

2019 INFORMS O.R. & Analytics Student Team Competition

PRELIMINARY PROBLEM STATEMENT

Important Note: This is a preliminary statement of the 2019 problem. Updates and changes may be made, with the final problem statement published on October 12. Be sure to check back on October 12 to see the final problem and to register your team's intent to compete. Teams must register in order to have access to company data, entry instructions and templates, links to complimentary software, and other competition materials. After October 12, there may be additional updates to the problem statement; all registered teams will be notified of any updates.

General Motors: Redefining Vehicle Delivery with Autonomous Cars

Contents

Background	2
Situation and Current Solution.....	2
Impact with Autonomous Vehicles	2
Static Routing vs Dynamic Routing	3
Problem Statement.....	6
Key Assumptions.....	7
Scope	7
Transportation.....	7
Vehicle Distribution Centers	7
Plant VDCs.....	7
Stand-alone VDCs.....	8
Flexible VDCs.....	8
Time.....	8
General.....	8
Datasets	9
Output Format	10
Evaluation Criteria.....	12

Appendix	13
Abbreviations	13
References.....	13

Background

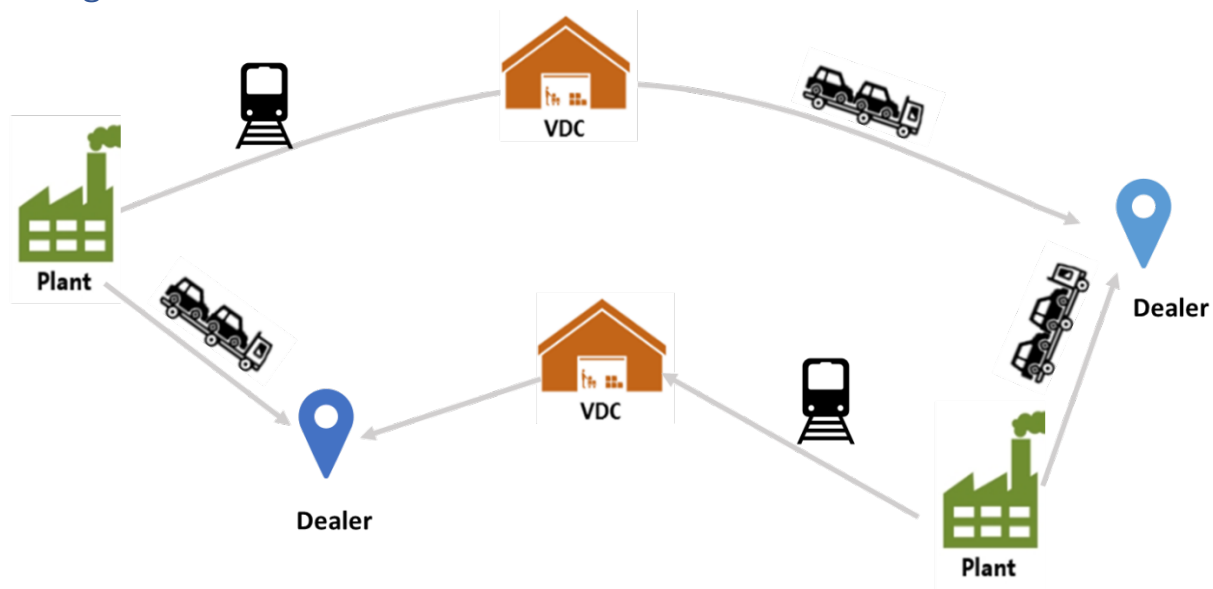


Figure 1 Idealized Outbound Logistics Network

Situation and Current Solution

Every day, General Motors (GM) ships out an average of 29,000 vehicles from approximately 70 assembly plants to approximately 20,000 dealers by truck, rail and ship. Geographical groups of dealers are assigned to vehicle distribution centers (VDC).

The traditional approach to designing the logistics network assumes constant manufacturing and demand volumes, and uses an optimization model to minimize the total shipping and operating cost. The model determines the lowest cost path (legs and modes) from each plant to each dealer, given their locations, existing VDC locations, forecasted demand by demand areas, availability, and the cost and duration of transportation between each node.

Vehicles must also be delivered to dealers in a timely manner. Target delivery times are established based on a dealer's distance from the plant, and GM must pay a penalty to the dealer if vehicles arrive late.

Impact with Autonomous Vehicles

Autonomous vehicles (AVs) may dramatically change the finished vehicle delivery and operating processes. For example, an AV could drive itself within the plant yard and VDC, and load itself

onto a trailer or rail car, which would significantly reduce the handling time. If the plant yard or VDC runs out of space, an AV could temporarily park itself in a nearby parking space. More importantly, an AV could drive itself to the dealer and reduce the last mile delivery costs. It could also drive itself to nearby hubs to consolidate trailer loads (for example, an AV could drive from the plant in Detroit, MI to the nearby plant in Flint, MI). In particular, an AV driving itself could be treated as a new transportation mode. This would enable more flexible logistics network operations and decrease order fulfillment time and cost. (Note: we assume that a central controller is able to send messages to AVs telling them when and where to move.)

