



Tecnológico de Monterrey

**INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY
CAMPUS QUERÉTARO**

Reporte

Actividad 6.2 (Regresión Lineal Múltiple)

Unidad de formación.

Analítica de datos y Herramientas de

Inteligencia Artificial II

Grupo 101

Integrantes.

Amairany Rodríguez A01702927

Maria Fernanda Martínez A01067198

Leonardo Isaias Guevara A01275327

Jose Ignacio Hernández Rodríguez A01703130

Profesor.

PhD. Alfredo García Suárez

Mayo 2023

En el presente reporte se aterrizan las conclusiones sobre el análisis de las regresiones lineales múltiples. Por lo que se realizará la comparación de los resultados obtenidos sobre las regresiones de Girona y California en comparación con el DF. Así mismo, se presentarán los coeficientes obtenidos en las regresiones mediante tablas para un mejor análisis.

Las variables a analizar son las siguientes:

- “host_acceptance_rate vs host_response_rate, Price, number_of_reviews , review_scores_rating”
- “Price” vs availability_365, host_response_rate, number_of_reviews, accommodates“
- “host_acceptance_rate vs number_of_reviews,review_scores_cleanliness,review_scores_communication ”

Coeficientes

Private Room			
Variables	DF	California	Girona
“host_acceptance_rate vs host_response_rate, Price, number_of_reviews , review_scores_rating”	D: 0.40628823796964 82 C: 0.63740743482457 76	D: 0.6048982748683 568 C: 0.7777520651649 578	D: 0.20860576050060 42 C: 0.45673379610075 3
“Price” vs availability_365, host_response_rate, number_of_reviews, accommodates“	D: 0.00060889751109 30539 C: 0.02467584874108 7993	D: 0.0276261564915 13565 C: 0.1662111804046 6942	D: 0.04448776802545 984 C: 0.21092123654449 743
“host_acceptance_rate vs number_of_reviews,review_scores_cleanliness, review_scores_communication ”	D: 0.14298700551364 285 C: 0.37813622613238 584	D: 0.0478579465413 97633 C: 0.2187645916079 6025	D: 0.03282011581832 678 C: 0.18116322976345 609

Hotel Room			
"host_acceptance_rate vs host_response_rate, Price, number_of_reviews , review_scores_rating"	D: 0.4062882379696482 C: 0.6374074348245776	D: 0.6048982748683568 C: 0.7777520651649578	D: 0.2086057605006042 C: 0.456733796100753
"Price" vs availability_365, host_response_rate, number_of_reviews, accommodates"	D: 0.0006088975110930539 C: 0.024675848741087993	D: 0.027626156491513565 C: 0.16621118040466942	D: 0.04448776802545984 C: 0.21092123654449743
"host_acceptance_rate vs number_of_reviews,review_scores_cleanliness, review_scores_communication "	D: 0.14298700551364285 C: 0.37813622613238584	D: 0.047857946541397633 C: 0.21876459160796025	D: 0.03282011581832678 C: 0.18116322976345609

Tomando en cuenta los resultados de los coeficientes de determinación y correlación de las 3 variables para ambos tipos de habitación, se puede identificar que los coeficientes de determinación ,que son aquellos que hacen referencia a la predicción de modelos futuros, miden la predicción entre las variables, las independientes, en este caso, con la independiente, es decir, miden qué tan bien se ajustan los datos con la regresión obtenida. Se puede observar que los coeficientes cambian dependiendo el país, ya que se usaron 3 bases de datos (Girona, D.F. y California) ,sin embargo, todos se encuentran dentro de un rango similar ya que se analizaron las mismas variables.

Los coeficientes de correlación van desde -1 a 1, mientras más cercano a 1 sea, mayor es la correlación, la más alta fue la de California con alrededor de 0.7 para el caso de "host_acceptance_rate vs host_response_rate, Price, number_of_reviews , review_scores_rating", seguido del DF con 0.6 y Girona con

0.4. Por lo tanto, al tener una relación directa con el coeficiente de determinación, mientras mayor fue el coeficiente de correlación, menor fue la variabilidad entre los datos y tuvieron coeficientes de determinación entre 0.2 a 0.6.

Para los siguientes casos que se analizaron, los coeficientes disminuyen, lo que nos dice que las variables independientes explican menor variabilidad de la variable dependiente, no obstante, no nula. Además, se observó que no hubo ningún cambio significativo entre los tipos de cuarto 'Private room' y 'Hotel room'. Estos coeficientes sólo miden el grado de relación lineal entre nuestras variables, en caso de existir otro tipo de relación entre los datos, tendría que utilizarse otro método.