Final Group Project

Smart City Analytics—Bike Sharing Operations

McGill University

MGSC 404-002 - Group 2

Setareh Farajollahzadeh

April 11, 2025

Laura Pantaleon

Ernest Bragg

Marc Pujol

William Fecteau

Tiphaine Levan

TABLE OF CONTENTS

1. Introduction

Bike-sharing has become a popular transportation option in cities worldwide, offering an eco-friendly alternative to cars and public transit. In Montreal, BIXI operates a network of over 500 docking stations, allowing users to pick up and return bikes across the city. Despite its growing popularity, BIXI faces several operational challenges, which our project aims to address—most notably, the imbalance of bikes across the network. If you've ever used BIXI—or any bike-sharing system—you've likely experienced the frustration of arriving at a station with no available bikes or trying to return a bike to a full station. These issues arise because travel patterns are not symmetric; bikes naturally accumulate in certain areas and become scarce in others throughout the day. To address this problem, we developed a mixed-integer programming (MIP) model using data from Chicago's Divvy bike-sharing system. The Divvy dataset includes more detailed information than what is publicly available for BIXI, such as station capacities and individual bike IDs, making it more suitable for our analysis. Although we use data from Chicago, our findings and recommendations are designed to be applicable to Montreal's BIXI system. This report outlines our data exploration process, the mathematical model we developed, and the insights gained from our analysis. We conclude with practical recommendations for BIXI to improve operations and enhance the user experience. Through this analysis, we aim to demonstrate how decision analytics concepts from the semester can be applied to solve real-world business problems.

2. Data Exploration & Cleaning

We began our analysis by cleaning and preparing the data to meet the specific requirements of our optimization model. The original datasets consisted of three primary files: Q3 trips, Q4 trips, and Stations datasets. Our initial data preparation involved removing duplicates and unnecessary customer-related fields that contained null values. The most significant transformation was merging the Q3 and Q4 datasets into a unified Trips dataset to facilitate comprehensive analysis across both quarters. For temporal analysis, we converted all timestamp fields (start_time and end_time) into *datetime* format and further decomposed them into separate date, hour, and time components. We also created day_of_the_week (e.g., "Monday") and day_of_week_num (e.g., "0" for Monday) variables to support day-specific analysis patterns. To ensure data quality, we removed outliers using a quartile-based technique, which proved more effective than standard deviation methods for handling the highly skewed distribution of trip durations (Appendix A)

2.1 Preliminary Insights from Visualization

2.1.1 Average Bike Journey

Analysis of the average bike journey duration revealed valuable patterns in system usage. The distribution of trip durations showed a right-skewed pattern (Appendix B), with most trips falling between 300 and 600 seconds (5 to 10 minutes). This suggests that the bike-sharing system is predominantly used for short-distance travel, likely quick commutes or errands rather than extended recreational rides. Despite the removal of outliers, the frequency of trips steadily declines as duration increases (shown by the long tail of the distribution) with very few trips lasting beyond 30 minutes, possibly for recreation or commuting through less connected areas. This pattern likely reflects the pricing structure of bike-sharing systems, which typically discourages extended use through increasing fees.

2.1.2 Evaluating Busiest and Most Idle Stations

To identify patterns of high and low demand, we analyzed stations based on their usage during peak hours and compared these values to their capacity. This approach allowed us to create a peak-hour demand-to-capacity ratio for each station (Appendix C). Stations with ratios below 1 typically have excess supply, while those with higher ratios experience significant demand pressure. Our analysis revealed several stations with exceptionally high demand. Station 287, for example, demonstrated a demand-to-capacity ratio of 56.8, with peak demand reaching 1534 bikes. Other stations, including 192, 35, 91, and 77, also showed substantial demand, making them prime candidates for inclusion in our optimization model. The geographical distribution of high-demand stations showed clear patterns, with most concentrated in downtown Chicago (latitude ~41.87 to ~41.90, longitude ~-87.61 to ~-87.64). The stark contrast between demand levels in central versus peripheral areas highlighted the need for strategic bike repositioning to ensure adequate service across the network. Additionally, we observed that many high-demand stations were located within a 1 km radius of each other in the downtown core, suggesting that bike reallocation could be logistically feasible within this concentrated area. This clustering pattern informed our approach to developing efficient repositioning routes in the subsequent optimization model.

2.1.3 Evaluating Periods of Activity

After identifying high-demand stations, we analyzed temporal patterns to determine optimal repositioning windows. Our analysis of trip volumes by season, day, and hour revealed distinct usage patterns (Appendix D). Summer months showed the highest overall activity, with July representing the peak usage period. Weekday patterns differed significantly from weekend usage, with weekdays showing pronounced bi-modal demand distribution corresponding to morning (7-9 AM) and evening (4-6 PM) commuting hours. Specifically, weekday mornings demonstrated high demand between 7 AM and 9 AM, reaching approximately 24,000 users at 7 AM and 30,000 users at 8 AM. Evening peaks were even more pronounced, occurring between 4 PM and 6 PM with demand ranging from 32,000 to 47,000 users. The highest single-hour demand was observed on Tuesday and Wednesday evenings at 5 PM, peaking at around 46,000 users. Weekend patterns showed a different distribution, with Saturday midday (around 2 PM) representing the highest weekend usage at approximately 29,000 users. System-wide activity was lowest between midnight and 5 AM, averaging below 2,000 users, which offered a potential window for operational activities such as bike repositioning.

3. Model Formulation & Solving

3.1 Model Overview and Decision Variables

At the core of our optimization model is the objective of minimizing the total cost associated with bike repositioning, while simultaneously ensuring that bike availability aligns with user demand across the network. The total cost function incorporates three primary components: fixed costs incurred per truck deployed, variable transportation costs determined by the distance traveled, and penalty costs associated with unaddressed station imbalances—namely, surplus or shortage of bikes. Specifically, it identifies which trucks should be deployed during each period, the sequence of stations each truck should visit, the number of bikes to be picked up or dropped off at each location, and the resulting inventory levels at all stations following the repositioning effort. To support these decisions, the model introduces a series of critical decision variables. Binary variables are used to indicate whether a truck is active during a given period and whether it visits specific stations on its route. Integer variables quantify the number of bikes moved at each station and track the truck's load following each stop. Additional variables represent the travel routes between stations and compute the total distance covered. To further account for service quality, the model includes variables that capture any remaining surplus or deficit at stations after repositioning operations are completed. This structured and comprehensive set of decision variables enables the model to generate a complete repositioning plan—from truck deployment schedules to specific bike movements—while balancing operational efficiency and service reliability.

