

Plowing for a Cornell Crowd

Cornell Mathematical Contest in Modeling, 2025

Logan Abramowitz (la437)

Maya Glick (mig46)

Eric Sullivan (ebs226)

November 9, 2025

Contents

1	Introduction	3
2	Model Assumptions	3
3	Parameter Values & Justification	3
4	Optimizing Plow Zones	5
5	Simulation Algorithm	6
6	Model Predictions	7
7	Validation & Robustness Testing	10
8	Strengths & Weaknesses	10
9	Future Work	11
10	Conclusion	11
	References	13
11	Appendix	14
11.1	Executive Summary	14
11.2	Source Code	14

Abstract

A major snowstorm is forecasted to hit Ithaca on the morning of Cornell’s Winter Graduation. Ithaca’s Department of Public Works needs to efficiently clear roads while ensuring emergency services access and allowing visitors to travel through Ithaca safely. Our goal is to determine the minimum number of snowplows required to keep roads safe and accessible across Ithaca and Cornell’s campus. To do this, we developed two complementary models: an optimization model that divides Ithaca’s road network into plow zones using K-Means clustering weighted by slope and traffic data, and a simulation model that builds and visualizes full plow routes using a weighted NetworkX graph that accounts for directionality, turn penalties, and road priorities. Across multiple analyses, both models consistently identified eight plows as the optimal number to balance total coverage time, fairness among drivers, and robustness to uncertainty. With eight plows, Ithaca’s roads can be fully cleared within roughly two hours per pass, leaving four additional plows available for backup or continued storms. Together, our models provide a realistic, data-driven approach for the DPW to plan plow deployment that prioritizes both efficiency and safety on one of Ithaca’s busiest winter days.

1 Introduction

Ithaca's Department of Public Works (DPW) is responsible for maintaining approximately 65 miles of roads. During snowstorms, they deploy truck-mounted snow plows and salt spreaders to keep the roads safe for civilians. The DPW's plow deployment routes prioritize clearing arterial roads, steep hills, and emergency access routes first, and then clearing smaller residential neighborhoods.

A 12-hour snowstorm is forecasted in Ithaca from 4:00AM to 4:00PM on the day of Cornell's Winter Graduation. Thus, in addition to the original priorities, the snow plow routes must also focus on clearing campus roads and key access points to ensure visitors can get to graduation safely and efficiently. While the DPW has 12 snow plows available, the goal is to minimize the number of plows deployed at once in case the snowfall continues in the following days, as plow drivers have time restrictions.

The challenge is to determine the minimum number of plows needed while still ensuring that all critical roads are cleared, travel around Ithaca remains safe, and visitors can get to and from graduation. Factors influencing the optimal routing and deployment plan include Ithaca's geography (especially road steepness and traffic patterns), parking rules (such odd/even parking), access routes from downtown hotels and parking lots to Cornell's campus, snowfall rate and temperature (which affect salt use), equity considerations for residents living in smaller neighborhoods, and over model robustness (e.g. accounting for continued snowfall or plow breakdowns).

2 Model Assumptions

We made several assumptions about the snowstorm and Ithaca's geography that inform the model. First, the DPW is only responsible for plowing public and university roads, not privately-owned roads. Moreover, due to Ithaca's borders, there are a small number of roads that are disconnected from the larger grid of Ithaca's roads, which will not be considered in the plow routes. For creating optimized plow routes during the snowstorm, the model assumes that there was no snow on the ground prior to 4:00AM on the day of Cornell's graduation. Additionally, we assume that certain areas in Ithaca will have higher traffic because of graduation (and therefore will be a higher priority to clear of snow). These include Cornell's campus, key access roads to campus, Collegetown neighborhood, and the Ithaca Commons (both areas with high volumes of hotels and restaurants). In terms of Ithaca's parking laws, because the storm only lasts within one 24-hour period, we assume that all of the cars parked on roads will stay on either the even or odd side of the road for the entire duration of the storm. Lastly, based on data about snowstorm temperatures historically in Ithaca during December, we assume that the temperature during the snowstorm is below 20°F, so using salt would be ineffective in mitigating ice-formation and snow melt.

