

# REPORTE DE ETAPA 3

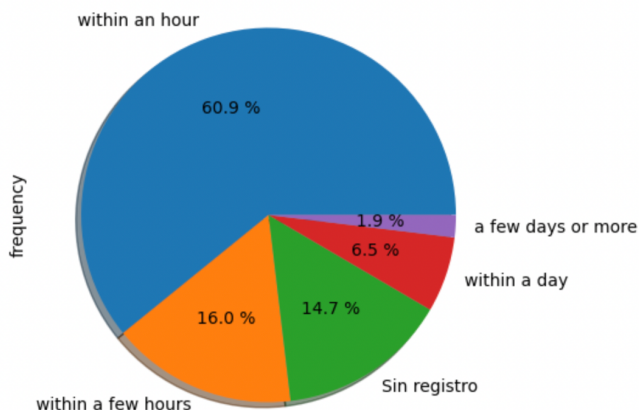
## Evaluación

---

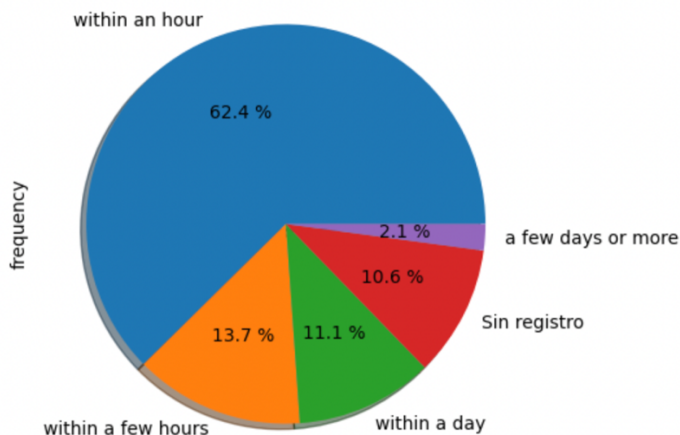
María Fernanda Martínez Ríos  
A01067198

Se hizo uso de 3 bases de datos acerca del servicio Airbnb en 3 lugares: DF, México; Girona, España y el estado de California, Estados Unidos. Anteriormente se realizó la limpieza de los datos, la eliminación de valores outliers para las columnas numéricas, así como la extracción de características de 5 variables de cada data frame, las cuales tuvieron los siguientes resultados:

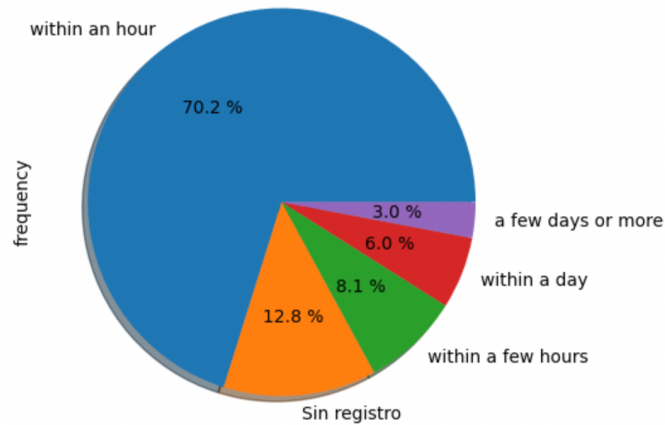
La primera variable fue la de 'host\_response\_time'. Para el caso de California tenemos esta gráfica de pastel y podemos observar que el 60% de los anfitriones de la plataforma contestan dentro de un hora, al contrario del 6.5% que tarda hasta un día.



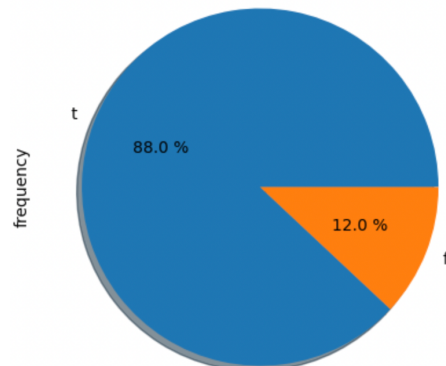
Ahora, si analizamos el de Girona, no hay cambios tan significativos, pero aumenta alrededor de 2% el tiempo de respuesta de una hora, disminuye a 13% en unas horas y a 11% en un día.



