Group Coursework and Project Cover Sheet

Third Year Undergraduate

Department of Civil and Environmental Engineering

Module CIVE60008: Transport Systems

Assignment Group Project 2: Freight Distribution Planning

Supervisor Dr. Panagiotis Angeloudis

Deadline Friday 17th December 2021

DECLARATION

I certify that I have read the definition of plagiarism given overleaf, and that the work submitted for this coursework assignment is my own work, except where specifically indicated otherwise. In signing this document I agree that this work may be submitted to an electronic plagiarism test at any time and I will provide a further version of this work in an appropriate format when requested:

Name: Iyanuoluwa Adewumi **CID:** 01492014 **Signature**: Date: 10/12/2021

Name: William Wang **CID:** 01719755 **Signature**: Date: 14/12/2021

Name: Hoong Hao Yap **CID:** 01728257 **Signature**: Date: 15/12/2021

Name: Sotos Lois **CID:** 01495893 **Signature**: Date: 15/12/2021

Name: Cheuk Kit Ian Tsang CID: 01700432 Signature: Date: 15/12/2021

Note: Until an assignment carries this completed front page it will not be accepted for marking. If the front page is absent, the delay in getting it added may result in a penalty for late submission.

TO BE COMPLETED BY THE MARKER

Grade awarded: ______________________________

Late penalty applied: ____________________________

PLAGIARISM

You are reminded that all work submitted as part of the requirements for any examination **(including coursework)** of Imperial College must be expressed in your own writing and incorporate your own ideas and judgements.

Plagiarism, that is the presentation of another person's thoughts or words as though they are your own, must be avoided with particular care in coursework, essays and reports written in your own time. **Note that you are encouraged to read and criticise the work of others as much as possible. You are expected to incorporate this in your thinking and in your coursework and assessments. But you must acknowledge and label your sources.**

Direct quotations from the published or unpublished work of others, **from the internet, or from any other source** must always be clearly identified as such. A full reference to their source must be provided in the proper form and quotation marks used. Remember that a series of short quotations from several different sources, if not clearly identified as such, constitutes plagiarism just as much as a single unacknowledged long quotation from a single source. Equally if you summarise another person's ideas, judgements, **figures, diagrams or software,** you must refer to that person in your text, and include the work referred to in your bibliography and/or reference list. **Departments are able to give advice about the appropriate use and correct acknowledgement of other sources in your own work.**

The direct and unacknowledged repetition of your own work, which has already been submitted for assessment, can constitute self-plagiarism. Where group work is submitted, this should be presented in a way approved by your department. You should therefore consult your tutor or course director if you are in any doubt about what is permissible. **You should be aware that you have a collective responsibility for the integrity of group work submitted for assessment.**

The use of the work of another student, past or present, constitutes plagiarism. Where work is used without the consent of that student, this will normally be regarded as a major offence of plagiarism.

Failure to observe any of these rules may result in an allegation of cheating. **Cases of suspected plagiarism will be dealt with under the College's Cheating Offences Policy and Procedures and may result in a penalty being taken against any student found guilty of plagiarism.**

Table of Contents

1.0 Introduction

The team has been hired by a rapid grocery delivery start-up to establish a network of hyperlocal last-mile distribution depots around London, with delivery times of 7 minutes or less. The objective of the brief is to develop a facility selection strategy that would allow testing of the concept before further expansion. The company would like to test the concept in the London Borough Hammersmith and Fulham, as part of a year-long trial.

Our assumptions in the selection are:

- The total initial budget that the company has is £6,000.
- Neighbourhoods are defined in terms of LSOAs (Lower Layer Super Output Areas) – a geographical hierarchy unit used by the Office of National Statistics in the UK to collect census data.
- Warehouses can only be established in the centroid of an LSOA/neighbourhood.
- Depots cost £2,000 per two months to rent and operate, with the amount paid in full at the beginning of each two months.
- Other neighbourhoods can also be served from an established warehouse, as long as their centroids are within a distance of 0.5km.
- The average order provides revenue of £20.
- 15% of the total profit obtained every two months can be set aside for further expansion.
- New warehouses can be established every 2 months (i.e. in months 0, 2, 4, 6, etc.), as long as enough funds are available to cover their cost of operation.

The team have been provided with the following datasets to perform the analysis:

- A geographical dataset that provides a breakdown of neighbourhood data across London, organised in terms of LSOAs units.
- A CSV file containing individual LSOA codes for London, alongside their centroid locations, population data and the borough they belong to.

2.0 Data preparation

2.1 Visualize warehouse locations

Prior to performing an analysis, a visualization of the borough is completed with Geopandas. The map aims to visualize the centroid, and the warehouse location, of each neighbourhood and the population density of each neighbourhood (Figure 2.1). From the map, we can conclude that the population is randomly distributed with no significant trend in population distribution.

