STATE UNIVERSITY OF CAMPINAS
FACULTY OF APPLIED SCIENCES

MULTI-WAREHOUSE VEHICLE ROUTING MAGAZINE LUIZA

Camila Delfini 232671
c232671@dac.unicamp.br

Eric Aparecido Moreira do Nascimento 167281
e167281@dac.unicamp.br

Filippo Tranchesi de Plato 173571
f173571@dac.unicamp.br

Flávio Campedelli Neto 234559
f234559@dac.unicamp.br

Guilherme Santos Conforto Melo 171145
g171145@dac.unicamp.br

Lucas Tramonte 182697
l182697@dac.unicamp.br

SUMMARY

The following paper presents a problem that uses data on the locations of 15 stores and 2 distribution centers of the company Magazine Luiza, in which we seek, with the help of Operations Research, to find effective solutions for vehicle routing from multiple depots (MDVRP). The main objective is to determine the best routes that minimize the total distance traveled to meet the demands of the stores, using the *PuLP* linear programming package, implemented via the *Python* computer language.

In this context, the computer tests were analyzed based on changes to the number of vehicles used, the maximum number of establishments that can be visited by each driver, and aspects linked to the mathematical formulation. The demand for each destination and the capacity of the trucks were kept fixed. Analysis of the results obtained is then provided, with the aim of discussing both the drivers' and the company's perspectives, proposing reflection on different aspects such as the total distance of the routes, travel time and the average distribution of mileage between vehicles. From the company's point of view, the solution that stands out was characterized by the use of 2 vehicles and did not have the restriction that limits the number of points that can be visited.

1. INTRODUCTION

The current economic context is marked by technological innovation and globalization. In this sense, logistics is proving to be a field of study and a tool capable of increasing business competitiveness. In order to generate better operational performance and, consequently, a competitive strategic advantage, it is necessary to quantify operations and measure their financial impact. To do this, it is prudent to analyze the transport operation in terms of logistics costs and the level of service offered by the organization.

Over time, the use of modeling and operational research through linear programming has gained prominence in scientific circles. Operational research is the application of scientific methods to complex problems to assist in the decision-making process, such as designing, planning and operating systems in situations that require efficient allocation of scarce resources (ARENALES, 2007). In short, operational research helps to optimize processes in critical situations of analysis and study, seeking competitive strategic planning.

Vehicle routing problems cover a broad category of operations research problems. They involve the design of minimum-cost delivery and/or collection routes from one or more depots to a given customer, subject to additional constraints such as a time window or multiple depots. In this respect, the problem of vehicle routing is fundamental in the area of distribution and logistics management.

In an attempt to maximize the positive strategic factors for the future growth of the Magazine Luiza company, we sought to implement operational research tools, encompassed within the activities of routing vehicles with multiple depots. This procedure was carried out in order to find possible routes for the delivery of goods and, subsequently, to find ways of reducing transportation costs. The aim of the research is to optimize Magazine Luiza's routes in order to make better use of transport resources, identifying the best routes to reduce the total distance travelled.

2. MATHEMATICAL MODELING

The problem that will be addressed is the *Multi-Depot* Vehicle Routing *Problem* (MDVRP). In this case, the MDVRP is an extension of *The Vehicle* Routing *Problem* (VRP). By definition, the VRP is a problem that consists of defining routes for a set of vehicles parked at a central depot that will serve a set of customers, minimizing transportation costs. Now, the MDVRP contains multiple depots, which makes it highly complex to carry out the mathematical modeling of the problem, as well as the construction of its computational solution for the optimal routes, among the countless possibilities for combining destinations and origins.

The mathematical modeling used is based on the formulation described in the article "Integer Programming Formulations of Vehicle Routing Problems" (KULKARNI, 2011). From this perspective, the problem to be solved can be modeled as:

$$minZ = \sum_{i=1}^{N+M} \sum_{j=1}^{N+M} c_{ij} x_{ij} \quad , \tag{1}$$

subject to

$$\sum_{i=1}^{N+M} x_{ij} = 1 \ para\ j = 1, 2, \ldots, N, \tag{2}$$

$$\sum_{j=1}^{N+M} {}_{ij} = 1 \ para\ i = 1, 2, \ldots, N, \tag{3}$$

$$\sum_{i=N+1}^{N+M} \sum_{j=1}^{N+M} {}_{ij} = V, \tag{4}$$

$$\sum_{j=N+1}^{N+M} \sum_{i=1}^{N+M} {}_{ij} = V, \tag{5}$$

$$x_{ij} = 0 \ ou\ 1 \ para\ todo\ i, j, \tag{6}$$

$$u_i - u_j + Px_{ij} \leq P - Q_i \ para\ 1 \leq i \neq j \leq N, \tag{7}$$

$$y_i - y_j + Lx_{ij} \leq L - 1 \ para\ 1 \leq i \neq j \leq N, \tag{8}$$

where the objective function Z seeks to find routes between N destinations and M origins that minimize the total cost for the carriers, so that c_{ij} is the distance from city i for city j. We define V as the number of vehicles, P as the vehicle capacity, Q_i as the demand for destination *i* and L as the maximum number of cities a vehicle can visit. It should be noted that the vehicle's capacity must be greater than the demand for a destination, so that all demands can be met. In this function, the programmed code generates the entire combinatorics of variables for the sum from city i to city j and the subsequent equations become constraints in search of a solution that meets the needs of the problem.

