

Evaluating location candidates for BYO Ltd. in Glasgow

Applied Data Science Capstone Project

Introduction

Background

BYO Ltd is a startup company, aiming to launch modern restaurants/bars/co-working spaces. Management is currently evaluating the options to open the first restaurant/bar/co-working space in Glasgow, Scotland.

Restaurants/bars/co-working spaces opened by BYO Ltd. target professionals between 20 and 45. During the day the location serve as co-working/event and meeting space. In the evening, it turns into a restaurant and bar. Although the cocktails are excellent, BYO Ltd. offers clients with a paid subscription the possibility to bring their own drinks to the restaurant/bar. As part of the business model, (paying) customers can use the BYO App to inform staff about the intent to bring wine/spirits to the bar. BYO Ltd.'s staff will schedule a pick-up time and fetch the wine/spirits from the client's house (or any other location requested). This makes it easier to enjoy their favorite drink in their favorite restaurant, without the hassle of having to carry it. Further premium options are available on request, e.g. requesting BYO Ltd.'s staff to source a specific bottle of wine or any other alcoholic beverage.

Problem

Selection of location candidates was (and partially still is) driven by suggestions of location scouts, realtors and business stakeholders. If at all, only a few, selective data points are included in the decision making process. Business logic ("rules") are rarely specified and the actual selection process remains a 'dark art'.

This project aims to combine the need for a reliable data basis for decision making with the application of (pre-defined) business rules/criteria. This should result in a ranking and recommendation of the most suitable location candidate in Glasgow.

Location candidates:

- Óran Mór
- Park Circus
- Merchant City
- Craington

Target Audience / Interest

Mainly, BYO Ltd.'s board of directors is the primary target group. But also other start-ups interested in opening a venue in Glasgow may take interest in this analysis. Besides that, residents and visitors may find it useful to use some of the heatmaps to plan their night out.

Data acquisition, preparation and cleaning

Data Sources

Foursquare [[link](#)] was used as main data source for this analysis. Rather than using the complete data set available, data extraction focussed on certain categories and sub-categories. For a complete list of the categories, please click [here](#).

Data preparation and cleaning

Data was downloaded from Foursquare via API calls on a per category group basis (coworking spaces, nightlife spots, restaurants, and parks/recreation). These were transformed into DataFrames and merged into one dataset.

Data cleaning was (surprisingly) straight-forward and only required removal of some duplicates, in case venues were listed in multiple categories. I decided to use only one category per venue, as the effort to address multiple categories per venue would not have been justified at this point.

Feature selection

I decided to extract two main features - location and category.

Based on the location, an additional feature - the distance to each location candidate - was calculated. In cooperation with business stakeholders, two distance features - called 'class 1' (500m) and 'class 2' (1000m) - around each location were also defined.

Categories were consolidated into an additional feature - category types. The category types could take four different values : competitor, nightlife, restaurant, and parks/recreation.

Exploratory Data Analysis

Mapping all available data out provided a colourful, yet inconclusive overview.