Static Routing vs Dynamic Routing

In “Static Routing,” the shipment of vehicles to the dealer is treated as two parts: shipment through one or more VDCs, and a “last leg” from the final VDC to the dealer. The first part is static; all vehicles from the same plant to a final VDC are delivered using the same transportation route – i.e., the path and transportation modes are fixed. The second part is optimized daily, based on available inventory. For this “last leg,” if there are enough vehicles going to the same dealer they are shipped directly (Figure 2a). Otherwise, a “milk-run” (multiple stop) route is used to deliver vehicles to multiple dealers. As shown in Figure 2b, milk-run is the name given to a logistics route that visits multiple locations in a single tour.

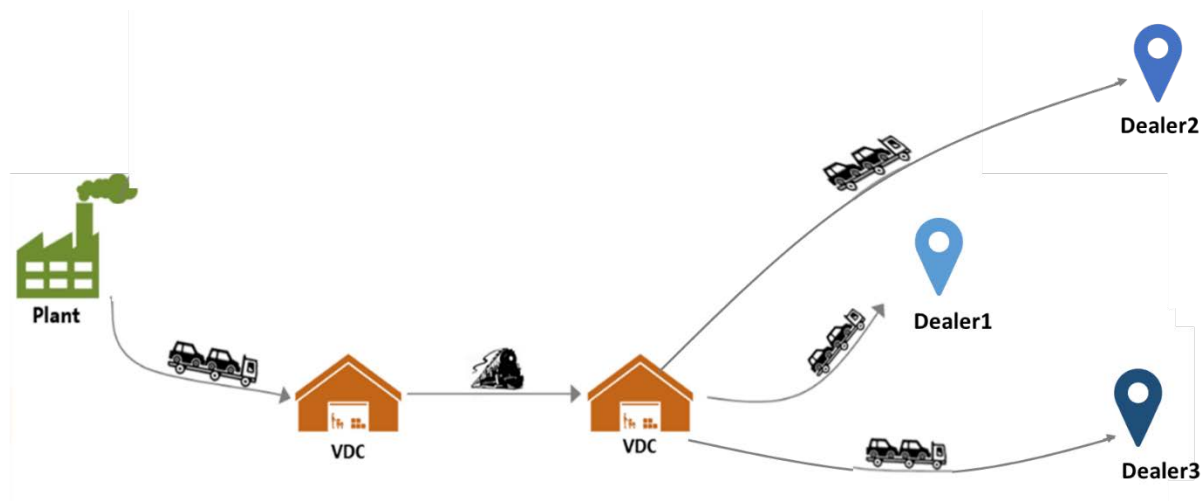


Figure 2a Static Routing Scenario 1

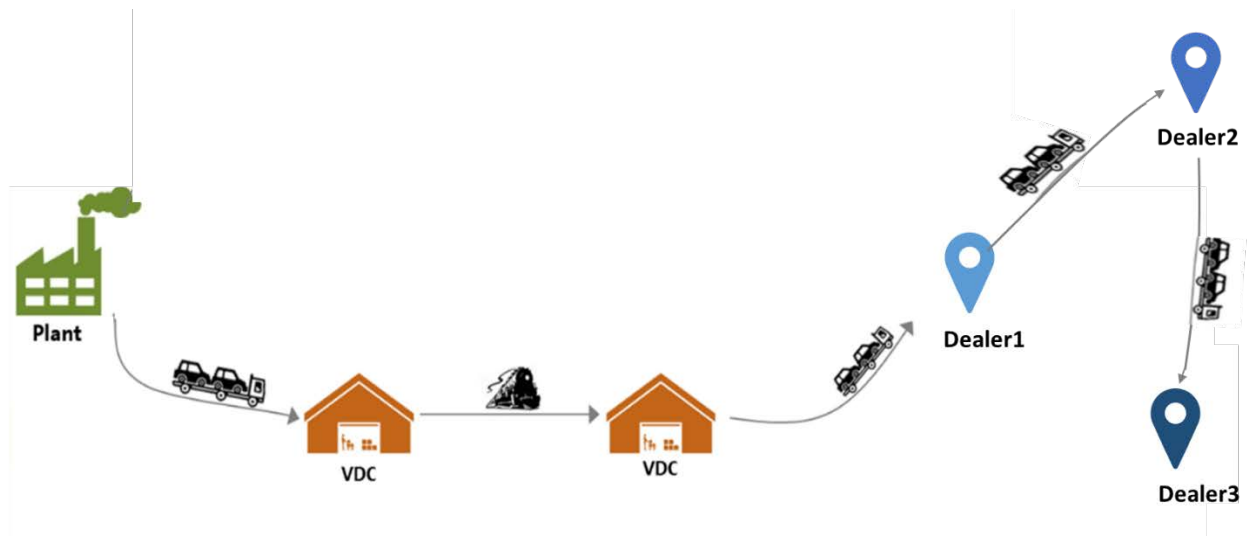


Figure 2b Static Routing Scenario 2

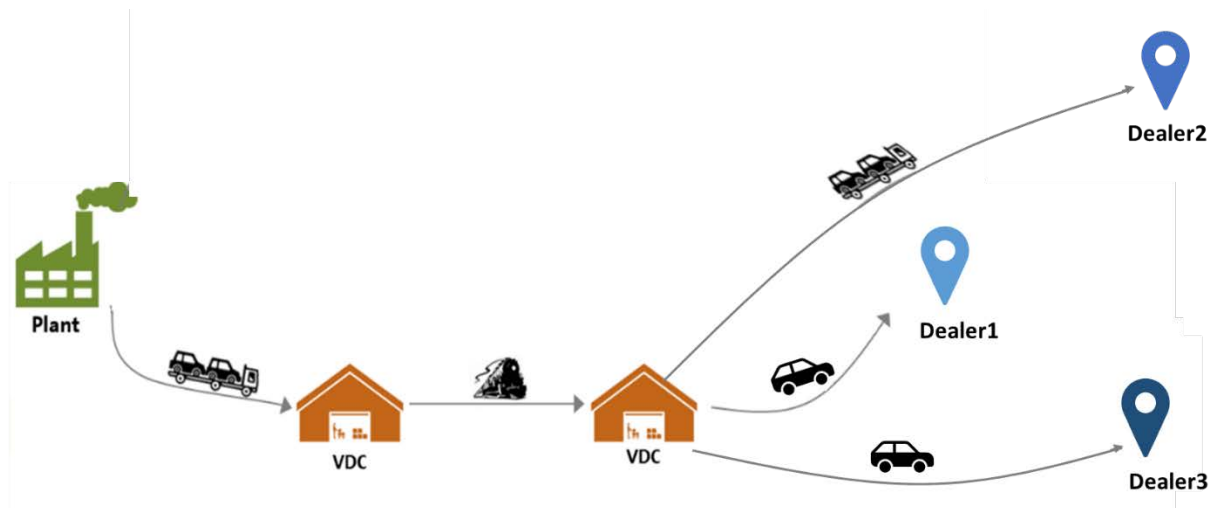


Figure 2c Static Routing Scenario 3 (with AV)

In “Static Routing”, as shown in Figure 2, the path and modes from the plant to the VDCs remain constant. Part or all of the network will only be redesigned when plant manufacturing volumes or dealer orders change dramatically -- for example, when a new vehicle model is introduced.

“Static Routing” simplifies the management of the vehicle delivery process. The vehicles are accumulated at plants and VDCs until a full load is made up to ship to the next leg in the route. To satisfy customer demand in time, some vehicles may be shipped with the truck or rail car not fully loaded. This is costly, as GM pays the same price for a shipment regardless of the number of vehicles shipped. Conversely, vehicles may be held to make up a full shipment, sacrificing lead time. Because of the fluctuation in dealer orders over time, this trade-off between lead time and logistics cost happens all the time.

What if vehicles could be shipped out with more flexible route options?

