

# AMES HOUSING DATA & NEXRES' NEXT STEPS



DSI-SG-33

Project 2: Group 1

Maryam, Priscilla, Rifqi, Shawn, Ting Wei

## Background: Ames, Iowa



### People

2,484 people  
per square mile

*41% are  
families*



### Income

26.60%  
poverty rate

*\$67,026  
average*



### Climate

humid  
continental

▲  $74.0^{\circ}\text{F}$  ( $23.3^{\circ}\text{C}$ )  
▼  $20.4^{\circ}\text{F}$  ( $-6.44^{\circ}\text{C}$ )

# Problem Statement

Existing Portal      Increased Competition      Develop Model      Deploy API      Maintain Status



But it lacks  
a unique feature!

*Differentiation  
needed!*

Forecast prices  
& features  
**80% accuracy**

**Empower agents**  
- negotiate for fairer  
valuations  
- advise on property  
optimisation for sale

**Provide top value**  
- Cement status  
as top RE  
agency



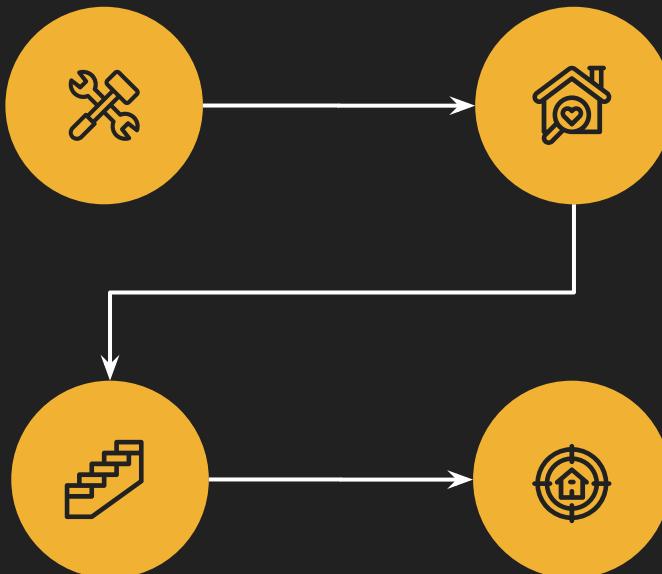
# METHODOLOGY

---

# Building a Predictive Model

## Data Cleaning

Dealing with "missing" values



## Modelling

Create models and evaluate performance

## Analysis

A closer look at features that relate to higher selling prices

## Insights

Interpretation and recommendations

# DATA CLEANING

---





# Process



## Imputation of null values

Median, Mode, "None", "0"



## Dropping of features

Identifiers



## Feature Engineering

Creation of new features  
using existing features

# Dataset

Obtained from the Ames Assessor's Office.

Record of sales for residential properties sold in Ames, Iowa from 2006 to 2010.

After cleaning and feature engineering:

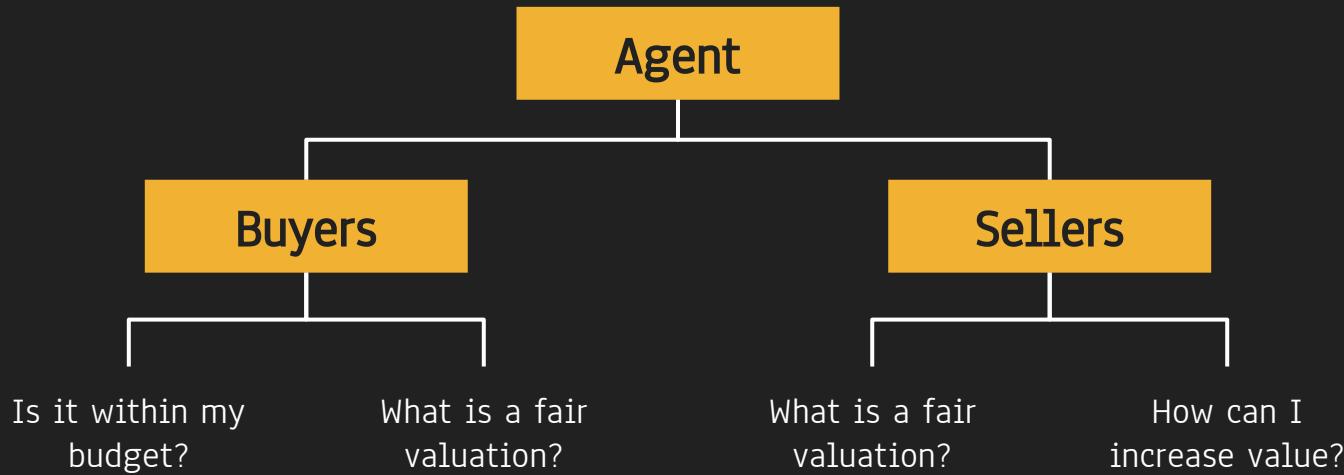
Valid House Observations	Original Features	New Features	Target
2049	77	2	Sale Price

