HS2 Curzon Street Terminal

Table of Contents

Introduction

Background to HS2

HS2 is a high speed rail link that is being built within the United Kingdom (UK) to improve the connection between the North and South of England. The first stage of this project is to build a high speed line between London and Birmingham. Subsequent stages will connect cities further North such as Leeds and Manchester.

London is the UK's capital located in the south and Birmingham is the UK's second largest city located in the centre, in an area known as the Midlands. The London terminal of HS2 will be at Euston and the Birmingham terminal at Curzon Street.

The aim of HS2 is to drive economic growth north of London. The initial business case determined that there would be a return of £2 on every £1 invested. That has been challenged due to the rising costs of the project potentially, doubling since the initial business case in 2013.

The Curzon Street Terminal

The Curzon Street terminal is located near the centre of Birmingham. Its location connects it to 2 other rail stations in in Birmingham, namely Moor Street and New Street. Allowing easy onward travel to other parts of the UK.

Location of the Curzon Street Terminal
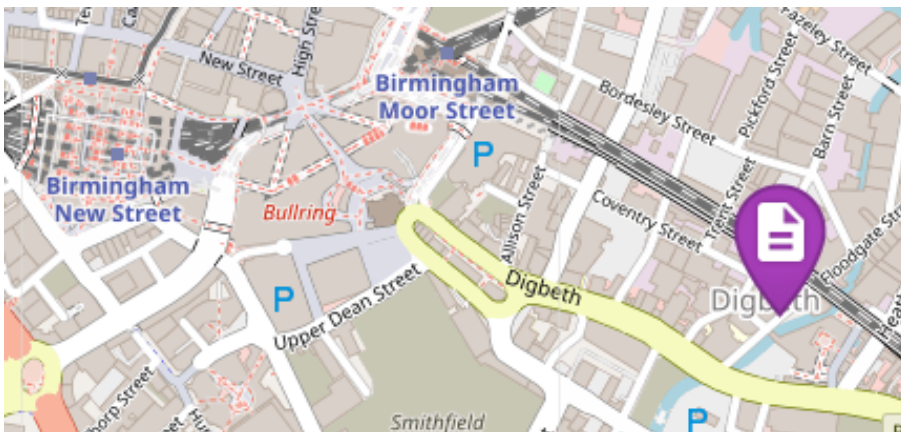
Part of the economic growth will come from increased spending within the vicinity of the Curzon Street terminal both by people passing through and it becoming a more desirable destination. The terminal is located by the iconic Bull Ring shopping centre.

The Bull Ring

The Problem

The full economic benefits to the area surrounding the Curzon Street terminal will only be available if the right venues are in place to service the traffic and tourists it generates. The first stage in this is understanding the existing venues in the immediate area. The immediate area, for this report being with 3k circa 1.9 miles, being a reasonable walking distance.

The Audience

This investigation will be of benefit to a number of different audiences:

HS2: to determine how they support the wider development of the area to drive traffic on the London Birmingham link

Birmingham City Council: to inform planning policy and decisions to encourage the most appropriate development of the surrounding area to promote the wealth and well-being of their residents

Developers & Businesses: to inform investment decisions in the area, both in new businesses and buildings.

The Data

Approach

This exercise will involve:

- Identifying venues within 3km of the Curzon Street terminal
- Analysing those venues

The analysis will involve:

- The density (number) of venues in the area
- The type of venues in the area

This will require grouping venues geographically to aid understanding (using statistical analysis, namely clustering)

Insight

The data will solve the problem by providing insight into the businesses in the vicinity of the Curzon Street terminal. This insight will include:

- The number of venues within the vicinity of the Curzon Street terminal
- The types of venues within the vicinity of the Curzon Street terminal
- The differences in venues (both numbers and types) between the areas surrounding the Curzon Street terminal

This information will help inform the audience which in turn will impact how the area surrounding the Curzon Street terminal in developed to help deliver the HS2 business case.

