Analyzing Tweets of Booking.com

by Favour ELEDU

02/02/2021

Introduction:

Booking.com is an online travel agency for making reservations of hotels, accommodations and stays. It is one of the most popular travel agency, competing closely with Airbnb. With tourism growing year on year, it's popularity too is growing, and their services are today offered throughout the world. Booking.com is owned and operated by and is the primary revenue source of United States-based Booking Holdings and is headquartered in Amsterdam. Website: booking.com.

As a group assignment, we decided to mine the tweets of bookingcom using the free version of Twitter's API. Booking.com uses only one twitter account (bookingcom), both for their marketing, announcements as well as support.

The goal of this project is preparing a detailed report in order to understand the entire landscape of Booking.com (both from the perspective of the company and its customers) by using their twitter account alone. Using rtweet we retrieved data, primarily focused for the entire year of 2020. In order to analyze the text using the Text Mining and Natural Language Processing we focused on using tidytext and tm.

Throughout this report we will discuss the information we gathered and how they are relevant for the company to better know their customers, and improve their presence, marketing and support on the Twitter platform

To better visualize our data we created a Shiny Application which can be used along with this report. Link for the shiny app: https://kiran-tl.shinyapps.io/BAT_OpenSource_GroupProject_ShinyApp/

By collecting the above datasets using the rtweet package and Twitter's API (free version), we were able to perform a comprehensive analysis of Booking.com based solely on their presence on Twitter!

This comprehensive analysis of Booking.com is divided into the following parts:

1. Getting the data
2. Understanding the landscape of the tweets throughout the world

   • Where are the tweets originating from
   • Top User Locations of the Tweets

3. Analyzing the Frequency of Tweets
4. Analyzing the Top Sources of Tweets
5. Top Tweeters (Screen Names)
6. Top Hashtags
7. Top Keywords

   • Unigrams (Single Words)
   • Bigrams (Two Words)

8. Sentiment Analysis

   • Method 1: Using the bing dictionary
   • Method 2: Using the afinn dictionary

9. Topic Modeling
10. Other Analyses

Importing the Data from Twitter

Using twitter's developer account's API option, we were able to retrieve data related to @bookingcom's twitter account. Being a group project, we were lucky enough to have at least 3 Twitter Developer Accounts with us in order to retrieve the maximum amount of data using rtweet's search_fullarchive function, as a workaround of the limitations of the free API access in order to retrieve as many tweets as possible.
.

In order to understand the landscape of Booking.com, we retrieved the following datasets:

   1. Tweets done by booking.com

This would provide us with the data to understand the type of tweets made by booking.com. For example, using these tweets, we can analyze the various topics that booking.com indulges in.

   2. Tweets made to booking.com

A very important aspect is to understand the tweets received by booking.com. This would help us understand the sentiment of the customers and the perception of the customers towards booking.com.

   3. Retweets of booking.com

Retweets is often another way of endorsing the tweets. This could help us understand further the perception of the customers towards booking.com

   4. Hasthtags of #booking.com

Hashtags are often used as keywords. Analyzing hashtags is another way of analyzing the sentiment of the tweet. By grouping hashtags, we could understand the general trend of the tweets and also the travel and tourism trends.

   5. Likes made by booking.com

Analyzing the tweets liked by booking.com could further provide information on how booking.com endorses its customers points of views. This can further provide more information on the positive environment bookin.com wishes to operate

   6. Followers of booking.com

One of the most important aspect of any business is the customers. Followers of twitter, can often be perceived as customers of the account they follow. Understanding the landscape of the customers, is of great importance in being able to provide the best services.

   7. Lists followed by booking.com

Lists in Twitter are like groups. Analyzing the lists that booking.com is part of can better help understanding the general trends in travel and tourism.

   8. Tweets with "booking.com" as keyword.

Analyzing the keywords where booking.com is mentioned (without specifically tagging booking.com) is an alternative in ensuring we capture the entire outlook.

Understanding the geographical spread

Where are the tweets originating from

With booking.com being part of the travel and tourism sector, understanding the geographical spread is of key importance. Knowing the answers to the questions such as, where the customers are from, where are the most complaints coming from, tourists from which countries travel the most, which are the most frequently visited cities and countries is of great importance.

Note: All tweets retrieved from twitter do not contain the location details. Only a subset of the data containing the location details were used for the analysis in this section.

