Data Science

Module 1 Project

Chelsea Power

King County Housing, USA (2015)

What factors influence the price of a house in this area?

What factors have little to no influence on the price of a house in this area?

How has the cost of housing changed over the years?

Methodology: OSEMiN

- **Phase 1 - Obtain**: Obtained and loaded the 2015 data from the kc_house_data.csv file
- **Phase 2 - Scrub**: Scrubbed data and conducted feature engineering by removing null/missing values and reset the types of many variables in the dataset
- **Phase 3 - Explore**: Generated heatmaps and histograms of the data to check for multicollinearity, removed outliers and then scaled/normalized the variables
- **Phase 4 - Model**: Generated OLS models, checked r-squared and p-values.
- **Phase 5 - iNterpret:** Results...

Model Results

Conclusion