



**SDAIA**  
الهيئة السعودية للبيانات  
والذكاء الاصطناعي  
Saudi Data & AI Authority



# **PREDICT THE RENTAL PRICE OF THE SHOPS IN RIYADH**

**By: Shahad Almutairi & Sahar Alharthi**  
**Instructor: Dr.Majdal Alqahtani**

## إستراتيجية تطوير مدينة الرياض

أعلن سمو ولي العهد خلال مشاركته في الدورة الرابعة لمبادرة مستقبل الاستثمار عن خطة المملكة في إعلان إستراتيجية تطوير مدينة الرياض كجزء من خططها لتنويع مصادر الدخل ونمو الاقتصاد.

### من تصريحات سموه

◀ "نستهدف أن تكون الرياض من أكبر عشر مدن اقتصادية في العالم، اليوم هي رقم أربعين، من أكبر أربعين اقتصاد في العالم كمدينة، نستهدف في الرياض أن نصل من 7.5 ملايين نسمة إلى ما بين 15 و 20 مليون نسمة في 2030."

◀ "كل الخصائص التي تمتلكها الرياض تعطي ممكّنات لخلق وظائف وخلق نمو في الاقتصاد وخلق استثمارات وخلق العديد من الفرص، لذلك ننظر للرياض بعين الاعتبار."

◀ "وبما أن المدن تشكل 85% من اقتصاد العالم، فلذلك التنمية الحقيقية تبدأ من المدن سواء في الصناعة أو الابتكار أو في التعليم أو في الخدمات أو في السياحة وغيرها من القطاعات."

### من مزايا مدينة الرياض

- ◀ تشكّل 45% من الاقتصادي السعودي
- ◀ تشكّل أكثر من 30% من سكان المملكة
- ◀ تشكّل 50% من الاقتصاد غير النفطي في المملكة
- ◀ تكلفة خلق الوظيفة فيها أقل 30% من بقية مدن المملكة
- ◀ تكلفة التطوير العقاري والبنى التحتية أقل بـ 29% من بقية مدن المملكة

”  
البنية التحتية في الرياض رائعة جداً بسبب ما قام به الملك سلمان  
فيما يزيد عن 55 سنة بإدارة مدينة الرياض والتخطيط لها.

– سمو ولي العهد

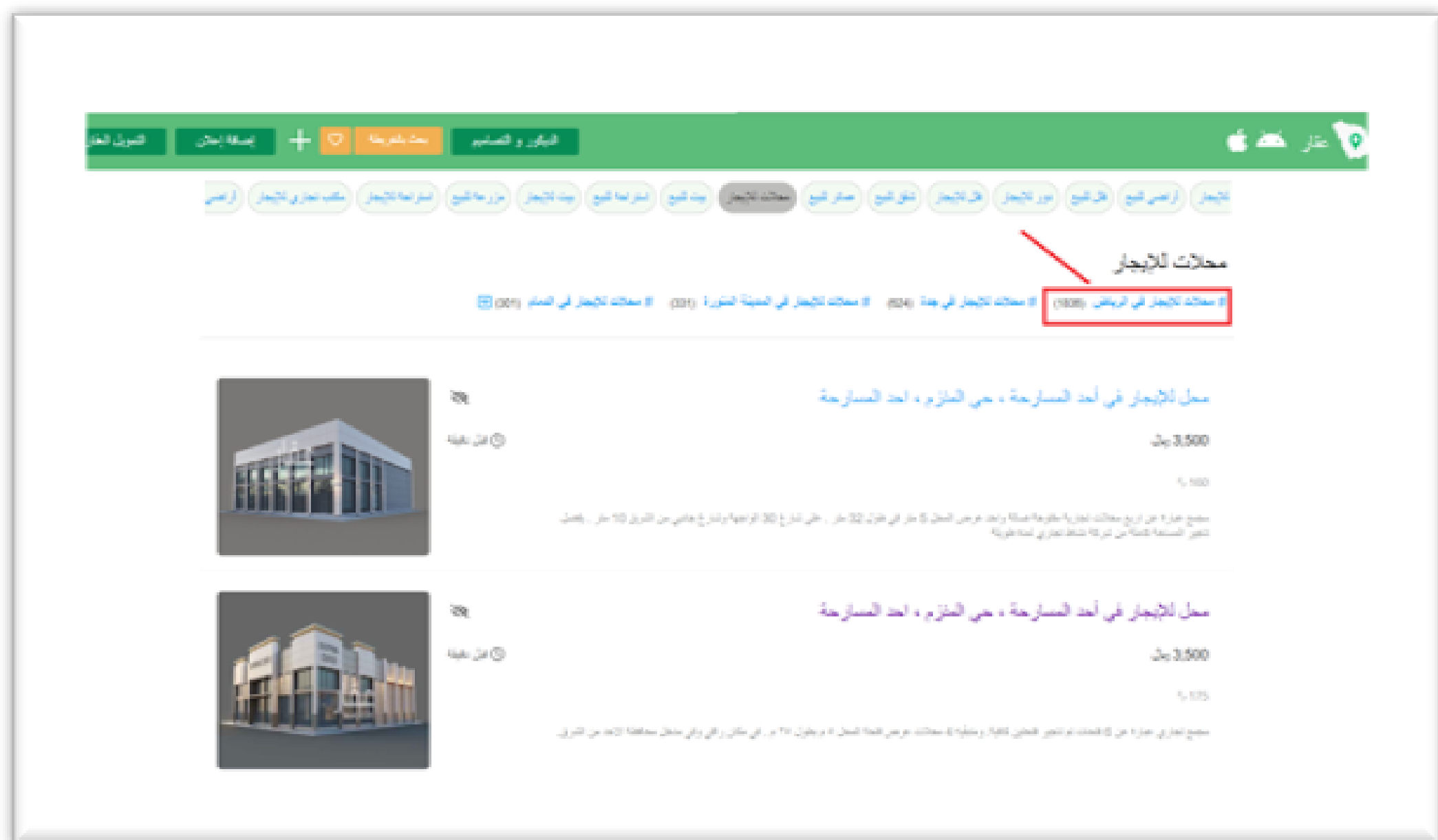
# Project Goal

Helping the investor to predict the rental  
price of the shop.

