

Airbnb Lisbon Dashboard: Prices, Availability, Listings

Room Type

(All)

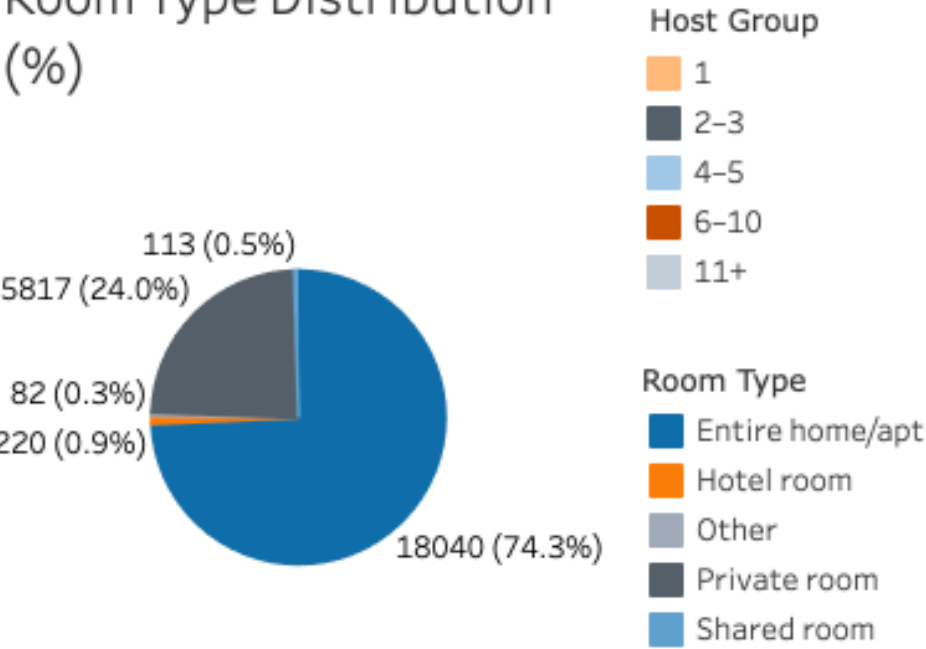
Host Group

(All)

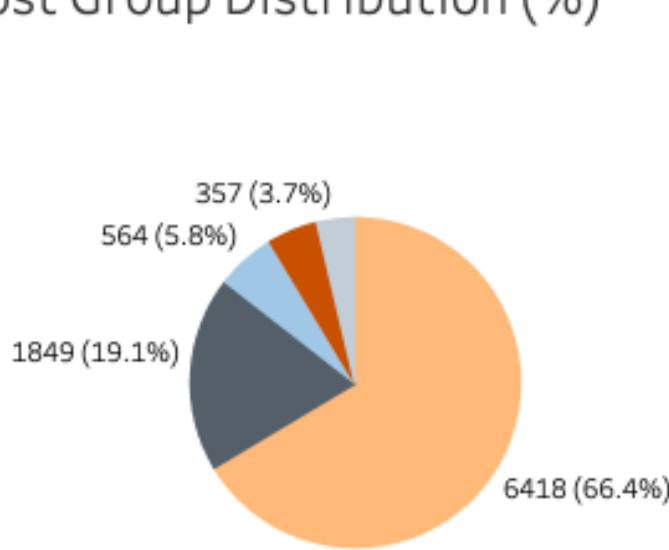
Price

11000

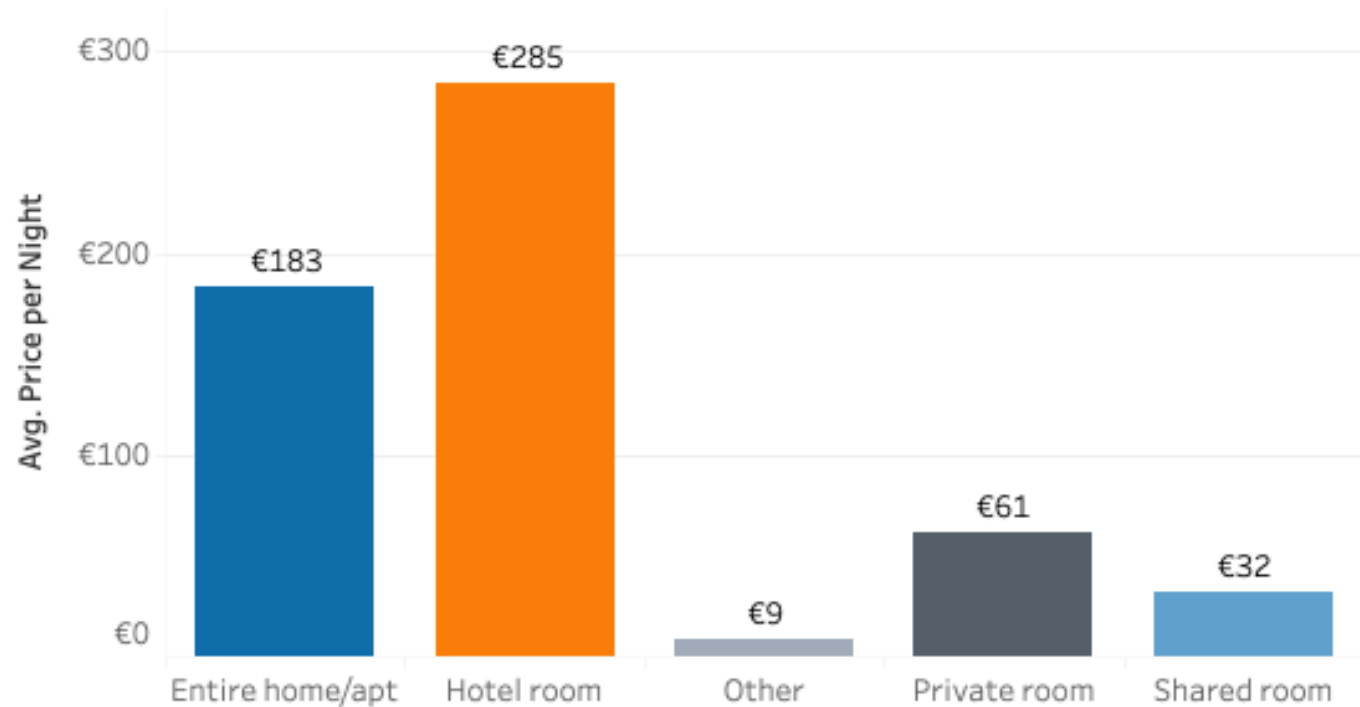
Room Type Distribution (%)



