

# Predicting AirBnB prices



**Sara Pazo Padrón**



# Data cleaning - Train

# 2

# 3

The heatmap displays the Pearson correlation matrix for the following features: id, log\_price, accommodates, bathrooms, host\_response\_rate, instant\_bookable, latitude, longitude, number\_of\_reviews, bedrooms, beds, room\_type\_Entire home/apt, room\_type\_Private room, room\_type\_Shared room, cancellation\_policy\_flexible, cancellation\_policy\_moderate, cancellation\_policy\_strict, cancellation\_policy\_super\_strict, bed\_type\_Airbed, bed\_type\_Couch, bed\_type\_Futon, bed\_type\_Pull-out Sofa, and bed\_type\_Real Bed. The color scale indicates the strength and direction of the correlation, ranging from -0.75 (dark red) to 0.75 (dark blue).

# Data cleaning - Test

1

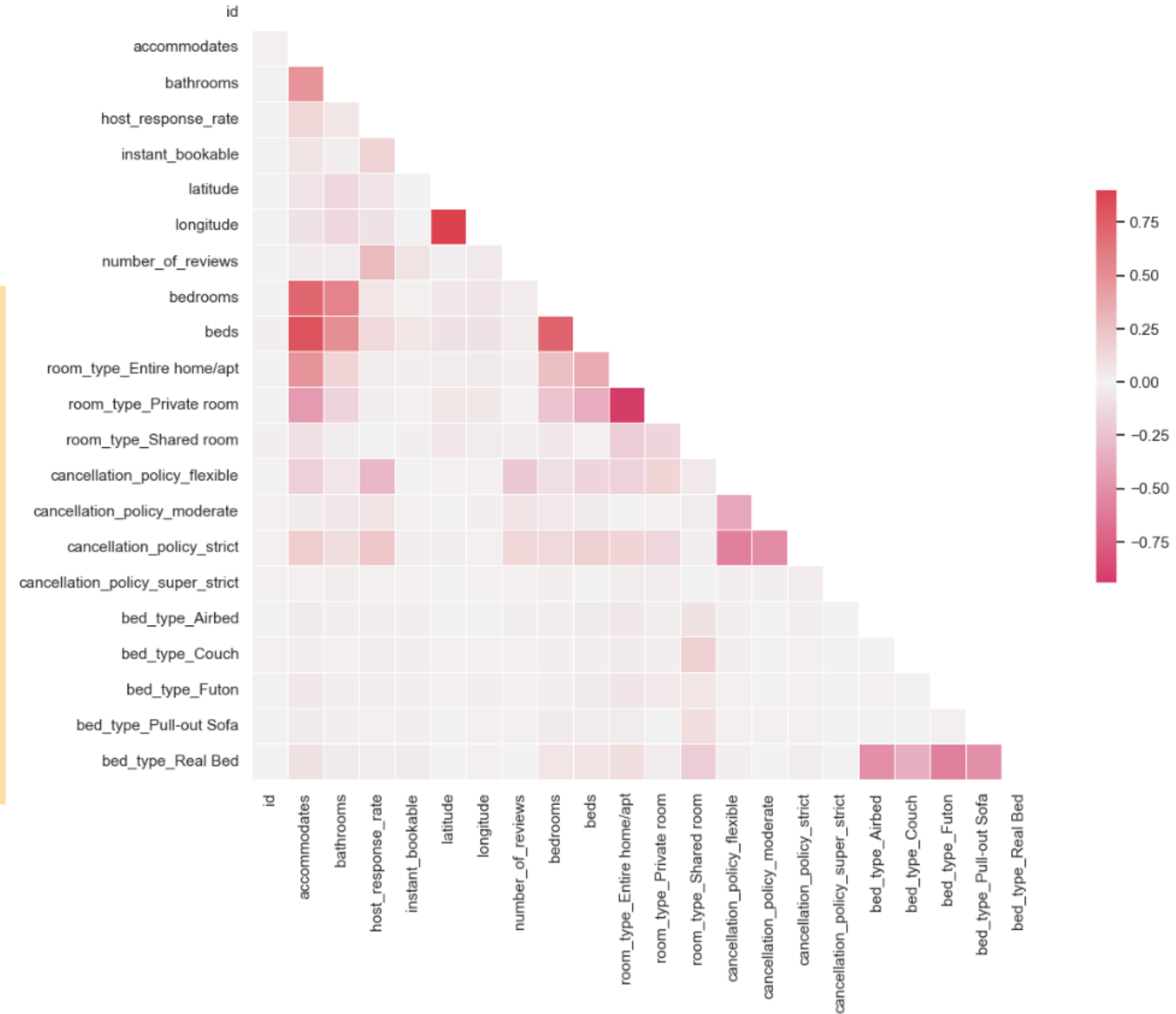
Repetimos la  
secuencia de  
transformación en  
el csv 'Test'