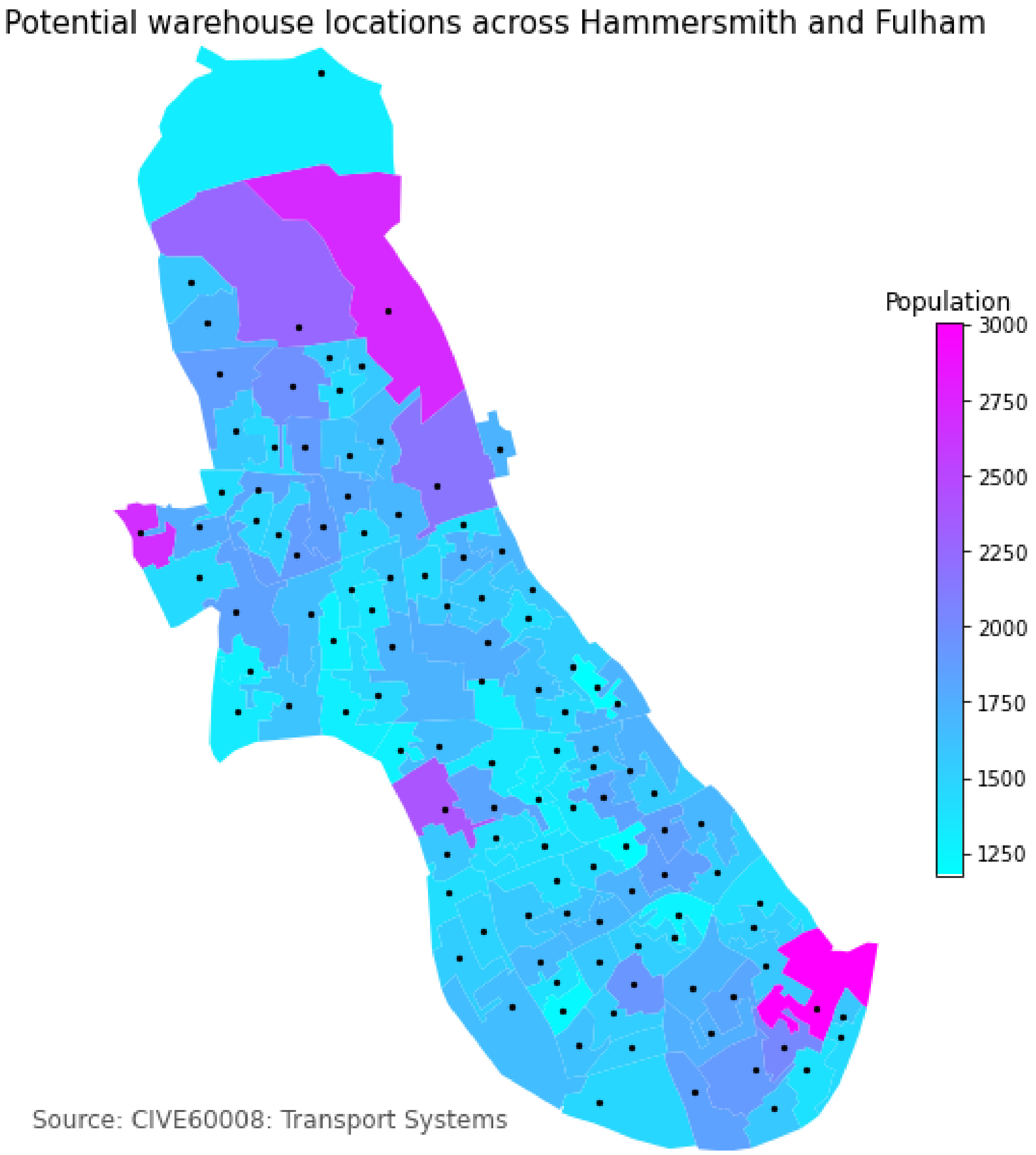

Figure 2.1: Map generated for Task 1.1

2.2 Neighbourhoods within range

This section aims to implement a code that will identify the neighbourhoods within a half-kilometre range from a centroid. This code will be used extensively in the upcoming tasks. Therefore, a function is prepared to evaluate the distance between a warehouse and the neighbourhoods in the borough. If a neighbourhood's centroid is within range, the entire neighbourhood is assumed to be within range and the centroid of the neighbourhoods is indicated red in colour on the map, with a blue line connecting the centroids. An illustration of the resulting map for neighbourhood E01001890 is shown in Figure 2.2.

