

An aerial photograph of a residential development project titled "AMES DEVELOPMENT PROJECT". The image shows a mix of completed houses with blue roofs and green lawns, and some areas under construction or planning. A large, semi-transparent circular interface is overlaid on the image, featuring a compass rose with numbers from 140 to 260 and letters A through F around its perimeter. Several dashed lines and arrows point from these numbers and letters to specific features in the image, such as a playground, a river, and a road. The background shows a valley with mountains in the distance.

# AMES DEVELOPMENT PROJECT

RAFI RAHMAN

# PROBLEM STATEMENT & RESEARCH

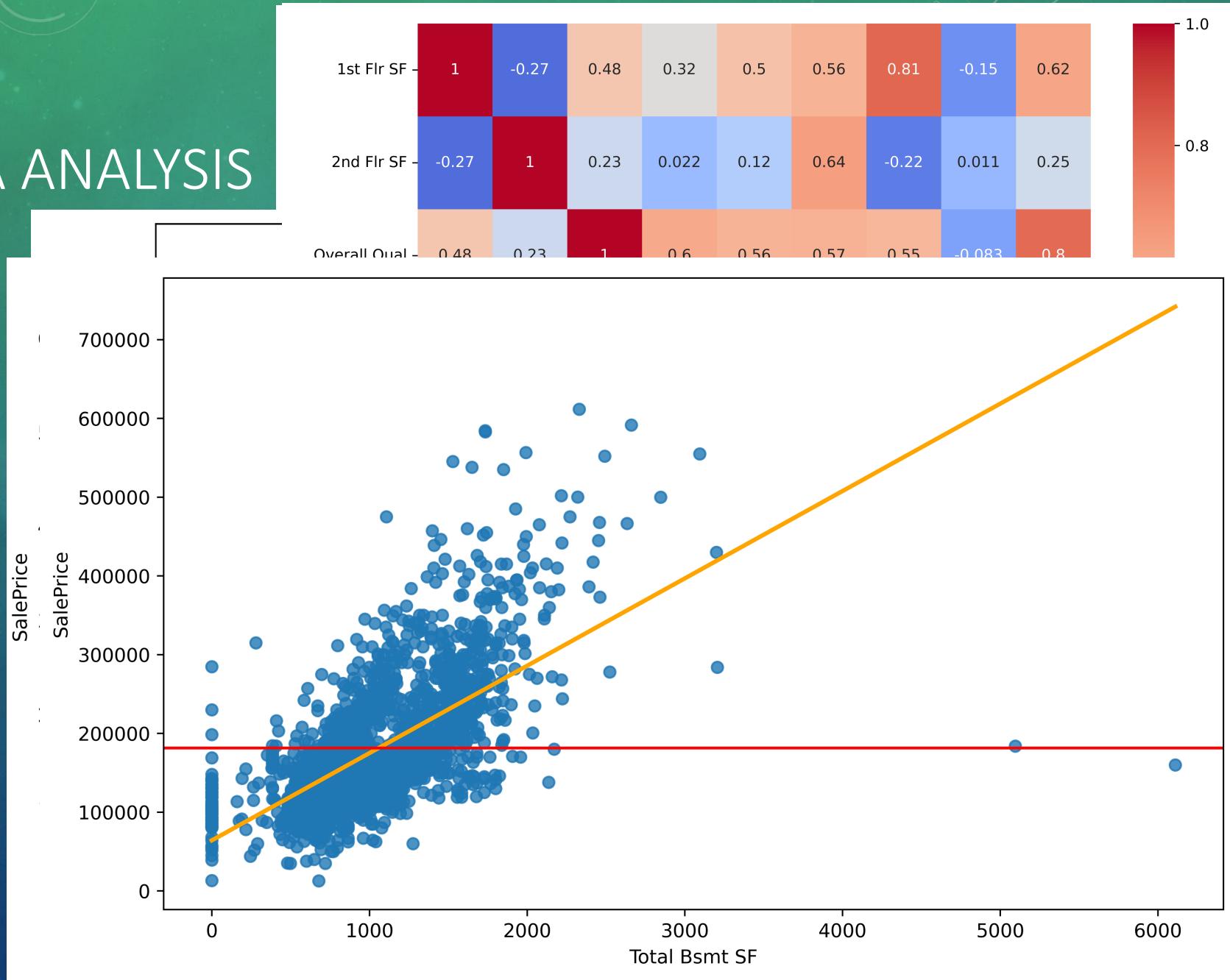
- Developing Ames the right way
  - Understanding initial cost of development, by calculating land cost
  - Fairly compensating home owners
- 
- “Neom, Saudi Arabia - US\$500 billion”
  - “California High-Speed Rail, US – US\$113 billion”
  - 6 months to 6 years in preplanning before ground is broken

# METHODOLOGY

- Identify most relevant parameter/house features
- Focusing on square footage and area
- Create baseline model
- Introduce Preprocessing and Feature Engineering techniques to improve models performance

