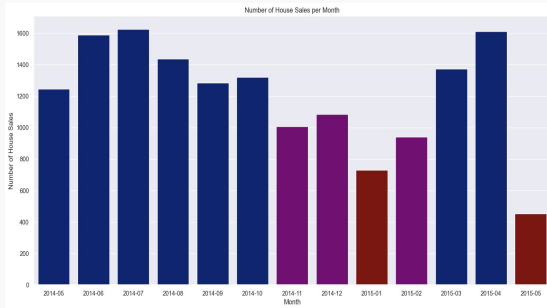


# EDA PROJECT



King County Real Estate Data  
by Dorothee Schmitt



# TABLE OF CONTENTS



**01** **THE TASK**  
Introducing the  
stakeholder and data

**02** **METHOD**  
Central vs Outskirts

**03** **INSIGHTS**  
Differences between the  
kinds of houses

**04** **STRATEGY**  
Recommendations on  
buying and selling for  
the stakeholder



# 01 THE TASK



## THE STAKEHOLDER AND HER INTERESTS

Amy Williams needs to buy inconspicuous houses in the outskirts to hide from the FBI. To finance that she is selling some central high end houses she owns.



## KING COUNTY REAL ESTATE DATA

Data about house sales in King County from mid 2014 to mid 2015.

# 02 METHOD: WHAT IS CENTRAL?

## POPULATION DENSITY AS PROXY

Not easy to get immediately from the government, but available by googling.

01

### ZIP CODES

Because US census data is (almost) zip code based.

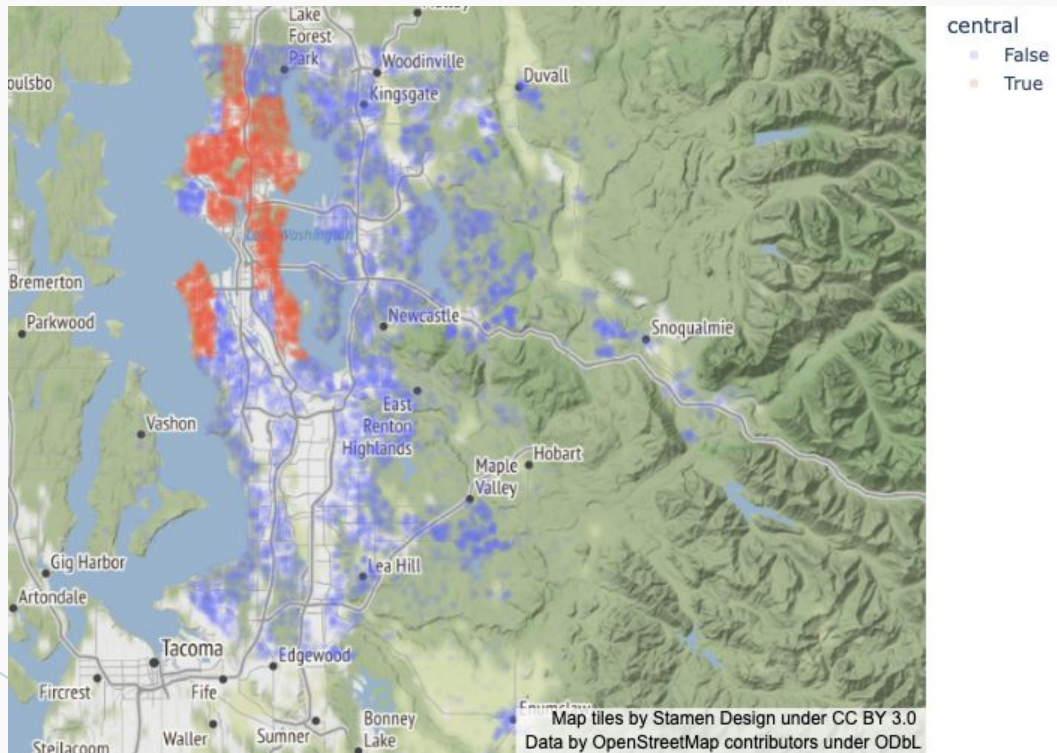
02

03

### CUTOFF

6000 people per square mile.  
That was a guess, is it usable?

## 02 METHOD: IS THE CUTOFF PLAUSIBLE?



### ● MAP OF SOLD HOUSES

The areas classified as central are mainly in and cover most of Seattle. Plausible.