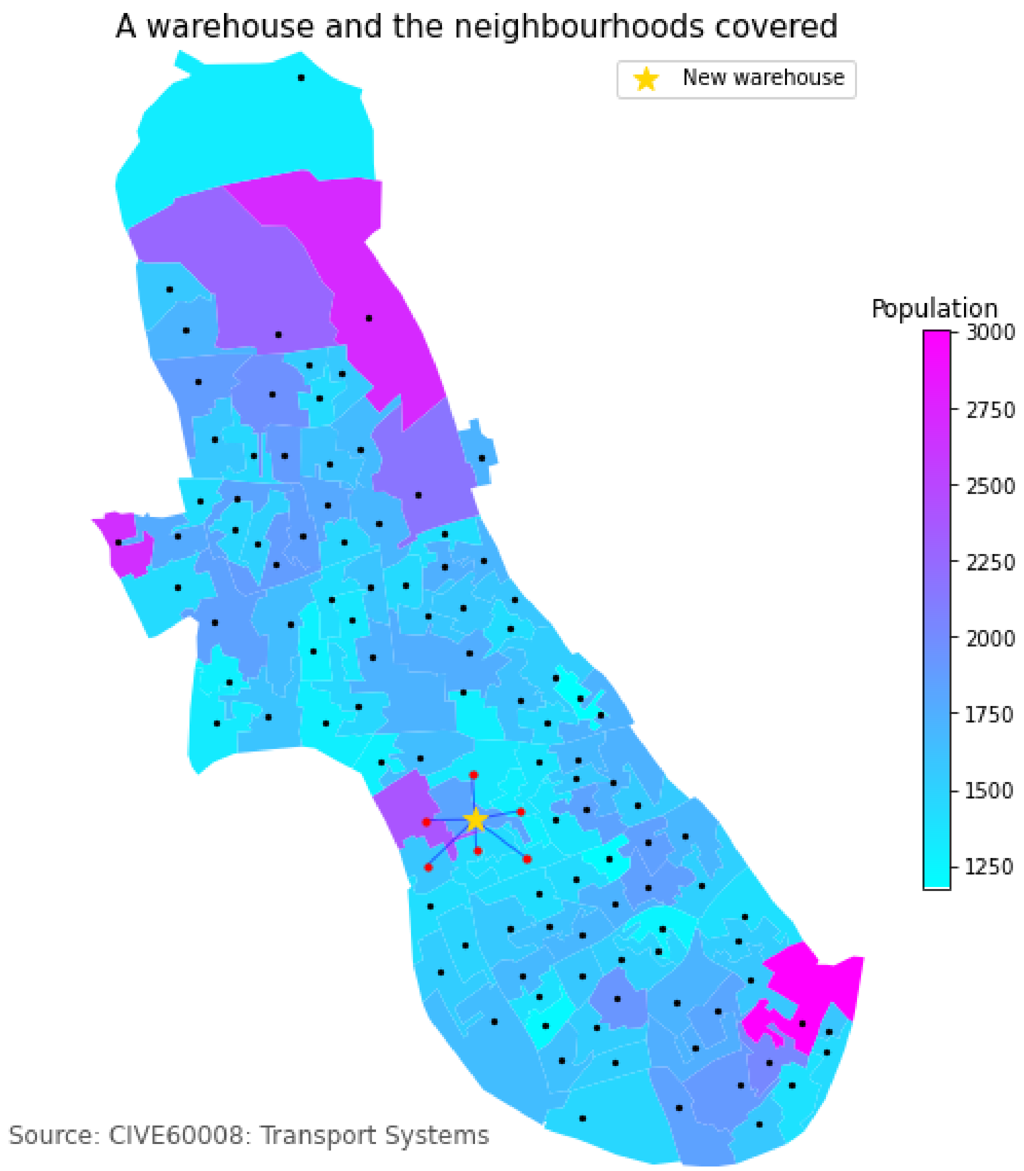

Figure 2.2: Map generated for Task 1.2

2.3 Warehouse locations with the highest coverage

With the code developed in the previous task, the warehouse locations with the highest population covered are identified in Table 2.1 and plotted in Figure 2.3. Based on the map, the selected centroids are not a good selection of warehouses in the first instance. It can be observed that the coverage population are overlapping and the uptake from the same population is not directly proportional to the amount of warehouse coverage. Instead of catering to existing customers, warehouses should be established with a goal to offer new customers its service and expand the business over a broader range of neighbourhoods. A more sophisticated selection strategy is required to account for overlapping coverage.

