CRIME HOTSPOTS

Using data and machine learning

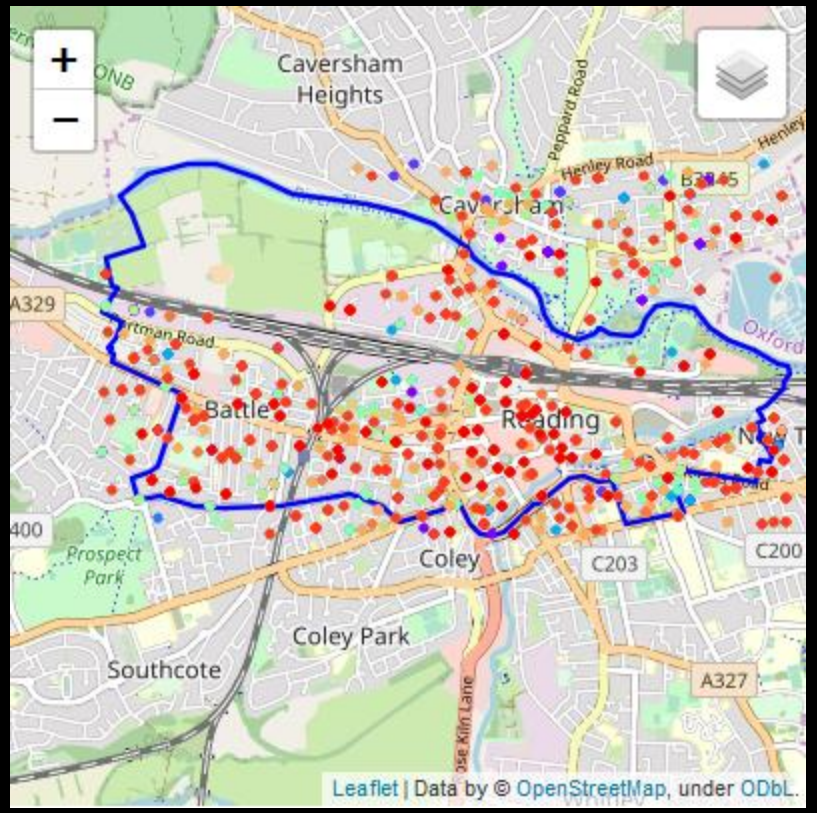
HOTSPOTS – WHERE IS CRIME GOING TO OCCUR

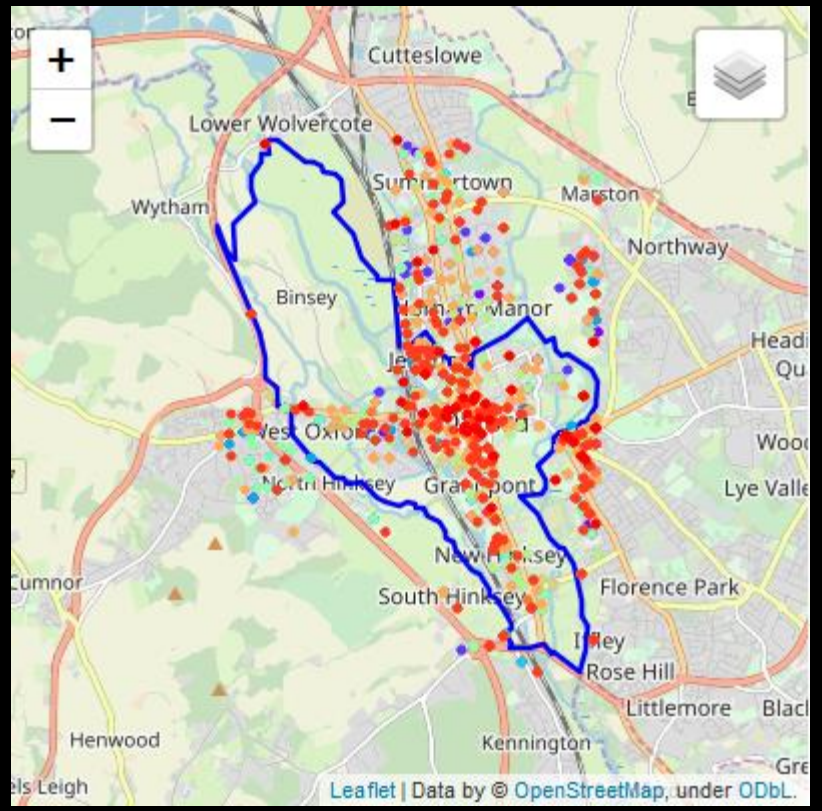
- We know where crimes have occurred
- We also know what venues such as bars, shops are near to those crime events
- What if we could gain insights into what areas will be prone to crimes based solely on the kinds of nearby venues?