3 Parameter Values & Justification

There are several parameters used within the model to inform the number of plow passes needed to effectively cover all of Ithaca's roads and the optimal routes. Based on the prompt guidelines, we assume that the average speed of each plow is 25 mph,

so for simplicity, every plow in our model runs at 25 mph. In the simulation, we assigned a penalty value of 1000 to u-turns, which helped discourage u-turns in the optimal routes. Research also indicates that a typical snow plow can remove 6 inches of snow per pass. To determine the total snowfall of the winter graduation snowstorm, we looked at data from a major snowstorm that occurred in Ithaca in 2010, which showed that there was a total of 18-24 inches of snow. Thus, taking the conservative value, we assume the winter graduation snowstorm will have 24 inches of snow. Therefore, it will take $24 / 6 = 4$ iterations of each plow route to fully clear Ithaca's roads during the duration of the storm.

Moreover, we determined road gradient (steepness) to be a major priority because Ithaca's steep hills can be incredibly dangerous when icy. In order to construct a priority weight for roads, we loaded the coordinate bounds of Ithaca's roads and pulled a 10 meter resolution elevation tile for these roads. We then calculated the slope of each road via euclidean distance to get a single parameter. Euclidean distance of the endpoints does limit the accuracy as some roads have curves, but we notice that this inaccuracy causes an overestimation of the danger of each curvy road, and since curvy roads are still dangerous on their own, we are satisfied with the approach. We grouped roads with similar parameters into 4 groups for gradient: < 5%, 5% – 10%, 10% – 15%, > 15%. Our results are graphed below:

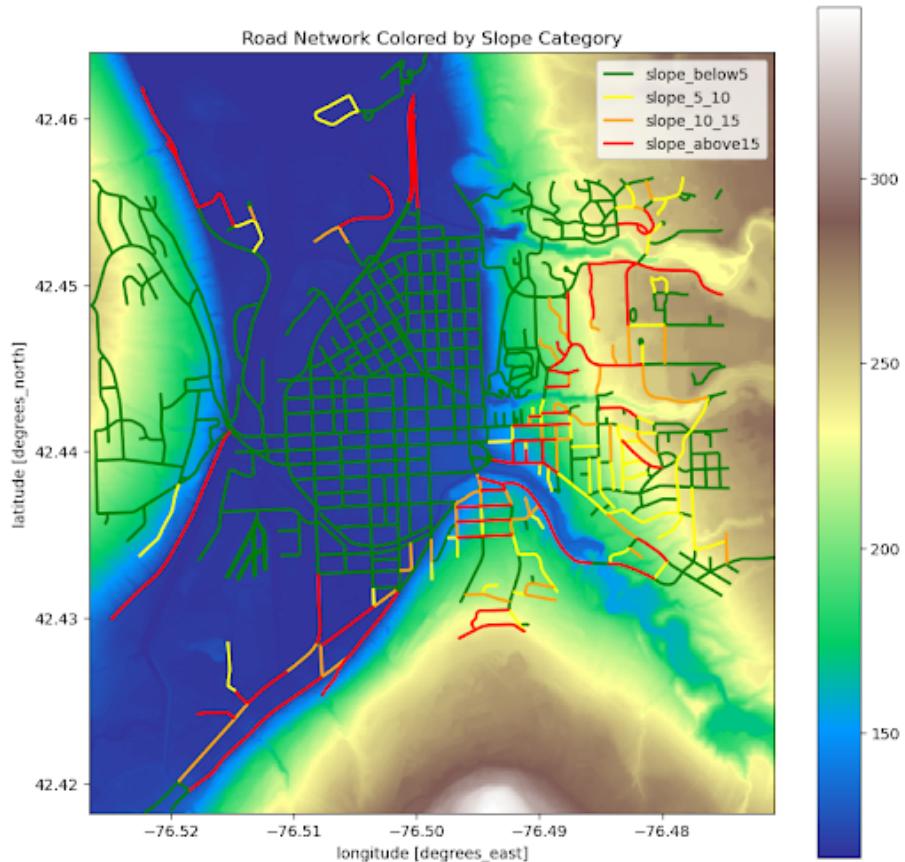


Figure 1: Road map of Ithaca overlaid on the surrounding elevation tile. Roads are color-coded based on slope.

