PRESCRIPTIVE ANALYTICS

THE FINAL FRONTIER *for* EVIDENCE-BASED MANAGEMENT *and* OPTIMAL DECISION MAKING

DURSUN DELEN

Prescriptive Analytics

Node 3 $$X_{1,3} + X_{2,3} + X_{4,3} + X_{5,3} + X_{6,3} + X_{7,3} - X_{3,1} - X_{3,2} - X_{3,4} - X_{3,5} - X_{3,6} - X_{3,7} = 0$$

Node 4 $$X_{1,4} + X_{3,4} + X_{6,4} - X_{4,1} - X_{4,3} - X_{4,6} = 0$$

Node 5 $$X_{2,5} + X_{3,5} + X_{7,5} - X_{5,2} - X_{5,3} - X_{5,7} = 0$$

Node 6 $$X_{3,6} + X_{4,6} + X_{7,6} - X_{6,3} - X_{6,4} - X_{6,7} = 0$$

Node 7 $$X_{3,7} + X_{5,7} + X_{6,7} - X_{7,3} - X_{7,5} - X_{7,6} = 1$$

To represent this algebraic formulation in an Excel sheet, we use square arrays for both the variables and the distances. As shown in Figure 2.20, on the main diagonal of the distance matrix, all the numbers are set to zero, and everywhere else the actual distance is used for all defined arcs. (The arc between nodes 1 and 3 has a distance of 40 meters.) For the undefined arcs, an arbitrarily large distance such as 999 is entered. The purpose of this high number is to discourage the algorithm from ever selecting the arc as part of the solution. In the flow matrix, the rows are summed in column I and the columns are summed in row 10. The net flow formulations are entered in column J. The Solver specification is shown in Figure 2.19. The objective function in cell K14 is computed by using the SUMPRODUCT function to multiply the two square matrixes (cells in B3:H9 and the corresponding cells in B13:H19). Solver minimizes cell K14 by changing variable cells B3:H9, subject to K3:K9 = M3:M9. The Excel representation of the network model and its optimal solution obtained using Excel Solver is shown in Figure 2.20.

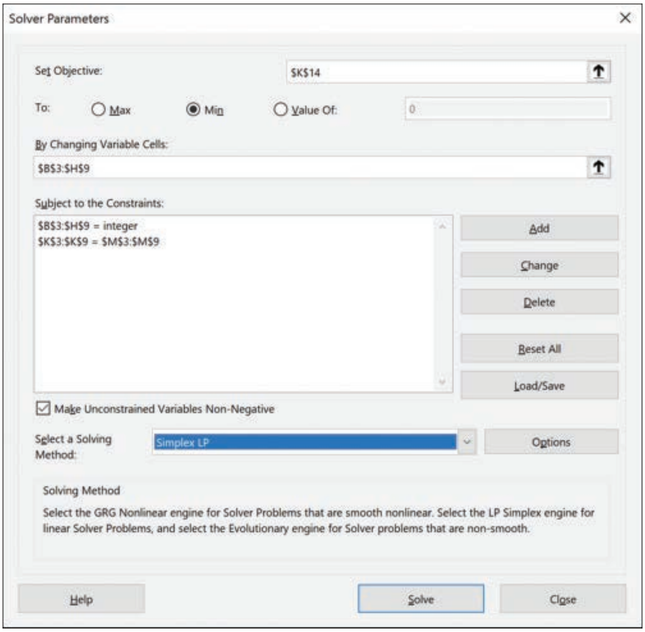
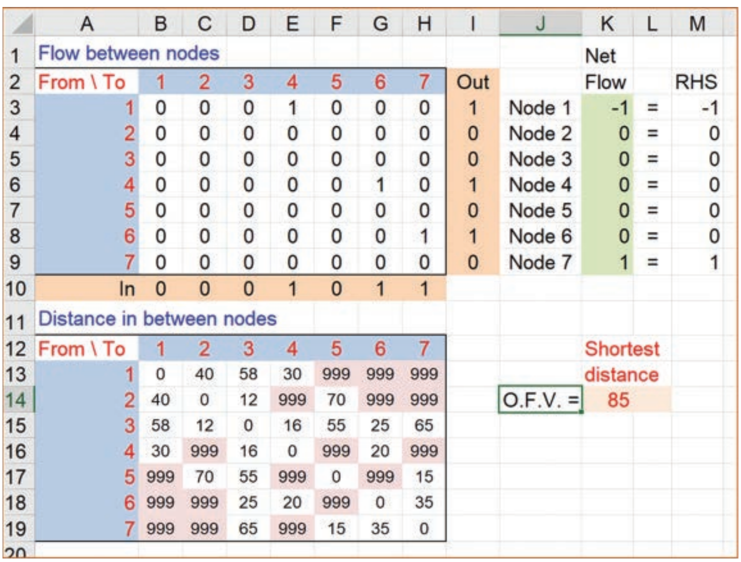


