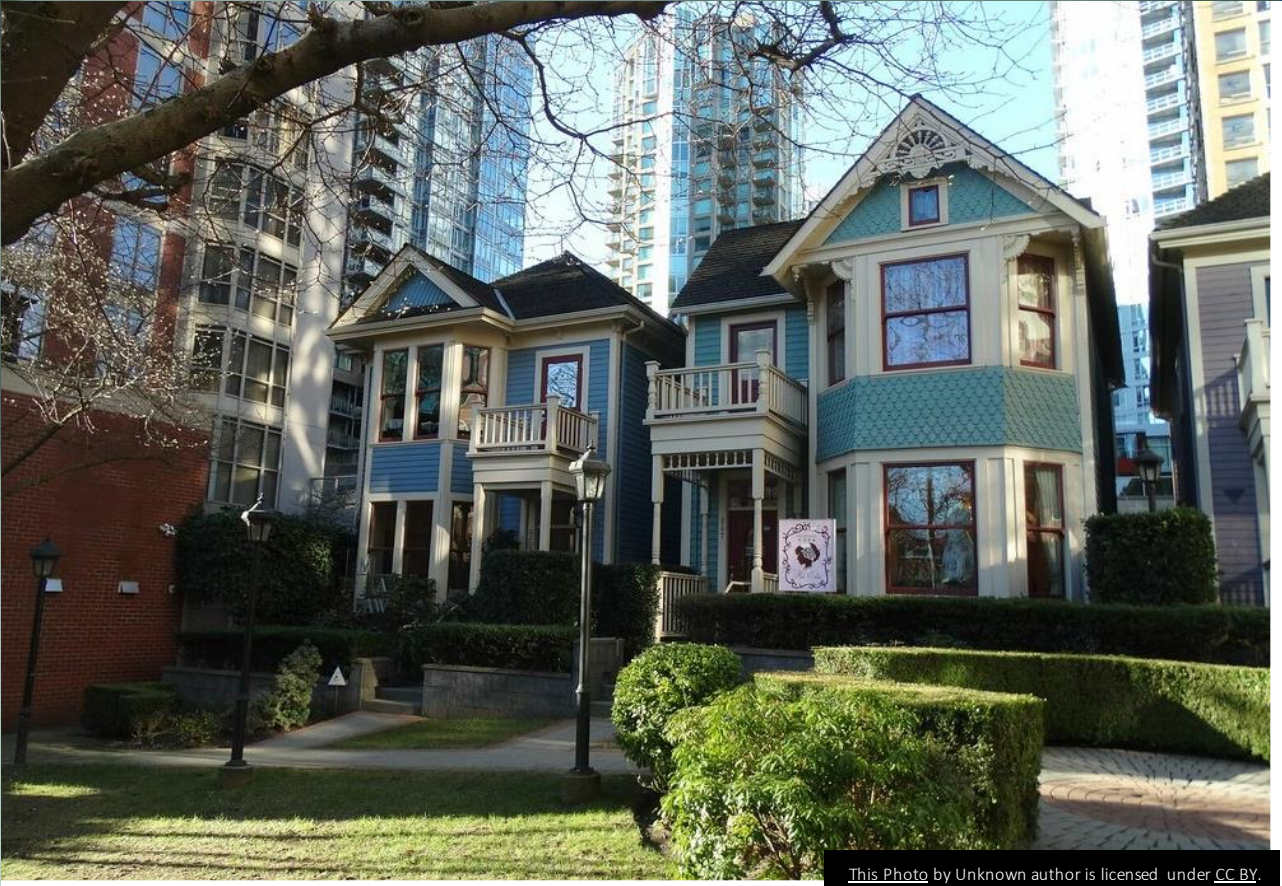
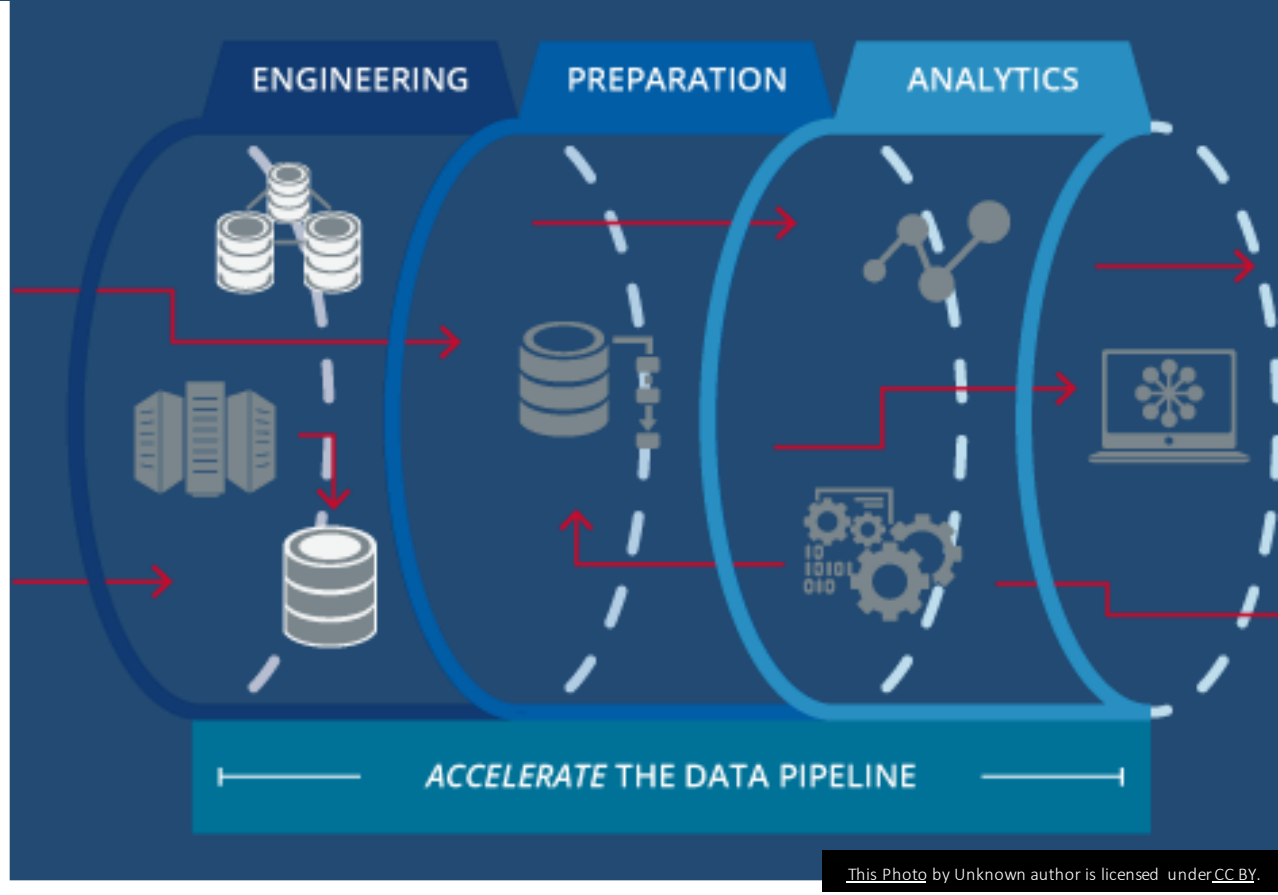
Kings County Housing Data

