# Impact of the Coronavirus in the Tourism and Relationship of visiting Museum, Restaurants and Coffees shops

Case of Study: London UK.

By Antonio BOSNJAK SEMINARIO

December 1, 2020

# REPORT

## 1) The problem and Discussion of the Background.

In the United Kingdom, the confirmed cases of coronavirus is 1,493,300, according to Public Health England. As of Saturday afternoon (November 21, 2020), 54,626 people had died. Due to a data entry error, nearly 16,000 people who tested positive between Sept. 25 and Oct. 2 were not recorded in the daily number of reported cases.

Everyone knows about the increase in Coronavirus cases. However, we are interested in finding out, which is the risk of contagion?, that is, there is really a greater risk of contagion if I dedicate to the tourism activities, such as visiting or working in a museum. Or in the other hand, I am frequently visit: Pubs, Coffee shop, Café, Restaurants, Supermarkets, or many other amenities as before the Coronavirus lockdown. What happened: If, Am I a owner of the Pub, Restaurant or a coffee shop?.

## Hypothesis:

Our hypothesis is the following: This virus (Covid-19) is affecting much more in highly populated regions, and very dense areas. We will analyze the neighborhoods of London and specifically where the main London Museums are located, in order to determine if someone who lives or travels to this area is more likely to contract this disease.

***The main question is:* Which are the most common business, or commerce located in the London Neighborhood that causes the new cases Coronavirus (Covid-19) to increase, and a high rate of contagium?.**

Macro studies have been carried out with data from all over the country, as can be seen in figure 1. World data has also been counted. However, there are a very few micro analysis of the neighborhoods of a City, to know if there are some areas where the contagion has decreased. And it is towards that micro-analysis of the city of London and its neighbourhoods' that we focus.

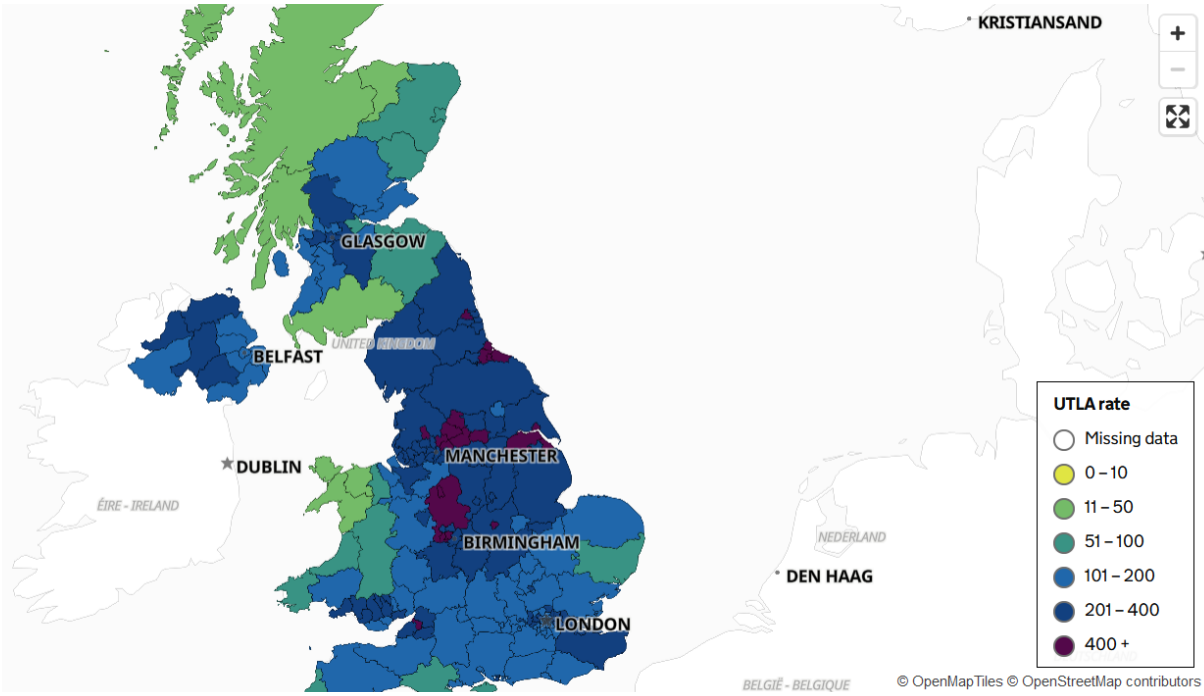


Figure 1. Map of the United Kingdom showing the areas most affected by Covid-19. The areas in dark blue are between 201 to 400 cases per 100,000 inhabitants, and those in dark red correspond to 400 cases per 100,000 inhabitants by October 2020.

## 2) Data Acquisition and Cleaning.

### 2.1) Data Acquisition:

In this case the data is distributed by four (4) Web sites:

1) For example, In Wikipedia we find the list of the most visited museums in the United Kingdom during the year 2017. This list can be found at the following link: https://en.wikipedia.org/wiki/List_of_most_visited_museums_in_the_United_Kingdom. In figure 2. One part of this list is shown.

# List of most visited museums in the United Kingdom

From Wikipedia, the free encyclopedia

This article **lists the most visited museums in the United Kingdom** (including art galleries). The list of 40 is based on the 2017 of the Association of Leading Visitor Attractions unless otherwise noted.[1]

| Rank | Museum | Location | Country | Visitor count |
|---|---|---|---|---|
| 1 | British Museum | London | England | 5,906,716 |
| 2 | Tate Modern | London | England | 5,656,004 |
| 3 | National Gallery | London | England | 5,229,192 |
| 4 | Natural History Museum | London | England | 4,434,520 |
| 5 | Victoria and Albert Museum | London | England | 3,789,748 |
| 6 | Science Museum | London | England | 3,251,000 |
| 7 | Royal Museums Greenwich | London | England | 2,607,099 |
| 8 | National Museum of Scotland | Edinburgh | Scotland | 2,165,601 |
| 9 | Tate Britain | London | England | 1,777,877 |
| 10 | Scottish National Gallery | Edinburgh | Scotland | 1,600,761 |
| 11 | Riverside Museum | Glasgow | Scotland | 1,355,359 |
| 12 | Kelvingrove Art Gallery & Museum | Glasgow | Scotland | 1,304,072 |
| 13 | National Portrait Gallery | London | England | 1,271,920 |
| 14 | National Museum of the Royal Navy | Portsmouth | England | 1,081,909 |
| 15 | Imperial War Museum | London | England | 992,690 |
| 16 | Horniman Museum and Gardens | London | England | 942,971 |
| 17 | Ashmolean Museum | Oxford | England | 937,568 |

Figure 2. List of most visited museums in the United Kingdom in 2017.

