

London's Healthiest Boroughs 2021

1. Introduction

1.1 Background:

London is a sprawling city with a population of roughly 8.6 million people. It is by far the largest city in the United Kingdom. Being a metropolitan city, it is home to a vast number of food venues that serve a variety of cuisines to suite London's incredibly diverse population. The availability and to some degree ease at which Londoners are able to order or dine out might explain why between 2017 and 2018 roughly 56% of adults were classified as overweight or obese.

Over the past several years obesity has been on the rise within the UK. Organisations, such as, Cancer Research UK declared obesity to be the second biggest cause of preventable cancer in the UK in early 2021. Obesity related fatalities has further been exacerbated by the appearance of the Covid-19 pandemic.

One study conducted by The Royal Society of Public Health estimated that for the severely obese, at a population infection rate of 10%, there is a direct impact of 240 and 479 excessive deaths in England and an indirect effect of 383 -767 deaths.

Furthermore, they found that, owing to BMI changes during Lockdown, 97,755 (5.4%: normal weight to overweight, 5.0%: overweight to obese and 1.3%: obese to severely obese) to 434,104 individuals (15%: normal weight to overweight, 15%: overweight to obese and 6%: obese to severely obese) would be at higher risk for COVID-19 over one year.

It is important to note here, that while this study found that the majority of BMI changes during the lockdown period increased. Many have utilised the extra time afford by the lockdown measures to workout at home and have seen an increase in their physical activity.

Nevertheless, it is clear that the prevention of obesity and promotion of physical activity are as important as physical isolation during the pandemic, particularly for those who have experienced a severe decline in physical activity. Therefore, it is advantageous for individuals to consider increasing their physical activity levels.

1.2 Problem:

The question then becomes what is the ideal way to become more physical active, or how can one stay in shape as lockdown restrictions are lifted. For many, this is to join a gym or fitness centre of some sort. Moreover, the potential of contracting Covid and the potential of future lockdown restrictions are a worry for many. Data, however, might prove effective in alleviating some of these worries.

By utilising the average health index of each borough, the percentage of fully vaccinated residents per borough and the foursquare API. We can aim to predict which areas are likely to improve on over health of their borough and which borough may be more resistant to a high reproduction rate of covid in the future. Thus, allowing for gyms to remain open and member to continue to visit them.

1.3 Interests:

There are several groups to whom this data will prove useful to. The, first and foremost, are those who are looking to move to a specific London borough. Whether or not they are looking at increasing their physical activities, knowing the overall healthy conditions of London's borough in comparison to the overall trend of London, will provided useful information on deciding where to live in London.

The second target audience, will be combination of those in the first group and those who are looking to increase their physical activity or maintain their current physical activities. Whether this is through joining a gym or by virtue of finding a borough where increasing physical activity is either supported or easier to achieve.

The third and most pertinent can be spilt into two. Firstly, those who are looking to switch or go to a gym wherein they can be somewhat assure that there is less chance of them contracting covid and that the area may be more resistant to the spread of covid and thus incredibly strict lockdown measures. And those who are looking to obtain

property or land with the intent of opening a gym or physical activity business, who are also looking for some assurances against the future spread of covid19. As we will see prospective business owners can be split even further into those who are looking to invest and see a return in the short term (less than 5 years) and those who are looking to invest and see a return in the long term (greater than 5 years).

2. Data Acquisition and Cleaning

2.1 Data Sources:

Most of the data on the health of a borough will be acquired from the National Office of Statistics' [Health Index Data](#). The health is an experimental statistic that is used to measure a broad definition of health, that can be tracked over time and compared between different areas. This will be used to offer an idea the varying levels of health between the boroughs. Please note, that measuring health in a form that allows for comparison is incredibly difficult and thus, this should only be taken as an estimate of the overall health levels in each borough.

The Health Index has been scaled to a base of 100 for England, with base year of 2015. Values higher than 100 indicate better health than England 2015, and values below 100 indicate worse health. The scale is such that a score of 110 represents a score one standard deviation higher than England 2015's score for that same indicator. In this way comparisons both over time and within a single year are simple to understand.

While this dataset is incredibly diverse and much of the data can be scoured for further insights. For the scope of this project, we will be looking at the three main domain scores

- Healthy People – focusing on health outcomes
- Healthy Lives – health-related behaviours and personal circumstances
- Healthy Places – wider determinants of health, environmental factors

And their overall scores between 2015 – 2018.

The data on vaccinations in London will be obtained from the [Nation Health Service statistics](#) on vaccination rates in the UK. Please note that these statistics are constantly changing as more people are vaccinated. For this project, the June 17th 2021 statistics will be used.

Covid Vaccination data in London is going to be used for two prominent reasons. Firstly, this data will become useful in the long run to businesses who are looking to open a gym. Gyms, saw a massive decline in profit and attendance due to the lockdown measures. As a preventative measure, having information about which areas in London have a higher covid vaccination rate may offer some peace of mind because it is possible for those area to have a lower covid reproduction rate in the future. Ensuring that it is less likely (but not certain) that these areas will see strict lockdown measures. Meaning gyms in these areas may be able to remain open for longer in comparison to other boroughs, where vaccination rates are lower.

The second reason, in a similar vein, address clients. With Covid-19 being a worry, many people have realised that going to gyms carries the risk of catching covid. So, with the knowledge of which areas of London have a lower vaccination rates may help gym-goers identify area which have a higher potential of covid – 19 reproduction rates. It is then possible to avoid either going to or join gyms in these areas and subscribing to gyms in boroughs where vaccinations rates are higher.

The dataset itself is quite comprehensive, it is broken into two major categories: 1st dose and 2nd dose. For this project, we will only be considering those who have had two doses as fully vaccinated. However, we still be considering how these boroughs will change in future as those who have the 1st dose will eventually get the second does.

2.2 Data Cleaning

Data is downloaded or scraped together from various sources online and combined into several different data frames. For the vast majority of the time both data sets downloaded contain several attributes that are irrelevant to this project. The scope of this project further applies limitations on the data gathered. If it were possible to extend the scope of this project. Then perhaps more of the data gathered could be used to gain insights into a borough's health and its trends. Perhaps an in-depth analysis simply on individual boroughs may be conducted in another project.

There were no major issues when building the data frame to visualise each borough. Some of the names have to be corrected from the web scraping and the areas code were gathered directly from the health index. The coordinates for each borough were gathered from a CSV file obtained from the Greater London Authority. Once the latitude and longitude attributes were imported. These were all compiled into a data frame and passed through the Geocoder API to generate an interactive map.