Data Requirements

The following data will be needed:

Object	Data Required	Use of Data	Source
Curzon Street Terminal Location	Latitude Longitude	To enable the area to be explored to be located	Internet: Latitude: 52° 28' 32.27" N Longitude: -1° 53' 6.55" W
Venue	Type of Venue	To enable the types of venues in the area to be identified	FourSquare: categories.name
Venue	Latitude Longitude	To enable venues within 3km of the Curzon Street terminal to be identified. To enable the location of each venue to be determined	FourSquare: venue.location.lat venue.location.lat

Venue	Postcode	To enable venues to be grouped by postal area	FourSquare venue.location.postalCode
Venue	District	To enable venues to be grouped by local district area	FourSquare venue.location.neighborhood

The FourSquare data used is set out in the Appendix.

Methodology

The investigation had 3 main stages:

1. Investigation on postcode and district (neighborhood)
2. Investigation by zones
3. Analysis of zones

Note the analysis utilised statistical analysis in the form of Clustering.

Investigation on Postcode and District

The following stages of analysis were undertaken:

1. The location of the Curzon Street terminal was plotted on the map
2. Venues in the neighbourhood of the Curzon Street terminal were retrieved from FourSquare
3. The venues were plotted on the map
4. A review of the data from FourSquare showed that :
   a. 100 records were recovered
   b. The neighborhood record was predominantly blank
5. Given the limited number of records there would be insufficient venues in each postcode to allow meaningful grouping
6. The postcode was split to identify the postal areas, however this again provided insufficient venues in each area to allow meaningful grouping

The conclusion of this investigation was that insufficient venues were retrieved by one call to FourSquare (100 venues) and that postcode and district did not provide a good method of grouping for analysis.

Investigation by Zones

The following stages of analysis were undertaken:

1. A set of zones were created covering an area 6km wide centred on the Curzon Street terminal
2. Each zone had a central location defined by it latitude and longitude and a radius of 300 meters
3. FourSquare data was retrieved for each zone, based on its central location

This retrieved a significantly larger number of venues around the Curzon Street terminals and the zones provided a meaningful way to group them to allow analysis.

Analysis of Zones

The zones were analysed in 2 ways:

1. Density of venues in each zone
2. Clustering of the zones

Density of Venues

The density of venues in each zone was calculated by:

1. Counting the number of venues in each individual zone
2. Colour coding the zones based on number of venues:

Number of Venues	Zone Colour
None	Zone not shown on map

Red	Less than 5
Orange	Less than 20
Green	Over 20

Clustering of Zones (Statistical Analysis)

The zones were clustered by:

- Taking the mean value of each category for venue for each location
- This data set was then used to identify the optimum number of clusters via an elbow test
- This data set was then used to cluster the zones
- The top 5 venue categories for each zone was then identified

The detail of the methodology can be found in the separate code documentation.
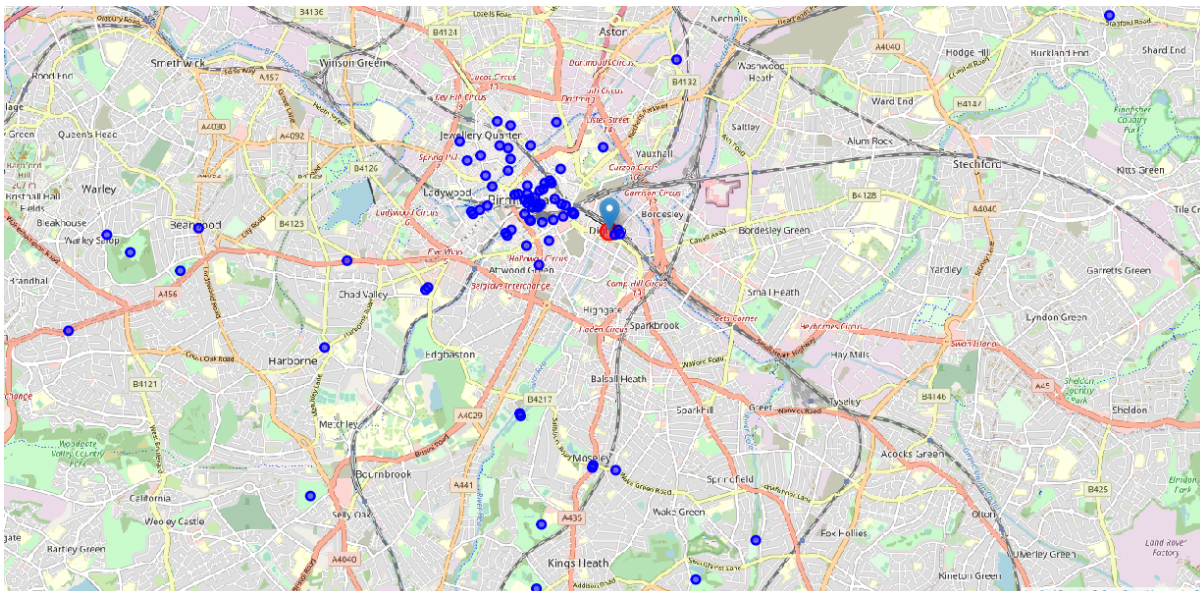
Results