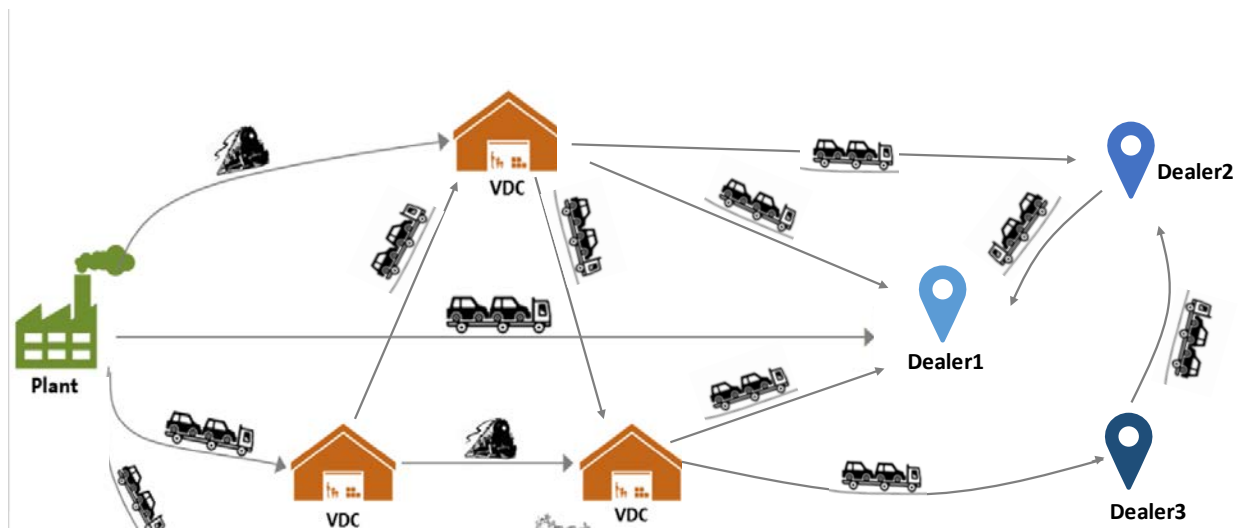


Figure 3a Dynamic Routing Without AV

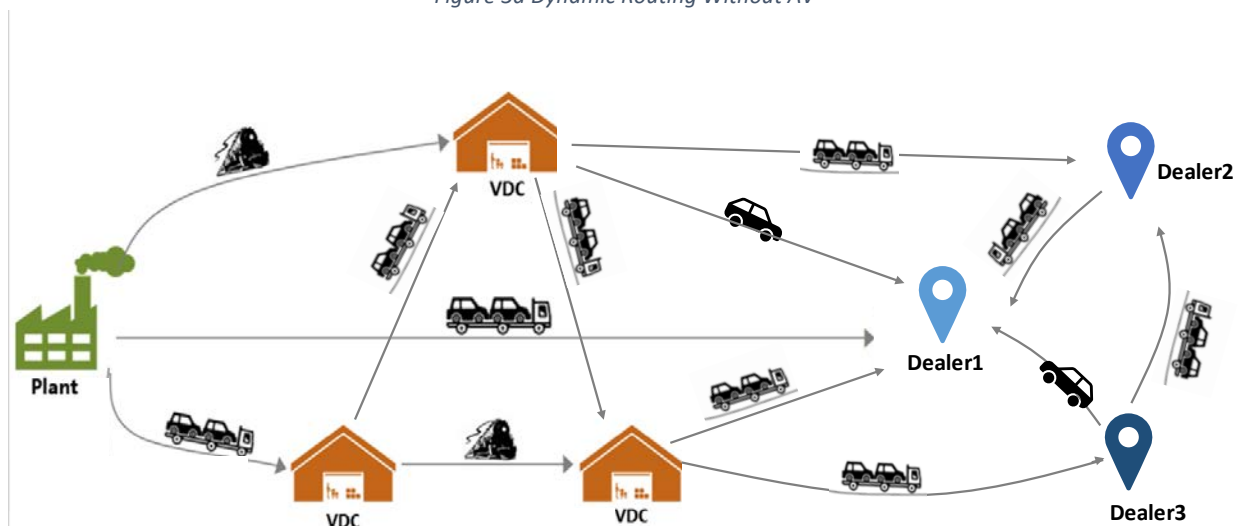


Figure 3b Dynamic Routing With AV

We call this “Dynamic Routing,” where a plant-dealer pair is no longer assigned a static route. Instead, the path and modes from a plant to a dealer may change from one load of vehicles to the next, as shown in Figure 3a.

Vehicle delivery involves resource (driver, equipment) scheduling at the plant yard, VDC, logistics carriers, dealers etc. AV delivery can significantly reduce the number of drivers in delivery: there is no need for a driver to drive an AV from the storage space to the staging area or load and unload the AV onto the trailer. In addition, an AV could park itself temporarily to an overflow lot, thus reducing the VDC space capacity limit and management needs (e.g., see

Figure 3b). With reduced human handling time and more flexible logistics network operations, “Dynamic Routing” is more plausible with AVs.

Problem Statement

This competition problem involves designing the vehicle delivery network (where to locate the VDC, and what is the capacity) and the routing from plant to the final VDC for each plant-dealer-vehicle combination. The objective is to minimize total cost (both shipping and yard operations), while satisfying the business rules, such as meeting delivery schedules.

Two key questions of concern are:

- When to utilize the AV feature in vehicle delivery?
- When to utilize the dynamic routing in vehicle delivery?

We ask you to provide answers to four scenarios:

Without AV, Static Routing	Without AV, Dynamic Routing
With AV, Static Routing	With AV, Dynamic Routing

GM will provide 2 years of data for training and 2 years of data for validation. You are not allowed to use the validation set to train the model. Your challenge is to redesign the vehicle delivery network in those four scenarios to minimize the total cost of shipping, yard operations, and late deliveries while meeting all business rules. You need to make decisions on:

- Given an existing set of VDCs, you could decide to add additional VDCs to the network with a certain location and a certain capacity. The decision of a certain VDC capacity will last for a year and will need re-examination on the expiration date. You decide the policies which trigger re-examination of the VDC options. The policies could be based on time (every year), demand change, or performance issues (VDC overflows).
- “Static Routing” scenarios: For all vehicles from a plant to a dealer, you need to decide which route to take – i.e., which VDCs to go through and which legs to use rail or truck or AVs. This decision is made every time you changed the VDC decision and valid until next time the VDC decision changes. Every day at the final VDC, for a given vehicle from a plant to a dealer, you need to decide whether to ship direct to the dealer that day, ship on a milk-run route to the dealer that day, or wait till the next day. Note: an algorithm for designing milk-run routes is provided in the References section; you may use this algorithm, or choose a different one.
- “Dynamic Routing” scenarios: Every day at each VDC, for all vehicles from a plant to a dealer, you need to decide which route to take: whether to ship out or wait till the next day. If shipped out, you also need to decide which VDCs to go through and which legs to use rail or truck or autonomous self-driving cars.
- For VDCs other than the final VDC (the last one visited before delivery to dealer), assume the vehicles are shipped out as soon as a full load is made up.

