

Kings county home sales predictive analysis

JOEY HUSNEY 10/01/2020



Outline

- About the king's county dataset
- About the Process
- Data cleaning
- Features used for predictions
- Prediction Analysis
- Business Recommendations
- ▶ Future Work

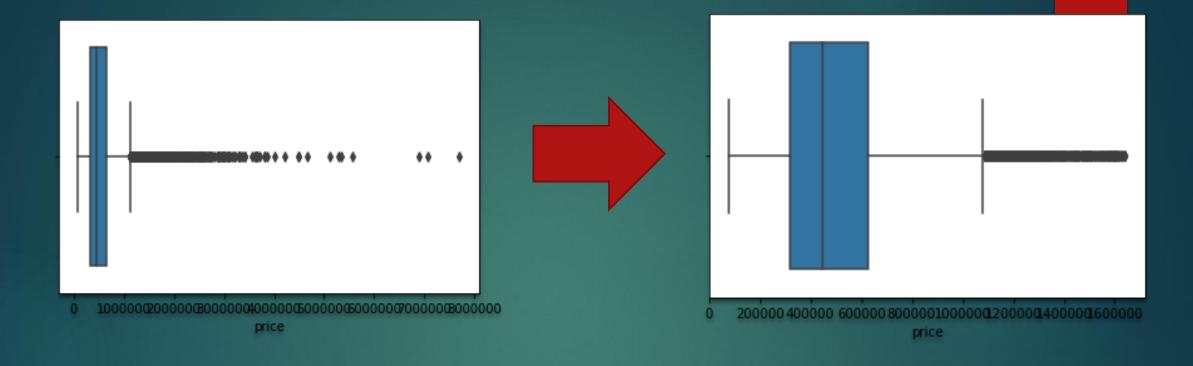
About the dataset

- Contains info for part of the population of the king's county in Washington
- ▶ Small dataset
- Hard to predict certain info with lack of variables that affect price
- Many incomplete columns to deal with

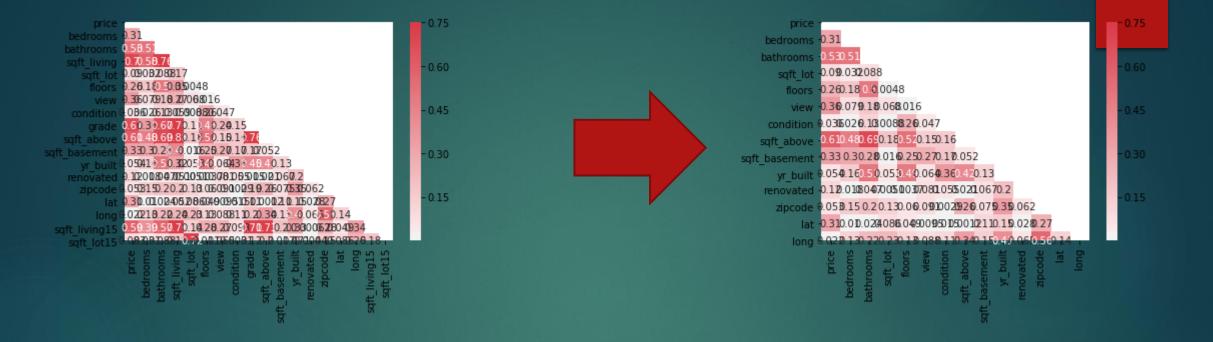
About the Process

Data Science Process SCRUB OBTAIN EXPLORE MODEL **INTERPRET** Clean data to formats Find significant patterns Gather data from Construct models to Put the results into relevant sources that machine and trends using predict and forecast good use statistical methods understands

Originally by Hilary Mason and Chris Wiggins



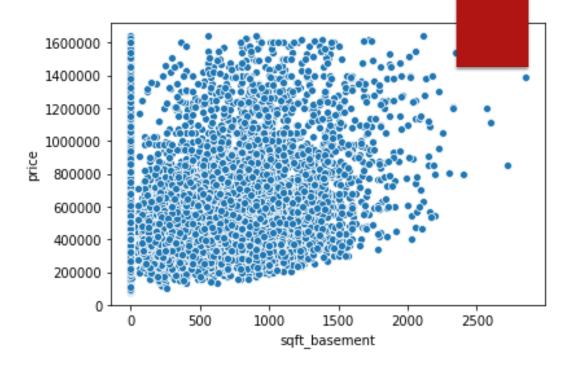
Data cleaning- Outlier removal

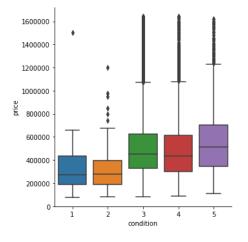


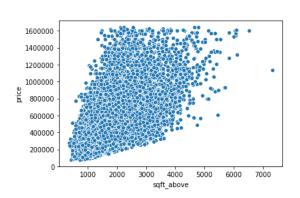
Data cleaning- Multicollinearity (Notice the lighter shade of red due to less columns that relate to each other)

Home sale predictors

- Sqft_basement seems to be correlated with sale price
- Sqft_above also seems to be correlated with the sale price
- The condition definitely has an effect on the sale price







Business Recommendations

Increase home square footage as much as possible

Home should be in very good condition

Future Work







COMPARE THE AREA SALE PRICE BASED ON DIFFERENT CULTURED AREAS



ANALYZE PRICES BASED ON AREA POPULATION

Thank you for listening!