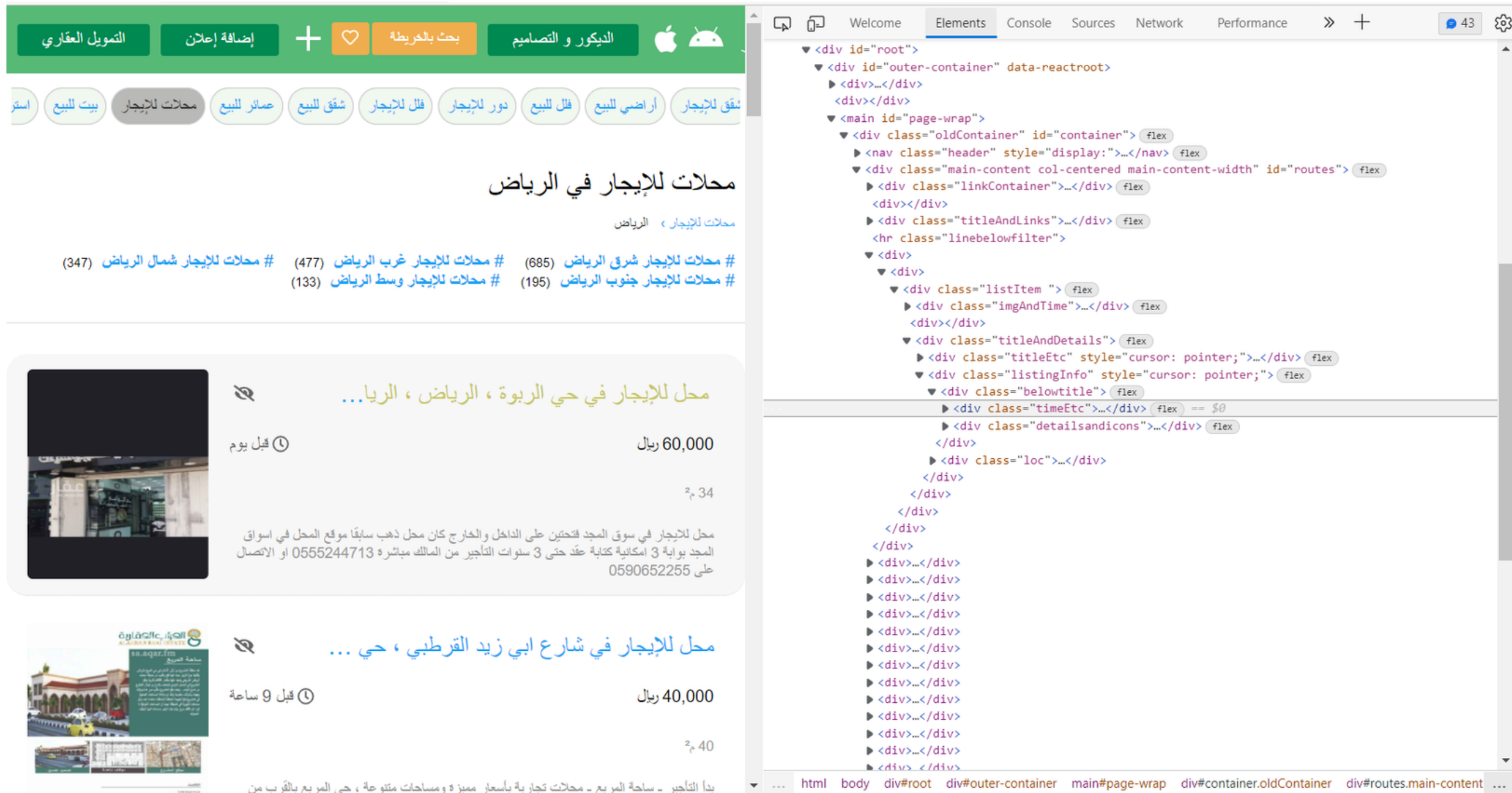


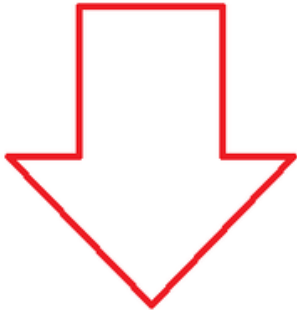
# Dataset

## Web Scraping



## Three overlapping hexagons in light blue, purple, and dark blue.





60,000 ريال سعودي

محل للايجار في سوق المجد  
فتحتين على الداخل والخارج  
كان محل ذهب سابقا  
موقع المحل في اسواق المجد بوابه 3  
امكانية كتابة عقد حتى 3 سنوات  
التاجير من المالك مباشرة  
0555244713  
او الاتصال على 0590652255

المساحة	34 م²
الواجهة	جنوب شرقي
عرض الشارع	100 م \ /
عمر العقار	20 سنة

#3682267

123



قبل يوم





# EDA

## Data Cleaning

460 rows x 7 columns

**Concat Dataframes:**

```
In [28]: df_Riyadh_shops = pd.concat([df_Center, df_North, df_West, df_South, df_East])
df_Riyadh_shops
```

