

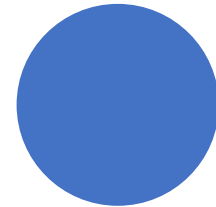


Analysis of The Most Expensive and Least Expensive London-based Hotels' Review Data

Hotel reviews are gold mine of customer insights for any hotel business. Also, its importance increases by many folds since majority of the future customers rely on the peer reviews while finalizing their stay.

In this study, we will analyze the reviews for the top 10 most expensive and least expensive hotels based out of London to compare various aspects of the rating and review text.

The review data was extracted from a leading travel site via PromptCloud's in-house [custom crawling solution](#).



The final number of reviews that we extracted is more than 27,000 with the following data fields:



Property
name



Rating



Review title



Review text



Location of
reviewer



Date of
review

You can download this for free by signing up for [DataStock](#) – one of our solutions to download pre-crawled web data sets instantly.



Goals



Comparison of overall rating



Comparison of the review count



Time series analysis of review count



Comparison of review length



Analysis of frequently used words



Trending words present in review text



Sentiment analysis to find out positive and negative words



Sentiment analysis for overall scores for polarity and emotions

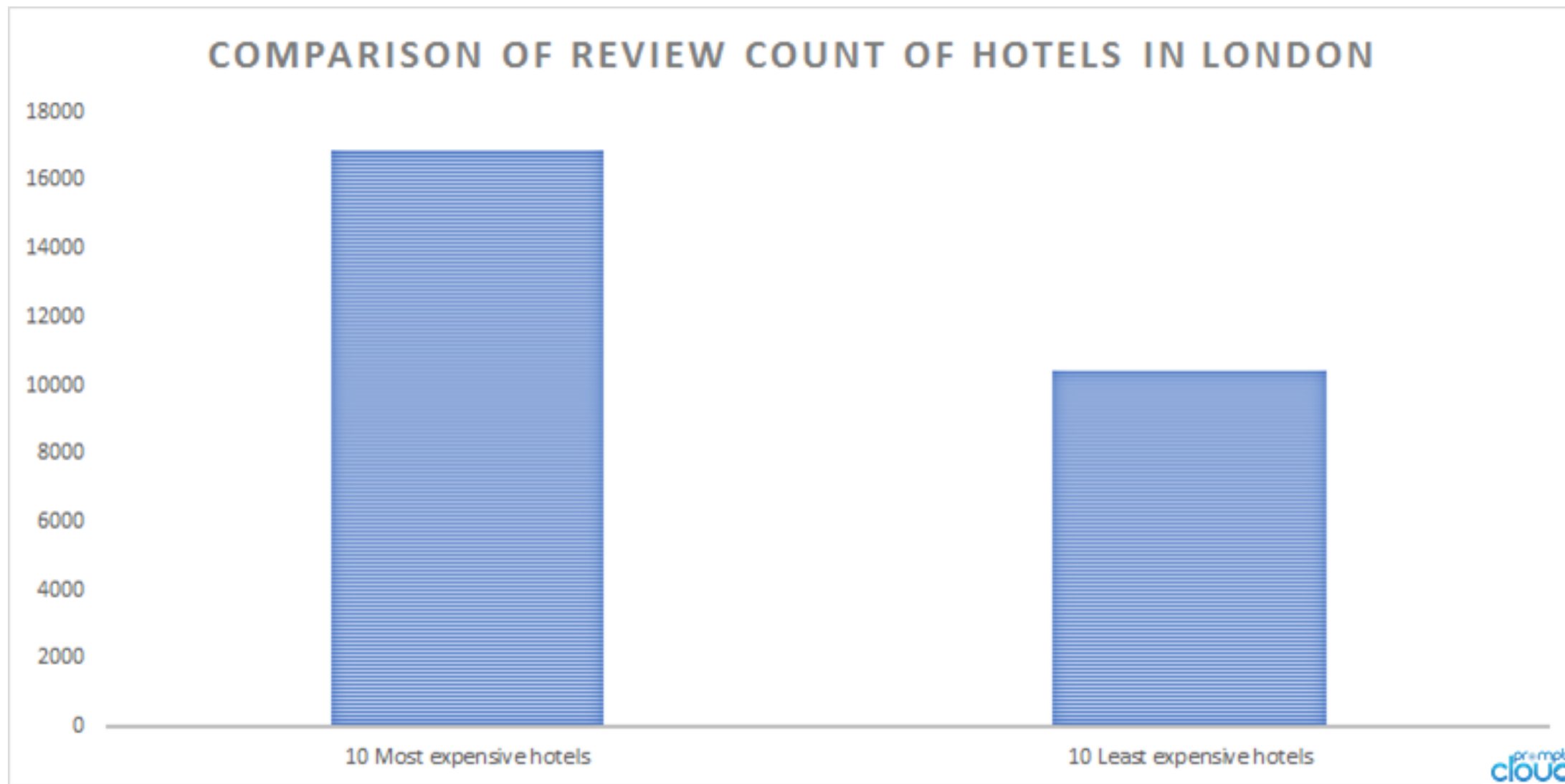
Comparison of overall ratings



The chart shows that the most expensive hotels have average rating of 4.62, which is negligibly higher than the rating of least expensive ones (4.28).



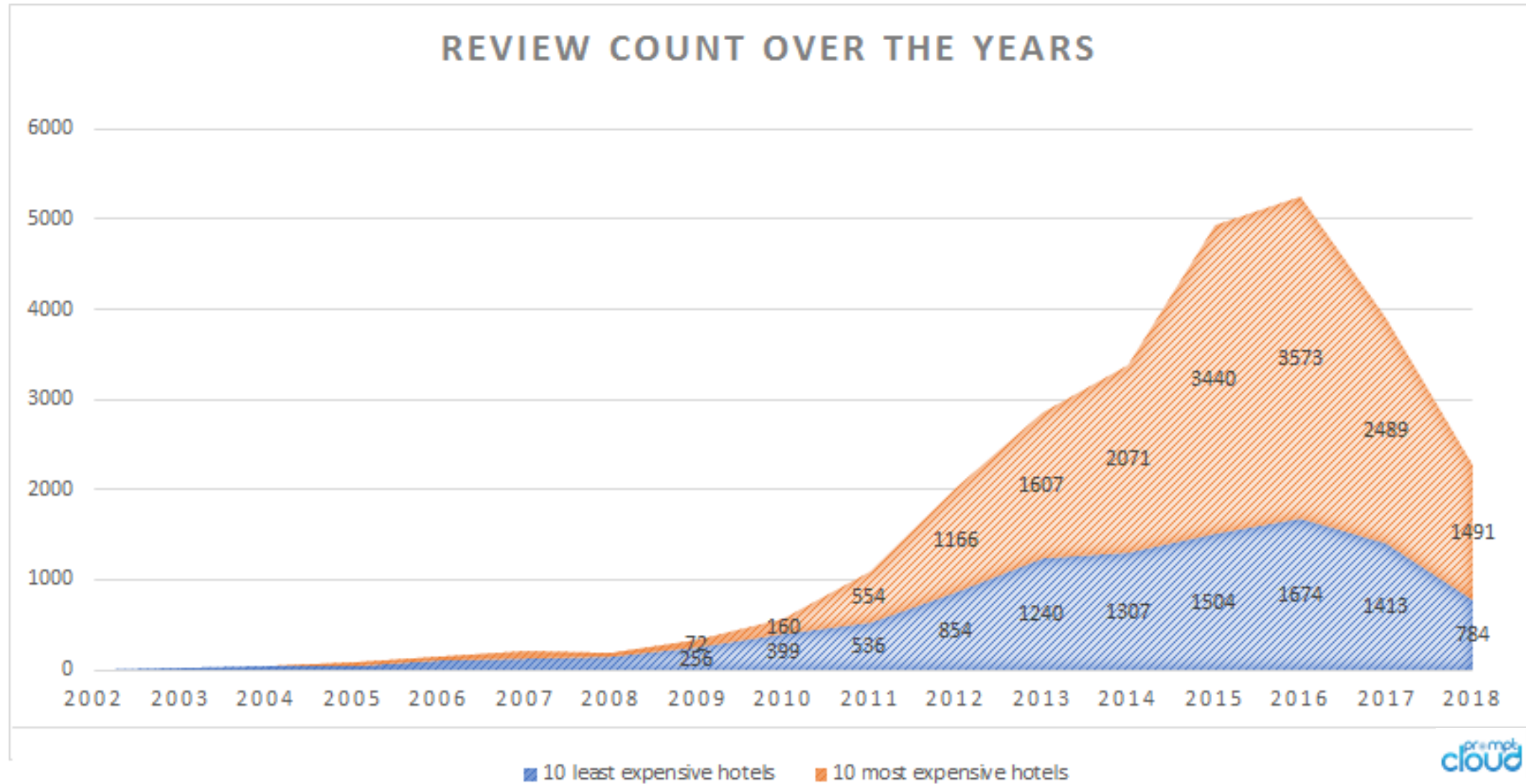
Comparison of the review count



This shows that the most expensive segment has received more number of reviews than the least expensive segment.



