

# ADEPAMENT PRICE IN SUDIA ARABIA PRICE PREDICTIONS



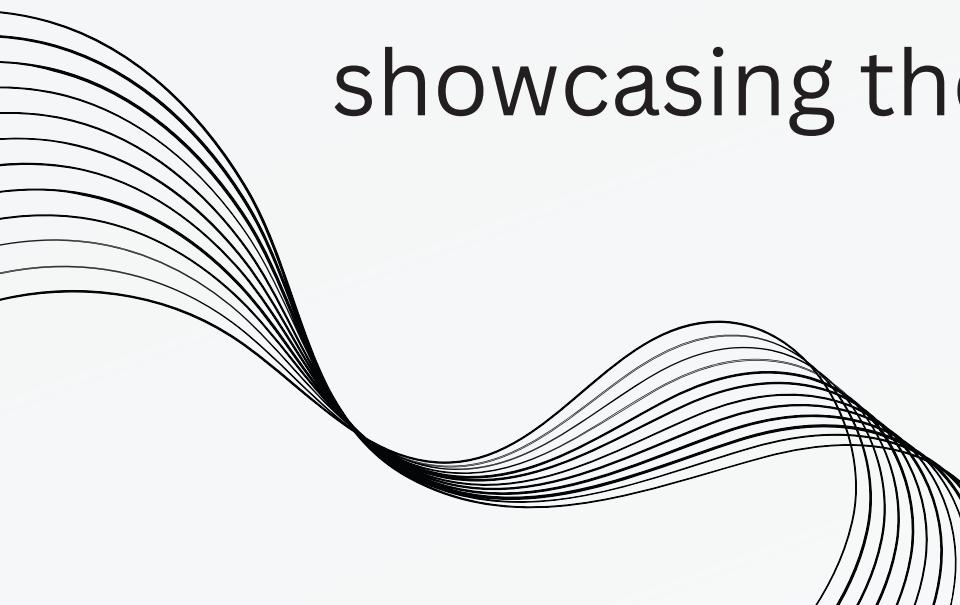
# CONTENT

- 01** INTRODTION
- 02** WEB SCRAPING
- 03** DATA CLEANING
- 04** DATA ANALYSIS
- 05** MACHINE LEARNING MODEL
- 06** RESULTS
- 07** CONCLUSION

# INTRO

our project revolves around extracting and analyzing data from the ArabNB platform, a leading hub for short-term accommodations. Using web scraping techniques with BeautifulSoup, we harvested crucial information such as property names, locations, bed and bedroom counts, pricing, ratings, and comment numbers. This raw data underwent rigorous analysis and cleaning to ensure its quality.

The core of our project lies in the application of machine learning models, specifically decision trees, random forests, and XGBoost. These models were employed to predict trends related to rental properties, achieving an impressive 86% accuracy score. Through our comprehensive approach to data science, we aim to offer valuable insights into the dynamics of short-term rentals on ArabNB, showcasing the transformative potential of data-driven decision-making in the realm of accommodation.

A decorative graphic in the bottom left corner consisting of numerous thin, black, wavy lines that curve and overlap, creating a sense of motion and depth.

# **WEB SCRAPING USING BEAUTIFULSOUP**

we extracting the follow features:

**location of department**

**price**

**numbers of beds**

**numbers of rooms**

**rate**

**numbers of comments**

# DATA CLEANING

*delete nulls value*

*delete unusfull columns*

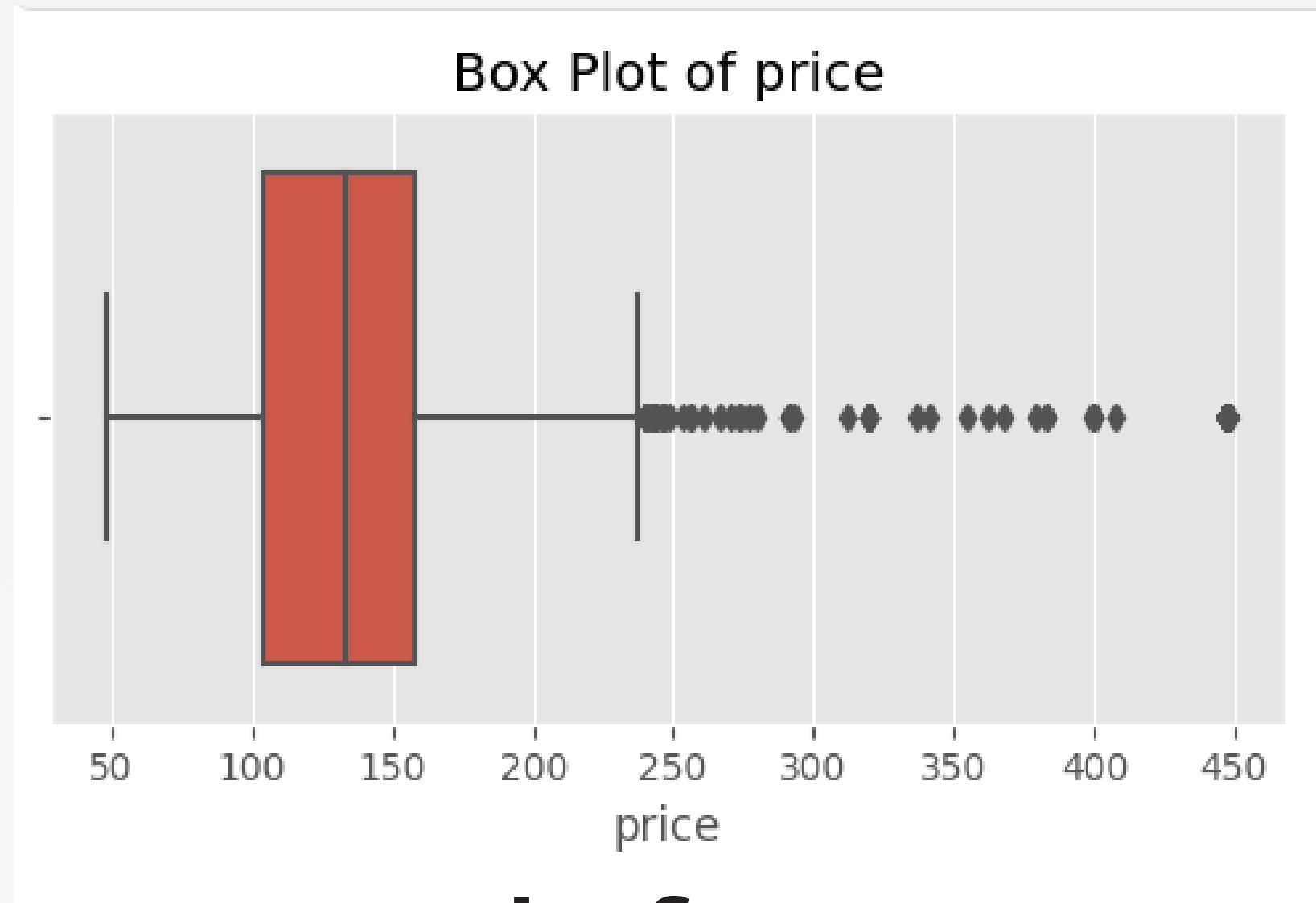
