Ames Housing Data

Presented by Jon Reynolds
DSI NYC-Nash

The Data Science Problem Statement

"Can we develop an algorithm to predict the price of residential homes based on immutable features, ie Location, Square Footage, Year Built, etc for a Real Estate Investment Firm with a portfolio of existing properties??

Additionally, can we identify features to recommend that are readily available to ameliorate?"

Exploring the Data

Cleaning the Data

- Handling missing values
- Identifying and dealing with outliers

Determining handling of different feature types

- Numerical (Continuous & Discrete)
- Categorical (Ordinal & Nominal)
  - Creating Dummies
  - Dictionary Assignment for Rankings

Examples of Missing Values

Pool QC	2042
Misc Feature	1986
Alley	1911
Fence	1651
Fireplace Qu	1000
Lot Frontage	330
Garage Qual	114
Garage Finish	114
Garage Cond	114
Garage Yr Blt	114
Garage Type	113
Bsmt Exposure	58
BsmtFin Type 2	56
Bsmt Cond	55
Bsmt Qual	55
BsmtFin Type 1	55
Mas Vnr Type	22
Mas Vnr Area	22
Bsmt Half Bath	2

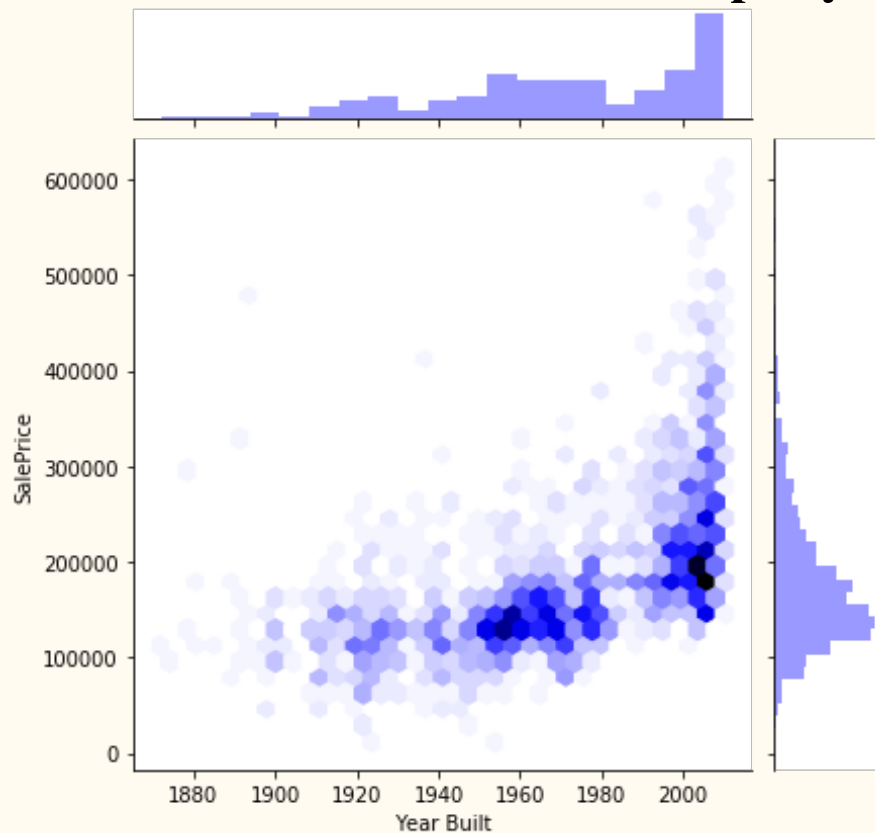# Visualizations

Sale Price of Properties
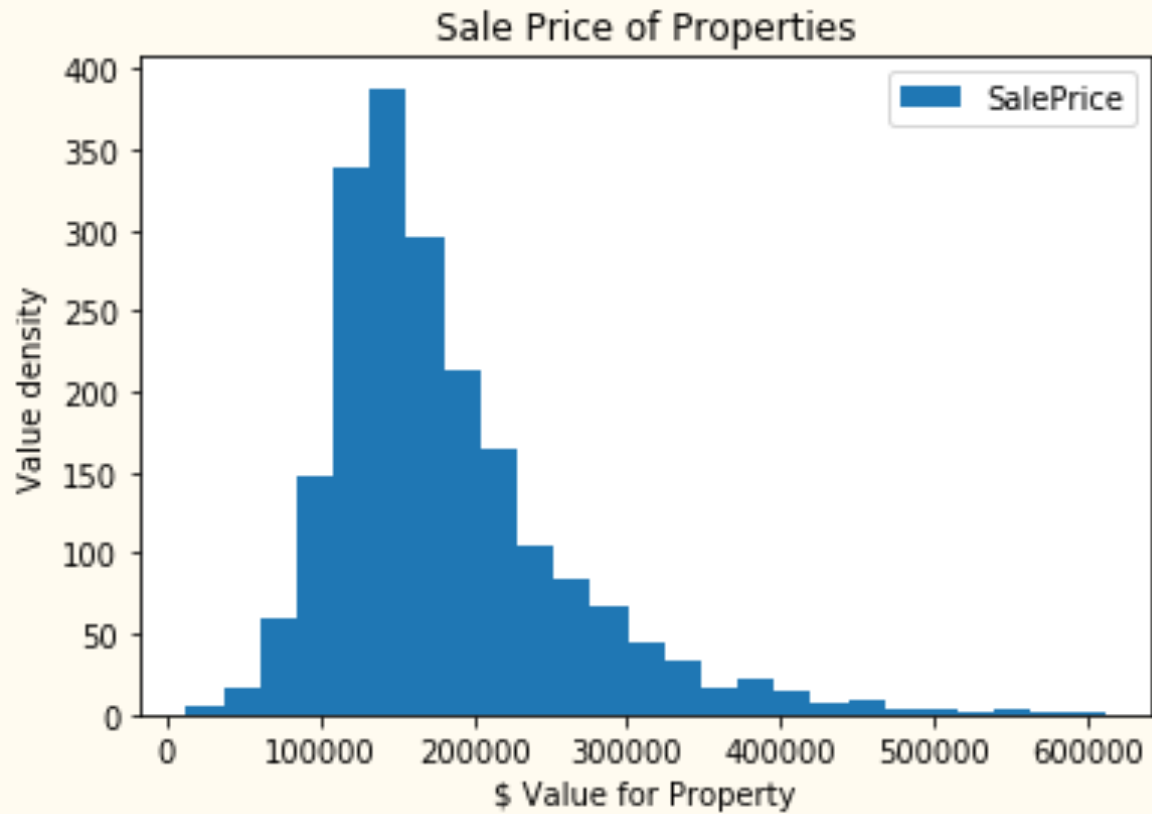
Key Takeaways:

Mean Price vs Median Sale Price

Mean Year Built / Median Year Built

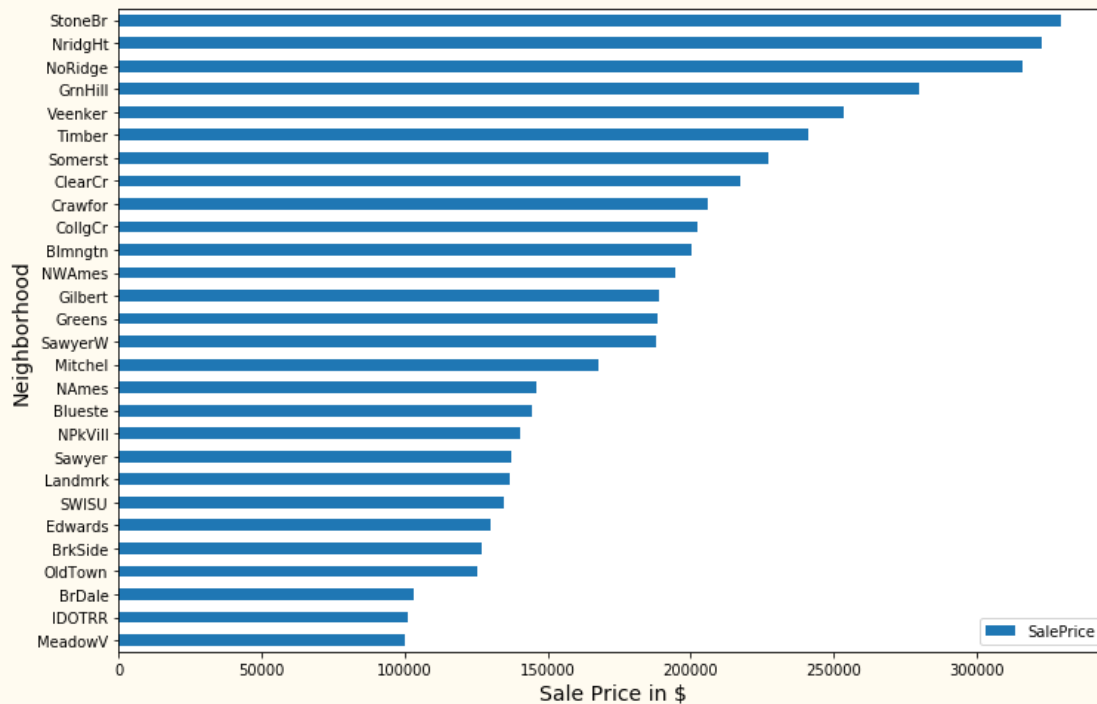# Visualizations

Sale Price vs Year Built of Property

Key Takeaways:

- Year Built and Sale Price are Positively Correlated

- Most of the data is in 1955 - 1980 and 1995 - 2010 for Year Built
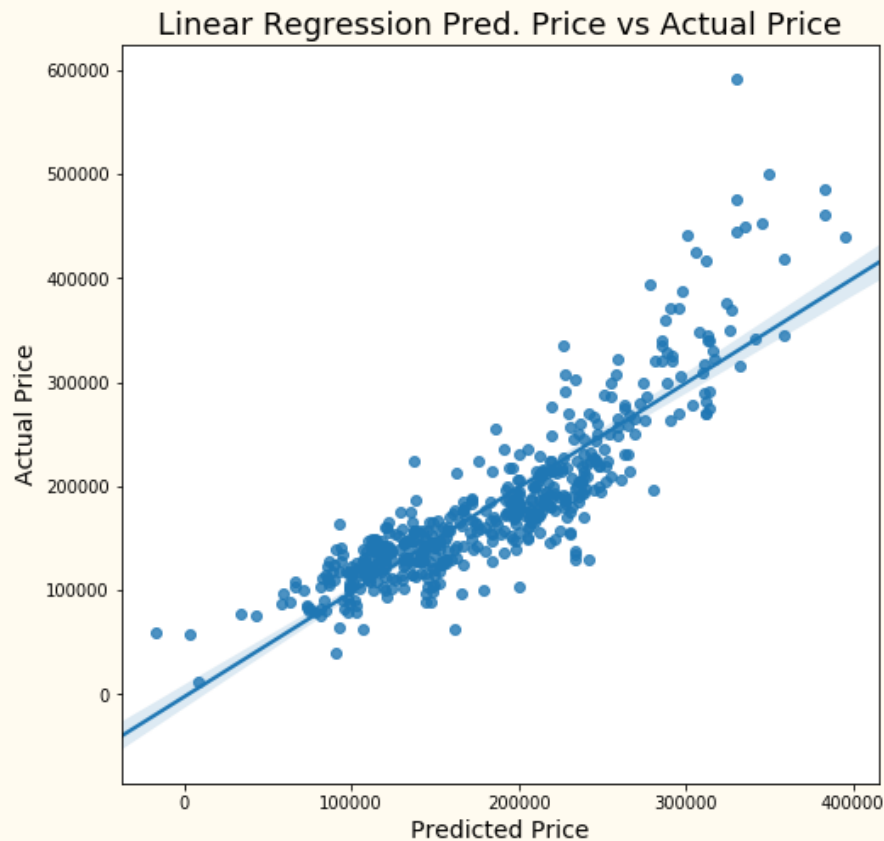
Visualizations

Mean Sale Price by Neighborhood

Key Takeaways:

- Top 3 Neighborhoods compared to Bottom 3 Neighborhoods

- Impact on the Distribution(Mean vs Mode)