# EXPLORATORY DATA ANALYSIS

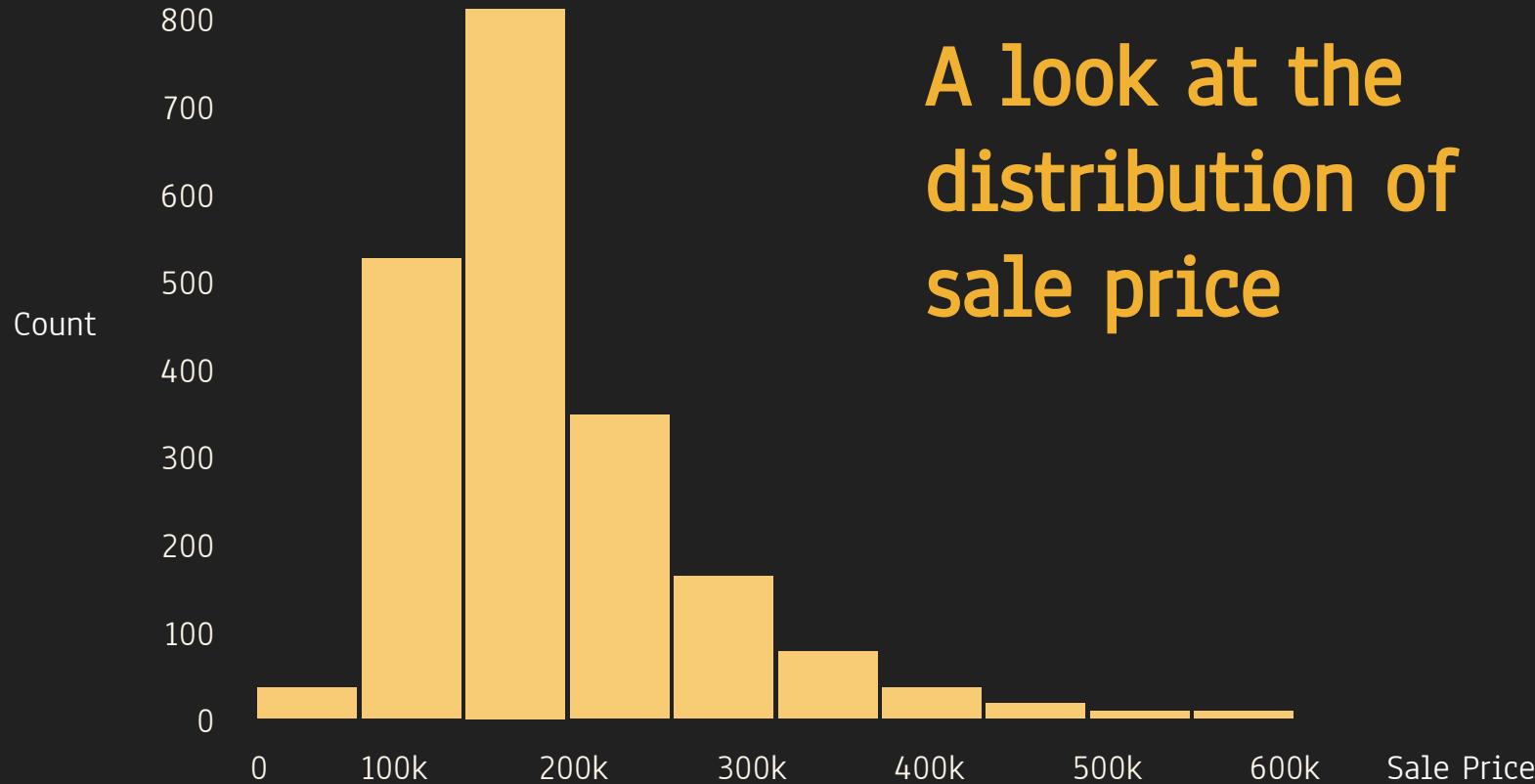
---

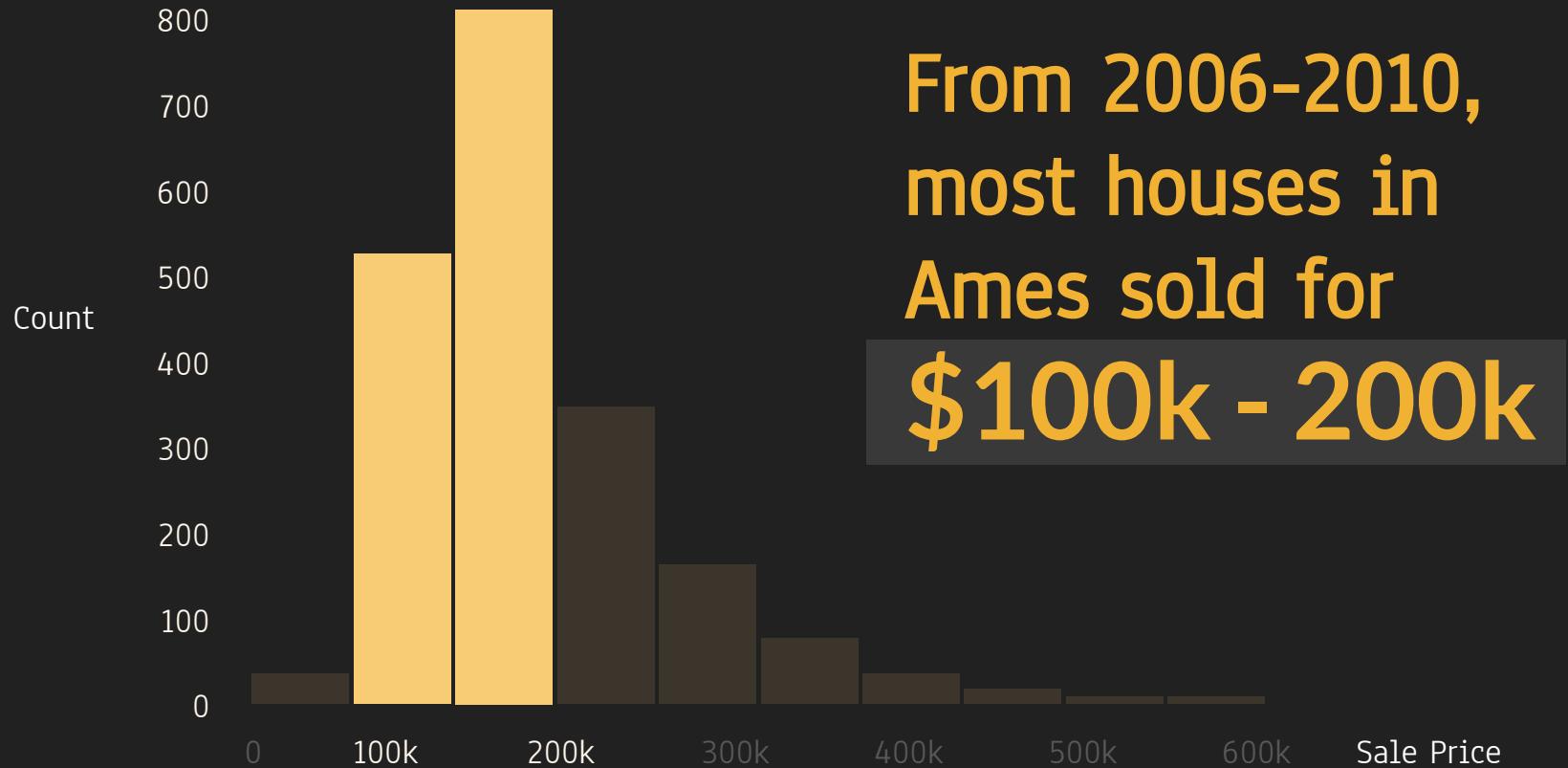


# Guiding Principle



A look at the  
distribution of  
sale price





From 2006-2010,  
most houses in  
Ames sold for  
**\$100k - 200k**

# Higher Selling Price - 3 Key Features



## Neighborhood

Houses