Inference Modeling for Housing Prices

This Photo by Unknown author is licensed under CC BY-SA.

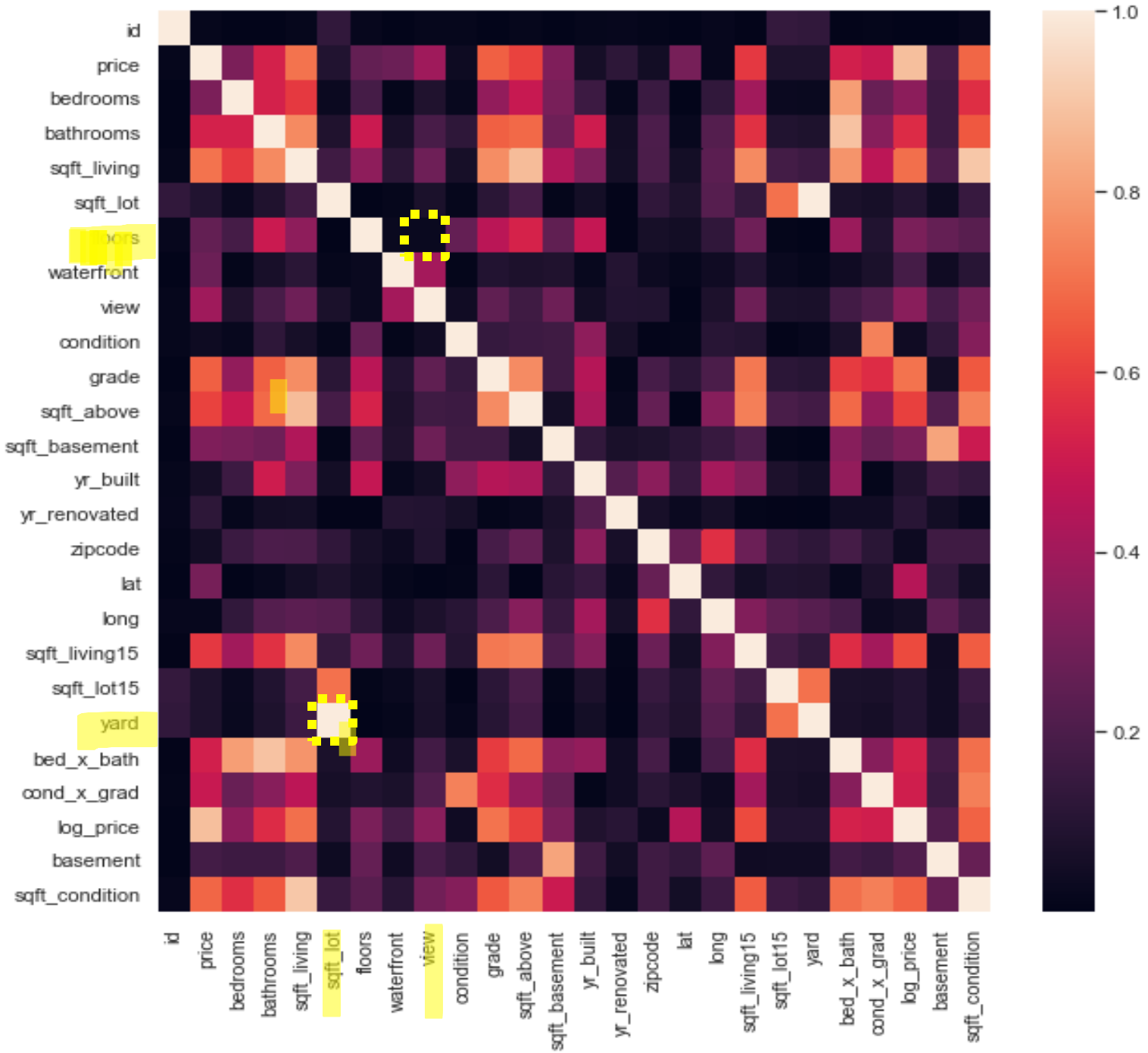


This Photo by Unknown author is licensed under CC BY.



This Photo by Unknown author is licensed under CC BY.