3.2 Model Approach and Scope

Our approach is based on a multi-stop truck routing optimization model that identifies both the most efficient redistribution routes and the optimal number of bikes to pick up or drop off at each station. Drawing on insights from our data analysis, we selected two high-contrast time windows for modeling: Wednesday afternoon (4 PM–6 PM), which captures the intensity of the evening commute, and Monday morning (7 AM–9 AM), representing typical morning commute patterns. These periods reflect distinct demand dynamics and enable us to test the model's responsiveness under varying conditions. Both windows were selected from July data, corresponding to peak annual bike usage in Chicago. To ensure computational feasibility, we limited our model's scope to the 20 stations with the highest activity volumes. This focused approach allows us to address the most significant imbalances in the network while maintaining model solvability within practical time constraints.

3.3 Key Parameters

We calibrated our model with several key parameters to reflect operational realities. Our truck fleet consists of a maximum of 20 vehicles, each with a capacity of 40 bikes. To maintain operational efficiency, we limited each truck to a maximum of 5 stops, and a total distance of 50 km per route. The cost structure includes a $50 fixed cost per truck deployed, a $3 per kilometer transportation cost, and a $50 penalty per bike for unfulfilled surplus or deficit. For inventory management, we established target levels based on station characteristics. Stations with higher arrivals than departures have their target inventory set at 30% of capacity, while stations with higher departures than arrivals have their target set at 70% of capacity. This approach ensures sufficient bikes are available where people tend to start trips while maintaining enough empty docks where trips typically end.

3.4 Implementation and Solution Approach

We implemented our model using the Gurobi optimization solver through its Python interface (gurobipy). The implementation process involved preprocessing the trip data to identify peak demand periods and station characteristics, calculating distances between stations using the Haversine formula, determining target inventory levels and initial imbalances, formulating the model with all necessary constraints, and solving to find optimal truck routes and bike movements. The mixed-integer programming formulation includes variables representing truck activities, station visits, bike pickups and dropoffs, and routing decisions. The complete mathematical formulation is provided in Appendix E.

3.5 Solution Results and Discussion

The model produced feasible and actionable repositioning strategies for both selected time periods. In the Wednesday afternoon scenario, five trucks were deployed at a total operational cost of $582.71 (Appendix F). In contrast, the Monday morning scenario required only four trucks, resulting in a significantly lower total cost of $228.27 (Appendix G). This cost disparity highlights the differing severity of station imbalances across the two periods, with the Wednesday evening peak exhibiting more pronounced demand asymmetries that necessitate broader redistribution efforts. The corresponding optimality gaps—13.05% for Wednesday and 5.04% for Monday—indicate how close each solution is to the theoretical optimum. Considering the model's complexity, these gaps represent an acceptable balance between solution accuracy and computational tractability. Importantly, the current output serves as an illustrative example of how the model operates. However, the model can be applied to any time window on any day of the week, offering flexibility for broader use.

3.6 Limitations and Practical Considerations

While the model yields valuable operational insights, it is important to recognize several limitations arising from practical constraints and necessary simplifications. To balance realism with computational feasibility, several assumptions were introduced. First, each station is assumed to begin at 50% capacity—an arbitrary but practical approximation that shifts focus to relative imbalances rather than exact inventory levels. Although more precise estimates could be derived from historical movement data, doing so would significantly increase computational complexity without a proportional improvement in model accuracy. Furthermore, trucks are assumed to start their routes at their first assigned stations rather than from a central depot, simplifying the routing process while potentially underestimating initial travel costs. A 50 km route length constraint is imposed, but vehicle speed and travel time between stations are not explicitly modeled. In practice, traffic conditions—especially during peak periods—could reduce the number of feasible stops within the two-hour repositioning windows. Additionally, the model treats time windows as independent, even though inventory levels at the end of one period may affect subsequent periods. The time required for loading and unloading bikes is also excluded, which would realistically extend route durations. Factors such as urban congestion, weather events, or station-level spatial limitations are not considered, despite their operational significance. For simplicity, we assume multiple trucks can operate at the same station concurrently. Finally, cost parameters—including per-truck fixed costs, fleet size, and per-bike movement costs—are estimated using publicly available information from BIXI and comparable bike-sharing systems. To maintain computational tractability, we limited the analysis to 20 high-volume stations and two peak-time windows, accepting near-optimal solutions where appropriate.

4. Sensitivity Analysis

Given the mixed-integer characteristics of our optimization model, which integrates both binary and integer decision variables for truck routing and fleet management, a conventional sensitivity analysis was not viable. Consequently, we employed a targeted 'what-if' analysis strategy, systematically adjusting pivotal parameters to evaluate their impact on operational outcomes. Initially, an exhaustive looping approach across all parameters—truck fixed cost, transport cost, surplus penalties, and deficit penalties—was considered. However, this was discarded due to computational intensity and inefficiency. We prioritized two parameters exhibiting the most significant operational leverage: fixed truck costs and deficit penalties. Transport cost per kilometer and surplus penalties demonstrated limited marginal impact and were thus excluded from the core sensitivity analysis. For example, variations in transportation costs from $0.50 to $10/km on Monday resulted in total cost fluctuations between $203.81 and $292.19 (Appendix I), without affecting truck deployment or unmet demand, highlighting that routing decisions were already optimized. Similarly, surplus penalties remained consistently at zero across tests, indicating minimal influence on operational decisions.