4 Optimizing Plow Zones

To inform the number of zones for our simulation analysis, we needed to determine the minimal number of plows we should deploy to achieve fast and balanced coverage of Ithaca's roadmap. To conduct this analysis, we used a road network graph built from a GIS shapefile of Ithaca's roads, a GeoDataFrame that contains traffic data (Average Annual Daily Traffic or AADT) and a slope priority score (1 - 4) for each road. Thus, each road segment carries a combined priority weight proportional to: Priority = AADT*Slope Score.

From this data, the algorithm tests a range of possible k clusters 3 - 14 (even though there are only 12 plows, we wanted to see how more would impact the results). For each k, our algorithm extracts all road-network nodes, applies K-Means clustering to partition those nodes into k groups representing potential vehicle zones, and then assigns each road segment (edge) to a cluster based on the cluster of its starting node. This process is meant to mimic what would happen if the city were divided into k operational zones for plows or maintenance crews.

For each cluster, the program computes a zone workload score:

$$\text{ZoneCost}_i = \sum_{e \in \text{zone}_i} \text{Length}(e) \times \text{Priority}(e)$$

According to this equation, longer roads require more time to plow and higher traffic or slope priority makes those roads more demanding. There is a ZoneCost for every vehicle, reflecting total "plowing effort" for its assigned area.

Once all the zones have been evaluated, the algorithm extracts three key metrics: max zone cost, mean zone cost, max time, and balance ratio. Max zone cost is the largest workload among all zones (slowest vehicle) and our main goal is to minimize this as a low max zone cost indicates no single plow is overloaded. The balance ratio is the ZoneCost's standard deviation \div mean workload, which overall represents fairness. A low balance ratio means workloads are evenly distributed across vehicles. Together, max zone cost and balance ratio show both efficiency and equity of resource allocation.

Max time is a single value we use to estimate the time taken to clear all of Ithaca (including multiple passes) in one go. The goal of this is to see if a certain number of plows can then be sustainable during different snowfall rates. Since we assume plow drivers all start work on their zone in relatively the same time, then we know the work will be done when the last driver finishes, so the time to clear Ithaca is the time of the final driver. To model an estimate of time we considered how drivers would handle multi pass lanes while still prioritizing emergency service routes. Ithaca DPW states that the first priority in clearing roads is to get a minimal service lane. We model this as one pass on all edges within a cluster before we branch out to finish the rest of the passes. The total time for this is then: $T(Z_i) = \sum_{e \in Z_i} L_e \cdot \frac{1}{V}$

Where we adopt the shorthand L_e for length of edge e, T_Z as the total time for each zone, and V as the average velocity of a snow plow. This is only the time for one pass, so to incorporate all passes we add on the summation for remaining passes needed per lane. We exclude parking to be conservative in the case that a road happens to have no cars on it. For an even more conservative estimation, we decrease the suggested speed of a snow plow from 25 mph to 15 mph to get a strong upper bound on time. Incorporating all the bases, using width to determine number of passes per road gives:

$$\begin{aligned}
T(Z_i) &= \sum_{e \in Z_i} L_e \cdot \frac{1}{V} + \sum_{e \in Z_i} \max \left(0, \frac{W_e}{W_{\text{plow}}} - 1 \right) L_e \cdot \frac{1}{V} \\
&= \sum_{e \in Z_i} \left(1 + \max \left(0, \frac{W_e}{W_{\text{plow}}} - 1 \right) \right) L_e \cdot \frac{1}{V}
\end{aligned}$$

Where W_{plow} is a constant for the width of a standard plow (about ten feet). We also notice that $\frac{W_{\text{edge}}}{W_{\text{plow}}}$ will be rounded up to account for passes that do not require the full length of the plow but are necessary to clear the roads. We see then that the time to clear Ithaca is $\max_i(T(Z_i))$. The optimization problem we then seek to solve is

$$\min_{Z_i} \left(\max_i(T(Z_i)) \right), \quad Z \subseteq G,$$

where we maximize over all zones Z_i , and we minimize over the possible zone structures in our graph G . Since this problem involves the NP-hard tactic of constrained graph optimization, which is not feasible given time limits for a ~ 7000 node graph, we explore a simpler yet meaningful simulation outlined below, that still achieves a strong feasible solution.