Number of reviews over the years



the number of reviews for the least expensive segment was higher than expensive ones from year 2002 to 2010 (except 2005). However, since 2011 the review counts have reversed.

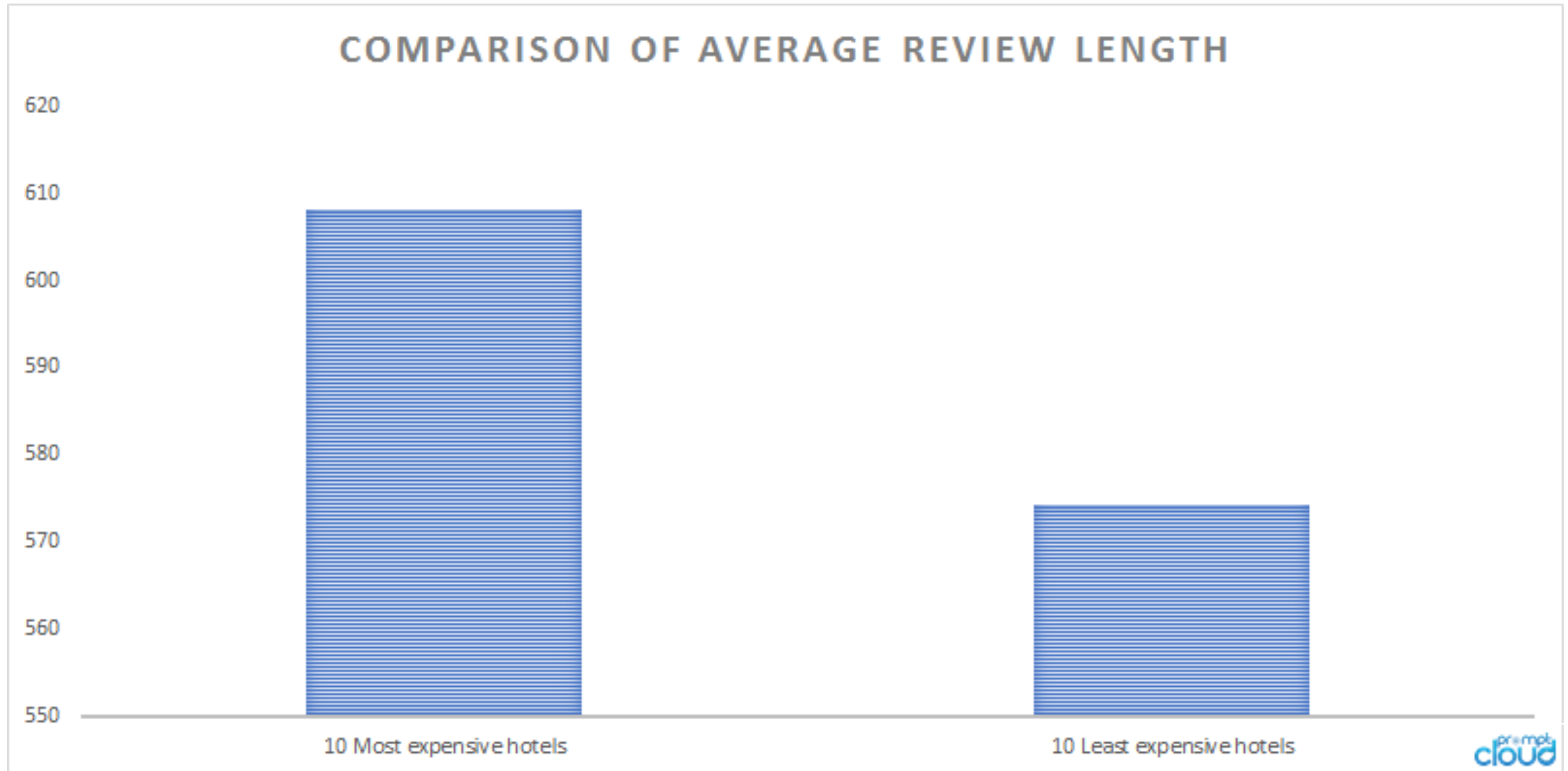


The beginning of reversal can be attributed to the following:

Renovation of Corinthia Hotel in 2011 which started receiving reviews from 2011.