Therefore, the restrictions in equations (2) and (3) mean that each destination city is only served once, and that all cities must be visited. In equations (4) and (5), it is assumed that all vehicles V are used to transport the product in question.

The variable x_{ij} is binary and takes on a value of 1 or 0, according to equation (6). As it is not part of of the problem a route that lists the same city as the point of departure and arrival, an additional constraint was assigned to the problem[1] :

$$x_{ij} = 0 \ , para\ todo\ i = j. \tag{9}$$

In addition, it is worth noting that more restrictions have been added to the model in order to meet the condition that a vehicle cannot leave one depot and go directly to another:

$$x_{ij} = 0 \ , para\ i = N + 1, N + 2, \ldots, N + M\ e\ j = N + 1, N + 2, \ldots, N + M. \tag{10}$$

[1] The constraints added are optional, since there is the possibility of not generating these variables in the implementation (which results in a computational gain, as there are fewer decision variables), or of assigning an infinite cost to them, so that the model doesn't choose them. However, as the computational cost was not high in this problem, these variables were set to zero due to the ease of implementing this option.

Constraints (7) and (8) have a dual role in the MDVRP model, they are capacity constraints and subroute elimination constraints[2] . They belong to a class of *Miller-Tucker-Zemlin* (MTZ) arc product flow constraints, which appear in polynomial quantity in the model, which is an advantage since they differ from typical exponentially growing subroute elimination constraints with respect to the input data. However, as a disadvantage, they produce a weak linear relaxation in the resolution of the relaxed model during the *branch-and-bound* process. The variables u and y must be non-negative, are assigned a value at each vertex except the depots, and do not allow disconnected sub-routes to occur in the optimal solution, since they define a certain order in which the flow of cargo in the vehicle must traverse the arcs along the customer visits (MILLER, 1960).

Finally, with the use of constraints (4) and (5), if it is established, for example, that a total of 2 vehicles must be used, there is the possibility of the model assigning no vehicles to one depot, and 2 vehicles to the other. Therefore, in order to carry out computational experiments, constraints were considered that do not allow this situation to occur, defining a fixed number of vehicles for each depot[3] :

$$\sum_{j=1}^{N+M} x_{ij} = V_i \text{, } para\ i = N + 1, N + 2, \ldots, N + M. \tag{11}$$

$$\sum_{i=1}^{N+M} x_{ij} = V_j \text{, } para\ j = N + 1, N + 2, \ldots, N + M. \tag{12}$$

3. COMPUTATIONAL EXPERIMENTS

After defining Magazine Luiza as the base company, it was necessary to choose the physical stores (destinations) and distribution centers to make up the problem. These establishments were collected manually using the Google search engine. 15 stores and 2 centers were selected, spread across the states of São Paulo, southern Minas Gerais, northern Paraná and western Rio de Janeiro. Two separate files were then built in Google Spreadsheets, the first containing the full addresses of the 2 distribution centers, and the other with the addresses of the 15 customers. From this, a code was created using the Google Apps Script extension, which provides, in a neighboring column, the latitude and longitude of any address entered in a given column of a spreadsheet. In this way, geographical coordinate data was obtained for all the addresses involved in the multi-purpose vehicle routing problem.

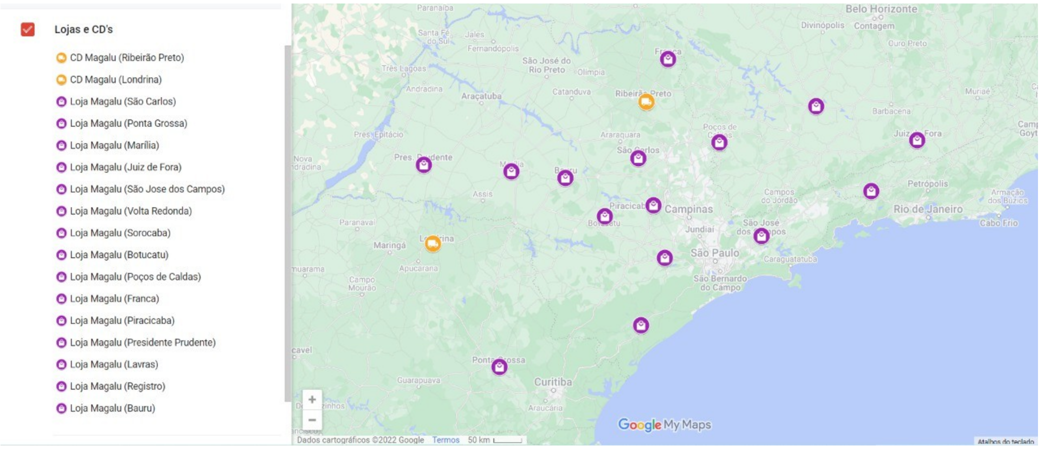
As far as the *Python* programming language is concerned, the Pandas library offers structure and tools for processing, organizing and displaying data sets of different proportions. Using this library, we read the two aforementioned Google Spreadsheet files - in .xlsx format - which stored the address, latitude and longitude information of the distribution centers and customers. Once the locations related to Magazine Luiza had been collected, the MyMaps program was used, a Google service aimed at creating and customizing maps, in order to obtain a visual-geographical overview of the locations.