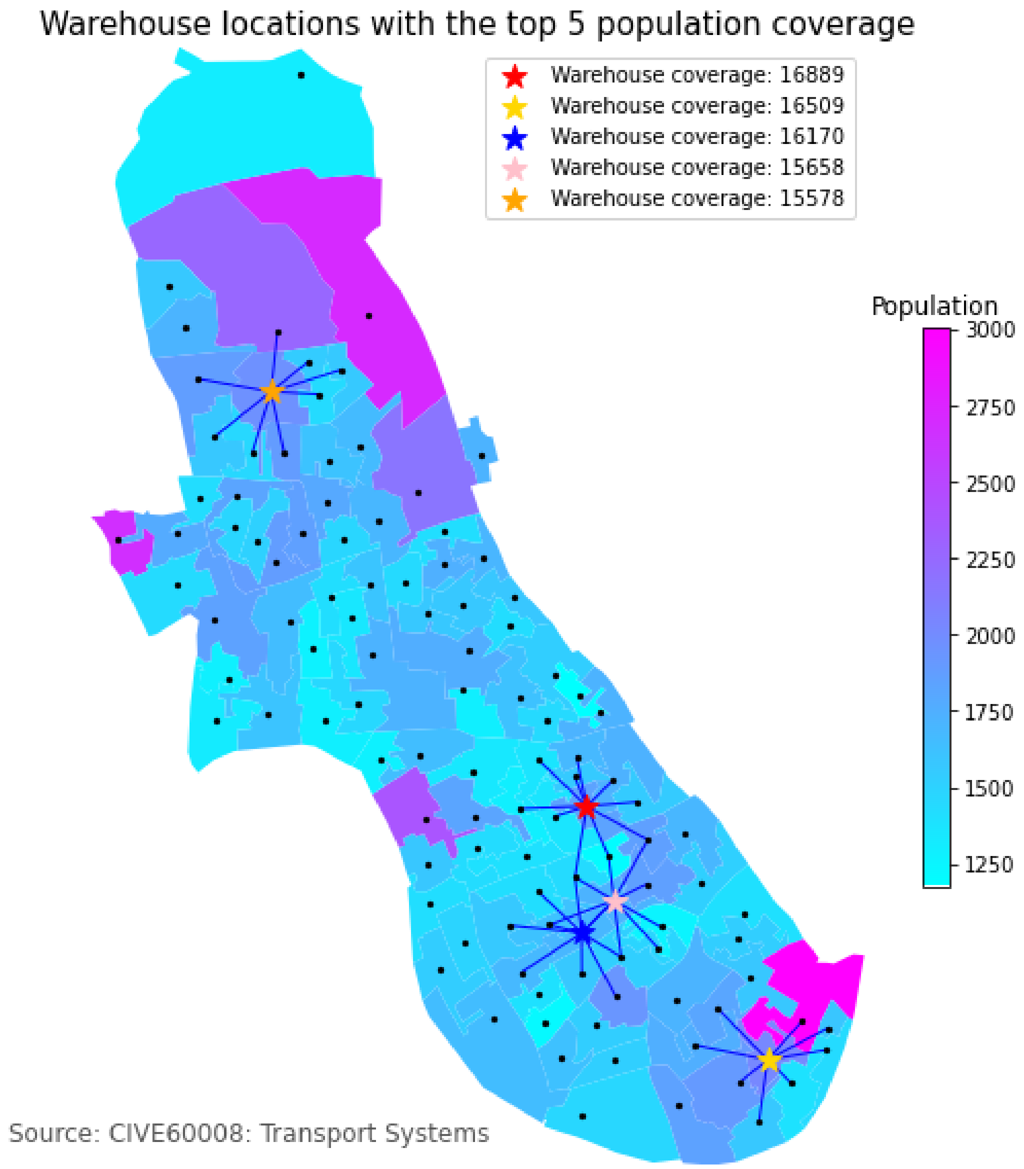

Figure 2.3: Map generated for Task 1.3

	Centroid name	Centroid LSOA	Centroids	Coverage area population
58	Hammersmith and Fulham 015C	E01001909	main59	16889
84	Hammersmith and Fulham 023A	E01001935	main85	16509
50	Hammersmith and Fulham 019B	E01001901	main51	16170
31	Hammersmith and Fulham 017C	E01001882	main32	15658
103	Hammersmith and Fulham 002A	E01001955	main104	15578

Table 2.1: Warehouses with the five highest population

3.0 Initial Location Selection

3.1 Mathematical model to maximise profits

L		$Set\ of\ depot\ locations$
N		$Set\ of\ neighbourhoods$
x_i	$i \in L$:	$Decision\ variable\ to\ open\ a\ depot$
y_j	$j \in N$:	$Decision\ variable\ to\ serve\ a\ neighbourhood$
a_{ij}	$i \in L, j \in N$:	$Parameter\ describing\ if\ a\ neighbourhood\ is\ served\ by\ a\ depot$
P_j	$j \in N$:	$Parameter\ for\ neighbourhood\ population$
D_j	$j \in N$:	$Parameter\ for\ a\ neighbourhood's\ expected\ demand\ (2\ months)$

$$maximise \sum_{j \in N} 20 y_j D_j - 2000 \sum_{i \in L} x_i$$

$$subject\ to$$

$$2000 \sum_{i \in L} x_i \leq 6000 \qquad \forall i \in L$$

$$D_j = P_j \frac{\log_{10} 3}{10} \qquad \forall j \in N$$

$$x_i \in \{0,1\} \qquad \forall i \in L$$

$$y_j \in \{0,1\} \qquad \forall j \in N$$

$$a_{ij} x_i \leq y_j \qquad i \in L, j \in N$$

$$y_j \leq \textstyle\sum_{i \in L} a_{ij} x_i \qquad \forall j \in N$$

<u>Discussion:</u>

$$maximise \sum_{j \in N} 20 y_j D_j - 2000 \sum_{i \in L} x_i$$

The objective function that seeks to maximise the profit made by selecting a configuration of depots at month 0. Profit is calculated as (expected revenue) – (initial set-up cost). If a neighbourhood is serviced by a depot ($y_j = 1$), its expected demand, D_j will be realised. Expected revenue set at £20 per order. The operating cost is set at £2000 for each depot established, x_i.

$$2000 \sum_{i \in L} x_i \leq 6000$$

Describes the cost restraint for the amount of money available for establishing a depot at month 0. Total budget set to £6000.

$$D_j = P_j \frac{\log_{10} 3}{10} \qquad \forall j \in N$$

Parameter for the expected demand, D_j for of each neighbourhood, as a function of its population, P_j.

$$x_i \in \{0,1\} \quad \forall i \in L$$

This constraint sets $x_i = 1$ for a depot to be established at a location, otherwise 0.

$$y_j \in \{0,1\} \quad \forall j \in N$$

This constraint sets $y_j = 1$ if a neighbourhood is serviced by a depot within range, otherwise 0.

$$a_{ij}x_i \leq y_j \quad i \in L, j \in N$$

This constraint ensures that if a neighbourhood is within range of an established depot, it must be set 1. The set of depot locations, L are at the centroids of the neighbourhoods within set N, and service that neighbourhood. A depot in an adjacent neighbourhood is also within range if the centroids are within $0.5km$.

$$y_j \leq \sum_{i \in L} a_{ij}x_i \quad \forall j \in N$$

This constraint ensures that if a neighbourhood does not have any established depots within range, it must be set 0.

3.2 PuLP model

The report uses the mathematical model formed in the previous section to identify the three warehouse locations that will earn the most profits. A PuLP problem was formed to represent the mathematical model. Each parameter is modelled as:

- L: 'centroid_prob', list of all possible warehouse locations
- N: 'neighbourhood_prob', list of all neighbourhoods
- a: 'cover_dict', dictionary describing the neighbourhood coverage of each warehouse location
- x: Integer LpVariable varying from 0 to 1 therefore modelling the decision variable $\in \{0,1\}$
- y: Integer LpVariable varying from 0 to 1 therefore modelling the decision variable $\in \{0,1\}$
- P: 'pop_dict', dictionary describing the population of each neighbourhood
- D: Calculated at the optimization equation using the population dictionary, for coding efficiency

Using these parameters, the PuLP optimization equation and subject equations were coded according to the mathematic formulation shown in Task 2.1. The list of selected warehouses, profit obtained from the warehouses and the warehouse positions are plotted in Figure 3.1. By solving the mathematical model using 'prob.solve()', the following optimal solution is found. The three selected stations are:

- Hammersmith and Fulham 015C, covering 16,889 residents
- Hammersmith and Fulham 023A, covering 16,509 residents
- Hammersmith and Fulham 002A, covering 15,578 residents

