Ames House Prices

Problem Statement: Given a set of inputs, can we accurately predict the prices of a house?

Dataset description

- 2051 entries, 81 unique features.

- 2006 – 2010

- Lots of null, empty and repeated values.

Model

The model predicts home prices within **<u>16%</u>** of the
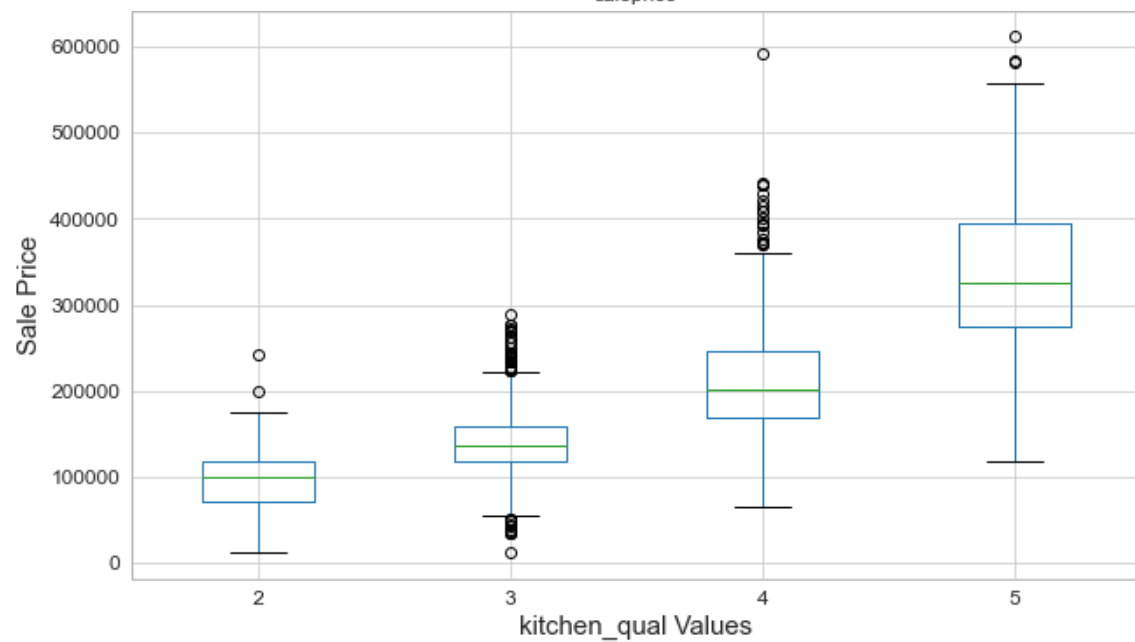average Ames home price

±$32,000

- RMSE: $ 32,000 (ave root squared error between actual model prediction)
- R^2: 0.84 (Goodness of fit between actual and model, 1.0 being a perfect model)
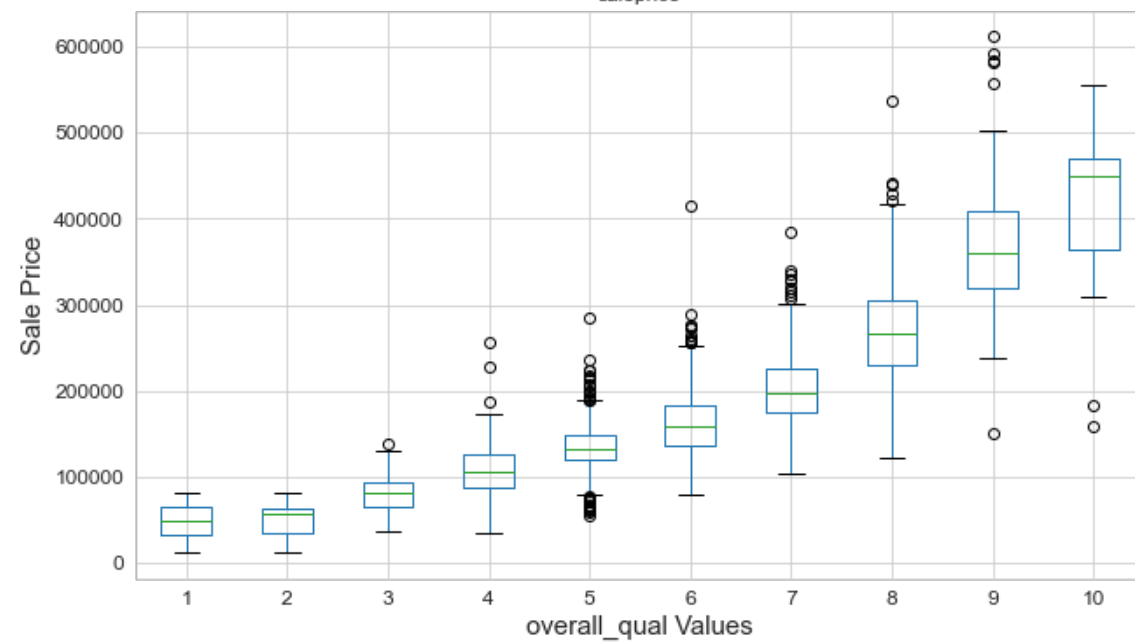
Factors that affected the model

