**Report: The battle of Neighborhoods**

**Introduction**

London is a multi-cultural city, characterized by several neighborhoods, each one with an history and populated by different ethnic groups. Among the others, the Italian population is one of the most numerous, and the Italian cuisine is one of the most appreciated.

In this project, we would like to understand, with a data-driven approach, what neighborhoods offers the highest business opportunities if a new Italian restaurant is opened. Specifically we would like to answer the following question: if someone is looking to open an italian restaurant, where would they open it?

In order to answer this question, we perform a segmentation and clusterization of the city based on its neighborhoods. This allows us to analyse the distribution of Italian restaurants in London, and thus to obtain the areas where this business is still in a preliminary stage.

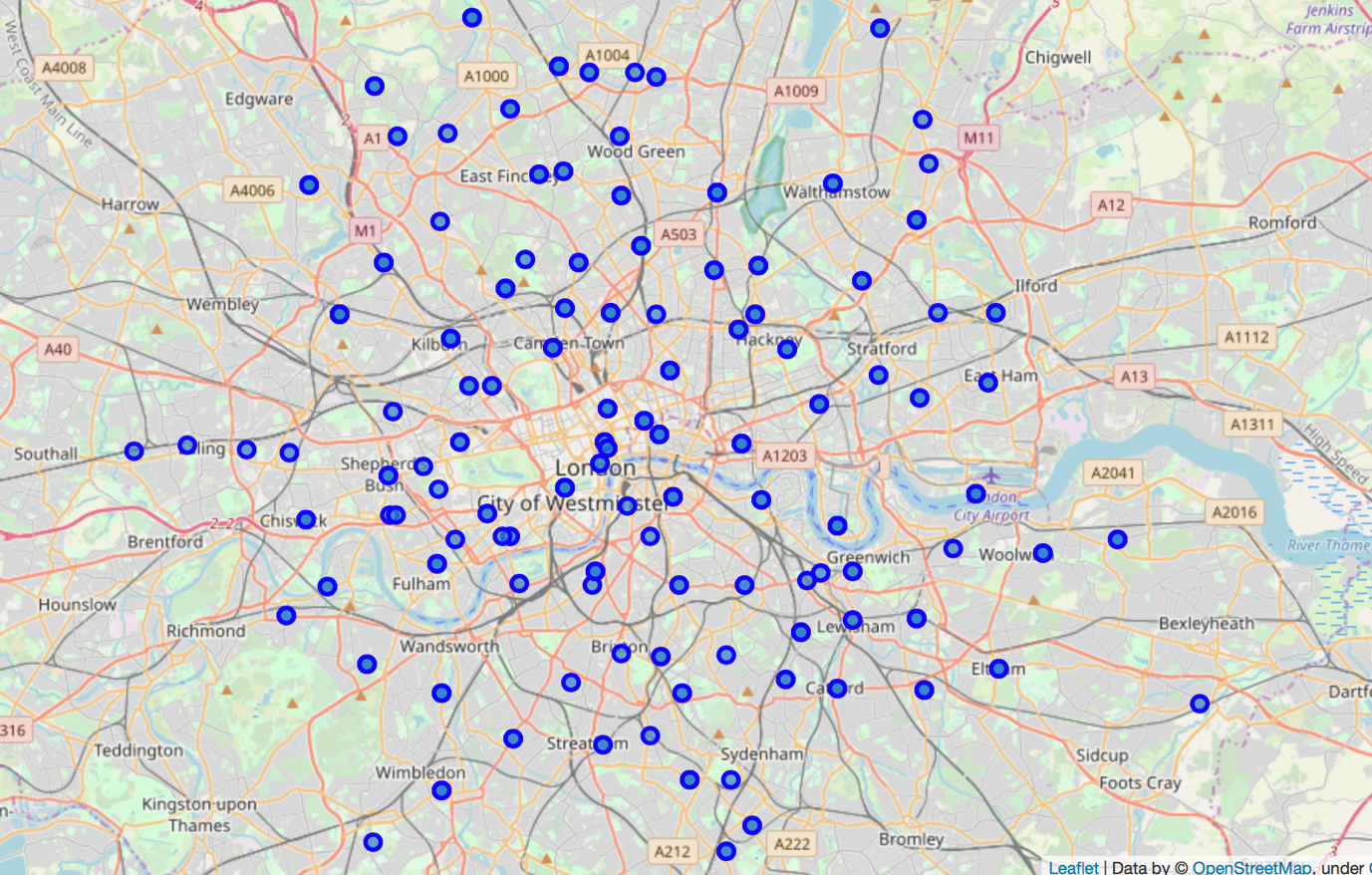
**Data**

In order to perform the analysis, we need to extract the data on London neighborhoods by using public available data from Wikipedia. Moreover, we will employ Foursquare to obtain data on the venues in each neighborhoods.