in certain locations command the higher average prices



## Overall Quality

Houses with better overall material and finish sell for higher



## Total Built Area

The amount of “useful” area above ground is a familiar valuation metric

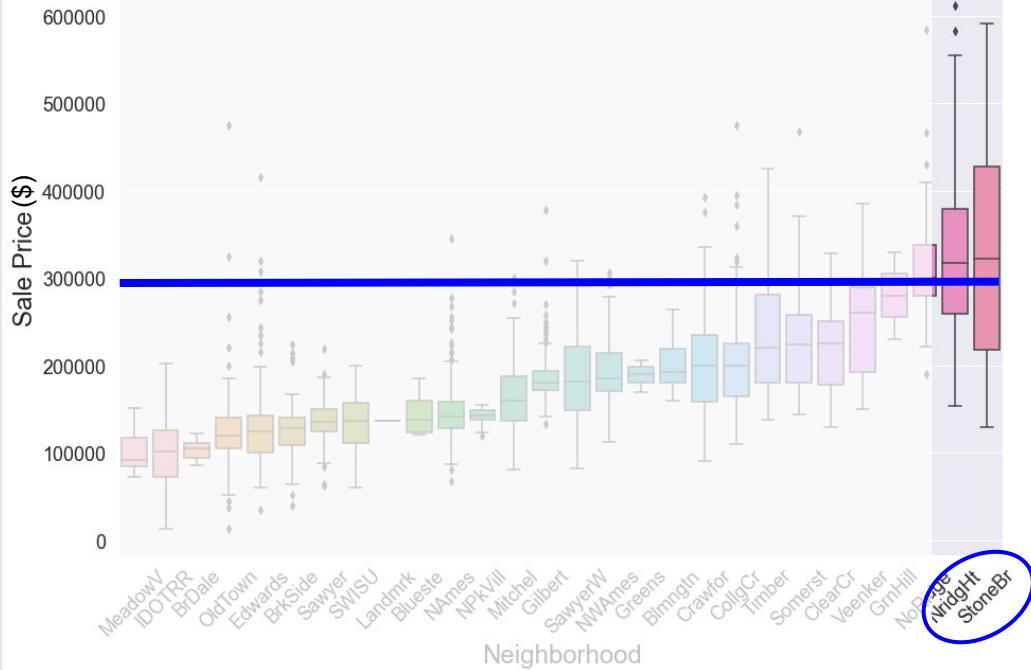
# Neighborhood

**Stone Brooke and Northridge Heights homes have a higher median sale-price of above \$300K**

📍 Proximity to Iowa state university and Gilbert school district <sup>1</sup>

📍 Full range of amenities <sup>2</sup>

Boxplot of Sale Price Across Neighborhoods



<sup>1</sup> About Ames, City of Ames

<sup>2</sup> Northridge Heights, Hunziker Development.