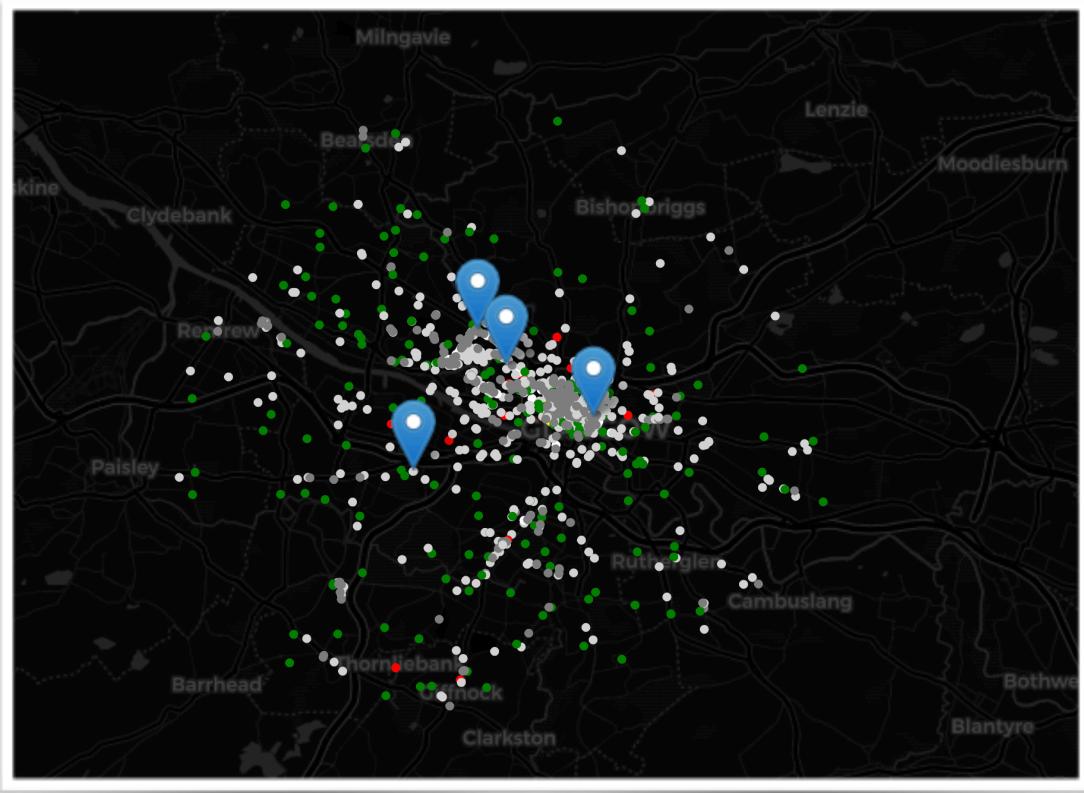


Figure 1: Map of Glasgow with all data

Therefore I began to breakout down by category type, to have a helpful visual tool when discussing the results with business stakeholders.

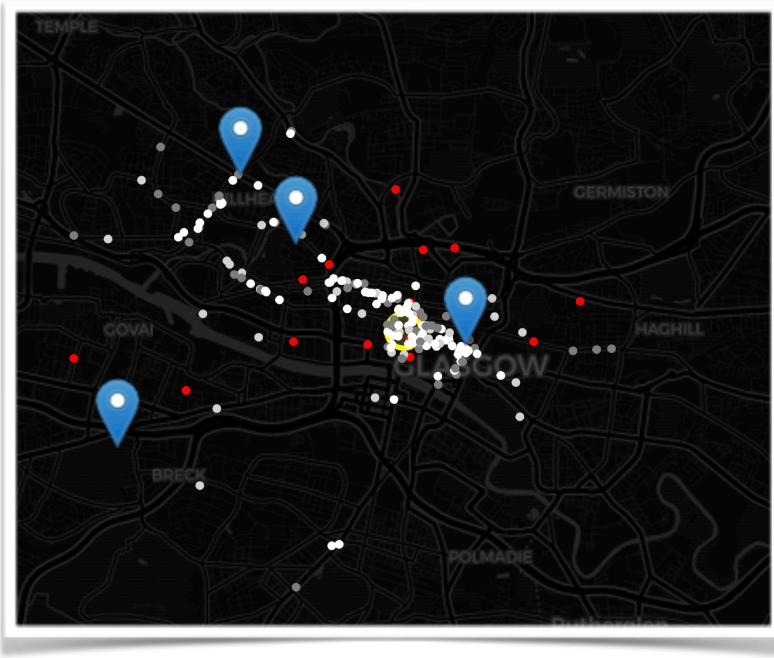


Figure 2: Map of coworking spaces (red) and other competitors (white/grey)

This is a slightly improved display of the information, but does not offer the needed guidance to have an informed conversation. Neither does it make use of all the features available. Hence I decided to display the information (on a per category type basis) with heatmaps, including the calculated feature of distance ‘class 1’ and ‘class 2’.