*rename same valeus of columns*

*delete value of space*

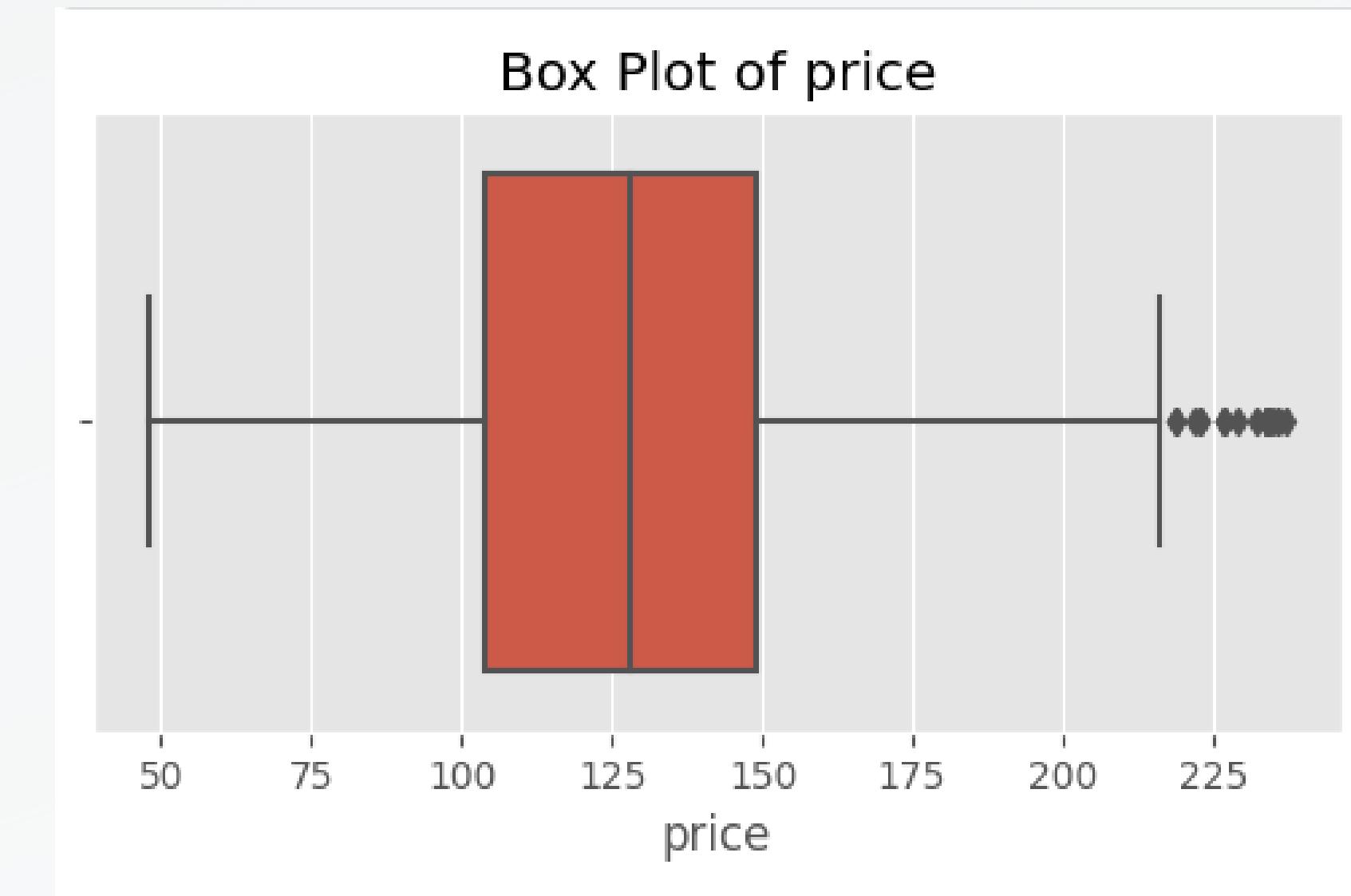


# DATA ANALYSIS

plot the price column and  
delete some outliers value



before

