

Exploratory Data Analysis

King County (Washington) real estate sales
May 2014 - May 2015

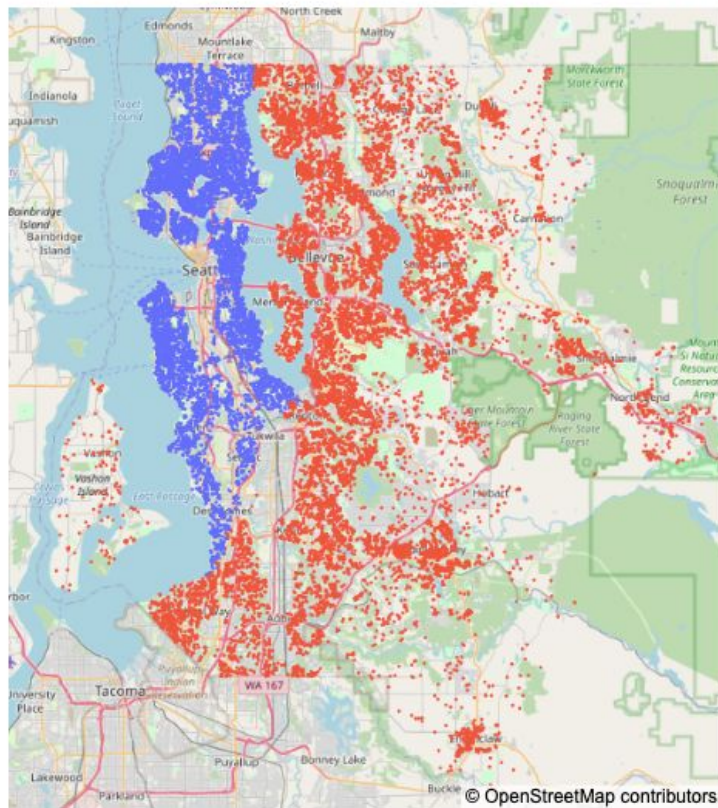
by Jakob Wegener

The Stakeholder

The target of our analysis is Timothy Stevens:

- a seller who owns expensive houses in the centre (Seattle).
- He needs to get rid of them fast and is interested in the best timing for selling within a year.
- He is open for renovation when profits rise