The initial investigation focused on using postcodes and districts to analyse the venues. This was found to be ineffective and focus shifted to using zones.

Investigation By Postcode and District

The Curzon Street Terminal and nearby venues were plotted on a map:
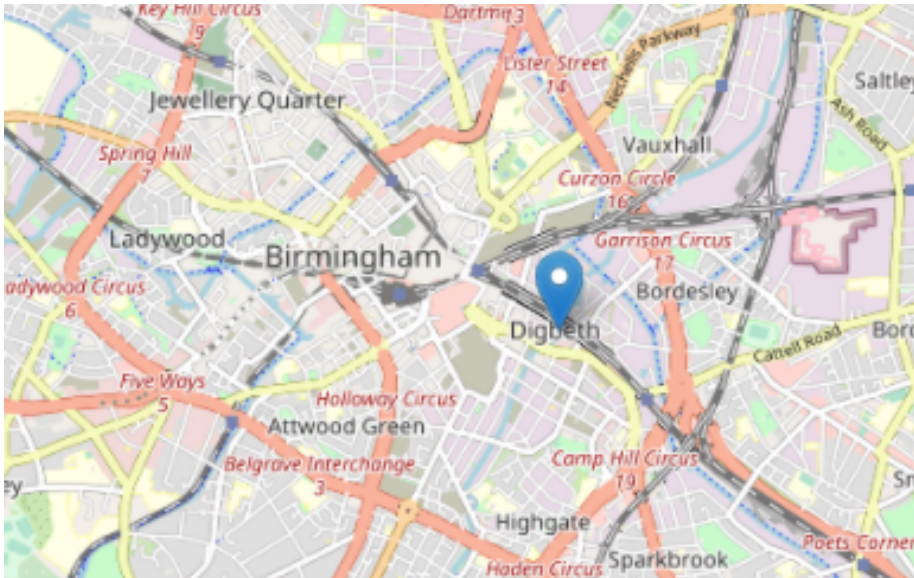
Data was retrieved from FourSquares and the locations plotted on the map.


This approach was not progressed due to:

- Limited number of venues retrieved from FourSquare
- The neighborhood not being provided for most venues
- The postcode and post district not providing an appropriate level of granularity for analysis (the post district was derived by splitting the postcode)