## Overall Quality (1-10 Scale)

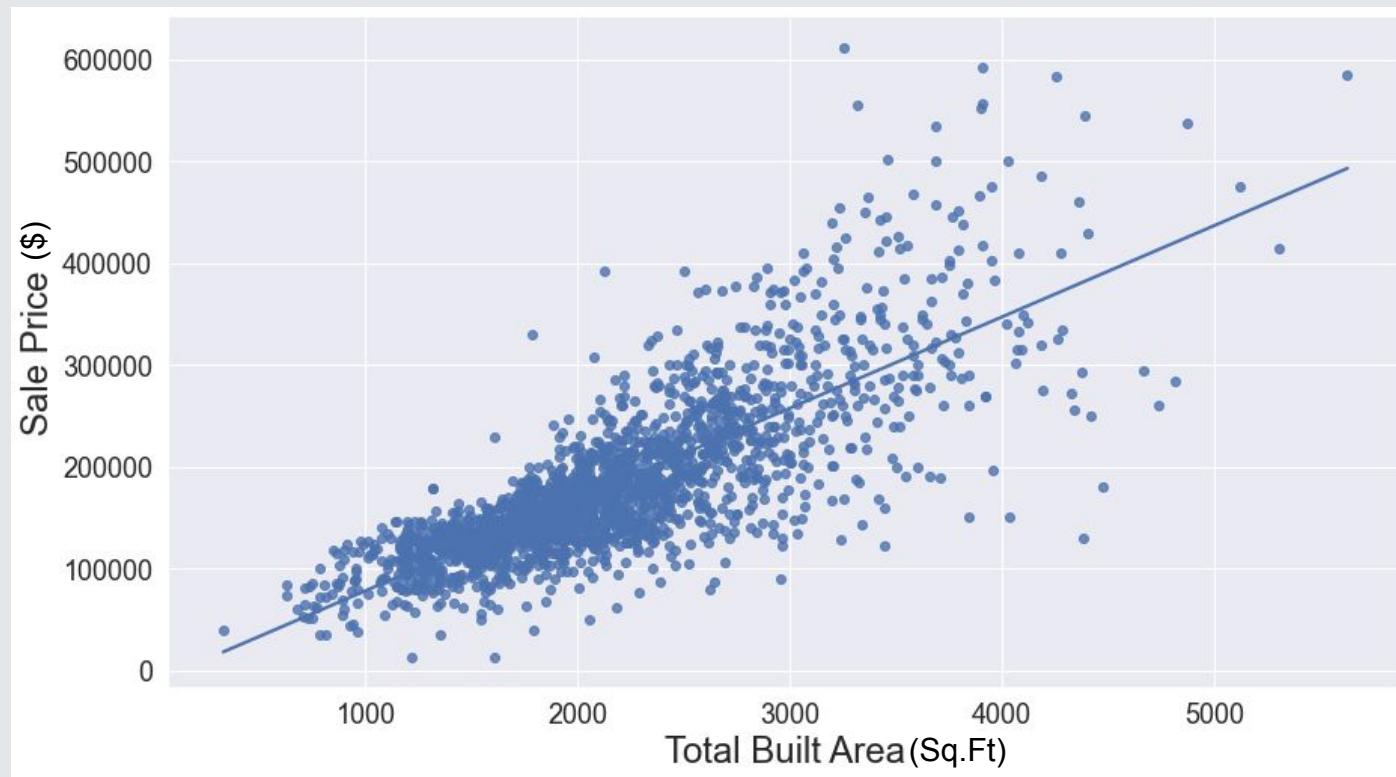
Houses with **better overall** materials and finish **sell for higher**

 An **increment in quality by 1 point** from 7-8 sees a marginal median sale price **increment of 35%**

 On average, the price of a house **rated as a 10** commands more than **2x** that of a 7



