Data Science Bootcamp 2022

House Pricing

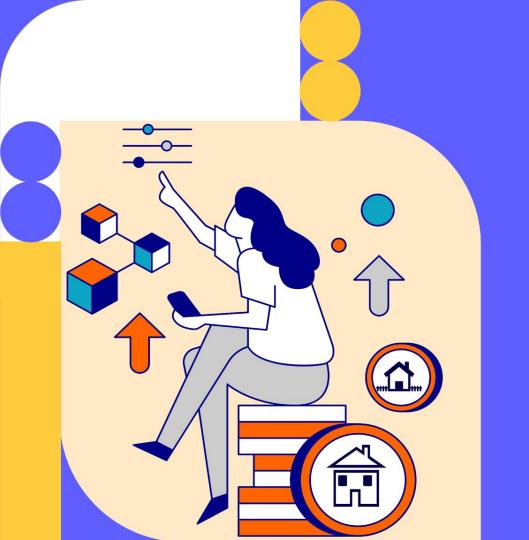
Prediction







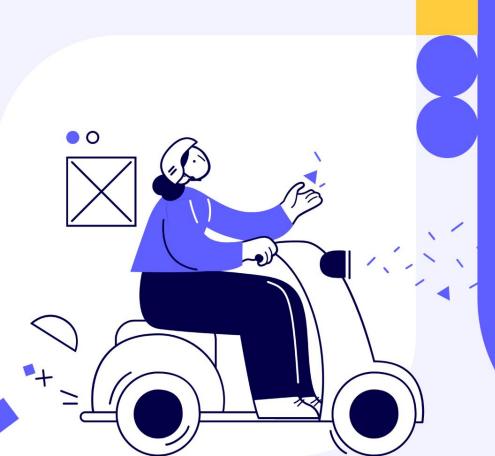
Rafael Rojas, Lydia Delgado, Yesenia Pérez



Selected Topic

House Pricing Prediction

The housing market's pricing everywhere depends on numerous factors, with this approach a curious client may have a place to start sensing the changes in price for each house depending on the characteristics that they are looking for.



Why this topic?

- 11 Avalaibility of information
- **02** Well-documented databases
- Real-life problem with reliable sources of investigation

Kaggle ••

The database is obtained from Kaggle, an online community of data scientists and machine learning practitioners, which for its nature it sources this type of datasets for educational purposes.

Dataset ••

The dataset consists on 80 labels pertaining to the characteristics of each house registered.

Entries ••

There are 1,460 rows for this dataset, sufficing the requirement for training a successful machine learning model of over 1,000 entries.

Description of data source



Questions to answer with the data

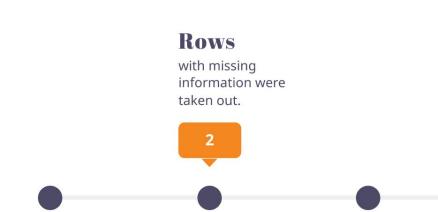
What is the price range of a house with the same characteristics that I am looking for?



Which variables account for a higher prediction weight?

Which variables can be neglected at the time of making a decision?

Data Exploration Phase



Analyze

data to look for null entries.

The target

value was classified as high or medium price.



were made of numeric labels, and labels turned to text

Statistics





Target Value

was selected as the price of the real estate property.

Data Types

were verified and changed when needed, depending on the type of analysis.

Analysis Phase

Label **Most Relevant** the properties of feature at the time higher price than the of making a decision mean as high price and price-wise is Kitchen those lower as medium Quality with 7% of price. the weight.

Price

of real state properties was converted to a numeric value to get the mean.

First Analysis

was made with a Balanced Random Forest Classifier which came back with 83.6% of accuracy.

Other

relevant feature came back as number of full bathrooms with 3.9% of the weight.

Dashboard

Navigation Bar

FILTERS

Drop down menus:

of Bedrooms

of Full Baths

of Half Baths

of cars in Garage

Range Slider:

Sale Price

Drop down menu:

ID of matching houses

Text:

Characteristics of matching houses

YEAR VS. PRICE GRAPH SQFT VS. PRICE GRAPH

ARTICLE ABOUT HOUSE MARKET IN 2022







For more information please visit

https://github.com/yempi/house-pricing

Data Science Bootcamp