Optimizing The Last Mile Delivery Operations: A Simulation Study for We-Doo's Warehouse Location Selection

Vipin Sharma
Dept. of Computing
National College of Ireland
Dublin, Ireland
x22207406@student.ncirl.ie

Abstract - The Delivery industry is growing very fastly to finding the best warehouse location is very important for every delivery company. In this report, the author finds the best optimal delivery center location for a We-Doo Start-up company by using Simulation. The main aim of the company implement a model where the local drivers using the electric cargo bikes deliver the Parcels from a centralized location. By using the Simulation and considering various factors like delivery route length, driver working time, etc. Also, the author takes the total 6 different warehouse locations and 20 days then finds the cost of all the warehouses for every day and gives the warehouse location with the lowest cost to the We-Doo company also by using the ANOVA and F-statistics performed the statistical analysis and the Python SimPy frame work play a very important role.

Index Terms: Simulation, e-Cargo bikes, Delivery, ANOVA, Warehouse

I. Introduction

The logistics industry has grown very Fastly in recent years due to e-commerce platforms. The Express delivery industry's competition is also very high. A growing field of express companies is focusing on ways to cut down costs while improving customer services. The main issues of optimizing the location of the distribution center have gained attention from both the academic and industrial sectors [1]. Everyone wants same-day delivery services of the products in their home whether it is vegetables, fruits, books, clothes, or any parcels. Currently in Asia and Africa 40% of the people live in urban areas which will increase by 60% in 2050 and because of the large population in urban areas shipping the parcels causes challenges. There are new technologies and transportation have been introduced to allow for more efficient last miles delivery in cities [2]. It is very challenging for the logistics company to deliver the product at the right time while taking the optimum route from the local shops to the customer's house. At this time most of the logistics companies establish their local delivery centers and then distribute the parcels to end customers through electric cargo bikes. Cargo bikes are much more better than the delivery trucks in the city and they also do not produce greenhouse emissions they are regarded as bikes with no emissions. There are many problems faced by the logistics company when they deliver the parcels in the urban areas and city center. One of the main problems in the city is the congestion problem and the pollution due to the strong increase in the number of parcels in recent years. There are a number of efforts have been put into place to break down the pollution levels and traffic congestion. Perhaps the most popular approach is the location of a warehouse center location close to the city center in order to combine goods delivery services [6]. The main goal of this research is to find the best optimum warehouse location for the We-Doo delivery center within a given township map while considering various factors such as operational cost, time to deliver the order, and other various factors. Traditionally the last-mile delivery phase involves the transportation of goods from the warehouse center to the customers and has been defined by various challenges such as high cost, inefficient routes, and congestion problems in city centers. To solve these problems various solutions have emerged such as crowd-shipping, drone delivery, and micro fulfillment centers. The idea of the startups company like We Doo are promoting, which optimizing the last-mile delivery operation. Many courier services will make individual trips to deliver the parcels to the customers who are dispersed throughout the region. For the logistics company, they cause inefficiencies, higher carbon emissions, and operating costs. Customers also faced problems like missed deliveries, taking too much time for the parcels to be delivered to customers and sometimes the package being broken, or stolen. To solve these issues We-Doo applies the centralized delivery models to collect the packages from the various different-different providers at the nearby best optimum local center. From that optimum location, the assigned drivers use the electric cargo bike or e-bike to deliver the parcels to people very quickly and effectively. Also the main aim of the We-Doo startup company is to improve last-mile delivery, reduce the environmental effects, and improve the delivery experience for customers. Also using the simulation important metrics such as the number of customers, total number of parcels delivered, and packing time of the parcels will be analyzed to find the total cost linked with each of warehouse location.

Assuming this situation, an author must provide the answer to the following question: What is the best optimum location of the warehouse location in the generated map? The document structure is broken down as follows. The first section provides an introduction to the research topic then section 2 provides the related and previous research done on this domain and in this field. In Section 3 author goes through the methodology and the next part of the report shows results and Interpretations in Section 4, and In the last section 5 mentions some reflections and future works that help to improve the proposed work in the future.