Out[28]:

	shoptitle	prices	areaa	front	street_width	age	Region
0	محل للتجارة في شارع شجرة العنبري - حي النور	60,000 ريال	100 م²	شرق	30 م	جديد	Center
1	محل للتجارة في طريق مصر بن عبد العزيز الغربي	276,000 ريال	460 م²	جنوب	60 م	جديد	Center
2	محل للتجارة في شارع الأمير متعب بن عبد العزيز	80,000 ريال	125 م²	غرب	36 م	سنة 6	Center
3	محل للتجارة في شارع احمد بن فارس - حي النور	16,000 ريال	50 م²	شرق	20 م	سنة 5	Center
4	محل للتجارة في شارع جريز - حي جريز - الرياض	17,500 ريال	70 م²	شمال	NaN	NaN	Center
...	...	...	...	...	...	...	...
455	محل للتجارة في شارع حفصة بنت عمر - حي الروضة	13,000 ريال	40 م²	شمال	NaN	NaN	East
456	محل للتجارة في شارع جريز - حي العسكر - الرياض	55,000 ريال	8 م²	جنوب	30 م	سنة 15	East
457	محل للتجارة في 7331-7307 شارع عبدالله بن عدا	25,000 ريال	72 م²	غرب	30 م	سنة 18	East
458	محل للتجارة في حي القادسية - الرياض - الرياض	250,000 ريال	1000 م²	شرق	40 م	جديد	East
459	محل للتجارة في شارع بصر الغرب - حي التنبلية	30,000 ريال	80 م²	شرق	40 م	سنة 2	East

1540 rows x 7 columns

```
In [23]: df_Riyadh_shops.to_csv("Riyadh_shopss.csv", encoding="utf-8")

In [24]: duplicate = df_Riyadh_shops.duplicated()
print(duplicate.sum())
```



prices areaa street\_width age

Out[81]:

```
df_Riyadh_shops['age'].unique()
array([ 1.,  6.,  5., nan, 10., 30., 12.,  7., 15., 26., 25., 13., 20.,
        17., 27.,  2., 28., 35.,  4., 23., 18., 14.,  9.,  8., 36.,  3.,
        24., 32., 19., 21., 11., 22., 29., 33., 16., 31.])
```

```
In [82]: df_Riyadh_shops=df_Riyadh_shops.reset_index(drop=True)
```

```
In [85]: df_Riyadh_shops
```

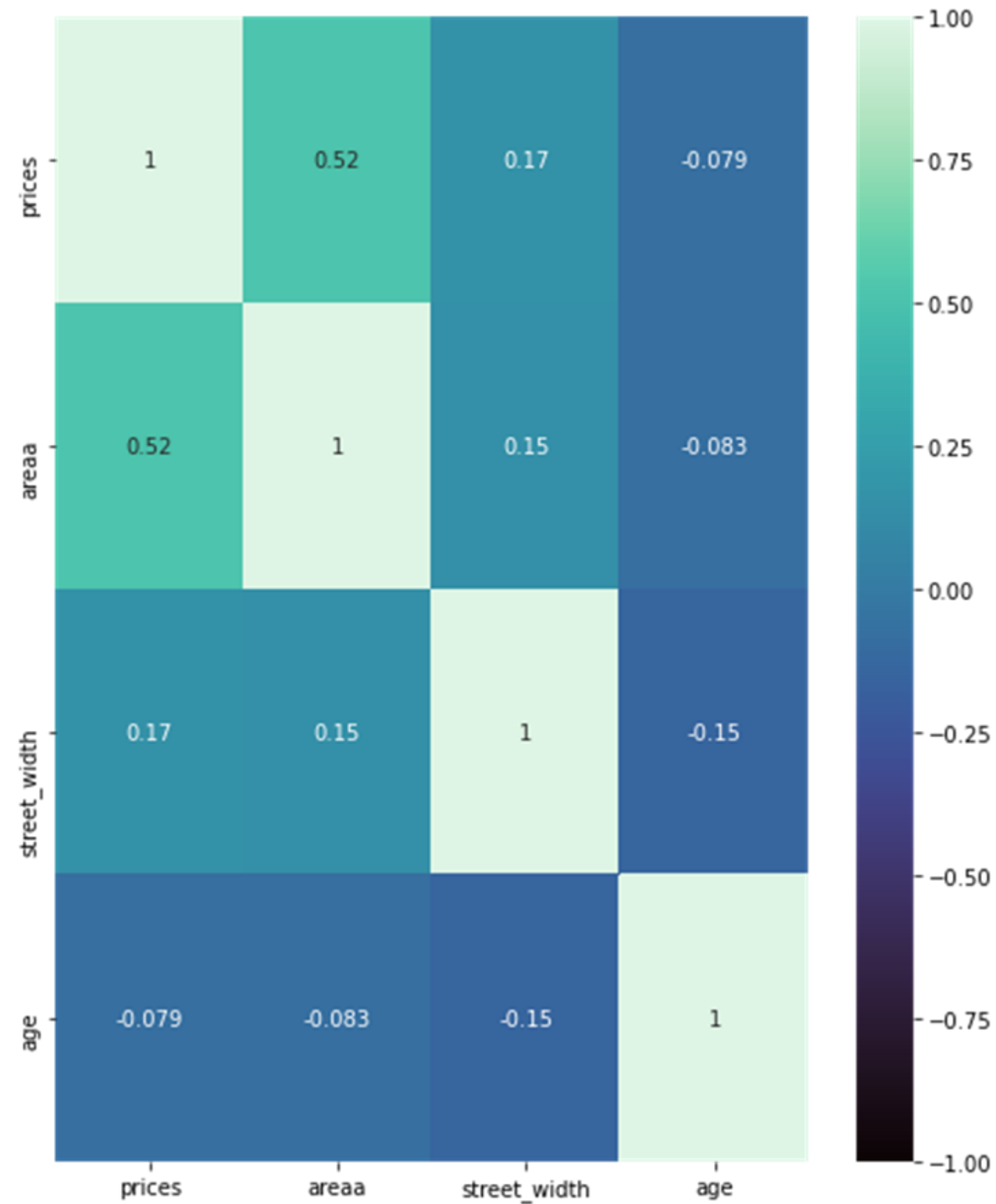
Out[85]:

	prices	areaa	front	street_width	age	Region
0	60000	100	شرق	30.0	1.0	Center
1	276000	460	جنوب	60.0	1.0	Center
2	80000	125	غرب	36.0	6.0	Center
3	16000	50	شرق	20.0	5.0	Center
4	17500	70	شمال	NaN	NaN	Center
...	...	...	...	...	...	...
1307	150000	233	شرق	NaN	NaN	East
1308	35000	56	شرق	NaN	NaN	East
1309	55000	8	جنوب	30.0	15.0	East
1310	250000	1000	شرق	40.0	1.0	East
1311	30000	80	شرق	40.0	2.0	East

