ActiveTrans Priority Tool Phase II: Prioritization

The goal of Phase I was to identify a project purpose and consider community/agency values, the availability of data, and technical resources. In Phase I, factors relevant to the identified purpose and variables to represent the applicable factors were also chosen. The goal of Phase II is to apply a prioritization framework that incorporates selected factors and variables and accomplishes the purpose identified in Phase I.

Prioritization is the process of scoring and ranking improvement locations based on identified criteria or variables. Prioritization is sometimes accomplished through an iterative process (Figure 7).

Iterative prioritization processes like these may save time and resources by limiting the data or inputs needed for each successive round. Agencies should evaluate how best to structure the prioritization process given their prioritization purpose, available staff time and resources, and other considerations. However the process is structured, it is likely to include the following key steps:

Step 7: Set Up Prioritization Tool.

Step 8: Measure and Input Data.

Step 9: Scale Variables.

Step 10: Create Ranked List.

These steps are discussed in greater detail below.

Step 7: Set Up Prioritization Tool

Having established the improvement locations, factors, variables, and required data, the next step is to set up a tool to implement the prioritization method. This tool will likely use one of the technological platforms discussed in Step 6 or a combination of these approaches. Regardless of what technological platform is used, the structure of the prioritization tool will be the same: all information is organized in a tabular form.

Starting with a blank table, agencies should set up the prioritization tool using these steps:

1. List all of the improvement locations down the left side of the table. Each improvement location should be given a unique identification number (typically the furthest-left column) and a unique name. For street segments, a unique name may be expressed across three columns: the actual street name in the first column, the intersecting street where the segment starts in the second column, and the intersecting street where the segment ends in the third column. For intersections, a unique name may be expressed in a single column (e.g., Main Street &

Figure 7. Iterative prioritization.

1st Street) or in a separate column for each intersecting street (e.g., Main Street in the first column and 1st Street in the second column).

2. List all of the variables across the top of the table, starting to the right of the unique name column. Each variable should be given a name that is easy to interpret (e.g., "NumCrashes" for the number of crashes, "PedVol" for pedestrian volume). Note that it is good practice to create a separate document (data dictionary) that relates the short variable names to detailed descriptions of the variables. The data dictionary should provide information about what each variable represents, how it was measured, measurement units, and any numeric codes used to represent qualitative characteristics (e.g., 1 = "Good," 2 = "Average," 3 = "Poor"). For each variable, there are likely to be two columns: one for the raw variable data value and one for the scaled variable data value (see Step 9). Variable columns may be grouped under headings for the associated factor category. For example, population density, employment density, and number of bus stops should be grouped together under the Demand factor.
3. If weighting will be applied at the factor level, add columns for the unweighted factor score, the factor weight, and the weighted factor score for each factor. The unweighted factor score is the sum of the scaled variable values divided by the number of variables for the factor. The weighted factor score is the unweighted factor score multiplied by the weight (see Step 3).
4. If weighting will be applied at the variable level, add columns for the variable weight, and the weighted variable value. The weighted variable value is the scaled variable value multiplied by the variable weight.
5. Add two final columns at the right side of the table for the prioritization score and prioritization ranking. The prioritization score will be calculated from the variable values entered for each improvement location. The final prioritization ranking is based on the sorted prioritization scores.

Once the prioritization table is organized it is possible to proceed to Step 8, which entails measuring and/or inputting data into the tool (Figure 8).

Tip: Programmed Spreadsheet

The programmed spreadsheet developed to accompany the APT illustrates the structure described in this section. Built-in factor and variable menus, relational functions, scaling formulas, and sorting capabilities of the spreadsheet make it easier for the user to accomplish much of what is needed in setting up a spreadsheet-based prioritization tool.

Step 8: Measure and Input Data

Once the scoping phase is complete and the prioritization tool has been set up, the next step is to input the data to be used in the prioritization process. Some variables have already been measured and can be imported directly into the spreadsheet from another source. For example, the number of travel lanes on all street segments may be available from an existing roadway network database, so these values can be entered directly into the spreadsheet for each street segment being considered during the prioritization process.

Other variables may need to be measured before they can be inputted into the spreadsheet. For example, an agency may wish to include "proximity to schools" as a variable under Demand. In this case, the agency would know

Tip: Data Completeness

A data value is needed for each variable (the value may be zero) and in some cases it may be necessary to "clean" data as it is inputted, for example, to ensure that blank values (missing data) are addressed in a consistent and appropriate manner.

		Stakeholder Input	Safety	Demand	
ID	LOCATION	Number of Requests/Comments	Total Bike Crash	Population Density	Proximity to Schools
1	CENTRAL AVE	15	8.0	9539.0	3.0
2	WASHINGTON/JEFFERSON	10	3.0	9068.0	2.0
3	3RD ST	23	5.0	4664.0	4.0
4	12TH ST	2	7.0	3018.0	4.0
5	15TH AVE	1	5.0	4505.0	4.0
6	ENCANTO BLVD	15	5.0	6586.0	0.0
7	OSBORN RD	21	6.0	8924.0	7.0
8	OAK ST	9	6.0	7426.0	7.0
9	20TH ST	5	8.0	7115.0	8.0
10	3RD/5TH	3	1.0	7084.0	8.0
11	DEER VALLEY DR	8	8.0	8382.0	6.0
12	UNION HILLS DR	12	3.0	9459.0	0.0
13	19TH AVE	14	3.0	6766.0	7.0
14	32ND ST	21	1.0	7858.0	3.0
15	40TH ST	8	4.0	4678.0	5.0

Figure 8. Example prioritization tool structure with inputted data.

school locations but would need to calculate the distance between each improvement location and the nearest school in order to determine the proximity.

The location type of the improvement locations is a key consideration when measuring variables. For example, if the location type is an intersection or crossing and the variable is "number of crashes," the variable would most likely be measured as the number of crashes within a certain distance of the intersection or crossing. If, on the other hand, the location type is a roadway corridor, it may be necessary to normalize the data by the length of the corridor in order to avoid giving more weight to longer corridors simply because they are longer. In this example, the variable would likely be measured as the number of crashes within a certain distance of the corridor divided by the length of the corridor.

The level of effort required for measuring data will depend on the complexity of data and the technological platform being used. For example, the task of counting pedestrian crashes within 50 feet of each intersection can be done quickly using GIS, but it takes much longer to do by hand. In general, spatial measurement tasks can be done most efficiently using a GIS platform. However, there are a variety of other tools that can be used to measure variables (see Step 6).

If there are very few improvement locations and/or variables, the task could simply require manually entering data into a spreadsheet. If existing datasets will be used, it may be possible to sort and then "cut and paste" data into one central location. If using GIS, data entry typically involves joining different datasets together through the use of an identification field or "common key." The common key is an attribute that two data sets have in common, such as a street segment ID number. Using a common key ensures that the data order and integrity are maintained as several sources are combined.

Step 9: Scale Variables

The purpose of Step 9 is to ensure variables, which are represented by different units, are comparable. Step 9 involves converting non-numeric values to numeric values, selecting a common numeric scale, and adjusting raw values to fit the common scale. Scaling should not be confused with weighting. Scaling is a more objective, technical function, while weighting is based on

Table 23. Example of converting non-numeric values to numeric values.

Non-Numeric Value	Numeric Value
Excellent	4
Good	3
Fair	2
Poor	1

community/agency values. In other words, agencies should not attempt to increase or decrease the influence of variables through scaling.

Scaling is necessary so that variables have a comparable impact on the prioritization score in the absence of weighting. Consider, for example, a prioritization process that includes both "Speed" and "Number of Transit Stops Within ¼ Mile" as variables. A typical value range for speed might be 15 to 70, while a typical value range for transit stops might be 0 to 6. In the absence of scaling, the top end of the "Speed" variable is 11 times greater than the top end of the "Transit Stops Within ¼ Mile" variable, which would result in the "Speed" variable having a far greater impact on the prioritization score than the "Transit Stops Within ¼ Mile" variable.

Assign Numeric Values to Non-Numeric Variables

Variables with non-numeric values must be converted to numeric values before they can be incorporated into the prioritization framework. Examples of non-numeric values include categories such as "yes" and "no"; "compliant" and "non-compliant"; "high," "medium," and "low"; or "excellent," "good," "fair," and "poor." Converting these values requires ranking them and assigning numeric values by rank. The highest numeric value should go to the non-numeric value with the highest rank, the next highest numeric value to the non-numeric value with the next highest rank, and so on. Table 23 illustrates this process.