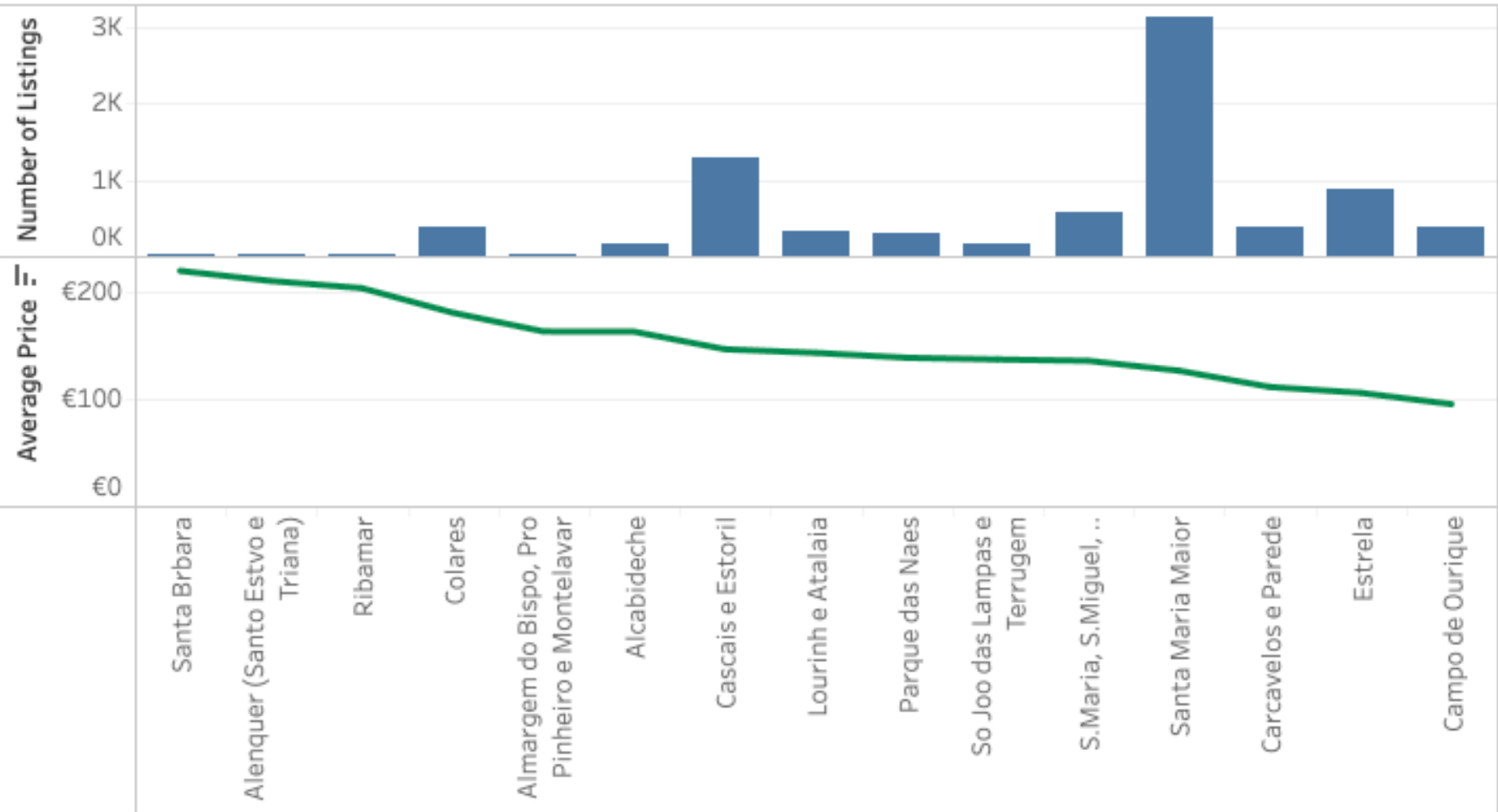
Host Group Distribution (%)



Average Price per Room Type



Average Price and Listings Count by Neighbourhood



Top 15 Lisbon Neighbourhoods: High Availability, Low Price

Mina de gua €72 228 days	So Domingos de Rana €101 212 days	Moscavide e Portela €73 199 days	So Joo das Lampas e Terrugem €198 199 days	Colares €294 197 days
Silveira €109 224 days	guas Livres €59 210 days			
Campolide €89 215 days	A dos Cunhados e Maceira €132 206 days	S.Maria, S.Miguel, S.Martinho, S.Pedro Penaferirim	Parque das Naes €215 194 days	Lourinh e Atalaia €204 194 days
Sacavm e Prior Velho €104 214 days	Santa Maria Maior €178 203 days	Alcntara €92 196 days		

Airbnb Lisbon — Ціни, Доступність, Пропозиція

1. Опис проєкту



1.1. Мета

Метою даного аналізу є дослідити пропозицію житла на Airbnb у місті Лісабон та виявити можливості для покращення ефективності платформи.

Аналіз спрямований на виявлення факторів, що впливають на вартість оренди, доступність житла протягом року та активність хостів.

Особлива увага приділяється:

- виявленню найпопулярніших форматів житла,
- визначенню районів із найвищою концентрацією пропозицій та попиту,
- аналізу взаємозв'язку між ціною, відгуками та кількістю бронювань.

Отримані інсайти можуть бути використані для:

- рекомендацій хостам щодо оптимального ціноутворення та підвищення конверсії,
- визначення цільових сегментів житла для маркетингових кампаній,
- покращення продуктового досвіду платформи для орендарів.

1.2. Джерело даних:

Завантажила дані з сайту <https://insideairbnb.com/get-the-data/>, а саме обрала для аналізу документ *listings.csv*.

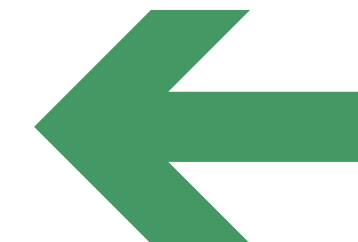
Дані взяті від 08 березня 2025 року.

Весь аналіз далі проводила за допомогою BigQuery.