# EXPLORATORY DATA ANALYSIS

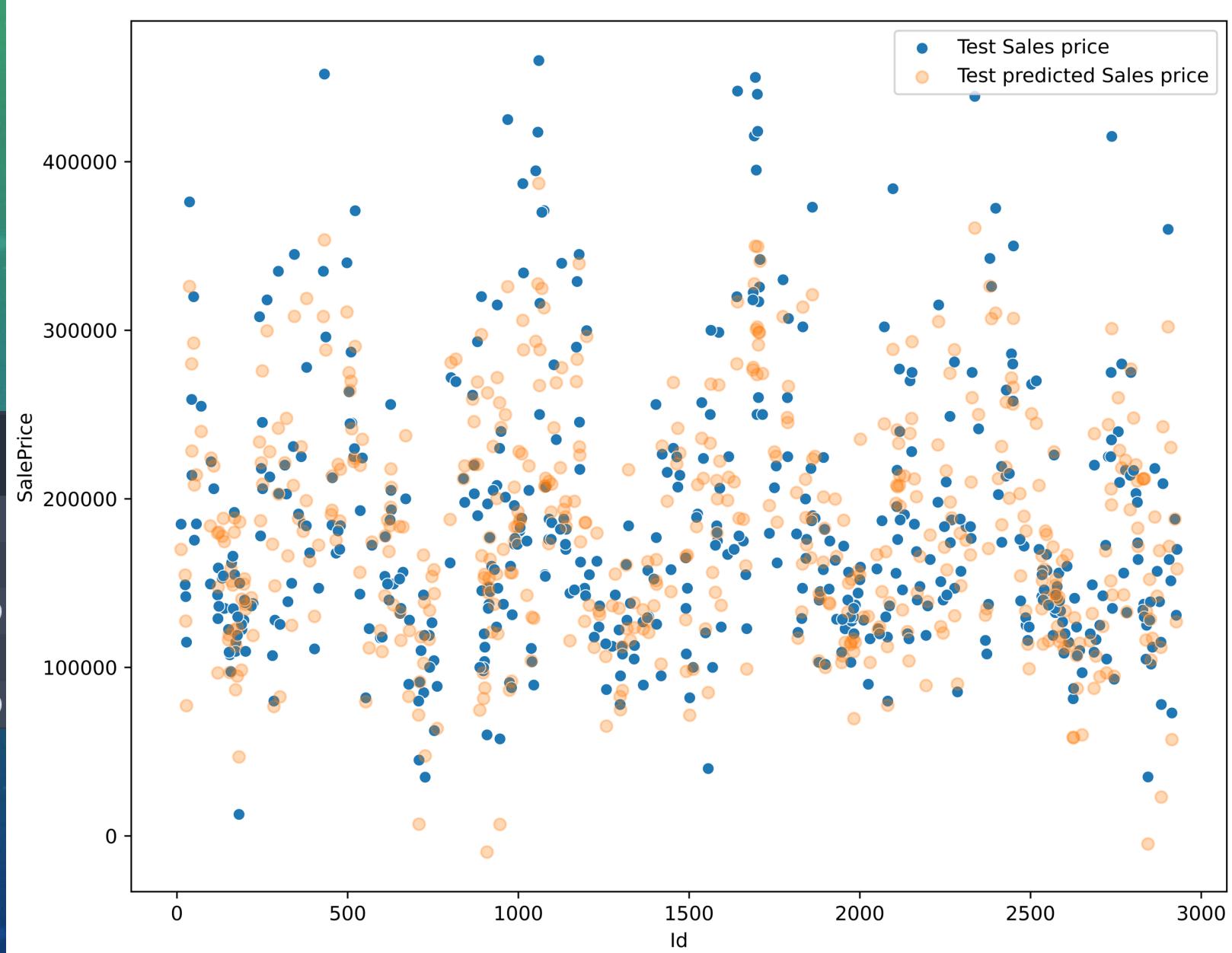
- No obverse correlation between sales price and lot area
- Overall Quality, Garage Area, Year Built & Full Bath showing higher correlation to the Sale Price.
- We can see how the Sales price increases as the overall quality increases.
- Total Basement SF, 1<sup>st</sup> Floor SF and Ground living area
- Good positive correlation with a few anomalies
- Helps show how other areas of the property affects Sales price



