Task Deliverable 4: Final Report

Contract BDV24 TWO 562-09

Project: Evaluating Community Building Effectiveness of Transportation Investments: Using Traditional and Big Data Oriented Analytical Approaches (Phase 4)

Naveen Eluru, Ph.D.

Samiul Hasan, Ph.D.

Sudipta Dey Tirtha, M.Sc.

Bibhas Kumar Dey, Ph.D.

Jiechao Zhang, M.S.

Tanmoy Bhowmik, Ph.D.

University of Central Florida

Department of Civil, Environmental & Construction Engineering

Orlando, FL 32816-2450

October 06, 2020

Table of Content

List of Tables

List of Figures

Chapter 1: Multi-criteria Decision Analysis

1.1 Introduction

According to Florida Chamber of commerce, Florida ranks number one in the US in terms of transportation infrastructure rankings. It is the third largest state by population, after California and Texas with a yearly growth rate of more than 1.5%. Orlando is the most thriving city of the Central Florida region; its growth being bolstered by its job creation rate (1,000 jobs are added per week). The economic and demographic trends suggest that Orlando has an expanding consumer market and these trends are set to drive increased demand for passengers and freight transportation in the coming years. To accommodate the future demand in an efficient and sustainable manner, several small and big transportation projects are underway in the region including second phase of SunRail commuter rail extension, I-4 expansion, and bicycle sharing system (Juice) introduction. The impacts of these investments can be classified into two broad categories: transportation system effects that result in direct benefits for system users (drivers, passengers, companies) and community (social and economic) effects that affect the community as a whole. There are well-defined performance measures, based on engineering and economic criteria, for assessing the direct system user benefits. For example, how a new facility leads to reduced journey time or reduced travel cost. On the other hand, such indicators are scarce for assessing the community impacts of transportation projects.

In the previous phase of research, to examine the community impacts of three transportation infrastructure investment projects – SunRail commuter rail extension, I-4 expansion and Juice bicycle sharing system - we proposed five community impact assessment measures or measures of effectiveness (MOE):

(1) property value change: computed as change in property value in the vicinity of the chosen projects by land use type for (1) Single family residential, (2) Multiple family residential, (3) Retail/Office area, (4) Institutional, and (5) Industrial land use.

(2) changes to job accessibility: estimated as the change in number of jobs accessible around the chosen investment projects,

(3) commuting time change: evaluated as change in commuting time in the vicinity of the chosen projects,

(4) land use type change: measured as change in vacant land use type to other land use patterns

(5) changes to travel patterns for zero car households: estimated as change in travel mode patterns for work travel

The development of these MOEs was a data intensive process. These indicators/measures were developed by collating appropriate data collected from different sources using the ArcGIS platform. While the MOE generation and evaluation provided visual and qualitative understanding of the impact of these investments, there was no overall quantitative finding from the analysis from the last phase.

In this phase of research, the research team focused on two objectives. parsing the comparison results to identify the community benefits to the region. The first objective of the research effort was to develop and implement a framework to compare the changes in MOEs across scenarios to identify benefits to the Central Florida region. Through this framework, a net positive, neutral or negative rating of a project can be provided for the three projects chosen above. The second objective of the research was on conducting an extensive knowledge transfer activity through the development and delivery of training to FDOT personnel. Specifically, through webinars and supporting manuals, we provide step-by-step guidance on the various data preparation, data download and data analysis tasks conducted for the project from all phases 1 through 4.

The report is organized as follows. Chapter 1 focuses on the framework development and implementation of multi-criterion decision analysis. Chapter 2 provides a summary of the knowledge transfer activities while supporting material is provided in the Appendices. The rest of Chapter 1 is organized as follows. Section 2 presents the outcome of literature review performed in this study. Section 3 provides a detailed overview of methodological approach. In the next section, detailed analysis is presented with final outcome of the analysis. Finally, section 5 concludes the chapter.

1.2 Literature Review

Multi-criteria Decision Making (MCDM) is a general class of operations research models that are associated with decision processes in the presence of a number of decision criteria. MCDM can be classified into two categories: (1) Multi objective decision making (MODM) and (2) Multi attribute decision making (MADM) (see Figure 1). The main distinction between the two groups of methods is based on the determination of alternatives. In MODM (also referred to as multi objective programming or a vector optimization/maximization/minimization problem) the alternatives are not predetermined but instead a set of objective functions is optimized subject to a set of constraints (Cristóbal, 2011). In MADM, where alternatives are predetermined, a small number of alternatives are to be evaluated against a set of attributes (Cristóbal, 2011). As MADM is relevant to our study context, we confine our review of literature to MADM models.

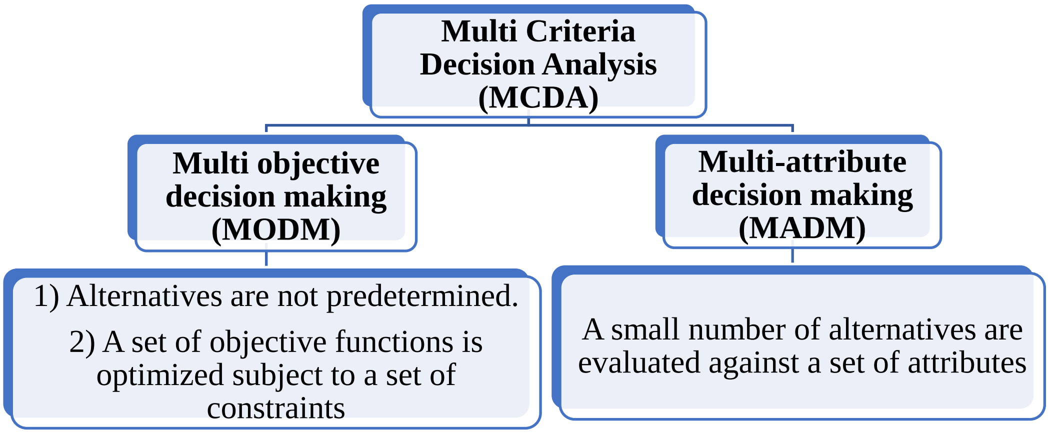


Figure 1: Classification of multicriteria decision making

1.2.1 Multi-Attribute Decision Making (MADM)

Multi-Attribute Decision Making (MADM) is a well-established branch of decision making. There are number of methods that are commonly employed as a part of MADM. While the methods differ in terms of their complexity, the overall architecture is consistent across these methods.

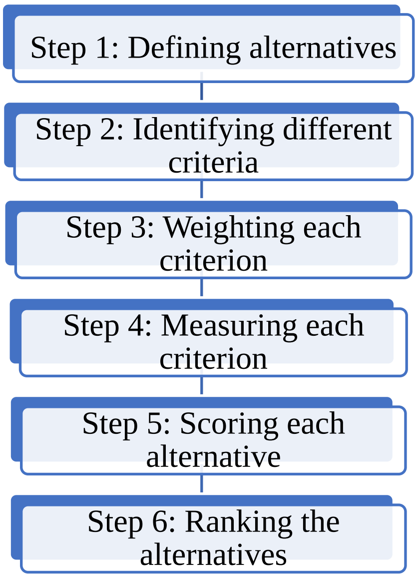


Figure 2: Hierarchical process of multi-attribute decision making

MADM models follow a hierarchical process presented in Figure 2. In the first step of the analysis, different alternatives are chosen by the decision makers. For example, expansion of existing roadway and improving transit system can be two alternatives for meeting increased transportation demand. In the second step, criteria or performance measures are defined based on project objectives by the decision makers such as travel time reduction, property value change, increasing job accessibility etc. In the third step, each criterion is assigned weightage by the stakeholders based on their judgment. For example: travel time reduction and property value change get 4 points and job accessibility gets 3 points. Then, in the fourth step, each criterion is measured for each alternative. For example: property value may change by 10% and 15% respectively for the above-mentioned alternatives. In the fifth step, each alternative is scored based on their criteria values and criteria specific weights. In the final step, alternatives are ranked based on their scores.

Table 1 provides a reverse chronological summary of the studies reviewed, specific MADM model employed, weighting method, and if the study is applied in the transportation domain. Several observations can be made from the summary. First, the various methods considered in the literature include:

(a) Weighted sum method (WSM)
(b) Weighted product method (WPM)
(c) Analytical hierarchy process (AHP)
(d) VIKOR Method
(e) Preference ranking organization method for enrichment evaluation (PROMETHEE)
(f) The elimination and choice translating reality (ELECTRE)
(g) The technique for order preference by similarity to ideal solutions (TOPSIS)
(h) Compromise programming (CP) and
(g) Multi-attribute utility theory (MAUT).

Based on the review, the most commonly employed approach is the Analytical hierarchy process. Approaches a through e have been employed multiple times in the literature. Second, the weighting approaches considered include: (a) Five points rating, (b) Point allocation, and

(c) Pairwise comparison. Of these approaches, the pairwise comparison method is the most commonly employed. Finally, the application of MADM methods in transportation literature, while prevalent, is still not very frequent. We provided details of the MADM models and weighting methods in the subsequent sub-sections.

Table 1: Reviewed Literature in MADM

Study	**Model**	**Weighting method**	**Applied for a transportation project?**
Bottero et al., 2018	PROMETHEE	Point allocation	Yes
Song and Kong, 2016	Analytical Hierarchy	Pairwise comparison	No
Til et al., 2014	--	Review of available methods	No
Cristóbal, 2011	VIKOR method	Pairwise comparison	No
Afshari et. al., 2010	Weighted sum method	Pairwise comparison	No
Da˘gdeviren, 2008	AHP and PROMETHEE	Pairwise comparison	No
Saaty, 2008	Analytical Hierarchy	Pairwise comparison	No
Macharis et al., 2007	Multi-actor Multi-criteria	--	Yes
Yan et al., 2007	PROMETHEE	Point allocation	No
Yang and Hung, 2007	TOPSIS and fuzzy TOPSIS	Five points rating	No
Haralambopoulos and Polatidis, 2003	PROMETHEE	Point allocation	No
Lai et al., 2002	Analytical Hierarchy	Pairwise comparison	No
Pohekar and Ramachandran, 2001	Review of available method	--	No
Sinuany-Stern et al., 2000	Combined Data Envelopment Analysis (DEA) and AHP	Pairwise comparison	No
Triantaphyllou and Sanchez, 1997	Review of available method	--	No
Zanakis et al., 1998	Simulation comparison between MCDM methods	Pairwise comparison	No
Saito, 1987	Analytical Hierarchy	Pairwise comparison	Yes
Saaty, 1980	Analytical Hierarchy	Pairwise comparison	No

1.2.2 Prevalent MADM Models

To describe the MADM models in detail we focus on the five most commonly adopted approaches (a-e). The models are illustrated through numerical implementations for ease of understanding.

