**Automatic Design Algorithm of Public Resources in Community Delivery: Take Bookmobiles in Madison as an Example**

Dingyi Zhou\* Zishan Bai\*

University of Wisconsin-Madison

**Executive Summary**

This paper introduces an automatic design algorithm aiming to provide the optimized way to deliver the public resources in the community and how it is applied to a case study of using public demographic data to design a mobile library route in Madison. The case study with the corresponding design algorithm follows two steps. First, we develop a method to generate a utility map and pick the most necessary block groups as the station candidate of the mobile library in the city of Madison. Second, we provide a tool to automatically set up the optimal schedule of the mobile library. By our algorithm packaged in Python, people could get a schedule by only inputting the original demographic information ultimately. The algorithm has two advantages over others. Firstly, users can explore the condition of each block group by visualization, so that it is much easier to justify the need in every area through our algorithm. Secondly, by the flexibility provided by the self-customed factor, users can adjust the layers and weights by themselves to meet the updated demand. Therefore, this algorithm introduced has the potential to make a further impact on other decision processes for public resources.

1. **Introduction** To better service for each community, the Madison Library System deploys a mobile library named “Dream Bus” besides stationary libraries. The Dream Bus visits a dozen locations throughout the Madison area every weekday. Patrons could access resources like a regular library on the Dream Bus, like browsing, checking out and returning items, placing and picking up holds, using the free WiFi, etc (“Dream Bus” n.d.). During Covid-19, the service of stationary libraries is limited whereas the Dream Bus as a mobile library provides more flexibility. However, the current schedule of “Dream Bus” without deliberate consideration of the demographic distribution limits the service for a larger group of people. Therefore, by the algorithm, this paper intends to improve the service of the mobile library system in the city of Madison, including its stations and its schedule, to serve more people in the special time better and hopefully reach more people in the long run. More importantly, this paper achieves the goal that people are able to regenerate the updated optimal schedule for each period based on the change of the external environment.
2. **Current Condition of Madison Library System**

The Madison Public Library System, which has enriched the City of Madison for more than 140 years, is governed by a nine-member Library Board (Anonymous 2011). In 2019, the circulation in the system is 2,389,461, the equivalent of 6545 circulations each day (“2019-Annual-Report” 2019). Compared with the average circulation per day in the US for each public library, which is about 230 circulation, the average circulation in Madison is more than three times, reaching about 727 (“The State of America’s Libraries 2019: A Report from the American Library Association” 2019). Thus, the public libraries in Madison provide immense support for readers and educators, and its optimization will enhance the fact further.

i) Stationary Libraries

The resources of stationary libraries are not evenly distributed in the city, where some people in the suburbs have less access than others living downtown. In Figure 1 (JELLYCAT [2020] 2020), those stars symbolize the locations of libraries, while the different layers represent different accessibility levels of library resources. The green one indicates people could reach a library within 15 minutes, while the yellow one indicates a 30-minute zone and the red one indicates 45 minutes. The red edge demonstrates that there should be other supplementary library stations to serve people in those broader areas.

At the same time, the Covid-19 pandemic also causes great trauma to the operation of the traditional library. To demonstrate how library visitors’ flow changes during COVID-19 and identify the busy time, we generate Figure 2 as a heat map about circulation drop following the COVID-19 outbreak. Each library is labeled on the y-axis, and the time is labeled on the x-axis. Across all time slots at all libraries, circulation dropped 81%-92% following the COVID-19 outbreak.

Another notable fact is that the pandemic has a larger influence on juveniles than adults. Figure 3 demonstrates the total monthly averagely circulated time in 2020, divided by that in 2019. According to the data, the book and video for adults have decreased by 51% and 62% respectively, while for juveniles have decreased by 55% and 72%. Thus, it is necessary to provide more access to library resources for juveniles during the pandemic via the mobile library (Bai and Zhou 2020).

ii) Mobile Libraries

The mobile library named “Dream Bus” is also a significant part of the Madison library system. In the first eight months, its staff issued 393 library cards and served more than 5,000 visitors who checked out 5,664 items (“2019-Annual-Report” 2019). In particular, the mobile library plays an important role during the Covid-19 pandemic, while it still reflects some issues during the operation. In Figure 1, the blue dots symbolize current Dream Bus stations visiting a dozen locations, especially Community Centers or Primary Schools throughout Madison, and the blue lines with arrows demonstrate their daily route.

In general, most Dream Bus routes are far away from stationary libraries, so to a certain extent, the circulation pressure of the library has been relieved. Nevertheless, current stations, as shown above, would have difficulty reaching more people. In Figure 1, only two of the Dream Bus stations cover the red area, while what most of them do cover is the green area. Therefore, people in the red area still cannot reach these services from Dream Bus efficiently. Adding that the distribution of the current routes is not evenly distributed, some routes have four stations, while others only have two stations. Such route design not only fails to allow people to enjoy the same resources but also does not conform to the idea of low carbon and environmental protection by moving a long distance within a day. Given that the investment in the mobile library is limited, how to arrange the route to optimize people as much as possible will be a question that needs to be considered.

1. **Methodology for Scoring Schedules**

Motivation

As mentioned above, the current mobile library routes are concentrated on primary schools or community centers, but they do not fully cover all areas with high demand for mobile libraries. Therefore, the current mobile library services fail to achieve the goal of larger coverage and a wider audience. The demand for mobile libraries is not just among teenagers, but also among the elderly, as well as low-income families with no cars thus cannot get to the library.

Refer to Figure 1, consider those factors, the current mobile library stations are unable to deliver a great mission, because not many communities that are far away from the city center have been covered by the current schedule, and most current stations are all within the green 15-minutes zone or yellow 30-minutes zone. In order to meet the needs of different groups for mobile libraries, we use the method of establishing a Utility Map to find areas with high demand for mobile libraries. The reason why Utility Map is used to find suitable mobile library sites is that it cannot only flexibly select the categories and preferences of audience groups but also intuitively see the demand level of each block group for mobile library and the demand situation of each audience group in this area.

Firstly, we set up an evaluation system to evaluate the current schedule by calculating total utilities based on the picked factors, the total “score” of each station, the percentage of driving time on service time, and the extent to match the daily circulation to get an “evaluation score”. After getting our optimal schedule, we will use this evaluation system and compare the “evaluation score” of these two.

