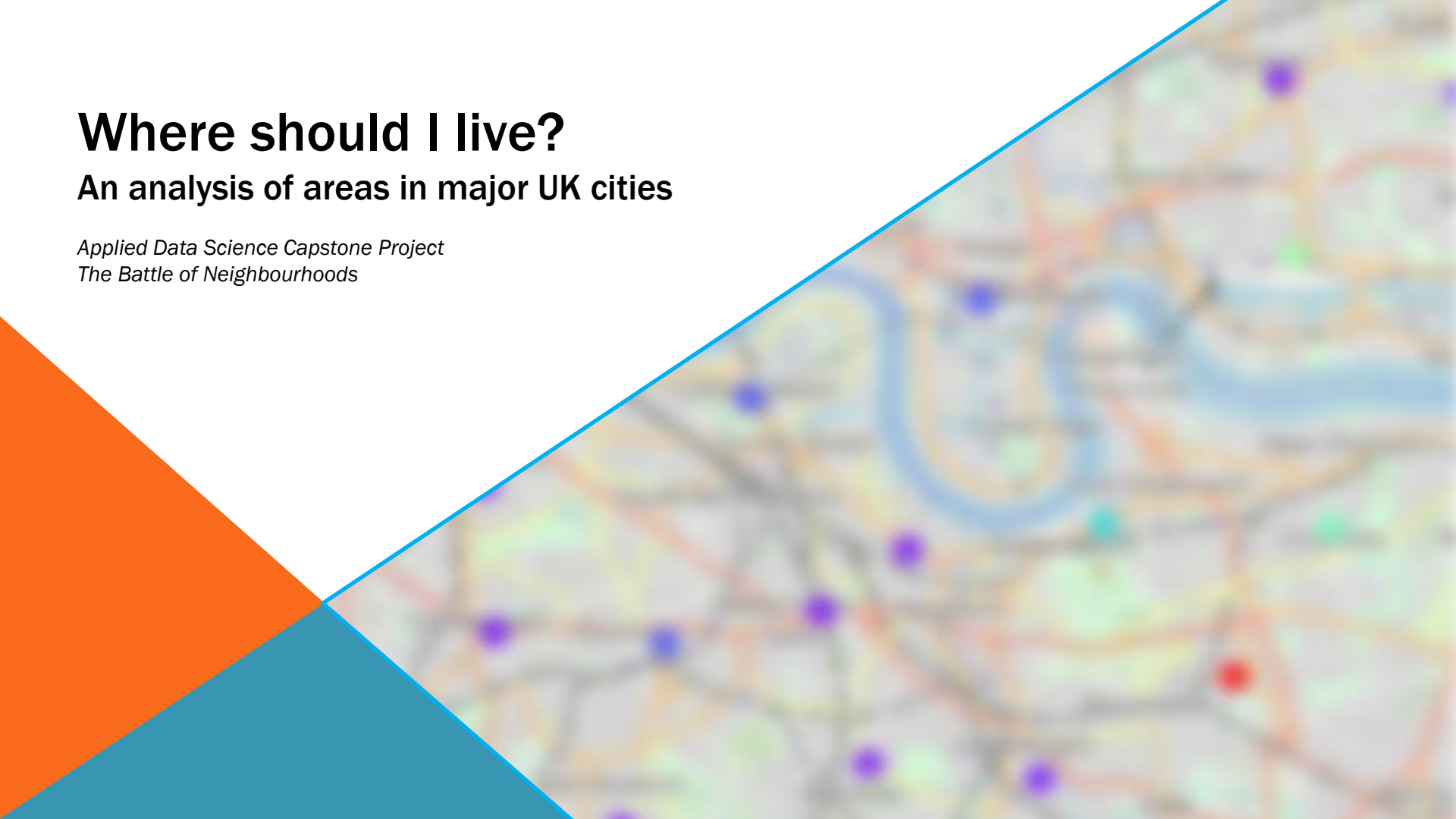
Where should I live?

An analysis of areas in major UK cities

Applied Data Science Capstone Project
The Battle of Neighbourhoods



HOUSES ARE EXPENSIVE

- It is well known that it is becoming increasingly difficult to afford a house in the UK.
- House prices are particularly high in southern cities such as Oxford.
- Young people living in these cities can often not afford to buy a property in their hometown.

