



Exploring areas of London

With machine learning and Python

Why London?

- One of the largest european cities
- 9 mil population
- Very diverse
- And why not?



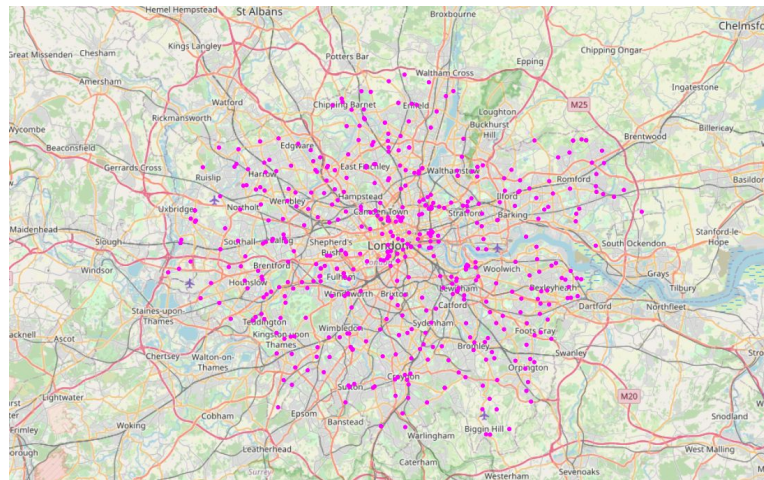


Data sources

- [Wikipedia](#) for areas of London
- [Geocoder](#) + Arcgis for coordinates
- [Foursquare](#) for venues

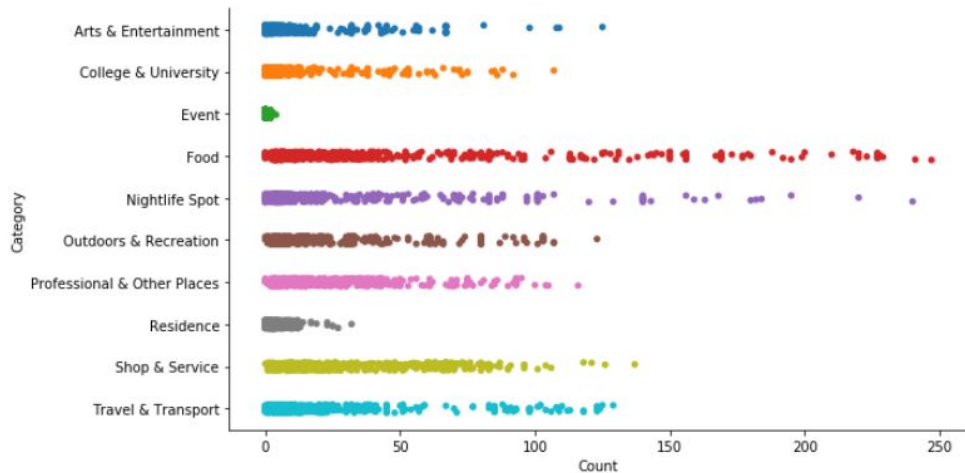
The areas

- Pandas `read_html` makes life easy
- Some data cleaning was necessary for geocoding addresses
- 533 locations

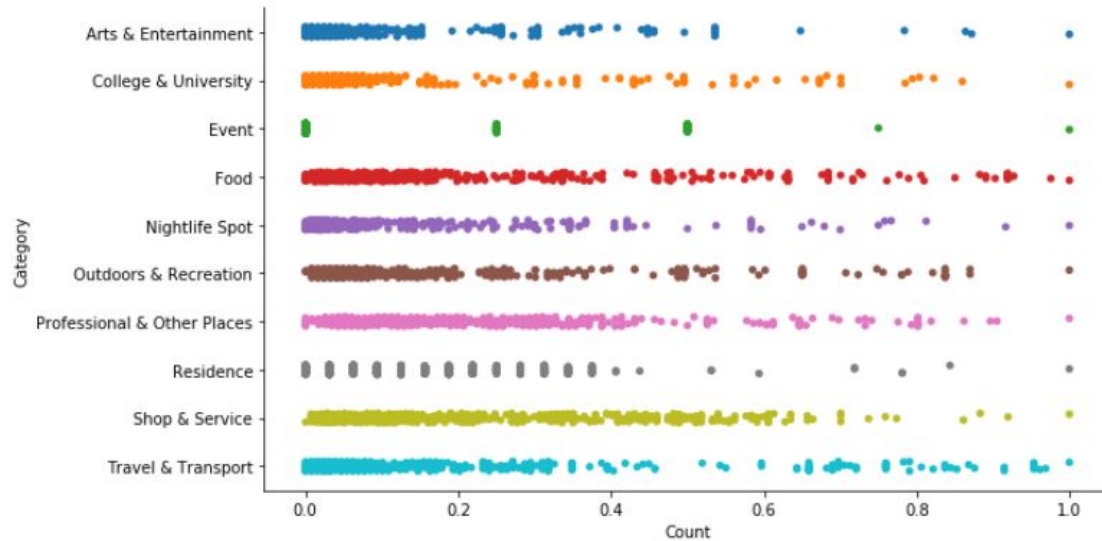


The venues

- Foursquare Places API
- Queried every location for main categories to get total counts

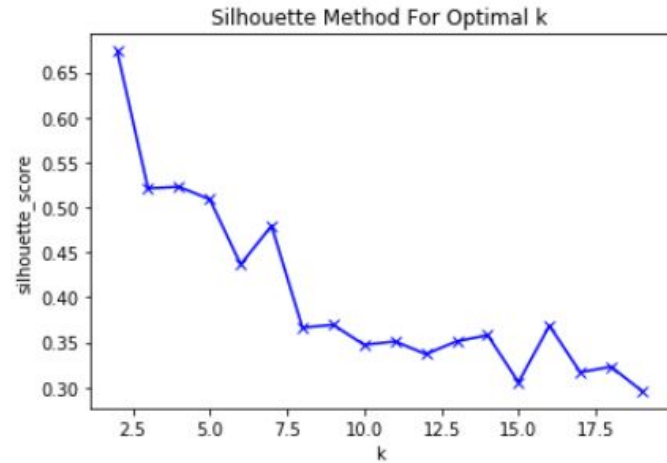


Data normalization



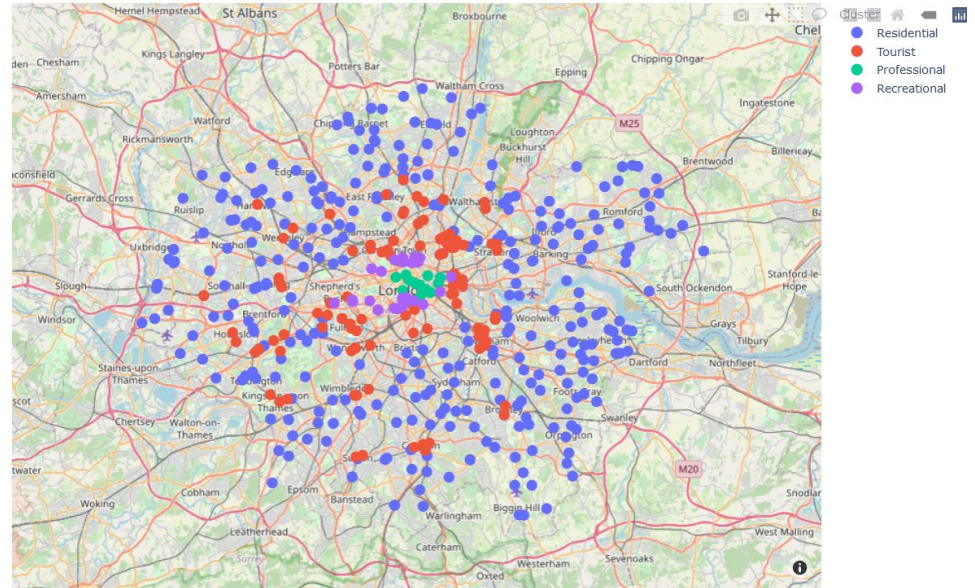
K-Means

- Silhouette method
- $k = 4$



Clusters found

- Touristic
- Recreational
- Professional
- Residential





What's next?

- Try a more tailored selection of Foursquare sub-categories
- Use venue “visits per day” as a factor
- Experiment with other possibilities for K

Thanks!