1312 rows x 6 columns



# EDA



• Heat Map



# Regression Models

**R-squared**  
Cross-Validation

**19%**

**Ridge**

**52%**

**Polynomial**

**19%**

**Linear Regression**



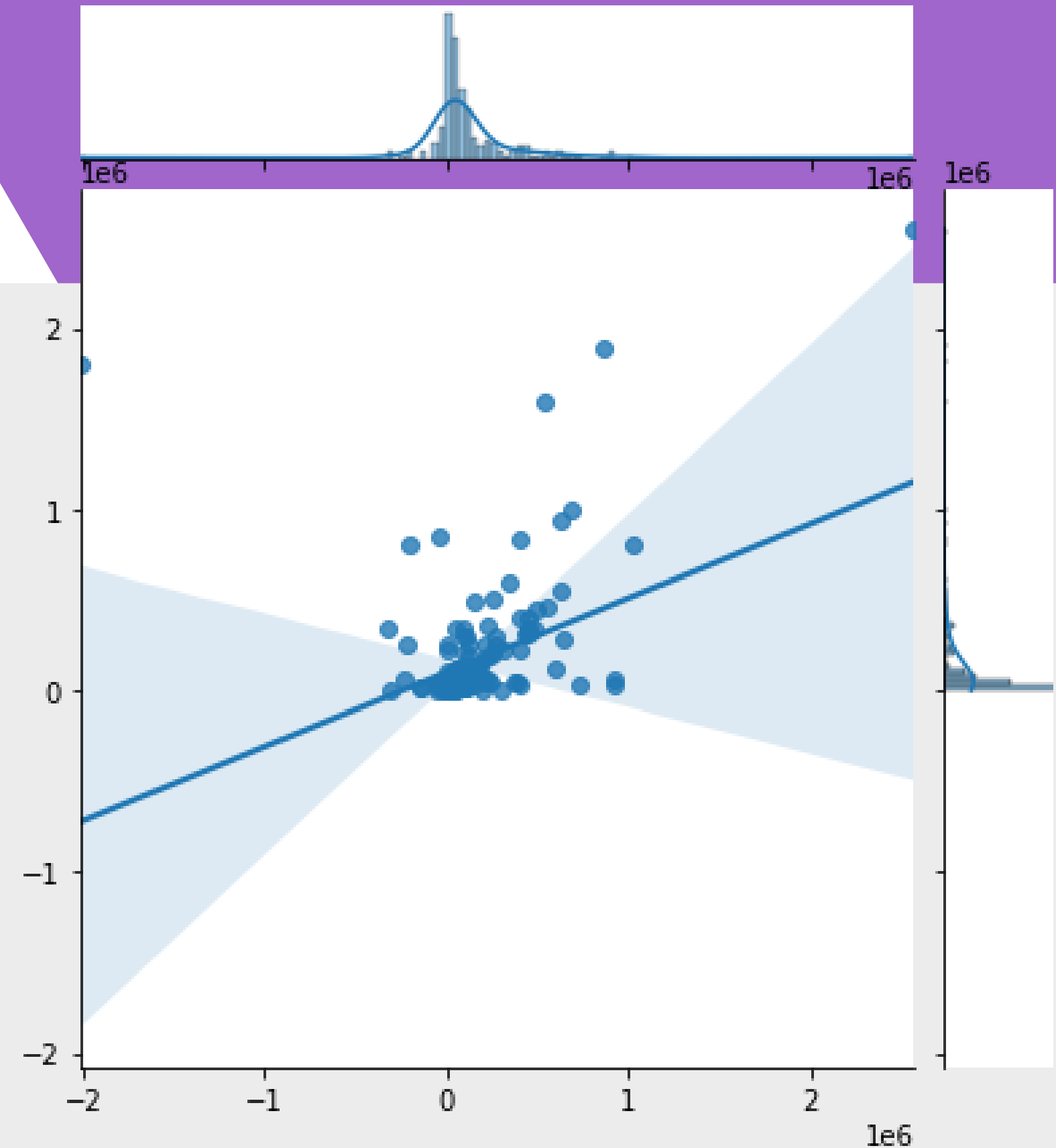
# Polynomial result:

R-squared for  
Training:

69%

R-squared for  
Testing:

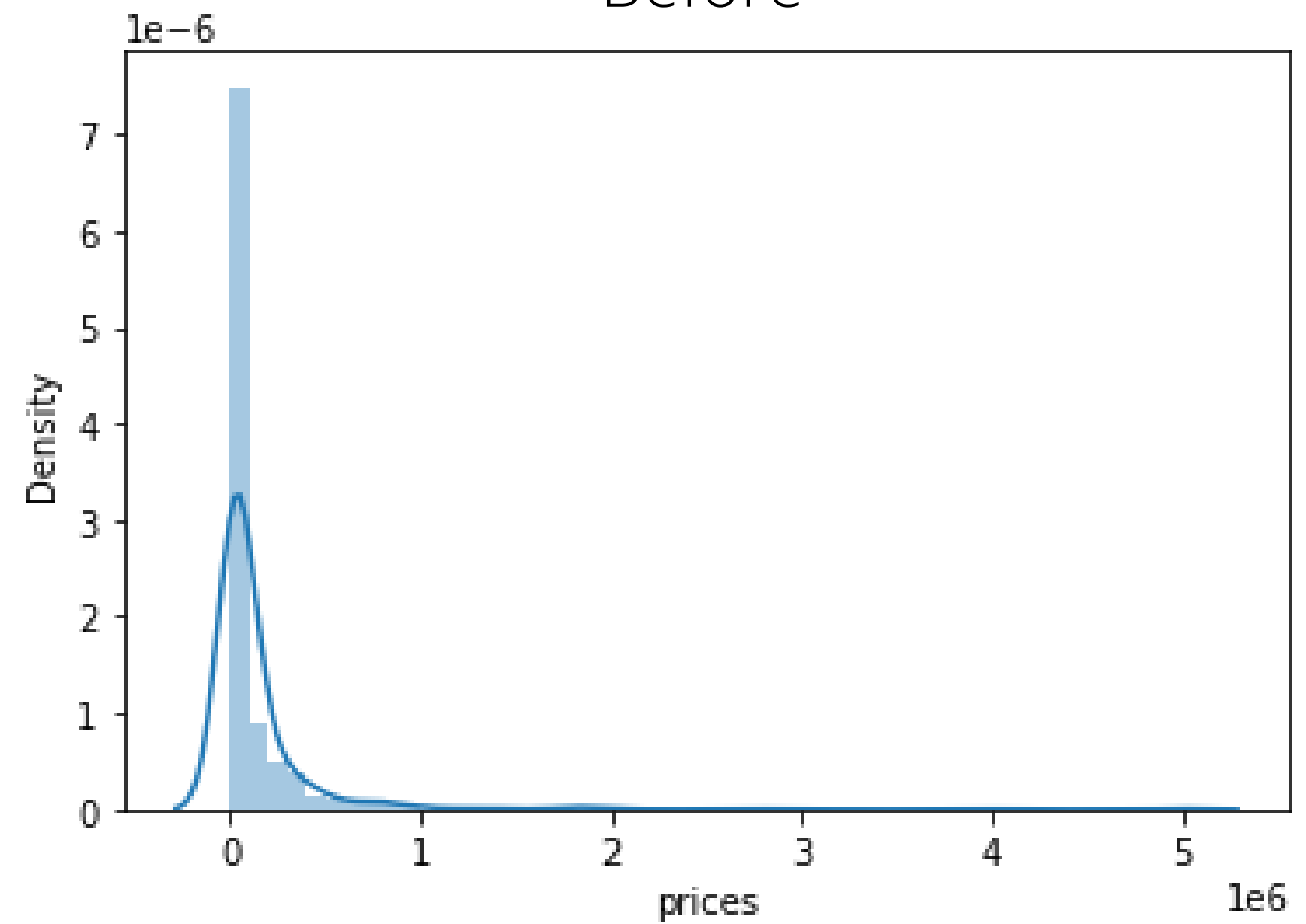
-17%



# Experiment

Use log (Prices)