2. Підготовка та аналіз даних

Огляну дані:

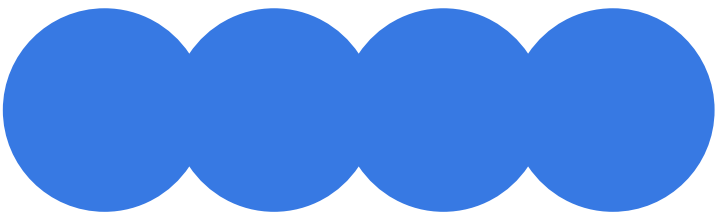
```
1 SELECT
2     column_name,
3     data_type,
4     is_nullable
5 FROM `my-test-project-445920`.INFORMATION_SCHEMA.COLUMNS
6 WHERE table_name = 'listings';
```



column_name	data_type	is_nullable	mean
id	STRING	YES	Основний ключ
name	STRING	YES	Назва оголошення
host_id	STRING	YES	Ідентифікатор хоста
host_name	STRING	YES	Ім'я хоста
neighbourhood_group	STRING	YES	Велика група районів
neighbourhood	STRING	YES	Район міста
latitude	FLOAT64	YES	Геокоординати(широта)
longitude	STRING	YES	Геокоординати(довгота)
room_type	STRING	YES	Тип житла (apartment, room)
price	INT64	YES	Ціна за ніч
minimum_nights	INT64	YES	Мінімальна кількість ночей
number_of_reviews	STRING	YES	Кількість відгуків
last_review	STRING	YES	Дата останнього відгуку
reviews_per_month	FLOAT64	YES	Середня кількість відгуків
calculated_host_listings_count	INT64	YES	Скільки оголошень має хост
availability_365	INT64	YES	Скільки днів у році доступне
number_of_reviews_ltm	STRING	YES	Відгуки за останні 12 місяців
license	STRING	YES	Ліцензія

Що ми бачимо:

Наші дані мають 18 колонок з різними типами даних.
Я додатково описала що означає кожна колонка.



Наступним запитом я хочу подивитись загальну інформацію про дані:

```
1 SELECT
2   COUNT(*) AS total_rows,
3   COUNT(DISTINCT id) AS unique_listings,
4   COUNT(DISTINCT host_id) AS unique_hosts,
5   COUNT(DISTINCT neighbourhood) AS unique_neighbourhoods,
6   COUNT(DISTINCT room_type) AS unique_room_types
7 FROM `my-test-project-445920.airbnb_lisbon.listings`;
```

Query results Save result

Job information	Results	Chart	JSON	Execution details	Execution graph
Row	total_rows	unique_listings	unique_hosts	unique_neighbourho	unique_room_types
1	24278	24270	9663	199	54

Тут ми бачимо, що :

- unique_listings ≈ total_rows — майже всі рядки унікальні
- 9,663 unique_hosts на 24,270 об'єктів
- 199 unique_neighbourhoods
- 54 unique_room_types- мені здається, що тут було щось позначено не вірно, це додатково потрібно перевірити, бо забагато типів кімнат

Розумію аномальність кількості типів кімнат тому хочу відразу подивитись детальніше, що там не так:

```
1 SELECT
2   room_type,
3   COUNT(*) AS total
4 FROM `my-test-project-445920.airbnb_lisbon.listings`
5 GROUP BY room_type
6 ORDER BY total DESC;
```



В даних я бачу багато різних не зрозумілих позначень.
Є тільки 4 основних типи кімнат.
Потрібно їх залишити, а інші прибрати.

Query results [Save results](#)

Job information	Results	Chart	JSON	Execution details	Execution graph
Row	room_type	total			
1	Entire home/apt	18040			
2	Private room	5817			
3	Hotel room	220			
4	Shared room	113			
5	null	25			
6	59	3			
7	60	2			
8	42	2			
9	50	2			
10	68	2			
11	94	2			
12	66	2			

Results per page: 50 1 – 50 of 55

Залишаємо основні типи кімнат:



```
1 CREATE OR REPLACE VIEW `my-test-project-445920.airbnb_lisbon.listings_cleaned` AS
2 SELECT
3   *,
4   CASE
5     WHEN room_type IN ('Entire home/apt', 'Private room', 'Shared room', 'Hotel room') THEN room_type
6     ELSE 'Other'
7   END AS room_type_cleaned
8 FROM `my-test-project-445920.airbnb_lisbon.listings`;
```

Перевіряю, що залишилось тепер:



Query results			
Job information		Results	Chart
		JSON	
Row	room_type_cleaned	total	
1	Entire home/apt	18040	
2	Private room	5817	
3	Hotel room	220	
4	Shared room	113	
5	Other	88	

Далі хочу подивитись скільки у нас пустих значень в таблиці. Обираю для дослідження тільки потрібні мені колонки. Адже не всі буду використовувати в запиті.

```
1 SELECT
2   COUNT(*) AS total_rows,
3   COUNTIF(room_type IS NULL) AS null_room_type,
4   COUNTIF(price IS NULL) AS null_price,
5   COUNTIF(minimum_nights IS NULL) AS null_minimum_nights,
6   COUNTIF(neighbourhood IS NULL) AS null_neighbourhood,
7   COUNTIF(availability_365 IS NULL) AS null_availability,
8   COUNTIF(reviews_per_month IS NULL) AS null_reviews_per_month
9 FROM `my-test-project-445920.airbnb_lisbon.listings`;
```

Query results

Save resultsOpen in

Job informationResultsChartJSONExecution detailsExecution graph

Row	total_rows	null_room_type	null_price	null_minimum_nights	null_neighbourhood	null_availability	null_reviews_per_mo
1	24278	25	3192	16	16	16	3559



Бачимо, що:

- Загально у нас 24,278 рядків.
- З них деякі категорії не значну кількість відсутніх значень мають (minimum_nights, neighbourhood, availability_365 - по 16ть).
- Багато відсутніх значень у відгуках (reviews_per_month) - що досить природньо.

Відразу прибираю NULL значення:

```
1 CREATE OR REPLACE VIEW `my-test-project-445920.airbnb_lisbon.listings_cleaned_final` AS
2 SELECT
3     *,
4     -- Заміняю NULL у reviews_per_month на 0
5     IFNULL(reviews_per_month, 0) AS reviews_per_month_cleaned
6 FROM `my-test-project-445920.airbnb_lisbon.listings_cleaned`
7 WHERE room_type IS NOT NULL
8     AND price IS NOT NULL
9     AND minimum_nights IS NOT NULL
10    AND neighbourhood IS NOT NULL
11    AND availability_365 IS NOT NULL;
```

Та перевіряю відразу чи прибрались пусті значення:

```
1 SELECT
2     COUNTIF(room_type_cleaned IS NULL) AS null_room_type,
3     COUNTIF(price IS NULL) AS null_price,
4     COUNTIF(minimum_nights IS NULL) AS null_minimum_nights,
5     COUNTIF(neighbourhood IS NULL) AS null_neighbourhood,
6     COUNTIF(availability_365 IS NULL) AS null_availability,
7     COUNTIF(reviews_per_month_cleaned IS NULL) AS null_reviews_per_month_cleaned
8 FROM `my-test-project-445920.airbnb_lisbon.listings_cleaned_final`;
```



Query results

[Save results](#)

Job information		Results	Chart	JSON	Execution details	Execution graph	
Row	null_room_type	null_price	null_minimum_nights	null_neighbourhood	null_availability	null_reviews_per_mo	
1	0	0	0	0	0	0	

Короткий висновок кроків, які були вище:



В дата сеті є **24,278 записів** про оголошення житла в Лісабоні.

