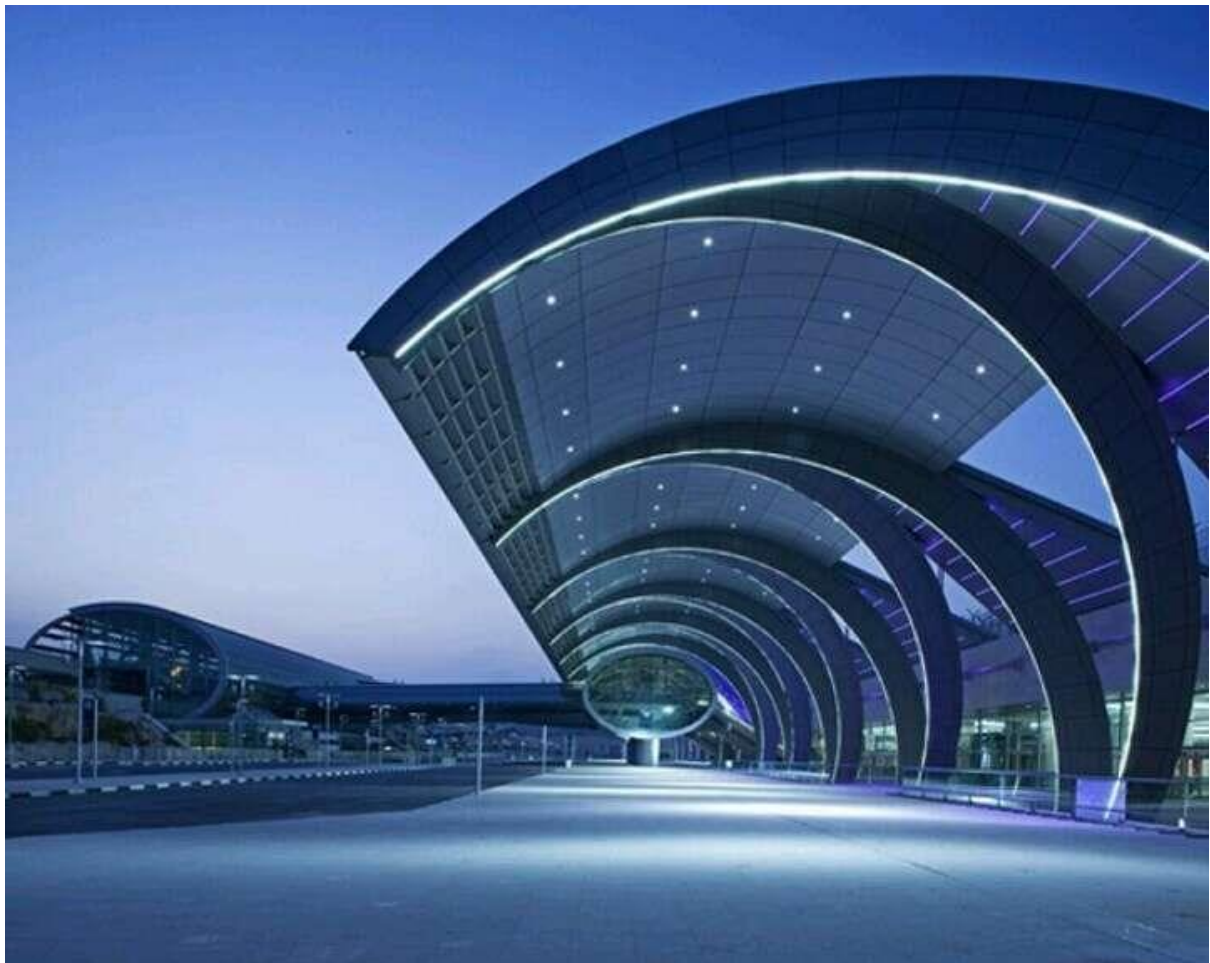


# CLUSTERING OF BUSIEST INTERNATIONAL AIRPORTS IN THE WORLD

*By Parikshit Verma*

Email: [parikshit.iitvaranasi@gmail.com](mailto:parikshit.iitvaranasi@gmail.com) [Linkedin](#) [GitHub](#)



## Contents

|   |          |
|---|----------|
| <b>1. Introduction &amp; Background .....</b>                         | <b>3</b> |
| <b>1.1. Target audience.....</b>                                      | <b>4</b> |
| <b>1.2. Problem statement .....</b>                                   | <b>4</b> |
| <b>2. Data Description .....</b>                                      | <b>4</b> |
| <b>2.1. FourSquare API.....</b>                                       | <b>4</b> |
| <b>2.2. List of busiest international airports in the world .....</b> | <b>6</b> |
| <b>2.3. Birds eye view of data .....</b>                              | <b>7</b> |