Method Description

To generate Utility Map, we first need to select necessary factors that influence the design of mobile library routes. The schedule designer can pick vital factors that are highly related to what they want to serve as the factors of the utility map, confirming the weighted average of selected factors depending on the local importance of this factor. There are two pieces of information needed. The one is the score of each block group, another is the total utilities of each block group based on selected factors. The final evaluation of potential sites is based on these two data. With the purpose of serving different kinds of people, we tried to cover the data of teenagers, seniors, and poverty families in our model. The detailed method is explained as follows.

*Step 1: Browse data from the United States Census Bureau and set up all factors.*

In our model, we picked seven factors including the population of seniors, teenagers, and the number of families under the poverty level, and the driving time to go to the stationary library. The unit of the data we used here is for each block group, and this data is available from the United States Census Bureau. Those data could be downloaded from the website directly. These data are regarded as the evaluation of each block group to express the degree of demand for mobile libraries. Except for the driving time to go to the stationary library, the absolute number and the percentage of other factors are both emphasized in our raw data, which means that for each factor such as seniors or teenagers. Our model includes the population of teenagers, the population of seniors, and the percentage of teenagers in each block group as well as the percentage of seniors for each block group. Also, the poverty families also have their raw values and percentage values. The reason for that is to avoid data misreading caused by too large or too small cardinality of the block group. For example, a group with a total population of 200 and 100 adolescents actually has a greater demand for mobile libraries than a group with a total population of 1000 and 200 adolescents.

In detail, the data related to teenagers will remain here since teenagers and students will always be the biggest service groups of mobile libraries in the city of Madison. Besides teenagers, seniors are also added in order to expand the scope of the audience. The poverty condition is also needed because the high demand for mobile libraries is from the low-income families who do not have cars to go to the stationary libraries. The more low-income families a block group has, the more families below the poverty level this block group has. Therefore, we set poverty as a factor and the driving time to the stationary libraries is also served for considering people who do have cars but live far away from the center of the city. Finally, the population density and the total circulation are also added in order to meet the requirement of the Covid-19 pandemic. Although the condition is becoming better, controlling the patron flow and social distancing are still very necessary. Therefore, using these two factors helps us design the optimal schedule during the pandemic.

*Step 2: Standardization of all data to get the “score”*

Based on those factors explained above, inventing a vehicle to represent the status of development in every community could serve our purpose of expressing the necessary. Considering this, a “score” is developed to represent the general level of each block group. The way to calculate the “score” is to sum all of the standardized factors up to get the final “score” of each block group.

The reason to standardize every category is to enable data in every category computable. Consider the fact we have the demographic data as well as the distance to libraries, those data are incalculable since they have different measurement units. Through standardization, manipulations between those data are legitimate. Another reason for the standardization is to fully focus more on the comparable level in the city rather than the absolute value, which benefits our decision for comparison. After we browse our data, the total “score” is predetermined and associated with its community, since those “scores” are transferred from the original data and those data are immobile. The total score is bigger than zero indicates the comprehensive level of this block group over the average level of all block groups.

*Step 3: Calculate total utilities and get the utility map*

After confirming which factors that we need to consider and how much proportion each factor covers, we can generate the utility map and select the planned number of stations. The utility map of each factor will serve as a layer on the total utility map. The Total Utilities are the sum of the utility maps of each factor in the following formula.

*Total Utilities = factor1 × weight1 + ··· + factori × weighti, where i =1, 2, 3, …, n*

In our model, there are seven factors and the ratio of preference Settings of these seven factors is 1:1:1:1:1:1:2, with a higher proportion of the distance to go to the stationary libraries. After plugging in all of the data and weight into the above function, the total utilities of our model were visualized as the final utility map in Figure 5. The red areas represent the utilities above the average level, and the blue areas represent the utilities below the average level. The shade of the color represents the level of demand for mobile libraries in the area and the darker the color, the greater the need for mobile libraries. Compare Figure 5 with Figure 1, it is consistent that the darkest areas are concentrated on or beyond the red 45-minutes zone in Figure 1, which explains the unreasonable design of the current schedule again. Therefore, it is necessary to use the total utility map to choose reliable sites as the mobile library stations again.

1. **Methodology for generating the schedule**

Motivation

In the last section, we select those targeting places based on our purposes and the data. Once those places are selected, how to arrange the schedule, i.e., making the bus route more efficient is another concern for our project. In terms of efficiency, we define it as minimizing the total traveled distance for the whole schedule, since it reduces the carbon dioxide emission and at the same time, spending less time on the routes, thus benefiting a larger population in the community. In addition, after calculating the total circulation of the historical data in 2019, we found the different total circulation of each weekday, so how to arrange the routes to meet for the different circulation is another concern. Therefore, this section talks about how to automatically generate the optimal schedule after selecting the sites.

Description of Method

After calculating the total utility, we need to select the potential good sites and the mobile library stations. We followed the original plan of the Madison Library System to select 12 stations as the final stations in our model. There are two steps to choose the location. The first step is to find the block groups where total utilities are bigger than zero. The reason is that it is not necessary to add new stations at block groups where their total utilities are below the average degree of demand, which is also a way to control the expense of mobile libraries' construction.

The potential stations are finally selected from these points whose total utilities are over zero. At this time, all of the data needed to choose potential stations are completed, and we make final decisions through the “score” calculated before after confirming the number of stations needed. The reason is that selecting block groups with the higher total score helps us to let mobile libraries satisfy different kinds of people as much as possible. Since total utilities over zero have selected points that perform well under chosen factors, choosing top n “scores” means that we find the best n stations that not only satisfy the preferences but also serve most kinds of people.

To find the optimal schedule of the mobile library in the city of Madison, we compared the different sites' grouping algorithms and selected K-medoids as the proper ones. The K-medoids algorithm gives the optimal grouping of the 12 block groups selected. With the idea of the minimum distance for each route, we design the algorithm that selects the optimal route for each day, and then we can calculate the total “score” for each day and match them with the total circulation. The detailed method is explained as follows.