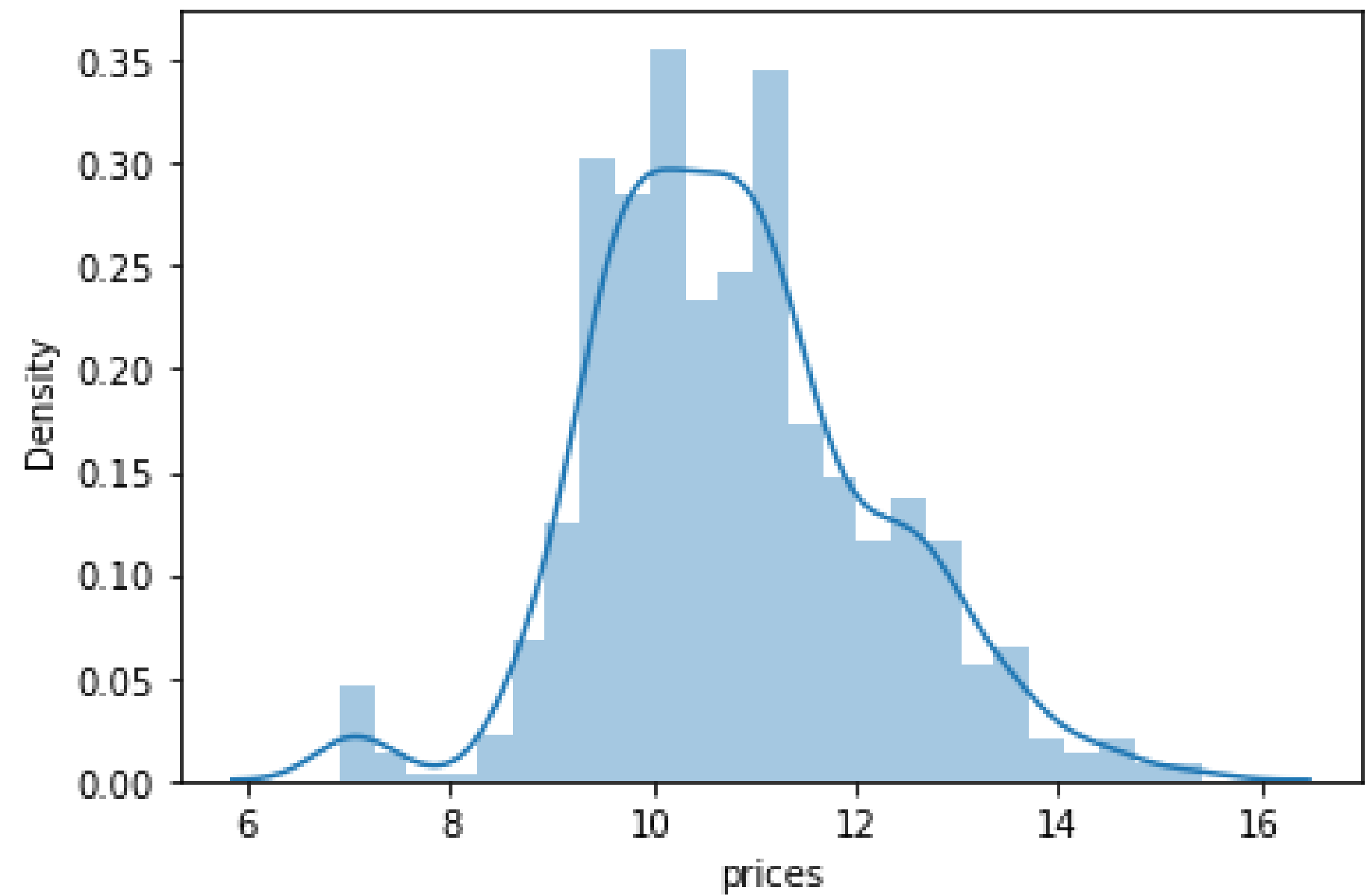
Before



# Experiment

Use log (Prices)

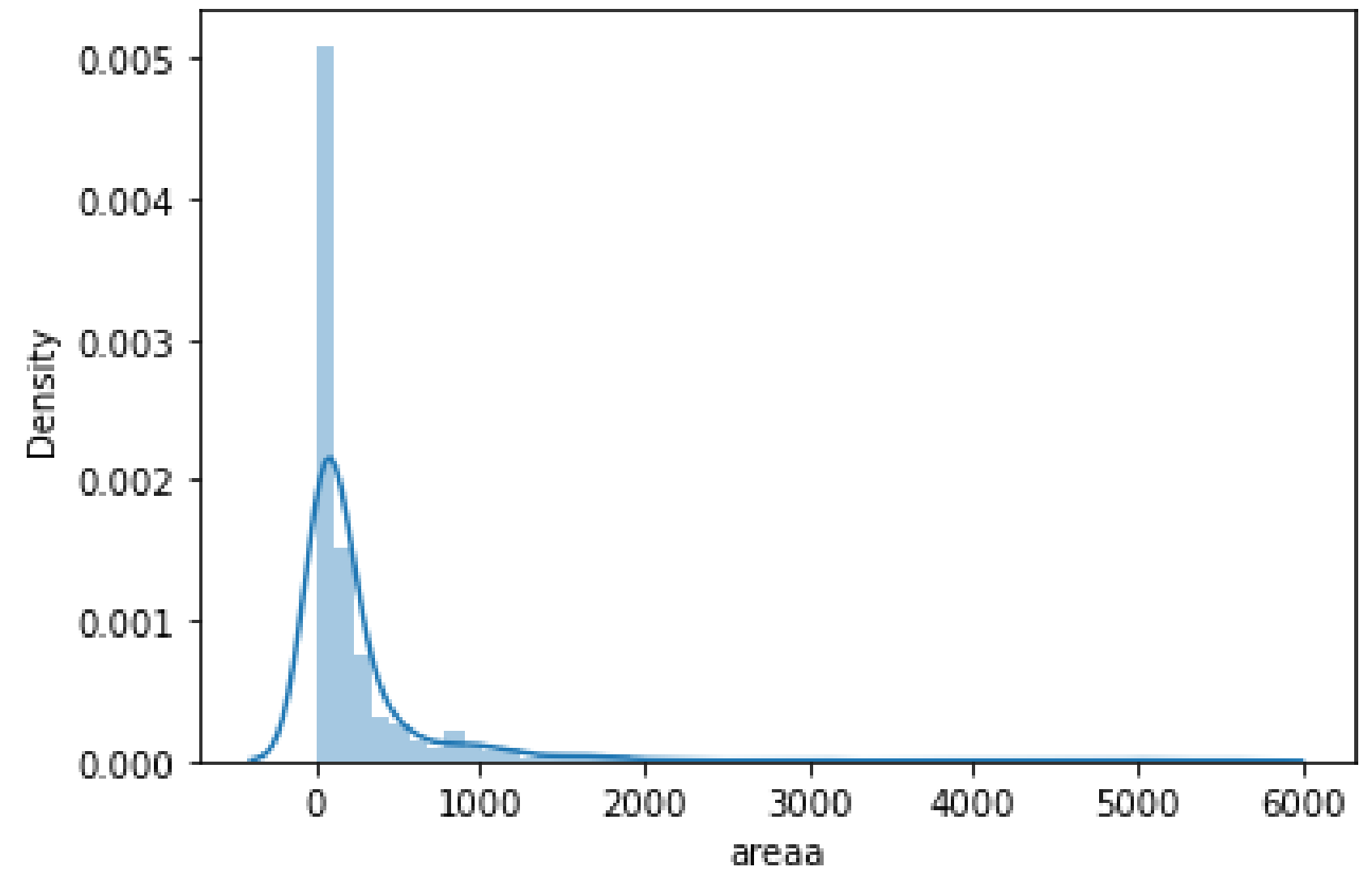
After



# Experiment

Use log (Area)

