# Housing prices King County - EDA

Data Timerange: 05/2014 - 04/2015

# The Dataset

-Consists of houses sold in the Seattle area between 05/2014 - 04/2015

-Gives date and price of sold house and a brief/objective specification of the house

-Two parameters that assert the sold house a certain quality: 'condition' and 'grade'

Not displayed (excerpt):

'condition', 'grade', yr_built','yr_renovated',' zipcode','sqft_basemen t', etc….

| date ▲ ▼ | price ▼ | bedrooms ▼ | bathrooms▼ | sqft_living▼ | sqft_lot ▼ | floors ▼ | waterfront▼ |
|---|---|---|---|---|---|---|---|
| 10/22/2014 | 295000 | 3 | 2.25 | 1330 | 7200 | 1 | 0 |
| 10/22/2014 | 458000 | 3 | 2.25 | 2150 | 9900 | 1 | 0 |
| 10/22/2014 | 700000 | 4 | 1.75 | 2420 | 520 | 1.5 | 0 |
| 10/22/2014 | 294950 | 4 | 2.25 | 1740 | 9600 | 1 | 0 |
| 10/22/2014 | 571000 | 5 | 1.75 | 2280 | 43560 | 1 | 0 |
| 10/22/2014 | 163500 | 3 | 1 | 940 | 7200 | 1 | 0 |
| 10/22/2014 | 442000 | 2 | 1.75 | 860 | 5535 | 1 | 0 |
| 10/22/2014 | 379950 | 3 | 1 | 1270 | 6900 | 1 | 0 |
| 10/22/2014 | 360000 | 3 | 1 | 1660 | 9600 | 1 | 0 |

# The Target Variable

Notes:

- Our target variable
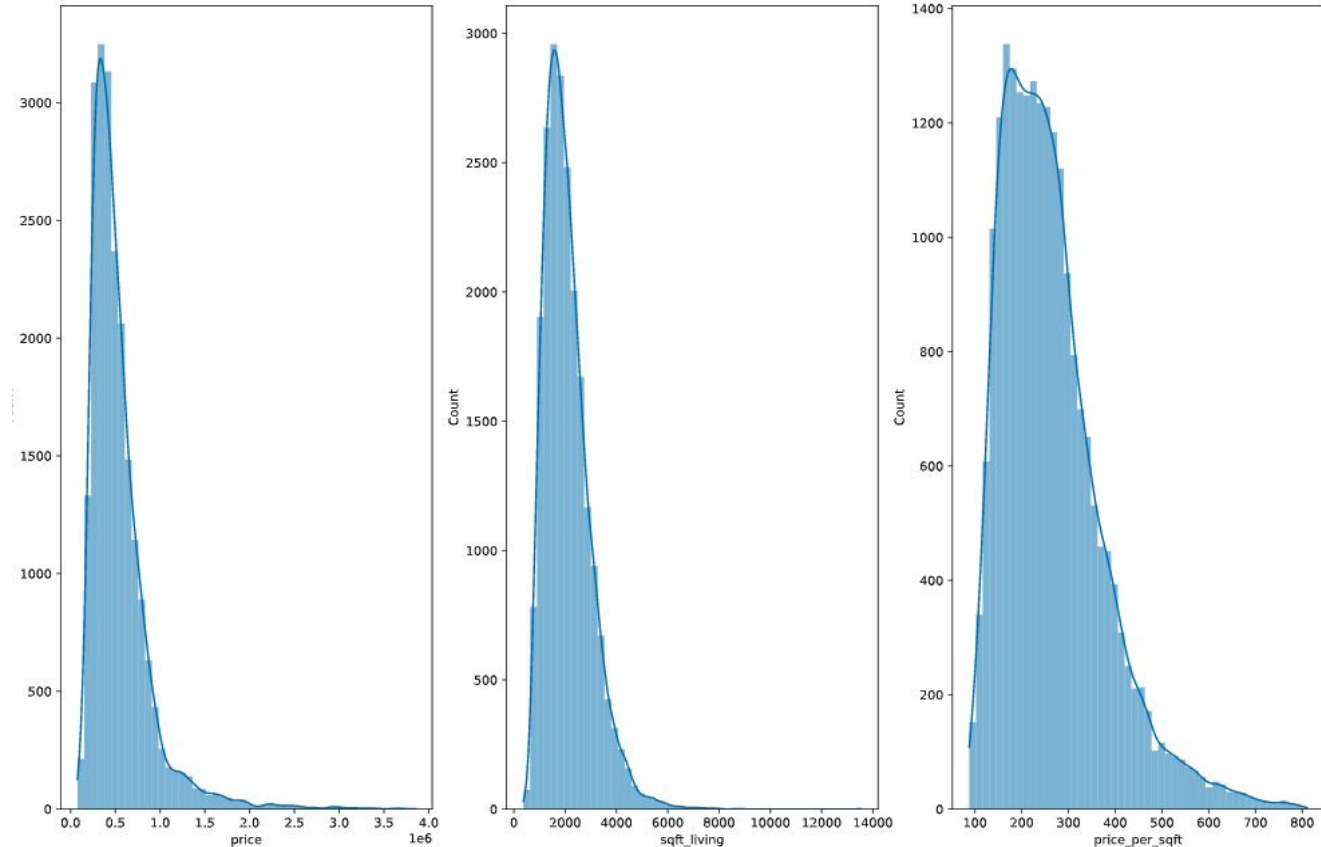  price = sqft_living *
      price_per_sqft
  is right skewed

