

Booking.com

European Luxury Hotel Market Analysis

Study Group 6: Magdalena, Agnieszka, Derek, Yichun, Malay, and Samy

Background & Structure of Analysis

We assisted the Data Insights Team at **Booking.com** with big data and sentiment analysis for market research of high-end hotels in Europe. The aim was to get a general overview of the +500K reviews on 1500 hotels in 6 major cities in Europe.

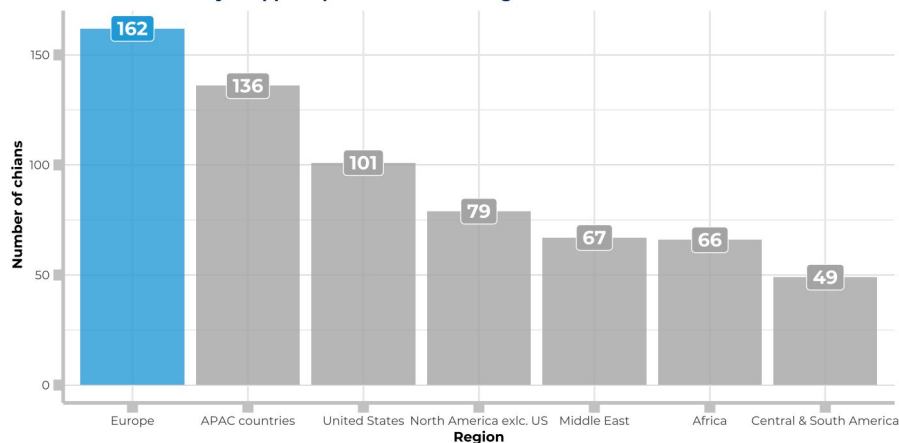
Key Issues to Report on:

- **Distribution** of Hotel Segments in Europe
- **Nationality** and Diversity of (Top) Reviewers
- **Clusters** of Hotels by Average Review
- **Sentiment Analysis** (general, by continent, by region)
- Investigation on the Presence of **Fake/Bot Reviews**
- Investigation on the **Predictive Power** of Reviews

European Hotels Are Market Leaders in Luxury and Upper Upscale

Europe is the most abundant region in Luxury & Upper Upscale Chains

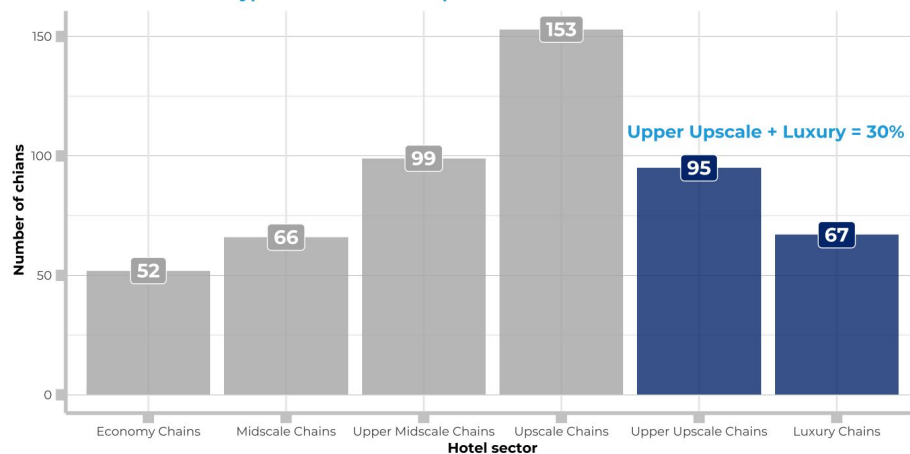
Number of Luxury & Upper Upscale across the regions



