



Locations for London Coffee Shops

IBM COURSERA CAPSTONE

Required Data Sources

- Geolocation data for all London train stations
- Passenger number data
- Foursquare API for nearby venues

Finding potential locations

There are 207 train stations in London zones 1 to 3. These are ideal candidate locations for new coffee shops because of the large number of passengers in the area each day.

These stations are dispersed over a large area with many different demographics, so further criteria for choosing suitable locations is necessary.



Data Cleaning

- Remove stations outside of zones 1 to 3
- Remove stations which we don't have passenger number data for

Using Foursquare and problems

Some of the Foursquare data appears to be incomplete

tube name	Coffee Shop
Acton Town	0.047619
Aldgate	0.120000
Aldgate East	0.110000
All Saints	0.076923
Angel	0.074074
Archway	0.171429
Arsenal	0.000000
Baker Street	0.051546
Balham	0.145161
Bank	0.110000
Barbican	0.054945
Barons Court	0.032258

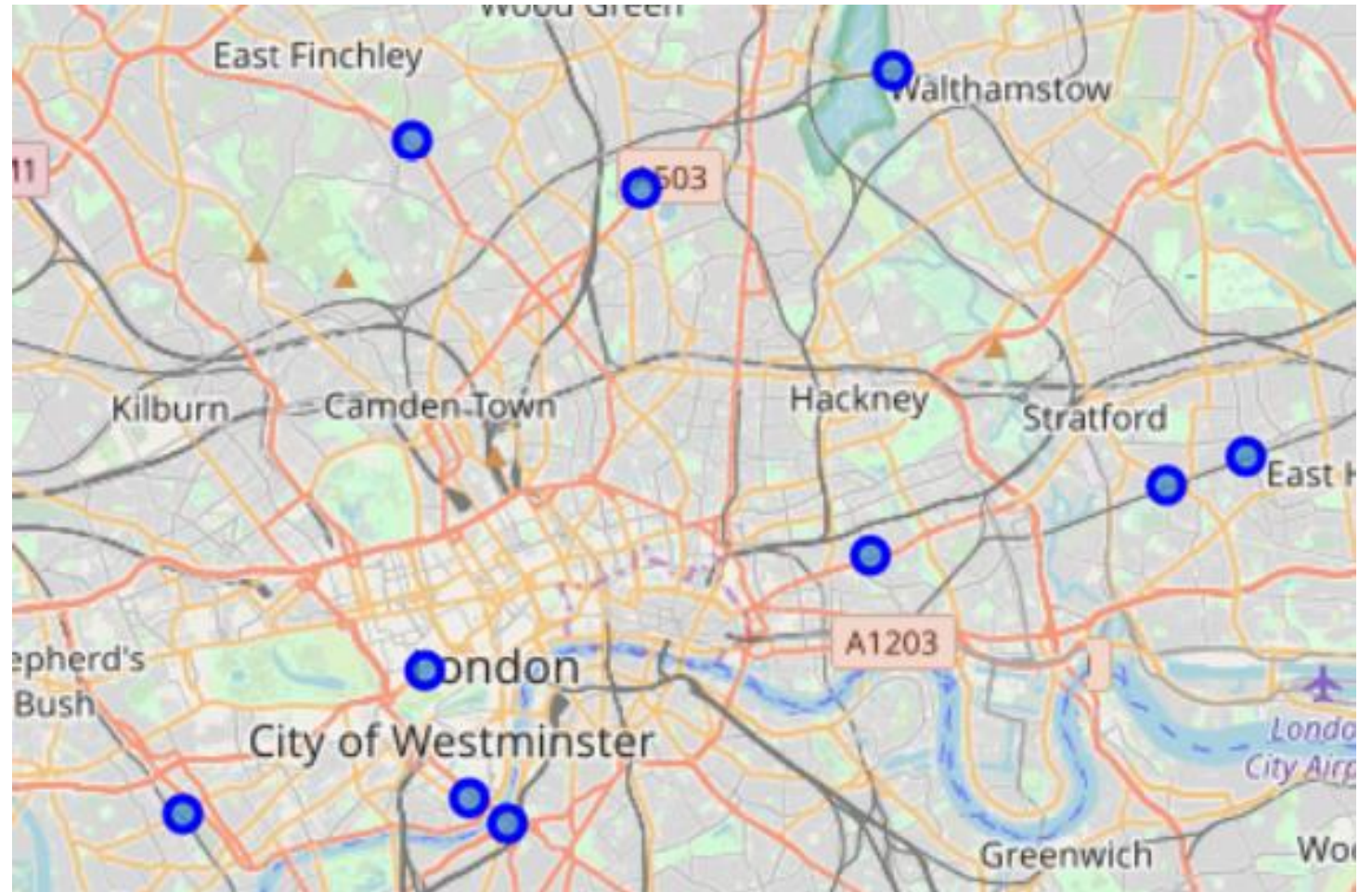
	tube name	coffee_shop	million
74	Green Park	0.0	39.338161
183	Vauxhall	0.0	30.833904
135	Pimlico	0.0	10.971039
182	Upton Park	0.0	9.593829
19	Blackhorse Road	0.0	8.999728
111	Manor House	0.0	8.688099
166	Stepney Green	0.0	6.341140
136	Plaistow	0.0	6.128691
191	West Brompton	0.0	5.877152
84	Highgate	0.0	5.873629