1. Tweets made by @bookingcom

Note: All tweets/retweets made by booking.com contain no location. With booking.com headquartered in Amsterdam, we believe most of the company's tweets/retweets would be originating from this city.

2. Tweets made to @bookingcom

Geographical spread of tweets made to @bookingcom
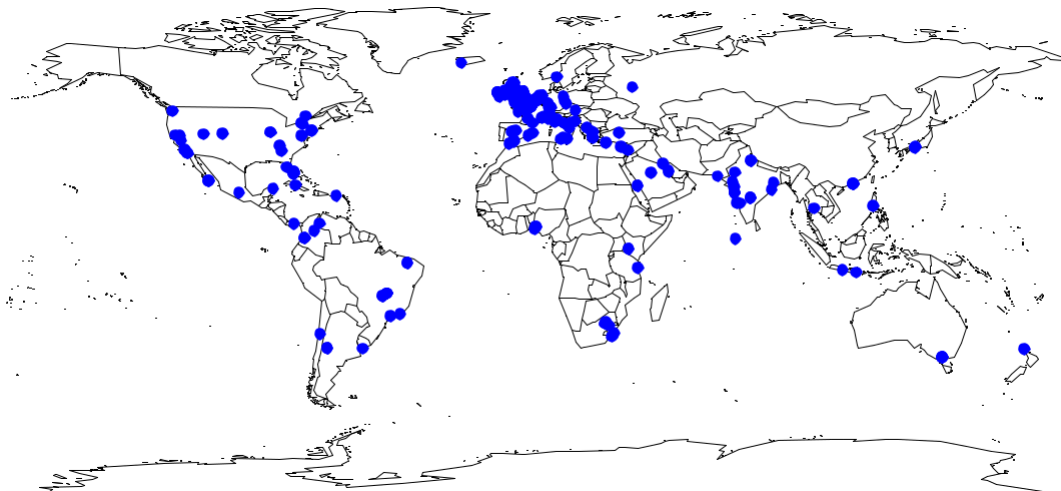
The above plot indicates that most of the tweets originate from Europe, with USA and India following.
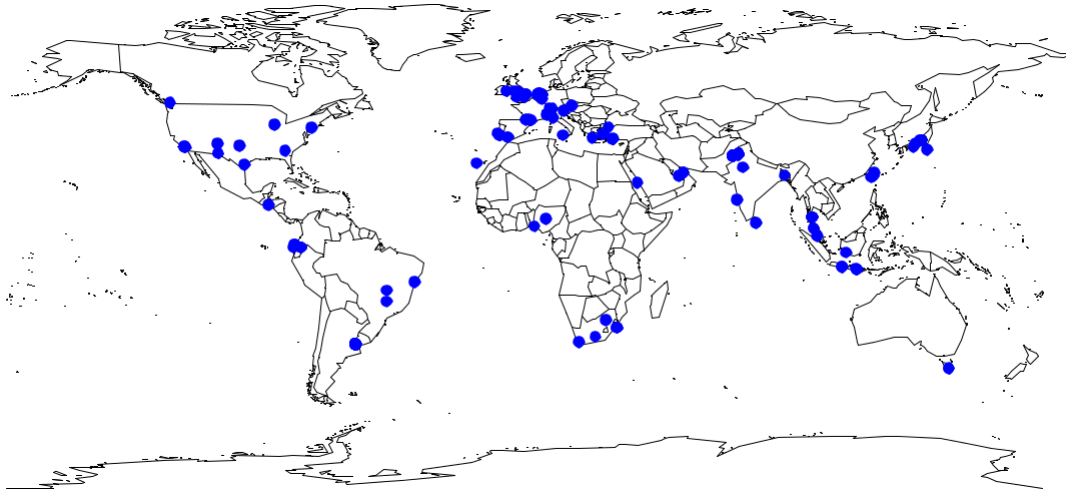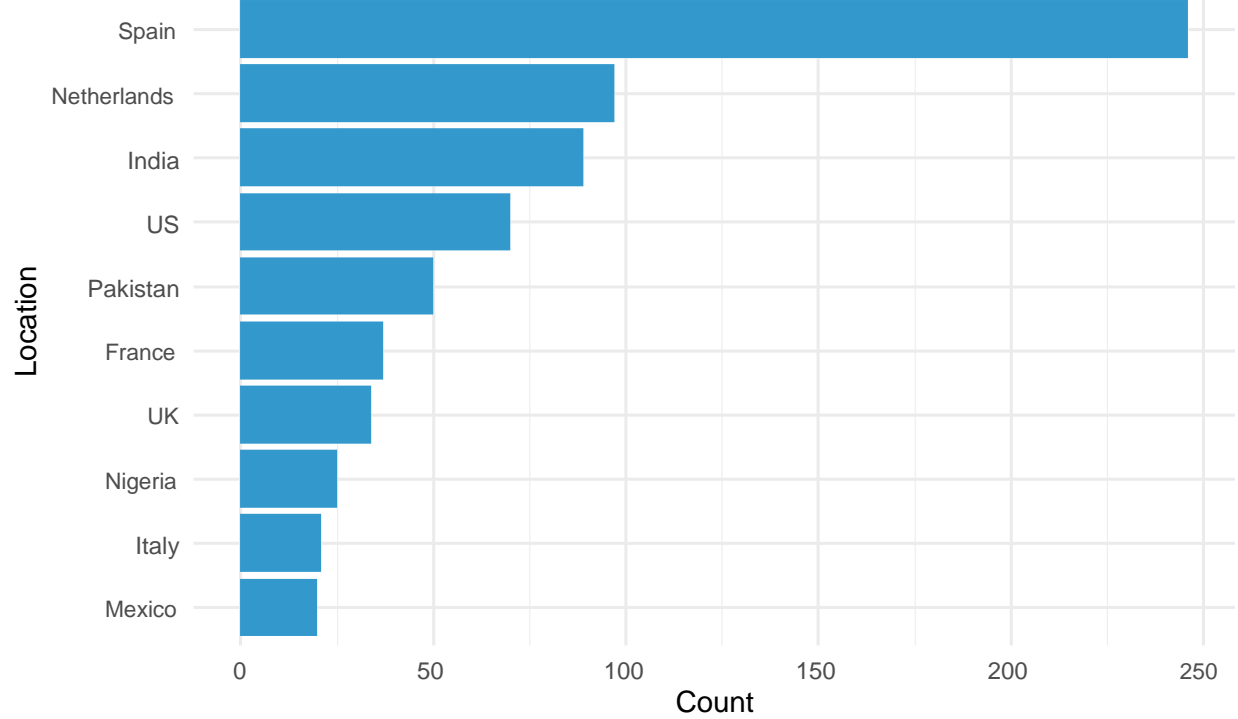
Top User Locations of Tweets made to @bookingcom
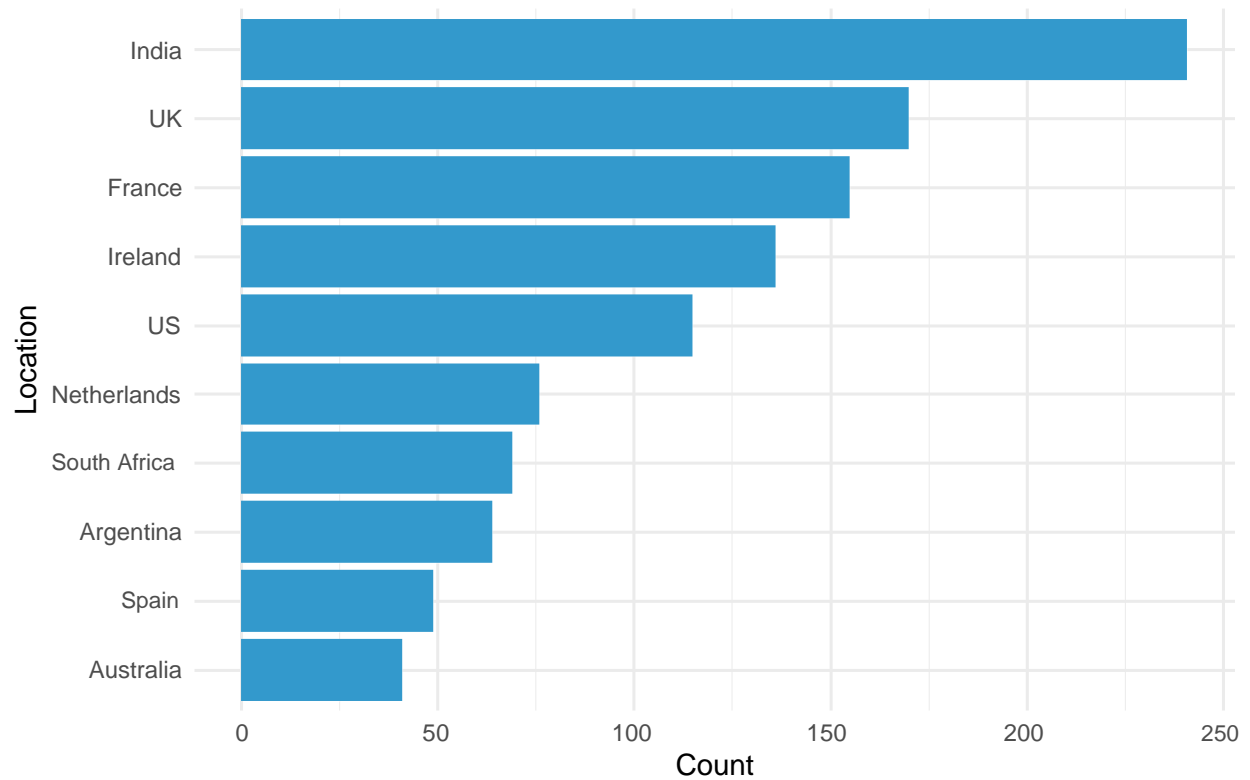
3. Retweets made by @bookingcom