*Step 1: Choose the Best sites Grouping Algorithm*

Considering the possible results, we propose several methods to try to minimize the total traveled distance. The first one is to generate 100 random routes and keep the best one for use. The second one is to use the K-means method, an unsupervised machine learning algorithm, to identify possible optimized routes. Among those possible optimized routes, we compare each other in the final result and keep the best one. And the third one is to use the K- medoids method, which is a revision of the K-means. And we would also use the same procedure to find the best of that as K-means did. All three methods are described below in detail.

a. Random Generate

The thought behind the random generate is that we could get the most optimized route by listing all the possibilities and comparing all of them to get the result. The algorithm of Random Generate applying here is:

Step 1: Randomly pick 100 combinations of those 12 locations

Step 2: Compute the total distance within every combination

Step 3: Return the minimized total distance with the schedule

Considering our case with 12 possible locations dividing into 5 days, the possible route would be

12! = 479001600

Because our tool is run by a personal computer, which means that computational resources are unable to deliver such tasks. Therefore, the workload is too big to complete. Thus, the expected possibility to get the result that is optimized is about 0, thus really rarely.

b. K-means

K-means is a centroid-based algorithm, or a distance-based algorithm, where we calculate the distances to assign a point to a cluster. In K-Means, each cluster is associated with a centroid. K-means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean (cluster centers or cluster centroid), serving as a prototype of the cluster (“K Means Clustering | K Means Clustering Algorithm in Python” n.d.). The algorithm of K-means applying here is:

Step 1: Run the K-means method 100 times

Step 2: Compute the total distance within every combination

Step 3: Return the minimized total distance with the schedule

K-means is an unsupervised machine learning algorithm, where the results are contingent. Meanwhile, the divide between those K-means methods could not divide those stations equally, i.e., it could have more stations in one day than that in another day. In other words, K-means can only be used when the number of stations can be equally divided by 5.

c. K-medoids

A k-medoids algorithm is a clustering approach related to k-means clustering for partitioning a data set into k groups or clusters. In k-medoids clustering, each cluster is represented by one of the data points in the cluster. These points are named cluster medoids. The term medoid refers to an object within a cluster for which average dissimilarity between it and all the other members of the cluster are minimal. It corresponds to the most centrally located point in the cluster (“K-Medoids in R: Algorithm and Practical Examples - Datanovia” n.d.).

In contrast to the k-means algorithm, k-medoids choose actual data points as centers and thereby allows for greater interpretability of the cluster centers than in k-means, where the center of a cluster is not necessarily one of the input data points. Because k-medoids minimize a sum of pairwise dissimilarities instead of a sum of squared Euclidean distances, it is more robust to noise and outliers than K-means. The algorithm of K-medoids applying here is:

Step 1: Run the K-medoids method 100 times

Step 2: Compute the total distance within every combination

Step 3: Return the minimized total distance with the schedule

Since the centroid of K-medoids is the point included, it avoids the problem that the algorithm only can be employed when the number of total sites is divided by 5. Therefore, as long as we set up the maximum and the minimum number of each group, K-medoids can give us the optimal group condition.

d. The result of K-medoids

Using the different methods, we could find that the K-medoids present the most usable and informative schedule since it doesn’t have any limitation of the number of planned total stations. Thus, we finally choose K-medoids as our best sites grouping algorithm. The results of K-medoids are shown in Figure 4 (Aflak [2019] 2021), which has a total of 4 results. Different colors represent different group situations, and the number near the points is the serial number of each block group. Noting that the result from K-medoids is not unitary, it is necessary to find the one with the minimum total distance of routes from the result. The details will be explained in the next step.

*Step 2: Schedule Generating Algorithm*

Schedule Generating Algorithm is the second algorithm in our model, which outputs the routes for each day whose total distance is the least. According to the previous step, we have found the appropriate groups, so in this step, we need to find the optimal route within each group. Schedule Generating Algorithm is to first calculate the distance from the center of all block groups to the other block groups and then exhaustively list all possible routes for each group. After that, it calculates the total distance of each possibility and selects the minimum one. For example, the first sub-plot in Figure 4 is one of the groupings, and block groups 1, 6, and 15 are in the same group. There are 3! =6 kinds of ways to design this piece of the route, and the Schedule Generating Algorithm will calculate the total distance of these 6 routes and select the minimum one as the final route. Each result from the K-medoids algorithm has five groups, and the Schedule Generating Algorithm will be used for each group and find the best route for each group.

*Step 3: Repeat Step 2 until all of the possibilities are covered*

As mentioned above, step 2 only generates the schedule of one grouping situation, but K-medoids gives us more than one possibility. Therefore, we need to redo step 2 for all of the possibilities and select the one with the smallest total distance as the final optimal route. In our model, the K-medoid gives us 4 different results, so we need to repeat Step 2 for four times and calculate the total distance of the five routes in each diagram as well as comparing them and getting the diagram with the minimum distance.

*Step 4: Match each route with the weekday based on the total circulation*

Following the requirement that only one route will be run every working day of the week, this step matches the order of total circulation with the total “score” of each day. Figure 6 is the total circulation for each weekday of the whole year in 2019. The day with the most circulation is Monday, and the day with the least circulation is Thursday. The day with the highest total circulation will match the route with the highest total utility. The highest utilities mean that the demand for the mobile library is the highest in this route, and the highest total circulation means that this route is the busiest. In this way, it serves people as much as possible. One thing is also worth mentioning since the time slots of the schedule are based on the busiest time of each stationary library. Referring to Figure 7 shows the busiest time of each stationary library, and the common busiest time is between 1 p.m. to 4 p.m. for each day. Therefore, we planned to set the stations between these time slots. For one thing, this way meets the assumption that people considering visiting a mobile library would prefer similar timings. For the other thing, it can release the stress from the stationary libraries so that people who live far away from the stationary libraries do not need to take a long time to enjoy the service.

The final output will be a completed schedule with the optimal routes and labels of block groups. Figure 8 is the final schedule with the minimum total distance, and Figure 9 is the visualized routes of the final schedule. The time in the schedule is based on the busiest time series of stationary libraries in the city of Madison. The different colors represent different routes and the numbers near the points are the labels for each block group, which are the same as the labels from the United States Census Bureau.

1. **Evaluation**

To prove that our new schedule is better than the current schedule, we evaluate both of them in calculating total utilities based on the picked factors, total “score” of each station, the percentage of driving time on service time, and the extent to match the daily circulation to get an “evaluation score” of these two schedules and compare them.

As evaluating the whole schedule, it is worth pointing out that some areas have a significantly higher score. The gap between those scores indicates another possible scenario, where the bookmobile could revisit the same location multiple times to supply the extra demand to benefit the whole community. Considering the inequality of supply and demand, classical economics thoughts provide several tools to analyze the situation. Among those tools, the “diminishing marginal returns” depict the dilemma and provide a robust way to search for equilibrium. Regarding the bookmobile supportive service, the law of diminishing returns describes that increasing visiting one location by one unit, while holding all other factors constant, will return a lower unit of output per incremental unit (Samuelson and Nordhaus 2001, 610). Therefore, in terms of the score and the utility of visiting one location, their benefit also known as the return of another same location will decrease. In our situation, when using utilities and score to evaluate these two schedules, we assume that the second visit will enable the benefit of the score and the utility to the power of one half, the third will be the power of one-third, etc. Thus, the final “evaluation score” with diminishing value is calculated by the following formula:

*Score = (1 - Driving proportion) × (utility2 + score + match)*

The driving proportion is the percentage of the mobile library bus time on the road in the total service time assuming the velocity of the bus is 13 miles/h (Daniel Hertz 2015), so one minus it represents the real service time. We used it times the sum of all other variables to get the final “evaluation score”, meaning that how many utilities truly realized finally. The reason why we squared utility is that this utility is the data that represents the demand for mobile libraries based on the 7 factors used in the above sections and it is much more important than the score. Therefore, squaring them means giving the utility more proportion. According to the final result, our new schedule has an “evaluation score” of 309.51, while the current schedule has an “evaluation score” of 115.71, indicating that our new schedule performs much better than the current one.

