Drone Delivery Kaggle Competition – Final Project

**Introduction**

Online shopping has increased in popularity over the last ten years. If you have not previously had experience in the online marketplace, the quarantine in 2020 should have expanded your palatability in the space. The traditional cycle of online shopping has a reliance on deliveries involving some form of human interaction. Over the years, there have been various efficiency improvements created to solve or mitigate the human variable. The evolutionary phenomenon has transformed the industry today, where GPS (global positioning system) and transportation route delivery optimizations are abundant. As automation technology and new inventions arise, the delivery industry previously bound in terms of human interfaced completion has become an obstacle to optimizing product delivery. As consumer expectations increase to same-day delivery, the capability and constraints of human interfaced completion also increase. ​ The expectations of consumers become increasingly difficult to satisfy.

Drones are a new technology rising in relevance and use that is seen as a potential replacement for human interfaced delivery systems, disrupting the current status quo. The Internet of Things Agenda states, “A drone, in technological terms, is an unmanned aircraft. Essentially, a drone is a flying robot that can be remotely controlled or fly autonomously through software-controlled flight plans in their embedded systems, working in conjunction with onboard sensors and GPS" (Rouse, 2019). These automated flying “robots” will require a focus on route optimizations to implement efficient strategies properly. This route optimization is still a problem discussed today, with many organizations setting up competitions to practice solving the eventual real implementation of drone delivery systems. The competition addressed in this paper is the Kaggle Hash Code Drone Delivery Competition. The idea is that a fleet of drones must efficiently complete a delivery scenario that might arise, which requires optimization ahead of time.

**History of Delivery Optimization**

Delivery services origination can be traced back to the Persian Empire in the sixth century BCE with a postal delivery service created by King Cyrus the Great. As the world evolved and developed, so did the need for transformative delivery methods. The early use of courier pigeons for delivery later became horse-based delivery. In the United States, the Pony Express was a rapid mail and package delivery system that reduced the time of package delivery to 10 days. When railroads were developed, they dramatically increased delivery speed and extended the volume of delivery capabilities. Despite the developments over the years, human interfaced delivery completion showed increasing constraints, increased costs (labor, transportation, maintenance, availability, etc.), transportation obstacles (traffic, roadway barriers, speed limitations, etc.), grid route dependency, time window maximization per delivery person, and more.

The origins of delivery route optimization are unclear. However, the concept had its earliest discussions in the 1800s in a German Handbook for traveling salespersons (Welch, 2015).  The official Traveling Salesperson Problem, or TSP, was first studied in Vienna and Harvard. Mathematicians were considering a bus routing problem searching for any possible mathematical solution (Welch, 2015). ​ When given a list of cities and distances between pairs of cities, the TSP asked the question of what the shortest possible route was for sending buses. This concept continued to evolve into different forms, such as the mTSP, a multiple TSP where more than one salesperson is present (Bektas, 2006). More recently, the VRP, or Vehicle Routing Problem, a derivation of the mTSP, has been the focus of optimization. A VRP involves fleets of vehicles being optimally routed to finish their specific tasks and focus on new automated vehicle technologies (Christofides, Mingozzi, and Toth, 1981). Currently drone VRPs are the topic of discussion due to their unrivaled potential. Drones can deliver packages for under 10 cents a package compared to the $1.20 per package that UPS (Welch, 2015). This is an enormous benefit that has caught the eyes of delivery companies worldwide.