2) A list with the names of the neighborhoods and their GSS_Code, plus their location in latitude and longitude. This list was found on Wikipedia and the link where it is found is: https://en.wikipedia.org/wiki/List_of_London_boroughs . The important part is to obtain the latitude and longitude for each of London's neighborhoods. The others columns can be discarded. Figure 3 shows that list:

List of boroughs and local authorities [edit]

| Borough | Inner | Status | Local authority | Political control | Headquarters | Area (sq mi) | Population (2013 est)[1] | Co-ordinates | Nr. in map |
|---|---|---|---|---|---|---|---|---|---|
| Barking and Dagenham [note 1] | | | Barking and Dagenham London Borough Council | Labour | Town Hall, 1 Town Square | 13.93 | 194,352 | 51.5607°N 0.1557°E | 25 |
| Barnet | | | Barnet London Borough Council | Conservative | Barnet House, 2 Bristol Avenue, Colindale | 33.49 | 369,088 | 51.6252°N 0.1517°W | 31 |
| Bexley | | | Bexley London Borough Council | Conservative | Civic Offices, 2 Watling Street | 23.38 | 236,687 | 51.4549°N 0.1505°E | 23 |
| Brent | | | Brent London Borough Council | Labour | Brent Civic Centre, Engineers Way | 16.70 | 317,264 | 51.5588°N 0.2817°W | 12 |
| Bromley | | | Bromley London Borough Council | Conservative | Civic Centre, Stockwell Close | 57.97 | 317,899 | 51.4039°N 0.0198°E | 20 |
| Camden | ✓ | | Camden London Borough Council | Labour | Camden Town Hall, Judd Street | 8.40 | 229,719 | 51.5290°N 0.1255°W | 11 |
| Croydon | | | Croydon London Borough Council | Labour | Bernard Weatherill House, Mint Walk | 33.41 | 372,752 | 51.3714°N 0.0977°W | 19 |
| Ealing | | | Ealing London Borough Council | Labour | Perceval House, 14-16 Uxbridge Road | 21.44 | 342,494 | 51.5130°N 0.3089°W | 13 |
| Enfield | | | Enfield London Borough Council | Labour | Civic Centre, Silver Street | 31.74 | 320,524 | 51.6538°N 0.0799°W | 30 |
| Greenwich [note 2] | ✓ [note 3] | Royal | Greenwich London Borough Council | Labour | Woolwich Town Hall, Wellington Street | 18.28 | 264,008 | 51.4892°N 0.0648°E | 22 |
| Hackney | ✓ | | Hackney London Borough Council | Labour | Hackney Town Hall, Mare Street | 7.36 | 257,379 | 51.5450°N 0.0553°W | 9 |
| Hammersmith and Fulham [note 4] | ✓ | | Hammersmith and Fulham London Borough Council | Labour | Town Hall, King Street | 6.33 | 178,685 | 51.4927°N 0.2339°W | 4 |

earch Desktop 10:17 PM 11/22/2020

Figure 3. List of names and Neighborhood of London and their latitude and longitude.

The Second source of data is scraped from a Wikipedia page that contains the List of London borough. This page contains additional information about the boroughs, the following are the columns:

- **Borough**: The names of the 33 London boroughs.
- **Inner**: Categorizing the borough as an inner London borough or an Outer London Borough.
- **Status**: Categorizing the borough as Royal, City or other borough.
- **Local Authority**: The local authority assigned to the borough.
- **Political Control**: The political party that control the borough.
- **Headquarters**: Headquarters of the Borough.
- **Area (sq mi)**: Area of the borough in square miles.
- **Population (2013 est)[1]**: The population of the borough recorded during the year 2013.
- **Co-ordinates**: The latitude and longitude of the Boroughs.
- **Nr. In map**: The number assigned to each borough to represent visually on a map.

3) We have a database with data from each of the London neighborhoods, their population, the number of coronavirus cases for the first week of October for each neighborhood; the number of coronavirus cases for the second week of October and its rate of increase per 100,000 inhabitants. This data was obtained from the UK government Web page using the following link: [https://www.gov.uk/government/publications/lower-tier-local-authority-watchlist-epidemiological-data-23-october-2020/case-rates-by-region](https://www.gov.uk/government/publications/lower-tier-local-authority-watchlist-epidemiological-data-23-october-2020/case-rates-by-region) . Figure 4 shows the .CSV file that I will analyse:

| | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Case numbers and rates for previous two weeks by local authority | | | | | | | | | |
| 2 | England | | | | | | | | | |
| 3 | | | | | | | | | | |
| 4 | Local Authority | Total population | Oct to 14 Oct 2020 | ct to 07 Oct 2020 | ct to 14 Oct 2020 | ct to 07 Oct 2020 | Rank | recent week | category [for map] | |
| 5 | Barking and Dagenham | 212906 | 236 | 160 | 110.8 | 75.2 | 14 | 35.6 | 100-199.9/100k | |
| 6 | Barnet | 395869 | 435 | 424 | 109.9 | 107.1 | 15 | 2.8 | 100-199.9/100k | |
| 7 | Bexley | 248287 | 183 | 165 | 73.7 | 66.5 | 31 | 7.2 | 50-99.9/100k | |
| 8 | Brent | 329771 | 319 | 299 | 96.7 | 90.7 | 23 | 6 | 50-99.9/100k | |
| 9 | Bromley | 332336 | 268 | 222 | 80.6 | 66.8 | 29 | 13.8 | 50-99.9/100k | |
| 10 | Camden | 270029 | 310 | 242 | 114.8 | 89.6 | 11 | 25.2 | 100-199.9/100k | |
| 11 | City of London | 9721 | 3 | 5 | 30.9 | 51.4 | 33 | -20.5 | 25-49.9/100k | |
| 12 | Croydon | 386710 | 295 | 270 | 76.3 | 69.8 | 30 | 6.5 | 50-99.9/100k | |
| 13 | Ealing | 341806 | 491 | 449 | 143.6 | 131.4 | 1 | 12.2 | 100-199.9/100k | |
| 14 | Enfield | 333794 | 334 | 269 | 100.1 | 80.6 | 21 | 19.5 | 100-199.9/100k | |
| 15 | Greenwich | 287942 | 212 | 206 | 73.6 | 71.5 | 32 | 2.1 | 50-99.9/100k | |
| 16 | Hackney | 281120 | 357 | 396 | 127 | 140.9 | 5 | -13.9 | 100-199.9/100k | |
| 17 | Hammersmith and Fulham | 185143 | 242 | 178 | 130.7 | 96.1 | 3 | 34.6 | 100-199.9/100k | |
| 18 | Haringey | 268647 | 330 | 271 | 122.8 | 100.9 | 6 | 21.9 | 100-199.9/100k | |
| 19 | Harrow | 251160 | 294 | 304 | 117.1 | 121 | 9 | -3.9 | 100-199.9/100k | |
| 20 | Havering | 259552 | 298 | 206 | 114.8 | 79.4 | 12 | 35.4 | 100-199.9/100k | |
| 21 | Hillingdon | 306870 | 353 | 303 | 115 | 98.7 | 10 | 16.3 | 100-199.9/100k | |
| 22 | Hounslow | 271523 | 319 | 276 | 117.5 | 101.6 | 8 | 15.9 | 100-199.9/100k | |
| 23 | Islington | 242467 | 261 | 190 | 107.6 | 78.4 | 16 | 29.2 | 100-199.9/100k | |
| 24 | Kensington and Chelsea | 156129 | 163 | 152 | 104.4 | 97.4 | 19 | 7 | 100-199.9/100k | |
| 25 | Kingston upon Thames | 177507 | 209 | 178 | 117.7 | 100.3 | 7 | 17.4 | 100-199.9/100k | |
| 26 | Lambeth | 326034 | 329 | 299 | 100.9 | 91.7 | 20 | 9.2 | 100-199.9/100k | |
| 27 | Lewisham | 305842 | 247 | 228 | 80.8 | 74.5 | 28 | 6.3 | 50-99.9/100k | |
| 28 | Merton | 206548 | 172 | 139 | 83.3 | 67.3 | 26 | 16 | 50-99.9/100k | |
| 29 | Newham | 353134 | 404 | 329 | 114.4 | 93.2 | 13 | 21.2 | 100-199.9/100k | |

Coronavirus_Cases_Neighborhood_ | North_East | North_West | South_East | South_West | East_of_England | East_M ...

Figure 4. Database with new cases of coronavirus in neighborhoods near the city of London. The name of the Neighborhood, the total population on each Neighborhood, Number of infections in the second week of October, etc. are shown.

4) Additionally we have managed to compile the map of the London neighborhoods excluding MHW. Figure 5 shows the London neighborhood map collected from the UK Government website and drawn with the QGIS program.

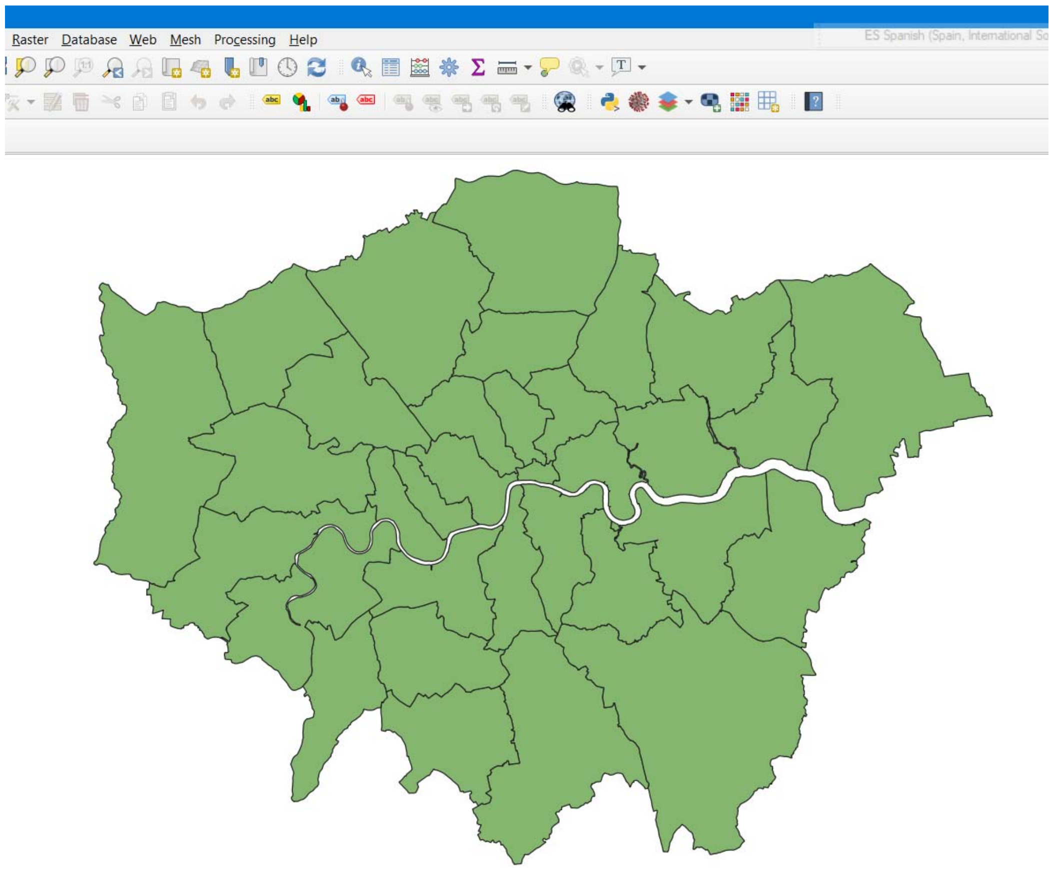


Figura 5. Mapa de los vecindarios de Londres y sus alrededores, dibujado con el Programa QGIS.