Data Analysis Process:



How do features interact with each other?

-Closer to 1: High correlation

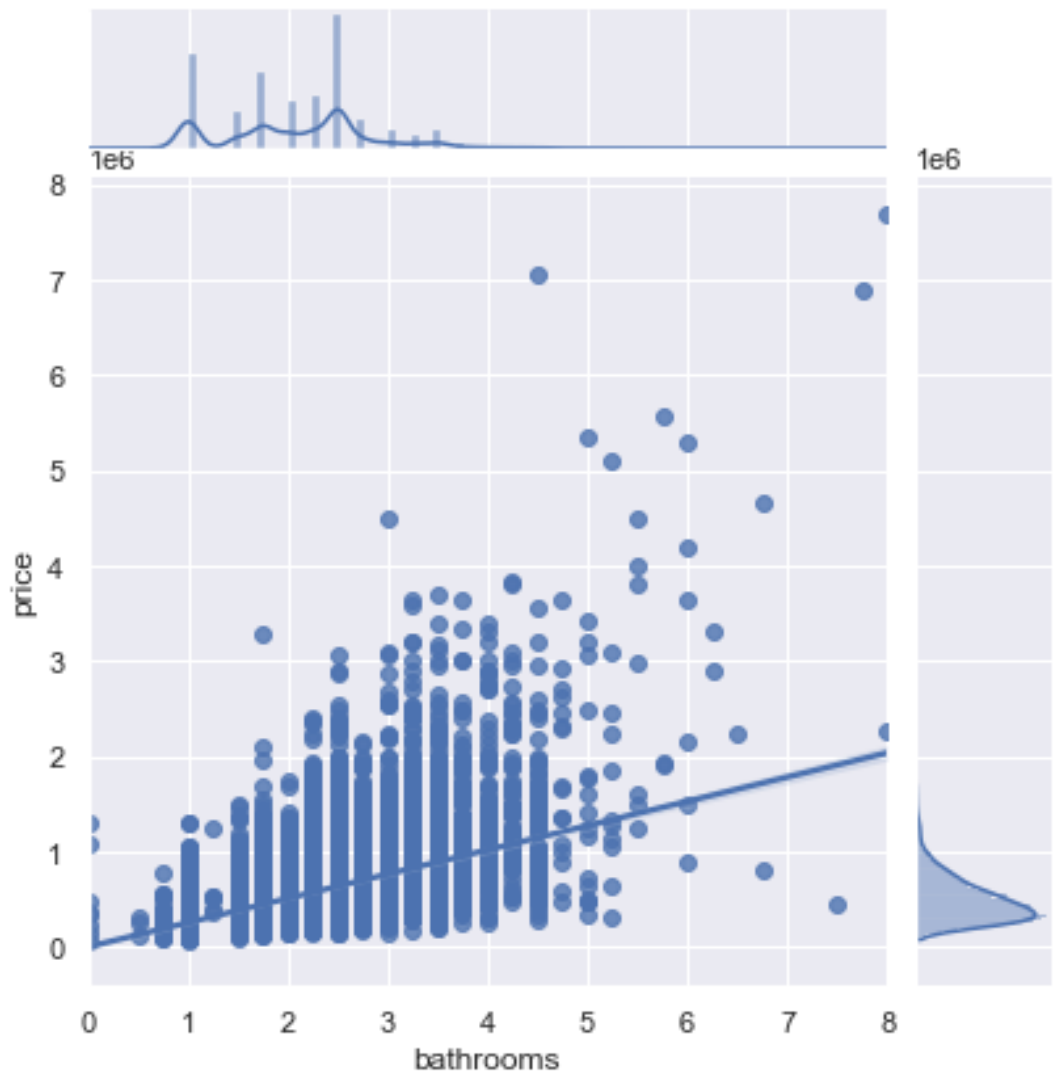
-Closer to 0: Low correlation

What impact do these features have on price?

price	1.000000
sqft_living	0.703540
grade	0.666215
sqft_above	0.607760
sqft_living15	0.585567
bathrooms	0.524335
view	0.395485
sqft_basement	0.323312
bedrooms	0.313775
lat	0.302394
waterfront	0.275779
floors	0.256985
yr_renovated	0.117770
sqft_lot	0.086803
sqft_lot15	0.081423
yard	0.080252
yr_built	0.056450
zipcode	0.048958
condition	0.036099
long	0.019327
id	0.015644

What features are strongly tied to the value of a house?

