



# **AIRBNB FINDINGS**

Barcelona & Atenas

Miroslava Arredondo  
Sarahí Gómez  
Alexa Gómez  
Mariana Torres

# Objetivo

Extraer las características más relevantes del archivo “listings.csv”, aplicando las acciones de preprocesamiento necesarias, como la limpieza de valores nulos y el manejo de outliers, para analizar y comparar datos de dos ciudades distintas obtenidos del sitio de Airbnb, con el objetivo de llegar a conclusiones significativas



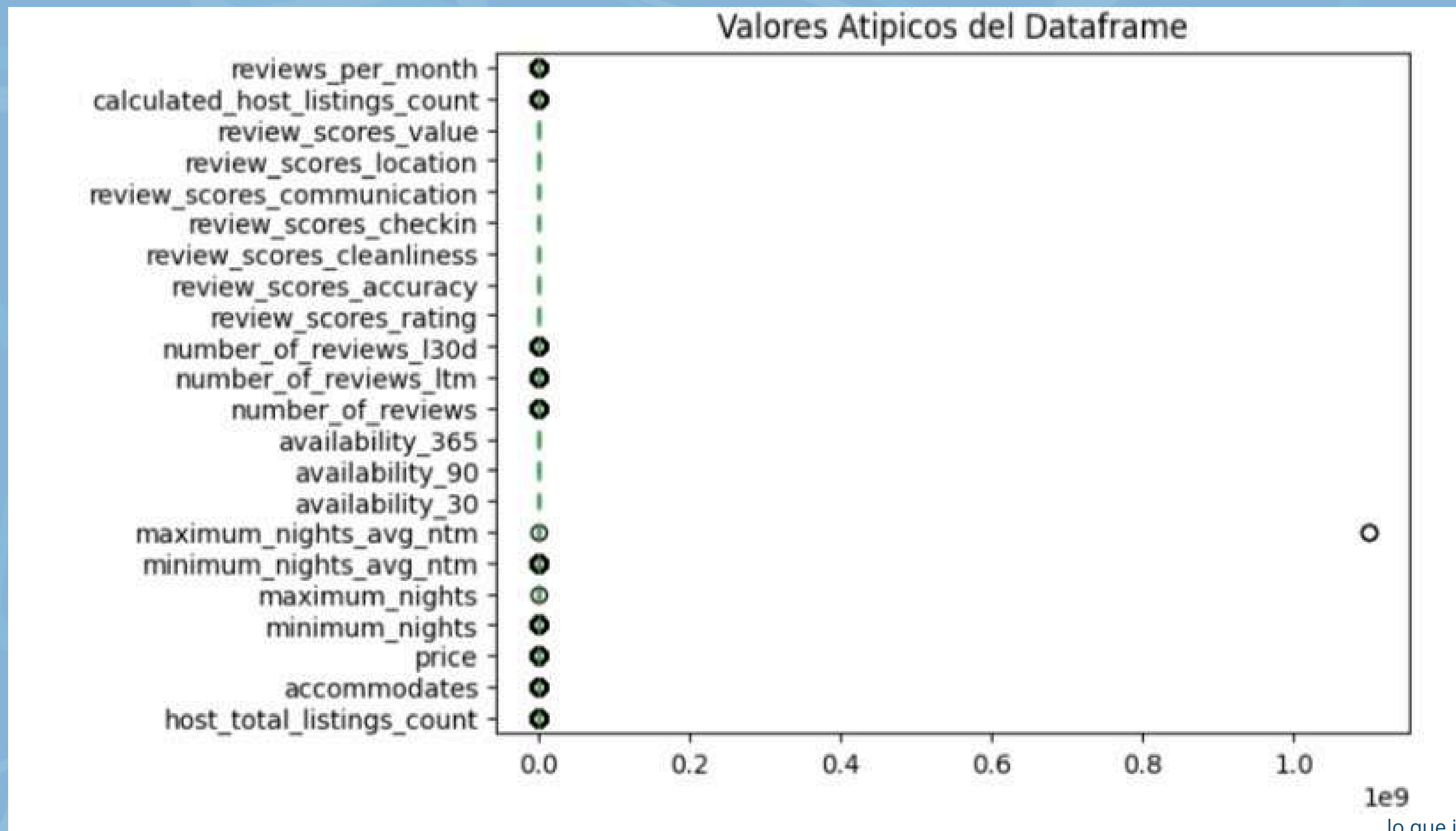
# CIUDADES ELEGIDAS



**BARCELONA**



**ATENAS**



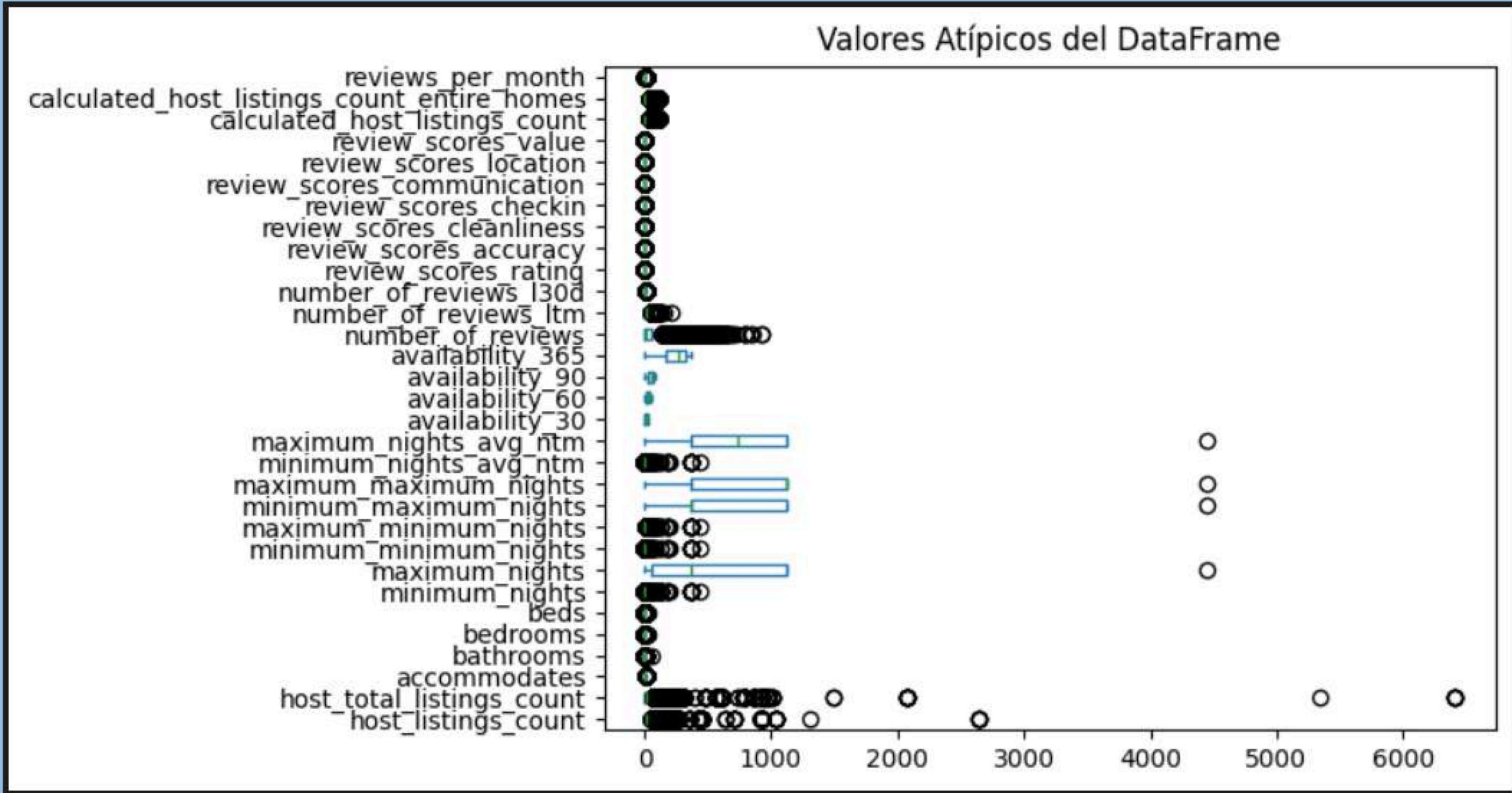
lo que indica una fuerte relación lineal positiva entre estas dos variables.  
Relación Lineal Perfecta en Visualización, pero no en Valores:

## OUTLIERS

La regresión puede ayudar a entender si existe una relación entre la percepción de la limpieza y el precio que los huéspedes están dispuestos a pagar.

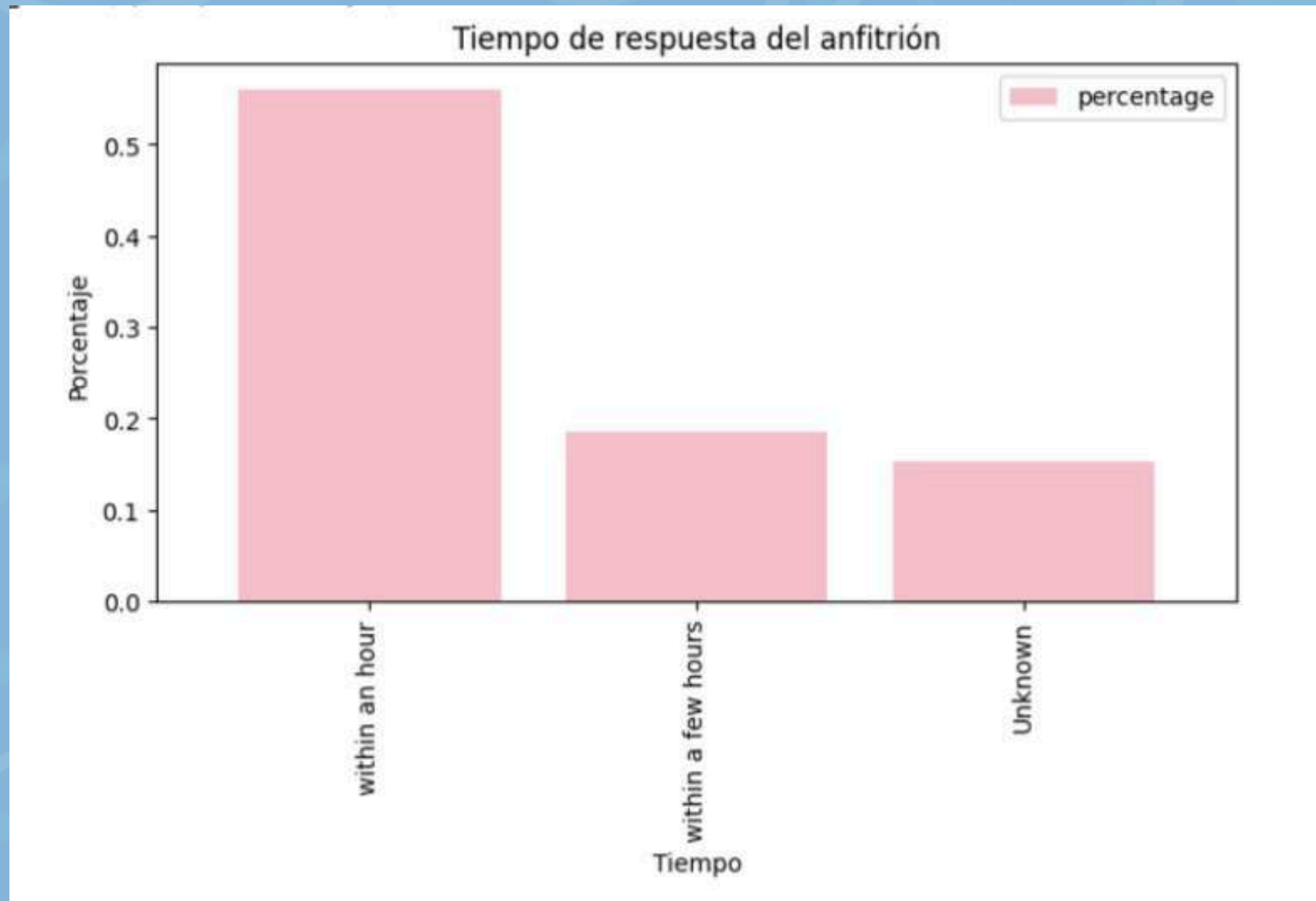
En resumen, el modelo de regresión lineal simple entre review\_scores\_cleanliness y price muestra que la puntuación de limpieza no es un predictor significativo del precio en el conjunto de datos analizado.

