



Instituto Tecnológico y de Estudios Superiores de Monterrey

**Analítica de datos y herramientas de
inteligencia artificial II (Gpo 501)**

Actividad 6.1

Said Arturo Flores Nieto | A01703477

Profesor
Alfredo García Suárez

04/05/2023

México DF

	Variables	Correlación Private room	Determinación Private room	Correlación Shared room	Determinación Shared room
0	host_acceptance_rate vs host_response_rate	0.618391	0.382407	0.717656	0.515030
1	review_scores_location vs review_scores_cleanl...	0.335194	0.112355	0.284996	0.081223
2	host_acceptance_rate vs price	0.068119	0.004640	0.311148	0.096813
3	availability_365 vs number_of_reviews	0.061028	0.003724	0.037257	0.001388
4	host_acceptance_rate vs number_of_reviews	0.213547	0.045603	0.236688	0.056021
5	reviews_per_month vs review_scores_communication	0.060777	0.003694	0.049467	0.002447

EUA California

	Variables	Correlación Private room	Determinación Private room	Correlación Shared room	Determinación Shared room
0	host_acceptance_rate vs host_response_rate	0.772220	0.596324	0.938324	0.880451
1	review_scores_location vs review_scores_cleanl...	0.309773	0.095959	0.052639	0.002771
2	host_acceptance_rate vs price	0.023758	0.000564	0.065773	0.004326
3	availability_365 vs number_of_reviews	0.062043	0.003849	0.145559	0.021187
4	host_acceptance_rate vs number_of_reviews	0.216926	0.047057	0.188627	0.035580
5	reviews_per_month vs review_scores_communication	0.002449	0.000006	0.073849	0.005454

Comparando las dos ciudades, así como los dos tipos de habitación, es notable que el comportamiento entre ambas ciudades en la mayoría de los casos es similar, sin embargo, hay otros donde si difiere bastante.

Por ejemplo, en la relación que tiene la tasa de respuesta del anfitrión sobre la tasa de aceptación del anfitrión, hay un comportamiento similar, ya que de las 6 relaciones entre variables, en ambas ciudades es la que tiene mayor correlación, siendo también en ambos mayores para el cuarto compartido que para el privado. Además, es importante mencionar, que aunque es similar correlación, es mayor en el caso de california, con un coeficiente de correlación de 0.77 para privado y uno de 0.94 para el compartido, supera al 0.62 de privado y al 0.72 del compartido en el DF.

Como segunda relación se quiere revisar si las puntuaciones de limpieza tienen influencia en las puntuaciones de localización, es decir que si la limpieza de la propiedad termina por afectar la percepción de los usuarios sobre la localización. En el caso del cuarto privado, el comportamiento en ambas ciudades es similar, ya que se maneja una correlación cercana al 0.3 y solo llegan a diferir en el cuarto compartido, al tener en el DF una correlación de 0.28 que aunque es baja, es mucho más baja la de California al ser de 0.53.

La tercera relación de variables es la influencia del precio en la tasa de aceptación del anfitrión, en esta relación la correlación en ambas habitaciones de California, así como en la habitación privada del DF es sumamente baja, ya que incluso es menor a 0.07, en el único caso que es considerablemente mayor a esto es la habitación compartida del DF que tiene una correlación de 0.31. En general con esto podemos identificar que estos modelos no son buenos, ya que realmente el precio no tiene influencia importante en la tasa de aceptación en ninguna de las ciudades.

Como cuarta relación se tiene el número de revisiones influyendo la disponibilidad en los 365 días. De igual manera que en el anterior, el comportamiento en ambas ciudades y ambas habitaciones es muy

similar, pues es evidente que la variable independiente no tiene gran poder explicativo, por lo que no son buenos modelos, ya que de las 2 habitaciones de las dos ciudades el R^2 más alto es de 0.021.

Ahora bien, como quinta relación es la influencia del número de revisiones en la tasa de aceptación del anfitrión, la cual de igual manera nos deja en evidencia que el número de revisiones no es una buena variable explicativa en ninguna de las ciudades y en ninguna de las habitaciones. Esto se puede identificar al tener coeficientes de determinación de 0.46 y 0.56 en el DF y de 0.47 y 0.36 en California.

Finalmente, como última relación se encuentra la influencia de las puntuaciones de comunicación en las revisiones por mes. Esto analizándolo por lógica, es notable que no debería haber una relación entre esas puntuaciones con la cantidad de revisiones. Lo anterior se ve reflejado a la hora de evaluar los modelos con el coeficiente de determinación, ya que en ambas ciudades y en ambas habitaciones los R^2 son incluso menores a 0.01, lo cual es bastante bajo.