Note: All tweets/retweets made by booking.com contain no location. With booking.com headquartered in Amsterdam, we believe most of the company's tweets/retweets would be originating from this city.

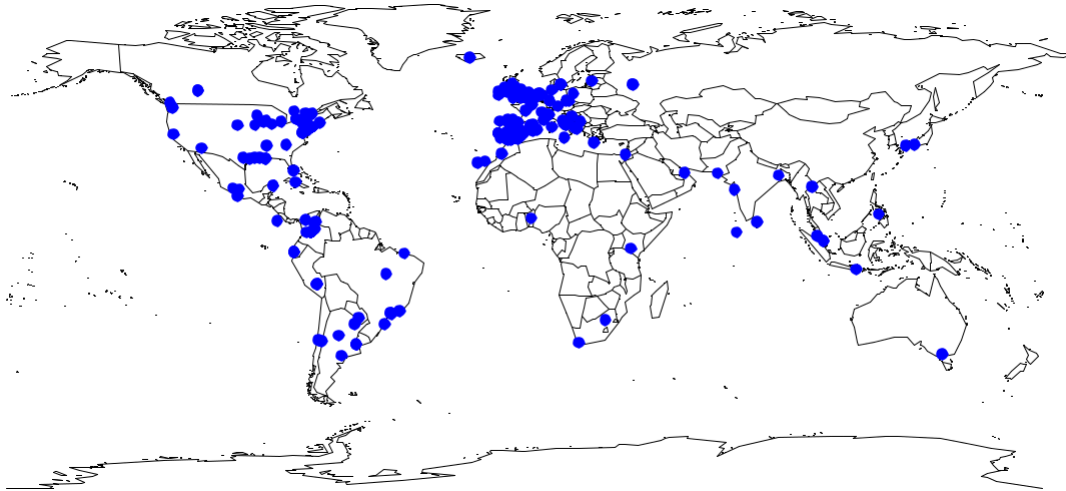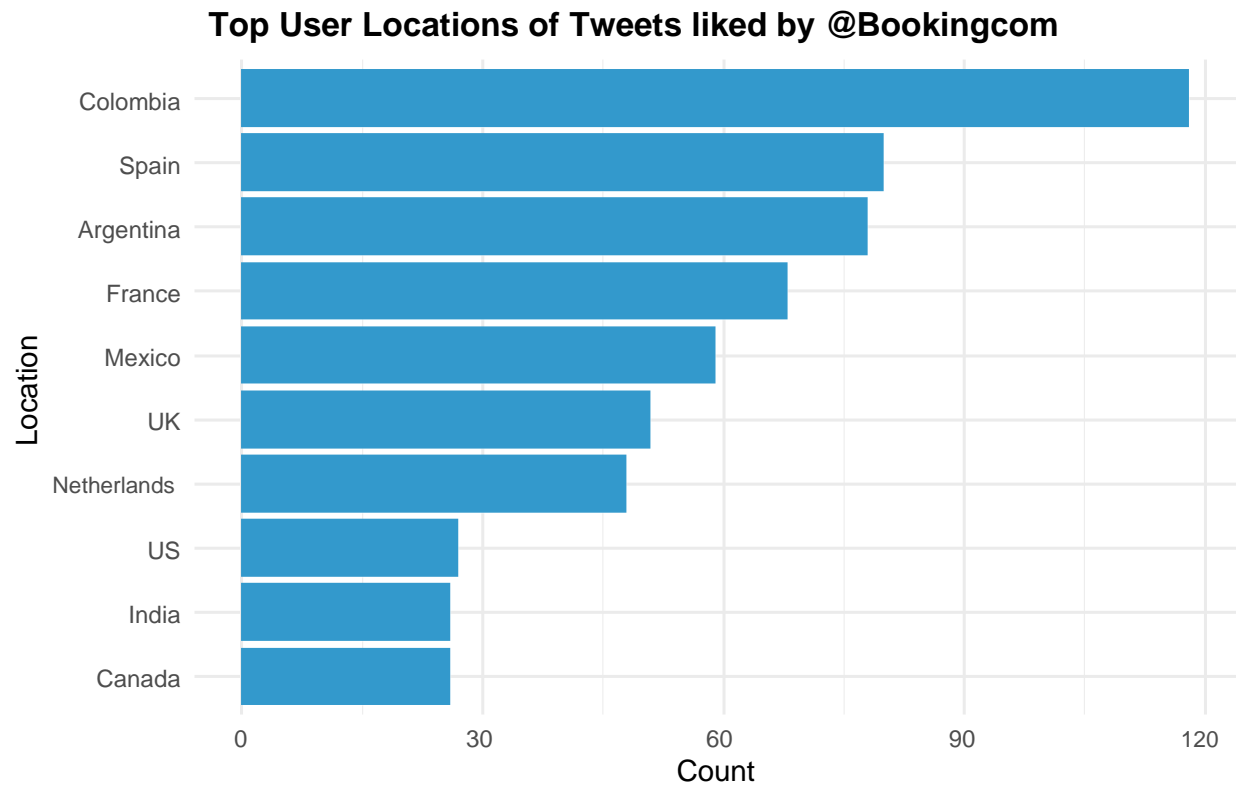**4. Tweets made with #bookingcom hashtag**