Analysis by Zones

The area within 3 km for the Curzon Street terminal was broken down into an hexagonal patter of circular zones 600 meters in diameter:
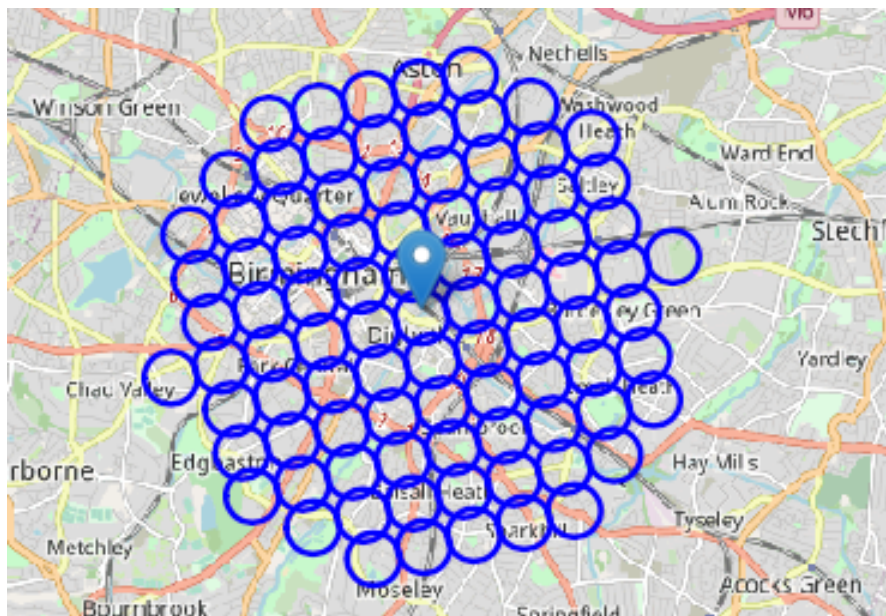