At the same time, Figure 10 visualized the current mobile library schedule based on these seven factors. Figure 10 shows that only three sites are in the red block groups, which means that it is not very efficient to serve people. Compare Figure 9 with Figure 10, it is convinced that most stations are in red block groups, and most of the sites are in the darker areas. Therefore, it is credible to use our model to get a reliable schedule.

1. **Scenario Exploration**

Besides the normal condition of the above case study, some scenarios are also worth considering.

i) Additional Buses

If the institution receives one more bus after designing the mobile libraries routes, they can back to “Methodology for Scoring Schedules” and add it into the routes which have the highest total utilities, since the highest total utility means the highest demand for the mobile library. Therefore, with one more bus in this route, it not only can serve more people but also can reduce the flow stress of this route.

ii) Winter Schedule

In the winter, the institution can reduce the number of stations for each week to get a new schedule. One possible solution is to run weekly and biweekly, and Figure 12. After confirming the stations weekly and biweekly, people can use “Methodology for Generating Schedules” separately to get the weekly schedule and biweekly schedule.

iii) Different value system

If the institution wants to change the concentrated factors later, they can input the new data into our model and get a new utility map. Changing factors or weights are flexible in our model. For example, Figure 11 is a new utility map with the factors including seniors, families, and Black or African Americans. The steps are the same as what we mentioned above.

iv) Different location selecting method

In our model, we just offer one of the possible methods to select potential locations as the mobile library stations. The utility map provides valuable information to select locations, so it is still available to use other ways considering the real conditions to choose mobile library stations.

v) Geographic Difficulty

The current mobile library routes in Figure 9 are the most efficient ones, but they did not rule out changes in routes due to physical geography. For example, the route marked on Thursday from Block 124 to Block 63 required crossing a lake, which was not realistic in the actual situation. Therefore, if such situations are encountered in practice, drivers need to choose their routes according to local conditions.

1. **Conclusion**

This paper basically introduces an algorithm following two steps to automatically generate the optical schedule of the mobile library in Madison. The “Scoring Schedule” gives people a utility map based on the flexible choice of factors, which helps people to find the potential good areas as the stations. The “Generating Schedule” generates the optimal schedule finally with the minimum total distance of routes. It also matches the order of utilities with the order of circulation for each day.

After the evaluation considering the service efficiency and diminishing return, we confirm that our new schedule is better than the current one with a higher score. Therefore, it is solid that our model realizes the goal to get the optimal schedule for a mobile library. Our research is not just limited to the city of Madison, so people can get different schedules for different cities as long as the original data is available.

Figures:

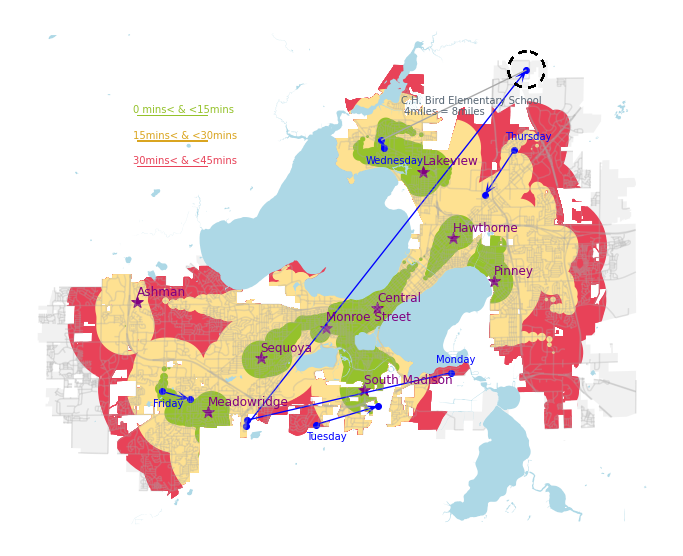


Figure 1 Current Condition of Madison Library System

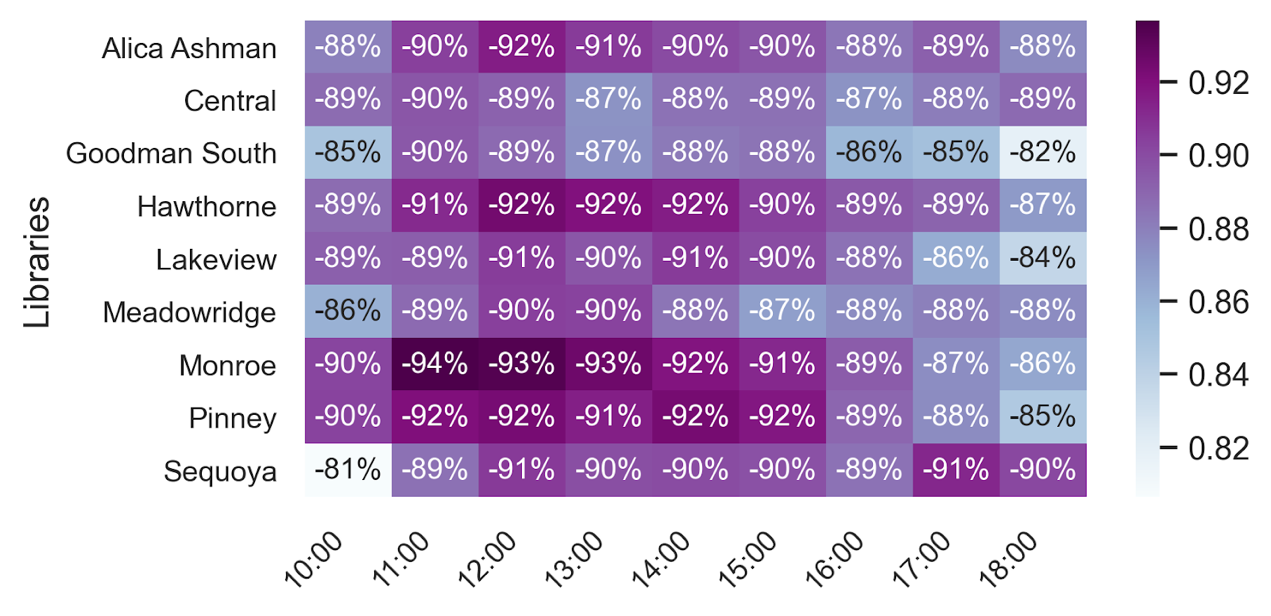


Figure 2 Circulation Drop Following Covid-19

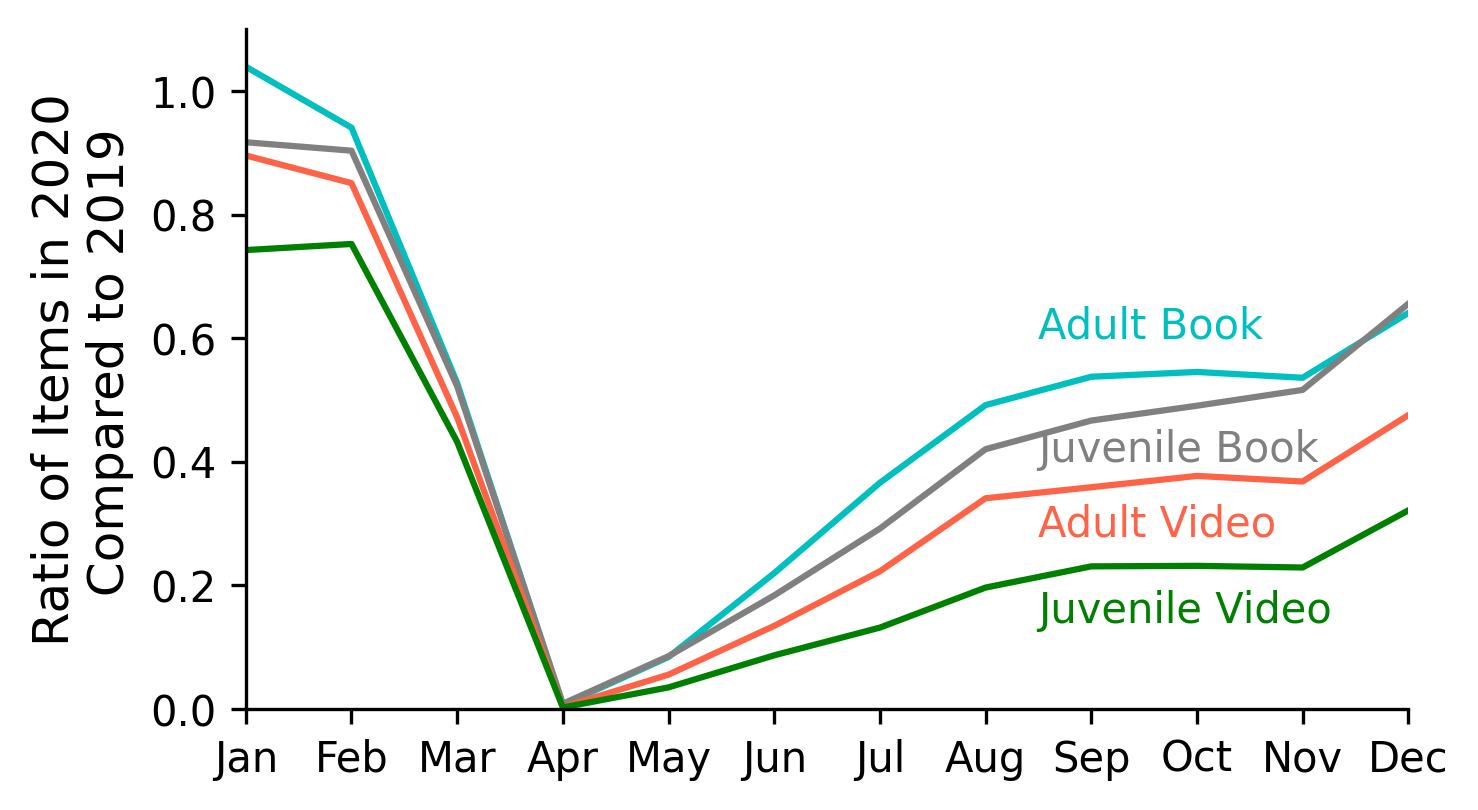


Figure 3 Circulated Time Proportion of 2020 in Different Age Group

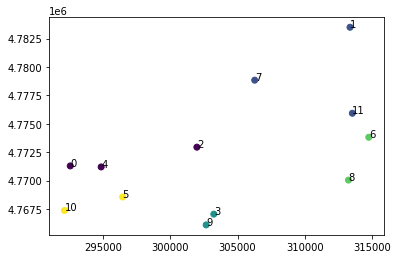
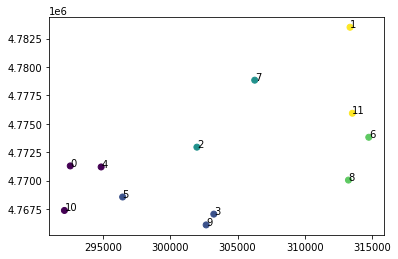
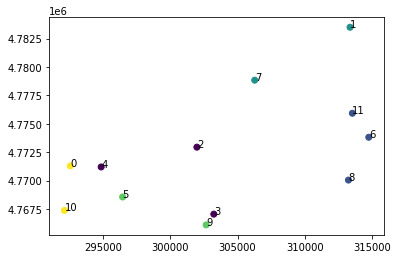
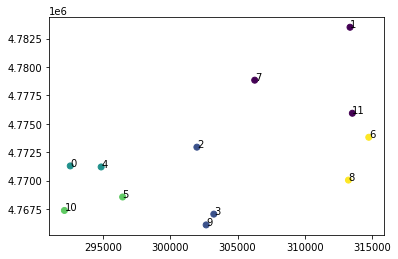