Select a Common Scale

There are many potential considerations that go into selecting a common scale; however, the main consideration is likely to be ease of calculation. For this reason, it is recommended that agencies select a common scale that is 0 to 1 or 0 to 10.

Tip: Adjusting the Common Scale for Specific Variables

For some variables, using a common scale may overemphasize the difference between the highest and lowest raw values for that variable. For example, consider an agency that is prioritizing improvements along arterial roadways and wants to consider 85th percentile speed as part of their analysis. The lowest 85th percentile speed in this dataset may still be quite high. Should an agency assign it a scaled value of 0? In such cases, it may be appropriate to establish a variable-specific scale with the same maximum scaled value as the common scale but with a minimum value that is higher, for example, 3 or 5 instead of 0.

Table 24. Common scaling methods.

Appropriate If Range of Raw Data Values *Does Not Include* Outliers	Appropriate If Range of Raw Data Values *Includes* Outliers
• Proportionate Scaling and Inverse Proportionate Scaling • Non-linear Scaling and Inverse Non-linear Scaling	• Quantile Scaling and Inverse Quantile Scaling • Rank Order Scaling and Inverse Rank Order Scaling • Jenks Natural Breaks Scaling and Inverse Jenks Natural Breaks Scaling

Adjust Values to Fit the Common Scale

Once a common scale is selected, it is necessary to adjust the raw values for each variable to fit the common scale. There are several ways to do this, depending on the distribution and relative importance of the values associated with each variable. Some common methods are shown in Table 24, which divides the methods into two categories based on their appropriateness for addressing outliers (i.e., minimum or maximum values that are much larger or much smaller than other values).

Whatever method is chosen, it should be evenly applied to all data values based on a consistent rule or formula, and the rule or formula should be documented.

Additional details regarding these methods and the issue of outliers is provided below.

Tip: Programmed Spreadsheet Scaling Formulas

The programmed spreadsheet that accompanies the APT allows users to select from a menu of scaling options, including:

- Proportionate scaling.
- Inverse proportionate scaling.
- Quantile scaling (4 quantiles).
- Inverse quantile scaling (4 quantiles).
- Quantile scaling (10 quantiles).
- Inverse quantile scaling (10 quantiles).
- Rank order scaling.
- Inverse rank order scaling.

Proportionate Scaling and Inverse Proportionate Scaling

If the range of values does not include outliers, then it is appropriate to adjust the raw numeric values proportionately to fit the common scale. Proportionate scaling involves assigning the highest value in the common scale to the maximum raw value for a particular variable and assigning the lowest value in the common scale to the lowest scaled value. Other raw values are scaled proportionately based on their relationship to the highest and lowest raw values. Inverse proportionate scaling is similar but involves assigning the lowest scaled value to the maximum raw value and the highest scaled value to the lowest raw value. The formula for proportionate scaling is:

$Y = (X - MIN)/(MAX - MIN) \times S$, where Y is the scaled value, X is the raw value, MIN is the minimum raw value, MAX is the maximum raw value, and S is the scale.

The formula for inverse proportionate scaling is:

$Y = (((X - MIN)/(MAX - MIN) \times S) - S)\ X - 1.$

In Table 25, the maximum raw value is 5, the scale is 0 to 10, and the raw values are adjusted using proportionate scaling.

Table 26 is the same as Table 25, except that the raw values are scaled using inverse proportionate scaling. An example for which inverse proportionate scaling might be applied is implementation costs: projects with lower implementation costs may be assigned higher scaled values if the prioritization purpose is to implement a greater number of lower cost projects.

Table 25. Example of proportionate scaling for a scale of 0 to 10.

Raw Value	Scaled Value
4	8
0	0
3	6
4	8
5	10
3	6
2	4
0	0
5	10
1	2

Note: *In this example, the minimum raw value is 0 and the maximum raw value is 5. Each of the raw values has been adjusted proportionately to fit a scale 0 to 10, and the resulting scaled values are shown in the right-hand column.*

Proportionate scaling and inverse proportionate scaling may not be appropriate if the range of values to be scaled includes outliers. In this case, proportionate scaling may result in a maximum or minimum scaled value that is much higher or lower than the next highest or lowest scaled value, which may be undesirable for a variety of reasons. One key reason is that it diminishes the level of differentiation between the majority of values and may skew the final prioritization rank for the outlier improvement location. There are several methods for addressing outliers when they are a concern, including quantile scaling and rank order scaling.

