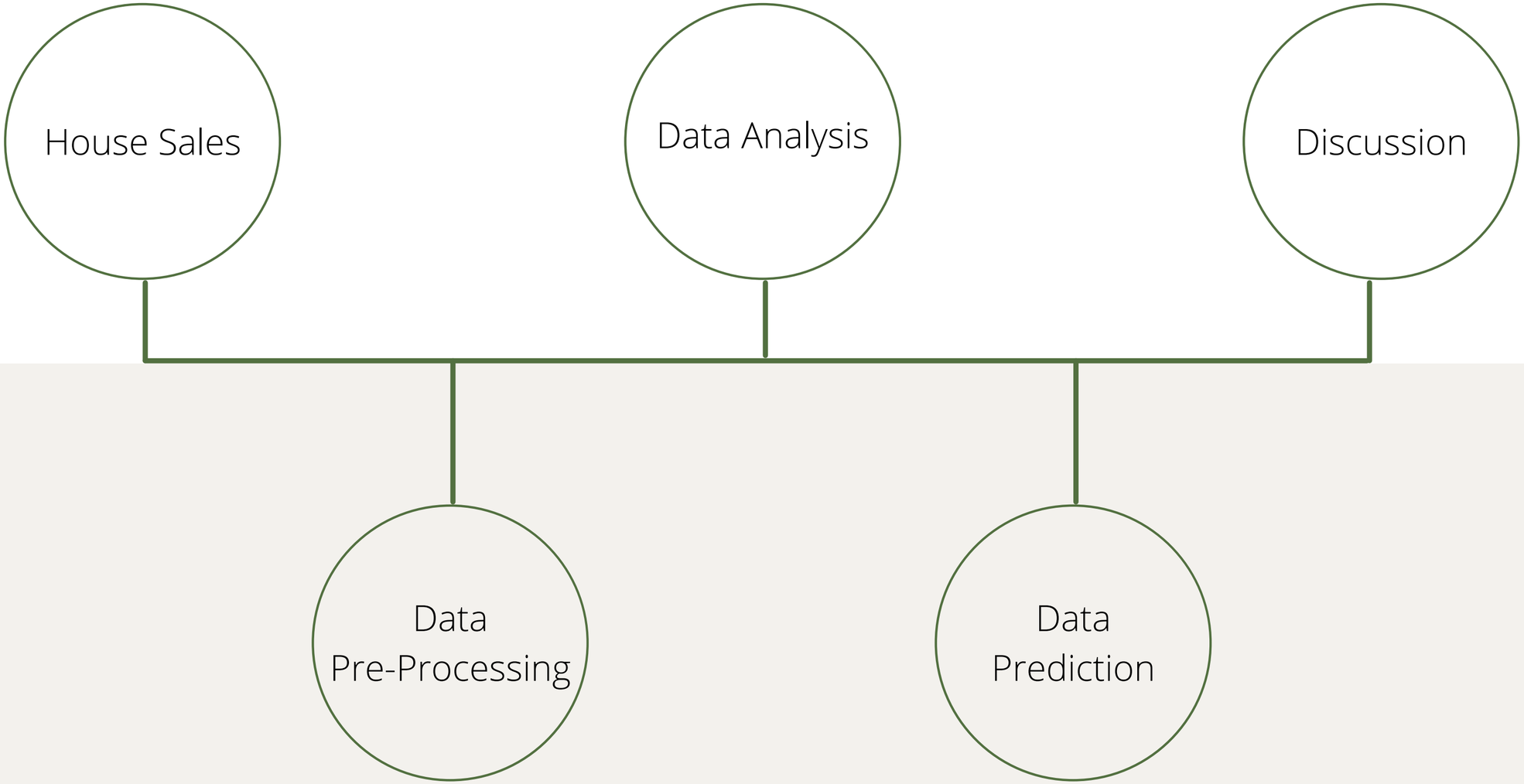
