Saturn's Housing Data: Predicting Sales & Evaluating Assessments:

Elisabeth Jones, M.S.

May 2023

Presentation Goals

01

Predicting Sales Prices

- What are the results of the sales price predictive model?
- What are the most important variables for predicting sales price?

02

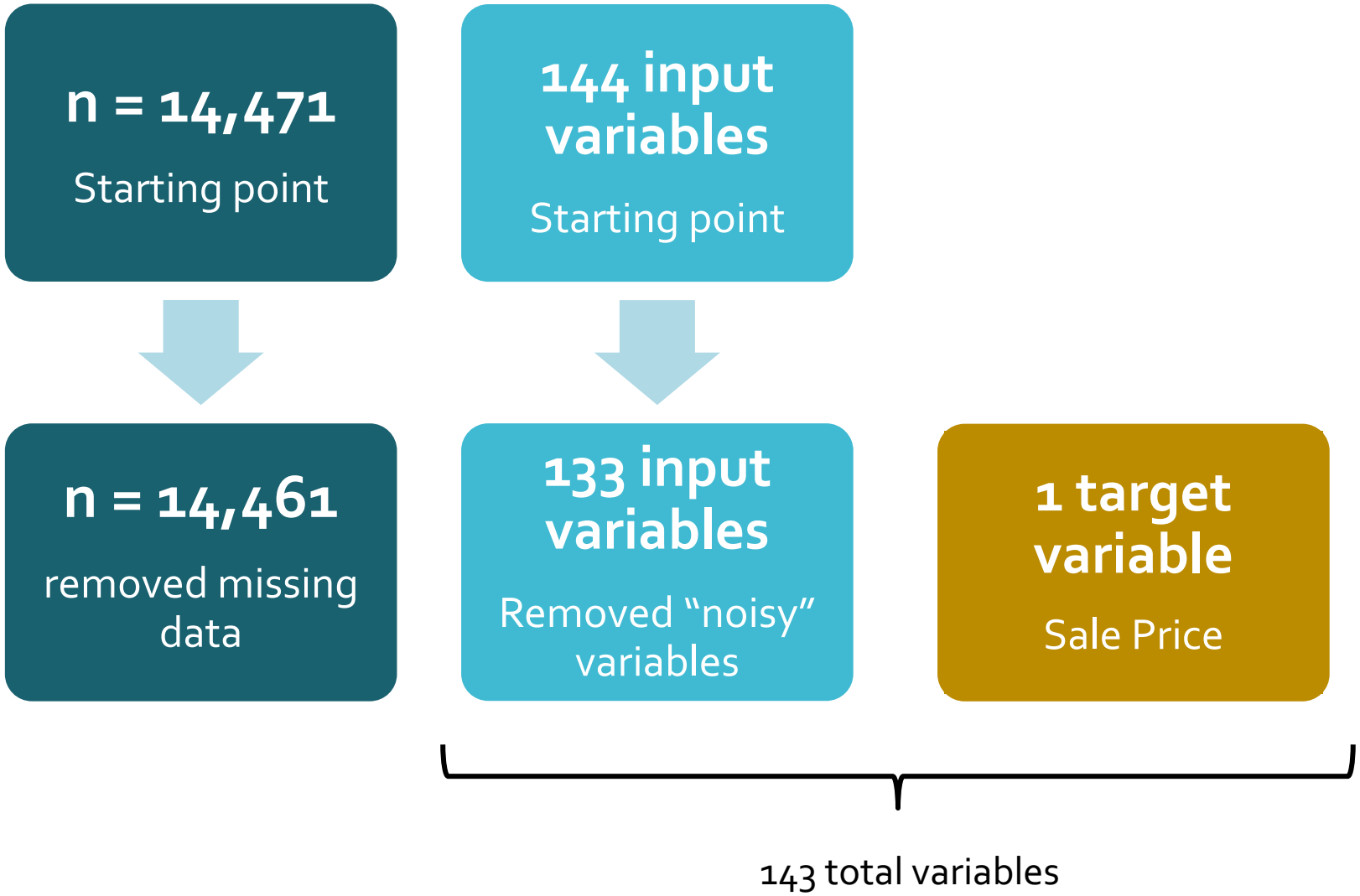
Evaluating Assessments

- Are there equity issues with the current assessment process?
- How can we use predictive modeling to address any issues?

