



Data Science Capstone Final Project

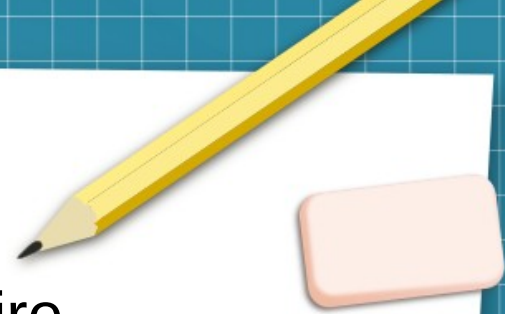
Clustering wards in Birmingham

Problem Definition

It is known that sewer blockages correlate to the amenities present in a local area. This is demonstrated in issues such as fatbergs as shown below, often downstream of restaurants.



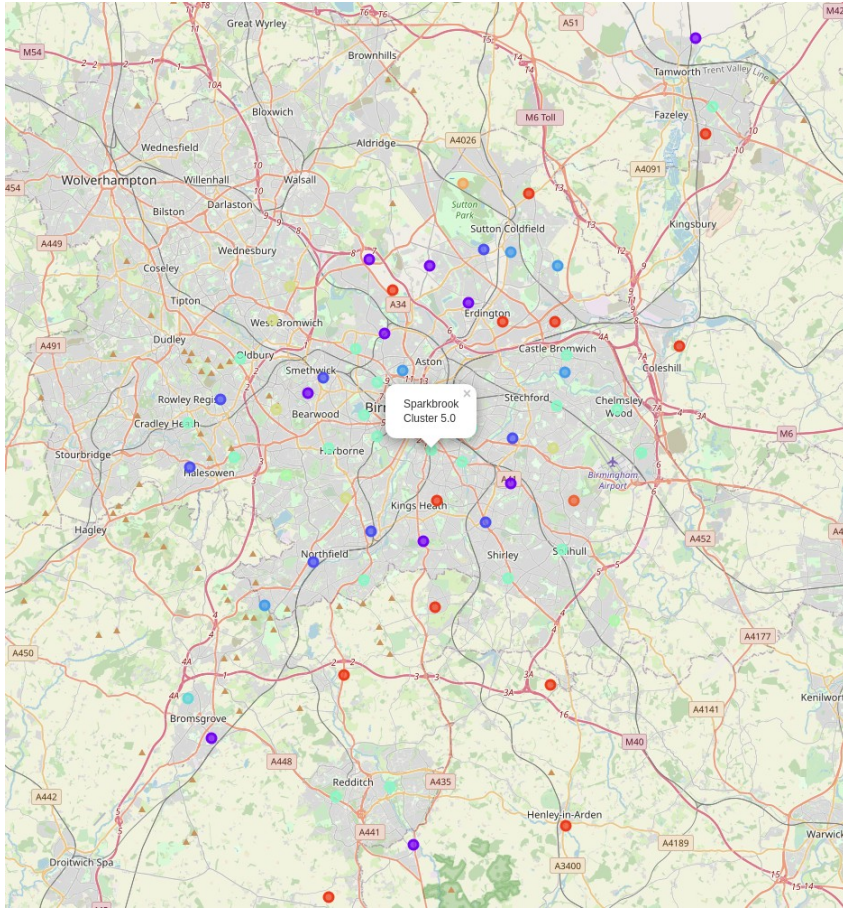
The question



In order to combat this issue and to prevent the dire consequences of blockages Sewage companies are looking to roll out a trial of blockage monitoring. We will use Birmingham as an example for this exercise.

The big question – Where should be put the monitors?

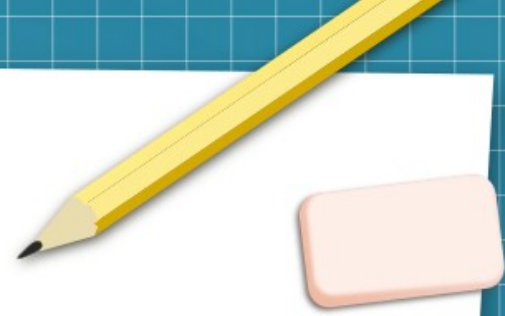
The Answers



The results from the clustering shows that areas known to be high in restaurants and takeaways tending to exist in cluster 5.0. A test case for this was the ward of Sparkbrook which is part of Birmingham's famous Balti Triangle. The more rural and suburban areas were also clustered together as red and purple. The image shows the categorisation of wards across the Birmingham area.

Further work

- Including actual Water Company data
- Including demographic and societal factors
- Using alternative forms of clustering
- Testing the clustering on a dataset from alternative areas





This work is licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.
It makes use of the works of Mateus Machado Luna.

