



Campus Puebla

Nombre:

Manuel Carrión Castañón - A01734873

Gerardo Aburto Cancino - A01552307

Actividad:

Actividad 2. Análisis exploratorio

Materia:

Gestión de proyectos de plataformas tecnológicas

Profesor:

Alfredo García Suárez

Fecha:

08/10/2022

En este reporte se abordará un análisis exploratorio de tres distintas bases de datos de tres ciudades diferentes, las cuales son acerca de México, Málaga y Hawaii; esto con el objetivo de encontrar algún tipo de correlación entre las variables de cada base de datos y observar si alguna de estas variables estudiadas tiene un impacto directo sobre el precio enlistado para el AIRBNB.

Las variables que se estudiarán son las siguientes:

- **Price vs minimum_nights**
- **Price vs number_of_reviews**
- **Price vs reviews_per_month**
- **Price vs calculated_host_listings_count**
- **Price vs availability_365**
- **Price vs number_of_reviews_ltm**

Se estudiarán una por una estas variables para identificar si realmente alguna de estas afecta el precio del AIRBNB, donde el primer paso para poder realizar este análisis exploratorio sobre las bases fue el hecho de la limpieza de estas bases. Donde con ayuda de la herramienta Visual Studio Code fue posible identificar y reemplazar valores nulos dentro de las bases y también identificar y sustituir aquellos valores outliers.

Una vez realizado esta limpieza de datos, se pasó a lo que es el análisis exploratorio de las variables de interés, esto se pudo obtener con ayuda de la herramienta tecnológica Minitab, la cual realmente es bastante útil y sencilla, pudiendo analizar datos de manera rápida, eficaz y sin ningún tipo de complicación, debido a que es bastante intuitiva y se pueden analizar diferentes tipos de datos de distintas maneras, dependiendo de la conveniencia del analista.

En esta herramienta se realizaron los diferentes subsets para cada ciudad, en donde los subsets fueron hechos con la variable “room_type” la cual tiene 4 entradas que son “Entire home”, “Hotel room”, “Shared room” y “Private room”; para cada subset se hizo el análisis de correlación para cada una de las variables antes mencionadas y ver su relación con el precio.

Dentro de las variables con mayor coeficiente de correlación se buscaron los modelos matemáticos que mejor correspondieran a nuestro programando, dándonos como resultado los siguientes modelos.

Modelos Matemáticos.

Ciudad	Variable	Tipo de habitación	Modelo Matemático
Hawái	Price VS Number of reviews ltm	Entire Home / Apt	$\text{price} = 568.87 - 8.926 \text{ number_of_reviews_ltm}$
Hawái	Price VS Number of reviews	Hotel Room	$\text{Price} = 537.52 - 2.180 \text{ number_of_reviews}$
Hawái	Price VS Calculated host listings count	Private Room	$\text{price} = 378.51 + 1.2291 \text{ calculated_host_listings_count}$
Hawái	Price VS Minimum nights	Shared Room	$\text{price} = 48.3 + 7.64 \text{ minimum_nights}$
Málaga	Price VS Number of reviews ltm	Entire Home / Apt	$\text{price} = 55.7 - 1.57 \text{ number_of_reviews_ltm}$
Málaga	Price VS Calculated host listings count	Hotel Room	$\text{price} = 273.0 - 39.0 \text{ calculated_host_listings_count}$
Málaga	Price VS Availability 365	Private Room	$\text{price} = 43.07 + 0.1002 \text{ availability_365}$
Málaga	Price VS Reviews per month	Shared Room	$\text{price} = 21.2 + 36.0 \text{ reviews_per_month}$
México	Price VS Minimum nights	Entire Home / Apt	$\text{price} = 1756.4 + 9.97 \text{ minimum_nights}$
México	Price VS Number of reviews ltm	Entire Home / Apt	$\text{price} = 1983 - 12.48 \text{ number_of_reviews_ltm}$
México	Price VS Availability 365	Hotel Room	$\text{price} = 872 + 3.47 \text{ availability_365}$
México	Price VS Calculated host listings count	Private Room	$\text{price} = 626.9 + 49.04 \text{ calculated_host_listings_count}$
México	Price VS Minimum nights	Shared Room	$\text{price} = 561 + 23.30 \text{ minimum_nights}$

