# Where will the next Waitrose be built?

Kevin Brown

*Disclaimer: Waitrose are unaware of this work. I have no connection with Waitrose or John Lewis.*

## Introduction

### Background

A box filled with different types of food

Description automatically generatedIt is well known that supermarkets choose their locations to fit with their target market – and the location of the supermarket is vital to the success of the shop. The ‘low cost’ supermarkets will target different areas to ‘premium’ supermarkets. But what makes the ideal location for supermarkets? Is it possible to look at where supermarkets are located, understand the neighbourhoods of which they are a part, and then find similar neighbourhoods across the country where they could be placed?

Waitrose is a supermarket in the UK. Compared with other supermarkets, it is perceived to be [more expensive](https://www.telegraph.co.uk/finance/newsbysector/retailandconsumer/10538240/Changing-the-upmarket-perception-of-Waitrose.html), associated with the [more well-off](https://www.theguardian.com/uk/2011/jun/29/waitrose-lidl-shopper-divided-britain) and linked to [Conservative-run councils](https://www.thesun.co.uk/news/3745561/waitrose-shoppers-are-tories-ukippers-prefer-aldi-and-sainsburys-customers-vote-labour-what-does-your-supermarket-choice-on-how-youll-vote-tomorrow/) - and can even [increase house prices](https://www.independent.co.uk/property/house-prices-latest-waitrose-effect-sainsburys-marks-and-spencer-uk-property-a7760926.html), more than other supermarkets.

In this project, I will aim to find the location of all current Waitrose stores and find the characteristics of the area to find which ones are most important, by:

* Finding the name and postcode of all stores
* Finding the lat/lon of all stores
* Finding the political constituency of all stores
* Finding the deprivation index of all stores
* Finding the local amenities for each of the stores
* Exploring if there is any correlation between the data (compared with National data) for Waitrose stores, and trying to find similar areas in the UK where new stores could be built.

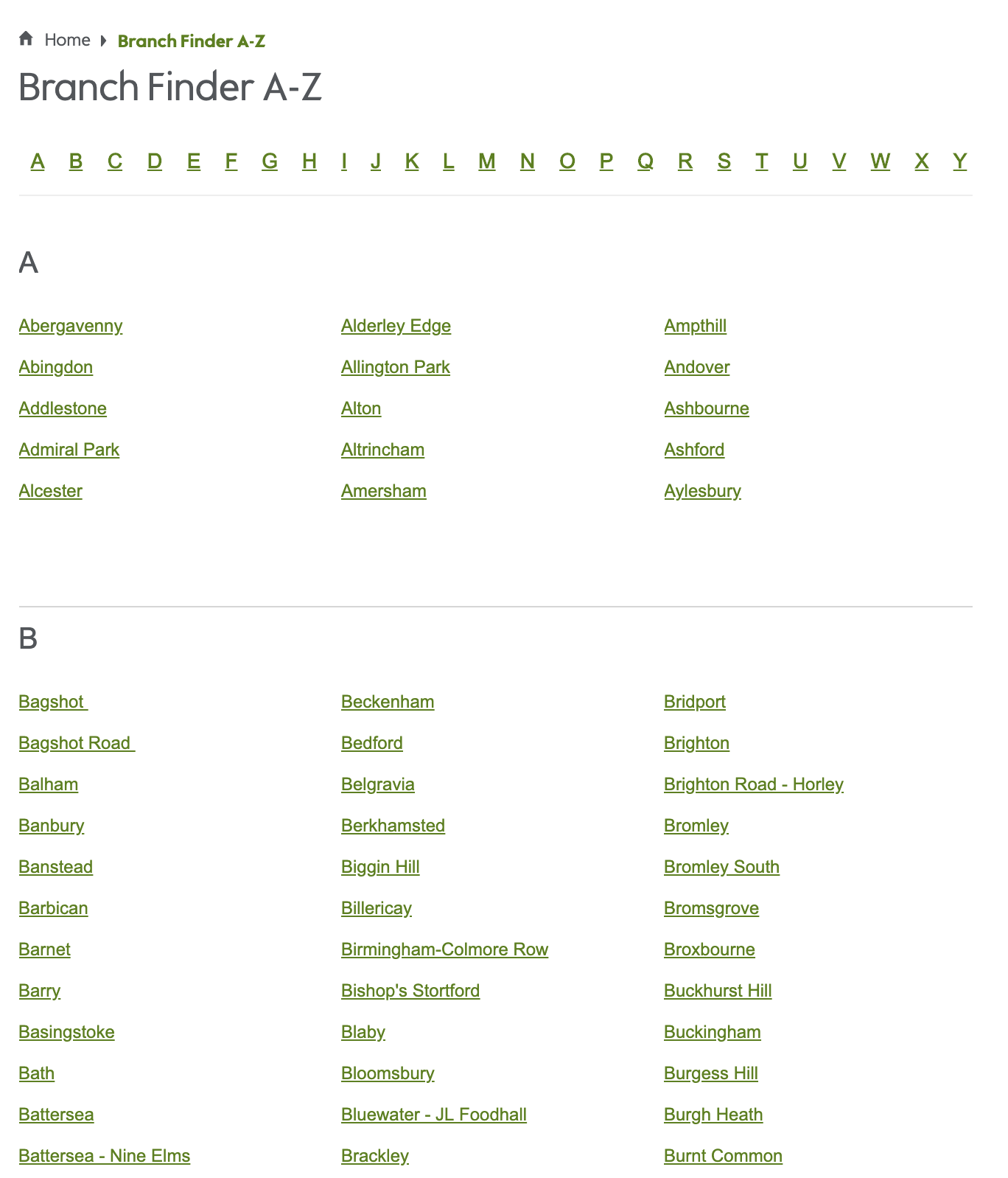
In order to predict where the best places are for Waitrose stores to be built I will:

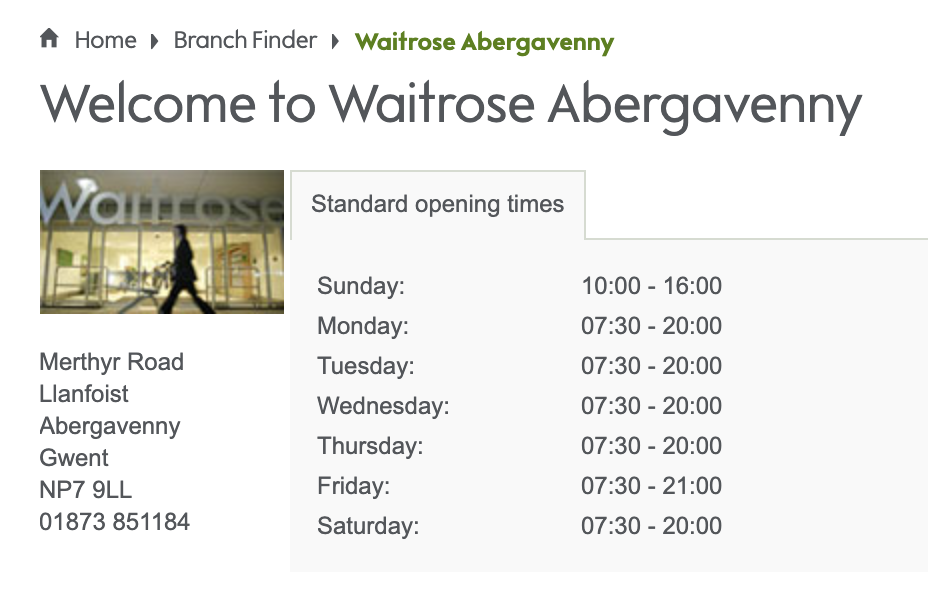
* Filter the list of postcodes by any correlating factors that are found, to potential neighbourhoods of interest
* Looking at the local amenities for those areas, use K-means to determine which of the non-waitrose areas are most closely aligned to the existing waitrose areas, and present some suggested locations.