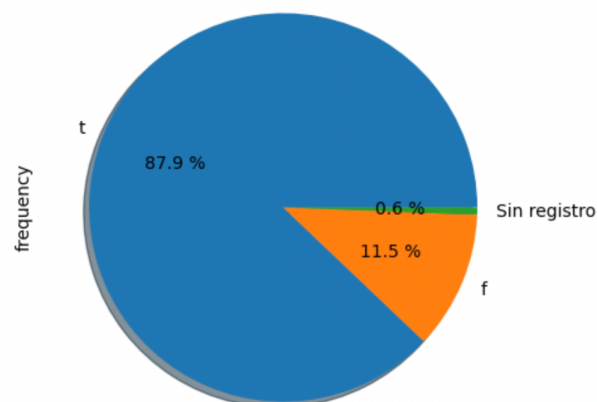
México en cambio es el que tiene el mejor tiempo de respuesta de los 3 lugares, siendo la respuesta en una hora para el 70% de sus registros y por tanto, los demas rangos tienen porcentajes menores, a excepcion de ‘a few days or more’, el cual tiene un 3% a comparación del 1.9 y 2.1% de California y Girona, respectivamente.



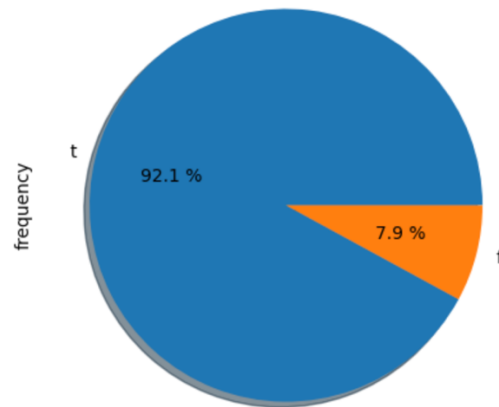
La siguiente variable analizada es una columna de tipo binario sobre la identificacion de los anfitriones en la plataforma, lo cual es un aspecto que suele darle mayor confianza a los clientes. California tiene un 88% de anfitriones identificados en Airbnb.



En Girona, los anfitriones identificados aumentan, la proporcion se ve de la siguiente manera:



Por ultimo, México también tiene en esta segunda variable el porcentaje mas alto con 92.1% de los anfitriones que se encuentran identificados.



Otra variable cualitativa es 'property\_type', la cual tiene muchos valores únicos y su análisis puede ser más complejo, pero solo mostraré aquellas de mayor frecuencia en los datos de cada lugar.

En esta tabla de los tipos de cuarto de California se muestran aquellos registros con frecuencia mayor o igual a 60 y 'Entire rental unit' es la que tiene mayor frecuencia de 1848 del total de este filtro, por lo que representa el 26%, pero en total hay un registro de 6936 filas, por lo que representa el 26%, seguido de 'Entire condo' con 12% y 'Private room in home' y 'Entire home' de igual forma con 12%.

	property_type	frequency	percentage	cumulative_perc
0	Entire rental unit	1848	0.266436	0.266436
1	Entire condo	879	0.126730	0.393166
2	Private room in home	853	0.122982	0.516148
3	Entire home	839	0.120963	0.637111
4	Entire guest suite	495	0.071367	0.708478
5	Private room in rental unit	471	0.067907	0.776384
6	Room in boutique hotel	352	0.050750	0.827134
7	Room in hotel	265	0.038206	0.865340
8	Private room in condo	192	0.027682	0.893022
9	Entire serviced apartment	145	0.020905	0.913927
10	Room in aparthotel	68	0.009804	0.923731
11	Entire townhouse	63	0.009083	0.932814
12	Entire loft	60	0.008651	0.941465