HOTSPOTS – CRIME DATA

- The country is split into police forces.
- A police force is split into areas defined by a boundary
- The crimes committed in that area can be obtained by Rest APIs from the central police database
- We download the crime data for a period of months for each area
  - For a period of 9 months over 6000 unique crime events per area
- The area is split into square cells – for each cell we search the crime events to see which occurred within a small distance from the cell's center

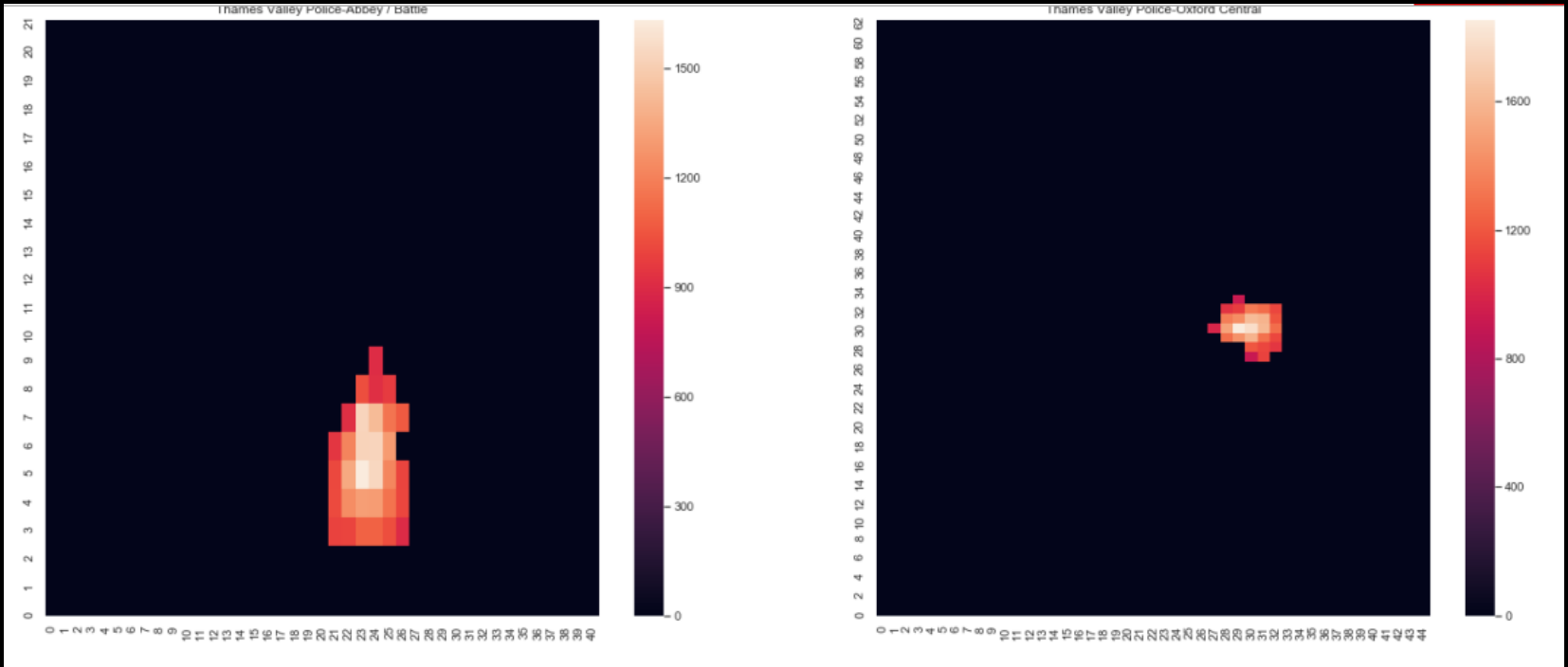
DETERMINING HOTSPOTS (1)





