



# Triple-A Ames Housing Price Model

DSI26 Project 2

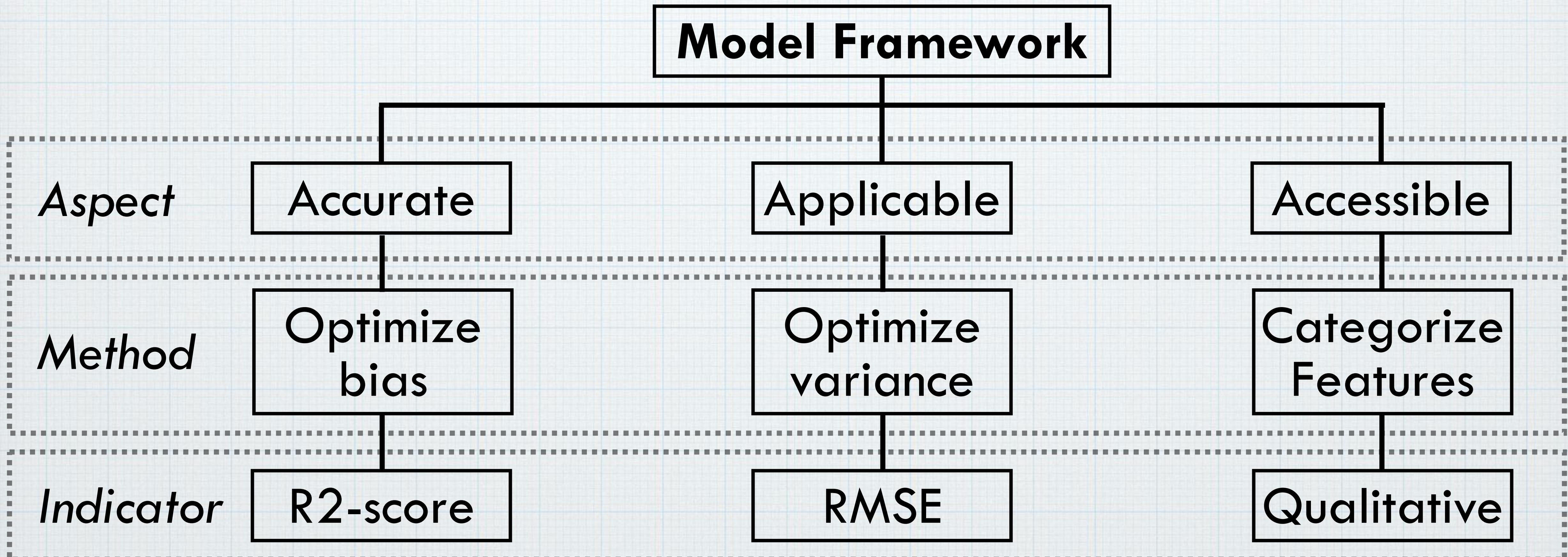


# Problem Statement

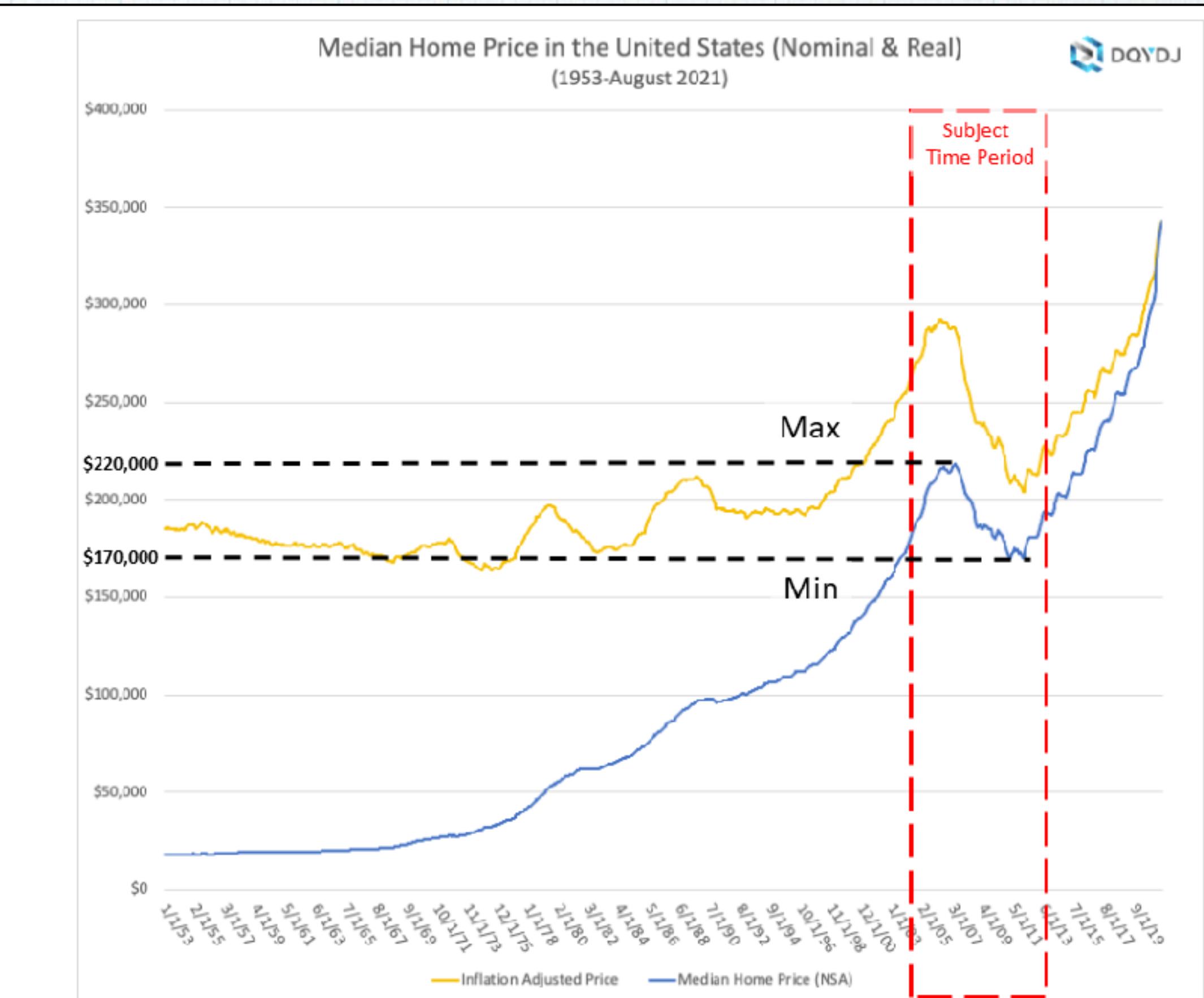
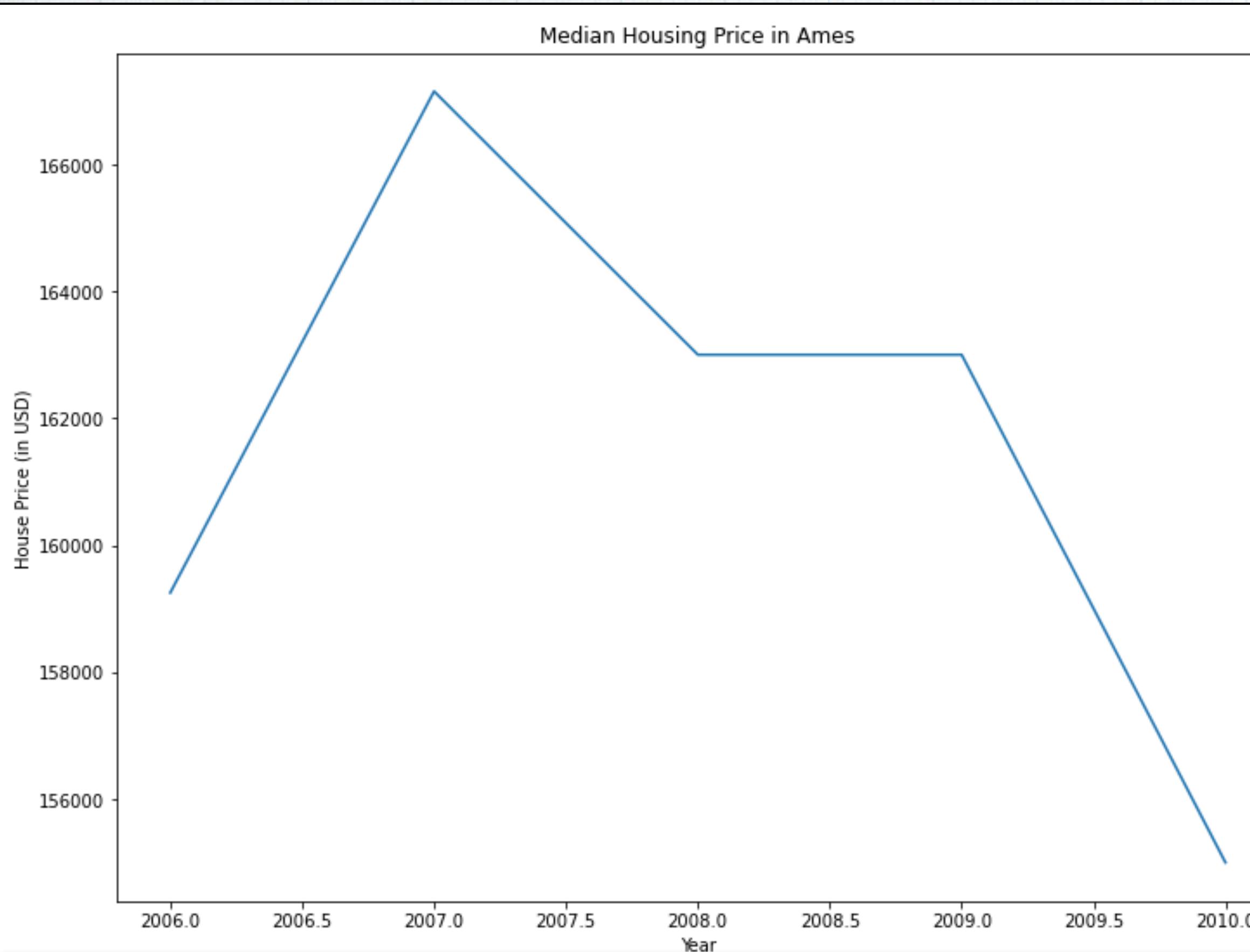
- \* The Ames City Council's digitalisation drive is underway. As the Planning Division's data scientist, develop a housing price model to forecast the sale price of properties in Ames.
- \* Utilise past data on Ames house sales sold between 2006 to 2010, which includes over 2000 records of houses sold and 80 related attributes.