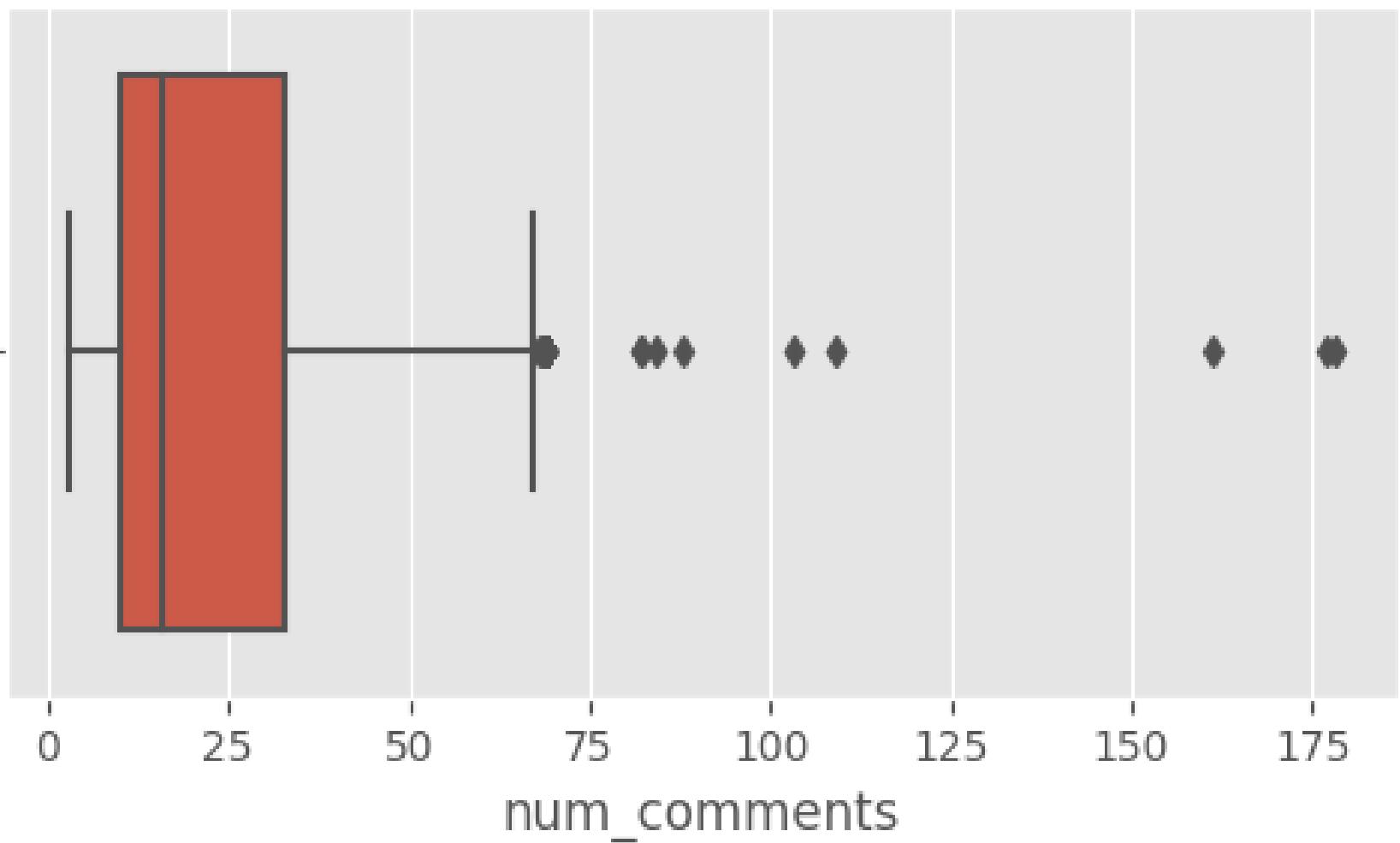


after delete outlliers

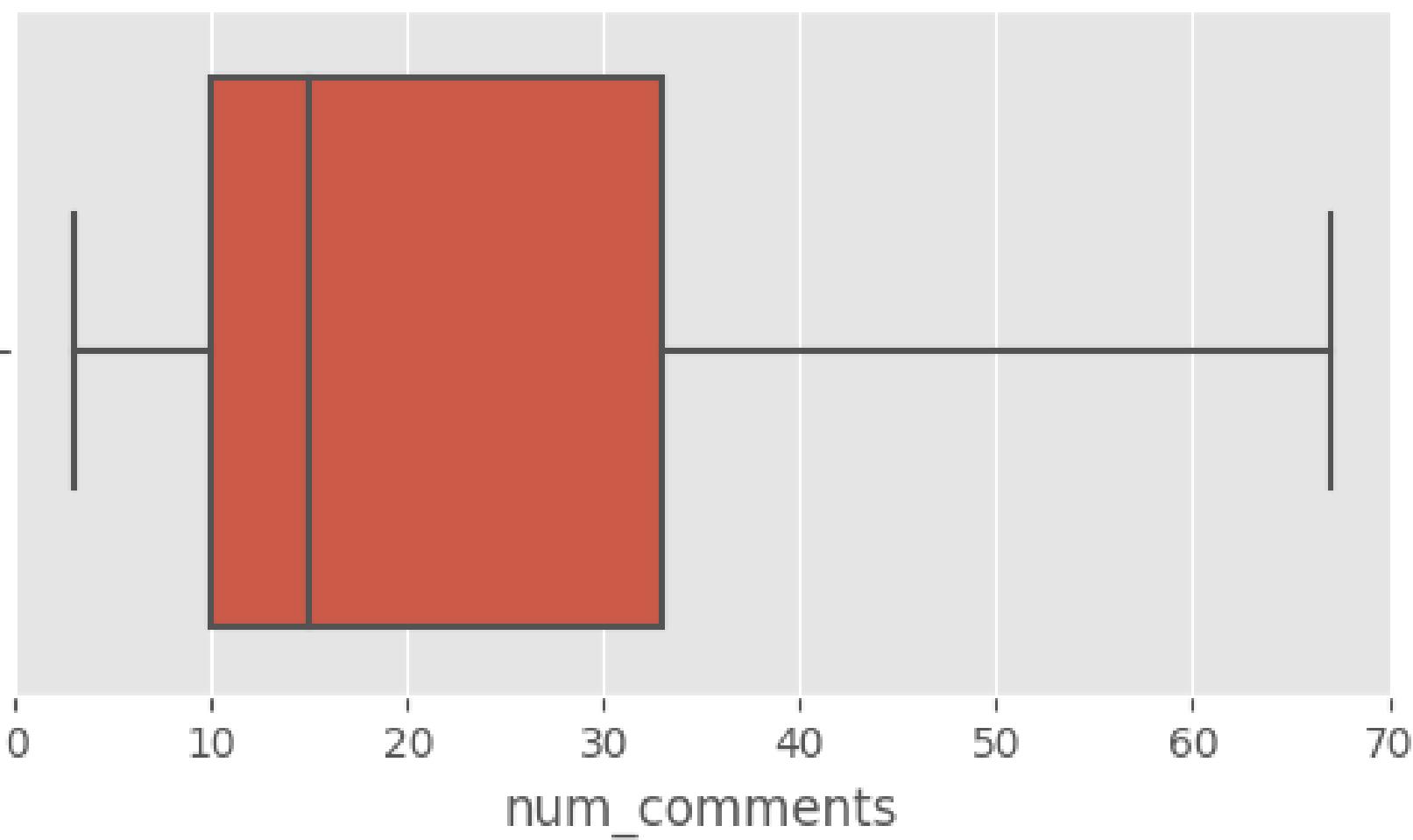
# DATA ANALYSIS

plot the num\_comments with outliers and after delete the outliers

Box Plot of num\_comments



