

CAPSTONE PROJECT :

PARIS VS NEW YORK

Introduction

In his *Paris vs New York*, Vahram Muratyan tries to expose a comparison between two diverse and international cities via a travel journal. In a particular section he tries to match the neighborhoods of the two cities. They are both financial and cultural capitals of their respective countries but offers some contrast in their organization and outdoor environment.

We are going to compare two major cities through their neighborhoods and see how close they are. Our goal will be to gather similar areas in the different cities to be able to cluster neighborhoods from these two. Therefore, we would be able to make a content-based recommendation and back (or not) the observations of Vahram Muratyan.

TARGET AUDIENCE

Everyone who wants to visit a new city as a tourist tends to compare based on the areas he knows. It could also be a tool for travel agency to recommend areas based on what you liked in the other cities or use appropriate comparison. This could also be seen as a tactic that applies to findings the best places to live. Do it before you start planning your move or to help narrow down your choices.

DATA SOURCE AND EXTRACTION METHODS

When you mention Paris as a city you only take the metropolitan area into account. This is where you have the biggest concentration of venues, monuments, offices.

Hence in this scenario we are only going to compare the metropolitan area of Paris with the county of Manhattan to have an appropriate area of comparison. In an extension of this study you could also define a broader area.

The zip code of Paris are available on the web. It can be retrieved easily from Wikipedia website like NYC (https://en.wikipedia.org/wiki/Arrondissements_of_Paris).

Also, we will use the Foursquare API to explore neighborhoods in Manhattan and Paris. You will use the ****explore**** function to get the most common venue categories in each neighborhood, number of trending venues and then use this feature to group the neighborhoods from both cities into clusters.

METHODOLOGY

Here is the following Method we are going to follow to analyse our data

- Retrieve ZIP code from each area
- Determine localization
- Use the Foursquare API to retrieve relevant data (Venues category, trending places)
- Use the K-Mean clustering method
- Analyse of the results and Conclusion

