

EXPLORING WeWORK LOCATIONS IN LONDON

MANDIE QUARTLY

Submitted 8th May 2019

TABLE OF CONTENTS

Introduction: Business Problem	1
Background	1
Business Problem	1
Interest	2
Data	2
Data Sources	2
Data cleansing	2
Methodology	4
Data exploration	4
Clustering	5
Results	6
Discussion	7
Conclusions	7

INTRODUCTION: BUSINESS PROBLEM

BACKGROUND

Given the recent years I've spent working with technology startups, I have had the pleasure to spend a large amount of time in WeWork offices around the world. I'm intrigued by the approach they've taken in creating spaces to allow many companies to work alongside one another, with the aim of creating local communities in the process. On that basis I've decided to focus on WeWork office locations in London in the UK for this project. The code used for this analysis can be found [here](#).

BUSINESS PROBLEM

[WeWork](#), currently preparing to be listed on the stock market, may have the potential to be [valued around \\$48bn \(£36bn\)](#). That kind of valuation puts them into the realm of high profile tech companies who are deemed to be disrupting their respective areas of business.

WeWork is an American company, founded in 2010 and headquartered in New York City. Essentially, they provide shared workspaces and have been very popular with technology startups, but also wider. They currently have 658 (open and coming soon) locations in 115 cities [around the world](#).

The question we aim to answer here is the following:

On the assumption that the current 47 London WeWork offices are successful, what links the neighbourhoods in which they are situated and could inform future choices for office locations in London?

Or asked another way, is there a particular type of neighbourhood in which WeWork offices should be focused?

INTEREST

The target audience for this work is WeWork themselves. Using this data and methodology WeWork could start to explore potential future locations for their offices new locations, based on previous successes in certain neighbourhood types.

Note: a key assumption here is that the current WeWork offices are "successful". Success would need to be defined by WeWork i.e. profitable / high occupancy rates etc.

DATA

DATA SOURCES

Three main sources of data will be used to explore this question in more detail:

1. Location data from the main WeWork website (as of 5th May 2019):
 - a. [Worldwide locations](#): in order to visualise locations around the world
 - b. [London Office locations](#): for use to explore specific neighbourhoods within London
2. [Geopy geocoder service](#): to turn city names, addresses and postcodes into latitude and longitude
3. Foursquare Places API:
 - a. The ["explore" endpoint](#): giving venue recommendations within a radius around a specific location

DATA CLEANSING

Initially we chose to visualise the cities with WeWork offices around the world, in order to provide a wider context for this specific investigation. This was done using the [Beautiful Soup](#) library in order to scrape city names from the WeWork website page listed at 1a in the previous section. Some locations were included but marked as "Coming Soon", so they were removed from the final listing. The final

dataset had 99 cities with offices and 16 marked “coming soon” removed. Latitudes and longitudes for all locations were found using the Geopy service based on city names. A map with all WeWork locations worldwide can be seen in Figure 1.



Figure 1: WeWork locations worldwide

We then focused on a subset of locations within a specific city, namely London in the UK. This city was chosen since it holds 47 offices within the central areas of the city. The addresses of the London offices were scraped using the same method as previously. Latitudes and longitudes for all locations were found using the Geopy service, primarily using postcode for most locations but falling back onto street address for those where the postcode was not found by the Geopy service. A map with all WeWork locations in London can be seen in Figure 2.

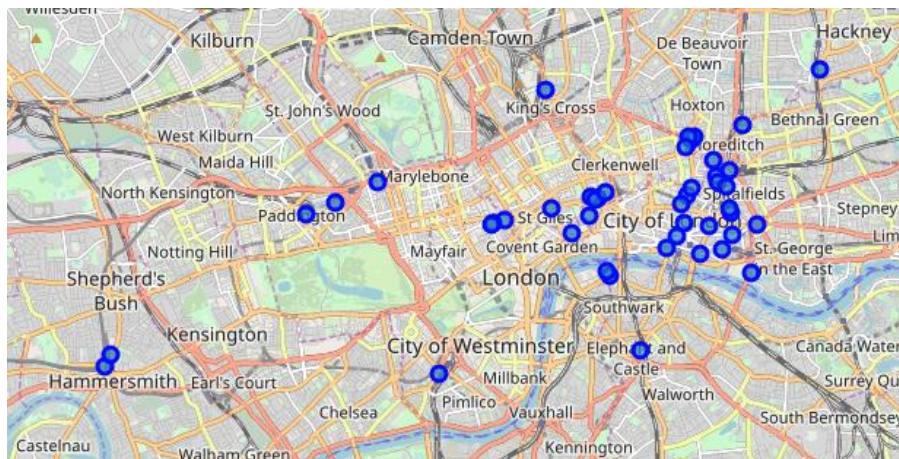


Figure 2: All WeWork locations in London, UK

Finally, a dataset was created using the Foursquare explore API to provide a listing of up to 100 popular venues in the immediate vicinity of each London office. A “[ped shed](#)” (short for pedestrian shed) radius of 400m was chosen, deemed as equivalent to 5 mins walk.