Figure 2.19 Excel Solver definitions for the network problem.



	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Flow between nodes										Net		
2	From \ To	1	2	3	4	5	6	7	Out		Flow		RHS
3	1	0	0	0	1	0	0	0	1	Node 1	-1	=	-1
4	2	0	0	0	0	0	0	0	0	Node 2	0	=	0
5	3	0	0	0	0	0	0	0	0	Node 3	0	=	0
6	4	0	0	0	0	0	1	0	1	Node 4	0	=	0
7	5	0	0	0	0	0	0	0	0	Node 5	0	=	0
8	6	0	0	0	0	0	0	1	1	Node 6	0	=	0
9	7	0	0	0	0	0	0	0	0	Node 7	1	=	1
10	In	0	0	0	1	0	1	1					
11	Distance in between nodes												
12	From \ To	1	2	3	4	5	6	7			Shortest		
13	1	0	40	58	30	999	999	999			distance		
14	2	40	0	12	999	70	999	999		O.F.V. =	85		
15	3	58	12	0	16	55	25	65					
16	4	30	999	16	0	999	20	999					
17	5	999	70	55	999	0	999	15					
18	6	999	999	25	20	999	0	35					
19	7	999	999	65	999	15	35	0					
20													

Figure 2.20 Description of the transportation model and optimal solution.

As shown in Figure 2.20, the shortest path has a distance of 85 meters. The shortest path itself is composed of the non-zero decision variables in the flow matrix, which are $X_{1,4}$, $X_{4,6}$, $X_{6,7}$. This is also the arcs 1 → 4 → 6 → 7. If we want to find the shortest path between a different pair of nodes, not much work needs to be done on the user's part. Suppose that we want to know the shortest path between nodes 4 and 5. There is no change to the objective function. In the constraints, node 4 rather than node 1 would have a –1 (minus 1) on the right-hand side, and node 5 rather than node 7 would have a 1 on the right-hand side.

Analytics Success Story: Boston Public Schools Use Optimization Modeling to Consolidate Stops, Improve Student Experience, and Save Money

With public school districts across the U.S. often underfunded, any money that can be redirected into core educational efforts is a boon for schools—and students. Cost savings can lead to more teachers, better facilities or new books, supplies, and technology.

Due to these pressures, Boston Public Schools (BPS) began to look for ways to reduce costs while improving educational outcomes. As part of this effort, the transportation department turned to its busing system to examine how it could make changes that could benefit the classroom, the traffic, and the budget. With an annual transportation budget of $120 million, compromising nearly 10 percent of the district's overall appropriation, any savings could have a significant effect.

BPS turned to SAS Analytics to optimize its bus routes, improving quality of service to students while using fewer buses. SAS used BPS data to optimize the best bus routes and stops to meet the needs of its students. As a result, the district has been able to redirect the money saved toward enhancing educational quality.

A Better Way to Get from Point A to B

The oldest public school system in America, BPS operates 125 schools serving 57,000 students from pre-K through 12th grade. In 2016, the district provided transportation for 25,000 students via 650 buses across 45,000 miles. This adds up to 20,200 unique stops at nearly 5,000 locations each day.

"We have a generous assignment process where students have a wide array of school choice," says John Hanlon, chief of operations for BPS. "Plus, because of the locations of special education programs or English as a Second Language programs, it adds to a complicated transportation system. Students are transported across the city in the morning and afternoon, close to an hour each way in some cases."