[2] In order to use equations (7) and (8) properly, it is necessary to define the desired analysis. Equation (7) should be used when the limiting factor for a vehicle is the ratio between its capacity and the sum of the demands of the cities it will visit. However, if the desired limiting factor is the maximum number of cities that each vehicle can serve, the use of equation (8) is the most appropriate option.

[3] Equations (11) and (12), if used, must replace equations (4) and (5).

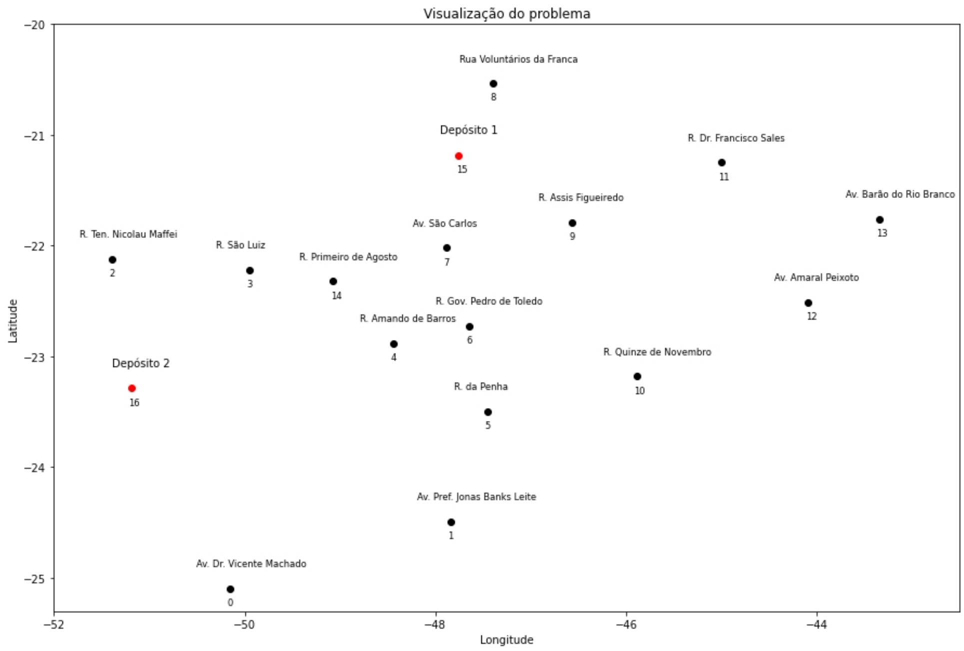
chosen points, as shown in Figure 1. The distribution of these points, along with their indices, was also generated using *Python* code, as shown in Figure 2.

Figure 1 - Geographical cross-section of the points chosen for the routing problem



Source: Own authorship.

Figure 2 - Schematic distribution of the points chosen for the routing problem

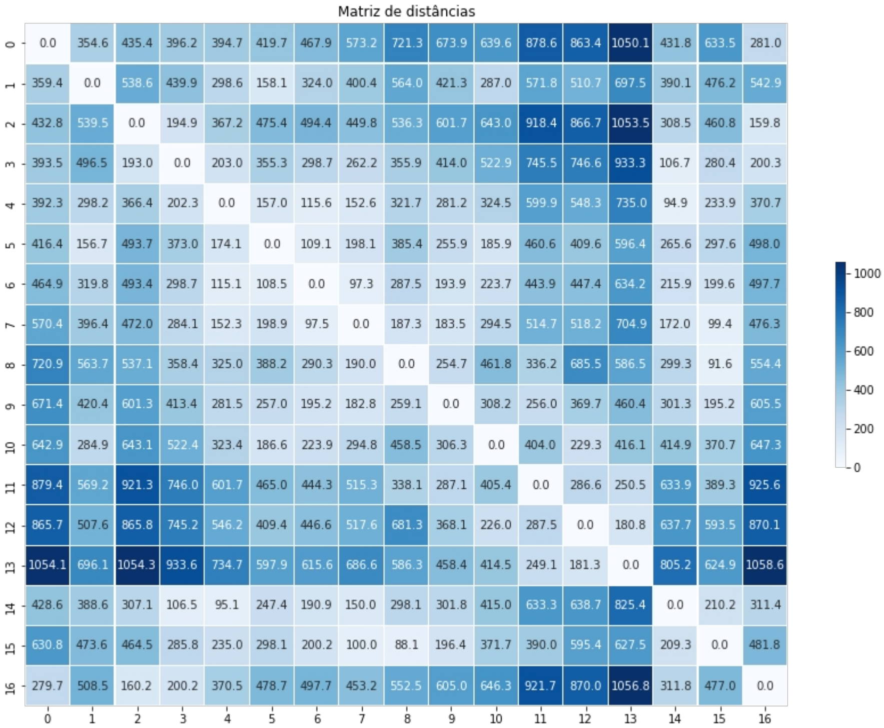


Source: Own authorship.