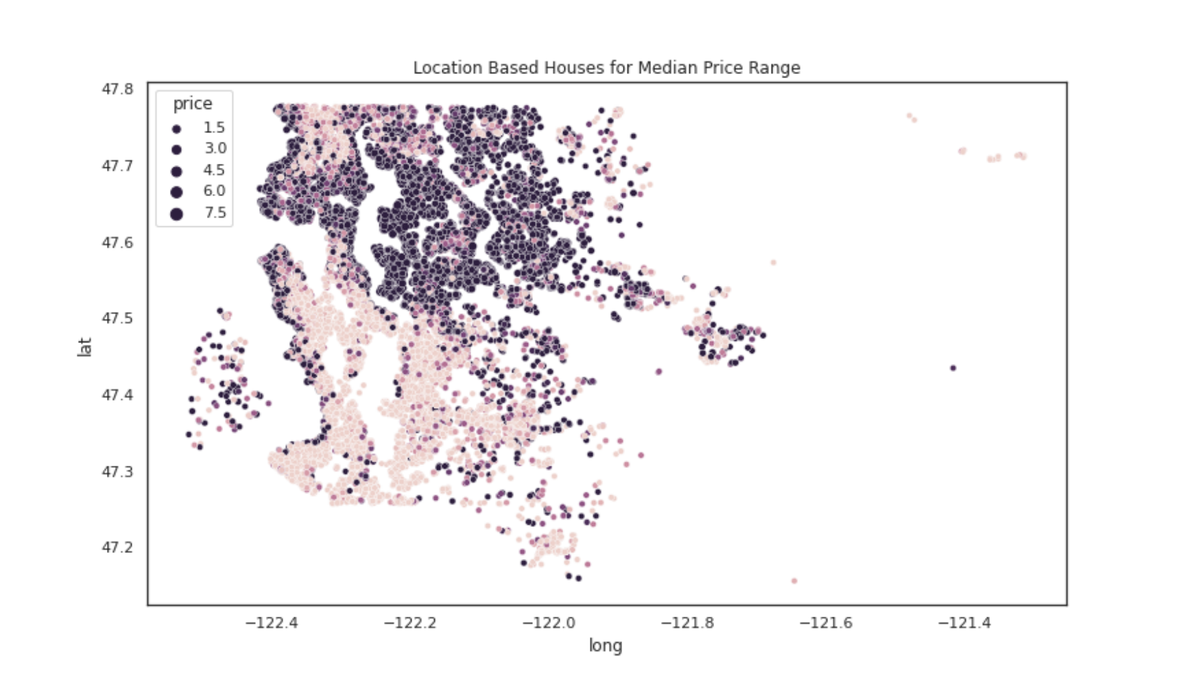
House Price Prediction Analysis