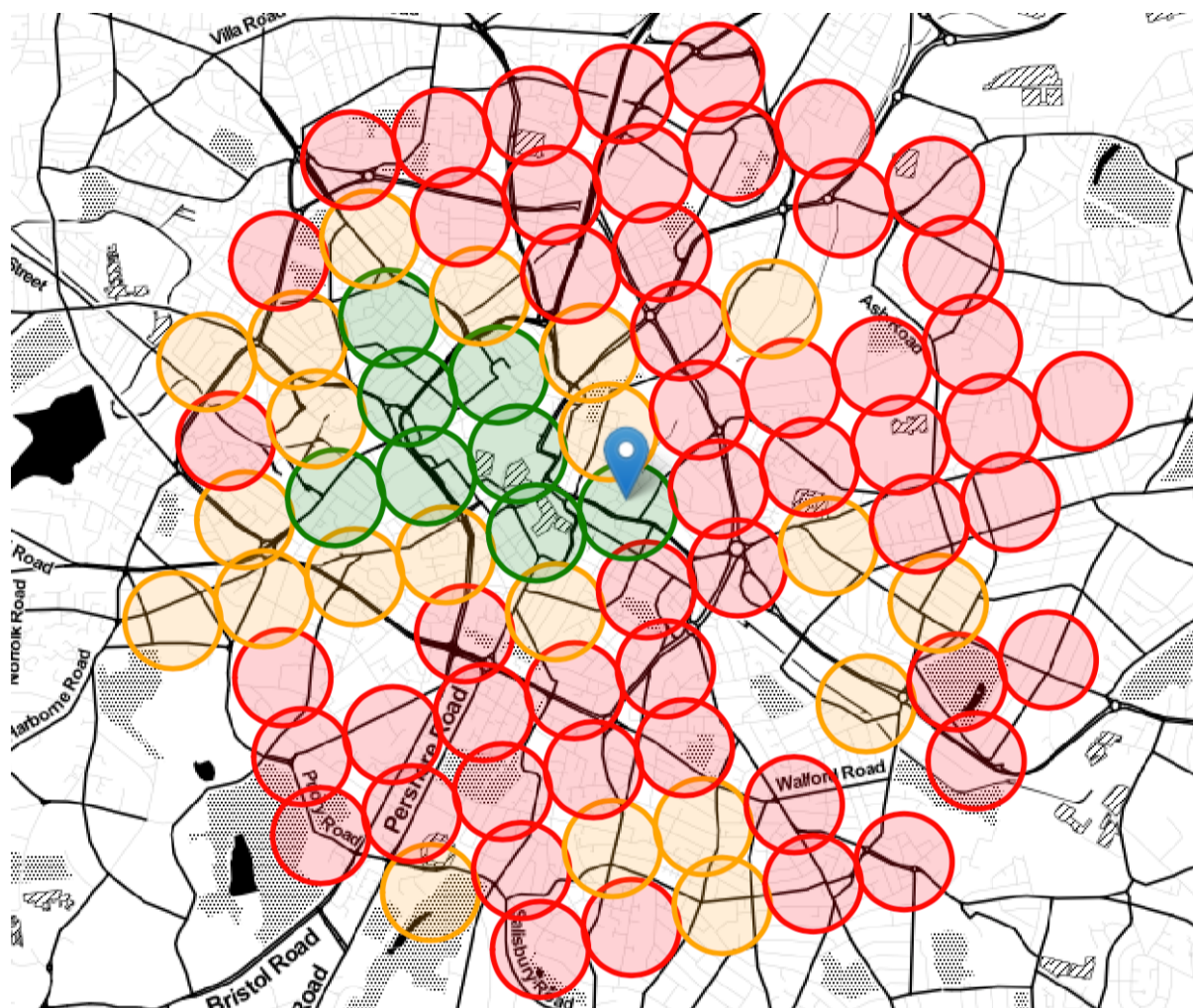Each of the 89 zones was given a unique reference number:
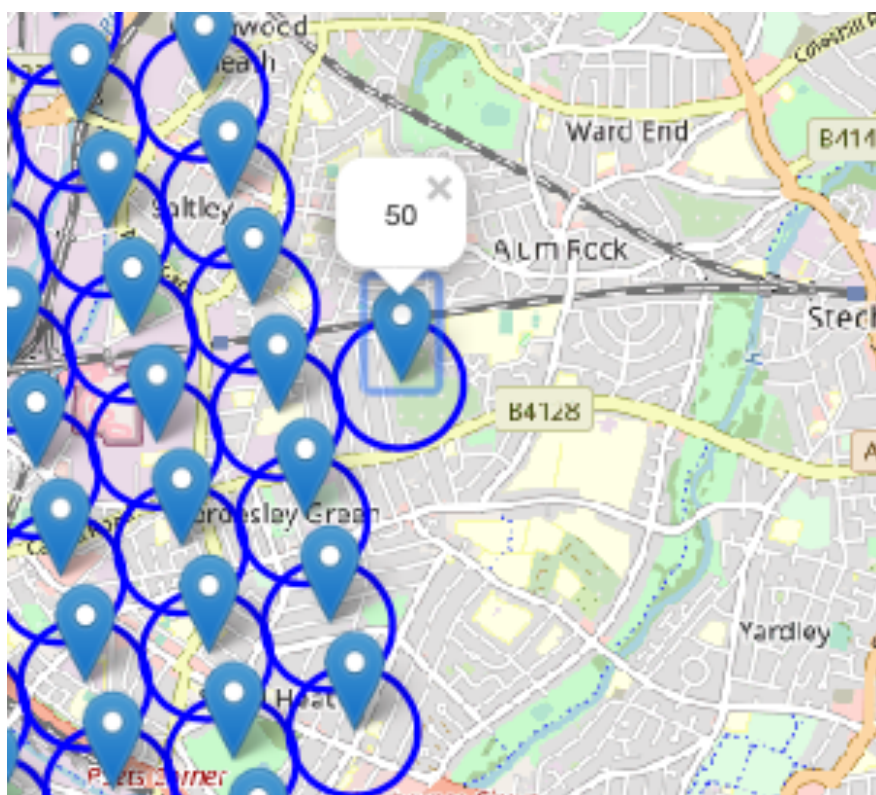
Venues were retrieved from FourSquare for each zone. In total 623 venues were identified.

