

INSIGHTS FOR LIVELIER LIBRARIES

1. INTRODUCTION

1.1 The background

The public libraries in the UK are suffering from the budget cuts of local authorities. [According to a Guardian's article](#), 2017 saw the closure of almost 130 libraries across the country while 3000 extra volunteers have been brought to run the services. This reliance on volunteers has not been helping to implement efficient and sustainable operation nor to bring impactful initiatives together.

There are some pointing out the technology development or changing society as the cause of the decline. However, there are a few countries which have seen an increased use and popularity of their public libraries.

1.1 The problem definition

The fundamental problem does not lie in the budget cut but in the decline in use which started long before the cuts in funding. To regain funding, it is necessary to put together a strategy for rejuvenate the use and its wider impact on the community.

1.2 The objective

In an attempt to support local authorities in drafting library regeneration strategies, this report aims to rediscover the value or the roles of libraries from the user-centered point of view.

2 DATA

For the purpose of rediscovering the value or the roles of libraries, it is a good idea to conduct UX research. However, due to the budget limitation, this report will deal with the existing data: Foursquare's Places API provides some comments left by their users.

A total of 182 libraries were found in the first 6 biggest cities (London, Manchester, Birmingham, Leeds, Glasgow, Liverpool) in the UK and 64 comments were extracted from the above-mentioned API.

Setting the parameters as 5 km radius from each of the city centre, a set of libraries was extracted with the features of:

| | Library_ID | Library_name | Latitude | Longitude | City |
|---|--------------------------|-----------------------------|-----------|-----------|--------|
| 0 | 4ceea4c0f368b1f761c481f7 | City Law School Library | 51.519018 | -0.114430 | London |
| 1 | 4c57f4ae744a0f47690bcacb | Barbican Children's Library | 51.519502 | -0.093278 | London |
| 2 | 4bb1cfd7f964a52044a43ce3 | Crowndale Centre | 51.534266 | -0.137782 | London |

- Comments:

The data will be analysed to gain insights on what the users liked about or what they would like to share about of the libraries. Comments were extracted per library with the features of:

The Likes /Dislikes features show the count of such votes from other users, which will be used to enhance the magnitude of the comments' value.

| | Library_ID | Library_name | Latitude | Longitude | City |
|---|--------------------------|-----------------------------|-----------|-----------|--------|
| 0 | 4ceea4c0f368b1f761c481f7 | City Law School Library | 51.519018 | -0.114430 | London |
| 1 | 4c57f4ae744a0f47690bcacb | Barbican Children's Library | 51.519502 | -0.093278 | London |
| 2 | 4bb1cfd7f964a52044a43ce3 | Crowndale Centre | 51.534266 | -0.137782 | London |

It is noted that those comments left on the Foursquare website, which usage is limited in the UK market, do not necessarily reflect the entire user group. However, it is a good starting point of an exploratory stage, while making most of the existing data.

3 METHODOLOGY - EXPLORATORY DATA ANALYSIS

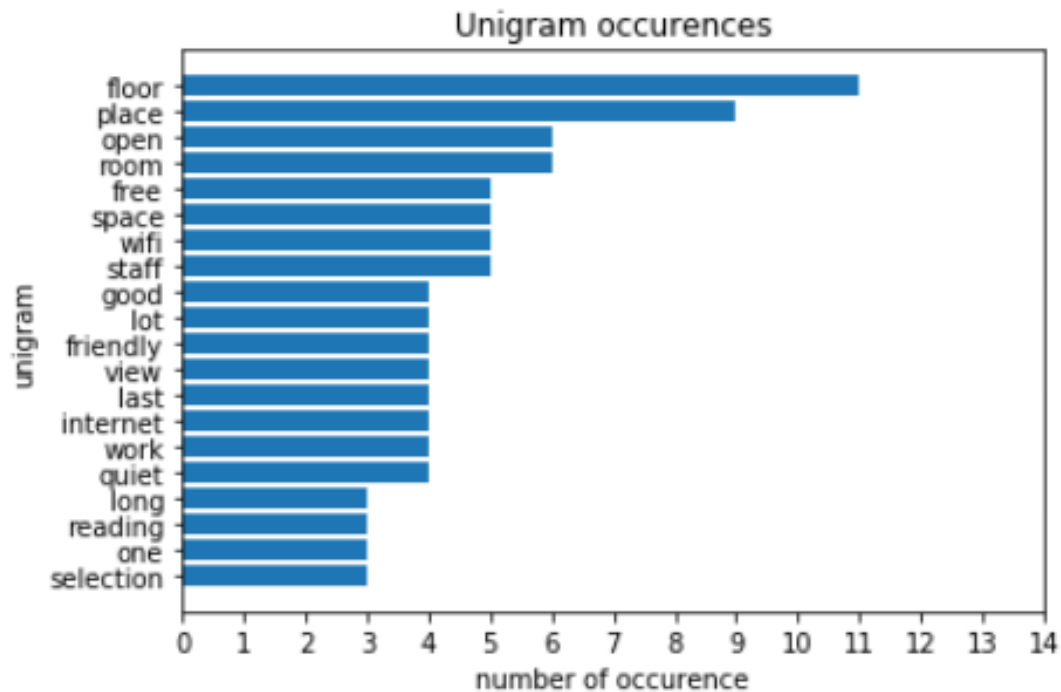
3.1. **WorldCloud:** I first made a visualization using WordCloud so that the common words used in the user comments become visible. It shows the words such as “floor” and “place” which do not bring any insight into what users wanted to tell about the library. Thus, a further data preparation / feature selection was necessary.



