EDA– Project
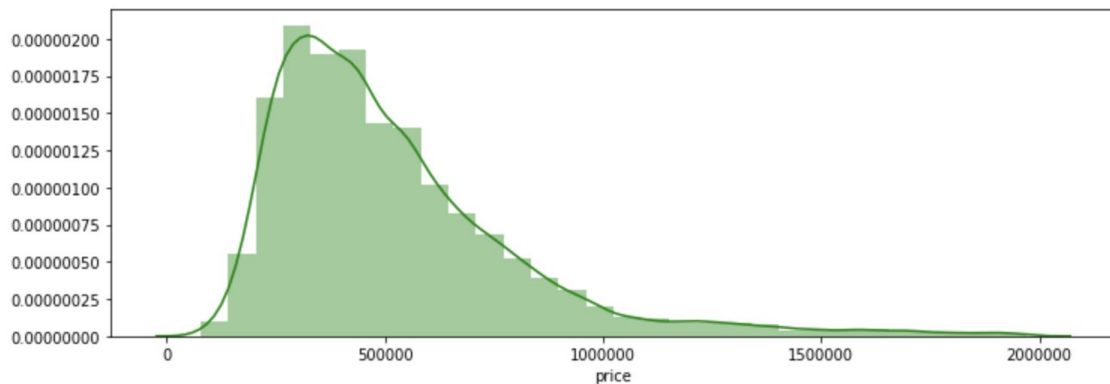# King County House sales

by Marvin Lorenzen
10.06.2020

The story of Peter's disaster

fixed values:

- he has now 450.000 $
- the house will be bought in july
- the house is in very good condition
- the house has been renovated in the last 10 years

changeable values:
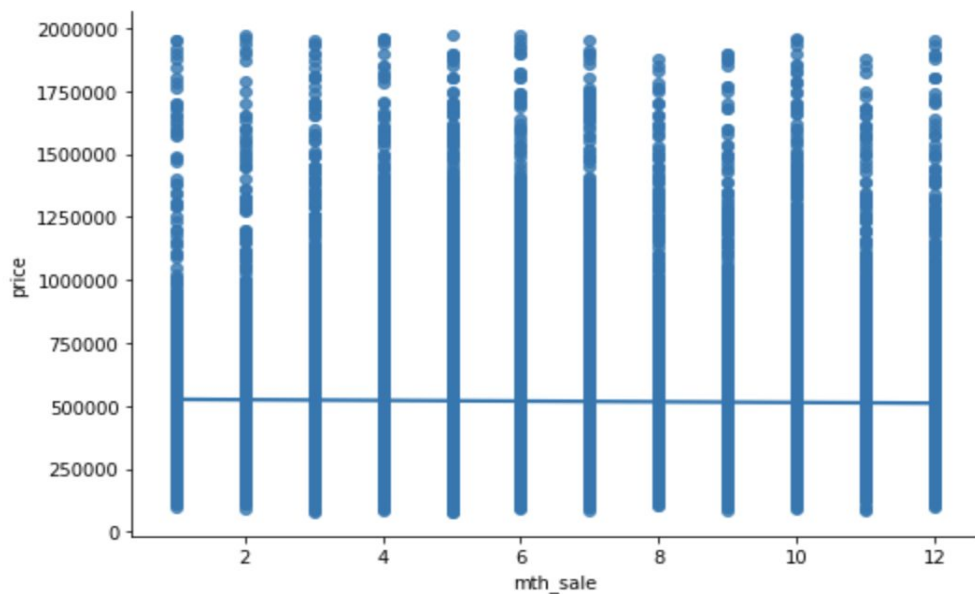
- grade of the house
- square footage of the house