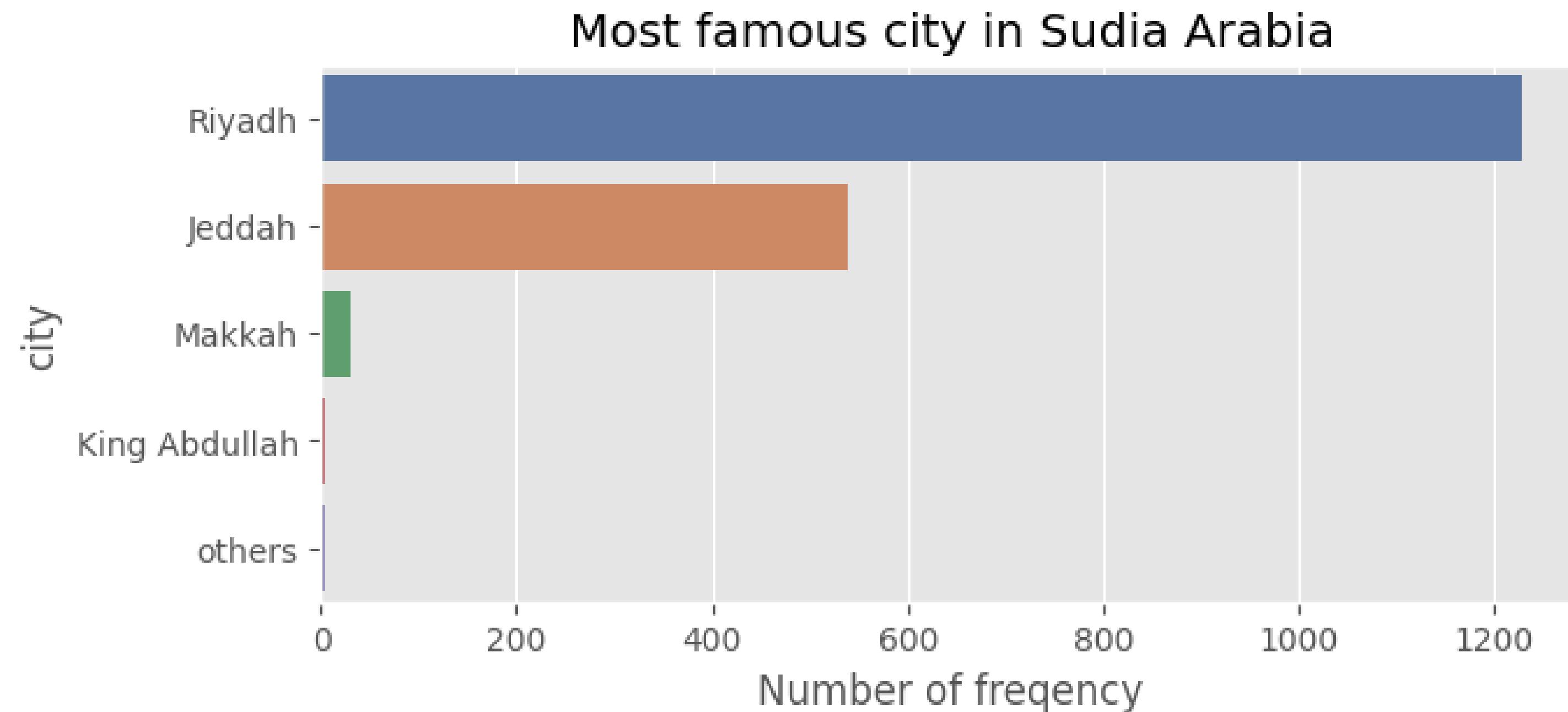
Box Plot of num\_comments



# City in Sudia Arabia

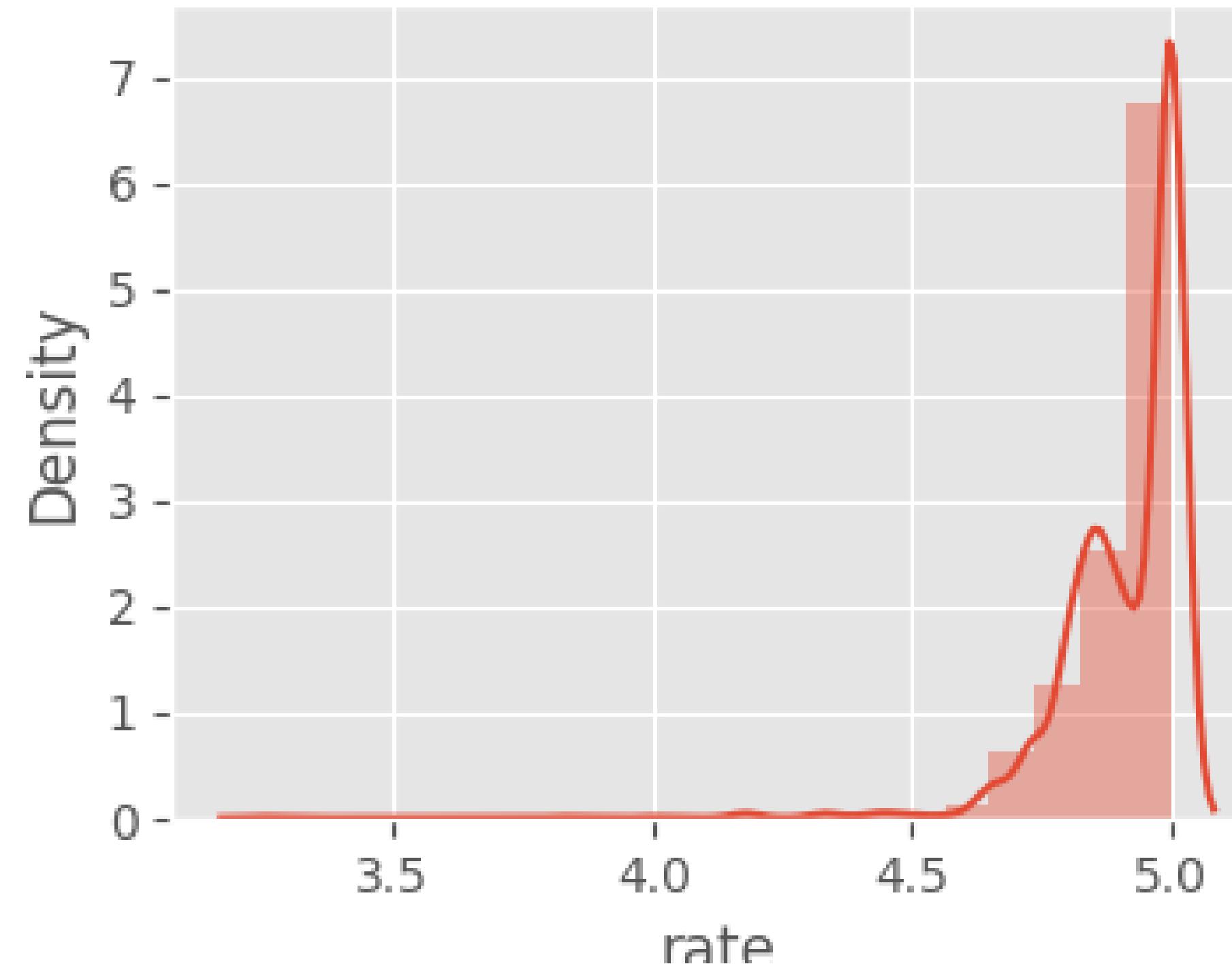
The most city frequency is Riyadh

Text(0.5, 0, 'Number of frequency')