5 Simulation Algorithm

Our simulation algorithm builds a multi-vehicle and coverage simulation for the city of Ithaca based on a GIS shapefile. The goal of the algorithm is to simulate snow-plows covering all roads in a given zone efficiently, subject to real-world constraints like one-way streets, turn penalties, slopes, and traffic intensity. The algorithm converts raw ArcGIS road geometries into a weighted NetworkX graph where nodes are intersections and endpoints and edges represent road segments. Each edge has metadata attached to it: travel time (length of road / 25mph), directional constraints, slope base priority weights, and traffic weights.

The first main stage of the algorithm is loading and preprocessing the data to create the graph. The shapefile's data is first read by GeoPandas. Then each road segment is decomposed into pairs of connected coordinates and direction attributes control whether or not edges are bidirectional. Next, each edge is weighted by an estimated travel time based on its physical length and the average speed of the plows (25 mph). Lastly, additional weights are added to each edge based on the slope priorities and the traffic data; steeper roads and higher traffic roads are prioritized.

The next stage of the algorithm divides the road network into clusters based on the number of plows allocated. Specifically, it uses k-means clustering to partition the road network into territories for each vehicle. Each cluster represents a rough “service zone” but the boundaries are soft so that plows can cross the borders when necessary. We also use frontier detection to identify edges near the cluster borders to allow for shared coverage where roads interconnect.

In the third stage, we begin route-building, one cluster at a time. Within each cluster, a greedy coverage heuristic (mainly using Dijkstra) generates a path over all roads. In this heuristic, u-turns are heavily penalized and sharp turns add small penalties to encourage smoother routes. Overall, edge traversal cost combines travel time, angular penalty, and also a repetition penalty to discourage plowing an already cleared road. Once all the routes are done, if there are still uncovered (unplowed)

roads, we do a “coverage cleanup” where nearby plows cover those roads to ensure complete network coverage.

To extract useful information from this simulation, we export the algorithmically generated routes. However, the algorithm uses the latitude and longitude coordinates of each node to generate the routes, so those routes don’t provide much value (those routes would provide a sequence of latitude-longitude coordinates). To solve this problem, we attached the name of each road to each road segment during graph formation so that we can follow the final route and translate the sequence of coordinates into human readable instructions. We also were able to calculate the angle of each turn in the route, so we could include whether a turn is a left, right, or u-turn.

Finally, to be able to visually understand the output of our algorithm, we plot the routes over a map of the Ithaca road network where each route is highlighted in a different color. We save this picture as a png file to be viewed later. To understand how the route works for each plow, we created an animation that displays each plow moving along its route where each plow is represented by a different colored circle and leaves a trail of its respective color behind showing its path. Once a plow completes its route, it stops to signify completion of its task.

6 Model Predictions

The optimization of zones of the Ithaca road map produced Figure 2, which shows the number of vehicles vs max zone cost and number of vehicles vs balance ratio. We can see that the max zone cost pretty steadily decreases as the number of plows increases, which is expected behavior as more plows means the work can be better distributed. However, when looking at the balance ratio, we can see that there is a significant dip at 7 and 8 plows, meaning that using 7 or 8 plows is more fair to plow drivers and smaller neighborhoods than their neighboring number of plows. Therefore, if we want to minimize the number of plows used, then we would recommend the use of 8 plows at a given time, as it gives low max zone cost and balance ratio, meaning that it is both optimal and fair.