Fixed truck costs significantly impact fleet deployment, with higher fixed costs discouraging fleet expansion, thus elevating the risk of unmet demand. For instance, increasing Wednesday's fixed truck cost from $50 to $100 escalated the total operational cost from $265.85 to $432.60 while maintaining constant truck usage (4 trucks) (Appendix H). This scenario saw transportation costs rise modestly from $25.85 to $32.60, while truck costs doubled, suggesting that elevated fixed costs can inflate total expenditures without service enhancement. Conversely, the Monday model demonstrated greater operational resilience; at a baseline fixed truck cost of $50, total cost was $228.27 using four trucks without penalties, reflecting balanced demand and optimized asset utilization (Appendix I).

Deficit penalties also emerged as influential, which was particularly evident in the Wednesday scenario (Appendix J). With penalties at $300 per unit of unmet demand, deficit penalties constituted over 50% of total operational costs ($582.71), compelling aggressive redistribution strategies. When penalties were reduced, the model permitted higher unmet demand to decrease overall costs, albeit sacrificing service reliability. For example, reducing the deficit penalty from $50 to $30 per unit decreased total cost from $492.63 to $461.34, demonstrating the model's willingness to tolerate service gaps under lower penalties. Thus, higher penalties effectively incentivize strategic resource allocation, even if associated transport expenses increase.

4.1 Managerial Insights

The comparative sensitivity analyses underscore that fixed truck costs had limited impact on Monday, but a significant one on Wednesday outcomes, meaning that there is no universal redistribution strategy applied across diverse demand scenarios. During periods of lower intensity, such as Monday mornings, achieving operational efficiency with minimal redistribution interventions is feasible. Station targets were achieved without penalties and minimal resource deployment, supporting strategies like late-night redistribution rounds utilizing fewer trucks and minimal crews, thereby reducing labor and fuel costs. This is especially practical for commuter-heavy stations like Station 91, characterized by predictable early morning demand.

Conversely, Wednesday's peak-hour analysis indicated substantial systemic stress, necessitating aggressive redistribution, increased truck utilization, and strategic prioritization. Under these conditions, maintaining uniform service across all stations proved costly and impractical. Therefore, adopting a tiered prioritization strategy focusing resources on downtown core stations (e.g., Stations 287, 192, and 35) with high demand-to-capacity ratios is recommended. Implementing guaranteed service zones serviced by smaller, agile transportation methods can facilitate just-in-time redistribution focused on critical nodes, reducing the strain on truck fleets. Furthermore, given deficit penalties' substantial contribution to total costs (50%), introducing cost-sharing arrangements with commercial partners (in exchange for prioritized access or branding) during surge hours could transform operational challenges into revenue-generating opportunities, aligning financial incentives with operational effectiveness. Overall, proactive adaptation to anticipated demand fluctuations, rather than reactive redistribution, represents the optimal managerial approach, leveraging insights derived from the sensitivity analysis of fixed truck costs and deficit penalties. These policies should not just react to demand patterns, but anticipate them before users even notice a shortage.

5. Conclusion

In summary, our analysis demonstrates that bike repositioning strategies must be tailored to distinct temporal demand patterns. Morning commutes require minimal intervention while evening peaks demand intensive rebalancing efforts. This asymmetry calls for time-differentiated approaches: proactive late-night repositioning for predictable morning demands, and targeted "guaranteed service zones" for evening rush hours. We've successfully developed a flexible optimization framework that balances operational efficiency with service quality, identifying when, where, and how to deploy limited resources for maximum impact. While applied to a subset of stations, our methodology offers BIXI a scalable solution for their entire network. Future refinements could include dynamic pricing to influence user behavior, real-time demand prediction, and integration with public transit systems, ultimately enhancing urban mobility in an increasingly interconnected smart city ecosystem.

APPENDICES

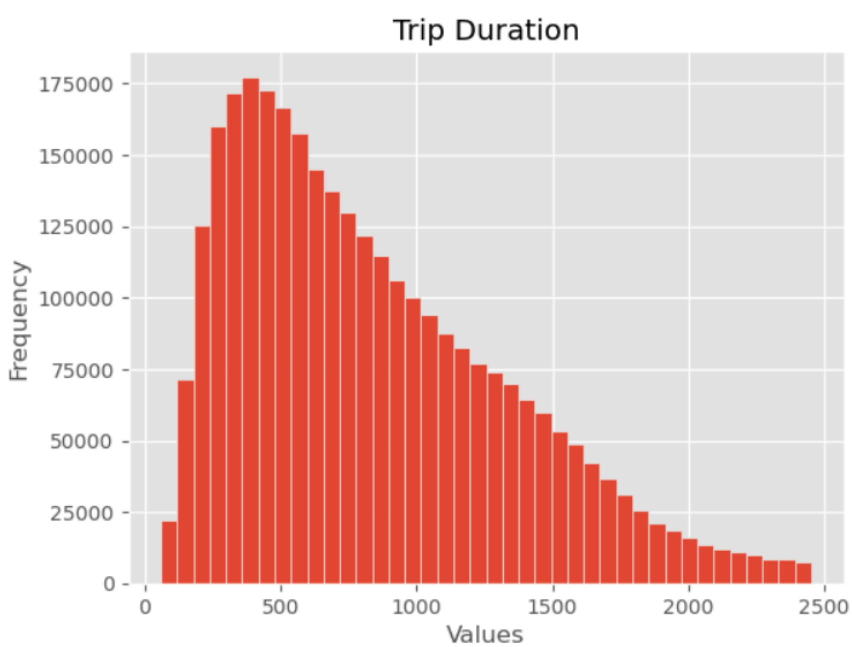
Appendix A: Final Dataset (Head) After Cleaning

	trip_id	bikeid	tripduration	from_station_id	from_station_name	to_station_id	to_station_name	start_date_hour_minute
0	16734065	1411	349	216	California Ave & Division St	259	California Ave & Francis Pl	2017-09-30 23:59:58

start_date	start_hour_minute	start_hour	end_date_hour_minute	end_date	end_hour_minute	end_hour	day_of_week	day_of_week_num
2017-09-30	23:59:58	23	2017-10-01 00:05:47	2017-10-01	00:05:47	0	Saturday	5