Predicting Sales Prices

Data Overview

Predicting Sales Price: **Dataset Overview**



Predicting Sales Price: **Dataset Overview**

Types of Input Variables

Direct House Metrics

- # of bedrooms
- year built
- Square footage

Building Utilities

- Water supply type
- Sewer type
- Central air

Building Classifications

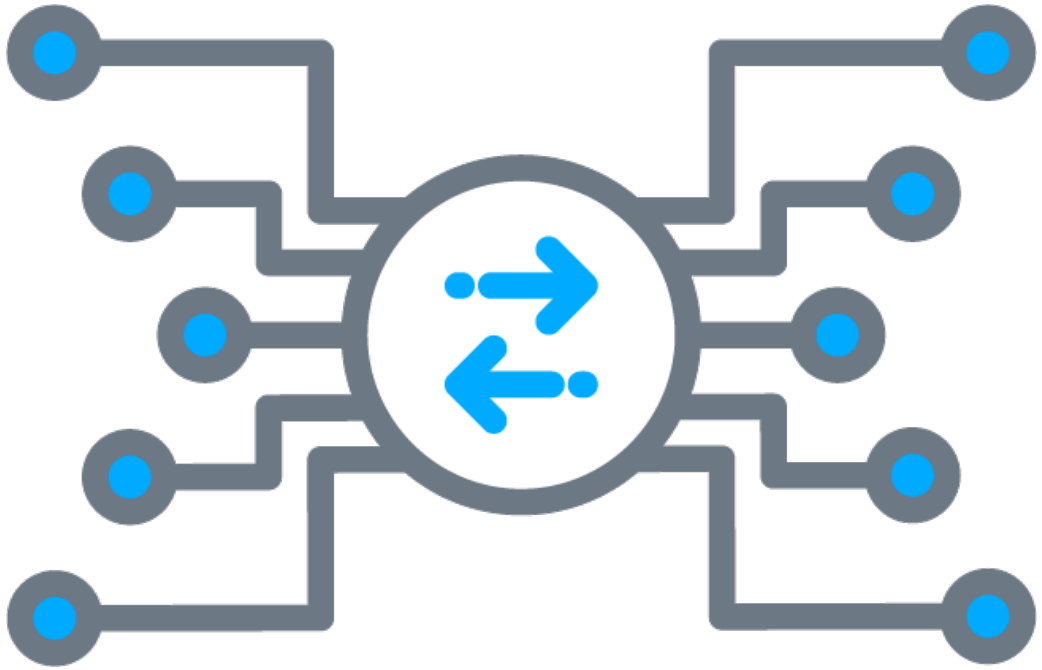
- Building styles
- Grade
- Property classifications