## 2.1) Data Cleaning:

The data preparation for each of the three sources of data is done separately. From the United Kingdom Museum data, I selected only the Museums located inside of London and their Boroughs. Figure 6, shows one 'Data Frame' table of 15 Museums located in London.

**Drop a row if it contains a Museum outside of London (in this case, “= London”)**

In [9]:
```
Museum_London = Museum_List[Museum_List.Location == 'London']
Museum_London.head(11)
```

Out[9]:

| | Rank | Museum | Location | Country | Visitor_count |
|---|---|---|---|---|---|
| **0** | 1 | British Museum | London | | 5,906,716 |
| **1** | 2 | Tate Modern | London | | 5,656,004 |
| **2** | 3 | National Gallery | London | | 5,229,192 |
| **3** | 4 | Natural History Museum | London | | 4,434,520 |
| **4** | 5 | Victoria and Albert Museum | London | | 3,789,748 |
| **5** | 6 | Science Museum | London | | 3,251,000 |
| **6** | 7 | Royal Museums Greenwich | London | | 2,607,099 |
| **8** | 9 | Tate Britain | London | | 1,777,877 |
| **12** | 13 | National Portrait Gallery | London | | 1,271,920 |
| **14** | 15 | Imperial War Museum | London | | 992,690 |
| **15** | 16 | Horniman Museum and Gardens | London | | 942,971 |

Figure 6. List of most visited Museums in London.

The second data is scraped from a wikipedia page using the **Beautiful Soup** library in python. Using this library we can extract the data in the tabular format as shown in the website. After the web scraping, string manipulation is required to get the names of the boroughs in the correct form (see figure 7). This is important because we will be merging this dataset with the third dataset contains the data of Coronavirus (Covid-19) using the same Borough names.

```
df_Borough_List.head(7)
```

Out[13]:

| | Borough | Inner | Status | Local_Authority | Political_Control | Headquarters | Area | Population_2013 | Coordinates | NrMap |
|---|---|---|---|---|---|---|---|---|---|---|
| **0** | Barking and Dagenham | | | Barking and Dagenham London Borough Council | Labour | Town Hall | 13.93 | 194,352 | 51.5607°N 0.1557°E | 25 |
| **1** | Barnet | | | Barnet London Borough Council | Conservative | Barnet House | 33.49 | 369,088 | 51.6252°N 0.1517°W | 31 |
| **2** | Bexley | | | Bexley London Borough Council | Conservative | Civic Offices | 23.38 | 236,687 | 51.4549°N 0.1505°E | 23 |
| **3** | Brent | | | Brent London Borough Council | Labour | Brent Civic Centre | 16.70 | 317,264 | 51.5588°N 0.2817°W | 12 |
| **4** | Bromley | | | Bromley London Borough Council | Conservative | Civic Centre | 57.97 | 317,899 | 51.4039°N 0.0198°E | 20 |
| **5** | Camden | Y | | Camden London Borough Council | Labour | Camden Town Hall | 8.40 | 229,719 | 51.5290°N 0.1255°W | 11 |
| **6** | Croydon | | | Croydon London Borough Council | Labour | Bernard Weatherill House | 33.41 | 372,752 | 51.3714°N 0.0977°W | 19 |

Figure 7. List of London Boroughs.

On this dataset the problem is to extract the Latitude and Longitude from the column “Coordinates” because their codification is “latºN longºW or ºE”. It is very important to obtain two columns: one for “Latitude” and another one for “Longitude”. A procedure of cleaning data and find only numbers for Latitude and Longitude was made. Figure 8, shows the following Python code:

```
In [14]: new = df_Borough_List["Coordinates"].str.split("N", n =1, expand=True)
df_Borough_List["Latitude"] = new[0]
df_Borough_List["Longitude"] = new[1]
# Dropping old Name columns
# df_Borough_List.drop(columns = ["Coordinates"], inplace=True)
df_Borough_ListA = df_Borough_List.replace('°', '', regex = True)
df_Borough_ListA["Latitude"] = df_Borough_ListA["Latitude"].astype(float)
df_Borough_ListA["Indexes"] = df_Borough_ListA["Longitude"].str.find('W')
df_Borough_ListA["Longitude"] = df_Borough_ListA["Longitude"].replace('E', '', regex = True)
df_Borough_ListA["Longitude"] = df_Borough_ListA["Longitude"].replace('W', '', regex = True)
df_Borough_ListA["Longitude"] = df_Borough_ListA["Longitude"].astype(float)
num = df_Borough_ListA["Indexes"]._get_numeric_data()
Longitude = df_Borough_ListA["Longitude"]._get_numeric_data()
j = -1
for i in num:
    j = j + 1
    if (i == 7):
        df_Borough_ListA["Longitude"] = df_Borough_ListA["Longitude"].replace(Longitude[j], -Longitude[j])

df_Borough_ListA.drop(columns = ["Inner", "Status", "Local_Authority", "Coordinates"], inplace=True)
df_Borough_ListA.head(10)
```

