



# Tecnológico de Monterrey

*Instituto Tecnológico y de Estudios Superiores de Monterrey*

**Analítica de datos y herramientas de inteligencia artificial**

## **Actividad 5 (Regresión Lineal Simple)**

**Integrantes:**

Natalia Cedillo Hernández -  
A01660022  
Elena Nivón Hernández - A01174666  
Jarlyn Loza Pacheco - A0176943  
José Jaime Ponce de León -  
A01552256  
Rodrigo Ruiz Teodoro- A01730322

**Profesores:**

Rigoberto Cerino Jiménez  
Candy Yuridia Alemán Muñoz  
Juan Manuel Ahuactzin  
Alfredo García Suárez

En este análisis, se ha utilizado una matriz de correlación para explorar cómo diferentes variables relacionadas con habitaciones de hotel se relacionan entre sí. La matriz de correlación permite visualizar las asociaciones entre estas variables y revelar posibles patrones subyacentes.

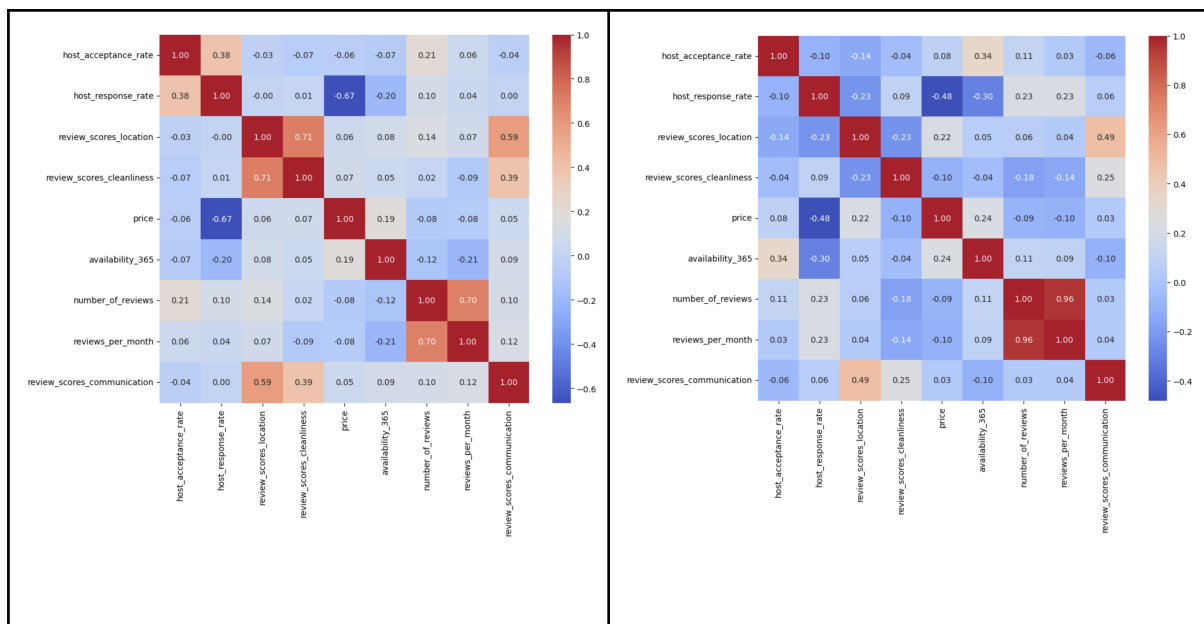
A través de un mapa de calor, se ha destacado las conexiones más fuertes y débiles entre las variables, lo que proporciona información valiosa sobre qué aspectos pueden estar influyendo en el éxito o el desempeño de las habitaciones de hotel. A medida que se examinan estas correlaciones, se podrían descubrir oportunidades para mejorar la experiencia del cliente, optimizar los precios, o tomar decisiones estratégicas que beneficien a la empresa en su conjunto.

En las siguientes secciones, se exploran los resultados de este análisis y se muestran las implicaciones de las correlaciones observadas en el contexto de la gestión hotelera para las ciudades de California y DF. Estos insights pueden servir como base para tomar decisiones más informadas y estratégicas en la industria hotelera en constante evolución.