Removed

- Latitude/Longitude
- Dates
- Walk Index
- Print key & misc. others

Individual input factors only give us **pieces** of the larger puzzle

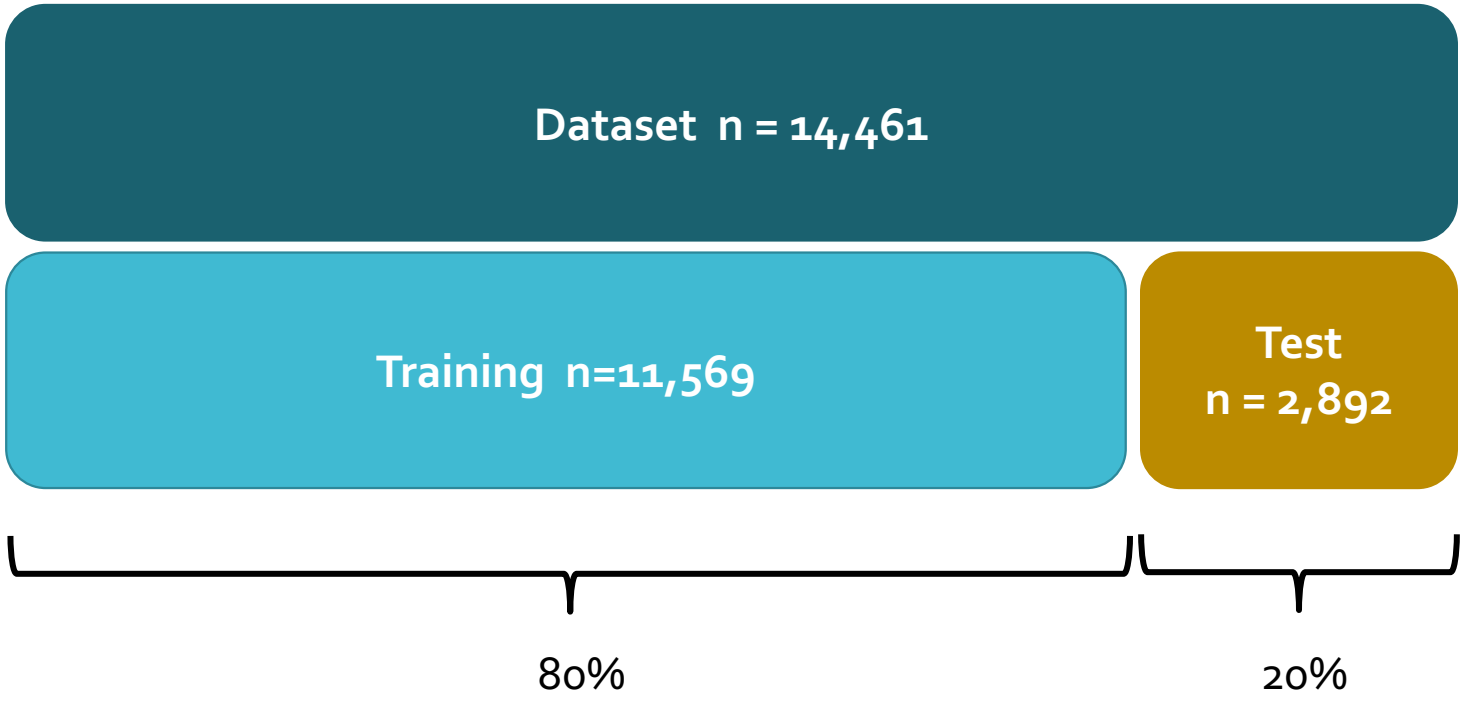
Predictive modeling allows us to understand data by identifying the **hidden relationships and patterns** among multiple variables and using them to **make predictions about future outcomes.**