### ● RATIO

The cutoff classifies 17 out of 70 zip codes as central. Plausible.

## 02 METHOD: ASSUMPTIONS AND MORE DATA SPLITTING



### SEASONALITY ASSUMPTION

Price and availability patterns by month are real. There is no way to check this with data for one year, but I want to analyze them, so they are hereby declared real.



### CENTRAL HOUSES TOP 10%

I interpreted the top 10% in the assignment as talking about price. This is somewhat fishy, since I'm also analyzing price. To check it's not too distorting, I checked that it doesn't result in wildly different numbers of sales in the category per month.

# 03 INSIGHT 1:

## LIVING AREA MORE EXPENSIVE IN CENTRE



- Prices rise faster with sq ft of living space in the centre.
- Especially particularly large houses in the outskirts are a lot less expensive.

## 03 INSIGHT 2:

# LOT AREA MORE EXPENSIVE IN CENTRE

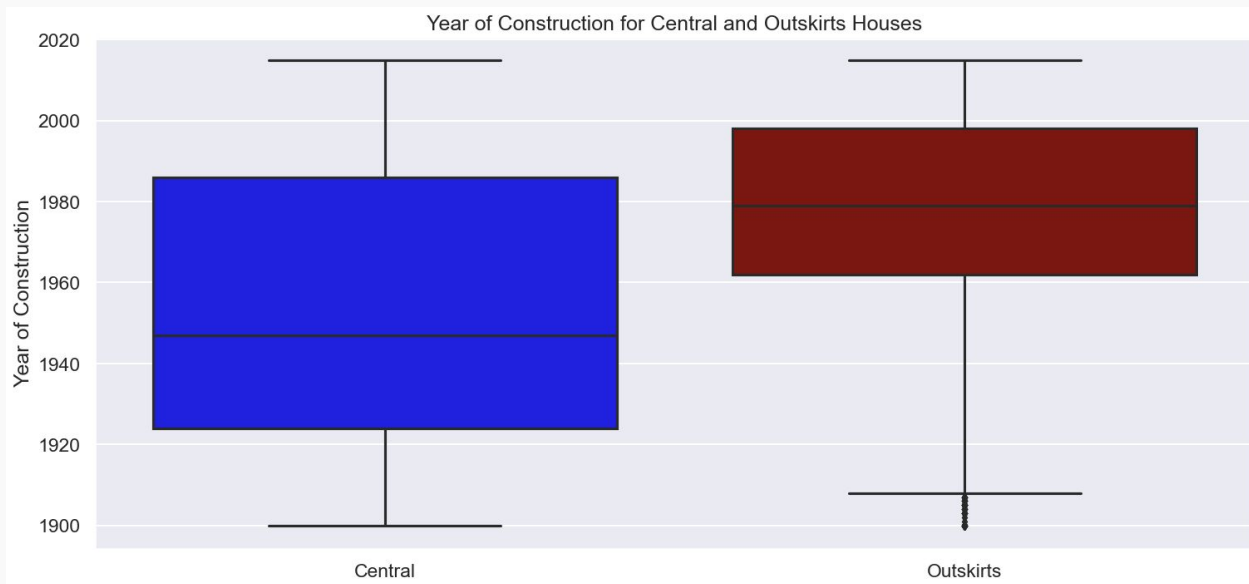


- Prices rise faster with lot sq ft in the centre.
- The outskirts have some houses with huge property for surprisingly low prices