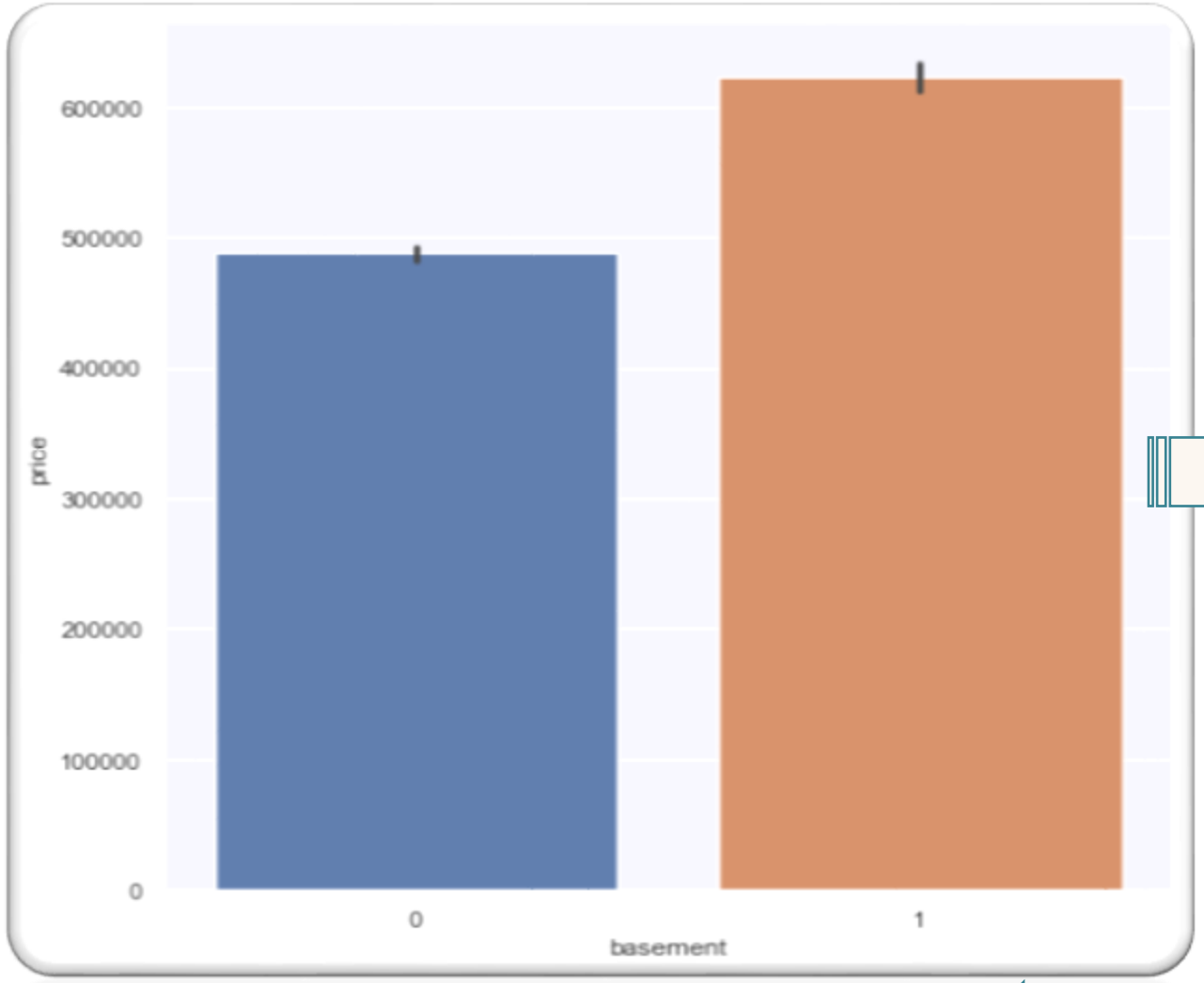
- Square foot living
- Bathrooms
- View
- Latitude



Bathrooms

- High correlation with price when modeled on its own.
- Scatterplot did not reflect this.
- Did not do well in the model as we hoped.