Dentro de esta tabla de modelos matemáticos podemos comparar que los mejores modelos de las ciudades analizadas son en Hawái, rentar una habitación compartida con una correlación del 27.32 % con un modelo matemático para realizar estadística y predicciones con una ecuación igual a $\text{price} = 48.3 + 7.64 \text{ minimum nights}$ que se compara con que la variable que más afecta al precio dentro de las habitaciones en Hawái es las noches de habitación. Dentro de nuestra segunda ciudad analizada, que es Málaga, España, la mejor correlación que nos resultó de 15.78% de las variables de precio, siendo la dependiente y la que más afecta al precio de las habitaciones en Málaga, es la variable independiente del número de reviews por habitaciones compartidas que son las más rentadas según nuestro análisis con una ecuación matemática que sería nuestro modelo igual a $\text{price} = 48.3 + 7.64 \text{ minimum nights}$, lo que nos dice que las personas se enfocan en las opiniones de los demás turistas para elegir su habitación. Por último, el modelo matemático de nuestra ciudad de México, resulto tener mayor correlación en una habitación compartidas de igual manera con un 14.01 % y su modelo matemático respectivo es junto con la variable de mínimo de noches igual a $\text{price} = 561 + 23.30 \text{ minimum_nights}$.

Comparación de coeficientes.

	MÉXICO				MÁLAGA				HAWAII			
	ENTIRE HOME/APT	HOTEL ROOM	PRIVATE ROOM	SHARED ROOM	ENTIRE HOME/APT	HOTEL ROOM	PRIVATE ROOM	SHARED ROOM	ENTIRE HOME/APT	HOTEL ROOM	PRIVATE ROOM	SHARED ROOM
Price VS Minimum nights	0.06%	0.00%	0.00%	14.01%	0.00%	3.79%	0.08%	0.00%	0.04%	2.03%	0.52%	27.32%
Price VS Number of reviews	0.05%	0.00%	0.05%	0.13%	0.57%	6.85%	2.40%	0.00%	0.93%	4.41%	1.26%	1.64%
Price VS Reviews per month	0.04%	0.00%	0.00%	0.49%	0.36%	4.86%	2.25%	15.78%	0.58%	4.38%	0.21%	1.87%
Price VS Calculated host listings count	0.00%	0.48%	1.20%	1.50%	0.37%	12.59%	0.29%	0.00%	1.01%	0.00%	2.33%	0.00%
Price VS Availability 365	0.03%	0.88%	0.06%	0.19%	0.13%	0.00%	2.85%	0.00%	0.00%	0.00%	0.22%	4.22%
Price VS Number of reviews itm	0.06%	0.00%	0.02%	0.17%	0.68%	7.80%	2.51%	1.18%	1.29%	1.30%	1.48%	1.38%

Dentro de esta tabla comparativa mostramos los coeficientes de todas las variables analizadas y de los 4 tipos de habitaciones de las tres ciudades analizadas dentro de esta actividad. Podemos notar que en la ciudad de México la mayor correlación es de un 14.01% y las habitaciones que más se alquilan en Airbnb son habitaciones compartidas y lo que más afecta al precio de estas habitaciones son el número mínimo de noches que las personas eligen quedarse dentro del establecimiento. Por otro lado, en la ciudad de Málaga, de igual manera, la mayor correlación de esta base de datos es dentro de las habitaciones compartidas, dando a entender que los turistas alquilan más habitaciones compartidas dentro de Airbnb en esta ciudad, de igual manera la variable independiente que más afecta al precio de la habitación es el número de reviews que tienen las habitaciones, es decir que el turista que viaja a Málaga se enfoca mucho en las opiniones de los antiguos huéspedes para poder rentarla por Airbnb, con una correlación del 15.78% en esta ciudad. Por último, dentro de Hawái, la correlación más grande de esta ciudad es de un porcentaje del 27.38% siendo la correlación más grande de las

3 ciudades analizadas donde la mayor variante que influye en el precio de una habitación compartida, que de igual forma es la que más se alquilan en el respectivo lugar, son de igual forma que en México los números de noches que te hospedas dentro de esa habitación, es decir que las personas se eleva el precio dependiendo del número de noches que se hospede en la habitación.

En conclusión, podemos notar que las personas les influye mucho el tema de cuantos días se hospedan para visitar partes como México y Hawái; sin embargo, la ciudad de México no tiene mucha correlación en las reviews que obtienen las instalaciones como si lo es en Málaga.