# GRECIA



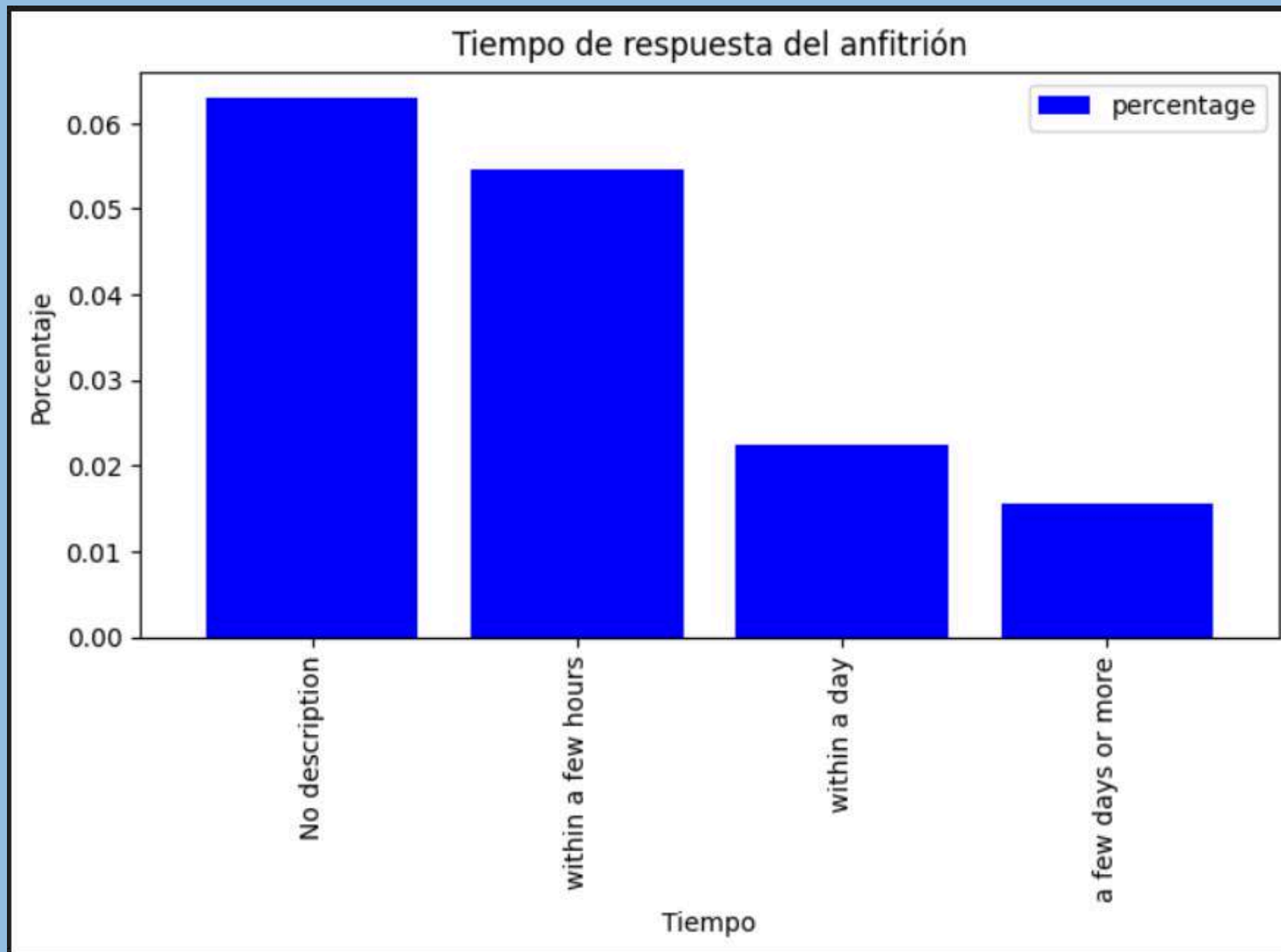


# VARIABLES CATEGÓRICAS



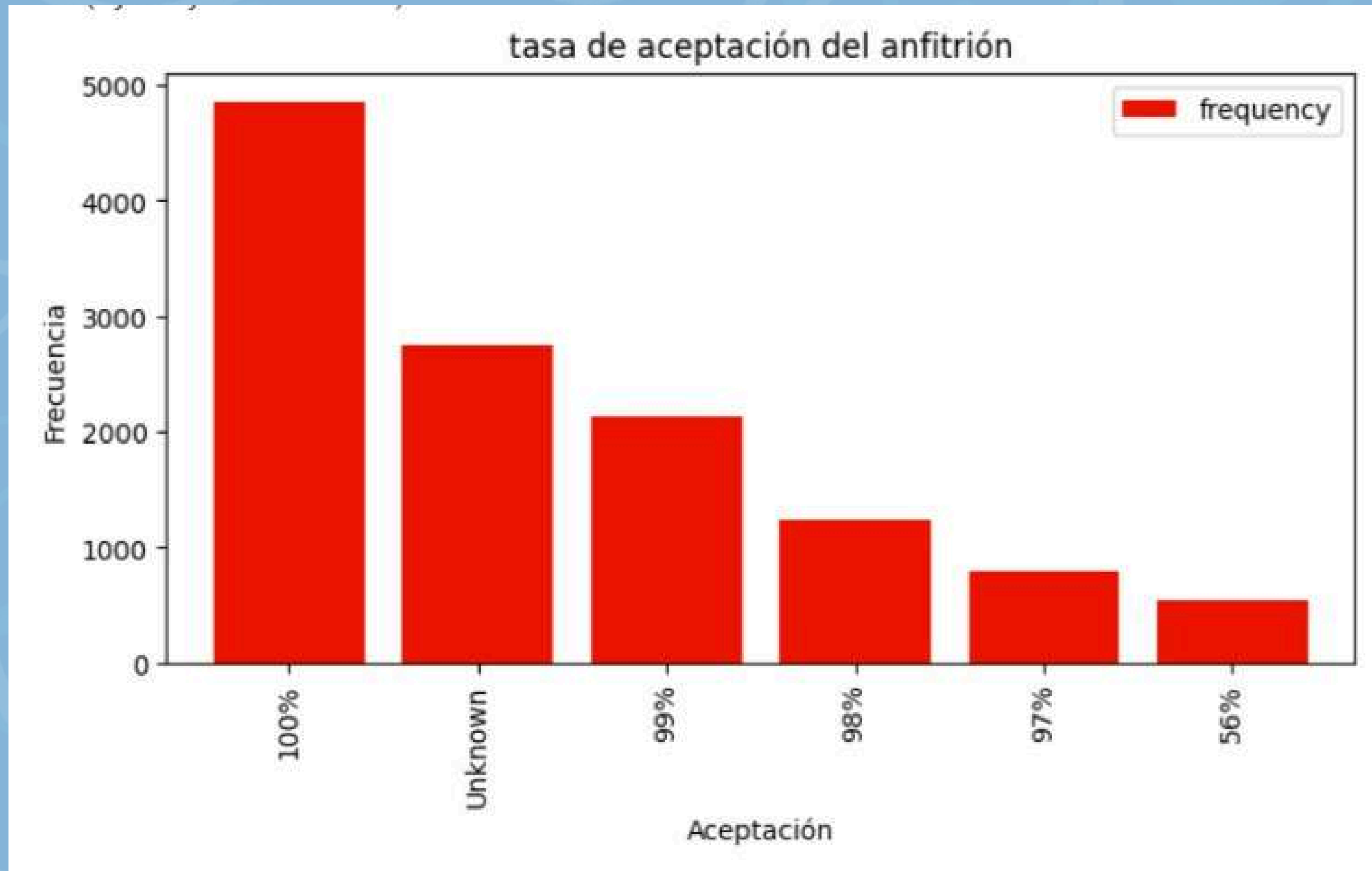
| percentage         |          |
|--------------------|----------|
| host_response_time |          |
| within an hour     | 0.560429 |
| within a few hours | 0.185964 |
| Unknown            | 0.152883 |

# GRECIA



| percentage         |          |
|--------------------|----------|
| host_response_time |          |
| No description     | 0.062905 |
| within a few hours | 0.054543 |
| within a day       | 0.022525 |
| a few days or more | 0.015594 |

