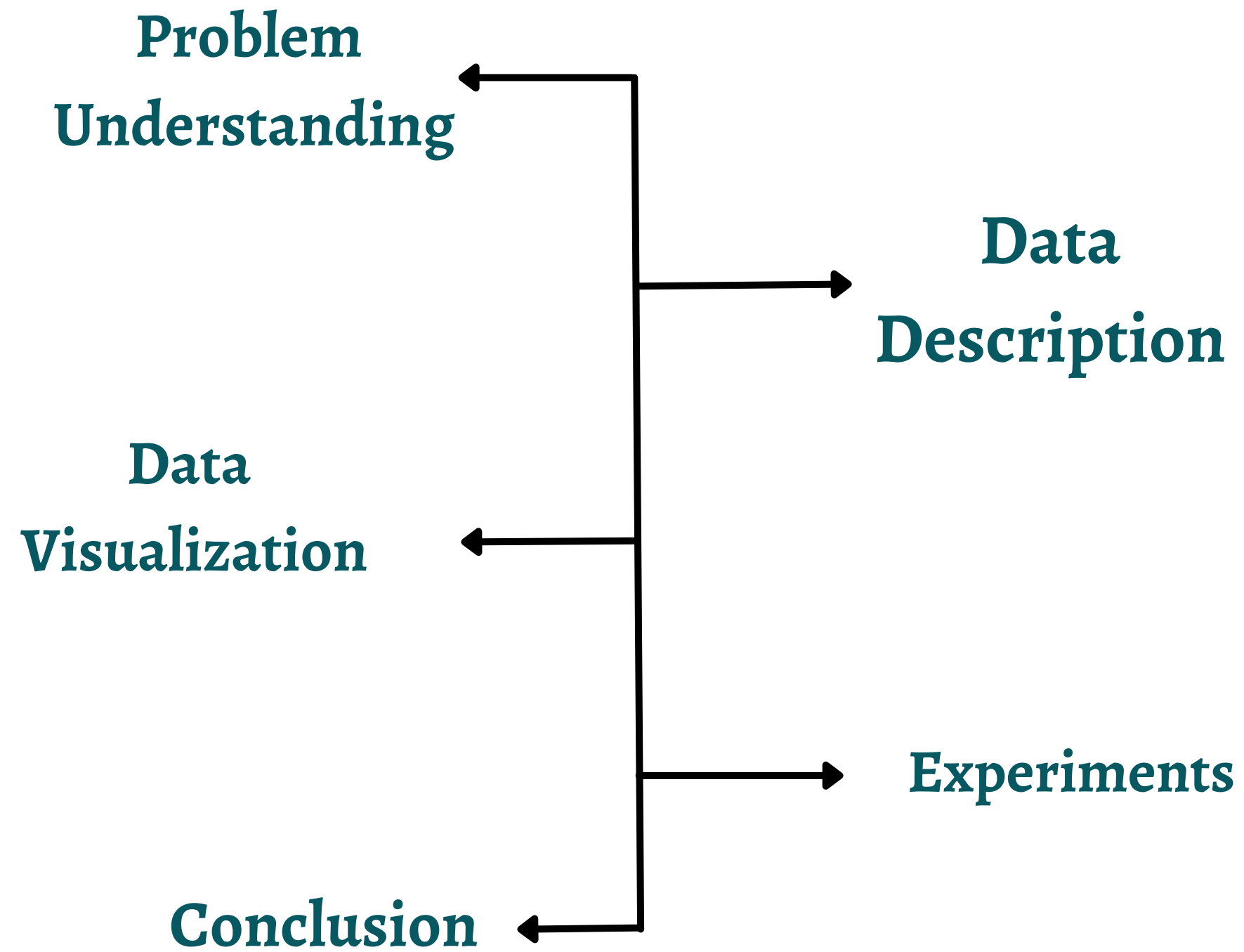


السعودية ريج

HOTELS PRICE PREDICTION IN SAUDI ARABIA



Methodology



- **Overview**

One of Saudi Arabia 2030 vision key initiatives is “Quality of life”. This initiative comes with the aim of diversifying and enriching the tourism and entertainment experience in the Kingdom.