# 03 INSIGHT 3:

## CENTRAL HOUSES TEND TO BE OLDER

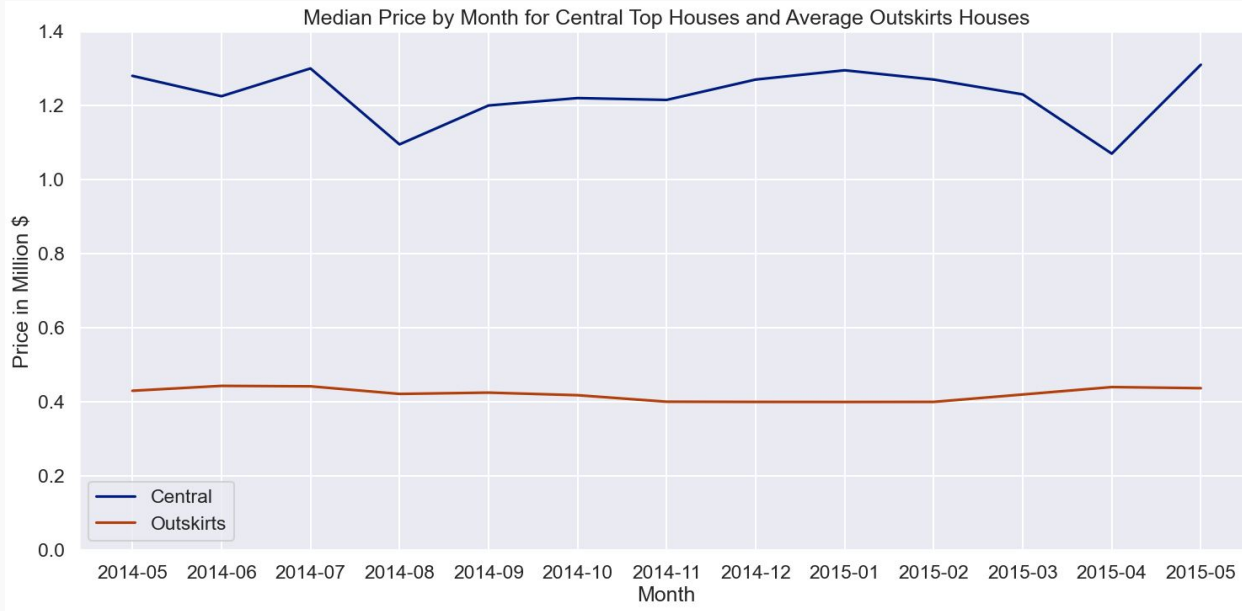


- Median age for outskirts houses is over 30 years younger.
- While there is some overlap, buying a central house as young as the median outskirts house should already get tricky.



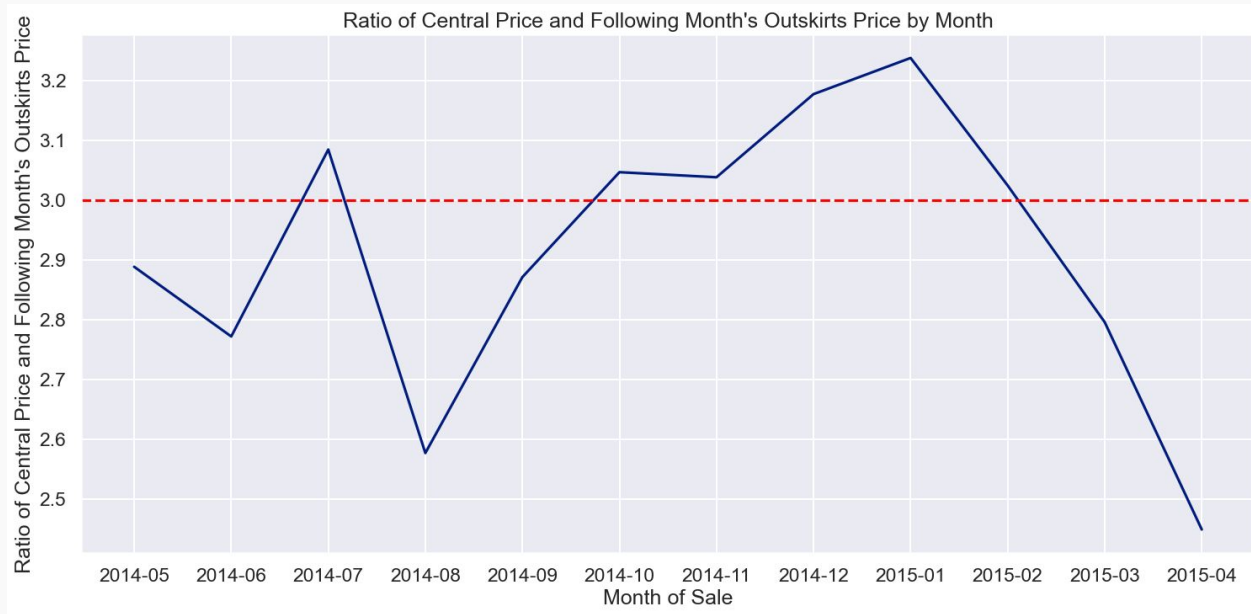
**GOOD FOR AMY!**  
**BIG, NEW HOUSES WITH**  
**PLENTY OF SPACE TO**  
**BURY THE BODIES!**