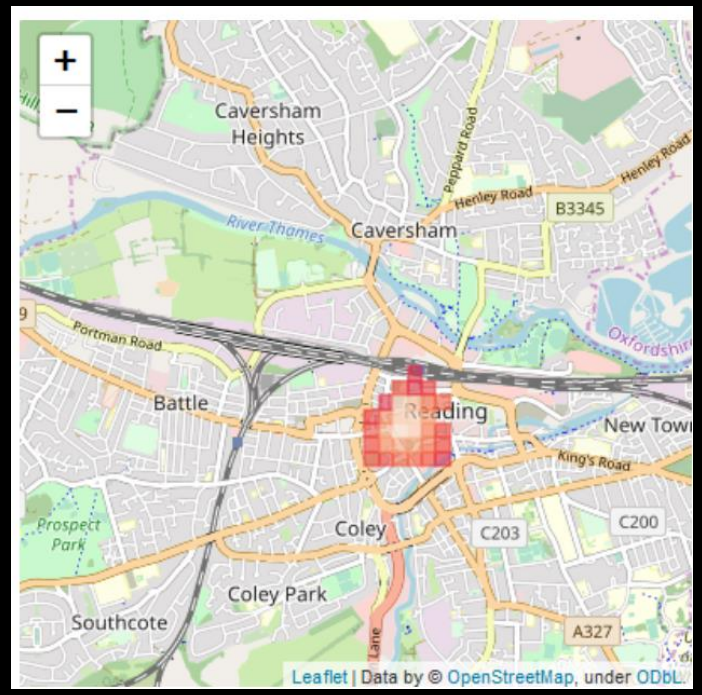
DETERMINING HOTSPOTS (2)

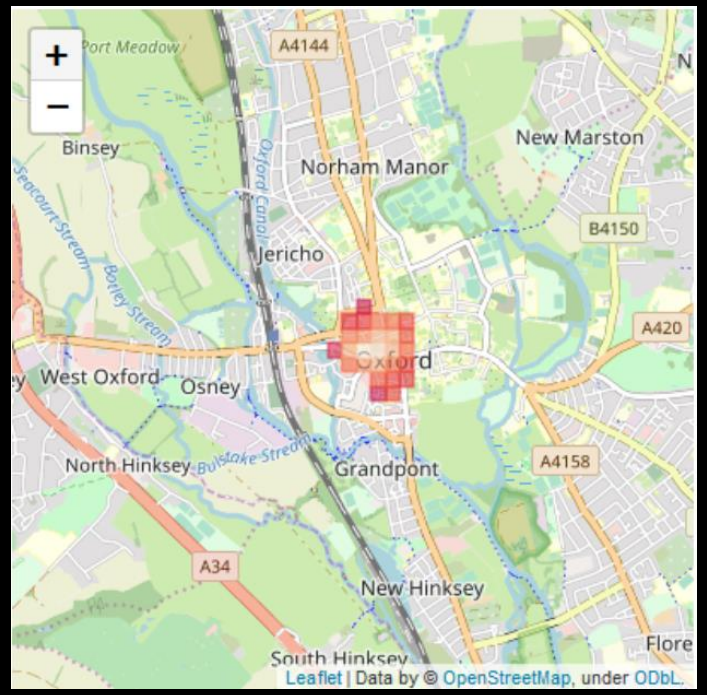
- For each cell add up the total crimes nearby.
- If the total is above a threshold then that cell is a hotspot



DETERMINING HOTSPOTS (3)

- For each cell add up the total crimes nearby.
- If the total is above a threshold then that cell is a hotspot