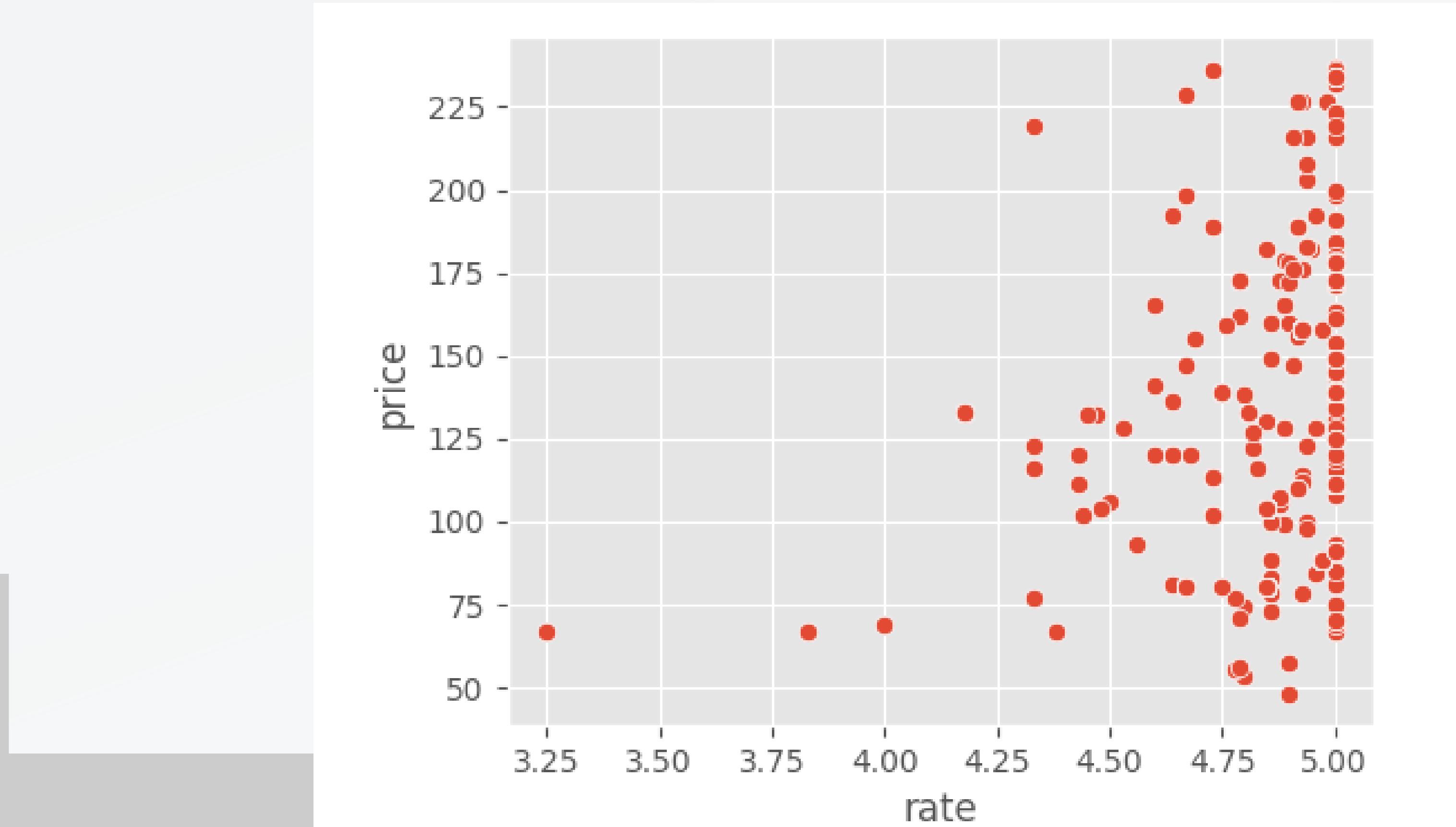


# Rate distribution

## the most rate is five

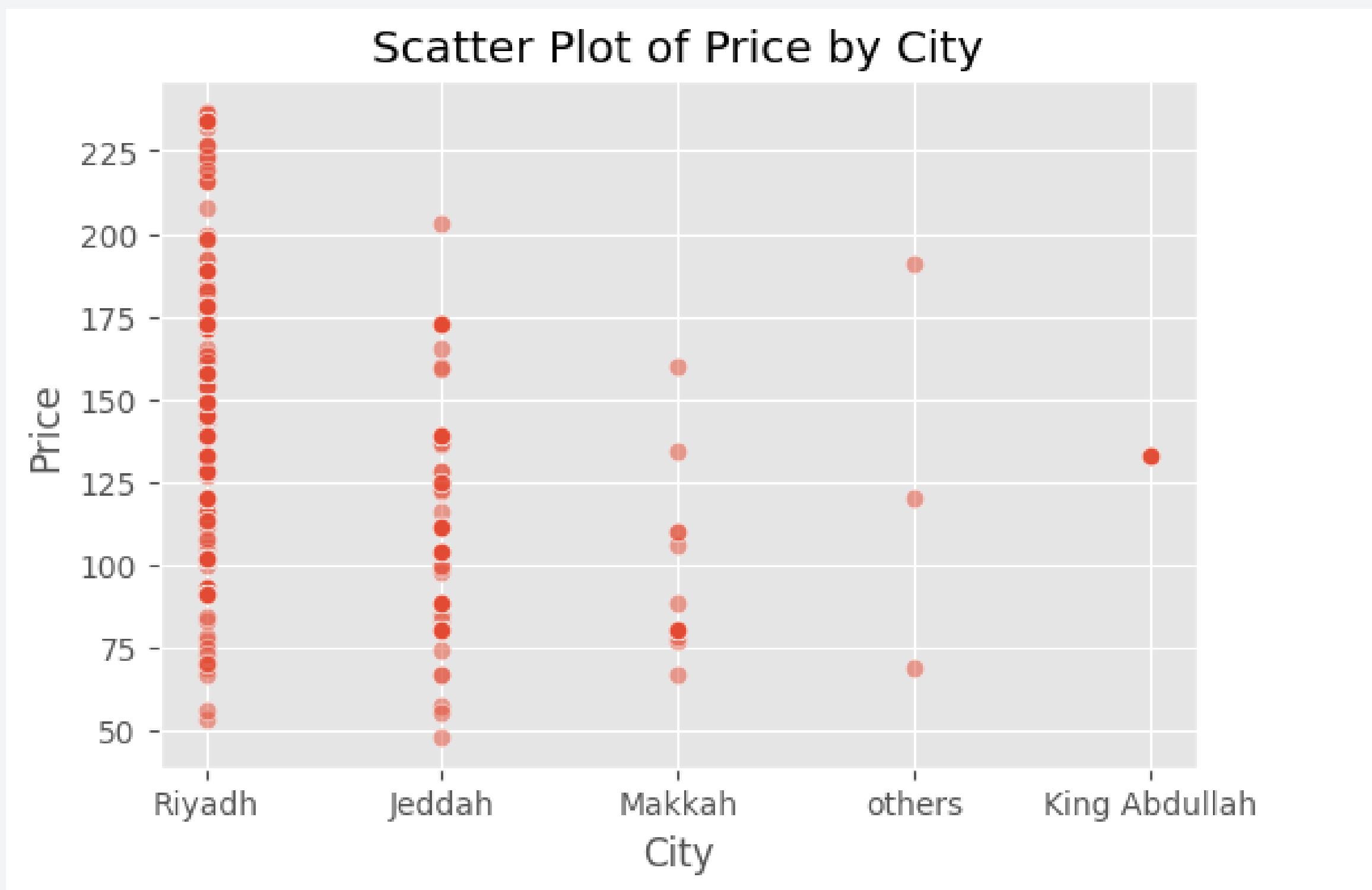


# CORRLATION BETWEEN THE PRICE