# Triple-A Framework (Accurate, Applicable, Accessible)



# Ames vs US Housing Sale Prices



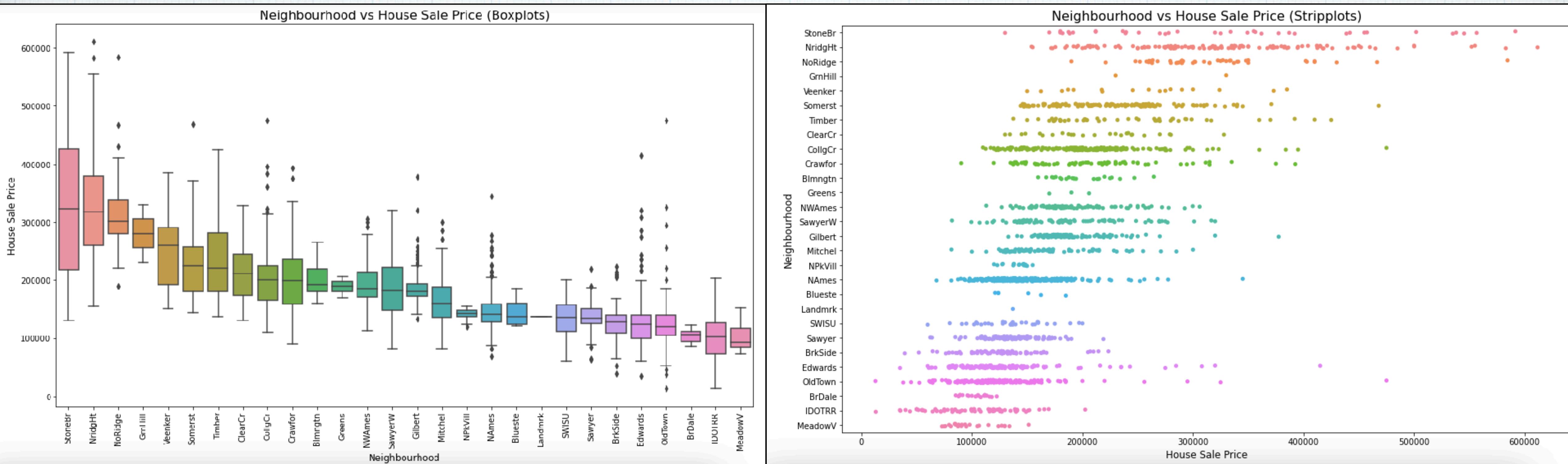
\* Data collection period affected by US subprime mortgage crisis.

\* Housing price bubble burst for Ames (2007) slightly delayed compared to rest of US.