The Foursquare data gave around 3,500 venues in total for 47 offices. Of these venues 235 unique categories were found. The data was shaped using one hot encoding, grouped for each office and a mean value found for each venue category.

METHODOLOGY

DATA EXPLORATION

As described in the Data section, a final dataset was created showing the types and mean number of venues occurring around each office location.

Since the offices are relatively close to each other, it is interesting to note how much overlap there is using a 400m radius around each office. The map in Figure 3 below shows that, in some cases, this overlap is substantial. We decided to continue with this value for radius since a larger number of offices in a specific area indicates that type of area could be a good place to site offices.

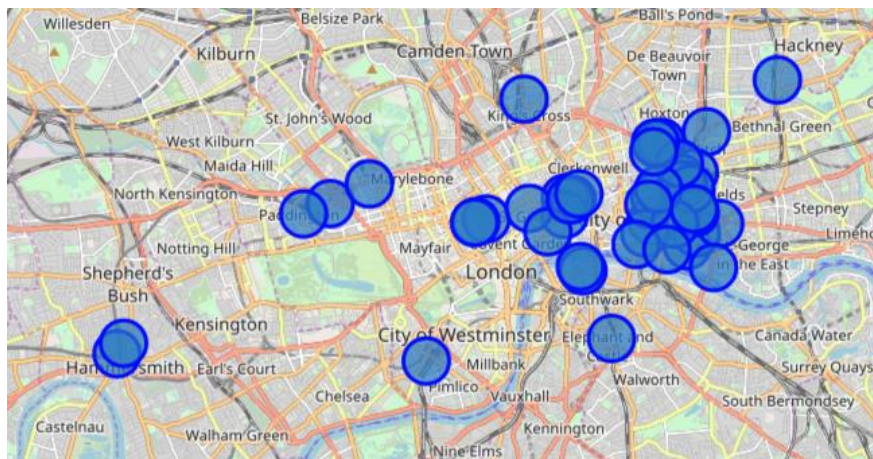


Figure 3: WeWork London offices with a 400m radius around each location

It is of interest to see how closely grouped the offices are, since this indicates areas which are not served currently by WeWork. As noted previously some offices are quite close to each other and it might make sense to avoid too high a density of offices in certain areas. This information, used in combination with the area types, could help in finding potential new areas for WeWork offices.



Figure 4: Heatmap of WeWork offices in London, UK

Figure 1 shows a heatmap of the office locations around London. Perhaps unsurprisingly, the highest density is seen around the City of London, around London Liverpool Street train station. Other hot spots are noted around the centre of the city, including in close proximity to areas more easily

accessed by commuters from outside the city such as London Paddington, London Bridge and Kings Cross train stations and the Hammersmith area (at the end of M4 motorway so more easily accessed by car). Also, to the west of the City around Tottenham Court Road / Chancery Lane.

CLUSTERING

The K-means clustering method was chosen to characterise the various areas around each office. This versatile method is often used for unlabelled data to find groupings.

The Elbow Method was used to explore the number of clusters to use. On the output of this, seen in Figure 5, we chose $k = 5$.

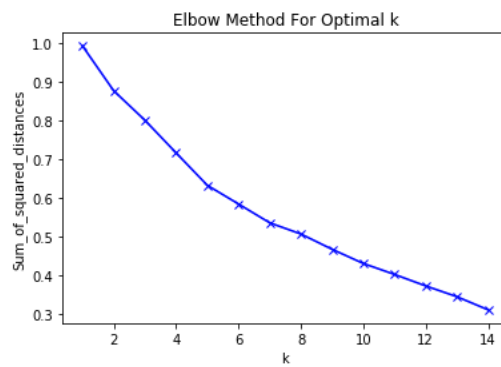


Figure 5: Elbow method results

The output of the final clustering can be seen in Figure 6. We note that we see three main groups, with two outlier groups with single members, as detailed in Table 1. The larger groups are the 4th, 1st and 2nd groups (in descending size order).



Figure 6: Results of clustering of London offices

Cluster Name	Number of members
Cluster 1 (label 0)	5
Cluster 2 (label 1)	10
Cluster 3 (label 2)	1
Cluster 4 (label 3)	30
Cluster 5 (label 4)	1

Table 1: number of members in clusters found

More of the specifics of the clustering can be seen in Figure 7, where the frequency of each venue type can be seen for the 5 clusters. The names of the venues aren't included (since there are so many), but in this figure they are plotted in alphabetical order on the x axis from left to right.