AND THE RATE COLUMN

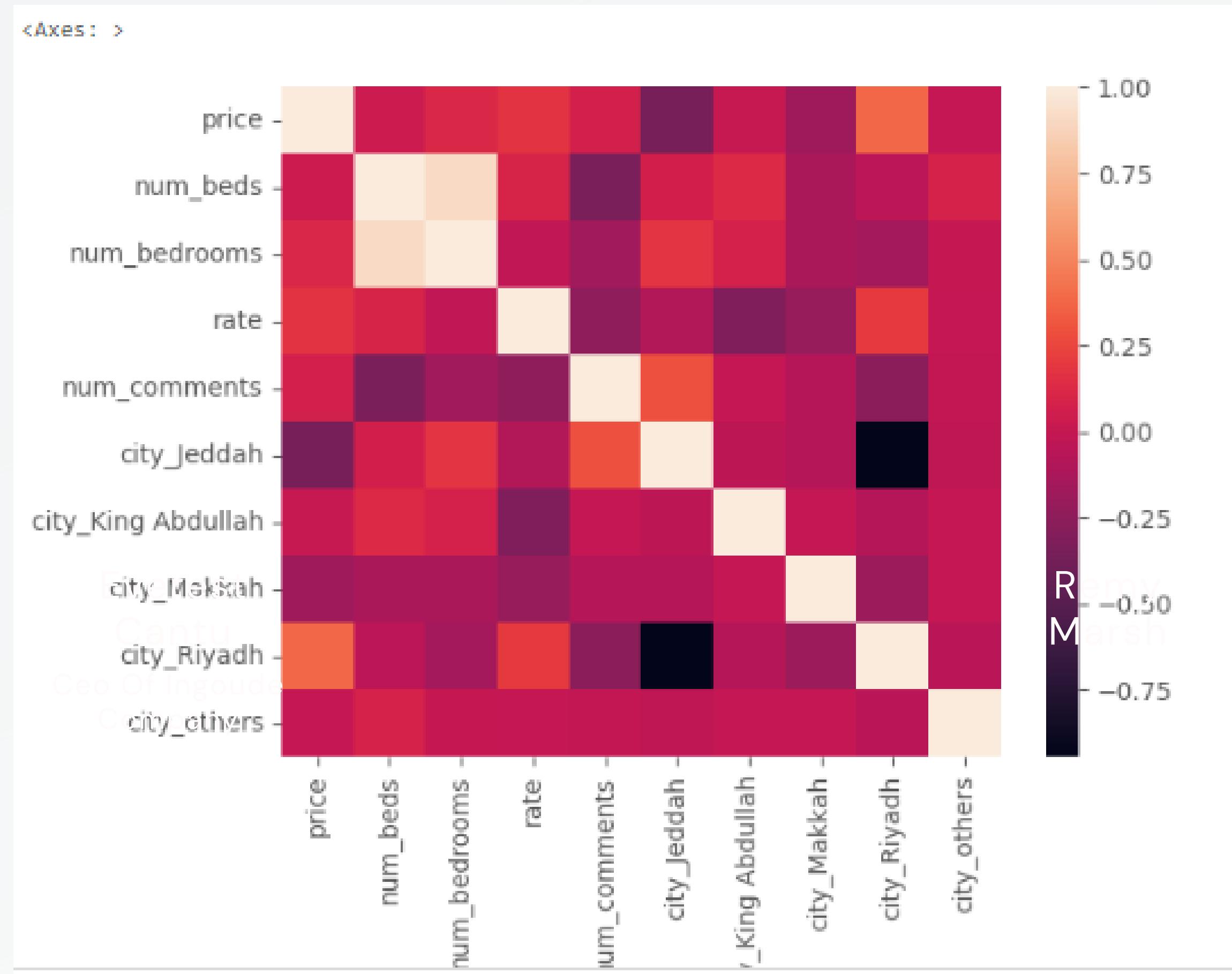


# THE CORRLATION BETWEEN THE CITY AND THE PRICE

Riyadh has the expensive price in Sudia Arabia