1.2.2.1 Weighted Sum Method (WSM)

The weighted sum method (WSM) is the simplest and the most widely used MADM method. If there are M alternatives and N criteria, then the best alternative is the one that satisfies the following expression (Triantaphyllou and Sanchez, 1997):

$$A_{wsm} = Max \sum_{j=1}^{N} a_{ij}\omega_j \qquad \text{Where, i = 1, 2, 3, ……, M}$$

A_{wsm} is the WSM score of the best alternative where N is the number of decision criteria. a_{ij} is the actual value of the i^{th} alternative in terms of the j^{th} criterion and ω_j is the weight of importance of the j^{th} criterion. For each criterion, the decision maker has to determine its importance, or weight, ω_j. It is also assumed that the following relationship is always true:

$$\sum \omega_j = 1$$

To illustrate the method, consider the following example. Three Projects A, B and C are three alternative projects to be considered by the decisionmakers. Hence,

Alternative 1, A_1 = Project A
Alternative 2, A_2 = Project B
Alternative 3, A_3 = Project C

For these three projects, five measure of effectiveness or five criteria have been estimated. So, five criteria are:

Criterion 1, C_1 = Property value by land use type
Criterion 2, C_2 = Land use changes over time
Criterion 3, C_3 = Accessibility to employment
Criterion 4, C_4 = Travel commuting time
Criterion 5, C_5 = Travel patterns for zero car households

In the next step, the above-mentioned criteria need to be weighted such that $\sum \omega_j = 1$. Assume that the assigned weights of the criteria are as follows:

Table 2: Weights of different criteria

Criteria	**C_1**	**C_2**	**C_3**	**C_4**	**C_5**
Weight, ω_j	0.3	0.2	0.2	0.2	0.1

Now, each of the alternatives needs to be scored based on each criterion to get the evaluation matrix. For demonstration, evaluation matrix of the alternative is shown below:

Table 3: Value of criteria obtained from detailed evaluation

Alternatives	**C_1**	**C_2**	**C_3**	**C_4**	**C_5**
Project A	10%	20%	10%	30%	5%
Project B	20%	30%	10%	10%	15%
Project C	30%	10%	20%	10%	20%

Based on the weights assigned and the value of the criteria, each alternative can be scored. Score of alternative 1 = 0.3 × 10 + 0.2 × 20 + 0.2 × 10 + 0.2 × 30 + 0.1 × 5 = 15.5. Thus, scores of all alternatives is shown in table below:

Table 4: Scoring of alternatives

Alternatives	**Scores**
Project A	15.5
Project B	17.5
Project C	19.0

On the basis of scores, alternatives are ranked and the alternative, having the highest score, is ranked 1^{st}.

Table 5: Rank of the alternatives

Alternatives	**Rank**
Project A	3
Project B	2
Project C	1

In our example, Project C would be the preferred project.

1.2.2.2 Weighted Product Model (WPM)

The weighted product model (WPM) is analogous to the WSM. The distinction between two methods is that instead of adding the scores for each criterion, scores are multiplied. Each alternative is compared to each one by multiplying a number of ratios, one for each criterion. Each ratio is raised to the power equivalent to the relative weight of the corresponding criterion (Triantaphyllou and Sanchez, 1997). In general, in order to compare the alternatives A_k and A_L the following product is obtained:

$$R\left(\frac{A_k}{A_L}\right) = \prod_{j=1}^{N} \left(a_{kj} / a_{Lj}\right)^{\omega_j}$$

N is the number of criteria, a_{ij} is the actual value of the ith alternative in terms of the j^{th} criterion, and ω_j is the weight of importance of the j^{th} criterion. If $R\left(\frac{A_k}{A_L}\right)$ is greater than one, alternative A_k is more desirable than alternative A_L (in the maximization case). From the previous example illustrated in section 2.2.1, weights and value of different criteria are summarized below:

Table 6: Weights and value of different criteria

Alternatives	**Criteria**				
	C_1	**C_2**	**C_3**	**C_4**	**C_5**
Weight, ω_j	0.3	0.2	0.2	0.2	0.1
Project A	10%	20%	10%	30%	5%
Project B	20%	30%	10%	10%	15%
Project C	30%	10%	20%	10%	20%

In the next step, alternatives can be compared to each other based on weights and criterion measure using above equation. For example: alternative 1 can be compared to alternative 2 by the functional value of $R\left(\frac{A_1}{A_2}\right)$.

$R\left(\frac{A_1}{A_2}\right)$ = $(10/20)^{0.3} \times (20/30)^{0.2} \times (10/10)^{0.2} \times (30/10)^{0.2} \times (5/15)^{0.1}$ = 0.836

Alternative 2 is preferred over alternative 1 as $R\left(\frac{A_1}{A_2}\right)$ is less than 1. Comparison between available alternatives is shown in following table:

Table 7: Table of preference by WPM

Comparing alternatives	$R\left(\frac{A_1}{A_2}\right)$	**Preference Rank**
Project A	0.836	2
Project B	0.778	3
Project C	0.933	1

Therefore, it can be concluded that Project C is the most preferred project among the alternatives considered. The result for the most preferred alternative is similar to what was obtained for WSM approach.

1.2.2.3 The Analytical Hierarchical Process (AHP)

Analytical Hierarchy Process (AHP) is developed by T.L. Saaty (1980). In this method, the decision hierarchy is structured from the top with the goal of the decision, then the criteria to be fulfilled for achieving the goal to the alternatives considered at the bottom. For example, improving existing transportation infrastructure is the main focus of a transportation project and land value change, travel commuting time and accessibility to employment are set as the criteria. At the bottom of hierarchy, two alternatives (e.g. Project A and Project B) are selected. Based on this scenario, hierarchy can be structured as presented in Figure 3:

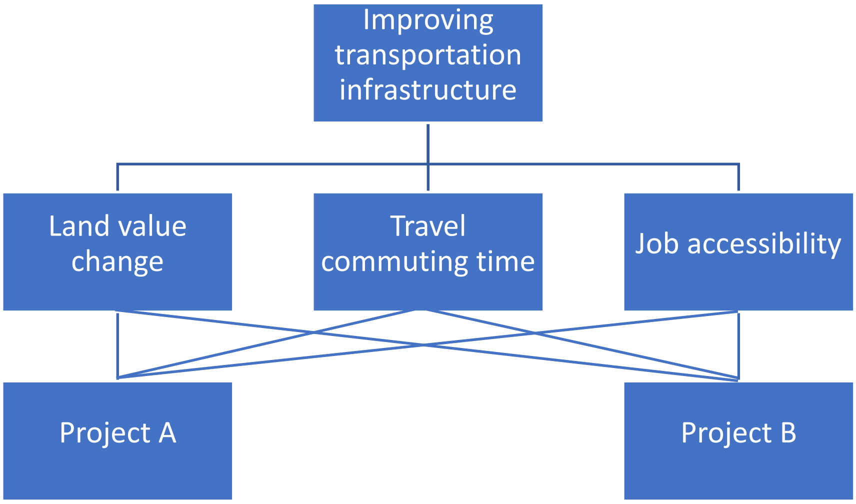


Figure 3: The analytical hierarchy process

In the following step, a set of pairwise comparison matrices is constructed. Each element in an upper level is used to compare the elements in the level immediately below with respect to it. Priority of each criterion can be derived by normalizing elements of each column and then, summing each row to get priority column. At the last step, elements of priority column also need to be normalized by their sum. For example: land value change is used to compare Project A and Project B as follows:

Table 8: Priority table by AHP

Alternatives	Project A	Project B	Priority
Project A	1	4	4/5
Project B	1/4	1	1/5

In the last step of the analysis, the priorities obtained from the comparisons and weights of criteria are used to get the overall priorities of alternatives.

Table 9: Ranking by AHP

Criteria	Land value change	Travel commuting time	Job accessibility	Overall Priority	Rank
Weights	0.2	0.4	0.4		
Project A	0.8	0.5	0.4	0.52	1
Project B	0.2	0.5	0.6	0.48	2

1.2.2.4 VIKOR Method

The Compromise Ranking method, also known as the VIKOR method, introduces the Multi-criteria ranking index based on the particular measure of “closeness” to the “ideal” solution (Cristóbal, 2011). Basic steps of VIKOR method with its application on previous example showed in section 2.2.1 are described below:

Step 1: Alternatives and criteria are defined. In this section, alternatives and criteria will be same as section 2.2.1. Table 2 and Table 3 are used for weights and value of criteria respectively.

Step 2: Determine the best f_i^* and the worst f_i^- values of all criterion functions, i = 1, 2, 3, ………, n. If the i^{th} function represents a benefit then f_i^* = max f_{ij} and f_i^- = min f_{ij} . In contrast, If the i^{th} function represents a cost then f_i^* = min f_{ij} and f_i^- = max f_{ij} . According to Table 3:

f_1^* = Max (Property value change) = 30

f_2^* = Max (Land use changes) = 30

f_3^* = Max (Job accessibility) = 20

f_4^* = Max (Travel commuting time change) = 30

f_5^* = Max (Travel pattern change) = 20

f_1^- = Min (Property value change) = 10

f_2^- = Min (Land use changes) = 10

f_3^- = Min (Job accessibility) = 10

f_4^- = Min (Travel commuting time change) = 10

f_5^- = Min (Travel pattern change) = 5

Step 3: Compute the values S_j and R_j for $j = 1, 2, 3, \ldots\ldots, J$ using the expression:

$$S_j = \sum_{i=1}^{n} \omega_i(f_i^* - f_{ij})/(f_i^* - f_i^-)$$

$$R_j = max[\omega_i(f_i^* - f_{ij})/(f_i^* - f_i^-)]$$

Thus, S_j and R_j for the three alternatives are presented in following table:

Table 10: Estimation of S_j and R_j

Alternatives	**Criteria**					S_j	R_j
	C_1	**C_2**	**C_3**	**C_4**	**C_5**		
Weight, ω_i	0.3	0.2	0.2	0.2	0.1		
f^*	30	30	20	30	20		
f^-	10	10	10	10	5		
Project A	10	20	10	30	5	0.70	0.30
Project B	20	30	10	10	15	0.58	0.20
Project C	30	10	20	10	20	0.40	0.20

Step 4: Determine Index, Q_j using the following equation and the alternative with lowest Q_j is considered to be the best alternative.

$$Q_j = \nu(s_j - s^*)/(s^- - s^*) + (1 - \nu)(R_j - R^*)/(R^- - R^*)$$

v is introduced as a weight for the strategy of maximum group utility and value of v ranges from 0 to 1 (Assumed 0.5 for this example). Here, $s^* = min_j S_j$ = 0.40, $s^- = maxS_j$ = 0.70, $R^* = min_j R_j$ = 0.20 and $R^- = maxR_j$ = 0.30. Using the above equation and Table 10, Q_j for j = 1, 2 and 3 are provided in Table 11:

Table 11: Ranking by VIKOR method

Alternatives	Q_j	**Rank**
Project A	1.00	3
Project B	0.30	2
Project C	0.00	1

1.2.2.5 Preference Ranking Organization Method for Enrichment Evaluation (PROMETHEE)

Preference ranking organization method for enrichment evaluation (PROMETHEE) performs a pair-wise comparison of alternatives (Bottero et al., 2018). Basic steps of PROMETHEE method is illustrated below:

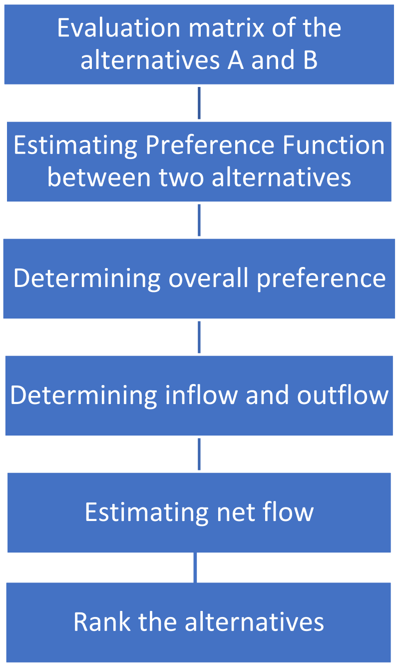


Figure 4: Steps of PROMETHEE method

The steps of PROMETHEE method is described briefly below using the example described in section 2.2.1:

Step 1: An evaluation matrix between two alternatives is constructed. This matrix accounts for deviations of evaluations on pairwise comparisons of two alternatives, A and B, on each criterion. For example: evaluation matrix shown in Table 3.