2

'Test':  
No tenemos datos del  
'precio', pero sí  
correlacionada con  
precio - 'accommodates'

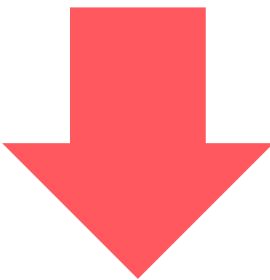
3

Guardamos el  
nuevo 'Test' a csv



# Modelo predictivo

RMSE



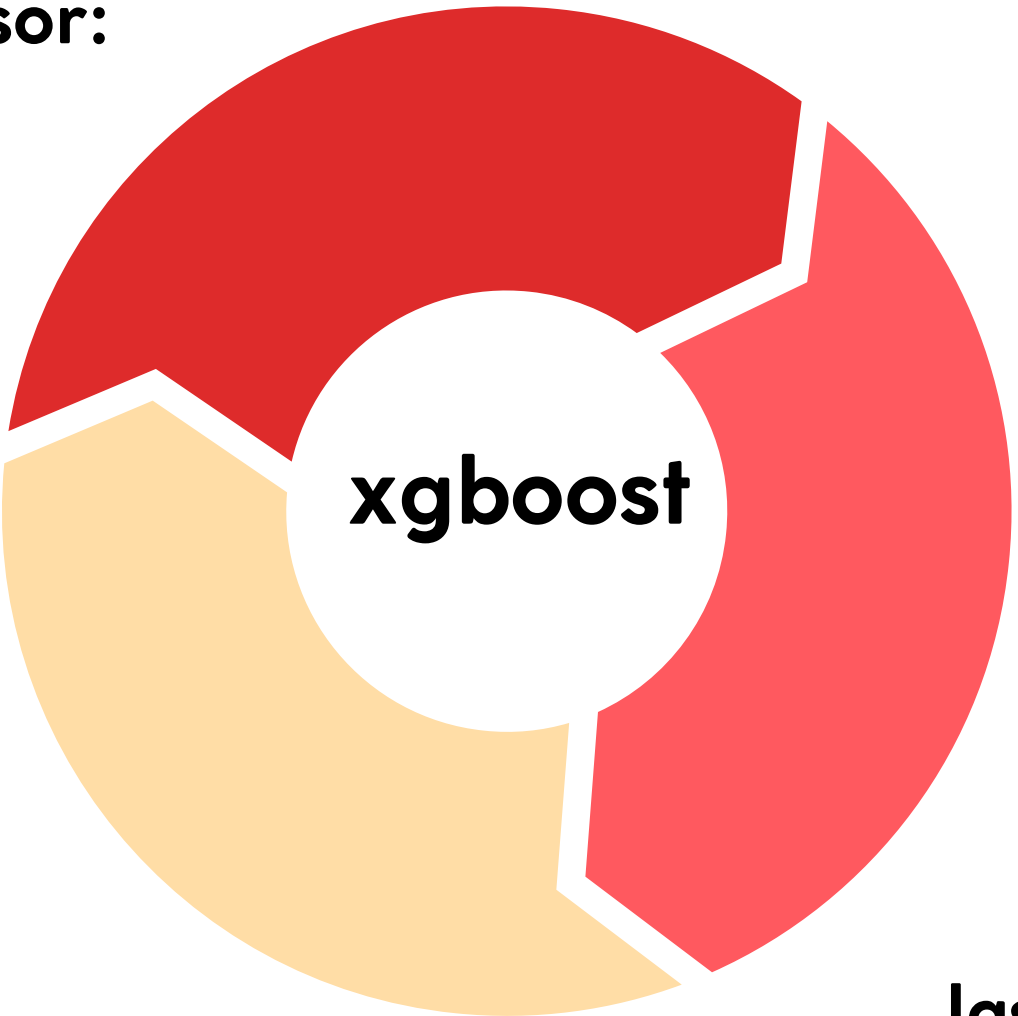
0.395559105

gradient\_boosting\_regressor:  
0.180491

xgboost\_regressor:  
0.161491

random\_forest\_regressor :  
0.16947

decision\_tree\_regressor:  
0.32350



linear\_regression:  
0.244605

ridge\_regression:  
0.24460

lasso\_regression:  
0.516482



airbnb

¡gracias!