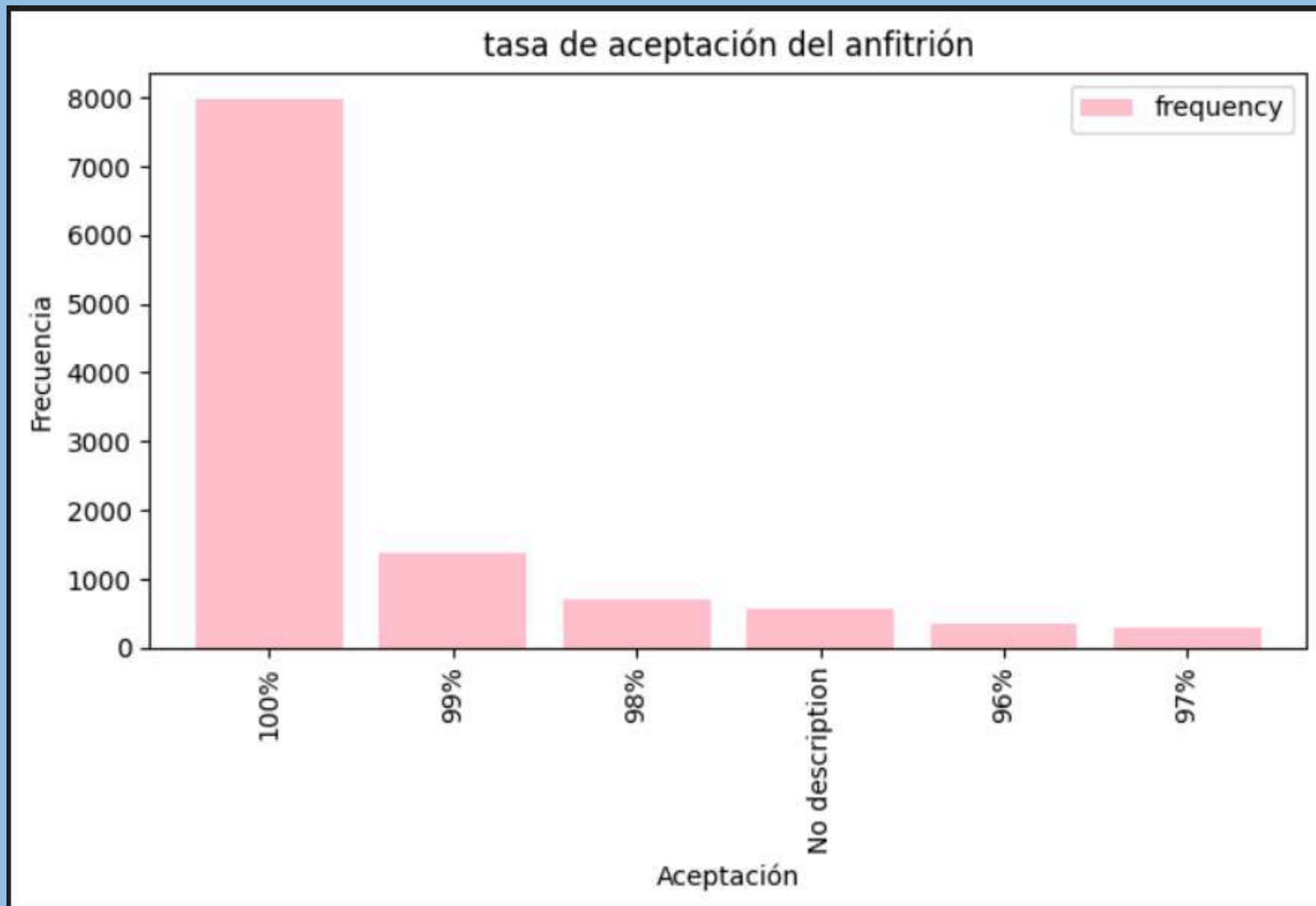
# VARIABLES CATEGÓRICAS



| frequency            |      |
|----------------------|------|
| host_acceptance_rate |      |
| 100%                 | 4855 |
| Unknown              | 2740 |
| 99%                  | 2125 |
| 98%                  | 1241 |
| 97%                  | 790  |
| 56%                  | 544  |

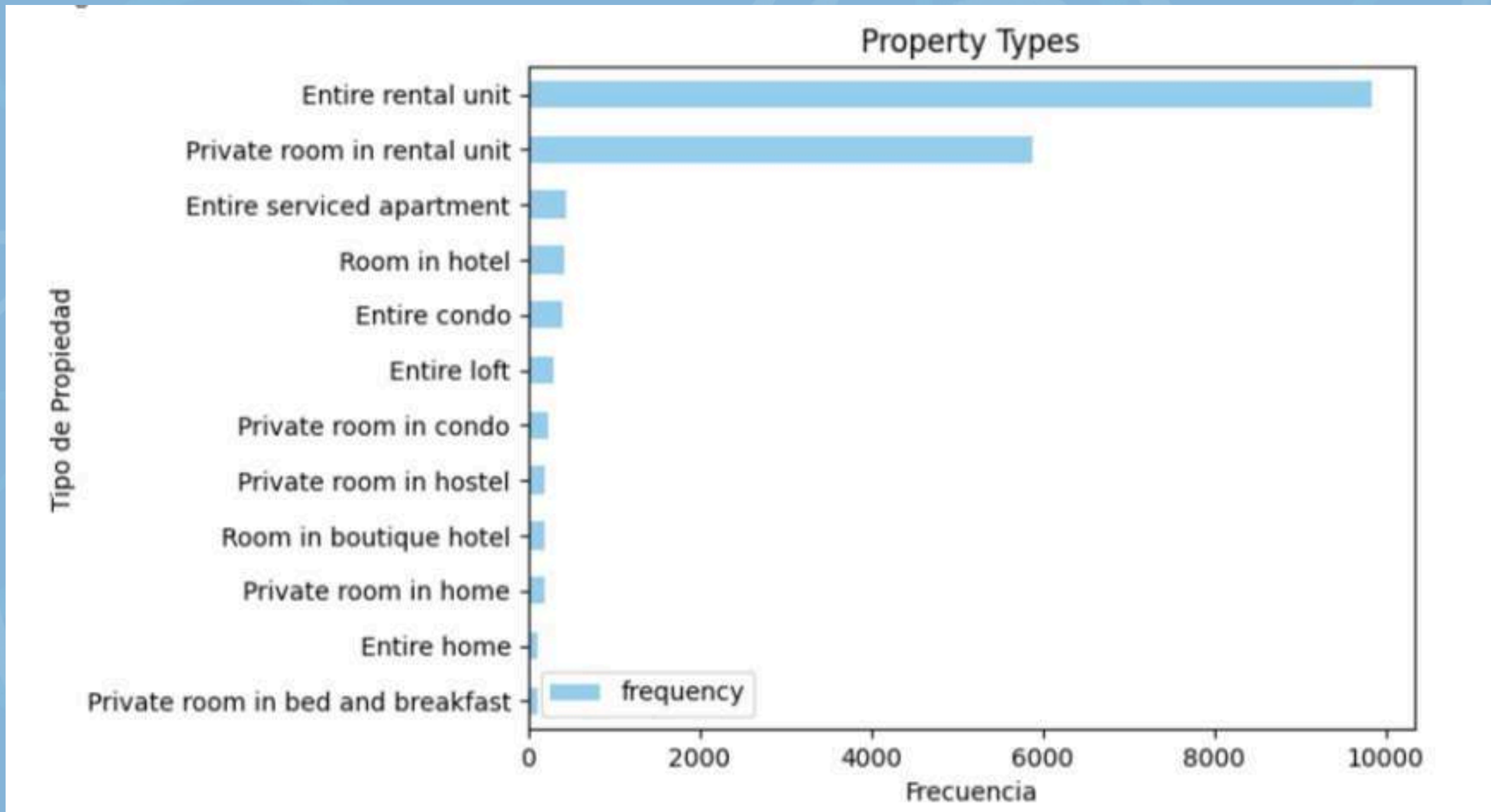


# GRECIA



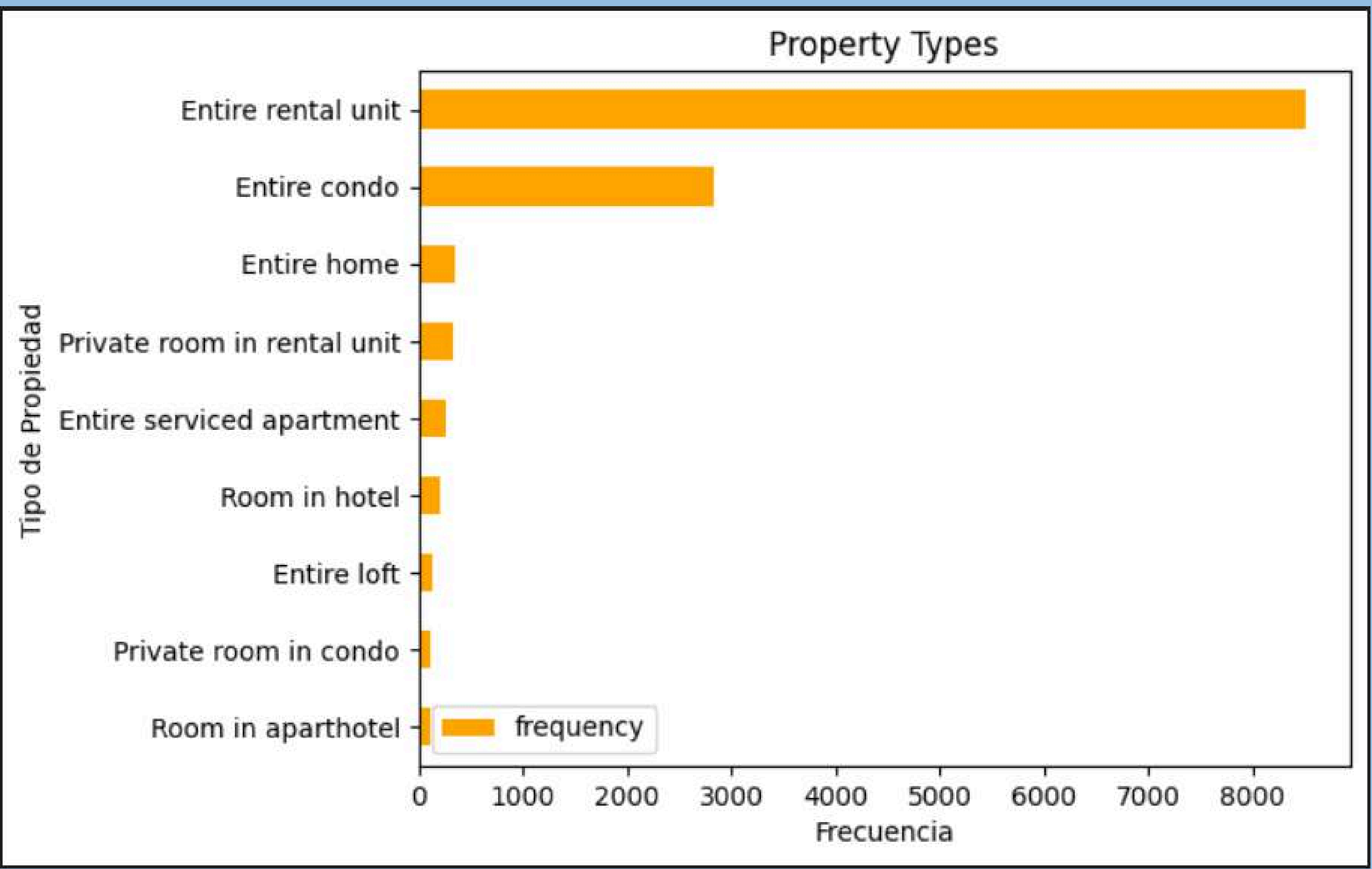
| frequency            |      |
|----------------------|------|
| host_acceptance_rate |      |
| 100%                 | 7967 |
| 99%                  | 1383 |
| 98%                  | 687  |
| No description       | 560  |
| 96%                  | 358  |
| 97%                  | 301  |