London is administered by the City of London and 32 London boroughs. Data on the boroughs can be found on Wikipedia at the following link:

<https://en.wikipedia.org/wiki/List_of_areas_of_London>

By using BeautifulSoup, we can extract info from this page (scraping) and we can look for the table class.

****

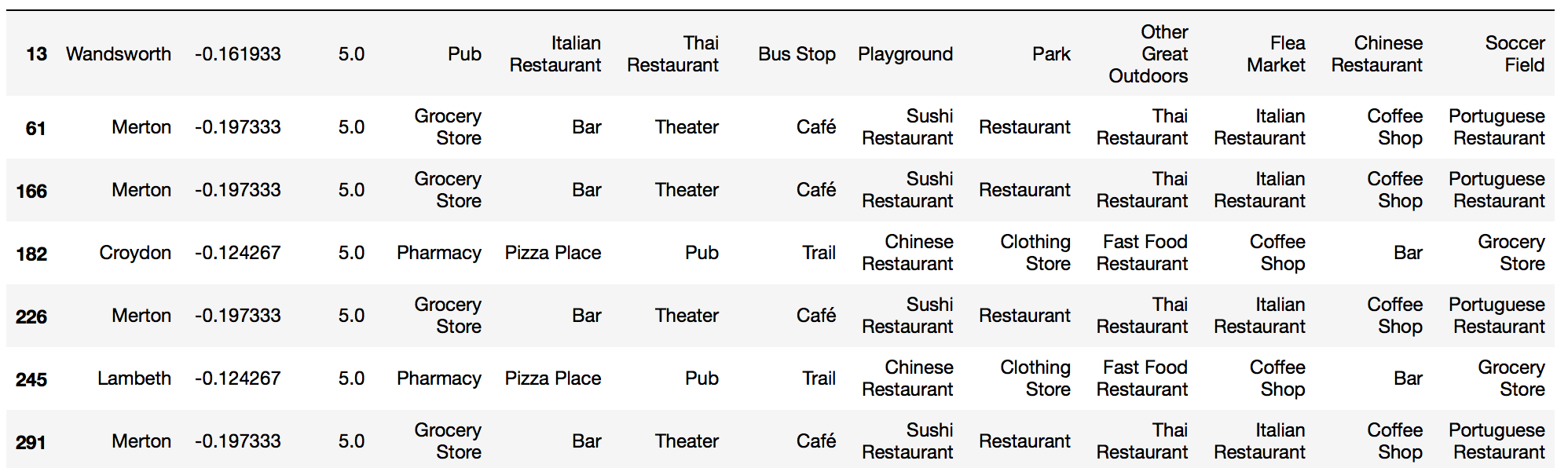