## List of figures

|   |   |
|---|---|
| Figure 1: First flight of the Wright Flyer, December 17, 1903. (Pic courtesy Wikipedia) ..... | 3 |
| Figure 2: Busiest airports in world .....   | 7 |

## 1. Introduction & Background

It has always been the dream of mankind from time immemorial to fly in the sky like the birds do. Umpteen number of unsuccessful attempts had been made by humanity in this quest upto the very beginning of 20<sup>th</sup> century. Finally in 1903 two American brothers made history when they flew over North Carolina sky first time in the history of mankind.



*Figure 1: First flight of the Wright Flyer, December 17, 1903. (Pic courtesy Wikipedia)*

Up until this date fascination of flying in the sky has not died but only has increased. The fact that is clearly established by the World Air Transport Statistics, 2018 by IATA. According to report, since 2004 the revenue growth of the airlines has outpaced world GDP continuously. Since the people are flying too frequently airports are not just the places to catch flight anymore. Airports have become more of a micro city offering all types of facilities & services that one can imagine and need. These facilities and services not only to promote businesses but also are necessary as many a times one has to spend a considerable amount of time on airport due to procedural requirements, non-availability of direct flight to intended destination etc.

Foursquare is an online platform that provides plethora of services including details about places of interest within a specified distance of a geographic location. As of the date of writing this article there were 941 different categories of places of interest covered by Foursquare. This huge information when combined with the airports data may reveal useful information for the stakeholders.

## 1.1. Target audience

In view of above it is quite useful from a commuter's point of view to have an idea that which airports in the world are likely to have the facilities he or she requires while at the airport. This may help people to choose from various flights options available to them. On the other side it may help business owners and multinational corporations to decide whether they should choose one airport or the other for their next store.

## 1.2. Problem statement

To classify busiest international airports of the worlds on the basis of certain set of establishments/ facilities/ services/ places of interest using foursquare APIs and any other relevant data.

## 2. Data Description

Data required for this project comes from different sources. Since our aim is to group airports on the basis of different categories of places of interest like different types of restaurants, services, places to stay etc. this shall constitute major part of our data and shall be obtained using foursquare API. Other data like list of busiest airports and their location shall be obtained by scraping Wikipedia pages and other relevant reports if necessary.

### 2.1. FourSquare API

Since one of the busiest international airports are quite huge hence to search for relevant places of interest a radius of 2km has been chosen. Four square API shall be used to get the number of places of interest of a particular category for each airport. But before that we required a list of relevant target categories. FourSquare [categories page](#) was scraped to accomplish this task giving us total 941 categories of places available in FourSquare data bases. Out of these 941 categories 442 categories are deemed to be relevant for our business problem and have been grouped into following 31 groups :-

**1 Eateries** Food BBQ Joint Bistro Breakfast Spot Diner Food Court Food Stand Food Truck Friterie Halal Restaurant Molecular Gastronomy Restaurant Restaurant Salad Place Seafood Restaurant Theme Restaurant Vegetarian / Vegan Restaurant Food & Drink Shop Airport Food Court

**2 Asian\_restaurant** Afghan Restaurant Asian Restaurant Korean Restaurant Bossam/Jokbal Restaurant Bunsik Restaurant Gukbap Restaurant Janguh Restaurant Samgyetang Restaurant Mongolian Restaurant Tibetan Restaurant Comfort Food Restaurant Gluten-free Restaurant

**3 African\_restaurant** African Restaurant Ethiopian Restaurant Moroccan Restaurant

**4 NAmerican\_restaurant(North American)** New American Restaurant American Restaurant Cajun / Creole Restaurant Caribbean Restaurant Cuban Restaurant Fried Chicken Joint Hawaiian Restaurant Poke Place Mac & Cheese Joint Mexican Restaurant Botanero Burrito Place Taco Place Tex-Mex Restaurant Yucatecan Restaurant Poutine Place Southern / Soul Food Restaurant Steakhouse Wings Joint

