

Fun City Rankings

Introduction

Which cities in the world are the most fun?

A lot depends on this question.

When tourists decide where to travel, their top question is “will I have fun?”. When planners invest in a city, their goal is to build a place where residents will enjoy their lives. Fun is at the center of a virtuous circle. Happy cities bring companies looking to hire productive workers.

Who is Ranking Cities?

Naturally the question is also of considerable interest to cultural journalists who run surveys and publish articles. Readers pay for this research by way of advertising, and tourism agencies pay for this research by way of sponsorship. It’s big business, which is why service providers like Time Out are listed on the London AIM stock exchange.

Time Out describe themselves as a “global media and entertainment business that inspires and enables people to explore and enjoy the best of the city”
[\[www.timeout.com/about/time-out-group\]](http://www.timeout.com/about/time-out-group).

The “Time Out Index” [\[www.timeout.com/things-to-do/time-out-index\]](http://www.timeout.com/things-to-do/time-out-index) is a survey designed to compare life in the greatest cities around the world. The Time Out Index 2019 surveyed 34,000 people about life in their city. They then ranked the 48 best cities in the world.

The top 5 ranked cities were:

1. New York
2. Melbourne
3. Chicago
4. London
5. Los Angeles

The bottom 5 ranked cities were:

44. Marseille
45. Bangkok
46. Kuala Lumpur
47. Beijing
48. Istanbul

Problem Statement: Replacing Surveys with Location Data

Surveying 34,000 people is an expensive project. The results are also prone to interpretation and strong cultural biases that cannot be completely scrubbed from the data.

Can location data be used instead to provide a cheaper and more objective “Fun City Rank”?

This project examines that question. The analysis pulls categorized venue location data from Foursquare [www.foursquare.com] to generate a “Fun City Ranking”. The analysis is then compared to the original “Time Out Index 2019” to see if pure data analytics can be used as a substitute for more personal surveys.

What is Fun?

Fun activities are things people *want* to do – as opposed to activities that people *need* to do. Hence venues that offer cultural activities were included in “fun”, while those which offered professional activities were excluded.

Included Venues	Excluded Venues
Museums	College Buildings
Art Galleries	Business Centers
Nature Preserves	Shopping Malls
Monuments	Hotels

Shopping venues and hotels were also generally excluded from this analysis. While shopping can be very fun, it is primarily an activity we *need* to do. Hotels, except for resorts, are usually not fun destinations in themselves, but are a means for tourists to access the fun they seek.

Fun was then divided into different ways a person could experience an activity. Seven (7) top-level descriptive verbs were used to classify fun:

Top Level Fun
Watch
Visit
Strive
Amuse
Indulge
Relax
Party

Some fun activities involve simply sitting and watching. Other fun activities involve visiting and exploring. While still others involved physically striving. Examples are provided below:

Watch	Visit	Strive
Circus	Exhibitions	Fitness Center
Movie Theatre	Memorial Sites	Bike Trail
Concert Hall	Botanical Garden	Fishing Spot
Stadium	Scenic Lookout	Golf Course

Activities that didn't quite fit into Watch, Visit, or Strive were allocated to "Amuse":

Amuse
Casino
Pool Hall
Racecourse
Theme Park

Food and drink were the most difficult to include in the definition of "fun". On the one hand, eating is clearly something that we *need* to do – not just something we *want* to do – and hence not fun. On the other hand, going to a restaurant is an important part of a fun social life.

But it is difficult to rank restaurants globally as a special activity designed to bring enhanced levels of joy. For example, going to a Szechuan Restaurant might be very special in Paris but very ordinary in Beijing.

As a result, only eating venues that were established for the sole purpose of fun and allowed patrons to "Indulge" were included in the analysis. Beverage venues were then generally split between "Relax" and "Party" (along with a few other venues).

