What 5,500 Hotels Teach Us About Value for Money

A data-driven analysis of Booking.com & TripAdvisor

3,290 Booking.com hotels | 2,248 TripAdvisor hotels | Real data

Giorgio Vernarecci - Data Analyst

The Question

Does paying more for a hotel
actually get you a better experience?

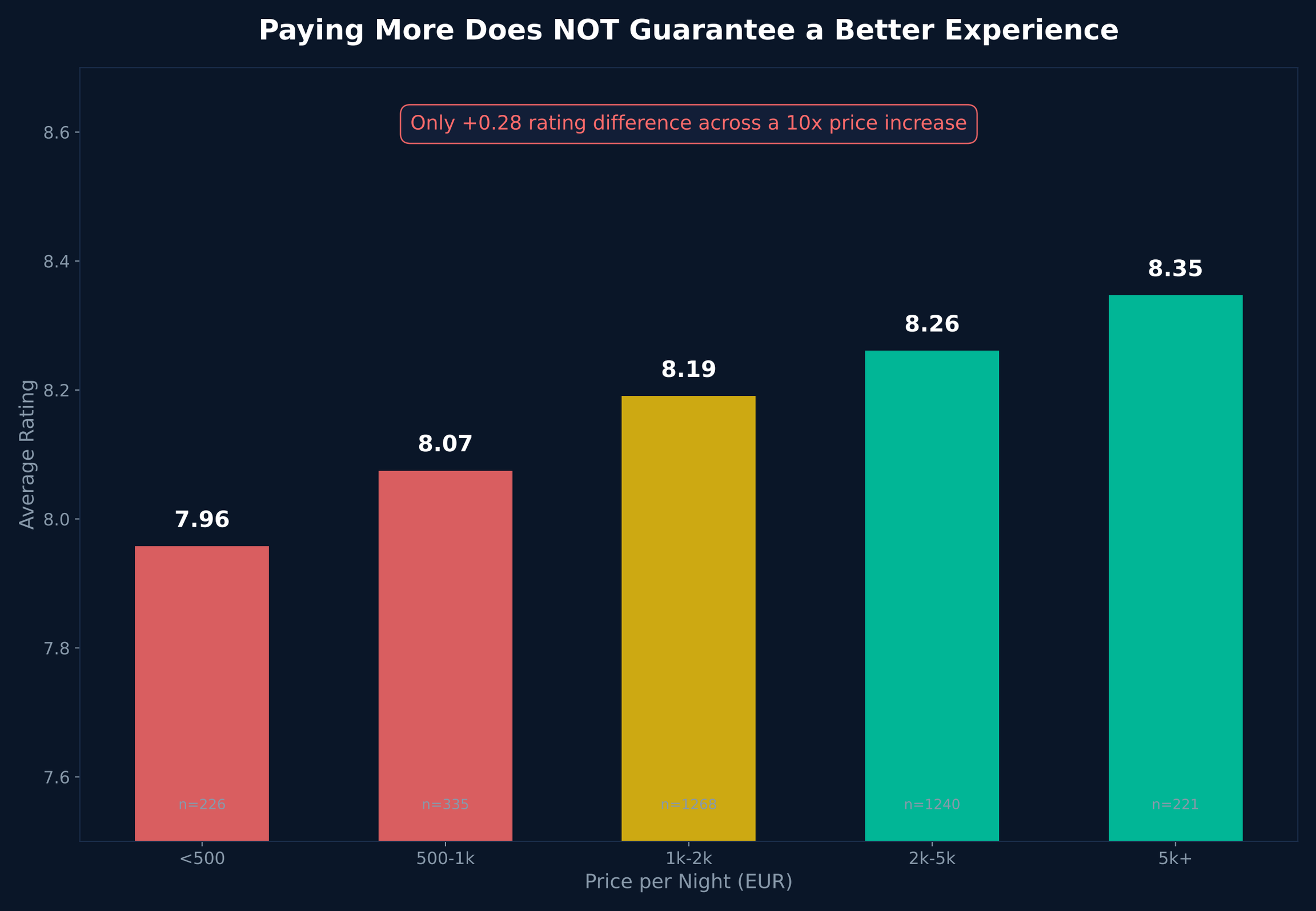
I analyzed 5,500+ hotels across Booking.com and TripAdvisor
to find out what really drives guest satisfaction
and where travelers get the best (and worst) value.