Out[14]:

| | Borough | Political_Control | Headquarters | Area | Population_2013 | NrMap | Latitude | Longitude | Indexes |
|---|---|---|---|---|---|---|---|---|---|
| 0 | Barking and Dagenham | Labour | Town Hall | 13.93 | 194,352 | 25 | 51.5607 | 0.1557 | -1 |
| 1 | Barnet | Conservative | Barnet House | 33.49 | 369,088 | 31 | 51.6252 | -0.1517 | 7 |
| 2 | Bexley | Conservative | Civic Offices | 23.38 | 236,687 | 23 | 51.4549 | 0.1505 | -1 |
| 3 | Brent | Labour | Brent Civic Centre | 16.70 | 317,264 | 12 | 51.5588 | -0.2817 | 7 |
| 4 | Bromley | Conservative | Civic Centre | 57.97 | 317,899 | 20 | 51.4039 | 0.0198 | -1 |
| 5 | Camden | Labour | Camden Town Hall | 8.40 | 229,719 | 11 | 51.5290 | -0.1255 | 7 |

Figure 8. Cleaning data of the List of London Boroughs. The objective is to obtain the numbers (float64) for Latitude and Longitude of each Borough.

The third source of data is acquired from the UK government. It has the number of new cases of Coronavirus (Covid-19) in each Borough of London area. This dataset has very long names for each column. They try to auto-explained each column, However for our research is better to use shorter names. Figure 9 depicts a list of new Coronavirus cases.

**Change the names of the headers with the purpose of processing of this DataFrame.**

```
[14]: headers = ["Borough", "Population", "Case2ndWeek", "Case1stWeek", "Rate2ndWeek", "Rate1stWeek", "Rar
print("headers\n", headers) # Change the header name and assigned into the columns
Coronavirus_df.columns = headers
Coronavirus_df.head(10)

headers
 ['Borough', 'Population', 'Case2ndWeek', 'Case1stWeek', 'Rate2ndWeek', 'Rate1stWeek', 'Rank', 'Ra
```

Out[14]:

| | Borough | Population | Case2ndWeek | Case1stWeek | Rate2ndWeek | Rate1stWeek | Rank | Rate_Change | Rate_Category |
|---|---|---|---|---|---|---|---|---|---|
| 0 | Barking and Dagenham | 212906 | 236 | 160 | 110.8 | 75.2 | 14 | 35.6 | 100-199.9/100k |
| 1 | Barnet | 395869 | 435 | 424 | 109.9 | 107.1 | 15 | 2.8 | 100-199.9/100k |
| 2 | Bexley | 248287 | 183 | 165 | 73.7 | 66.5 | 31 | 7.2 | 50-99.9/100k |
| 3 | Brent | 329771 | 319 | 299 | 96.7 | 90.7 | 23 | 6.0 | 50-99.9/100k |
| 4 | Bromley | 332336 | 268 | 222 | 80.6 | 66.8 | 29 | 13.8 | 50-99.9/100k |
| 5 | Camden | 270029 | 310 | 242 | 114.8 | 89.6 | 11 | 25.2 | 100-199.9/100k |
| 6 | City of London | 9721 | 3 | 5 | 30.9 | 51.4 | 33 | -20.5 | 25-49.9/100k |
| 7 | Croydon | 386710 | 295 | 270 | 76.3 | 69.8 | 30 | 6.5 | 50-99.9/100k |
| 8 | Ealing | 341806 | 491 | 449 | 143.6 | 131.4 | 1 | 12.2 | 100-199.9/100k |
| 9 | Enfield | 333794 | 334 | 269 | 100.1 | 80.6 | 21 | 19.5 | 100-199.9/100k |

Figure 9. Second wave of Coronavirus, from 01 October 2020 until 14 October 2020, for each Borough and computing the rates of the increment of Coronavirus cases.

The third source of data codified by a new headers for each column:

- **Borough**: The names of the 33 London boroughs.
- **Population**: A population for each Borough by 2019.
- **Case2ndWeek**: A number of new Cases of Coronavirus; Week from 08 October until 14 October 2020.
- **Case1stWeek:** A number of new Cases of Coronavirus; Week from 01 October until 07 October 2020.
- **Rate2ndWeek**: Case rate per 100.000 inhabitants. Recent week from 08 October until 14 October 2020.
- **Rate1stWeek**: Case rate per 100.000 inhabitants. Previous week from 01 October until 07 October 2020..
- **Rank**: Rank.
- **Rate_Change**: Rate change between previous and recent week.
- **Rate_Category**: Categorizing of each Borough using for graphics of UK government.

# METODOLOGY

As we did with the city of New York and the City of Toronto we will use the Foursquare application to obtain the missing data and solve this Problem. Foursquare will allow me the link between the databases of point (1) List of the names of the names of the most visited museums with the databases (2) and (3). Thus, we can determine the location of the museum on the map, the closest Pub, Coffee Shops, Café, Restaurants, Supermarkets, Clothing store, Fast food, and other stores and to compare them with the increases in cases of Coronavirus.

Many of the economic activities of the year 2019, as is the case of Tourism, have fallen drastically. However, a scientific study has not been done to determine if the lockdown really prevents contagion or on the contrary we are lengthening the problems by causing the economy to end up falling below the global economic recession observed in the 1930s.

I merged these three database in only one dataset, and using the Latitude and Longitude of each Borough, with the purpose to graphics using “folium” the Case rate per 100.000 inhabitants for the second week from 08 October until 14 October 2020, that correspond the beginning of the second wave of Coronavirus. Figure 10 shows the map of London using the dataset of Covid-19, if the circle is bigger the rate of Coronavirus contagium is highest.