Step 2: Identify the preference function $P_j(a, b)$ for each criterion j. The method uses preference function $P_j(a, b)$ which is a function of the difference between two alternatives a and b for any criterion j. Based on threshold values for the difference, preference function is assigned values of 1 and 0 where "0" represents the indifferent preference value between the two alternatives, a and b.

$$P_j(a, b) = 1 \qquad \text{if} \quad f_{aj} - f_{bj} \geq \theta_j$$

$= 0 \quad \text{if} \quad f_{aj} - f_{bj} < \theta_j$

For the example in section 2.2.1, values of $P_j\,(a, b)$ using threshold value 10 is shown in table below:

Table 12: Preference matrix

Criteria	P (1, 2)	P (1,3)	P (2, 1)	P (2, 3)	P (3, 1)	P (3, 2)
1	0	0	1	0	1	1
2	0	1	1	1	0	0
3	0	0	0	0	1	1
4	1	1	0	0	0	0
5	0	0	1	0	1	0

Step 3: Calculate the overall preference index $\Pi(a, b)$. The overall preference index $\Pi\,(a, b)$ represents the intensity of preference of a over b and it is calculated using the following formula:

$$\Pi(a, b) = \sum_{j=1}^{k} \omega_j P_j(a, b)$$

Using above equation, Table 12 and Table 2, overall preference index for all combination of alternatives is tabulated below:

Table 13: Overall Index Calculation

Overall Preference Index	Π (1, 2)	Π (1,3)	Π (2, 1)	Π (2, 3)	Π (3, 1)	Π (3, 2)
Estimated value	0.2	0.4	0.6	0.2	0.6	0.5

Step 4: Calculate the outranking flows, i.e., positive flow $\phi^+_{(a)}$ and negative flow $\phi^-_{(a)}$. In PROMETHEE method, two flow measures can be determined for each alternative. There is a positive flow (it expresses how alternative a is outranking all the others):

$$\phi^+_{(a)} = \frac{1}{n-1} \sum_{b \varepsilon A} \Pi(a,\ b)$$

And negative flow (it expresses how alternative a is outranked by all the others):

$$\phi^-_{(a)} = \frac{1}{n-1} \sum_{b \varepsilon A} \Pi(b,\ a)$$

Positive and negative flows for three alternatives are shown in following table:

Table 14: Outranking Flow Calculation

Alternatives	**Positive flow**	**Negative flow**
Project A	½*(0.2+0.4) = 0.30	½*(0.6+0.6) = 0.60
Project B	0.40	0.35
Project C	0.65	0.30

Step 5: Compare the outranking flows to define the alternatives complete ranking. The higher the net flow, the better the alternative. The net flow:

$$\phi_a = \phi^+_{(a)} - \phi^-_{(a)}.$$

Net flow of three alternatives for the given example is tabulated below:

Table 15: Net Flow and Ranking by PROMETHEE

Alternatives	Net flow	Rank
Project A	-0.30	3
Project B	0.05	2
Project C	0.35	1

1.2.3 Weighting the Criteria

As described earlier, an important element of the MADM approaches is the development of weights for the evaluation criteria. Several weighting methods have been discussed in literature (see van Til et al., 2014). We describe three common methods that are useful for our research.

1.2.3.1 The five points rating method

This exercise is a technique in which all criteria are rated on a five-point scale ranging from (1) not important to (5) very important. From the scores obtained, weights of each criterion can be computed by normalizing the points by total sum of the points.

1.2.3.2 Point allocation (PA)

In this method, a budget of 100 points is allocated across different criteria to reflect their relative importance and these assigned points can be used as weights (by dividing them with 100).

1.2.3.3 The pairwise comparison (PC) technique

In this approach, we compare all possible pairs of criteria (e.g. Criteria A and Criteria B) on a reciprocal numerical rating scale ranging from 1/9 (extreme preference for criteria B) to 9 (extreme preference for criteria A). A description of the numerical scale for preference rating is given below:

Table 16: Preference rating point table

Preference rating	Definition
1	Equal importance
2	Weak or slightly important
4	Moderate importance
5	Moderate plus
6	Strong importance
7	Strong plus
8	Very, very strong
9	Extreme importance

* T.L. Saaty (2008)

An example of pairwise relationships for four criteria is provided in Table 17.

Table 17: Pairwise comparison Matrix

Criteria	1	2	3	4
1	1	4	3	6
2	1/4	1	3	5
3	1/3	1/3	1	2
4	1/6	1/5	1/2	1

Priorities of the criteria can be computed from the pairwise comparison matrix. First, elements of each column are normalized by sum of the column (See Table 18). Then elements of each row are summed to get priority column. Elements of priority column are normalized at the last step (see Table 18 column 6 and 7).

Table 18: Determining priorities from Pairwise comparison matrix

Criteria	**1**	**2**	**3**	**4**	**Priorities**	**Normalized Priorities**
1	0.57	0.72	0.40	0.43	2.12	0.53
2	0.14	0.18	0.40	0.36	1.08	0.27
3	0.19	0.06	0.13	0.14	0.53	0.13
4	0.10	0.04	0.07	0.07	0.27	0.07

1.3 Methodology

1.3.1 Weighting of Criteria

In this study, we will follow pairwise comparison method for determining weights of the criteria. We compare all possible pairs of criteria (e.g. Criteria A and Criteria B) on a reciprocal numerical rating scale ranging from 1/9 (extreme preference for criteria B) to 9 (extreme preference for criteria A). Numerical scale for preference rating is given below:

Table 19: Preference Rating Point Table

Preference rating	Definition
1	Equal importance
2	Weak or slightly important
3	Moderate importance
4	Moderate plus
5	Strong importance
6	Strong plus
7	Very strong
8	Very, very strong
9	Extreme importance

Source: T.L. Saaty (2008)

To illustrate the use of the process from Table 19, consider the following example. Between criteria A and B, an individual prefers A very strongly. In that case the rating for A versus B will be scored as 7 and B versus A will be scored as 1/7. Once a ranking for A versus B is available, the ranking for B versus A is simply computed as it's reciprocal. An example of pairwise relationships for four criteria is provided in Table 20.

Table 20: Pairwise Comparison Matrix

Criteria	1	2	3	4
1	1	4	3	6
2	1/4	1	3	5
3	1/3	1/3	1	2
4	1/6	1/5	1/2	1

Priorities of the criteria can be computed from the pairwise comparison matrix. First, elements of each column are normalized by sum of the column (See Table 21). Then elements of each row are summed to get priority column. Elements of priority column are normalized at the last step (see Table 21 column 6 and 7).

Table 21: Determining Priorities from Pairwise Comparison Matrix

Criteria	1	2	3	4	Priorities	Normalized Priorities
1	0.57	0.72	0.4	0.43	2.12	0.53
2	0.14	0.18	0.4	0.36	1.08	0.27
3	0.19	0.06	0.13	0.14	0.53	0.13
4	0.1	0.04	0.07	0.07	0.27	0.07

1.3.2 Scores of the Criteria

Scores of the criteria are one of the basic inputs for evaluating the projects considered. Five MOEs for the three projects in consideration are set as the evaluation criteria in this analysis. These criteria are scored based on percentage changes of them in case and control area from year 2011 to 2017. For example, commuting travel times in case area of project A are 25 minutes and 20 minutes in 2011 and 2017, respectively. Commuting travel times in control area of project A are 30 minutes and 28 minutes in 2011 and 2017, respectively. Therefore, commuting travel time reduction in case and control are 20% and 6.67%. We will take the difference of these percentage changes, therefore, 13.33% as the score of commuting travel time for project A. Similarly, other criteria are scored with justified modification in few cases.

1.3.3 Scoring and Ranking the Projects

In the final part of the analysis, three projects are scored and ranked using scores of the criteria and respective weights. While scoring the projects, we will use weighted average methods. To provide an example, scoring of project A using assumed values of weights and criteria is presented in Table 22.

Table 22: Scoring Example of the Projects

Criteria	1	2	3	4	Scores	Rank
Weights	0.53	0.27	0.13	0.07		
Project A	10	20	15	15	13.7	3
Project B	15	30	5	10	17.4	1
Project C	25	5	15	10	17.25	2

1.4 Analysis and Results

1.4.1 Weighting of Criteria

To weight the criteria, pairwise comparison method has been adopted in this study. To get stakeholder's (FDOT) assigned values, a survey was performed asking for FDOT officials' judgmental weights for the criteria in pairs in a scale of 1/9 to 9. Total 21 responses were collected and average values of the pairwise weights were taken for computing the final normalized weights. Average pairwise weights of different criteria are presented in following Table 23:

Table 23: Pairwise Comparison of Criteria (MOEs) Based on Survey

Criteria	Property value change	Job accessibility	Commuting time	Land Use Change	Travel Pattern
Property value change	1.000	0.198	0.178	0.236	0.221
Job accessibility	5.048	1.000	0.187	0.274	0.258
Commuting time	5.619	5.360	1.000	0.280	0.235
Land Use Change	4.238	3.644	3.573	1.000	0.250
Travel Pattern	4.524	3.878	4.251	4.000	1.000

Weights of the criteria can be computed from the pairwise comparison matrix. First, elements of each column are normalized by sum of the column (See Table 24). Then elements of each row are summed to get weights column. Elements of weights column are normalized at the last step (see Table 24 column 7 and 8).

Table 24: Weights of the Criteria

Criteria	Property value change	Job accessibility	Commuting time	Land Use Change	Travel Pattern	Weights	Normalized Weights
Property value change	0.049	0.014	0.019	0.041	0.113	0.236	0.047
Job accessibility	0.247	0.071	0.020	0.047	0.131	0.517	0.103
Commuting time	0.275	0.381	0.109	0.048	0.120	0.933	0.187
Land Use Change	0.207	0.259	0.389	0.173	0.127	1.155	0.231
Travel Pattern	0.221	0.275	0.463	0.691	0.509	2.160	0.432

From the normalized weights, it can be found that zero household travel pattern change is the most important and property value change is the least important evaluation criteria.