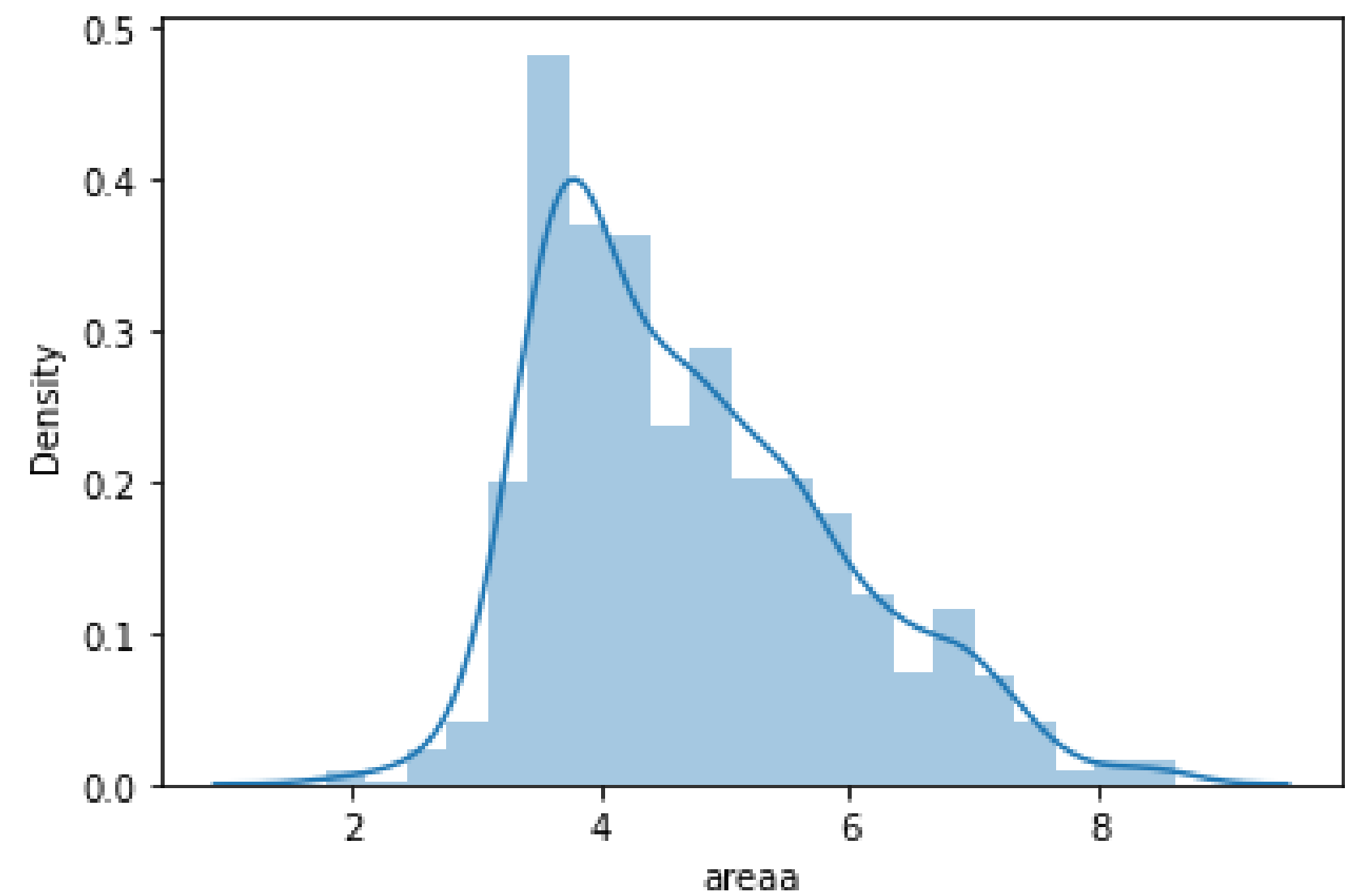
Before



# Experiment

Use log (Area)