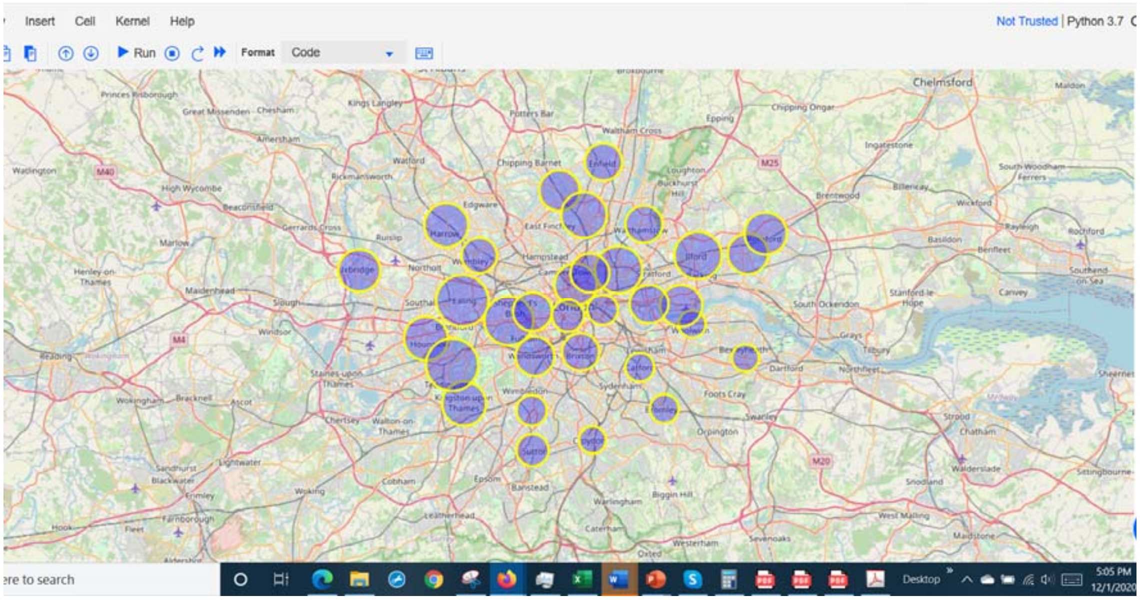


Figure 10. Visualization of the number of rate cases per 100.000 inhabitants on each Borough. If the circle is bigger the risk of contagium is highest.

With the objective to find which Borough and Neighborhood has the highest risk of Contagium of Coronavirus, we select all the Borough with Rate of Coronavirus cases per 100.000 inhabitants in the second week of October are more than 117.5. (A list of seven (7) Borough are shown in figure 11).

**However, for illustration purposes, let's simplify the above map and segment and cluster only the with a high rate of increment of Co more than 117.5 cases by 100000 inhabitants. So let's slice the original dataframe and create a new dataframe of the Westminster.**

In [22]:
```
Westminster_data = Coronavirus_Merge_df[Coronavirus_Merge_df['Rate2ndWeek'] > 117.5].reset_index(drop=True)
Westminster_data
```

Out[22]:

| | Borough | Population | Case2ndWeek | Case1stWeek | Rate2ndWeek | Rate1stWeek | Rate_Category | Latitude | Longitude |
|---|---|---|---|---|---|---|---|---|---|
| 0 | Ealing | 341806 | 491 | 449 | 143.6 | 131.4 | 100-199.9/100k | 51.5130 | -0.3089 |
| 1 | Hackney | 281120 | 357 | 396 | 127.0 | 140.9 | 100-199.9/100k | 51.5450 | -0.0553 |
| 2 | Hammersmith and Fulham | 185143 | 242 | 178 | 130.7 | 96.1 | 100-199.9/100k | 51.4927 | -0.2339 |
| 3 | Haringey | 268647 | 330 | 271 | 122.8 | 100.9 | 100-199.9/100k | 51.6000 | -0.1119 |
| 4 | Kingston upon Thames | 177507 | 209 | 178 | 117.7 | 100.3 | 100-199.9/100k | 51.4085 | -0.3064 |
| 5 | Redbridge | 305222 | 388 | 377 | 127.1 | 123.5 | 100-199.9/100k | 51.5590 | 0.0741 |
| 6 | Richmond upon Thames | 198019 | 282 | 286 | 142.4 | 144.4 | 100-199.9/100k | 51.4479 | -0.3260 |

Figure 11. List of the Borough with highest risk of Contagium of Coronavirus per 100.000 inhabitants.

# RESULTS:

One hot encoding is done on the venues data (One hot encoding is a process by which categorical variables are converted into a form that could be provided to ML algorithms to do a better job in prediction). The Venues data is then grouped by the Neighborhood and the mean of the venues are calculated, finally the 10 common venues are calculated for each of the neighborhood.