For “Dynamic Routing,” to reduce the time and effort required to manage the network, GM has found that rule-based policies or automated machine learning algorithms are preferred to continuous re-optimization. However, optimization can be used to develop policies or algorithms – for example, you could select specific points in time and optimize the network, then use the results as training data for a machine learning problem. You may want to consider any available data at the decision moment on the current and past state of the network, such as existing orders, historical order trends, and order fulfillment time in developing “Dynamic Routing” policies/algorithms.

Key Assumptions

Scope

To reduce the complexity of the problem, consider a single vehicle model sold in a single geographic region. This vehicle will be built in multiple plants. Note that each vehicle is individually configured with colors and options, and individual dealers may receive vehicles from more than one plant. Assume that these vehicles cannot be substituted (i.e., vehicles scheduled for delivery from one VDC cannot be substituted with vehicles from another); as a result, the volume provided between plant and dealer pairs must be maintained.

Transportation

Since all plants, VDCs and dealers are in a single geographic region, consider only rail, truck and AV shipping (i.e., do not consider ocean shipping). We will provide latitude and longitude for each location, distances between locations, and available transportation modes between locations. For the purpose of the competition, you may ignore geographical constraints such as mountains, rivers, lakes, and oceans when proposing new locations. Road distances involving any new locations must be estimated as 1.2* great circle distance, using the Haversine formula, as discussed in the Appendix.

Transit time may be treated as constant, based on distance and average speed. We will provide average speed by mode (truck, rail, AV). For simplicity, do not consider the variation in transit time caused by traffic, weather, scheduled pickup and delivery windows, etc.

For the competition, assume that there is sufficient capacity available for all transportation modes.

Vehicle Distribution Centers

Plant VDCs

There is a VDC located at each plant. The transportation time and cost from the plant to this VDC is negligible, and should be ignored. For each VIN, the data starts with the available time at the plant VDC, until the VIN delivers to the dealer. Assume that the plant VDC has full knowledge of shipping demand for the current and next two days (variation in production sequence makes it difficult to predict beyond two days). Assume the network has full knowledge of what happens – i.e., once vehicles are shipped from the plant VDC, the rest of the

network knows that they are in transit, when they will arrive at the next destination, etc. Note that Plant VDCs operate in the network like other VDCs, and may be used as intermediate nodes.

Stand-alone VDCs

All VDCs serve as cross-docks for logistics legs that use different transportation modes, or that require consolidation. For example, a VDC may receive vehicles by truck and ship them out by rail, or receive by rail and ship them out by truck. Stand-alone VDCs are any VDCs not associated with a plant.

Flexible VDCs

Flexible VDCs are only possible with AVs, since the vehicles could drive themselves to a parking lot and load themselves onto a trailer. If a flexible VDC is right next to an existing VDC, it is functioning as an overflow site except that the shuttle fee is ignored with AV.

GM pays yearly to lease a fixed capacity at each VDC. In addition, we pay a per-vehicle handling cost at VDCs. If volume at any time exceeds the leased capacity, vehicles are shuttled to an overflow site. We assume the overflow space is available for any VDC, however we must pay a fixed cost to shuttle the vehicles, and a storage cost per vehicle per day.

Inside the VDC, vehicles are parked such that any individual vehicle can be accessed without having to move other vehicles. This minimizes the labor required to store and retrieve vehicles, but a significant amount of space is dedicated to aisleways. Vehicles could be stored more densely (by double parking, triple parking, etc.). However, this would impact the labor required and hence the handling cost per vehicle. Note that AVs that can drive themselves inside the VDC may affect this tradeoff. Teams may change some of the provided costs and capacities based on changes they propose in operating the network, however any changes must be clearly documented and defended. The judges will evaluate these changes, and may reject them or modify their impact.

Time

We ignore all holidays and weekends; the distribution network tends to operate 24 hours a day, 7 days a week, 365 days a year. Assume that daily decisions are made at 12:00 am (midnight), with perfect information for that day, and that vehicles are available to be shipped on the same day that they arrive -- in other words, ignore the dynamics of what happens within an individual day. All timestamps in the data are based on the same time zone. For simplicity, assume that all locations are in the same time zone.