- median 450.000 $
- the distribution of the price is skew
- use the logarithm to distribute it normally
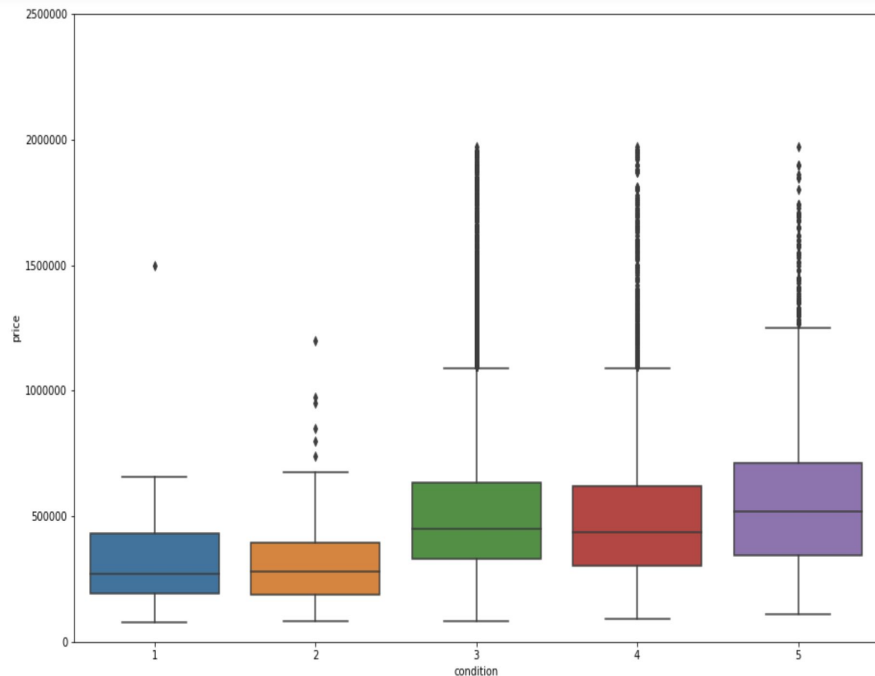
The month of buying the house

relationship between price and the month of buying the house

```
1  sns.lmplot("mth_sale", "price", data=df, aspect=1.5);
```
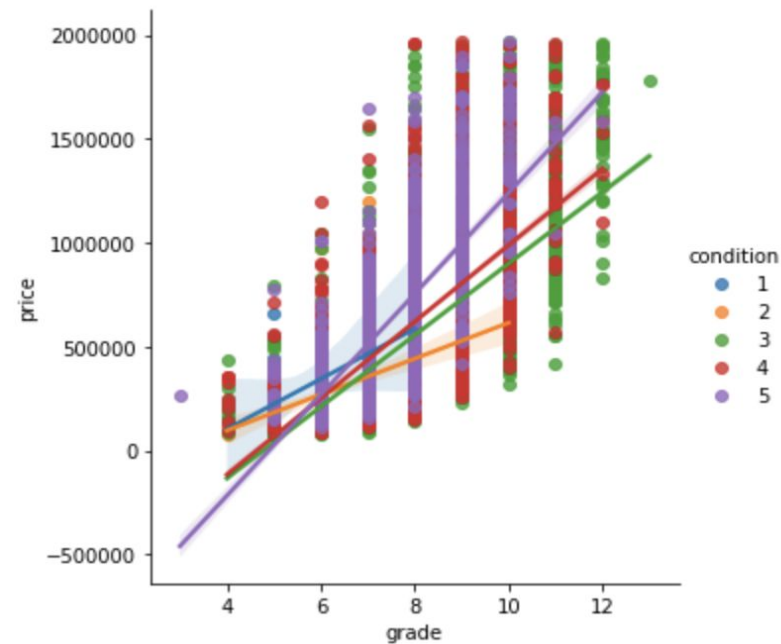
The condition of the house
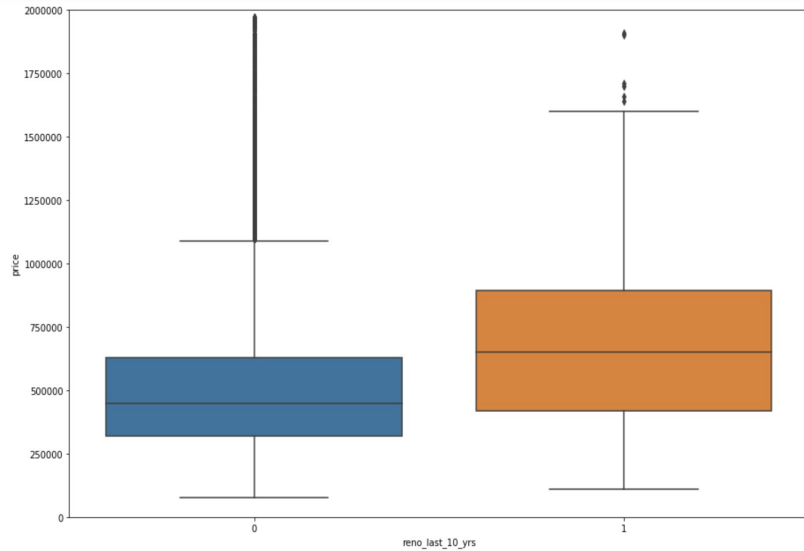
relationship between condition and price
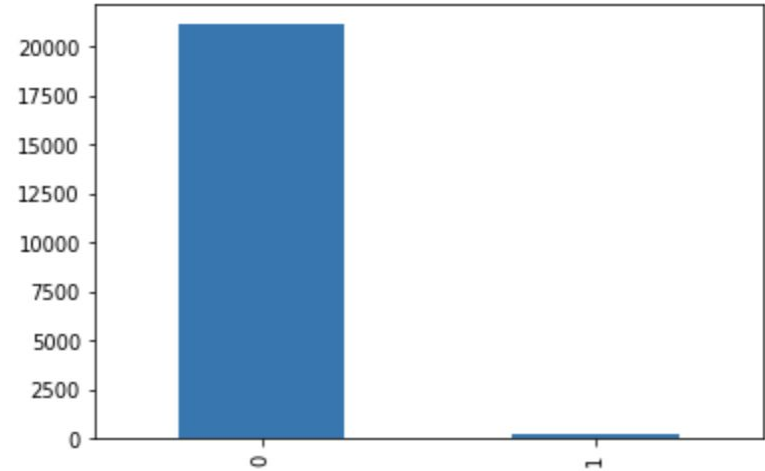
impact of the grade