**Analysis and Models**

**About the Problem**

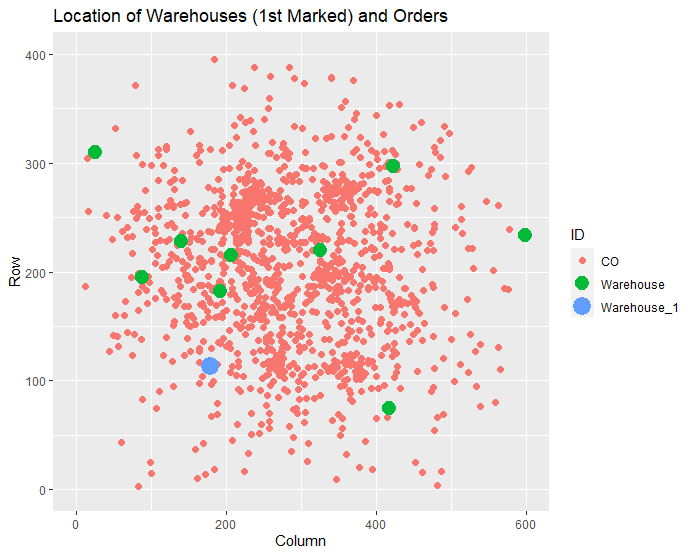
The Kaggle Drone Delivery Problem requires optimizing the delivery of products of customer orders from a warehouse in the most efficient time. The goal is to meet the scenario completion deadline as efficiently as possible to ensure a high score. Solving such a scenario would, in the real world, increase profits and decrease operation costs. Current considerations for this project include the weights of the individual and collective products per delivery, turn-based time requirements, product availability, available drone commands, and entity coordinates, which are the locations of warehouses and orders.

**About the Data**

The provided scenario, a .in file, comprises multiple lines with a pattern of scenario parameters, followed by product information, warehouse information, and last but not least, customer order information. The data is loaded by utilizing loops, and the readLines() function within R. Generation of random .in files for testing and training can also be used for differential scenarios. Loading the data results in multiple tables of transactions and locations. In creating the optimization models, non-relevant aspects like drone lifespan and drone maintenance are not considered. The Kaggle scenario provides a static fleet of 30 drones, a list of 1,250 customer orders for fulfillment, and assigned availability of products in each of the ten warehouses.

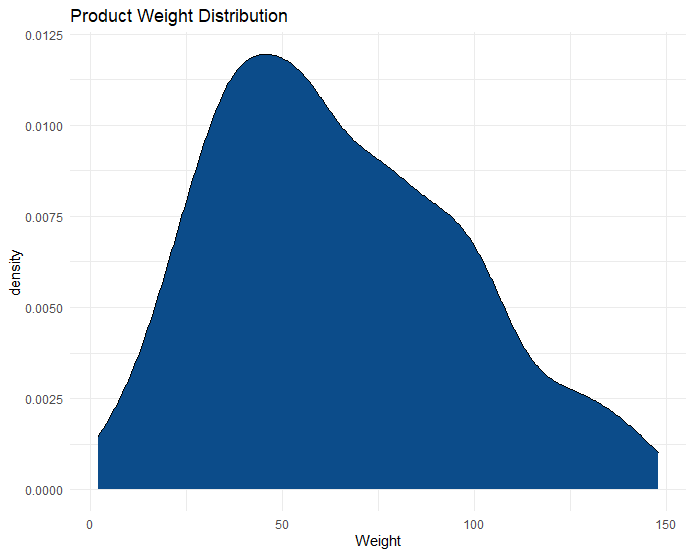
**Descriptive Analysis**

Location of all Entities with Drone Starting Location Marked:



The customer orders were plotted as orange circles first to inspect where customer orders were for the simulation parameter. Next, the warehouses were plotted and designated as green. Finally, the original warehouse where all the drones begin from was colored blue. The customer orders appear to have a large cluster near the 200th column and 250th row. There are some other customer orders clustered together, and some spread out. The warehouses show that none appear directly in the middle of the largest cluster of customer orders, and two of the warehouses are not near many orders. The warehouses on the edges will require much more time to deliver products for centrally located orders and vice versa for central warehouses to deliver products to edge customer orders.

Density Distribution of Product Weights:



The next step in the analysis is to view the product weights. By reviewing the graph, none of the products weighed the same as the maximum drone capacity of 200u. The minimum product weight is 2u, and the max product weight is 148u. Most of the products weigh less than 100u, and the average product weight is 65u while the median is 60u. Since none of the drones weigh maximum capacity, the drones can carry multiple products at once (average of 3 products of the same type loaded at once). This leads to numerous possible combinations of products that can be carried by the drone that must be considered during optimization.