II. Literature Review

The delivery business has grown very fast in some years. Everyone uses delivery services whether it is a customer, E-Commerce business, cloud kitchen, distributor etc. So It is very important for delivery companies to reduce the cost of delivery by doing the delivery on time. Basically, It means finding the optimal path for the customers, and distributors so the delivery companies reduce the cost and time for the delivery of the products

The research is done by Y. Ji, H. Yang et al [1], to establish an optimization model of the regional express distribution center location based on the reality of the road network. There is one big distribution center in every region that supplies the product to the smaller shops. Every small local shop has there own requirements and the capacity of trucks is limited so due to various kinds of goods and the complex distribution route conditions the lack of direct demand for each point it becomes very difficult to establish and analyze the best optimizations model therefore the researcher uses the fuzzy C means clustering algorithm to solve this problem and optimize the location of the distribution center also Fuzzy algorithm using the shortest path algorithm to calculate the shortest distance point between the two places.

This research paper by V. Naumov and M. Pawluś [2] describes methods which developed in order to solve a combination of various problems of cargo bike logistics, including routes, packing, speed constraints, etc. By applying various methods like the Traveling Salesman Problem, Knapsack Problem, and Traveling Thief Problem based on the data which is generated synthetically. The models described in the paper are implemented using the Python programming language, which will used to solve the electric cargo bike routing under real-world conditions. There are two main issues Routing problem and the packing problem which are solved in the paper. The routing problem is solved by using the Travelling salesman problem (TSP) to create an optimal route that visits all the given nodes in a minimum distance. The another problem which is the packing problem addressed by the Knapsack Problem (KP). In the Knapsack the main goal is to find a good packing plain that does not exceed the maximum cargo bike load. The paper also tells about how Cargo bikes is efficient for deliveries basically the cargo bikes are divided into several categories according to their frame build, maximum load capacity, number of wheels, and electric motor which are used to convert our normal cargo bikes into the electric-cargo bikes or e-cargo bikes. There are many advantages to using cargo bikes for delivery as compared to trucks for short and last-mile delivery, mainly in the city center which is very congested and lots of people moving on the road. The size of the cargo bikes is small therefore there is no chance of traffic, it takes fewer parking areas and for many of the countries there is no need for a license to use cargo bikes. Various algorithms used in this paper like Bin-Pack-3D, MTSP-First algorithm, and CVRP-First algorithm all used to find the best feasible route with the lowest travel time. When testing the algorithm it is proved that the CVRP approach is able to give the best optimal results which deliver the 10 consignments with 4 bicycles in 2s and the MTSP is not able to give the best optimal results. The research was done by Tanja Niels et al. [3] in Munich, Germany how package delivery is carried out by a big Courier, Express and Parcel (CEP) company by electric bikes and cargo bikes. In this project, two containers and one truck are places in a given area in the Munich city center. The main work of the containers function as depots for the parcels to be delivered to nearby locations in the city centers. There are some main issues which is addressed in this paper first what are the optimal places for the containers? secondly, what are the optimal routes for the cargo bikes, and how much time it will take to deliver the parcels to the customers and Lastly what was the total vehicle mileage of diesel trucks? The heuristic approach used in this method to find the optimal routes and at the end find the best container locations and routes for a CEP service system. The average time and distance also decrease by diesel trucks per day from 180 km to 45 km approximately.