The data for the health index, however, was an entirely separate issue. There were several issues with this dataset. To understand how the data was cleaned and utilised it is important to understand how the health index scores were developed. Firstly, the overall health index score for each borough is developed from the average score of its three subdomain scores for that year. If a borough has an overall health score of 98.3 for 2015. The 98.3 is the mean between all of its subdomains. Similarly, so, each one of the boroughs three subdomain scores is developed from the mean of its attribute scores. Which are as follows:

- **Sub-domain Healthy People**, as the name suggests, focuses on health outcomes. The attributes used to develop this index score are: Mortality, Physical Health Conditions, Difficulties in daily life, Personal wellbeing, and Mental Health. The factors used to develop this score are mainly ones that people generally don't have much control over and are typically the inherent health indicators used to evaluate someone's health.
- **Sub-domain Healthy Lives**, follows from above, but places more of a factor on the social factors that influence the health of people. This Index score focuses on the healthy behaviours of people within each neighbourhood. This subdomain index was developed using the following attributes: Psychological risk factors, behavioural risk factors, unemployment, working conditions, risk factors for children, children and young people's education, and protective measures.

This arguably is the most important sub-domain for businesses and those who are looking to move into a borough. A lower score here informs people that there are social issues within the borough that can have negative effects upon your health.

- **Sub domain Healthy Places**, examines the quality of the environment. The attributes used to develop the healthy places index are: Access to green space, local environment, access to services, access to housing, and crime.

This factor is rather important to many members of our target audience. A borough with a high crime rate is a risk and a negative factor for each member of the target audience. Gym-goers would be less inclined to attend a gym in a specific borough over another if it meant endangering themselves.

Furthermore, factors like access to services and housing will influence how likely people are to attend a gym. A gym that is in a borough that has a higher access to daily services, like shops, banks, restaurants and work environments is advantageous.

Separately, for members of the target audience areas that are more concerned with outdoor physical activity, the assessment of access to green spaces may be more important.

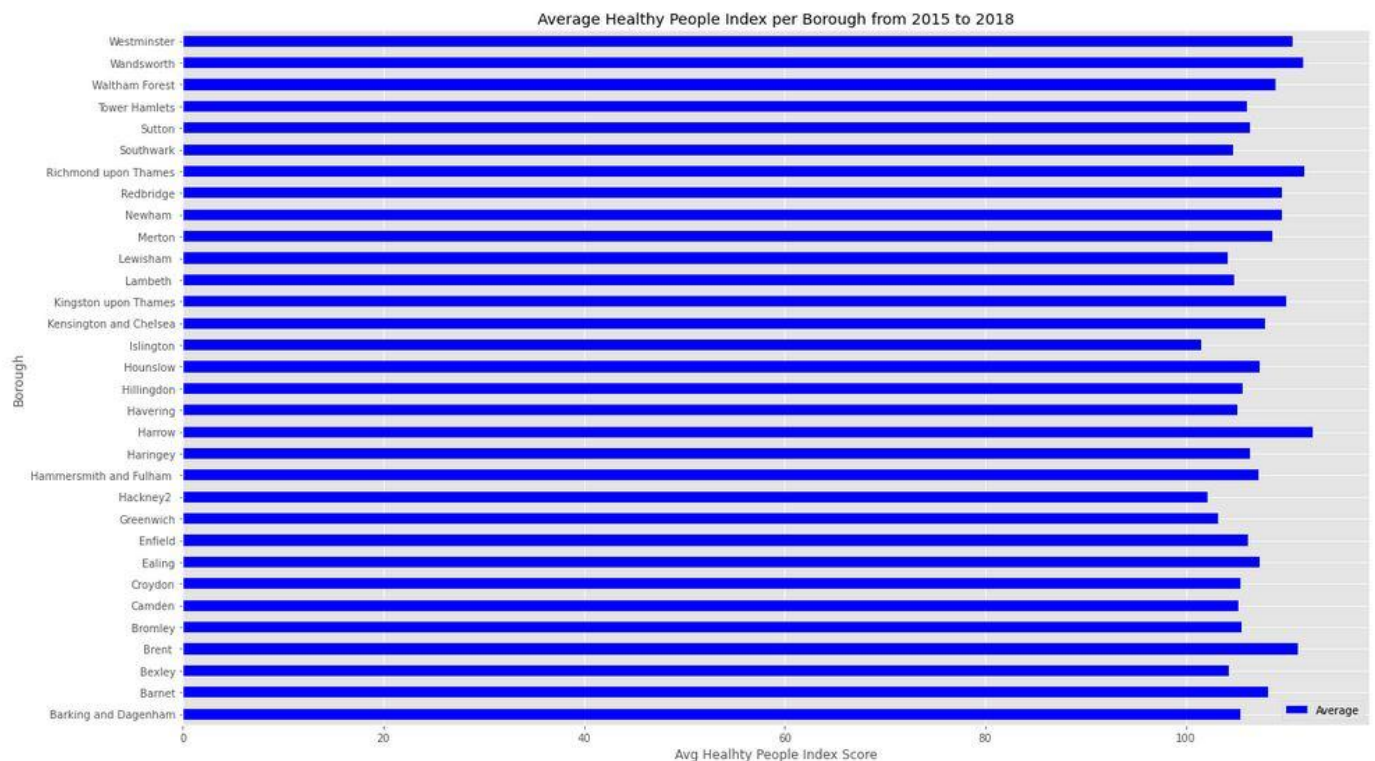
Due to the enormity of the data for each sub-domain and the fact that it was developed for each year. A large portion of the data clean took part in the excel. As, otherwise, it would have resulted in an incredibly long Jupyter Notebook before even any data exploration could take place. Beyond this there was not much data cleaning to be done. From here on out the data was simply portioned in to different data frames for exploration.