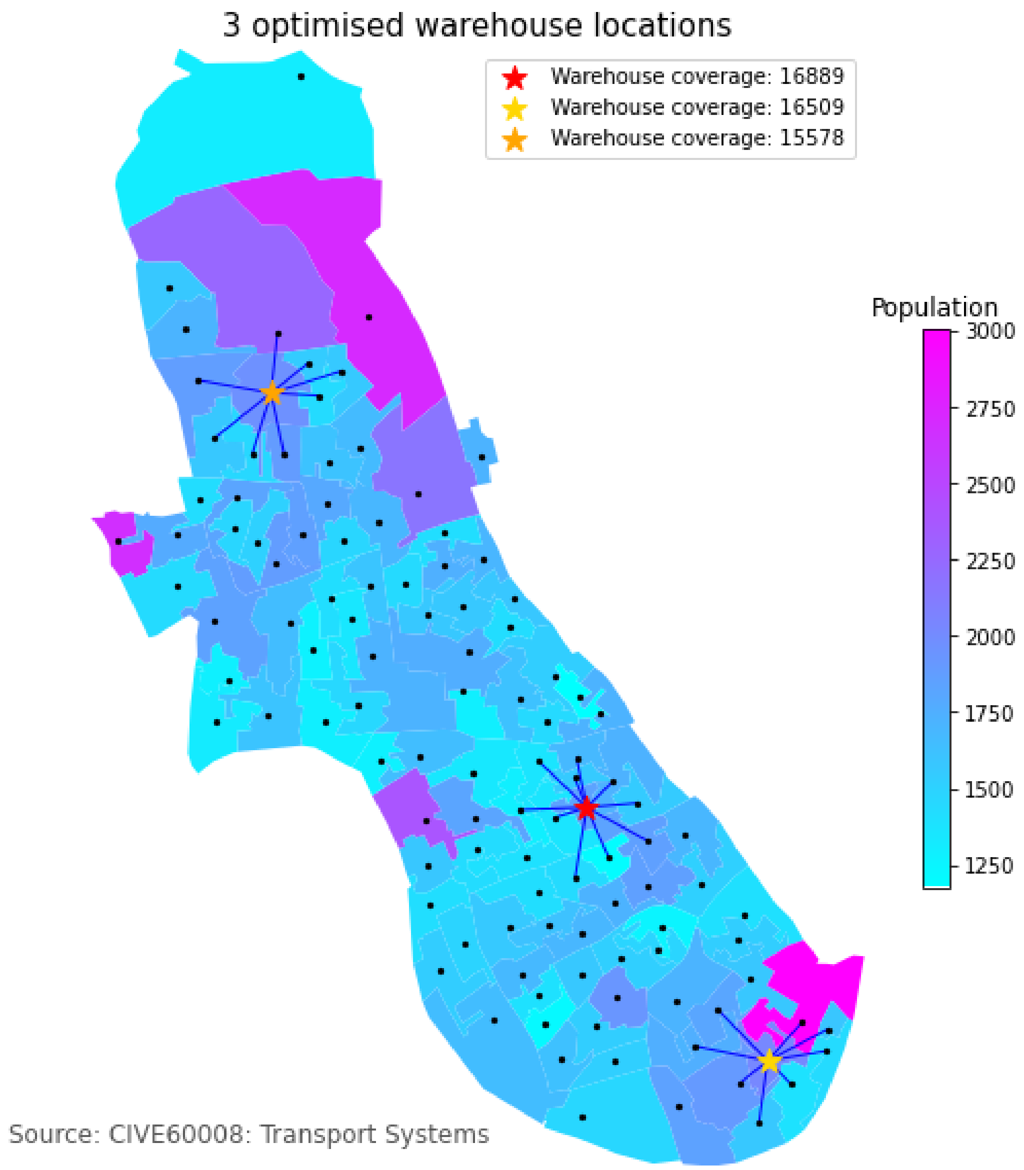

Figure 3.1: Optimal warehouse locations and coverage area of each warehouse

The solution includes three of the 5 depot locations with the highest coverage found in Task 1.3. However, the warehouse location with the 5th highest coverage is chosen over the warehouses with the 3rd and 4th highest coverage. This can be explained by them having a high percentage of common coverage area with the first location chosen (which has the most population coverage). Therefore, it is supported that the subject equation: $y_j \leq \sum_{i \in L} a_{ij} x_i, \forall j \in N$ (preventing depots with common coverage area to be chosen), was coded efficiently. It can be concluded that the combination of the three depots shown in Figure 3.1 would earn the maximum profit, since they are the three depot locations with the highest coverage population (therefore demand) and no overlapping area between them.

4.0 Expansion Modelling

4.1 Mathematical model for the facility expansion plan

The model used in Section 3.2 is repeated through multiple iterations to obtain an expansion plan for the next 12 months. The initiation of a new warehouse can only occur every two months and the capacity for expansion is calculated at the month of expansion. Based on the brief, the demand for warehouses will vary based on the duration of expansion, not the duration of operation of the warehouses. Therefore, the mathematical model is modified to account for the budget from the previous iteration and new variables are created to store or carry forward values to the next iteration. They are specified as follows:

month_trial:	Month of expansion (0, 2, 4 …… ,10)
initial_budget:	To store the initial budget for expansion every two months
savings:	To store the money not used for expansion every two months
exist_facility:	To store the number of new warehouses
number_facility:	To carry forward the cumulative increase in number of warehouses
population_covered_in_phase:	To store the population covered at each phase of expansion
expansion_percentage:	To store the percentage covered by each phase of expansion
expansion_status:	To store the coverage status of each neighbourhood (used for plotting)
left_over_money:	To store the money allocated but not used in expansion
carry_forward:	To carry forward the money allocated but not used in expansion at each iteration
exist_profit:	To store the profit from existing warehouses

<u>Discussion on mathematical modelling:</u>

- The objective function is modified to account for the increased profit and maintenance cost of warehouses from the previous iterations. Given *E* is the number of stations already opened, the new objective function is:

$$maximise \sum_{i \in L, j \in N}^{t=n} 20 y_j D_j - 2000 x_i - 2000\, E_n \quad \forall\, n \in 2N$$

$$where\ D_j = P_j \frac{\log_{10}(3+3t)}{10} \quad \forall\, j \in N$$

- The population dictionary, *P*, is also modified so the maximisation algorithm only accounts for neighbourhoods without coverage at each iteration. This is done by changing the population of the neighbourhood to zero if a warehouse is already available. The demand associated with the already opened warehouses is calculated separately, under the variable name *exist_profit*. The profit is then added back into the objective function to give the same effect as the above-mentioned equation for D_j. This is done such that each iteration only identifies new stations and improves running efficiency. This also eases the second analysis in this section where demand is dependent on the duration of warehouse operation.