The Savoy's review count increased significantly from 2011.

Comparison of review length

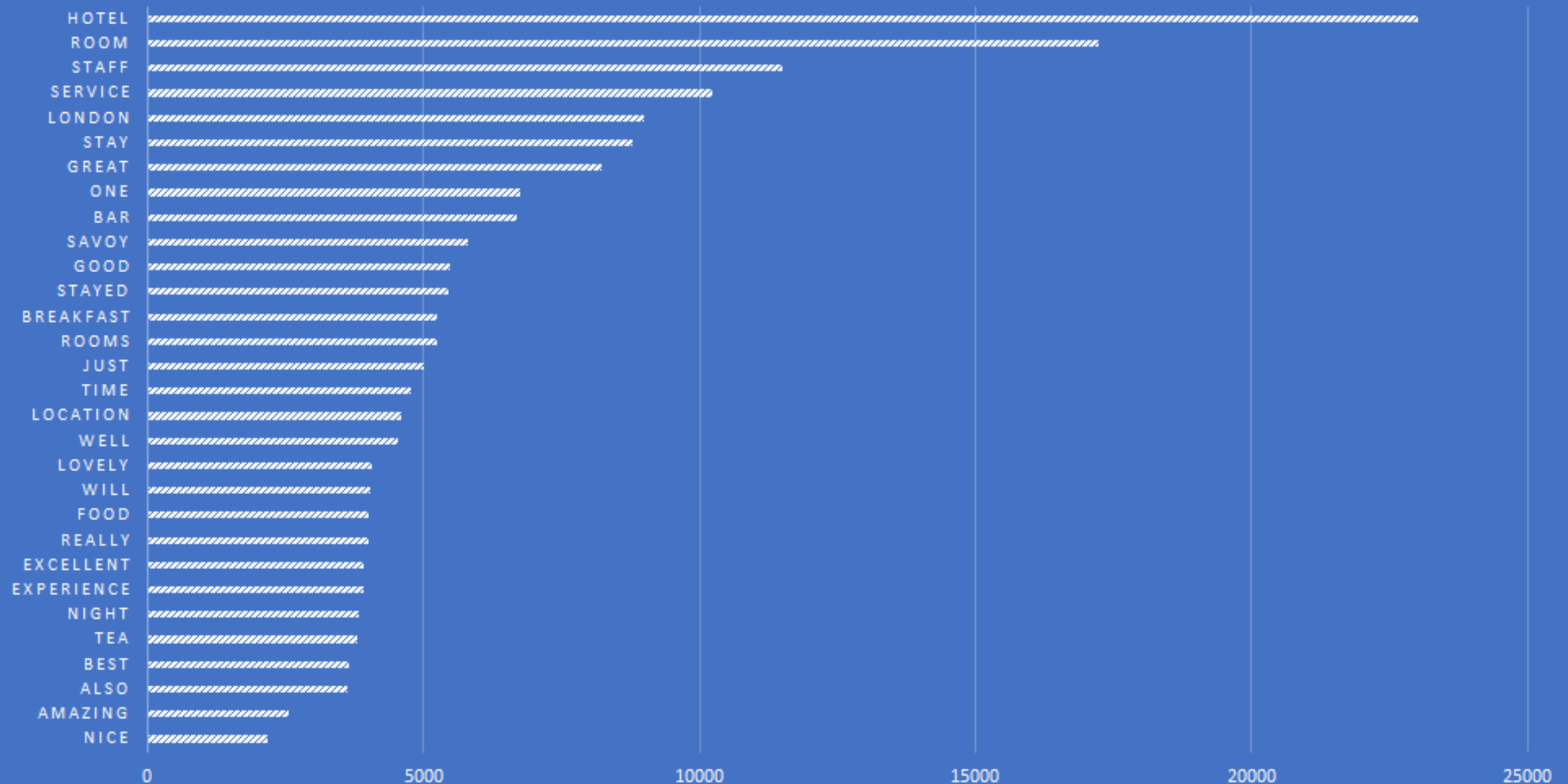


Let's now find out the average review length (excluding whitespace) of the two types of hotels to see which one receives lengthier reviews.