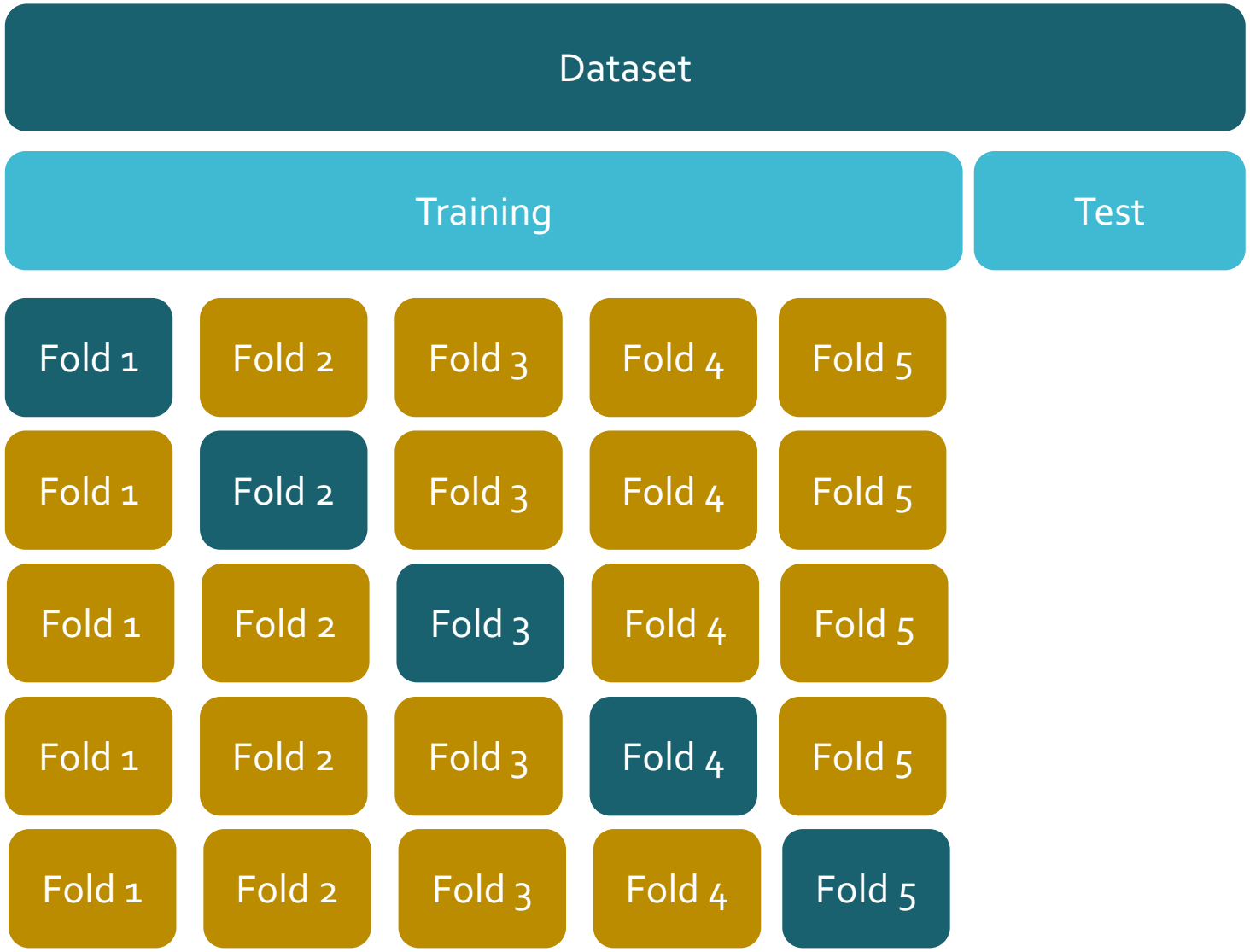
Predicting Sales Prices

Methods

Predicting Sales Price: **Methods**



Predicting Sales Price: **Methods**



Predicting Sales Price: **Methods**

Three models tested:

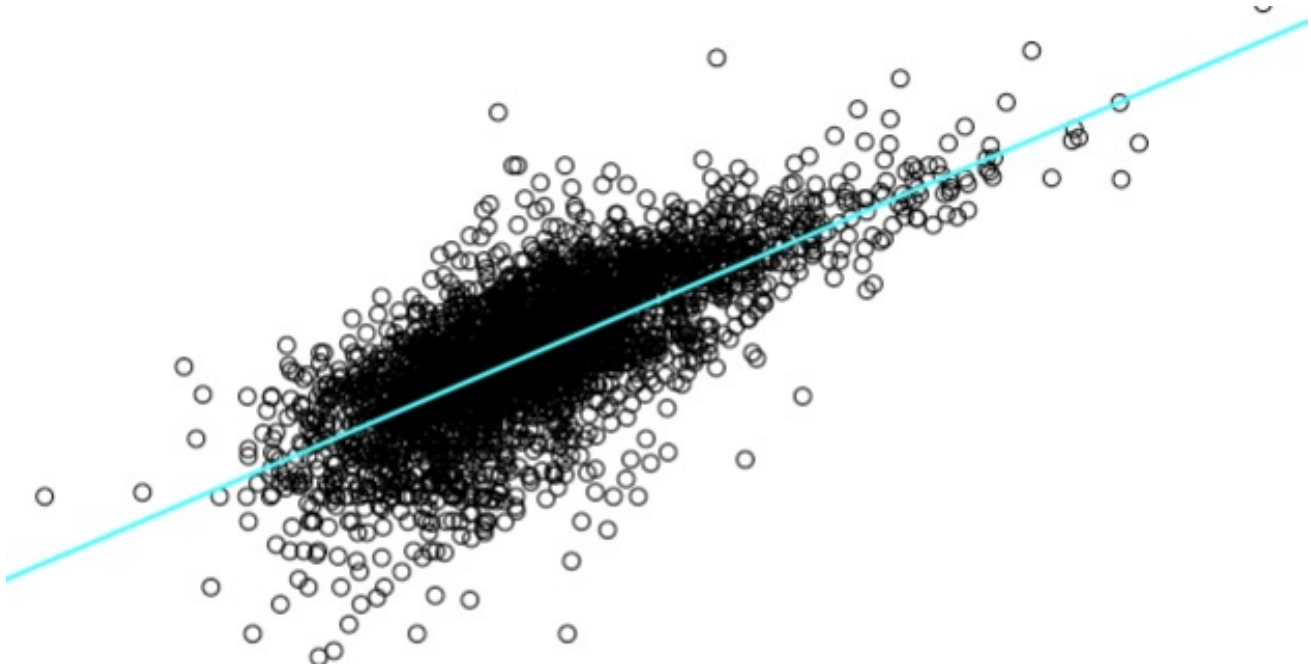
1. Linear Regression
2. Linear Regression +
3. Lasso Regression