3. Data Exploration

3.1 Health Trends:

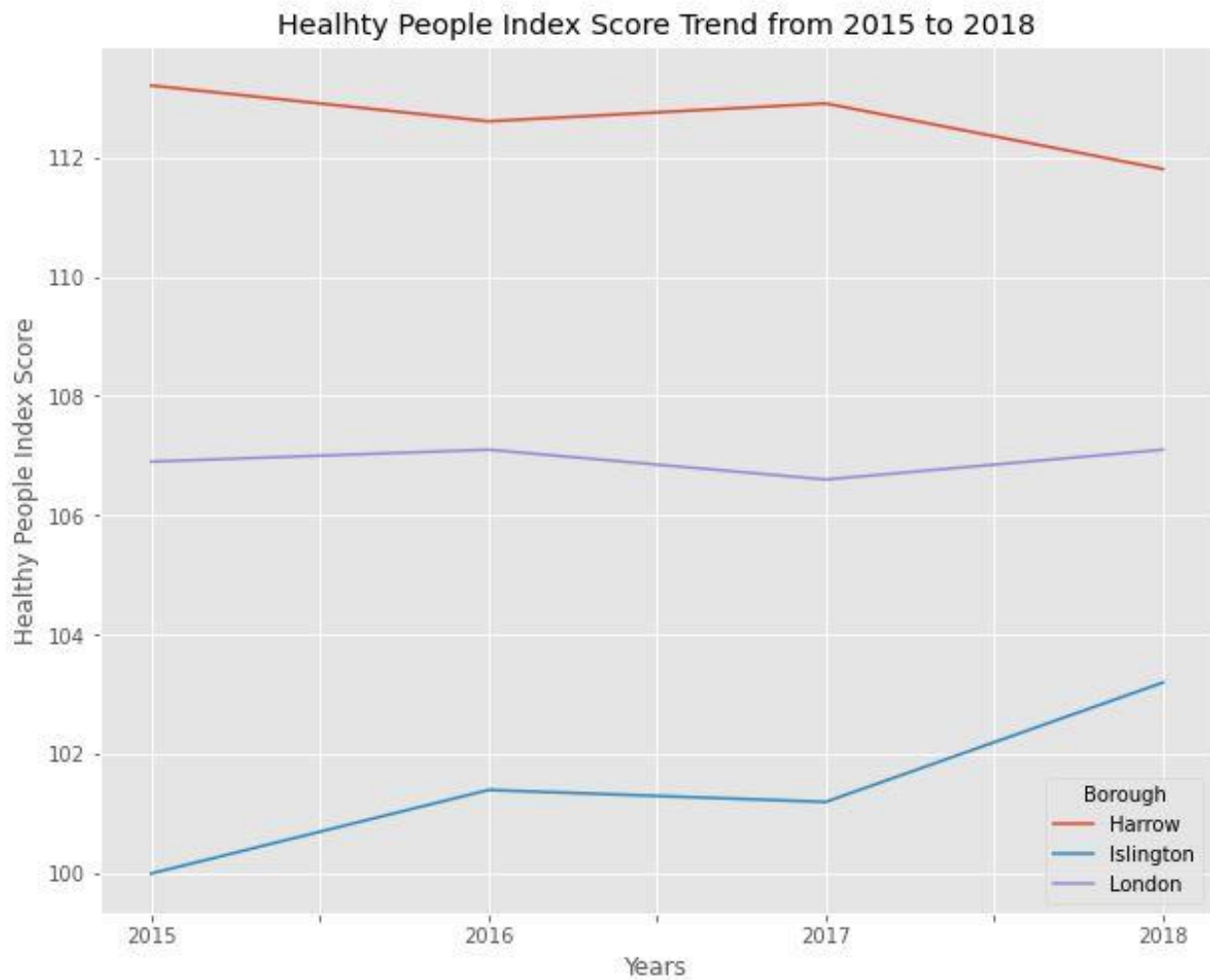
At first, the data exploration began by compiling a data frame of each borough respective scores for each year and then proceeding to plot them. Unfortunately, as there are 32 different boroughs this proved to be unnecessary as the output is an incredibly incomprehensible graph with lines for each borough. Similarly, a bar graph of this effect is also very unclear and difficult to comprehend.

So, to understand how the data could be useful in evaluating each borough. I choose to examine the averages of each borough's sub-domain from 2015 -2018. This proved insightful, initially, as it illustrated contrasting positions in comparison to the overall health index averages. It allowed an insight to see where each borough was either performing well in a subdomain and within which sub-domain it was underperforming. It also gave a good understanding where each borough was situated in comparison to others.



This bar graph although a little cluttered, visualises the average improvements that each borough has seen over between 2015 – 2018 for the Healthy People Index. We can see that Harrow has the highest average, with 112.625, and that Islington has the lowest, 101.75.

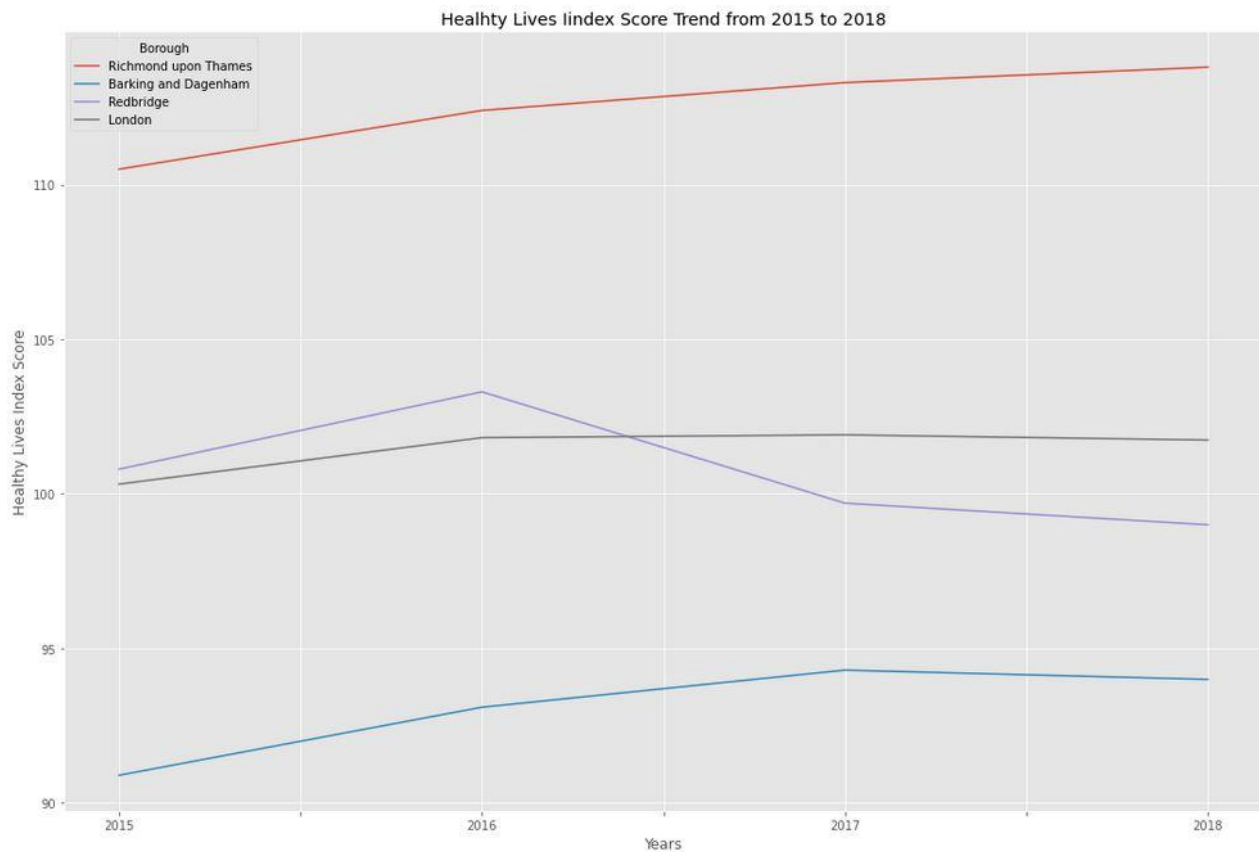
While this does allow one to easily grasp which borough is better than another, it doesn't necessarily reflect whether these boroughs are improving or deteriorating. It just tells the average state over the years. So, to better visualise the changes the borough has been through over the years would be to taking the average of the changes between each year, rather than just the average of the scores.