# Categorizing Features

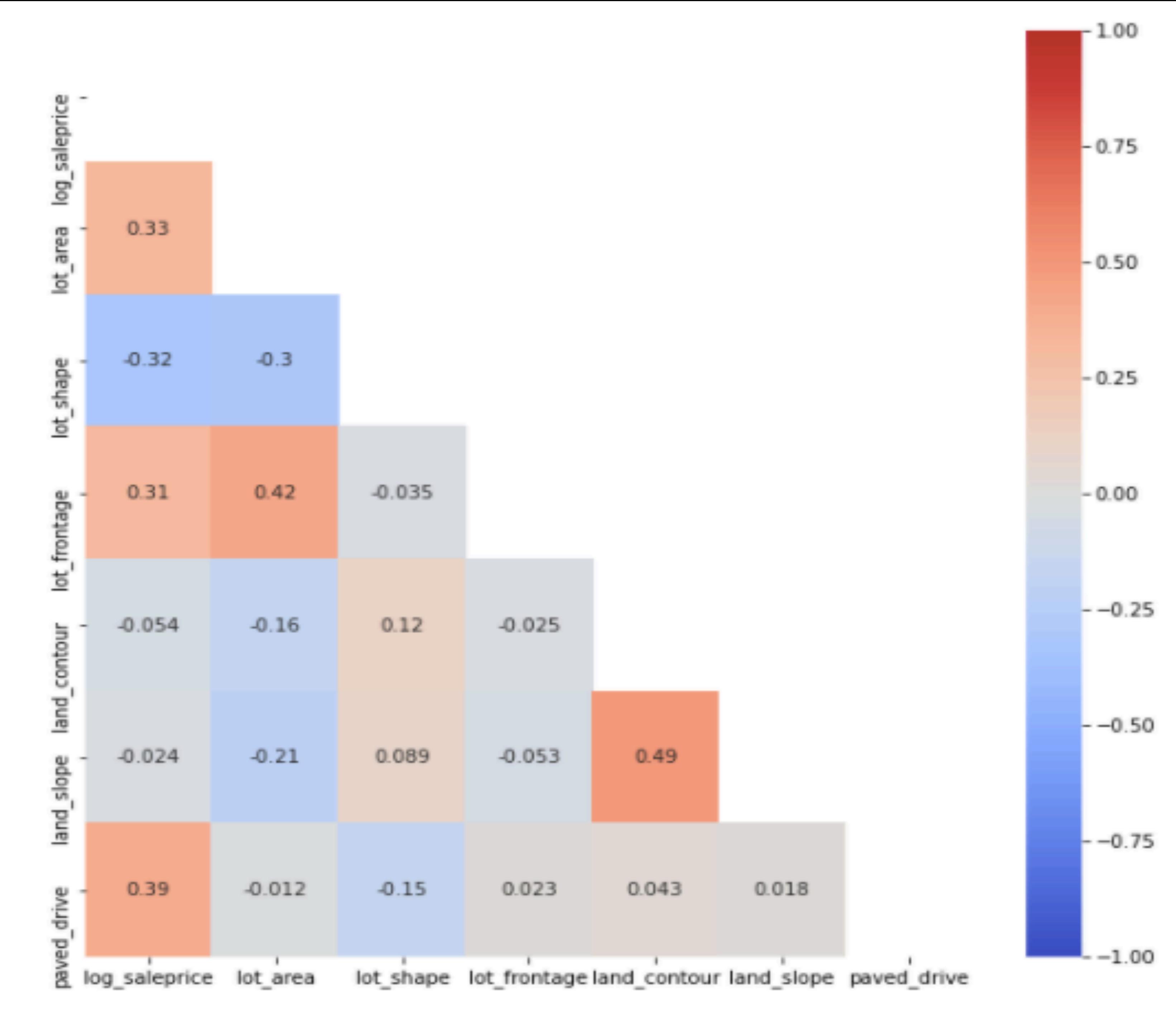
Geography	Parcel	Internal Built Environment	External Built Environment	Miscellaneous
	PID	Lot Area	Gr Liv Area	Exter Qual ID
Neighbourhood	Lot Shape	1st Flr SF	Exter Cond	Year Built
Overall Qual	Lot Config	2nd Flr SF	Foundation	Year Remod/Add
Overall Cond	Lot Frontage	Low Qual Fin SF	Garage Area	Yr Sold
Condition 1	Land Contour	TotRms AbvGrd	Garage Cars	Mo Sold
MS Zoning	Land Slope	Bedroom AbvGr	Garage Type	Utilities
MS SubClass	Paved Drive	Kitchen AbvGr	Garage Qual	Heating QC
Building Type		Kitchen Qual	Garage Cond	Electrical
House Style		Full Bath	Garage Finish	Functional
		Half Bath	Roof Style	Misc Val
		Fireplaces	Exterior 1st	Sale Type
		Fireplace Qu	Exterior 2nd	
		Total Bsmt SF	Mas Vnr Area	
		BsmtFin SF 1	Mas Vnr Type	
		BsmtFin SF 2	Wood Deck SF	
		Bsmt Unf SF	Open Porch SF	
		Bsmt Qual	Enclosed Porch	
		Bsmt Cond	3Ssn Porch	
		Bsmt Exposure	Screen Porch	
		BsmtFin Type 1	Pool Area	
		BsmtFin Type 2		
		Bsmt Full Bath		
		Bsmt Half Bath		