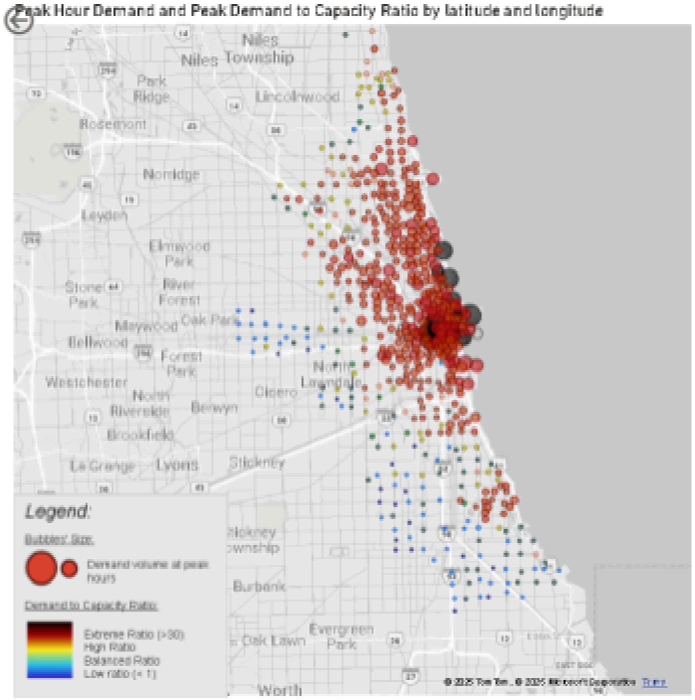
3053574 rows × 17 columns

Appendix B: Average Bike Journey



Appendix C: Evaluating Busiest and Most Idle Stations

Peak Hour Demand and Peak Demand to Capacity Ratio by latitude and longitude



Legend Explanation:

Bubble Size: Represents the demand volume for bikes during peak hours. Larger bubbles indicate higher demand at the peak hour. Peak hours can be different for each station.

Peak Demand to Capacity Ratio: This ratio compares the demand during peak hours to the station's capacity. The ratio helps identify stations with high demand relative to their capacity.

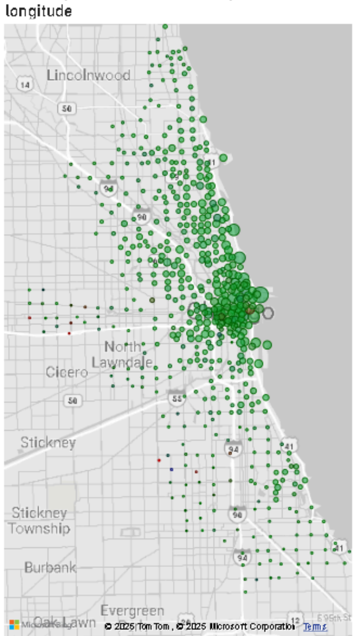
Color Coding:

Blue (Low Ratio): Indicates that the station has low demand compared to its capacity.
Green (Ratio ≈ 1): Represents stations where demand is balanced with capacity (demand ≈ capacity).
Orange/Red: Indicates stations where demand exceeds capacity, and there may be a strain on resources.
Black (Extreme Ratio ≥ 30): Indicates an extremely high demand relative to capacity.

Example:

Station 35: A Peak Demand Ratio of 35 means that during peak hours, the station has 1651 bike demand, but its capacity is much lower. While not all 1651 people will be at the station simultaneously (some will pick up bikes, others will drop off), the map highlights potential bottlenecks in areas of high demand.
City center stations are likely to experience high demand, while stations on the outskirts of the city may have surplus bikes.
The chart reveals areas where adjustments, such as rebalancing bikes or adding docks, are needed to meet demand efficiently.

Total Trips and Total Arrivals by latitude and longitude



id	Departure to Arrival Ratio
594	2.58
538	2.33
544	2.20
556	2.00
539	1.87
412	1.78
536	1.76
195	1.71
529	1.50
376	1.49
385	1.43
191	1.40
593	1.39
392	1.32
531	1.31
518	0.80
591	0.79
488	0.79
220	0.79
541	0.78
279	0.78
409	0.78
275	0.77
478	0.76
437	0.76
450	0.73
431	0.73
443	0.72
535	0.71
543	0.71
369	0.64
551	0.62
586	0.62
103	0.60
550	0.59
Total	1.20

Legend Explanation:

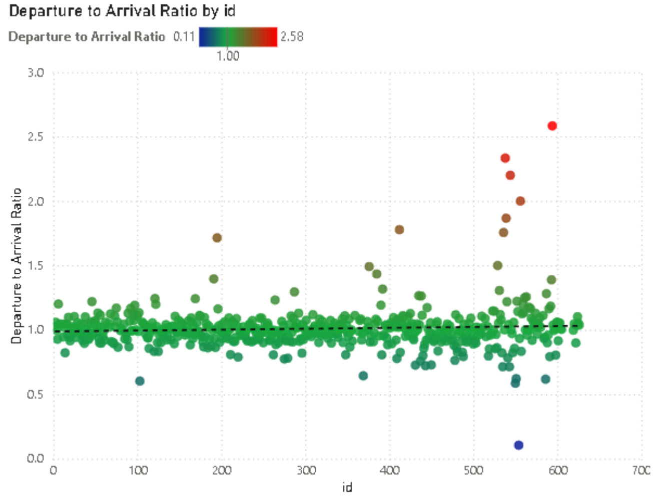
- **Ratio > 1:** *Station frequently loses bikes* (more departures than arrivals).
- **Ratio < 1:** *Station accumulates bikes* (more arrivals than departures).
- **Ratio ≈ 1:** *Station is balanced* (departures and arrivals are approximately equal).

Bubble Size: Represents the total number of trips (both departures and arrivals) at the station.