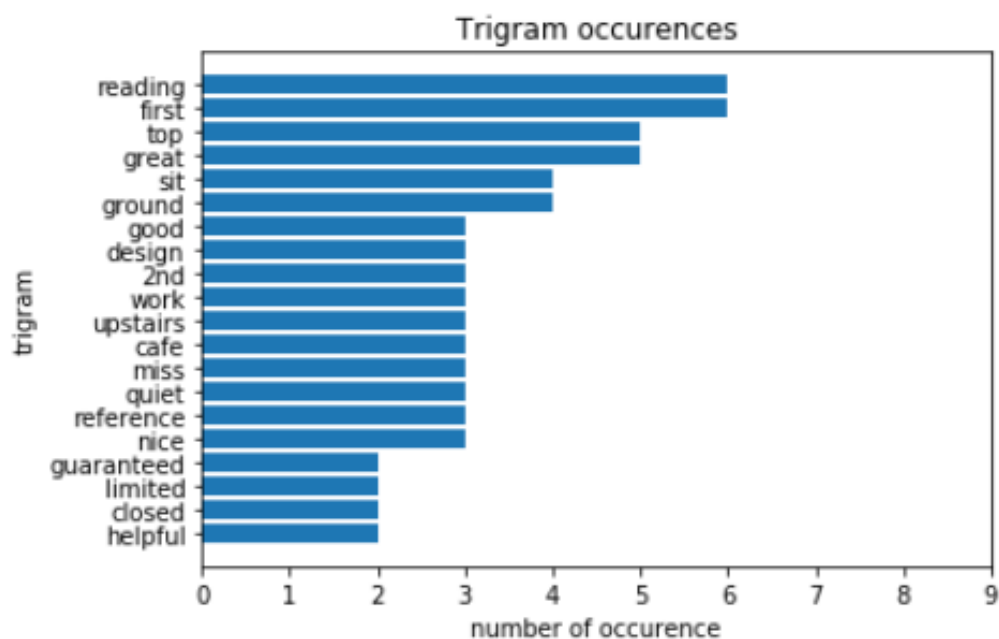
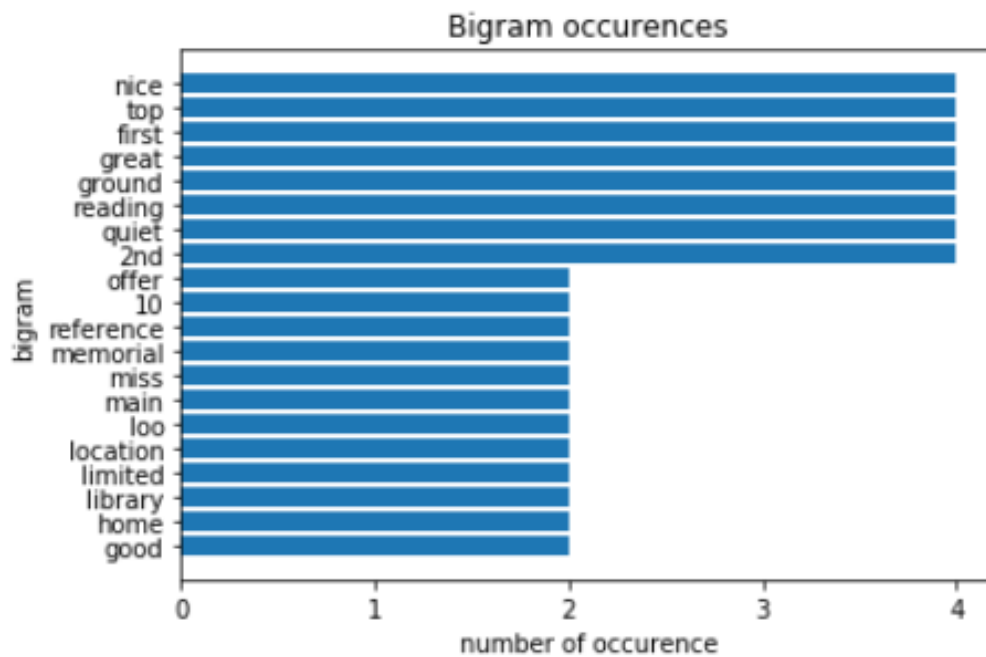
3.2. Natural Language Processing: Secondly, from NPL techniques, N-grams method was applied to pull out the combinations for the words., using libraries such as TextBlob and NLTK. Out of 64 comments, 660 bigrams, 597 trigrams and 676 unigrams are extracted.

