



# Riyadh real-estate

by: Samaher Alhadi, Tia Alzanan



AQAR HOME  
real estates

# INTRODTION



## Problem:

In these days, the real estates in Riyadh is a hot topic and the prices changes monthly because of the high demand. 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 low.

## Solution:

will develop a linear regression model to predict the property prices according to its characteristics and area. Which will help the owners to set the best price.

# WorkFlow

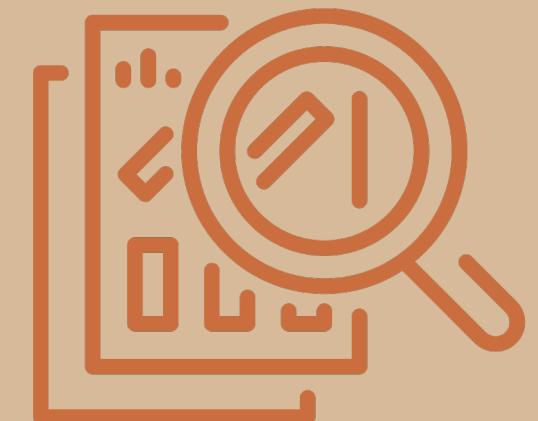
Web Scraping



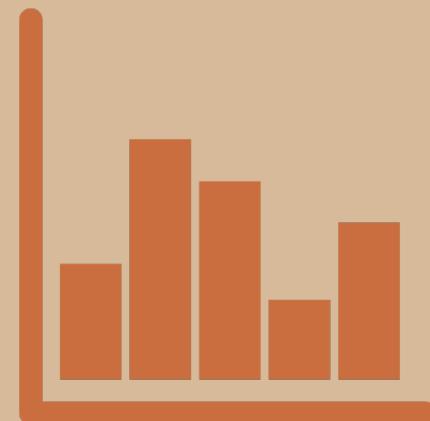
EDA



Feature Engineering



Regression Models



# Web Scraping

The screenshot shows the Aqar website's search interface at the top, featuring filters for location (الرياض), type (شقق للبيع), and price range (4,000,000). Below the search bar, a specific listing is displayed for a villa in West Riyadh. The listing includes the address (فلايللبيع في حي طويق، الرياض)، the price (4,000,000 ريال)، and details about the property (494 م²، 4 غرف نوم، 7 حمامات). A thumbnail image of the villa is also shown.

## Data Source

the prices of villas for sale in West Riyadh from Aqar website which is a famous real estate website in Saudi Arabia. it consists of 17,963 rows and 9 columns

## Data Columns

- Price
- Area
- Direction
- Bedroom
- Bathroom
- Living room
- Street width
- House\_age
- Views

