# Battle of Neighbourhoods

## Introduction/Background

Mrs A has decided to open a new Italian restaurant in London, having moved from Italy but she has not yet decided on the location. Mrs A wants to find the neighbourhood with successful existing restaurants to ensure there is high enough footfall, however she does not want there to be too much direct competition from other nearby Italian restaurants.

This data science problem will be addressed with Mrs A as the target audience to help her decide where the best location for her new Italian restaurant will be.

## Data

There are multiple factors to consider in the approach to this problem. These include the number of similar restaurants in the immediate area (here we will use a 2 mile radius), the success of other food businesses and the affluence of the area in general.

A recent report found that "full-service restaurant revenues are directly correlated with the available level of customers' disposable income". (*Parsa, H. & Kreeger, Jeff & van der Rest, Jean-Pierre & Xie, Karen & Lamb, Jackson. (2019). Why Restaurants Fail? Part V: Role of Economic Factors, Risk, Density, Location, Cuisine, Health Code Violations and GIS Factors. International Journal of Hospitality & Tourism Administration. 10.1080/15256480.2019.1598908.*). This should be a contributing factor in deciding upon the most suitable neighbourhood for a new Italian restaurant, therefore, I will bring in an external dataset showing the regional gross disposable household income (GDHI) in London to identify the borough with the highest GDHI.

An API call to Foursquare will be made to obtain details of all restaurants in the borough and to establish clusters, K-means clustering will be used. The aim is to identify a suitable area where there are no Italian restaurants within a two mile radius but where there are other highly rated restaurants in the vicinity.