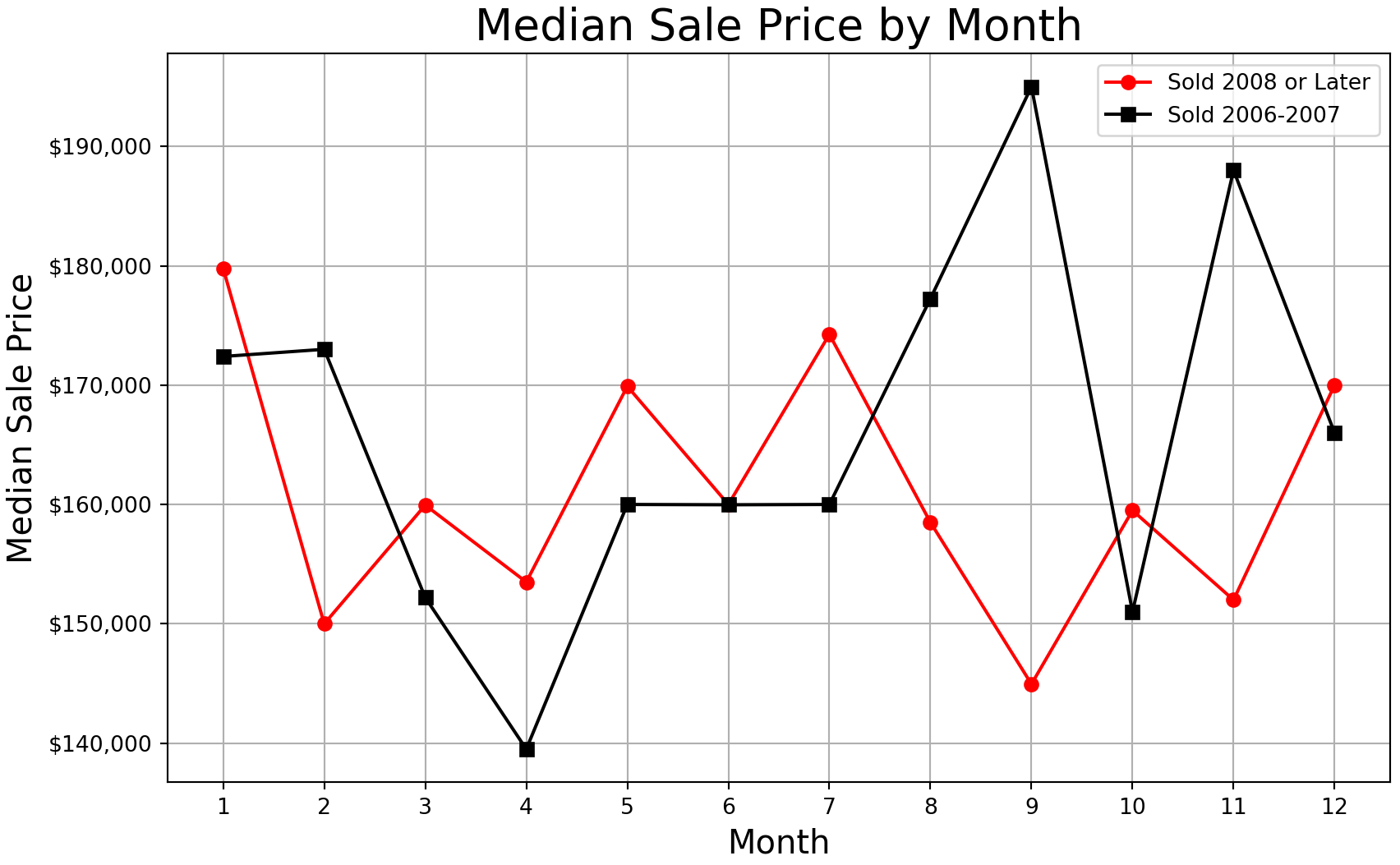
- Population for *196* countries from the year 1800 projected to the year 2100 (population.csv)
- Gross National Income (GNI) per capita for *191* countries from the year 1800 projected to the year 2100 (gni_per_cap_atlas_method_con2021.csv). GNI is reflected in USD dollars.
- Despite being numeric data, both data sets are formatted as strings with population size formatted as k for thousands, M for millions, B for billions, and GNI formatted so GNI over $10,000 represented as $10k.
- Removed missing NULL data from both data sets and merged two data sets together, resulting in *190* countries to analyze.
- Since the goal is to look at projections over the next 10 years (as projections further out may be unreliable due to unforeseen events), grabbed data for just the years 2024 to 2034, converted strings to numeric values, and calculated GDP (GNI per capita * population) for each country for each year.
- Calculated GDP growth in both raw dollars (2034 GDP – 2024 GDP) and percentage growth ((2034 GDP – 2024 GDP) / 2024 GDP).
- Final data set: country_gdp_pop_data.csv
- ***Key question: What countries will see the greatest increase in GDP over the next 10 years?***
  - Imagine countries as index funds. Some countries start out with a higher valuation, so even if they have growth potential, starting costs are higher and less ability to diversify funds in case an investment fails. ***What are some low-cost countries with high growth potential?***


