

1. Introduction

1.1 Background

"UK restaurant market facing fastest decline in seven years"

A headline from last year [i] prior to the coronavirus.

MCA's UK Restaurant Market Report 2019 [ii] indicated that *"large falls in the sales value and outlet volumes of independent restaurants is the cause of the overall decline of the UK restaurant market. It attributes this to a "perfect storm" of rising costs, over-supply, and weakening consumer demand."*

London's restaurant scene changes week on week, with openings and closures happening on a regular basis; it must be hard to keep up. The hyper-competitiveness of London's restaurant scene make it one of the toughest cities in the world to launch a new venture.

"With business rates up and footfall down, a winning formula is worth its weight in gold and although first-rate food is inevitably the focus, other factors can also affect a restaurant's success. Atmosphere is frequently cited in customer surveys as second only to food in an enjoyable restaurant visit and getting the vibe right is crucial." [iii]

Due to the coronavirus most businesses have suffered even greater losses. As restrictions lift businesses will be looking for ways to make up for lost time and earnings. Reopening a restaurant once lockdown is over is one thing, but knowing what to put on the menu if you haven't been in contact with a punter in months is another.

"Are there any grounds for hope? A wild optimist might point to some encouraging data about the overperformance of small chains while everyone else loses their shirts; a realist might make coughing noises about small sample sizes and growth from a low base. The queues snaking out of Soho's recently opened Pastaio suggest one genuinely viable route to salvation - concepts may need to follow its lead and amp up the comfort food factor while dialling down prices.

And while home delivery is a source of confidence for some parties (Deliveroo, for instance, recently listed its shares on the stock market) it may well end up a false friend: the increased volume of so-called "dark kitchens" presage a sinister vision of the future, where restaurants don't exist to serve customers onsite at all, but just pump out takeaway meals for us to consume on our sofas. A little far-fetched, perhaps, but with lights going out at a faster rate than many can remember, it can't be too long before whole tranches of the market do indeed go dark, one way or another." [iv]

[i] <https://www.bighospitality.co.uk/Article/2019/09/24/UK-restaurant-market-facing-fastest-decline-in-seven-years-according-to-MCA-Insight>

[ii] <https://www.mca-insight.com/market-reports/uk-restaurant-market-report-2019/597394.article>

[iii] <https://www.standard.co.uk/go/london/restaurants/how-londons-top-restaurants-soundtrack-their-spaces-a3737991.html>

[iv] <https://www.theguardian.com/global/2017/nov/26/who-killed-londons-restaurant-scene>

1.2 Business problem

The task is to identify a new, on trend hospitality business opportunity in a thriving location in London. However, with the country currently under lockdown it is much harder to understand what types of food and drink businesses are popular. A novel method will need to be deployed to analyse the current situation with food businesses in London during the coronavirus.

The project originally planned to use foot traffic data at different times to identify what food venues appear to be trending and where. However, given that the whole country is under lockdown due to the Coronavirus there is no trending data to analyse.

Another flawed idea would be to assume whatever food venues are most common are most in demand/popular. This is not ideal since it would be using data too set in the past (it takes time to build a restaurant) and trends move more quickly. This method would not provide near-real-time visibility of what is actually trending to make more accurate predictions, ahead of the curve. In other words, market trends and tastes change all the time and whatever shows up using the mode of venue results is only what was popular months prior. **Note:** This is true of this method, regardless of lockdown!

Instead, data that might contribute to determining restaurant improvements might include performance metrics during lockdown; hours, venue likes, volume of recommendations, quality of recommendations, content of recommendations (word densities), as well as clusters of food businesses that remained operational during lockdown. This project aims to predict pandemic proof food enterprises and what the industry might look like after restrictions are lifted.

1.3 Interest

Obviously, any restaurateur or leisure and hospitality entrepreneur/enterprise would be very interested in accurate prediction of food trend data for competitive advantage and added business value. These such data could be used to inform new menu creation or concepts for new boutique restaurants or street food vendor pop-ups - proving valuable to food and retail parks, such as Boxpark.

1.4 Desired outcome

The ideal outcome of this notebook would be to:

- Create a dataframe of London districts, postcode centroids, and coordinates
- Identify the top 5 (most common) food venues by cuisine
- Plot all food venues to map
- Use venue Hours Endpoint of FS API to see what venues are still operating during lockdown
- Create word clouds of venue Tips using Foursquare API to identify trending menu items. Since the Foursquare trending feature won't return any results at this time due to the coronavirus lockdown, I will use the Tips Endpoint in the Foursquare API to try and identify patterns in the reviews i.e. what menu items get the most positive mentions (**Note:** This may require the use of sentiment analysis, which is out of scope for this notebook)
- Use choropleth maps to highlight food vendor densities per London district by different cuisines (optional: choropleth map by density of venues open during lockdown)

- Use k-means clustering to cluster food venues in London to identify restaurant hotspots and prime locations as suggestions for the client (optional: cluster by venues open during lockdown)
- Map commercial venues that are available to rent using either Zoopla API or Rightmove webscraper
- List suitable commercial venues for further analysis

2. Data Acquisition and Cleaning

2.1 Data sources

To explore the problem we can use the data listed below:

- Wikipedia page for London Postal Districts [1] to get an initial high-level view of what we are working with.
- For cross-referencing postcodes with districts and longitudes and latitudes I will use a combination of the Office for National Statistics [2] and London Datastore [3]. I also checked "A Guide to ONS Geography Postcode Products" [4] to make sure I was using the correct postcode system for statistics (NSPL).
- I found the Second-level Administrative Divisions of the United Kingdom from NYU Spatial Data Repository [5]. The .json file [6] has coordinates and boundaries of the all the cities of the UK. This will be cleaned and reduced to London where I will use it to create a choropleth map of food vendor densities for different cuisines using the Foursquare API.
- Foursquare API [7] will be used to get the most common food venues of London. **Note:** This may be reduced further to City of London and Westminster (or just EC and WC postcodes) to reduce the number of API calls. The FS API's hours and tips Endpoints [8] will be used to get venue operating hours (to see which venues are still operating during lockdown) and to get user recommendations, which will be used for word clouds to try and identify trending menu items.
- I will then use either Zoopla API [9][10] or Rightmove webscraper [11] to pull in commercial properties on the market as options for the client.

[1] https://en.wikipedia.org/wiki/London_postal_district

[2] <https://geoportal.statistics.gov.uk/datasets/national-statistics-postcode-lookup-may-2020>

[3] <https://data.london.gov.uk/london-area-profiles/>

[4] <https://www.ons.gov.uk/methodology/geography/geographicalproducts/postcodeproducts>

[5] <https://geo.nyu.edu/catalog/stanford-wj438mh2295>

[6] <https://earthworks.stanford.edu/download/file/stanford-wj438mh2295-geojson.json>

[7] <https://developer.foursquare.com/>

[8] <https://developer.foursquare.com/docs/places-api/endpoints/>

[9] <https://developer.zoopla.co.uk/>

[10] <https://github.com/AnthonyBloomer/zoopla>

[11] https://github.com/toby-p/rightmove_webscraper.py