- **Problem Statment**

In this time of the year, hotels and resorts are in high demand in the Kingdom. In this project we build a model to predict the prices in riyals, given the rating score, and the number of reviews.

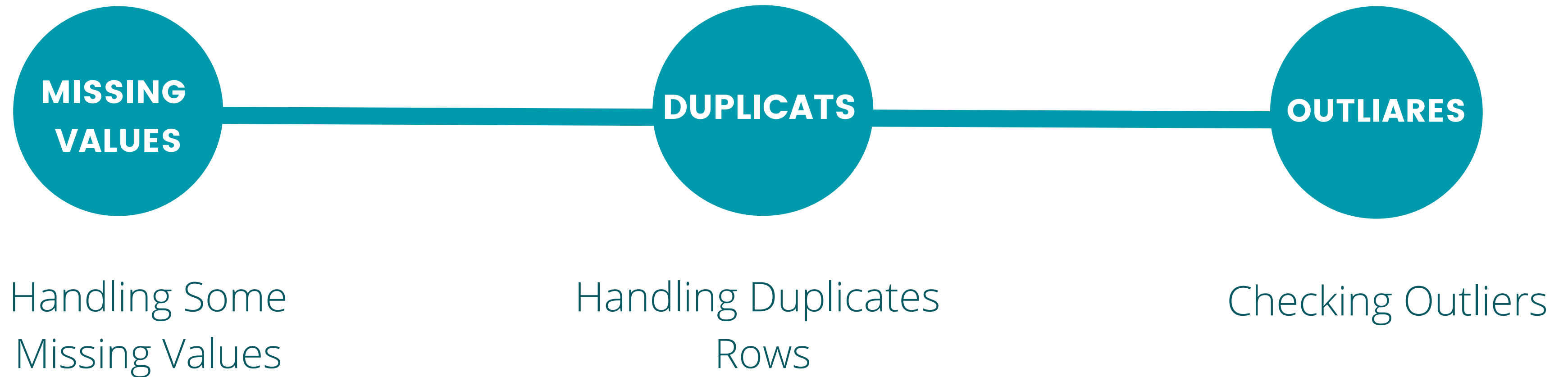
Methodology

Data Description



They data was token by web-scraping
(booking) website and combine it with
data from kaggle. it is about data of 1224
hotels in Saudi Arabia in the current
month (December 2021),

A screenshot of a hotel booking website's search results page for Riyadh. The page has a dark blue header with navigation links like "rentals", "Attractions", and "Airport taxis". It also includes a currency selector set to "SAR", a US flag, a help icon, and buttons for "List your property", "Register", and "Sign in". Below the header, the search results for "Riyadh" show "423 properties found". There are filter tabs for "Our Top Picks", "Homes & apartments first", "Price (lowest first)", and "Stars (highest first)". Two hotel listings are visible: "Radisson Blu Hotel, Riyadh" and "InterContinental Riyadh, an IHG Hotel". Each listing includes a photo, a heart icon, star ratings, a location pin, a brief description of the room, the number of reviews, a rating of 7.8, the price for 1 night for 2 adults, and a "See availability" button. The InterContinental listing also features a "Promoted" badge and a "FREE cancellation" note.