Ці дані охоплюють:

- **9,663 унікальних хостів**
- **199 районів**
- **54 різні категорії типів житла**, з яких **4 основні** (Entire home/apt, Private room, Shared room, Hotel room).

Що зроблено для зручності подальшого аналізу:

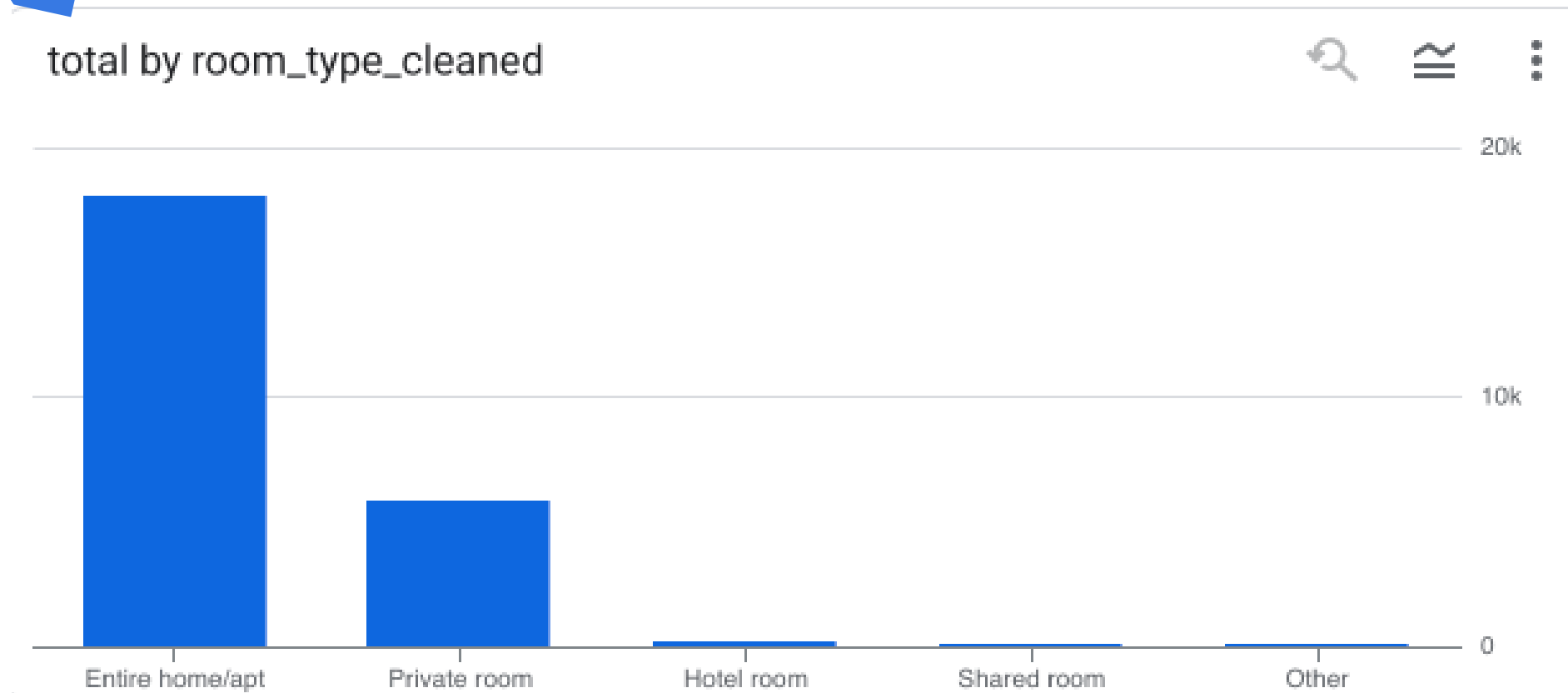
- **Очищено поле room_type**, виділено 4 логічні категорії
- Видалено записи з NULL у критично важливих колонках (price, room_type, availability, тощо)
- reviews_per_month очищено від NULL шляхом заміни на **0**, що дозволяє врахувати всі оголошення, навіть без активності

3. SQL-аналіз

Питання № 1

Які типи житла є найпопулярнішими серед хостів у Лісабоні?

```
1 SELECT
2   room_type_cleaned,
3   COUNT(*) AS total
4 FROM `my-test-project-445920.airbnb_lisbon.listings_cleaned`
5 GROUP BY room_type_cleaned
6 ORDER BY total DESC;
```



Бачимо, що найпопулярніший тип житла — **"Entire home/apt"**.

Далі йдуть **"Private room"**, також доволі відрізняється кількість від інших.

А **"Shared room"** та **"Hotel room"** мають досить маленькі показники

Отже Airbnb може орієнтувати маркетинг на "Entire home" як основну категорію.

Також можна провести додаткові заходи з девелоперами, та заохочувати їх на відкриття "Hotel room" форматів у туристичних районах.

Питання № 2

У яких районах Лісабона найбільше оголошень?