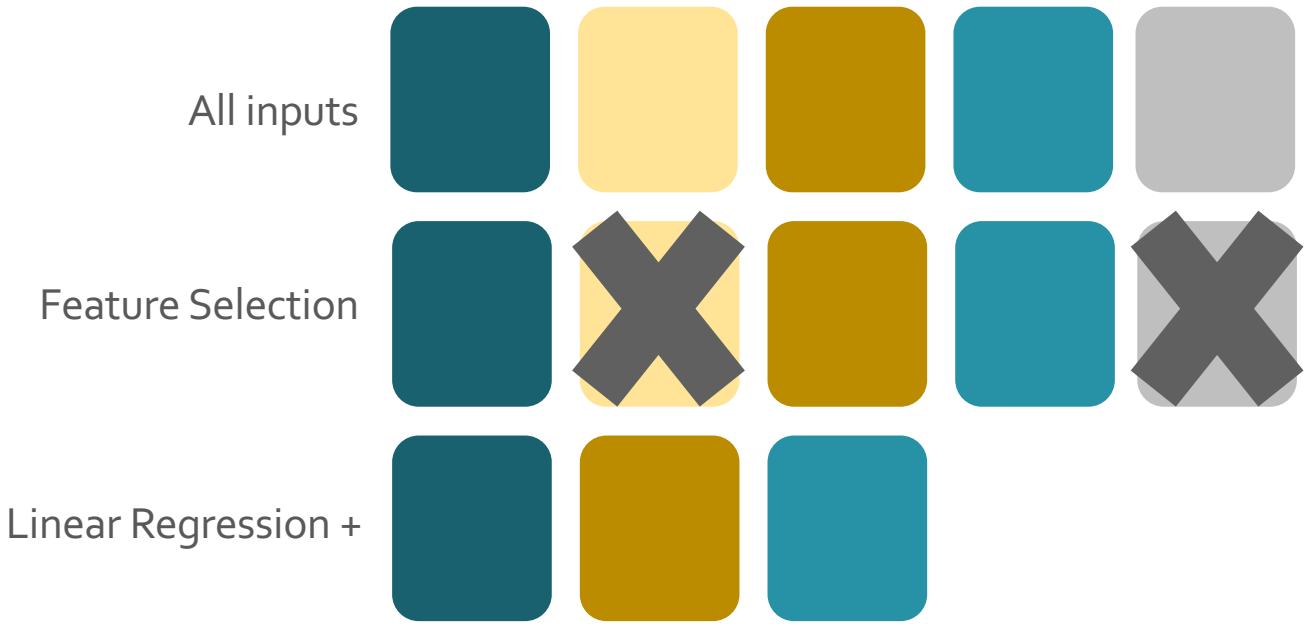
Linear Regression

Linear regression is a tool used to find the **relationship between two things** (i.e, number of kitchens in a house and sale price) .

It does this by drawing a straight line through a group of data points that shows the trend of the relationship.

Once the line is drawn, it can be used to make predictions about what might happen in the future.





Linear Regression +
(aka - feature selection)

By first using a simple linear regression model, you can **figure out which factors are most important in predicting sale price.**

The linear regression+ model eliminates any factors that aren't very helpful in making accurate predictions, which ideally would make the predictions even more precise.

Lasso Regression

Lasso regression is a type of linear regression that can also help you choose which inputs to include in your model.

