

# EDA project

neue fische - data science bootcamp

Max Langer

# The stakeholder

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



zdf.de

# The data set

- information on houses in King County (Washington, USA)
- includes houses with sales dates from 2014 and 2015
- contains information such as
  - location
  - 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
  - whether the house has a view on the water
  - the year of the last renovation
- a few outliers could also be found



iqint.org

# The challenge

- What is a central, what an outskirts house?
- What makes up the top 10% of houses?
- What makes an average outskirts house?



istockphoto.com

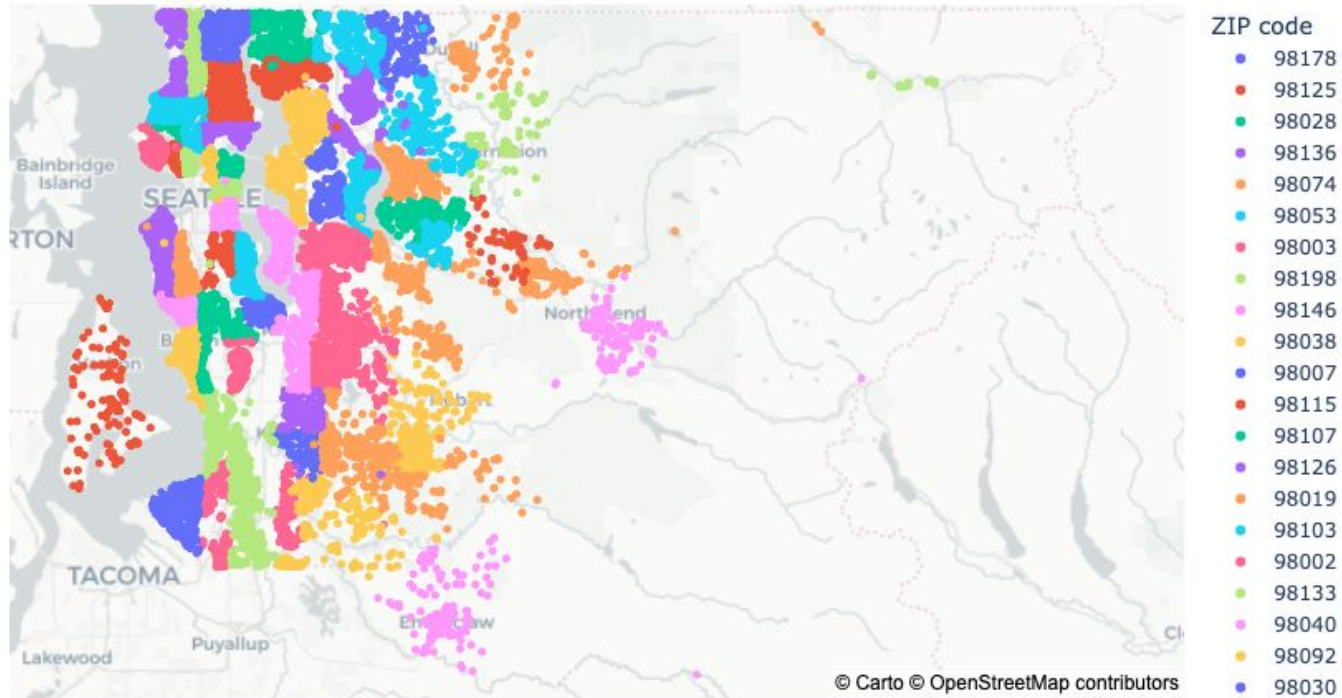
# What is a central, what an outskirt house?