Here we can see that despite the fact that Harrow had a much higher average score than Islington in the HP index, Harrow has experienced a deterioration over the years and is on a downwards trend. Whereas, Islington, on the other hand, has seen incredible improvement from 2015 to 2018. This might be considered a more important factor than the simple average as it suggests that Islington is improving from its lower scores.

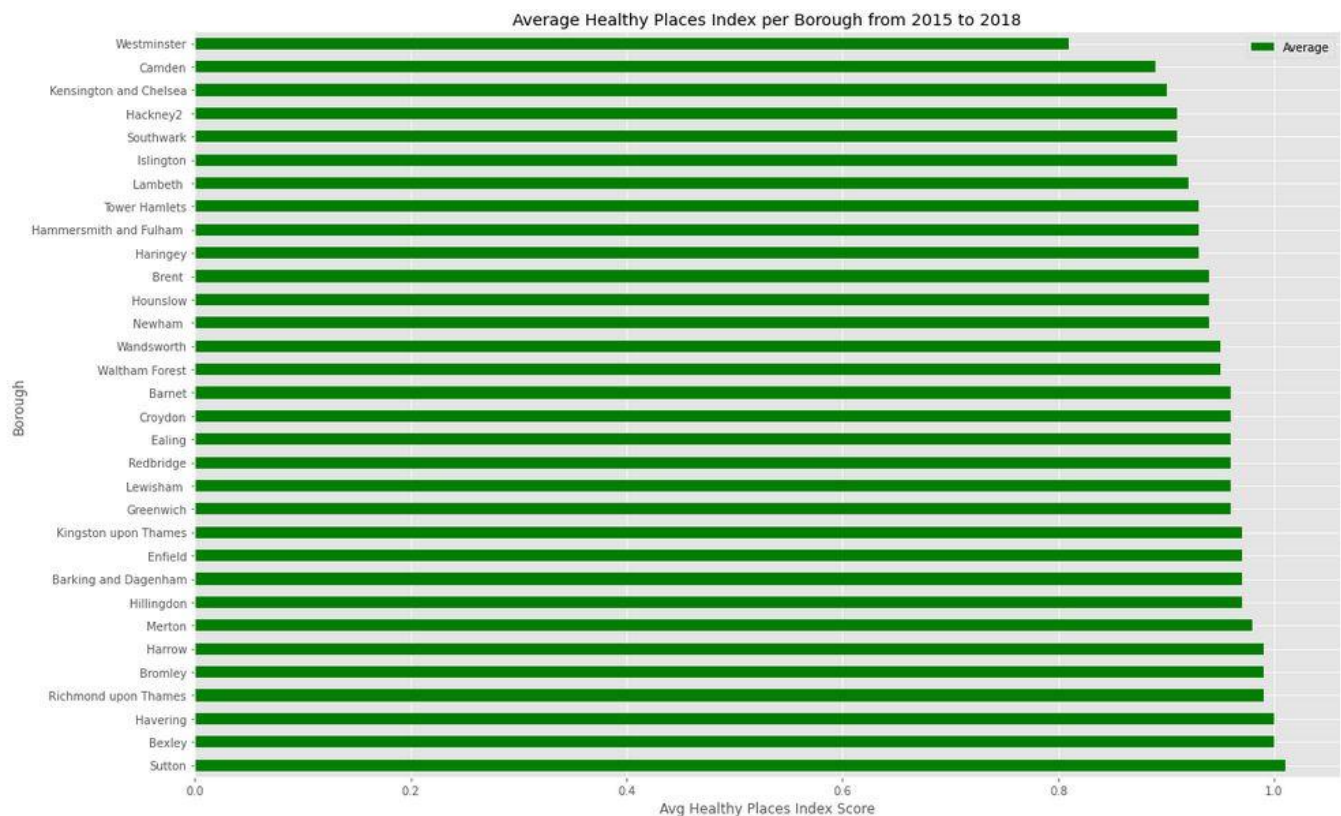
Moreover, by examining the trend of these two boroughs against the trend for the whole of London over the same period, we are able to see that Islington, like that of London, is on an upwards trend, albeit that London's overall increase is marginal in comparison to Islington. This informs us that Harrow is rather the outlier as its development is contradictory to London's overall trend. It seems then, to have a better picture of whether or not a borough is improving, we should take the average of the changes between each year rather than just the average scores. This is not to say that taking the average scores is useless, but rather that it only offers part of the picture.

As such we can move on to the other sub-domains. In comparison to the healthy people index winners and losers, the healthy lives index has different winners and losers. Kingston Upon Thames and Richmond upon Thames have the highest average score in comparison to the rest of the borough for healthy lives. While, boroughs like Barking and Dagenham, Newham and Brent are at the lower end. It is unclear as to why these boroughs rank where they do without conducting a deeper analysis of the data. However, as we have already seen, this score does not necessarily tell the entire story. As with the previous index, the borough on the lower end may have seen improvement from where they started.



Unfortunately, unlike the previous trend graph, we can see that there is a clear difference between the borough with the highest score and the borough with the lowest score. While they have both seen an improvement over the last 3 years. Barking has experienced some stagnation and a slight decline in improvement from 2017 to 2018. Whereas, Richmond Upon Thames has simply continued to improve in this aspect. What is interesting, however is to see how these boroughs have performed in comparison to the overall trend of London.

