Predicting Apartment Rentals in Riyadh Using Linear Regression

Presented by Feras Alyahya ,Ahmed Alonaizi ,Hussain Alhadab and Mohammed Alhamoud

Outline :

Introduction

- **Problem** :
  - Many of apartments owners struggle to set the optimal price for rent due to their lack of experience in rent prices. We are developing a linear regression model to predict the best prices that fits the specifications and location of the apartment.

- **Solution :**
  - In this project, we will develop a linear regression model to predict apartment rent prices according to their characteristics. Which will help the owners to determine the best price.

Data :

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

---

Data Columns :

PRICE- BED - BATH - LIVING - AGE - IMAGES - FLOOR - TARGET-
FURNISHED - KITCHEN - ANNEXE - PARKING - ELEVATOR - AC

Data cleaning & engineering

Before Data cleaning & engineering

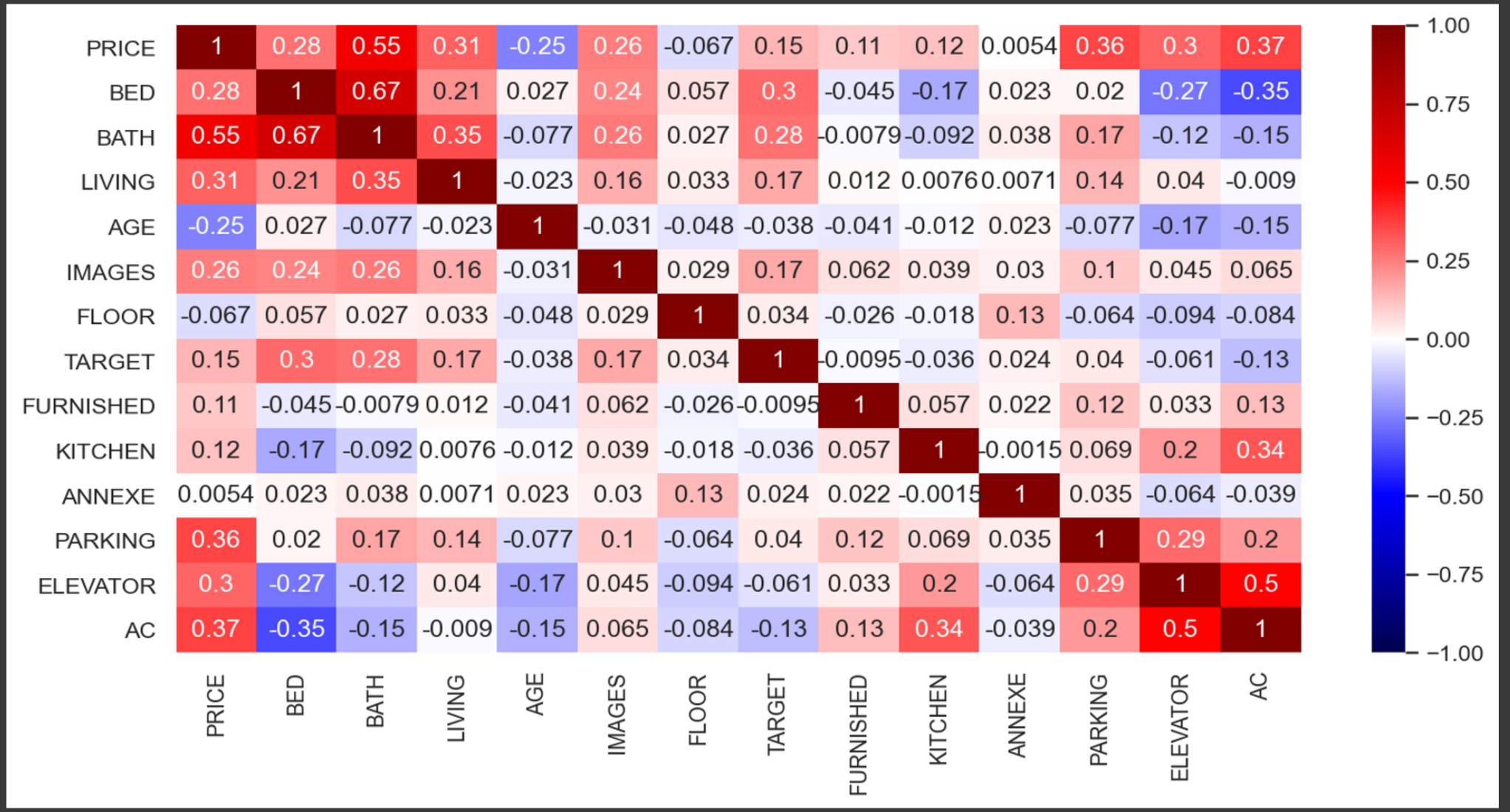
Number of rows = 6978

- **Remove empty values**
- **Remove redundancy**
- **Remove the outlier values in the (Price) and (BED) columns**
- **Convert the variable type to numeric**
- **Add a column with the name (Amanat)**
- **Add a column with the name (Regions)**

After Data cleaning & engineering

Number of rows = 5146

Correlation :



Result

		Linear Regression	Ridge Regression	Polynomial Regression
X1 =	• All features • Using Amanat (Baladiat)	Train R^2 = 0.71 Validation R^2 = 0.71 Test R^2 = 0.72	Train R^2 = 0.71 Validation R^2 = 0.71 Test R^2 = 0.72	Train R^2 = 0.83 Validation R^2 = 0.76 Test R^2 = 0.78
X2=	• All features • Using Neighborhoods	Train R^2 = 0.83 Validation R^2 = -4.0 Test R^2 = -4.2	Train R^2 = 0.83 Validation R^2 = -4.0 Test R^2 = -4.2	Train R^2 = 0.91 Validation R^2 = -900 Test R^2 = -800
X3=	• All features • Using Regions	Train R^2 = 0.70 Validation R^2 = 0.70 Test R^2 = 0.71	Train R^2 = 0.70 Validation R^2 = 0.70 Test R^2 = 0.71	Train R^2 = 0.78 Validation R^2 = 0.75 Test R^2 = 0.76
X4=	• Scaled features • Log10 Price • Using Amanat	Train R^2 = 0.784 Validation R^2 = 0.78 Test R^2 = 0.77	Train R^2 = 0.785 Validation R^2 = 0.78 Test R^2 = 0.71	Train R^2 = 0.82 Validation R^2 = -111 Test R^2 = -087

Prediction:

	Actual Price	LinearRegressionModel	LinearRegressionScaledLog10	PolynomialRegression
0	20,000	20,809	21,193	24,478
1	60,000	48,129	48,993	54,305
2	20,000	12,224	11,997	14,990
3	22,000	16,347	16,278	11,240
4	35,000	23,744	23,309	29,736

Conclusion

•After the results of the model appeared, and the comparison between the results of each model, we concluded that the best model is the model with the highest R^2.

We seek to improve the accuracy of the model in the future by adding additional features such as the average income of the population in each neighborhood or the number of residents in each neighborhood , which helps us in developing the project and determining the rental prices of apartments based on specifications