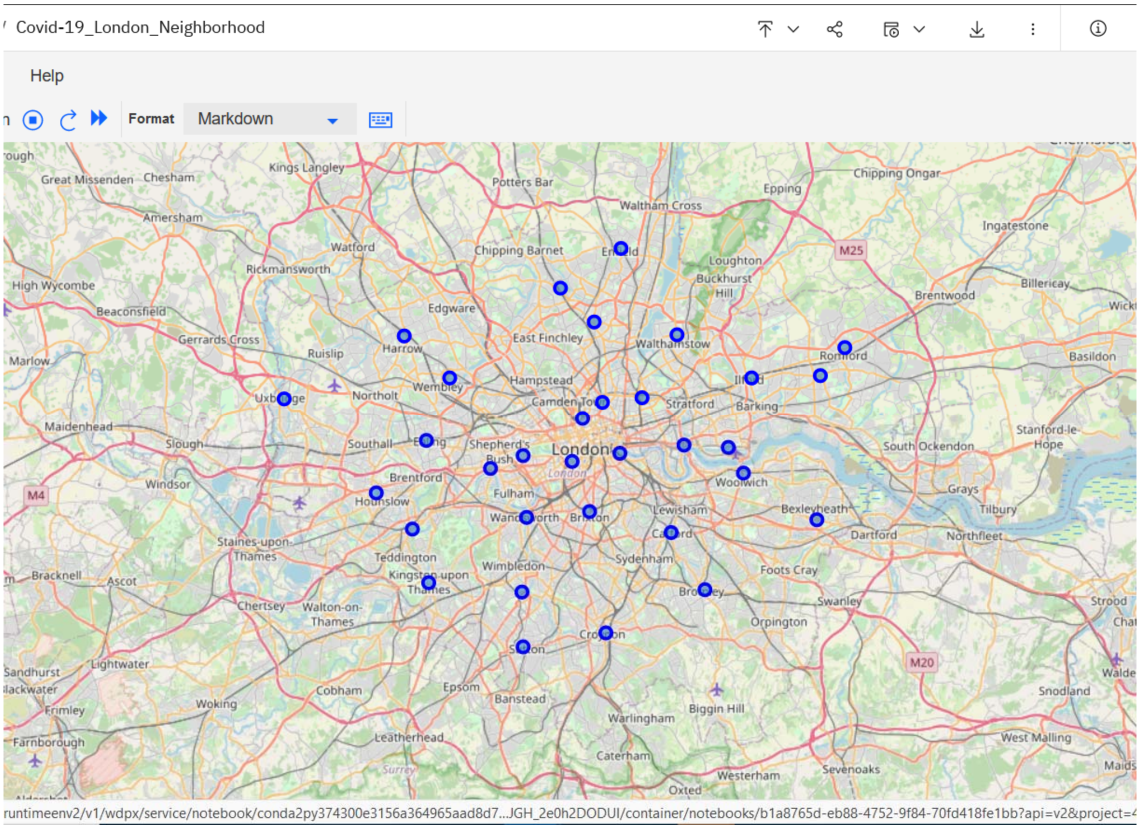


Figure 12. Position of each Borough in Latitude and Longitude for London, United Kingdom.

To help to the people find a similar neighborhood where we found high level of Contagium of Coronavirus, we will be clustering similar neighborhood using K – means clustering which is a form of unsupervised machine learning algorithm that clusters data based on predefined cluster size. We will use a cluster size of 4 for this project that will cluster the 8 neighborhoods into 4 clusters. The reason to conduct a K-means clustering is to cluster neighborhoods with similar venues together so the people can shortlist the area of their interest based on the venues/amenities around neighborhood.

After running the K-means clustering we can access each cluster created to see which neigborhood were assigned to each of the four (4) clusters. Looking the list of figure 13.

```
neighborhoods_venues_sorted.head(10)
```

]:

| | Neighborhood | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue | 6th Most Common Venue | 7th Most Common Venue | 8th Most Common Venue | 9th Most Common Venue | 10th Most Common Venue |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | Ealing | Coffee Shop | Clothing Store | Italian Restaurant | Bakery | Park | Pub | Café | Hotel | Burger Joint | Pizza Place |
| 1 | Hackney | Pub | Coffee Shop | Brewery | Cocktail Bar | Café | Bakery | Grocery Store | Modern European Restaurant | Vegetarian / Vegan Restaurant | Organic Grocery |
| 2 | Hammersmith and Fulham | Pub | Café | Coffee Shop | Indian Restaurant | Italian Restaurant | Gastropub | Clothing Store | Grocery Store | Hotel | Japanese Restaurant |
| 3 | Haringey | Park | Café | Portuguese Restaurant | Supermarket | Grocery Store | Hotel Bar | Indian Restaurant | Italian Restaurant | Mediterranean Restaurant | Fast Food Restaurant |
| 4 | Kingston upon Thames | Coffee Shop | Café | Clothing Store | Pub | Italian Restaurant | Sushi Restaurant | Burger Joint | Bakery | Department Store | Sandwich Place |
| 5 | Redbridge | Supermarket | Grocery Store | Fast Food Restaurant | Clothing Store | Turkish Restaurant | Sandwich Place | Bakery | Department Store | Coffee Shop | Sports Bar |
| 6 | Richmond upon Thames | Pub | Coffee Shop | Italian Restaurant | Bus Stop | Grocery Store | Pharmacy | Steakhouse | Indian Restaurant | Sandwich Place | Fish Market |

Figure 13. List of Most Common Venues in Neighborhood of London.

| Rate2ndWeek | Rate1stWeek | Rate_Category | Latitude | Longitude | Cluster Labels | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue | 6tl Co |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 143.6 | 131.4 | 100-199.9/100k | 51.5130 | -0.3089 | 2 | Coffee Shop | Clothing Store | Italian Restaurant | Bakery | Park | |
| 127.0 | 140.9 | 100-199.9/100k | 51.5450 | -0.0553 | 2 | Pub | Coffee Shop | Brewery | Cocktail Bar | Café | |
| 130.7 | 96.1 | 100-199.9/100k | 51.4927 | -0.2339 | 1 | Pub | Café | Coffee Shop | Indian Restaurant | Italian Restaurant | Gas |
| 122.8 | 100.9 | 100-199.9/100k | 51.6000 | -0.1119 | 3 | Park | Café | Portuguese Restaurant | Supermarket | Grocery Store | Ho |
| 117.7 | 100.3 | 100-199.9/100k | 51.4085 | -0.3064 | 2 | Coffee Shop | Café | Clothing Store | Pub | Italian Restaurant | Res |
| 127.1 | 123.5 | 100-199.9/100k | 51.5590 | 0.0741 | 0 | Supermarket | Grocery Store | Fast Food Restaurant | Clothing Store | Turkish Restaurant | Sa |
| 142.4 | 144.4 | 100-199.9/100k | 51.4479 | -0.3260 | 1 | Pub | Coffee Shop | Italian Restaurant | Bus Stop | Grocery Store | Pha |

Figure 14. Clustering of the data using the K-means equal 4.