Top User Locations of Tweets made using #bookingcom hashtag

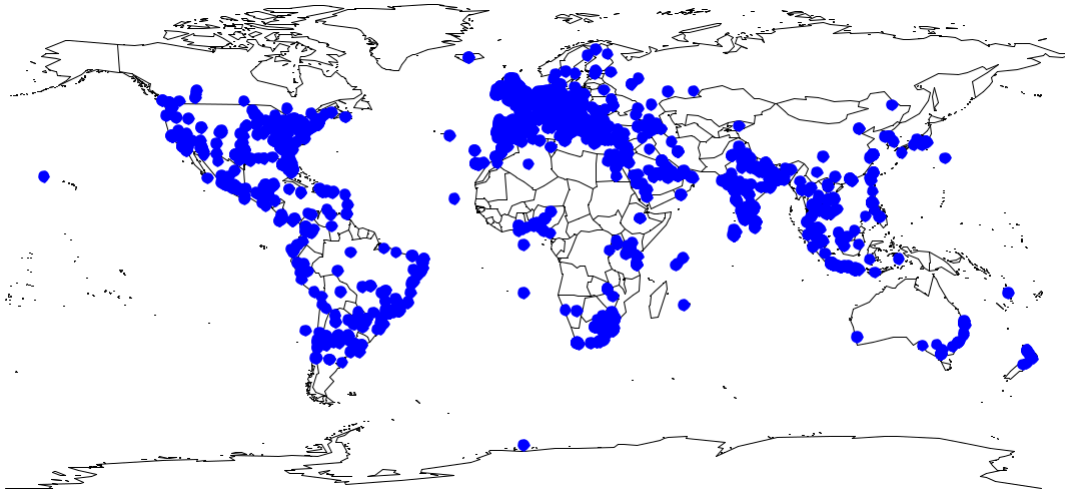Source: Data collected from Twitter's REST API via rtweet
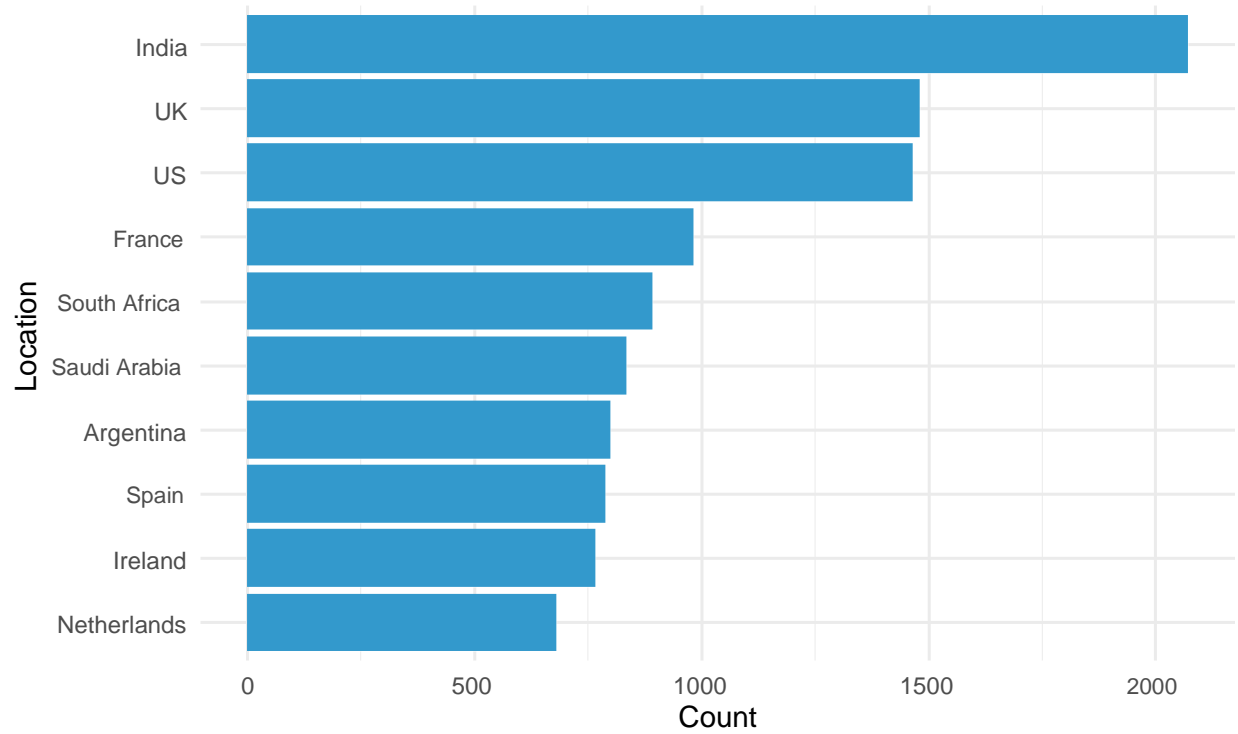