This plot allows a high level visualisation of the differences and similarities for the cluster groups. Some key differences can be noted from this, for example the largest peak seen is for Cluster 5 (label 4) for number of cafes within 400m. The second largest peak is for Cluster 1 (label 0) for Coffee Shops.

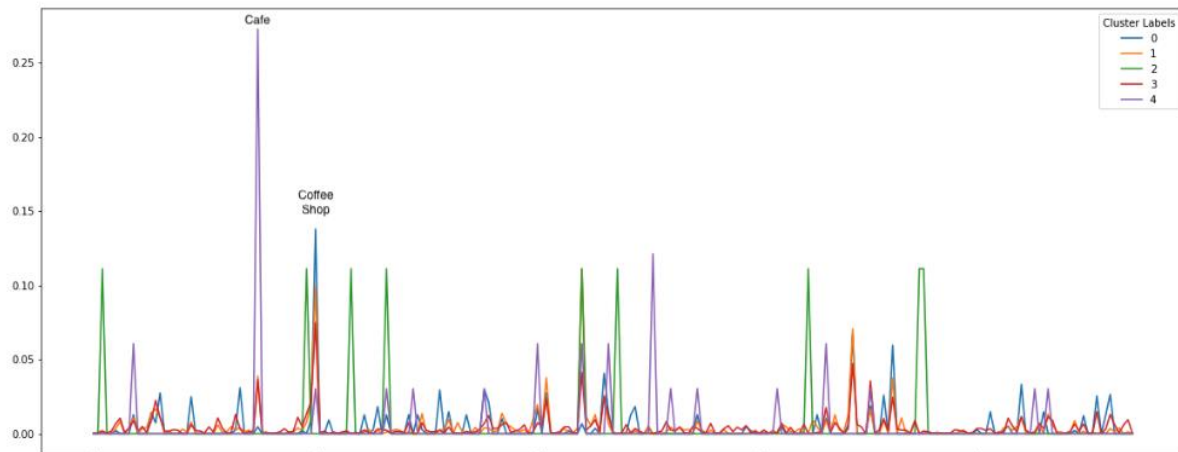


Figure 7: Specifics of clusters found

Finally, since the visualisation above has limited value (due to the large number of venue types), further exploration of the clusters was achieved by sorting to show the 10 most common venues for each. Based on this, the area types for clusters can be characterised by venues close to them.

RESULTS

In this analysis we have explored the locations of the WeWork offices in London. Using clustering based on venues within walking distance (defined as 400m radius) we have clustered the offices into five groups. Of these groups, three have more than one member.

Comparing the three largest clusters (clusters 4, 2 and 1 in descending order) we note the following:

- The three largest clusters share a number of common features. These include coffee shops and pubs in the top 5 of all three.
- Being specific about the type of each area beyond these shared features is difficult but we note a few differences using the top ten venue types for each cluster:
 - Cluster 4 (label 3) has the "core" venue types one might expect but has less variety in other areas i.e. types of restaurants and grocery stores.
 - Cluster 2 (label 1) has a relatively large number of coffee shops compared to the other clusters. But a wide range of food offerings and grocery stores.
 - Cluster 1 (label 0) shows fewer hotels and grocery stores than 4 and 2. But quite a variety of food providers.

- Clusters 3 and 5 only have one member each and the reason for this is evident in their characteristics. Cluster 3 shows construction and shopping areas. Cluster 5 shows a large range of Asian themed eating venues, plus a light rail station.

DISCUSSION

Based on our results we have identified three types of areas where WeWork offices appear to have been successful. As noted at the beginning of this project, a key assumption here is that the current WeWork offices are "successful". This assumption would need to be validated with data from WeWork highlighting offices which, by their own metric, are defined as successful. Potential metrics to use here could be profitability or high occupancy rates. Or potentially even other metrics measuring the outcomes desired from such working spaces, potentially harder to quantify, such as startup success rates, innovations or partnering successes.

In theory, using these results, WeWork could characterise other areas where they might be considering new offices to see if they are similar to those seen here, influencing their decision to create a new office there or not.

One thing noted during the exploration of the data and in particular a heat map of the offices was the higher incidence of offices around key commuter links into and out of the capital. In further work it would make sense to quantify this and add into the clustering methodology.

CONCLUSIONS

The aim of this project was to explore the following question:

On the assumption that the current 47 London WeWork offices are successful, what links the neighbourhoods in which they are situated and could inform future choices for office locations in London?

Or asked another way, is there a particular type of neighbourhood in which WeWork offices should be focused?

We have been able to show that the current offices fall into three main types of neighbourhood based on the types of venues in those neighbourhoods. This analysis might allow the key stakeholder here, WeWork themselves, to rate potential locations for future offices.

Key things to explore further would be how to define / subset the more successful offices to better inform the clustering process. Also, to explore the distances from commuter train stations / key roads which will most likely have impact here also.