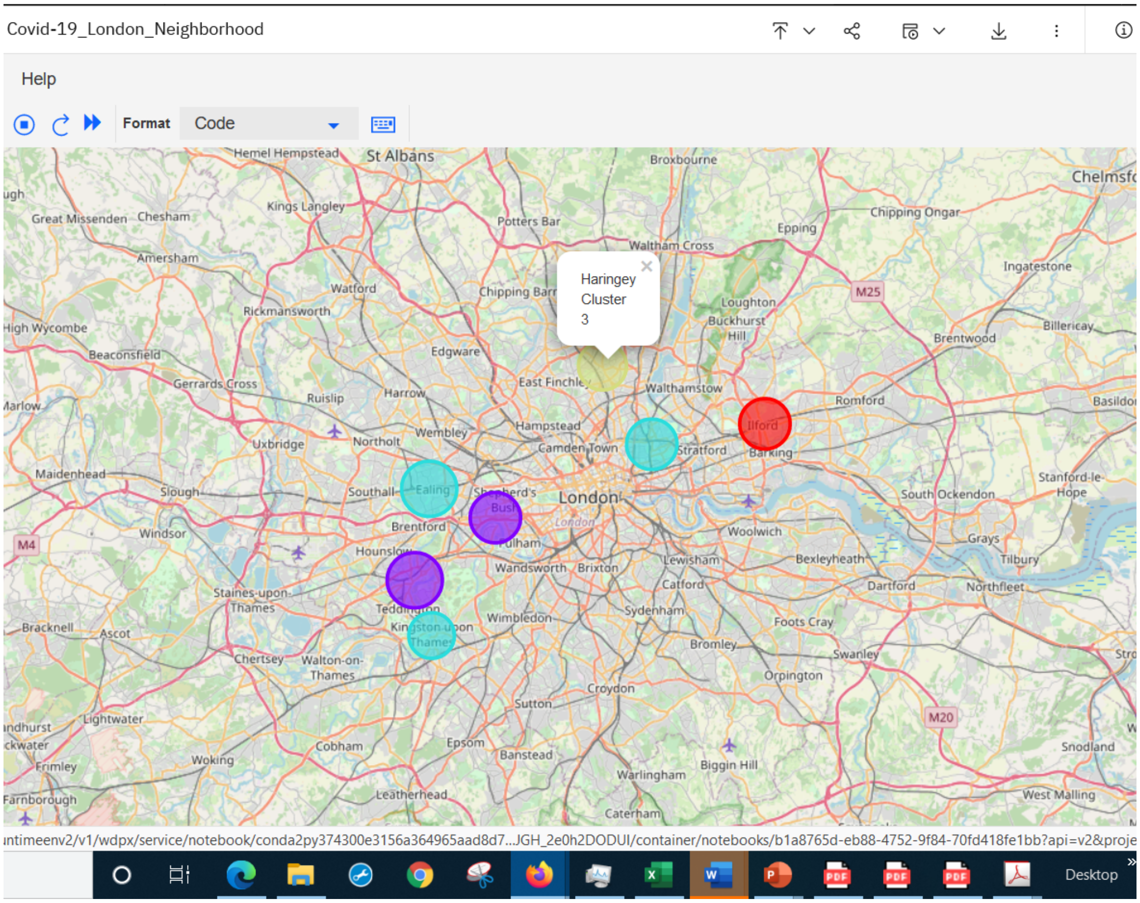


Figure 15. Clustered neighborhoods with highest risk of Contagium of Coronavirus (Covid-19).

Each cluster is color coded for the ease of presentation, we can see that three of Neighborhoods are Blue which corresponds to second cluster. Cluster three corresponding to Haringey is represented in yellow. (Color is similar to background color of this map).

# CONCLUSION:

When analyzing the data of the cases of Coronavirus (Covid-19) in each of the London neighborhoods, we can conclude that the hypothesis that Museums constitute a high-risk place for the contagion of Coronavirus is not true; or there are very few data to validate it for various reasons : (a) There is a very low population around the neighborhoods where the Museums are located. (b) Visitors to museums are generally foreign tourists or tourists from distant cities, generally from other countries, while the local population has no interest in attending the museums. (c) On the other hand, due to political measures, these buildings have already been closed and are easier to subject to strict biosecurity controls. Therefore in practice, these institutions do not exist, since they were closed. As the date these conclusions are written, the Museums are closed for the second time due to the medical emergency corresponding to the second wave of the Coronavirus. The government orders the closure of the museums from November 5, 2020 to December 3, 2020.

After a scientific study, we have found that the highest risk of contagion by Coronavirus are the ‘Pubs’ that are located in the Hackney neighborhoods, where there are a total of 5 ‘Pubs’, in Hammersmith and Fulham A total of 11 ‘Pubs’ were found, as well as in the Richmond upon Thames neighborhood there are a total of 8 ‘Pubs’ according to information obtained through the ***“Foursquare API”***. In total there are 24 Pubs in three neighborhoods. The ‘Pubs’ are closed bars where people go to drink, they meet some strangers, many people smoke making the environment very loaded with nicotine. It is well known that drinking excess liquor, and having a happy life at night lowers the defenses of the immune system, making it easier for the Coronavirus to attack. Some of these ‘Pubs’ are temporarily closed; some of them works until 18:00 hours. They are expected to full reopen on December 2020.

Another of places that constitute a high risk of contagion of Coronavirus are the Coffee Shop and the cafes, in the Ealing neighborhood : there are a total of 8 coffee shops in Kingston upon Thames : there are a total of 9 Coffee shops in Hammersmith and Fulham : there are 6 coffee in Haringey : there are 2 cafes and in Kingtom upon Thames : 6 cafes in Hackney : there are 4 Coffee shops. That means there are a total of 35 establishments that sell coffee and light meals or sweets, for tourists and for local people (people who live or work near the Center of London).

## For the moment we can do:

The World Health Organization WHO has mentioned the following directives. However, I consider that it is very little. Since the WHO is not leading any study with the depth that is required.

- Keep your distance from others. Stay at least six feet away from people outside your household as much as possible.
- Wear a mask outside your home. A mask protects others from your germs, and it protects you from infection as well. The more people who wear masks, the more we all stay safer.
- Wash your hands often. Anytime you come in contact with a surface outside your home, scrub with soap for at least 20 seconds, rinse and then dry your hands with a clean towel.