- Quality
- Size
- Location
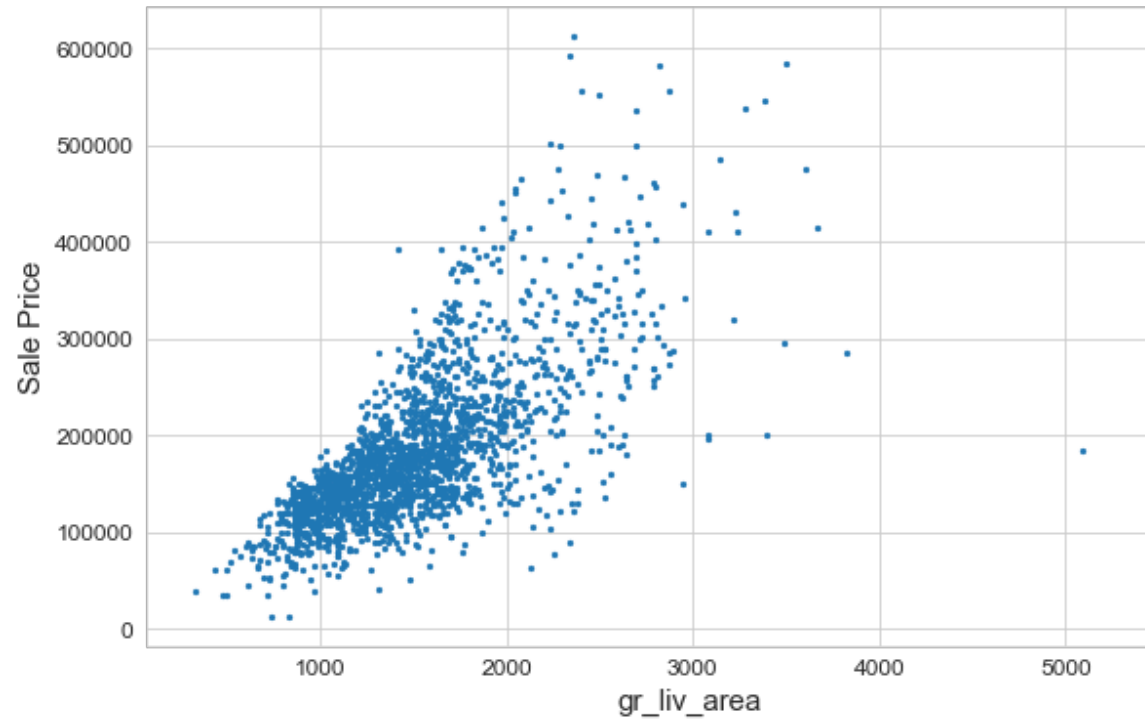- Age

Quality

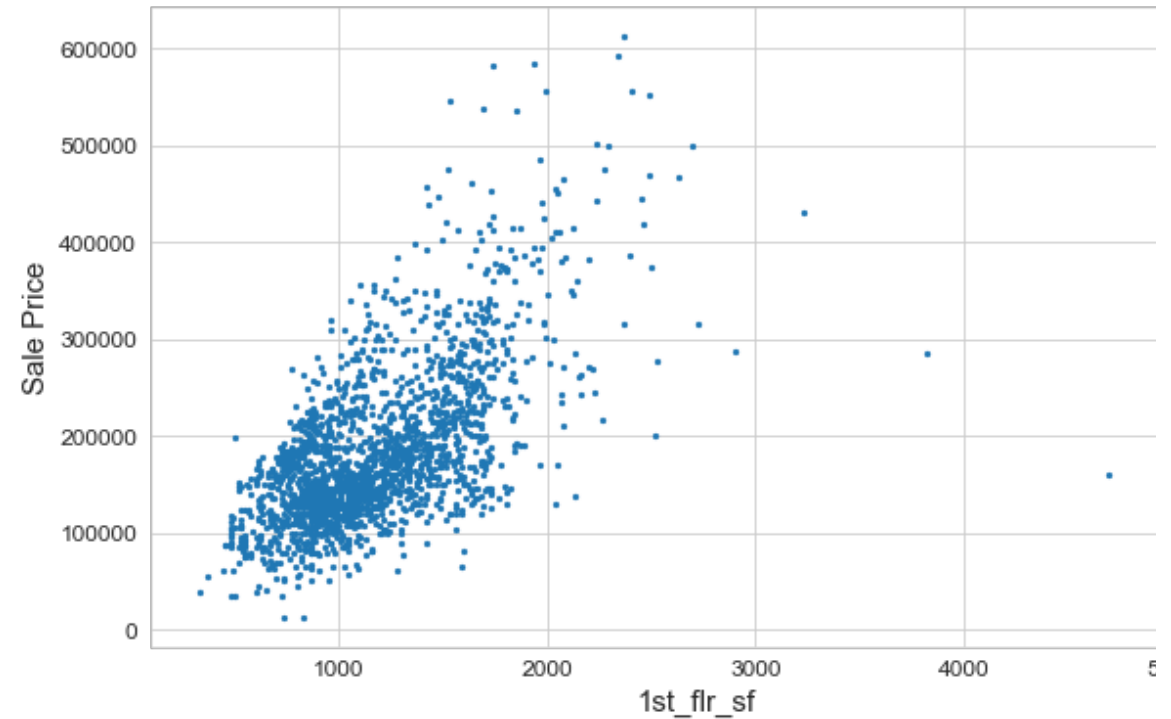Boxplot grouped by kitchen_qual

Boxplot grouped by overall_qual

Size

Sales Price Scatter Chart
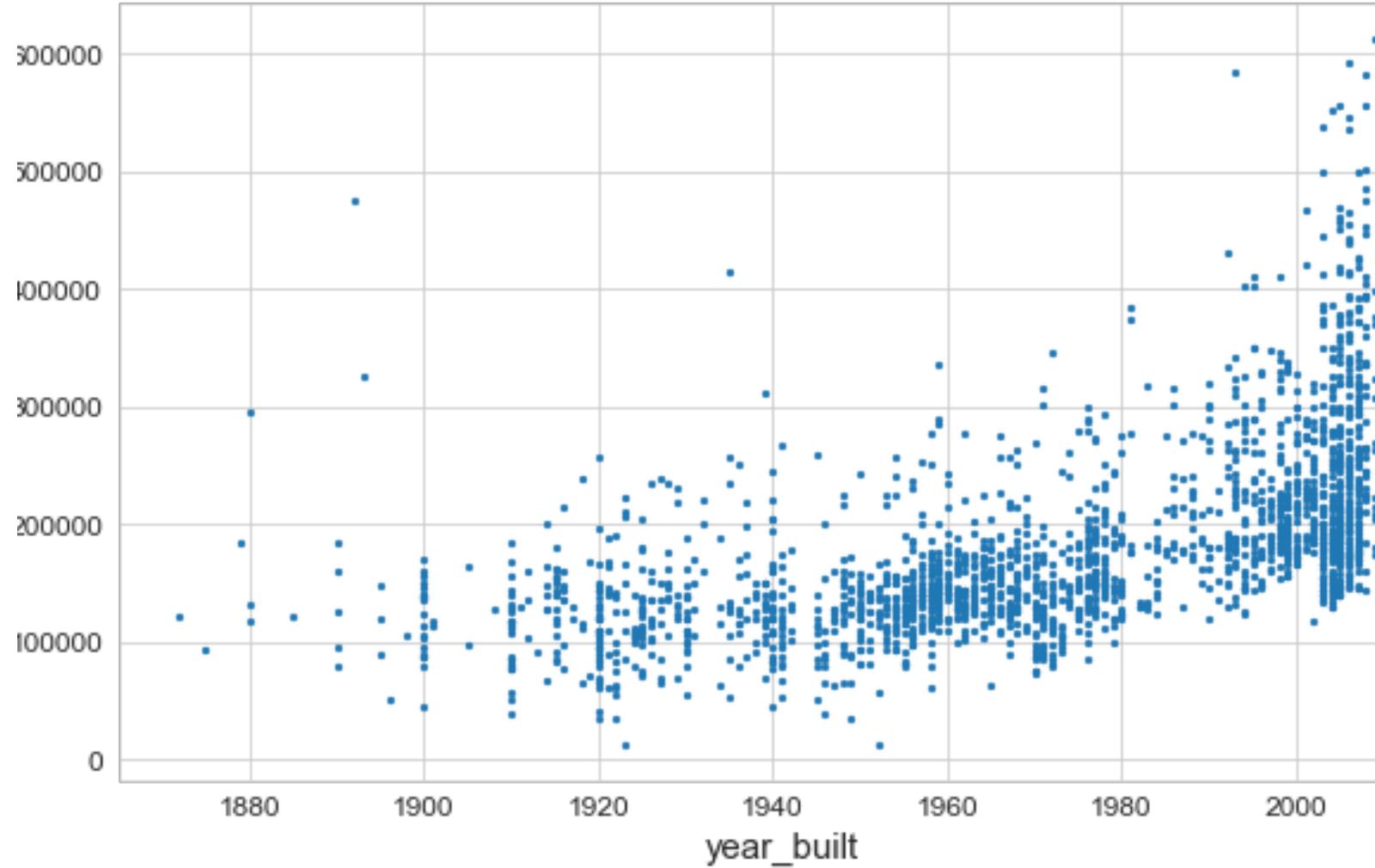
Age

Boxplot grouped by neighborhood

Location

Location