- presumably originates from 'price_per_sqft' (ppsqft) that does <u>not</u> resemble a normal distribution

→ 'price' is a better suited target for a linear regression than 'price_per_sqft'
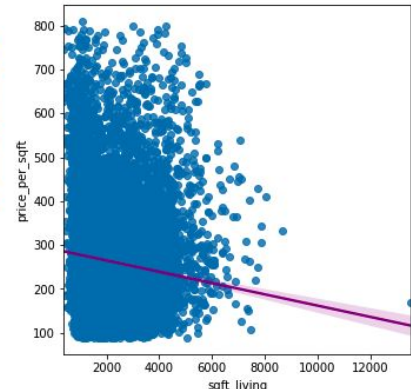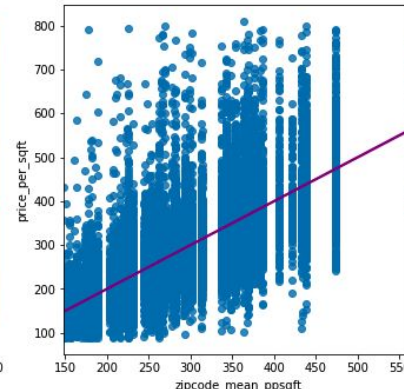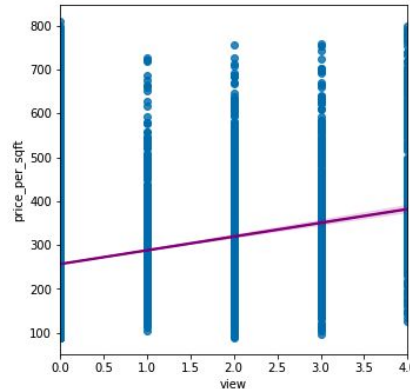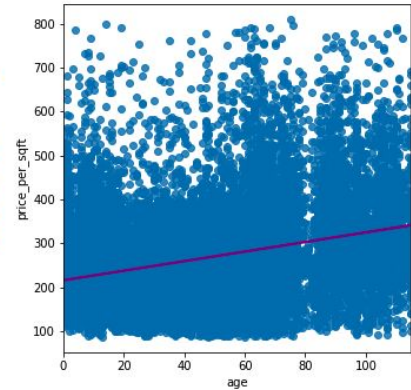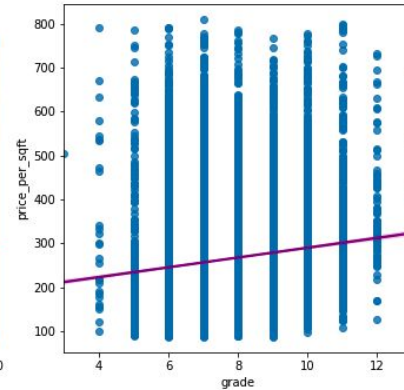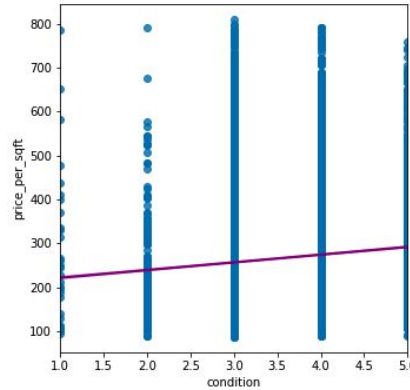


