Modeling Voter Turnout in Local Elections With Housing Data

Elliott Day

November 20, 2020

Odd Year Elections

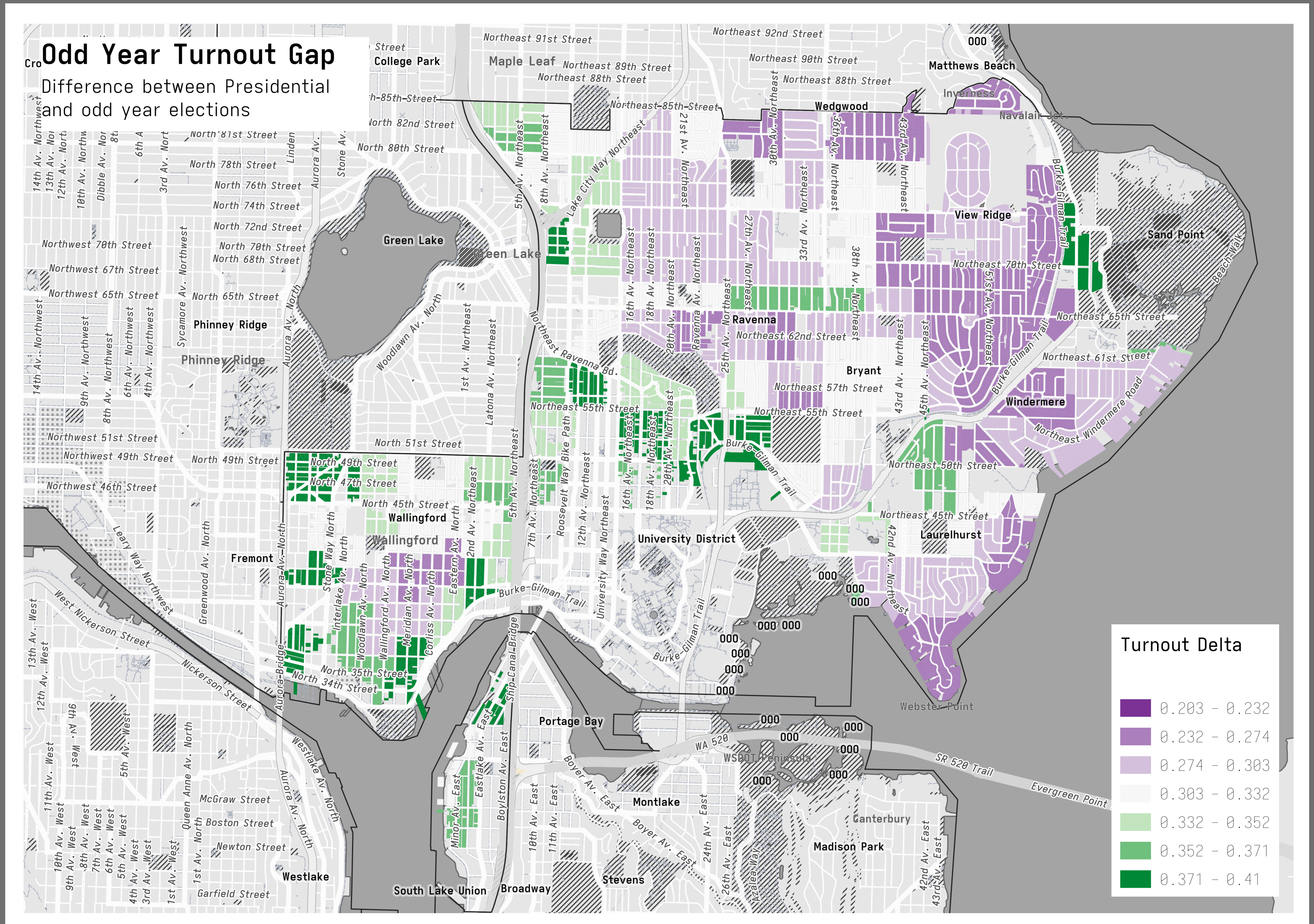
In the State of Washington, elections for local office are held in odd-numbered years.