### **Análisis de gráficos:**

A continuación, se presenta una tabla comparativa con los mapas de calor correspondientes a las ciudades DF y California para el tipo de cuarto, cuarto de hotel:

<b>Comparación con el tipo 'Hotel room'</b> <b>room_type= 'Hotel room'. Para la</b> <b>ciudad de DF</b>	<b>Comparación con el tipo 'Hotel room'</b> <b>room_type= 'Hotel room'. Para la</b> <b>ciudad de California</b>
---	---

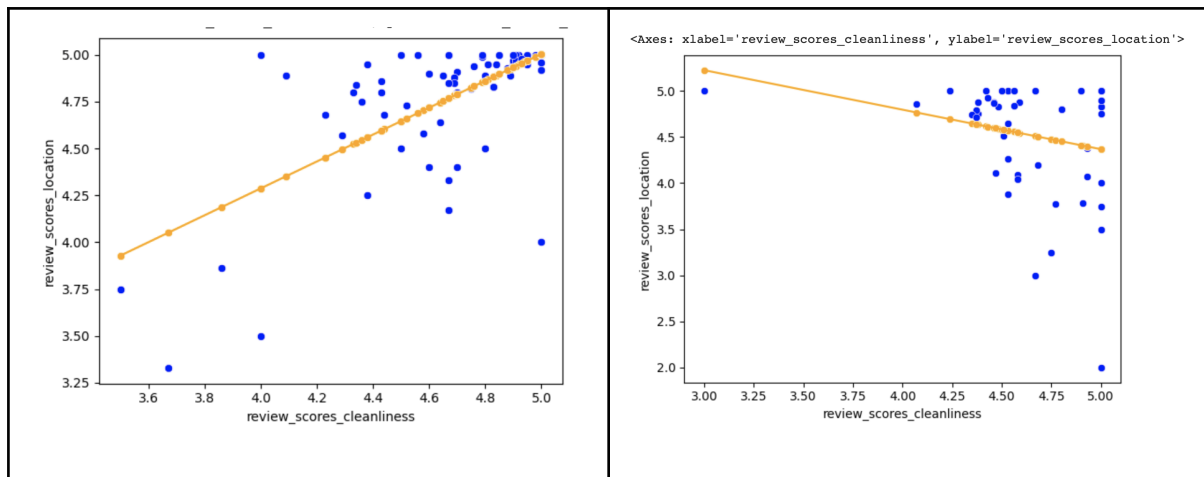


Los dos mapas de calor, muestran patrones diferentes en las relaciones entre las variables. La primera matriz muestra una alta correlación positiva entre "Predicciones\_scores\_cleanliness" y "review\_scores\_cleanliness", mientras que la segunda muestra una correlación negativa perfecta entre estas mismas variables. Además, algunas correlaciones entre otras variables varían significativamente entre las dos matrices, como la relación entre "host\_response\_rate" y "host\_acceptance\_rate". Estas diferencias en las correlaciones sugieren que los datos o el contexto de análisis pueden haber cambiado entre las dos matrices, lo que podría tener implicaciones importantes en la interpretación y las decisiones basadas en estos datos. Se recomienda una revisión cuidadosa de los datos y del proceso de análisis para comprender estas diferencias y sus implicaciones.

Así mismo, se presentan las gráficas comparativas correspondientes a las ciudades DF y California para el tipo de cuarto, cuarto de hotel:

**Gráfica comparativa con el tipo 'Hotel room'**  
**room\_type= 'Hotel room'. Para la ciudad de DF**

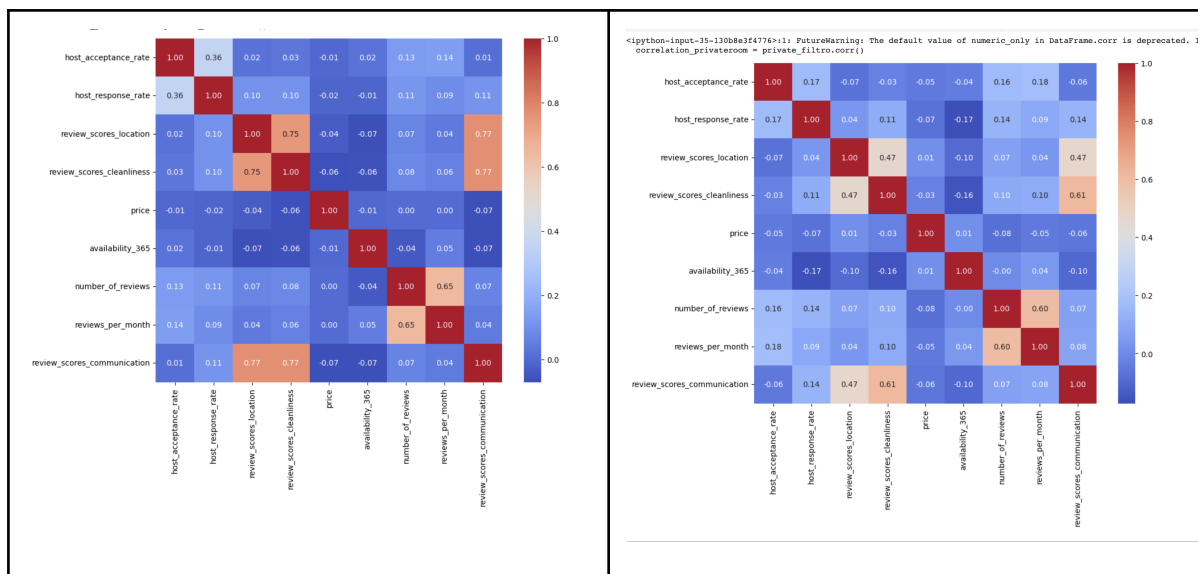
**Gráfica comparativa con el tipo 'Hotel room'**  
**room\_type= 'Hotel room'. Para la ciudad de California**