After generating these maps, we calculated the distances between the two warehouses and each of the stores, as well as between the stores themselves. We did this using Bing's geolocation API (*Application Programming Interface*). This tool makes it possible to locate devices anywhere on the earth's surface with great precision and detail. The API was used to precisely find the distance between all possible pairs of points in the problem in question, in an extremely efficient manner.

Thus, a distance matrix was generated using *Python* with the information provided by the API (data in kilometers), as shown in Figure 3. It should be noted that this matrix is classified as non-Euclidean, i.e. the matrix is not symmetrical (the x values$_{i,j}$ are not necessarily the same as the x values$_{j,i}$). This is due to the fact that the use of the geolocation API guarantees accuracy that is so faithful to reality that the path and journey time between one destination and another does not necessarily have the same distance and return time. Factors such as road direction, alternative routes, traffic lights and even real-time traffic are taken into account in this process.

Figure 3 - Distance matrix



Source: Own authorship.

In order to complete the basic information of the problem, it was established that the number of vehicles for transportation will be varied between 2 and 5, seeking to analyze how this variable alters the total distance traveled, as well as the total travel time. Each vehicle has a homogeneous theoretical capacity to meet 20 demands[4] , and it was defined that each demand is equivalent to 4 cardboard boxes of 120L (60cmx40cmx50cm), which are loaded with various products to supply the stores' stock.

Once we knew the distances between all the locations defined for the routing problem, the vehicle capacities and the demands of the destinations, we ran the code in *Python* to determine the routes that minimize the total distance travelled, using the open-source linear programming package *PuLP*, which allows us to implement the variables, the objective function and the constraints of the problem to determine the optimal solution.

Variations were made to the number of vehicles in order to see what the benefits of using a larger fleet would be, especially in view of the average time dedicated by each driver to completing their respective routes. Thus, the problem was initially approached from 3 perspectives: meeting the demands of destinations with 2, 3 and 4 vehicles. In addition, in order to obtain an estimate of the total travel time, it was

[4] It was decided to keep the values for demand and vehicle capacity fixed, due to the high sensitivity of these parameters in the model. The values determined for demand were first drawn from a range of 1 to 5, and after successive tests, the values that provided the best responses to the model were chosen, which can be found in the code presented in Appendix A. It is also worth noting that complementary sensitivity analyses were carried out at the end of this same code.

A hypothetical average speed of 70 km/h was established for all possible vehicles and routes. Table 1 shows the results obtained.

Table 1 - Comparison of results considering 2, 3 and 4 vehicles.

No. of vehicles	Objective Function (km)	Travel time (h)	Distance per vehicle (km)
2	3190,070	45,572	1595,035
3	3223,344	46,048	1074,448
4	3386,440	48,378	846,610

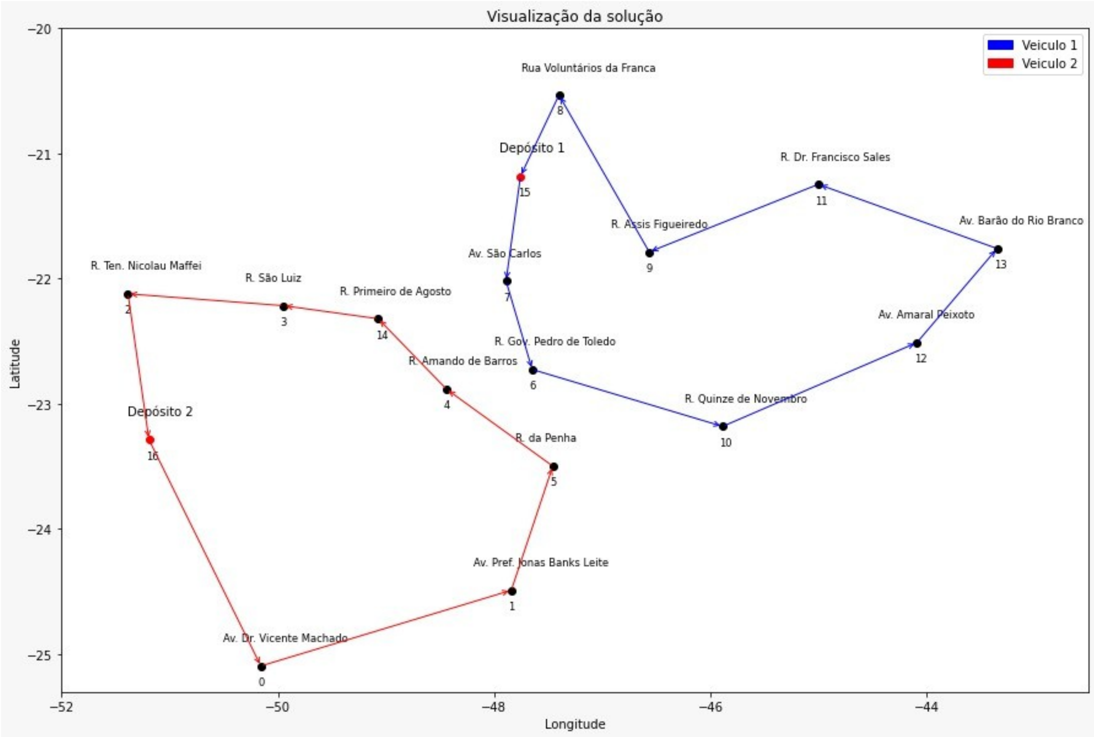
Source: Own authorship.