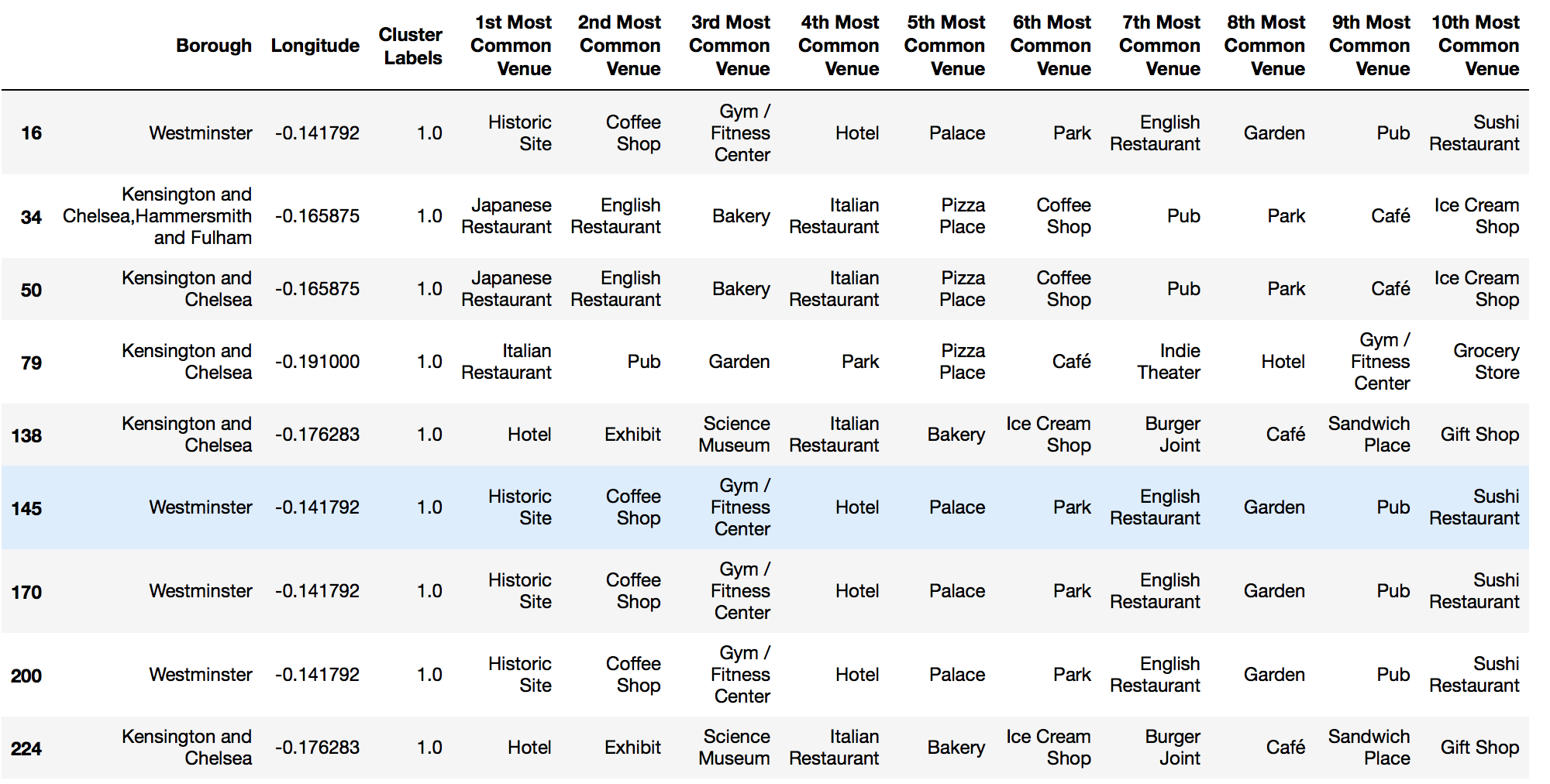
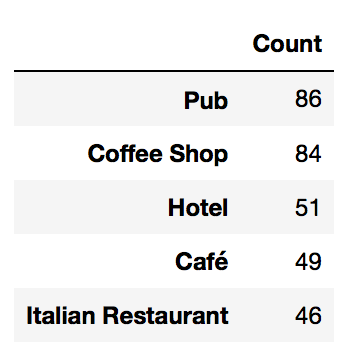
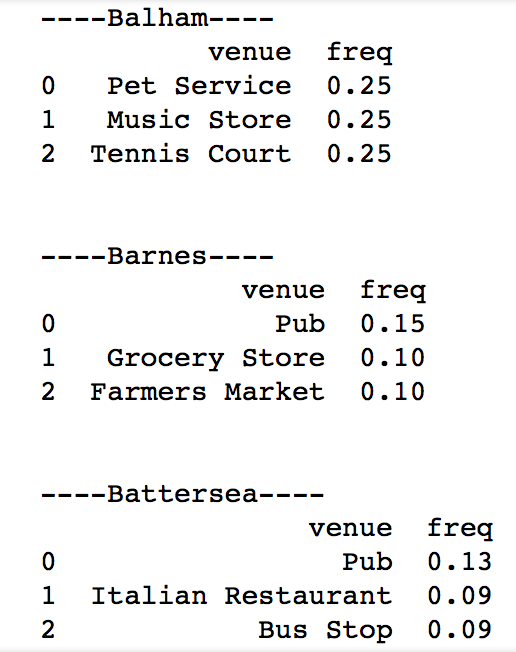
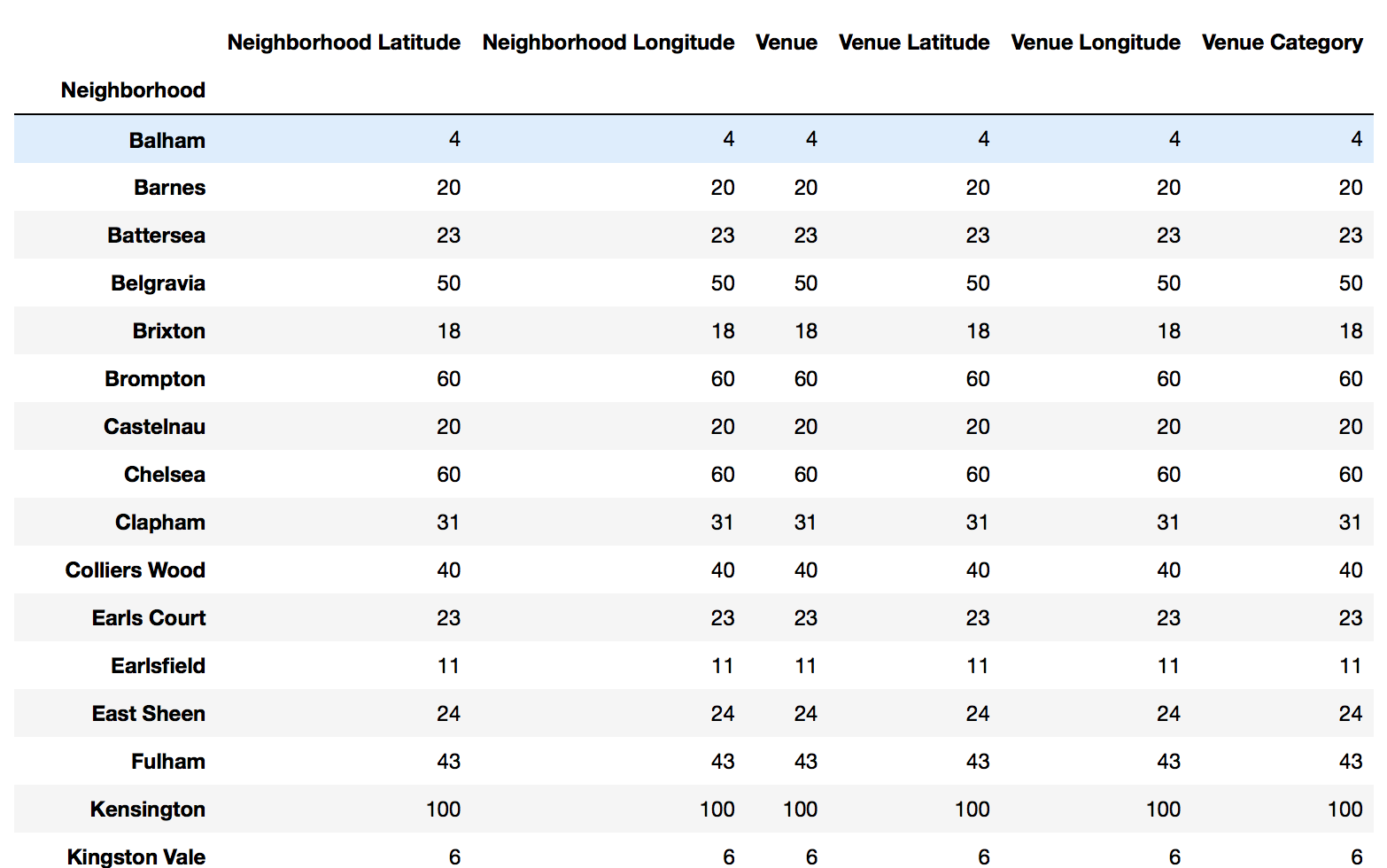
**Methodology**

In order to study the venues distribution in London, we emply the following techniques:

1. we compute The total number of business in each Neighborhood
2. for every neighbourhood, we check what are the most important venue categories
3. we perform a clasterization of the South-West neighbourhood of London

**Results and discussion**

From the clustering results, we could detect three main clusters with different characteristics

****

**Conclusion**

Our analysis can be used to evaluate any opportunity present in the South-West London.