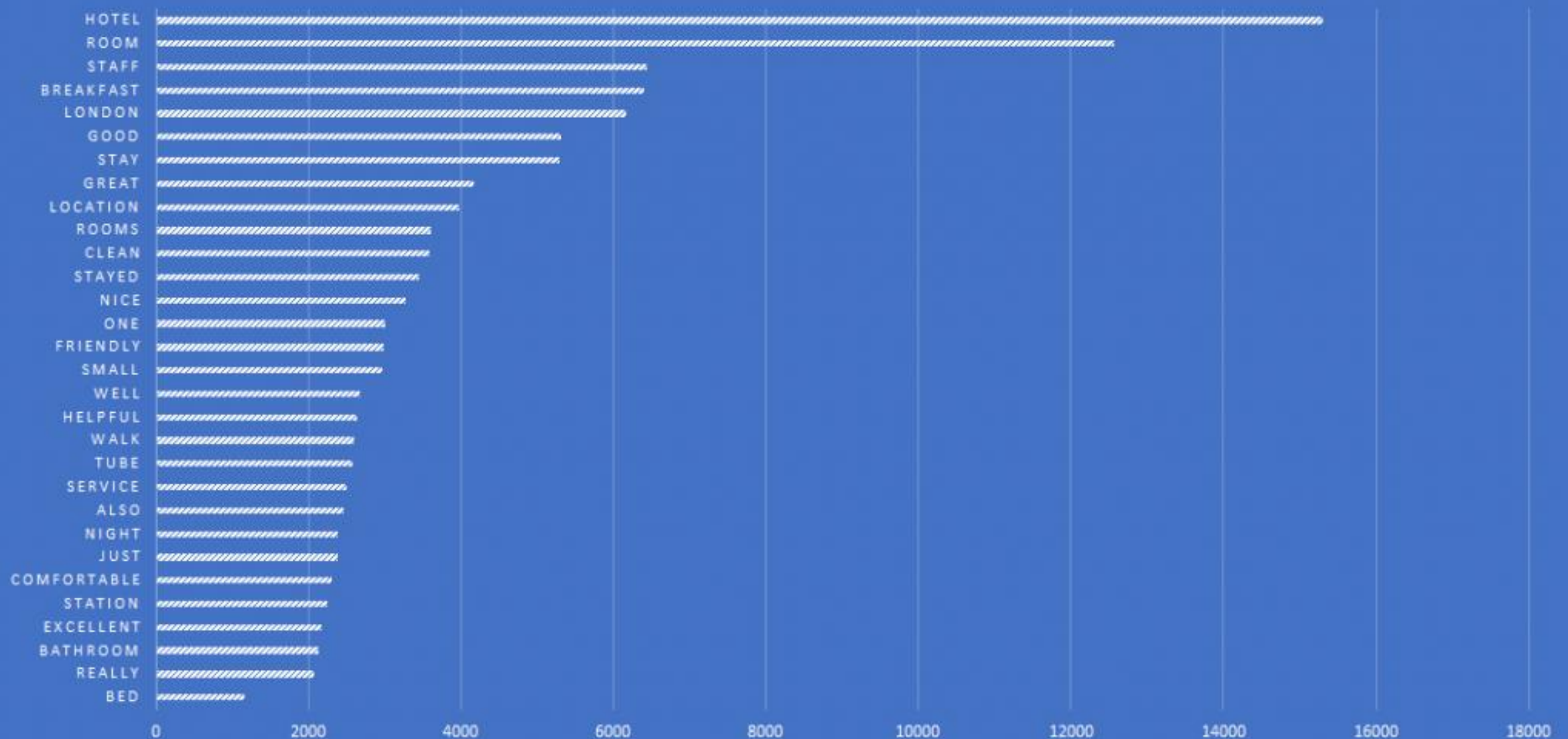
The background features a series of concentric circles in light gray, some solid and some dashed, creating a sense of depth and movement. A large, solid green oval is positioned in the center, serving as a container for the text. A dark gray, curved shape is visible behind the green oval on the left side.

