



Tecnológico de Monterrey

INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY

CAMPUS PUEBLA

ANALÍTICA DE DATOS Y HERRAMIENTAS DE INTELIGENCIA ARTIFICIAL I

REPORTE COMPARATIVO REGRESIÓN

ALUMNOS:

Ángel Roberto González Angulo | A01735880

Sebastián Conteras Hernández | A01736115

Jesús Miguel García Guízar | A01736213

PROFESOR:

Alfredo García Suárez

GRUPO:

101

FECHA DE ENTREGA:

08 - 10 - 2024

Introducción

En este reporte comparativo, se tomaron en cuenta los resultados obtenidos por el análisis individual de las tres ciudades escogidas como Ottawa (Canadá), Ciudad de México (México) y Seattle (Estados Unidos) en cuanto a las regresiones lineales que aplicamos en nuestro código.

Tipos de Habitación

Entire Home / Apt

Ottawa:

Filtro 1: 'host_acceptance_rate vs host_response_rate'

Determinación: 0.14490285210893816

Correlación: 0.38066107248960745

Filtro 2: 'host_acceptance_rate vs price':

Determinación: 0.004915037976859327

Correlación: 0.0701073318338341

Filtro 3: 'host_acceptance_rate vs number_of_reviews':

Determinación: 0.028701315703891006

Correlación: 0.16941462659372422

Filtro 4: 'review_scores_location vs review_scores_cleanliness':

Determinación: 0.16607122787794937

Correlación: 0.4075183773499661

Filtro 5: 'availability_365 vs number_of_reviews':

Determinación: 0.0028664296368576814

Correlación: 0.05353904777690467

Filtro 6: 'reviews_per_month vs review_scores_communication':

Determinación: 0.013340421917210188

Correlación: 0.11550074422794941

CDMX:

Filtro 1: 'host_acceptance_rate vs host_response_rate'

Determinación: 0.216083675534681

Correlación: 0.4648480133706941

Filtro 2: 'host_acceptance_rate vs price':

Determinación: 0.0019703110141576774

Correlación: 0.04438818552450277

Filtro 3: 'host_acceptance_rate vs number_of_reviews':

Determinación: 0.034018897917829305

Correlación: 0.18444212620176906

Filtro 4: 'review_scores_location vs review_scores_cleanliness':

Determinación: 0.16880650223135274

Correlación: 0.41086068469902637

Filtro 5: 'availability_365 vs number_of_reviews':

Determinación: 0.0003168144609013712

Correlación: 0.017799282595132065

Filtro 6: 'reviews_per_month vs review_scores_communication':

Determinación: 0.015270538439908332

Correlación: 0.12357402008475864

Seattle:

Filtro 1: 'host_acceptance_rate vs host_response_rate'

Determinación: 0.19573594698382513

Correlación: 0.4424205544318947

Filtro 2: 'host_acceptance_rate vs price':

Determinación: 0.0026302101779714926

Correlación: 0.051285574755202776

Filtro 3: 'host_acceptance_rate vs number_of_reviews':

Determinación: 0.02486603259661757

Correlación: 0.15768967181339927

Filtro 4: 'review_scores_location vs review_scores_cleanliness':

Determinación: 0.16799694613870075

Correlación: 0.4098743052921234

Filtro 5: 'availability_365 vs number_of_reviews':

Determinación: 0.011605314990562166

Correlación: 0.10772796754122008

Filtro 6: 'reviews_per_month vs review_scores_communication':

Determinación: 0.029980752637688957

Correlación: 0.17314950949306485

Private Room

Ottawa:

Filtro 1: 'host_acceptance_rate vs host_response_rate'

Determinación: 0.15818594060470337

Correlación: 0.39772596169310265

Filtro 2: 'host_acceptance_rate vs price':

Determinación: 0.011932025172673

Correlación: 0.10923380965924881

Filtro 3: 'host_acceptance_rate vs number_of_reviews':