While, each borough is fairly far away from the overall trend of London. They both share a similar increasing gradient from 2015 to 2017 with the overall trend of London. In fact, Barking seems to follow the overall trend of London far closer than Richmond. This is become increasingly peculiar when we look at the third borough: Redbridge. Although, its score throughout the years is much closer to that of London's overall trend. There are much steeper kinks in its trend line. It is difficult to truly understand why this is the case without devilling deeper into the data, but it does imply that despite all being a part of London, each borough experiences completely different economic and social changes.



Unlike the previous subdomain indexes, the Healthy Places index is far easier to understand. From the graph above we can see a very clear divide between the boroughs that correlate directly to their average healthy places index scores. The lowest ranked boroughs are Westminster, Camden, and Kensington and Chelsea respectively. Whereas the top scorers are Havering, Bexley, and Sutton respectively.

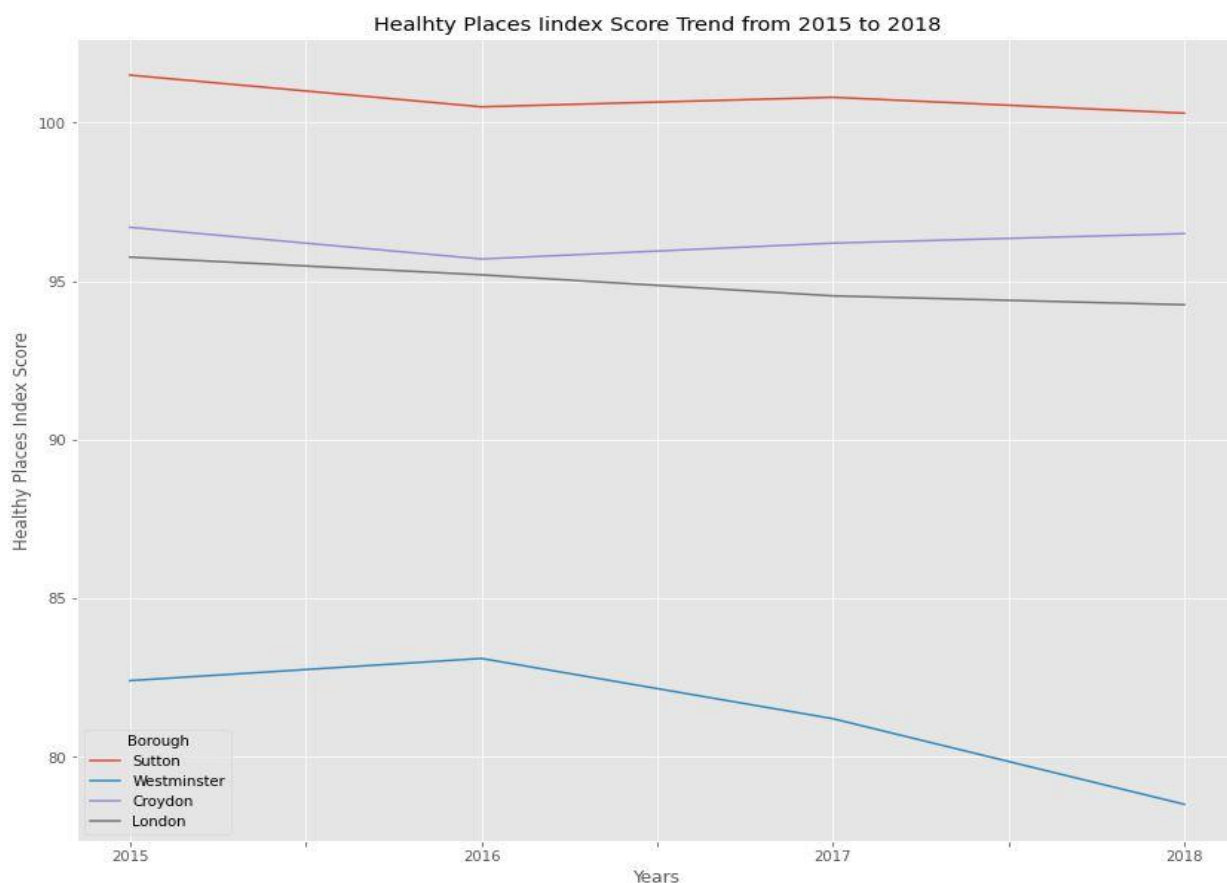
We can understand why these boroughs are ranked where they are by simply looking at the visualisation of the boroughs. The boroughs ranked at the lower end are all situated deep within the centre of the city, where there is less likely to be lots of greener spaces. While the boroughs on the higher end are on the outskirts of London, where there are more likely to be large areas of undeveloped land. In fact, this can easily be seen in the map itself.

The graph above illustrates the trend of Sutton, the borough with the highest Healthy place index, Westminster, which is the borough with the lowest average healthy place index score and Croydon, which is in the middle. As well as, the overall trend of London as a whole.

As the lines depict has been a decline in the Healthy places index score. While Westminster has seen a very steep decline from 2016, Sutton has also seen a decline since 2016, although of a much lower gradient. What's more is that both of these boroughs are in line with the overall trend of London, which has also declined over the years. As this line is the average scores of every borough together, the gradient or the trending decline is more telling. Unlike with the earlier indexes there are no kinks in the line. It is a steady decline.

An in-depth answer to why there is a steady decline would require a much further analysis of the data. One which is outside the scope of this project. However, one explanation or presumption is that this score will always trend downhill as population increase. The attributes that compose the Healthy Places index would be heavily influence by an increase population. This is one possible explanation of this trend.

Oddly enough, however, we can see that Croydon is defying the overall trend and has seen a steady improvement from 2016 onwards. This could be due to an increase in green areas, more housing being developed or a decline in crime. It is unclear what has caused this change.