- The initial budget available for expansion is £6000. Where *L* is the leftover money from the previous expansion, it is carried forward to the next iteration. Every subsequent iteration has an initial budget of:

$$L_{n-1} + 0.15 \times max \sum_{i \in L, j \in N}^{t=n-1} 20 y_j D_j - 2000 x_i - 2000\, E_{n-1} \quad \forall\, n \in 2N$$

The expansion plan produced, if demand varied with duration of expansion, is attached in Table 4.1, Table 4.2, and Figure 4.1. From the analysis, one can observe that the entire borough is covered by the expansion in month 6 and the there exists a total of 24 warehouses. The results are validated by comparing the cost of expansion (£2000 multiplied with the number of new warehouses) and the available budget – there exists sufficient funds to cover the costs at each iteration. Besides that, the map is analysed to ensure that the overlaps occur at a later stage of the expansion, to validate that the maximisation objective function. The results are also compared with Section 3.2 to ensure that the first three optimised warehouses are the same.

Despite obtaining valid and reasonable results, to correlate expansion duration with demand is a severe overestimate of the potential profit from the borough. The demand (and revenue) associated with the warehouses should consider the operation duration of individual warehouses instead (warehouses opened in month 4 would not generate the same profit as warehouses operating for 4 months). Therefore, a second analysis is supplemented to account for the overestimation. In this version of the analysis, the *exist_profit* variable is altered to account for the operation duration of each individual warehouse and the results are attached as Table 4.3, Table 4.4 and Figure 4.2. A key observation is that not the entire borough is covered. This is attributed to the fact that the demand generated by the uncovered population is not sufficient to cover the expansion cost. Therefore, these boroughs are left without coverage instead. In contrast to the 24 warehouses in the previous analysis, 19 warehouses are opened in this analysis. Lesser or the same number of warehouses are opened at each iteration in comparison to the previous analysis.

In the first two months, enough profit is made to create a budget for expansion of £6110, corresponding to 15% of the total profit. There is no leftover budget from the initial warehouse expansion plan so this is the total budget available. As explained before, the population of the nodes that are covered are set to 0. Using this, the next 3 warehouses to be located are chosen based on the maximum profit made. From the figure showing the expansion plan for month 2, the 3 stations are successful in covering 23 more nodes and a larger area of the borough. There is only one overlapping node compared to the initial expansion plan. From the budget given, a leftover budget of £110 is achieved and can be used for the next expansion plan. In the next 2 months a total budget of £18309 is available. To maximise profit in this case, 9 new depots are built, shown in the expansion plan for month 4. The individual warehouses will be connected to less nodes, but the area of the borough covered will be increased and 40 nodes will be connected. In month 6, 4 new depots are created. Two of these depots only cover one node each. Despite this, the plan is still maximising profit, and these warehouses will still make a profit. After this, only 5 nodes are not connected to the warehouse system in the borough. As shown in the expansion plan for months 8 and 10, no profitable depots can be created. The resulting nodes are spaced away from each other, and new warehouses will need to be built at the nodes to serve the residents. Despite a high budget, none of the nodes are worth upgrading when considering the profit. This leaves a large, disconnected area at the North of the borough but as shown in Figure 2.3, the population density of this area is low.

Month	Location of new warehouse
0	• Hammersmith and Fulham 015C • Hammersmith and Fulham 023A • Hammersmith and Fulham 002A
2	• Hammersmith and Fulham 010A • Hammersmith and Fulham 009A • Hammersmith and Fulham 019B
4	• Hammersmith and Fulham 008B • Hammersmith and Fulham 012A • Hammersmith and Fulham 001B • Hammersmith and Fulham 017A • Hammersmith and Fulham 016C • Hammersmith and Fulham 013A • Hammersmith and Fulham 018E • Hammersmith and Fulham 025A • Hammersmith and Fulham 011B • Hammersmith and Fulham 004A • Hammersmith and Fulham 004D • Hammersmith and Fulham 024D
6	• Hammersmith and Fulham 006B • Hammersmith and Fulham 012B • Hammersmith and Fulham 001A • Hammersmith and Fulham 001D • Hammersmith and Fulham 020D • Hammersmith and Fulham 021F
8	No warehouses to open
10	No warehouses to open

Table 4.1: Expansion plan if demand varies with expansion duration

	Expansion month	Budget available for current expansion (£)	Budget not used for expansion (£)	Number of new facilities	Percentage of population covered (%)	Leftover budget for next expansion (£)
0	0	6000.000000	0.000000	3	26.683520	0.000000
1	2	6110.247171	34624.733971	3	49.441006	110.247171
2	4	24288.354400	137009.274295	12	94.612191	288.354400
3	6	56158.592677	316598.016905	6	100.000000	44158.592677
4	8	109764.218147	371765.210995	0	100.000000	109764.218147
5	10	181379.687366	405820.992239	0	100.000000	181379.687366