# Positive Correlation between Total Built Area and Sale Price

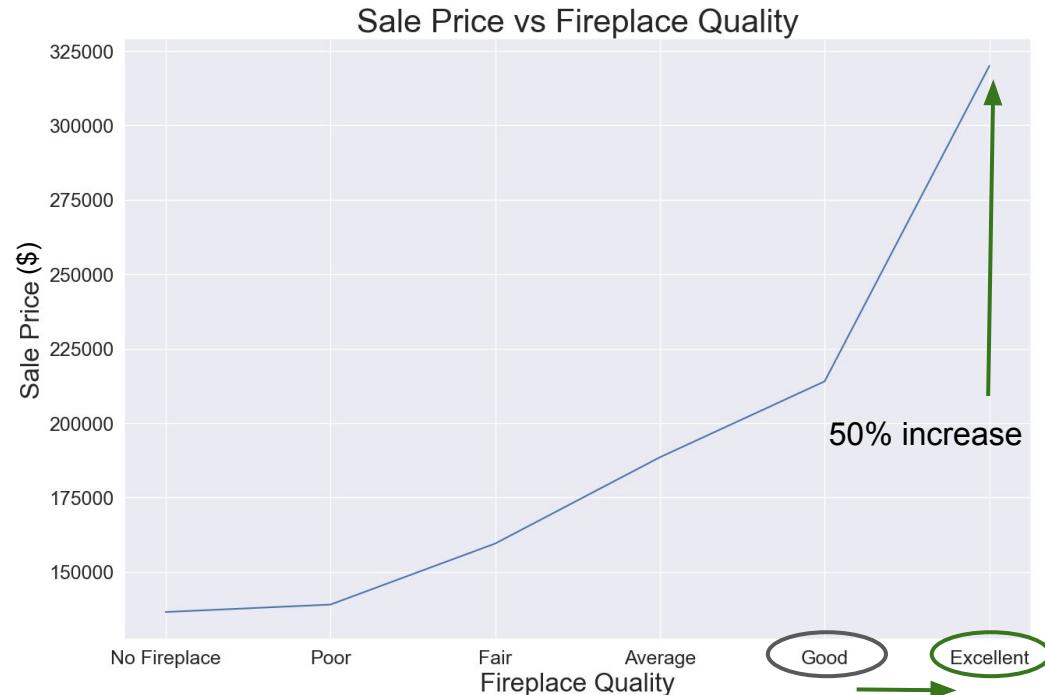


# Fireplace Quality

**Sub-zero (°C) temperatures**  
during winter

 **Better quality fireplaces → Higher prices**

 **Excellent fireplace → 50% higher median sale price!**

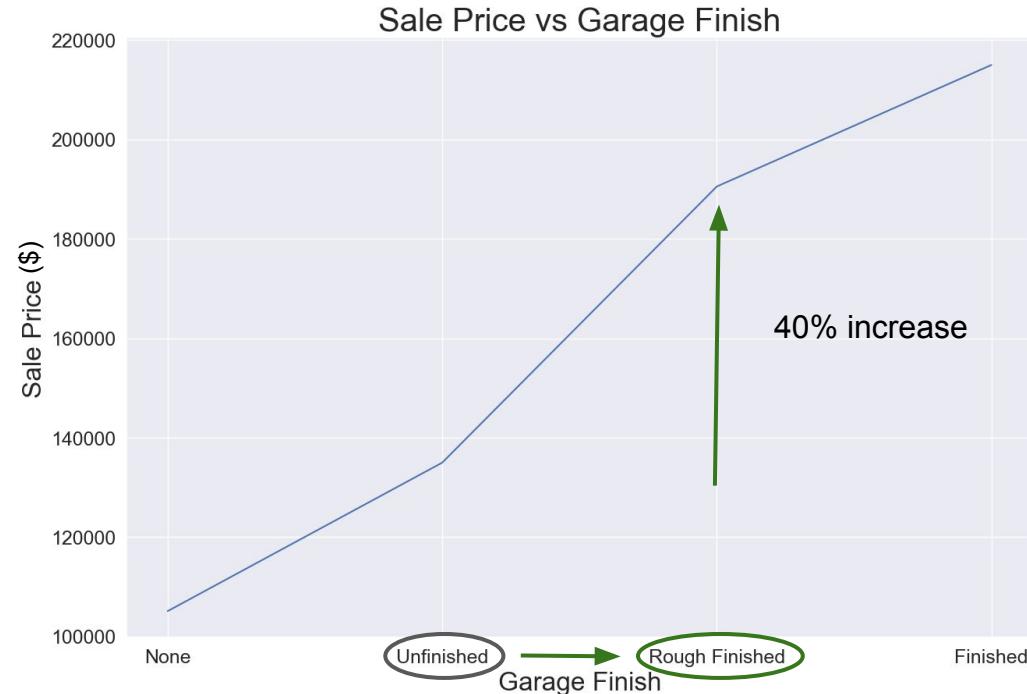


## Garage Finish

Average car ownership is relatively high at 2 per household



**Biggest jump** in median sale price of **40%** between an **unfinished** and **roughly finished** garage.

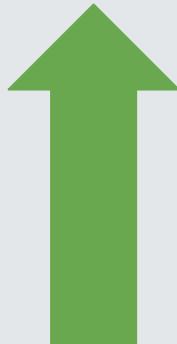