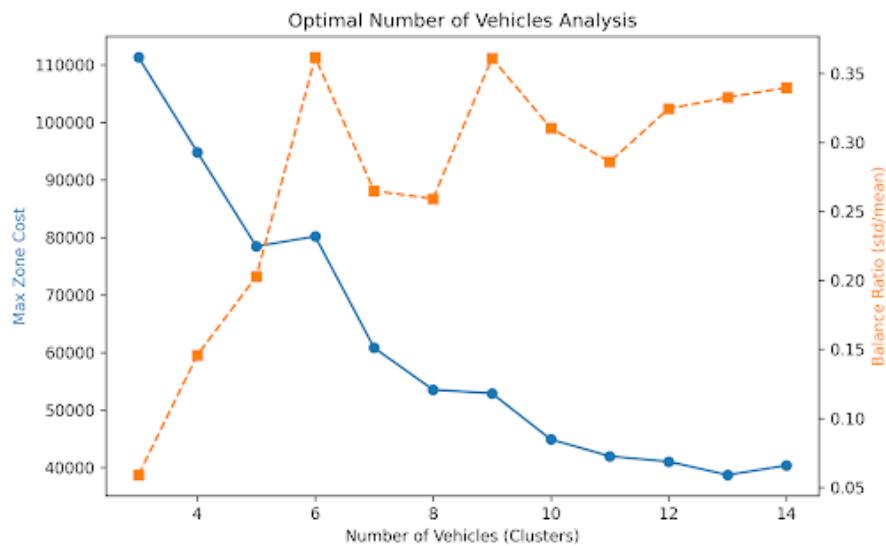


Figure 2: Optimal Number of Vehicles Analysis with Max Zone Cost and Balance Ratio.

To reinforce our previous analysis’ results, we ran a similar algorithm, but this

time compared the number of plows to the maximum time it would take one of them to clear their cluster, and thus the time it would take for all of Ithaca to be cleared. It can be seen in Figure 3 that once again 8 plows seems to be a very good choice, as the total time it would take to clear Ithaca is less than if there were 7 or 9 plows. Of course, 10 plows or more would also be effective, but we want to minimize the number of plows. Thus, we hold firm that 8 plows is the optimal number to use. Based on Figure 3, it would take about 1.85 hours to clear all of Ithaca with 8 plows, which is fast enough to outpace the snowstorm - we need 4 passes ($24/6$) to clear all the snow, which would take 7.4 hours. This means that starting at 4:00AM, we would finish plowing by 11:30AM and have all roads fully clear in time for graduation.

Maximum Times to clear all of Ithaca with k Plows:

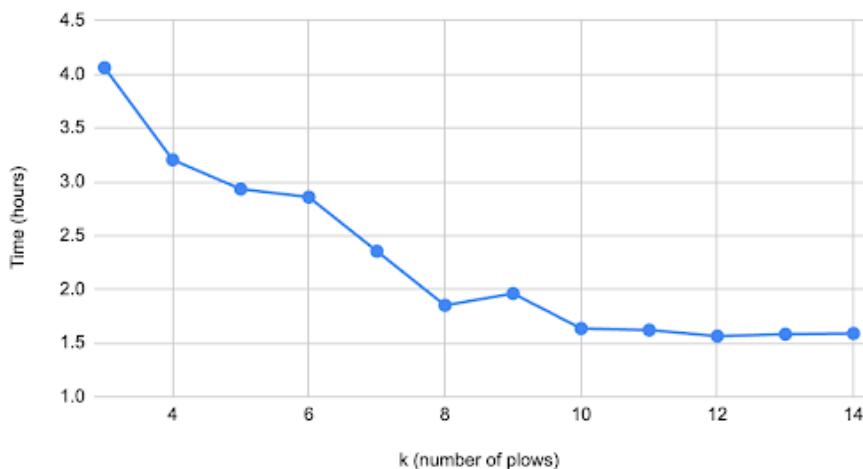


Figure 3: Maximum times (hours) it would take to clear all of Ithaca with k plows, with k ranging 3 - 14

With the optimal number of clusters, we ran the simulation and divided Ithaca into 8 fairly even zones. The simulation provided the graph seen in Figure 4, which shows the 8 distinct zones that should be covered by the drivers. It can be seen that there is some overlap between zones which is meant to promote smoother routes. While this is a good baseline, the drivers would still be able to contribute to other zones in an effort to clear Ithaca as fast as possible.

