22BSP415

Logistics Modelling and Operations Analytics Coursework

‘Determining Distribution Centre Locations for delivery to online shopping customers’



Group: Loughborough’s Finest

Coursework Deadline: Wednesday 3rd May

Word Count: 3103

# **Executive Summary**

A major international online retailer intends to establish several regional distribution centres (DCs) in Greater London to deliver products to online customers and has asked us to help them find an optimal solution regarding the allocation of each distribution centre. The software used to create the models and find the optimal solution for each problem has been done in Spyder. The data gathered was consolidated in Excel and then prepared to be placed into the software. After, finding the solutions, a series of results analysis has been conducted to find which approach would be the best for the international online retailer. Although, the 2nd model would have a better impact on customer service as the customers would receive their packages in a shorter period of time. The recommended option would be to choose the 1st model, as it has the lowest annual operating cost, hence it aligns with the company’s objective of minimising costs.

**Table of Contents**

[Executive Summary 2](#_Toc133994722)

[1.0 Introduction 4](#_Toc133994723)

[1.1 Factors for consideration 4](#_Toc133994724)

[2.0 Data Collection and preparation 5](#_Toc133994725)

[3.0 Model and Implementation 7](#_Toc133994726)

[3.1 Mathematical Modelling 9](#_Toc133994727)

[4.0 Results Analysis 10](#_Toc133994728)

[4.1 1st Model Results 11](#_Toc133994729)

[4.2 2nd Model Results 12](#_Toc133994730)

[4.3 Comparisons between the models 13](#_Toc133994731)

[5.0 Recommendations 13](#_Toc133994732)

[6.0 Conclusion 14](#_Toc133994733)

[7.0 References 14](#_Toc133994734)

# **1.0 Introduction**

A major international online retailer intends to establish several regional distribution centres (DCs) in Greater London to deliver products to online customers. The business seeks to reduce total annual costs, including delivery and DC running expenses. A residential area's population can be thought of as proportionate to the number of items to be delivered there. Therefore, it is necessary to only consider the region's largest cities and towns as potential DC locations and demand locations when deciding where to locate DCs. Our team has been asked to assist the company to choose where to locate the distribution centres in Greater London and must create a location-allocation model to solve the DC location problem.

To obtain the optimal solution, we have created a model in Python. The report's next sections cover our assumptions, data collection and preparation, model and implementation (which includes our mathematical formulation), results analysis, followed by our recommendations and conclusions.

As a result of model 1's three DCs satisfying the objective function of reducing total costs, we think it to be the optimum choice.

# **1.1 Factors for consideration**

Finding optimal locations for the DCs is a challenge that an international online retailer may experience while setting up many regional distribution centres (DCs) in Greater London. Before choosing a location, a variety of factors must be considered, including:

* Proximity to customers: This is significant because it may influence the service provided by the online retailer in terms of latency and network performance. The retailer may experience a slower network connection if the DCs are too far away from buyers, which will have a negative effect on customer experience, potentially resulting in a loss of customers.
* Transportation costs: The cost of transporting goods between the DC and its clients will vary depending on where the DCs are located. A more central position can aid in lowering transportation costs and improving supply chain effectiveness.
* Transportation infrastructure: Location can affect transportation infrastructure, which can affect how quickly the supply chain responds. Reduced lead times and faster order fulfilment can be achieved with a DC close to suppliers and transportation hubs.

The following are things we already know and should consider when creating the model and implementation:

* Each distribution centre’s capacity is assumed to be sufficient for servicing all regional demands.
* For the first task, at most three DCs are to be set up in the region.
* The annual operational costs for DC are £120,000, while the per-mile transportation cost is £1.50

# **2.0 Data Collection and preparation**

Information about interesting factors is gathered through the process of data collection. We obtained the data for the population estimates from city population, where we chose the top 20 most populated locations in Greater London region. After gathering the data, we have created a table in Excel and placed the data there to make it more comprehensible. Data preparation is crucial as it enables effective data analysis and lowers the possibility of processing errors. After determining the 20 most popular locations, we determined the annual demand for each of them, and we input this information into an Excel table, shown in Figure 1. As the annual demand is one unit for every 100 people, we divided each population by 100 to get this figure. Since each possible DC will have enough resources to meet all area demand, the overall capacity is then determined by adding up all annual demand, which comes to 63,210 units.