5. Tweets liked by @bookingcom

Top User Locations of Tweets liked by @Bookingcom

Source: Data collected from Twitter's REST API via rtweet
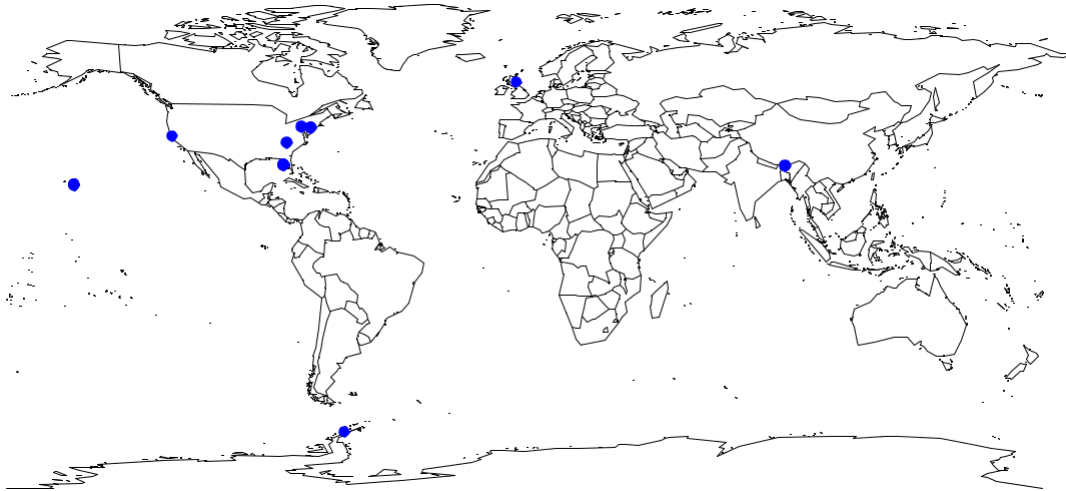
6. Followers of @bookingcom