Source: <https://www.kaggle.com/playyyywa/global-hotel-chain-presence>

Europe has a large proportion of luxury and upper upscale chains

Number of all the types of chains in Europe

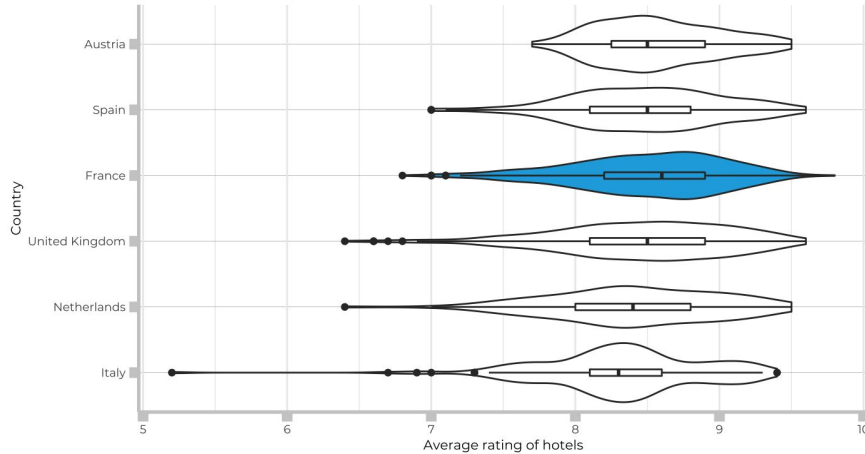


Source: <https://www.kaggle.com/playyyywa/global-hotel-chain-presence>

France and UK Luxury Chains Guarantee Highest Standards and Customer Satisfaction

France has the highest median rating

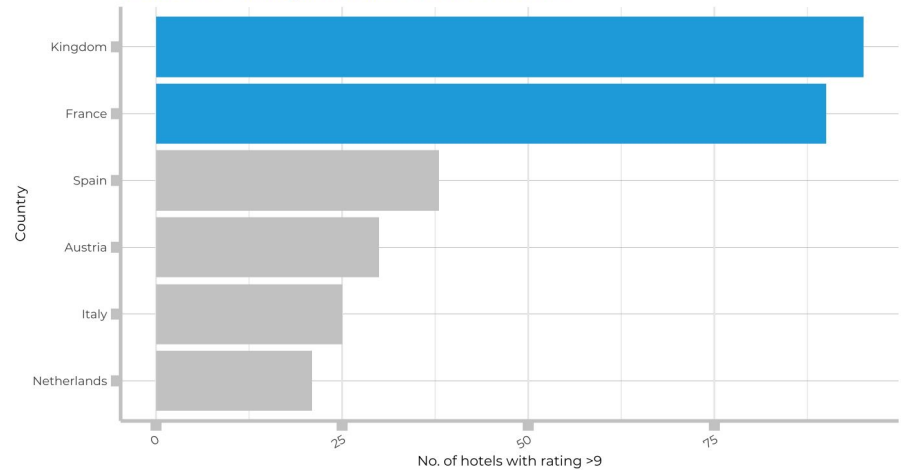
Distribution of average ratings of hotels across 6 selected countries



Source: <https://www.kaggle.com/playyyywa/global-hotel-chain-presence>

UK & France dominate hotels with >9 average rating

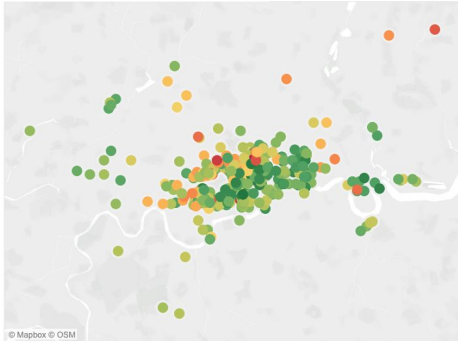
Number of >9 rating hotels in 6 selected countries



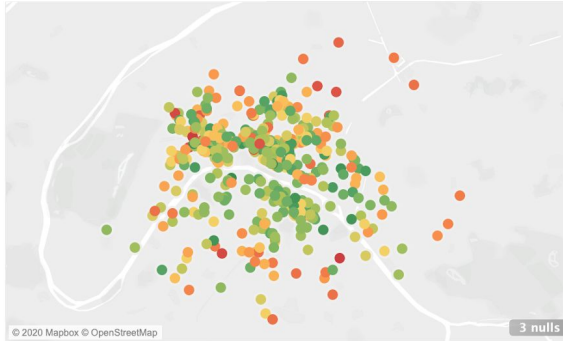
Source: <https://www.kaggle.com/playyyywa/global-hotel-chain-presence>

