



The Airbnb in Sydney

Introduction

After moving to Sydney, I stayed in an Airbnb during my first month. This experience piqued my interest in analysing Airbnb listings in Sydney. In this project, I used MySQL to explore, process and analyse data and then presented the result using Tableau.

Inside Airbnb is a mission-driven project that provides data and advocacy about Airbnb's impact on residential communities. There were seven files provided for each city. I downloaded "listings.csv" in Sydney for the fourth quarter of 2023. According to the description, it summarised information and metrics which are handy for analysis.

Sydney, New South Wales, Australia

12 December, 2023 ([Explore](#))

Country/City	File Name	Description
Sydney	listings.csv.gz	Detailed Listings data
Sydney	calendar.csv.gz	Detailed Calendar Data
Sydney	reviews.csv.gz	Detailed Review Data
Sydney	listings.csv	Summary information and metrics for listings in Sydney (good for visualisations).
Sydney	reviews.csv	Summary Review data and Listing ID (to facilitate time based analytics and visualisations linked to a listing).
Sydney	neighbourhoods.csv	Neighbourhood list for geo filter. Sourced from city or open source GIS files.
Sydney	neighbourhoods.geojson	GeoJSON file of neighbourhoods of the city.

Data Preparation

I imported the dataset into TablePlus, a user-friendly tool for managing relational databases like MySQL. To start, I checked the shape of data. There are 18 columns and 25,480 rows in the dataset.

#	column_name	data_type
1	id	varchar(255)
2	name	text
3	host_id	varchar(255)
4	host_name	text
5	neighbourhood_group	text
6	neighbourhood	text
7	latitude	float
8	longitude	float
9	room_type	text
10	price	int
11	minimum_nights	int
12	number_of_reviews	int
13	last_review	date
14	reviews_per_month	float
15	calculated_host_listings_count	int
16	availability_365	int
17	number_of_reviews_ltm	int
18	license	text

```

1 # Check the number of rows imported: 25480
2 SELECT COUNT(*)
3 FROM listings;

```

	COUNT(*)
1	25480

Data Analysis

There were several questions I wanted to analyse and explore:

- How many listings and hosts are there in Sydney?
- Which suburb has the most listings in Sydney and what is the price range?
- What is the most common room type of listings?
- Do hosts use Airbnb as a platform for long-term rental or short-term rental?

1. How many listings and hosts are there in Sydney?

As I explored the data, I realised that there might be inactive listings. I considered listings as inactive when both 'availability_365' and 'number_of_reviews' columns are equal to zero because they show no signs of recent engagement or future availability. There are 3,517 inactive listings, accounting for 16% of the total. Since it will impact the result of the analysis, I will not consider them in the following analysis. After excluding the inactive listings, the number of listings in Sydney is 21,963.

```

18 # Get the number of inactive listings - 3517
19 SELECT COUNT(id) AS inactive_listings
20 FROM listings
21 WHERE availability_365 = 0 AND number_of_reviews = 0;
22
23 # Get the number of active listings - 21963
24 SELECT COUNT(id) AS active_listings
25 FROM listings
26 WHERE NOT (availability_365 = 0 AND number_of_reviews = 0);
27
28 # Create a new table with active listings only
29 CREATE TABLE active_listings AS
30 SELECT *
31 FROM listings
32 WHERE NOT (availability_365 = 0 AND number_of_reviews = 0);
33
34 # Check active listings - 21963
35 SELECT COUNT(id)
36 FROM active_listings;
37

```

	COUNT(id)
1	21963

There are 13,813 hosts in Sydney.

```

38 # Get the number of hosts - 13813
39 SELECT COUNT(DISTINCT host_id) AS total_host
40 FROM active_listings;
41

```

	total_host
1	13813

We can also take a look at how many listings these hosts have.

```

42 # Get the number of listings per host
43 SELECT DISTINCT host_id, host_name, calculated_host_listings_count AS num_listings
44 FROM active_listings
45 ORDER BY num_listings DESC;
46

```

	host_id	host_name	num_listings
1	279001183	MadeComfy	245
2	7409213	L'Abode Accommodation Specialist	164
3	288743418	Ken	143
4	15739069	The Apartment Service	142
5	16357713	PropertyProviders	117
6	2450066	Hotelesque	101
7	301753450	Tim	95
8	23717928	Palm Beach Holiday Rentals	93
9	91961414	Kimi	91
10	382384057	360 (Hidden By Airbnb)	67

With a quick search on Google, the top 2 hosts, MadeComfy and L'abode Accommodation Specialist, are property management agencies. 6 hosts are property management agencies among the top 10 hosts with the most listings. This was an interesting finding, as all the Airbnbs I've stayed in were hosted by individual owners, so I never realised that property agencies have a significant presence in the Airbnb market.