Find the best warehouse location for Online retailers is very difficult In this paper Can Chen et al. [4] provides the data-smart approach for define the connected capacitated warehouse location problem (CCWL) which is used to search for the minimum total transportation cost of the warehouse network including various things like warehouse customer delivery cost, cost of warehouse, supplier warehouse shipping cost, etc. Next using the data mining techniques performed a computational study on the capacitated warehouse location problem (CCWL) and Used the EM algorithm to find the best optimum warehouse location with the less computation cost and The Hierarchical clustering used to reduce grouping cities with close locations also using the Artificial Neural Network (ANN) to predict the sales distribution for all the online retailers based on logistics service factors and various demographics data. The goal of this paper is to find the optimal warehouse design and F. Yener and H. R. Yazga [5] addressed two main problems optimum warehouse design, and assignment problem these problems were solved by a new mathematical model and a Multidimensional Scaling Algorithm (MDS). The routes of selecting picking are sorted by a linear mathematical model which is based on the Vehicle problem (VRP). Simulation is also used to evaluate the best warehouse design location. There are three Stages used in this paper Stage 1 to proposes a new design to the existing one, using the association rules to obtain confidence and support measures via data mining techniques, and lastly, find the relationship matrix of all the items. The second stage uses MDS and a mathematical model to assign the products to the slots using the distance formula and the VRP to minimize the travel distance of order-picking routes. In stage 3 Simulation and hypothesis testing to get an effective proposed approach. The research was done by Andrés Muñoz-Villamizar et al. [6] to create a simulation-based algorithm for location and routing problems. The paper addresses the issue of urban distribution center placement and vehicle routing, for this issue a hybrid algorithm uses Monte Carlo simulation to introduce the biased randomness into multiple optimization stages to solve the combined problem. Doing this reduces the noise and traffic problems within urban areas. For the depot location using the stochastic heuristic that performs the biased selection clustering using Monte Carlo simulation and for the assignment and vehicle routing two algorithms SR-GCWS and SR-GCWS-CS are used. The interesting feature of this algorithm it also used Monte Carlo to generate a random points distribution. The results using the simulation-based algorithm are good but the complexity of the problems makes the method inefficient.

III. Methodology

In this section explains a detailed description of the process that how the simulation and optimization are used to get the optimum results. To develop a model which is based on the simulation is not easy so it is very important to use the ''SimPy" package which comes with Standard Python. By using the SimPy package the simulation performed fast and it is easy to create a simulation environment in the jupyter notebook by using the python programming language and various python packages like numpy, pandas, matplotlib etc. The Simpy package is very versatile tool for simulating complex systems and its allows many researchers to analyze the behavior of dynamic systems in a compact environment. The problem explanation that was provided, the author needs to give the best delivery center location which takes the lowest cost value to deliver the parcels and also give the statistical analysis method to compare the cost of different warehouse location. The words ''Recorder", ''Parcel", ''Customer", ''Driver", ''Delivery Center" are the important components of a simulation.

The 'Recorder' class was the first class which is declared as the reference points for capturing data during the simulation. The class have also define the data frame which records per working day outcomes and after that I also define the function called calculateDailyCost which calculates the daily cost operation based on three factors operational cost per km, pay per hour, and minimum daily pay. The values of these factors give in the problem so I have

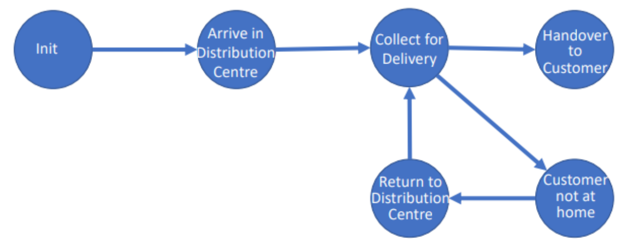


Fig. 1. Event Graph for Parcel class

to put those values in the above three constraints. This method iterates every day in the simulation and calculates the operational cost, driver pay for a particular day and also the total daily cost and at the end updates the cost column with the daily calculated cost. There are also many methods in the class that record the many events which is used to record the various events during the simulation.

The second class is the Parcel class which is very important for simulating individual parcels and their movement in a delivery system. Figure 1 show that the Parcel class easily depicts the movement of parcels from one location to another location or from the warehouse to the delivery center and using the Parcel class it is easier to implement the features. The third class is the Customer class graph shown in figure 2, In the delivery simulation, the customer class is very important for modeling customer behavior and interaction with delivery personnel. In the simulation of customer class, First call the customer at the door if the customer accepts the call and take the parcel then sign off and if the customer does not take the call or not answer the door bell sign off and leave the house. Basically, the customer class tracks all the Parcels received by each customer or not and evaluates the delivery performance matrices.