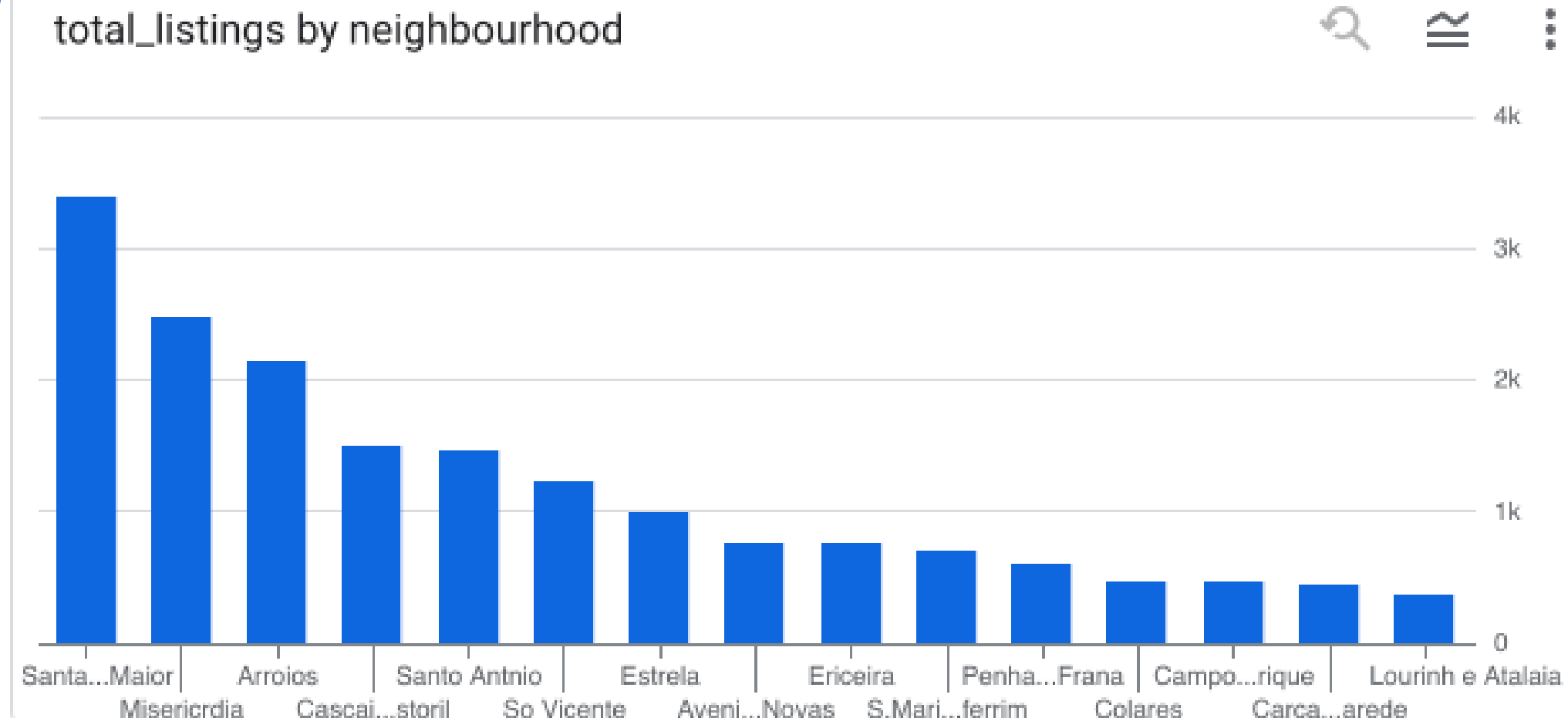
```
1 SELECT
2   neighbourhood,
3   COUNT(*) AS total_listings
4 FROM `my-test-project-445920.airbnb_lisbon.listings_cleaned`
5 GROUP BY neighbourhood
6 ORDER BY total_listings DESC
7 LIMIT 15;
```

Тут бачимо, що найбільше оголошень зосереджено у таких 5-ти районах:

- Santa Maria Maior - 3397 оголошення
- Misericrdia - 2474
- Arroios - 2137
- Cascais e Estoril - 1506
- Santo Antnio - 1465

Варто відмітити, що ці райони мають дійсно найбільшу популярність серед туристів через розсташування всіх історих пам'яток, близькість до транспортних розв'язок та найкрасивіші природні місця в Лісабоні. Райони, які показують нам меншу кількість оголошень не такі зручні за інфраструктурою та не такі цікаві для туристів, які шукають житло на декілька днів.

Отже можна адаптувати стратегію розвитку та підтримувати найпопулярніші райони або додатково розробити маршрут до визначних місць з менш популярних районів, щоб збільшити їх популярність.



Питання № 3

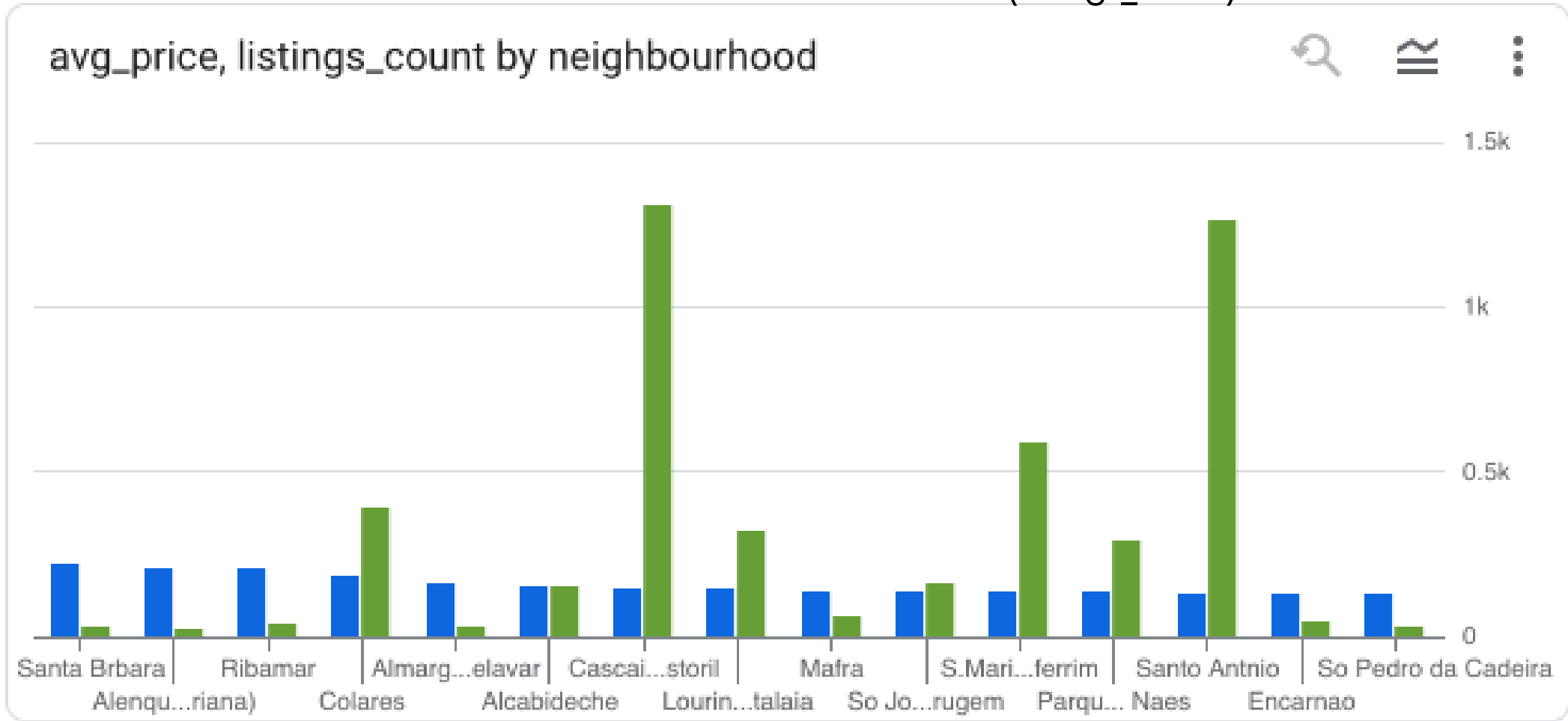
Яка середня ціна оренди в кожному районі Лісабона?

```
1 SELECT
2   neighbourhood,
3   ROUND(AVG(price), 2) AS avg_price,
4   COUNT(*) AS listings_count
5 FROM `my-test-project-445920.airbnb_lisbon.listings_cleaned`
6 WHERE price < 1000 -- прибираю аномально дорогі варіанти
7 GROUP BY neighbourhood
8 HAVING listings_count >= 20 -- беру лише райони з достатньою кількістю оголошень
9 ORDER BY avg_price DESC
10 LIMIT 15;
```

Нотатка:

Синій стовпчик — середня ціна (avg_price)

Зелений стовпчик — кількість оголошень (listings_count)



Ми бачимо, що ціни майже у всіх районах не йдуть поруч з кількістю оголошень:

Наприклад, у районах **Santa Brbara** чи **Alenquer**, середня ціна **висока (€218.96 та €209.56)**, але кількість оголошень — **низька (25–27)**.