**5 SAmerican\_restaurant(South American)** Latin American Restaurant Arepa Restaurant Empanada Restaurant Salvadoran Restaurant South American Restaurant Argentinian Restaurant Brazilian Restaurant Acai House Baiano Restaurant Central Brazilian Restaurant Churrascaria Empada House Goiano Restaurant Mineiro Restaurant Northeastern Brazilian Restaurant Northern Brazilian Restaurant Pastelaria Southeastern Brazilian Restaurant Southern Brazilian Restaurant Tapiocaria Colombian Restaurant Peruvian Restaurant Venezuelan Restaurant

**6 Australian\_restaurant** Australian Restaurant

**7 European\_restaurant** Austrian Restaurant Bagel Shop Belgian Restaurant Czech Restaurant Dutch Restaurant English Restaurant Fondue Restaurant German Restaurant Bavarian Restaurant

Bratwurst Joint Currywurst Joint Franconian Restaurant German Pop-Up Restaurant Palatine Restaurant Rhenisch Restaurant Schnitzel Restaurant Silesian Restaurant Swabian Restaurant Hungarian Restaurant Modern European Restaurant Polish Restaurant Portuguese Restaurant Scandinavian Restaurant Scottish Restaurant Slovak Restaurant Spanish Restaurant Paella Restaurant Tapas Restaurant Swiss Restaurant

**8 IndianSub\_restaurant(Indian sub-continent)** Himalayan Restaurant Bangladeshi Restaurant Indian Restaurant Andhra Restaurant Awadhi Restaurant Bengali Restaurant Chettinad Restaurant Dhaba Dosa Place Goan Restaurant Gujarati Restaurant Hyderabad Restaurant Indian Chinese Restaurant Irani Cafe Jain Restaurant Karnataka Restaurant Kerala Restaurant Maharashtrian Restaurant Mughlai Restaurant Multicuisine Indian Restaurant North Indian Restaurant Northeast Indian Restaurant Parsi Restaurant Punjabi Restaurant Rajasthani Restaurant South Indian Restaurant Udupi Restaurant Pakistani Restaurant Sri Lankan Restaurant

**9 Chinese restaurant** Chinese Restaurant Anhui Restaurant Beijing Restaurant Cantonese Restaurant Cha Chaan Teng Chinese Aristocrat Restaurant Chinese Breakfast Place Dim Sum Restaurant Dongbei Restaurant Fujian Restaurant Guizhou Restaurant Hainan Restaurant Hakka Restaurant Henan Restaurant Hong Kong Restaurant Huaiyang Restaurant Hubei Restaurant Hunan Restaurant Imperial Restaurant Jiangsu Restaurant Jiangxi Restaurant Macanese Restaurant Manchu Restaurant Peking Duck Restaurant Shaanxi Restaurant Shandong Restaurant Shanghai Restaurant Shanxi Restaurant Szechuan Restaurant Taiwanese Restaurant Tianjin Restaurant Xinjiang Restaurant Yunnan Restaurant Zhejiang Restaurant

**10 MiddleEastern\_restaurant** Falafel Restaurant Jewish Restaurant Kosher Restaurant Kebab Restaurant Middle Eastern Restaurant Egyptian Restaurant Iraqi Restaurant Israeli Restaurant Kurdish Restaurant Lebanese Restaurant Persian Restaurant Ash and Haleem Place Dizi Place Gilaki Restaurant Jegaraki Tabbakhi Shawarma Place Syrian Restaurant Yemeni Restaurant

**11 Japanese\_restaurant** Japanese Restaurant Donburi Restaurant Japanese Curry Restaurant Kaiseki Restaurant Kushikatsu Restaurant Monjayaki Restaurant Nabe Restaurant Okonomiyaki Restaurant Ramen Restaurant Shabu-Shabu Restaurant Soba Restaurant Sukiyaki Restaurant Sushi Restaurant Takoyaki Place Tempura Restaurant Tonkatsu Restaurant Udon Restaurant Unagi Restaurant Wagashi Place Yakitori Restaurant Yoshoku Restaurant Noodle House

**12 SouthEastAsia\_restaurant** Burmese Restaurant Cambodian Restaurant Filipino Restaurant Hotpot Restaurant Indonesian Restaurant Indonesian Restaurant Balinese Restaurant Betawinese Restaurant Indonesian Meatball Place Javanese Restaurant Manadonese Restaurant Padangnese Restaurant Sundanese Restaurant Malay Restaurant Mamak Restaurant Satay Restaurant Thai Restaurant Som Tum Restaurant Vietnamese Restaurant