# VARIABLES CATEGÓRICAS



| property_type                     | frequency |
|-----------------------------------|-----------|
| Entire rental unit                | 9843      |
| Private room in rental unit       | 5891      |
| Entire serviced apartment         | 448       |
| Room in hotel                     | 428       |
| Entire condo                      | 392       |
| Entire loft                       | 289       |
| Private room in condo             | 233       |
| Private room in hostel            | 201       |
| Room in boutique hotel            | 192       |
| Private room in home              | 188       |
| Entire home                       | 101       |
| Private room in bed and breakfast | 100       |

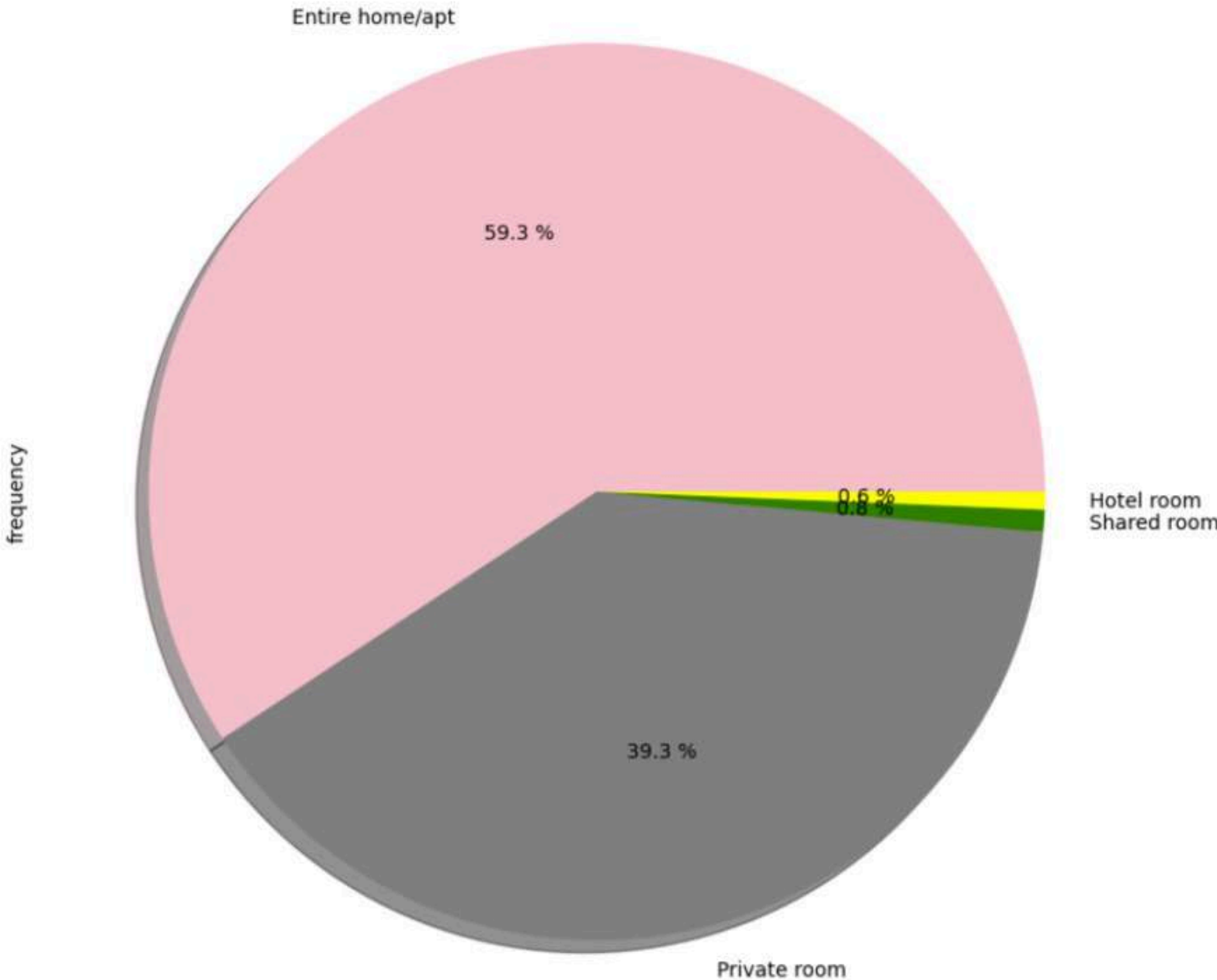
# GRECIA



| property_type               | frequency |
|-----------------------------|-----------|
| Entire rental unit          | 8512      |
| Entire condo                | 2832      |
| Entire home                 | 353       |
| Private room in rental unit | 334       |
| Entire serviced apartment   | 265       |
| Room in hotel               | 200       |
| Entire loft                 | 125       |
| Private room in condo       | 115       |
| Room in aparthotel          | 114       |

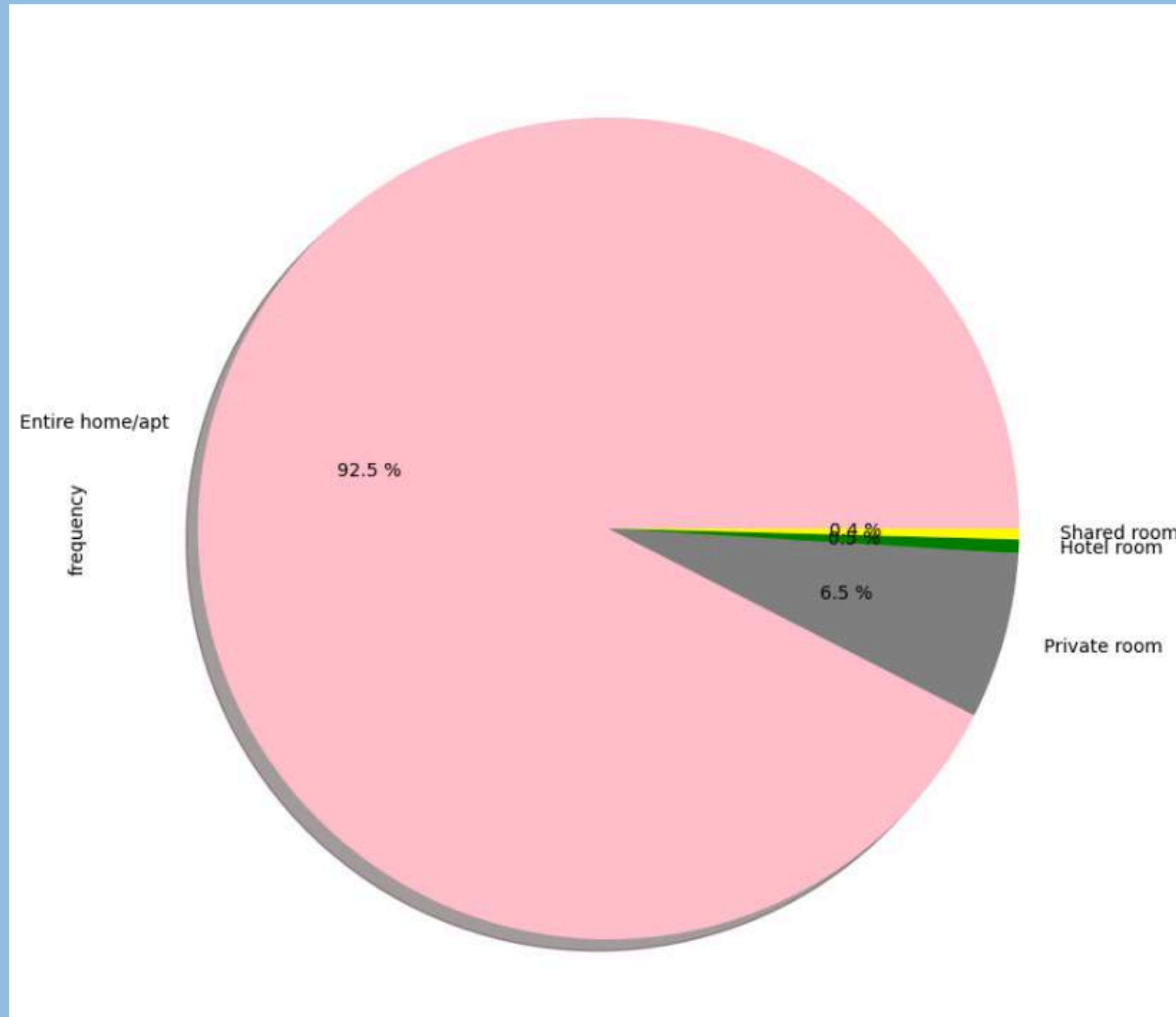


# VARIABLES CATEGÓRICAS



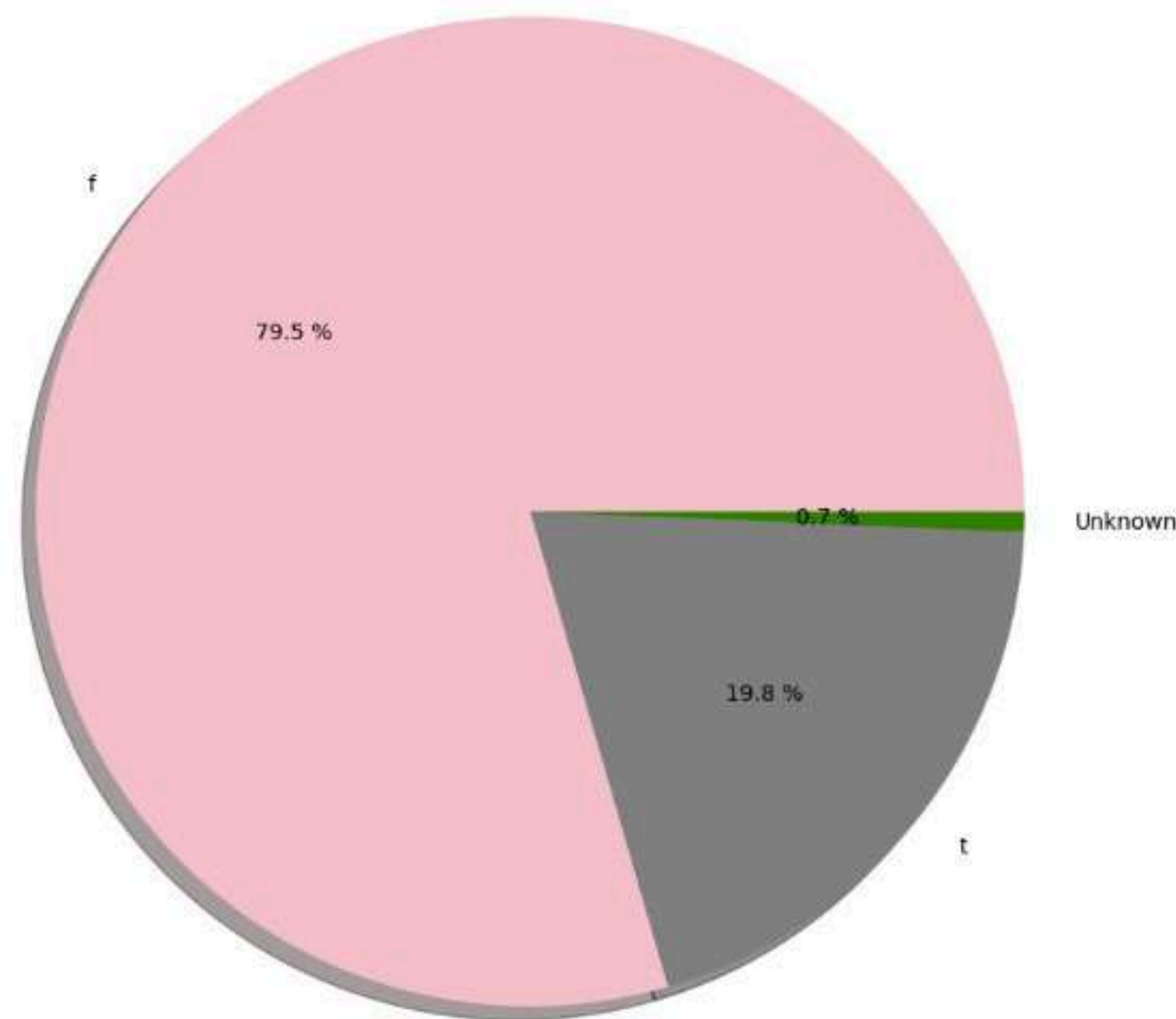
| frequency       |       |
|-----------------|-------|
| room_type       |       |
| Entire home/apt | 11217 |
| Private room    | 7434  |
| Shared room     | 151   |
| Hotel room      | 121   |