Las dos gráficas comparativas muestran diferencias notables en las calificaciones de limpieza (review\_scores\_cleanliness), las calificaciones de ubicación (review\_scores\_location) y las predicciones de calificaciones de limpieza (Predicciones\_scores\_cleanliness) entre ellas. La primera gráfica presenta calificaciones más coherentes y cercanas entre sí, con predicciones que se asemejan a las calificaciones reales. En contraste, la segunda tabla muestra una variabilidad más amplia en las calificaciones y predicciones, con algunas desviaciones significativas en comparación con las calificaciones reales. Estas discrepancias pueden sugerir diferencias en los conjuntos de datos, en la calidad de las predicciones o en el contexto subyacente de las calificaciones, lo que subraya la importancia de considerar cuidadosamente las fuentes y el contexto de los datos en el análisis y la toma de decisiones.

Ahora, se presenta una tabla comparativa con los mapas de calor correspondientes a las ciudades DF y California para el tipo de cuarto, cuarto privado:

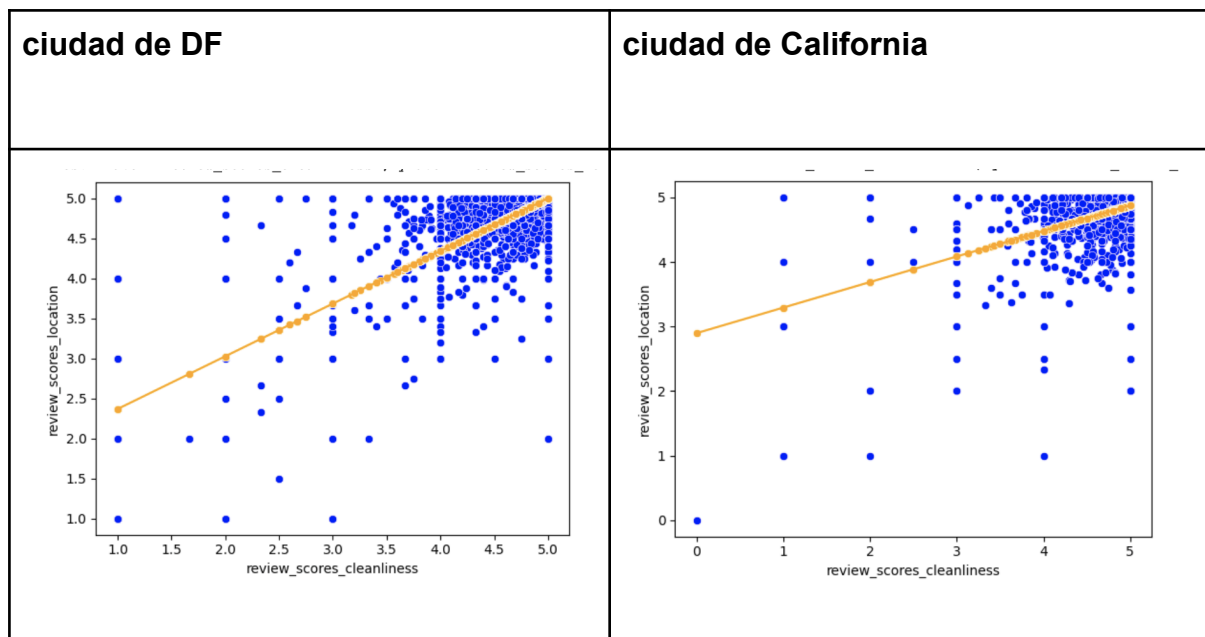
<p><b>Comparación con el tipo 'Private Room'</b></p> <p><b>room_type= 'Private Room'. Para la ciudad de DF</b></p>	<p><b>Comparación con el tipo 'Private Room'</b></p> <p><b>room_type= 'Private Room'. Para la ciudad de California</b></p>
--	--



Los dos conjuntos de matrices de correlación comparadas revelan diferencias notables en las relaciones entre las calificaciones de limpieza, calificaciones de ubicación y predicciones de calificaciones de limpieza en dos contextos distintos. En la primera matriz, relacionada con habitaciones privadas, destaca una alta correlación entre las calificaciones reales de limpieza y las predicciones de calificaciones de limpieza (0.77), así como una fuerte correlación entre las calificaciones de ubicación y ambas (0.75). Esto sugiere que las predicciones están estrechamente vinculadas a las calificaciones reales de limpieza y ubicación. En contraste, la segunda matriz, que parece representar otro conjunto de datos o contexto, muestra correlaciones ligeramente más bajas, aunque aún positivas, entre estas variables (0.61 entre predicciones y calificaciones de limpieza, 0.47 entre calificaciones de ubicación y calificaciones de limpieza). Estas diferencias sugieren que la relación entre estas métricas puede variar según el contexto, enfatizando la importancia de interpretar las correlaciones considerando el contexto específico de los datos.