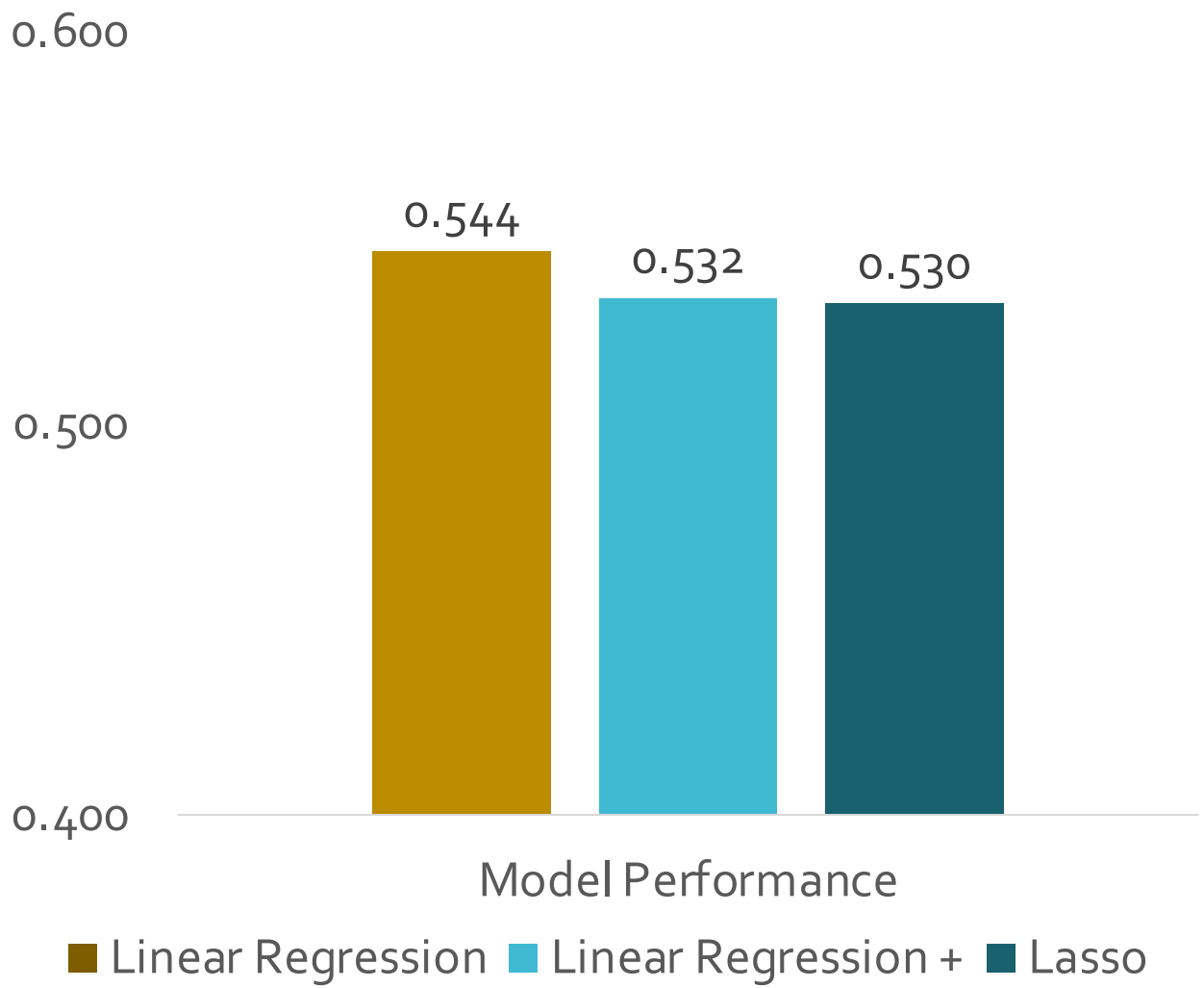
It works by adding a penalty to the regression equation, which can **shrink the influence of less important input variables to zero.**

This makes the model simpler and easier to interpret, while still making accurate predictions.