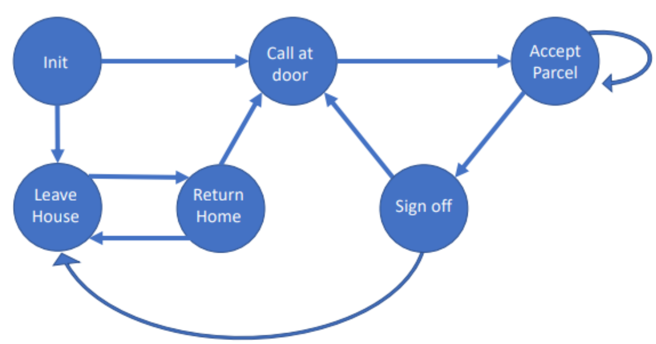


Fig. 2. Event Graph for Customers Class

The Next class is the Driver class which is very essential to delivering the Parcels to the customer. In figure 3 event for delivery class represents the delivery process and defines the methods like arriving for work and leaving for Delivery. The complete flow of this class driver arrives for work and then leaves the delivery point once reaches the customer's house rings the customer bell if accepted hands over the Parcel and signoff if the customer is not at home return the parcel to the delivery center.

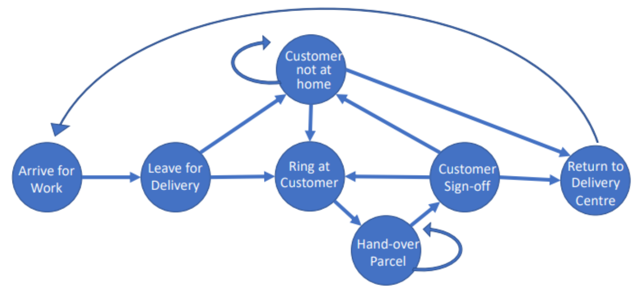


Fig. 3. Event graph for Driver Class

The last class in the simulation is the Delivery Center class shown in figure 4 which is important to managing parcel distribution within the simulation delivery system. The Flow graph shows once the parcel is accepted it goes for the preparation of delivery and then returns from delivery if it was not delivered successfully then after this it will again go to the accepted parcel and the loop is going.

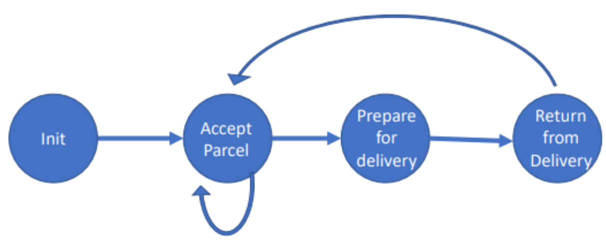
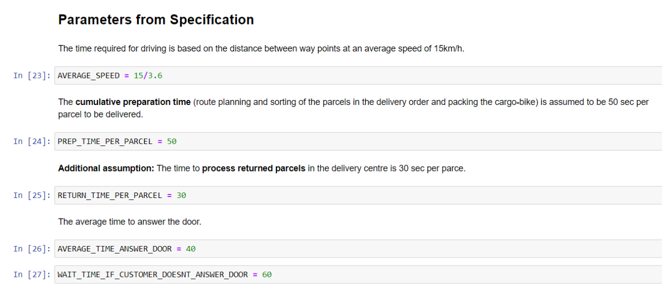


Fig. 4. Event graph for Delivery Class

A. Part 1: Baseline Simulation

In the baseline simulation to take the optimum simulation, the author needs to take some parameters assumptions like average speed, preparation time, return time per parcel, average time to answer the door and wait time if the customer does not answer the door which shown in figure 5. The average speed is 15km/h and the cumulative preparation time which includes route planning, sorting of the parcels in the delivery and packing the cargo bikes we have assumed it takes 50 seconds per parcel to be delivered. There are some additional assumptions which the author makes, the time taken to complete the process of returned parcels in the delivery center is 30 seconds per parcel and the last two assumptions average time answer the door by the customer is 40 seconds and if the customer to not give the answer we assume that the wait time is 60



Parameters from Specification

The time required for driving is based on the distance between way points at an average speed of 15km/h.

```
In [23]: AVERAGE_SPEED = 15/3.6
```

The **cumulative preparation time** (route planning and sorting of the parcels in the delivery order and packing the cargo-bike) is assumed to be 50 sec per parcel to be delivered.

```
In [24]: PREP_TIME_PER_PARCEL = 50
```

Additional assumption: The time to **process returned parcels** in the delivery centre is 30 sec per parce.

```
In [25]: RETURN_TIME_PER_PARCEL = 30
```

The average time to answer the door.

```
In [26]: AVERAGE_TIME_ANSWER_DOOR = 40
In [27]: WAIT_TIME_IF_CUSTOMER_DOESNT_ANSWER_DOOR = 60
```

Fig. 5. Parameter for Assumptions

seconds. Once author put all those parameters then Generate the input data using the generateDeliveryData function which take the four parament p which is the average number of parcels per day to per customer, c which is the total number of customer is served, days and seed number which means author take the last four digit of student Id as a seed number. The author Student ID (222207406) and the seed number is 7406. Using this seed number the result is very distinct from the other students and also generates a different different delivery routes for the delivery.