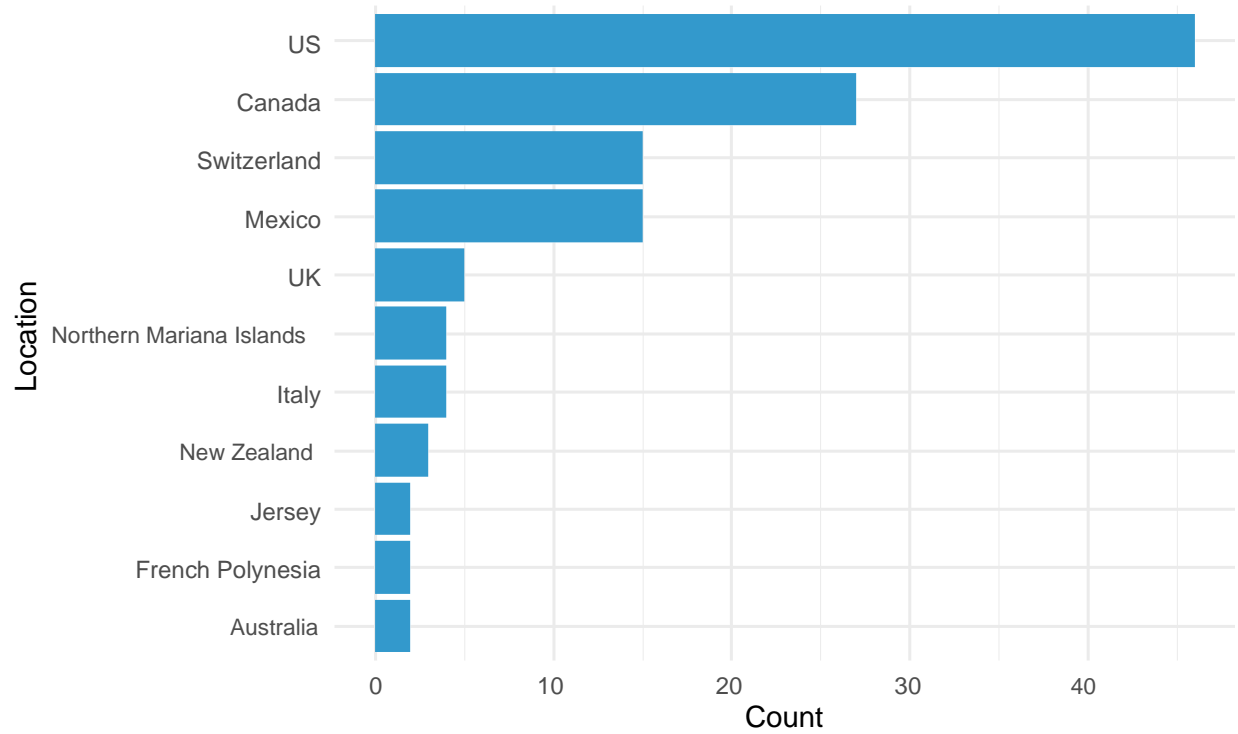## Top User Locations of Followers of @Bookingcom