3.2 London's Healthiest Boroughs

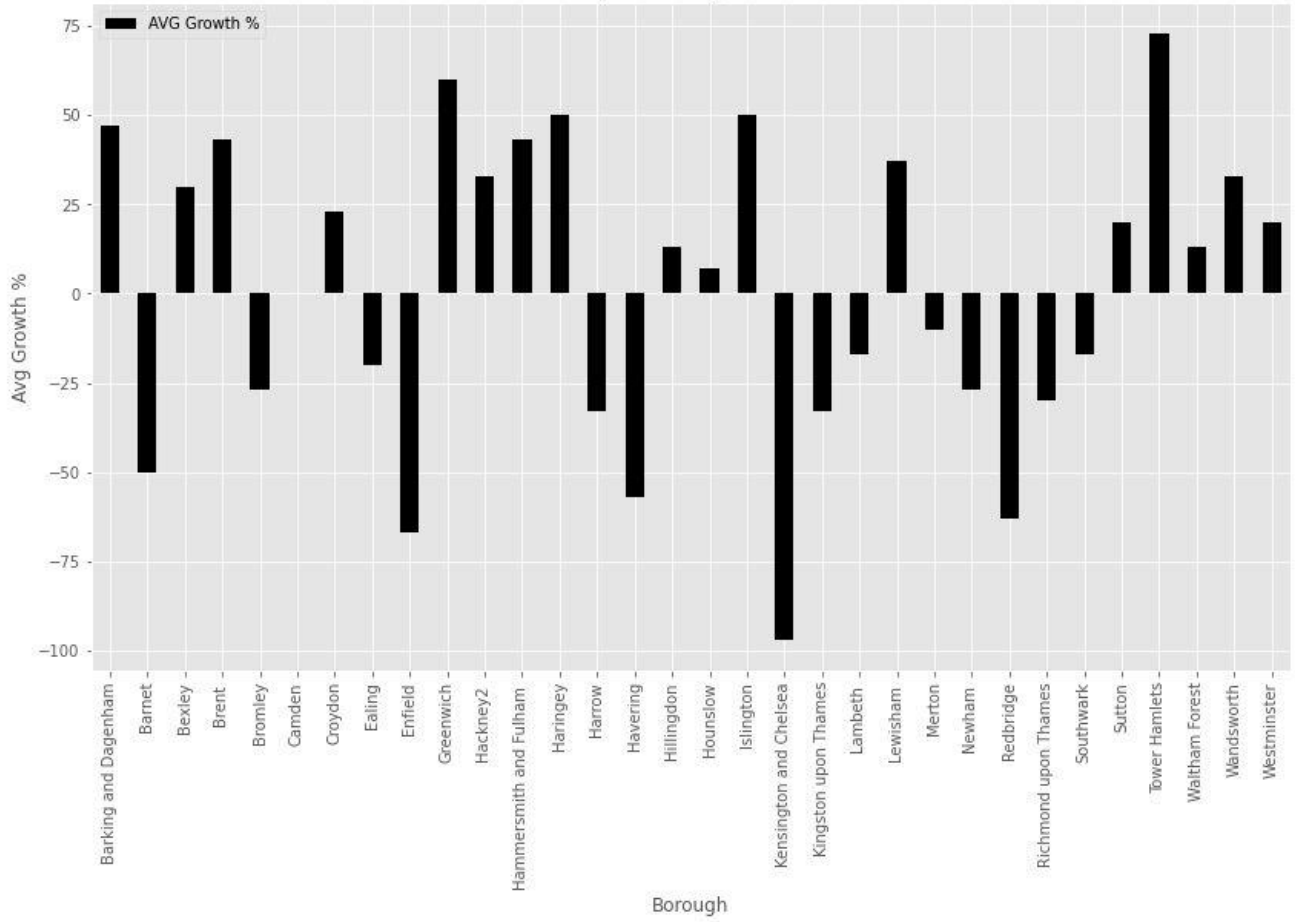
With the individual subdomains briefly explored, some understanding hopefully has been conveyed as to how the overall score has been developed. The overall score of each borough is the average of all of its subdomains each year. Thus, a boroughs Health index for 2015 is the average of its Health people, Healthy Lives and Healthy places scores. We can now move on to examining and exploring the overall health index in a similar manner as before.

Crucially, though, the following analysis will go a step further and look at the average change that each borough has seen over the years. This is important to members in the target audience that are looking at capitalising on boroughs that are developing and will have better scores in the future.

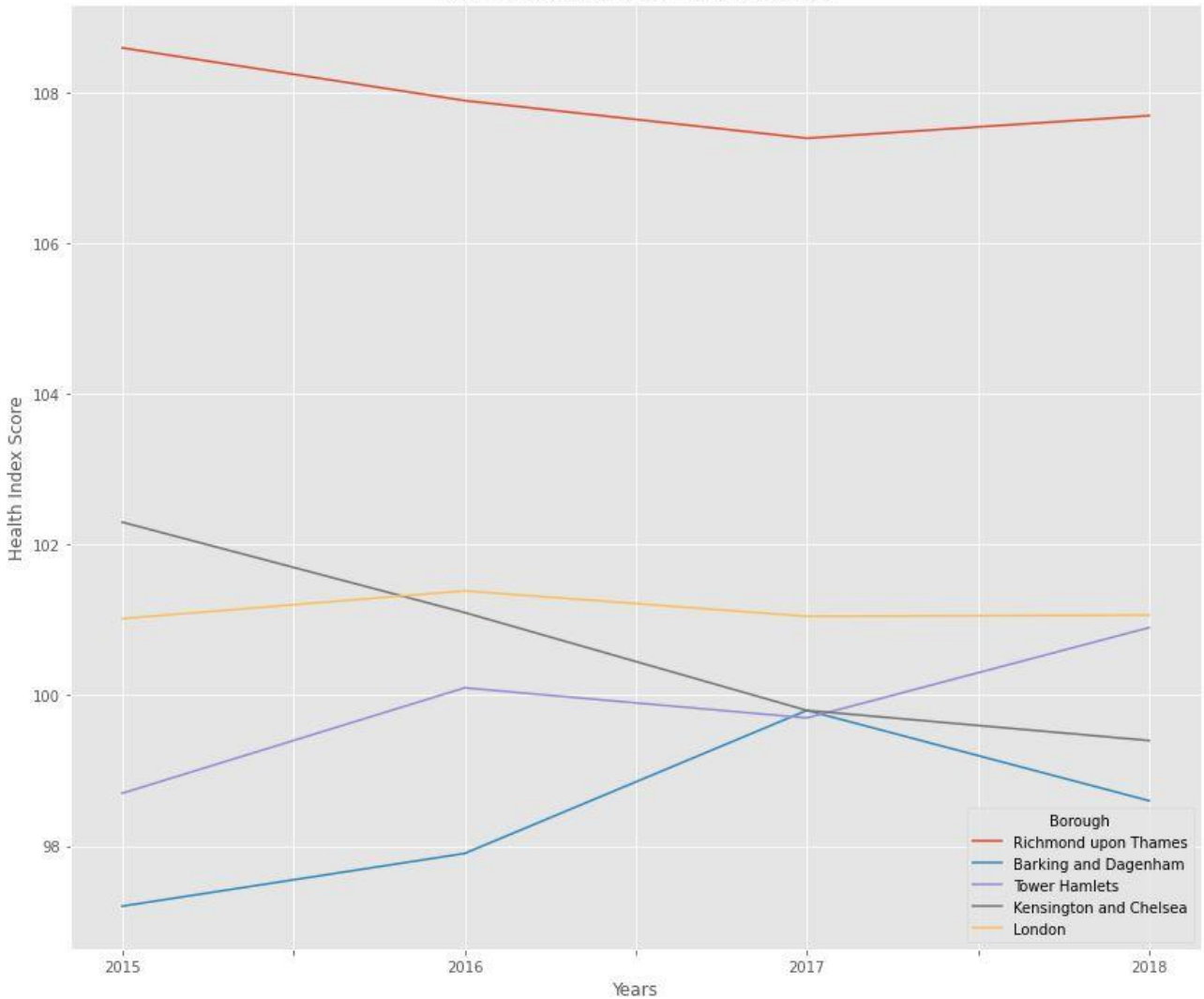
The table below is the data frame used to evaluate the overall health index for each borough. We can not only see how each borough ranks by their average but also by their average growth.

