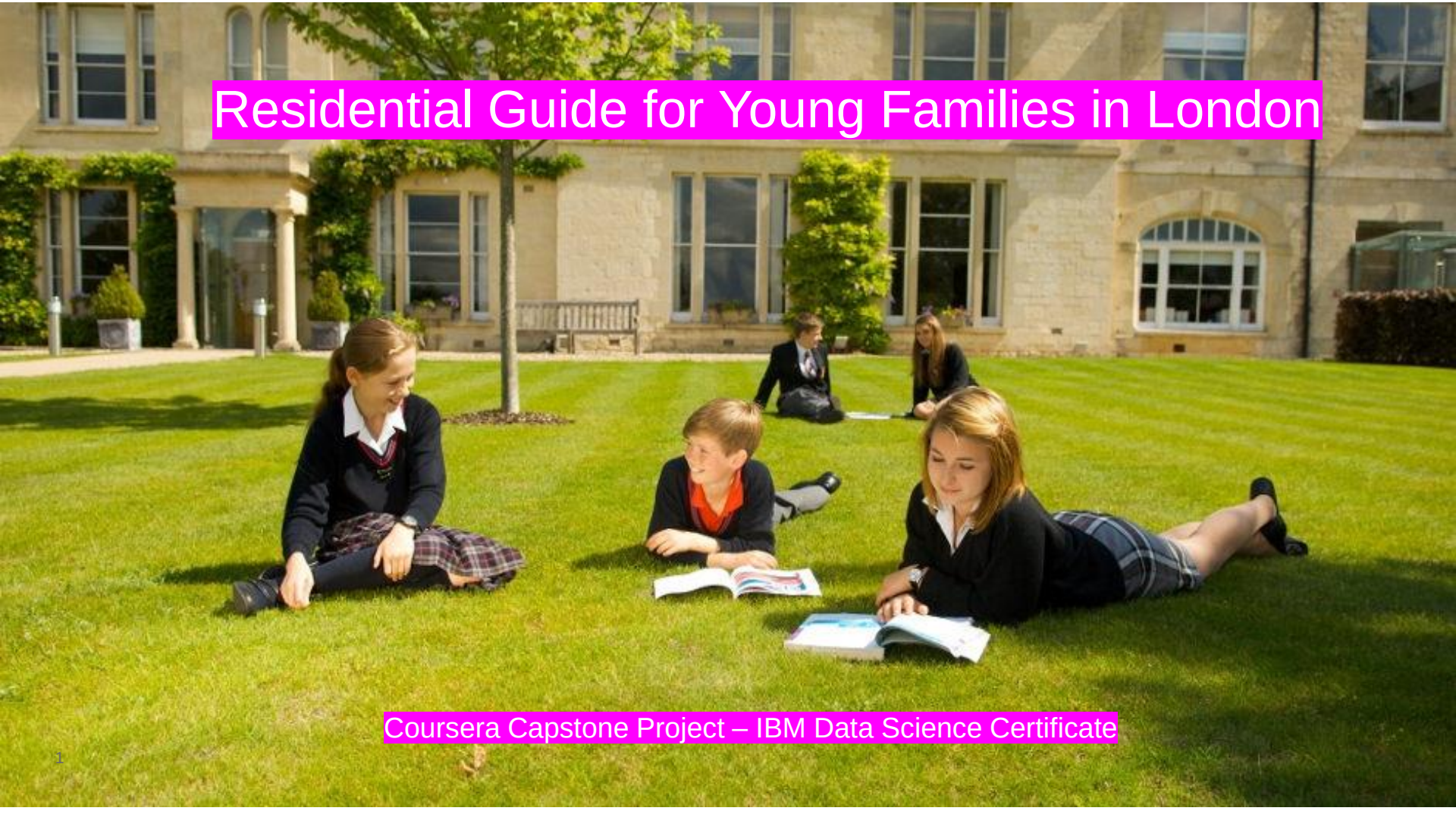
Residential Guide for Young Families in London



Coursera Capstone Project – IBM Data Science Certificate

Contents

Project Background

- London, England is made up of **33 boroughs** with relatively important differences between areas. The analysis aims to assist young families to help identify suitable residential areas if they wish to move to the city.

- As part of the project, we have identified a number of key indicators and requirements that a certain area would have to possess in order to be considered “attractive” to a young family. Though preferences as to these indicators may vary among families, we assume that there are certain basic elements that would be shared across families with young children. We have identified ***primary education opportunities*** as well as the availability of ***child-friendly recreational activities*** as being of the utmost importance for young families.

- The project could certainly be further enhanced down the line with the inclusion of crime statistics or residential real estate prices.

