Optimization of citi bike

Introduction of CitiBike

CitiBike is the largest bike sharing system in the US, located in New York City. This bike sharing system consists of specially designed durable bikes that are locked into docking stations across the city, intended to provide people an easy and environmentally friendly option for traveling. Customers who are members of the CitiBike system can unlock bikes at any station and return them at any station, depending on their traveling needs.

Motivation

In 2013, there were more than 30,000 CitiBike members. More and more people started to use CitiBike for its convenience and accessibility. Such increasing demand also caused problems for the company. Popular stations around New York City got filled up very quickly at certain times throughout the day. For example, during rush hour, docking stations near subway stations are filled quickly and become unavailable for further use. Also, for instance, docking stations in the Central Business District of lower Manhattan get crowded during business peaks.


REUTERS

Analysis & Opinion | Felix Salmon

The one big problem with NYC's bikeshare

By Felix Salmon | June 5, 2013

DAILY NEWS

Citi Bike racks up complaints from riders over operational issues that continue to linger

Riders are still dealing with a slew of issues, although the Citi Bike operator says they've made significant improvements in recent months. Earlier this year, city officials urged the bike sharing system to address its troubles.

Ten Percent of Citi Bike Docks Appear to Fail Each Day

These inconveniences caused people to file angry reports. To solve the problem, the NYC government threatened to fine CitiBike if too many angry reports were received in any given day. CitiBike needs solutions to improve customer satisfaction.

Problem Statement

CitiBike needs a system to rebalance bike docking stations so that they are neither empty nor full.

Solution

CitiBike's current solution incorporates the use of three central hubs that store bikes and transporting trucks. These trucks travel through NYC to rebalance full and empty stations as needed. An image of the hubs and stations are seen in the map below:

citi bike

Since the data for average demand for bikes at each station was unavailable, we estimated these demands using numbers we gathered from the CitiBike website at noon on a Monday. We chose to create a model that would find the optimal hub to base Citibike's operations and to find the optimal path the trucks must take to rebalance these station nodes.

To solve this optimization problem, we used a Pick Up and Delivery problem to set up a mixed integer non-linear programming model. The objective function of this model is to minimize the total cost of travel, while constrained to create a path for trucks to pick up and deliver bikes along.

Programs Used

Python to parse data and create distance matrix

GAMs to run model

Excel to sort results

Results

The GAMS output showed that there was an optimal starting hub that would allow CitiBike to minimize the distance, and thus time, needed to travel to all the bike stations that need rebalancing. The optimal hub in our set of nodes was the Delancey Street hub.

The optimal path for the trucks would be:

Optimal Path
Vehicle 1: 3→14→8→12→11→10→20→18→15→21→22→3
Vehicle 2: 3→23→7→6→16→4→13→5→19→17→9→3

If CitiBike takes this path, its minimum cost would be $15.69. This was the cheapest path found.

Summary of Interesting Runs

Under this model we only visited one hub. The first series of what-ifs that we looked at was what-if we started with different hubs? Under this premise, the transportation costs based on the starting hubs are as follows:

Starting Hub	Transportation Cost
1	$17.78
2	$18.72
3	$15.69

Another what-if we assessed was "what if we could visit other hubs to restock (eliminating the need to visit all the nodes)?" The optimal path then became:

Optimal Path
Vehicle 1: 1→6→13→4→7→16→5→14→19→17→9→1
Vehicle 2: 1→2→8→12→11→10→20→3→18→21→15→23→22→1

Using this model the minimum transportation cost was $16.98. Thus, visiting all 3 hubs provided a comparatively more efficient transportation than starting at hub 1 and 2.

We examined other scenarios, but another scenario we found interesting was allowing the models to adjust the number of bikes a station provides. If a certain docking station had more bikes in general than others, for example, it would supply more bikes to be picked up. The resulting paths were:

Optimal Path
Vehicle 1: 1→14→8→12→11→22→20→10→21→5→16→23→1
Vehicle 2: 1→6→13→4→7→15→18→19→17→9→1

This resulted in a transportation cost of $15.96. This was another interesting result since balancing both demand and supply did not give the cheapest outcome, but rather balancing just supply did. Thus, sometimes an intuitive solution is not always more optimal when the model is actually run.

Final Recommendations and Expanding the Model

Using our model, we demonstrated that we could optimize the path a fleet of vehicles could take in order to rebalance the subset of chosen bike stations. After reviewing these what-ifs, we recommend the company chose the lowest transportation cost of $15.69. This was a result of CitiBike starting at Delancey Station.

As we proved through the interesting examples given, the most intuitive solution isn't always what is the most cost efficient. That is what makes modeling the problem so important. In fact

there are many ways this model can be used to test out other different scenarios. For instance, based from the results of allowing multiple hubs, we predict that if CitiBike expands the model to not only adjust the supply it can take from each station but also incorporate the use of all three hubs, it may realize greater savings. This is because when they leverage these hubs we recognize that it decreased travel costs comparatively to starting at just hub 1 or hub 2 alone for instance. The parameters needed to be changed are included in the Appendix (Interesting Runs, part 8) but was not a scenario we tested ourselves. It is just one of the many possibilities for changes in parameters that are possible with this model. CitiBike can use the model we created and expand the number of nodes, number of hubs its trucks visits, number of vehicles used, or try different parameters by leveraging its abundant computing resources.

Lastly, in future iterations, it would be beneficial to incorporate a function that more accurately forecasts demand. For instance, CitiBike could estimating demand by using the average customer needs during a wider time period across multiple months, instead of taking data just one snap shot in time as was done in this report. Powerful use of data and application of MINLP programming could take these better averages, predict bike docking trends by doing so continuously, and use the data to predict demand and deliver to bikes likewise.