This results in an electorate for local offices that is different than the one for congressional seats and president.

An Example

An Example

Shaun Scott, candidate for City Council in Seattle Council District 4



An Example

Shaun Scott, candidate for City Council in Seattle Council District 4

Lost by more than 4 points in an odd year

If precincts had voted with the same preference (conservative assumption) at the rate of the 2020, that margin shrinks to less than a percent

2019 Seattle City Council District 4

	Actual	Presidential Turnout
Shaun Scott	15568	24548
	45.84%	47.33%
Alex Pedersen	16954	24982
	49.92%	48.16%
Votes Cast	33962	51871
Registered Voters	61367	61367

Why does this matter?

Why does this matter?

Odd year elections distort the signal from local elections and result in less representative and less effective local government.

Why does this matter?

Odd year elections distort the signal from local elections and result in less representative and less effective local government.

(And either way if you want to win local elections you need to be prepared for the electorate that will be participating.)

What are we doing here?

What are we doing here?

- Using housing data to build a durable estimator of hypothetical voter turnout scenarios based on previous turnout patterns.
- Of interest to campaigns, election administrators, researchers, and nonprofits
- The dataset is extensive and has many potential other applications

What are we doing here?

— Using housing data to build a durable estimator of hypothetical voter turnout scenarios based on previous turnout patterns.

— Of interest to campaigns, election administrators, researchers, and nonprofits

What are we doing here?

— Using housing data to build a durable estimator of hypothetical voter turnout scenarios based on historical turnout patterns and future landscapes.

— Of interest to campaigns, election administrators, researchers, and nonprofits

— The dataset is extensive and has many potential other applications

The Data

King County Assessor

A record for every residence in the county (~989,000)
Seattle and surrounding area; relatively diverse demographically and in terms of housing across the county

King County Assessor

A record for every residence in the county (~989,000)

The Voter File

A record for every voter in the county (~1.4 million)

The Targets

— Registered Voters Per Parcel

The Targets

— Registered Voters Per Parcel
— Votes cast in 2019 Per Parcel

The Targets

— Registered Voters Per Parcel
— Votes cast in 2019 Per Parcel
— Votes cast in 2020 Per Parcel