2. Data and Methodology

- For the purposes of analysing the ***primary education opportunities,*** we have used the following methodology.
  - The Mayor of London office makes available a number of key data sources free-of-charge on the ***data.london.gov.uk*** site. In order to commence the analysis, we have obtained the following databases:

    - **1. School Database:** The database provides a Borough-by-Borough breakdown of key characteristics of the schools. The database contains invaluable information that helps us assess and rank the quality of education in each of the areas of the city. The analysis calculates a number of key performance indicators (% of students in independent schools; % of students in highly-rated schools, etc.) and uses a number of other available variables (number of highly-rated schools, etc.) to rank the educational "attainment" of each Borough.

    - **2. Economic Indicators:** We have alimented the initial analysis by adding in a snapshot of

2. Data and Methodology (cont'd)

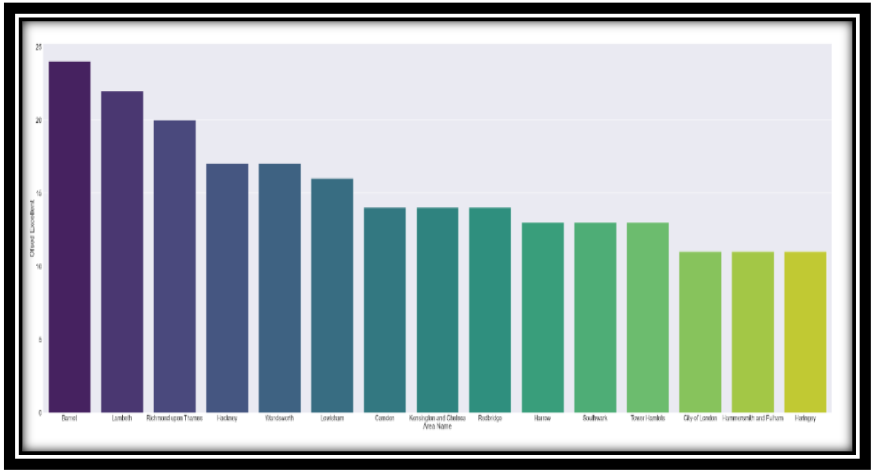
- For the purposes of analysing the ***child-friendly recreational activities,*** we have used the following methodology.
  - We have used the Foursquare API tool to get a breakdown of available recreational / child-friendly venues within each Borough.
  - We have identified around 10 venue types (e.g. Park, Market, Performing Arts Venue, Garden, etc.) that we believe young families would be particularly keen on having nearby.

Data Analysis

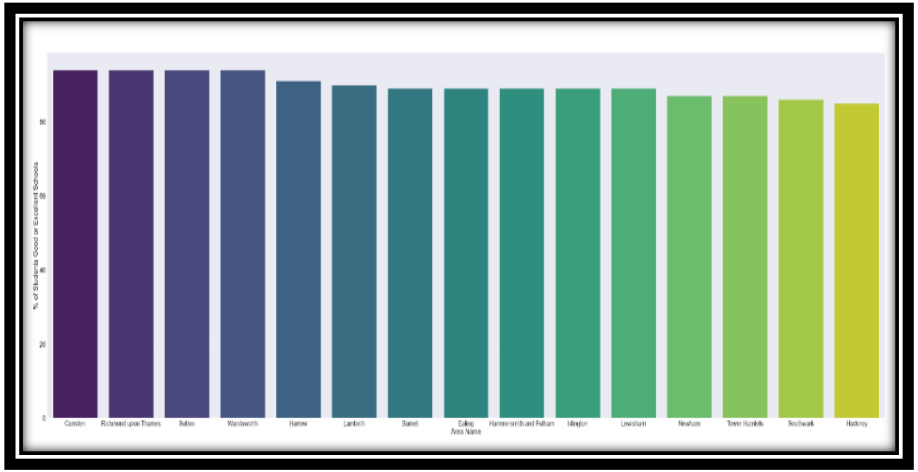
Data input and data wrangling

- We were able to download and use .csv data files for our project. These were obtained from a dedicated data server hosted by the City of London. We have also obtained and data coordinates of each Borough from Wikipedia to help facilitate our analysis.
- We have performed exploratory data analysis by first calculation key indicators on the School database and subsequently performing a ranking of Boroughs based on educational attainment.
- The below graphs are just two examples of the detailed exploratory data work.