8-Vehicle Coverage (v7 frontier-aware)



Figure 4: 8-Plow Zone Coverage of Ithaca

The combined results of these two algorithms provide a planning framework where the simulation algorithm predicts how plows traverse road networks and how long full coverage will take for a given number of plows. Overall, the models form a prediction system where the optimizer suggests an ideal number of plows and the simulator confirms the resulting performance through route-based execution. This dual-model approach will allow the DPW to visualize trade-offs between efficiency and fairness. Below, Figure 5 shows a snapshot of the simulation in progress through the animation.



Figure 5: Snapshot of 8-Plows Coverage Simulation

7 Validation & Robustness Testing

We ran several analyses of different cluster sizes, with the number of zones ranging from 3 to 14 (again, even though there are only 12 plows, we wanted to see how more would impact the results). This allowed us to visualize the optimality of different clusters, and validated that $k=8$ gives optimal and fair zones. Additionally, because we are only using 8 out of 12 plows, this contributes to the robustness of our model, as there are 4 remaining plows that can be used if one breaks down or the snow continues beyond the expected end time. Overall, we believe our optimal value of $k=8$ is sound, as all of our analyses showed 8 as a cost-effective and fair number. Moreover, our calculation of time is fairly robust. We round up the number of passes a plow must take per road based on road width and plow width, and we bring the speed down to 15mph to account for any shock factors slowing down progress, as well as turns and harsh storm conditions. This gives a relaxed upper bound on the maximum time it will take to traverse all of Ithaca's roadways.

8 Strengths & Weaknesses

Our optimal plow zone model has several strengths that inform the model with real-world scenarios. Including road-specific traffic and steepness data in the priority score and doing a comparison of fairness vs the ZoneCost parameter help ensure

that the zones are partitioned in an equitable way (so no plow driver has all of the highest priority roads in their zone). Moreover, our model analyzes many different potential zones which helps ensure optimal clustering. Some weaknesses include that the model doesn't consider roads in Ithaca that are disconnected from the larger grid (only a small number roads) and that the priority score doesn't take into account traffic data specifically during a snowstorm or graduation, but rather the average annual data. Additionally, our simulation exhibits several notable strengths as well as a few limitations that define its effectiveness as a planning and visualization tool. One of its greatest strengths is its ability to visualize multi-vehicle snowplow coverage across the city's road network. It not only partitions the map into distinct operational zones using a K-Means clustering algorithm, but also provides animated route playback, allowing users to observe how plows traverse their territories. Another strength lies in its edge weighting system, which incorporates both traffic volume (AADT) and slope-based priorities to influence routing decisions. This ensures that steeper or higher-priority roads are serviced earlier, reflecting realistic operational constraints. Furthermore, because the algorithm enforces complete network coverage, every road segment receives attention, meaning that after a full iteration, all streets are guaranteed to have a cleared or traversable emergency lane. The main weakness of the current implementation is that it can occasionally exhibit local looping behavior, where a plow repeatedly covers the same short segment even when nearby roads remain unvisited. This issue arises from local cost-optimization behavior inherent to the greedy traversal heuristic, however, these loops tend to be limited in scope and duration.

9 Future Work

One area of future work is within the simulation algorithm as it sometimes produced repetitive loops near borders or visited the same road an excessive number of times. We want to fix this by implementing global cooperation and stronger anti-looping logic. We could do this by having a shared set that tracks roads covered by any plow. To prevent repetitive visits we can introduce a hard limit on the number of times a plow is allowed to visit a road. Lastly, once a plow finishes its local zone, we want that plow to enter “cooperative mode” where it searches for nearby uncovered roads even if it’s outside its cluster to help other plows complete their zones.