**Frequently used words
in reviews**

TOP 30 MOST FREQUENTLY USED WORDS IN THE REVIEWS FOR THE TOP 10 MOST EXPENSIVE HOTELS IN LONDON



TOP 30 MOST FREQUENTLY USED WORDS IN THE REVIEWS FOR TOP 10 LEAST EXPENSIVE HOTELS IN LONDON

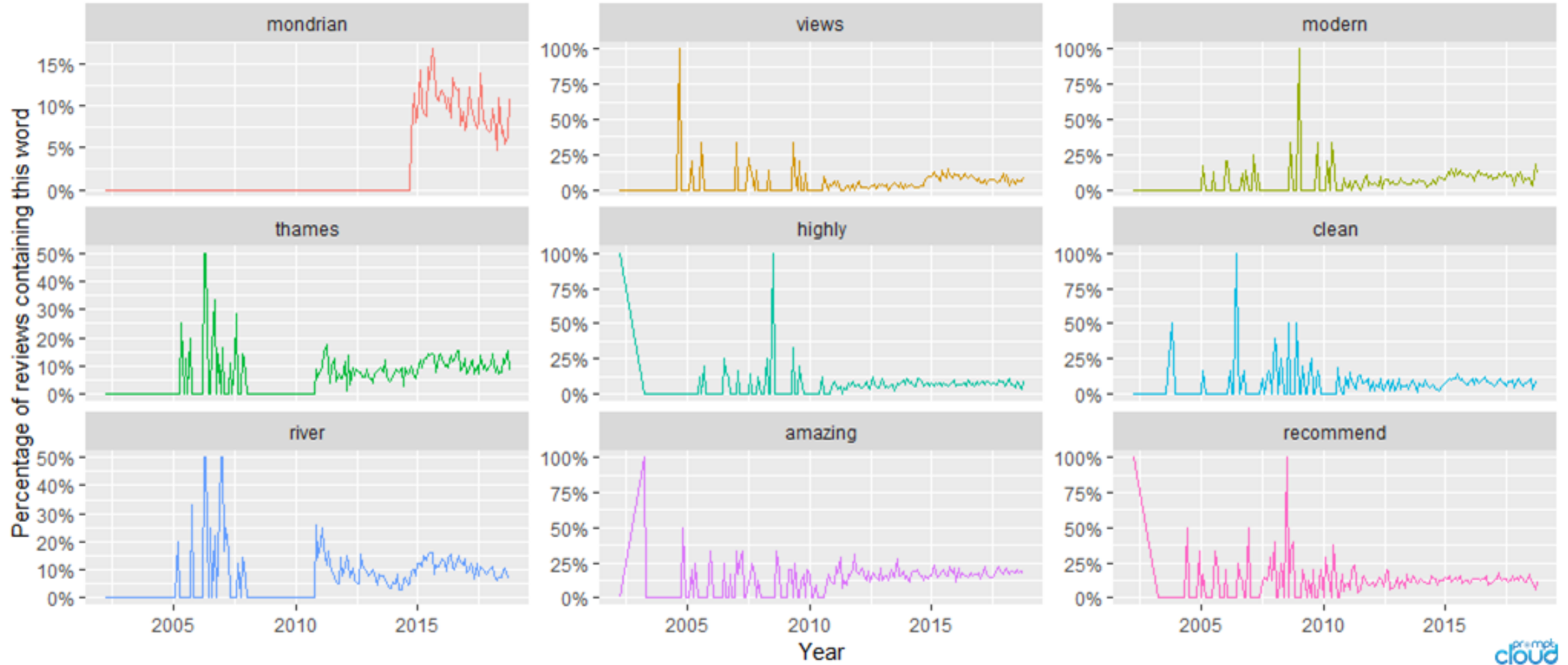


We see that in both of the cases, staffs are major factor. The other most frequent words are 'breakfast', 'stay', 'location', 'rooms', 'clean', 'small', 'station', 'bathroom' and 'bed'. It is interesting to see that 'breakfast' has higher count here in comparison to the reviews for expensive hotels.