Boxplot grouped by condition_1

saleprice
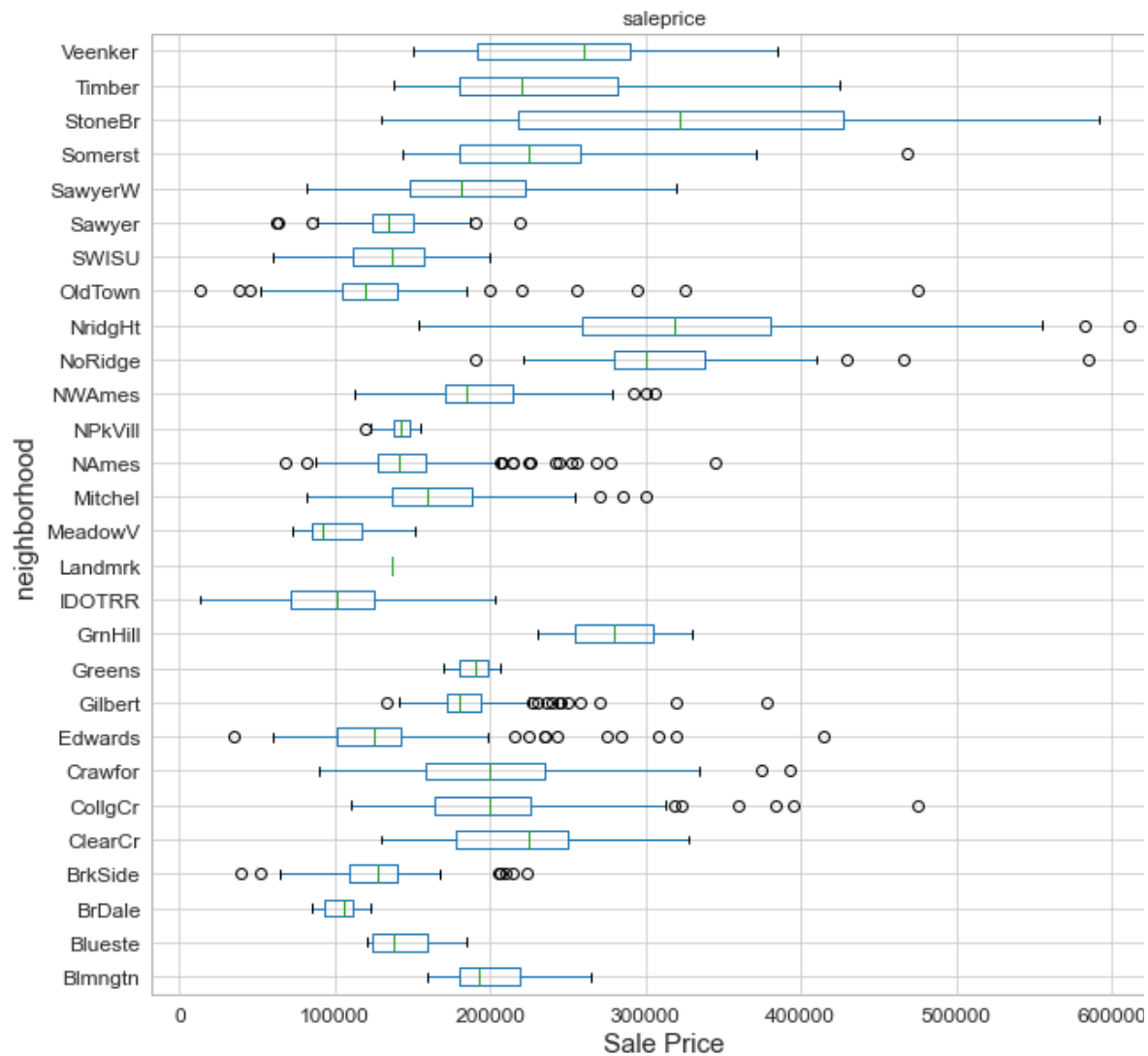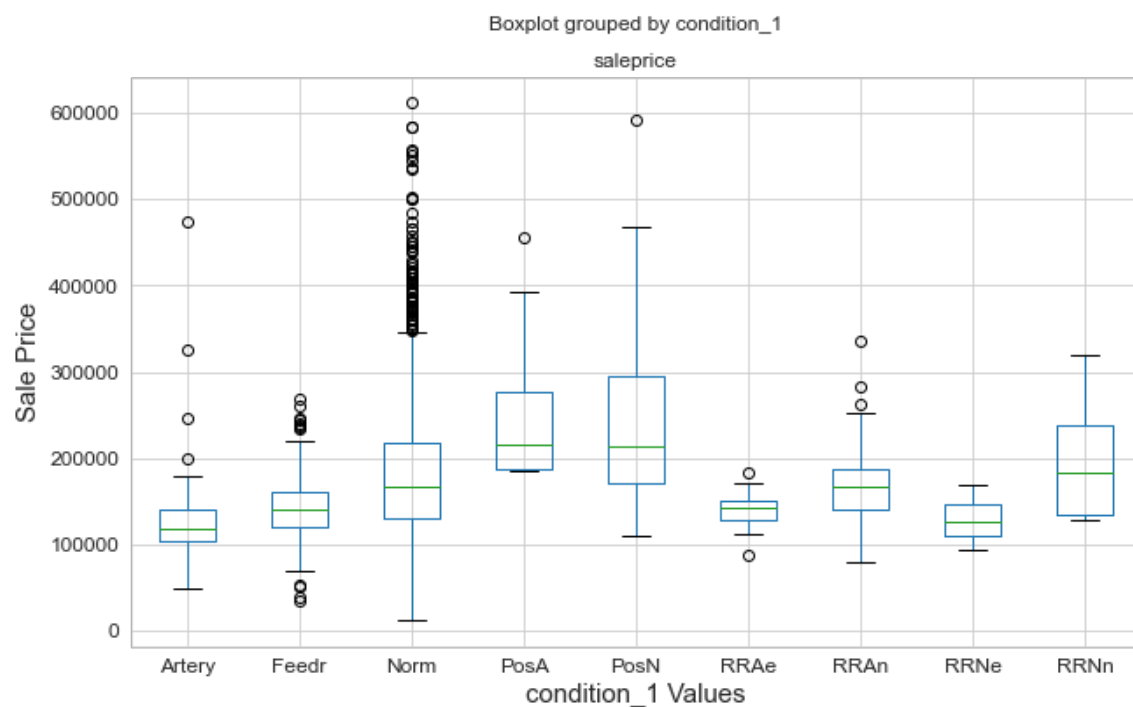
Artery	Adjacent to arterial street
Feedr	Adjacent to feeder street
Norm	Normal
RRNn	Within 200' of North-South Railroad
RRAn	Adjacent to North-South Railroad
PosN	Near positive off-site feature--park, greenbelt, etc.
PosA	Adjacent to postive off-site feature
RRNe	Within 200' of East-West Railroad
RRAe	Adjacent to East-West Railroad

Who will find the model useful?

- Buyers : Will be able to figure out if they are overpaying or underpaying

- Sellers: Will know what to do to improve the value of their home

- Investors: Will be able to know what locations to invest in

Limitations of model

- Data is only 4 years long and out of date

- Factors such as inflation may not have been taken into account for over the years

- Model is only applicable to Ames, Iowa and cannot be generalized to other regions.

Thank you very much