Source: Data collected from Twitter's REST API via rtweet

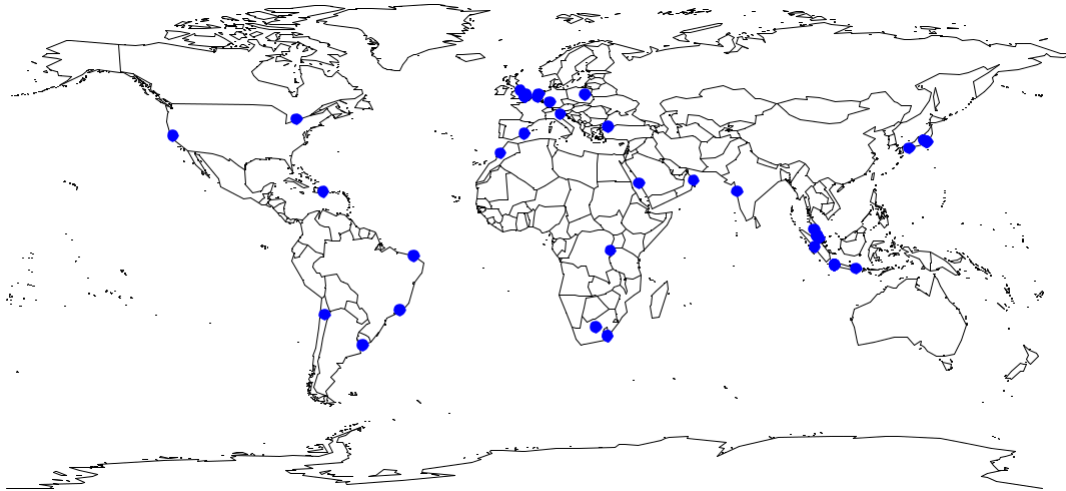**7. Tweets of subscribers associated with @bookingcom lists**

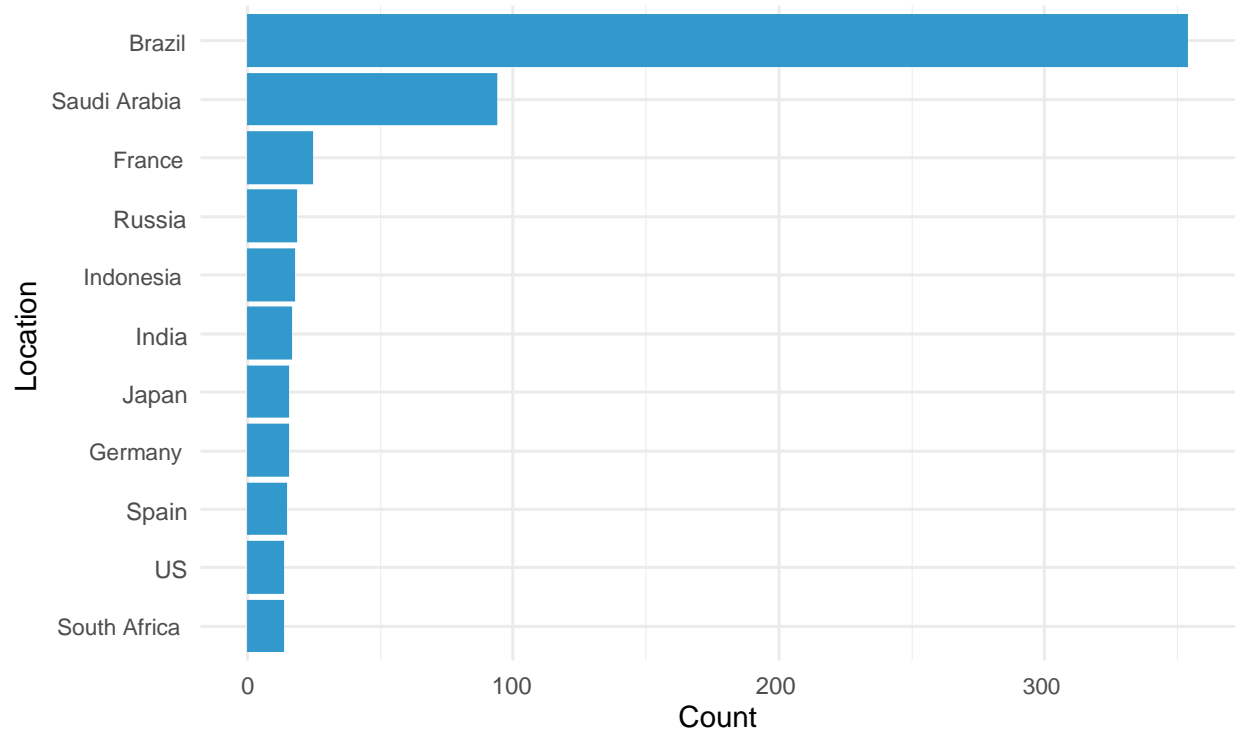**Top User Locations of @Bookingcom's common lists subscribers**

Source: Data collected from Twitter's REST API via rtweet

8. Tweets made with booking.com as keyword

Top User Locations of Tweets made with Booking.com as keyword

Source: Data collected from Twitter's REST API via rtweet