SEATTLE METRO

HOME PRICE PREDICTIVE MODEL

PREPARED FOR SEATTLE REALCO

MARCH 15, 2021

Flatiron Data Science Bootcamp Project

insight CONSULTING

TODAY'S MODEL REVIEW

Outline

- Business Objective
- Data and Methods
- Model Progression
- Model Fit Plot
- Model Prediction Sample Home
- Next Steps

EXPECTED PRICE OF A HOME?

Business Objective

Business Objective: Seattle Realco wants to empower agents with a tool to ballpark expected home prices when working with buyer/seller clients

Requirements

1. Find best initial prediction model for Seattle area home prices
2. Model performance significantly better than using "averages"
3. Predictor inputs into model commonly available for new home buyers/sellers
4. Explain relative influence that each predictor has on overall model predictions

DATA AND METHODS

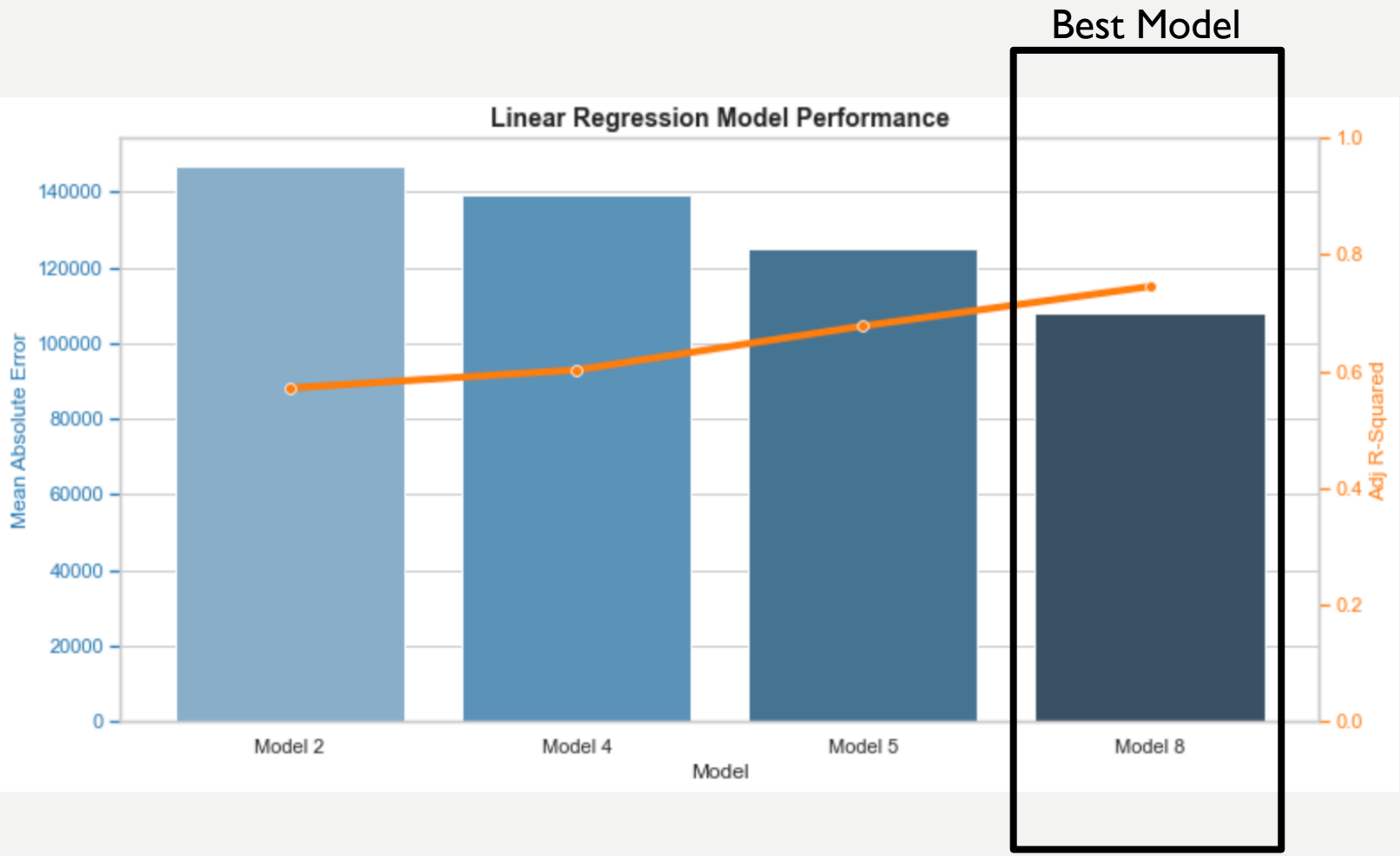
Analysis Background