→ Це може бути пов'язано з тим, що там загалом маленька кількість забудов, тому цінник вище.

По кількості оголошень лідирують **Cascais e Estoril** і **Santo Antnio** — мають **високу кількість оголошень (1306 та 1266)**, але середню ціну на рівні **€144.64 та €133.26**
→ Це можуть бути через більшу територію, яку охоплює район.

Більш зрозумілу динаміку мають райони **Colares, Alcabideche, Loures і Atalaia** — показують **помірну ціну (140–170)** і **достатню кількість пропозицій (150–400)**.
→ Саме ці райони можуть бути **цікаві для оптимізації бізнесу**, оскільки вони більше мають потенціал.

Питання № 4

Яка середня ціна на житло залежно від типу кімнати?

```
1 SELECT
2   room_type_cleaned,
3   ROUND(AVG(price), 2) AS avg_price,
4   COUNT(*) AS listings_count
5 FROM
6   `my-test-project-445920.airbnb_lisbon.listings_cleaned_final`
7 GROUP BY
8   room_type_cleaned
9 ORDER BY
10  avg_price DESC;
```



Query results

Job information					Results	Chart	JSON	Execution details
Row	room_type_cleaned		avg_price	listings_count				
1	Hotel room		284.92	144				
2	Entire home/apt		183.26	16224				
3	Private room		60.98	4556				
4	Shared room		32.03	90				
5	Other		9.92	63				

Бачимо, що:

Hotel room має найвищу середню ціну — **~285 €**, але представлений лише в **144 оголошеннях**.



Найпоширеніший тип — **Entire home/apt** з **~16,200 оголошеннями** і середньою ціною **183 €**, що робить його ключовим сегментом ринку.

Private room — помітно дешевший варіант (**61 €**), але займає **друге місце за популярністю (~4,500 оголошень)**.

На основі цього вже можна будувати стратегії по маркетингу, або розвитку самого додатку на території Лісабону.

Питання № 5

Наскільки довго гості зазвичай залишаються в орендованому житлі?

```
1 SELECT
2 CASE
3   WHEN minimum_nights BETWEEN 1 AND 3 THEN '1-3 nights'
4   WHEN minimum_nights BETWEEN 4 AND 7 THEN '4-7 nights'
5   WHEN minimum_nights BETWEEN 8 AND 14 THEN '8-14 nights'
6   WHEN minimum_nights BETWEEN 15 AND 30 THEN '15-30 nights'
7   ELSE '30+ nights'
8 END AS stay_length_group,
9 COUNT(*) AS listings_count
10 FROM `my-test-project-445920.airbnb_lisbon.listings_cleaned`
11 GROUP BY stay_length_group
12 ORDER BY listings_count DESC;
```

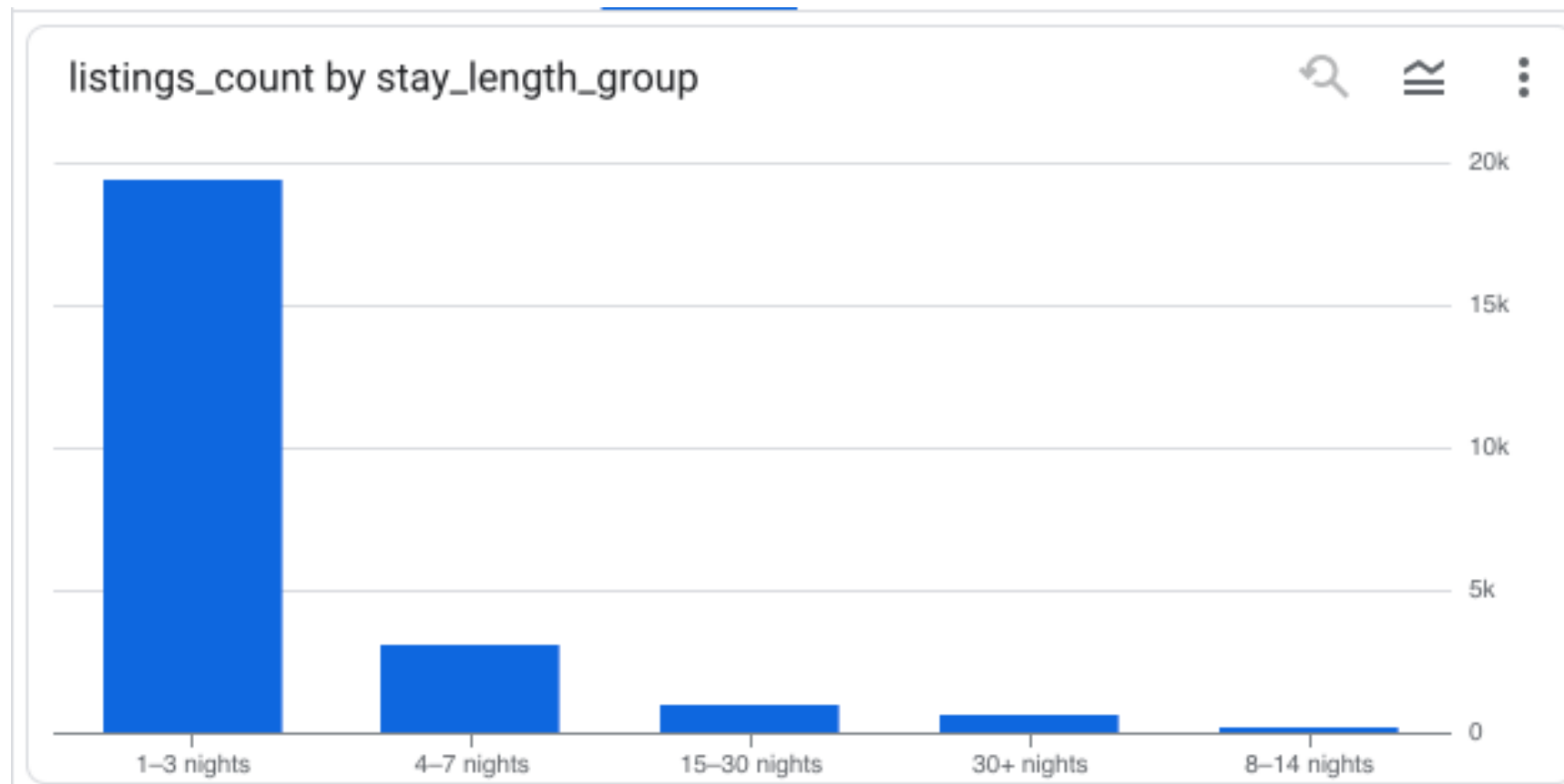
Ми бачимо, що

Короткі терміни **(1–3 дні)** - **найбільш популярні**.
Значну частку займають також середні терміни **(4–7днів)**.