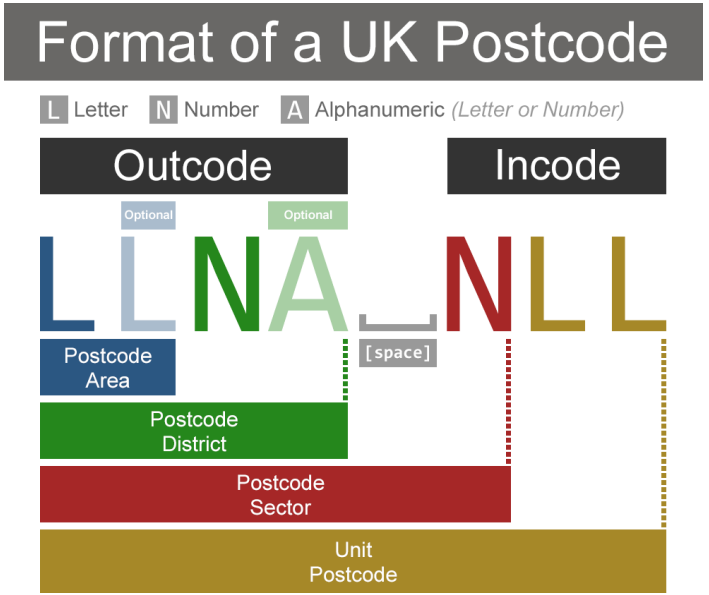
PROBLEM:

Where could someone move that would offer a similar lifestyle to their current location?

- This project aims to group areas of different cities across the UK based on access to venues, in order to recommend cheaper places to buy houses than the user's current location.

DATA ACQUISITION

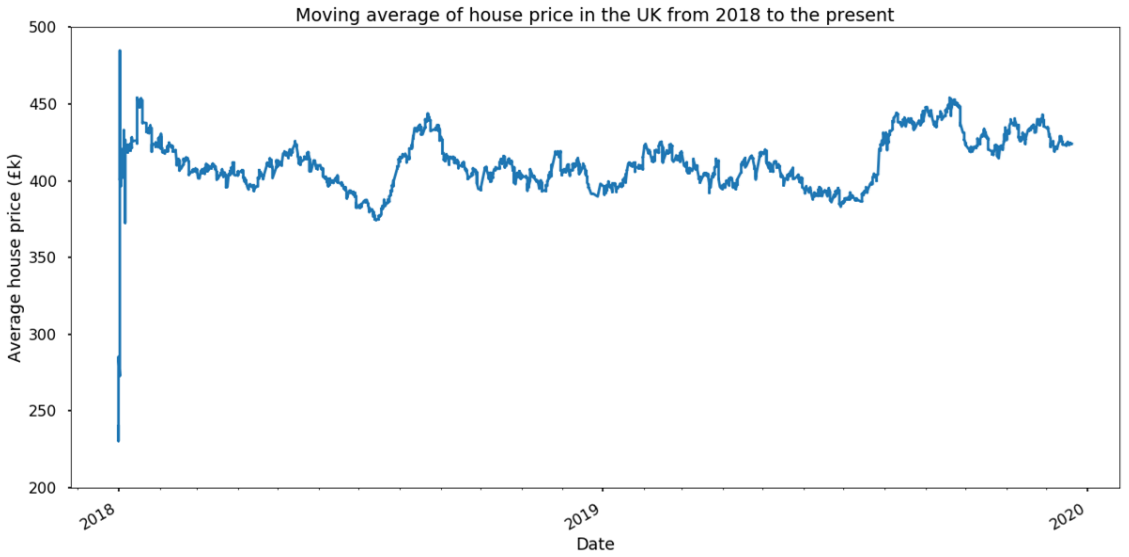
- Price data were obtained from the HM land registry open data Price Paid dataset.
  - All transactions between January 1 2018 and February 22 2020 for a selection of 11 cities.[1]
  - Freehold only.
  - Prices were averaged by **postcode district.**
- Positional coordinate data for each district were obtained from the Code-Point Open postcode dataset.
- Venue data were obtained using FourSquare.



[1]BIRMINGHAM, BRISTOL, CARDIFF, LEEDS, LIVERPOOL, LONDON, MANCHESTER, MILTON KEYNES, NEWCASTLE, OXFORD, YORK

AVERAGE PRICE OVER THE DATASET PERIOD

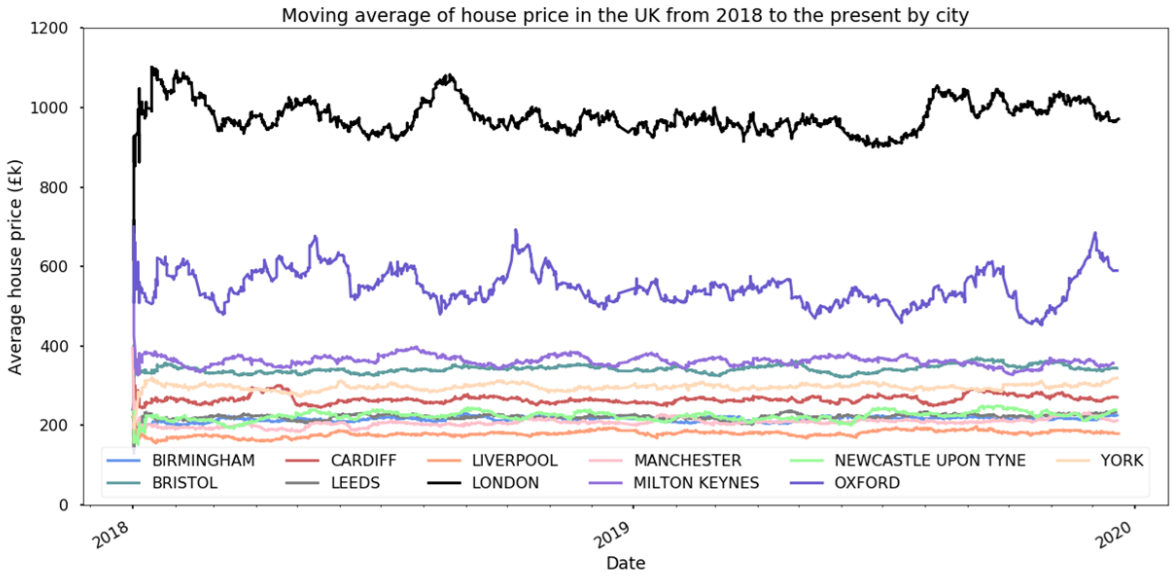
The 30-day moving average price was relatively stable over the period of the dataset, with a spike around August of 2019.



AVERAGE PRICE OVER THE DATASET PERIOD

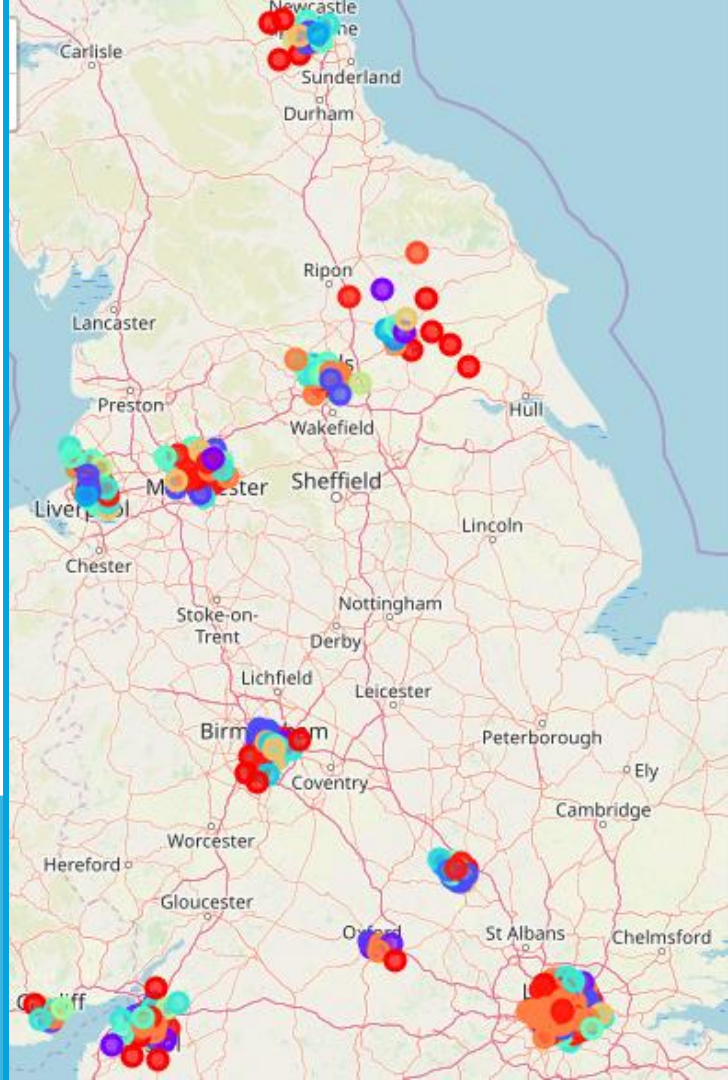
A comparison of cities shows that London is the most expensive, with an average house price of **£1,000,000.**

Oxford is also an outlier, with the other cities coming in between £200,000 - £400,000.



