**Introduction**

In the 21st century mostly every single process we do is automated. Starting from ordering our favourite food or requesting a taxi and ending with some advanced processes such as executing millions of millions transactions by simply pressing one button. And it all thanks to a major progress of Information Technologies.

Nowadays if we need to know where we can find some lovely dinner places in foreign country, what we need to do is simply open maps and by analyzing our geolocation the advanced machine can predict our preferences and provide us with most rated nearby restaurants in no time. But not that it is user-friendly and really reliable, it is, also, can be extremely precious for commercial use. For example, if we would get a project to build a new venue (office building, department store, grocery store, restaurant and so on.) in a city we have never been to, by using some machine learning algorithms we can cluster our whole city’s dataset and then visualize it on map to predict what would be the most efficient spot for a new start up. That is exactly what we will be doing today, imagine, that we got a business project to build a new office building in City of London the most advanced Borough in London and we need to find the best neighbourhood to proceed with. Please do take your seat ant let me take you through this fascinating journey where will be exploring and clustering different neighbourhoods in London to find which one best suits our needs.

# Data

First thing first in order for us to start analyze neighbourhoods in City of London we need to find dataset which one would include boroughs and neighbourhoods’ names and coordinates, because we will need them later on for data visualization. We will be using dataset provided by doogal.co.uk which is almost ideal for our project. Then we will just need to clean our data a little bit by dropping all insufficient columns and rename district to a borough and ward to neighbourhood columns, so our whole data would look more appealing and user-friendly.

Afterwards we will be calling foursquare api to find top 20 venues in every neighbourhood so we can cluster them and decide which one will be the most suitable for our project.

# Methodology

In this section first of all we will be cleaning our dataset, to prepare it for visualization and clustering.

Secondly when data is cleaned, we will be visualizing it on the map to see how it is distributed.

Thirdly we will be categorizing our dataset by using k-means clustering algorithm.

Fourthly we will be analyzing what cluster and what neighbourhood within a cluster is the nearest one to the station.

And lastly we will be transferring all of our processed dataset on the map, so that way we would be able to predict, what will be the most sufficient place for our project to begin.

# Results

By analyzing and clustering dataset of City of London dataset, we were able to find out, that the best area for our project to build an office building will be within the first cluster and Ideally in Dowgate Neighbourhood.

In the discussions section we will sum up why it is the best option to begin our work with.

# Discussions

Let me explain you, why previously suggested solution is the best for our project. By examining our exploratory data analysis, we can see, that the first cluster is the most diverse cluster of all, having: coffee shops, restaurants, gyms, lounges, grocery stores and even a park, which is in fact really important, because study shows, that people working in a greener environment tend to be much productive than others, also in stressful situations they feel more relaxed and are able to make quicker decisions, which is extremely crucial in business related problems.

Furthermore, the first cluster is the closest to the station and has a wonderful public transport communication system, which is really important for people who does not drive a car.

Lastly this cluster has hotels within its sector included, which is highly appreciable when you are having international representatives of the company for a business convention.

# Conclusion

In the conclusion, to sum up my report, I would like to say, that by analyzing dataset of City of London, we were able to find out, that the best area to build an office building will be within first cluster and ideally in Dowgate neighbourhood. Because it has the most developed public transport communication system and there will be no hassle to get to work if you live outside of a Central London. Furthermore, it is the most progressive area of all, having everything you need (coffee shops, restaurants, gyms, lounges, grocery stores) and even more, to enjoy your everyday life at work.