Figure 3: Heatmap of coworking spaces with class 1(green) and class 2 (yellow)

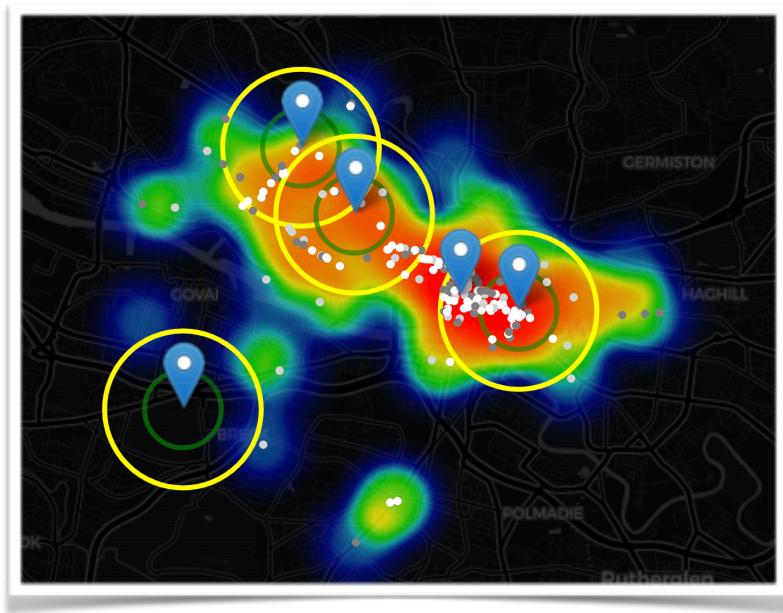


Figure 4: Heatmap of nightlife spots with class 1(green) and class 2 (yellow)

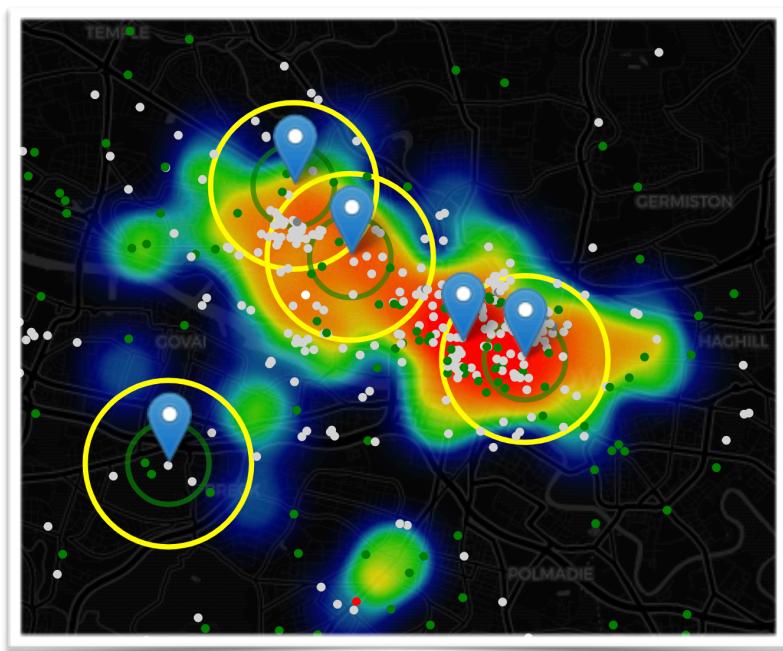


Figure 5: Heatmap of parks/recreational venues with class 1(green) and class 2 (yellow)