Всі інші займають незначну частку.

Мені здається продукт міг би покращити кількість довготривалих термінів та створити можливо окрему категорію та розвивати її також.

Також ця інформація може бути корисна для маркетинг відділу, щоб створити певну сегментацію по пропозиціям та можливо ввести новий, зручний для користувачів фільтр.



Питання № 6

Яка кількість оголошень припадає на одного хоста?

```
1 SELECT
2   host_id,
3   COUNT(*) AS listings_per_host
4 FROM `my-test-project-445920.airbnb_lisbon.listings_cleaned`
5 GROUP BY host_id
6 ORDER BY listings_per_host DESC;
```

Запустила такий запит і побачила, що у мене 50 результатів, що важко для аналітики. Тому вирішила трішки змінити запит та згрупувати власників за кількість оголошень.



```
1 SELECT
2   host_group,
3   COUNT(*) AS host_count
4 FROM (
5   SELECT
6     host_id,
7     MAX(calculated_host_listings_count) AS listings_count,
8     CASE
9       WHEN MAX(calculated_host_listings_count) = 1 THEN '1'
10      WHEN MAX(calculated_host_listings_count) BETWEEN 2 AND 3 THEN '2-3'
11      WHEN MAX(calculated_host_listings_count) BETWEEN 4 AND 5 THEN '4-5'
12      WHEN MAX(calculated_host_listings_count) BETWEEN 6 AND 10 THEN '6-10'
13      ELSE '11+'
14    END AS host_group
15 FROM `my-test-project-445920.airbnb_lisbon.listings_cleaned`
16 GROUP BY host_id
17 )
18 GROUP BY host_group
19 ORDER BY host_count DESC;
```

Query results

Job information		Results	Chart	JSON
Row	host_group	host_count		
1	1	6418		
2	2-3	1849		
3	4-5	564		
4	6-10	475		
5	11+	358		

Бачимо, що **найбільшу кількість займають хости тільки з одними апартаментами.**

Значить платформою користуються більше приватні особи. 6 – 11+ оголошень у дійсно малої кількості хостів. Це можуть бути бізнеси або професійні компанії.

Для продукту це можна використати наступним чином:

- **Професійним хостам** можна запропонувати спеціальні програми лояльності або підтримки.
- **Індивідуальним хостам** — полегшення старту (гайди, шаблони, підтримка).

Питання № 7

Яка середня ціна залежно від кількості оголошень у хоста?

```
SELECT
  host_group,
  ROUND(AVG(price), 2) AS avg_price,
  COUNT(DISTINCT host_id) AS host_count
FROM (
  SELECT
    host_id,
    price,
    CASE
      WHEN calculated_host_listings_count = 1 THEN '1'
      WHEN calculated_host_listings_count BETWEEN 2 AND 3 THEN '2-3'
      WHEN calculated_host_listings_count BETWEEN 4 AND 5 THEN '4-5'
      WHEN calculated_host_listings_count BETWEEN 6 AND 10 THEN '6-10'
      ELSE '11+'
    END AS host_group
  FROM `my-test-project-445920.airbnb_lisbon.listings_cleaned`
)
GROUP BY host_group
ORDER BY avg_price DESC;
```

Query results

Job information	Results	Chart	JSON	Execution detail
Row	host_group	avg_price	host_count	
1	11+	181.59	357	
2	6-10	175.48	475	
3	1	142.18	6418	
4	2-3	138.0	1849	
5	4-5	118.2	564	



Бачимо, що **хости з більшою кількістю апартаментів встановлюють вищі ціни.**
Власники одних апартаментів - мають переважно середню вартість.

Тут бізнес може вирішити для себе, які кроки та групи хостів можна підтримувати.
Можна **допомагати та навчати хостів з невеликою кількістю апартаментів**, щоб вони могли розвивати свої послуги та підвищили якість і середню ціну пропозицій.
Або заохочувати та підтримувати власників великої кількості апартаментів.

Питання № 8

Які типи житла мають більшу кількість відгуків?

```
1 SELECT
2   room_type_cleaned,
3   ROUND(AVG(reviews_per_month), 2) AS avg_reviews_per_month
4 FROM `my-test-project-445928.airbnb_lisbon.listings_cleaned`
5 WHERE reviews_per_month IS NOT NULL
6 GROUP BY room_type_cleaned
7 ORDER BY avg_reviews_per_month DESC;
```



Query results

Job information				Results	Chart	JSON
Row	room_type_cleaned			avg_reviews_per_mo		
1	Other			3.19		
2	Entire home/apt			1.56		
3	Private room			1.24		
4	Shared room			0.94		
5	Hotel room			0.43		



Бачимо, що **найактивніше залишають відгуки про Entire home/apt**. Але вони і були найбільш популярними на основі 1-го нашого питання.

А от **Hotel room** —мають найменшу кількість відгуків. Думаю це пов'язано з тим, що готелі більше бронюють на booking.com.

Тут варто визначити чи хоче бізнес залучати аудиторію для готелів більше та надавати якісь бонуси за залишений відгук чи все ж не цільова дія.

Питання № 9

Чи відрізняється доступність житла (availability_365) для різних типів оголошень?

```
1 SELECT
2   room_type_cleaned,
3   ROUND(AVG(availability_365), 2) AS avg_availability
4 FROM `my-test-project-445920.airbnb_lisbon.listings_`
5 GROUP BY room_type_cleaned
6 ORDER BY avg_availability DESC;
```



Query results

Job information		Results	Chart	JSON
Row	room_type_cleaned	avg_availability		
1	Shared room	197.87		
2	Entire home/apt	192.67		
3	Hotel room	183.72		
4	Private room	169.93		
5	Other	12.25		



На основі цього аналізу, бачимо, що **Shared room** і **Entire home/apt** мають найвищу доступність — ~198 та ~193 днів.

Hotel room та **Private room** — дещо менше (~184 і ~170 днів).

Категорія **Other** майже не використовується — лише ~12 днів на рік, що підтверджує її нерелевантність.

Можемо зробити висновок, що **найпопулярніші типи житла активно доступні**. Тому платформа може планувати свої маркетингові дії на тривалий період часу.