**13 French\_restaurant** Creperie French Restaurant Alsatian Restaurant Auvergne Restaurant Basque Restaurant Brasserie Breton Restaurant Burgundian Restaurant Catalan Restaurant Ch'ti Restaurant Corsican Restaurant Estaminet Labour Canteen Lyonese Bouchon Norman Restaurant Provençal Restaurant Savoyard Restaurant Southwestern French Restaurant

**14 Italian\_restaurant** Italian Restaurant Abruzzo Restaurant Agriturismo Aosta Restaurant Basilicata Restaurant Calabria Restaurant Campanian Restaurant Emilia Restaurant Friuli Restaurant Ligurian Restaurant Lombard Restaurant Malga Marche Restaurant Molise Restaurant Piadineria Piedmontese Restaurant Puglia Restaurant Romagna Restaurant Roman Restaurant Sardinian Restaurant Sicilian Restaurant South Tyrolean Restaurant Trattoria/Osteria Trentino Restaurant Tuscan Restaurant Umbrian Restaurant Veneto Restaurant Pizza Place

**15 Turkish\_restaurant** Turkish Restaurant Borek Place Cigkofte Place Doner Restaurant Gozleme Place Kofte Place Kokoreç Restaurant Kumpir Restaurant Kumru Restaurant Manti Place Meyhane Pide Place Pilavcı Söğüş Place Tantuni Restaurant Turkish Home Cooking Restaurant Çöp Şiş Place

**16 Greek/Mediterranean\_restaurant** Greek Restaurant Bougatsa Shop Cretan Restaurant Fish Taverna Grilled Meat Restaurant Kafenio Magirio Meze Restaurant Modern Greek Restaurant Ouzeri Patsa Restaurant Souvlaki Shop Taverna Tsipouro Restaurant Mediterranean Restaurant

**17 Russia/East Europe\_restaurant** Eastern European Restaurant Belarusian Restaurant Bosnian Restaurant Bulgarian Restaurant Romanian Restaurant Tatar Restaurant Russian Restaurant Blini House Pelmeni House Ukrainian Restaurant Varenyky restaurant West-Ukrainian Restaurant Caucasian Restaurant

**18 Desserts/Bakery** Bakery Dessert Shop Cupcake Shop Frozen Yogurt Shop Ice Cream Shop Pastry Shop Pie Shop Donut Shop Indian Sweet Shop Candy Store Chocolate Shop

**19 Cafe** Bubble Tea Shop Cafeteria Coffee Shop Tea Room Turkish Coffeehouse

**20 Fast food/Snack Burger Joint** Deli / Bodega Dumpling Restaurant Fast Food Restaurant Fish & Chips Shop Hot Dog Joint Sandwich Place Snack Place Soup Place

**21 Pub/Bar/Brewries** Gastropub Apple Wine Pub Irish Pub Bar Beer Bar Cocktail Bar Pub Whisky Bar Wine Bar Champagne Bar Brewery

**22 Beverages/juices/shakes** Juice Bar Beer Store Liquor Store Wine Shop

**23 Waiting/Stay** Lounge Airport Lounge Hotel Bed & Breakfast Boarding House Hostel Inn Motel Vacation Rental Travel Lounge

**24 General\_Shopping** Shop & Service Baby Store Convenience Store Department Store Cheese Shop Gourmet Shop Grocery Store Leather Goods Store Music Store Toy / Game Store

**25 Gifts/ Souvenirs** Antique Shop Flower Shop Gift Shop Souvenir Shop Thrift / Vintage Store

**26 Electronics\_stores** Camera Store Electronics Store Mobile Phone Shop Mobility Store

**27 Lifestyle** Clothing Store Accessories Store Boutique Kids Store Lingerie Store Men's Store Shoe Store Women's Store Costume Shop Cosmetics Shop Fabric Shop Health & Beauty Service Jewelry Store Watch Shop Salon / Barbershop

**28 Books** Bookstore Comic Shop

**29 Tour & Travel** Tour Provider Travel Agency Bus Station Bus Line Bus Stop Metro Station Rental Car Location Taxi Stand Tourist Information Centre Train Station Train Transportation Service

**30 Duty Free shops** Duty-free Shop

**31 Services** ATM Currency Exchange Baggage Locker Bike Rental / Bike Share

Only getting the names of categories was only half the work done as for API calls we require code of category also arranged in a similar manner as the groups. For arranging names in group as mentioned, excel was used. Similarly, a corresponding sheet was populated using 'Index' and 'Match' function of excel to find the corresponding code for each member of the newly formed group. Both these excel sheets were loaded in two different data sets for further processing.