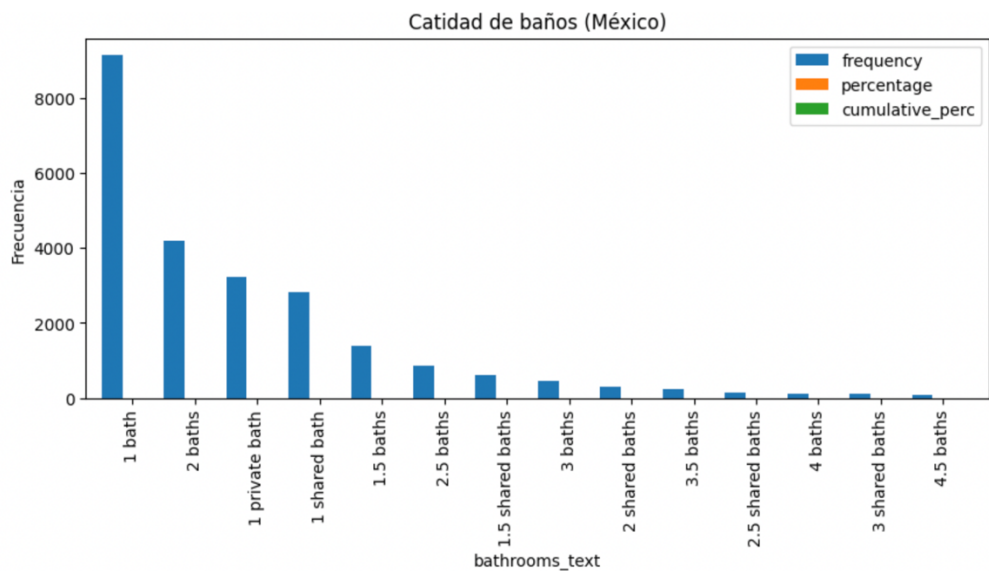
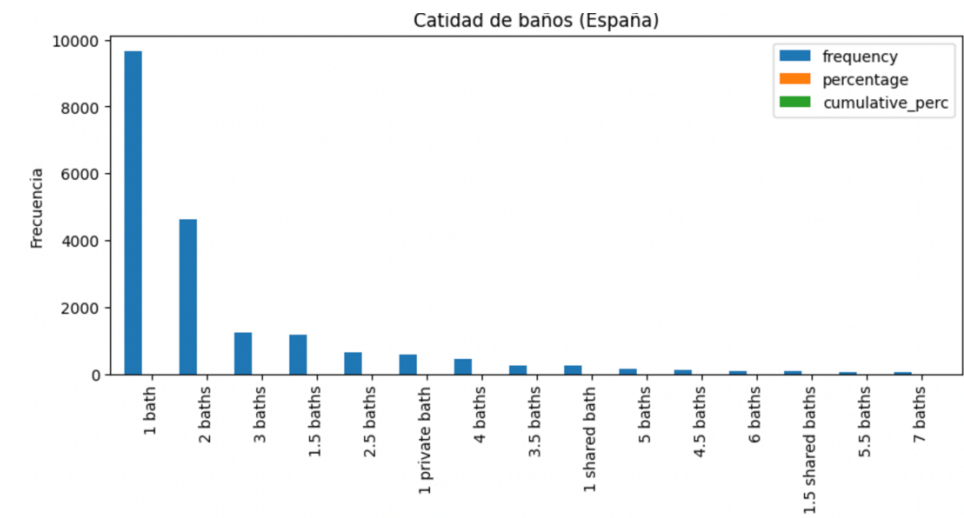
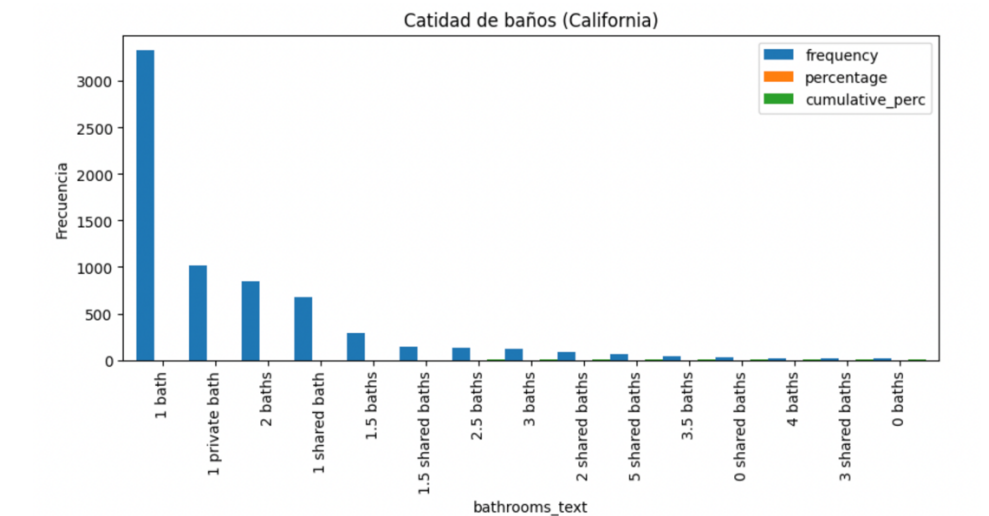
Para Girona, España, por otro lado, la tabla muestra aquellos registros con frecuencia mayor a 160. 'Entire rental unit' representa el 47%, es decir, 21% más que California, seguido de 'Entire Home' con 18% y 'Entire villa' con 8%.