# GRECIA



| frequency       |       |
|-----------------|-------|
| room_type       |       |
| Entire home/apt | 12277 |
| Private room    | 869   |
| Hotel room      | 72    |
| Shared room     | 56    |

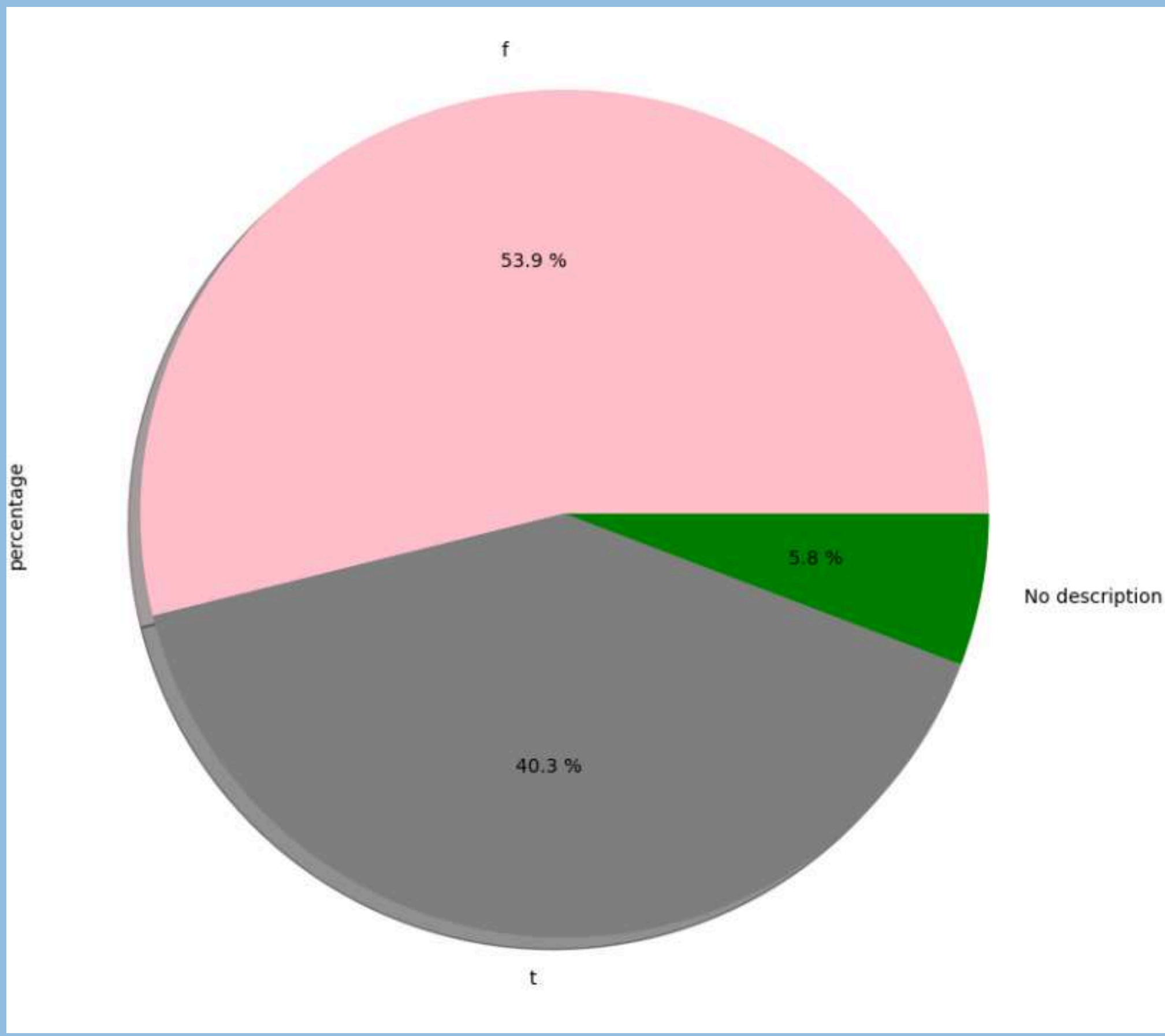
# VARIABLES CATEGÓRICAS



|                   | percentage |
|-------------------|------------|
| host_is_superuser |            |
| f                 | 0.795011   |
| t                 | 0.198436   |
| Unknown           | 0.006553   |