- Data Source
  - King County homes sales (May 2014 - May 2015)
  - Target Variable: Sales Price
- Raw Source Features
  - Date Sold, Year Built, Floors, Bedrooms, Bathrooms, Condition, Grade
  - Living sq. ft., Basement sq. ft, Lot sq. ft., Closest 15 neighbors average living and lot sq. ft.
  - Zip code, Latitude-Longitude, Waterfront, View indicators
- Features Engineered
  - Log transformations, Renovation indicator, Basement indicator, Grade group, Miles from Seattle, Zip price-per-sqft decile, Zip price-per-sqft median
- Model Type
  - Multiple linear regression model with cross-validation

technical details available on request and at github

MODEL PREDICTIONS

Model Summary



Model 8 performed best with 74% of variance explained

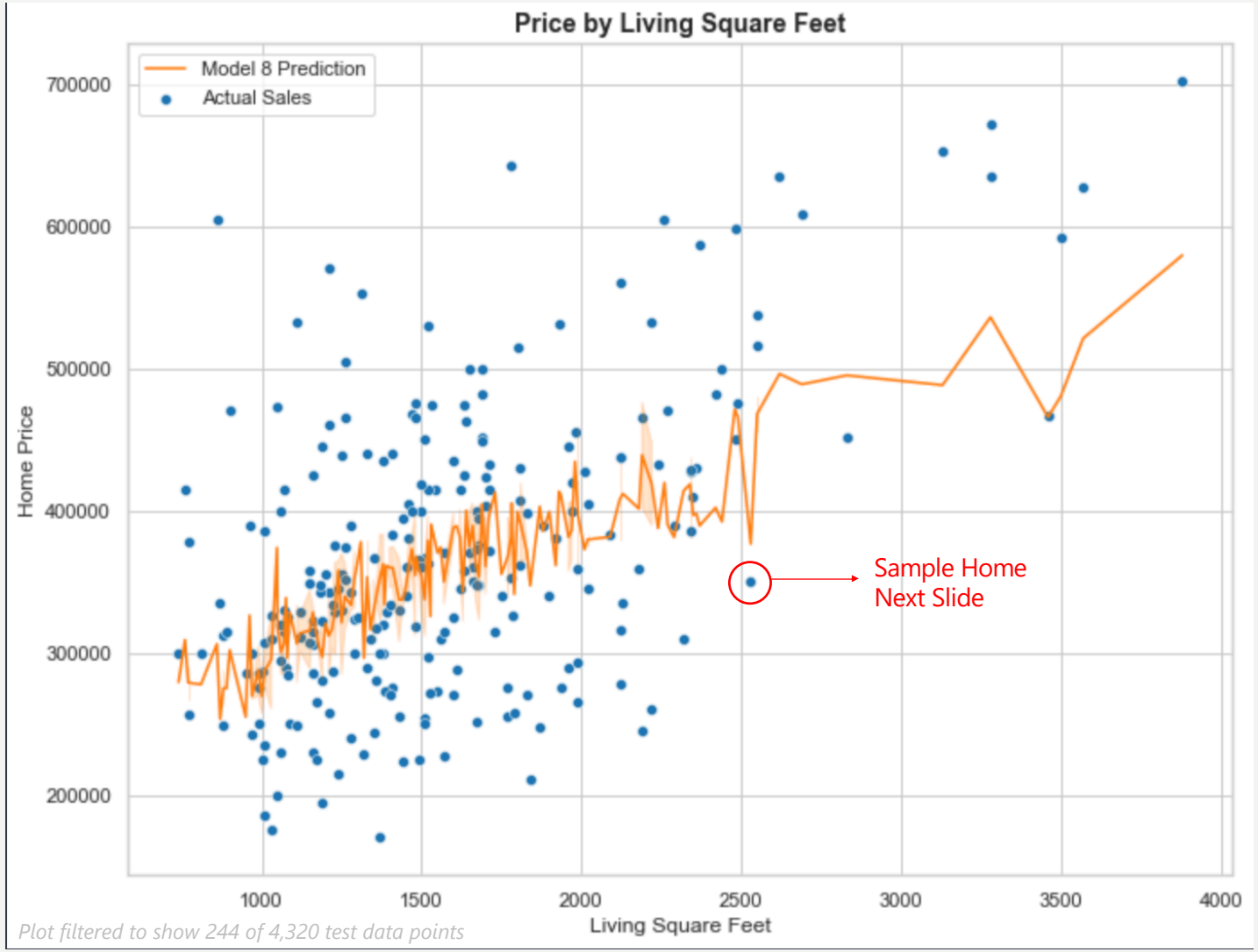
- 0.745 Adjusted R-squared
- $108K Mean Absolute Error
- $182K Root Mean Sq. Error
- 0.00 P-Value