impact of the price if the house has been renovated in the last 10 years
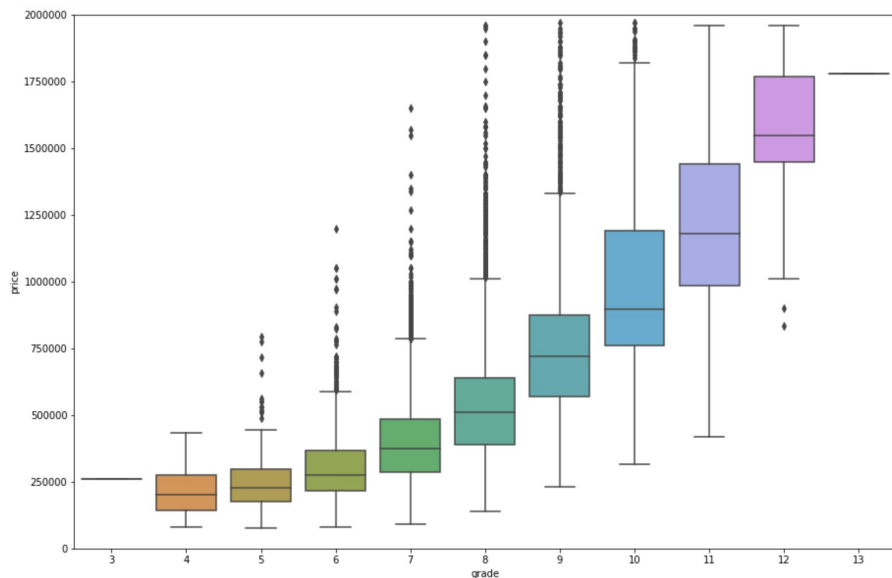
Number of houses that have been renovated in the last 10 years
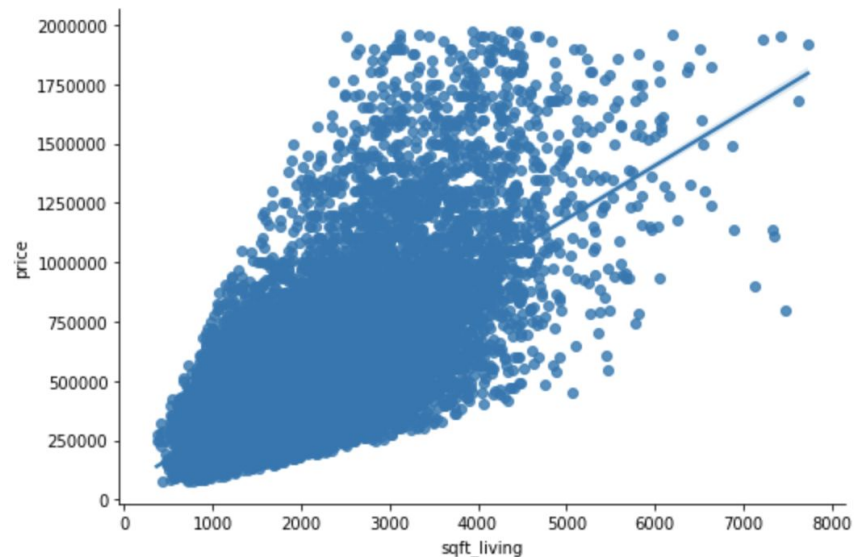