Quantile Scaling and Inverse Quantile Scaling

If the range of values includes outliers, it may be more appropriate to calculate scaled values based on quantiles. Quantile scaling involves assigning each raw value to a quantile (i.e., equal groups containing the same number of values) and scaling the quantile values proportionately to fit the selected scale. In Table 27, raw values for a variable are divided into six equal groups. Then, the quantile values are scaled proportionately to fit on a 0 to 10 scale. Note that there are two data values for each quantile. Most spreadsheet tools contain functions that allow computation of quantiles.

Table 26. Example of inverse proportionate scaling for a scale of 10.

Raw Value	Scaled Value
4	2
0	10
3	4
4	2
5	0
3	4
2	6
0	10
5	0
1	8

Note: *In this example, the minimum raw value is 0 and the maximum raw value is 5. Each of the raw values has been adjusted proportionately.*

Table 27. Example of quantile scaling using 6 quantiles.

Raw Value	Quantile	Scaled Value
16	1	0
17	1	0
22	2	2
24	2	2
26	3	4
32	3	4
33	4	6
36	4	6
37	5	8
41	5	8
48	6	10
150	6	10

Note: *In this example, the minimum raw value is 16 and the maximum raw value is 150. 150 is also an outlier, since it is more than three times larger than the next highest raw value. To address this, the raw values are sorted from low to high and divided into 6 quantiles of 2 values each. The quantile values are then adjusted.*

Table 28 is the same as Table 27 except that the raw values are scaled using inverse quantile scaling.

Please note that quantile scaling is not appropriate when multiple instances of the same data value would have to be separated into more than one quantile. For example, if there are 20 data values for a variable and 10 of them are 0, dividing the data into 10 quantiles results in two 0s being classified in the first quantile, two 0s being classified in the second quantile, and so on through the fifth quantile. In such cases, the methods described below may be more appropriate.

Rank Order Scaling and Inverse Rank Order Scaling

Rank order scaling is another method for addressing outliers. Rank order scaling involves calculating the rank of each value in the range and then scaling the rank values proportionately

Table 28. Example of inverse quantile scaling using 6 quantiles.

Raw Value	Quantile	Scaled Value
16	1	10
17	1	10
22	2	8
24	2	8
26	3	6
32	3	6
33	4	4
36	4	4
37	5	2
41	5	2
48	6	0
150	6	0

Note: *In this example, raw values are scaled using deciles. There are 20 raw values and breaking these values into deciles involves dividing them into ten equal groups. Consequently, there are two records per decile.*

Table 29. Example of rank scaling.

Raw Value	Rank	Scaled Value
0	1	0
0	1	0
0	1	0
0	1	0
5	2	2
7	3	4
9	4	6
10	5	8
32	6	10

Note: *In this example, the minimum raw value is 0 and the maximum raw value is 32. 32 is also an outlier, since it is more than three times larger than the next highest raw value. To address this, the values are ranked from low to high (i.e., the lowest value gets a rank of 1, next lowest value gets a rank of 2, and so on). The ranked values are then scaled proportionately.*

to fit the selected scale. In Table 29, the raw values for a variable are ranked from low to high. Then the ranked value is adjusted proportionately to fit a 0 to 10 scale.

Table 30 is the same as Table 29 except that the raw values are scaled using inverse rank scaling.

There are several other scaling methods that agencies may wish to consider if they are confronted with the issue of outliers. One possible approach is to give the outlier either the maximum or minimum scaled value, and then to scale the remaining values proportionately based on a range that excludes the outlier. Another approach is to use Jenks natural breaks. Jenks natural breaks are determined using a mathematical formula that assigns classes so that the average deviation from the class mean average is minimized while each class's deviation

Table 30. Example of inverse rank scaling.

Raw Value	Rank	Scaled Value
0	1	10
0	1	10
0	1	10
0	1	10
5	2	8
7	3	6
9	4	4
10	5	2
32	6	0

Note: *In this example, the minimum raw value is 0 and the maximum raw value is 32. 32 is also an outlier, since it is more than three times larger than the next highest raw value. To address this, the values are ranked from low to high. The ranked values are then scaled inverse proportionately.*

Table 31. Example of non-linear scaling.

