

---

---

# Kings County Housing Project

Arnold Chovu Kalage • 01.10.2022

---

# Overview

## Kings County Housing

- This is data that describes the houses in the market of King County and their prices
- Proper understanding of the data enhances stakeholder involvement in the sector

---

---

# Data Understanding

<ul style="list-style-type: none"><li>• <b>id</b> - unique identifier for a house</li></ul>	<ul style="list-style-type: none"><li>• <b>sqft_lotsquare</b> - footage of the lot</li></ul>
<ul style="list-style-type: none"><li>• <b>dateDate</b> - house was sold</li></ul>	<ul style="list-style-type: none"><li>• <b>floorsTotal</b> - floors (levels) in house</li></ul>
<ul style="list-style-type: none"><li>• <b>pricePrice</b> - is prediction target</li></ul>	<ul style="list-style-type: none"><li>• <b>waterfront</b> - House which has a view to a waterfront</li></ul>
<ul style="list-style-type: none"><li>• <b>bedroomsNumber</b> - of Bedrooms/House</li></ul>	<ul style="list-style-type: none"><li>• <b>view</b> - Has been viewed</li></ul>
<ul style="list-style-type: none"><li>• <b>bathroomsNumber</b> - of bathrooms/bedrooms</li></ul>	<ul style="list-style-type: none"><li>• <b>condition</b> - How good the condition is ( Overall )</li></ul>
<ul style="list-style-type: none"><li>• <b>sqft_livingsquare</b> - footage of the home</li></ul>	<ul style="list-style-type: none"><li>• <b>grade</b> - overall grade given to the housing unit, based on King County grading system</li></ul>

---

---

# Data Understanding

<ul style="list-style-type: none"><li>• <b>sqft_above</b> - square footage of house apart from basement</li></ul>	<ul style="list-style-type: none"><li>• <b>lat</b> - Latitude coordinate</li></ul>
<ul style="list-style-type: none"><li>• <b>sqft_basement</b> - square footage of the basement</li></ul>	<ul style="list-style-type: none"><li>• <b>long</b> - Longitude coordinate</li></ul>
<ul style="list-style-type: none"><li>• <b>yr_built</b> - Built Year</li></ul>	<ul style="list-style-type: none"><li>• <b>sqft_living15</b> - The square footage of interior housing living space for the nearest 15 neighbors</li></ul>
<ul style="list-style-type: none"><li>• <b>yr_renovated</b> - Year when house was renovated</li></ul>	<ul style="list-style-type: none"><li>• <b>sqft_lot15</b> - The square footage of the land lots of the nearest 15 neighbors</li></ul>
<ul style="list-style-type: none"><li>• <b>zipcode</b> - zip</li></ul>	

---

---

---

# Business Need

- Individuals selling houses need to understand the market
  - Different features of the houses affect their market value
  - Stakeholders need to understand the relationship between the features
  - What affects the price of the houses
-