# Geography (Neighbourhood)

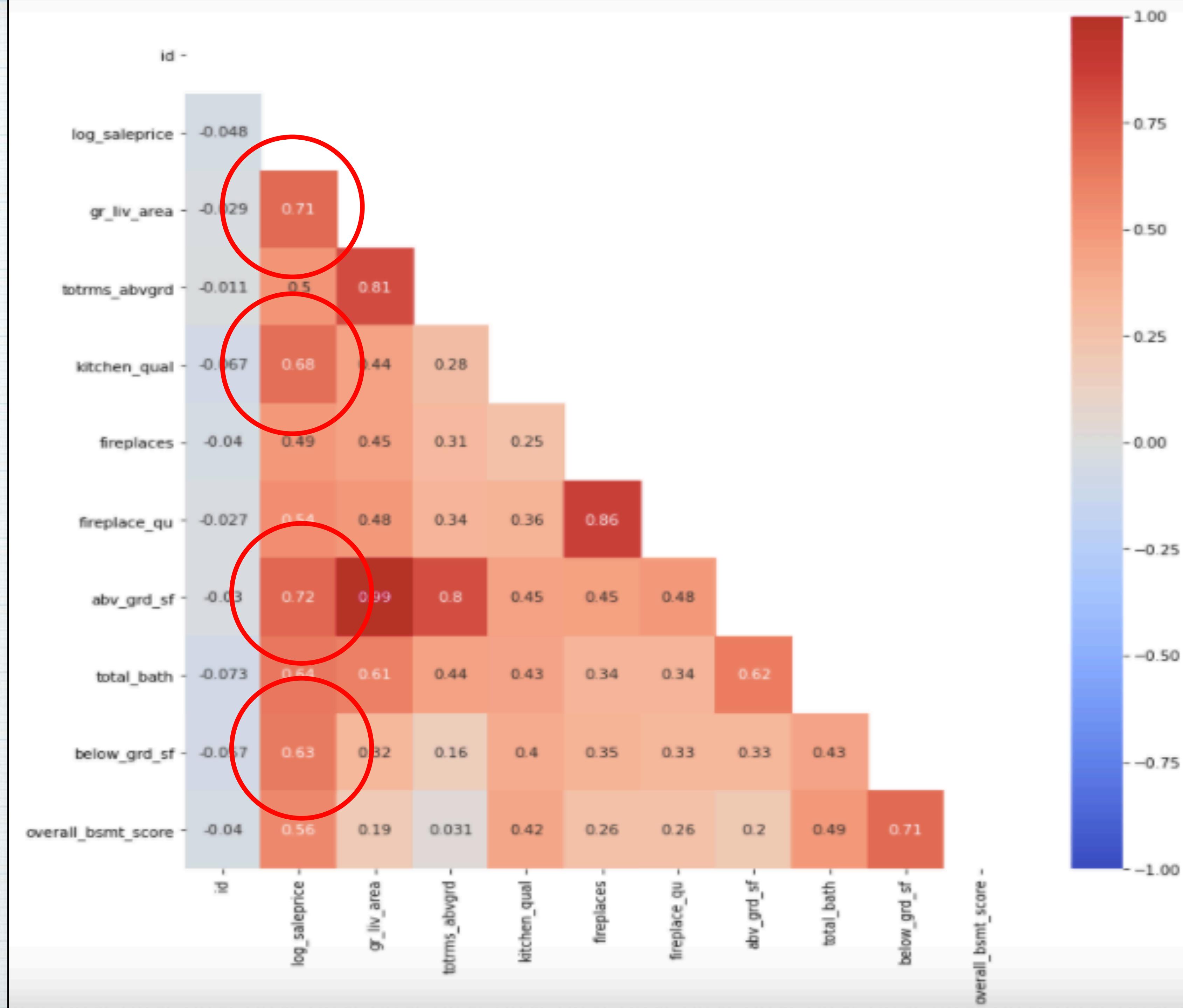


- \* Visible ranked variation in median house sale price by neighbourhood
- \* Lowest (MeadowV): Within airport protection zone
- \* Highest (StoneBr): Near green space/waterbody and high-end recreation

