# TRAVEL HANDBOOK


## ABSTRACT

The work presented here buckets top 50 most beautiful cities of the world into mutually exclusive clusters. The cluster labels are self-explanatory and depict the most prominent venue category in that particular cluster of cities.



Trehan, Rahul

Applied Data Science Capstone Project

# Table of Contents

**Capstone Project - The Battle of the Neighborhoods (Cities)**

**Applied Data Science Capstone by IBM/Coursera**

## Data

I have taken the list of "THE WORLD'S 50 MOST BEAUTIFUL CITIES" from the website https://www.flightnetwork.com/blog/worlds-most-beautiful-cities/

The website contains a numbered list of the top 50 cities globally. As such, ***BeautifulSoup*** package has been used to scrape the website and get the top 50 cities in a tabular format.

The output after this step is as shown below:

```
In [23]: top_50_df
```

Out[23]:

| | City_name |
|---|---|
| 0 | PARIS |
| 1 | NEW-YORK |
| 2 | LONDON |
| 3 | VENICE |
| 4 | VANCOUVER |
| 5 | BARCELONA |
| 6 | CAPE-TOWN |
| 7 | SAN-FRANCISCO |
| 8 | SYDNEY |
| 9 | ROME |
| 10 | SINGAPORE |

In order to find out the geographical coordinates of these cities, the ***geopy*** package was used.

Next, ***Foursquare*** was used to explore the vicinity of these top 50 cities.

A **radius** of 20 km from the city center was chosen and **limit** of 1000 results per city was applied.

Upon completing the above steps, the data thus obtained looked like:

In [36]:
```
nearby_venues
```

Out[36]:

| | City | Latitude | Longitude | Venue | Venue Latitude | Venue Longitude | Venue Category |
|---|---|---|---|---|---|---|---|
| **0** | PARIS | 48.856697 | 2.351462 | Cathédrale Notre-Dame de Paris | 48.853124 | 2.349561 | Church |
| **1** | PARIS | 48.856697 | 2.351462 | Fleux' | 48.858763 | 2.354161 | Furniture / Home Store |
| **2** | PARIS | 48.856697 | 2.351462 | Place de l'Hôtel de Ville – Esplanade de la Li... | 48.856925 | 2.351412 | Plaza |
| **3** | PARIS | 48.856697 | 2.351462 | Shakespeare & Company | 48.852568 | 2.347096 | Bookstore |
| **4** | PARIS | 48.856697 | 2.351462 | Miznon | 48.857201 | 2.358957 | Israeli Restaurant |
| **5** | PARIS | 48.856697 | 2.351462 | La Maison d'Isabelle | 48.850007 | 2.348443 | Bakery |
| **6** | PARIS | 48.856697 | 2.351462 | Centre Pompidou – Musée National d'Art Moderne | 48.860730 | 2.351660 | Art Museum |
| **7** | PARIS | 48.856697 | 2.351462 | Comme à Lisbonne | 48.856767 | 2.356462 | Café |

This marks the end of the data gathering stage. Now we will proceed with describing the overall methodology employed in carrying out the analysis.