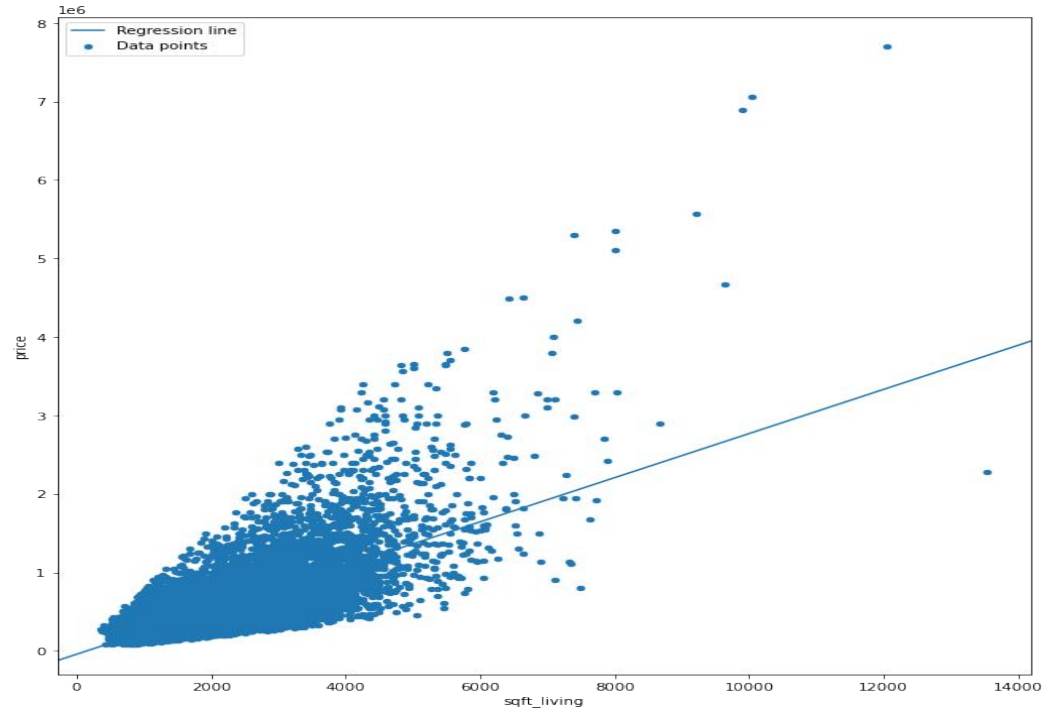
---

# Modelling

---

---

# Simple Linear Model(Base)



---

# Simple Linear Model

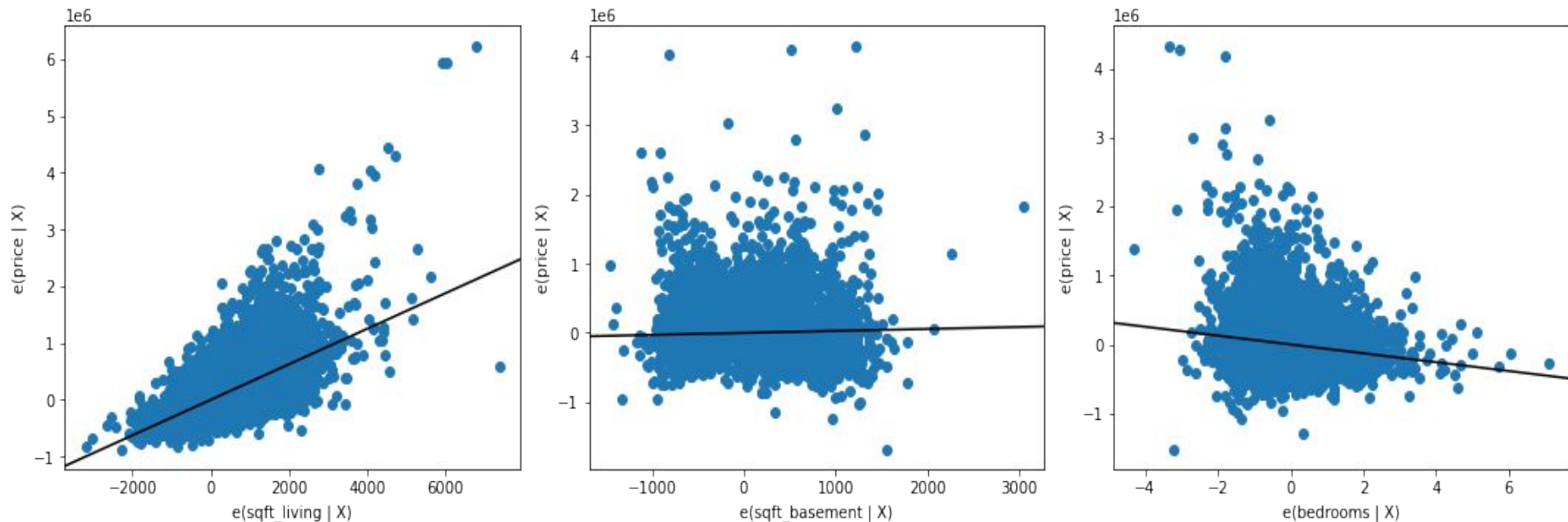
- The area of the living are showed a linear relationship with price
  - The linear relationship created the base model (previous slide)
  - The simple linear model can be used to predict the price on its own
-



---

# Multilinear Regression Model

Partial Regression Plot



---

# Multilinear Regression Model

- The model incorporated the area of the living area, the area of the basement and the number of bedrooms
  - The area of the living area is the best predictor for house prices
  - The area of the basement does not heavily impact the price of the houses.
  - As the number of bedrooms increase, the price goes down
-

---

# Recommendations

- Real estate dealers should focus on the size of the living area
  - The size of the living area is a predictor of the price
  - Real estate dealers should not focus on the number of bedrooms
-

---

# Limitations of the Model

- The model can work only on data from Kings County
  - The difference in pricing by region limits the model
  - Some of the features of the data do not satisfy regression assumptions
  - To use the model, the data requires to go through all the steps outlined in this analysis
-