The table shows that, in the cases explored, a greater number of trucks represents a slight increase in total travel time, considering the sum of the time spent by each of the vehicles. In practice, however, the time is less than that shown, since the journeys are made simultaneously. In addition, the average distance covered by each driver is lower the more vehicles are used, so this is a way of prioritizing worker well-being and contributing to ergonomics at work. On the other hand, the case of 2 trucks, despite having a higher ratio between total distance traveled and number of vehicles, represents a more balanced solution compared to the other cases, due to the similarity of the mileages traveled by each driver, as can be seen in Figures 4, 5 and 6.

Before analyzing the routes obtained by the model according to the variations in the number of vehicles, it should be clarified that, due to a feature of the *Python* language, the indexing of the terms *i* and *j* (referring to the points of departure and arrival of a trip, respectively), which define the variables x_{ij} of the problem, start at 0. Thus, the 15 destination cities correspond to the indexing from 0 to 14, while the cities of the distribution centers correspond to the indexes 15 and 16.

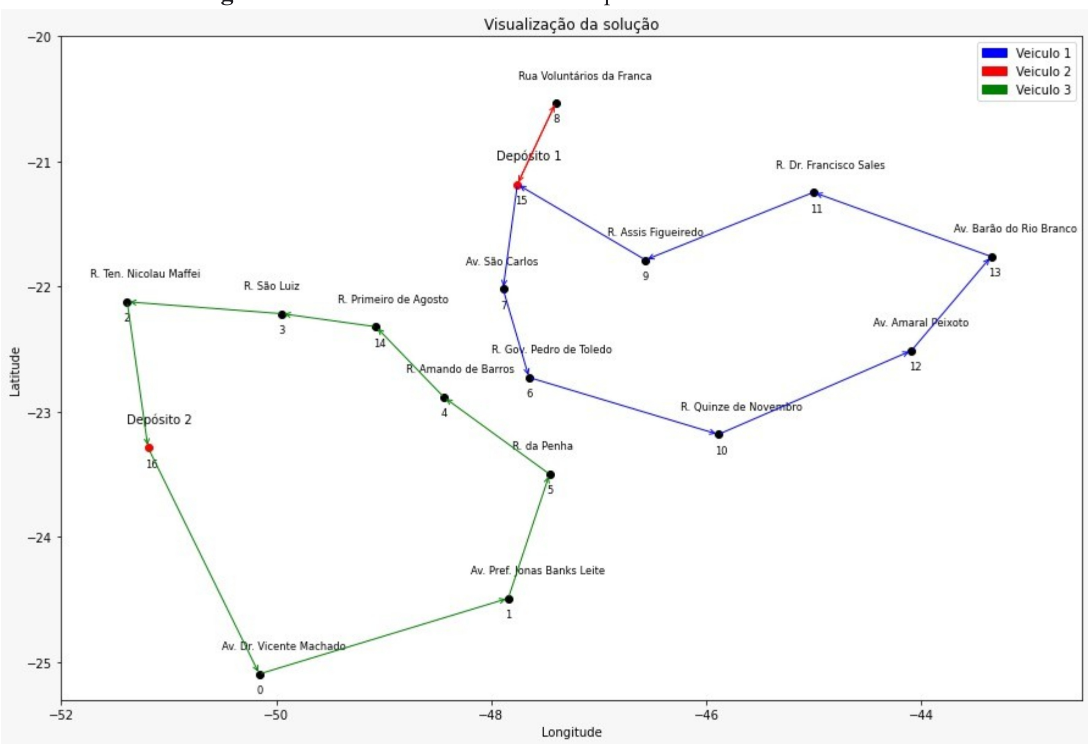
In all 3 cases, the constraints of equations (4) and (5) are used in the modeling, so the model itself chooses the optimal distribution of vehicle departure points, in order to minimize the sum of the route distances. In the case of 2 vehicles, the model provides a homogeneous initial distribution, i.e. 1 truck leaves from each depot. When 3 vehicles are used to meet the demands of the destination cities, the model allocates 2 of them to depart from distribution center 1, while the other departs from center 2. Finally, the case of 4 vehicles also results in a uniform initial distribution, with 2 trucks departing from each depot. The ordered pairs that define the best route found by the model have the following configuration: (departure, arrival). Figures 4, 5 and 6 illustrate schematic maps with the optimal routes - in different colors according to the route of each truck - for the cases with 2, 3 and 4 vehicles, respectively.