The Best Reviewed Hotels in General are Located in City Centers

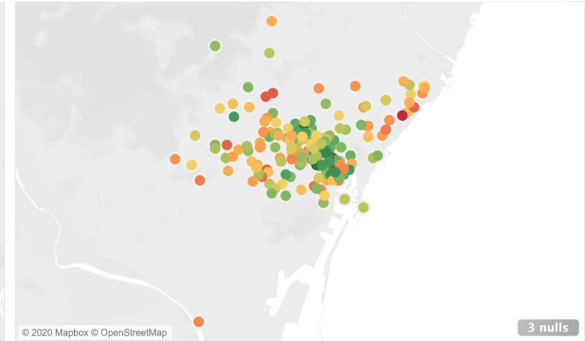
London, UK



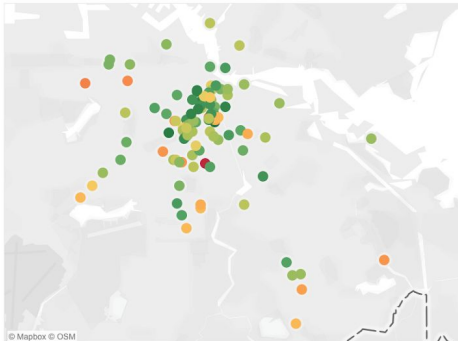
Paris, France



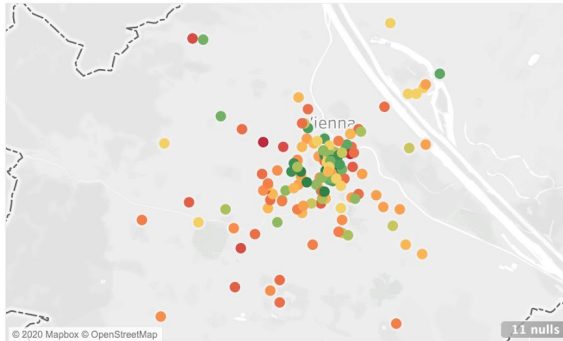
Barcelona, Spain



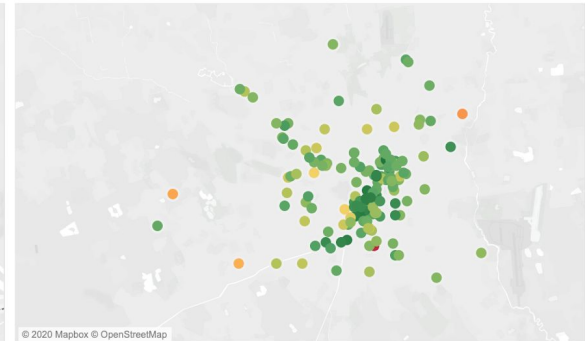
Amsterdam, Netherlands



Vienna, Austria

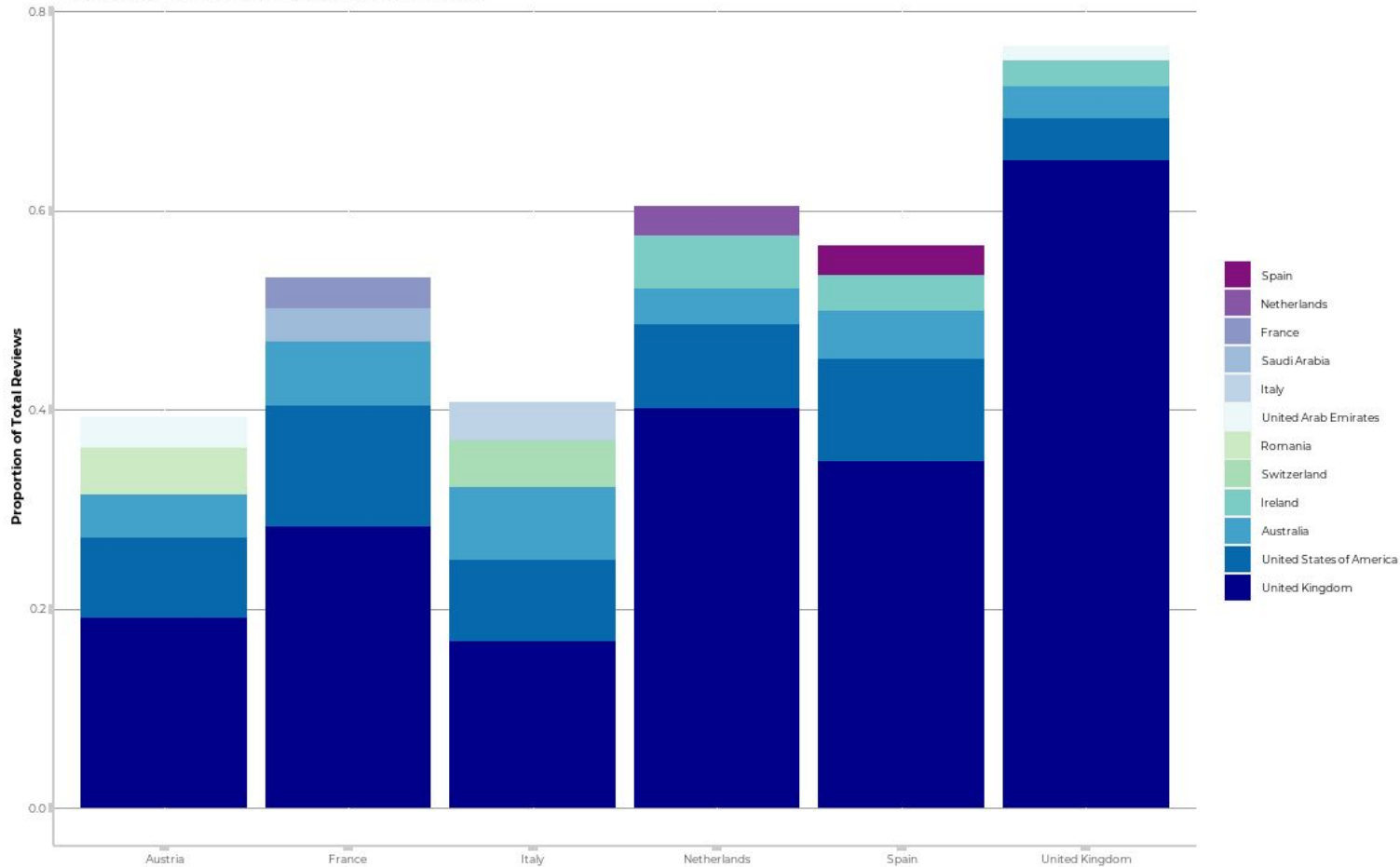


Rome, Italy