Determinación: 0.03398203108366815

Correlación: 0.18434215764080703

Filtro 4: 'review_scores_location vs review_scores_cleanliness':

Determinación: 0.2652214003551193

Correlación: 0.514996505187287

Filtro 5: 'availability_365 vs number_of_reviews':

Determinación: 0.015592464939040984

Correlación: 0.12486979193960797

Filtro 6: 'reviews_per_month vs review_scores_communication':

Determinación: 0.01708197230820896

Correlación: 0.13069801952672794

CDMX:

Filtro 1: 'host_acceptance_rate vs host_response_rate'

Determinación: 0.20414266700941075

Correlación: 0.45182149905622104

Filtro 2: 'host_acceptance_rate vs price':

Determinación: 0.00023146902592019458

Correlación: 0.015214106149235143

Filtro 3: 'host_acceptance_rate vs number_of_reviews':

Determinación: 0.047445932687816605

Correlación: 0.21782087293879024

Filtro 4: 'review_scores_location vs review_scores_cleanliness':

Determinación: 0.28585617014130227

Correlación: 0.5346551880804135

Filtro 5: 'availability_365 vs number_of_reviews':

Determinación: 0.002709409490651127

Correlación: 0.05205198834483777

Filtro 6: 'reviews_per_month vs review_scores_communication':

Determinación: 0.01469131946425506

Correlación: 0.1212077533174139

Seattle:

Filtro 1: 'host_acceptance_rate vs host_response_rate'

Determinación: 0.17754050910163632

Correlación: 0.42135556137499397

Filtro 2: 'host_acceptance_rate vs price':

Determinación: 0.0063582150423336925

Correlación: 0.07973841635205513

Filtro 3: 'host_acceptance_rate vs number_of_reviews':

Determinación: 0.041578591379144436

Correlación: 0.2039082915899803

Filtro 4: 'review_scores_location vs review_scores_cleanliness':

Determinación: 0.2540802854683133

Correlación: 0.5040637712316898

Filtro 5: 'availability_365 vs number_of_reviews':

Determinación: 0.010904867809836016

Correlación: 0.10442637506796842

Filtro 6: 'reviews_per_month vs review_scores_communication':

Determinación: 0.019982596675173436

Correlación: 0.14135981280114032

Shared Room

Ottawa (Solo 7 registros):

Filtro 1: 'host_acceptance_rate vs host_response_rate'

Determinación: 1.0

Correlación: 1.0

Filtro 2: 'host_acceptance_rate vs price':

Determinación: 1.0

Correlación: 1.0

Filtro 3: 'host_acceptance_rate vs number_of_reviews':

Determinación: 0.066666666666666654

Correlación: 0.2581988897471609

Filtro 4: 'review_scores_location vs review_scores_cleanliness':

Determinación: 1.0

Correlación: 1.0

Filtro 5: 'availability_365 vs number_of_reviews':

Determinación: 0.0374999999999999756

Correlación: 0.1936491673103702

Filtro 6: 'reviews_per_month vs review_scores_communication':

Determinación: 0.0

Correlación: 0.0

CDMX:

Filtro 1: 'host_acceptance_rate vs host_response_rate'

Determinación: 0.14412264795412

Correlación: 0.37963488769358383

Filtro 2: 'host_acceptance_rate vs price':

Determinación: 0.0006610143809677727

Correlación: 0.025710199940252754

Filtro 3: 'host_acceptance_rate vs number_of_reviews':

Determinación: 0.046074064121143676

Correlación: 0.21464869932320502

Filtro 4: 'review_scores_location vs review_scores_cleanliness':

Determinación: 0.36384652476498347

Correlación: 0.6031969203875162

Filtro 5: 'availability_365 vs number_of_reviews':

Determinación: 0.0002410196258190167

Correlación: 0.015524806788460096

Filtro 6: 'reviews_per_month vs review_scores_communication':