Figure 4 - Schematic distribution of optimal routes for 2 vehicles



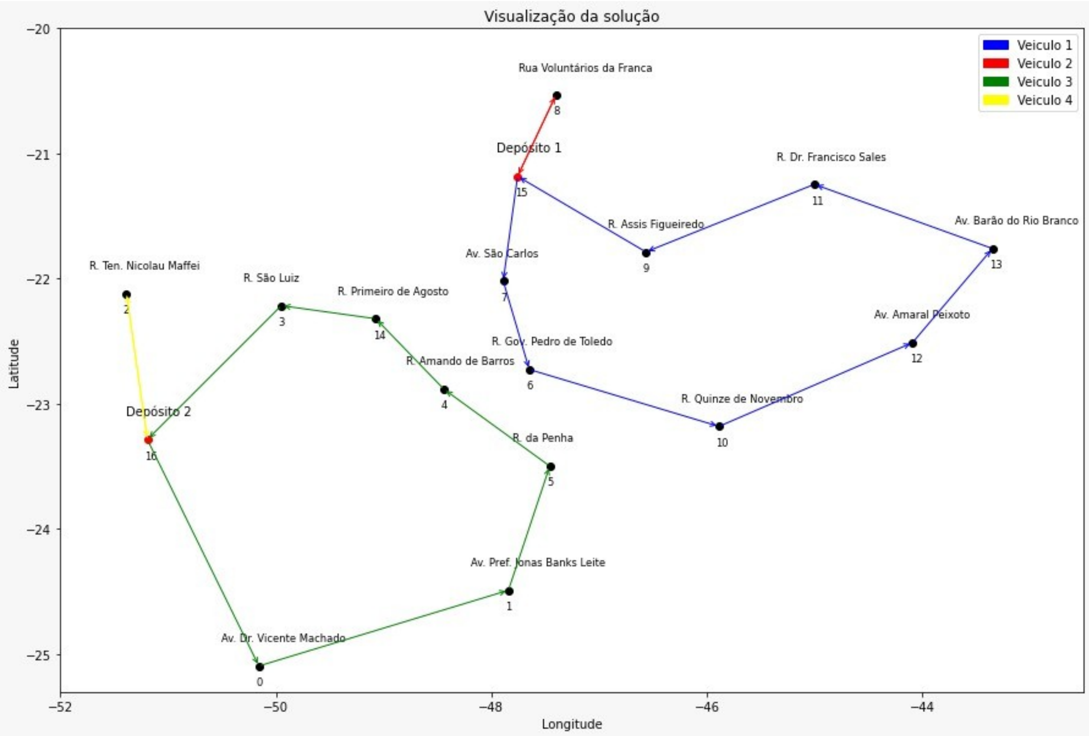
Caption: Starting from warehouse 1: [(15, 7), (7, 6), (6, 10), (10, 12), (12, 13), (13, 11), (11, 9), (9, 8), (8, 15)].
Starting from warehouse 2: [(16, 0), (0, 1), (1, 5), (5, 4), (4, 14), (14, 3), (3, 2), (2, 16)].
Source: Own authorship.

Figure 5 - Schematic distribution of optimal routes for 3 vehicles



Caption: Starting from warehouse 1: [(15, 7), (7, 6), (6, 10), (10, 12), (12, 13), (13, 11), (11, 9), (9, 15)] and [(15, 8),
(8, 15)]. From warehouse 2: [(16, 2), (2, 3), (3, 14), (14, 4), (4, 5), (5, 1), (1, 0), (0, 16)].
Source: Own authorship.

Figure 6 - Schematic distribution of optimal routes for 4 vehicles



Caption: Starting from warehouse 1: [(15, 7), (7, 6), (6, 10), (10, 12), (12, 13), (13, 11), (11, 9), (9, 15)] and [(15, 8),
(8, 15)]. From warehouse 2: [(16, 0), (0, 1), (1, 5), (5, 4), (4, 14), (14, 3), (3, 16)] and [(16, 2), (2, 16)].
Source: Own authorship.