I also wanted to find out how many of the hosts have more than one listing. It showed that 17.03% of the hosts have more than 1 listing. This indicates that a portion of the market is treating their Airbnb operations more like a business rather than Airbnb's original concept of sharing a space with people who are visiting the area.

```

47 # Get the number of hosts who have more than 1 property - 2352, 2352/13813 = 17.03%
48 SELECT SUM(IF(num_listings > 1,1,0)) AS num_more_than_1,
49        ROUND(SUM(IF(num_listings > 1,1,0))/COUNT(*)*100, 2) AS percentage
50 FROM (
51     SELECT DISTINCT host_id, host_name, calculated_host_listings_count AS num_listings
52     FROM active_listings
53     ) AS distinct_host;

```

	num_more_than_1	percentage
1	2352	17.03

2. Which suburb has the most listings in Sydney and what is the price range?

I am interested in identifying the popular areas for Airbnb listings. There are 38 areas with listings in Sydney and the average daily price is \$343.76.

```

55 # Get the number of areas with listings - 38
56 SELECT COUNT(DISTINCT neighbourhood) AS num_area
57 FROM active_listings;
58
59 # Get the average price - 343.76
60 SELECT ROUND(AVG(price),2) AS average_price
61 FROM active_listings;

```

	num_area		average_price
1	38	1	343.76

Here are the top 10 areas with the most listings in Sydney.

	neighbourhood	num_listing	average_price
1	Sydney	5043	264.85
2	Waverley	2448	394.27
3	Randwick	1566	303.70
4	Warringah	1122	420.95
5	Pittwater	1086	976.97
6	Manly	1040	489.86
7	North Sydney	838	320.81
8	Woollahra	737	617.47
9	Marrickville	658	192.02
10	Leichhardt	541	304.52

The area of Sydney has the highest number of listings, with a total of 5,043. Waverley and Randwick are also popular areas with 2,448 and 1,566 listings, respectively, suggesting they are preferred destinations within Sydney.

The average daily prices across these areas vary significantly. Pittwater stands out with the highest average daily price of \$976.97, which is substantially more than the others.

Marrickville appears to be the most affordable area among the top 10, with an average daily price of \$192.02.

3. What is the most common room type of the listings in Sydney?

Continuing from the previous question, I would like to further explore the prices of different room types. Entire homes or apartments are the most prevalent room type listed, constituting 75.03% of all listings. Most of the Airbnb hosts rent out their entire properties.


```

69 # Get the number of listings of each room type - Most common/Most expensive: Entire
70 SELECT room_type, COUNT(id) AS num_listings,
71 ROUND(
72     COUNT(id) / (SELECT COUNT(id)
73         FROM active_listings) * 100, 2) AS percentage,
74 ROUND(avg(price), 2) AS average_price
75 FROM active_listings
76 GROUP BY room_type
77 ORDER BY num_listings DESC;

```

	room_type	num_listings	percentage	average_price
1	Entire home/apt	16479	75.03	414.57
2	Private room	5222	23.78	123.20
3	Shared room	167	0.76	99.41
4	Hotel room	95	0.43	277.32

How about the price in different areas? The highest-priced areas are not in the most listed areas. The most expensive area listed is Pittwater, with an average price of \$1,008 for an entire home/apt, indicating it may be a premium location. Besides, for the top 3 highest-priced areas, the room type is all Entire home/apt, which is reasonable considering that the average price for Entire home/apt is significantly higher than that of the other room types.

```

79 # Get the price by area and room type - Pittwater, Entire home/apt, 1008
80 SELECT neighbourhood AS suburb, room_type, ROUND(AVG(price), 0) AS average_price
81 FROM active_listings
82 GROUP BY suburb, room_type
83 ORDER BY average_price DESC;

```

	suburb	room_type	average_price
1	Pittwater	Entire home/apt	1008
2	Mosman	Entire home/apt	857
3	Woollahra	Entire home/apt	699
4	Strathfield	Shared room	586
5	Manly	Entire home/apt	527
6	Sutherland Shire	Hotel room	489
7	Warringah	Entire home/apt	457
8	Waverley	Entire home/apt	456
9	The Hills Shire	Entire home/apt	452
10	Sutherland Shire	Entire home/apt	428

Then I checked areas and prices for each room type.
For Entire room,

```

85 # Get the average price of Entire home/apt by areas – Pittwater 1008
86 SELECT neighbourhood, room_type,
87 ROUND(avg(price), 0) AS average_price
88 FROM active_listings
89 WHERE room_type = "Entire Home/apt"
90 GROUP BY neighbourhood
91 ORDER BY average_price DESC;

```

	neighbourhood	room_type	average_price
1	Pittwater	Entire home/apt	1008
2	Mosman	Entire home/apt	857
3	Woollahra	Entire home/apt	699
4	Manly	Entire home/apt	527
5	Warringah	Entire home/apt	457
6	Waverley	Entire home/apt	456
7	The Hills Shire	Entire home/apt	452
8	Sutherland Shire	Entire home/apt	428
9	City Of Kogarah	Entire home/apt	406
10	Randwick	Entire home/apt	392

For hotel room,

```

93 # Get the average price of Hotel room by areas – Sutherland Shire: 489
94 SELECT neighbourhood, room_type,
95 ROUND(avg(price), 0) AS average_price
96 FROM active_listings
97 WHERE room_type = "Hotel room"
98 GROUP BY neighbourhood
99 ORDER BY average_price DESC;