Raw Value	Scaled Value
20	0
25	1
30	2.5
35	5
40	10

Note: *In this example, raw values are scaled in a non-linear fashion to represent the relationship between motor vehicle speed and the risk of pedestrian death in collisions involving pedestrians and motor vehicles.*

Table 32. Example of inverse non-linear scaling.

Raw Value	Scaled Value
20	10
25	5
30	2.5
35	1
40	0

from the means of other classes is maximized. The number of classes assigned depends on the chosen scale.

Non-linear Scaling and Inverse Non-linear Scaling

Finally, for some variables the importance of the raw numeric values may increase in a non-linear fashion. For example, the risk that a pedestrian will be killed in collision with a motor vehicle is 7 to 9 times higher at 30 mph than at 20 mph. Agencies may wish to incorporate this relationship in their scaling process, as in Table 31.

Table 32 is the same as Table 31 except that the raw values are scaled using inverse non-linear scaling.

Step 10: Create Ranked List

The goal of Step 10 is to create a prioritized list of projects. This involves summing the weighted values for each factor (or variable) to derive a prioritization score for each improvement location. The improvement locations are then ranked based on the prioritization score. Completion of Step 10 includes reviewing the calculations and clearly communicating the results and the process.

Calculate Prioritization Scores

At the beginning of Step 10, each variable should be scaled to a common scale and each factor should have a designated weight. To calculate the prioritization scores, follow the steps below for each improvement location:

- Calculate the *unweighted* score for each factor by summing the scaled variable values and divide by the number of variables used to get the factor score (Table 33).

Table 33. Example of calculating unweighted factor score for three improvement locations and two variables.

Improvement Location	Scaled Variable Value 1	Scaled Variable Value 2	Sum of Scaled Variable Values	Unweighted Factor Score
Location 1	3	5	8	4
Location 2	2	4	6	3
Location 3	6	4	10	5

Table 34. Example of calculating weighted factor score for three improvement locations.

Improvement Location	Factor 1 Score	Factor 1 Weight	Factor 1 Weighted Score
Location 1	4	8	32
Location 2	3	8	24
Location 3	5	8	40

Table 35. Example of calculating prioritization score for three improvement locations and two factors.

Improvement Location	Factor 1 Weighted Factor Score	Factor 2 Weighted Factor Score	Prioritization Score
Location 1	32	12	44
Location 2	24	10	34
Location 3	40	18	58

- Calculate the *weighted* score for each factor by multiplying the *unweighted* factor score by the factor weight (Table 34).
- Sum the *weighted* factor scores to get the prioritization score for each improvement location (Table 35).

Develop Ranked List

Once the prioritization scores have been calculated, the improvement locations can be ranked based on their score. The simplest approach is to give the improvement location with the highest

Tip: Scale Prioritization Scores

It can be helpful to scale the final prioritization scores using proportionate scaling to better understand the relationships between them. For example, if a common scale of 0 to 10 is selected, the improvement location with the highest prioritization score would get a scaled prioritization score of 10, and the improvement location with the lowest prioritization score would get a scaled prioritization score of 0. It's much easier to understand the significance of an improvement location that scores a 9.0 on a 10 point scale than it is to understand the significance of an improvement location that scores a 238 on scale of 135 to 250.

Table 36. Example of developing ranked list for three improvement locations.

Improvement Location	Prioritization Score	Rank
Location 3	58	1
Location 1	44	2
Location 2	34	3

score a rank of 1, the improvement location with the next highest score a rank of 2, and so on (Table 36). However, this should be done with an awareness of the relationships between prioritization scores (see *Tip: Scale Prioritization Scores*) and an understanding that small differences in prioritization scores can be the result of measurement errors.

Review Ranked List

It is important for practitioners to review the results of any prioritization scoring and ranking process carefully to understand how weighting, scaling, correlation of variables, and other issues may affect the results. The level of review should be proportional to the level of complexity of the process (i.e., the more factors and variables used, the more scrutiny the process demands). Recommended review steps include:

- Review the ranked list and/or a visual representation of the ranked list on a map. Do some improvement locations rank unexpectedly high or unexpectedly low? If so, do the raw data values make sense? Have the weighting and scaling calculations been done correctly?
- Review the scaled values for each variable to understand the impact of scaling and verify that data values are scaled appropriately.
- Review the unweighted and weighted scores for each factor to understand the impact of weighting and verify that weighting is having the intended effect.
- Review the factors and variables used. Are key policy objectives or community values being fully represented by the chosen factors or variables? Agencies have the ability to use factors and variables that are not presented in this methodology.