1.4.2 Value of the Criteria (MOEs)

1.4.2.1 Property Value Change

(a) SunRail Project

In this step, values of the criteria were determined by the changes of property value in case and control areas in the following order:

- At first, changes of the property value for a particular land use are estimated for case and control areas.
- Secondly, the changes are normalized by the sum of the changes in case and control and expressed in percentages or in a scale of hundred.
- Thirdly, difference between the normalized percentages of case and control is determined. This difference in percentages is then weighted by the land use share in respective region of the projects to find a single scoring for each region.
- Finally, scores for different regions are weighted by the project allocations of respective region to find a single score of property value change for the project.

Table 25: Property Value Evaluation for SunRail Project

Region	Land Use	Property Value Change in Case ($ per Acre)	Property Value Change in Control ($ per Acre)	% Change (Case)	% Change (control)	% change (Case-Control)	% Area	Weighted Property Value Change (%)	Total Invest-ment (million)	Property Value Change (%/$100 million)
Downtown	Single Family	456,558	315,503	59.13	40.87	18.27	53.89	35.40	176.47[1]	1.99
	Multi-Family	1915,108	381,978	83.37	16.63	66.74	5.67			
	Retail	570,925	163,763	77.71	22.29	55.42	21.20			
	Industrial	237,311	76,954	75.51	24.49	51.03	13.03			
	Institutional	340,343	100,801	77.15	22.85	54.30	6.21			
Outside Downtown Phase I	Single Family	221,968	201,879	52.37	47.63	4.74	53.89	14.15	529.41	
	Multi-Family	767,571	369,499	67.50	32.50	35.01	5.67			
	Retail	192,578	144,824	57.08	42.92	14.15	21.20			
	Industrial	170,936	96,988	63.80	36.20	27.60	13.03			
	Institutional	110,475	38,303	74.25	25.75	48.51	6.21			
Phase II	Single Family	309,987	322,017	49.05	50.95	-1.90	53.89	20.99	294.12	
	Multi-Family	1459,670	163,480	89.93	10.07	79.86	5.67			
	Retail	955,465	98,549	90.65	9.35	81.30	21.20			
	Industrial	54,423	115,120	32.10	67.90	-35.80	13.03			
	Institutional	792,467	91,732	89.63	10.37	79.25	6.21			

[1] = Total project investment, 1,000 million (Source: Railway Technology) is allocated considering number of stations

For example, consider property value change in phase 2 of SurRail project. Total property value change per acre for single family residential in case and control of phase II are $309,987.42 and $322,017.15, respectively. Then, these changes are normalized by the total changes in case control and expressed in percentages. Therefore, normalized changes for case and control are 49.05% and 51.95% respectively. So, difference between percentage changes is -1.90%. Thus, percentage changes for each type of land uses are computed and weighted average of them are estimated using land use share. Similar analysis has been performed for Phase I Outside Downtown and Downtown area. Weighted averages of property value change of three phases are then re-weighted by corresponding project costs and final score was estimated as 1.99% per $100 million of investment. Therefore, final result indicates that SunRail project has improved overall property value near SunRail stations. Detailed calculations are presented in Table 25.

(b) I-4 Expansion Project

Similar to SunRail Project, property value was scored for I-4 expansion project. Final result shows that property value has increased by 1.4% near I-4 expansion project area for each $100 million of investment from 2011 to 2017. Detailed calculation has been presented as tabulated form in Table 26.

Table 26: Property Value Evaluation for I-4 Expansion Project

Region	Land Use	Property Value Change in Case ($ per Acre	Property Value Change in Control ($ per Acre)	% Change (Case)	% change (control)	% change (Case-Control)	% Area	Weigh-ted Proper-ty Value Change	Total Invest-ment(m illion)	Property Value Change (%/$100 million)
Attraction	Single Family	500,104	269,953	64.94	35.06	29.89	49.94	30.49	773.53[1]	0.88
	Multi-Family	771,574	1038,100	42.64	57.36	-14.73	7.82			
	Retail	869,466	492,299	63.85	36.15	27.70	31.34			
	Industrial	544,392	91,069	85.67	14.33	71.34	6.18			
	Institutional	537,273	69,961	88.48	11.52	76.96	4.71			
Downtown	Single Family	220,217	230,254	48.89	51.11	-2.23	49.94	29.93	569.97	
	Multi-Family	836,234	269,488	75.63	24.37	51.26	7.82			
	Retail	519,365	72,903	87.69	12.31	75.38	31.34			
	Industrial	152,030	118,220	56.26	43.74	12.51	6.18			
	Institutional	649,855	183,579	77.97	22.03	55.95	4.71			
Ivanhoe	Single Family	354,816	143,717	71.17	28.83	42.34	49.94	55.66	664.97	
	Multi-Family	1509,037	449,386	77.05	22.95	54.11	7.82			
	Retail	270,338	30,586	89.84	10.16	79.67	31.34			
	Industrial	212,890	-192	100.09	-0.09	100.18	6.18			
	Institutional	85,756	124,795	40.73	59.27	-18.54	4.71			
Altamonte	Single Family	141,806	148,119	48.91	51.09	-2.18	49.94	-5.86	868.53	
	Multi-Family	588,380	699,996	45.67	54.33	-8.66	7.82			
	Retail	68,099	102,032	40.03	59.97	-19.95	31.34			
	Industrial	142,913	231,700	38.15	61.85	-23.70	6.18			
	Institutional	43,924	5,763	88.40	11.60	76.80	4.71			

[1] = Total project investment, 2877 million (Source: U.S. DOT) is allocated considering length in each region

(c) Juice Bikeshare Project

Analysis of property value change for Juice Bikeshare project indicates that property value near Juice Bikeshare stations has increased by 1016.7% from 2011 to 2017. Detailed calculation of property value change for Juice Bikeshare project has been presented as tabulated form in Table 27.

Table 27: Property Value Evaluation for Juice Bikeshare Project

Land Use	Property Value Change in Downtown	Property Value Change in Outside Downtown	% Change (Down-town)	%Change (Outside Downtown)	% Change (Downtown-Outside Downtown)	% Area	Weighted Property Value Change	Investm-ent Differ-ential (million)	% Change/ $100 million
Single Family	838233.48	712860.72	54.04	45.96	8.08	53.89	4.27	0.42[1]	1016.7
Multi-Family	1147255.82	3442447.78	25.00	75.00	-50.01	5.67			
Retail	1256561.00	937687.26	57.27	42.73	14.53	21.20			
Industrial	402773.13	448302.90	47.33	52.67	-5.35	13.03			
Institutional	1396720.68	1242913.07	52.91	47.09	5.83	6.21			

[1] = Calculated considering $3 million investment per 50 stations/500 bikes (Source: bikeshare system feasibility study in Redmond, Washington)

1.4.2.2 Job accessibility

(a) SunRail Project

Scoring of job accessibility has been performed based on difference of percentage change of number of accessible jobs in case and control area. For example, increase in number of accessible jobs from 2011 and 2017 in case and control of Downtown region are 66,958.60 and 80,805.96 per unit area, respectively. Then, increases of job accessibility in case and control areas are normalized by their sum and expressed in percentages. So, percentage changes of job accessibility in case and control area are estimated as 45.31% and 54.69%, respectively. So, difference of percentage change in case and control of Downtown region is -9.37%. Similarly, differences in percentage change for other two phases were estimated and weighted based on corresponding investments. Thus, final score for job accessibility has been calculated as -2.38% per $100 million of investment.

Table 28: Job Accessibility Evaluation Change for SunRail Project

Region	Change in Case per unit area	Change in Control per unit area	% Change in Case	% Change in Control	% Change	Investment in $million	% Change/ $100 million
Downtown	66958.60	80805.96	45.31	54.69	-9.37	176.47	-2.38
Outside Downtown Phase I	12733.35	38325.90	24.94	75.06	-50.12	529.41	
Phase II	68855.06	50796.25	57.55	42.45	15.09	294.12	

The reader would note that reduced job accessibility due to a transportation project is not a valid assessment and is a manifestation of the challenges associated with generating these MOEs. In this event, for plausibility reasons, we consider the impact of SunRail project on Job accessibility as 0 i.e. no impact.

(b) I-4 Expansion Project

Similar to SunRail Project, job accessibility was scored for I-4 expansion project. Detailed calculation has been presented as tabulated form in Table 29.

Table 29: Job Accessibility Change Evaluation for I-4 Expansion Project

Region	Change in Case per unit area	Change in Control per unit area	% Change in Case	% Change in Control	% Change	Investment in $million	% Change/ $100 million
Attraction	75803	2992.12	96.20	3.80	92.41	773.53	1.23
Downtown	57355	17113.78	77.02	22.98	54.04	569.97	
Ivanhoe	42301	11921.67	78.01	21.99	56.03	664.97	
Altamonte	5830	14611.18	28.52	71.48	-42.96	868.53	

(c) Juice Bikeshare Project

Detailed calculation of job accessibility change for Juice Bikeshare project has been presented as tabulated form in Table 30.

Table 30: Job Accessibility Change Evaluation for Juice Bikeshare Project

Change in Case per unit area	Change in Outside Downtown per unit area	% Change in Case	% Change in Control	% Change	Investment Differential in $million	% Change/ $100 million
474.48	672.12	41.38	58.62	-17.24	0.42	-4104.8

As discussed earlier, for plausibility reasons, we consider the impact of Juice project on Job accessibility as 0 i.e. no impact.

1.4.2.3 Commuting Travel Time Change

(a) SunRail Project

Commuting travel time change is scored based on percentage changes of commuting time in case and control from 2011 to 2017. Difference between commuting time changes in case and control area of Downtown region from 2011 to 2017 is -1.61%. For outside downtown phase I and phase II, the differences are -7.14% and -3.91%, respectively. These percentages are then weighted to find a single score, -0.52% per $100 million of investment.

Table 31: Commuting Travel Time Change Evaluation of SunRail Project

Region	% Reduction in Case	% Reduction in Control	% Reduction (Case - Control)	Investment in $million	% Change/$100 million
Downtown	-2.66	-1.05	-1.61	176.47	-0.52
Outside Downtown Phase I	-11.76	-4.62	-7.14	529.41	
Phase II	-6.86	-2.95	-3.91	294.12	

(b) I-4 Expansion Project

Similar analysis is performed for I-4 expansion project and final score is found to be -0.11% per $100 million of investment. Detailed calculation is presented in Table 32.

Table 32: Commuting Travel Time Change Evaluation of I-4 Expansion Project

Region	% Reduction in Case	% Reduction in Control	% Reduction (Case - Control)	Investment in $million	% Change/$100 million
Attraction	-3.42	-3.40	-0.02	773.53	-0.07
Downtown	-6.55	-2.50	-4.05	569.97	
Ivanhoe	-8.31	-7.18	-1.13	664.97	
Altamonte	-7.07	-3.99	-3.08	868.53	

(c) Juice Bikeshare Project

Detailed calculation of zero car household travel pattern change for Juice Bikeshare project has been presented as tabulated form in Table 33.