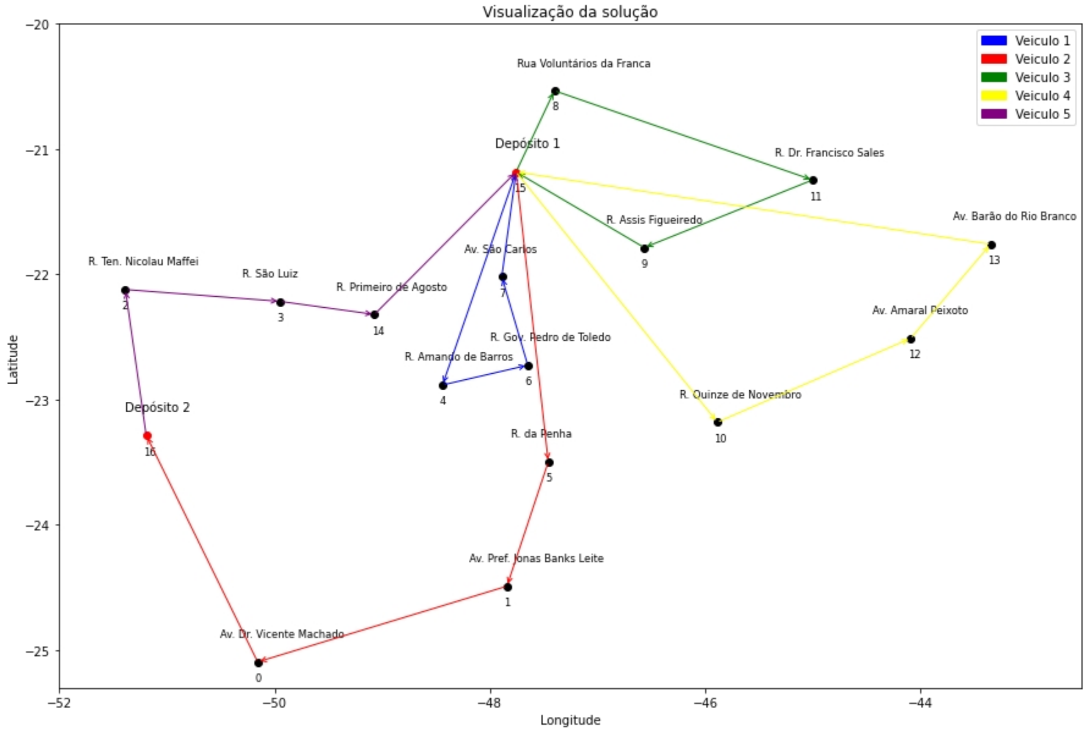
4. ANALYSIS OF RESULTS

After exploring the most comprehensive optimization cases of the problem, which do not consider, for example, restrictions on the maximum number of locations that can be visited by the same vehicle, it was decided to analyze this new scenario. By varying the number of vehicles in these cases, we sought optimal solutions that were more faithful to the reality of the problem, which provide a better balance between the number of destination cities served by each driver. In order to observe the model's new behavior under the desired conditions, the constraint of equation (8) was incorporated, whose contribution is to establish a maximum limit L of points (nodes) that can be visited by each vehicle, in place of equation (7).

From the workers' point of view, the solutions proposed below are much more satisfactory, since they now travel routes with significantly closer mileages. It should also be noted that, in the optimal solution of one of the cases, the departure and arrival depots of some routes do not coincide, i.e. the model provides routes that start at distribution center 1 and end at distribution center 2, and vice versa. These possibilities, although they correspond to minimizing the total distance travelled, may be of concern to the company, since its preference may be to start and end the routes in the same place.

Firstly, using a total of 5 vehicles, the maximum possible number of destination cities to be served by a single truck was defined as L = 3. Figure 7 shows the schematic map with the ideal routes. It is worth noting that the routes in red and purple do not have the same arrival and departure points.

Figure 7 - Schematic distribution of optimal routes for L = 3 and 5 vehicles



Caption: Starting from warehouse 1: [(15, 4), (4, 6), (6, 7), (7, 15)], [(15, 5), (5, 1), (1, 0), (0, 16)], [(15, 8), (8, 11), (11, 9), (9, 15)] e [(15, 10), (10, 12), (12, 13), (13, 15)]. From warehouse 2: [(16, 2), (2, 3), (3, 14), (14, 15)].
Source: Own authorship.

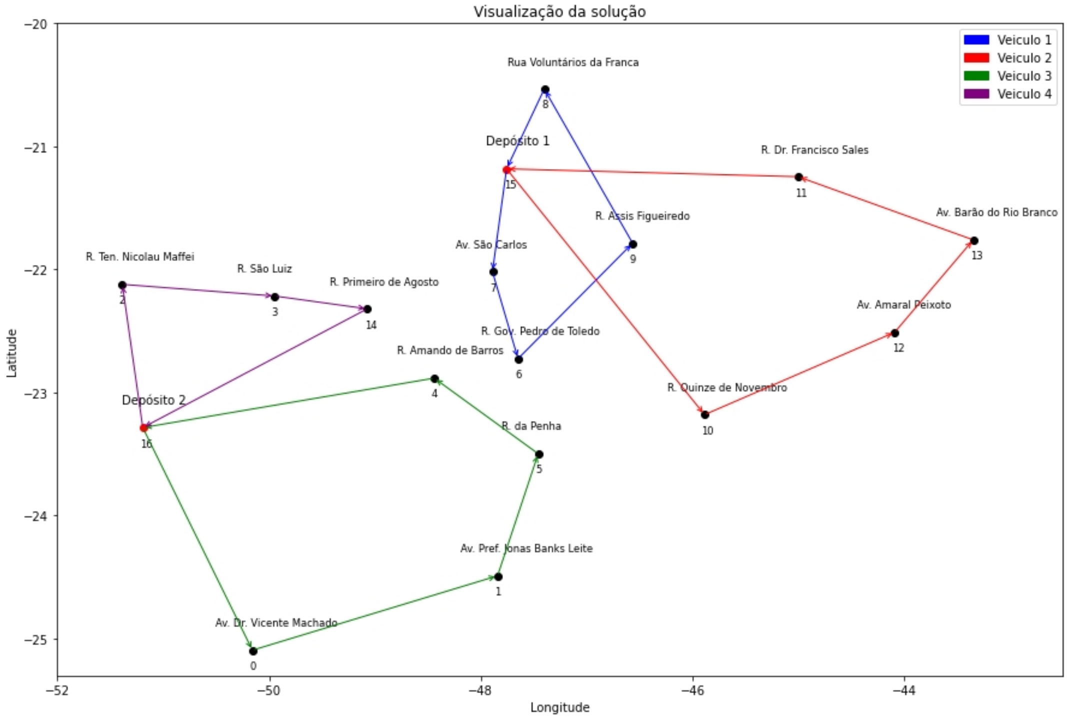
In another application, L = 4 was set and 4 vehicles would be used. To solve the problem under these conditions, instead of the constraints in equations (4) and (5), constraints (11) and (12) were incorporated, which require each depot to be the starting point for two trucks, guaranteeing a homogeneous initial distribution. Figure 8 shows the schematic map with the optimal routes. Table 2 provides a comparison between the 2 cases studied.