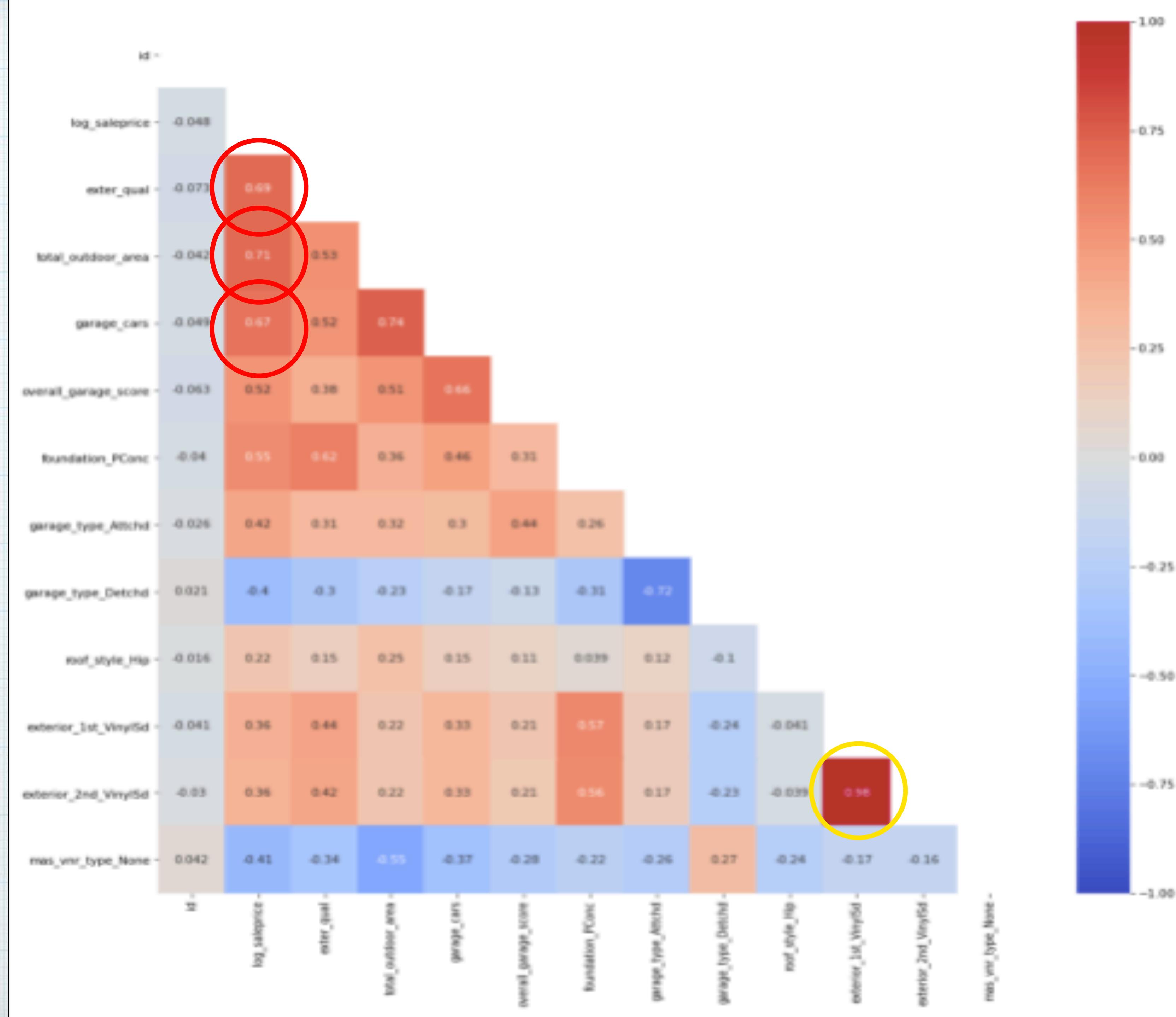
# Parcellation



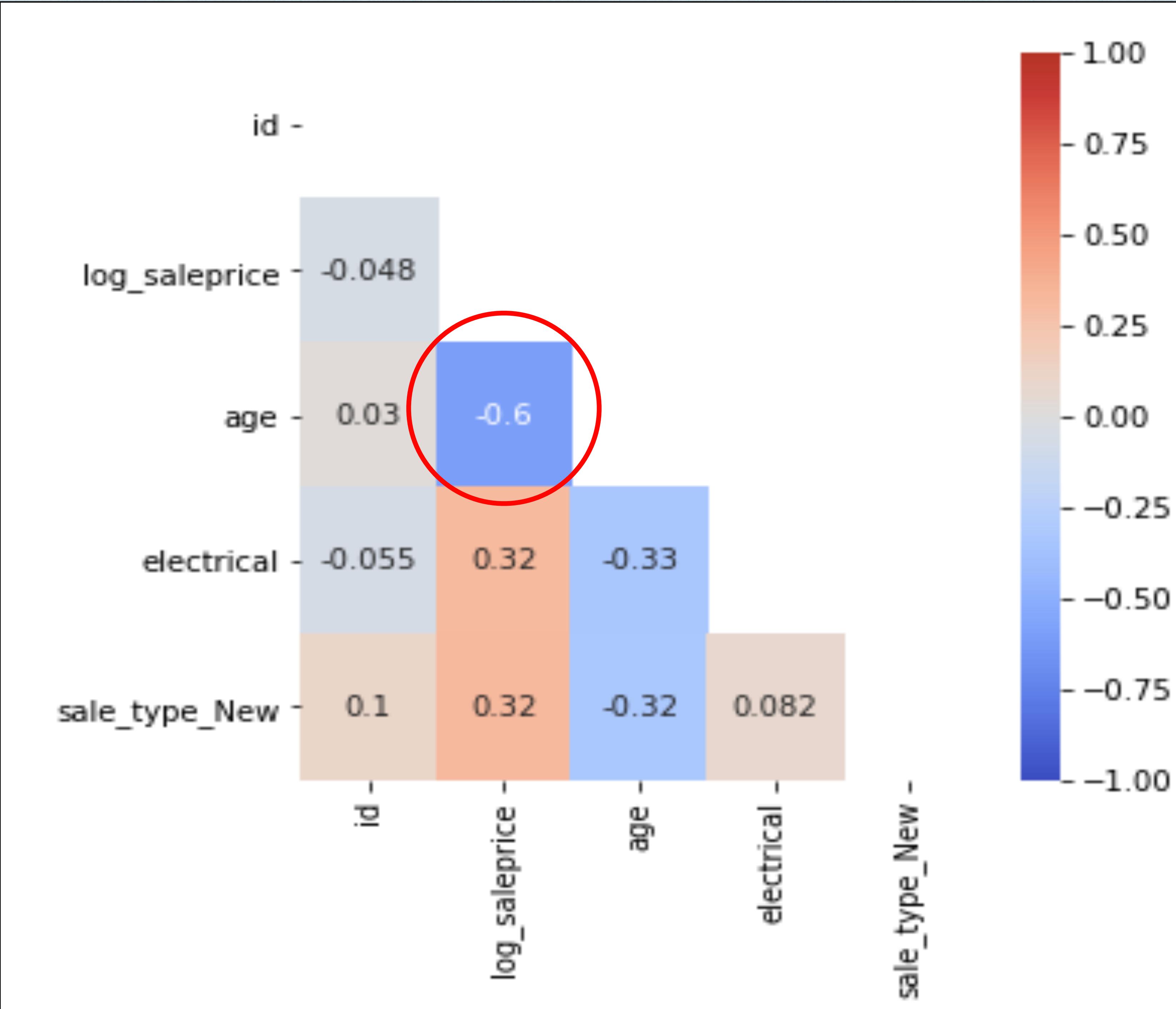
# Internal Built Environment



# External Built Environment



# Miscellaneous



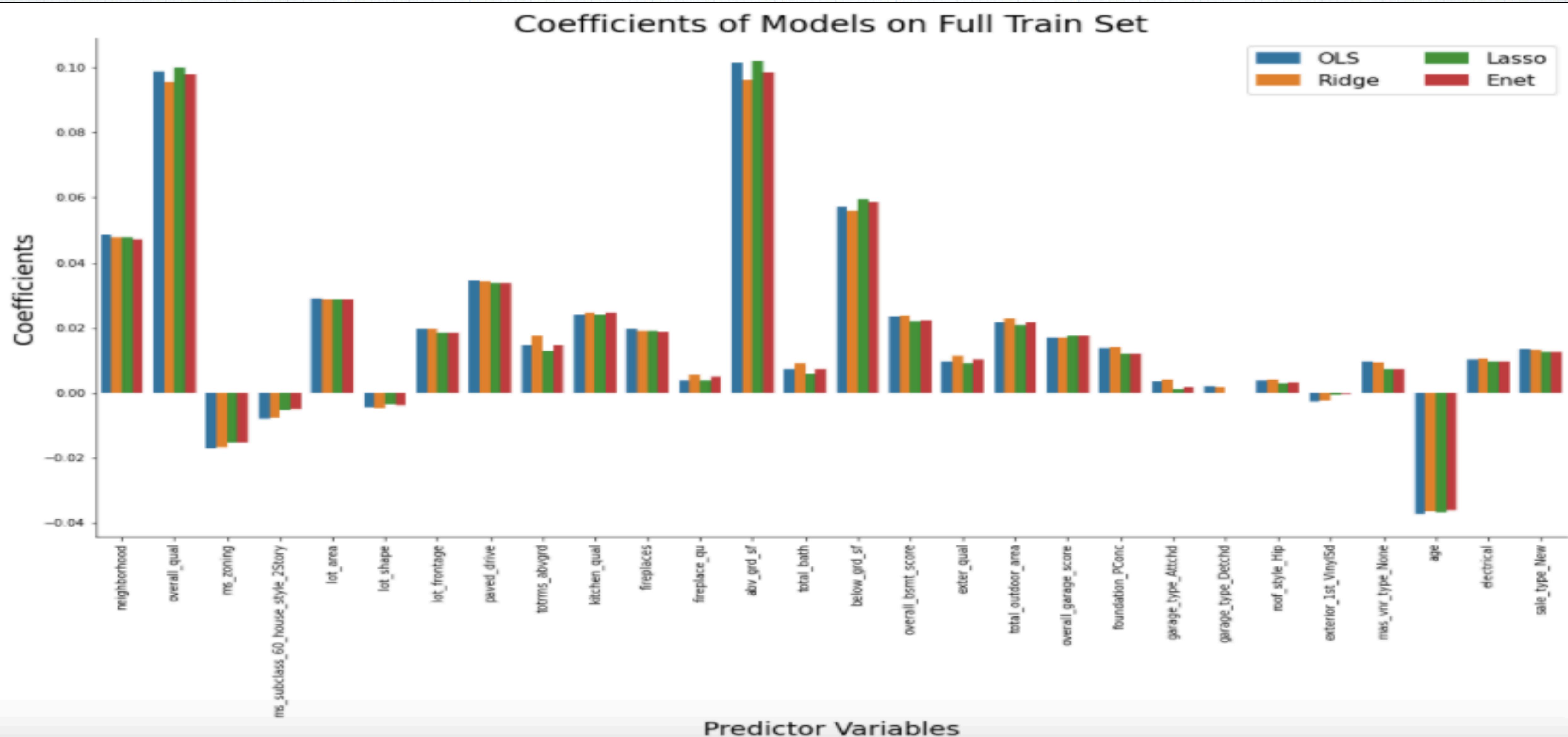
# Production Model Features

```
X_test_final = test_raw_df[['neighborhood',
                             'overall_qual',
                             'ms_zoning',
                             'ms_subclass_60_house_style_2Story',
                             'lot_area',
                             'lot_shape',
                             'lot_frontage',
                             'paved_drive',
                             'totrms_abvgrd',
                             'kitchen_qual',
                             'fireplaces',