# Machine Learning Models

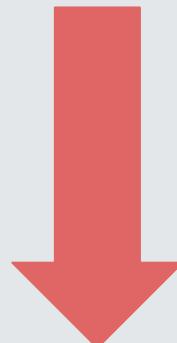


# What is a Regression Model?

Altitude



Temperature



# Relationship between Sale Price and Features



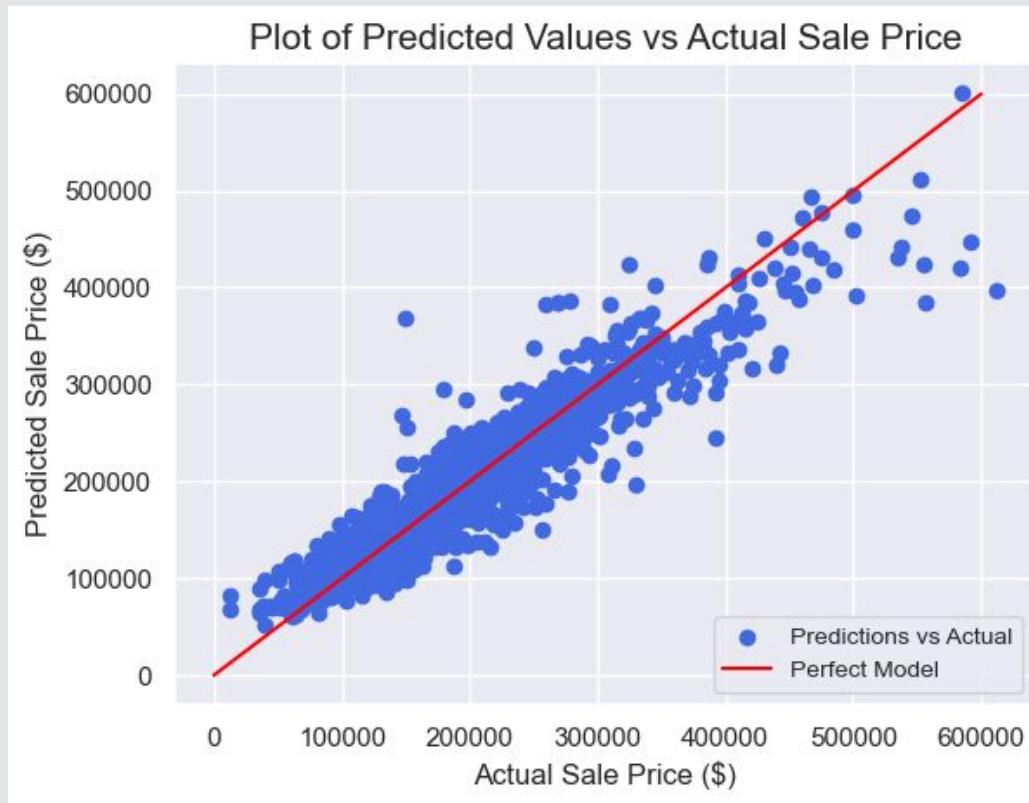
## 10 Features

1. Overall Quality
2. Total Built Area
3. No. of Bathrooms
4. House Style
5. Type of Foundation
6. Exterior Covering
7. Type of Masonry Veneer
8. Fireplace Quality
9. Type of Garage
10. Garage Finish

# Regression Model Performance

	R <sup>2</sup>	RMSLE	% Error	Selection
Linear	83.96%	0.2353	17.71%	
Ridge	84.30%	0.2326	17.49%	
Lasso	<b>84.44%</b>	<b>0.2316</b>	<b>17.42%</b>	✓