Table 33: Commuting Travel Time Change Evaluation of Juice Bikeshare Project

% Reduction in Downtown	% Reduction in Outside Downtown	% Reduction (DT - Outside DT)	Investment in $million	% Change/ $100 million
-1.00	-5.87	4.88	0.42	1161.9

1.4.2.4 Land Use Change

(a) SunRail Project

To estimate the score for land use change, changes of total vacant area from 2011 to 2017 in case and control area were used. For example, total changes in vacant area in case and control from 2011 to 2017 were 17.36 and 3.41 acre per square mile, respectively. These changes are then normalized by their total and expressed in percentages or in a scale of 100. The difference of two percentages, therefore, 67.14% is taken as the score for Downtown region. Similarly, other two regions are scored, and scores are then weighted by respective investments to derive a single score which is found as 6.62% percent per $100 million of investment.

Table 34: Land Use Change for SunRail Project

Region	Change of Vacant Area in Case (Acre/sqmile)	Change of Vacant Area in Control (Acre/sqmile)	% Change in Case	% Change in Control	% Change (Case-Control)	Investment in $million	% Change/ $100 million
Downtown	17.36	3.41	83.57	16.43	67.14	176.47	6.62
Outside Downtown Phase I	24.29	6.53	78.80	21.20	57.61	529.41	
Phase II	62.93	6.57	90.54	9.46	81.08	294.12	

(b) I-4 Expansion Project

Similar analysis has been performed for I-4 expansion project and final score was found to be 1.55% per $100 million of investment. Detailed calculation is presented in Table 17.

Table 35: Land Use Change for I-4 Expansion Project

Region	Change of Vacant Area in Case (Acre/sqmile)	Change of Vacant Area in Control (Acre/sqmile)	% Change in Case	% Change in Control	% Change	Investment in $million	% Change/ $100 million
Attraction	23.62	21.49	52.36	47.64	4.72	773.53	0.97
Downtown	31.66	20.96	60.17	39.83	20.34	569.97	
Ivanhoe	21.70	24.47	47.00	53.00	-6.01	664.97	
Altamonte	81.31	9.23	89.81	10.19	79.62	868.53	

(c) Juice Bikeshare Project

Detailed calculation of land value change for Juice Bikeshare project has been presented as tabulated form in Table 36.

Table 36: Land Use Change for Juice Bikeshare Project

Total Change (Acre) in Downtown	Change of Vacant Area in Downtown (Acre/sqmile)	Total Change (Acre) in Outside Downtown	Change of Vacant Area in Outside Downtown (Acre/sqmile)	% Change in Down-town	% Change in Outside Down-town	% Change	Investment Differential in $million	% Change per $100 million
51.35	32.28	105.24	99.19	21.44	78.56	-57.12	0.42	-13600

1.4.2.5 Zero Car Household Travel Pattern Change

(a) SunRail Project

To score zero car households travel pattern change, public transportation share changes in case and control area were used for SunRail project. For example, public transportation usage increase from 2011 to 2016 in case and control area of Downtown region are found as -0.40 and -4.95% respectively. Therefore, difference of the increases is 4.55%. Thus, the increases for Downtown phase I and Phase II regions are determined as 1.68% and 10.37%, respectively. These three changes are then weighted by respective investments to find a single score of 0.47% per $100 millions of investment.

Table 37: Travel Pattern Change Evaluation of SunRail Project

Region	% Change of Public Transport Share in Case	% Change of Public Transport Share in Control	% Change (Case-Control)	Investment in $million	% Change/ $100 million
Downtown	-0.40	-4.95	4.55	176.47	0.47
Outside Downtown Phase I	1.01	-0.66	1.68	529.41	
Phase II	11.24	0.87	10.37	294.12	

(b) I-4 Expansion Project

Similar analysis has been performed for I-4 expansion project and final score was found to be -0.17% per square mile. Detailed calculation is presented in Table 38.

Table 38: Travel Pattern Change Evaluation of I-4 Expansion Project

Region	% Change of Public Transport Share in Case	% Change of Public Transport Share in Control	% Change (Case-Control)	Investment in $million	% Change/ $100 million
Attraction	1.15	-0.65	1.8	773.53	-0.11
Downtown	-4.61	-4.8	0.19	569.97	
Ivanhoe	-5.56	-3.84	-1.72	664.97	
Altamonte	-7.38	3.44	-10.82	868.53	

(c) Juice Bikeshare Project

Percentage changes in share of bike/walk has been utilized in case of Juice Bikeshare project to score travel pattern change criterion. Detailed calculation of zero car household travel pattern change for Juice Bikeshare project has been presented as tabulated form in Table 39.

Table 39: Travel Pattern Change Evaluation of Juice Bikeshare Project

% Change of Bike/Walk Share in Downtown	% Change of Bike/Walk Share in Outside Downtown	% Change (Case-Control)	Investment Differential in $million	% Change/ $100 millions
6.19	-14.44	20.63	0.42	4911.9

1.4.3 Overall Scoring

Final step of the multicriteria decision analysis is overall scoring of the projects and rank them based on their scores. Overall scoring of the projects is performed by weighting the scores of the criteria. In this section, overall scores of the three projects will be performed and they will be ranked on the basis of their individual scores. Detailed calculation of scoring has been shown in following Table 40. Final result shows that among the three projects considered, SunRail project has maximum scores followed by I-4 expansion project. Juice Bikeshare project has the minimum score.

Table 40: Scoring and Ranking of the Projects

Criteria	**Property value change**	**Job accessibility**	**Commuting time**	**Land Use Change**	**Travel Pattern**	**Overall Score**	**Rank**
Weights	0.047	0.103	0.187	0.231	0.432		
SunRail	1.99	0.00	-0.52	6.62	0.47	1.729	1
I-4 Expan.	0.88	1.23	-0.07	0.97	-0.11	0.332	2
Juice Bike	1016.7	0.00	1161.9	-13600	4911.9	-754.599	3

1.5 Conclusion

The report discussed multi-criteria analysis methodology adopted for this study and presented analysis steps in details. At the end of this report, three projects were scored on the basis of criteria scores and their respective weights. Results show that SunRail project is the highest ranked project among these three projects. In contrast, Juice bikeshare project is the lowest ranked project. Of the three projects SunRail and I-4 project offered positive overall values indicating they offer a positive return. It is important to note that the results for Juice system need to be considered with an abundance of caution as the spatial distribution is smaller (relative to the other projects) and there is substantial variation in the results across MOEs. The definition of the MOEs also has significant influence on the findings. For instance, the job accessibility measured for SunRail project offered negative values, indicating that job accessibility has reduced due to SunRail project. In analyzing data, it is possible to arrive at non-plausible results due to the inherent complexity of the process being considered. In such events, it is important that we evaluate the result as engineers and possibly ignore the MOE or consider alternative MOEs. In our case, we considered SunRail impact on job accessibility as 0 for further computations. For the land use type change MOE, we considered changes from vacant to other land use types. It is possible to consider changes at a finer resolution such as single family to multi-family (if any) and so on. However, in our context these changes were minimal. Broadening the definition, could potentially affect the results. The main contribution of the research conducted is to develop a useful and customizable framework for evaluating the impacts of transportation projects. Every project/study is unique and will benefit from analyst's experience and judgement.

Chapter 2: Knowledge Transfer

The current research effort was envisioned as a mechanism to provide hands on experience to FDOT personnel and other stakeholders on the state of the art approaches for evaluating the impact of transportation infrastructure projects on community building. Towards this objective, the research team developed a knowledge transfer plan through a mix of webinars, step-by-step tutorials and webpage designed to host the research material.

The research team organized a webinar series for FDOT personnel and stakeholders. Three webinar sessions were held in July and August, 2020. Dr. Naveen Eluru, Dr. Samiul Hasan,Sudipta Dey Tirtha and Jiechao Zhang presented in the webinar sessions. The webinar series included the following 3 modules:

1. *Measures of Effectiveness (MOE) Data Preparation and Analysis*: The development of the MOEs is a data intensive process. These indicators/measures can be developed by collating appropriate data collected from different sources using the ArcGIS platform. In the 1st webinar held on July 16, we discussed the data preparation steps of MOE computation process with examples and visualizations. In this webinar, we discussed the steps for estimating changes of property values by land use types in case and control areas. The steps for job accessibility change estimation are also discussed in a separate presentation (video) which is separately uploaded for easy access to webinar participants.

2. *Social Media data download and analysis for Transportation Projects:* The second webinar, held on July 23, provided instructions for running the codes and scripts for social media data download and analysis. The webinar consisted of four topics: i) data collection, ii) sentiment analysis, iii) visualization of the sentiment analysis results, iv) visualization of the topic analysis results. In the presentation, all the codes and required steps were presented with detailed instructions, examples and visualizations.

3. *Project evaluation in the presence of multiple MOEs using multi-attribute decision making:* The last webinar, held on August 6, described the proposed framework to compare the changes in MOEs across scenarios to identify benefits to the region. In this presentation, we discussed how each of these projects were scored by MOEs by comparing their changes in case and control areas. We also discussed how the criteria/MOEs were weighted based on the responses from the FDOT officials. Based on the analysis, we presented a net positive, neutral or negative rating for the three projects considered. The webinar was concluded with a discussion of the implications of the findings.

The research team also prepared two supporting documents to facilitate the knowledge transfer activity. The two tutorials include: 1) Estimation Procedure of Various Measures of Effectiveness (MOE) for Transportation Investments, and 2) Social Media Data Analysis. The tutorial provides detailed steps of MOE computation process in ArcGIS and social media data analysis with examples and visualizations. The tutorials are included in Appendices A and B respectively.

Finally, in addition to webinar series, the research team also prepared an online resource (a webpage) to facilitate the easy access of the webinar videos, presentation material and tutorials. The URL of the webpage is http://www.people.cecs.ucf.edu/neluru/CommunityBuilding.html. The webpage content is organized along three modules reflecting the structure of the three webinars.

References

Afshari, A., Mojahed, M., Yusuff, R.M. (2010). Simple Additive Weighting approach to Personnel Selection problem. International Journal of Innovation, Management and Technology, 1(5): 511 – 515.

Angelis, A., Kanavos, P. (2017). Multiple Criteria Decision Analysis (MCDA) for evaluating new medicines in Health Technology Assessment and beyond: The Advance Value Framework. Social Science & Medicine, 188: 137-156.

Bottero, M., Alpaos, C.D., Oppio, A. (2018). Multicriteria Evaluation of Urban Regeneration Processes: An Application of PROMETHEE Method in Northern Italy. Advances in Decision Sciences, 1 – 12.

Da˘gdeviren, M. (2008). Decision making in equipment selection: an integrated approach with AHP and PROMETHEE. J Intell Manuf, 19: 397–406.

Final report, Bikeshare feasibility study, Redmond, Washington. https://altaplanning.com/wp-content/uploads/Redmond-Final-Report.pdf Last accessed on 26 December 2019.

Franek, J., Kresta, A. (2014). Judgment scales and consistency measure in AHP. Enterprise and the Competitive Environment 2014 conference, ECE 2014.

Hajkowicz, S., Higgins, A. (2008). A comparison of multiple criteria analysis techniques for water resource management. European Journal of Operational Research, 184(1): 255-265.

