

REPORTE Actividad 6.1 (Regresión Lineal Simple)

Nicole Aryam Rodríguez A00831569

DF, MÉXICO

Coeficientes de correlación

	Tipo Habitación	dependiente	independiente	Coef Determinación	Coef Correlación
0	Entire home/apt	host_acceptance_rate	host_response_rate	0.479756	0.692644
1	Entire home/apt	host_acceptance_rate	price	0.000555	0.023561
2	Entire home/apt	host_acceptance_rate	number_of_reviews	0.017849	0.133599
3	Entire home/apt	review_scores_location	review_scores_cleanliness	0.806106	0.897834
4	Entire home/apt	availability_365	number_of_reviews	0.002782	0.052747
5	Entire home/apt	reviews_per_month	review_scores_communication	0.140989	0.375485
6	Hotel room	host_acceptance_rate	host_response_rate	0.581784	0.762748
7	Hotel Room	host_acceptance_rate	price	0.123291	0.351128
8	Hotel Room	host_acceptance_rate	number_of_reviews	0.044381	0.210669
9	Hotel room	review_scores_location	review_scores_cleanliness	0.858691	0.926656
10	Hotel Room	availability_365	number_of_reviews	0.007228	0.085018
11	Hotel Room	reviews_per_month	review_scores_communication	0.098431	0.313737

CALIFORNIA, EUA

	Tipo Habitación	dependiente	independiente	Coef Determinación	Coef Correlación
0	Entire home/apt	host_acceptance_rate	host_response_rate	0.394167	0.627827
1	Entire home/apt	host_acceptance_rate	price	0.001412	0.037574
2	Entire home/apt	host_acceptance_rate	number_of_reviews	0.024557	0.156706
3	Entire home/apt	review_scores_location	review_scores_cleanliness	0.960796	0.980202
4	Entire home/apt	availability_365	number_of_reviews	0.010570	0.102809
5	Entire home/apt	reviews_per_month	review_scores_communication	0.123237	0.351052
6	Hotel room	host_acceptance_rate	host_response_rate	0.951898	0.975652
7	Hotel Room	host_acceptance_rate	price	0.009217	0.096003
8	Hotel Room	host_acceptance_rate	number_of_reviews	0.050857	0.225515
9	Hotel room	review_scores_location	review_scores_cleanliness	0.685443	0.827915
10	Hotel Room	availability_365	number_of_reviews	0.000414	0.020343
11	Hotel Room	reviews_per_month	review_scores_communication	0.063164	0.251325

En estas dos tablas podemos ver la correlación que hay entre una variable dependiente con una independiente en cada tipo de habitación. Se muestran 2 tablas correspondientes a cada ciudad, DF, México y California, Estados Unidos. Se eligió de tipo de habitación: entire home/apt y hotel room, para las cuales analizamos 6 conjuntos de variables por cada habitación.

De las correlaciones correspondientes al dataset de México vemos que la mejor correlación se encuentra en la habitación Hotel Room, con las variables review_scores_location y review_scores_cleanliness. Esta correlación de 0.92 significa que cuando una de las

variables aumenta, es muy probable que la otra variable también aumente. Es decir, si un huésped califica un alojamiento con una puntuación alta en limpieza, es muy probable que también le dé una puntuación alta en ubicación. Esta misma correlación alta entre estas 2 variables podemos encontrarla en la habitación Entire room/apt de México (coef= 0.89) y en el dataset de California con un coeficiente de 0.82 para Hotel Room y 0.98 para Entire Room/apt. De igual manera podemos encontrar en el dataset de México y aún más en el de California en especial para la habitación "Hotel Room", una correlación fuerte entre host acceptance rate y host response rate. Esto quiere decir que los anfitriones o en este caso hoteles que aceptan más reservas también tienden a responder más rápidamente a las consultas de los huéspedes, guiándonos con el coeficiente de 0.97.

Por el otro lado tenemos variables donde no hay una correlación fuerte aunque esta sea positiva, esto lo podemos ver en ambos datasets, con las variables host acceptance rate y price, al igual que con availability 365 y number of reviews teniendo coeficientes debajo de 0.1.