Determinación: 0.016343289951090023

Correlación: 0.1278408774652694

Seattle (Solo 10 registros):

Filtro 1: 'host_acceptance_rate vs host_response_rate'

Determinación: 0.5265919995095771

Correlación: 0.7256665897707962

Filtro 2: 'host_acceptance_rate vs price':

Determinación: 0.05508493373927492

Correlación: 0.23470179747772474

Filtro 3: 'host_acceptance_rate vs number_of_reviews':

Determinación: 0.08883953723189753

Correlación: 0.2980596202639625

Filtro 4: 'review_scores_location vs review_scores_cleanliness':

Determinación: 1.0

Correlación: 1.0

Filtro 5: 'availability_365 vs number_of_reviews':

Determinación: 0.2499206198694811

Correlación: 0.4999206135672754

Filtro 6: 'reviews_per_month vs review_scores_communication':

Determinación: 0.11329297859700405

Correlación: 0.3365902235612378

Hotel

CDMX:

Filtro 1: 'host_acceptance_rate vs host_response_rate'

Determinación: 0.014181707906701257

Correlación: 0.11908697622620727

Filtro 2: 'host_acceptance_rate vs price':

Determinación: 0.12102933588274001

Correlación: 0.34789270742966144

Filtro 3: 'host_acceptance_rate vs number_of_reviews':

Determinación: 0.02211737628906285

Correlación: 0.14871911877449667

Filtro 4: 'review_scores_location vs review_scores_cleanliness':

Determinación: 0.27297443921484843

Correlación: 0.5224695581704722

Filtro 5: 'availability_365 vs number_of_reviews':

Determinación: 0.007364541847907247

Correlación: 0.08581690886944861

Filtro 6: 'reviews_per_month vs review_scores_communication':

Determinación: 0.08169115708371177

Correlación: 0.28581664941656526

Seattle (Solo 21 registros):

Filtro 1: 'host_acceptance_rate vs host_response_rate'

Determinación: 0.8013425537988388

Correlación: 0.8951773867780837

Filtro 2: 'host_acceptance_rate vs price':

Determinación: 0.03697917548768759

Correlación: 0.19229970225584747

Filtro 3: 'host_acceptance_rate vs number_of_reviews':

Determinación: 0.12996983294863618

Correlación: 0.36051329094589035

Filtro 4: 'review_scores_location vs review_scores_cleanliness':

Determinación: 1.0

Correlación: 1.0

Filtro 5: 'availability_365 vs number_of_reviews':

Determinación: 0.000400590134396972

Correlación: 0.020014747922393928

Filtro 6: 'reviews_per_month vs review_scores_communication':

Determinación: 0.0

Correlación: 0.0

Conclusiones:

Para Entire Home / Apt y para Private Room se tienen los registros suficientes para hacer un análisis robusto con el cual trabajar con los datos respecto a las regresiones, sin embargo, en cuanto a Shared Room este análisis al menos en la ciudad de Canadá se desvirtúa debido a la cantidad de registros donde los campos y regresiones son más volátiles y para nada eficaces al comparar con otras ciudades.

En el caso de Hotel Room y Shared Room, para CDMX, se cuenta con los suficientes registros para hacer una visualización de las correlaciones entre las distintas variables, sin embargo, Entire / Apt sigue siendo el tipo de habitación que más registros tiene por lo que se tiene una mayor precisión en el cálculo de las regresiones y por lo tanto, se puede hacer una mejor comparativa con Seattle y Ottawa.

Hablando de Seattle, específicamente de Entire Home y Private Room, cuentan con los suficientes registros para observar el comportamiento de las correlaciones entre las variables analizadas. De igual forma, para Hotel Room y Shared Room, al tener un número reducido de registros, la interpretación de los resultados es más complicada, ya que cualquier conclusión puede ser poco representativa (regresiones).