Table 2 - Comparison of results considering the restriction of the maximum number of nodes.

No. of vehicles	Maximum number of nodes per vehicle	Objective Function (km)	Distance per vehicle (km)
5	3	4592,503	918,5006
4	4	4249,801	1062,450

Source: Own authorship.

Figure 8 - Schematic distribution of optimal routes for $L = 4$ and 4 vehicles



Caption: Starting from warehouse 1: [(15, 7), (7, 6), (6, 9), (9, 8), (8, 15)] and [(15, 10), (10, 12), (12, 13), (13, 11), (11, 15)]. From warehouse 2: [(16, 0), (0, 1), (1, 5), (5, 4), (4, 16)] and [(16, 2), (2, 3), (3, 14), (14, 16)].
Source: Own authorship.

The solution with 4 vehicles shows a more balanced distribution of mileage per vehicle than the case with 5 vehicles. Compared to the solution in Figure 4, which considers 2 vehicles and does not impose a maximum limit on the number of points each truck can serve, the relationship between the total distance traveled and the number of vehicles in the case in Figure 8 is lower. From the company's point of view, the best solution may in fact be the one in Figure 4, due to the possibility of paying a higher price for just 2 drivers to make a journey of an average of 1595 km each, than paying for the services of 4 drivers to make a journey of an average of 1062 km each.

5. CONCLUSION

In order to find routes that minimize the total distance travelled to supply the stores in the Magazine Luiza chain, data was collected on two distribution centers and fifteen stores used, so that the distances between all the points in the problem were established. In addition, the number of vehicles available, their capacities and the demands of the destinations were defined. From this perspective, the solutions found for the Multiple Depot Vehicle Routing Problem in question were satisfactorily obtained from the *PuLP* linear programming package using the *Python* language.

The implementation of the code was tested by varying the number of vehicles, the maximum number of points that each vehicle can visit, as well as some aspects of the mathematical modeling, but keeping the capacity of the vehicles and the demand of the destinations fixed. Thus, different optimal solutions were found for the problem presented. Among these solutions, it was possible to see that the solution shown in Figure 4 obtained the shortest total distance traveled, with the use of only 2 vehicles and no restrictions on the maximum number of us that each truck could visit.

Finally, it can be concluded that, with the help of operational research and linear programming, the optimum results were presented and analyzed, considerably improving the construction of routes and promoting a reduction in the total distance traveled. From this perspective, the importance of companies dedicating themselves to the details of the logistics area is confirmed, because through careful analysis of the locations to be served and the distribution centers, it is possible to reduce costs and optimize operations, which favours the company's competitive factor in the market.

BIBLIOGRAPHICAL REFERENCES

[1] ARENALES, M.; ARMENTANO, V.; MORABITO, R.; YANASSE, H, **Research Operations for Engineering Courses**. Rio de Janeiro: Elsevier. 2007.

[2] KULKARNI , R.V.; BHAVE, P. R, **Integer programming formulations of vehicle routing problems**. India. 2011.

[3] MILLER, C. E.; TUCKER, A. W.; ZEMLIN, R. A. **Integer programming formulation of traveling salesman problems**. Journal of the ACM (JACM), ACM New York, NY, USA, v. 7, n. 4, p. 326-329, 1960.

[4] NumFOCUS. pandas: pydata, 2015. Home home page. Available at: <https://pandas.pydata.org>. Accessed on: July 1, 2022.

[5] MITCHELL, S.; O'SULLIVAN, M.; DUNNING, I. PuLP: A Linear Programming Toolkit for Python. Auckland, Sept. 2011.

APPENDIX A

The project, carried out through the *Google Collaboratory* platform, is available at: https://colab.research.google.com/drive/1i1h9AtouwsJryBiOHoIJuO7SbodXU_lR?usp=sharing

APPENDIX B

Duties performed by each of the group members:

- Summary Writer: Guilherme Melo
- Introduction: Camila Delfini
- Writing mathematical modeling: Lucas Tramonte, Eric Nascimento, Camila Delfini
- Writing the computer experiments: Filippo de Plato, Flávio Campedelli, Guilherme Melo
- Writing the analysis of the results: Filippo de Plato
- Conclusion Writing: Camila Delfini
- Code implementation: Lucas Tramonte
- Presentation slides: Camila Delfini, Eric Nascimento, Filippo de Plato