To get the number of places of interest corresponding to each group, every element of the group has been queried using 'Explore' function of FourSquare API for each of the airport. And the number so obtained is stored in separate excel sheets. Building this data for number of places of interest for each of the group member required 62255 API calls. However before making API calls we needed to have list of airports of our interest.

## 2.2. List of busiest international airports in the world

To find the list of busiest airports in the world two Wikipedia pages were scraped. First pages '[List of cities by international visitors](#)' and second one '[List of international airports by country](#)'.

List of cities by international visitor was the basis to decide which airports shall be part of our project and provided us with the data 'Country' and 'City'.

List of international airports by country was then used to find the 'Name' and 'IATA Code' of airports of our interest. However, it was observed that as per wikipedia list, one airport may have multiple city names and hence same had to be resolved before comparing the same to list of cities by international visitors.



Finally the both the lists were merged to form one pandas dataframe and rearranged, resulting in a data frame containing 'Country', 'City', 'Airport' and 'IATA Code' for 155 airports around the world.

To the list of airports so obtained, latitude and longitude for each airport was added using Geopy library. For some of the airports geopy didn't seem to work and same had to be done manually.

Note:- All the data available is for the year 2018.

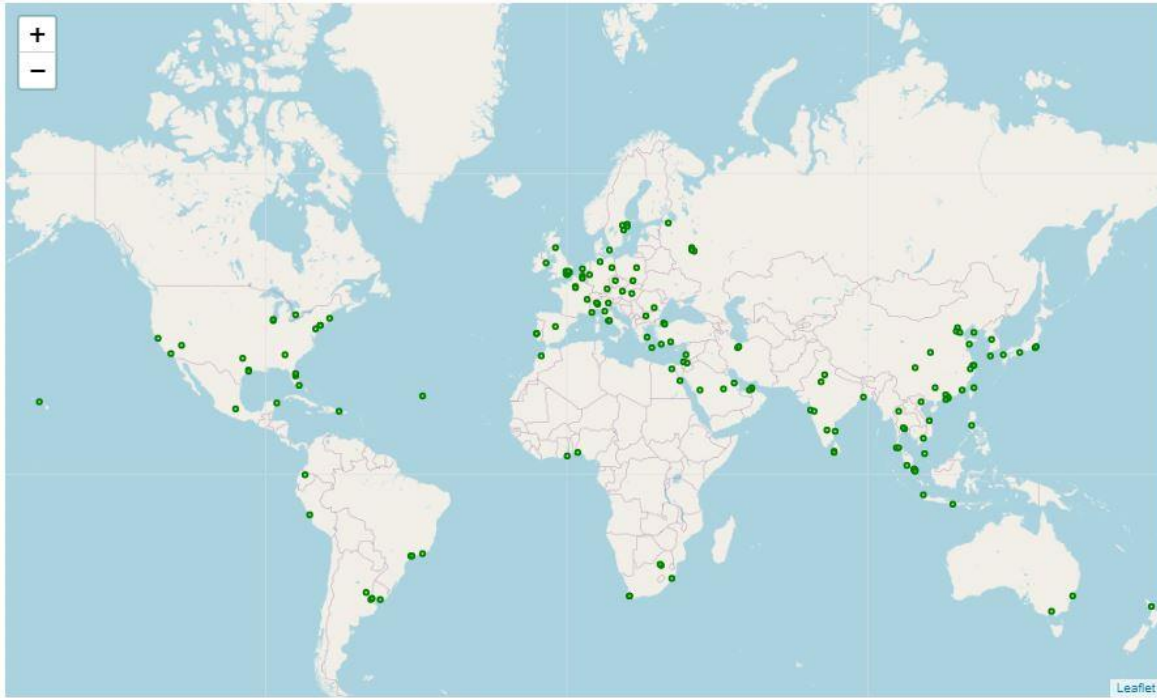


Figure 2: Busiest airports in world

### 2.3. Birds eye view of data

For our final data set 'IATA Code' column from section 2.2 Data was taken and corresponding to each airports all 31 categories of section 2.1 were added containing count of each category group within 2 km radius of the location (latitude and longitude) obtained using Geopy library. Additionally a column 'distribution' is added to the end of this dataframe to get an idea of distribution of facilities in terms of average distance from central point.

Our final data base comprised of 'IATA Code', 31 groups of categories and finally a distribution column for each of the 155 airports. And same was populated with integer values representing number of a places of that particular category located within 2km from the centre (latitude longitude location form Geopy) of airport.