

MODELING THE DC AREA REAL ESTATE MARKET

WALTER TYRNA | METIS



BACKGROUND & GOALS

The market is the craziest it's been in 15 years, with buyers competing against dozens of other offers and waiving every contingency.

-Washington Post, March 2021

Model goals:

Can we build a model that determines the main factors that contribute to a property's ultimate sale price?

Will this model be able to reasonably predict new home sale prices?

Potential clients:

Home-sellers, real estate agencies, investors

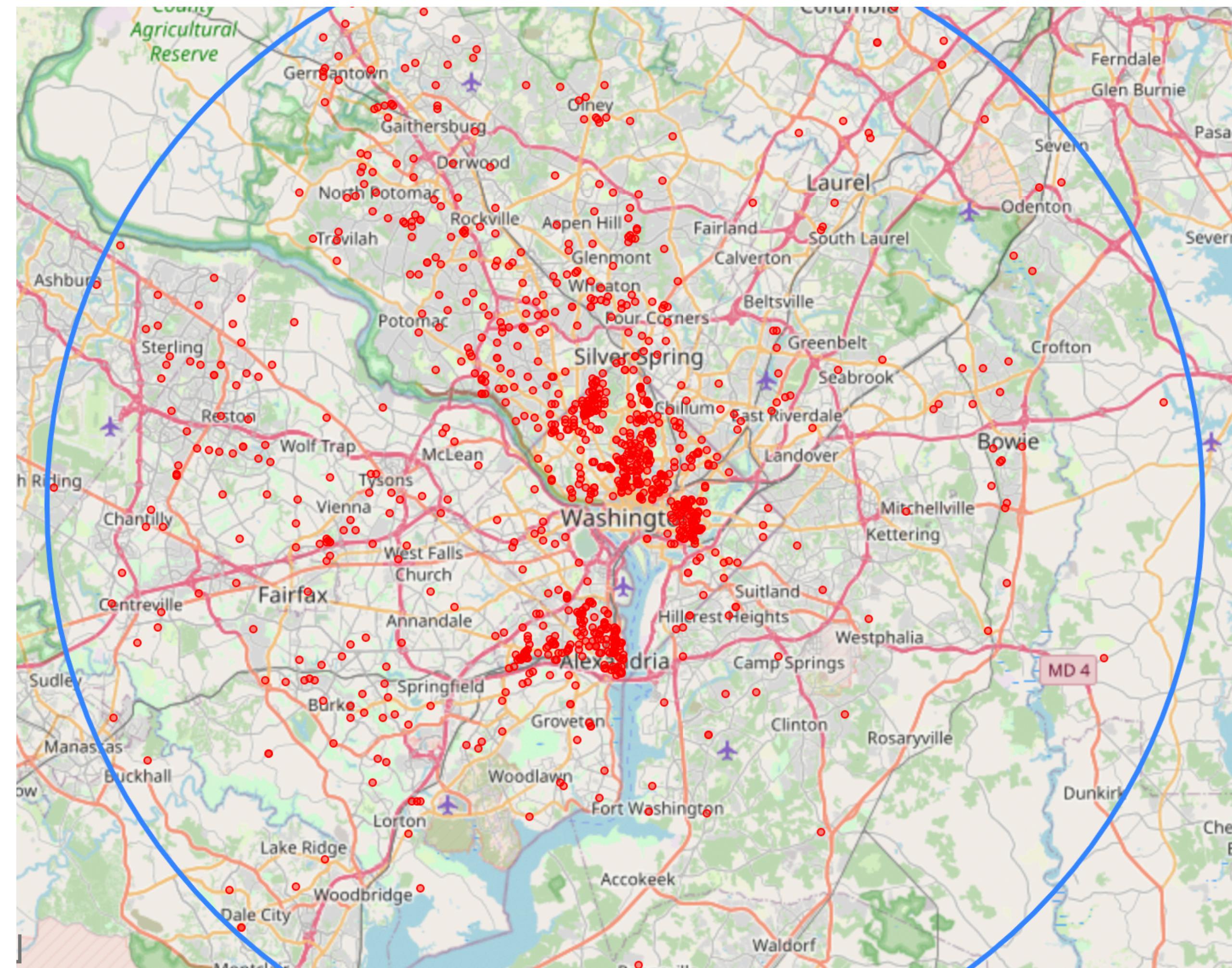
METHODOLOGY - DATA

Data:

- 2840 real estate listings from Compass (*beautifulsoup*)
- DC Metro (WMATA) locations (*beautifulsoup*)
- Coordinates for properties and metro locations (*geopy*)

Scoping:

- Single-family homes & townhomes
- 25 mi of downtown DC (Farragut Sq)
 - Post scoping data -> ~1000 listings



METHODOLOGY - MODEL & FEATURE SELECTION

A preference for interpretability:

- Linear Regression
- Lasso Regression

Initial features (15):

Physical:

- # Beds
- # Baths
- Lot size (sqft)
- Home size (sqft)
- # Stories
- Overall condition
- # Parking Spaces
- Does it have a garage?
- Does it have A/C?
- Property age
- Property type
- Assessed tax value**

Location:

- General location (DC/MD/VA)
- Distance from downtown
- Distance from metro station

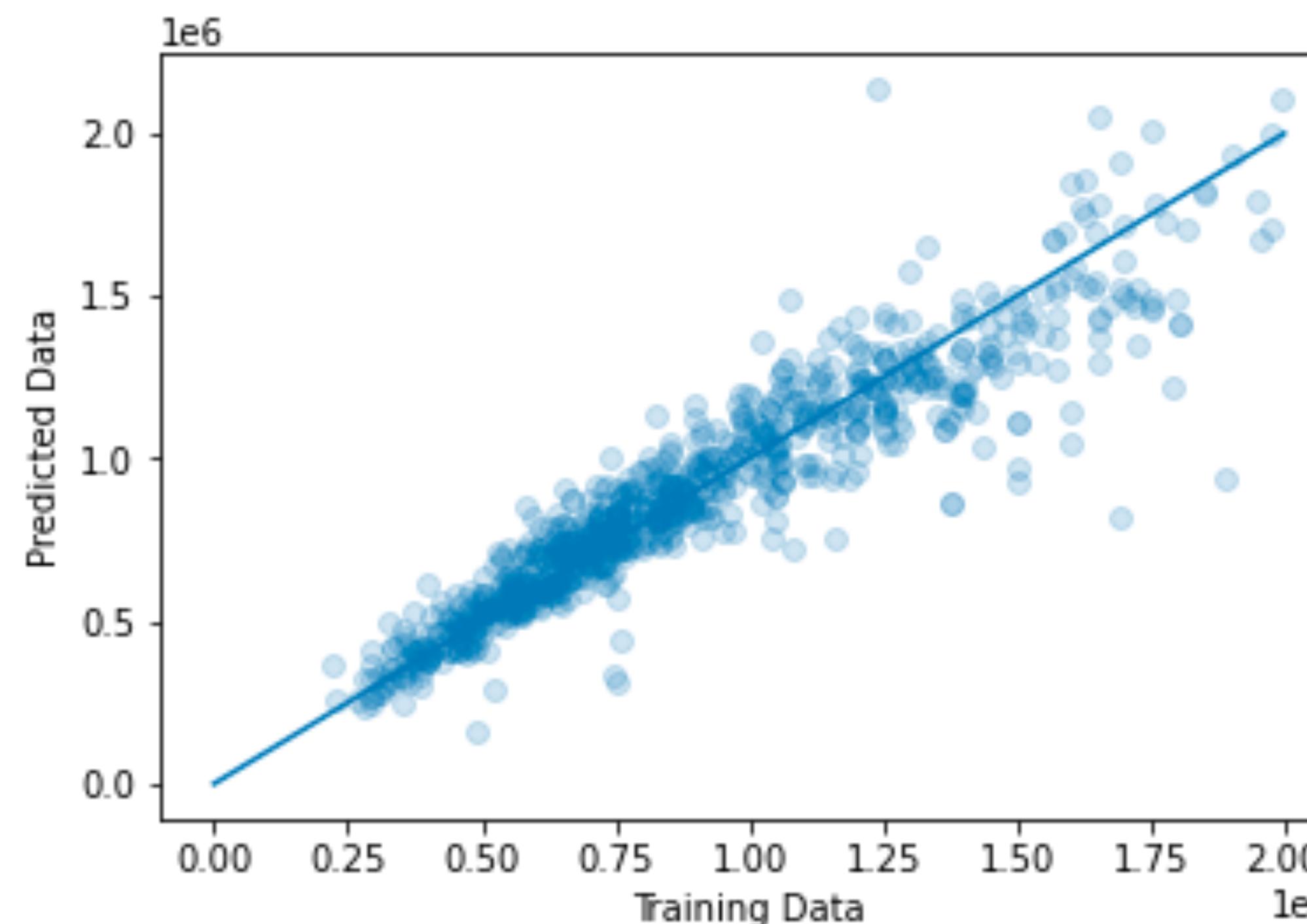
MODEL & FEATURE RESULTS

Best results: Lasso Regression

R²: 0.862

Mean absolute error (MAE): \$89,836

Predicted Sales Price versus Actual Sales Price using Lasso Model



Geographic Ranking:

- 1) DC (west of Anacostia River)
- 2) DC (east of Anacostia River)
- 3) Maryland
- 4) Virginia

Key Property Features:

- Property Type - Single Family Home
- Property Condition - Excellent
- Number of stories
- Having a Garage
- Number of Parking Spaces

MODEL & FEATURE RESULTS (CON'T)

Residual Errors for Lasso Model on DC Real Estate Data



811 Plum St SW, Vienna, VA 22180

\$1,887,045 5 6 5,582
Sold Price Beds Baths Sq Ft



8321 Forrester Blvd, Springfield, VA 22152

\$540,000 3 3 1,970
Sold Price Beds Baths Sq Ft

MODEL IMPROVEMENT

- Including seasonality of real estate sales
- Including data sources outside of Compass to broaden dataset



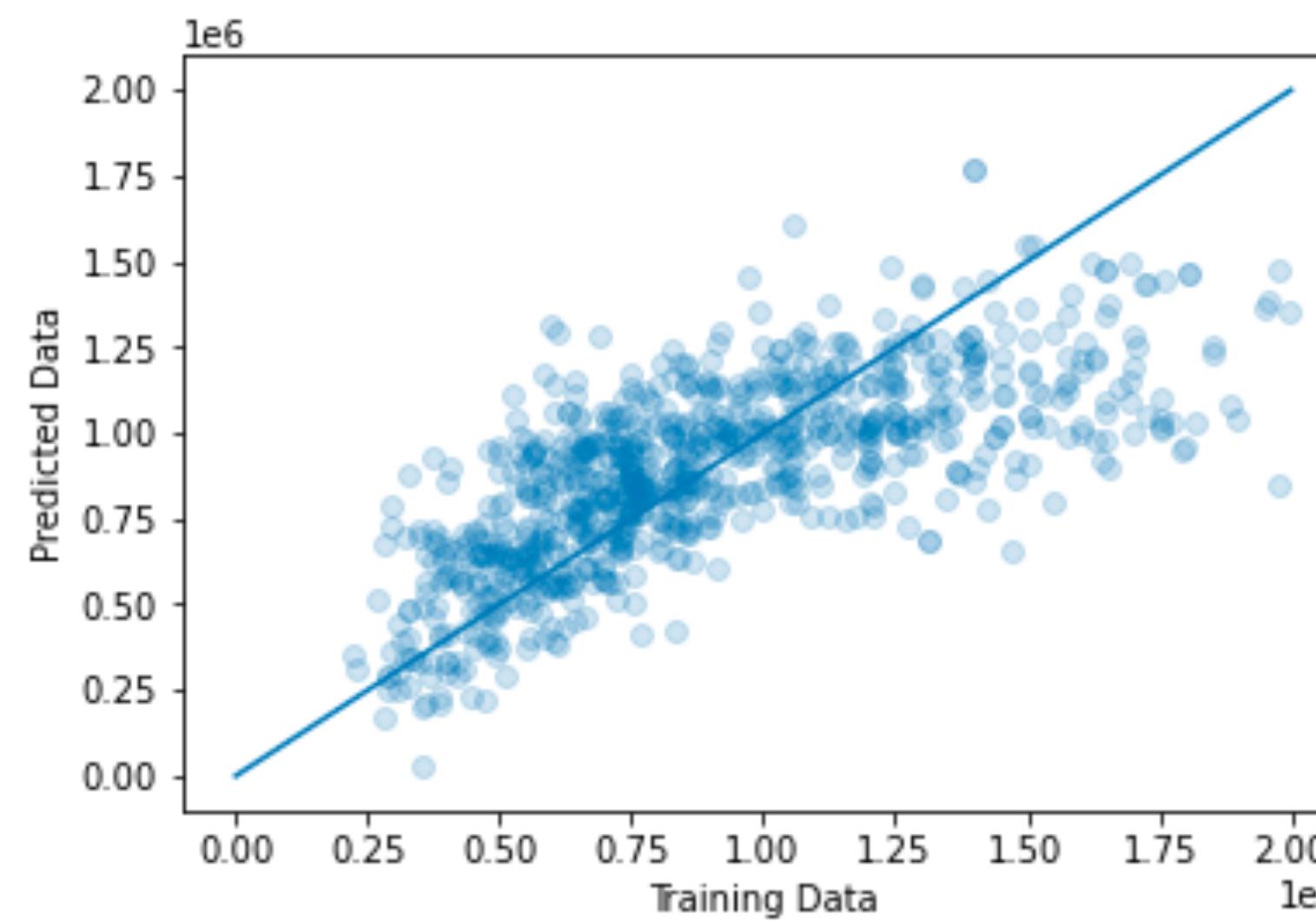
APPENDIX

Model Results Excluding Assessed Tax Value

Linear Regression:

- R² Train: 0.494
- R² Test: 0.454

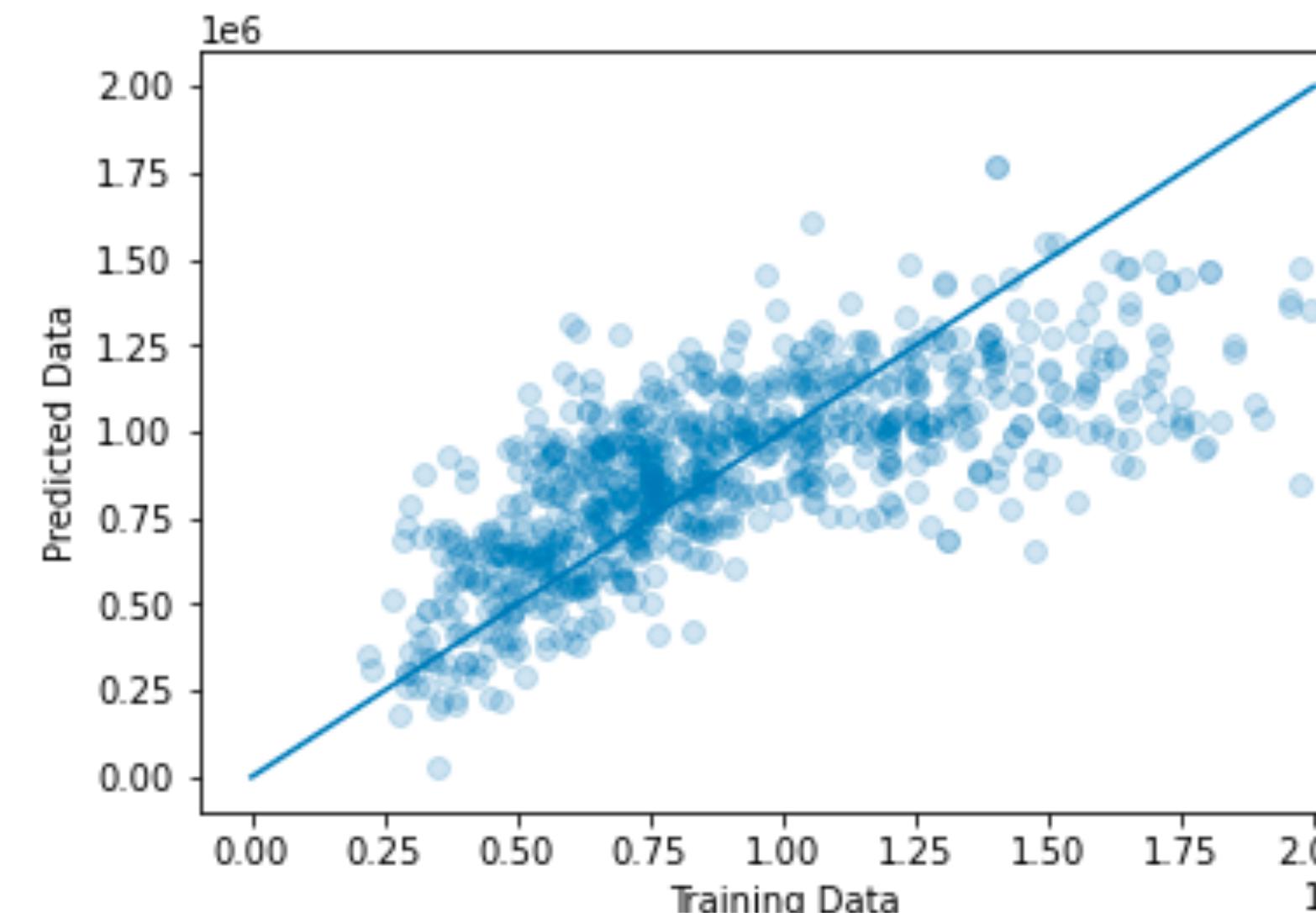
Fitted Sales Price versus Actual Sales Price using Linear Regression Model



Lasso Regression:

- R² Train: 0.494
- R² Test: 0.454

Predicted Sales Price versus Actual Sales Price using Lasso Model



APPENDIX (CON'T)

Model Results Excluding Assessed Tax Value

Lasso Regression:

- MAE: \$197,647

Residual Errors for Lasso Model on DC Real Estate Data