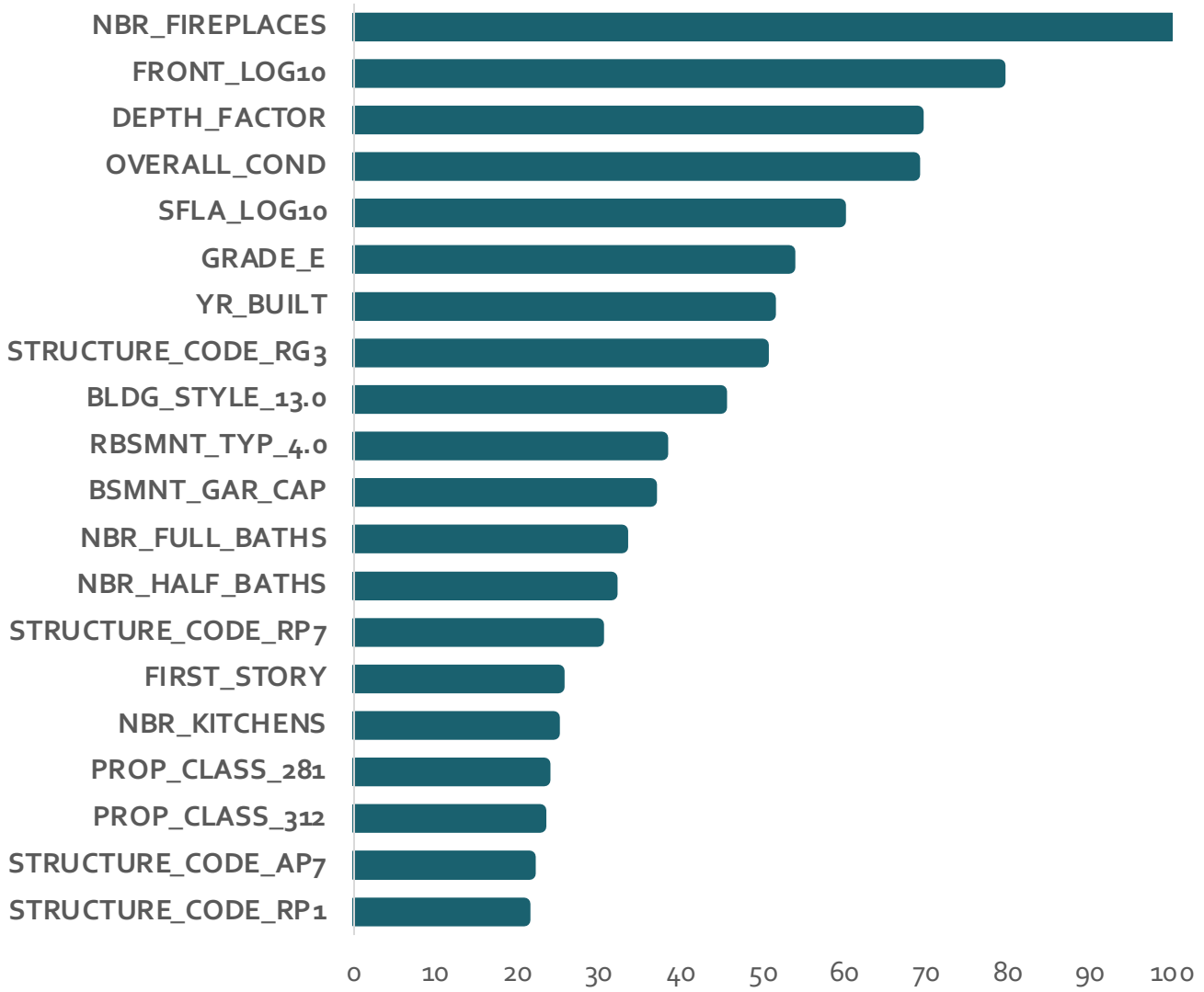
Predicting Sales Prices

Results

Simpler is usually better!



What are the **most important** factors when predicting sale price in Saturn?