Top 10 store locations

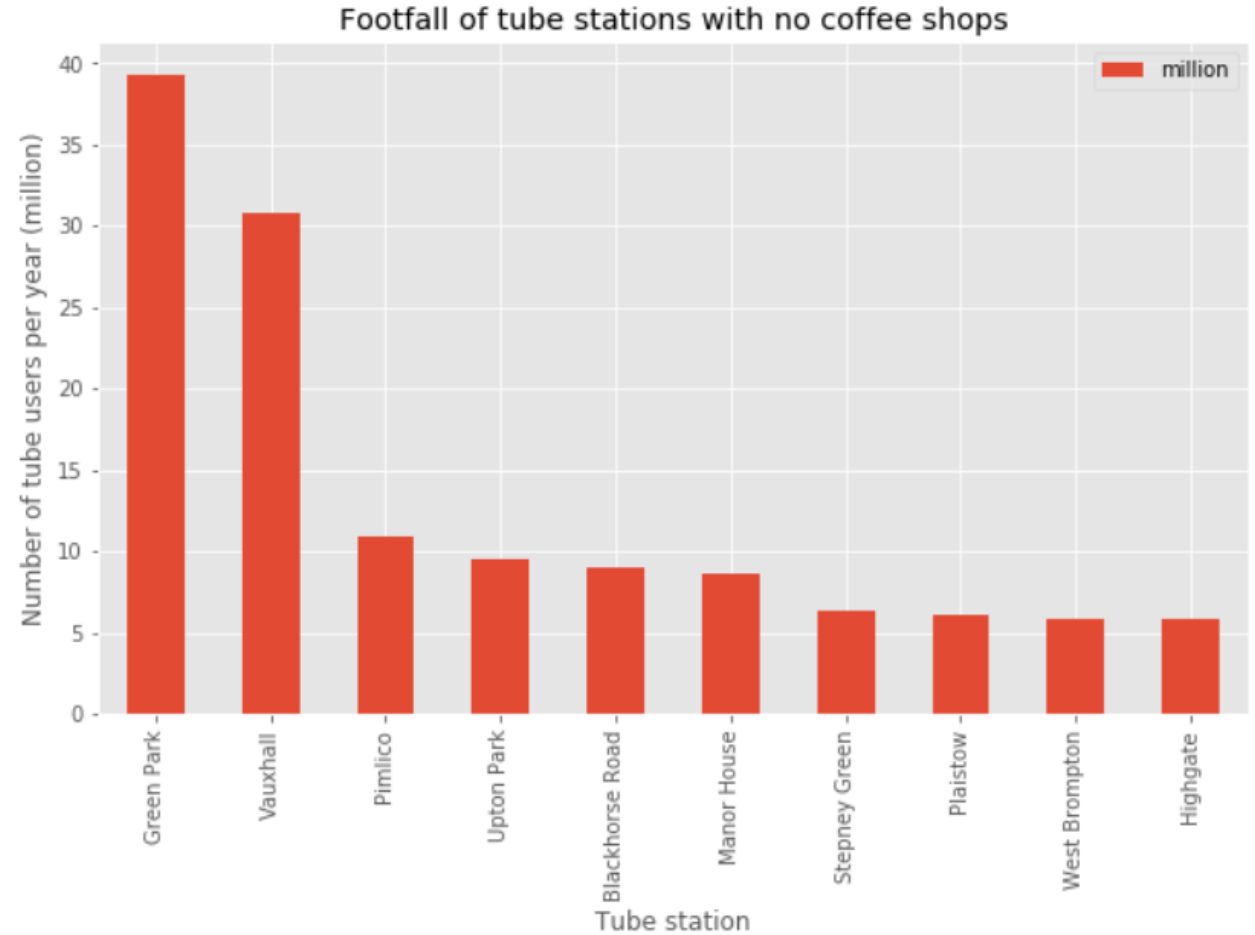
Between them, the top 10 tube locations have passenger numbers of 133 million each year.

However, the Foursquare data may be incorrectly reporting that there aren't already coffee shops in the area.