# 04 STRATEGY - RECOMMENDATION 1: FOCUS ON SELLING WELL, OVER BUYING WELL



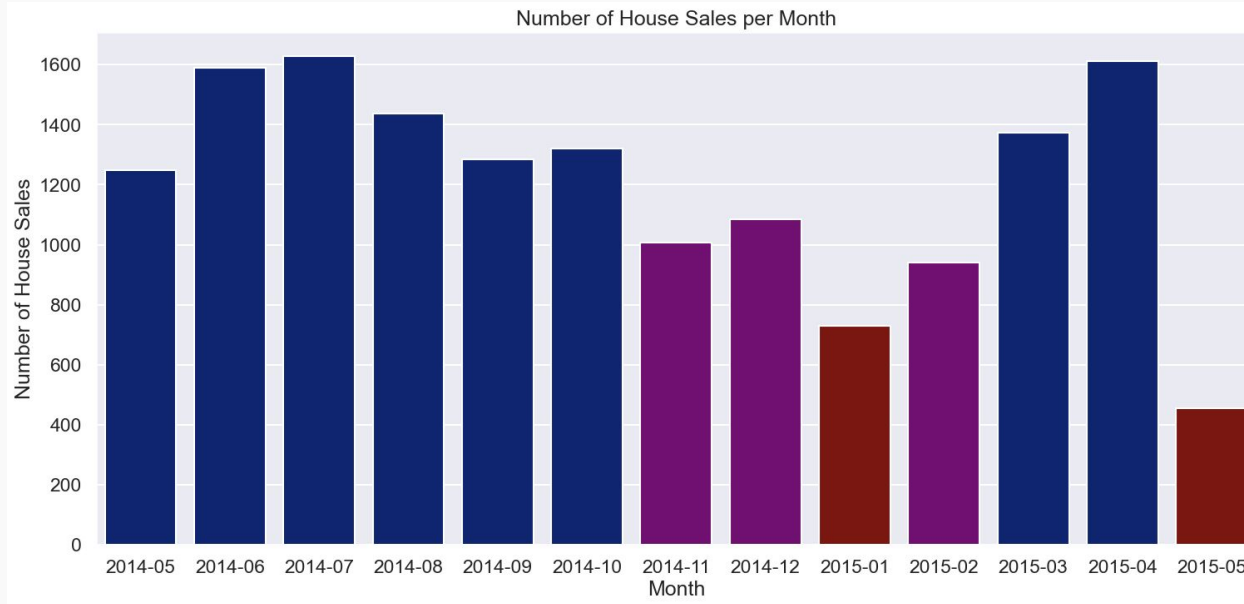
- Median price by month for high end central houses varies a lot more than for average outskirts houses
- For Amy that means that good timing for a sale will make a much larger difference than optimizing buying.

## 04 STRATEGY - RECOMMENDATION 2: SELL IN JULY AND BETWEEN OCT AND FEB



- The ratio between a month's high end central price and the following month's normal outskirts price is above three in quite a few months.
- For Amy that means she can buy three outskirts houses (and have money left over) after selling one central one, if she sells in the right months.

## 04 STRATEGY - RECOMMENDATION 3: PLAN MORE TIME FOR PURCHASES NOV-FEB



- Fewer houses get sold during the winter months. (And in March 15, maybe missing data.) Probably fewer houses are on the market then.
- For Amy that means finding suitable houses in the winter might be harder, she especially should not rely on finding houses in January.



# FURTHER RESEARCH

## RELATIONSHIP BETWEEN AVAILABILITY AND PRICE

In the winter months, fewer houses get sold and prices for high end houses are higher. This suggests there might be a relationship to investigate.

## SEASONALITY

Do the (here assumed to be real) patterns related to month of sale hold up in the next year?



# ADDITIONAL RESOURCES USED



- CSV with population density by ZCTA, based on 2010 US Census data from here:

<https://blog.splitwise.com/2013/09/18/the-2010-us-census-population-by-zip-code-totally-free/>

# THANKS

Do you have any questions?

[dkschmitt@web.de](mailto:dkschmitt@web.de)

<https://www.linkedin.com/in/dorothee-schmitt-394b83242/>

CREDITS: This presentation template was created by Slidesgo,  
including icons by **Flaticon**, infographics & images by **Freepik**