

The impact of different venue types on the prevalence of COVID-19 cases in London

Felix Mantz

3 January 2021

1. Introduction

a. Background

The global COVID-19 pandemic has dominated much of 2020, with countries enacting lock-downs and other socio-economic measures to contain the spread of the virus. Nevertheless, the virus has continued to spread throughout the world. The virus is primarily spread through droplets (e.g. of saliva) from a contagious person to another non-infected person. Venues with close proximity between customers and/or staff may therefore be contributing to the spread of COVID-19.

b. Problem

This project aims to identify if specific venue types are more prevalent in London boroughs which have seen faster increases in COVID-19 cases than other boroughs.

c. Interest

The results of this analysis may be of interest to policymakers considering if specific venues should be closed to help curb the spread of COVID-19. It may also be of interest to people aiming to avoid infection with COVID-19 who may chose to avoid certain London boroughs or neighbourhoods.

2. Data

a. Data sources

- i. Daily number of new and current total COVID-19 cases for London boroughs from 11 February 2020 to 27 December 2020 is provided by Public Health England and can be accessed [here](#)
- ii. The population of each London borough can be found [here](#). This is not presented in a format that is readily downloadable or scrapeable for free, so has to be transcribed manually
- iii. To analyse this data spatially, a GeoJSON file with outlines of London boroughs can found [here](#)
- iv. Lastly, information on venue types data can accessed using the Foursquare API, by requesting venues 'near' each of the London boroughs found above

b. Data cleaning

Initial cleaning of the COVID case data included formatting the dataframe (unstacking, index setting), and expressing the absolute number of new cases in each borough per 100k inhabitants of said borough.

COVID case data was combined for two boroughs, Hackney and the City of London. The City of London borough only has c10k residents, so my assumption is the COVID data was grouped with the larger neighbouring Hackney borough as it would otherwise be insignificant. This has knock on implications:

- i. The GeoJSON file to create a map of cases needed to have the geometries of these two boroughs combined. This was done by individually extracting the

geometries for these boroughs, merging them using the shapely library, and then updating the GeoJSON file

- ii. As the combined “Hackney and the City of London” does not exist, the Foursquare API was not able to resolve it as a viable geocode and did not deliver any venues. As a proportionate fix for this, for the purpose of querying Foursquare I simply looked for venues in Hackney.

c. Feature selection

I opted to focus on the number of new cases each day rather than total cases, as I wanted to get investigate where the virus spread most quickly. The change in total cases day to day is further muddled as it includes recoveries/deaths, and so is not a good indicator of new cases.

I did not trust the Foursquare API to be a comprehensive list of all venues in a borough. So, instead of working with absolute venue type frequencies, I worked out the proportion of each type for each borough. This gives an indication of the composition of venues in the borough, assuming those provided by Foursquare are a representative sample.

3. Methodology

- a. [Geospatial mapping]
- b. [Visual analysis]
- c. [Linear regression]

4. Results

- a. []

5. Discussion

- a. [See Expansion A]

6. Conclusion

- a. []

Expansion A:

Many venues have been shut at various times over the course of 2020 due to government restrictions in place for London/the UK, so future analysis could focus on time periods where venues were largely open (15 June 2020 to 5 November 2020 – see timeline below).

Timeline of key policy events relevant for the analysis

- 15 June 2020: Non-essential shops re-opened on 15 June following the first wave. Begin of the analysis period
- August 2020: “Eat out to help out” scheme encouraging restaurant use
- 14 September 2020: “Rule of six” capping social meetings at 6 people
- 15 October 2020: London moved to “Tier 2”, further restricting meeting
- 5 November 2020: Begin of second UK national lockdown, including closing of all non-essential venues