**Coursera Capstone Foursquare API Project**

Introduction

The purpose of this project is to allow you to easily compare different cities based on their top venues. This is particularly useful if you are planning a holiday or moving to a new city and you want to be sure the best places in a city align with your interests.

A lot of the information available about city venues is anecdotal (bloggers, travel guides etc.) and be time consuming to search through. Taking a more data-based approach you could ask what category is most popular among a cities top 100 venues? If you love parks/recreation you may be interested in London where 25 of it’s top 100 venues are parks. If you’re a foodie you may enjoy Berlin where 35 of their top 100 venues are food spots.

Further to just stating numbers, it should be easy to visualize and compare how popular each category of venue is. An easy way to do this is with radar charts, if you’ve played video games such as FIFA you may recognize these when you compare player attributes. So the final result is an easy to way to compare the attributes of each city.. as if they were players on FIFA.

Data

The data used is solely provided using the Foursqaure API

Method

The first step in the process is to do an example on just one city, then combine steps in that process into a function to use on other cities. My example case is London, UK:

1. Pass London, UK along with the number of top venues required into Foursquare API call
2. Drop relevant venue information into a dataframe. Info such as:
   1. Venue name
   2. Venue category
   3. Venue location
3. Define a function to create a radar plot for each category type. This is done by:
   1. Plotting a polar coordinate axis
   2. Setting the number of points using the number of categories
   3. Then plotting the shape on the axis based on counts of each category
4. There are visibly too many categories to make this readable – so the categories must be aggregated
5. I’ve then combined steps 1-2 into a function and passed several cities into this function to create a long list of categories and removed duplicates
6. I’ve created 8 distinct category groups, each venue category will assigned to 1
7. A few keywords are listed for each category group, if a venue’s category contains a keyword it will be assigned to that group
   1. Categories which are not assigned a group are assigned to ‘Other’
8. Category groups can now be assigned to our London dataframe
9. The radar chart is replotted with grouped categories and now looks much better
10. Now we can combine the previous steps into a function and pass any city name (or list of cities to compare) into the function and compare attributes
11. The final result is an easy to visualize method of comparing city venues!

Results

The results are based on what the user would like to input and compare. However, we can generally see that the smaller a city gets the more likely it is to have ‘Food’ as the most popular category. You could say this is expected as larger cities generally have more things to do and are more likely to capitals of their country where popular cultural and nightlife based venues are more common. It is interesting to see how capital cities differ. Paris’s top 100 venues are spread out evenly across 4 or 5 categories whereas Amsterdam is largely towards food (coffeeshops are included here so not too surprising maybe..)

Conclusion

The project achieves an easy way to compare cities based on their venues, and specifically for holidaymakers it can be particularly useful and provide a different approach to choosing a location or simply to find out more about a city that isn’t from a conventional travel guide.

**This project will be presented to colleagues!**