De igual manera, se presentan las gráficas comparativas correspondientes a las ciudades DF y California para el tipo de cuarto, cuarto privado:

<p><b>Gráfica comparativa con el tipo 'Private Room'</b></p> <p><b>room_type= 'Private Room'. Para la</b></p>	<p><b>Gráfica comparativa con el tipo 'Private Room'</b></p> <p><b>room_type= 'Private Room'. Para la</b></p>
---	---



En este caso las dos gráficas comparativas muestran una considerable similitud en las calificaciones de limpieza (`review_scores_cleanliness`), las calificaciones de ubicación (`review_scores_location`) y las predicciones de calificaciones de limpieza (`Predicciones_scores_cleanliness`) entre ellas, la única diferencia radica en que la primera parece tener mayor cantidad de datos y por ende, más datos dispersos, pero en su mayoría presenta calificaciones más coherentes y cercanas entre sí, con predicciones que se asemejan a las calificaciones reales. En el caso de la segunda tabla, es muy similar a excepción de que tiene menos datos y por ende menos dispersión.

## Análisis de equipo:

Ahora bien, a partir de esto, se decidió analizar los gráficos con las variables "review scores location" y "review scores cleanliness" y se obtuvo que, tienen una alta correlación positiva, esto significa que existe una relación lineal fuerte y positiva entre estas dos variables. A continuación se desglosan algunos insights que se pueden derivar de esta alta correlación:

**1. Relación entre ubicación y limpieza:** La alta correlación positiva sugiere que, en general, las propiedades con una puntuación alta en ubicación tienden a tener una puntuación alta en limpieza. En otras palabras, las ubicaciones que se califican positivamente suelen estar asociadas con propiedades que se consideran limpias.

**2. Importancia de la ubicación:** La ubicación puede ser un factor importante en la percepción de la limpieza de una propiedad. Si los huéspedes consideran que la ubicación es buena, es más probable que tengan una opinión positiva sobre la limpieza de esa propiedad.

**3. Calidad general:** Una alta correlación entre estas dos variables también puede indicar que los huéspedes valoran mucho tanto la ubicación como la limpieza en sus experiencias de alojamiento. Esto puede ser útil para los anfitriones de propiedades de alquiler o hoteles, ya que muestra la importancia de mantener tanto la ubicación como la limpieza a un alto nivel para satisfacer a los clientes.

**4. Decisiones de mejora:** Los propietarios o administradores de propiedades pueden usar esta información para tomar decisiones sobre dónde enfocar sus esfuerzos de mejora. Si estas dos variables tienen una correlación fuerte, podrían considerar que mejorar la limpieza de la propiedad puede tener un impacto positivo en las reseñas de ubicación, y viceversa.

**5. Análisis adicional:** Aunque una alta correlación sugiere una relación fuerte, no necesariamente implica causalidad. Para comprender completamente la relación entre estas dos variables, puede ser útil realizar un análisis adicional, como regresión lineal, para evaluar si una variable realmente afecta a la otra o si hay otros factores en juego.

En general, la alta correlación entre "review scores location" y "review scores cleanliness" proporciona información valiosa sobre cómo los huéspedes perciben y evalúan las propiedades, destacando la importancia de la ubicación y la limpieza en la satisfacción general de los clientes.

En conclusión, el análisis de la matriz de correlación de los datos relacionados con las habitaciones de hotel ha proporcionado una visión valiosa sobre las interacciones y relaciones entre diversas variables. Este enfoque analítico ha permitido identificar patrones y asociaciones que pueden tener un impacto significativo en la gestión y el rendimiento de las operaciones hoteleras.

Los insights obtenidos de este análisis han revelado conexiones tanto fuertes como débiles entre las variables estudiadas, lo que puede guiar a los responsables de la industria hotelera en la toma de decisiones estratégicas. Estas decisiones pueden incluir la optimización de precios, la mejora de la experiencia del cliente o la identificación de áreas de mejora en la gestión de habitaciones de hotel.

En un entorno competitivo como la industria hotelera, la capacidad de comprender y utilizar datos de manera efectiva es esencial para mantener y mejorar la calidad de los servicios ofrecidos. Los resultados de este análisis ofrecen una base sólida para la toma de decisiones informadas y estratégicas, lo que puede traducirse en un mayor éxito y rentabilidad en el sector hotelero.