The Models

The Models

— RandomForestRegressor

The Models

— RandomForestRegressor

— Keras Neural Network

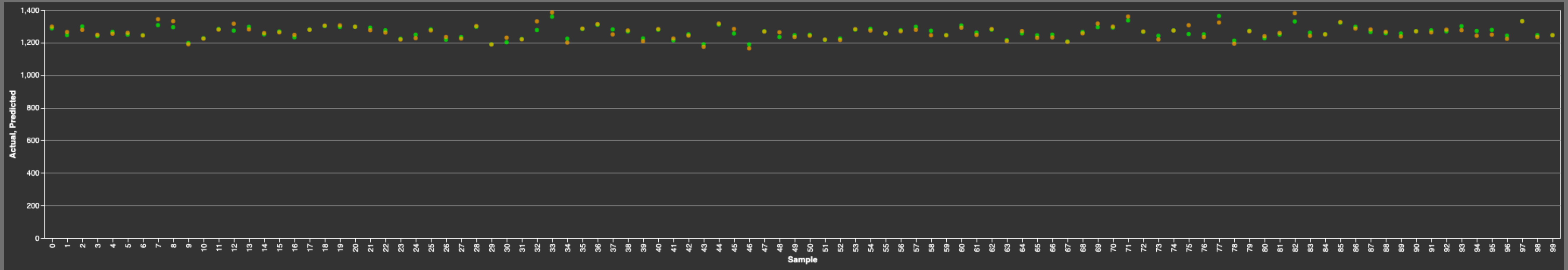
Random Forest

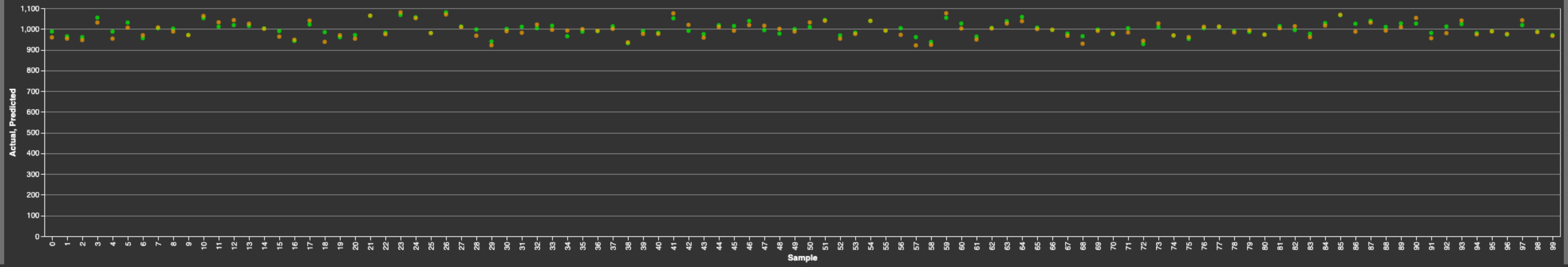
— RandomizedSearchCV
— Then trained on full dataset