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

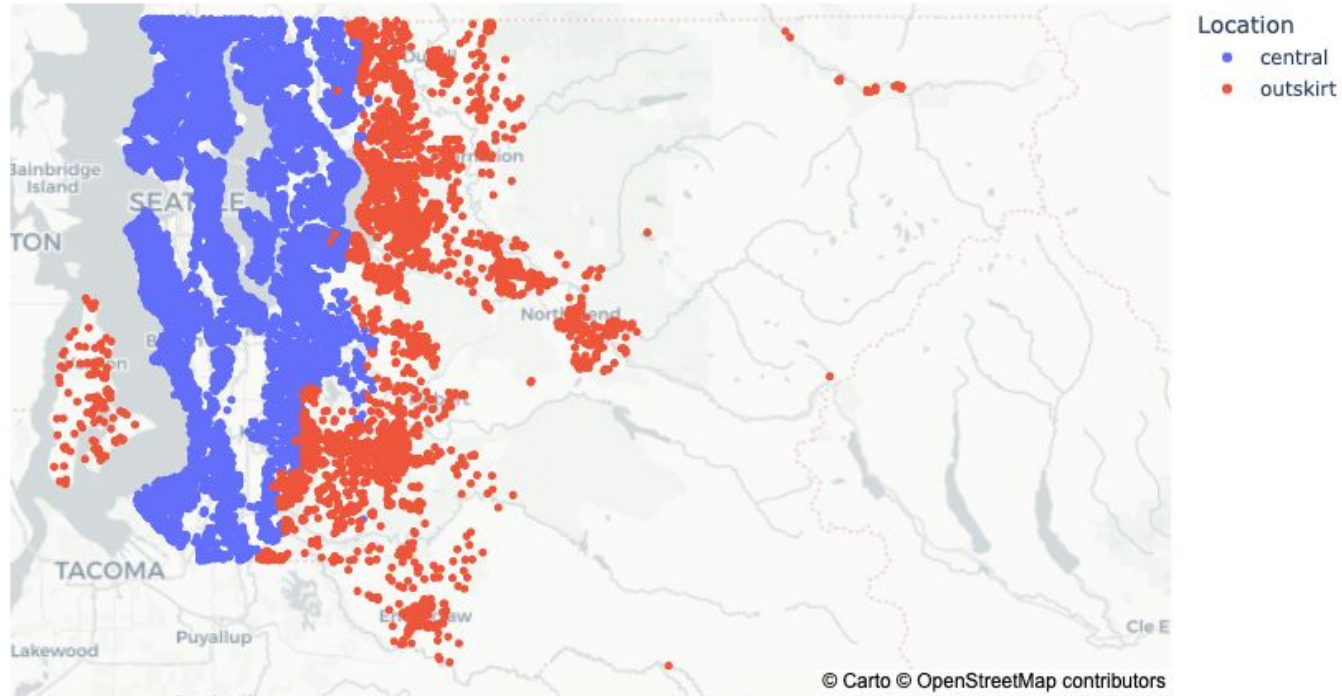


res.cloudinary.com

# Does the zip code define a central or outskirts house?



# Central and outskirt houses





# What is a central, what an outskirt house?

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



res.cloudinary.com

# What is a central, what an outskirt house?

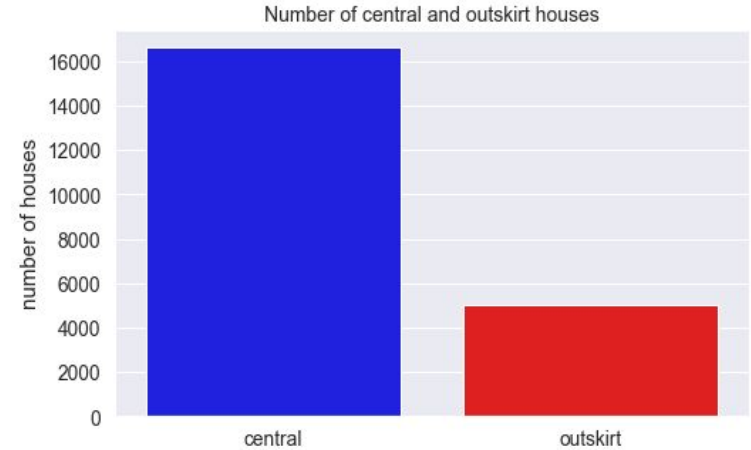
- Is the number of houses in the central area higher than in the outskirt area?
  - **Hypothesis:** There are more houses in the central area



