Riyadh Houses prediction

By: Ebtesam Sultan - Nuha Aljohani

Abstract

The goal of this project is to build a regression model to predict sale prices for houses in Riyadh.

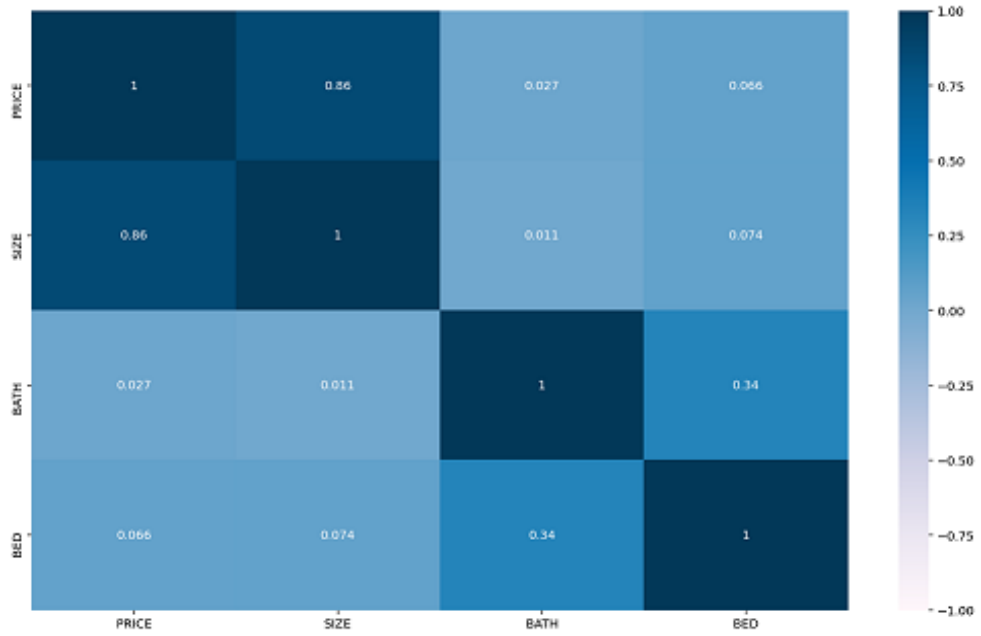
Data Description

- We scraping our data set from Aqar website with 1575 rows and 6 columns.
- The Columns include (Location, Price, Size , Bed , Bath and link) .
- Data Type: String -Integer.
- Data set hasn’t null values.

Tools

- **Technologies**: Python, Jupyter Notebook, PowerPoint
- **Libraries**: Pandas, NumPy, Matplotlib, Seaborn, Requests, BeautifulSoup, Sklearn

Communication:



- The figure represented in heatmap plot, the above figure shows the relations between the features in range between -1 to 1

- 1- the plot show that price has a correlation with size

- 2- the plot show that rooms and bath have a lower correlation with price

Experiments:

Model	Train	validation
Base line	0.77	0.28
Label Encoder Regression	0.76	0.59
Adding Polynomial terms	0.82	-0.09
Adding interaction terms	0.77	0.64
Lasso Regression	0.77	0.64
Ridge Regression	0.77	0.72

- We noticed during the project, that model Ridge Regression gave us the best results among the models.
- Ridge Regression model can help the real estate companies to predict Riyadh housing prices.