                             'fireplace_qu',
                             'abv_grd_sf',
                             'total_bath',
                             'below_grd_sf',
                             'overall_bsmt_score',
                             'exter_qual',
                             'total_outdoor_area',
                             'overall_garage_score',
                             'foundation_PConc',
                             'garage_type_Attchd',
                             'garage_type_Detchd',
                             'roof_style_Hip',
                             'exterior_1st_VinylSd',
                             'mas_vnr_type_None',
                             'age',
                             'electrical',
                             'sale_type_New']]
```

# Model Scores

	<b>Train R2 Score</b>	<b>Train Cross-validation Mean Score</b>	<b>Train Cross-val RMSE Score</b>	<b>Test R2 Score</b>	<b>Test RMSE Score</b>	<b>alpha / l1 ratio</b>
OLS	0.9057	0.9014	26351.9597	0.857	20504.0885	
Ridge	0.9057	0.9015	26336.9819	0.8571	20458.6934	27.0496
Lasso	0.9056	0.9017	26351.9593	0.8577	20452.2196	0.0010
Enet	0.9056	0.9017	26339.8375	0.8578	20421.308	0.0113, 0.0925

# Model Evaluation



# Conclusion

- \* Production Model Scores (based on Enet):
  - Train Cross-val Mean R2: 0.9
  - Test RMSE: USD26,705
- \* The envisioned goal to develop a regression-based predictor model guided by the triple-A framework (accurate, applicable, accessible) for Ames housing sale prices was achieved and thresholds established duly complied with.