res.cloudinary.com

# What is a central, what an outskirt house?

- 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



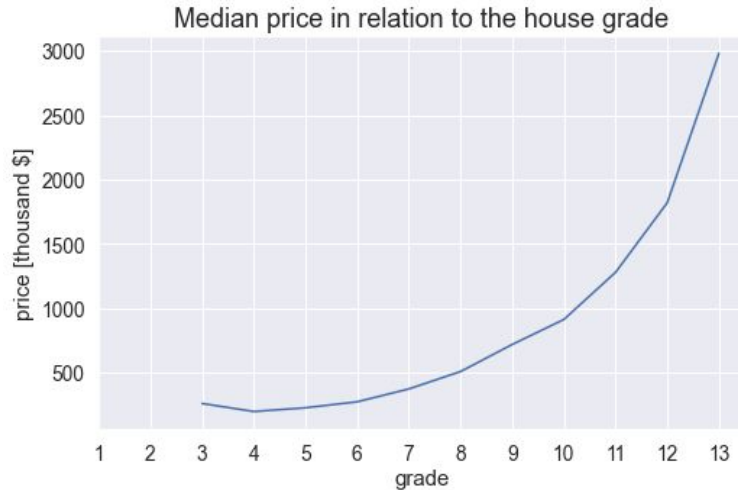
# What makes up the top 10% of houses?

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

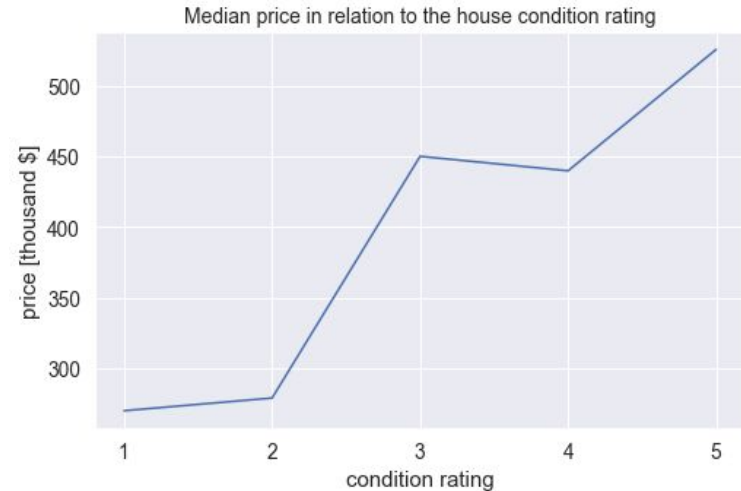


[luxuryhomemagazine.com](http://luxuryhomemagazine.com)