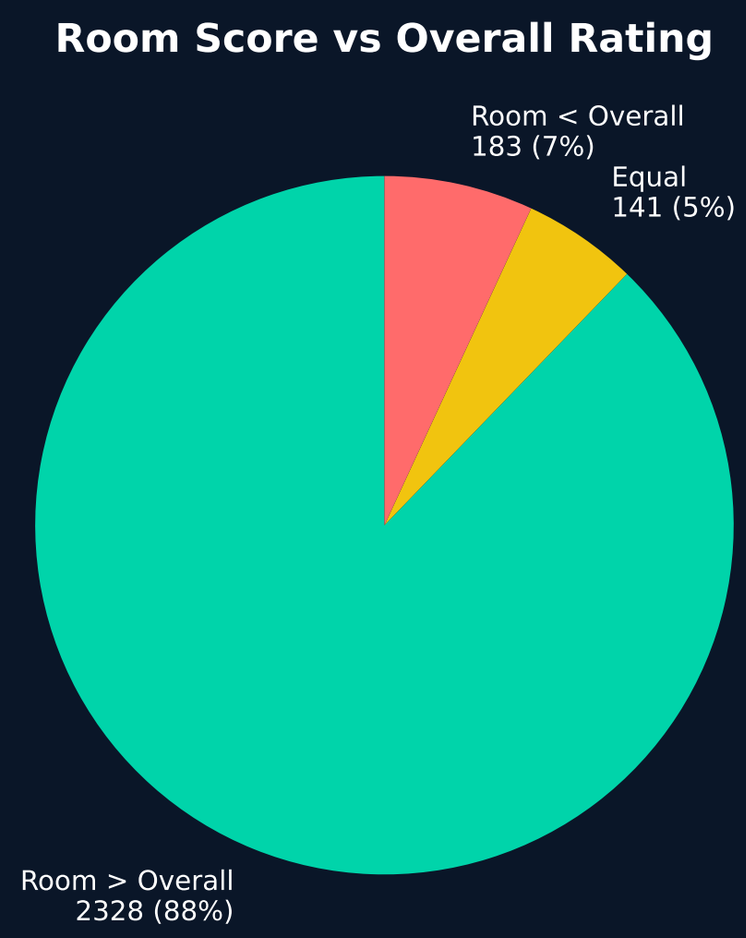
The Short Answer

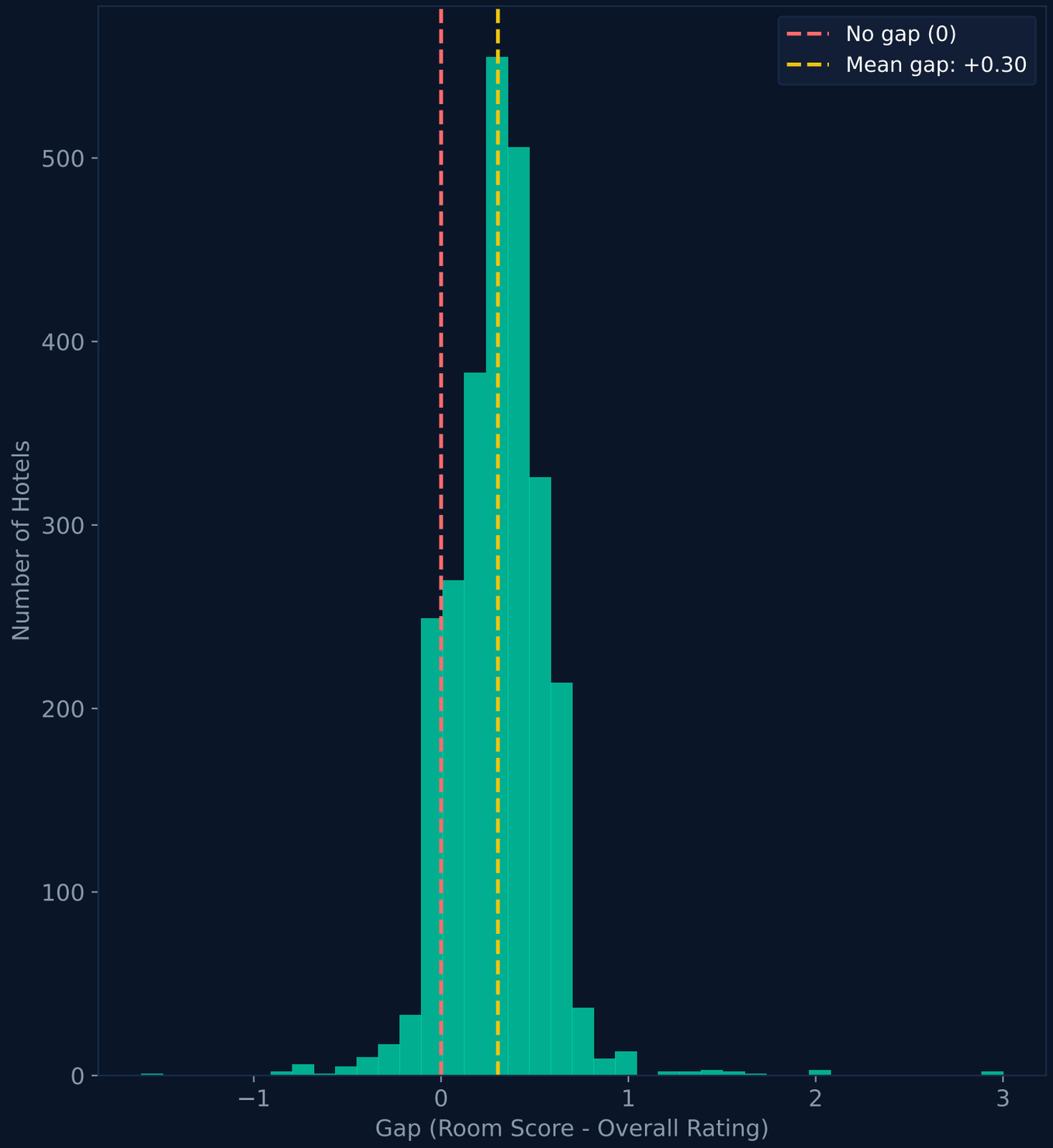
+0.28

That's the rating difference between
