



Linear Regression Property Price

By

Mashari Almuhanha

Sohaib Albakri



Data Science Bootcamp

SDAIA Academy

Riyadh – Saudi Arabia

[September 2021]

Problem

Most of the property owners those trying to set a selling price for their property does not know what is the best price which make them set a high price or to low.

Solution

In this project, we will develop a linear regression model to predict the property prices according to its characteristics and location. Which will help the owners to set the best price.

Dataset

The dataset we will use contains the number of rooms, number of baths the size of the property, price, and the district. The data are taken from the [Aqar website](#). The following table explains the dataset in detail:

Column	Data type
Rooms	Integer
Baths	Integer
Size	Integer
Price	Integer
District	String

Features

First, We need to grab the districts names in Riyadh from [Wikipedia](#) because the names in Aqar Are not consistent so we need to use a consistent names. Second step is to grab the dataset from Aqar.

Tools

To grab the data and build the model we will use the following tools:

Tool	Description
Jupyter notebook	Contains cells of Python code and human-readable text
pandas	Library is written in Python for data manipulation and analysis
sklearn	Software machine learning library for the Python programming language
beautifulsoup	Python package for parsing HTML documents. It creates a parse tree for parsed pages that can be used to extract data from HTML

Conclusion

We expect the regression model will predict the property price after receiving the number of rooms, number of baths, size, and price. The model will help the property owners to set the optimal price for their property.