Subselection of the data based on geography

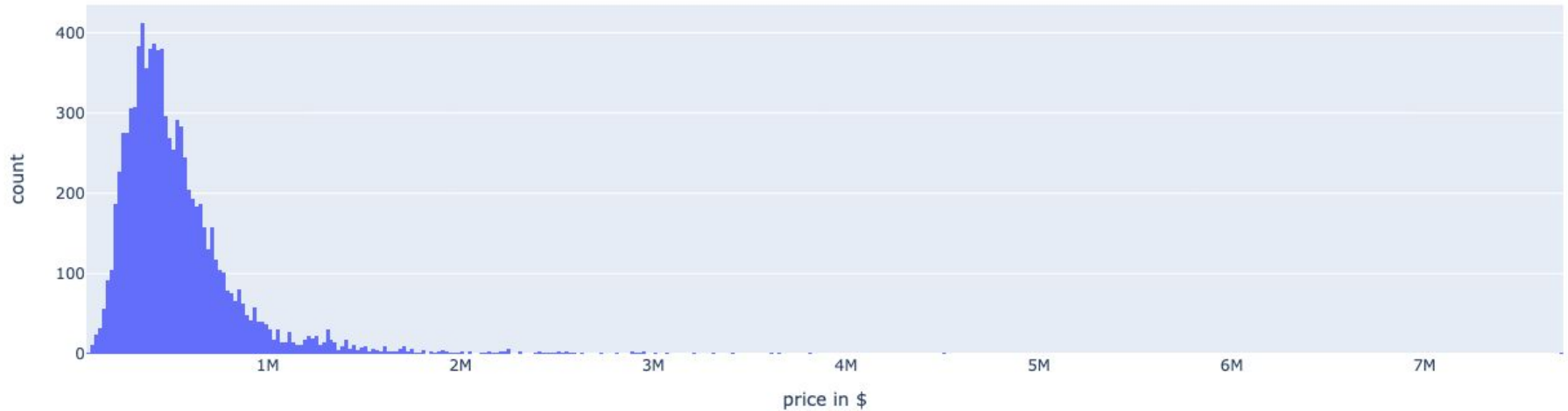


is the datapoint in Seattle

- True
- False

further subselection based on price

histogram of house prices in Seattle



only prices higher than the 75th percentile: > \$ 630000

research questions and hypotheses

- when is the best time to sell during the year?
 - if there is a point at time with significantly lower amount of sales, lower prices should be accepted to sell before that point
- when does it makes sense to renovate?
 - after a certain time period between construction year and current date a renovation should increase the selling price
 - raising the building grading by renovation should also increase the selling price

direct influence of renovation on price?

- hypothesis: houses that were sold more than once in dataset and had renovations in between can show what influence renovation has on price
- Analysis showed that no such data points exist in the entire dataset

Create new features in data selection

- time between construction and renovation in years
- time between renovation and sale in years
- time between construction and sale in years in houses in which no renovation took place

when to sell?

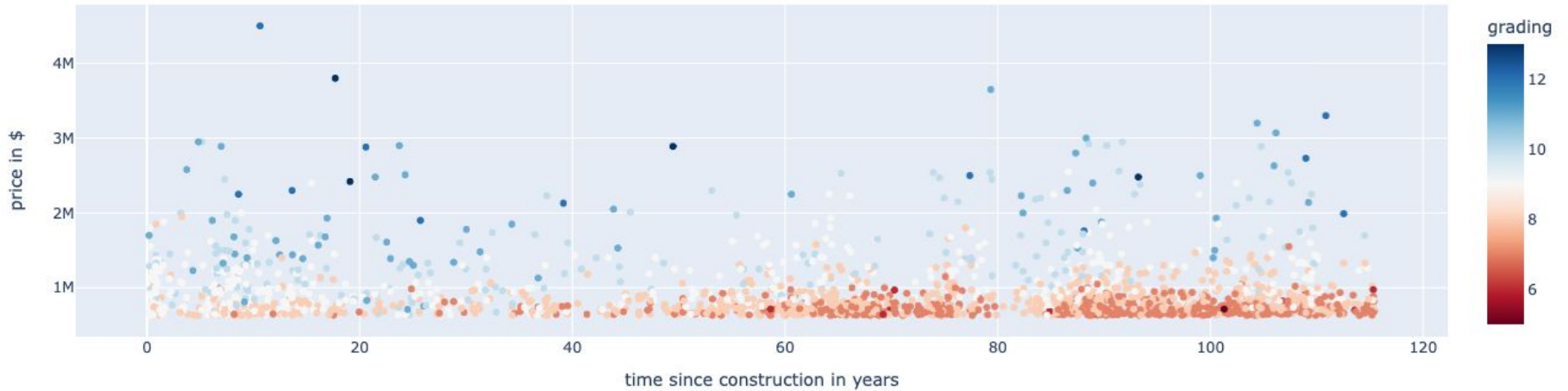
distribution of house sales between May 2014 and May 2015



=> slump of sales around thanksgiving and after christmas/until february

when to renovate?

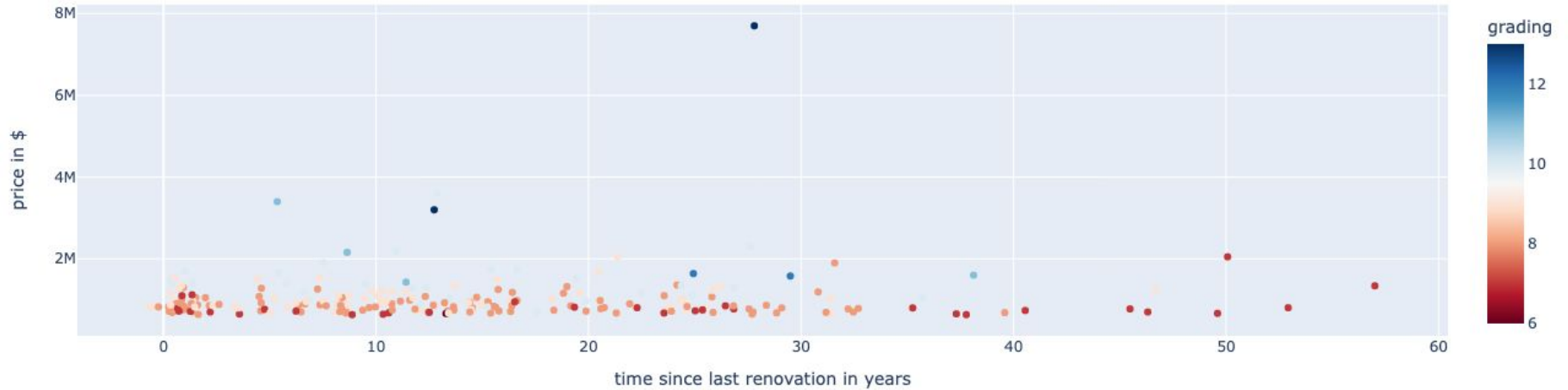
price plotted against time since construction in unrenovated houses



