## Y AndreiMklv / Phase\_2\_Project Public

forked from learn-co-curriculum/dsc-phase-2-project-v2-3

**11** Pull requests Projects Security <> Code Actions ✓ Insights

Codebout រុះ main ▾ Go to file

This branch is 2 commits ahead of learn-cocurriculum:main.

ໃ¹\ Contribute ▼

Д c51401c 9 minutes ag 18 commits revisions for v... data 13 months ago

- .giti... Initial commit 2 years ago
- LICE... Initial commit 2 years ago
- REA... Final Version 17 minutes ago
- pres... Final Version 17 minutes ago
- Final Version 17 minutes ago proj...
- **Final Version** 17 minutes ago proj...

**README.md** 

# **Phase 2 Project Description**

No description, website, or topics provided.

- M Readme
- থা View license
- ☆ 0 stars
- 0 watching
- 252 forks

#### Releases

No releases published

#### **Packages**

No packages published

#### Languages

**Jupyter Notebook** 100.0%

#### **Overview**

During this project, we will be investigating real estate datasets using multiple linear regression analysis. The customer is a real estate agency, who want to give better price recommendations to their customers. For the sake of research, we will use the following dataset (with formats):

 King County, WA Real Estate, CSV

We will employ Pandas, Statsmodel, NumPy, MatPlotLib, and SeaBorn libraries.

Considering the task and dataset natures where we need to predict price against explicitly and a relatively limited number of parameters, we decided to use linear regression analysis due to its simplicity, transparency, and versatility which makes it appropriate for our task.

### **Business Problem**

The customer, a real estate agency, wants to have a tool for home price prediction based on different parameters, like Living Area, Number of Bedrooms, or house conditions.

## **Methods**

For the analysis, we used OLS Linear Regression method. The following libraries have also been used.

- Pandas
- Statsmodel
- Matplotlib
- Scipy
- Seaborn
- Numpy

# **Findings**

The overall regression was statistically significant (R2 = .552, p < .000).

It was found that the Square Footage of the Living Area significantly predicted price ( $\beta$  = 272.5097, p < .000). It was also found that:

- Number of bathrooms significantly predicted price (β = 1.069e+04, p < .000)</li>
- Number of bedrooms
  significantly predicted price (β = -4.638e+04, p < .000).</li>

We suggest further investigation how features as 'Grade' and 'View' affect home prices.

# **Repository Structure**