Table 4.2: Descriptive information of expansion plan if demand varies with expansion duration

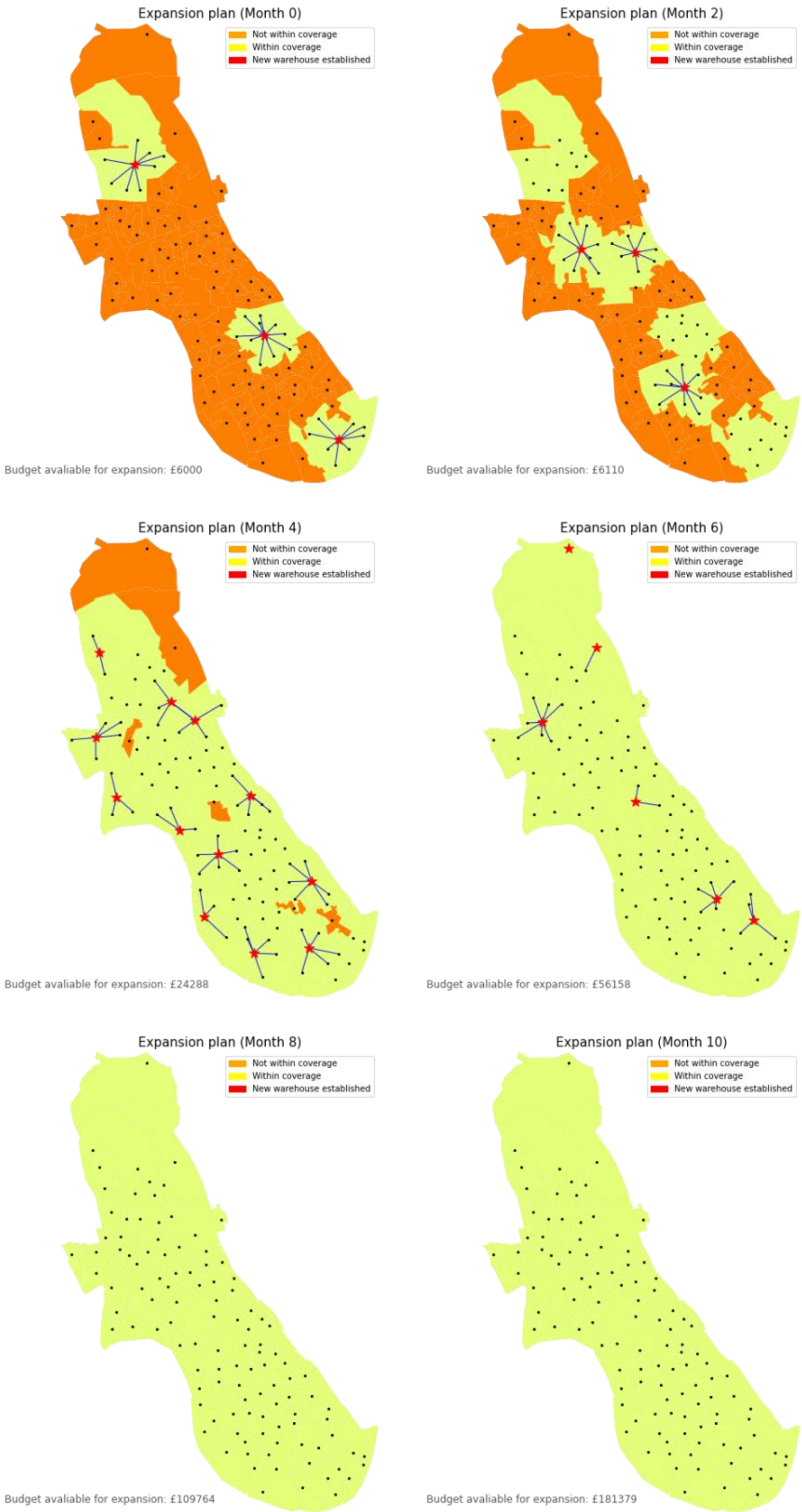

Figure 4.1: Expansion plan if expansion varies with expansion duration

Month	Location of new warehouse
0	• Hammersmith and Fulham 015C • Hammersmith and Fulham 023A • Hammersmith and Fulham 002A
2	• Hammersmith and Fulham 010A • Hammersmith and Fulham 009A • Hammersmith and Fulham 019B
4	• Hammersmith and Fulham 008B • Hammersmith and Fulham 012A • Hammersmith and Fulham 017A • Hammersmith and Fulham 016C • Hammersmith and Fulham 013A • Hammersmith and Fulham 018E • Hammersmith and Fulham 011B • Hammersmith and Fulham 004A • Hammersmith and Fulham 024D
6	• Hammersmith and Fulham 001A • Hammersmith and Fulham 001E • Hammersmith and Fulham 025A • Hammersmith and Fulham 004D
8	No warehouses to open
10	No warehouses to open

Table 4.3: Expansion plan if demand varies with warehouse operation duration

	Expansion month	Budget available for current expansion (£)	Budget not used for expansion (£)	Number of new facilities	Percentage of population covered (%)	Leftover budget for next expansion (£)
0	0	6000.000000	0.000000	3	26.683520	0.000000
1	2	6110.247171	34624.733971	3	49.441006	110.247171
2	4	18309.547957	103129.371118	9	89.018982	309.547957
3	6	35445.090286	199101.406530	4	96.088676	27445.090286
4	8	78562.842711	289667.263743	0	96.088676	78562.842711
5	10	139807.428648	347052.653641	0	96.088676	139807.428648

Table 4.4: Descriptive information of expansion plan if demand varies with warehouse operation duration