Figure 4 Dot Map through K-Medoid Algorithm

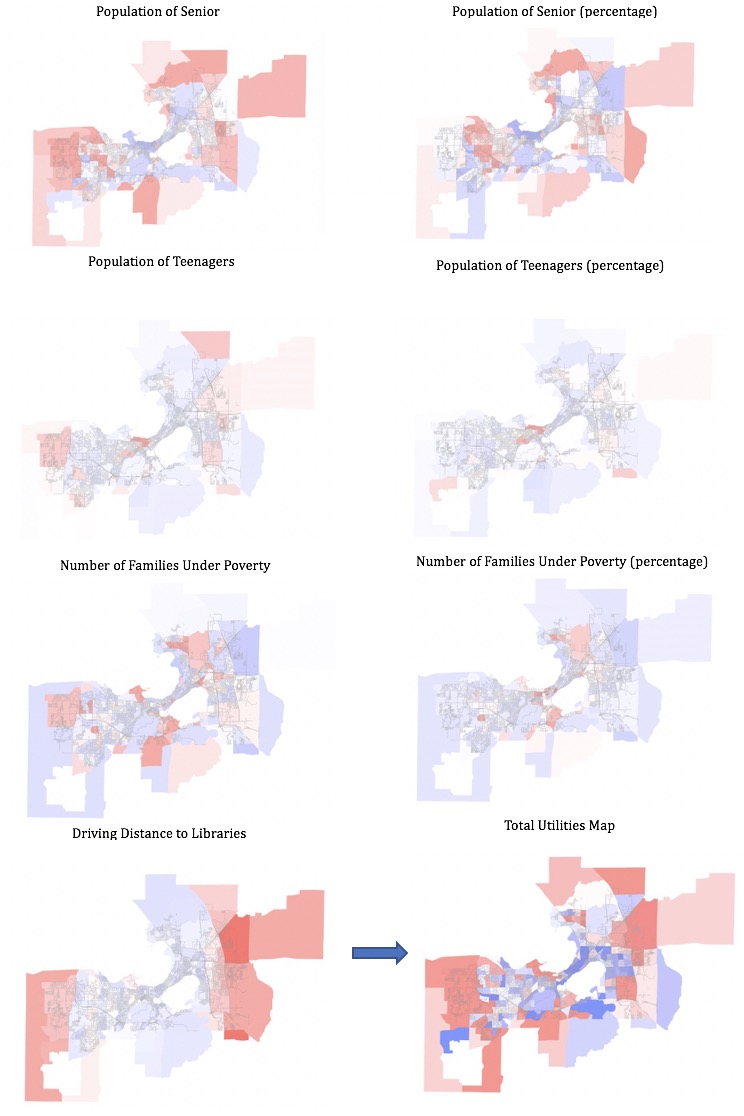


Figure 5 Process to Get the Total Utility Map

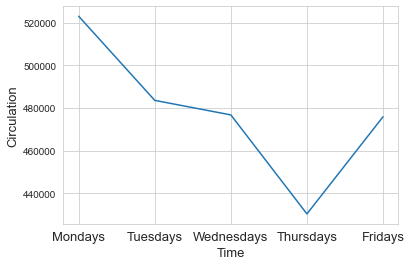


Figure 6 Total Circulation for Each Weekday in 2019

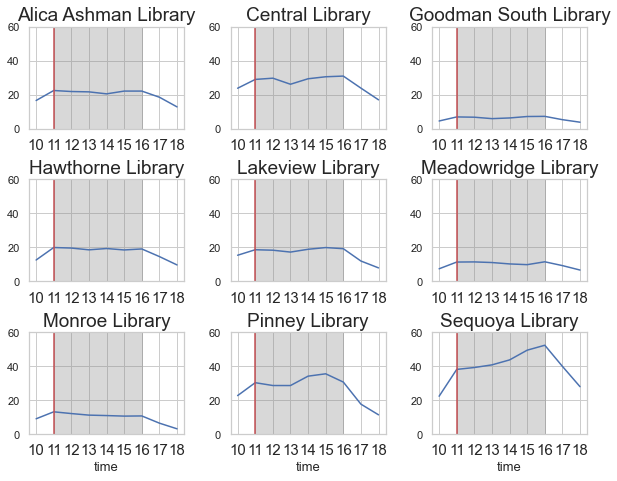


Figure 7 the Busiest time of Each Library

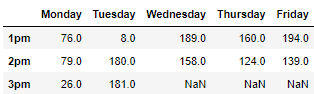


Figure 8 Final Optimal Schedule

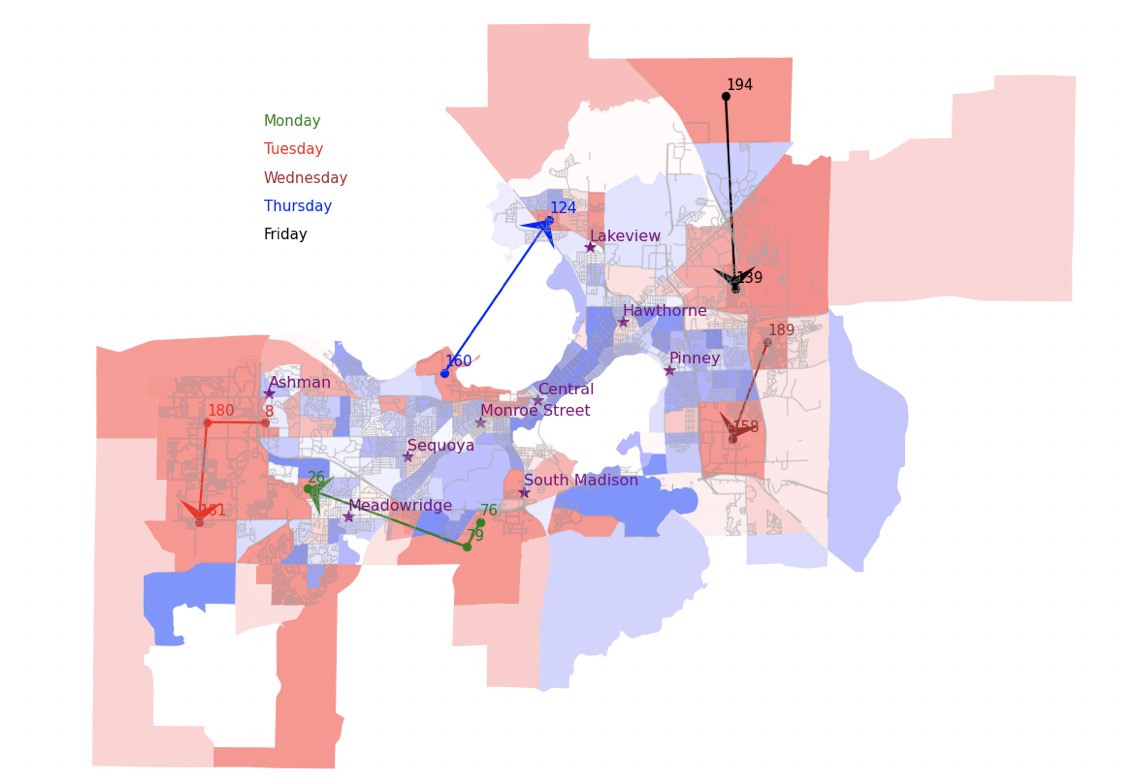


Figure 9 New Mobile Library Routes

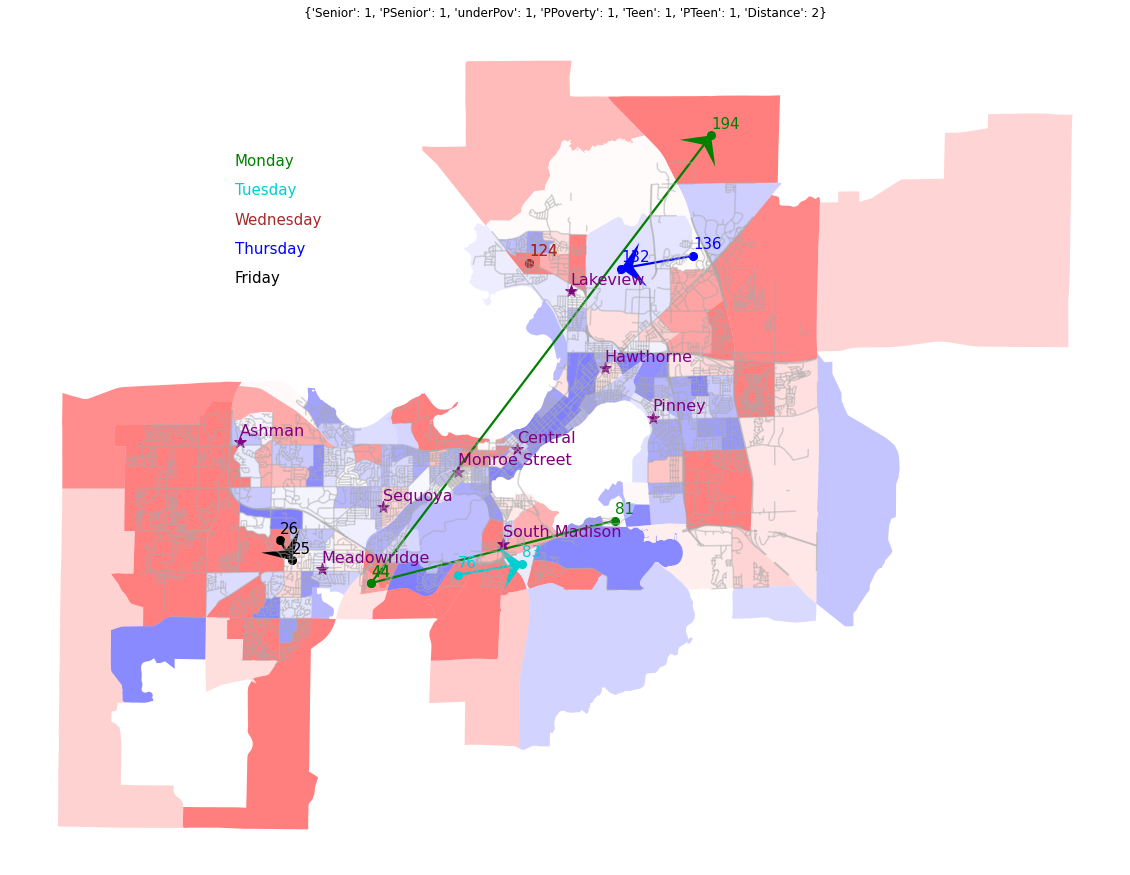


Figure 10 Current Schedule under Picked Factors

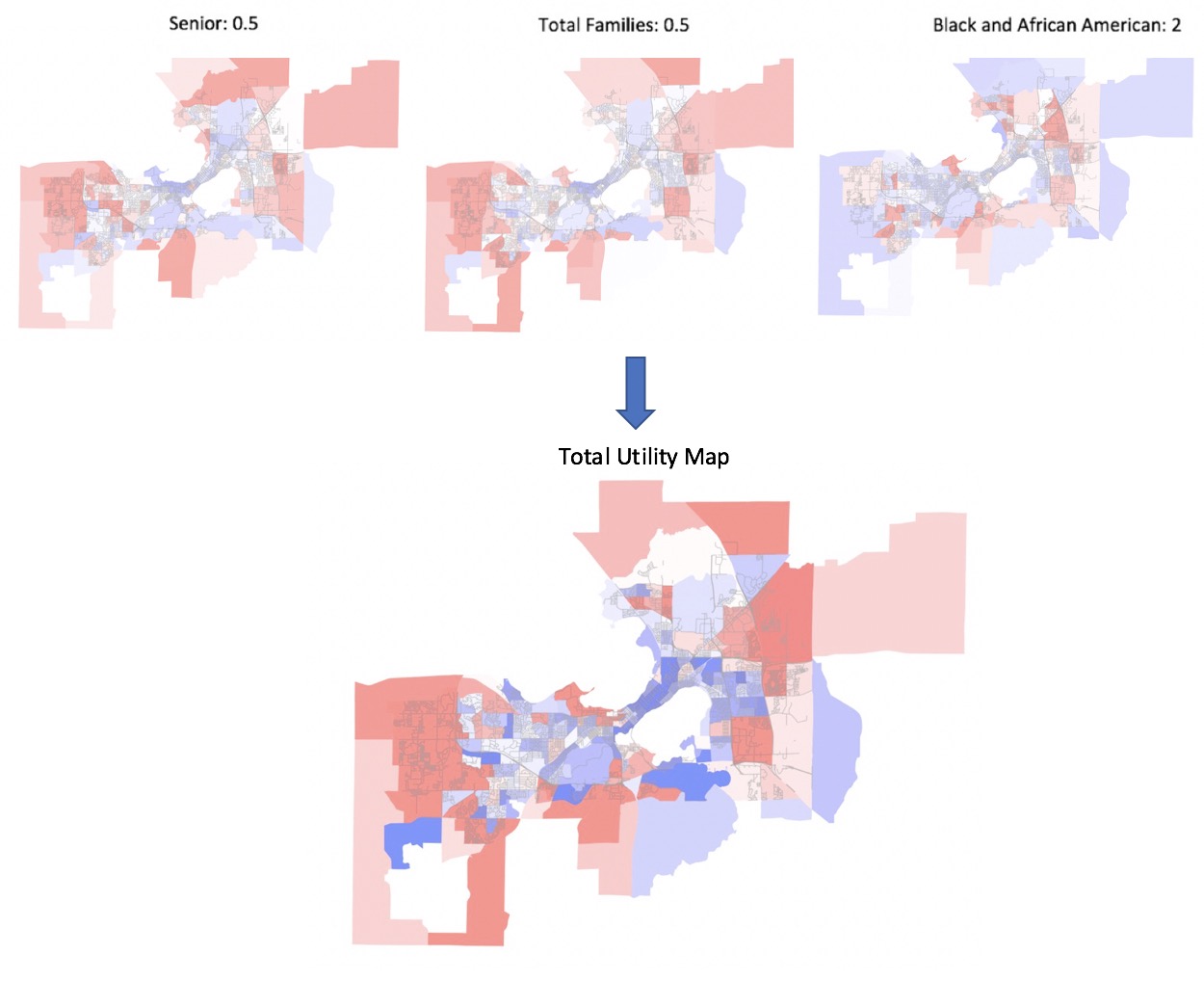


Figure 11 Utility Map Changing Factors

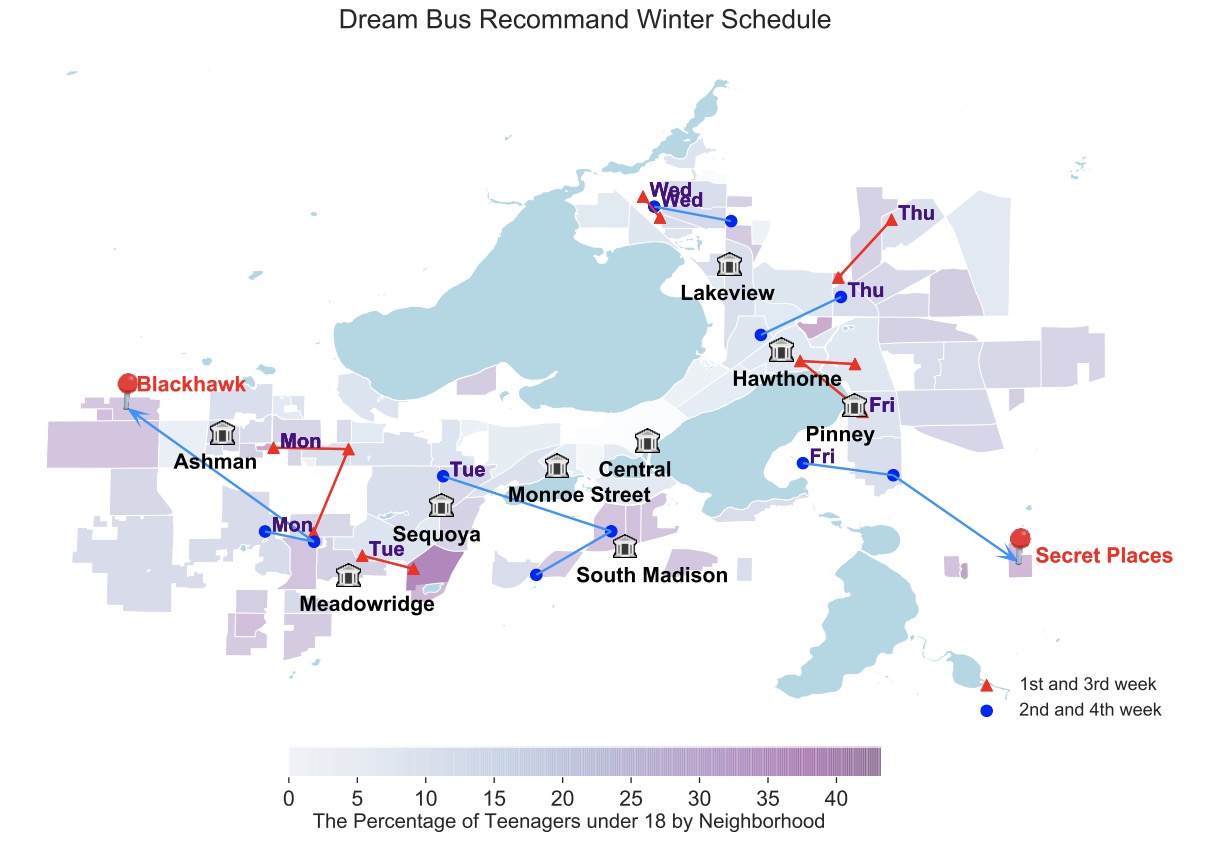


Figure 12 Dream Bus Old Winter Schedule (Bai and Zhou 2020)

