



# Amsterdam v/s New York

SEGMENTATION AND CLUSTERING  
OF  
DIFFERENT VENUES IN BOTH  
CITIES

Kumar Shivam | Capstone Project | 20/10/2019

# □ Introduction

- Few cities in Europe pair style and dilapidation like Amsterdam. With a downright raunchy [red light district](#) on the one hand and some of the world's very best museums on the other, cosmopolitan Amsterdam takes its contradictions in stride.
- New York City, officially the City of New York, historically New Amsterdam, the Mayor, Alderman, and Commonality of the City of New York, and New Orange, byname the Big Apple, city and port located at the mouth of the [Hudson River](#), southeastern [New York](#) state, northeastern U.S. It is the largest and most influential American metropolis, [encompassing](#) Manhattan and Staten islands, the western sections of [Long Island](#), and a small portion of the New York state mainland to the north of Manhattan
- This project is comparison of different venues in Amsterdam and New York city. The most popular venues in both cities according to the ratings.
- Using Foursquare API we can find the popular venues in both the cities. Clustering is applied on the venues of each city.

## ❑ Source of data

- We will download the dataset from Kaggle and perform some data cleaning of the data. Both the datasets are available on Kaggle.
- To download the Dataset for Amsterdam:
- <https://www.kaggle.com/dushyantsapre/amsterdam-neighbourhood-and-venues>
- To download the dataset for New York city:
- <https://www.kaggle.com/dgomonov/new-york-city-airbnb-open-data>

# □ Data Cleaning

Data Obtained from different sources needs to be cleaned. The New York data need to be cleaned as it contains unnecessary columns that needs to be deleted.

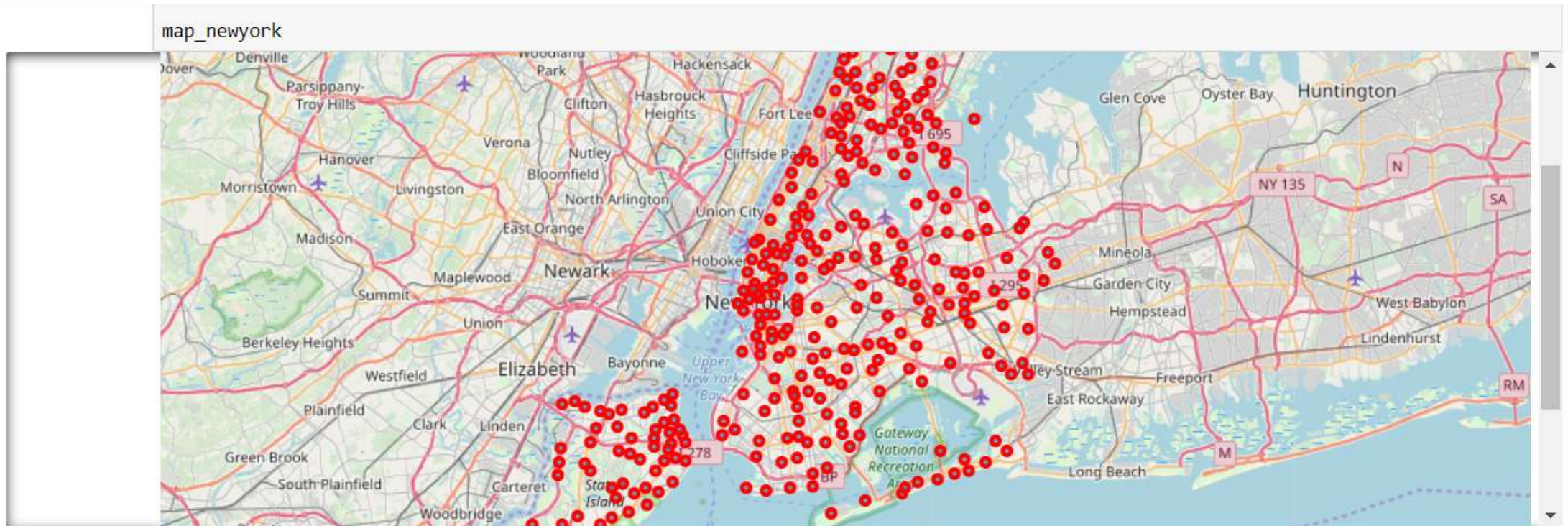
The New York data is obtained from the json file and the data frame is created from the json file and obtained data frame needs to be cleaned

```
In [11]: new_york_data
```