Distribution of Housing Prices/ Sqft of Sold Houses/ Price per Sqft of Sold Houses
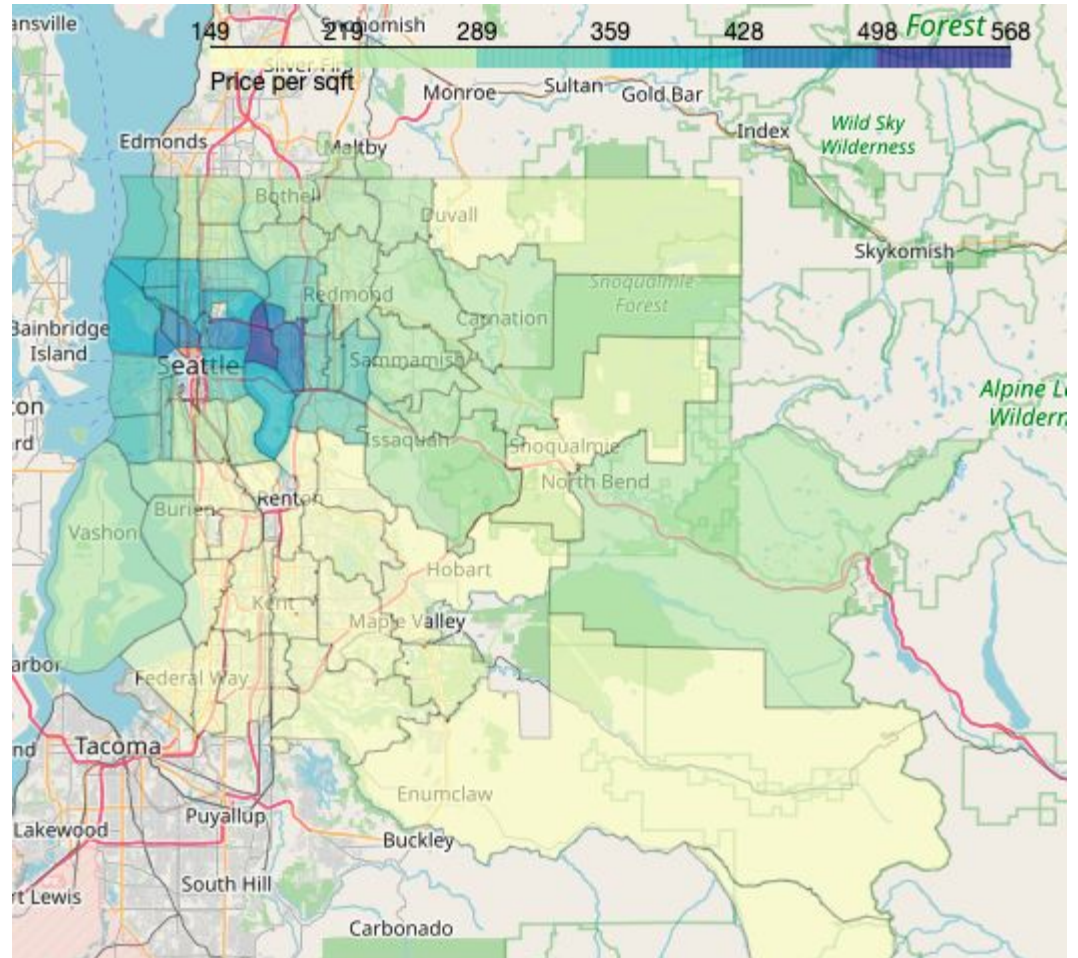
# Variables influencing Price per Sqft

Some insights:

- Older houses sell for more

- Larger house → lower ppsqft

- 'Condition' and 'grade' have a meaningful impact on the house price

- more 'views' = number of times the house has beeen viewed
            → higher ppsqft

- Price per sqft of neighborhood/zipcode heavily influences ppsqft

- Price per sqft of neighborhood/zipcode varies pretty strongly across the dataset

→ Very important for analysis

# Conclusions for Buyers/Investors

- The grade in the King County grading system and the condition of the house heavily influence the price per sqft

    → Check for houses in worse condition to lower budget


- 'Views' of houses increase the price

    → Try to find houses by different channels than major sellers/housing platforms (newspaper, facebook groups?)

-Looking for an investment?
There seems to be a
relationship between time
and ppsqft



Monthly price per sqft over time