**References**

“2019-Annual-Report.” 2019. Madison Public Libraray. https://mplfoundation.org/wp-content/uploads/2020/09/2019-Annual-Report.pdf.

Aflak, Omar. (2019) 2021. *OmarAflak/Kmedoids*. Python. https://github.com/OmarAflak/kmedoids.

Anonymous. 2011. “About Madison Public Library.” Text. Madison Public Library. March 28, 2011. https://www.madisonpubliclibrary.org/about.

Bai, Zishan, and Dingyi Zhou. 2020. “Dream Bus: Improving Teenagers’ Accessibility of Service via Mobile Library.” https://wisc-ds-projects.github.io/f20/projects/1-dream-bus.pdf.

Daniel Hertz. 2015. “Urban Buses Are Slowing down | City Observatory.” May 8, 2015. https://cityobservatory.org/urban-buses-are-slowing-down/.

“Dream Bus.” n.d. Text. Madison Public Library. Accessed May 27, 2021. https://www.madisonpubliclibrary.org/locations/dream-bus.

JELLYCAT, RaenonX. (2020) 2020. *RaenonX/Madison-Metro-Sim*. Jupyter Notebook. https://github.com/RaenonX/Madison-Metro-Sim.

“K Means Clustering | K Means Clustering Algorithm in Python.” n.d. Accessed May 27, 2021. https://www.analyticsvidhya.com/blog/2019/08/comprehensive-guide-k-means-clustering/.

“K-Medoids in R: Algorithm and Practical Examples - Datanovia.” n.d. Accessed May 27, 2021. https://www.datanovia.com/en/lessons/k-medoids-in-r-algorithm-and-practical-examples/.

Samuelson, Paul A., and William D. Nordhaus. 2001. *Microeconomics*. 17th ed. McGraw-Hill.

“The State of America’s Libraries 2019: A Report from the American Library Association.” 2019. American Library Association. http://www.ala.org/news/state-americas-libraries-report-2019.