BY: TEAM 4
KATIE HU, BRYAN FLORES, MAYEL ESPINO

OVERVIEW


House Sales
Data Analysis
Discussion
Data
Pre-Processing
Data
Prediction

House Sales



Dataset from Kaggle

House Sales in King County, Washington between
May 2014 - May 2015

Method

Python via Jupyter Notebook, VS Code, Deepnote

Objective

To predict the house sale prices for the future quarter

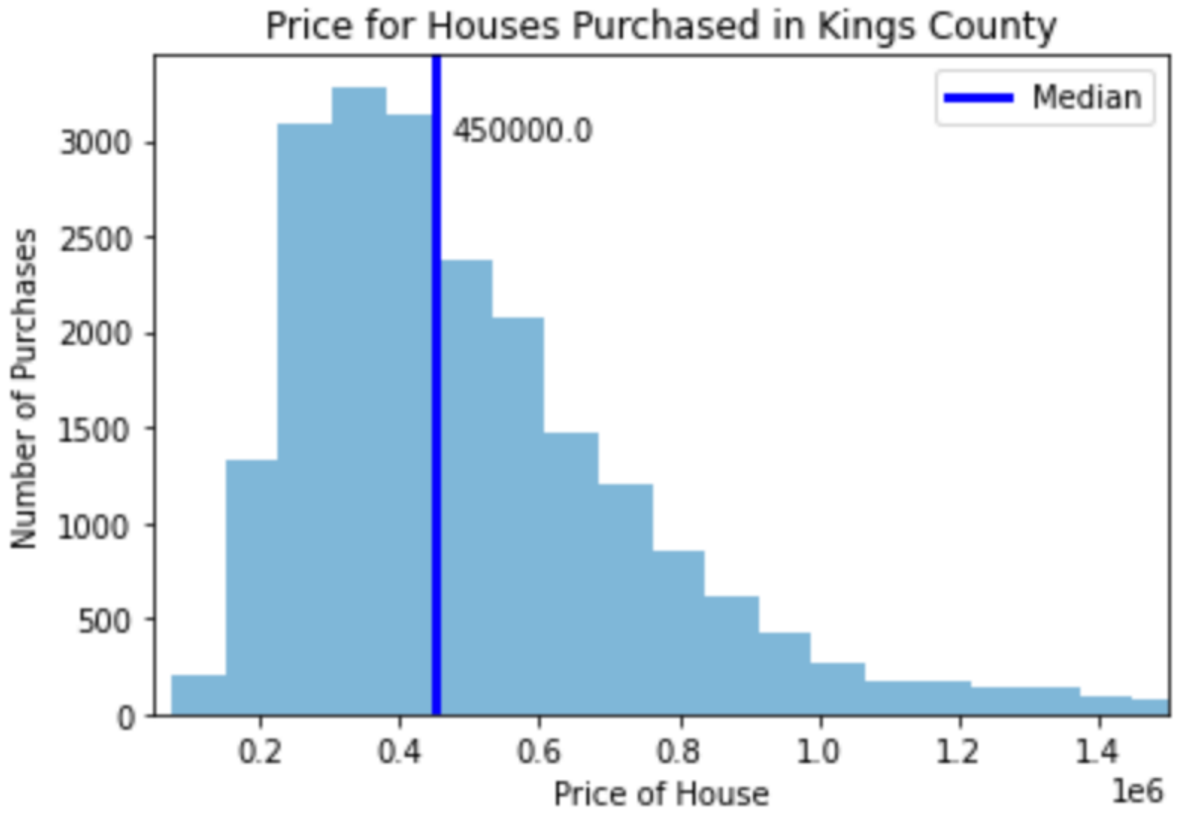
Data Pre-Processing

Understanding the Data

21,613 House Sales
21 House Features

Data Cleaning

- Date Adjustment
- Remove Outliers
- Missing Values
    - Bedrooms, Bathrooms, Square Foot Lot, Square Foot Living



Data Analysis

Highest Correlation with Price: Grade, Sqft_living, Bathrooms

Dependent Variable (y) : Price
Predictor Variables (X) : Grade, Sqft_living, Bathrooms

	price	bathrooms	sqft_living	grade
price	1.00	0.53	0.70	0.67
bathrooms	0.53	1.00	0.76	0.66
sqft_living	0.70	0.76	1.00	0.76
grade	0.67	0.66	0.76	1.00