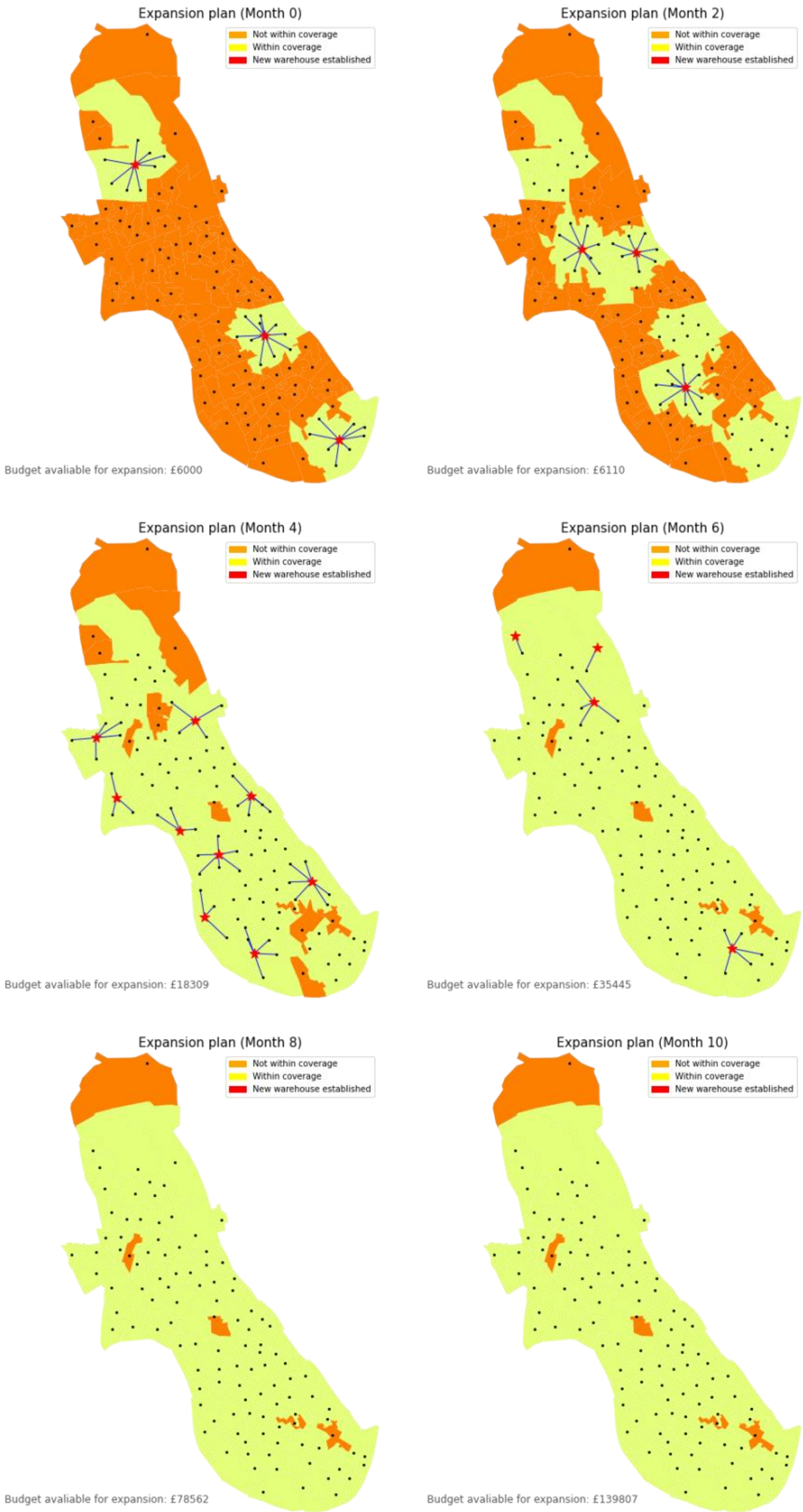

Figure 4.2: Expansion plan if expansion varies with warehouse operation duration

4.2 Other boroughs with higher end-of-year sales

To check if Hammersmith and Fulham are the best places to start an expansion, all the calculations in the previous section are repeated for each borough. The warehouse-dependent method is modelled instead for higher accuracy. The money left at the end of the trial period (assuming no more expansions occur) is measured and each borough is ranked. This is illustrated in Table 4.2. The rankings show that there are 11 other boroughs that are more profitable in comparison to Hammersmith and Fulham. However, Hammersmith and Fulham require the minimum number of warehouses.

	Borough	Total leftover budget by the end of the trial	Number of new warehouses
29	Tower Hamlets	1.938839e+06	30
24	Newham	1.790166e+06	35
32	Westminster	1.716041e+06	25
6	Camden	1.566570e+06	28
27	Southwark	1.551771e+06	34
11	Hackney	1.547277e+06	27
18	Islington	1.475953e+06	23
21	Lambeth	1.474002e+06	36
31	Wandsworth	1.300002e+06	40
30	Waltham Forest	1.186182e+06	31
13	Haringey	1.136588e+06	34
12	Hammersmith and Fulham	1.113383e+06	19

Table 4.2: 12 boroughs with the highest leftover budget in London

5.0 Conclusion

Based on the analysis above, it can be concluded that advanced analysis is required in managing a covering set problem. Whilst targeting the most populated neighbourhood is essential, an expansion plan must account for the highest coverage in the shortest amount of time. Besides that, more research needs to be conducted to validate whether the demand of a neighbourhood varies with the expansion duration or the warehouse operation duration. Nonetheless, the first six warehouses are the same in both scenarios and the variation in duration only impacts the possibility of expanding across the entire borough – some neighbourhoods do not have sufficient population to cover the initial start-up cost.