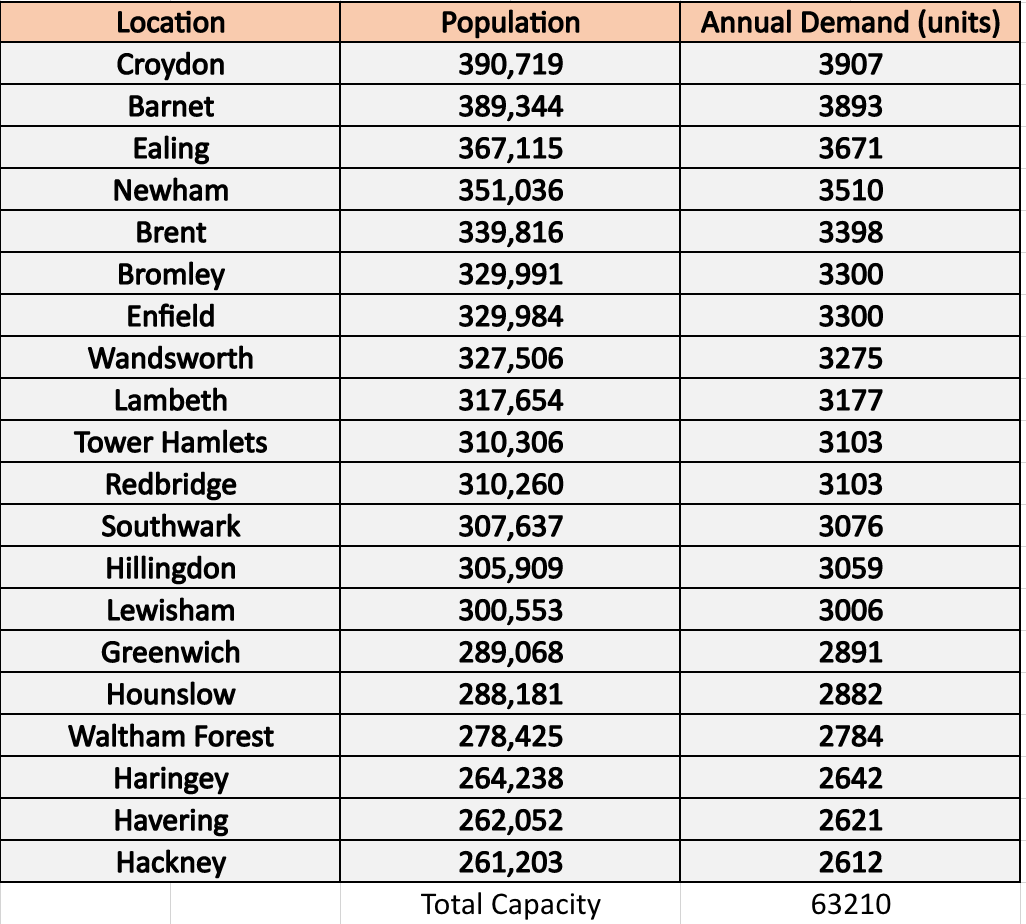


Figure 1: shows the population and annual demand for the demand locations.

Following completion of this, we used Google Maps to determine the distances between each potential DC location and each demand location. We assume that this is the best route, one without detours and heavy traffic. With this information, we created a new Excel table (Figure 2) that made it simple to see the distances.