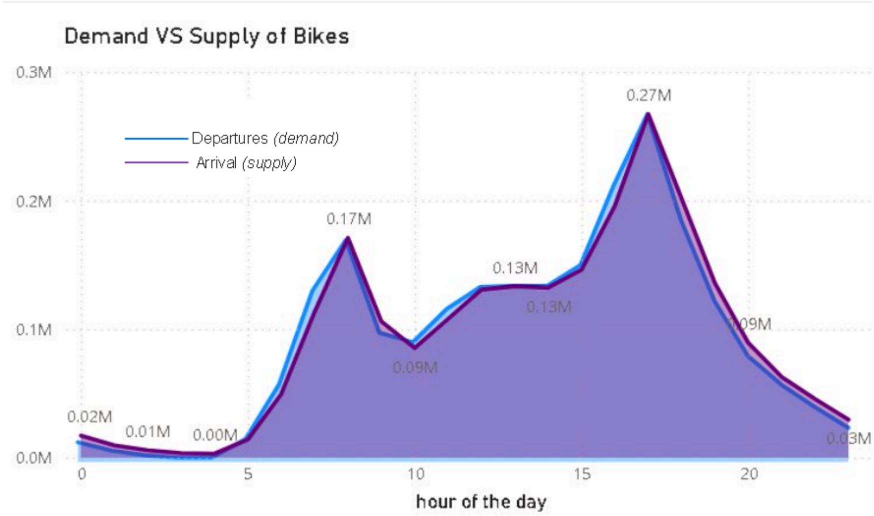
Bubble Color:

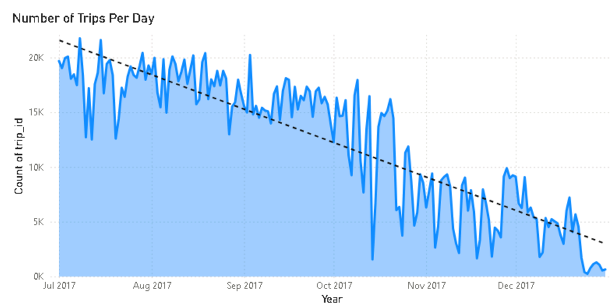
- **Red:** Station needs bikes (Ratio > 1, high departures).
- **Blue:** Station needs docks (Ratio < 1, high arrivals).
- **Green:** Station is balanced (Ratio ≈ 1, departures ≈ arrivals).

Departure to Arrival Ratio by id



Appendix D: Evaluating Periods of Activity





Demand at Days and Hours of the week.

start_hour	Friday	Monday	Saturday	Sunday	Thursday	Tuesday	Wednesday	Total
0	2066	1395	3260	4070	1474	1340	1391	14996
1	1023	789	2256	2645	718	686	600	8717
2	534	415	1361	1725	369	371	319	5094
3	364	265	767	982	315	202	280	3175
4	498	427	440	658	422	467	459	3371
5	3003	2782	763	827	3226	3385	3229	17215
6	10061	10085	2028	1747	10992	12126	11722	58761
7	23364	23375	4348	3632	23987	26846	24785	130337
8	29166	29766	8662	7288	30677	32436	30981	168976
9	14588	14315	14187	13057	14193	14171	13842	98353
10	11251	11694	19808	18519	10159	10305	9317	91133
11	14995	14878	25110	23201	12910	13380	11802	116276
12	18269	17645	28151	25440	14836	14762	14053	133156
13	18744	17667	28836	25470	14706	14763	13856	134042
14	18901	17415	29289	25254	14283	14876	13981	133999
15	23343	19714	28317	26175	17418	17841	16944	149752
16	33662	32761	26207	25191	30428	31884	30744	210877
17	35792	48044	23445	22138	45329	46038	44155	264941
18	23404	31758	20514	19370	29539	30769	28982	184336
19	15074	20362	14886	14749	18325	20160	19120	122676
20	9862	12616	10332	9881	12053	13558	12391	80693
21	7216	8867	8052	7230	8938	9515	9011	58829
22	5902	5164	6902	5034	6393	6429	5967	41791
23	4482	2856	5685	2893	3803	2998	3275	25992
Total	325564	345055	313686	287176	325493	339308	321206	2257488

Appendix E: Bike Repositioning Optimization: MIP Formulation

Bike Repositioning Optimization: MIP Formulation

Indices and Sets

- $u \in \{1, \ldots, U\}$: Trucks (with U being the maximum number of trucks).
- $i, j \in \{1, \ldots, N\}$: Stations (with N as the total number of stations).
- $s \in \{0, 1, \ldots, S-1\}$: Stop indices along a truck's route (with S being the maximum number of stops per truck).
- $t \in T$: Time periods, where

$$T = \{\text{Mon}, \text{Wed}\}.$$

Note: Monday corresponds to the time window $[7$ AM, 9 AM$)$ and Wednesday to $[4$ PM, 6 PM$)$.

Parameters

- $C = 40$: Truck capacity.
- $S_{\max} = 5$: Maximum number of stops per truck.
- $F = 50\$$: Fixed cost per truck used.
- $c_d = 3\$$: Transportation cost per kilometer.
- $D_{\max} = 50km$: Maximum distance a truck can travel in a time period.
- $p_s = 50\$$: Surplus penalty (for bikes left over at stations with a surplus).
- $p_d = 50\$$: Deficit penalty (for unmet bike demand at stations with a deficit).
- d_{ij}: Distance (in km) between station i and station j.
- Δ_i: Inventory imbalance at station i, defined as

$$\Delta_i = \text{Target Bikes}_i - \text{Current Bikes}_i.$$

Define:

- $\mathcal{S} = \{i : \Delta_i < 0\}$ (stations with surplus bikes),
- $\mathcal{D} = \{i : \Delta_i > 0\}$ (stations with bike deficits).