After

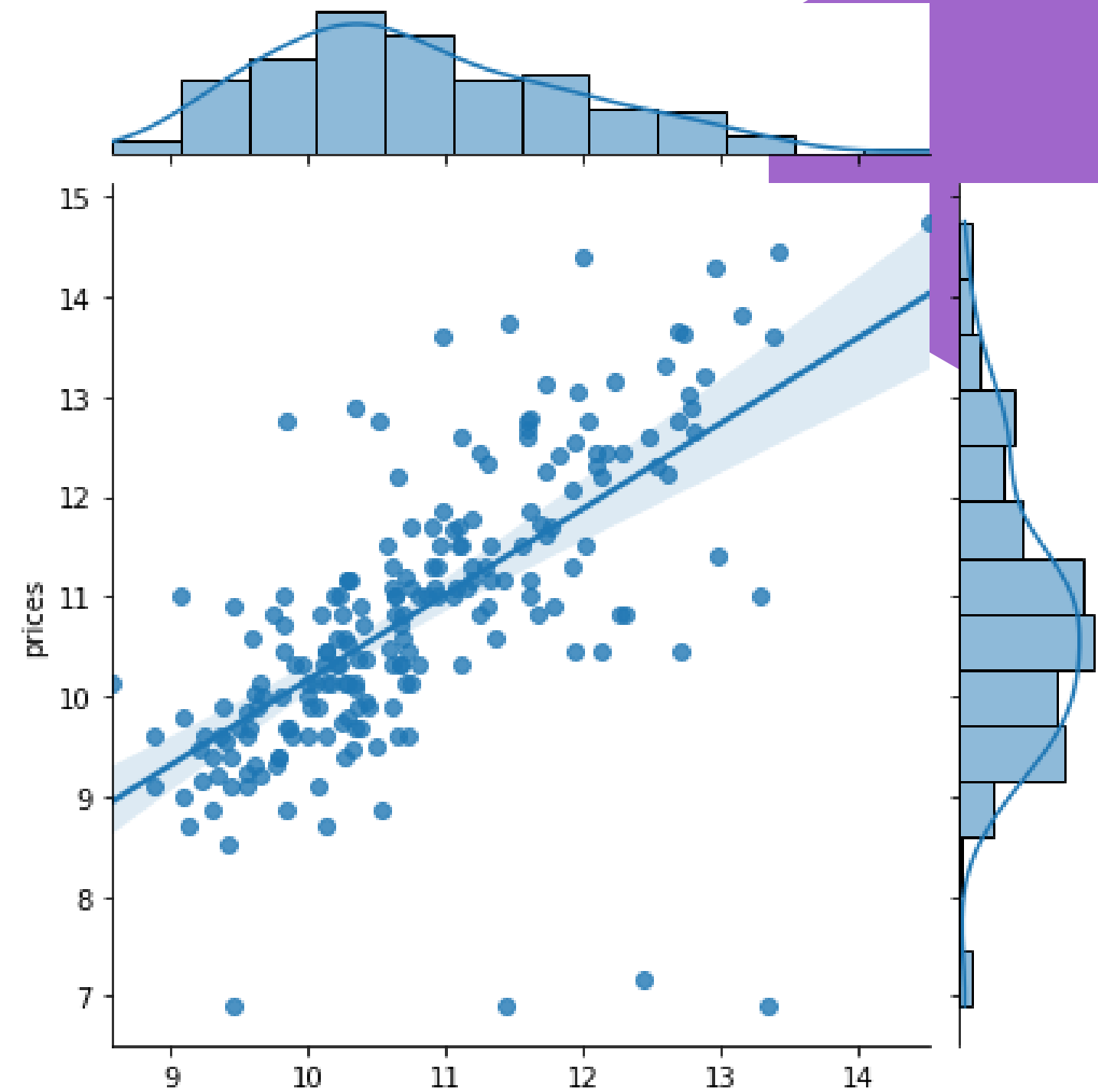


# Prediction after the experiment

Training: 61%

.....

Testing: 43%

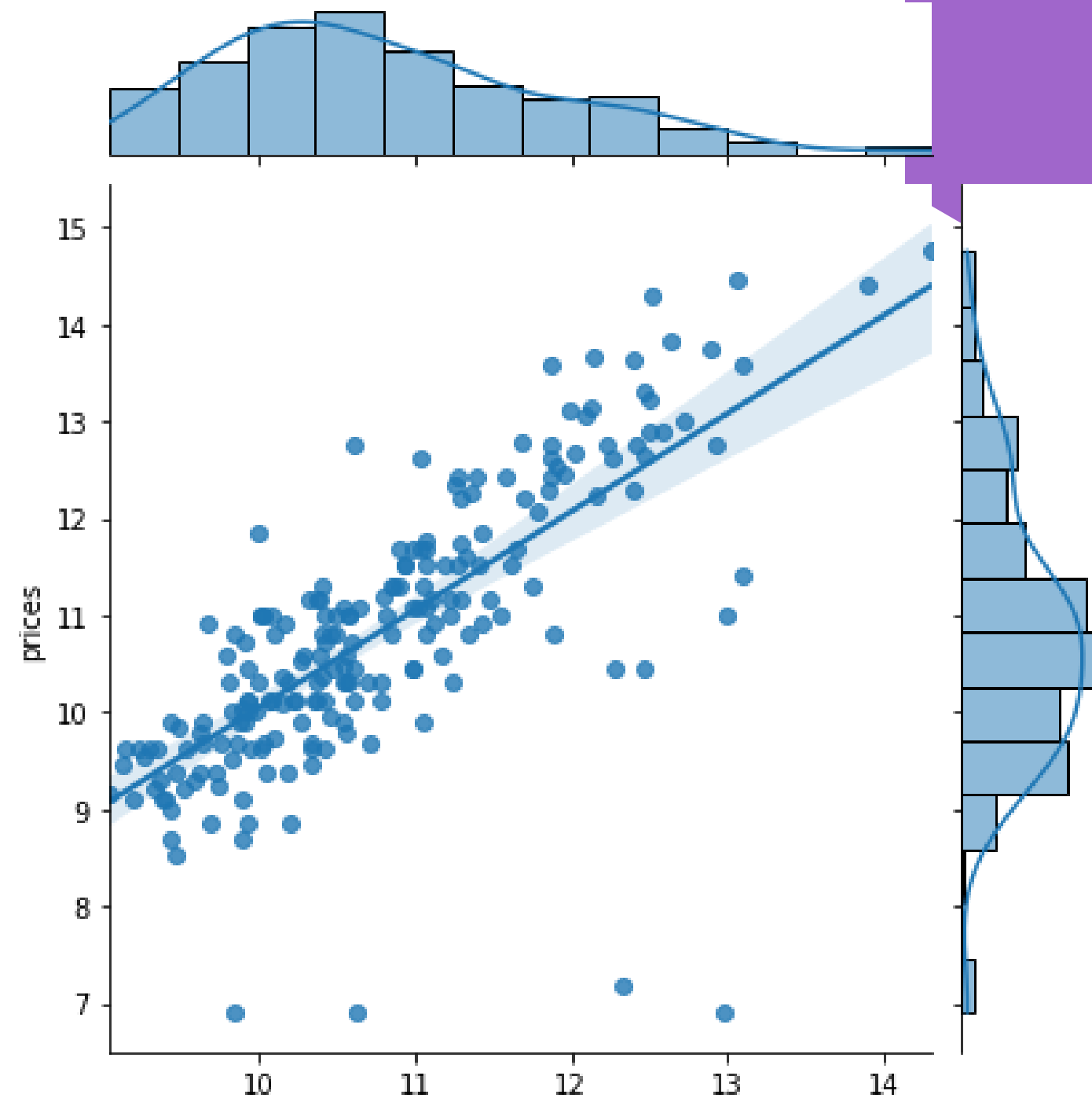


# Prediction after the experiment(Use Ridge Model)

Training: 56%

.....

Testing: 54%





# Conclusion

When increase the number of data...  
In the future this model will help the investor to achieve  
a careful feasibility study with the least possible effort  
and with high accuracy

# Thank you!