B. Simulation Study

The total number of days is 20 which is denoted as (N=20) and we have to take the total of 6 different warehouse locations map. There are 20*6 =120 maps generated with different warehouse locations and then by performing the Annova test we compare all the warehouse location's mean of operational cost with each other also compare the F-statistics and p-value. After that run the Monto Carlo optimization code to get the best optimal warehouse location which takes less cost with the less p-values to complete all the deliveries. For the comparative analysis, we compared the p-values, f-statistics, and cost values of all the warehouse locations to find the best one warehouse location.

This result section will outline the conclusion reached after running the simulation model along with the changes we made in the code like changing the parameter, taking the 6 warehouse locations, and conducting the statistical analysis. By performing the statistical analysis of the 6 warehouses author

IV. Result and Interpretation

makes two graphs that show the daily working time. In figure 6 above graph x-axis shows the working time and the y-axis shows the density and in the below images the x-axis shows the number of days and the axis shows the working Time (minutes). Here we calculate the Mean, Median, and standard deviation of all the warehouses which is shown in figure 6.

By comparing all the graphs the author finds that on the 10 day the working time is less in all the warehouses and on 6 day the working time is very high. The mean value of warehouses 2, 3, and 4 are (171.5, 170.25, 171.7) similar

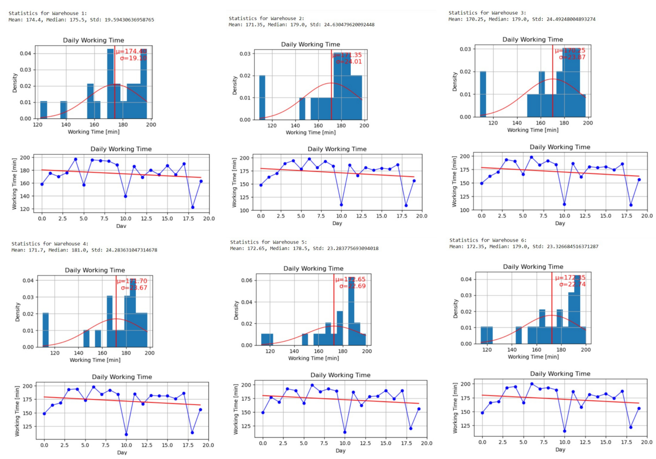


Fig. 6. All 6-warehouse Daily working time

not too much difference and warehouses 5 and 6 mean value (172.65, 172.35) also no difference in the mean value but the mean value of the first warehouse is 174.4 which is higher than others warehouses.

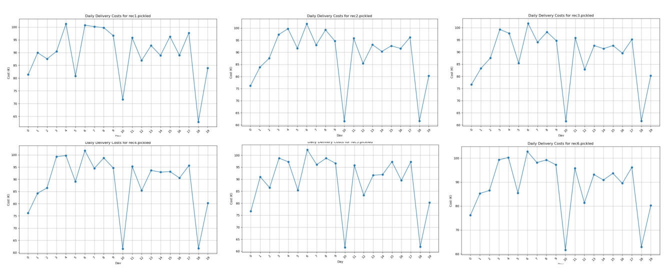


Fig. 7. All 6-warehouse Daily Delivery Costs

After this author calculates the Daily Delivery cost of all the warehouses and makes the line chart for all 6 warehouses In Figure 7 the x-axis shows the total number of days and the y-axis shows the total cost of the

warehouse. If comparing all the images author concludes that on the 10 and 18th days the cost is very low approximately 62 and 63 euros but on the 6th day the cost will be very high and for all other days the cost is between 80 to 95 euros. To Optimize the warehouse location we use the Monte Carlo Optimization techniques. In the figure 8 shows all the warehouse candidate positions based on the data and given parameters. Each location undergoes an evaluation by a function called the objective function and calculates the mean tour length after that by employing the Monte Carlo approach the algorithm iterates and find the best warehouse positions which shows in the figure 9.