	property_type	frequency	percentage	cumulative_perc
0	Entire rental unit	9339	0.475146	0.475146
1	Entire home	3703	0.188400	0.663546
2	Entire villa	1715	0.087255	0.750801
3	Entire condo	1400	0.071229	0.822030
4	Entire cottage	549	0.027932	0.849962
5	Entire townhouse	453	0.023048	0.873009
6	Entire loft	217	0.011040	0.884050
7	Room in hotel	184	0.009361	0.893411
8	Private room in bed and breakfast	179	0.009107	0.902518
9	Entire serviced apartment	175	0.008904	0.911422
10	Private room in rental unit	172	0.008751	0.920173
11	Entire vacation home	169	0.008598	0.928771
12	Entire chalet	169	0.008598	0.937370

Por ultimo, en el DF, ‘Entire rental unit’ tiene un 33%, menor que Girona, pero mayor que California, el segundo tipo de propiedad con mas frecuencia fue ‘Entire condo’ y ‘Private room in rental unit’. Para los 3 lugares, ‘Entire rental unit’ fue el mayor, aunque en realidad todas las propiedades de tipo ‘Entire’, es decir, entero, que no son compartidas, son las de frecuencia más alta.

	property_type	frequency	percentage	cumulative_perc
0	Entire rental unit	8145	0.336237	0.336237
1	Entire condo	3363	0.138829	0.475066
2	Private room in rental unit	2860	0.118065	0.593131
3	Private room in home	2009	0.082934	0.676065
4	Entire loft	1584	0.065390	0.741455
5	Entire serviced apartment	1067	0.044047	0.785502
6	Private room in condo	917	0.037855	0.823357
7	Entire home	661	0.027287	0.850644
8	Room in hotel	381	0.015728	0.866372
9	Private room in guest suite	297	0.012261	0.878633
10	Private room in serviced apartment	295	0.012178	0.890811
11	Entire guest suite	252	0.010403	0.901214
12	Room in boutique hotel	246	0.010155	0.911369

Ahora, hablaremos de la cantidad de baños que tienen las unidades de renta, en los 3 casos, la mayoría solo tiene un baño y en California hay incluso algunos pocos lugares que rentan sin baño. El maximo es de 7 baños en Girona y algunos se registran como baños compartidos o medios baños.





La última variable que se analizó fue 'host\_acceptance\_rate', en la cual dan los registros como porcentaje ya que mide el nivel de aceptación de los anfitriones de las unidades de renta. Hubo datos únicos porque hay quienes ingresaron cantidades de porcentajes específicos, pero en general, los de las tablas son los que se repitieron más para cada país.

En California, el 32% tiene un nivel de aceptación del 100% de los 6936 totales, por lo que 2252 tienen ese nivel, seguido de 99% con 7.7% y en si los que tienen mayor frecuencia son porcentajes arriba de 80%

	host_acceptance_rate	frequency	percentage	cumulative_perc
0	100%	2252	0.324683	0.367074
1	99%	540	0.077855	0.455094
2	97%	451	0.065023	0.528606
3	98%	258	0.037197	0.570660
4	91%	216	0.031142	0.605868
5	96%	158	0.022780	0.631622
6	95%	143	0.020617	0.654931
7	94%	138	0.019896	0.677425
8	0%	121	0.017445	0.697148

En Girona, el porcentaje de 100% fue en un 38% de los registros y el de 99% en un 11%, por lo que hay mayor aceptación que en los de California, no obstante, hay un 2% que tiene nivel de aceptación 0 y todos los demás están casi en 90%.

	host_acceptance_rate	frequency	percentage	cumulative_perc
0	100%	7478	0.380463	0.404173
1	99%	2221	0.112999	0.524214
2	98%	1049	0.053371	0.580910
3	96%	753	0.038311	0.621608
4	97%	735	0.037395	0.661334
5	91%	518	0.026355	0.689331
6	0%	407	0.020707	0.711329
7	89%	391	0.019893	0.732461
8	95%	291	0.014805	0.748189
9	94%	267	0.013584	0.762620
10	93%	233	0.011854	0.775213

	host_acceptance_rate	frequency	percentage	cumulative_perc
0	100%	11617	0.479566	0.523737
1	99%	2704	0.111625	0.645643
2	98%	1031	0.042561	0.692124
3	97%	711	0.029351	0.724178
4	0%	701	0.028938	0.755782
5	96%	575	0.023737	0.781705
6	94%	430	0.017751	0.801091
7	93%	279	0.011518	0.813669
8	92%	268	0.011063	0.825752
9	95%	264	0.010898	0.837654
10	50%	233	0.009619	0.848158

México tiene el nivel de aceptación de 100% mas alto (47%), por lo que podemos decir que en DF la mayoría de los clientes se siente muy comodo con su renta o servicio. Al igual que en Girona, tiene un 11% dio un 99% de aceptación. Pero a la vez, en México hay el porcentaje mas alto de aceptación cero, con 701 registros así y California con solo 121.