Communicate Results

Agencies can build confidence among stakeholders by clearly communicating the prioritization results and the process that led to them. Graphics are a highly effective form of communication. For example, pie charts can be used to show the respective weights of selected factors or variables (as in Figure 2). Maps can also effectively communicate results, allowing stakeholders to better understand how results correspond to locations of the improvement locations. Maps may also be used to further explain the prioritization process. For example, each factor category can be mapped separately and then combined into a composite map (Figure 9), giving stakeholders a better sense of how each factor category influences the prioritization results. The same can be done for individual variables.

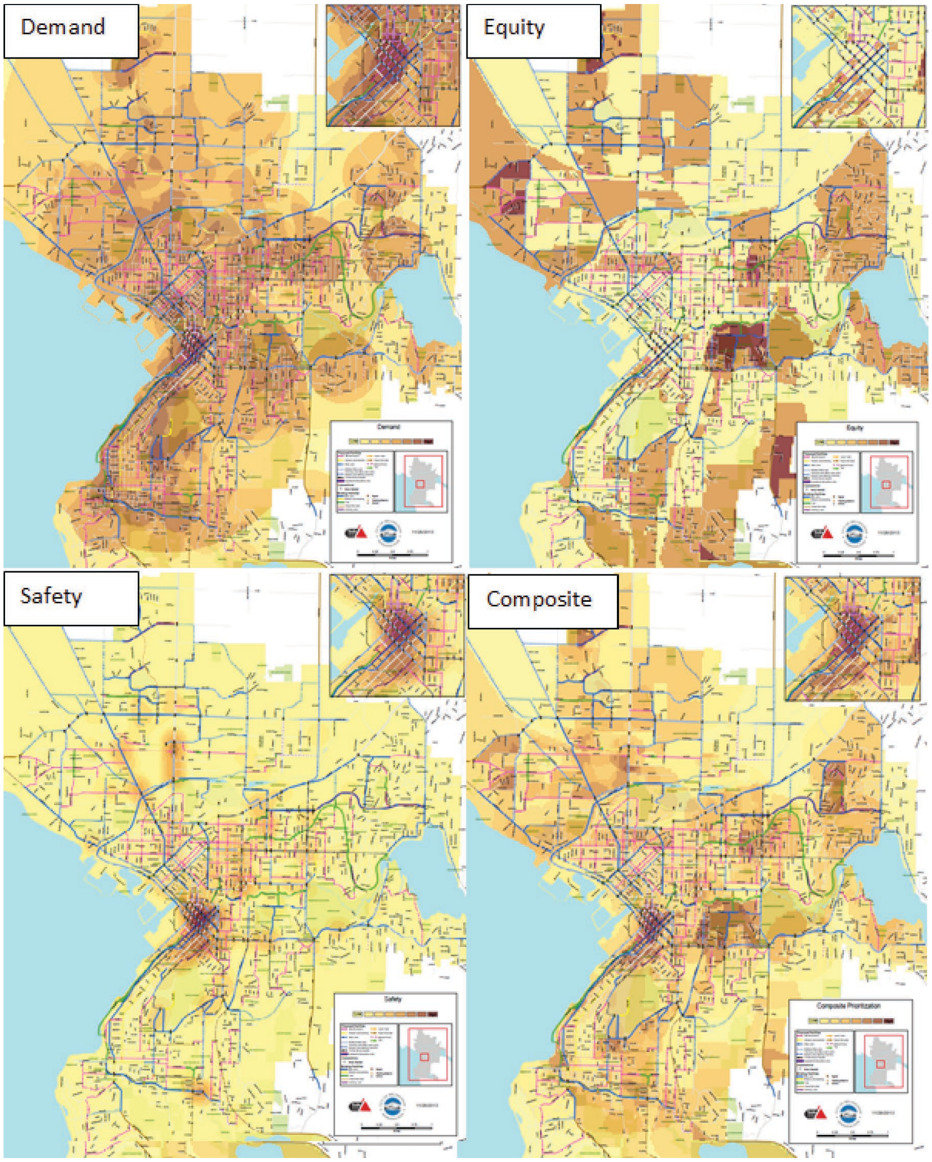

Figure 9. Communicating prioritization process by mapping selected factors—example from City of Bellingham, Washington, bicycle master plan.