Visualizations: Simple Linear Regression Model

Linear Regression Pred. Price vs Actual Price
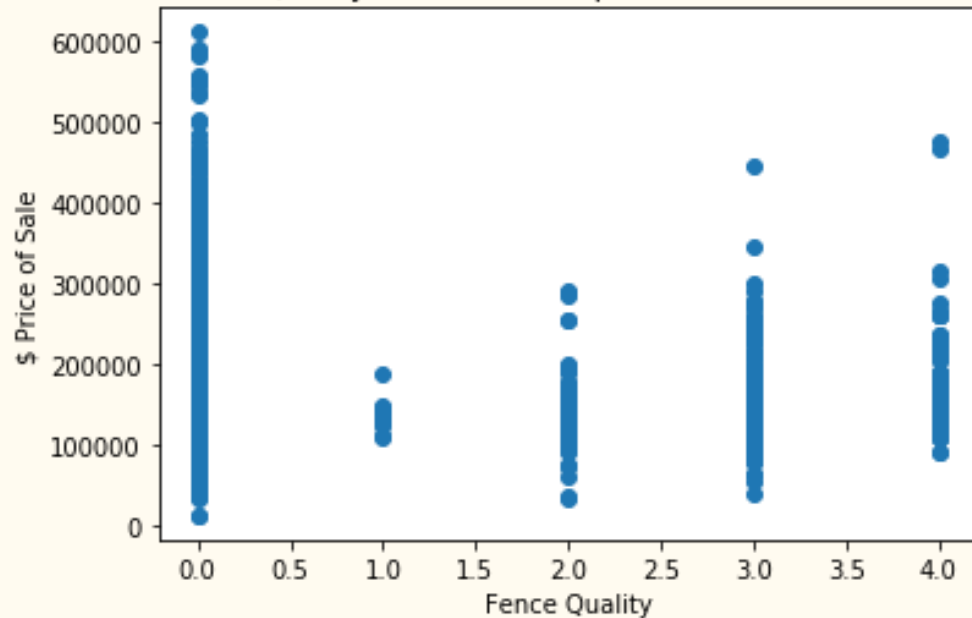
Key Takeaways:

Simple Model Features:
- Overall Quality
- Year Built
- Above Grade Living Area
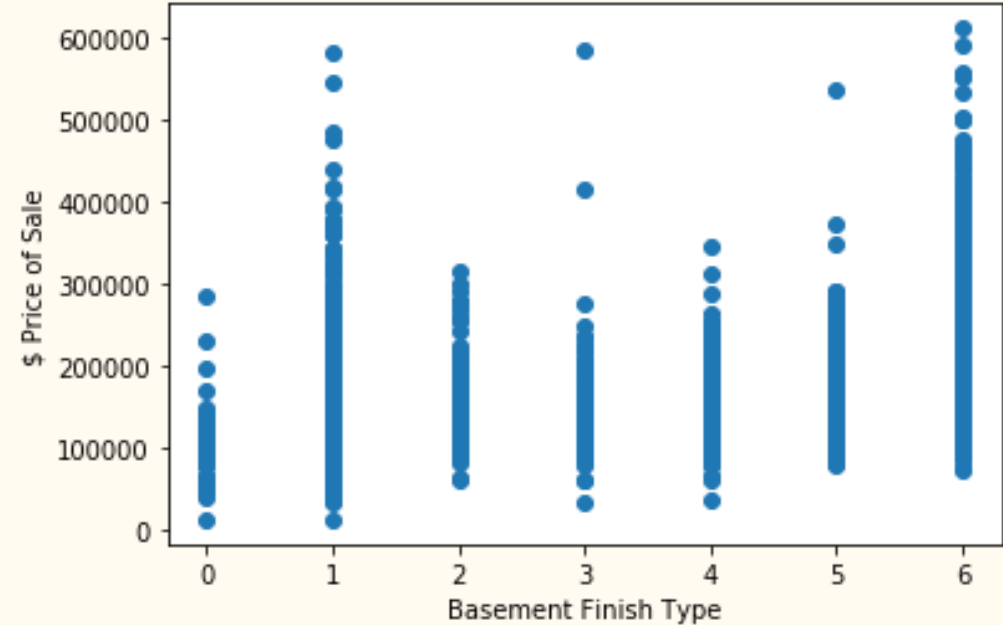- Total Rooms Above Ground

Associated Impacts:
- 25,787.48
- 496.26
- 65.86
- -1,960.09

Visualizations: Improvement Areas

Conclusions and Recommendations

Observable Trends:

   More recently built houses, higher Sale Prices

   Location, Location, Location. Neighborhoods matter!

Actionable Items:

   Fence Removal: it's worth it!

   Basement Finishing: more expensive, but potentially valuable!