	2015	2016	2017	2018	Average	Growth from 2015 to 2018	Average Growth	AVG Growth %
Borough								
Barking and Dagenham	97.2	97.9	99.8	98.6	0.98	1.4	0.47	47.0
Barnet	103.6	103.0	103.0	102.1	1.03	-1.5	-0.50	-50.0
Bexley	101.6	102.6	101.7	102.5	1.02	0.9	0.30	30.0
Brent	100.0	101.0	100.5	101.3	1.01	1.3	0.43	43.0
Bromley	104.0	103.7	103.1	103.2	1.03	-0.8	-0.27	-27.0
Camden	99.7	101.0	98.4	99.7	1.00	0.0	0.00	0.0
Croydon	100.0	100.6	99.5	100.7	1.00	0.7	0.23	23.0
Ealing	101.5	100.0	101.1	100.9	1.01	-0.6	-0.20	-20.0
Enfield	100.7	100.0	100.1	98.7	1.00	-2.0	-0.67	-67.0
Greenwich	97.8	98.2	99.3	99.6	0.99	1.8	0.60	60.0
Hackney2	97.3	97.3	98.0	98.3	0.98	1.0	0.33	33.0
Hammersmith and Fulham	100.3	100.8	100.2	101.6	1.01	1.3	0.43	43.0
Haringey	98.2	99.7	99.2	99.7	0.99	1.5	0.50	50.0
Harrow	105.0	105.1	104.5	104.0	1.05	-1.0	-0.33	-33.0
Havering	102.9	102.5	101.9	101.2	1.02	-1.7	-0.57	-57.0
Hillingdon	100.5	101.8	101.0	100.9	1.01	0.4	0.13	13.0
Hounslow	100.2	101.1	99.1	100.4	1.00	0.2	0.07	7.0
Islington	97.7	98.7	98.1	99.2	0.98	1.5	0.50	50.0
Kensington and Chelsea	102.3	101.1	99.8	99.4	1.01	-2.9	-0.97	-97.0
Kingston upon Thames	106.0	106.3	105.1	105.0	1.06	-1.0	-0.33	-33.0
Lambeth	98.8	99.5	100.2	98.3	0.99	-0.5	-0.17	-17.0
Lewisham	99.7	99.1	99.8	100.8	1.00	1.1	0.37	37.0
Merton	103.7	103.8	103.7	103.4	1.04	-0.3	-0.10	-10.0
Newham	100.0	100.0	99.0	99.2	1.00	-0.8	-0.27	-27.0
Redbridge	103.2	103.0	101.0	101.3	1.02	-1.9	-0.63	-63.0
Richmond upon Thames	108.6	107.9	107.4	107.7	1.08	-0.9	-0.30	-30.0
Southwark	99.8	100.1	100.4	99.3	1.00	-0.5	-0.17	-17.0
Sutton	101.8	103.3	103.6	102.4	1.03	0.6	0.20	20.0
Tower Hamlets	98.7	100.1	99.7	100.9	1.00	2.2	0.73	73.0
Waltham Forest	100.2	101.4	101.4	100.6	1.01	0.4	0.13	13.0
Wandsworth	103.3	104.3	104.6	104.3	1.04	1.0	0.33	33.0
Westminster	98.3	99.5	99.4	98.9	0.99	0.6	0.20	20.0

AVG Growth % per Borough from 2015 to 2018



Health Index Trend from 2015 to 2018



We can see in the graph above that there is a varying degree of difference between boroughs. The graph reflects mainly the outliers of all of the boroughs. This is intentional, as it can be used to assess whether there is a borough that is improving and could be potential areas to open a gym. This is particularly the case with Tower Hamlets. We can see in comparison to the yellow line, London's overall trend, Tower Hamlets has seen a very drastic improvement in its Health Index from 2015 to 2018. This could potentially mean that the borough is seeing a lot of development and could continue to improve further down the line.

As a business owner or potential business owner this is can be incredibly attractive. As it could mean investing in opening a gym in a borough now and, hopefully, capitalising on future development and customers.

Unfortunately, the Health Index data doesn't include any data concerning how each borough has been affected by Covid19 and the restrictions that have come into place. Without this we cannot state that the trends shown above have or will be consistent. In fact, the trends and growths that some of these boroughs has seen in the previous years may have been either completely halted or eroded by the effects of Covid.

As such, safe guarding against Covid and Lockdown restrictions is an important factor for those looking to attend or run a gym. And to attempt to do so we need to look at the vaccine data per borough.

3.3 Vaccine Data

The intended use of the vaccine data is to analyse which boroughs have a higher level of vaccination than others. In doing so, it provides our target audience another set of information on each borough to help them perhaps choose where a borough which may see a slower rate of infection.

