SDAIA Academy
أكـاديـمـيـة سـدايـا

Project report:Linear regression model

Predict land real estate in Saudi Arabia

By

Amal Alshihabi

Shoug Alkhathran

Dr.Mejdal Alqahtani

Data Science Bootcamp SDAIA Academy

Summary:

In this project, we will use data from (Aqar) website, which provides lands for sale and their prices. Our goal with this project is to build a linear regression model that predicts land prices.

Data Description:

We have used Web Scrapping to collect our data and the source was Aqar site and Wekipedia

| Columns	Data type	Description
City	object	city where house locate in
District	object	district where house locate in
Price	int	show the price of the house
Area	int	size in m^2

Algorithms:

- Problem understanding
- Data collection
- Data cleaning
- Explore data
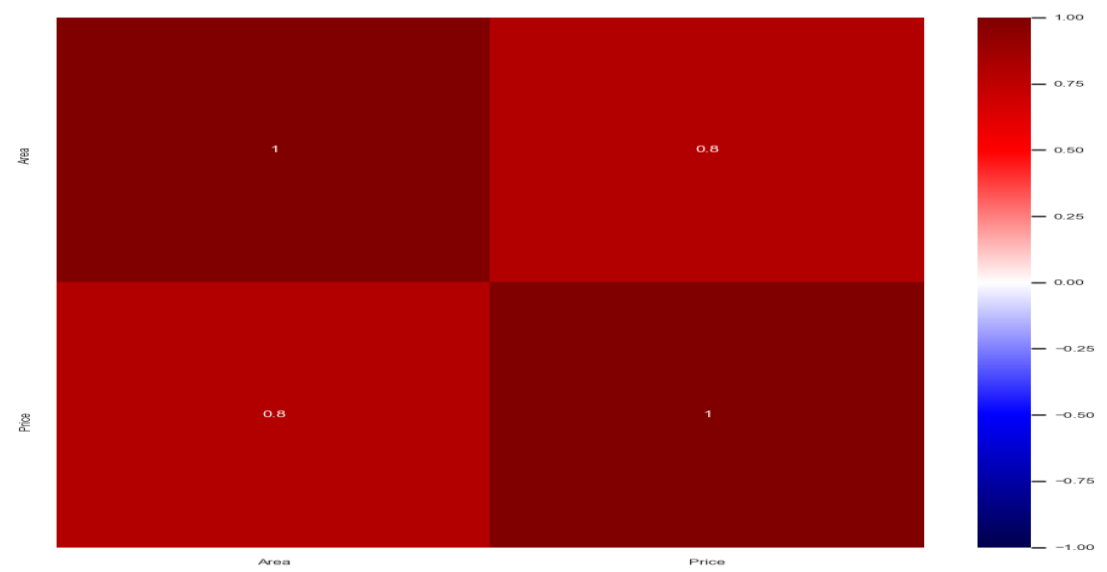- Finding and insights
- Model build

Tools:

- Web scraping tools (of Beautiful Soup, Selenium)
- Libraries: Pandas, Matplotlib, Seaborn and NumPy.
- Plots: Bar graph, forecast, bar chart, and stock price prediction.

Communication:
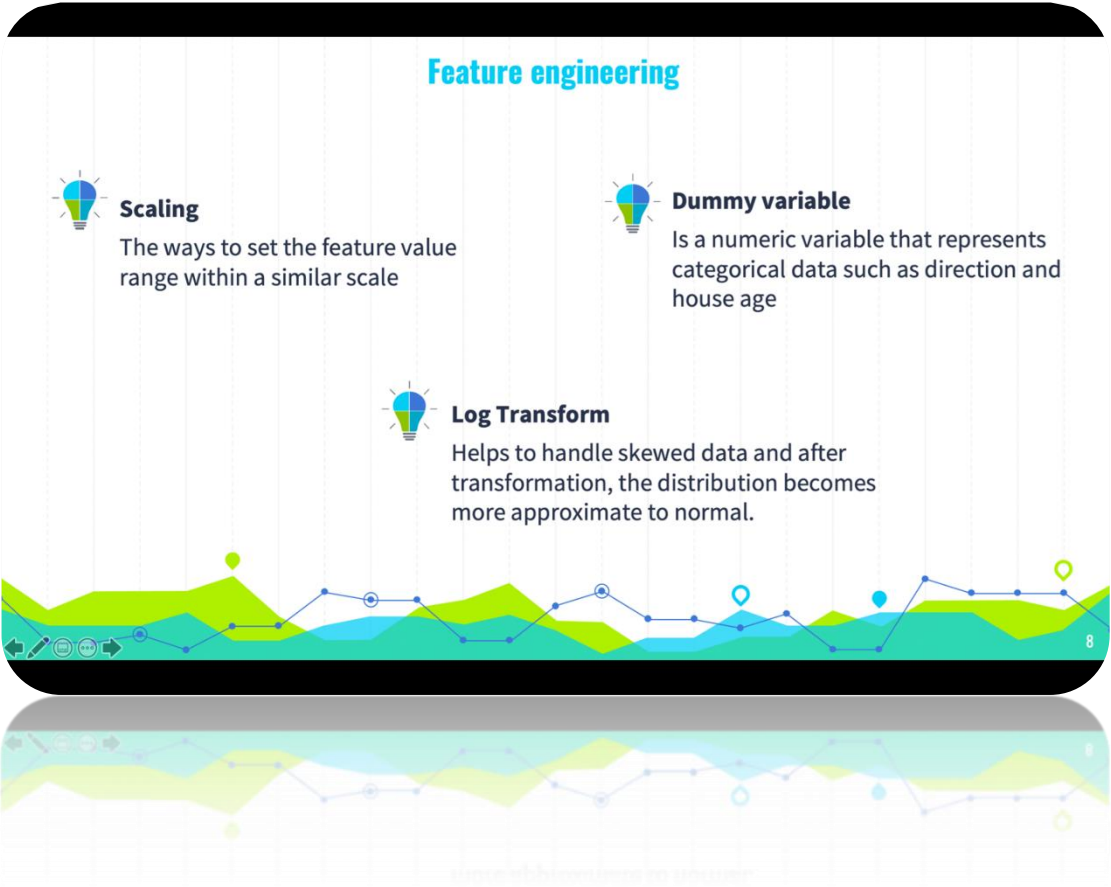
The relationship between the price column and the area

Model Scores

Model	Training	Test
Linear Regression	0.968239	0.968238
Ridge Regression	0.968239	0.967057
Lasso Regression	0.688971	0.688971
Polynomial Regression	0.969674	0.969674
Random Forest Regressor	0.956914	0.956913

Error Values