```

	neighbourhood	room_type	average_price
1	Sutherland Shire	Hotel room	489
2	Auburn	Hotel room	378
3	Sydney	Hotel room	315
4	Waverley	Hotel room	221
5	Botany Bay	Hotel room	219
6	Randwick	Hotel room	214
7	Marrickville	Hotel room	185
8	Manly	Hotel room	179
9	Parramatta	Hotel room	145
10	North Sydney	Hotel room	74

For private room,

```

101 # Get the average price of Private room by areas – Pittwater 375
102 SELECT neighbourhood, room_type,
103 ROUND(avg(price), 0) AS average_price
104 FROM active_listings
105 WHERE room_type = "Private room"
106 GROUP BY neighbourhood
107 ORDER BY average_price DESC;
108

```

	neighbourhood	room_type	average_price
1	Pittwater	Private room	375
2	Woollahra	Private room	310
3	Willoughby	Private room	273
4	North Sydney	Private room	184
5	Penrith	Private room	165
6	Warringah	Private room	163
7	Manly	Private room	144
8	Sydney	Private room	139
9	Waverley	Private room	128
10	Lane Cove	Private room	127

For shared room,

```

109 # Get the average price of Shared room by areas – Strathfield 586
110 SELECT neighbourhood, room_type,
111 ROUND(avg(price), 0) AS average_price
112 FROM active_listings
113 WHERE room_type = "Shared room"
114 GROUP BY neighbourhood
115 ORDER BY average_price DESC;
116

```

	neighbourhood	room_type	average_price
1	Strathfield	Shared room	586
2	Woollahra	Shared room	256
3	Blacktown	Shared room	210
4	Camden	Shared room	198
5	North Sydney	Shared room	194
6	Botany Bay	Shared room	186
7	Leichhardt	Shared room	152
8	Ryde	Shared room	143
9	Warringah	Shared room	123
10	The Hills Shire	Shared room	120

The average price of the shared room in Strathfield is 586 which is surprisingly high.

4. Do hosts use Airbnb as a platform for long-term rental or short-term rental?

A rental is considered long-term if it is 28 days or longer. To identify long-term listings, I looked at the column 'minimum_nights'. The result shows that the majority of listings (57.51%) are available for short stays of 1 week. Additionally, a significant portion of the listings (38.72%) are available for more than 1 month, suggesting that a substantial number of hosts are offering their properties for long-term stays.

In terms of price, the average price for listings available for more than 1 month is \$255, which is lower than the weekly price. There could be a long-term rental discount provided. From my experience, I received a 20% discount for a one-month stay in Surry Hills.

```
117 # Get night range and average price - More than 1 month 39%
118 SELECT night_range, COUNT(id) AS num_listings,
119        ROUND(
120        COUNT(id) / (SELECT COUNT(id)
121                     FROM active_listings) * 100, 2) AS percentage,
122        ROUND(AVG(price), 0) AS average_price
123 FROM (
124        SELECT id, minimum_nights,
125        CASE WHEN minimum_nights BETWEEN 0 AND 7 THEN "1 week"
126             WHEN minimum_nights BETWEEN 8 AND 14 THEN "2 weeks"
127             WHEN minimum_nights BETWEEN 15 AND 21 THEN "3 weeks"
128             WHEN minimum_nights BETWEEN 22 AND 28 THEN "4 weeks"
129             ELSE "more than 1 month" END AS night_range,
130        price
131        FROM active_listings
132        ) AS sub_active_listings
133 GROUP BY night_range
134 ORDER BY night_range;
```

	night_range	num_listings	percentage	average_price
1	1 week	12631	57.51	401
2	2 weeks	402	1.83	468
3	3 weeks	201	0.92	274
4	4 weeks	224	1.02	223
5	more than 1 month	8505	38.72	255

Conclusion

In this report, I used MySQL to get a better understanding of the ecosystem of Airbnb in Sydney. The primary stakeholders for this analysis are guests like myself who prefer to use Airbnb to find temporary accommodation. The goal is to help these users make informed decisions and enhance their Airbnb experience.

To sum up the findings:

- Listings and Hosts in Sydney

- 1) There are 21,963 listings and 13,813 Airbnb hosts in Sydney. Property management agencies are prominent among the top hosts, with 6 out of the top 10 hosts being agencies.
- 2) 17% of hosts have more than one listing.

- Suburbs with Most Listings and Pricing

- 1) The Sydney area has the most listings, with 5,043 properties available. Waverley and Randwick follow as popular suburbs.
- 2) The average daily price in Sydney is \$344, but prices vary widely among areas, with Pittwater being the most expensive at \$977 on average per day.

- Most Common Room Type and Pricing

75% of listings are entire homes or apartments. The top 3 most expensive areas all offer Entire home/apt types, with prices notably higher than other room types.

- Most Common Room Type and Pricing

Most listings (57%) are for short-term rentals of one week. 39% of listings are available for long-term stays exceeding one month.

I visualised the findings by [Tableau](#).