Decision Variables

- $y_{u,t} \in \{0, 1\}$: Equals 1 if truck u is active in time period t, and 0 otherwise.
- $x_{u,i,s,t} \in \{0, 1\}$: Equals 1 if truck u visits station i as its s-th stop in period t, and 0 otherwise.
- $z_{u,i,j,s,t} \in \{0, 1\}$: Equals 1 if truck u travels from station i to station j immediately after stop s in period t, for $i \neq j$.
- $p_{u,i,s,t} \in \mathbb{Z}_{\geq 0}$: Number of bikes picked up by truck u at station i during stop s in period t.
- $d_{u,i,s,t} \in \mathbb{Z}_{\geq 0}$: Number of bikes dropped off by truck u at station i during stop s in period t.
- $\ell_{u,s,t} \in \mathbb{Z}_{\geq 0}$: Load (number of bikes) carried by truck u after stop s in period t, with $\ell_{u,s,t} \leq C$.
- $D_{u,t} \geq 0$: Total distance traveled by truck u in period t.
- For surplus stations ($i \in \mathcal{S}$): $r_i^s \in \mathbb{Z}_{\geq 0}$ represents the number of surplus bikes not moved.
- For deficit stations ($i \in \mathcal{D}$): $r_i^d \in \mathbb{Z}_{\geq 0}$ represents the number of deficit bikes not fulfilled.

Objective Function

Minimize the total cost comprising the truck activation cost, transportation cost, and penalties for surplus and deficit:

$$\min \quad \sum_{t \in T} \sum_{u=1}^{U} \Big(F\, y_{u,t} + c_d\, D_{u,t} \Big) + \sum_{i \in \mathcal{S}} p_s\, r_i^s + \sum_{i \in \mathcal{D}} p_d\, r_i^d.$$

Explanation: This objective function consists of three main components:

- **Fixed truck activation costs:** Each truck that is used in a time period incurs a fixed cost F.
- **Transportation costs:** Trucks incur a variable cost c_d proportional to the distance traveled $D_{u,t}$.

- **Penalty costs:** These are incurred when inventory imbalances at stations are not fully resolved — p_s penalizes uncollected surplus bikes, and p_d penalizes unmet demand (deficits).

The optimization model seeks to minimize the sum of these costs across all trucks and time periods while repositioning bikes efficiently to meet station demands.

Constraints

(1) Truck Activation Constraints:

$$\sum_{i=1}^{N}\sum_{s=0}^{S-1} x_{u,i,s,t} \leq S_{\max}\, y_{u,t}, \quad \forall u,\ t \in T. \tag{1}$$

Explanation: This constraint ensures that if truck u is not activated in time period t (i.e., $y_{u,t} = 0$), then it cannot visit any station (all $x_{u,i,s,t}$ must be 0). If truck u is activated (i.e., $y_{u,t} = 1$), then it is allowed to make at most $S_{\max}$ stops.

(2) Load Management Constraints:

$$\ell_{u,0,t} = \sum_{i=1}^{N}\Big(p_{u,i,0,t} - d_{u,i,0,t}\Big), \quad \forall u,\ t \in T, \tag{2}$$

$$\ell_{u,s,t} = \ell_{u,s-1,t} + \sum_{i=1}^{N}\Big(p_{u,i,s,t} - d_{u,i,s,t}\Big), \quad s = 1,\ldots,S-1,\ \forall u,\ t \in T, \tag{3}$$

$$\ell_{u,s,t} \leq C, \quad \forall u,\ s,\ t \in T. \tag{4}$$

Explanation:

- The first equation defines the truck's load after the initial stop (stop 0) as the net number of bikes picked up minus the bikes dropped off.
- The second equation updates the load for each subsequent stop by adding the net change (pickups minus dropoffs) from the current stop to the load from the previous stop.
- The third equation ensures that the load carried by the truck never exceeds its capacity C.

(3) Visit-Operation Linkage:

$$p_{u,i,s,t} \leq C\, x_{u,i,s,t}, \quad \forall u,\ i,\ s,\ t \in T, \tag{5}$$

$$d_{u,i,s,t} \leq C\, x_{u,i,s,t}, \quad \forall u,\ i,\ s,\ t \in T, \tag{6}$$

$$\sum_{i=1}^{N} x_{u,i,s,t} \leq y_{u,t}, \quad \forall u,\ s,\ t \in T. \tag{7}$$

Explanation:

- The first two constraints link the operational decisions (picking up or dropping off bikes) to the visit decision. They ensure that if the truck does not visit station i at stop s (i.e., $x_{u,i,s,t} = 0$), then no bikes can be picked up or dropped off there.
- The third constraint ensures that a truck can only perform an operation at a stop if it is active in that time period (i.e., if $y_{u,t} = 0$, then no station visits occur).

(4) Route Continuity Constraints:

$$\sum_{i=1}^{N} x_{u,i,s,t} \geq \sum_{i=1}^{N} x_{u,i,s+1,t}, \quad \forall u,\ s = 0,\ldots,S-2,\ t \in T, \tag{8}$$

$$\sum_{\substack{j=1\\ j\neq i}}^{N} z_{u,i,j,s,t} = x_{u,i,s,t}, \quad \forall u,\ i,\ s = 0,\ldots,S-2,\ t \in T, \tag{9}$$

$$\sum_{\substack{i=1\\ i\neq j}}^{N} z_{u,i,j,s,t} = x_{u,j,s+1,t}, \quad \forall u,\ j,\ s = 0,\ldots,S-2,\ t \in T. \tag{10}$$

Explanation:

- The first constraint ensures that if a truck makes a stop at position $s+1$, then it must have also made a stop at position s. In other words, the sequence of stops is cumulative.
- The second constraint ties the visit decision at stop s to the route decision: if the truck visits station i at stop s, then it must depart from that station to some other station immediately after.
- The third constraint ensures that if the truck arrives at station j as the $(s+1)$-th stop, it must have come directly from some other station at the previous stop.