Haralambopoulos, D.A., Polatidis, H. (2003). Renewable energy projects: structuring a multicriteria group decision-making framework, Renewable Energy, 28: 961–973.

https://www.fhwa.dot.gov/ipd/project_profiles/fl_i4ultimate.aspx Last accessed on 26 December 2019.

https://www.railway-technology.com/projects/sunrail-project/ Last accessed on 26 December 2019.

Lai, V.S., Wong, B.K., Cheung, W. (2002). Group Decision Making in Multi Criteria Environment: A Case Using the AHP in Software Selection. European Journal of Operational Research, 137: 134-144.

Macharis, C., Witte, A.D., Ampe, J. (2008). The Multi-Actor, Multi-Criteria Analysis Methodology (MAMCA) for the Evaluation of Transport Projects: Theory and Practice. Journal of Advanced Transportation, 43(2): 183-202.

Pohekar, S.D., Ramachandran, M. (2004). Application of multi-criteria decision making to sustainable energy planning—A review. Renewable and Sustainable Energy Reviews, 8: 365–381.

Saaty, T.L. (1980). The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation. McGraw-Hill, New York.

Saaty, T.L. (2008). Decision making with the analytic hierarchy process. Int. J. Services Sciences, 1(1).

Saito, M. (1987). Application of the analytic hierarchy method to setting priorities on bridge replacement projects. Transportation Research Record, 1124: 26–35.

Sinuany-Stern, Z., Mehrez, A., Hadad, Y. (2000). An AHP/DEA methodology for ranking decision-making units. Intl. Trans. in Op. Res. 7: 109 – 124.

Song, B., Kang, S. (2016). A Method of Assigning Weights Using a Ranking and Non-hierarchy Comparison. Advances in Decision Sciences.

Triantaphyllou, E. (1997). A Sensitivity Analysis Approach for Some Deterministic Multi-Criteria Decision-Making Methods. Decision Sciences, 28(1).

van Til, J., Groothuis-Oudshoorn, C., Lieferink, M., Dolan, J. and Goetghebeur, M., 2014. Does technique matter; a pilot study exploring weighting techniques for a multi-criteria decision support framework. Cost effectiveness and resource allocation, 12(1), p.22.Cristóbal, J.S. (2011). Multi-criteria decision-making in the selection of a renewable energy project in spain: The Vikor method. Renewable Energy, 36: 498-502.

Yan, J., Dagang, T., Yue, P. (2007). Ranking environmental projects model based on multicriteria decision-making and the weight sensitivity analysis. Journal of Systems Engineering and Electronics, 18(3): 534-539.

Yang, T., Hung, C.C. (2007). Multiple-attribute decision making methods for plant layout design problem. Robotics and Computer-Integrated Manufacturing 23: 126–137.

Zanakis, S.H., Solomon, A., Wishart, N., Dublish, S. (1998). Multi-attribute decision making: A simulation comparison of select methods. European Journal of Operational Research, 107: 507 – 529.

Appendix A: Tutorial for Measures of Effectiveness Estimation

A.1 Background

A.1.1 Transportation Infrastructure

According to Florida Chamber of commerce, Florida ranks number one in the US in terms of transportation infrastructure rankings. It is the third largest state by population, after California and Texas with a yearly growth rate of more than 1.5%. Orlando is the most thriving city of the region; its growth being bolstered by its job creation rate (1,000 jobs are added per week). The economic and demographic trends suggest that Orlando has an expanding consumer market and these trends are set to drive increased demand for passengers and freight transportation in the coming years. To accommodate the future demand in an efficient and sustainable manner, several small and big transportation projects are underway in the region including second phase of SunRail commuter rail extension, I-4 expansion, pedestrian and bicycling facility installation, and bicycle sharing system (Juice) introduction. The proposed research effort is geared towards examining the community impacts of three transportation infrastructure investment projects: SunRail, I-4 expansion, and JUICE Orlando bikeshare system (see Figure 1).

A.1.2 MOE Computation

The development of the MOEs is a data intensive process. The process involves collection of appropriate data from different sources, extracting data for the geographic regions under study, and eventually combining layers of data as needed. Informed from the literature review, we propose five MOEs to evaluate the community building effects of the major transportation investment projects currently underway in the Central Florida Region. The proposed MOEs are:

- Property value change
- Changes to job accessibility
- Commuting time change
- Land use type change
- Changes to travel patterns for zero car households

The proposed changes will be evaluated for the time period 2011-2017. For sake of brevity, we present the layer preparation steps for 2012. The procedure was repeated for the entire time period of analysis. For job accessibility, commuting time and zero car household pattern based MOEs, data for 2017 was unavailable and the analysis was conducted from 2011-2016.

The development of these MOEs is a data intensive process. These indicators/measures can be developed by collating appropriate data collected from different sources using the ArcGIS platform. In this deliverable, we discuss the data preparation steps, MOE computation process.

A.1.3 Current Study

The development of above MOEs is a data intensive process. These indicators/measures can be developed by collating appropriate data collected from different sources using the ArcGIS platform. In this report, we discuss the data preparation steps, MOE computation process using ArcGIS and SPSS.

The proposed MOE changes will be evaluated for the time period 2011-2017 and also for all three transportation infrastructure (SunRail, I-4 ultimate and JUICE Orlando Bikeshare system). For sake of brevity, we present the layer preparation steps for 2012 and for SunRail stations only.

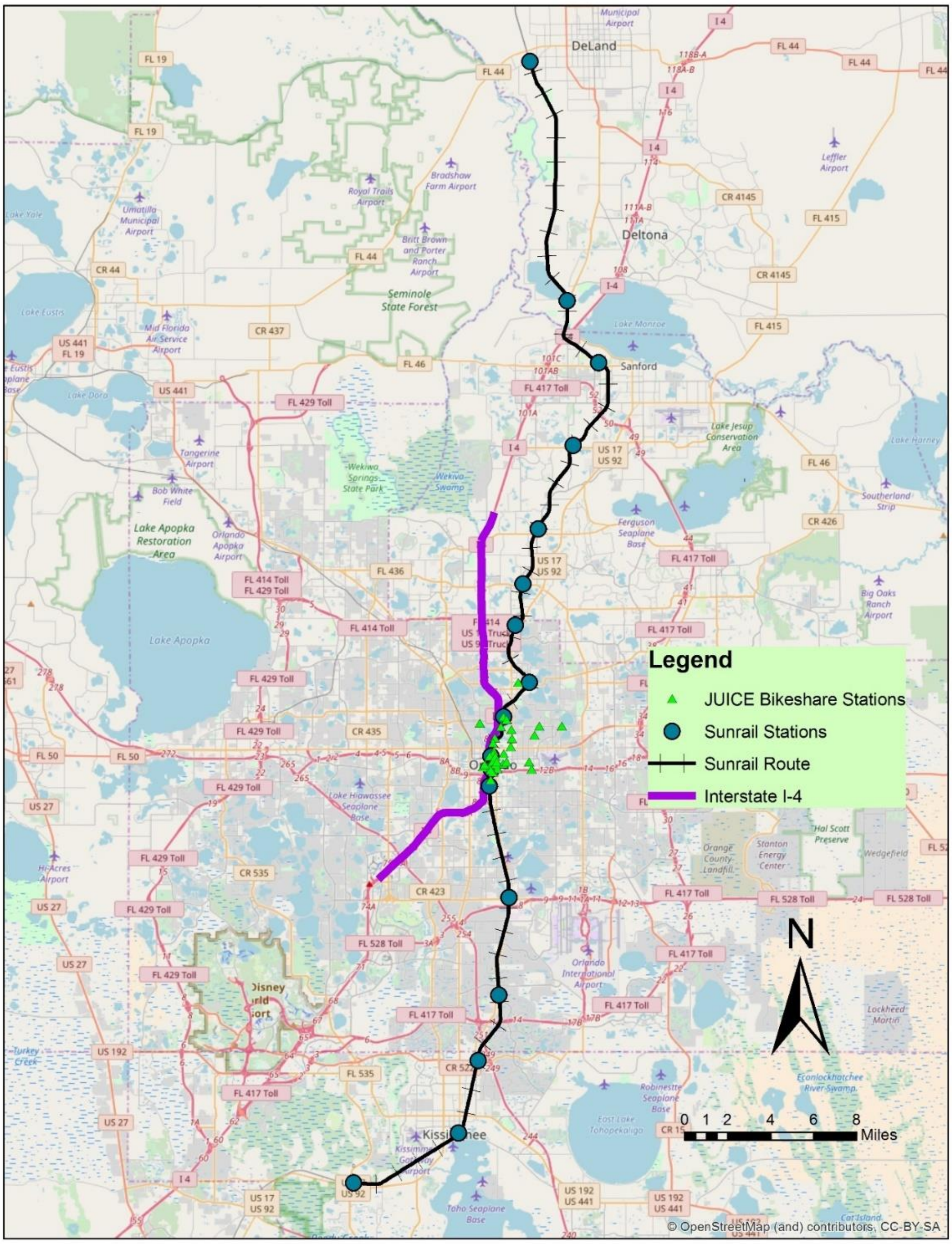


Figure A.5: Major Transportation Investment Projects (SunRail, I-4 Expansion and JUICE Bikeshare) in Central Florida Region

A.2 Property Value Estimation by Land Use Type

To estimate the property value for different use type, county 'Parcel' data were used. Several data preparation steps were followed for estimating the property value by using GIS. Here, we will give a brief description of estimation steps of property value using GIS for SunRail stations only.

A.2.1 Parcel Data Preparation

A.2.1.1 County Parcel Shapefile

- To capture the change in property value, parcel data for (2011-2017) obtained from Florida Department of Revenue (FDOR) were utilized (ftp://sdrftp03.dor.state.fl.us/Map%20Data/).
- County parcel shapefile contains unique parcels within each county that indicated by unique 'Parcel No' together with each parcel's length and area (Figure A.2).

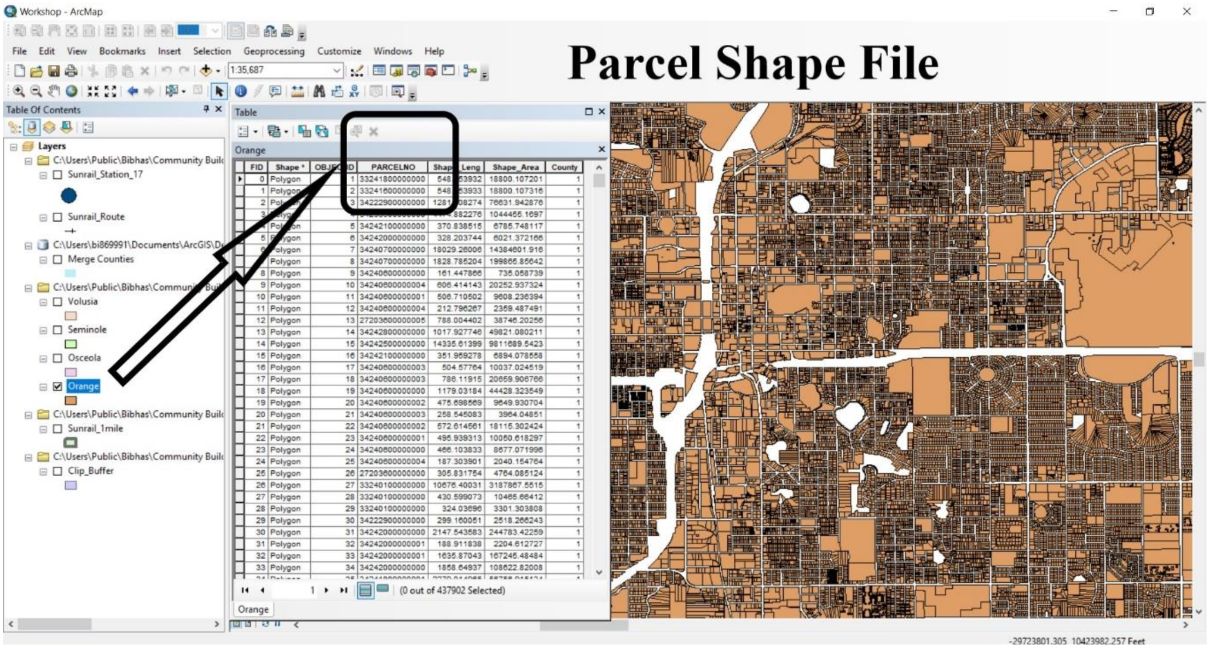


Figure A.6: Parcel Shapefile

- The transportation infrastructure projects considered in our research passes through four counties: Orange, Osceola, Seminole and Volusia (See Figure A.3).