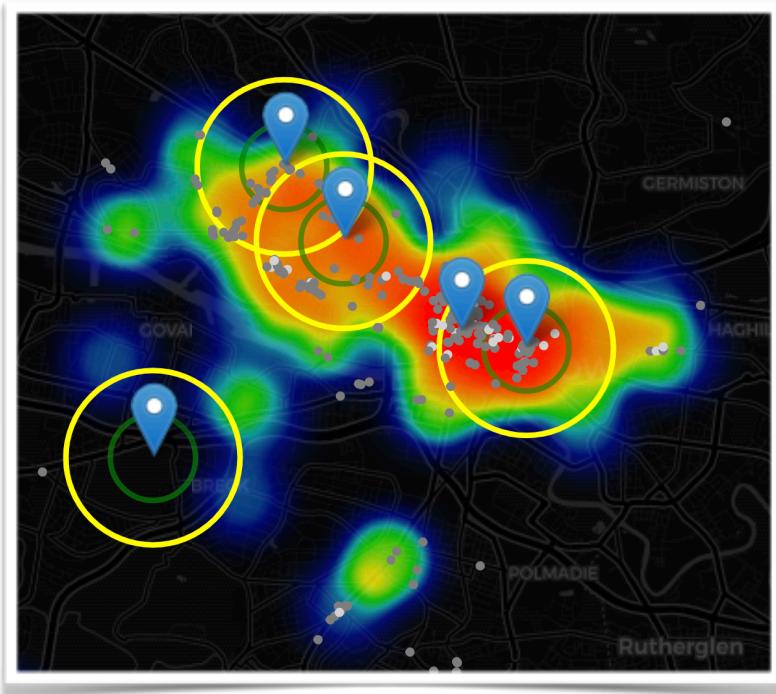


Figure 6: Heatmap of restaurants with class 1(green) and class 2 (yellow)

With these heatmaps, conversations with business stakeholders can be simplified and - using the visual guidance - individual perceptions about the 'benefit of a specific location candidate' immediately verified.

Business rules

To recap, goal of this analysis is to evaluate the four given location candidates for suitability to open a BYO Ltd. venue.

Several business rules and criteria were defined, in cooperation with key stakeholders and subject matter experts. The ideal candidate location will match all of the criteria defined in the business rules.

Business rules & criteria - Class 1 (within 500m radius from location candidate)

- Number of competitors : [the smaller the better]
- Number of suppliers : > 3
- Number of restaurants : > 2 && < 10
- Number of parks & recreation: > 2

Business rules & criteria - Class 2 (within 1000m radius from location candidate)

- Number of competitors : [the smaller the better]
- Number of suppliers : [the higher the better]
- Number of restaurants : [the higher the better]
- Number of parks & recreation: [the higher the better]

Per as ‘agreement’ with the stakeholders, the business rules work also as exclusion criteria for the final ranking. This means, if a location candidate does *not* match the business rule, its name will not be included in the ranking for the specific rule.

Application of business rules

The business rules were applied to each location candidate, resulting in a final ranking of the location candidates.

| Class | Business Rule | #1 | #2 | #3 | #4 |
|--------|---------------|---------------|-------------|---------------|-----------|
| class1 | br1 | Òran Mór | Park Circus | Merchant City | NaN |
| class1 | br2 | Merchant City | Òran Mór | Park Circus | Craighton |
| class1 | br3 | Park Circus | NaN | NaN | NaN |
| class1 | br4 | Òran Mór | Park Circus | Merchant City | NaN |
| class2 | br1 | Òran Mór | Park Circus | Merchant City | NaN |
| class2 | br2 | Merchant City | Òran Mór | Park Circus | Craighton |
| class2 | br3 | Park Circus | NaN | NaN | NaN |
| class2 | br4 | Òran Mór | Park Circus | Merchant City | NaN |

Figure 7: Final result table

Based on the given criteria and business rules, Òran Mór is the best location candidate. The location fulfills two (2) of the class 1 business rules and two(2) of the class 2 business rules.

Conclusion

The defined task was to identify the most suitable of the four location candidates, and provide supporting data, to empower stakeholders to make a data driven decision.

Pre-defined business rules (aka criteria the location needs to fulfill) needed to be applied and the results evaluated.

Òran Mór proved to be the best location, within the frame-set of this analysis. A final decision can be made by the board of BYO Ltd., in cooperation with local subject matter experts, once available properties around Òran Mór are identified and the financial aspect is evaluated.

Further directions / Opportunities for improvement

It would be interesting to include additional, socio-economic or transportation data, e.g. crime statistics, public transport or Uber travel times, into this analysis. This could provide further insights. Statistical correlations between location and e.g. Uber travel times, may provide an insightful indicator for the suitability of the location candidate.