Brits take over as Majority Reviewers

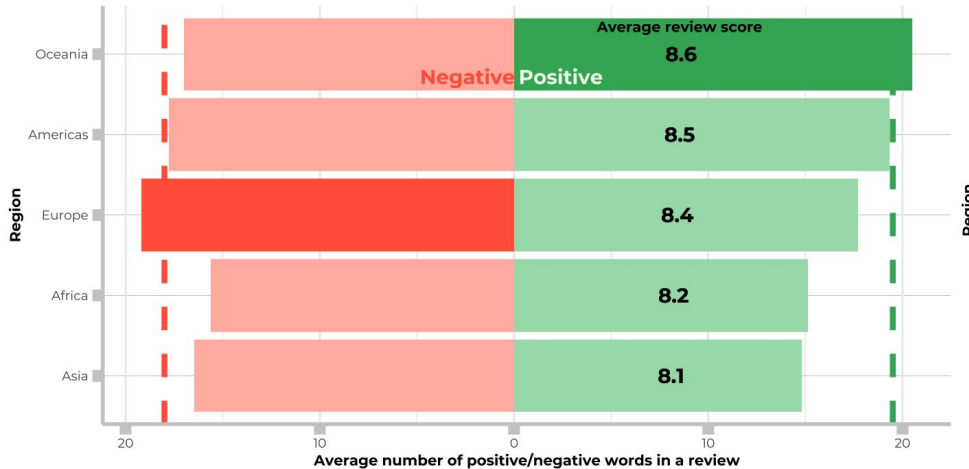
Proportion of Top 5 Reviewer Nationalities per Country



Easy Going Australians & Fussy Europeans

The most criticising guests come from Europe, the most positive come from Australia

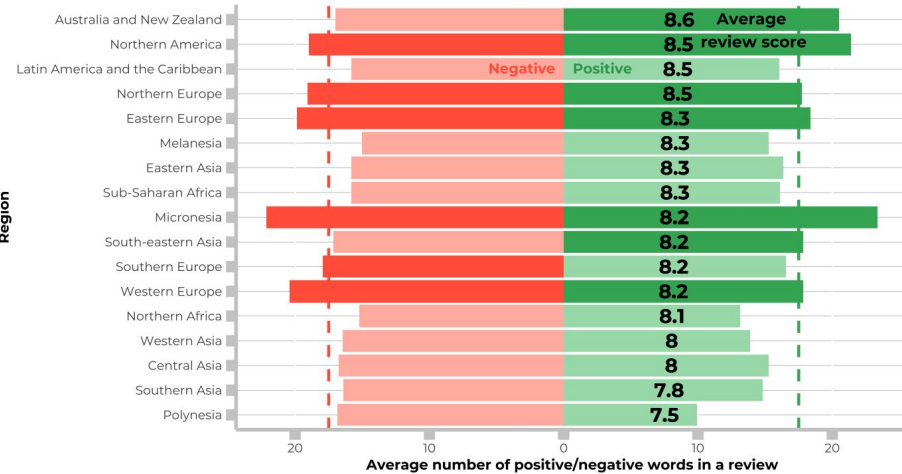
Number of positive and negative words per review per continent and average review score