10 Conclusion

Faced with a large snowstorm during Cornell’s Winter Graduation, we designed a system of algorithms to designate zones and routes for snow plow drivers in order to effectively clear Ithaca’s roads safely with certain priorities. From Figures 2, 3, and 4, we see that with 8 plows, we can effectively route emergency services with a low relative zone cost, balance work per zone effectively for the drivers, and plow the entire city with an upper bound of just under two hours. This allows for multiple passes to guarantee roads are clear, while reserving four drivers for shock scenarios and additional nearby storms.

Our solution is a K-Means clustering into zones based on distance and zone cost. The higher a zone cost, the more and the higher priority work there is, which we aim to split among the drivers evenly in our K-Means. Through an analysis of the 8 plow balance ratio comparing the mean and standard deviation zone cost for all 8 zones,

we achieve a relatively balanced outcome. With a greedy algorithm we can then find a fast route for covering all edges in a zone with one pass to prioritize emergency services. With our time computation, we are able to get an understanding of how the zone structure affects our efficiency, and the cost of removing all the snow from the roads in Ithaca. To best prepare for a storm, we follow guidelines of re-plowing every 2 hours, which also ensures that snow will not remain on the ground long enough to form ice. With our zone distribution prioritizing based on available traffic data and road gradient, we can also ensure that the most important roads to clear will be taken care of the soonest, and be split up evenly enough to handle.

In our future work, we outline an improvement in routing, as our goal is to get the quickest possible clearance. We understand that there will be a diminishing return in time saved with improvements to our greedy algorithm, as the Ithaca roadmap is a finite solvable system, but we hope to add reliability and robustness.

References

- [1] “Ithacating in Cornell Heights.” *The Great ‘Snowicane’ of 2010*, 27 Feb. 2010. Available at: <https://ithacating.com/2010/02/27/the-great-snowicane-of-2010/>.
- [2] Ithaca–Tompkins County Transportation Council. *2024 Traffic Count Report*. Tompkins County, 2024. Available at: <https://www.tompkinscountyny.gov/files/assets/county/v1/2024-traffic-count-report-final.pdf>.
- [3] U.S. Climate Data. “Climate Ithaca – New York.” n.d. Available at: <https://www.usclimatedata.com/climate/ithaca/new-york/united-states/usny0717>.
- [4] Data–IthacaNY Open Data. “Roads.” n.d. Available at: <https://data-ithacany.opendata.arcgis.com/datasets/roads/explore>.
- [5] “K-means Clustering.” *Wikipedia, Wikimedia Foundation*, n.d. Available at: https://en.wikipedia.org/wiki/K-means_clustering.
- [6] “Dijkstra’s Shortest Path Algorithm (Greedy Algo-7).” *Geeks-forGeeks*, n.d. Available at: <https://www.geeksforgeeks.org/dsa/dijkstras-shortest-path-algorithm-greedy-algo-7/>.

11 Appendix

11.1 Executive Summary

To Mr. Potter (Superintendent of Department of Public Works),

We have developed two models that will aid in your snowplow operation during the upcoming snowstorm that is predicted to occur during Cornell's Winter Graduation. The first model aids in finding the ideal number of plows needed to safely and efficiently clear Ithaca. Based on this model's partitioning of Ithaca into designated plow zones, we found it is optimal to deploy 8 out of your 12 plows at a given time to effectively clear the roads. The second model aids in understanding the routes that the plow drivers should take in their respective zones. This model simulates the complete coverage of Ithaca by the routes taken by the plow drivers and displays it in an animation. Therefore, we recommend that during the upcoming snowstorm, the Department of Public Works deploys 8 plows along the routes outlined in Figures 4 and 5. We hope our research proves helpful in effectively plowing Ithaca's roads so that all residents are safe and visitors can attend graduation smoothly.

Sincerely, Logan Abramowitz, Maya Glick, and Eric Sullivan

Source Code

11.2 Source Code

[https://drive.google.com/drive/folders/17e89jUN8qdf5hQzXI8T21u05Co0kkmbm?
usp=sharing](https://drive.google.com/drive/folders/17e89jUN8qdf5hQzXI8T21u05Co0kkmbm?usp=sharing)