Parks and Recreation

By Joshua Montgomery

Designing for demand

If you were given the money and resources to build a new park, how would you design the park to get the largest possible demand?

This presentation will demonstrate how location and size of a park correlate with the number of visitors.

Modelling Choices

- London was chosen as the city for to analyse to its broad spectrum of parks both publicly and privately owned. Cities like New York too few parks to make this analysis worthwhile, but still have a similar amount in area.

- This model is largely aimed at a park to be constructed by a local government, hence focusing on parks that are free to enter, and mainly consist of open green spaces.

- To utilise the Foursquare API, a centre point had to be defined, in this case Buckingham Palace due to its central location and recognisability

- The parks included in the data set or all within 10km of Buckingham Palace, hence parks in the greater London area, such as Hampstead Heath, have been neglected.

Foursquare Output

After collecting the data from Foursquare, the following map was plotted of parks included in the data set
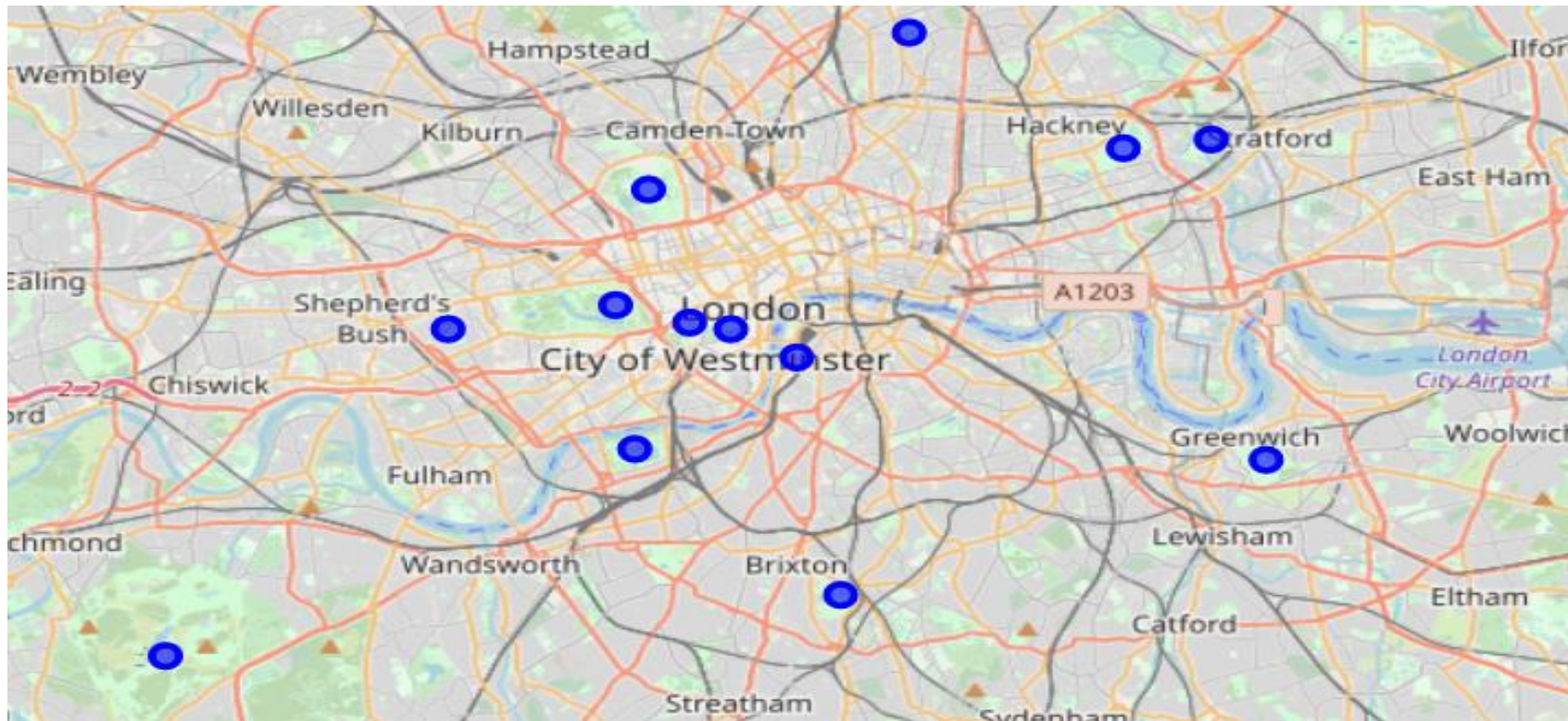
Multiple Linear Regression

▶ The chosen model was multiple linear regression, with the independent variables being distance and size. The dependent variable was the number of visitors per annum.

▶ The data was split into testing and training data, so that the model could be analysed.

▶ When fit was then evaluated with the test data and found to have a residual sum of squares of 6.8 and a variance score of 0.03. Hence this model may be inappropriate for data sets in the extreme ranges of either independent variable.

Proposed Parks

The model was then used to predict visitor numbers of several hypothetical parks

- ▶ If the grounds of Buckingham Palace became a public park

- ▶ A small park situated centrally

- ▶ A large park on the outskirts of London

Evaluating Proposed Parks

Name	Size (Acres)	Distance (m) to centre of London	Predicted Visitors Per Year (million)
Buckingham Palace	39	0	8,81
Small Park	10	100	8.72
Large Park	800	10000	4.34

The model shows a clear preference for parks that are centrally located as opposed to larger parks that are further away.

Problems With the Model

- Availability of data on visitor counts constrained the scope of this model
- The model is only applicable for London, but could be easily adapted for other cities.
- The data for visitor counts had to be collected manually.
- Data was only gathered for parks that are free to access

Future Analysis

▶ Gathering more data through freedom of information requests would lead to a superior model, and more independent variables.

▶ Performing the analysis again but with profit as the dependent variable would be more commercially useful

▶ The analysis could be expanded to take into account safety by analysing crime statistics for each park

▶ The model could be updated for the current environment by considering that in smaller parks there may have to be a capacity limit to ensure effective social distancing.

Thanks for reading this presentation!

I hope you found it interesting.