Density of Venues
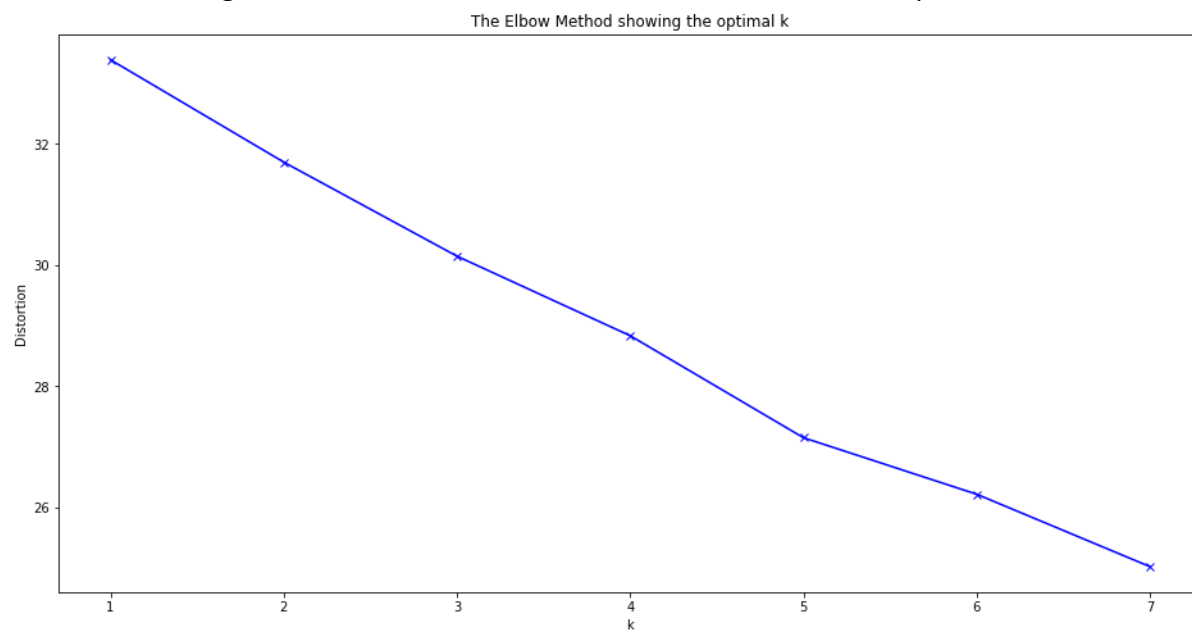
The density of venues in each zone was plotted on the map:

Number of Venues	Zone Colour
None	Zone not shown on map
Red	Less than 5
Orange	Less than 20
Green	Over 20

The map style was changed to remove any confusion caused by colour overlay.

Clustering by Zones

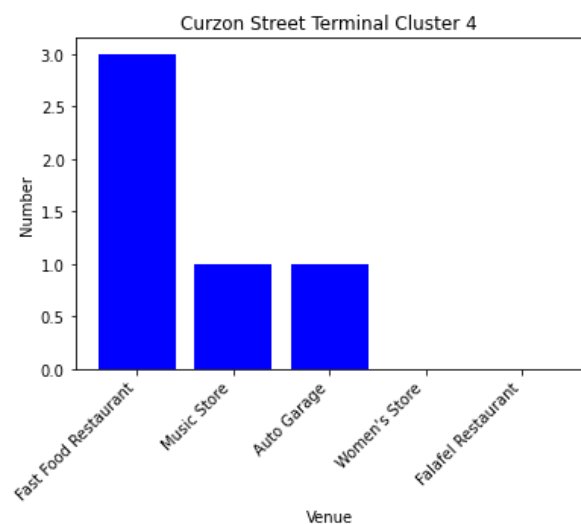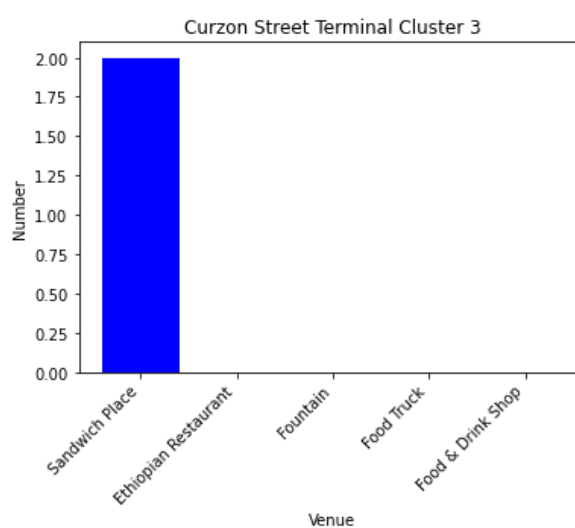Prior to clustering the zones an elbow test was run to determine the optimum number of clusters>
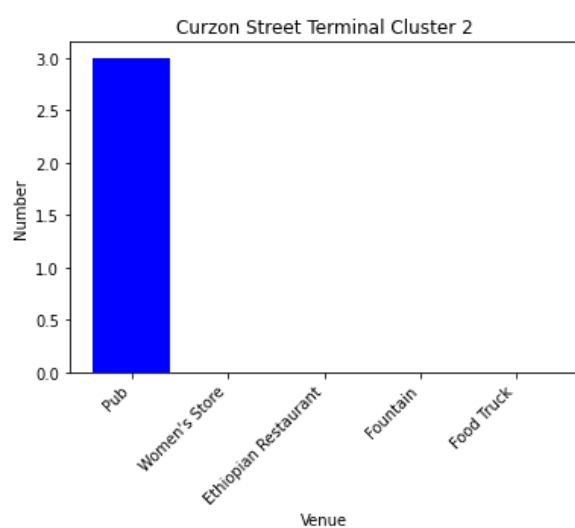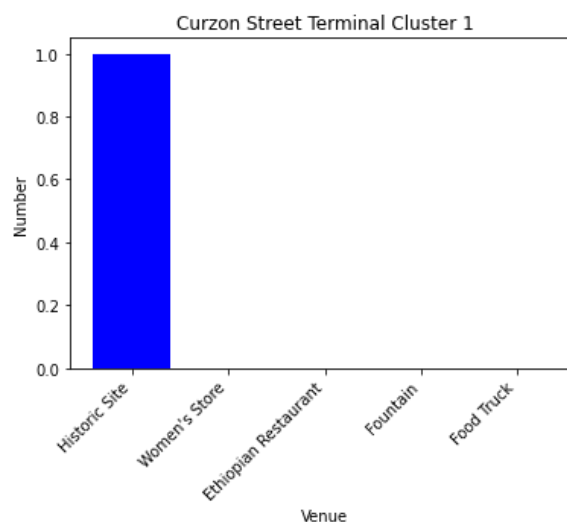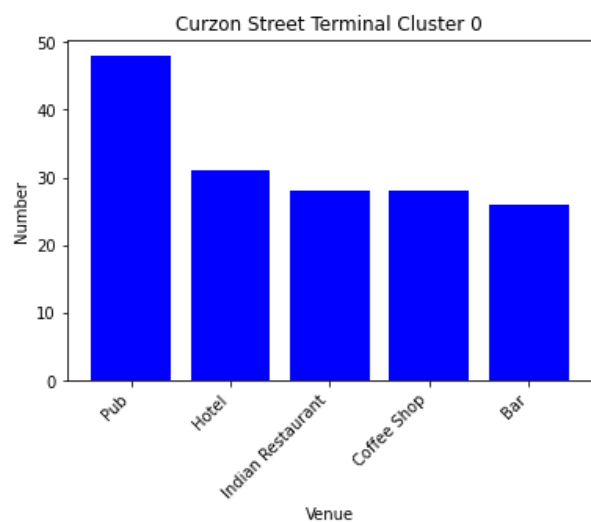

The optimum number of clusters (K) is where the slope levels out. In this instance it is 5.

The clusters were plotted on the map:

The characteristics of each of the clusters was then investigated by the top 5 venues in each.

Cluster	Colour on Map
0	Blue
1	Cyan
2	Yellow
3	Orange
4	Red

Observations and Recommendations

The analysis highlights the following:

1. The City Centre to the West of the Curzon Street terminal is well served with venues.
2. The areas to the North, East and South has more potential for investment in venues.
3. That some zones have no venues. The reason for this should be investigated.
4. That zones in cluster 0 already seem well places to serve the Curzon Street terminal.
5. The potential for development of zones in cluster 2,3 and 4 should be investigated.
6. The possibility of promoting local historic sights shown by cluster 1 should be investigated.

