

# **IBM Applied Data Science Capstone**

Finding the most suitable area in London for a  
new residential building

Gilberto Agostinho  
May 2020

# Introduction

Business problem:

- housing development company
- best area in London to construct a new building
- 3-bedroom apartments targeted to middle-class families
- ample amenities such as supermarkets, parks, restaurants, cafes, etc.
- they also prefer a more central area

# Data

The data used in this project comes from several sources:

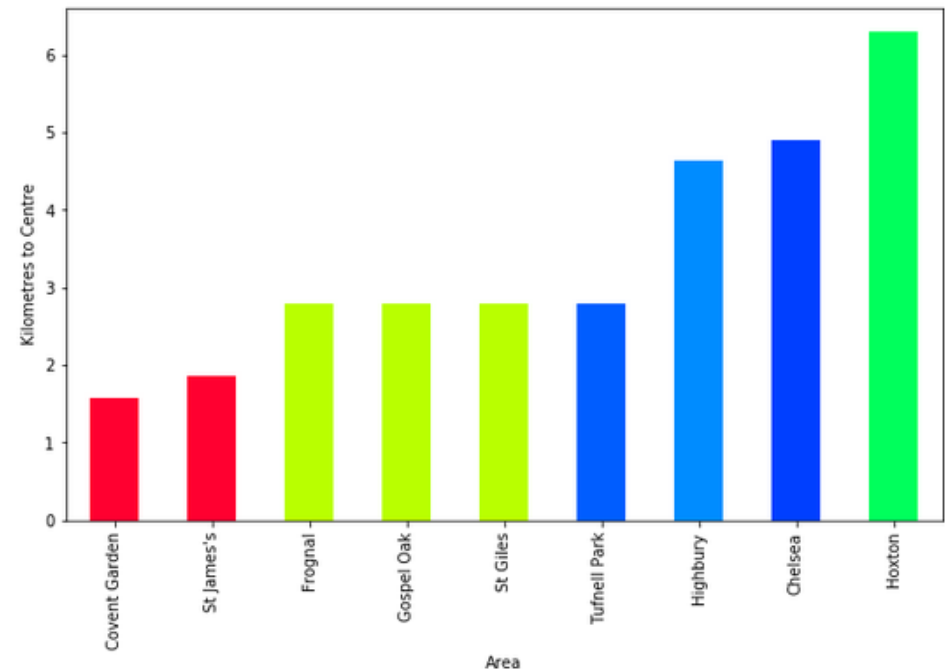
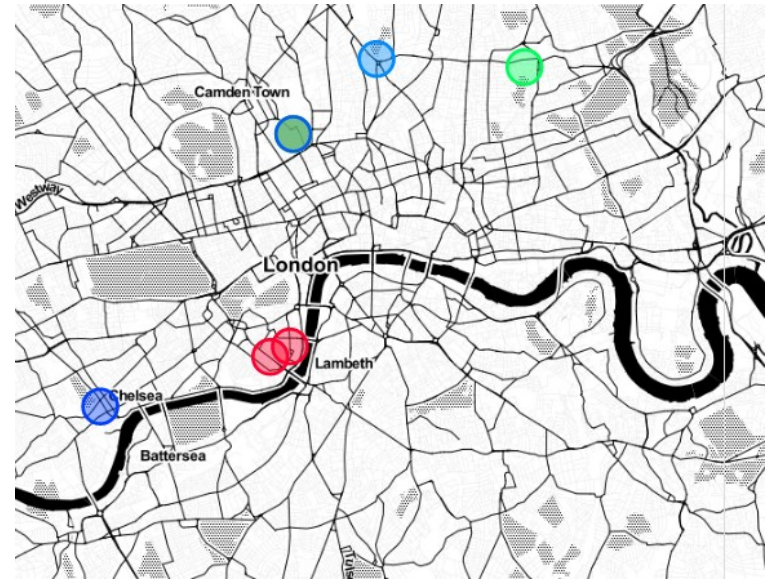
- The list of London areas and their boroughs is obtained from a table from Wikipedia
- The latitude and longitude of these boroughs is obtained using the `geocoder` package. This information will be used to search for venues within those boroughs
- Venue data is obtained using the FourSquare API

# Methodology

- scrapping Wikipedia using `BeautifulSoup`
- cleaning the data using regex
- obtaining location for each area using `geocoder`
- acquiring a list of venues using FourSquare
- groupping the data by area and filtering out the ones that do not conform to the required venue categories
- look for the closest ones to the centre of London

# Results and Discussion

- there are 9 areas in London which conform to our client's needs: Bethnal Green, Holborn, St James's, and Wapping.
- Convent Garden and St James's are the closest ones to the centre of the city.



# Conclusion

- Given our findings above, we can strongly recommend our client to consider the areas of Convent Garden or St James's. Both are very central and have all basic amenities required within a radius of 500m.
- Three areas in the borough of Camden could also be good choices but are a little further away from the centre of the city than the selected two.