

APPLIED DATASCIENCE CAPSTONE

FINAL ASSIGNMENT REPORT

Business problem

For this capstone project, I came up with an idea to leverage the Foursquare location data to explore the city of Paris, where I live. Paris is a big city and there is so much things to do and to visit that tourists maybe lost at first sight. Where to stay ? What are the most famous places to visit ? Where are the best restaurants to eat ? These are the questions most people ask when arriving in Paris. To answer these questions, we have to know precisely what people are looking for : historical places, trending places, shopping malls,... so that we can recommend the best borough to go depending on what they search. As you guess, the target here is the tourists visiting Paris. In 2018, they were about 50 millions.

So, here is the problem I would like to solve :

If your friends are coming to Paris, which places would you recommend them to visit ? What are the most common venues in each of the 20 boroughs of Paris ?

Answering this problem would allow people to gain time and make sure that they do not miss famous and interesting places in Paris !

Data

To solve this problem, we will need different kinds of data.

First, we will rely on a dataset present on a Wikipedia page (https://fr.wikipedia.org/wiki/Liste_des_quartiers_administratifs_de_Paris). In this HTML page, we can find a table containing 5 columns : Borough, Neighbourhood, Population, Area and Density. Here, we will only focus on two of the five columns : all the neighbourhoods of Paris and their corresponding borough. We will work on the data to group the neighbourhoods by borough so that we can easily find the most common venues in each borough. For example, the borough « 1^{er} arrondissement » contains the following neighbourhoods : Palais-Royal, Place Vendôme, Halles, Saint Germain l'Auxerrois.

In combination to the data mentioned above, we will use Foursquare location data as it will allow us to explore the city of Paris and access the venues located in the different boroughs. Foursquare is important because it will help us to find the trending places of the capital and all kinds of venues (restaurants, cultural places, hotels,...) that we will then recommend to the tourists.