Also at the end author performed the ANOVA test, The ANOVA test is used to compare the means of multiple things here we used ANOVA for calculating the means of all the warehouses. By applying the ANOVA the F-statistic is 0.0443 and the p-value is approximately 0.998 shown in the figure 12. the author assumed that the p-value is less than 0.05 but by

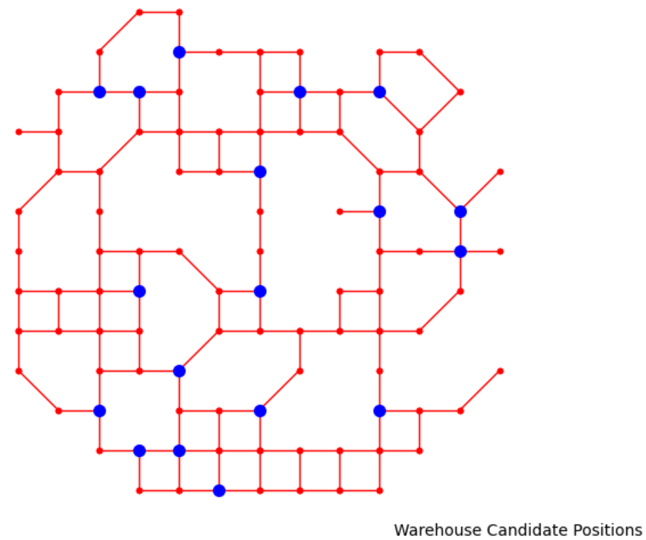


Fig. 8. All Warehouse Candidate Positions

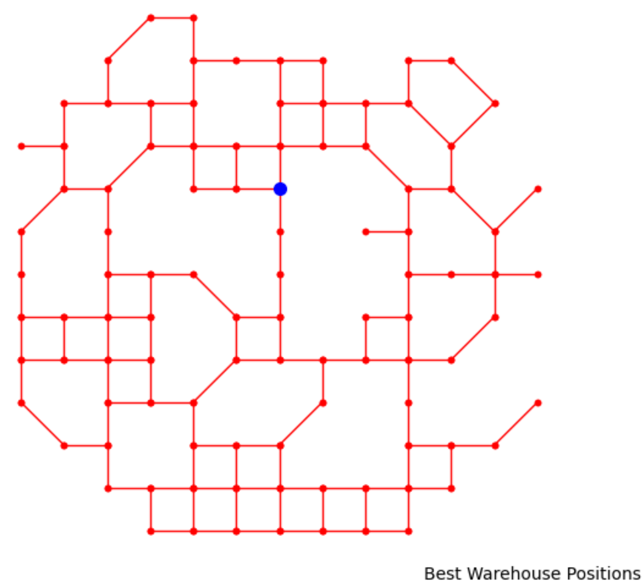


Fig. 9. Best Warehouse Position for Seed Number=7406

doing the Annova test the p-value is higher than 0.05 which means that there is no such significant difference across the warehouses.

V. Reflection and Future Work

By performing all the analysis and taking 6 different warehouse locations found that there is no such difference between the mean values. Also in the report, we calculate the Cost of each day, the total left Parcel, the total distance traveled, start time and end time of the work. By applying the ANOVA test compare costs across various warehouses. This work is a step toward an in-depth understanding and optimization of ware-