# Exploratory Data Analysis



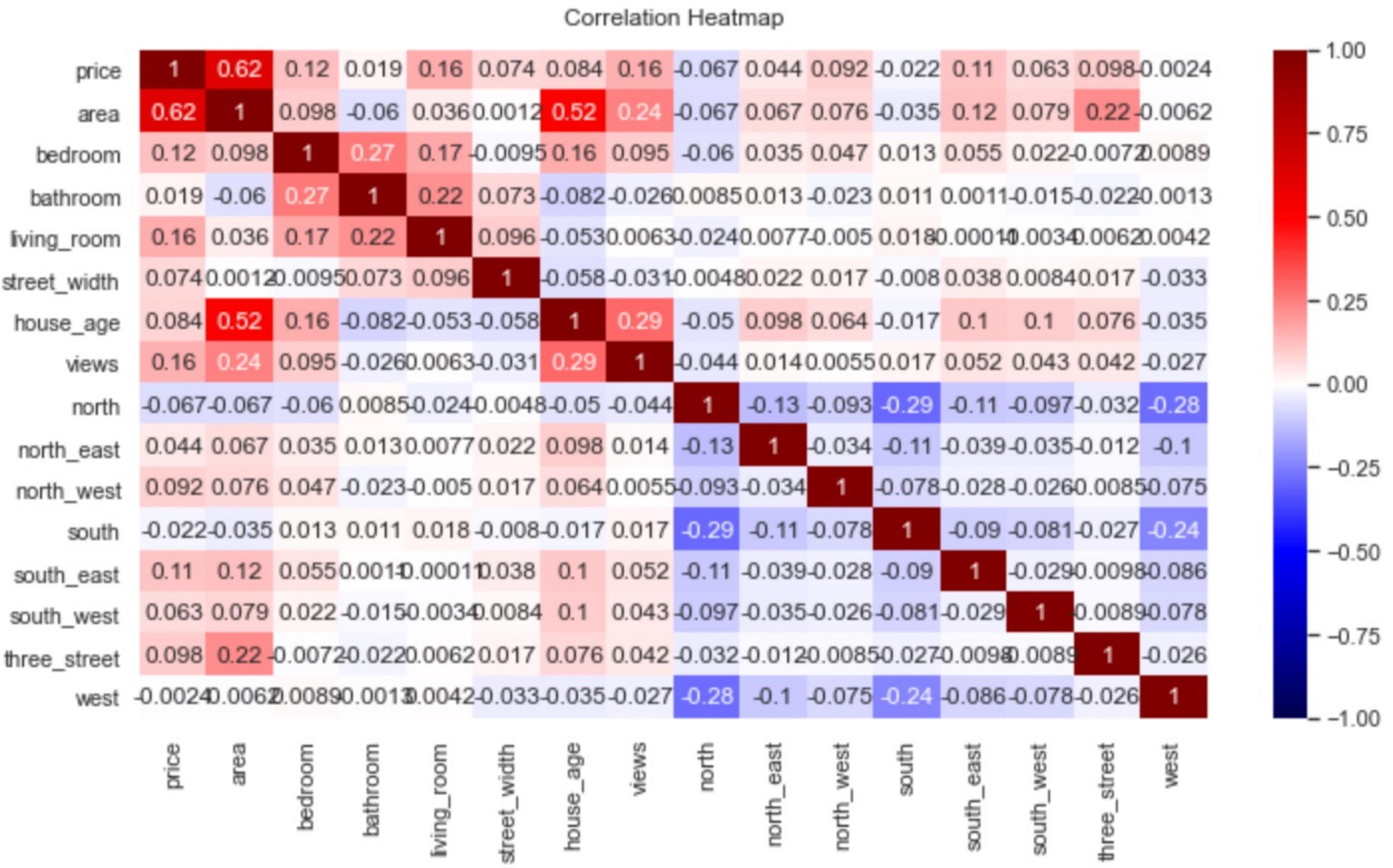
Convert data to DataFrame

## Data Cleaning:

- missing values (NANs).
- duplicate rows.
- outliers.
- remove (commas in price,  
squared meter in area,  
meter in street\_width,  
سنہ in house\_age)
- Change Arabic values to English in  
direction column.

# Feature Engineering

- Correlation of data.
- Dummy variable (61 columns).



# Regression Models

## Models Scores

Models	Training	Validation
Linear Regression	<b>0.53</b>	<b>0.52</b>
Polynomial Features (2)	<b>0.81</b>	<b>0.18</b>
Lasso	<b>0.53</b>	<b>0.52</b>
Ridge	<b>0.53</b>	<b>0.53</b>
Random Forest Regressor	<b>0.97</b>	<b>0.75</b>

# Regression Models

## Testing Scores

Models	R <sup>2</sup> Test	R <sup>2</sup> Train
Linear Regression	<b>0.5277</b>	<b>0.5303</b>
Ridge	<b>0.5268</b>	<b>0.5296</b>
Lasso	<b>0.5278</b>	<b>0.5303</b>

