Houses for the mafia

EDA project - neue fische

Max Langer

The stakeholder

- Amy Williams is an italian mafiosi
- works in King County (Washington, USA)
- sells central houses
 - o is interested only the top 10% of the houses
- needs average houses in the outskirts
 - o to hide from the FBI from time to time



zdf.de

The data set

- information on houses in King County (Washington, USA)
- includes 21500 houses with sales dates from 2014 and 2015
- contains information such as
 - location
 - o prices
 - number of rooms
 - square footages
 - grade and condition rating of King County



govtech.com

Quality of the data

- the data quality was ok for most of the information
- missing entries could be filled in by understanding the data
- such information were
 - viewings by buyers
 - o whether the house has a view on the water
 - the year of the last renovation



iqint.org

What makes up the top 10% of houses?



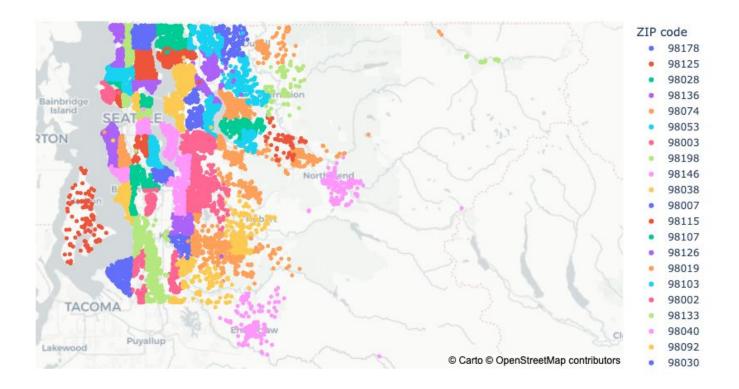
istockphoto.com

- Does the ZIP code define an outskirt or central house?
 - Hypothesis: Outskirt and central houses can be distinguished by their ZIP code

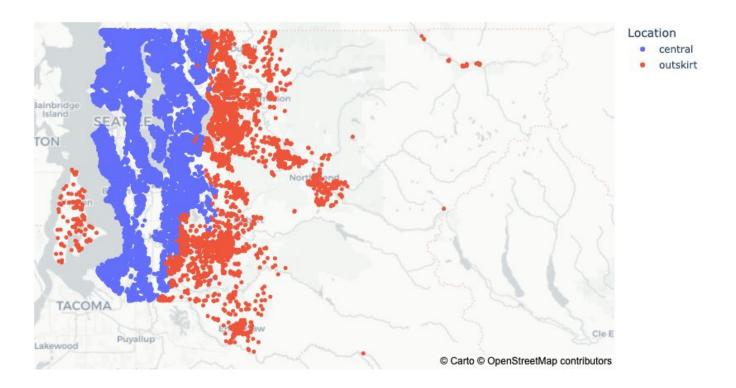


res.cloudinary.com

Does the zip code define a central or outskirt house?



Central and outskirt houses



- Does the ZIP code define an outskirt or central house?
 - Answer: Yes, they can be distinguished by ZIP code



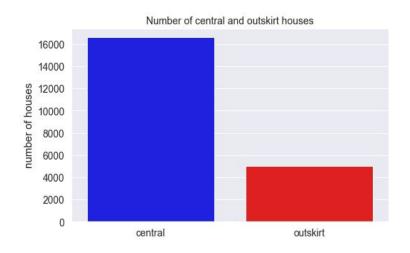
res.cloudinary.com

- Is the number of houses in the central area higher than in the outskirt area?
 - Hypothesis: There are more houses in the central area



res.cloudinary.com

- Is the number of houses in the central area higher than in the outskirt area?
 - Answer: Yes, there are more houses in the central area



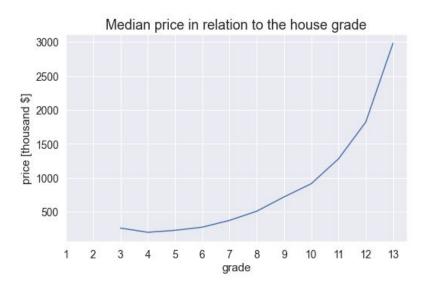
What makes up the top 10% of houses?

- Do the grading, condition rating and living space impact the prices?
 - Hypothesis: A better grade, condition rating and larger living space increase the price



luxuryhomemagazine.com

Impact of grade and condition on the price



correlation coefficient between price and grade: 0.66



correlation coefficient between price and condition rating: 0.018

Impact of living space on the price



correlation coefficient between price and living space is:

- 0.717 for central houses
- 0.782 for outskirt houses

What makes up the top 10% of homes?

 Do the grading, condition rating and living space impact the prices?