| Bigram | | Trigram | | Unigram | Occurence | |
|--------|---------------------|---------|-------------------------------|---------|-----------|----|
| 0 | [well, utilized] | 0 | [well, utilized, space] | 0 | floor | 11 |
| 1 | [utilized, space] | 1 | [utilized, space, especially] | 1 | place | 9 |
| 2 | [space, especially] | 2 | [space, especially, liked] | 2 | open | 6 |
| 3 | [especially, liked] | 3 | [especially, liked, music] | 3 | room | 6 |
| 4 | [liked, music] | 4 | [liked, music, library] | 4 | free | 5 |
| 5 | [music, library] | 5 | [great, quiet, space] | 5 | space | 5 |
| 6 | [local, library] | 6 | [quiet, space, sit] | 6 | wifi | 5 |
| 7 | [great, quiet] | 7 | [space, sit, read] | 7 | staff | 5 |
| 8 | [quiet, space] | 8 | [sit, read, wellcomes] | 8 | lot | 4 |
| 9 | [space, sit] | 9 | [read, wellcomes, book] | 9 | friendly | 4 |

Unfortunately, neither the bigram nor the trigram has a combination occurring multiple times, so those table could not be used for statistical analysis. However, Unigram shows a significant tendency:



It became quite clear that the users of libraries mention a lot about 'floor', 'place', 'room', 'space'. As those words have a similar meaning, they are grouped together, converted to a one hot encoding, so that Bi/Trigrams containing either of those key words can be extracted. Instead, the related bigrams and trigrams were assembled according to the top unigrams:



4. RESULT

During the first Unigram analysis, the words such as “floor”, “space”, “room”, “place” stand out. When I looked the co-occurrence of those words, via Bigram and Trigram, the words “top”, “first”, “ground” occupy the top ranking. Those are specifying the floor the use would like to mention.

Apart from the adjective such as “nice”, “great”, “quiet”, the verb “reading”, “work” are interesting key words implying the activities of the users in the library.

During the analysis around those key words “floor”, “space”, “room”, “place”, however, the unigrams such as “free”, “wifi”, “staff”, “friendly” have been lost. However, those words also indicate some important facilities that users value about libraries.

5. DISCUSSION

In this report, no statistical analysis was used. Some might point out that I could have looked at:

- The number of libraries per person in each city
- Correlations with different socio-economic indices such as higher-education rate, free-meal student rate (in the UK context) etc.

However, when thinking about the defined problem of already declining number of libraries, it will not give any new explanation by showing that one city has got less libraries proportionally than another. It is also tentative to look in to the socio-economical indices, however, it will only show the correlation not the causation. It will not answer the questions such as “a city with more libraries become better-off” or “a city with a bigger number of higher tax-payers will have more libraries”, thus it cannot give a further insight into what to do with a declining number of libraries.

Instead, some qualitative data on what people value of libraries, what people like or dislike about will give insights into how to improve / enhance the existing libraries, when especially the budget is limited. And this will be able to generate an explanation on how valuable libraries are to the community, in what way, to be used in funding applications.

6. CONCLUSION

In this report, some qualitative data was explored in order to re-evaluate the libraries’ quality that users value. Although the resource was limited, it is interesting to observe that the word “book” was not appearing on the top ranking. Instead, the words related to “space” were most mentioned, together with the activity of “reading” and “working” as well as the facilities such as “wifi” and “staff”. This might be an indication that a library’s value is not only providing the book loan services but also to provide space for working and reading. This aspect could be then be assessed from the perspective of the gig economy and start-up fostering.

This project’s data collection was based on Foursquare, which didn’t provide much volume for the UK market. As a future direction, it is necessary to expand the data through Google map, twitter as well as some in-house data of the libraries. Also, more in-depth knowledge and exploration with the latest NLP is necessary.

At last as a personal note, I enjoyed my first Data Science project. I’ve never learned Data Science nor coding before. However, after leaning the basic of Data Science through this course, I enjoyed exploring further myself. With this project, I needed to use Natural Language Processing, which was not covered in the course. However, there are so much open source materials available out there. I wish to continue this new endeavor by keep learning.