# 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](http://luxuryhomemagazine.com)

# What makes up the top 10% of houses?

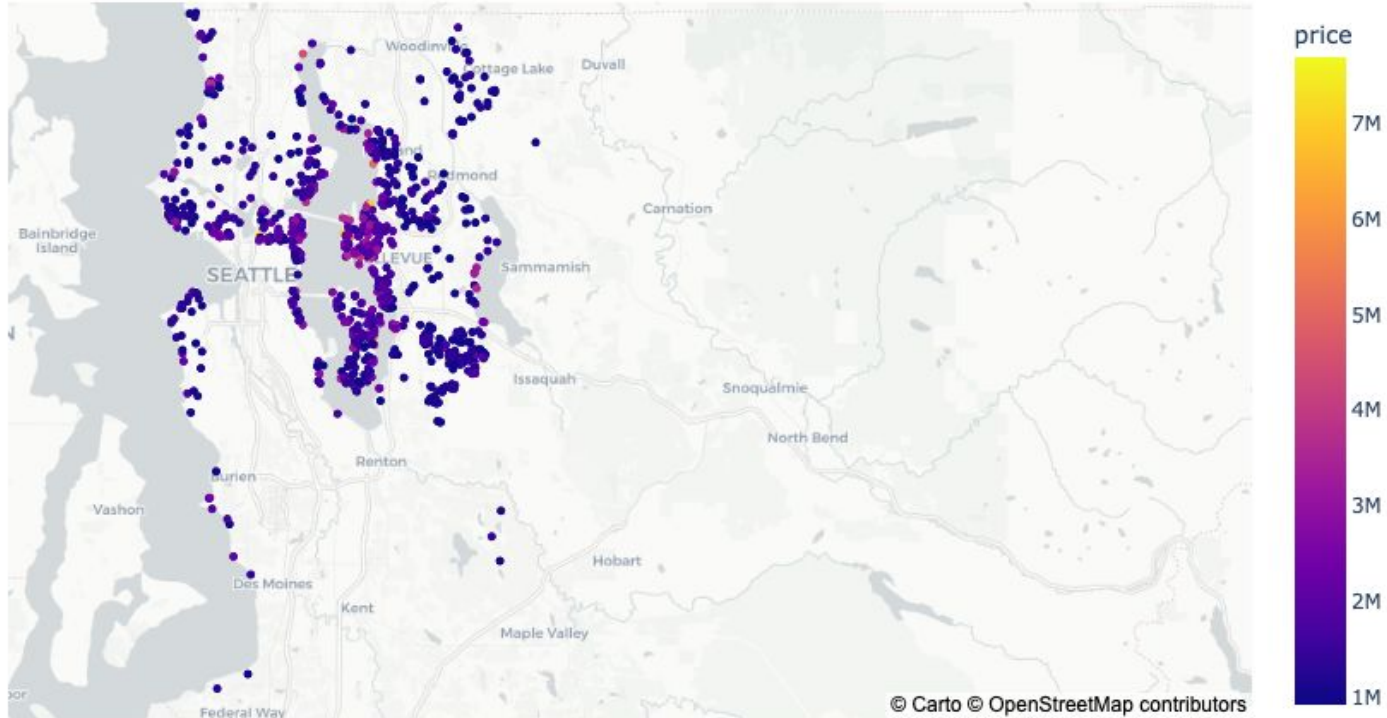
- Three recommendations to find central houses that are in the top 10%
  - house prices higher than 920000\$
  - grades of 9 or higher
  - a minimum of 3140 sqft living space



[luxuryhomemagazine.com](http://luxuryhomemagazine.com)