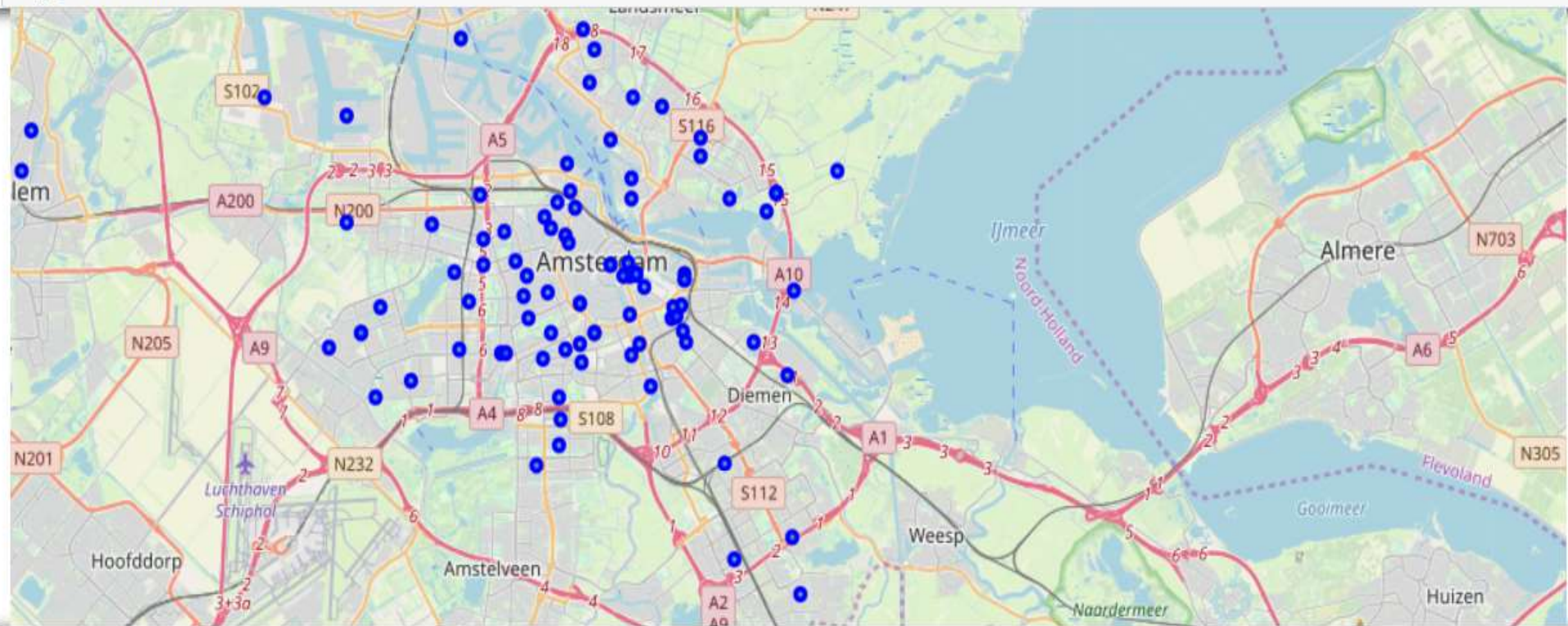
```
Out[11]: {'type': 'FeatureCollection',  
  'totalFeatures': 306,  
  'features': [{'type': 'Feature',  
    'id': 'nyu_2451_34572.1',  
    'geometry': {'type': 'Point',  
      'coordinates': [-73.84720052054902, 40.89470517661]},  
    'geometry_name': 'geom',  
    'properties': {'name': 'Wakefield',  
      'stacked': 1,  
      'annoline1': 'Wakefield',  
      'annoline2': None,  
      'annoline3': None,  
      'annoangle': 0.0,  
      'borough': 'Bronx',  
      'bbox': [-73.84720052054902,  
        40.89470517661,  
        -73.84720052054902,  
        40.89470517661]}},  
    {'type': 'Feature',  
      'id': 'nyu_2451_34572.2'
```



# Different venues are marked in the map using Folium



```
map_amsterdam
```





# □ Result

map\_clusters\_ams

out[42]:

