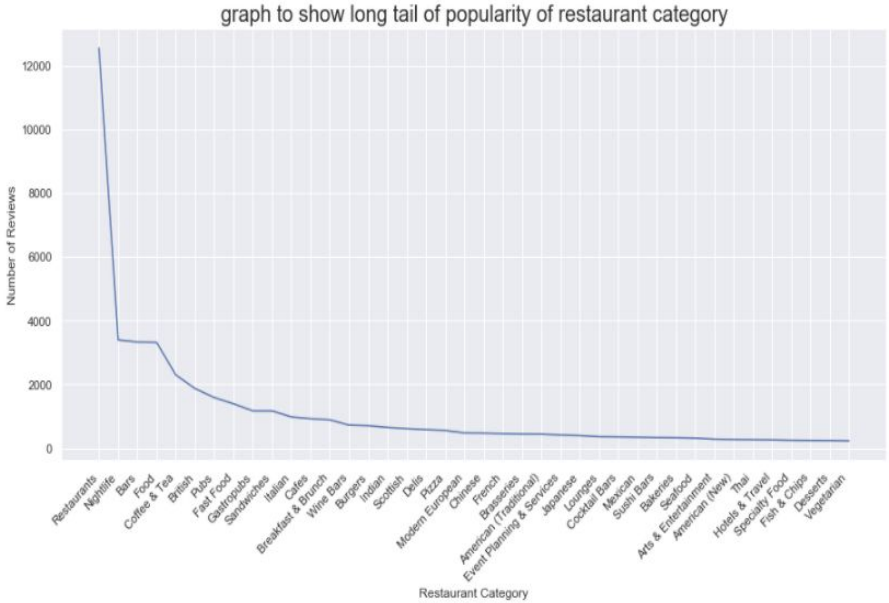
Restaurant Recommendation System

Capstone Data Science Project - James Moulds

Aims and Objectives

- Produce a restaurant recommendation system for a new lifestyle and food startup.
- The system should be suitable for new and existing users to engage with.
- The startup is based in Scotland so would like to run a test product in Edinburgh.
- Use a variety of techniques to tailor recommendations to a specific users preferences



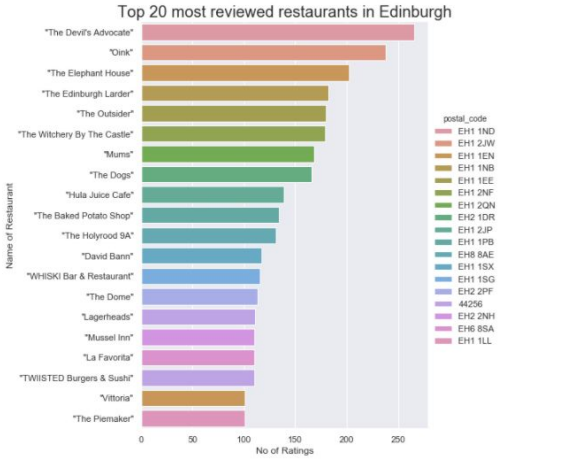
The Data

We have data extracted from Yelp reviews and filtered for the Edinburgh area. The key numbers are as follows:

- 28407 reviews
- by 9162 active users
- For 1605 restaurants

Data challenges include:

- Only 4% of users have racked up more than 10 reviews.
- How to access less popular but well rated restaurants - the hidden gems.



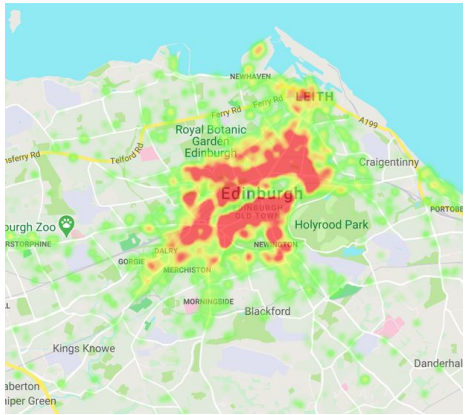
The Product

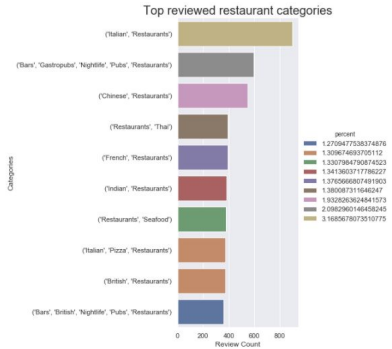
- For our MVP we have used a variety of differing text and ratings based models to achieve a basic product.
- Our current error is around 1 which is best illustrated by the table left which shows our models predictions for existing users.
- The current model is suffering from sparse data which we hope to improve upon in the next phase by focussing on different user groups and preferences.

Table to show best model predicted ratings against actual user ratings

	user_ref	business_ref	stars	Prediction	Difference
0	246	68	5	4.03	0.97
1	146572	68	4	4.22	0.22
2	211139	68	3	4.01	1.01
3	260271	68	4	3.86	0.14
4	491834	68	5	3.60	1.40
5	1168645	68	4	3.76	0.24
6	35710	380	5	4.01	0.99
7	77088	380	4	3.99	0.01
8	438436	380	5	4.06	0.94
9	590862	380	4	3.78	0.22

The Next Steps





Use the geography of the city to locate restaurant hotspots to direct people based on location.

Look deeper into the review text available for each restaurant.

Use keyword weightings to recommend based on user preferences.

Deploy the recommender on a user friendly and accessible interface such as Google Streamlit.