Evaluating Assessments

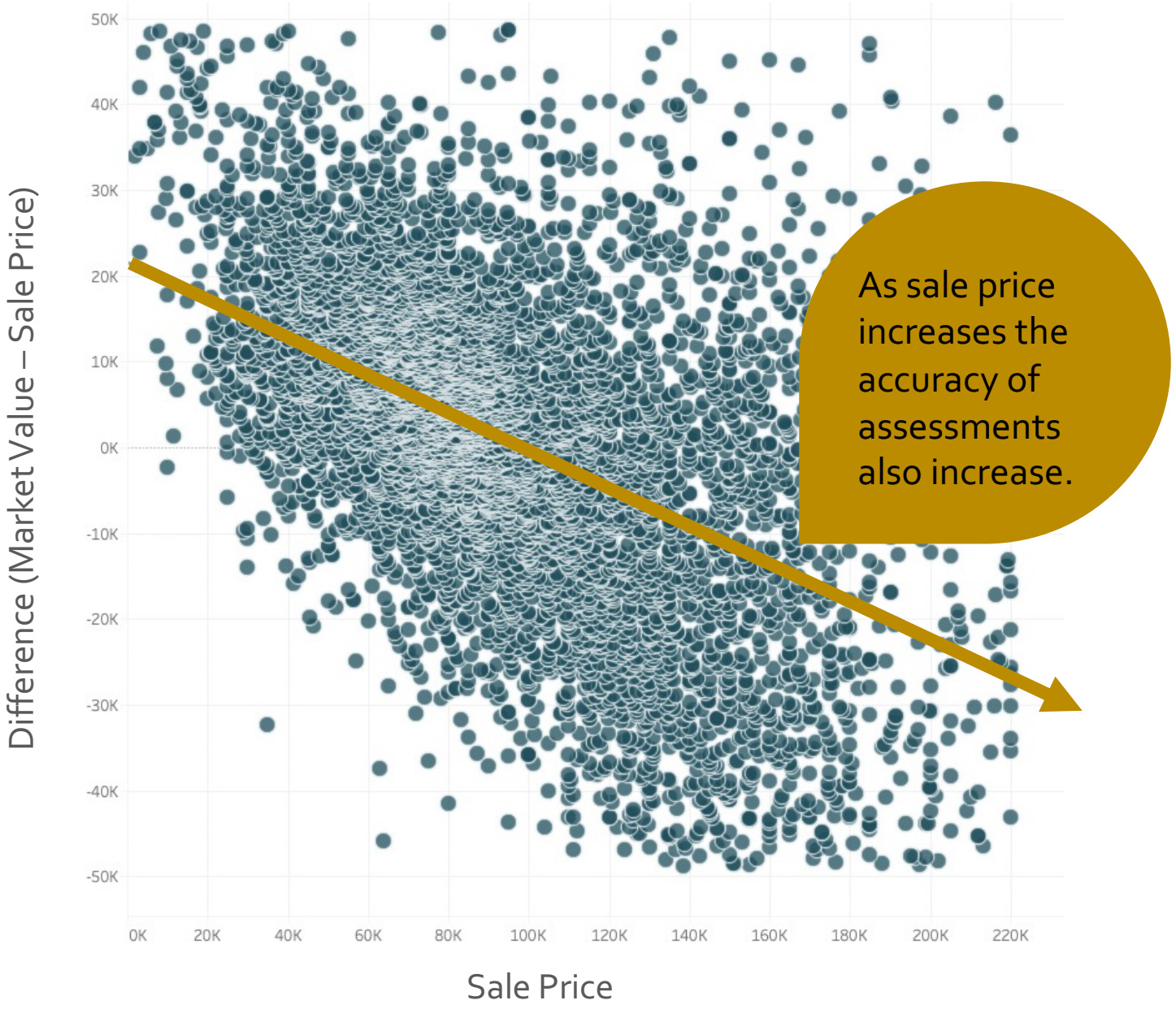
Evaluating Assessments: **Methods**

1. Calculate the difference between market value assessments and actual sale price

2. Build box plots to determine where outliers lie in the data

3. Remove outliers from both sale price and difference

4. Graph sale price vs difference using a scatter plot



Evaluating Assessments: **Results**

If assessments become more accurate as sale prices increase, properties in more affluent neighborhoods may be assessed more accurately than properties in lower-income neighborhoods.

Accurate property assessments are crucial for funding public services and accurate property tax assessments.

This can lead to disproportionate burdens for lower-income and historically marginalized homeowners, who may already struggle to afford their housing costs.

Next Steps

Predictive Modeling

- Gain further insight on removed variables by consulting with a Saturn housing expert.
- Continue to tune existing models and try new models to optimize sale price predictions.

Evaluation Assessment

- Utilize predictive modeling to more accurately assess market value in order to increase equity within Saturn's housing market.

Questions?