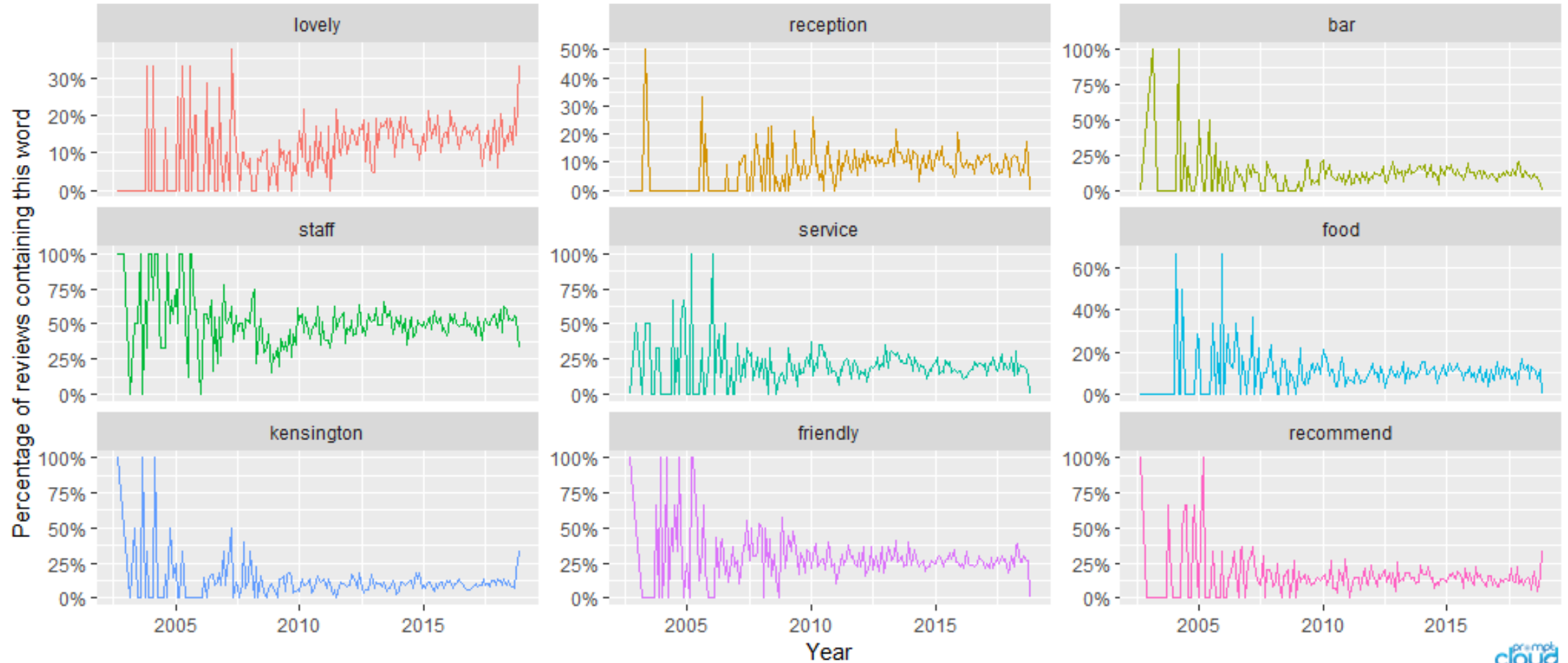
The background features a series of concentric circles in light gray, some solid and some dashed, creating a sense of depth and movement. A large, solid green oval is positioned in the center, containing the text. A dark gray, curved shape is visible behind the green oval on the left side.