Answer:

- Yes, a higher grade and larger living space correlate with the price of a house
- No, condition is only a minor factor



luxuryhomemagazine.com

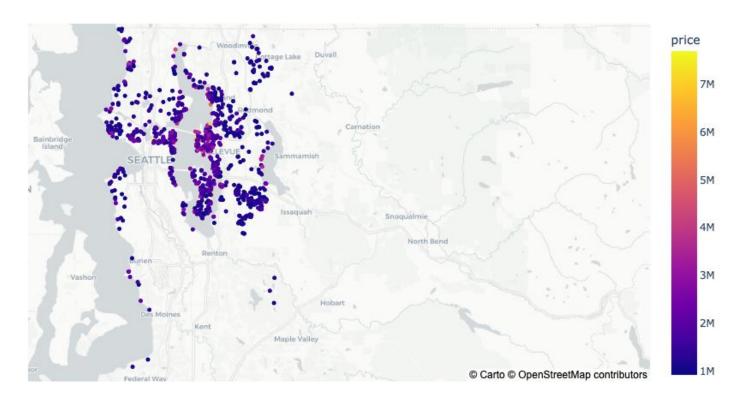
What makes up the top 10% of houses?

- Three recommendations to find central houses that are in the top 10%
 - house prices higher than 920000\$
 - o grades of 9 or higher
 - a minimum of 3140 sqft living space



luxuryhomemagazine.com

These are the top 10% of homes

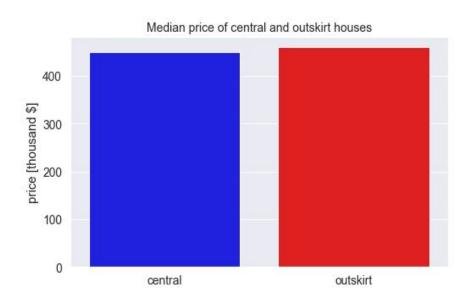


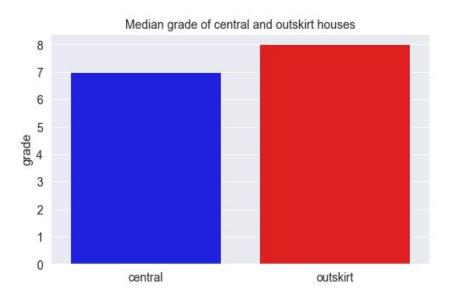
- Does the average price differ between outskirt and central houses?
 - Hypothesis: Houses in central areas are in average pricier
- Is the average grade lower in the outskirts?
 - Hypothesis: Houses in the outskirt areas have overall lower grades



minimalblogs.com

Price and grade between central and outskirt





- Does the average price differ between outskirt and central houses?
 - Answer: No, house in outskirt and central areas are in the same price range
- Is the average grade lower in the outskirts?
 - Answer: No, outskirt houses are in general one grade higher



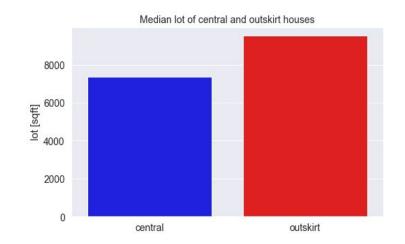
minimalblogs.com

- Do outskirt houses have more lot than central houses?
 - Hypothesis: Houses in the outskirt areas have more lot



minimalblogs.com

- Do outskirt houses have more lot than central houses?
 - Answer: Yes, houses in the outskirt have more lot

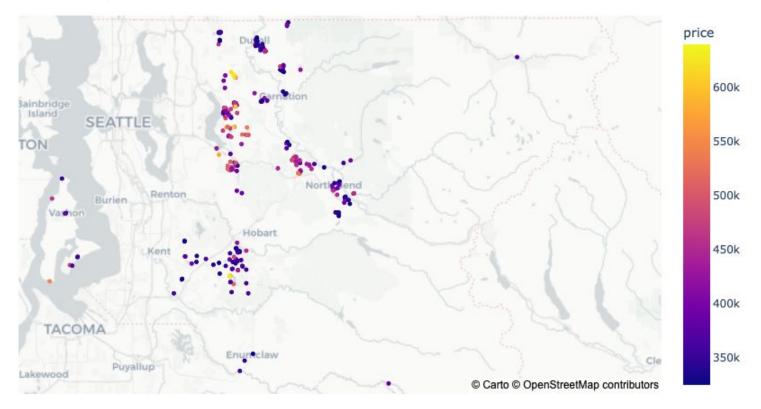


- Three recommendations to find an average outskirt house?
 - house prices of around 460000\$
 - o grades of 8
 - o lot of around 9500 sqft



minimalblogs.com

The average outskirt houses



Conclusion and outlook

- helped Amy making offers you can't refuse
 - regarding top houses and good hideouts

- a different approach could be used for
 - distinguishing between outskirt and central
 - based on the latitudinal and longitudinal data
- the view of water has not been considered as a factor yet
 - but seems to play a role for the top 10% houses



tripadvisor.com