Indulge	Relax	Party
Desert Shop	Coffee Shop	Bar
Creperie	Massage Studio	Nightclub
Candy Store	Sauna	Music Venue
Vineyard	Resort	Dance Club

Data Source

The Foursquare API was used to collect the location data for the analysis. Two Foursquare API's were used:

1. Search for Venues
2. Get a Venue's Photos

Search for Venues

Search is a Regular API Endpoint that can be called up to 99,500 times per day. Both user and userless authentication is accepted. The GET request looks like:

GET <https://api.foursquare.com/v2/venues/search>

The 'near' field can be passed to specify a starting location. In this case, the downtown location of each of the 48 cities from the Time Out Index 2019 were used.

A 'categoryId' can be passed to the call to focus on those venues classified within a "Top Level Fun" category.

The 'radius' field is set so that fun venues within 2 km of each downtown city are returned.

The search API was called for 48 cities x 123 categories = 5,904 times.

Each call returns up to 50 venue results. A total of 71,283 venues were collected.

Get a Venue's Photos

Photos is a Premium API Endpoint so it can only be called up to 500 times per hour. Both user and userless authentication is accepted. The GET request looks like:

GET https://api.foursquare.com/v2/venues/VENUE_ID/photos

The 'VENUE_ID' collected from the Search API is passed to the call. But only those venues that are likely to return good photos were used. Hence venues that fall within the following categories were selected:

- Historic Site
- Monument / Landmark
- Intersection (an extra category searched for the purpose of photo collection)

Analytic Challenges: Targeting Tourists

This analysis focused primarily on tourists coming to enjoy a city. It was assumed that tourists would arrive in the downtown heart of the city and be interested in just those activities within a short 2 km journey.

While a broader "Fun City Rank" analysis would also involve the long-term residents of a city, focusing just on tourists eliminates a lot of practical problems. For instance, a full analysis would involve the data collection of all venues across the entire city – not just those within a 2 km circle of downtown. This would, in turn, require an analysis of the size of the city as well as population density so that some sort of "fun per capita" metric could be generated.

But tourists are less concerned with a city's size, population, and density. Their primary concern is whether fun venues are easily accessible nearby.

Methodology: Calculating a "Fun City Rank"

The Fun City Rank of each city is calculated by comparing both the *number* and *variety* of fun venues accessible to tourists.

Those cities that scores well in each “Top Level Fun” category – the cities that offer the most ways to *Watch* and *Visit* and *Strive* and *Amuse* and *Indulge* and *Relax* and *Party* – will be ranked among the world’s most fun cities.

Cities that do not offer a wide variety of fun are therefore punished. For example, a city that has dozens of casinos but nothing else will score low overall.

Detailed Methodology

This project will:

1. Use the Foursquare Venue Search API to collect city venue data [Regular Endpoint]
 - a. Data is collected for the same Top-48 cities identified in the Time Out Index
 - b. The search covers venues fitting the 7 “Top Level Fun” categories (these cover 123 Foursquare Categories and Sub-Categories)
 - c. The search radius covers 2 km around each downtown city center
 - d. The maximum of 50 venues per query is collected for each [City x Category]
2. Calculate a “Fun Rank” for each city
 - a. Parse the Foursquare data to calculate the number of accessible fun venues
 - b. Group the individual categories by “Top Level Fun” categories
 - c. Sort each city by each Top-Level Fun category and allocate points (this normalizes the results so that a city with 2 large stadiums won’t lose out to a city with 3 small salsa clubs)
 - d. Allocate ‘stars’ to each city
The best city in each Top-Level Fun category receives 5 stars
The worst city will receive 0 stars.
 - e. Sum the total points given to each city
 - f. Sort and rank all the cities
3. Compare the “Fun Rank” with the “Time Out Index 2019”
 - a. Create a correlation Scatter Plot comparing the two results
 - b. Run a regression to fit the Fun Rank against the Time Out Index
 - c. Discuss sources of discrepancies
4. Map Fun Venues in each City
 - a. Randomly select 10% of venues for plotting on each city map
 - b. Use the Folium library to map the selection of venues for select cities
 - c. Color code the venues by Top-Level Fun category
5. Use the Foursquare Venue Photos API [Premium Endpoint]
 - a. Collect city photos taken from nearby intersections and historic sites
6. Prepare a blog posting discussing the project findings
 - a. Include rankings, charts, maps, and photos