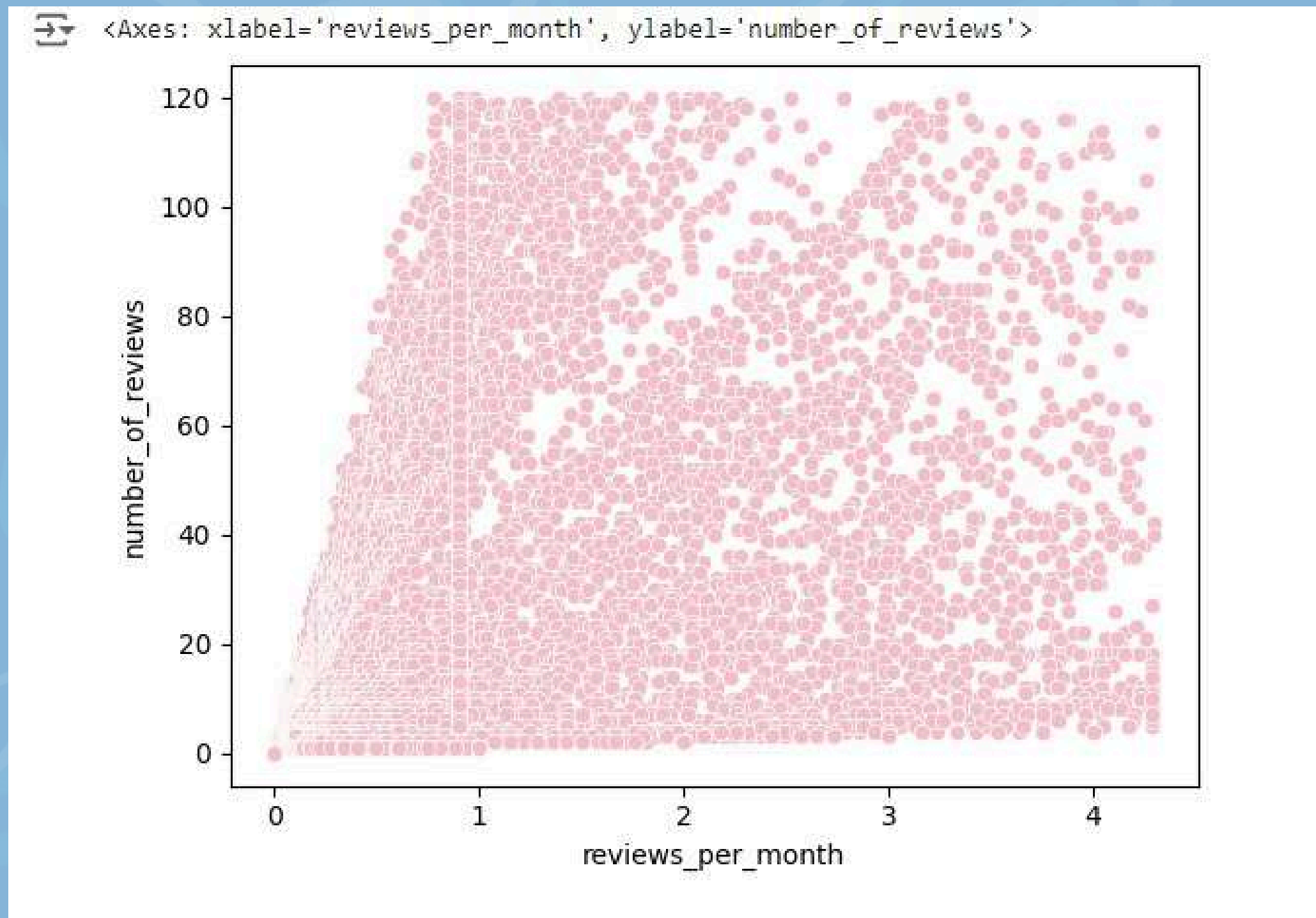


# GRECIA



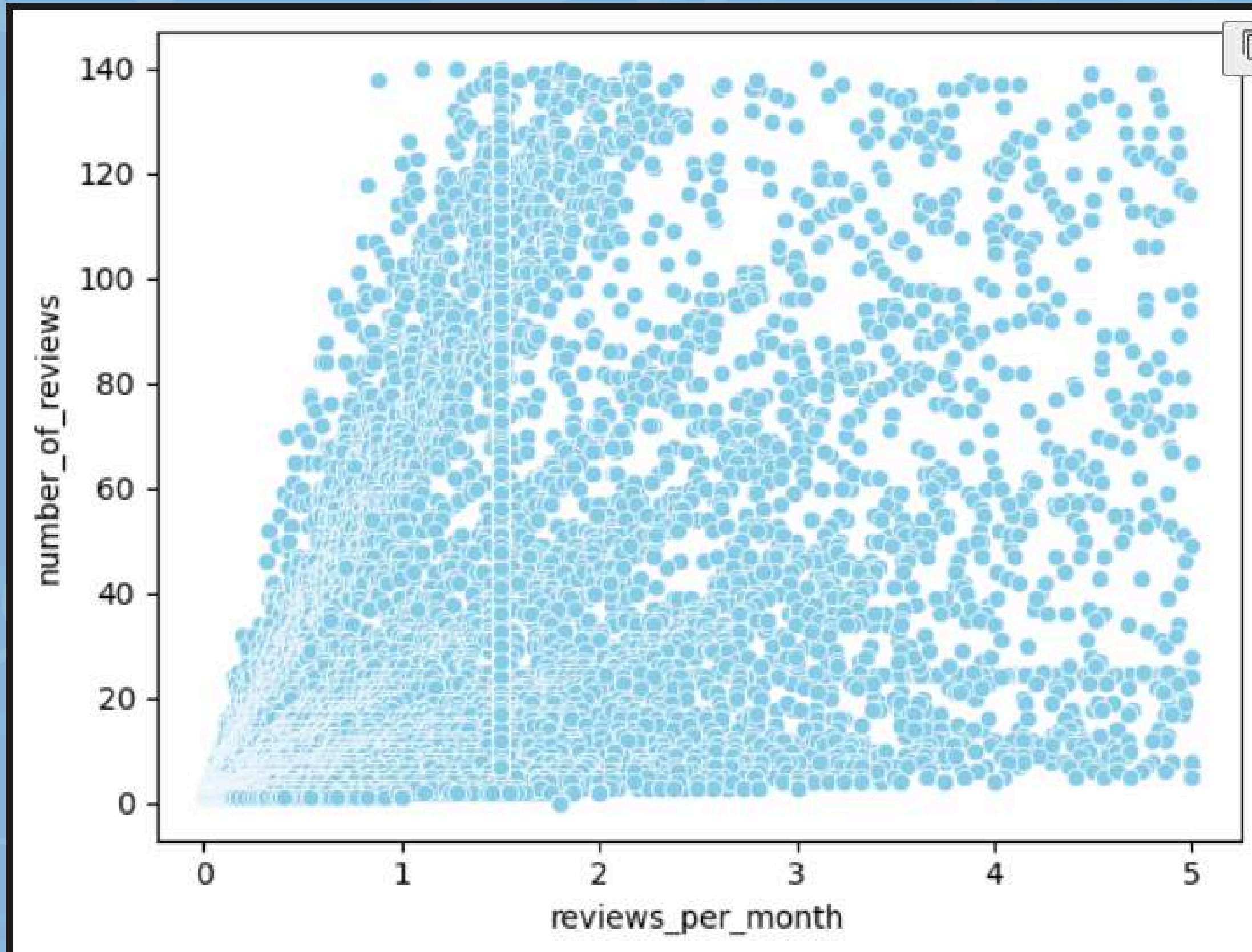
| percentage        |          |
|-------------------|----------|
| host_is_superuser |          |
| f                 | 0.538798 |
| t                 | 0.403270 |
| No description    | 0.057933 |

# MODELOS DE REGRESIÓN LINEAL SIMPLE



Al analizar la relación entre `reviews_per_month` (reseñas por mes) y `number_of_reviews` (número de reseñas), puedes determinar si un aumento en las reseñas por mes está asociado con un aumento o una disminución en el número total de reseñas.

# MODELO REGRESIÓN LINEAL SIMPLE GRECIA



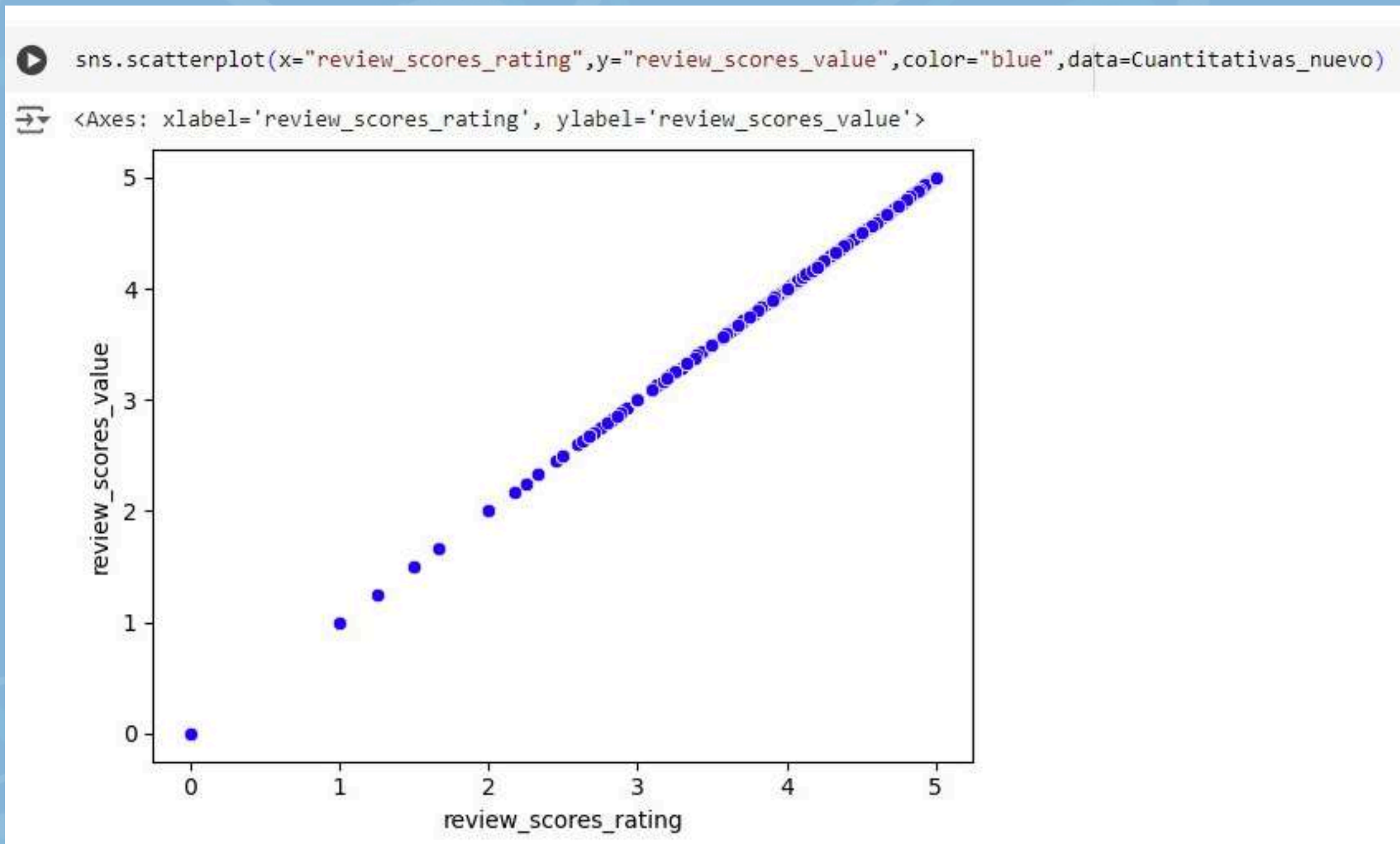
Conforme aumenta el número de reseñas por mes, observamos que también tiende a aumentar el número total de reseñas.

El patrón no es perfectamente lineal y hay dispersión, indicando variabilidad entre los datos.

Los puntos concentrados en la parte baja señalan que la mayoría de los listados tienden a tener menos reseñas por mes.



# MODELOS DE REGRESIÓN LINEAL SIMPLE

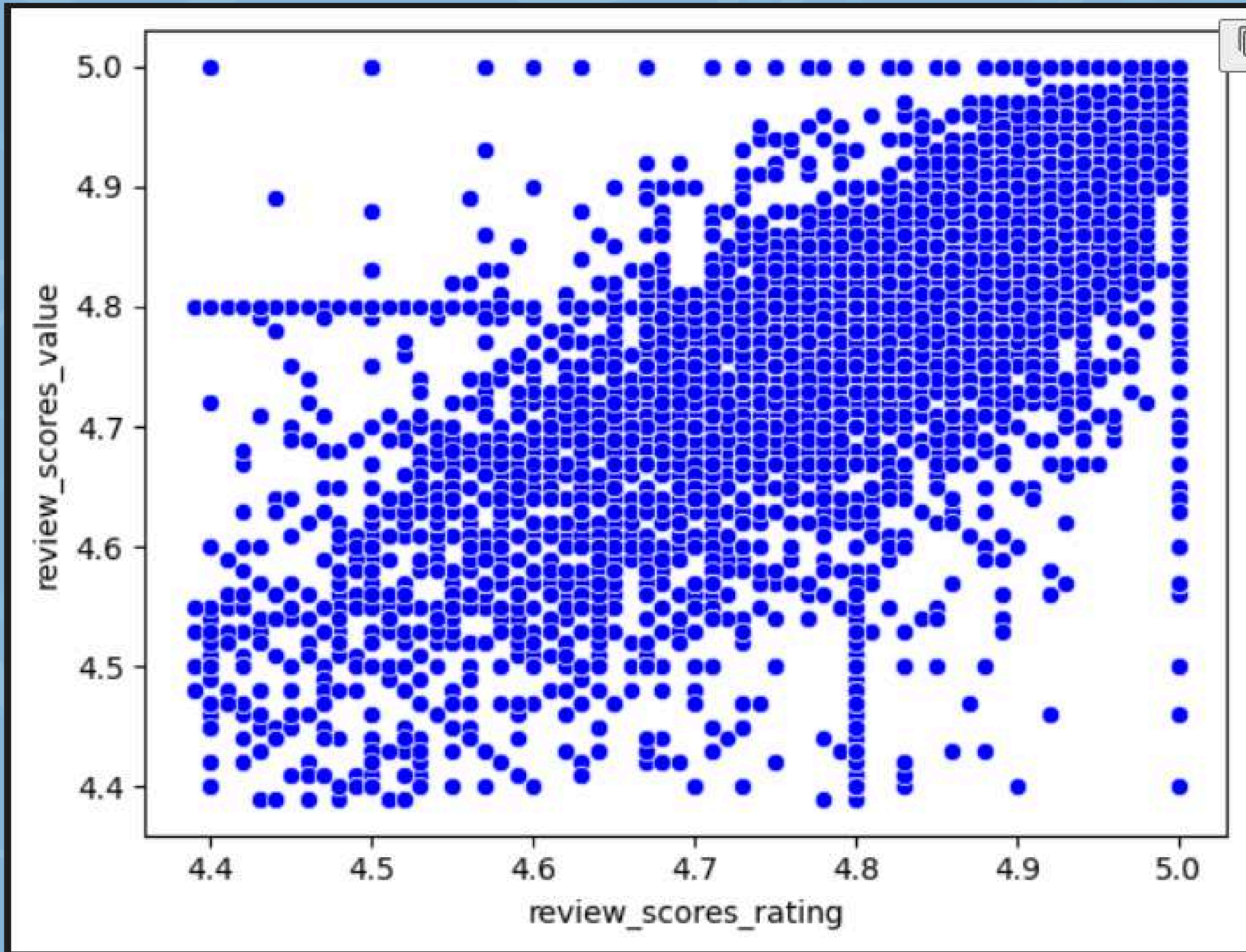


La regresión entre `review_scores_rating` y `review_scores_value` puede ayudar a evaluar si las puntuaciones de "valor" de las reseñas están alineadas con las calificaciones generales.