General

As stated above, the flow of vehicles from plant to dealer must be maintained. The location of existing sites is fixed, and cannot be changed (although existing VDCs could be closed, and/or new ones added). You may (optionally) modify other constraints, costs, or other parameters. If you choose to do so, you must clearly describe any changes made, and provide a detailed

justification in the final report. The judging committee will evaluate these changes, and choose whether to accept them. If the judges choose not to accept them, the final design will be evaluated using the original assumptions. An example of a change with the expected level of explanation is provided in the Appendix.

Datasets

We will use data for a single vehicle model (e.g. pickup truck). You will need to deal with the data as is (there may be some data quality issues). Note that the actual data may include details that occasionally violate the rules provided – for example, the “static” routing of vehicles from Plant A to Dealer B goes through VDC 1, but ~2% of vehicles are routed differently. These are operational exceptions that occur due to day-to-day variation, and are not intended as hidden “tricks,” and are representative of the type of data issues that occur in real world applications. It is up to individual teams to decide how to handle any “bad” data. Some data will be transformed to protect sensitive dealer information.

1. Historical shipment transactions.

VIN	Plant	Location	Lat	Long	Arrival_Time	Depart_Time	Transp_Mode_CD
000001EB-8395-498A-9134-E550B7521276	3A	3A	32.48	181.64	3/30/14 15:33	3/30/14 15:34	R
000001EB-8395-498A-9134-E550B7521276	3A	3A	32.45	181.65	3/31/14 14:00	4/1/14 21:00	R
000001EB-8395-498A-9134-E550B7521276	3A	UL	20.76	167.56	5/8/14 0:05	5/14/14 13:28	T
000001EB-8395-498A-9134-E550B7521276	3A	2EADCB1-5E65-4273-9549-317F629D871F	19.93	162.75	5/15/14 13:15	5/15/14 13:15	T
00000372-3BAE-4512-910D-55CAD1D5509A	FF	FF	19.57	170.72	8/1/14 22:09	8/8/14 8:56	T
00000372-3BAE-4512-910D-55CAD1D5509A	FF	0BF273E4-FF96-45C6-B37E-D248A963D232	21.35	168.55	8/8/14 16:03	8/8/14 16:03	T

- For each vehicle, this data will show the time of all events: release to carriers, delivery to dealer etc. You could calculate how long a vehicle stays in the plant yard, VDC; how long is transit between locations etc. You may use this to make general assumptions, for example, the distribution of vehicle dwell time in the network.
- This data shows how a vehicle delivery network is operated: which leg is truck, which leg is rail, and how different modes are connected.
- This data shows how many orders each dealer makes over time, how many vehicles a plant builds over time, how many vehicles are held at the plant yards or VDCs at any time.

2. Geographic locations of the plants and dealer demand areas (geographical group of dealers).

Location	Lat	Long
VDC		
Plant		
Dealer		

3. VDC locations and cost models. The following information will be provided:

VDC	Fixed Cost	Annual Cost/VehCapacity	Handling Cost/Veh (Non AV)	Handling Cost/Veh (AV)	OverflowVariable Cost/Veh/Day	Overflow Shuttle Cost/Veh (Non AV Only)	Rail Available
Existing Cross Docking VDC		\$ 730	\$ 50	\$ 40	\$ 4	\$ 30	Varies by Location
Potential Cross Docking VDC	\$ 1,000,000	\$ 730	\$ 50	\$ 40	\$ 4	\$ 30	YES
Flexible VDC (AV)				\$ 40	\$ 4		NO

Location	Capacity Upbound	Rail Available
RO	3,464	TRUE
3A	1,878	TRUE
FF	1,531	TRUE
MN	1,097	TRUE
NZ	646	TRUE
LM	536	TRUE
DZ	482	FALSE
DI	461	FALSE

4. Shipping + VDC + Penalty cost (historical) by VIN.

VIN	Total \$
1	\$ 900
2	\$ 1,201

5. Logistics (example data only; use numbers provided in data files for actual costs)

Mode	Equivalent Speed (mile/hour)	Load Factor (max veh/load)	Fixed Cost (\$/load)	Variable Cost (\$/mile/load)
AV	50	1		0.55
Truck	50	10	200	4
Rail	20	20	2000	3

- Assume static load factor (i.e., capacity of shipping equipment).
- Fixed cost and cost per mile by transportation mode (including AV).
- Target delivery times model based on plant-dealer distance.

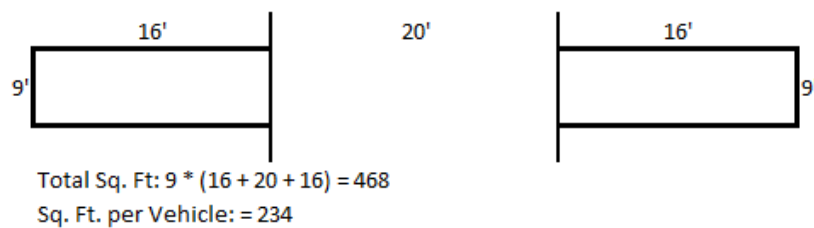
6. Business Rules include

- An AV cannot accumulate more than 300 miles before delivery to dealers and still be sold as new.
- Each demand area must receive 100% of its demand.
- Manufacturing day to delivery day must be within a time limit. GM must pay a penalty to the dealer for each day a vehicle exceeds the time limit.

