**London districts and schools catchment areas**

**Introduction**

Despite Brexit, London is a great city to live. It is one of the most prominent financial, cultural, historical and social centres of the world, which offers many exciting job opportunities for young people.

However, as most of the new parents in their late twenties and early thirties will tell you is how high the housing and childcare costs are here. The schools are divided into very expensive private schools and free state schools. Good quality free schools are scarce and hard to get into. A child must typically leave within one mile radius from a school to get accepted and it is not uncommon for families to move closer to the school of their choice to increase their admission chances.

This Data Science capstone project covers the following:

* Analysing London postal districts venue types and clustering the districts by their respective venues’ popularity.
* Analysing London school venues and locating state primary schools on a map, only mixed or girls only schools are selected and only the ones which have an ‘outstanding’ OFSTED (Office for Standards in Education, Children's Services and Skills) rating. The resulting map should display a 1 mile radius around the schools. It will be of help to parents to research their local schools and to plan their move if needed.

**Data**

For the purposes of the project the following data has been used:

1. A list of the London’s postal districts, their names and postal IDs. The list was obtained from [1]. Python Nominatim functionality was used to obtain each district’s latitude and longitude coordinates.
2. The most popular venues and their types (e.g. a coffee shop or a theatre) for each of the London’s postal districts. These were sourced using the Foursquare API functionality. The popularity of each venue type was determined by the number of occurrences of that particular venue type in a district.
3. List of schools for each of the London’s district was initially sourced with the Foursquare API. However, this approach appeared to be not ideal: a) it displayed all the venues which included the word ‘school’ in the name, even locations such as bus stops or driving schools, which are irrelevant for the purpose of this project; b) it didn’t not display information about the schools like their ratings, if a school is private or state, primary or secondary, etc.
4. Data on London schools, their URNs, names, types, OFSTED ratings and postcodes. The data was taken from [2]. Nominatim functionality was used to obtain each schools’ latitude and longitude coordinates.

**References**

1. London postcodes: <https://www.doogal.co.uk/london_postcodes.php>
2. Data on London schools: <https://www.compare-school-performance.service.gov.uk/schools-by-type?step=default&table=schools&region=all-england&for=ofsted>