# These are the top 10% of homes



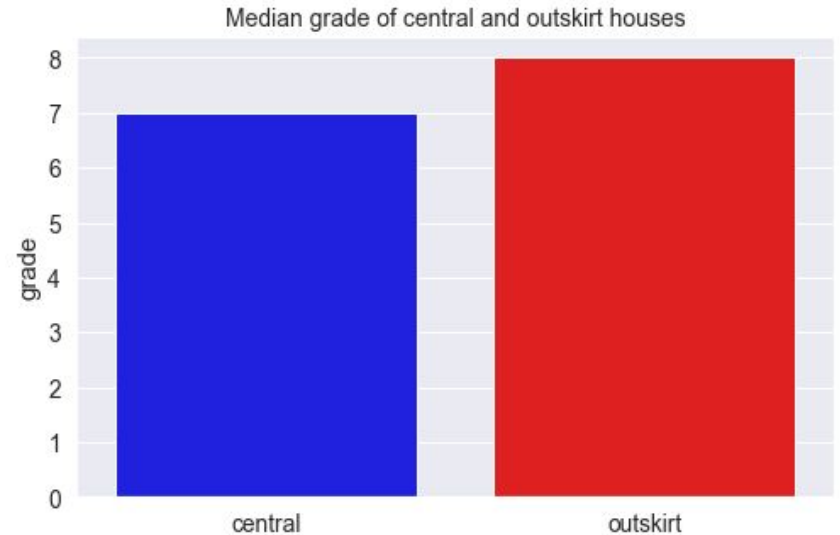
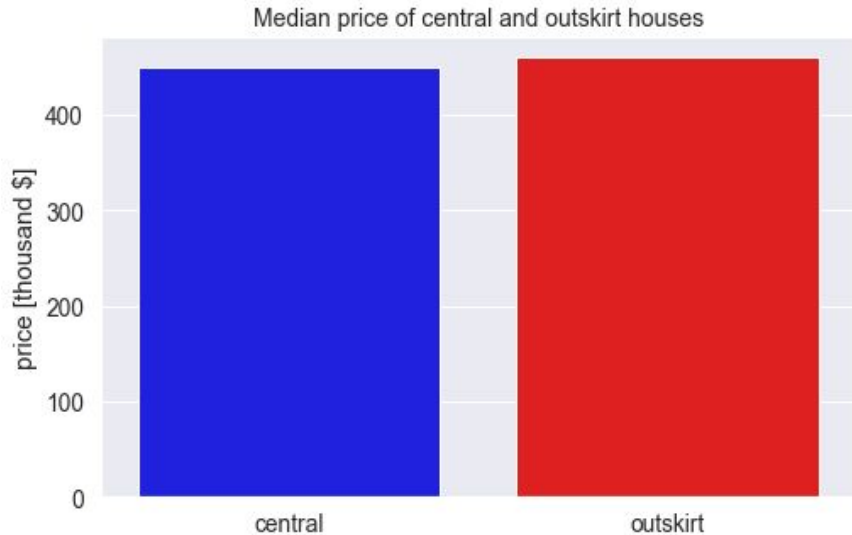
# What makes an average outskirts house?

- Does the average price differ between outskirts and city houses?
  - **Hypothesis:** Houses in central areas are in average pricier
- Is the average grade lower in the outskirts?
  - **Hypothesis:** Houses in the outskirts areas have overall lower grades
- Do outskirts houses have more lot than central houses?
  - **Hypothesis:** Houses in the outskirts areas have more lot



[minimalblogs.com](http://minimalblogs.com)

# Price and grade between central and outskirt



# What makes an average outskirts house?

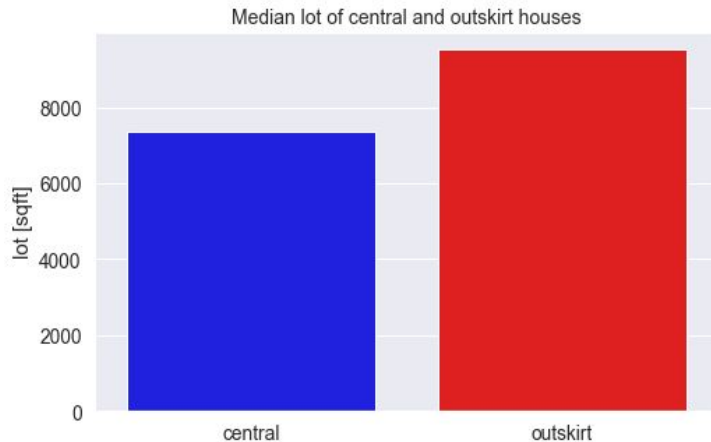
- Does the average price differ between outskirts and central houses?
  - **Answer:** No, house in outskirts and central areas are in the same price range
- Is the average grade lower in the outskirts?
  - **Answer:** No, outskirts houses are in general one grade higher



[minimalblogs.com](http://minimalblogs.com)