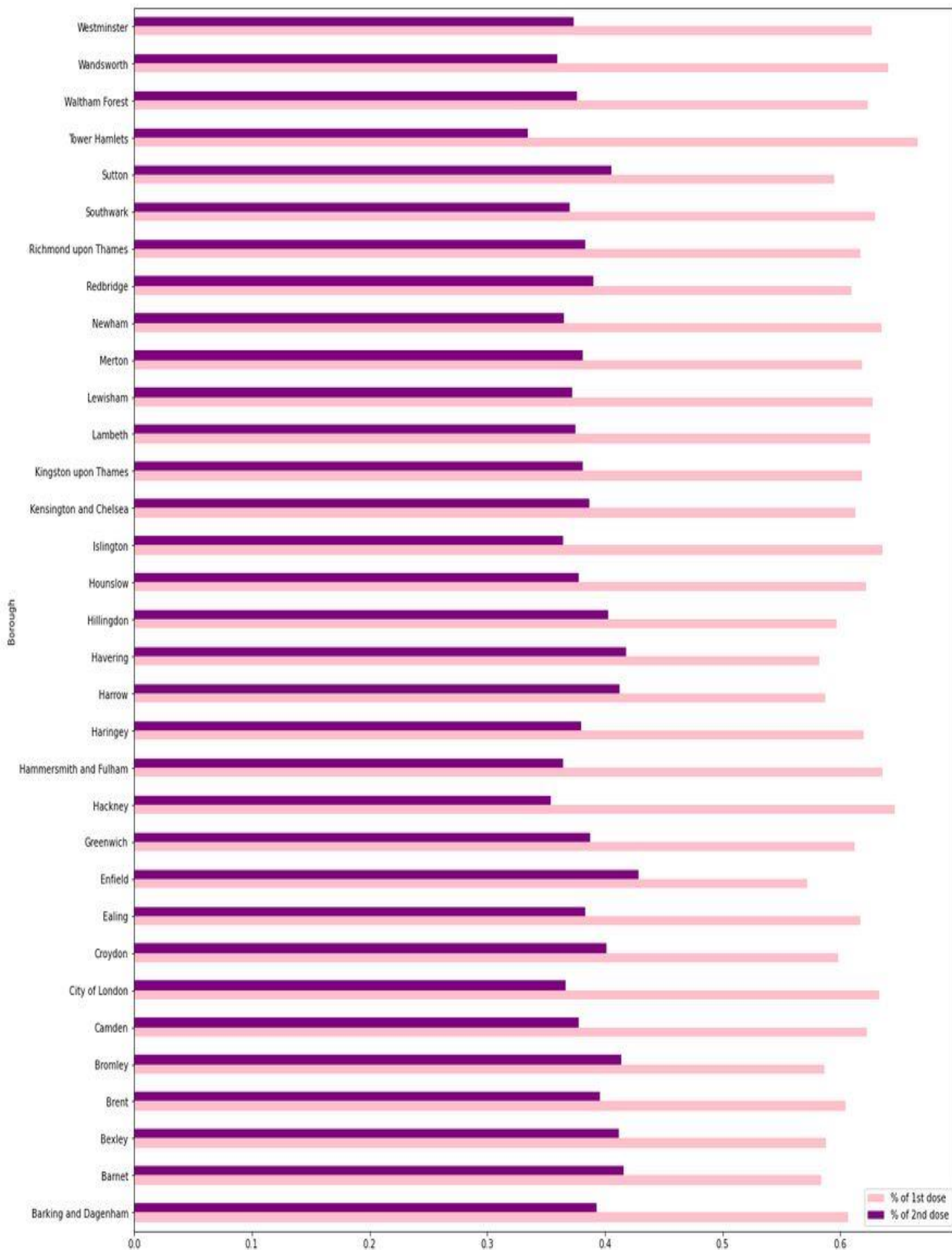
Furthermore, as gyms are considered a hotspot for Covid 19, areas that have a high R reproduction rate of the virus are more likely to see lockdowns which shut gyms down. While, inversely area with higher vaccine rates will have a lower R reproduction rate and are likely to see less restrictions. This could mean being able to keep gyms open in one borough, which may mean maintain monthly subscriptions from customers. While in another borough all gyms are closed

Enfield, at the time of this data set was produced has the highest number of fully vaccinated residents. This means that Enfield could potentially be one of the more robust boroughs in London to future lockdowns and covid case.

That being said, we cannot discount the number of residents with only one dose of the vaccine. From the graph, it is clear that Tower Hamlets current has the highest number of residences who have received one does of the vaccine. In the future then, once these residents receive the 2nd dose. Tower Hamlets will have the highest rate of vaccinated residence, provided they maintain an efficient vaccination rate.

Thus, the argument could be made that Tower Hamlets will become the most robust area in London to the spread of covid19 and future lockdowns. Furthermore, as we saw with the Overall Health Index trends, Tower Hamlets has had the highest average level of improvement from 2015 to 2018. Although, they might not be the highest rated borough in terms of overall health. They are a borough that it on the rise and for prospective gym owners or business owners Tower Hamlets tend could be very enticing. As it currently may offer a lower cost to opening a business in comparison to other boroughs ranked higher on the index, with the potential for a higher rate of returner in the future. i.e., Tower Hamlets might be a better borough to look into investing for the future.

On the other hand, this may not apply for customers who won't be looking for a long-term return on an investment (+5 years) but rather a gym that they could join now and can safely guard against lockdown measures and Covid19 cases. For customers areas like Havering, Harrow, and Barnet which ranks on the higher end of the health index and has one of the highest numbers of fully vaccinated residents.



4. Conclusions

4.1 Joining a gym:

With the top % most common venues of we can see that there are two boroughs that have Gyms as there 4th most common type of venue. These two being Southwark and Bromley. For Customers who are looking to join a gym these are incredibly attractive boroughs. Firstly, as these types of venues are among the 4th most common venues, there is a high chance that there is an increased amount of price competition between gyms. As such, they will have more competitive price ranges.

Additionally, as Bromley ranks as one of the higher boroughs on the health index, as well as, having a higher % of fully vaccinated residents. Bromley will have access to a housing and plenty of services. In addition to plenty of green spaces and a lower crime rate than other Boroughs.

While Southwark has a lower ranking on the health index it as well as % of fully vaccinated residents. So it many not have as much access to housing and services as Bromley. It does have one of the higher number of residents who have has at least one vaccination. So in the near future, there is a high probability that Southwark has a high number of fully vaccinated residents and is hopefully more robust to the spread of covid19.

Thus, for people looking to join a gym who either will live or live London, Southwark and Bromley seem like a good choice.

4.2 Opening a gym:

For those who are looking to start a business particularly, those who want interested in a longer-term return then areas such as Tower Hamlets might be a good investment now. As it seems like the local authority has been working to improve the overall standing of the borough between 2015 to 2018. If we presume this trend will continue, then perhaps purchasing property now might yield higher returns in the long terms.

However, a shorter-term solution may be Havering. It ranks rather high on both the health index which means it is an area with access to housing and plenty of services. In addition, to plenty of green spaces and a lower crime rate than other Boroughs. Havering, is also one of the boroughs with a higher number of fully vaccinated residents so it too may be more resistant to the spread of covid. Unfortunately, by looking at Havering's top 5 venues there are none that indicate that health nor fitness is a high priority in Havering. This could mean that either there is very small supply of gyms however there is still a high demand for gym. Thus, opening a gym could be profitable in Havering. Or, there isn't a high demand for gyms in Havering. To develop a much clearer answer we would have to do further market research on gyms withing each borough.