This effort can cause on-time performance challenges that affect students and their families, as well as an expensive transportation system for the school district to maintain. The district used the same legacy bus stops year after year.

The logistics behind these stops largely weren't adjusted over time, which was driving up costs. The district wanted to consolidate and streamline routes. The software used to plan bus routes required significant manual adjustments, making it impossible to rapidly evaluate the systemwide impacts of any changes.

"When assigning stops, the district had applied blanket rules to all students," says Will Eger, strategic project manager. "For example, every student should walk less than half a mile to his or her bus stop. In reality, individual students were walking a wide range of distances. Nor did the stops take into account important factors like the student's age or neighborhood safety. We needed a way to automatically account for all of those factors without dramatically increasing costs."

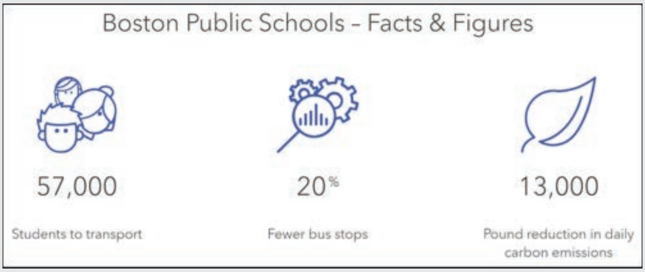


Figure 2.21 Summary of improvements. *(© 2019 SAS Institute Inc. All Rights Reserved.)*

Strategic Route Consolidation Quickly Yields Benefits

The district needed a method to more strategically designate bus stops. SAS Analytics analyzes numerous factors to reduce the number of bus stops to cut costs while better serving the needs of students. The transportation group can enter different sets of constraints into SAS Analytics and see the impact on the total number of stops across the entire system.

For example, the transportation group looked at how many stops it could cut under various route consolidation scenarios. "SAS enables us to understand the policy implications of various tradeoffs throughout the system," Hanlon says. "Before, we didn't really control for the average distance that students walked by their grade or age, and we didn't control for neighborhood safety. Now, we can integrate that along with location and other information to better serve the students."

BPS has begun rolling out individualized walk-to-stop maximums for each student—all while reducing the number of bus stops. Strategic stop placement has been a big part of the story for a district that last year eliminated nearly 50 buses (about 8 percent of the total) for a long-term cost savings that's expected to top $5 million.

The reduction in force has also had significant environmental benefits. "We've saved about 13,000 pounds of carbon emissions a day, which is huge," Eger says.

"Because of SAS, we've found new ways to consolidate bus stops, which leads to savings for the school districts without putting students in unsafe situations, and keeping them close to home in terms of the distance to their bus stops," Hanlon says. "It also allows us to think differently about the power of analytics and what it can bring to the transportation system as a whole."

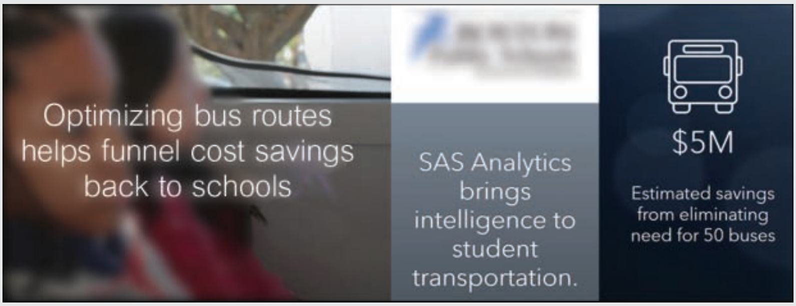


Figure 2.22 SAS Customer Success Story, 2018 *(© 2019 SAS Institute Inc. All Rights Reserved.)*

Optimization Modeling Terminology

As you may have already noticed, optimization modeling has its own language, special terms, and definitions. Following are the most common optimization-related terms and their brief definitions:

- **Decision variables** are the key components of any optimization modeling problem. The optimization modeling and analysis are conducted to identify the values of these variables. In a product mix problem, the decision variables are the optimal quantities of production quantities of all product types.
- **Objective function** is the mathematical representation of the objective of a given optimization problem. The goal of the optimization function is either to minimize or to maximize the objective function. It is usually represented using decision variables and numerical parameters.