Number of Houses Sold and Average Price by Year
Count of Houses Sold
Average Sales Price
Year Sold
438
497
434
446
234
$179,472
$185,528
$181,800
$182,455
$174,180
0
100
200
300
400
500
2005
2006
2007
2008
2009
2010
2011
$174,000
$176,000
$178,000
$180,000
$182,000
$184,000
$186,000

SLIDE 3


Number of Houses Sold
Count of Houses Sold
Month Sold
Year
2006
2007
2008
2009
2010
10
20
30
40
50
60
70
80
1
2
3
4
5
6
7
8
9
10
11
12


Median Price of House Sold by Month/Year
Median House Price Sold
$200,000
$180,000
$160,000
$140,000
$120,000
$100,000
1
2
3
4
5
6
7
8
9
10
11
12
Month Sold
Year
2006
2007
2008
2009
2010


Median Sale Price by Month
Sold 2008 or Later
Sold 2006-2007
$190,000
$180,000
$170,000
$160,000
$150,000
$140,000
Median Sale Price
1
2
3
4
5
6
7
8
9
10
11
12
Month

Features Included	Type	Model Name	Test R2	Cross Val Score	RMSE
All numeric data	Linear	lr_numeric	0.894110441	0.881697695	25255.1008
All numeric data with correlation above 0.3 with sale price	Linear	lr	0.875672017	0.873416735	28068.84285
All numeric data with correlation above 0.3 with sale price and standard scaler applied	Linear	lr_scaled	0.88023	0.870541025	26827.90016
Just categorical variables	Linear	lr_dum	0.698970723	-1.86E+22	43082.6478
7 variables in categorical data	Linear	lr_fewer_dum	0.632994818	-1.25E+23	48346.82812
All numeric data with correlation above 0.3 with sale price and 7 categorical variables	Linear	lr_comb	0.883459199	0.895960703	27389.97766
All numeric data with correlation above 0.3 with sale price and 7 categorical variables	Ridge	ridge_cv	0.884026321	0.896286349	309891.0271
All numeric data with correlation above 0.3 with sale price and 7 categorical variables	Lasso	lasso_cv	0.883497096	0.896037228	27385.52394
All numeric data using polynomial features and ridge regularization	Ridge	ridge_overfit_cv	0.920055403	0.923993778	24960.53399
All numeric data and 7 categorical variables	Linear	lr_filt_num_cat	0.904482985	0.886431246	24633.78271
All numeric data and 7 categorical variables, without "Sold in Crisis"	Linear	lr_drop_crisis	0.893127823	0.897592051	26429.31064
All numeric data and 7 categorical variables using a standard scaler	Linear	ols	-1.5726E+14	-1.70E+21	9.99541E+11

Conclusions and Recommendations

- To really see the impact of the financial crisis on housing prices, would want to look at a longer time window – before 2006 and after 2010, to look at impact of financial crisis on housing prices.

- Would also want to adjust housing prices to reflect changes in CPI / purchasing power. For example, 2009 prices slightly higher than 2008 and 2006 (but lower than 2007), but were prices higher than accounting for inflation.