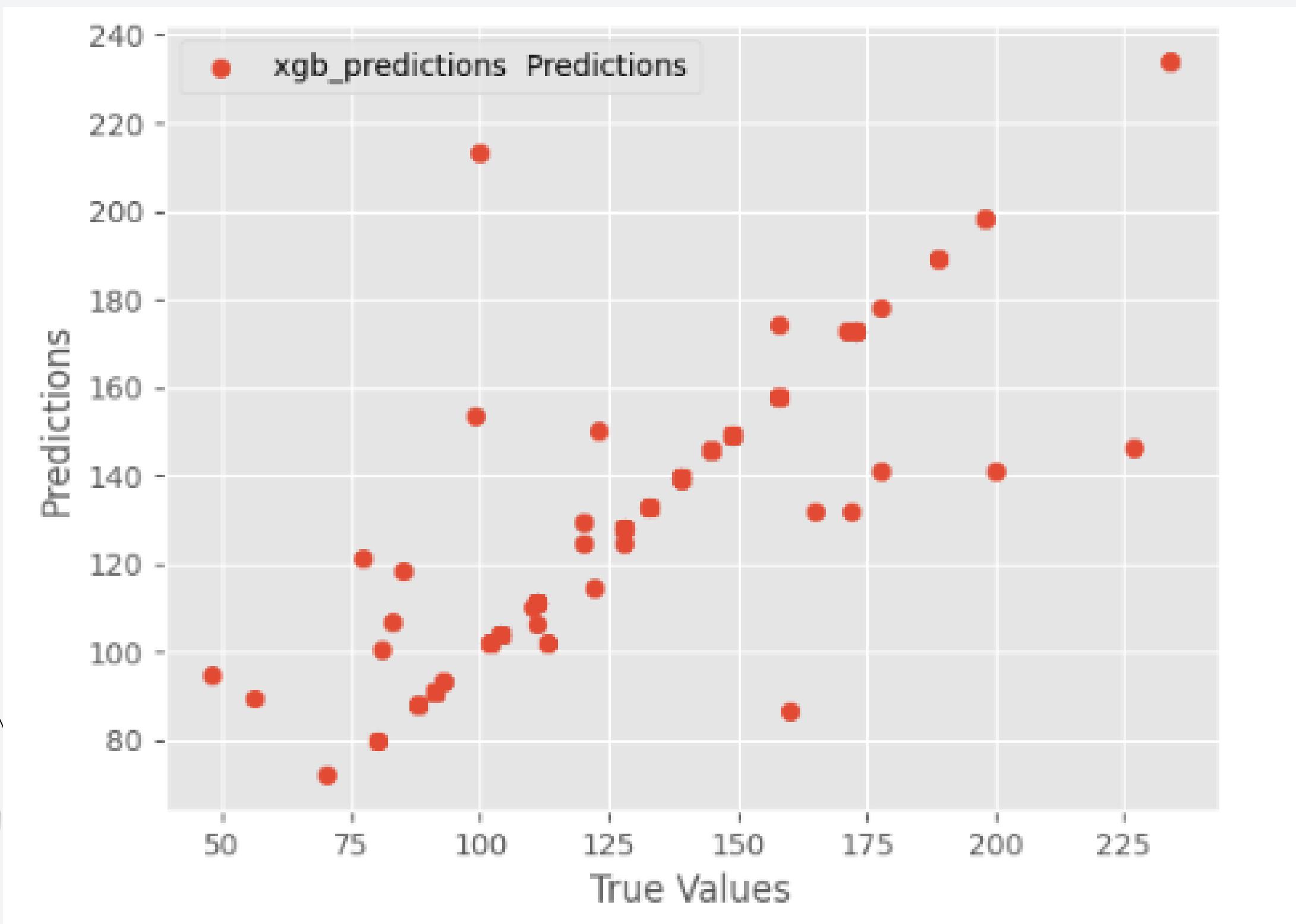


# HEATMAP



# THE MACHING LEARNING MODEL

WE USE MULTI MODEL THE BEST MODEL IS XGBOOST THE  
ACCURACY IS 86 %



**finish**