5

the greatest impact of the price

grade

square footage of the house

multivariate linear regression

```python
# all the variables that i put in my regression
variable_set = [
    "grade", "sqft_living", "condition", "reno_last_10_yrs", "mth_sale",
    "bedrooms", "view_yes", "sqft_basement2", "sqft_above", "bathrooms",
    "waterfront", "floors"
]
```

- the model shows that is possible to get a house that costs less than 450.000 \$ and meets the fixed values
- R2: 0,894
- Root Mean Squared Error: 186101,66

visualization
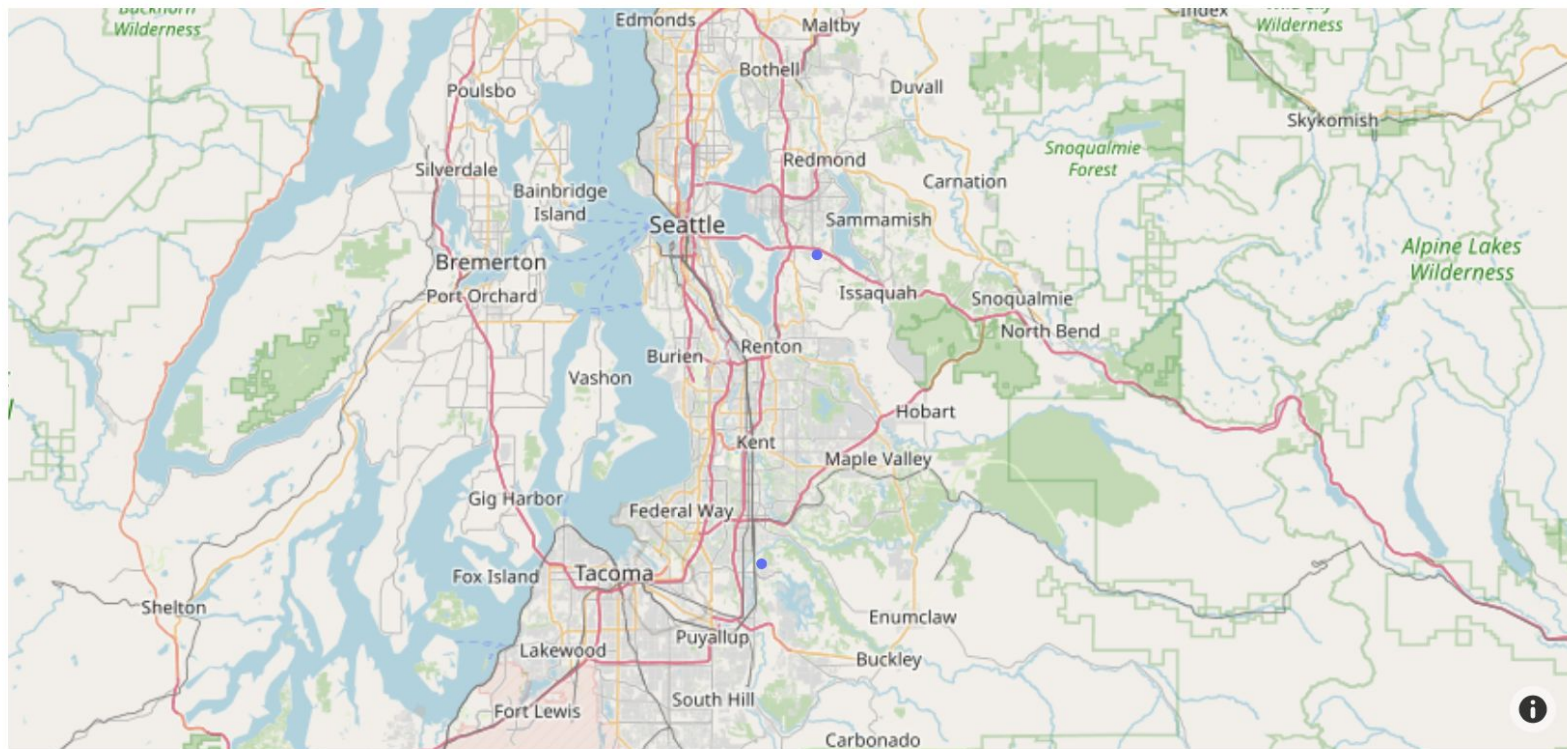
Thank you for attention
are there any questions ?