a EUR 500/night hotel and a EUR 5,000/night hotel

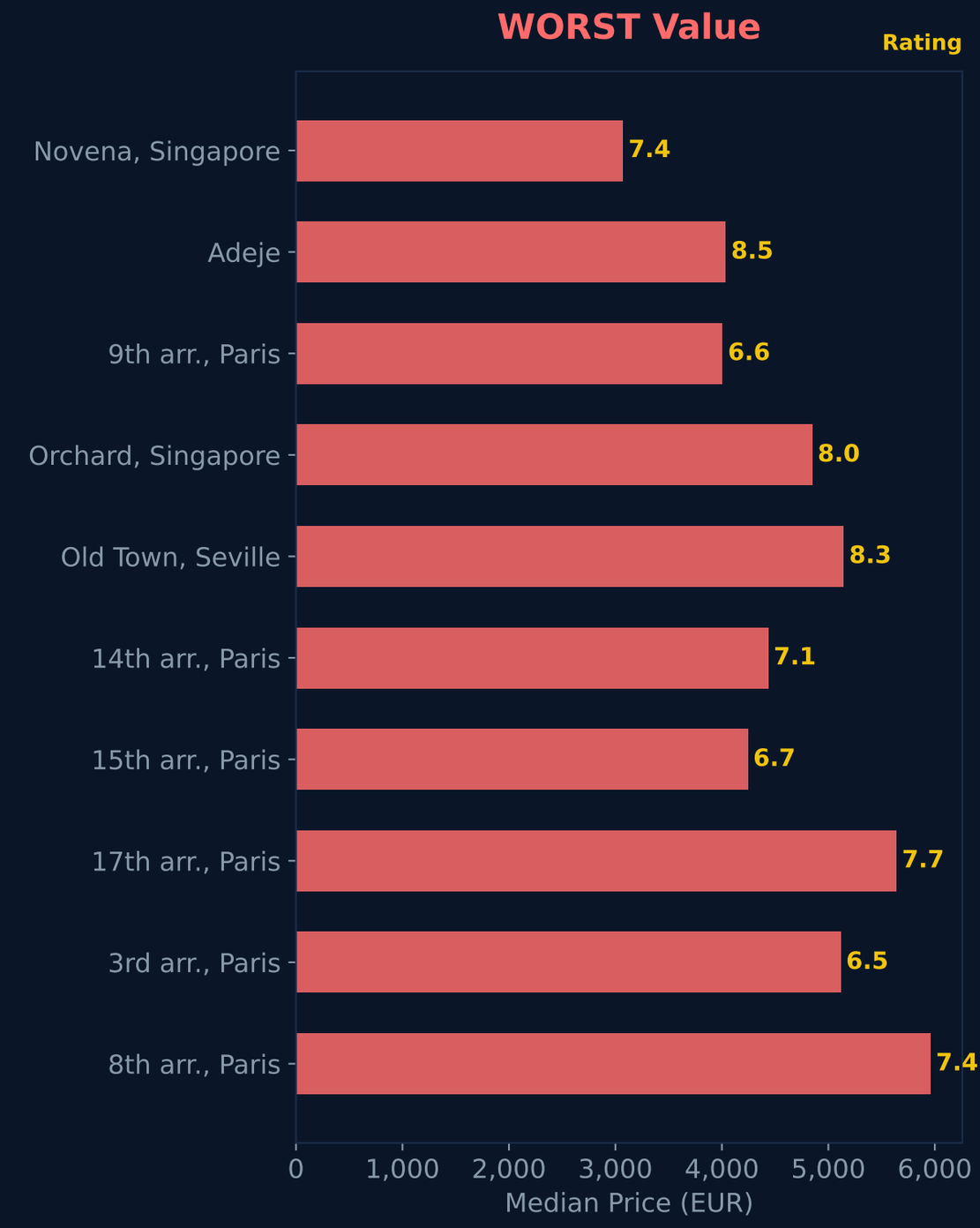
Spearman correlation: r = 0.19 (weak positive)