Density
The area to the West of the Curzon Street terminal is already well served with venues. The opportunity for investment in development and new businesses is to the North, East and South. As shown by the density of venues in the zones.

Zones with No Venues
A number of zones have no venues.

1. The characteristics of these zones should be investigated. This may be due to current usage e.g. school, park or large manufacturing plant.
2. The completeness of the FourSquare data should be investigated, potentially enriching or replacing it with other sources

Cluster 0 (Blue)
This cluster contains the largest number of zones, widely spread around the Curzon Street terminal. The venues all relate to hospitality and indicate that both people in transit, business travellers and tourists will be well served.

Cluster 1 (Cyan)
This consists of one historic site in one zone and is on the edge of the area under investigation. It does however highlight that the area around the Curzon Street terminal does have historic interest and this may be investigated further to promote tourism.

Cluster 2 (Yellow)
The two zones in this cluster only contain pubs. These zones may represent opportunities for development and new businesses to better reflect the range of businesses within the area of the Curzon Street terminal, as reflected in Cluster 0.

Cluster 3 (Orange)
Like cluster 2 this cluster consists on 2 zones, this time only containing sandwich shops. These zones may represent opportunities for development and new businesses to better reflect the range of businesses within the area of the Curzon Street terminal, as reflected in Cluster 0.

Cluster 4 (Red)
The 3 zones in this cluster might again represent areas for development and investment. Particularly in the zone adjacent to the terminal. The current venues including a music store and garage, would appear to serve the local community, rather than those using the Curzon Street Terminal.

Conclusion

The analysist shows that the area around the Curzon Street terminal is already well suited to support those in transit, tourists and business travellers. This shown by the large number of zones in cluster 0 which are focused on hospitality.

The zones in clusters outside zone 0 may provide investment opportunities for developers and businesses. This is because appropriate venues to exploit the opportunities presented by the Curzon Street terminal may be underrepresented in them.

Cluster 1 also highlights Birmingham's historic importance and this might be further investigated to encourage visitors.

Finally the quality of the FourSquare data should be considered and if appropriate augmented or replaced by different sources.

Appendix 1: Caveats

The following caveats should be applied to the analysis

1. The radial nature of the zones means some venues may not be covered by a zone
2. The radial nature of the zones means some venues may fall into more than one zone
3. The completeness of the FourSquare venue data for the area is not known

Appendix 2: FourSquare Record

Sample record from FourSquare, required data items highlighted

```
'reasons': {'count': 0,
   'items': [{'summary': 'This spot is popular',
      'type': 'general',
      'reasonName': 'globalInteractionReason'}]},
  'venue': {'id': '51d70128498ef30e94e9f569',
   'name': 'Digbeth Dining Club',
   'location': {'address': 'Spot*light Unit 2',
    'crossStreet': 'Lower Trinity St',
    'lat': 52.475385298837345,
    'lng': -1.8829814431038692,
    'labeledLatLngs': [{'label': 'display',
       'lat': 52.475385298837345,
       'lng': -1.8829814431038692}],
    'distance': 149,
    'postalCode': 'B5 6DY',
    'cc': 'GB',
    'neighborhood': 'Digbeth',
    'city': 'Birmingham',
    'state': 'West Midlands',
    'country': 'United Kingdom',
    'formattedAddress': ['Spot*light Unit 2 (Lower Trinity St)',
     'Birmingham',
     'West Midlands',
     'B5 6DY',
     'United Kingdom']},
   'categories': [{'id': '53e0feef498e5aac066fd8a9',
     'name': 'Street Food Gathering',
     'pluralName': 'Street Food Gatherings',
     'shortName': 'Street Food Gathering',
     'icon': {'prefix': 'https://ss3.4sqi.net/img/categories_v2/food/streetfood_',
      'suffix': '.png'},
     'primary': True}],
   'photos': {'count': 0, 'groups': []}},
  'referralId': 'e-0-51d70128498ef30e94e9f569-0'},
```