Variables in Model 8

- Target: Price
- Predictors (10):
  - Sq. Ft. Living Space
  - Sq. Ft. Per Bedroom
  - View
  - Grade
  - Condition
  - Waterfront
  - Basement Indicator
  - Renovation Indicator
  - Miles From Seattle
  - Zip code Median Price-Per-Sq. Ft.

MODEL FIT AGAINST LIVING SQ. FT.

Example Fit



- **Model 8 includes 10 predictor variables**
- **Plot shows predictions against test data**
  - X = Sq. Ft. Living Space
  - Y = Home Price
- **…filtered by these 7 predictors (244 sales)**
  - View = 0 (none)
  - Grade = 7 (average)
  - Condition = 3 (average)
  - Waterfront = 0 (none)
  - Basement Indicator = 0 (none)
  - Renovation Indicator = 0 (none)
  - Zip code Median Price-Per-Sq. Ft. between $200-$300
- **…and allowing any value on these final two variables**
  - Miles From Seattle
  - Sq. Ft. Per Bedroom

LET'S REVIEW A REAL TEST CASE

Prediction Example

Model Feature	Rainy Paradise	Feature Contribution
Model Constant	N/A	23%
Living Sq. Ft.	2,530	26%
Sq. Ft. Per Bedroom	506	5%
View	0 (None)	0%
Grade	7 (Average)	9%
Renovation	0 (No)	0%
Waterfront	0 (No)	0%
Basement	0 (No)	0%
Condition	3 (Average)	2%
Miles From Seattle	6.95	0.2%
Zip Code Price Per Sq. Ft.	$201.82	36%
Predicted Home Price		**$375,586**
Actual Sales Price		$349,900
Model Variance		$25,686

Sample Test Home

- Home Name: "Rainy Paradise"
- Home ID: 2780910100
- Zip Code: 98038
- Year Built: 2004
  Date Sold: 12/18/2014
- Sales Price: $349,900

Model Sensitivity

- If Waterfront +$189K
- If Renovation +$54K
- If Basement +$9K
- If Grade 10 +$240K

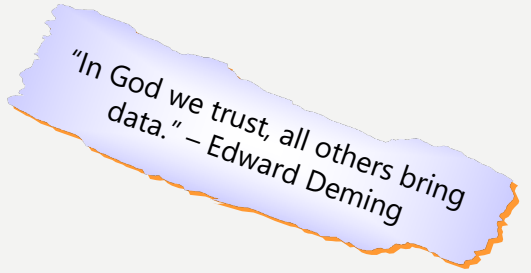
Due to nature of model, multiple features changes would not be directly additive

NEXT STEPS

Next Steps

- Suggest half-day technical workshop deep-dive
- Begin socializing to broader team
- Capture requirements for sales tool integration
- Continue iterating model for improvements

- Questions?



Insight Consulting Group is a boutique consulting firm specializing in analytics, business intelligence, strategy, data visualization, and data warehousing for Fortune 500 client accounts.

Visit www.insightconsultinggrp.com for more information.

Please contact us at:

Chuck Utterback, Principal 770.298.3169
chuck.utterback@insightconsultinggrp.com

https://github.com/cutterback
https://www.linkedin.com/in/chuckutterback/

THANKS