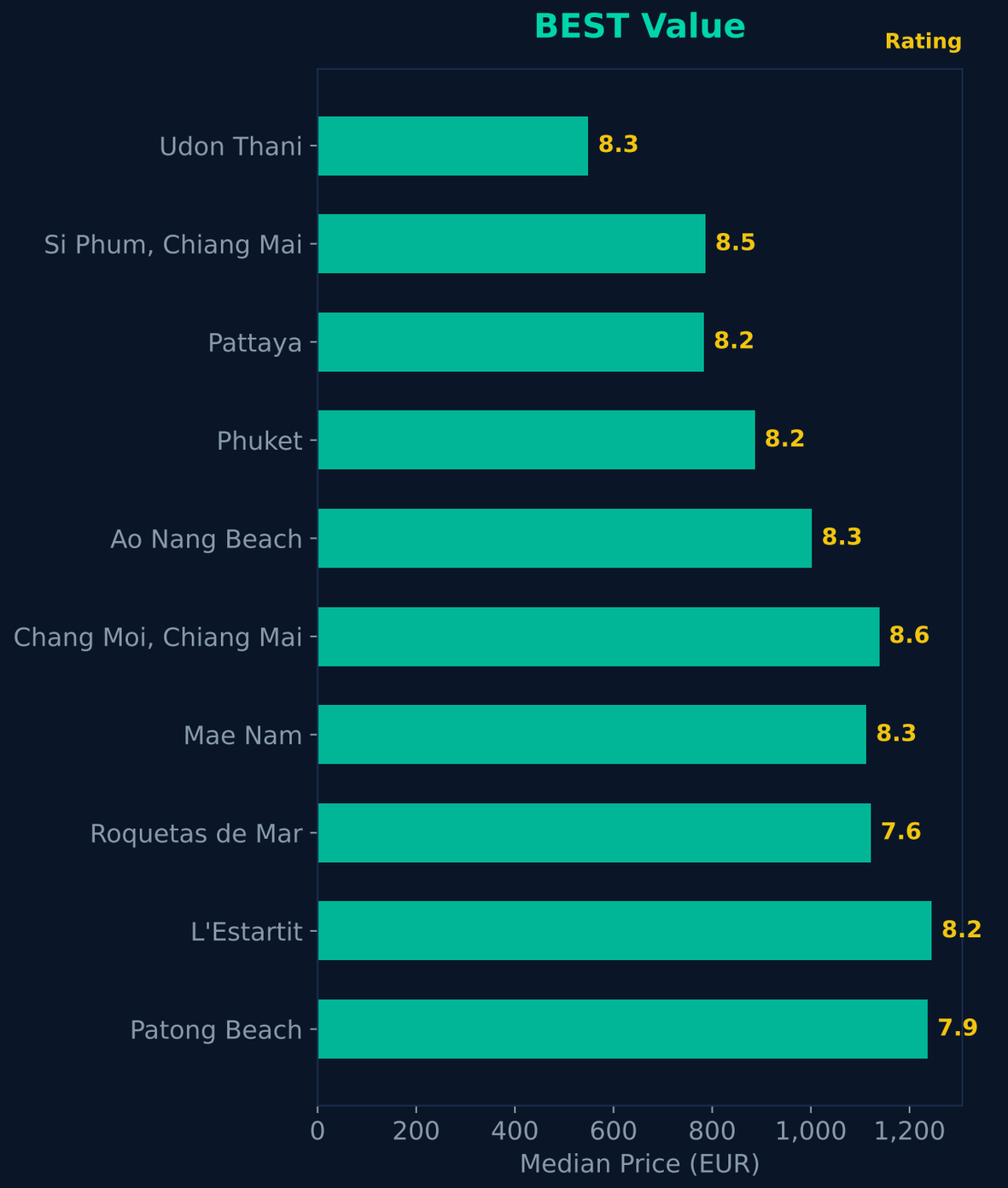

Paying More Does NOT Guarantee a Better Experience
Only +0.28 rating difference across a 10x price increase
Average Rating
8.6
8.4
8.2
8.0
7.8
7.6
7.96
8.07
8.19
8.26
8.35
n=226
n=335
n=1268
n=1240
n=221
<500
500-1k
1k-2k
2k-5k
5k+
Price per Night (EUR)