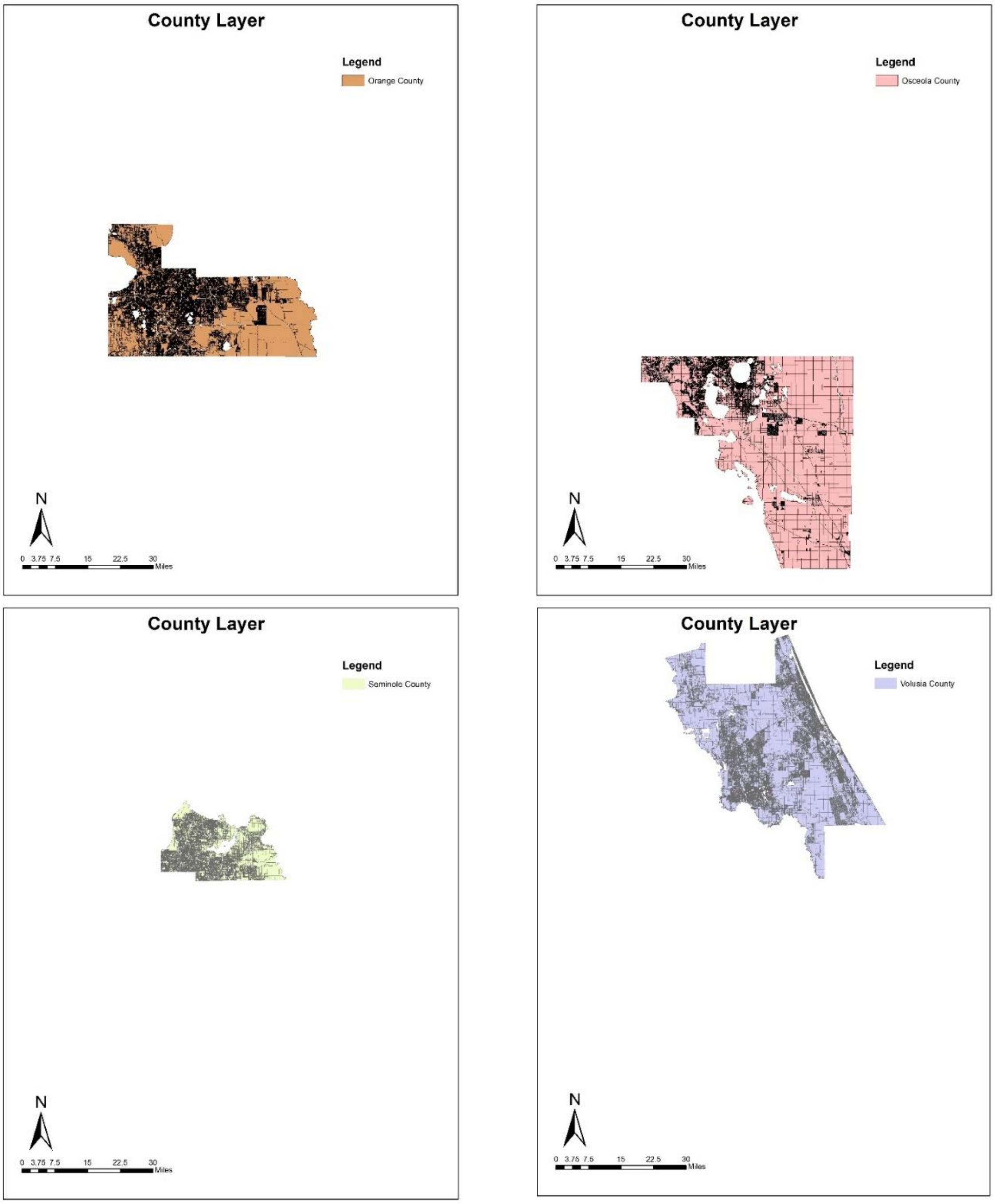
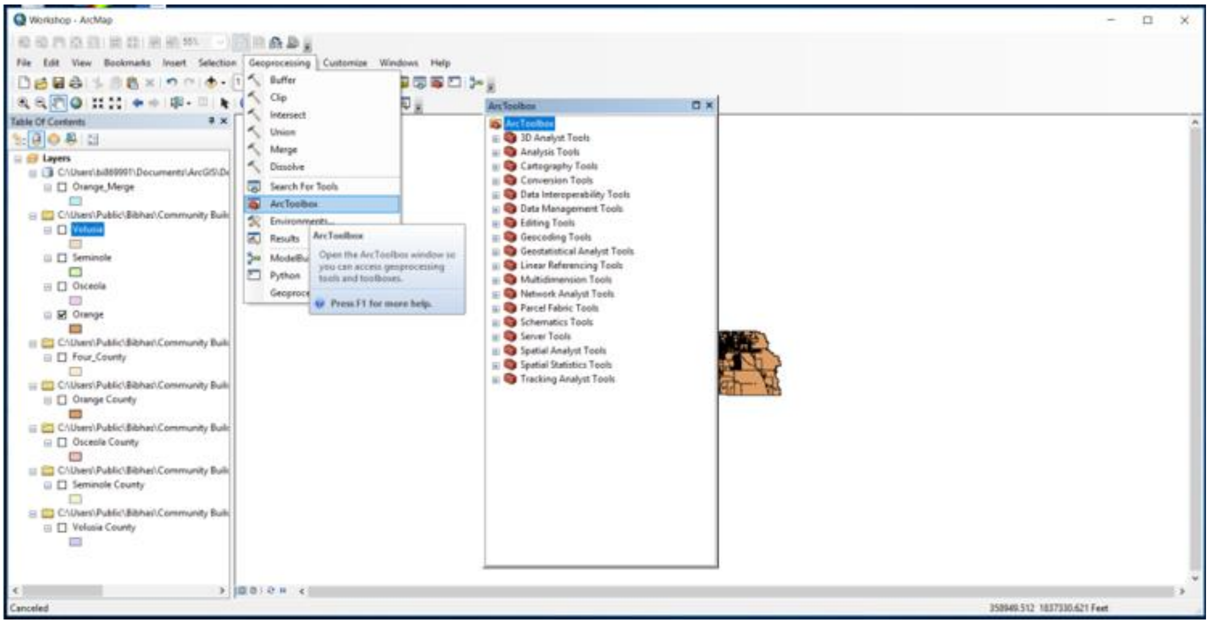


Figure A.7: Counties Parcel Shapefile

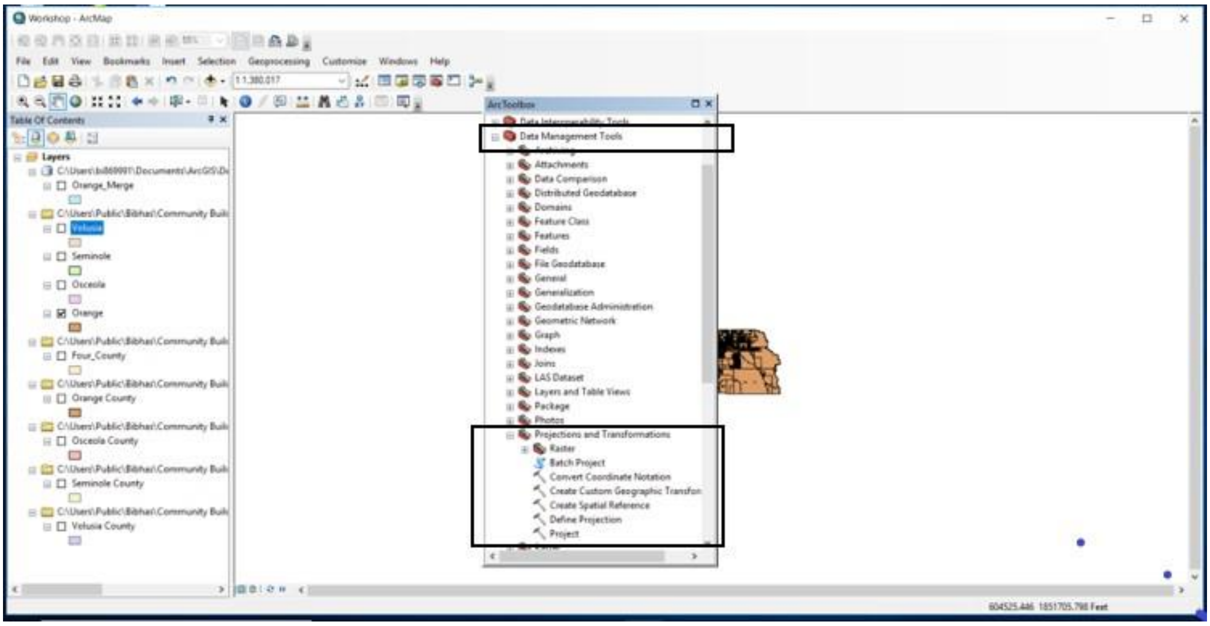
A.2.1.2 Shapefile Co-ordination System Projection

- The 1st step among all is to project all the parcel shapefile to same coordination system.
- Following are the steps to project all parcel shapefile to same coordination system:
  - ✓ 1st select Geoprocessing
  - ✓ Choose Arc Toolbox bar in Geoprocessing (See Figure A.4(a))
  - ✓ Select 'Data Management Tools' (See Figure A.4(b))
  - ✓ Select 'Projections and Transformations' from 'Data Management Tools'
  - ✓ Then click to 'Project' to select coordinate system
  - ✓ Select county shapefile in as input features (See Figure A.4(c))