Top 10 potential coffee shop locations in London

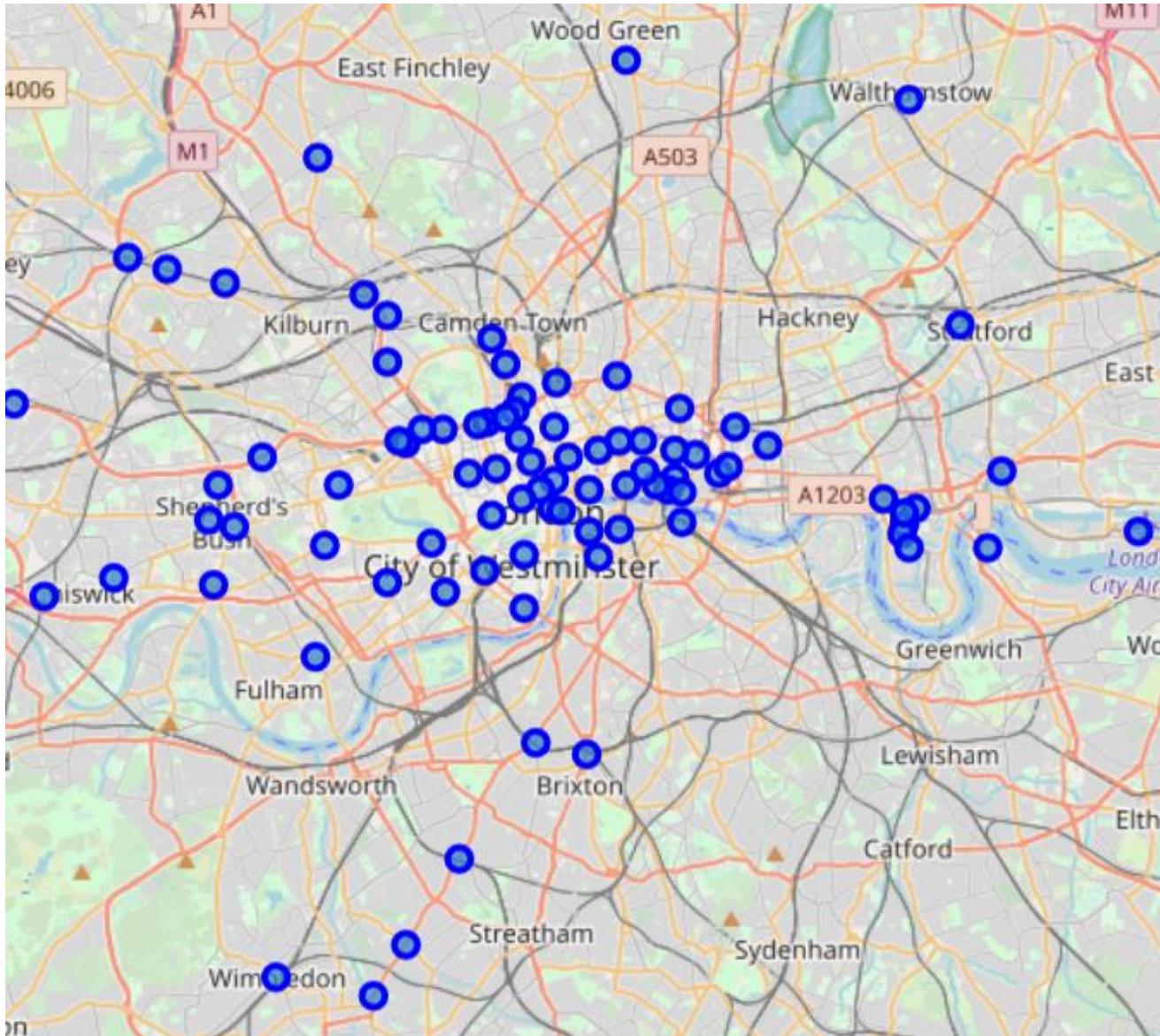


Annual passenger throughput of top 10 locations



K-Means clustering on model location

- Green Park was identified in the data as the best candidate location for a new coffee store.
- K-Means clustering was used to expand the analysis and use Green Park as a model for selecting other tube stations.
- This takes into account additional factors other than just the existing number of coffee stores, and looks at the make-up of the whole area to determine whether a new coffee shop might be viable



90 potential coffee
shop locations
within the same
cluster as Green
Park

Further Analysis

The Foursquare data doesn't appear to be a true reflection of the number of coffee shops. The research should be carried out again using a different API such as Google

New data sources could be found, such as information about the wealth of residents in a particular area

It is also important to know the rent and business rate costs. Suitable vacant premises will need to be identified and compared

We have also not considered the type of coffee shop which would be best, such as a big chain, or a niche artisan store