	Warehouse1	Warehouse2	Warehouse3	Warehouse4	Warehouse5	Warehouse6
0	81.34944	76.15344	76.67616	76.15344	76.68800	76.15344
1	89.95056	83.78960	83.27136	84.28960	90.98064	85.27136
2	87.52336	87.52336	87.52336	86.47872	86.47872	86.47872
3	90.49888	97.28672	99.28672	99.28672	98.78672	99.28672
4	101.25448	99.71480	97.69032	99.71480	97.19032	100.26544
5	80.82280	91.67056	85.44528	89.08096	85.42832	85.45808
6	100.74216	101.75976	101.75976	101.75976	102.25976	102.76720
7	100.17688	92.99656	93.99656	94.49656	96.05432	98.13432
8	99.76296	99.23720	98.23720	98.73720	98.73720	99.23720
9	96.65120	94.62912	94.62912	94.62912	96.63072	97.21872
10	71.61992	61.58424	61.58424	61.58424	61.64328	61.67384
11	95.77840	95.77840	95.77840	95.26176	95.78944	95.77840
12	86.93592	85.40744	82.81784	85.39296	83.31784	81.31784
13	92.72992	93.14032	92.64032	93.64032	91.64032	93.14032
14	88.88920	90.34456	91.39448	92.93416	91.93416	90.86344
15	96.27624	92.65816	92.65816	93.15816	97.26352	93.65816
16	88.95328	91.58752	89.49792	90.54288	89.48864	89.48864
17	97.71408	96.16944	95.16944	95.66944	97.25904	96.16944
18	62.83280	61.63024	61.63024	61.70040	61.84512	62.89608
19	83.92936	80.27832	80.28576	80.27832	80.27984	80.27832

Fig. 10. Cost of all 6-warehouses

Out[42]:

	begin work at	end work at	dist	left	working time	cost
0	64800	73693	26918	0	148	76.15344
1	151200	161218	28392	0	166	85.27136
2	237600	247707	30984	0	168	86.47872
3	324000	335587	34834	0	193	99.28672
4	410400	422111	34568	1	195	100.26544
5	496800	506764	30726	0	166	85.45808
6	583200	595228	34590	3	200	102.76720
7	669600	681078	32929	0	191	98.13432
8	756000	767632	34215	1	193	99.23720
9	842400	853758	33984	0	189	97.21872
10	928800	935711	20923	4	115	61.67384
11	1015200	1026386	34730	1	186	95.77840
12	1101600	1111111	28973	1	158	81.31784
13	1188000	1198907	33004	0	181	93.14032
14	1274400	1285077	29543	0	177	90.86344
15	1360800	1371731	33227	0	182	93.65816
16	1447200	1457665	31108	0	174	89.48864
17	1533600	1544866	33368	0	187	96.16944
18	1620000	1627376	23701	0	122	62.89608
19	1706400	1715769	28479	0	156	80.27832

Fig. 11. Best warehouse location reuslts

house operations. For future work applying the data analysis, machine learning techniques, and Predictive modeling also increases the number of days and total number of warehouses. By using the various machine learning techniques forecast the total cost, identify the cost of drivers, and make the future map and strategies on how to deliver the parcels in less time by choosing the optimum routes and saving money.

```
ANOVA Test Results:
F-Statistic: 0.04430605927978627, P-Value: 0.9988309182132318
There is no significant difference between the warehouses.
```

Fig. 12. Annove test Results

References

[1] Y. Ji, H. Yang, Y. Zhang, and W. Zhong, "Location Optimization Model of Regional Express Distribution Center," Procedia - Social and Behavioral Sciences, Nov. 01, 2013. https://www.sciencedirect.com/science/article/pii/S1877042813022416.

[2] V. Naumov and M. Pawluś, "Identifying the Optimal Packing and Routing to Improve Last-Mile Delivery Using Cargo Bicycles," Energies, Jul. 08, 2021. https://www.mdpi.com/1996-1073/14/14/4132.

[3] "Design and Operation of an Urban Electric Courier Cargo Bike System," IEEE Conference Publication — IEEE Xplore, Nov. 01, 2018 https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=arnumber=8569606.

[4] "Warehouse Site Selection for Online Retailers in Inter-Connected Warehouse Networks," IEEE Conference Publication — IEEE Xplore, Nov. 01, 2017. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=arnumber=8215559.

[5] F. Yener and H. R. Yazgan, "Optimal warehouse design: Literature review and case study application," Computers Industrial Engineering, Mar. 01, 2019. https://www.sciencedirect.com/science/article/pii/S0360835219300130.

[6] "A simulation-based algorithm for the integrated location and routing problem in urban logistics," IEEE Conference Publication — IEEE Xplore, Dec. 01, 2013. https://ieeexplore.ieee.org/abstract/document/6721581.