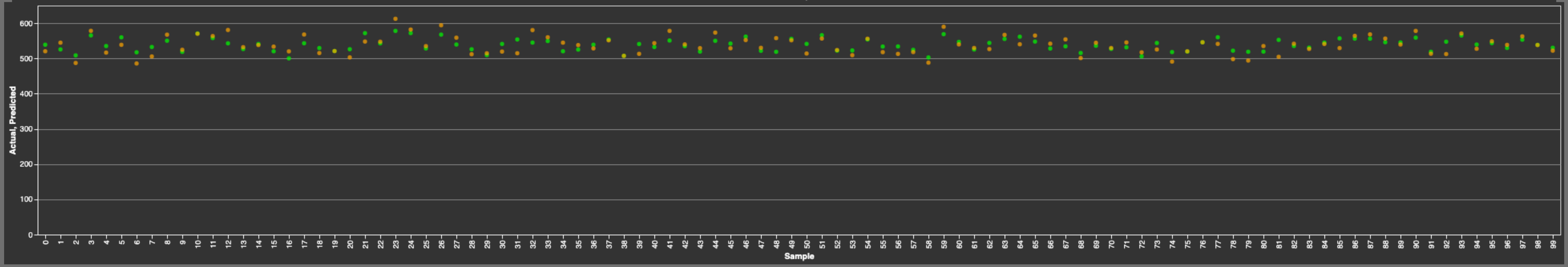
Neural Regressor

— Early stopping
— Lots of trial and error

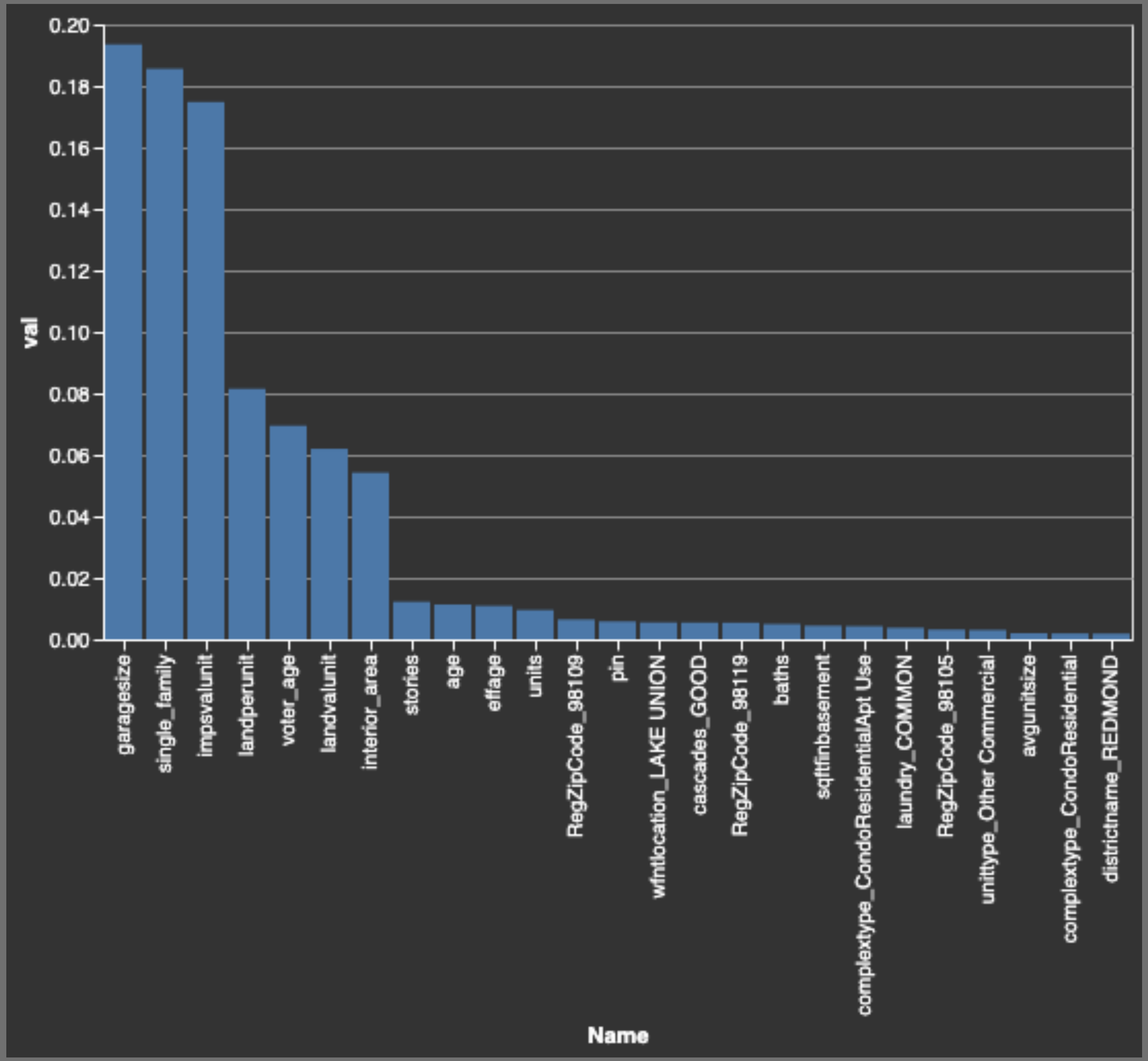
Aggregate Accuracy







Feature Importances



Evaluation

Evaluation

Goal:

Evaluation

Goal: Compare model performance on different forms and sizes of aggregation

Evaluation

	Random Forest Regression - 2020 Voters				Neural Regression - 2020 Voters		
	Actual	Predicted	RMSE		Actual	Predicted	RMSE
senior_housing	2032.964824	2033.171594	0.000102	fourplex	1567.817949	1557.450270	0.006613
fourplex	1567.817949	1569.035525	0.000777	apartment	41987.903225	42298.559992	0.007399
apartment	41987.903225	41882.981701	0.002499	single_family	202247.673392	197756.910856	0.022204
single_family	202247.673392	202818.624523	0.002823	duplex	2966.606277	3036.821827	0.023669
duplex	2966.606277	2991.836718	0.008505	senior_housing	2032.964824	1982.716823	0.024717
triplex	1210.083333	1230.755636	0.017083	triplex	1210.083333	1242.919038	0.027135
townhouse	5618.548988	5489.541046	0.022961	townhouse	5618.548988	5289.885259	0.058496
mobile_home	259.421984	251.634855	0.030017	mobile_home	259.421984	219.275274	0.154754

What's Next

— Error analysis

— New Features (e.g. gated communities)

— Analyzing and modeling candidate preference, including measuring for when such a preference is expressed in patterns related to housing