Source: <https://www.kaggle.com/playyyywa/global-hotel-chain-presence>

Guests from Europe & North America seem to be most influential ones and shaping the opinion

Average number of positive/negative words in a review and average review score



Source: <https://www.kaggle.com/playyyywa/global-hotel-chain-presence>

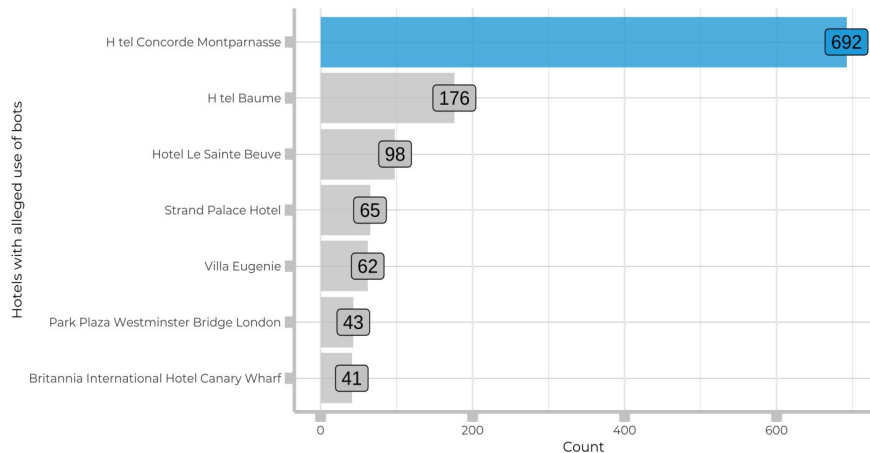
Most Frequent Words in Negative Reviews



Warning: Beat the Bots!

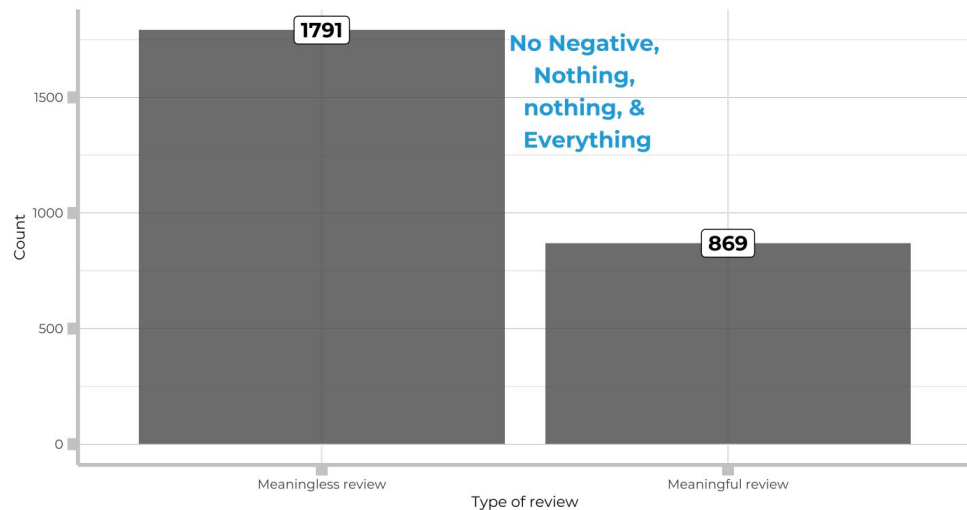
Improve the Security System

Hotels reviewed by highest number of alleged bots



Source: <https://www.kaggle.com/ployyyywa/global-hotel-chain-presence>

Two-thirds of bots produce meaningless reviews



(We define repetitive reviews as exactly one of 'No Negative', 'Nothing', 'nothing' and 'Everything')

Discussion and Critique

Reviews are in English, Introducing an inherent bias in data.

Linear regression did not work as well as predicted as our R-squared was only 0.23 : more variables needed.

Flow Maps seemed to be too messy and too hard to read.

Bot Reviews highlight the importance of control mechanisms to ensure authentic, trustworthy ratings.

Suggested Improvements are prohibitively expensive to do in most cases.

Nationality is the only published demographic