Graph 1: Top 15 Boroughs with highest school ratings



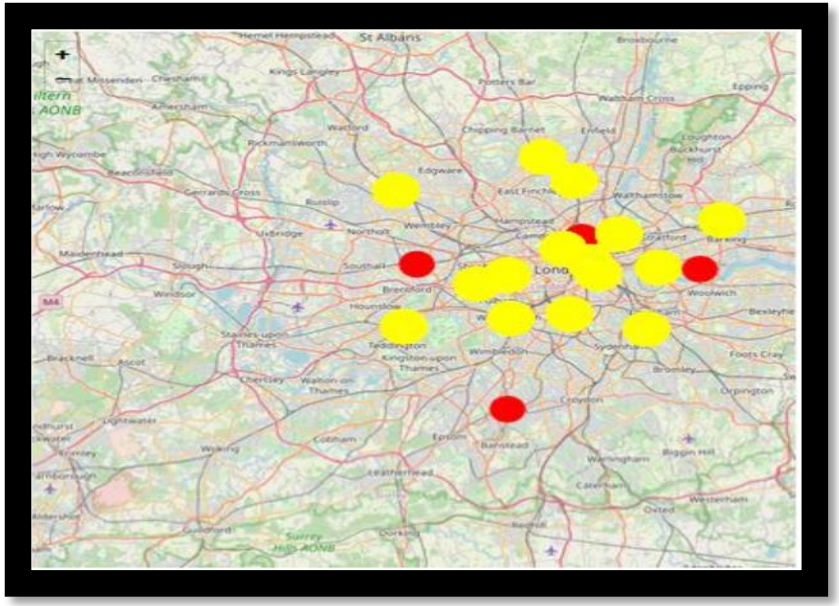
Graph 2: Top 15 Boroughs with highest % of students in highest rated schools



Data Analysis (cont'd)

Data input and data wrangling

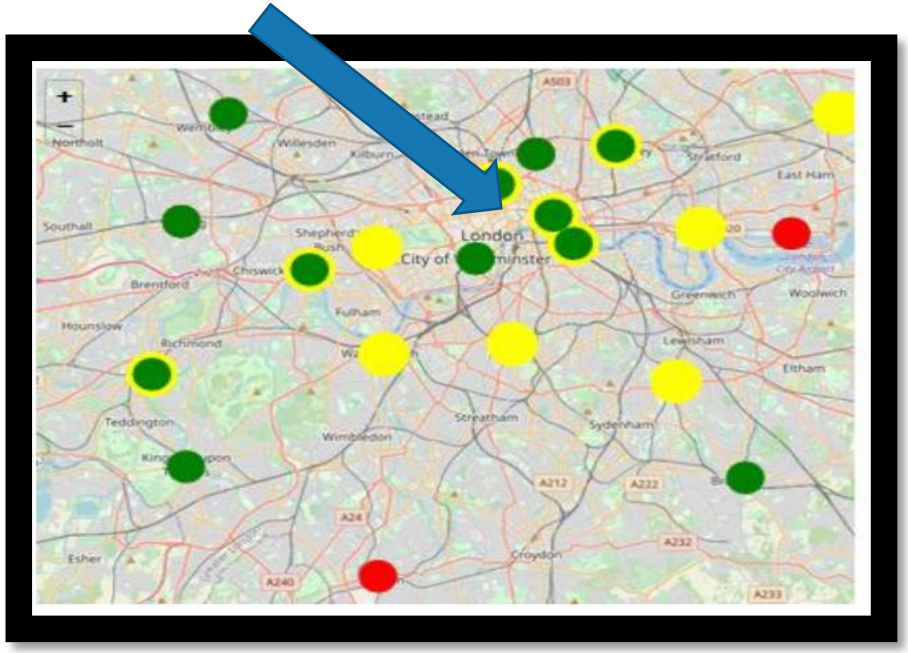
- Using the FOLIUM mapping solution, we are able to create a visualisation of the 15 boroughs with the highest educational attainment ***(in yellow)***. **These areas represent the base data that we will use to cross-check and map against the Foursquare API data.**



Data Analysis (cont'd)

Data input and data wrangling

- Using the Foursquare API data, we are able to identify and geographically visualise datapoints for Boroughs with the highest number of "family-friendly" venues. ***These datapoints (in green) are overlaid the school statistics (previously identified) to show areas of potential interest, i.e. a combination of high-performing schools with high number of family-friendly recreational venues.***



4. Results and Conclusions

We have identified the following 2 key observations and recommendations. We have also identified the top 5 residential areas young families moving to London may wish to explore further.

1. Not surprisingly, we noted that the ***proportion of students in independent schools tends to be highest in the wealthiest boroughs in London***. Perhaps more surprisingly, our second finding shows that ***school performance is not correlated with the respective borough's economic status.*** We would have expected that wealthier areas may also house the best schools, however this is not necessarily the case.

2. As we have identified the top-performing Boroughs from an educational standpoint and as we overlaid the availability of family-friendly amenities, we have identified and ranked the Boroughs to come up with the recommended residential choice for young families, namely: ***City of London, Hammersmith and Fulham, Richmond, Hackney and Camden.*** As a long-time resident of London, based on anecdotal evidence, the above conclusion seems to be in line with the results of this analysis.