Appendix F: Model Output: Wednesday (4pm-6pm)

(I) Wednesday Truck Routes

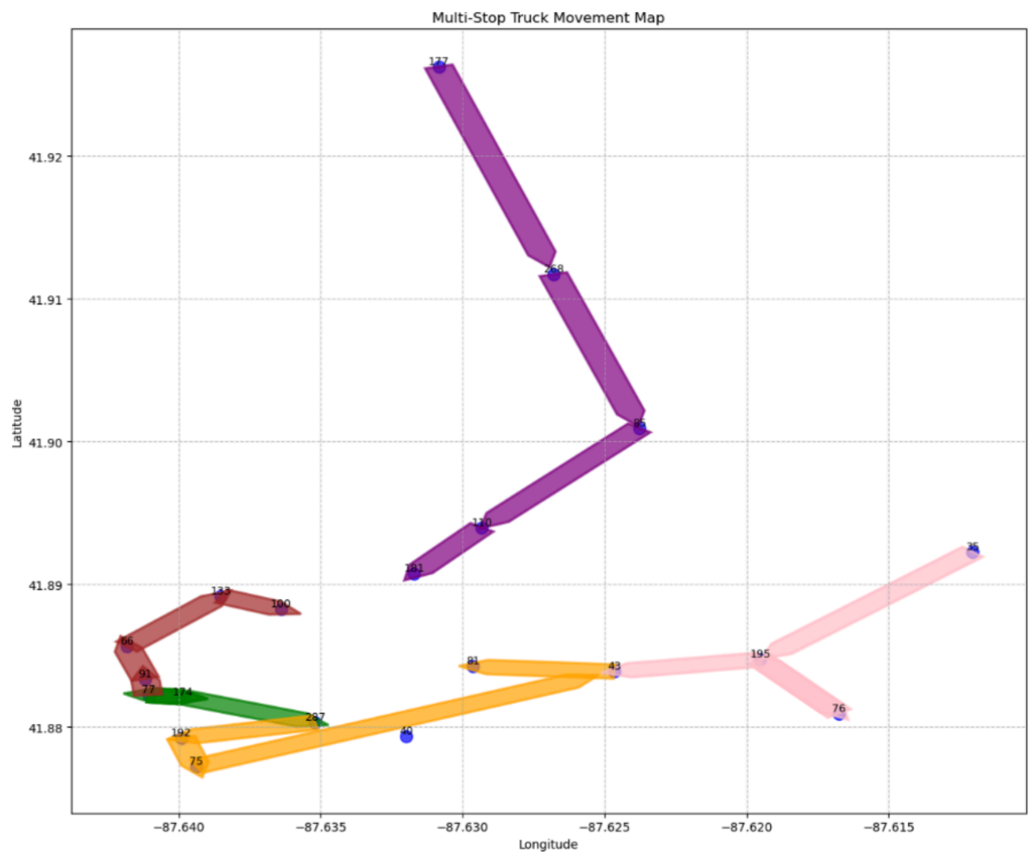
Truck	Route	Total distance (km)	Total Cost ($)
1	77 → 174 → 77 → 174 → 287	0.74	52.23
2	287 → 192 → 75 → 43 → 81	2.48	57.43
3	177 → 268 → 85 → 110 → 181	4.19	62.56
4	77 → 91 → 66 → 133 → 100	1.07	53.20
5	35 → 195 → 76 → 195 → 43	2.43	57.30

(II) Wednesday Cost Breakdown

Total Cost ($)	Transportation Cost ($)	Truck Fixed Cost ($)	Surplus Penalty ($)	Deficit Penalty ($)
582.71	32.71	250.0	0.0	300.0

(III) Wednesday Station Inventory

Bike Movement Map Wednesday (4pm-6pm)



Appendix G: Model Output (Monday, 7am-9pm)