## Data

The data required will come from a variety of sources, and be useful in different ways:

### Store Data

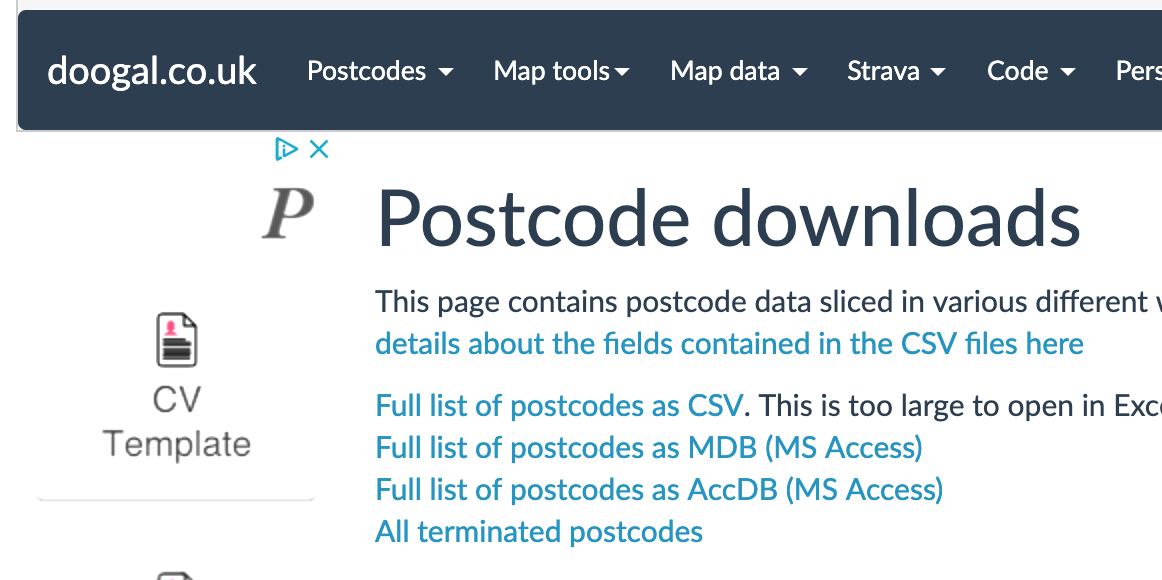
Waitrose [A-Z store finder](https://www.waitrose.com/content/waitrose/en/bf_home/branch_finder_a-z.html) has a list of stores, which I can scrape, in order to get the full list of stores, and the URL of that store’s web page.

Each Waitrose store has its own [web page](https://www.waitrose.com/content/waitrose/en/bf_home/bf/632.html). I’ll need to use the URL to visit the web page for each of the stores, and scrape the postcode and store details.

The features from the A-Z store page will be store name, and URL.

The features required from the store details will primarily be Postcode – but I will capture the Phone number and county, in case it’s needed for display purposes later.

### Postcode Data

There is a python library call postcodes, which I explored for mapping postcodes to Lat/Lon, however it relied on a web service which is no longer active, so the python library does not work.

Therefore, I will use the site doogal.co.uk to download a csv [mapping](https://www.doogal.co.uk/PostcodeDownloads.php) of postcodes to Lat/lon, Parliamentary Constituency, Deprivation index or crime index for that postcode

As described, the features required from the Postcode csv will be looked up using the Postcode of the store, and will consist of the Latitude, Longitude, Constituency ID, Index of Multiple Deprivation. The [field description page](https://www.doogal.co.uk/PostcodeCsvFields.php) describes how the deprivation rank for the postcode, is a number where 1 is the most deprived. The range of values is as follows:

1 - 32844 = England

1 - 1909 = Wales

1 - 6976 = Scotland

1 - 890 = Northern Ireland

### Political Data

Having retrieved the Constituency ID for the postcode, I will need the last election results (2017) to give me the current political party for the constituency under which the Waitrose store currently sits. This is available from the [UK Parliament](https://researchbriefings.parliament.uk/ResearchBriefing/Summary/CBP-7979), as a downloadable csv.

### Neighbourhood / Venue data

From experience, it is usual for Waitrose to be located in towns with lots of local amenities, rather than in a trading estate, or in the countryside etc., Therefore, we’ll review the profile of existing Waitrose stores to see what venues they have around them, and then compare them to some proposed sites, to see which are most similar. The top local amenities around the supermarket, or any UK location are available from Foursquare.