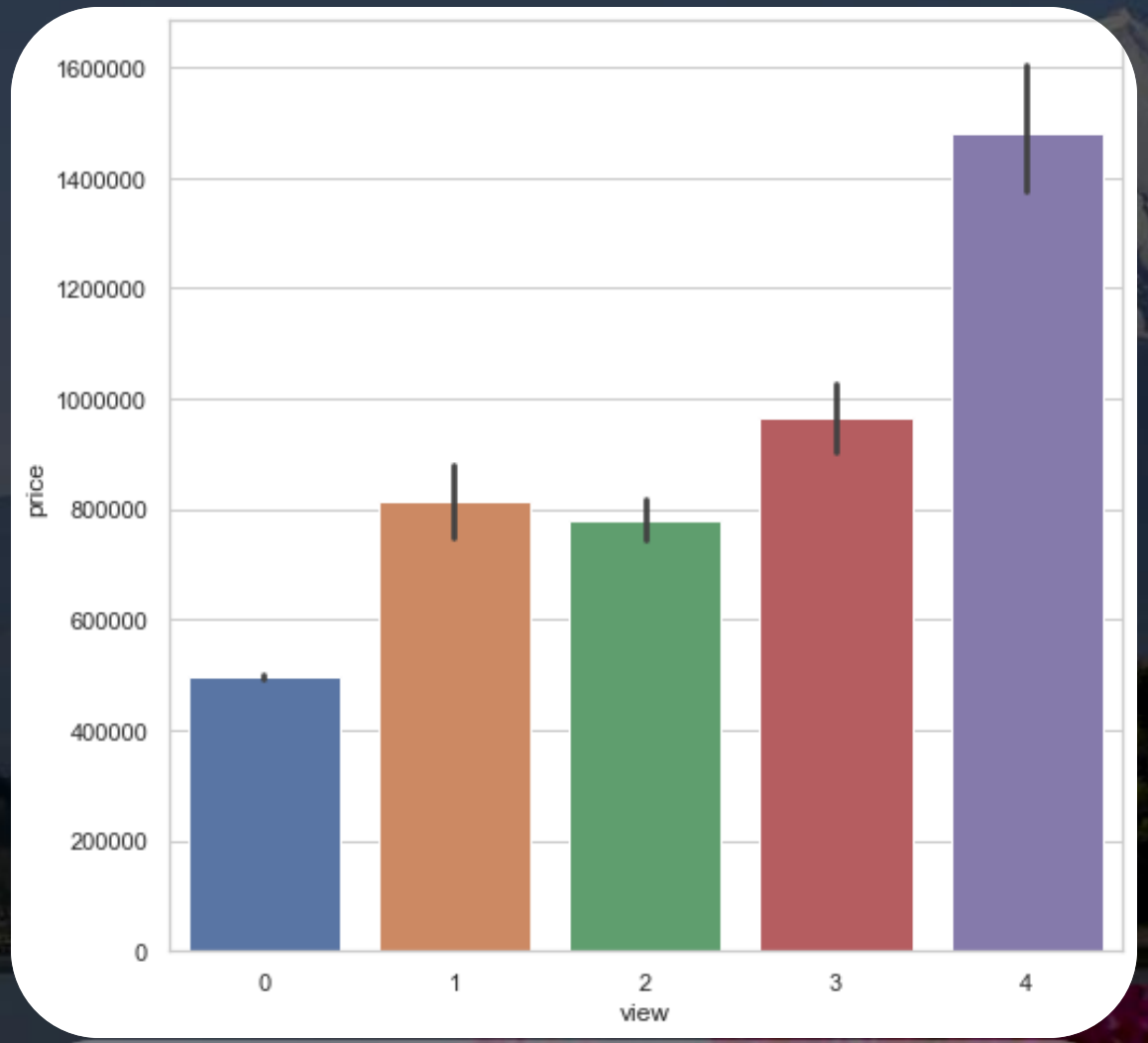
Basements



- 40% of our houses in the data set have basements.

-Does having a basement have any statistical impact on the price of a house?

YES



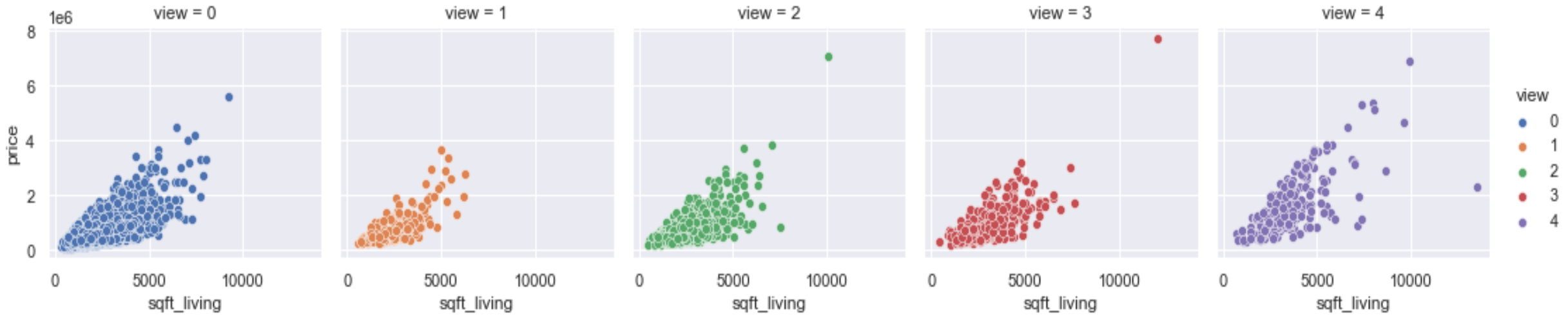
View

-What is the relationship to Price?

-We can see that better views bring higher home values by a marginal difference

This Photo by Unknown author is licensed under CC BY-NC-ND.

Examining View with respect to Square Feet:



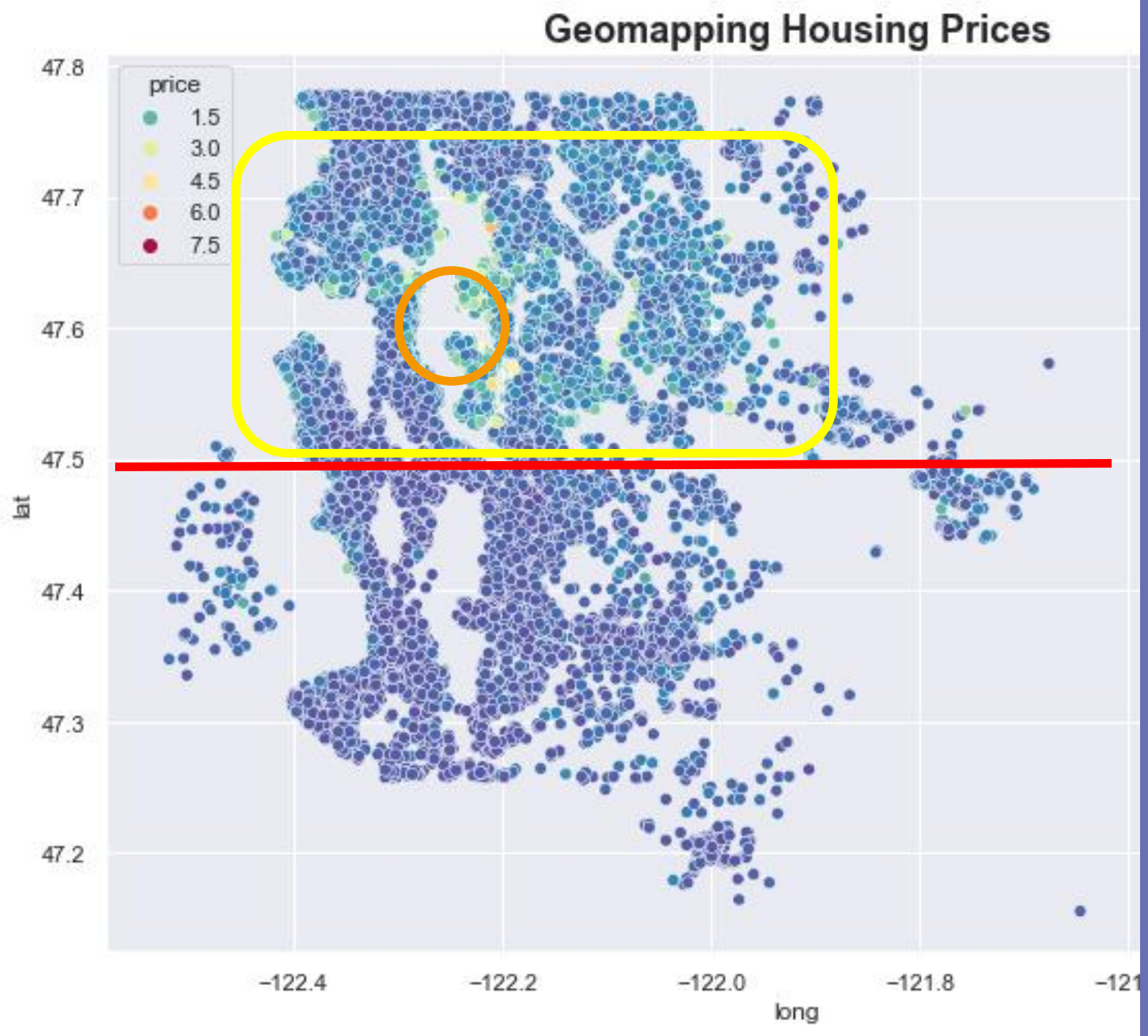
Square Foot Living

Created a visualization to better understand the relationship between Square Foot Living & Price

```
1 sm.formula.ols(formula = 'price ~ sqft_living',
```

OLS Regression Results

Dep. Variable:	price	**R-squared:**	0.495
Model:	OLS	**Adj. R-squared:**	0.495
Method:	Least Squares	**F-statistic:**	1.694e+04
Date:	Tue, 01 Jun 2021	**Prob (F-statistic):**	0.00



Location:

- Latitude & Longitude coordinates indicated there is a relationship to home value

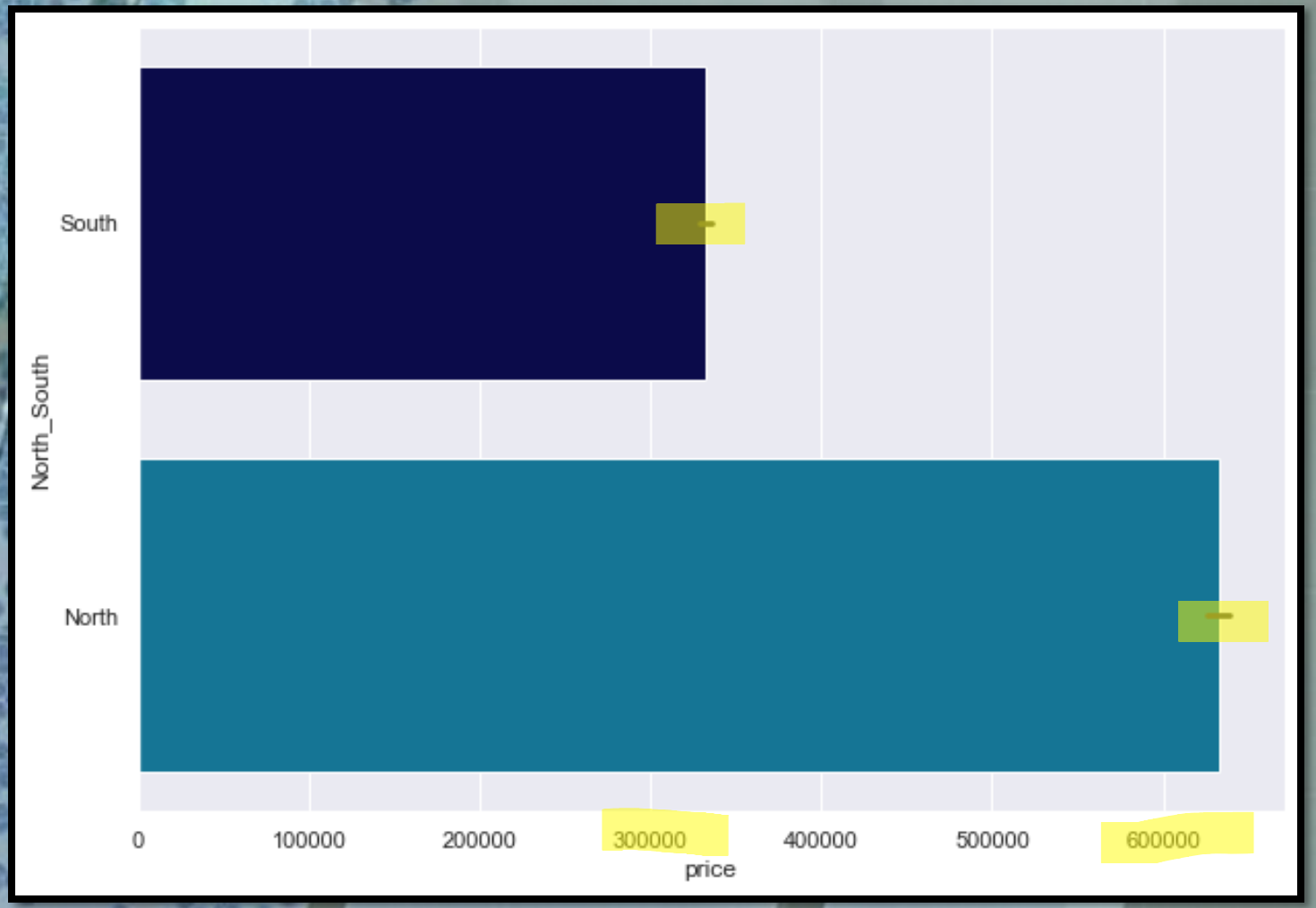
-Distinction around 47.5 degrees latitude