**Models and Techniques**

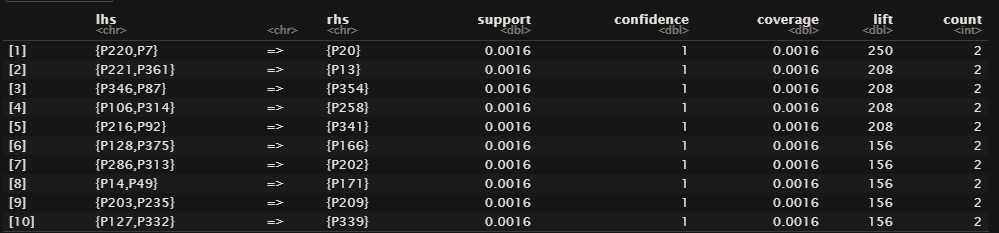
Cluster Plot of a K-Means model with 7 Centers:

Chart, scatter chart

Description automatically generated

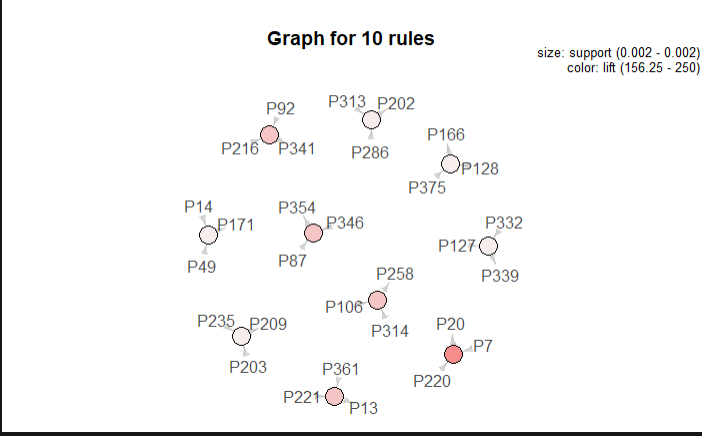
The above cluster plot is based on a simple K-Means for seven centers. This plot shows that products seem to have identifiable groupings, including a few clusters which contain only a few products inferring some unique product associations. This is useful to consider for shipping products together for multiple orders. When optimizing the best route, clusters could be utilized for designating products to be shipped together with more efficiently.

Association Rules:



Association Rule Mining was used to determine if relevant associations existed between various products. This was done due to further understand product relationships in customer orders that were partially identified during clustering. After transforming the data into transaction format, the Apriori algorithm was used on the product transactions using support = .001, confidence = 0.9, and max length = 3. This generated 1982 rules. By sorting by lift, it can be seen that the top 5 rules had a lift greater than or equal to 200. However, the majority of the 1982 rules had a confidence of 1 and a support of 0.0016. When checking the count, the max count for any rule was 4. This indicates the relationships between products in the given Kaggle Scenario are either relatively weak or non-relevant due to the lack of orders they apply to.

Top 10 Rules by Lift:

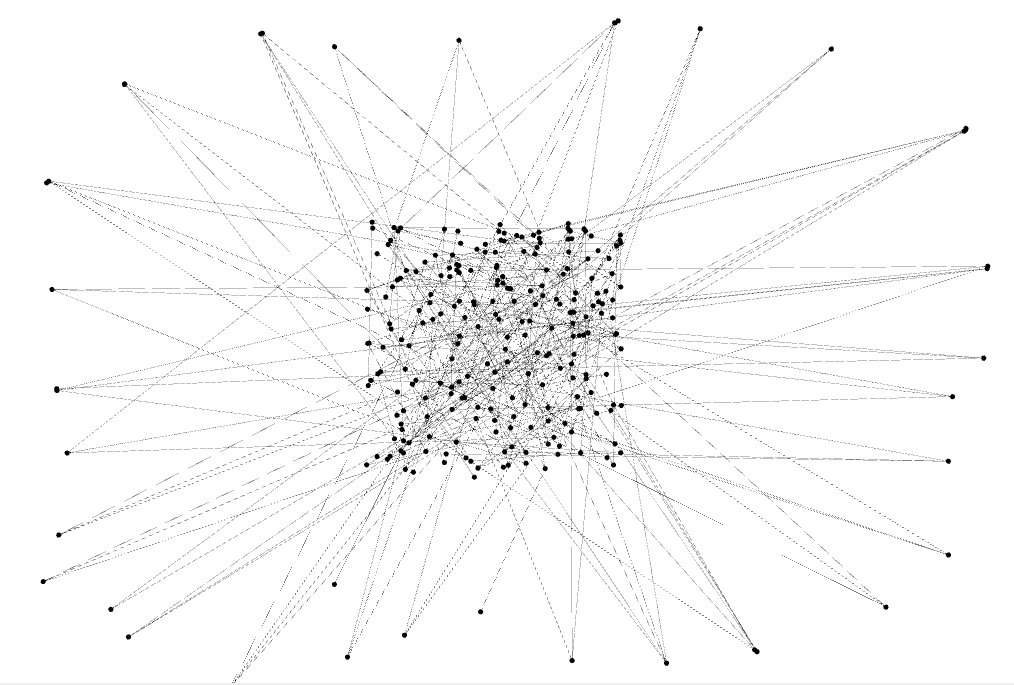


Top 50 Rules by Lift:

Chart, bubble chart

Description automatically generated

Top 100 Rules by Lift:



**Drone Optimization**

A variety of ways of assigning orders to drones can be utilized in attempting to increase the scores of drones. These assignment variations can be basic algorithms for optimizing the delivery scenario, such as assigning orders to drones according to warehouses or by the drones themselves. More complex algorithms such as genetic sweep, dynamic evolution, or gravitational search algorithms can be applied. Unfortunately, most of these algorithms are not available for use in an R package requiring that this algorithm be built by hand, so only basic algorithms for assigning orders were utilized.

Score Obtained with Order Assignment by Drones: **3389790**

The score obtained shows that the optimization while successful still has room for further improvement that more complex algorithms would likely drastically improve the result. However, due to the complexity of said algorithms, manually coding them would take quite a while and therefore could not be done during this project.

**Conclusion**

Times are changing with new technology and better automation, disrupting long-standing business fields just like the delivery industry. Drones are set to be a considerable disrupter that is already starting to influence several sectors. Drones will provide huge cost reductions and new benefits that existing transportation techniques and technologies do not offer. Trailblazers of drone delivery such as Zipline and Wingcopter have had wild success in Africa and Europe. Now drone delivery is coming to the United States. Big companies like Google (Alphabet), Amazon, and UPS are rushing to implement drone delivery systems due to the foreseen benefits and cost reductions. The delivery industry and others are about to be drastically shaken up as drone technology continues to catch on.

There are clear improvements that can be made to the optimization techniques, especially when considering more complex algorithms. Advances in routing technology are coming out quickly with new discoveries every month. An automated flying delivery system's benefits are apparent with the reduction in human labor costs, increased delivery flexibility, easier routing due to lack of terrain constraints, and more direct non-Manhattan routes. The dramatic difference in price for shipping packages and the potential to ship to places that could not be shipped to previously demonstrates why so many people are looking into solving vehicle routing problems for drones. Drones are the future of the delivery industry, and it will be interesting to see how quickly companies expand their drone infrastructure once they start truly implementing drone delivery systems.

**References**

Christofides, N., Mingozzi, A. & Toth, P. Exact algorithms for the vehicle routing problem, based on spanning tree and shortest path relaxations. *Mathematical Programming* **20,**255–282 (1981). <https://doi.org/10.1007/BF01589353>

Rouse, Margaret. “What Is a Drone?” *IoT Agenda*, TechTarget, 15 July 2019, internetofthingsagenda.techtarget.com/definition/drone.

Tolga Bektas, The multiple traveling salesman problem: an overview of formulations and solution procedures, *Omega*, Volume 34, Issue 3, 2006, Pages 209-219, ISSN 0305-0483, https://doi.org/10.1016/j.omega.2004.10.004.

Welch, Adrienne, "A cost-benefit analysis of Amazon Prime Air" (2015). Honors Theses.