Observamos que los puntos están alineados perfectamente en una línea diagonal.

A pesar de que el gráfico de dispersión muestra una relación lineal perfecta, los valores de  $R^2$  y el coeficiente de correlación sugieren que esta relación no es tan fuerte como parece.

# MODELO REGRESIÓN LINEAL SIMPLE GRECIA



Tendencia positiva entre las dos variables.

La concentración de puntos en la parte superior derecha del gráfico, sugiere que muchos de los elementos evaluados tienen calificaciones altas tanto en el valor general como en el valor específico, pudiendo indicar que los encuestados suelen estar satisfechos en ambos aspectos.

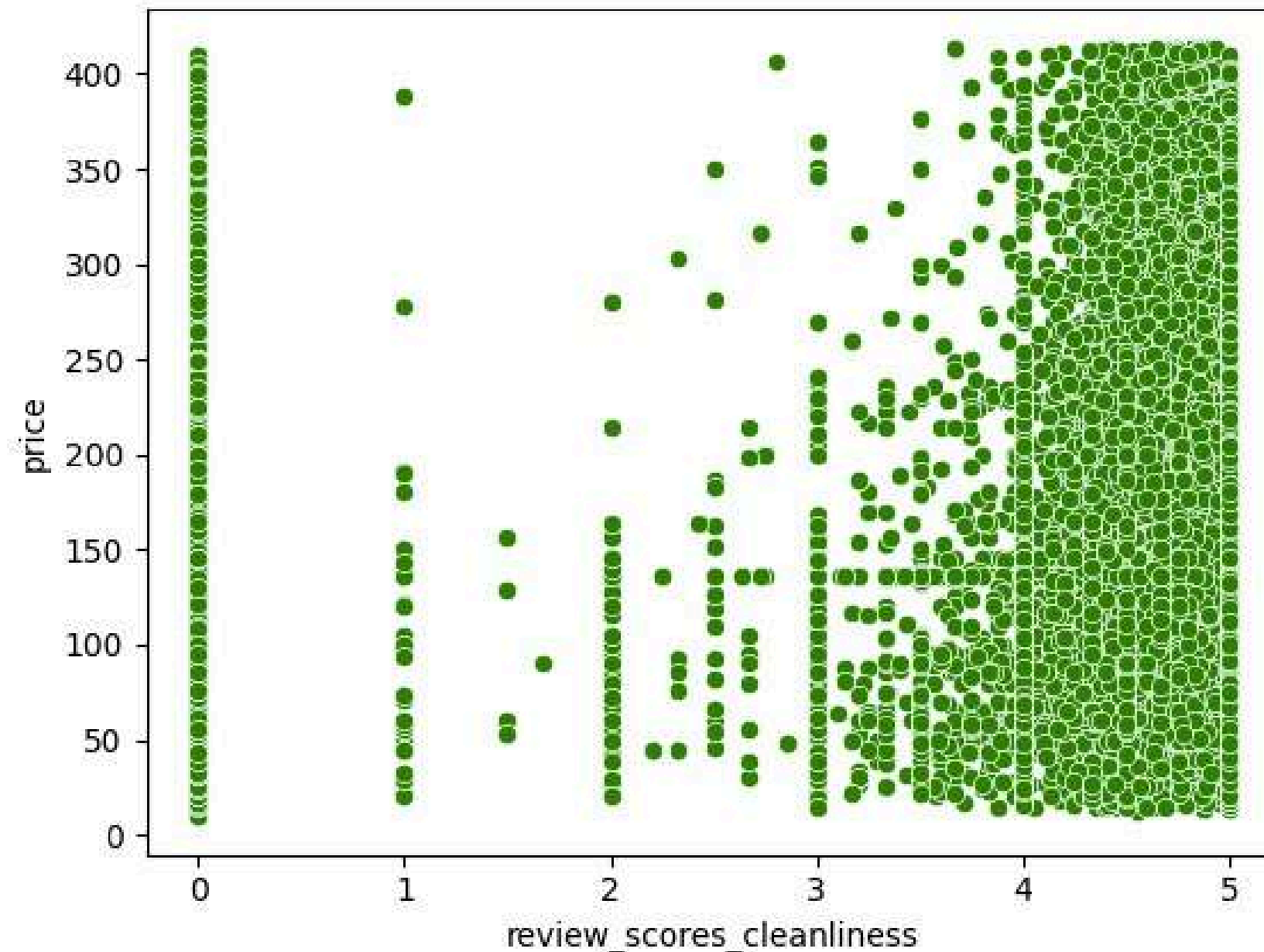
La dispersión baja en el rango superior se ve menos dispersión debido a que hay menos variabilidad en las calificaciones altas.



# MODELOS DE REGRESIÓN LINEAL SIMPLE

```
sns.scatterplot(x="review_scores_cleanliness",y="price",color="green",data=Cuantitativas_nuevo)
```

```
<Axes: xlabel='review_scores_cleanliness', ylabel='price'>
```

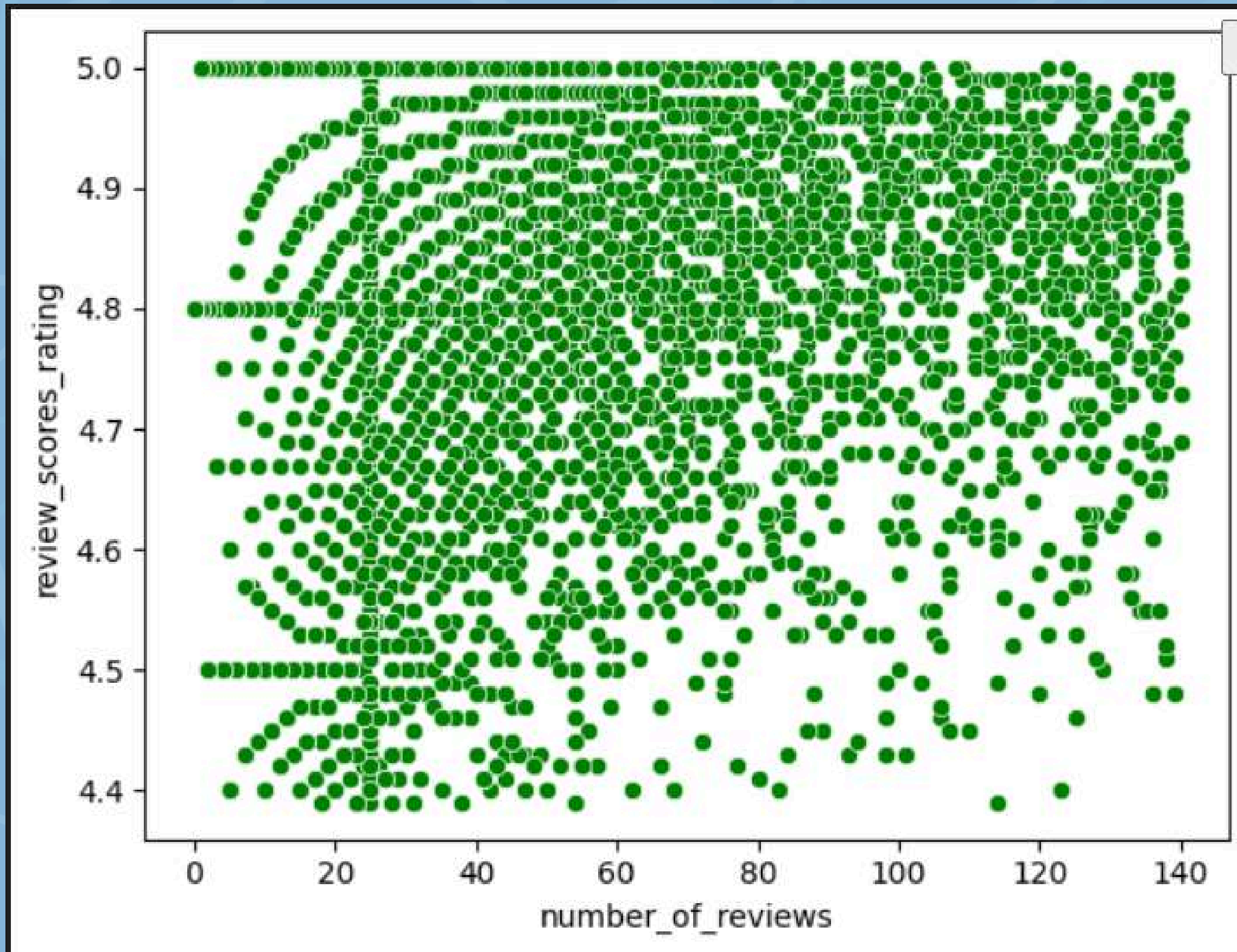


La regresión puede ayudar a entender si existe una relación entre la percepción de la limpieza y el precio que los huéspedes están dispuestos a pagar.

El modelo de regresión lineal simple entre review\_scores\_cleanliness y price muestra que la puntuación de limpieza no es un predictor significativo del precio en el conjunto de datos analizado.



