Home Insurance Premiums:

Determining Relevant Home Features

Regression Analysis on King County, WA Homes





Martin Reyes

Home Insurance Premiums

Determine the validity of Goal insurance premiums Which home features **Question** matter? Models and statistics that analyze **Results** feature-price relationships





Model and Analysis Process



01 Obtain



Source King County House Data from May 2014 to May 2015

O2 Scrub



Clean the data

O3 Explore



Analyze statistics of the house features and price

04 Model



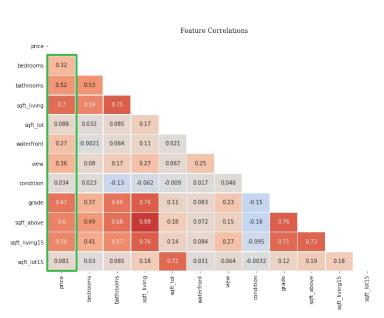
Analyze feature-price relationships with linear regression models

05 INterpret



Gather and organize our findings

Results: Which Features Matter?





Relevant Features

- Sqft.
- KC grade
- Sqft of 15 nearest neighbors
- Bathrooms



Irrelevant Features

- Lot sqft.
- Condition (unknown source)
- Day/month sold



Results: Model Suggestions

Best Performing Model

71%

71% of the variation in price is explained by the features

Most Trusted Model

47%

Assumptions met

Takeaways

Strong relationship with price:

- Total Sqft
- KC grade
- 47.5° N



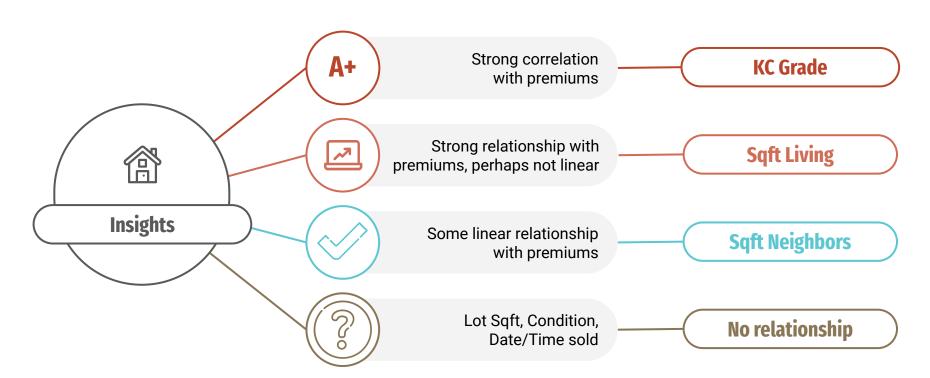
Takeaways

Some linear relationship with price:

- Sqft of 15 nearest neighbors
- KC grade



Summary of Insights



Next Steps



Feature Engineering And Selection

Advanced techniques to efficiently improve models



Exploratory Data Analysis

More EDA on the home features



Data Gathering

Source more data for better analyses and models



Machine Learning

Develop a ML model to accurately predict house prices



Thank You

martinreyes.eng@gmail.com

GitHub

github.com/MartinReyesDS

LinkedIn

www.linkedin.com/in/martin-reves-ds