Питання №10

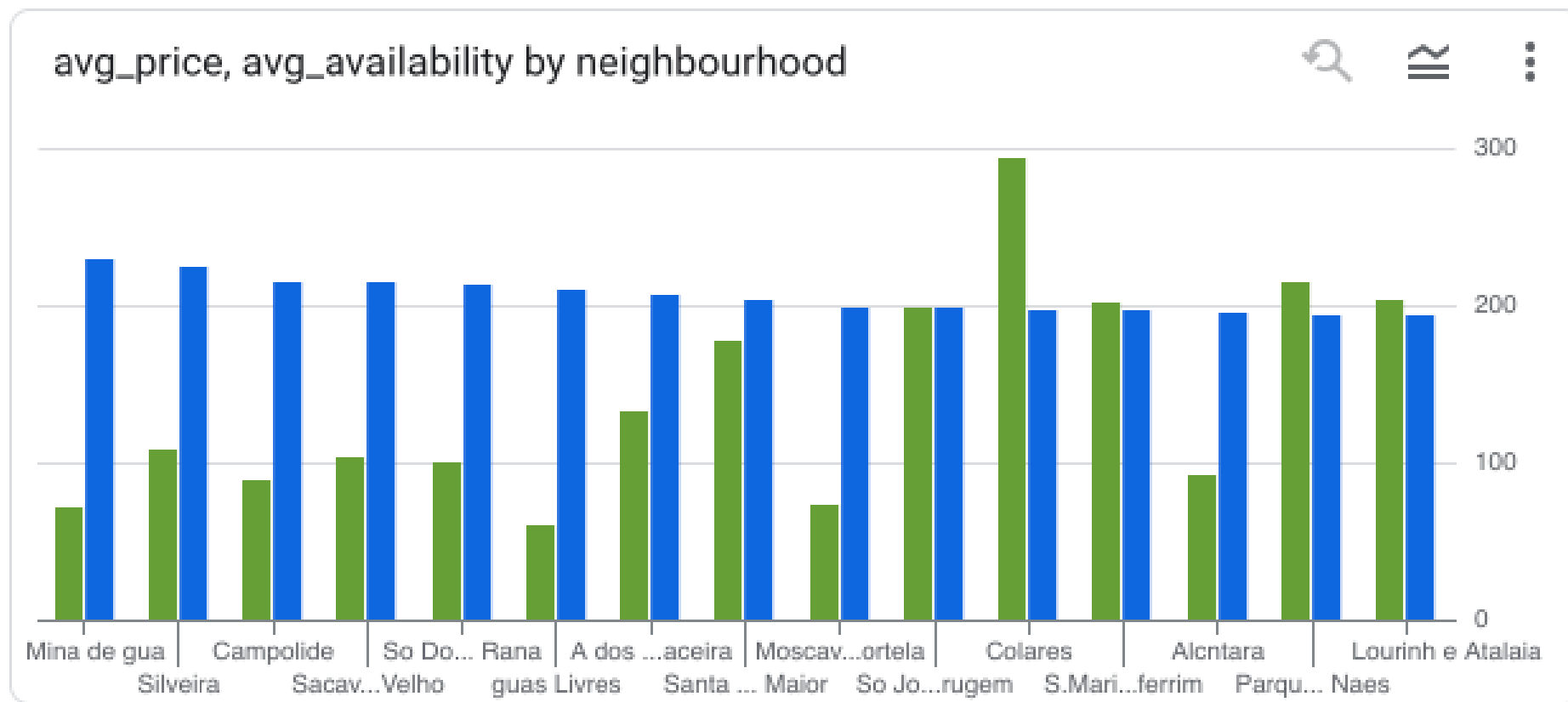
Які райони мають найвищу доступність житла, але при цьому низьку середню ціну?

```
1 SELECT
2   neighbourhood,
3   ROUND(AVG(price), 2) AS avg_price,
4   ROUND(AVG(availability_365), 2) AS avg_availability
5 FROM `my-test-project-445920.airbnb_lisbon.listings_cleaned`
6 GROUP BY neighbourhood
7 HAVING COUNT(*) > 50 -- виключаю малопопулярні райони
8 ORDER BY avg_availability DESC, avg_price ASC
9 LIMIT 15;
```

Нотатка:

Зелений колір - avg_price

Синій колір - avg_availability



Райони як *Colares*, *Parque das Nações* та *Alcântara* демонструють **високу середню ціну оренди при низькій доступності житла**. Це може свідчити про обмежену кількість доступних днів для бронювання або про сезонний характер пропозицій у цих районах.

Для бізнесу це — **сигнал про потенціал для залучення нових хостів у ці райони**, оскільки попит може перевищувати пропозицію.

І навпаки район **Mina d'Água** - Висока доступність житла, але середня ціна одна з найнижчих.

Águas Livres - Знову бачимо високу доступність при помірній ціні.

So Domingos de Rana - Баланс між високою доступністю та низькою вартістю.

Airbnb може просувати райони з низькою вартістю як «доступну альтернативу». Або стати самотійно інвестом у нерухомість в цих районах.



4. Дашборд: Airbnb Lisbon Dashboard: Prices, Availability, Listings

З повним інтерактивним дашбордом ви можете ознайомитись за посиланням:

<https://public.tableau.com/>

5. Висновки



1. **Цілі квартири (Entire home/apt)** — найпопулярніший тип житла на платформі, займає основну частку ринку.
2. **Ціни значно коливаються залежно від району:** преміальні локації (Colares, Alcântara) мають найвищу вартість, тоді як інші райони виглядають недооціненими.
3. **Хости з великою кількістю оголошень** встановлюють вищі ціни та, ймовірно, працюють на рівні малого бізнесу.
4. **Доступність житла в середньому висока (170–200 днів на рік)** — це вказує на орієнтацію хостів на стабільний дохід. Або що туристичний сезон в Португалії значний по часовим рамкам.
5. **Рівень активності (відгуки)** — вищий у приватних кімнатах і цілих квартирах, що може вказувати на більший попит.



6. Рекомендації

1. Сегментувати хостів за кількістю оголошень:

- Індивідуальні хости — надати їм навчальні кампанії, запропонувати підтримку в старті.
- Професійні хости — розробити програми лояльності, та сформулювати рекомендації з ціноутворення.

2. Організувати новий напрями маркетинг стратегії по регіонах:

- Підсвітити *недооцінені райони* з високою доступністю та низькою ціною.
- Якщо є проблема того, що люди переглядають сайт але так і не бронюють апартаменти то ввести перенаправлення частини попиту із перенасичених/дорогих районів в більш доступні.

3. Можна вводити рекомендації для користувачів більш прицільні:

- Надавати більше інформації де підсвічувати доступність і середню активність житла.

4. Очистити категорії житла (**room_type**) — є кількість даних, які ніяк не допомагають проводити аналіз. Їх я згрупувати в категорію "Other". Їх потрібно якось усунути або переформатувати.