# MODELO REGRESIÓN LINEAL SIMPLE GRECIA

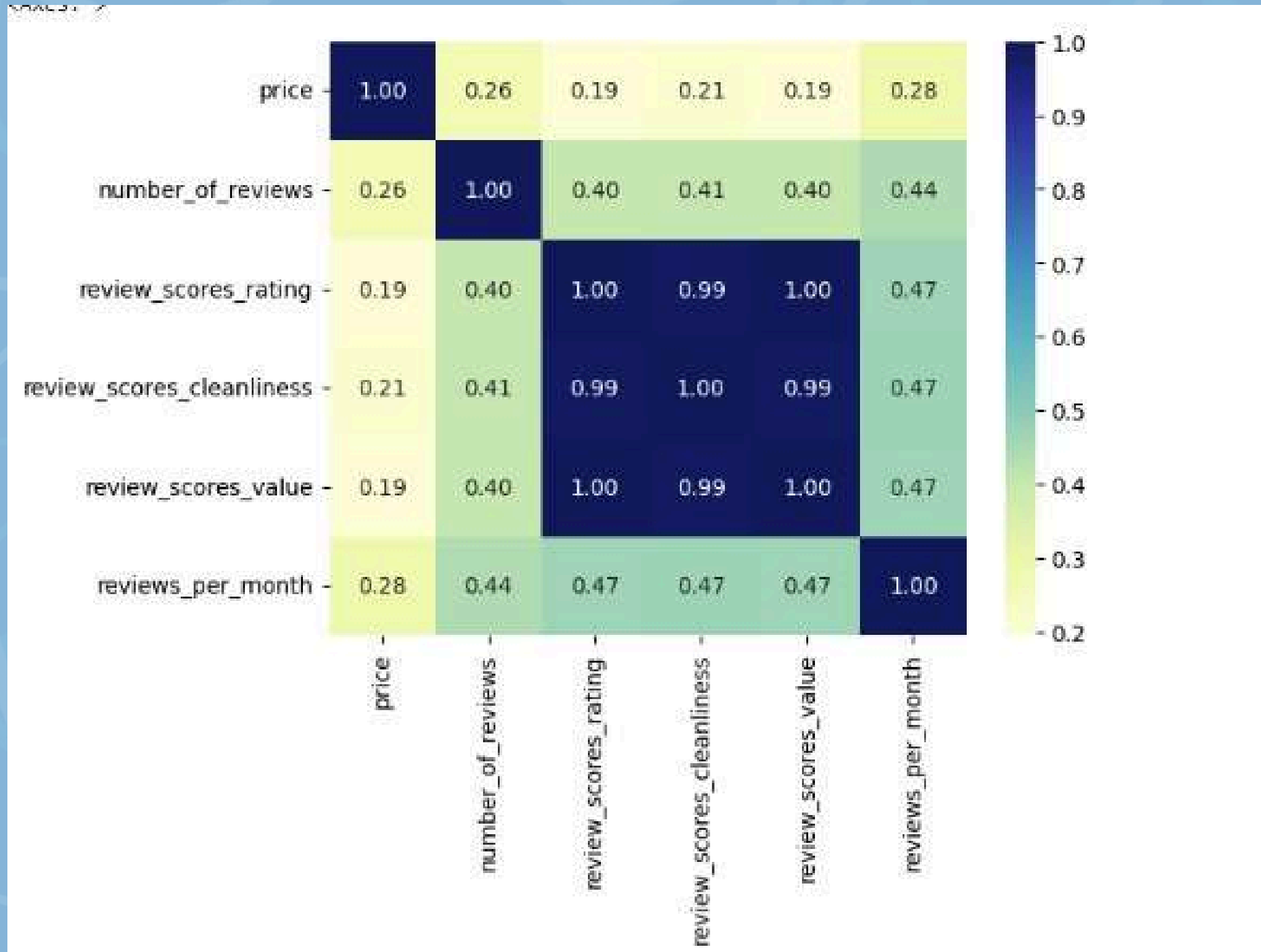


A medida que aumenta el número de reseñas, la calificación promedio tiende a concentrarse en los valores altos, por lo que aquellos con mayor número de evaluaciones suelen ser quienes tienen mejores calificaciones.

La dispersión de datos podría indicarnos que el número de reseñas no es el único factor que determina la calificación.

# MAPA DE CALOR

## BARCELONA



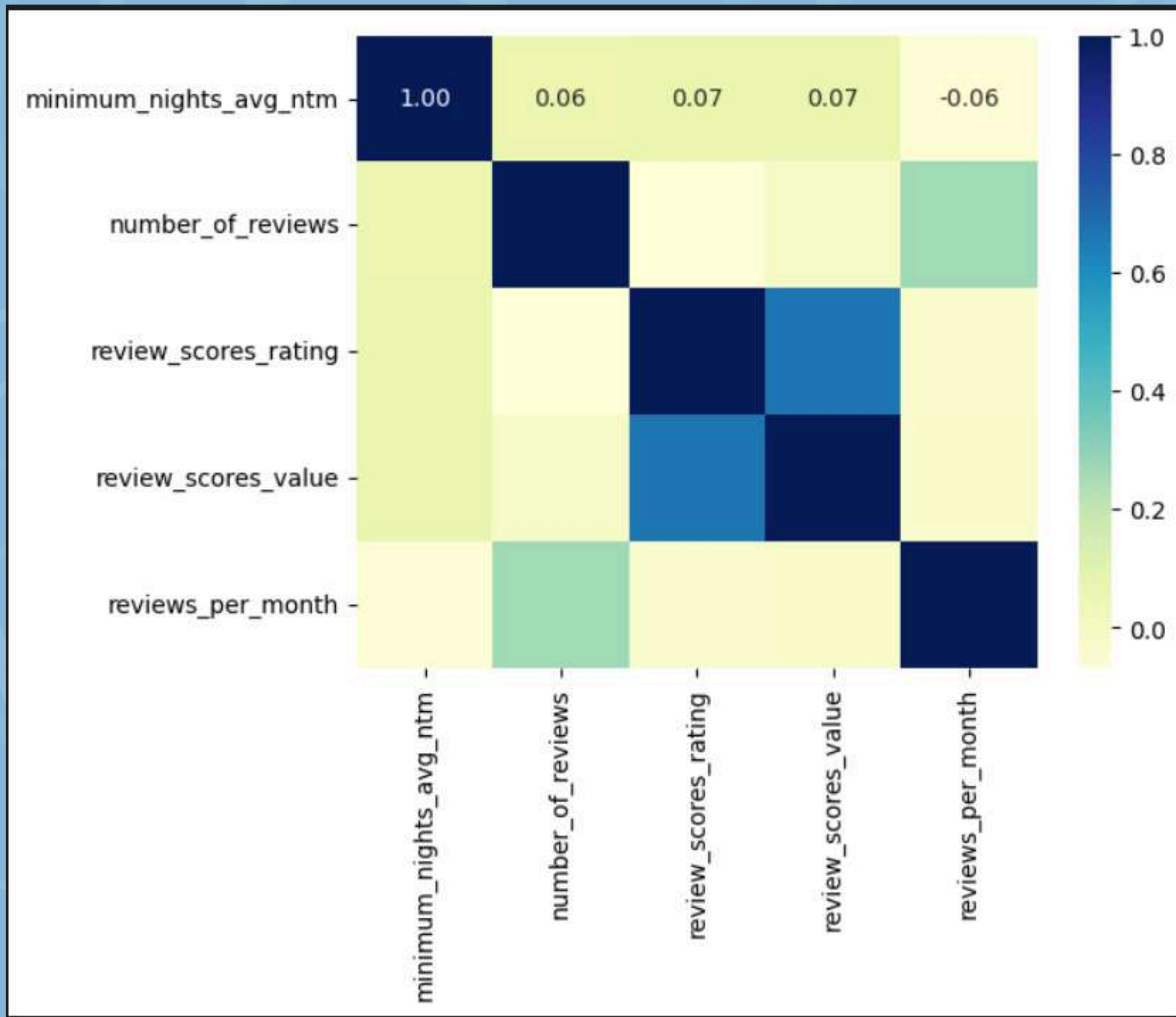
**Alta correlación entre puntuaciones**  
(calificación, limpieza y valor)

**Correlación positiva con número de reseñas**  
los alojamientos con más reseñas tienden a ser  
más caros y mejor evaluados.

**Correlación moderada con reseñas por mes**  
Esto sugiere que los alojamientos con más  
reseñas recientes tienden a tener mejores  
puntuaciones.



# MAPA DE CALOR GRECIA



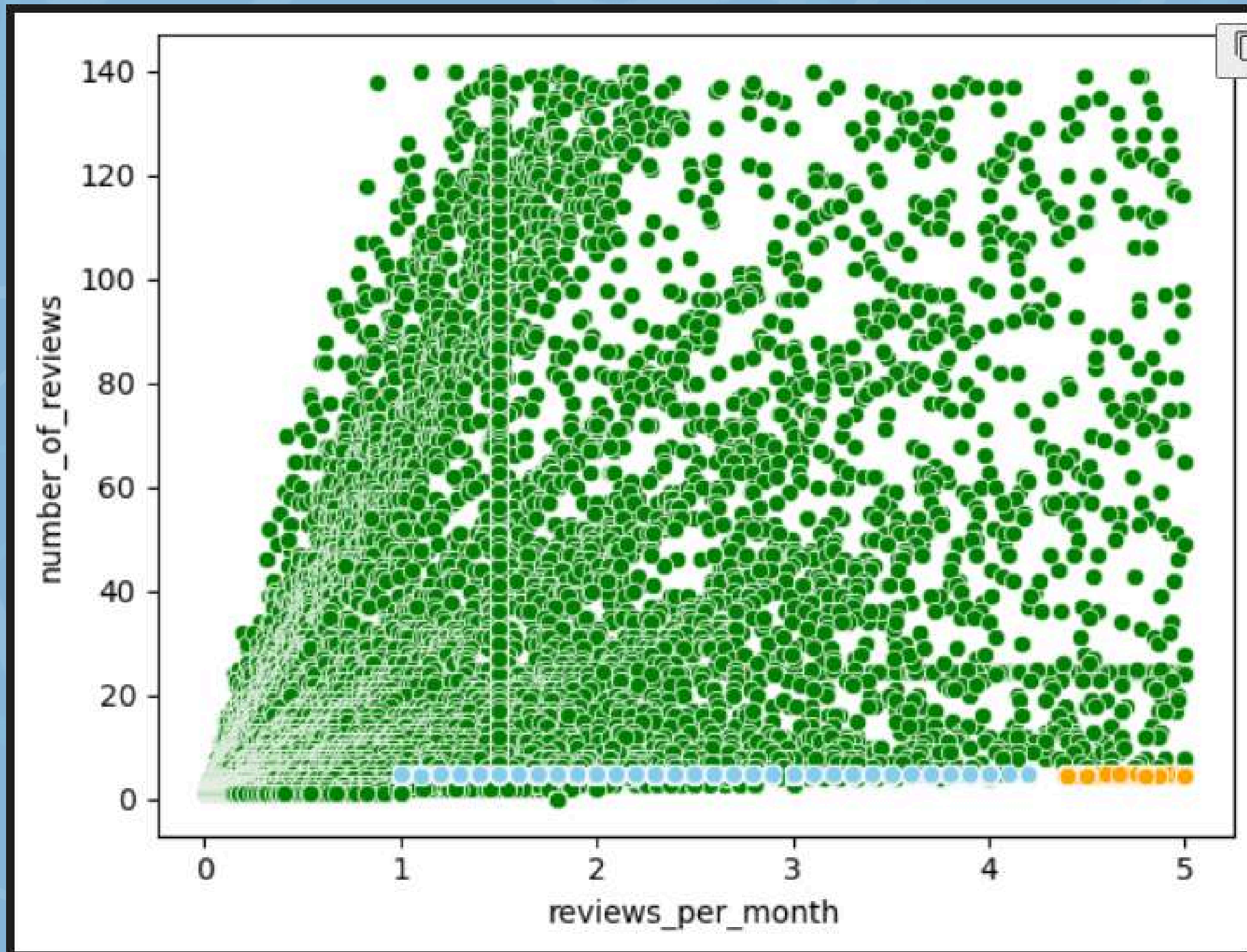
**Correlaciones positivas:**  
review\_scores\_rating y  
review\_scores\_value

Ejemplo: A medida que aumenta una,  
la otra también tiende a hacerlo.

**Correlaciones negativas**  
minimum\_nights\_avg\_ntm y  
reviews\_per\_month

Ejemplo: A medida que el promedio de  
noches mínimas aumenta, las reseñas por  
mes tienden a disminuir ligeramente.

# REGRESIÓN LINEAL MÚLTIPLE GRECIA



## Verde

Relación entre “reviews\_per\_month”  
y “number\_of\_reviews”

## Naranja

Relación entre “review\_scores\_rating”  
y “review\_scores\_value”

## Azul

Relación entre “minimum\_nights\_avg\_ntm” y  
“review\_scores\_rating”





**¡GRACIAS!**