Methodology

Data Description



Features

Hotel Name

Rating

Rating Title

Number of Ratings

Room Type

Place

Target

Price

Heatmap Correlation



Baseline Model

Train Score : 0.096

Validation Score : 0.130

Experiments

1

**Dummy
Variables**

Train Score
0.1248

Validation Score
0.1024

2

**Count Value
for Location**

Train Score
0.1671

Validation Score
0.1325

3

polynomial Equ

Train Score
0.1694

Validation Score
0.1329

4

**Features
Engineering
(Adding Interaction
Terms)**

**Train Score
0.225**

**Validation Score
0.175**

5

Box Cox

**Train Score
0.6436**

**Validation Score
0.629**

6

Lasso

$\lambda = 2$

Train Score

0.640

Validation Score

0.631

7

Ridge

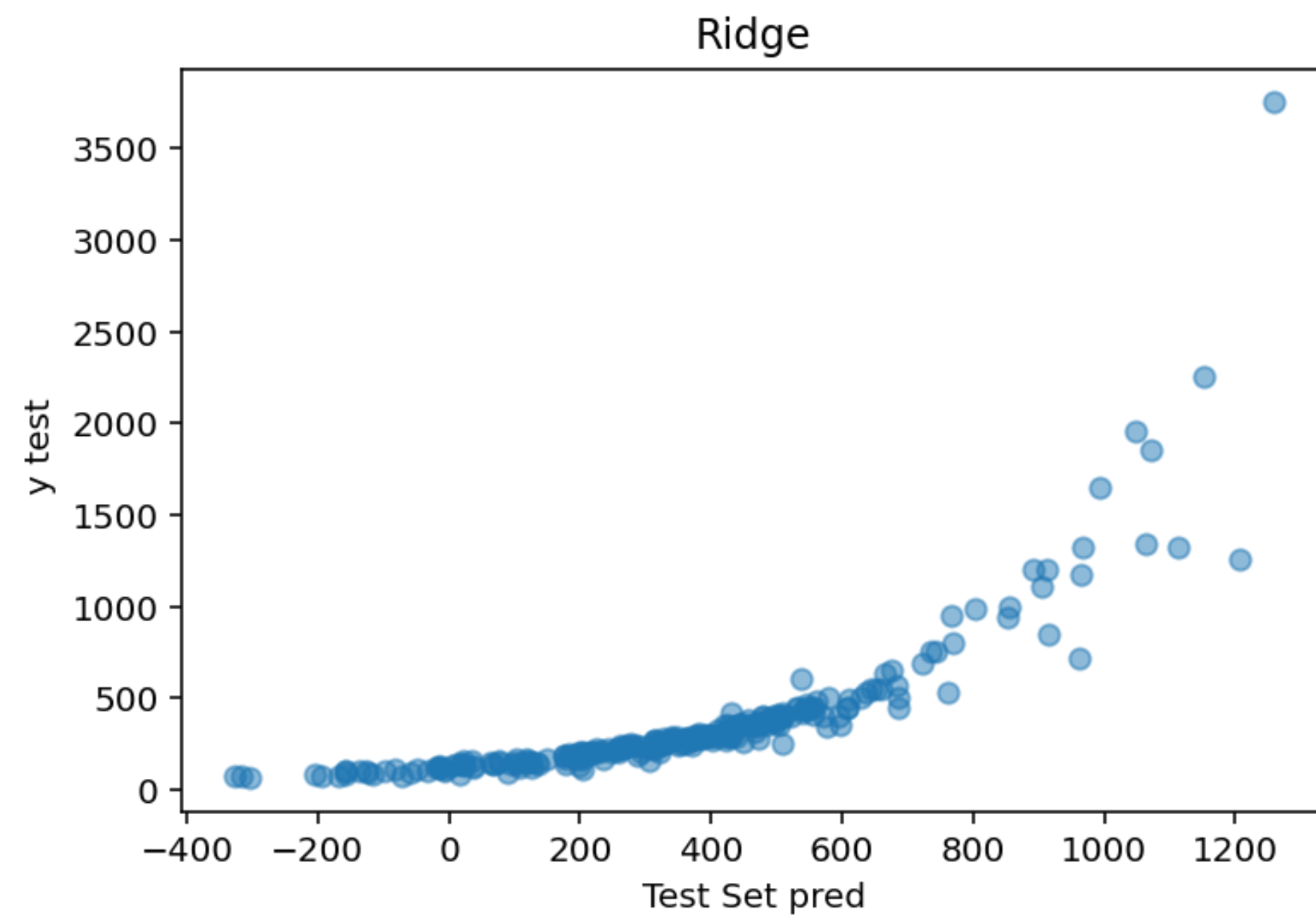
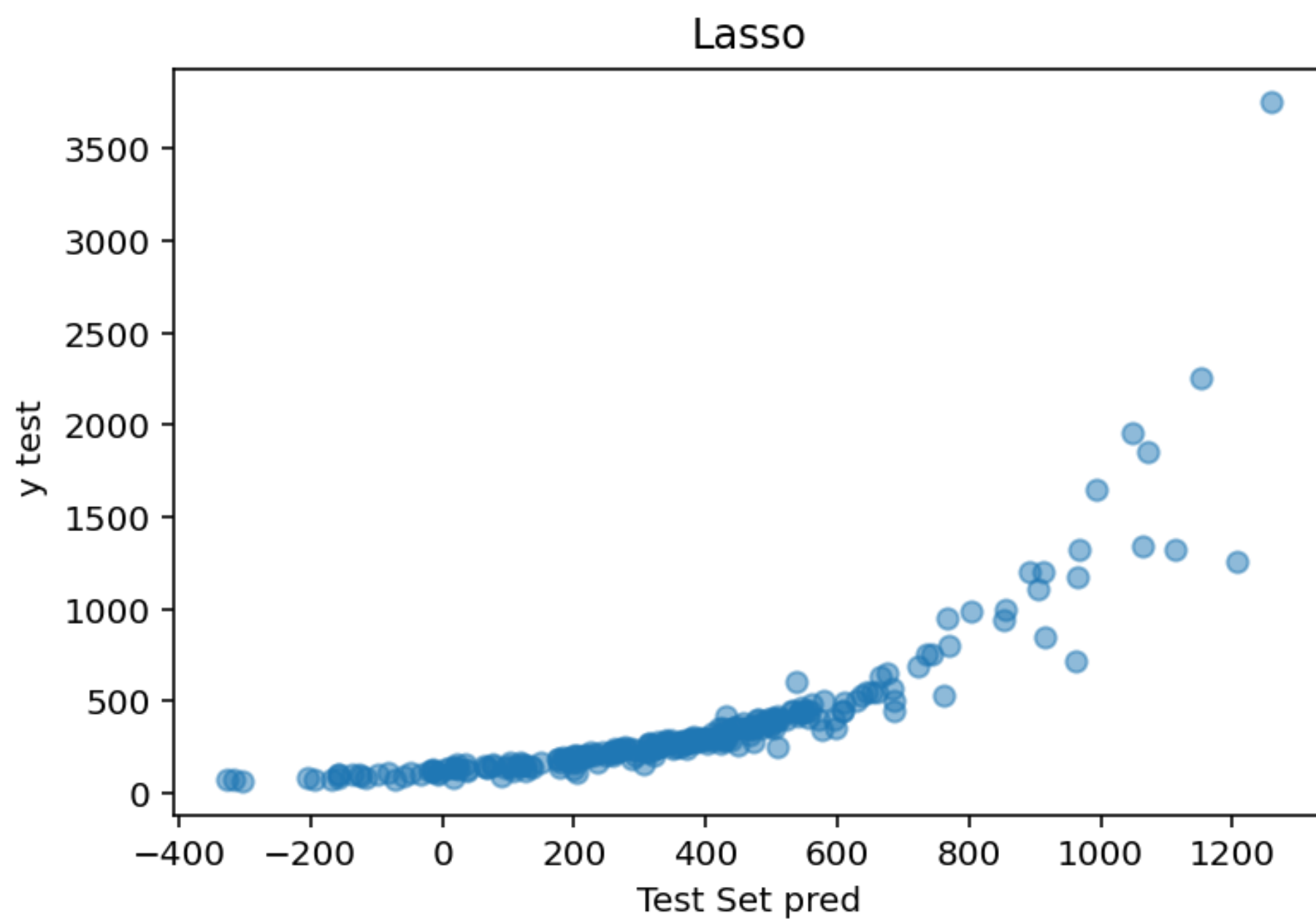
$\lambda = 6$

Train Score

0.643

Validation Score

0.631



SSE = 1.316065e+07

Conclusion

	Baseline	Dummy	Count	Polynomial Equ	Group-by price mean per location	Adding Interaction Terms	Box-Cos	LASSO	RIDGE
Train Score	0.096	0.124	0.1671	0.1694	0.169	0.225	0.643	0.64	0.64
Validation Score	0.1308	0.102	0.1325	0.1329	0.132	0.175	0.629	0.63	0.63

Best Model Train Score = 0.64

Best Model Validation Score = 0.63

Best Model Test Score = 0.63

Thank You ..

Raghad Albarrak

Maryam Aljasham