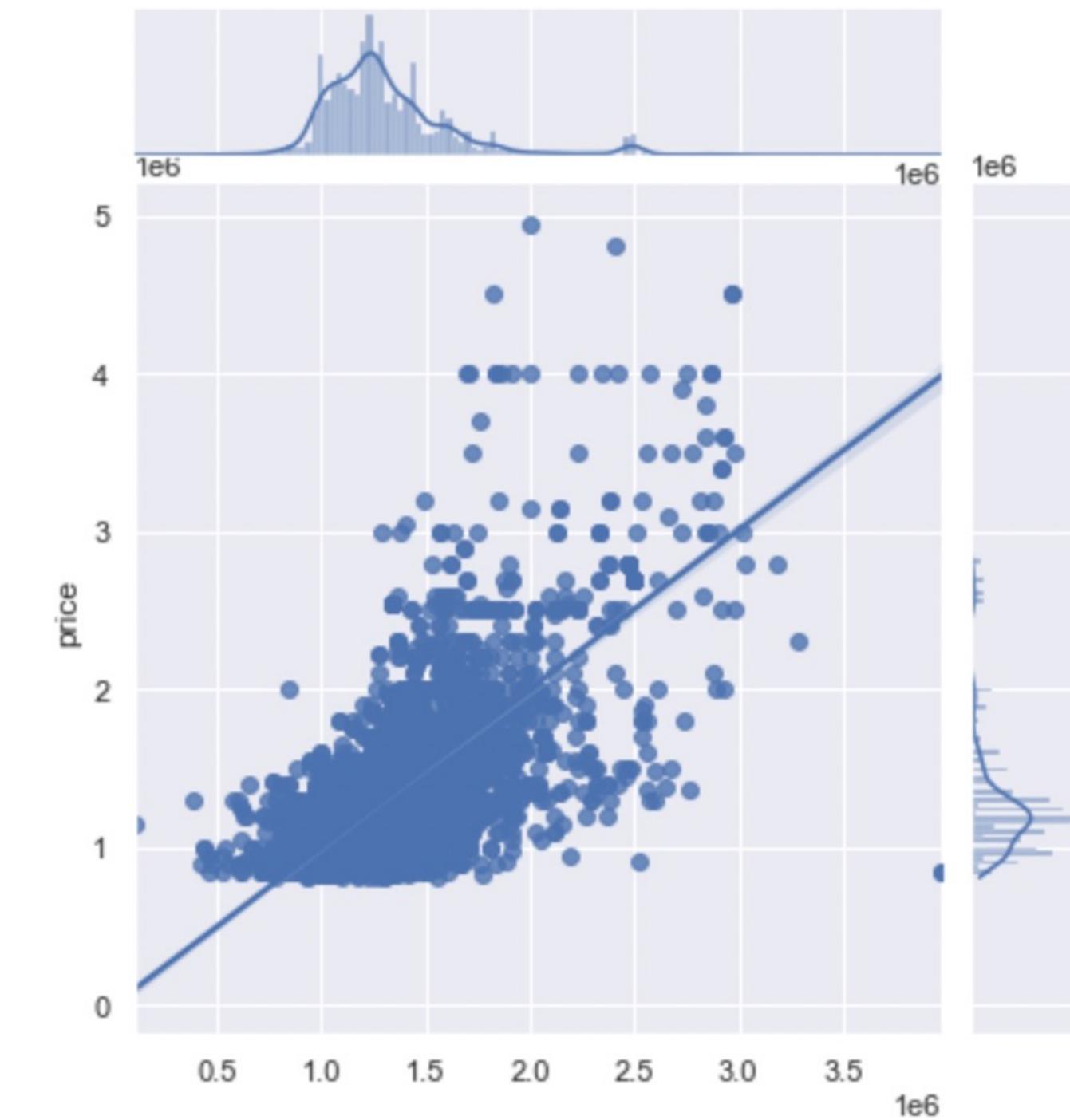


AQAR HOME  
real estates

# Regression Models



**Lasso Model:**  
**Train\_R<sup>2</sup>: 0.530**  
**Test\_R<sup>2</sup> : 0.527**





# Future Works

- Enhance the performance.
- Improve R Square.
- Add the ad date column.



Thank you !