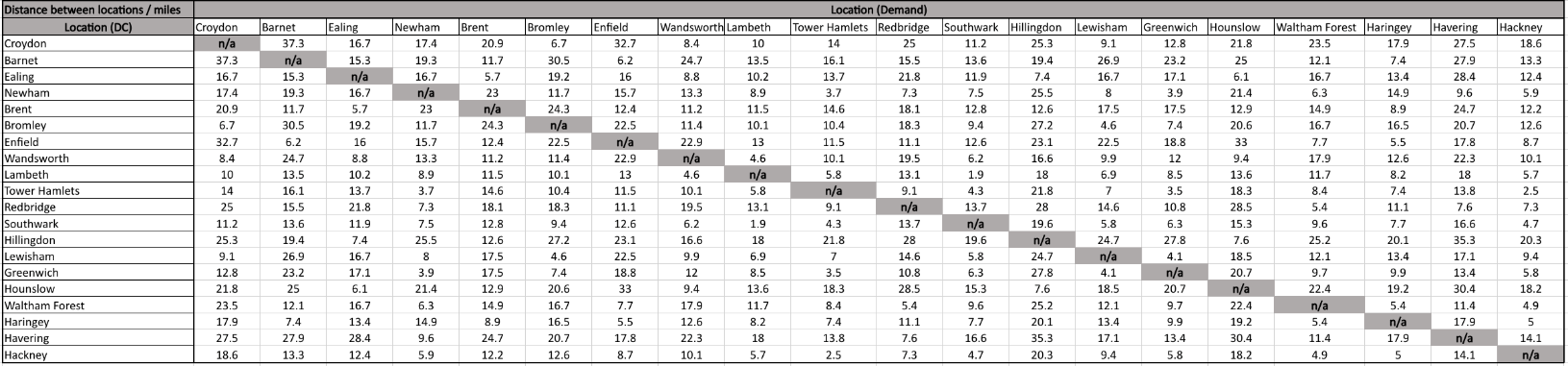


Figure 2: shows the distances in miles between each potential DC and demand location

The final step was combining the distances, capacity, annual demand, transportation cost, and annual operational cost into a single table from which Spyder could be used to extract the optimal solution. We added a column on the right and entered the information on the DC capacity which was previously calculate by adding all the annual demand. The annual operational cost has been identified as £120,000, and a column is added to the right of the capacity to reflect this. The following stage involved adding a row at the bottom that displayed the annual demand for each location as determined in the previous phase. We added another row below this for the transportation cost, which was calculated as the transportation cost per unit demand per mile by multiplying the annual demand by 1.5. The final table, shown in figure 3, was then finished and prepared for usage in Spyder.

Calendar

Description automatically generated

Figure 3: shows the final table

In addition, the raw data above is computed using Microsoft excel software package. There are five parameters captured in Excel sheet for python (Spyder) to read and run the model. Distribution centres and demand locations are denoted by ‘i’ and ‘j’ respectively for ease of computation and listed down in column A and row 1 of sheet 1. Annual demand and operating costs are designated by variables ‘ki’ and ‘fi’ that take up column’s ‘V’ and ‘W’. Lastly, the location wise annual demand and transportation demand is captured using variable rj and wj.

# **3.0 Model and Implementation**

To determine the optimal locations for the DCs, we developed a location allocation model that can be solved in Spyder. Spyder is a free and open-source scientific environment for Python that combines data exploration with sophisticated analysis, debugging, editing, and profiling. A multilingual editor window is available in Spyder to generate, open, and edit source files (Wilfong, 2022).

Importing the Xpress, Pandas, and OS libraries that we needed to write our code came first. With Xpress, we can create and solve optimisation problems; with pandas, we can utilise a quick, adaptable expressive data structure; and with OS, we can run a command from a Python script. The following step was creating the working directory in the chosen folder. We created a shared folder where we could place our work and use it as our working directory because this was a collaborative effort.

The parameters we were going to use such as potential DC locations, the number of user locations, and the maximum number of DCs, were to be specified after that. Due to the number of locations in our data file, the number of user locations and prospective DCs was set to 20, respectively. We were informed while completing the first exercise that the maximum number of DC was set at 3.

The distance matrix for the DC and users had to be obtained after that. To do this, we used a code that we placed into Spyder to read the excel document we had created and to obtain the data related to the columns between columns B and U. The capacity and yearly operational costs of the DC had to be established next. To do this, we wrote two lines of code: one to read the data from only column W, which had the annual operating cost data, and another to read the data from column V, which contained the capacity data. The skip row's function was then used to determine the unit cost and transportation demand. To read the following rows transport demand and the unit cost row, we utilised skip rows to skip every location for the transport cost and every location plus the next row (transport demand) to get the unit cost.

Following this, we named the problem to ‘A Location Allocation Problem’. After successfully importing data onto Spyder, the decision variables were identified as ‘z’ and ‘x’, the prior to decide the variable for each candidate DC location and the later to decide the variable for each candidate DC location serving a particular demand location. The decision variables were added onto the model using the add variable() function. It is imperative to define the set of constraints to arrive at an accurate feasible solution. For the first model 3 constraints were set in place, the first constraint was to ensure that each DC location was serving at least one demand location which would then meet the requirements of serving all the demand locations. Constraint two was put in place to ensure the capacities of each distribution centre were not exceeded. Finally, constraint three was specifically to address the first model which was to limit the maximum number of DC’s to 3.

The first component sums up the product of the operating cost for each distribution centre (DC) (found in the operating cost data frame at row i, column 0) and the binary decision variable z[i]. which has a value of 1 if the DC is open and a value of 0 otherwise, from the first DC to the last. This determines the selected DCs' overall operating costs. The second component calculates the sum of the unit transportation costs from each DC (retrieved from the unit cost dataframe at row 0, column j), the distance between each DC and each demand location (retrieved from the distance dataframe at row i, column j), and the number of goods shipped from each DC to each demand location (represented by the decision variable x[i,j]). This determines the overall shipping expense for all shipments made from the chosen DCs to the demand destinations.