=> houses with grading 9+ have higher prices

when to renovate?

price plotted against time since last renovation



=> no clear increase in grading because of recent renovation

Insights and recommendations

- slump of sales at holidays (november: thanksgiving, december: christmas) and at the start of the year

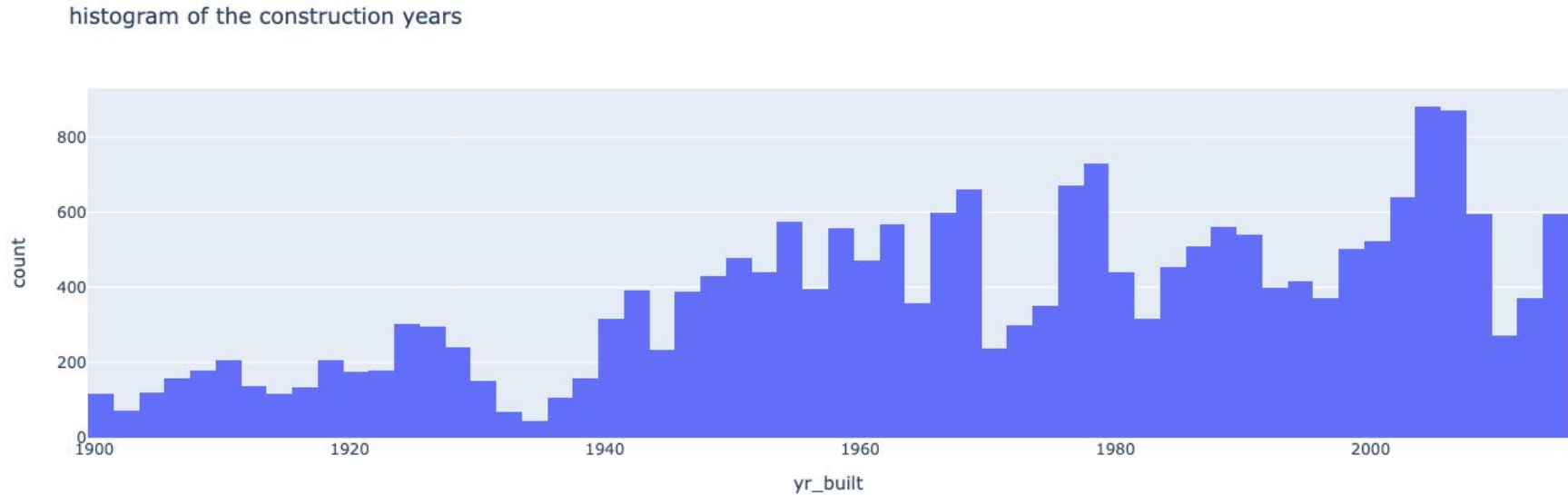
=> if property has to be sold this year, accept lower prices to sell before end of november

- higher grading is correlated with higher sale prices
- recent renovation is **not** correlated with higher grading
- no data points in the dataset which can show direct influence of renovations on price

Outlook

- analyze bigger dataset to show direct link between construction and renovation
- find ways to normalize the influence of size, view rating, waterfront access to see clearer if renovation has an influence on the price

Bonus plot!



=> clear influence of economic crisis (1930: great depression, 1970: oil crisis, 2008: real estate crash) on house construction!