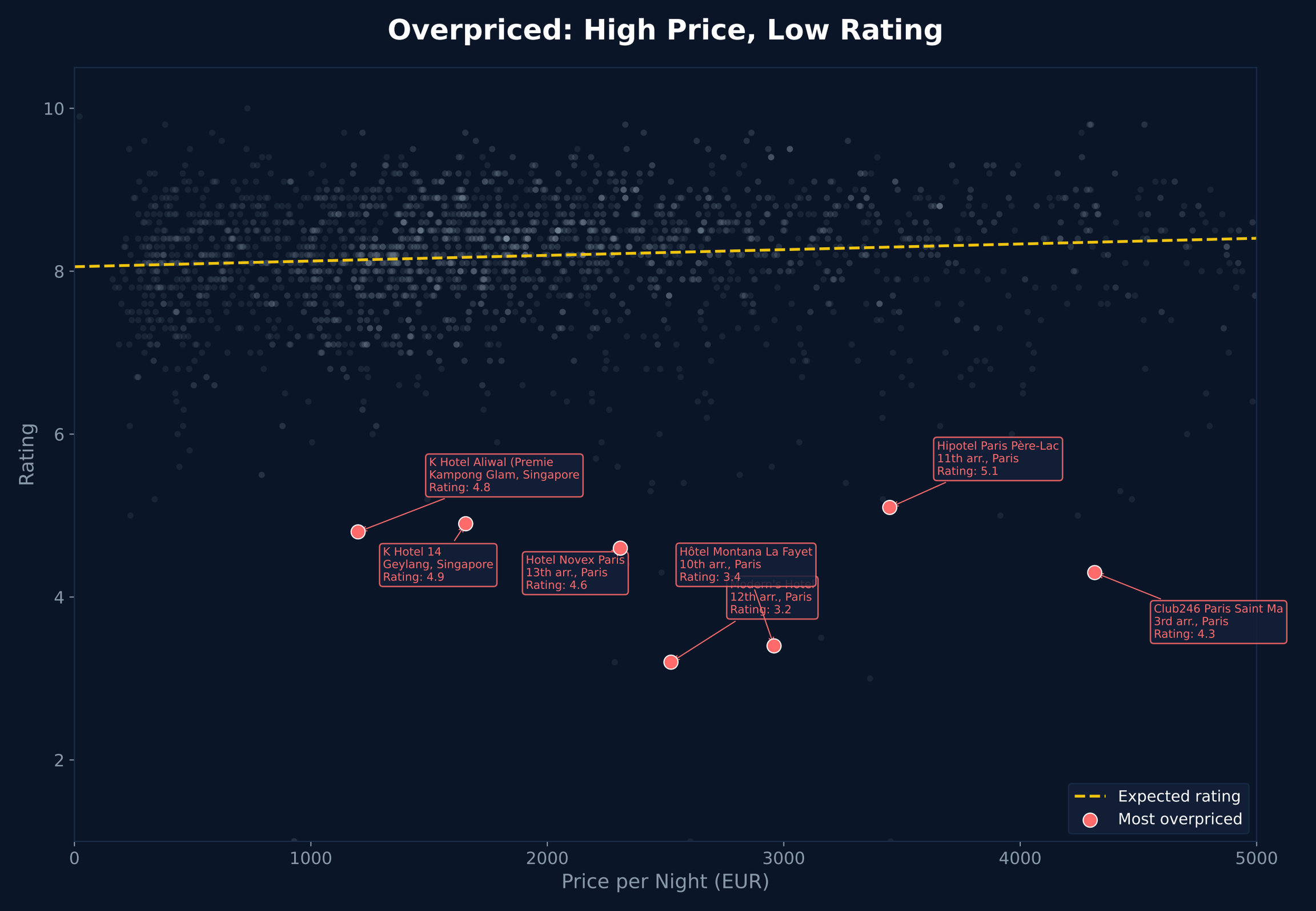

Room Score vs Overall Rating
Room < Overall
183 (7%)
Equal
141 (5%)
Room > Overall
2328 (88%)


88% of Hotels: Room Is the Strongest Point
No gap (0)
Mean gap: +0.30
Number of Hotels
0
100
200
300
400
500
−1
0
1
2
3
Gap (Room Score - Overall Rating)

Paris vs Thailand: Where Does Your Money Go?






Overpriced: High Price, Low Rating
Rating
Price per Night (EUR)
10
8
6
4
2
0
1000
2000
3000
4000
5000
K Hotel Aliwal (Premie
Kampong Glam, Singapore
Rating: 4.8
K Hotel 14
Geylang, Singapore
Rating: 4.9
Hotel Novex Paris
13th arr., Paris
Rating: 4.6
Hôtel Montana La Fayet
10th arr., Paris
Rating: 3.4
12th arr., Paris
Rating: 3.2
Hipotel Paris Père-Lac
11th arr., Paris
Rating: 5.1
Club246 Paris Saint Ma
3rd arr., Paris
Rating: 4.3
Expected rating
Most overpriced


Cheaper Hotels Get Longer Reviews
Budget guests write 2x more words than luxury guests | Spearman r = -0.43
Average Words per Review
0
5
10
15
20
25
30
35
27 words
21 words
18 words
16 words
12 words
n=426
n=531
n=345
n=349
n=370
<30
30-60
60-100
100-200
200+
Price per Night (EUR)

Key Takeaways

1. **Price is a poor predictor of quality**
   +0.28 rating difference across 10x price range

2. **Rooms are the strongest asset**
   88% of hotels score higher on rooms than overall

3. **Paris is the worst value destination**