# What makes an average outskirts house?

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



# What makes an average outskirts house?

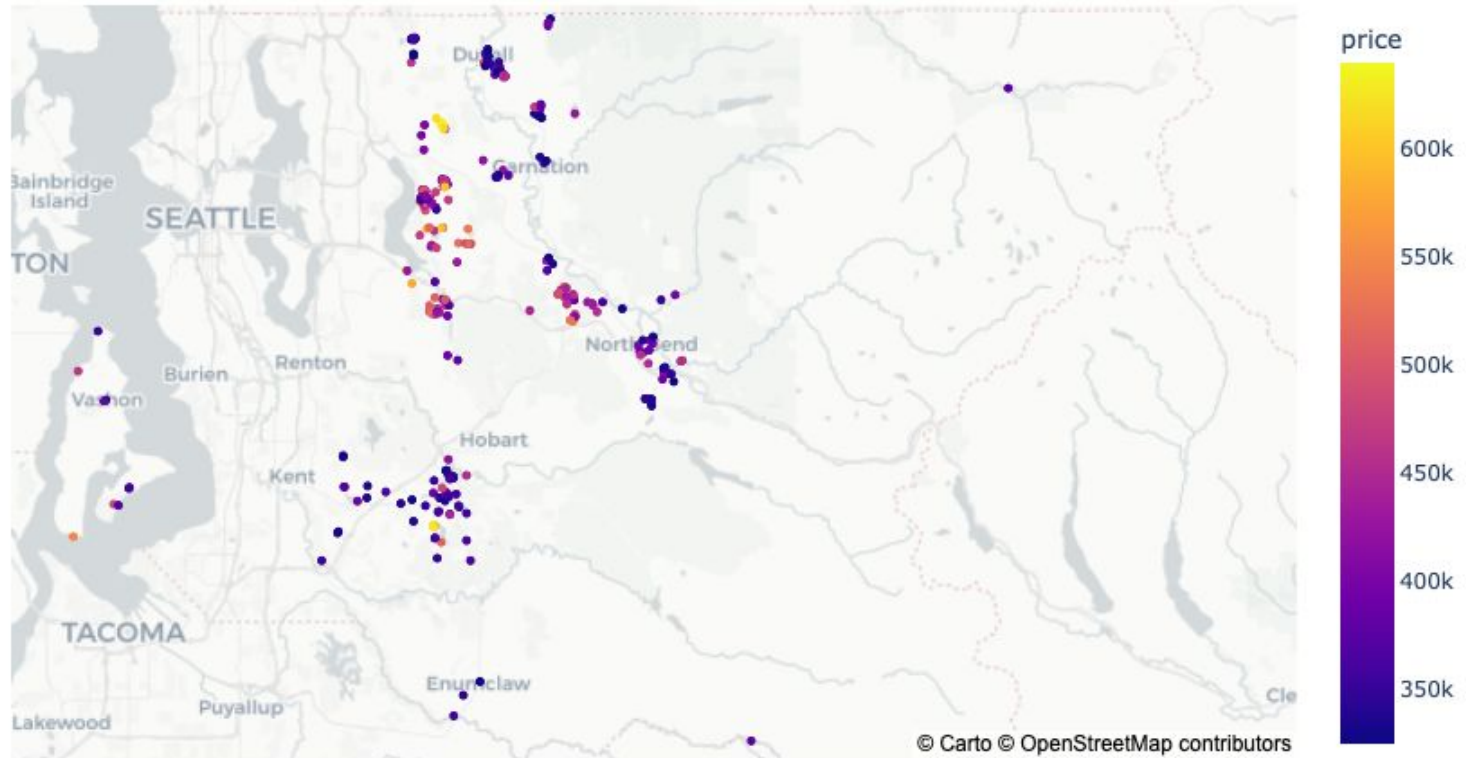
- Three recommendations to find an average outskirts house?
  - house prices of around 460000\$
  - grades of 8
  - lot of around 9500 sqft



[minimalblogs.com](http://minimalblogs.com)



# The average outskirts houses



# Conclusion and outlook

- could help Amy to find objects to buy/sell and also to hide in
- a different approach could be used for
  - distinguishing between outskirts 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](https://www.tripadvisor.com)