# High Accuracy of Model in Predicting Sale Price



# Does it actually work?

## Stone Brooke

Predicted Price  
\$334,523

- Overall Condition - 8
- Foundation - P Concrete
- Fireplace - Good

Actual Price  
\$349,265

4% lower

## NW Ames

Predicted Price  
\$169,320

- Overall Condition - 6
- Foundation - C Block
- Fireplace - Typical

Actual Price  
\$181,000

6% lower

## Edwards

Predicted Price  
\$82,544

- Overall Condition - 4
- Foundation - Brick & Tile
- No Fireplace

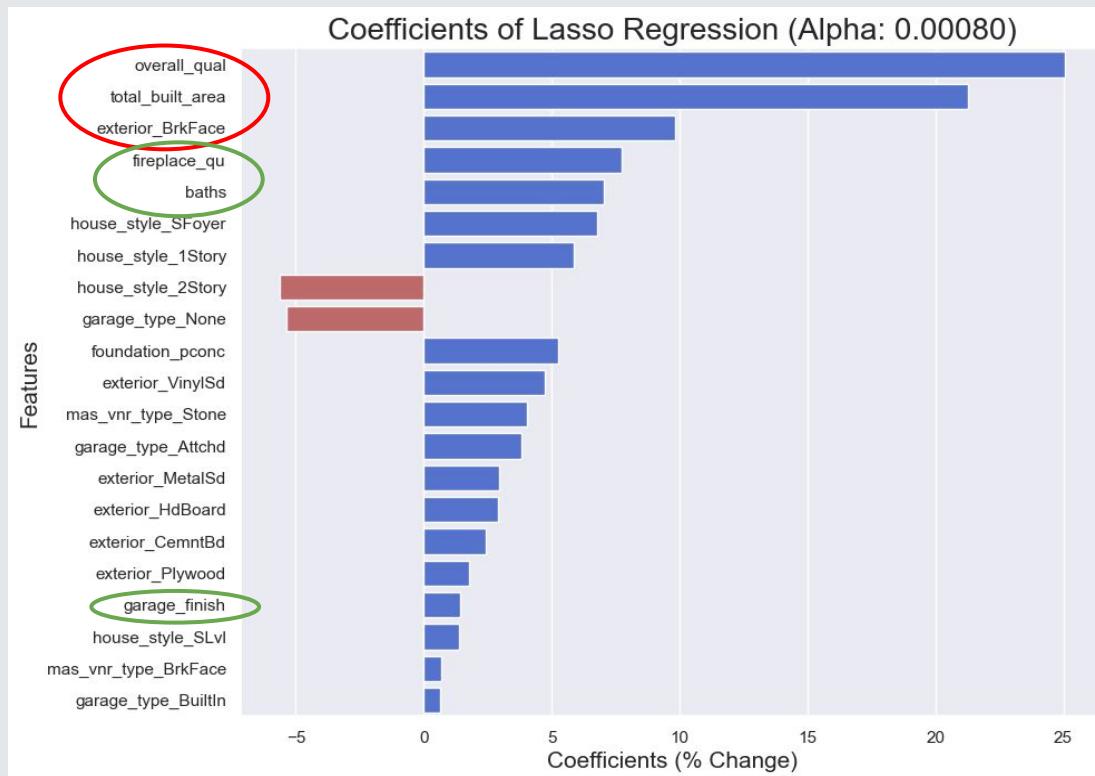
Actual Price  
\$86,000

4% lower



# RECOMMENDATIONS

# % Change in Sale Price per Feature



The top 3 **positive** features

Improve the value of a home

Negotiate for better prices

# FUTURE WORK



# Limitations

## Outdated Dataset

Significant changes to neighborhoods since



## Socio-Economic Factors

Recession periods from  
2007 - 2009



## Lack of Certain Features

Amenities, nearby  
schools

# Future Uses



Build an in-house app (for Agents only)

Advise clients on house features



# Conclusions

---

We managed to:

- Build an effective model
- Understand which and to what extent various features affect sale price
- Advise agents to negotiate better transactions

---

# Thank you!



---

CREDITS: This presentation template was created by [Slidesgo](#), and includes icons by [Flaticon](#) and infographics & images by [Freepik](#)

# Number of Bathrooms

As number of full bathrooms increase

- from 1-2 and from 2-3 bathrooms
- sale price of properties almost double!