# BASELINE

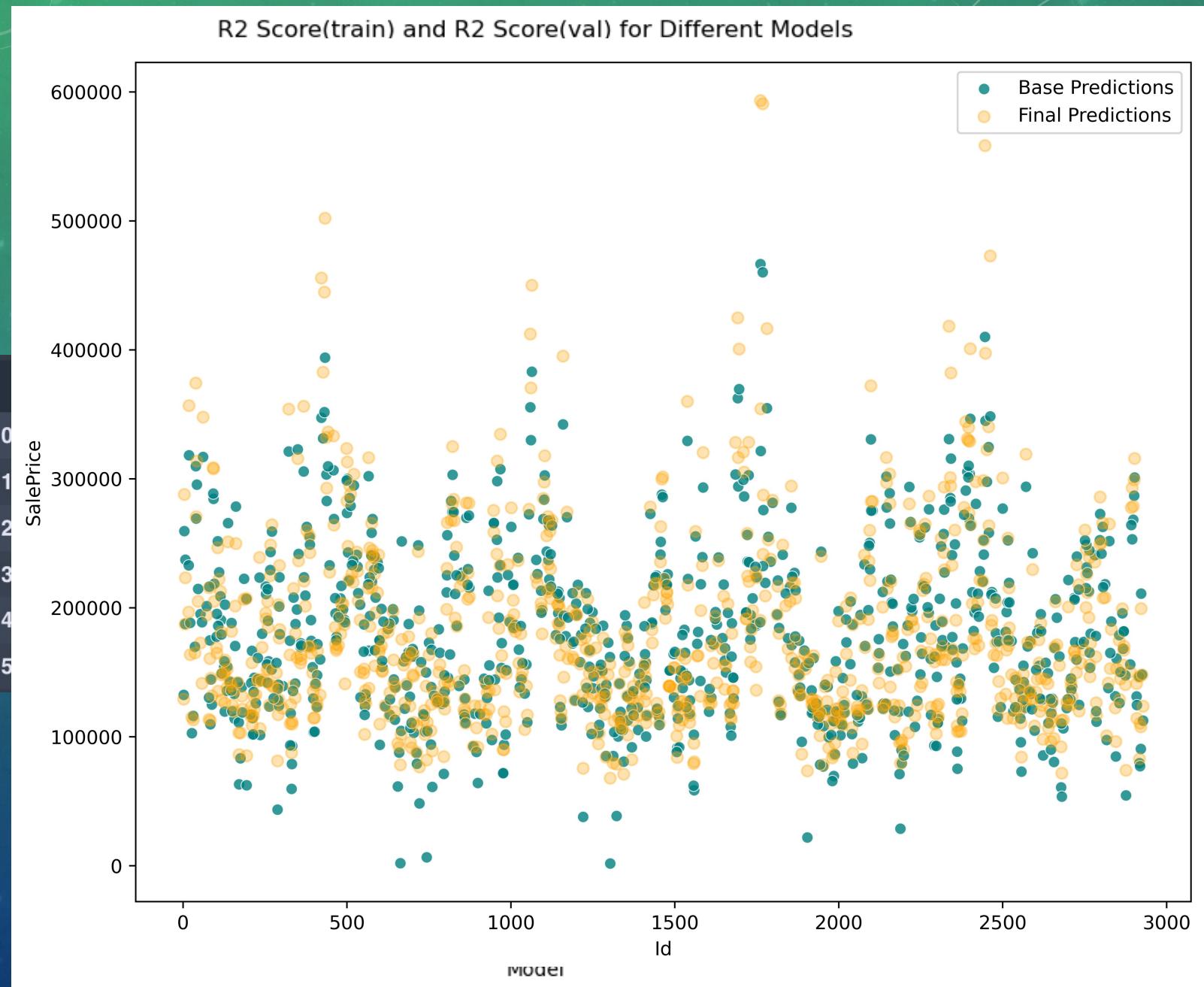
- Let's have a look at the data that we will be using.
- Metrics Metrics Metrics!!!!
- Comparing the given provided Sales prices with the predicted

Id	Overall Qual	Year Built	Full Bath
<b>Baseline</b>			
$R^2$ Train score	0.771023		
$R^2$ Train score	0.835513		
MAE	23307.513959		
RMSE	31779.748064		



# MODEL TUNING

- Polynomial Features
- Standard Scaling
- Both
- $\approx \$4000k$  drop in MAE
- Linear Regression vs Ridge CV
  - Regularization
  - Handling Multicollinearity



# CONCLUSION & RECOMMENDATIONS

- Predicted total costing of buying all the properties = \$157,135,272.02
- Variation of = \$19523.526618
- $\approx \$17,141,656.37$
- Recommendations
  - Introducing more variables
  - Deeper dive into Neighborhood and House style
  - Investigate if Year of remodel and additional constructions changes the model



# THANK YOU FOR LISTENING ANY QUESTIONS?

## CITATIONS AND THANKS

Thanks to Rowan and Tim for answering all my endless questions.

Background picture: <https://www.aimircg.com/3d-architectural-rendering-for-a-real-estate-development-in-montana-usa-case-study/>

7 Most Expensive Construction Projects: <https://www.international-construction.com/news/7-of-the-world-s-most-expensive-construction-projects/8026711.article>

Real Estate Development Timeline: <https://lev.co/blog/assets/real-estate-development-timeline/>