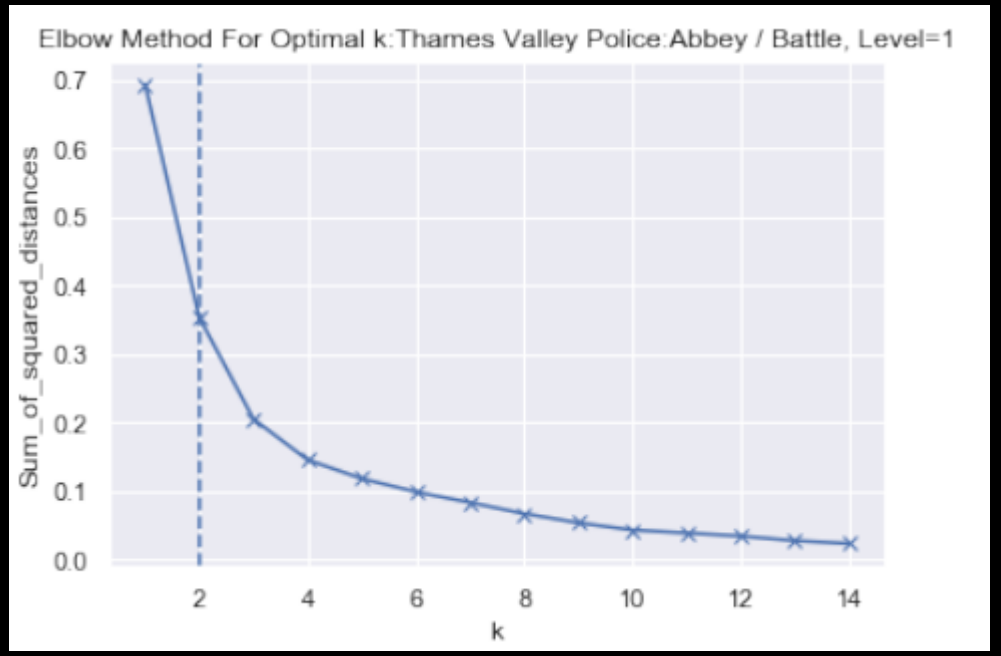
VENUE DATA

- Each cell has a set of venues nearby.
- Use the FourSquare API to find these.
- Assign numerical levels to each venue category – enables an extract of categories for each level

```
1 Outdoors & Recreation
|-- 2 Plaza
1 Food
|-- 2 Bakery
|-- 2 French Restaurant
|   |-- 3 Burgundian Restaurant
|-- 2 Indian Restaurant
|   |-- 3 Andhra Restaurant
|   |-- 3 Awadhi Restaurant
|-- 2 English Restaurant
```

OPTIMAL K-MEANS

- Across the possible venue categories, use one-hot encoding and frequencies to create a table for the hotspots
- Use K MEANS for clustering these venue frequencies across a range of cluster numbers
- Use the elbow method, convexity, to determine the optimal K (cluster number) for each level



ANALYSE THE CLUSTERS

Region 0, Level 1

Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Food	Nightlife Spot	Shop & Service	Travel & Transport	Outdoors & Recreation
0	Food	Nightlife Spot	Travel & Transport	Shop & Service	Outdoors & Recreation
0	Food	Shop & Service	Nightlife Spot	Outdoors & Recreation	Travel & Transport
0	Food	Shop & Service	Nightlife Spot	Travel & Transport	Outdoors & Recreation
0	Food	Shop & Service	Travel & Transport	Nightlife Spot	Outdoors & Recreation
1	Food	Shop & Service	Nightlife Spot	Travel & Transport	Outdoors & Recreation
1	Food	Shop & Service	Travel & Transport	Nightlife Spot	Outdoors & Recreation
1	Food	Travel & Transport	Nightlife Spot	Shop & Service	Outdoors & Recreation
1	Food	Travel & Transport	Shop & Service	Nightlife Spot	Outdoors & Recreation

Region 1, Level 1

Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Food	Nightlife Spot	Shop & Service	Outdoors & Recreation	Arts & Entertainment
0	Food	Shop & Service	Nightlife Spot	Arts & Entertainment	Outdoors & Recreation
0	Food	Shop & Service	Nightlife Spot	Arts & Entertainment	Travel & Transport