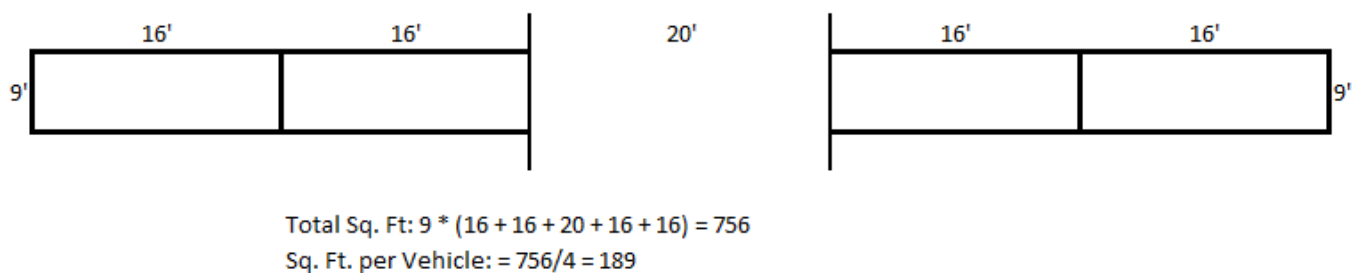
Output Format

The judging committee will evaluate the results based on orders from the validation data. The committee will verify that the business rules are met, for example, all vehicles ordered are delivered, and vehicle arrivals into the network each day are maintained. The final report must clearly list any changes to costs, business rules, or assumptions that were provided, and provide an explanation. An example is shown below. Note that this example is provided to show the level of explanation expected, but student teams are encouraged to explore alternatives.

Example: VDC X currently holds up to 2,000 singly parked vehicles with a 20' access aisle between rows of cars:



By parking autonomous vehicles 2 deep, the same VDC could hold $2,000 / (189 / 234) = 2,476$ vehicles, as shown below.



The final report should also include a description of the computing platform used, and the running time of each component. The report should discuss the performance of the algorithm, including both memory requirements and time, including the time to train the model, if a training algorithm is used.

An important aspect of performance analysis is to understand how algorithm performance varies with parameters such as problem size (# of vehicle, # of plant, # of dealer, # of available VDC), processor count, and message startup cost. Since GM's actual network operates on many different vehicle models, and a larger number of plants, we are particularly interested in how the problem scales. In addition, we are interested in the applicability of parallelization, that is, how effectively can an increased number of processors be used for larger problems.

There is no specific limit on the time to train the model, or decision time, however the decision time must be appropriate for the frequency that it is used. For example, if the recommended approach is to run an optimization algorithm weekly, a run time of several hours is acceptable; however, if the algorithm is to be run every day, the run time should be less than 1 hour. In general, we prefer faster running time if the solution quality is similar.

You are encouraged to discuss the space complexity as well, the amount of working storage an algorithm needs (how much memory, in the worst case, is needed at any point in the algorithm). As with time complexity, we're mostly concerned with how the space needs grow, in big-O terms, as the size N of the input problem grows.

The teams should report the following solution data as an output for all four scenarios and for both the training and validation data set. All data should be submitted as comma separated text files, with double quotes used for text fields. Refer to sampleoutput.xlsx file for clarification.

1. Network Design Details (File Name: NetworkDesign.csv)

VDC	Lat	Long	Capacity	Total Cost (\$)	From Time	To Time

2. Routing Details File Name: RoutingDetails.csv)

VIN	LOC	Arrive Time	Depart Time	Depart Mode
1	3A	6/16/2017 1:47	6/16/2017 1:48	Truck
1	RR	6/16/2017 10:10	6/22/2017 19:04	Rail
1	CE	6/23/2017 0:05	6/27/2017 0:05	AV
1	347	6/27/2017 7:30		

3. The Lead Time for each vehicle (File Name: Leadtime.csv)

VIN	Lead Time

4. Shipping Details (File Name: ShipDetails.csv)

From LOC	To LOC	Transit Distance (Mile)	Mode	Load	Total Cost (\$)	Depart Time	Transit Time
			Truck				
			Rail				
			AV				

Your calculation of transit distance should be consistent with the Great Circle Mile method described in Key Assumptions section.

Evaluation Criteria

The solution will be the evaluated based on the following network characterizations on validation set:

- The total cost of running the network including logistics, distribution center, penalty cost
- Distribution of the lead time between the order time and delivery time

Additional credit will be given for innovative solutions when it comes to the use of AVs in the vehicle delivery network design area, as well as suggestions for utilizing data dynamically to determine the routing considering existing orders, historical order trends etc.

