

1. DATA

Se tienen tres tablas con información de Airbnb Nueva York: reseñas de los huéspedes, características de los alojamientos y costos en diferentes periodos del tiempo.

OBJETIVO

Nuestra pregunta de interés es: ¿Podemos encontrar listados similares de Airbnb basados en lo que ofrecen?

Tabla Calendar

	listing_id	date	available	price	adjusted_price	minimum_nights	maximum_nights
0	2595	2021-12-05	f	\$150.00	\$150.00	30.0	1125.0
1	65615	2021-12-05	f	\$85.00	\$85.00	10.0	1125.0
2	65615	2021-12-06	f	\$85.00	\$85.00	10.0	1125.0
3	65615	2021-12-07	f	\$85.00	\$85.00	10.0	1125.0
4	65615	2021-12-08	f	\$85.00	\$85.00	10.0	1125.0
...
13964530	51781100	2022-11-30	t	\$300.00	\$300.00	1.0	365.0
13964531	51781100	2022-12-01	t	\$300.00	\$300.00	1.0	365.0
13964532	51781100	2022-12-02	t	\$300.00	\$300.00	1.0	365.0
13964533	51781100	2022-12-03	t	\$300.00	\$300.00	1.0	365.0
13964534	51781100	2022-12-04	t	\$300.00	\$300.00	1.0	365.0

Listado

id	listing_url	scrape_id	last_scraped	name	description	neighborhood_overview
2595	https://www.airbnb.com/rooms/2595	20211204143024	2021-12-05	Skylit Midtown Castle	Beautiful, spacious skylit studio in the heart...	Centrally located in the heart of Manhattan ju...
3831	https://www.airbnb.com/rooms/3831	20211204143024	2021-12-05	Whole flr w/private bdrm, bath & kitchen(gls ...	Enjoy 500 s.f. top floor in 1899 brownstone, w...	Just the right mix of urban center and local n...
5121	https://www.airbnb.com/rooms/5121	20211204143024	2021-12-05	BlissArtsSpace!	 The space HELLO EVERYONE AND THANK...	NaN
5136	https://www.airbnb.com/rooms/5136	20211204143024	2021-12-05	Spacious Brooklyn Duplex, Patio + Garden	We welcome you to stay in our lovely 2 br dupl...	NaN
5178	https://www.airbnb.com/rooms/5178	20211204143024	2021-12-05	Large Furnished Room Near B'way	Please don't expect the luxury here just a bas...	Theater district, many restaurants around here.
...
53662330	https://www.airbnb.com/rooms/53662330	20211204143024	2021-12-05	HUGE PRIVATE ROOM, AWESOME LOCATION	Must See! SPECTACULAR ONE FAMILY DUPLEX Home!	NaN
53662542	https://www.airbnb.com/rooms/53662542	20211204143024	2021-12-05	LARGE PRVT ROOM, QUIET RESIDENCE PERFECTLOCATION	Must See! SPECTACULAR ONE FAMILY DUPLEX Home!	NaN
...

Tabla Reviews

	listing_id		id	date	reviewer_id	reviewer_name	comments
	0	2595	17857	2009-11-21	50679	Jean	Notre séjour de trois nuits. Nous avons ...
	1	2595	19176	2009-12-05	53267	Cate	Great experience.
	2	2595	19760	2009-12-10	38960	Anita	I've stayed with my friend at the Midtown Cast...
	3	2595	34320	2010-04-09	71130	Kai-Uwe	We've been staying here for about 9 nights, en...
	4	2595	46312	2010-05-25	117113	Alicia	We had a wonderful stay at Jennifer's charming...

891959	53609786	508714999062086277	2021-12-02	56267284	Mathew	Fantastic experience. Thank You, Sam! This apa...	
891960	53609786	509368869955553014	2021-12-03	184485667	Tamara	Immaculate unit in a great neighborhood. The a...	
891961	53612732	509403683179062336	2021-12-03	65072448	Divya	Room is very small though it's close to subway...	
891962	53614162	508723153064804388	2021-12-02	60399002	Andy	Great stay! I wasn't sure what I was expecting...	
891963	53648074	510101726158799786	2021-12-04	416539996	Melissa	amazing beach getaway private the pictures don...	

La descripción de cada variable la podrá encontrar en el excel.

Se debe ocupar la información de las tres tablas para generar los grupos

PUNTOS A CUBRIR:

- Calidad de datos
- EDA - Descripciones
- Reducción de dimensiones
- Número de clusters óptimo
- Clustering
- Visualización de grupos
- Perfilamiento

ENTREGABLE:

- Código en python (Notebook)
 - Ordenado
 - Comentarios en cada resultado obtenido, desde transformaciones, análisis, visualizaciones, método de clustering empleado, definición del número de grupos óptimo , ...