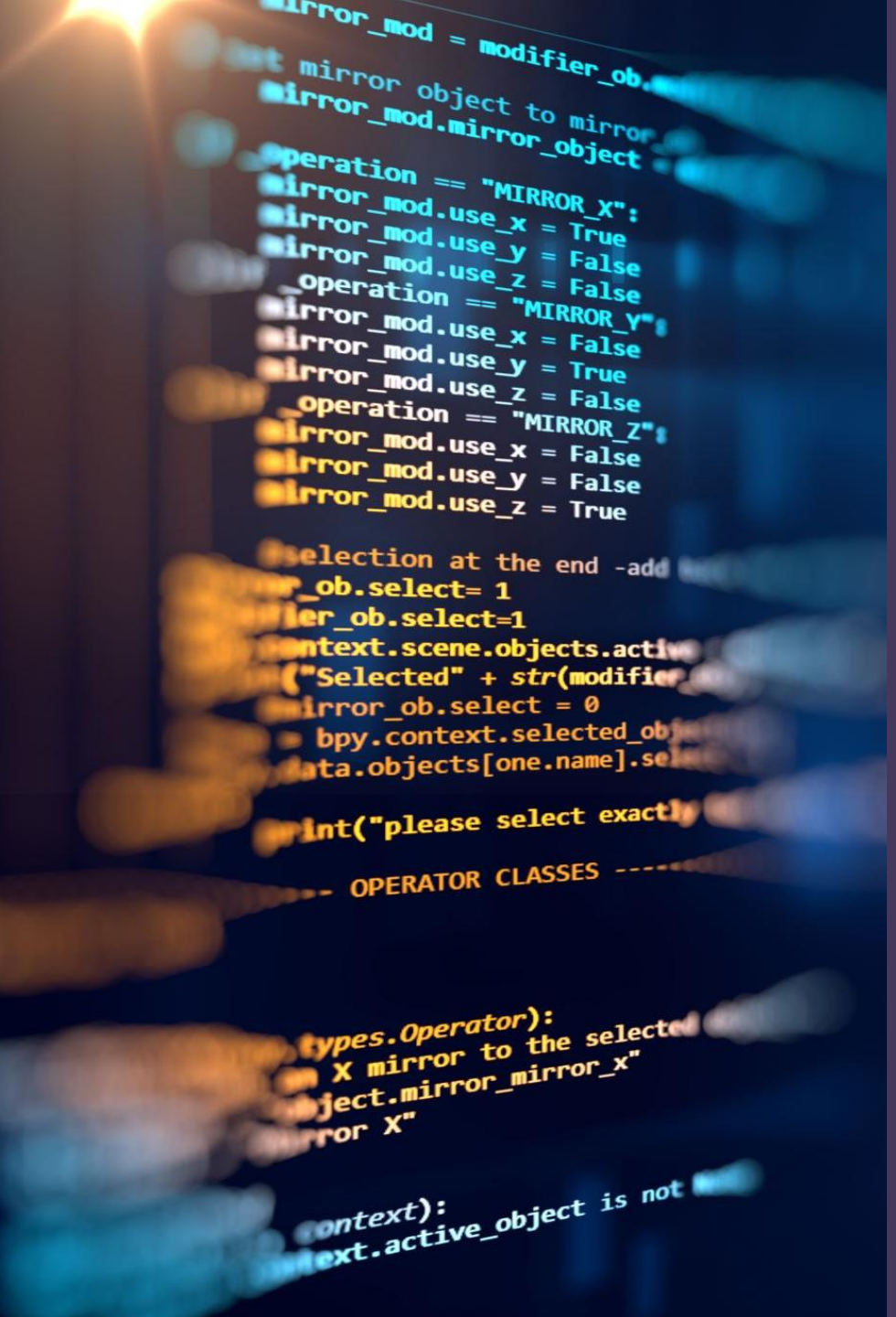
Latitude: Dividing by Region

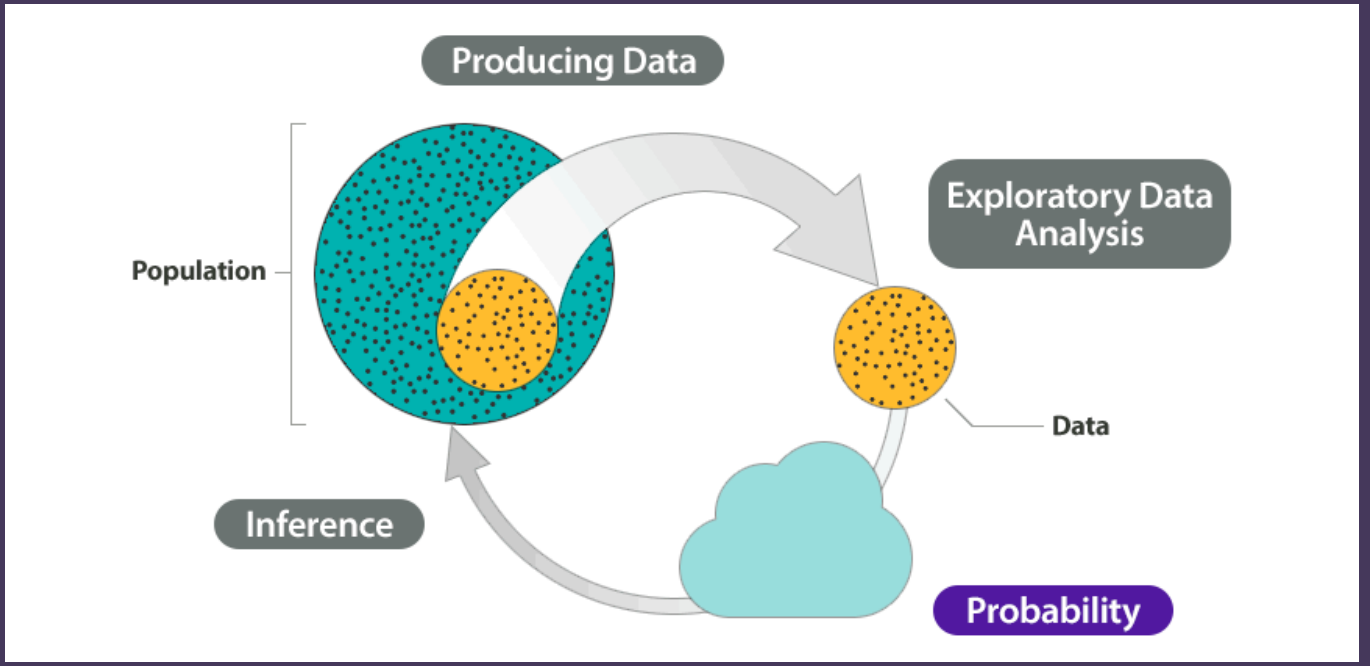
South< 47.5 Latitude

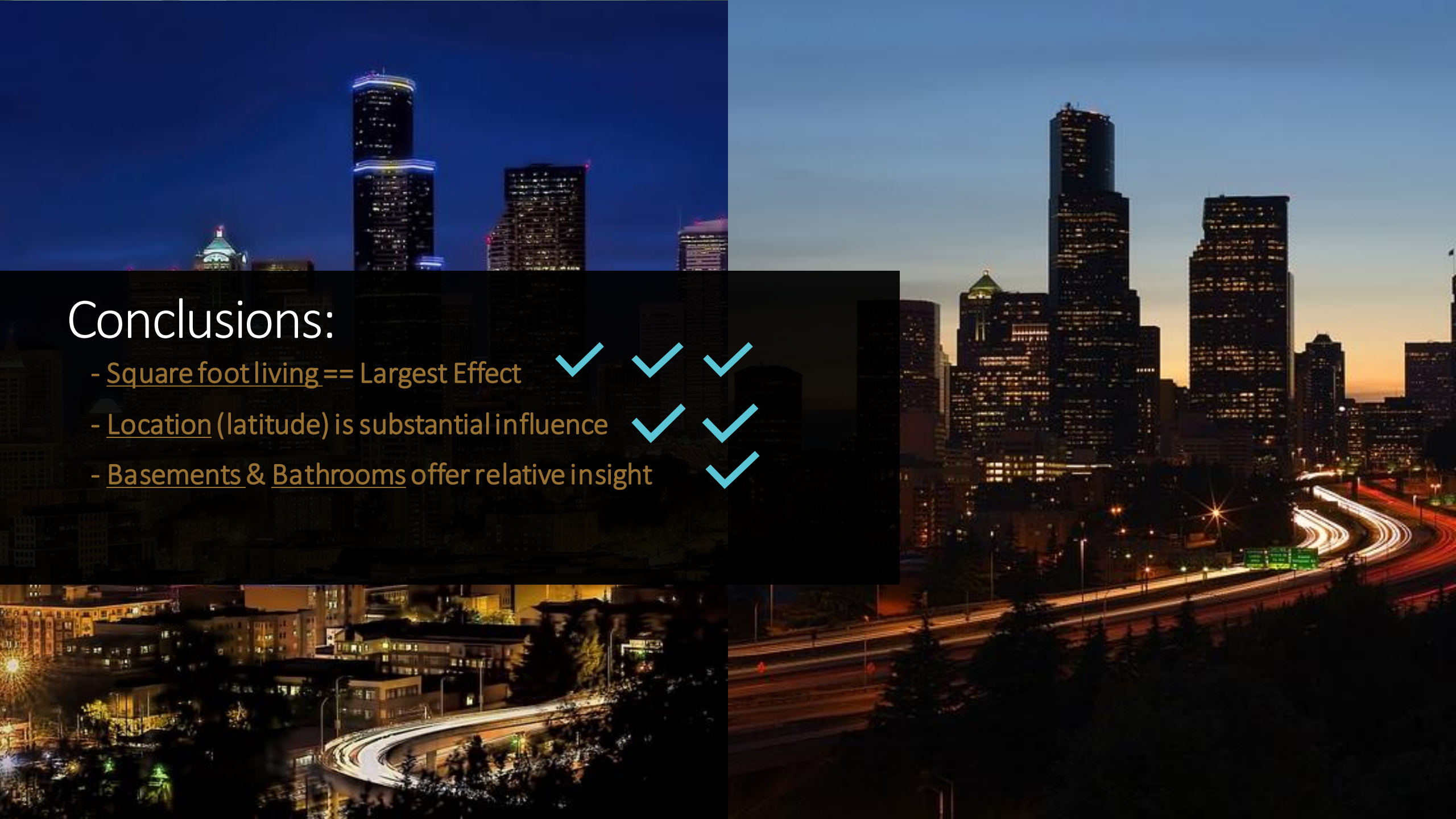
North > 47.5 Latitude

Nearly 2x the impact on housing prices




mirror_mod = modifier_ob.
mirror object to mirror
mirror_mod.mirror_object
operation == "MIRROR_X":
irror_mod.use_x = True
irror_mod.use_y = False
irror_mod.use_z = False
operation == "MIRROR_Y":
irror_mod.use_x = False
irror_mod.use_y = True
irror_mod.use_z = False
operation == "MIRROR_Z":
irror_mod.use_x = False
irror_mod.use_y = False
irror_mod.use_z = True
election at the end -add
ob.select= 1
er_ob.select=1
ntext.scene.objects.activ
("Selected" + str(modifier
irror_ob.select = 0
bpy.context.selected_ob
ata.objects[one.name].se
int("please select exactly
OPERATOR CLASSES
types.Operator):
X mirror to the selected
ject.mirror_mirror_x"
ror X"
context):
ext.active_object is not


Producing Data
Population
Exploratory Data Analysis
Data
Inference
Probability



Conclusions:

- Square foot living == Largest Effect
- Location (latitude) is substantial influence
- Basements & Bathrooms offer relative insight