

# Using Demographics and Location-Based Social Network Data to Understand the Spread of COVID-19

---

Thomas Hall

Monday, 04 May 2020

## 1. Introduction

Since it was first discovered in December 2019, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has had a massive impact as it has spread across the globe. The virus causes the disease known as COVID-19 and due to its novel appearance, there is great uncertainty surrounding its epidemiology, such as risk factors, transmission characteristics, and possible control measures. With such uncertainty and rapid onset, governments the world over have had no other option but to shut down their entire economies, confine their citizens under lockdown, and bide their time until an exit strategy can be found.

As of May 4<sup>th</sup>, 2020, there have been 24,988 total cases of COVID-19 in London, representing a significant outbreak of the disease [1]. This report seeks to use demographic profile data, venue location data and coronavirus case data to examine outbreak trends across London's 32 boroughs. Example questions to be answered are:

- Are deprived areas at more risk of developing COVID-19?
- Are areas with an older population or those with higher BAME populations at greater risk of having more COVID-19 cases?
- Is prevalence increased by the number of certain venue types, e.g. pubs, restaurants, or parks?

Findings from the investigation could be useful in informing how the capital defines and manages its exit from lockdown.

## 2. Data

Data for this investigation is acquired from the following sources:

- **Demographics:** data is sourced from the Greater London Authority [2], which has comprehensive profiles of each borough. Example data for each borough includes:
  - population and population density,
  - average age,
  - migrant and BAME populations,
  - employment rates,
  - household incomes,
  - house ownership,
  - carbon emissions,

- transport usage,
  - life expectancy,
  - and life satisfaction amongst other.
- **Venue Location Data** is sourced from the Foursquare API [3] versioned to March 15<sup>th</sup>. The top 200 locations are acquired for each borough to create a venue profile.
- **Coronavirus** data is sourced from Public Health England [1], which publishes daily updates for each upper tier local authority in England – corresponding to the boroughs in London.

## References

- [1] Public Health England, “Coronavirus (COVID-19) cases in the UK,” 4 May 2020. [Online]. Available: <https://coronavirus.data.gov.uk/>.
- [2] Greater London Authority, “London Borough Profiles and Atlas,” [Online]. Available: <https://data.london.gov.uk/dataset/london-borough-profiles>. [Accessed 4 May 2020].
- [3] Foursquare, “Places API,” 4 May 2020. [Online]. Available: <https://developer.foursquare.com/docs/places-api/>.