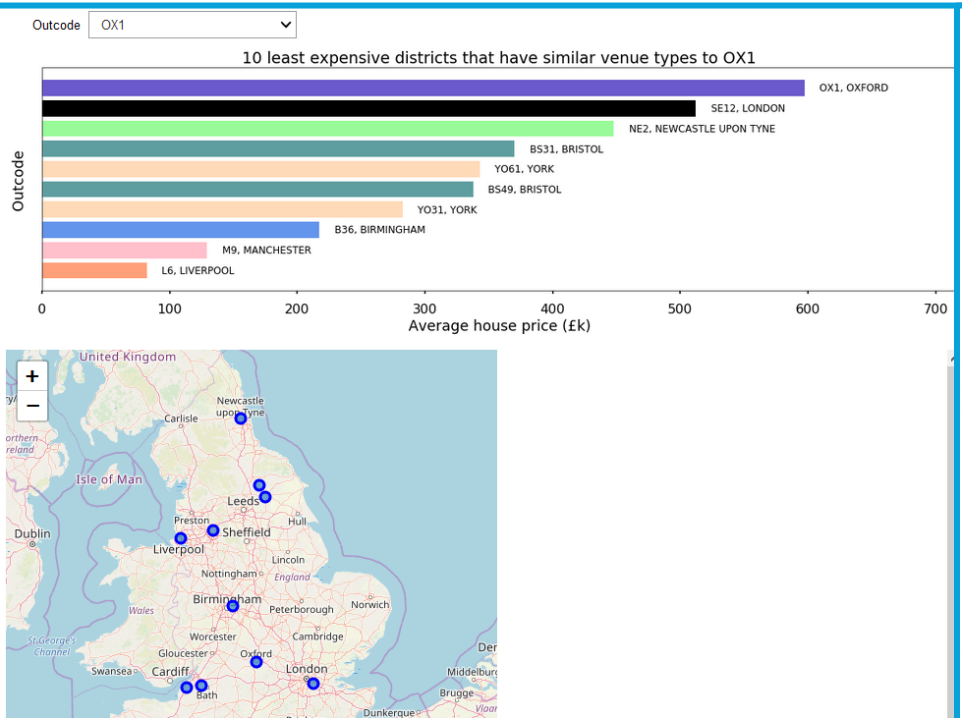
CLUSTERING

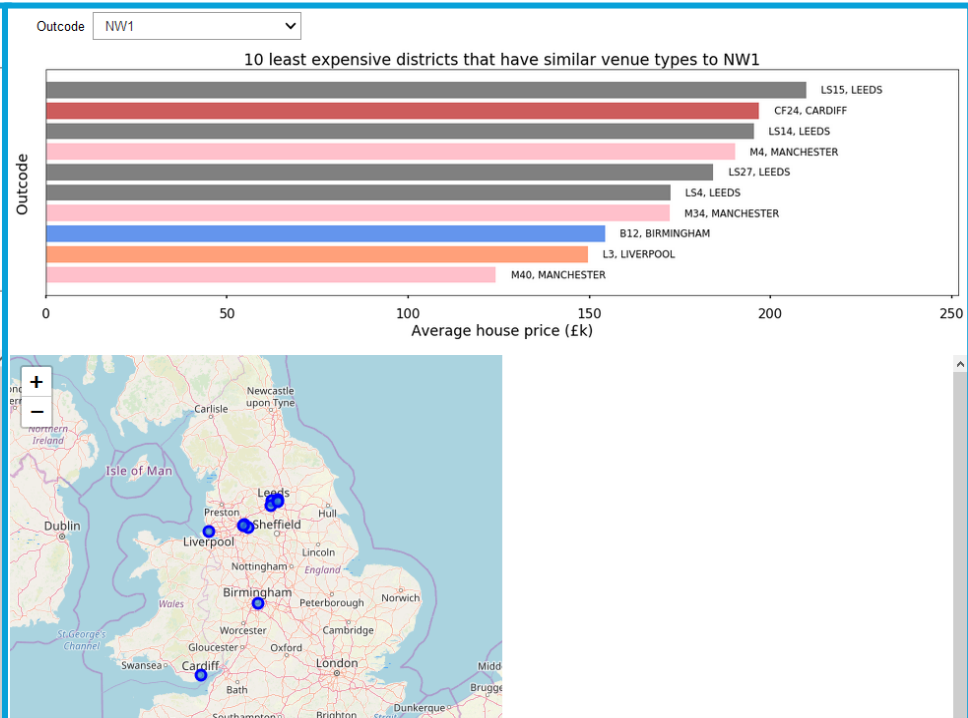
- For each district, the top 100 venues in a 1 km radius were obtained using Foursquare.
- The types of venue were analyzed and used to cluster the districts using a K-means clustering algorithm.



INTERACTIVE TOOL

- The clustered data were then input into an iPython widget to create a tool to allow the user to input their postcode district, and obtain a graph displaying the 10 least expensive districts that shared a cluster





DISCUSSION AND CONCLUSIONS

- Our initial exploration of the data confirmed the premise of the analysis, i.e. southern cities are significantly more expensive places to buy houses
- The tool generated is a good ‘proof of concept’, suggesting cheaper districts to users based on their home district.
- Clusters are likely not that accurate due to varying size of postcode districts.
  - Future improvements to the tool would use the Code-Point with polygons dataset, allowing density of venue to be calculated.
  - The full UK transaction dataset would also be used.
- If you want to be able to afford a house, you should probably **move North.**