   6 of the 10 worst value locations are in Paris

4. **Overpriced hotels are identifiable**
   Residual analysis reveals consistent underperformers

5. **Budget guests leave richer feedback**
   2x longer reviews at cheap hotels (r = -0.43)

Methodology

Data Source	Kaggle - Hotel Dataset: Rates, Reviews & Amenities (6k+) by joyshil0599 - CC0 License
Datasets	Booking.com (3,465 rows) + TripAdvisor (5,330 rows)
Cleaning	Encoding fixes, currency conversion (BDT to EUR at 1:120), outlier removal, string normalization, duplicate handling
After Cleaning	Booking: 3,290 hotels | TripAdvisor: 2,248 hotels
Analysis	Pearson & Spearman correlations, linear regression for residual analysis, descriptive statistics
Tools	Python (pandas, numpy, scipy, matplotlib)

Note: BDT/EUR conversion rate is approximate. The two datasets cover different market segments and are analyzed separately where appropriate.

Giorgio Vernarecci

Data Analyst

SQL | Python | R | Tableau | n8n

Former hospitality professional turned data analyst.
I combine operational experience with analytics
to find insights others miss.

Let's connect - follow me for more data stories.

Sources & Links

Dataset: kaggle.com/datasets/joyshil0599/hotel-dataset-rates-reviews-and-amenities5k

GitHub: github.com/logiop

LinkedIn: linkedin.com/in/giorgio-vernarecci-4b5a8a23b

Thank you for reading.