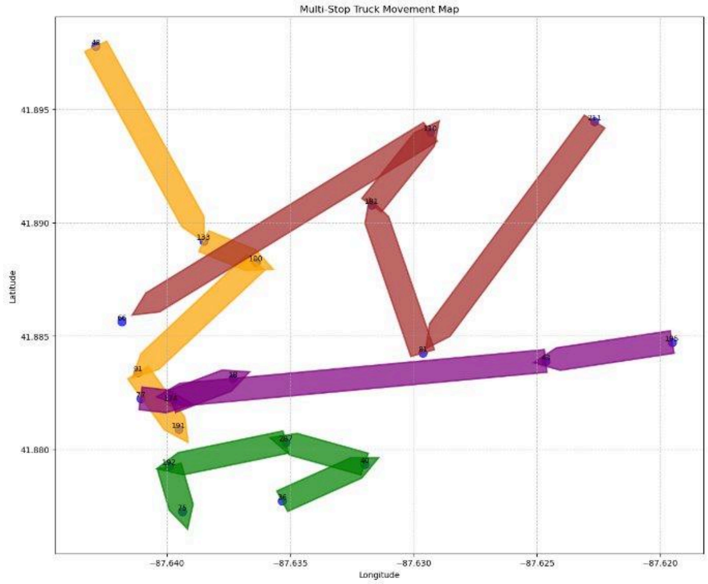
```
  Nodes    |    Current Node    |     Objective Bounds      |     Work
 Expl Unexpl |  Obj  Depth IntInf | Incumbent    BestBd   Gap | It/Node Time

     0     0   34.25388    0  210 6300.00000   34.25388  99.5%     -    5s
   ….
H    0     0                      228.2694500  216.77191  5.04%     -  533s

Cutting planes:
  Gomory: 1
  Implied bound: 4
  MIR: 40
  Flow cover: 26
  Flow path: 28
  Zero half: 26
  Network: 18
  RLT: 6

Explored 1 nodes (113487 simplex iterations) in 550.19 seconds (117.37 work units)
Thread count was 8 (of 8 available processors)
```

```
Solution count 10: 228.269 234.809 234.809 ... 276.43

Time limit reached
Best objective 2.282694499814e+02, best bound 2.167719117631e+02, gap 5.0368%

========================================
========== FOUND 10 SOLUTIONS ==========
Solution 0: Cost $228.27
Solution 1: Cost $234.81
Solution 2: Cost $234.81
Solution 3: Cost $234.81
Solution 4: Cost $275.39
Solution 5: Cost $275.39
Solution 6: Cost $275.44
Solution 7: Cost $276.11
Solution 8: Cost $276.39
Solution 9: Cost $276.43

========================================
========= OPTIMIZATION RESULTS =========
Total Cost:          $228.27
Trucks Used:         4

----------------------------------------
------------- TRUCK ROUTES -------------

Time period 7 - Truck 2:
  Total distance: 2.20 km
  Total cost: $56.61
  Route details:
    Stop 1: Station 48 - Pickup 8 bikes (Load after: 8)
    Stop 2: Station 133 - Pickup 6 bikes (Load after: 14)
    Stop 3: Station 100 - Pickup 7 bikes (Load after: 21)
    Stop 4: Station 91 - Dropoff 6 bikes (Load after: 15)
    Stop 5: Station 191 - Dropoff 5 bikes (Load after: 10)

Time period 7 - Truck 3:
  Total distance: 2.15 km
  Total cost: $56.44
  Route details:
    Stop 1: Station 195 - Pickup 9 bikes (Load after: 9)
    Stop 2: Station 43 - Pickup 9 bikes (Load after: 18)
    Stop 3: Station 77 - Dropoff 6 bikes (Load after: 12)
    Stop 4: Station 174 - Dropoff 7 bikes (Load after: 5)
    Stop 5: Station 18 - Pickup 4 bikes (Load after: 9)

Time period 8 - Truck 1:
  Total distance: 1.25 km
  Total cost: $53.76
  Route details:
    Stop 1: Station 36 - Pickup 8 bikes (Load after: 8)
    Stop 2: Station 40 - Pickup 3 bikes (Load after: 11)
    Stop 3: Station 287 - Pickup 5 bikes (Load after: 16)
```

```
    Stop 4: Station 192 - Dropoff 9 bikes (Load after: 7)
    Stop 5: Station 75 - Dropoff 6 bikes (Load after: 1)

Time period 8 - Truck 4:
  Total distance: 3.82 km
  Total cost: $61.46
  Route details:
    Stop 1: Station 211 - Pickup 4 bikes (Load after: 4)
    Stop 2: Station 81 - Pickup 8 bikes (Load after: 12)
    Stop 3: Station 181 - Pickup 6 bikes (Load after: 18)
    Stop 4: Station 110 - Dropoff 5 bikes (Load after: 13)
    Stop 5: Station 66 - Dropoff 5 bikes (Load after: 8)

----------------------------------------
---------- STATION INVENTORY -----------
    station_id  capacity  current_bikes  target_bikes  final_bikes  demand  \
8           18      19.0              9             5            5   248.0
14          36      39.0             19            11           11   162.0
17          40      15.0              7             4            4     4.0
4           43      43.0             21            12           12   112.0
9           48      39.0             19            11           11    90.0
13          66      23.0             11            16           16   228.0
12          75      31.0             15            21           21   244.0
2           77      31.0             15            21           21   532.0
5           81      39.0             19            11           11    68.0
1           91      31.0             15            21           21   762.0
6          100      35.0             17            10           10    66.0
18         110      23.0             11            16           16   184.0
11         133      31.0             15             9            9   220.0
7          174      35.0             17            24           24   380.0
19         181      31.0             15             9            9    64.0
15         191      28.0             14            19           19   230.0
0          192      47.0             23            32           32   802.0
10         195      47.0             23            14           14   178.0
16         211      19.0              9             5            5    56.0
3          287      27.0             13             8            8    50.0

    supply  net_flow  surplus_remaining  deficit_remaining  \
8    294.0      46.0                  0                  0
14   266.0     104.0                  0                  0
17   362.0     358.0                  0                  0
4    634.0     522.0                  0                  0
9    432.0     342.0                  0                  0
13   206.0     -22.0                  0                  0
12   202.0     -42.0                  0                  0
2    432.0    -100.0                  0                  0
5    562.0     494.0                  0                  0
1    326.0    -436.0                  0                  0
6    514.0     448.0                  0                  0
18   182.0      -2.0                  0                  0
11   228.0       8.0                  0                  0
7    192.0    -188.0                  0                  0
19   296.0     232.0                  0                  0
15   172.0     -58.0                  0                  0
```

```
0    358.0    -444.0                 0                   0
10   302.0     124.0                 0                   0
16   314.0     258.0                 0                   0
3    698.0     648.0                 0                   0

    final_occupancy_rate
8               0.263158
14              0.282051
17              0.266667
4               0.279070
9               0.282051
13              0.695652
12              0.677419
2               0.677419
5               0.282051
1               0.677419
6               0.285714
18              0.695652
11              0.290323
7               0.685714
19              0.290323
15              0.678571
0               0.680851
10              0.297872
16              0.263158
3               0.296296

----------------------------------------
------------ COST ANALYSIS -------------
Transportation cost: $28.27
Truck fixed cost:    $200.00
Surplus penalty:     $0.00
Deficit penalty:     $0.00
Total cost:          $228.27
```

Bike Movement Map Monday (7am-9am)



Appendix H: Wednesday - Fixed Truck Cost Variation

Fixed Truck Cost ($)	Total Cost ($)	Trucks Used	Transportation Cost ($)	Deficit Penalty
30	492.63	5	42.63	300.0
40	549.47	5	49.47	300.0
60	632.54	5	32.54	300.0
70	682.74	5	32.74	300.0
100	834.48	5	34.48	300.0

Appendix I: Monday - Fixed Truck Cost Variation

Fixed Truck Cost ($)	Total Cost ($)	Trucks Used	Transportation Cost ($)	Deficit Penalty
40	187.25	4	27.25	0
60	265.85	4	25.85	0

70	303.37	4	23.37	0
100	432.6	4	32.6	0

Appendix J: Wednesday : Change in Deficit Penalty

Deficit Penalty ($)	Total Cost ($)	Trucks Used	Transportation Cost ($)	Truck Fixed Cost ($)
30	461.34	5	31.34	250.0
40	526.9	5	36.9	250.0
60	653.21	5	43.21	250.0
80	762.66	5	32.66	250.0