**Trending words
present in review text**

9 FASTEST TRENDING WORDS IN REVIEWS POSTED FOR THE TOP 10 MOST EXPENSIVE HOTELS



9 FASTEST TRENDING WORDS IN REVIEWS POSTED FOR THE TOP 10 LEAST EXPENSIVE HOTELS



Most expensive hotels

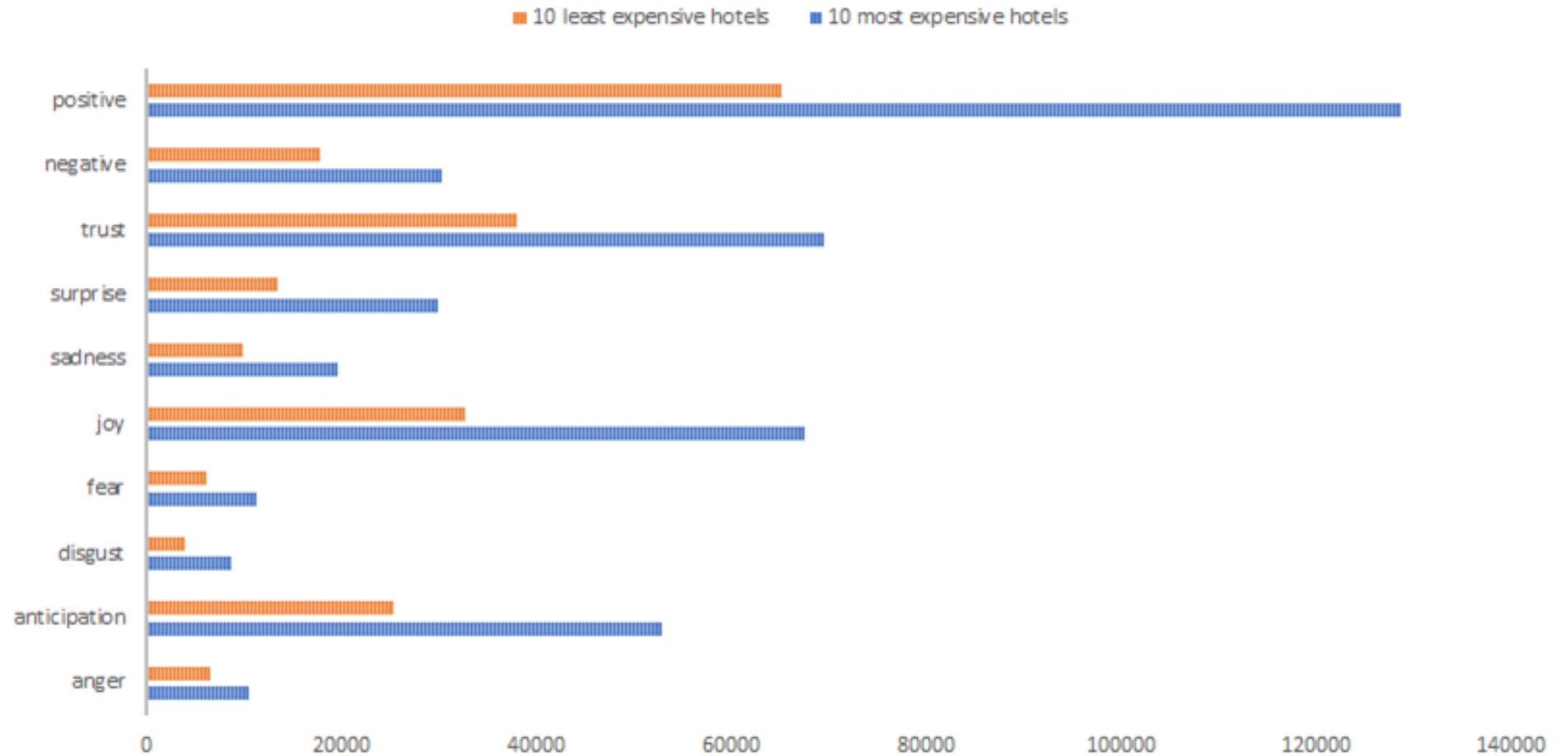
Discussions around 'views' and 'thames' picked up around 2005, while 'clean' and 'modern' spiked in between 2005 and 2010. The word 'river' has also registered significant interest after 2005 and 2010. Also, we see that Mondrian London started trending since its inception in 2014.

Least expensive hotels

Discussions around 'reception', 'bar', 'staff', 'service' and 'food' have been trending. We also see that the location Kensington has captured significant conversation which is an upscale location with regal charm and quite near to London's important attractions.

Sentiment analysis

SENTIMENT AND POLARITY SCORES



There is not much difference if we look at the common pattern for the scores. Overall, they are largely positive for both of the hotel types. In case of emotions, 'trust', 'joy', and 'anticipation' are the prevalent ones.

negative



negative



Most expensive hotels

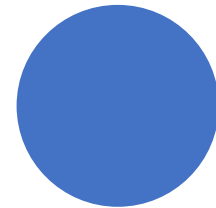
The irony is that in the negative words for this category of hotels, we see 'expensive' as the most frequent word. In the positive words, we see 'friendly', 'comfortable', 'attentive' as some of the important terms. How does this change for the least expensive category?

Least expensive hotels

While 'expensiveness', 'die' (both are common with the two categories) and 'noise' have emerged as the most frequent terms in the negative words, in case of positive terms, 'clean', 'friendly', 'helpful' and 'comfortable' are the most notable ones.

Now it's time for you to [download the data](#) set and perform analysis.

To start off, create network graph of bigrams, perform topic modelling techniques, find out the terms associated in positive/negative words, and compare the importance of various factors (e.g., location vs. service) considered during decision making process.



The image features the PromptCloud logo in a bold, blue, sans-serif font. The word 'prompt' is on the top line, and 'cloud' is on the bottom line. The 'o' in 'prompt' is replaced by a target symbol, and the 'd' in 'cloud' contains a white cloud icon. The background is white and filled with numerous small, light gray icons representing various data and technology concepts, such as databases, charts, documents, and network elements.

prompt cloud

A pioneer in custom and large-scale web data extraction.

www.promptcloud.com | sales@promptcloud.com