The goal of the objective function is to add up the costs associated with operating and transporting associated with the chosen DC’s. Finding the values of the decision variables x and z that minimise this objective function is the goal of the optimisation problem. When everything was completed, we could run the code to solve the problem and print the solution.

To solve the second part of the client’s requirement we identified the largest distance captured from the results of the first solution, assigned it to a variable (max\_distance) and implemented a 4th constraint. The 4th constraint would allow us to limit the maximum distance between the distribution centres and the demand location by 80% of the largest distance identified earlier, which was 14.1 miles. All 4 constraints were embedded with two for loops for an effective iterative process to recognise the most optimal solution. The first for loop transcends along the various distribution centres taken into consideration the latter progressed across the various demand locations. The constraints were added onto the model so the objective function could deliver solutions based on the set of constraints.

## **3.1 Mathematical Modelling**

The following is the mathematical formulation for this location-allocation problem with general distances:

Parameters and notations:

|  |  |
| --- | --- |
| Dij | Distance from distribution centre (DC) to demand location |
| i | Distribution Centre |
| j | Demand Location |
| ki | Annual demand |
| fi | Annual operating cost |
| rj | Location wise Annual demand |
| wj | Transportation demand |

Our Objective function

The first part of the objective function accounts for the annual operating cost and the second part depicts the annual transportation cost. ‘wj’ represents the weight of the transport demand.

Subjected to the following constraints

Constraint 1: each customer is only served by one facility

Constraint 2: the capacity constraint at each facility should not exceed the annual demand.

Constraint 3: maximize the number of DC’s to 3.

Constraint 4: minimize the largest distance identified in solution 1.

is a binary variable that represents a “selection” decision and can take up values ranging from 0 to 1.

xij is a binary variable that represents a “assignment” or “allocation” decision.

I and J can take up values ranging from 1 to n and 1 to m respectively. Where n represents the last DC location and m represents the last demand location.

# **4.0 Results Analysis**

After the two models have been created, we ran the code to solve the problems. A table (Figure 4) was then created to record our findings after finding the optimal solution for both models. This section will discuss the findings in each approach and will conclude with a comparison of both models.

Table

Description automatically generated

Figure 4: Table of results for 1st and 2nd model.

## **4.1 1st Model Results**

Map

Description automatically generated

Figure 5: DC Locations (Model 1)

During the process of running our model, we have found that the optimal solution for the company was to set up 3 DC locations which would supply to all 20 demand locations. The annual operating cost of running these 3 DC locations would be £919,735. The model has chosen Ealing, Bromley and Hackney as the distribution centres, the reasoning behind the decision was the distance between the locations. The task was to minimise the operating costs as well as the transportation costs, which explains the DC positioning in Figure 5. This way the distribution centre in Bromley can reach the locations in the East and South, whilst the DC in Ealing is able to reach the demand locations in the West, which leaves Hackney to reach the centre of London as well as the North/East region.

One concern with this choice would be the distribution of centres between the 3 DC locations, as shown in Figure 4, the distribution centre in Hackney is distributing to 12 demand locations. Therefore, if a problem would occur, those 12 locations would have been attended by the nearest DCs - Ealing and Bromley (which are moderately far from the demand locations), hence the transportation cost would increase immensely.

## **4.2 2nd Model Results**

Map

Description automatically generated

Figure 6: DC Locations (Model 2)

The second model was constructed to abide to a new restriction which limits the distance between any residential area and the supplying DC to be at most 80% of the longest distance identified in the 1st Model. Due to this new constraint more DCs have been required.

After revising the 1st model, we have been able to solve the problem and get an optimal solution for this new situation. 4 DCs have been identified: Ealing; Newham; Enfield and Lambeth. The annual operating cost for this model is £952,951. The distribution of demand locations per each distribution centre is evenly distributed as shown in Figure 4. The DC placement is perfect in this scenario as each DC is in one of the four cardinal directions. Enfield supplies the demand locations which are in the North; Newham supplies the demand locations in the East; Lambeth supplies demand locations in the South and Ealing supplies the West region.

Furthermore, due to the new restriction all the distances between the distribution centre and the demand locations are under 11.28 miles, with the longest distance being 10.1 miles. Shorter distance between the DCs and demand locations will improve the delivery times.