Entries will be judged on use of the full analytics process – incorporating not only the technical analysis and solution, but also understanding of the business problem, team organization, and the clarity and effectiveness of communication. Think of this report as a presentation to GM's upper management on the critical elements in the analytics decision process.

Entries will be judged by the clarity of the solution, the technical strength of the methodology, the uniqueness of the approach, the degree to which the data support your conclusions, and written presentation.

Appendix

Abbreviations

General Motors (GM)

Autonomous vehicle (AV)

Vehicle distribution centers (VDC)

References

To understand better about vehicle logistics networks, please check:

- Solving a Network Optimization example:
<https://www.coursera.org/lecture/wharton-operations-analytics/network-optimization-example-3awtd>
- GM demand and supply news:
<https://automotivelogistics.media/intelligence/general-motors>
- Network Optimization for vehicle delivery
Eskigun, E, R. Uzsoy, P.V. Preckel, G. Beaujon, S. Krishnan and J.D. Tew, Outbound Supply Chain Network design with Mode Selection, Lead Times and Capacitated Vehicle Distribution Centers, European Journal of Operational Research, Vol. 164, pg. 182-206, 2005.
<https://www.sciencedirect.com/science/article/pii/S0377221704000633>
- Vehicle delivery
https://www.researchgate.net/publication/247835472_Automotive_distribution_network_design_A_support_system_for_transportation_infrastructure_decision_makers
- Milkrun Design Algorithm

Gillet B.E., Miller L.E., Johnson J.G. (1979) Vehicle dispatching — Sweep algorithm and extensions. In: Ritzman L.P., Krajewski L.J., Berry W.L., Goodman S.H., Hardy S.T., Vitt L.D. (eds) Disaggregation. Springer, Dordrecht

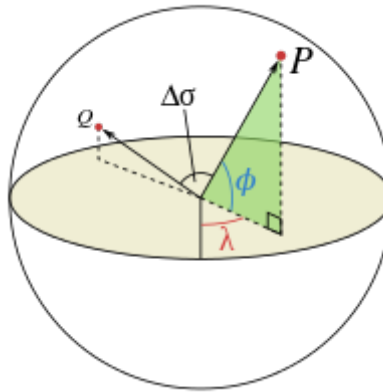
The Sweep Algorithm divides the problem into two subproblems: one of assigning locations to routes, and the other of minimizing the length of each route using a traveling salesman algorithm.

More specifically, the Forward Sweep Algorithm orders the locations in ascending order according to the size of their polar coordinate angle with respect to the terminal and then starts forming routes by collecting locations while moving around the terminal in a sweeping fashion. The first route consists of the location with the smallest angle, the location with the second smallest angle, etc. up through the location added just before the maximum load of the vehicle is exceeded. At this point a traveling salesman algorithm (Lin's 3-opt) is used to determine the optimal path to visit the locations in the first route. If the maximum distance capacity of the vehicle is exceeded, the last location added is removed and the traveling salesman algorithm is executed again. This process is repeated until a feasible route is found.

- Great Circle Distance

https://en.wikipedia.org/wiki/Great-circle_distance

The arc length between two points on a sphere (P and Q as shown below) can be calculated as $d = r\Delta\sigma$, where r is the radius (use 3,959 miles for the radius of the Earth), and $\Delta\sigma$ is the central angle (in radians).



The central angle, $\Delta\sigma$, can be calculated using the Haversine formula as:

$$\Delta\sigma = 2\arcsin \sqrt{\sin^2\left(\frac{\Delta\phi}{2}\right) + \cos\phi_1 \cdot \cos\phi_2 \cdot \sin^2\left(\frac{\Delta\lambda}{2}\right)}$$

Where ϕ_1, λ_1 and ϕ_2, λ_2 are the latitude and longitude of two points in radians, and $\Delta\phi, \Delta\lambda$ are their absolute differences.

We have found that great circle distance plus ~20% is a good approximation for over the road distance, and request that teams use this method for calculating distances to ensure consistency. Below are example great circle and road distances between four locations:

			Great Circle Distance (miles)				Road Distance (miles) (Google)				Ratio			
			Washington Monument	Grand Canyon	Empire State Building	Disney World	Washington Monument	Grand Canyon	Empire State Building	Disney World	Washington Monument	Grand Canyon	Empire State Building	Disney World
Location	Latitude	Longitude												
Washington Monument	38.889484	-77.035279	-	1,921	207	771	-	2,291	230	867	-	1.19	1.11	1.12
Grand Canyon	36.106965	-112.112997		-	2,072	1,854		-	2,403	2,151		-	1.16	1.16
Empire State Building	40.748441	-73.985664			-	956			-	1,102			-	1.15
Disney World	28.385233	-81.563874				-				-				-