**Peer-graded Assignment: Capstone Project**

**The Battle of Neighborhoods**

Introduction:

Now that you have been equipped with the skills and the tools to use location data to explore a geographical location, over the course of two weeks, you will have the opportunity to be as creative as you want and come up with an idea to leverage the Foursquare location data to explore or compare neighborhoods or cities of your choice or to come up with a problem that you can use the Foursquare location data to solve.

You will required to submit the following:

WEEK1:

1. A description of the problem and a discussion of the background
2. A description of the data and how it will be used to solve the problem.

WEEK2:

1. A link to your Notebook on your Github repository, showing your code
2. A full report consisting of all of the following components
   1. Introduction where you discuss the business problem and who would be interested in this project
   2. Data where you describe the data that will be used to solve the problem and the source of the data
   3. Methodology section which represents the main component of the report where you discuss and describe any exploratory data analysis that you did, any inferential statistical testing that you performed, if any, and what machine learnings were used and why.
   4. Results section where you discuss the results.
   5. Discussion section where you discuss any observations you noted and any recommendations you can make based on the results
   6. Conclusion section where you conclude the report
3. Your choice of a presentation or blogpost.

**WEEK1**

1. **A description of the problem and a discussion of the background**

We are a digital start up in the process of launching a new revolutionary app.

Our product = **“Pub Quiz Champions”**

An App that will help you to organize any Pub Quiz like a professional.

Thanks to the apps you can play with your friend in face to face at home or in a pub or you can play virtually from any distance which is quite convenient during this difficult time of COVID lockdown.

But first think first : What is a Pub Quiz ?

**Pub quiz**

*From Wikipedia,*

*A pub quiz is a quiz held in a pub or bar. These events are also called quiz nights[1], trivia nights[2], or bar trivia[3] and may be held in other settings. Pub quizzes may attract customers to a pub who are not found there on other days. The pub quiz is a modern example of a pub game. Although different pub quizzes can cover a range of formats and topics, they have many features in common. The pub quiz was established in the UK in the 1970s by Burns and Porter and became part of British culture.[4] The Great British Pub Quiz challenge is an annual event.[4] In continental Europe, pub quizzes are a staple event at Irish pubs, where they are usually held in English.*



We are now in the last phases of the projects and are preparing our communication plan for the launch of the product.

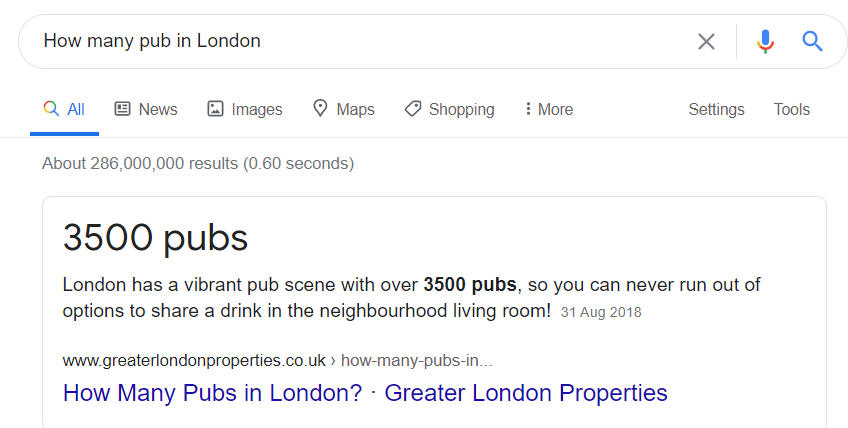
The customer audience we want to target is :

* Fan of quiz , pub aficionados , pub owners
* Our test market is London UK

Our communication plan will be composed of :

* Organization of live events in different Pub where we will demonstrate the added values of our products
* Digital advertising campaigns ( social media + display ads ) that will be geo targeted on the areas where we find the most frequented pub.

**The problem :**



We don’t have a huge communication budget so it’s not possible for us to target the 3500 London pubs.

To make our communication plan a success:

We need to identify and create a selection of London areas where we will start our communication plan.

We want to identify which areas of London are distinguished by the frequency of pub visits.

1. **Data where you describe the data that will be used to solve the problem and the source of the data**

We will use the following data :

1. List of London Borough & their GPS coordinates : Longitude - Latitude that i have available in an excel format from a previous project.
2. Venue data that will be extracted from Foursquare API and will be used for the clustering of the neighbourhood

**Methodologie :**

1. Visualize the Borough of London
   1. Download and read coordinates data
   2. Use **Folium package** to create a map of London with Borough superimposed on top
2. Explore neighborhood
   1. **Foursquare API**
   2. Define Foursquare credentials and version
   3. Explore the neighbourhood in our dataframe and extract the top 50 venues in a 5000 radius.
   4. Convert the venues list in a new dataframe
   5. Add the venues to the london map
3. Analyse neighborhood
   1. Proceed a one hot encoding
   2. Group by rows
   3. Define top venues
4. Cluster the neighborhood
   1. Create the clusters with with **k-means**
   2. Visualize the clusters by adding them to the map
5. Examine cluster & conclusion
   1. Identify the cluster with “pub” as most common venues

---------------------------------------------------------------------------------------------------------------------------