## **4.3 Comparisons between the models**

After analysing both models we have found some similarities and differences between them:

* Model 1 has a lower annual cost (**£919,735**) compared to Model 2 (**£952,951**).
* Model 1 contains less DC locations (**3**) in the optimal solution compared to Model 2 (**4**).
* Both models contain **Ealing** as one of the DC locations in the optimal solution.
* Model 2 has a lower ‘longer distance’ (**10.1 miles**) value compared to Model 1 C
* The distribution of demand locations per each DC is more equal in the 2nd Model compared to 1st Model.

Although both models have been executed perfectly as we have managed to find the optimal solution for both problems, there are a few limitations. For example, the major limitation in this task was that the gathered data based on the distance between DC and demand locations are time-specific, which means that if we have tried to collect them on another day, the values that we would get, would be different. Therefore, this would have a direct impact on the total cost, resulting in a different outcome.

# **5.0 Recommendations**

After considering the discussions above, we suggest that if the company prioritises minimising costs, then Model 1 will be the best option. However, if they prioritise customer satisfaction, then Model 2 could be considered, as there is a more even distribution of demand. For example, in Model 1, Hackney which is one of the DCs selected is assigned to 12 demand locations where as the other 2 DCs are assigned to 5 or less demand locations. Therefore, if an issue with their systems were to occur, the other DC locations would have to attend to extra demand locations which would slow down the process and increase the transportation cost due to the distance. Whereas in Model 2 it would be easier for the other DC to pick up the slack, resulting in less extra cost compared to Model 1.

We recommend spreading the demand more evenly among the DCs and ensuring an appropriate number of DCs are used. This strategy would help avoid any disruptions in service if one or more DCs experience incidents. Although the DCs are flexible enough to accommodate the new demand allocation, the cost of adjusting might increase significantly due to varying distances from the DCs to online shoppers. Interestingly, the recommended model has a more diverse allocation than the previous one, but one DC is still responsible for almost twice the number of demands as the others.

# **6.0 Conclusion**

In conclusion, our team developed two effective models to address the location-allocation problem by using python(Spyder) and significantly minimising costs, which include both fixed and variable costs. Based on our thorough analysis and assessment, we recommend selecting the first model as it is the best option for achieving the company's cost-saving objective. The cost of implementing the first model is lower at £919,735 compared to the second model, which costs £952,951 annually. The first model has fewer distribution centres (DCs), resulting in a longer travel distance of 14.1 miles. However, the second model addresses this issue by adding more DCs of the annual operating cost of £120,000 which exceeds the cost saving from the previous one, resulting in a shorter travel distance of 10.1 miles. Although this has increased the annual cost, the advantage of the second model is that the demand locations are more evenly distributed among the DCs, which reduces the risk of disruption in service in case one DC is experiencing issues.

# 

# **7.0 References**

*Europe* (no date) *Greater London (United Kingdom): Boroughs - Population Statistics, Charts and Map*. Available at: <http://www.citypopulation.de/en/uk/greaterlondon/> (Accessed: 02 May 2023).

Hübner, A., Holzapfel, A. and Kuhn, H., 2016. Distribution systems in omni-channel retailing. *Business*  *Research*, *9*, pp.255-296. Available at: [https://link.springer.com/article/10.1007/s40685-016-](https://link.springer.com/article/10.1007/s40685-016-0034-7)0034-7 (Accessed: 02 May 2023)

*Package overview#* (no date) *Package overview - pandas 2.0.1 documentation*. Available at: <https://pandas.pydata.org/docs/getting_started/overview.html#:~:text=pandas%20is%20a%20Python%20package,world%20data%20analysis%20in%20Python> (Accessed: 02 May 2023).

Shubham (2022) *Python OS module*, *DigitalOcean*. Available at: <https://www.digitalocean.com/community/tutorials/python-os-module> (Accessed: 02 May 2023).

Wilfong, S. (2022) *PyCharm vs. Spyder: Which should you use?: BestColleges*, *BestColleges.com*. Available at: <https://www.bestcolleges.com/pycharm-vs-spyder-which-should-you-use/#:~:text=What%20Is%20Spyder%3F,open%2C%20and%20modify%20source%20files> (Accessed: 02 May 2023).

*Xpress* (no date) *PyPI*. Available at: <https://pypi.org/project/xpress/#:~:text=The%20Xpress%20Python%20interface%20allows,MISOCP%2C%20together%20with%20general%20nonlinear> (Accessed: 02 May 2023).