Exploratory Analysis

Bathrooms

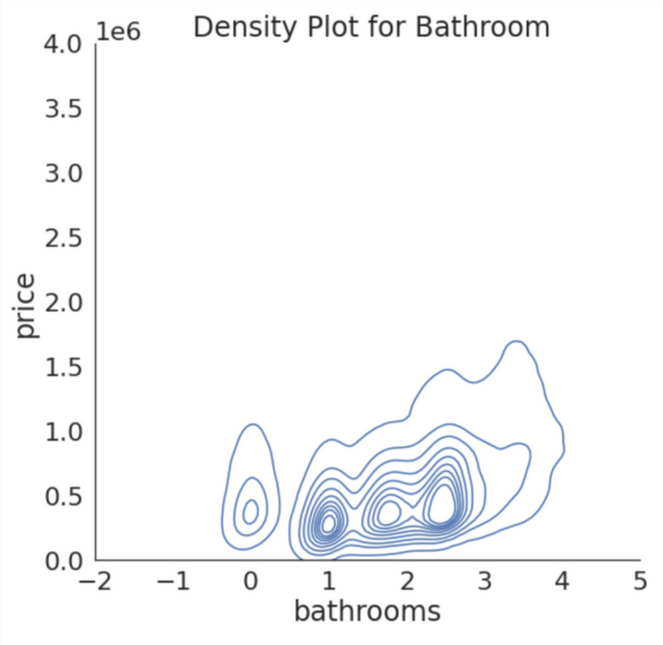
- Common Sizes

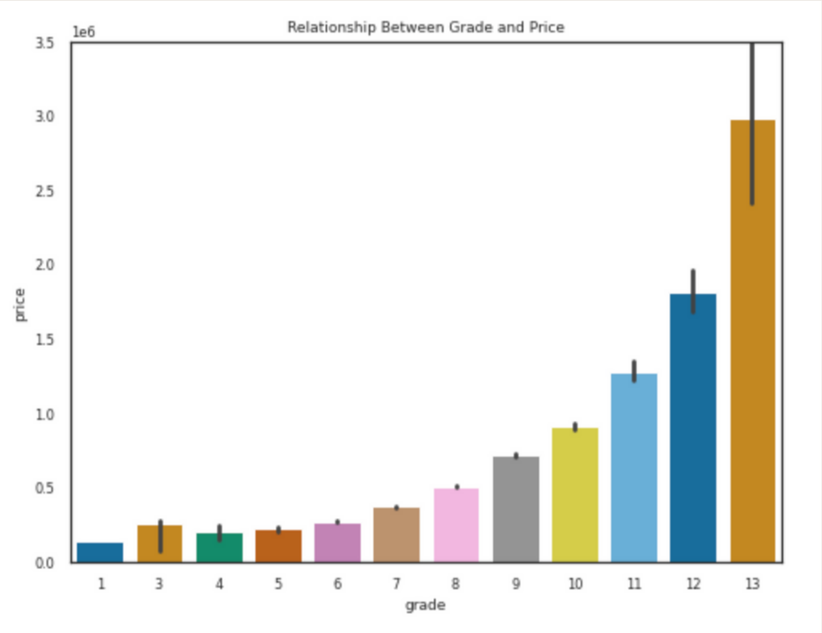
Grade

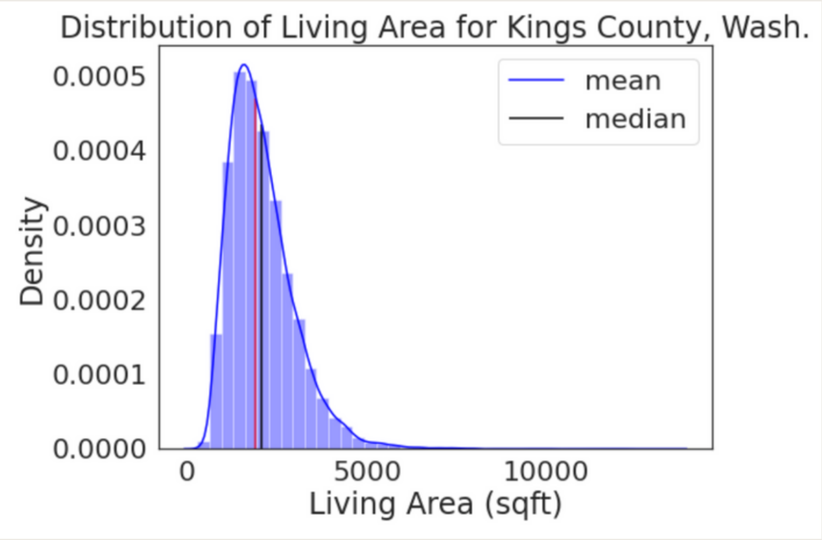
- Direct Relationship

Sqft_Living

- Positive Skew

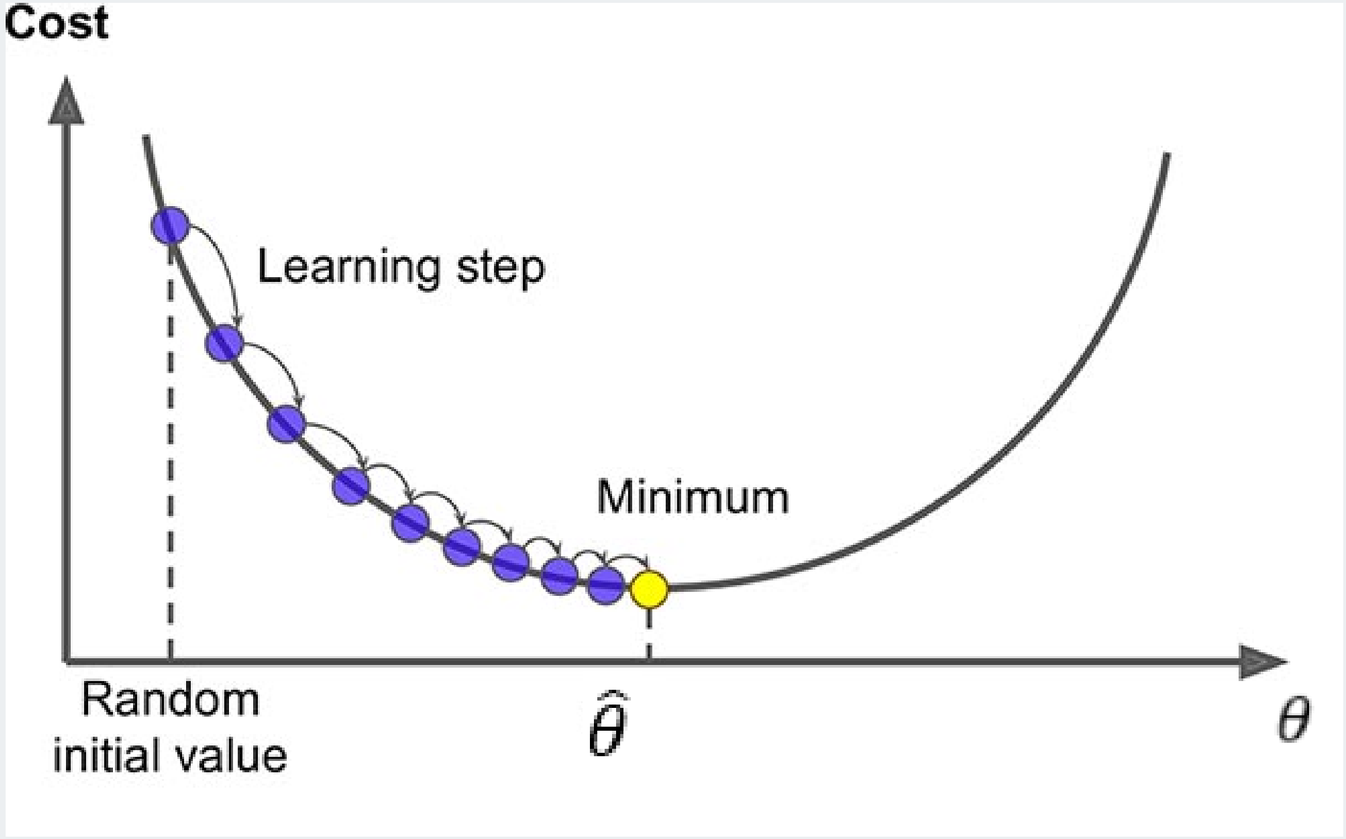






Prediction Method

Method: **Batch Gradient Descent**



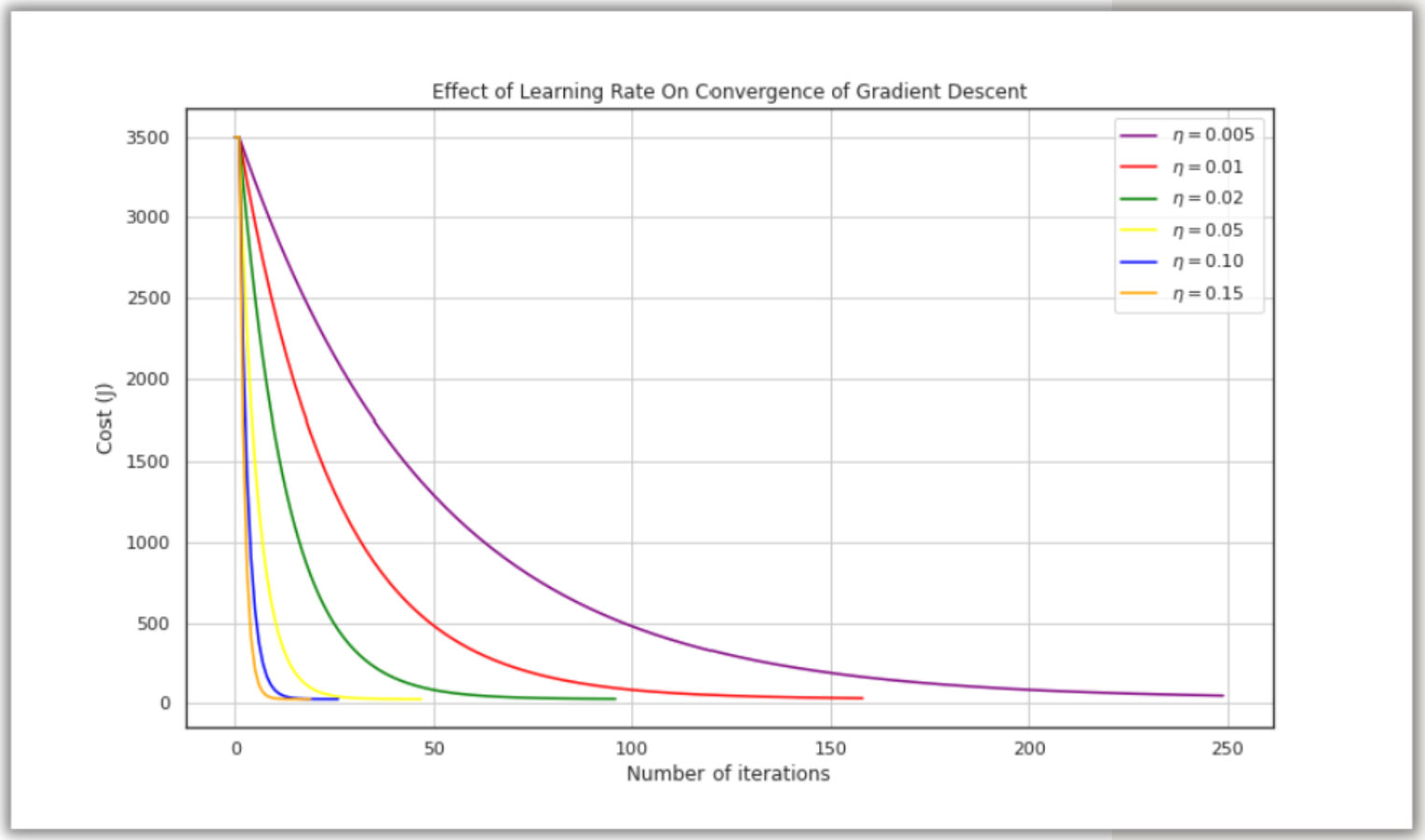
Gradient Descent Step

$$\theta^{next\ step} = \theta - \eta\ \nabla_\theta MSE(\theta)$$

Prediction Model

$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3$$

Results



Method: **Batch Gradient Descent**

Optimal Learning Rate: **0.15**

Metric: **Pearson**

Accuracy: **70%**

P-value < 0.05

Predicted House Sale Prices for King County, WA		
Constant: Grade as 7		
Bedrooms	Square Feet	Predicted Price
1	900	$ 250,985.93
2	1400	$ 347,977.84
3	1850	$ 435,293.07
4	2400	$ 541,961.64

DISCUSSION

Expectations of House Features

- Initial Expectations
- Data Cleaning
- Location, Location, Location

Additional Regression Analysis

- Random Forest
  - Decision Tree Regression
  - Decision Tree Classifier
- Standard Linear

THANK YOU

Citations

Harlfoxem. (2016). *House Sales in King County, USA* (Version V1) [Data set]. https://www.kaggle.com/harlfoxem/housesalesprediction

Bhattarai, S. (2018). *What is Gradient Descent in Machine Learning?* [Photograph]. https://saugatbhattarai.com.np/wp-content/uploads/2018/06/gradient-descent-1.jpg