Assessing childhood obesity through geospatial placement of fast food restaurants

A Coursera Applied Data Science Capstone project

Introduction: Business Problem

The World Health Organization terms childhood obesity as "one of the most serious public health challenges of the 21st century", with prevalence increasing globally and affecting low and middle-income countries, as well as low and middle-income areas in more developed countries. In 2016, there were about 41 million children under the age of five that were categorized as overweight [1].

Although there are different ways to measure obesity within different age brackets (i.e. it is measured in a different way for 0-5 years, 5-19 years, and 19+), one thing is for certain: obesity has detrimental effects on the health of the individual in addition to putting strain on the healthcare system of the community.

There are several contributing factors to obesity. This report will analyse one aspect only: the possible impact of energy-dense foods that are high in fat and sugars and low in healthy vitamins and minerals. One particular source of such foods is fast food restaurants that offer foods such as fried chicken, burgers, pizza, etc. While there are many ways that a possible association between such restaurants and childhood obesity can be analysed, this report will investigate only whether there is a correlation between childhood obesity and the proximity of fast food restaurants to schools.

The results of such an analysis and others within this topic are very important because such results can make a strong case for policy intervention in ensuring that areas close to schools are kept free of such eateries where children can have quick and easy access to unhealthy foods either with their parents after school or (if they are old enough) by themselves on the way back after school. While there are bound to be other factors that impact the obesity rate, the geospatial placement of fast food restaurants within proximity of schools is certainly one that should be analysed to help policymakers in reducing childhood obesity in areas of high incidence.

This report will analyse the 32 boroughs in London, UK. Each borough in London is a local authority district that makes up Greater London (thereby excluding the City of London), and is governed by a London borough council. The selection of London for this report is because it is one of the most culturally-diverse cities in the world whilst being one of the most important economic and financial hubs globally. Additionally, it is also the city of residence of this author. The prevalence of childhood obesity in London, as last measured in 2015/16, was 23.2% compared to a global average of around 18% [2]. The following map shows the different boroughs of London (with each borough centred around its respective borough council address) in which all boroughs with childhood obesity prevalence higher than the global average are marked with red and those with lower are marked in blue. This indicates the severity of the problem within London and is representative of how high rates of childhood obesity are prevalent in urban centres in developed countries.
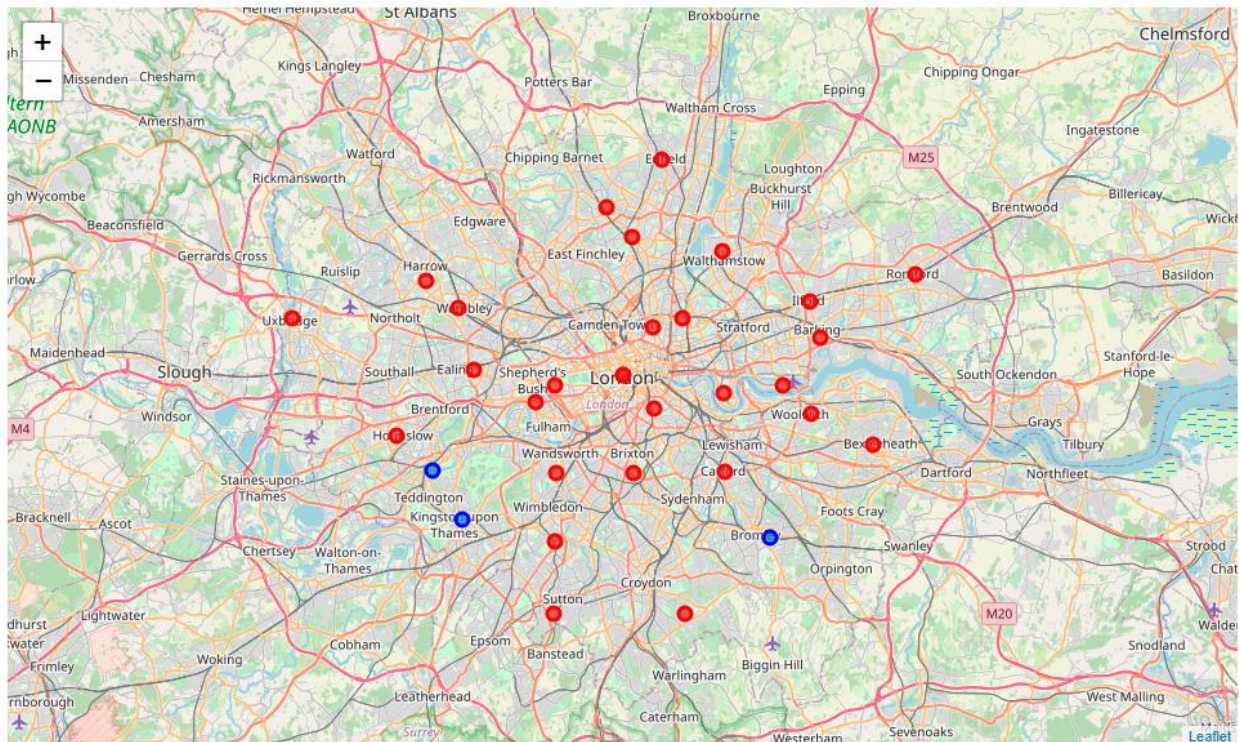
The null hypotheses for this report are :

H01: There is at least a moderate correlation between the prevalence of childhood obesity in a borough and the average distance between a fast food restaurant and a school in the borough.

H02: There is at least a moderate correlation between the prevalence of childhood obesity in a borough and the average number of fast food restaurants within a 1 km radius of a school in the borough.

Data

To evaluate the possible correlation, the variables of interest will be:

- Prevalence of childhood obesity in each London borough (%)
- Average distance of a fast food restaurant from a school in the borough (km)
- Average number of fast food restaurants within a 1 km radius of the school

The main data set that will be used is the London Borough Profiles and the London Borough Atlas from the London Datastore (www.data.london.gov.uk). These two files are available in Excel format and they will provide data on prevalence of childhood obesity in each of the boroughs.

The other two variables, i.e. the average distance of a fast food restaurant from a school in the borough, and the average number of fast food restaurants within a 1 km radius of the school, are measured using the Foursquare API and are explained in more detail in the Methodology section.

Further, the following assumptions are taken for the purpose of analyzing the data:

- The coordinates representing each borough are those of the respective borough council office (or town hall).
- Only 10 schools within a 2 km radius of each borough council are taken as a representative sample for this analysis.
- Only those fast food restaurants are considered that are within a 1 km radius of each sample school.

References:

1. https://www.who.int/dietphysicalactivity/childhood_why/en/
2. https://www.worldobesity.org/about/about-obesity/prevalence-of-obesity