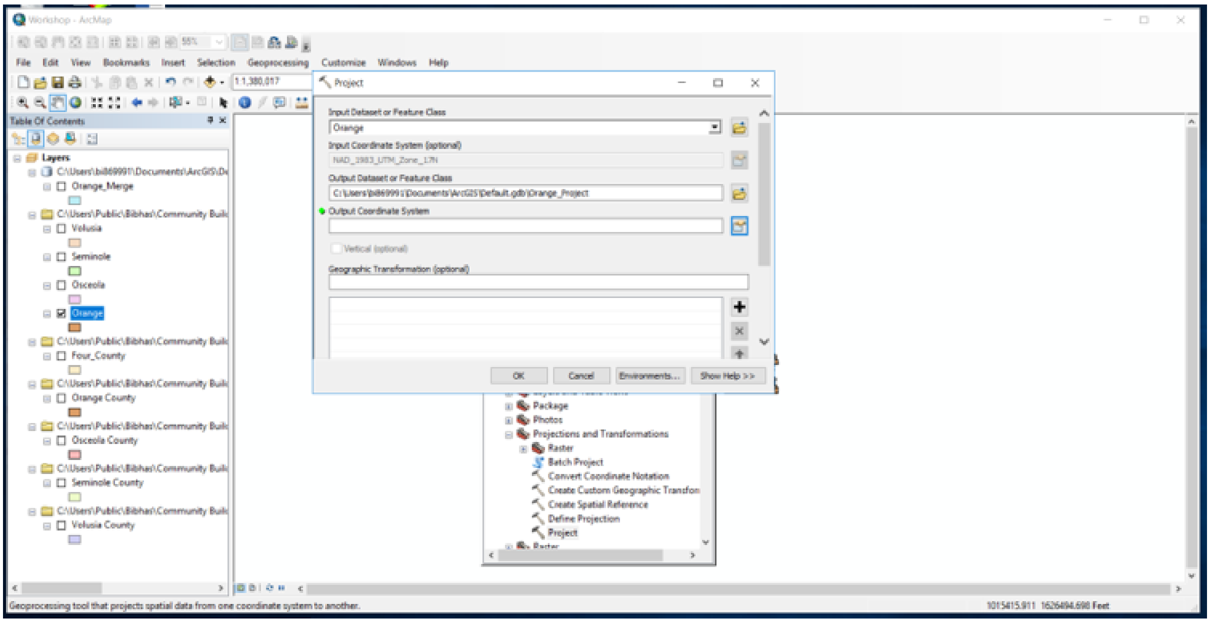
- ✓ Click ‘Output Coordinate System’
- ✓ There are two coordinate systems available in Output Coordinate System – Geographic and Projected Coordinate System (See Figure A.4(d))
- ✓ Select NAD 1983 within Universal Transverse Mercator (UTM) bar (See Figure A.4(e))
- ✓ Finally Select NAD 1983 UTM Zone 17N that represents the infrastructure zone (See Figure A.4(f))



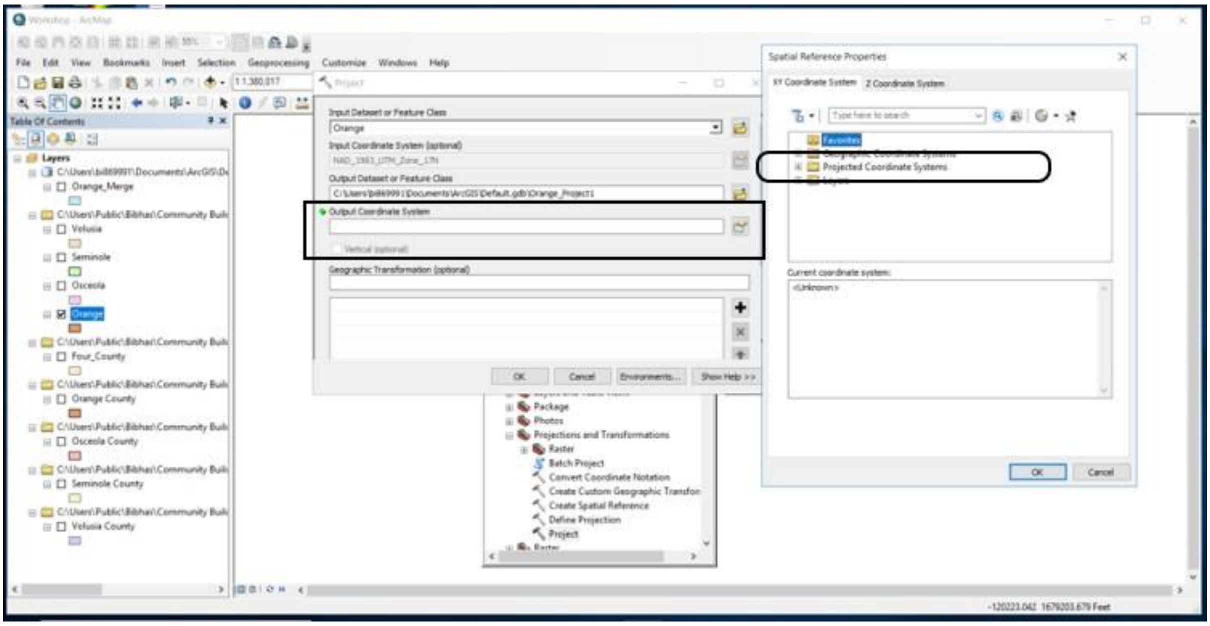
(a)



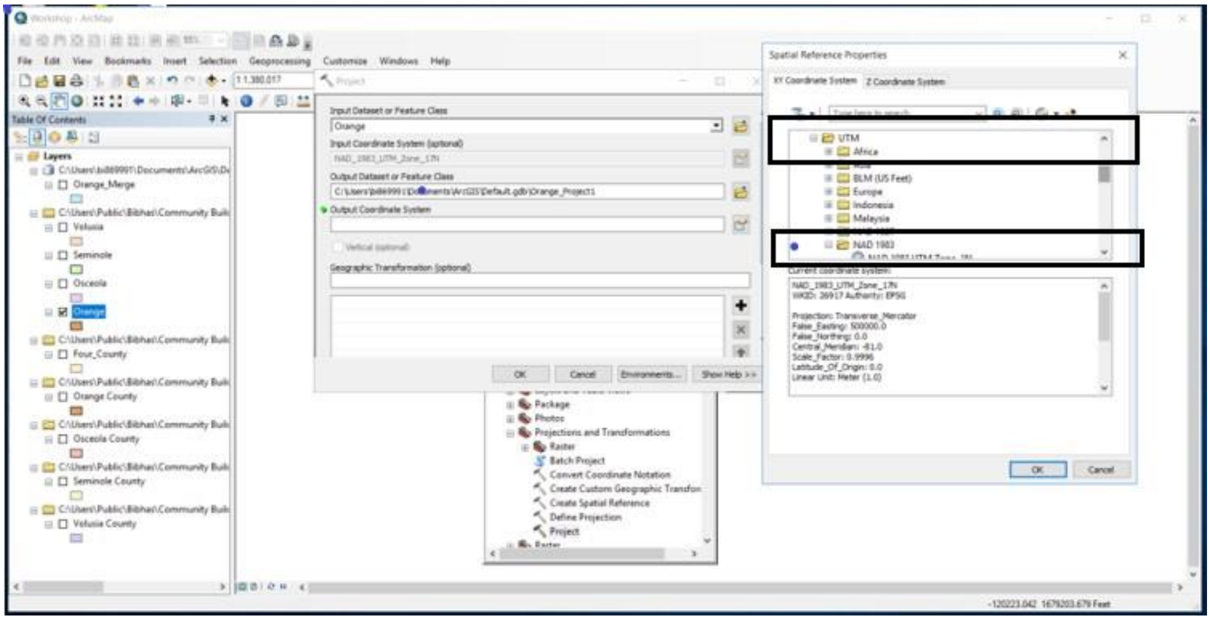
(b)



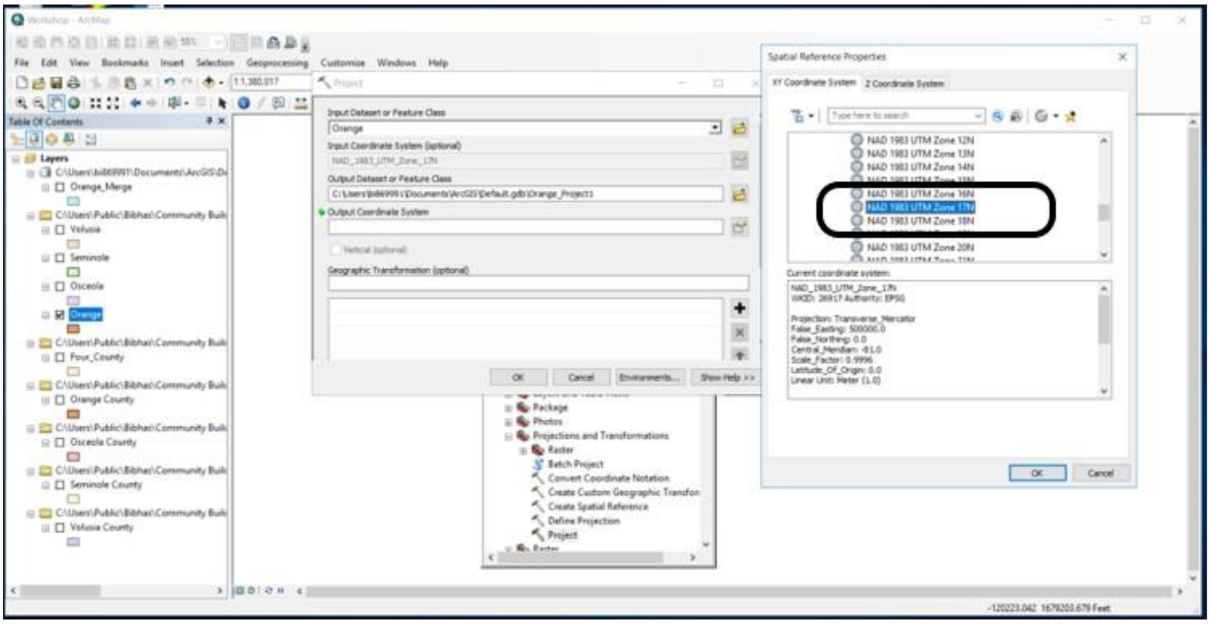
(c)



(d)



(e)



(f)

Figure A.8: Projected Coordination System

A.2.1.3 Parcel Data Layer Preparation

- The Name-Address-Legal (NAL) file for 2011-2017 was used together with parcel level county shapefile to get the property value information for parcel level (ftp://sdrftp03.dor.state.fl.us/Tax%20Roll%20Data%20Files/).
- NAL file has unique parcel ID with equivalent parcel level attribute information such as property/feature value, land value, land area in square feet, land use codes (DOR-UC), owner name, owner address, physical address, physical zip code, building details and so on (Figure A.5).
- Please note that Just Value (land just value, building value, and special feature value) of a property includes: present cash value; use; location; quantity or size; cost; replacement value of improvements; condition; income from property; and net proceeds if the property is sold. The net proceeds equal the value of the property minus 15% of the true market value. This accounts for the cost of selling the property. In calculating the change in property values, we consider Just Value reported by DOR as a surrogate measure for direct property value and in the following sections, we will refer to this value as the property value for simplicity.

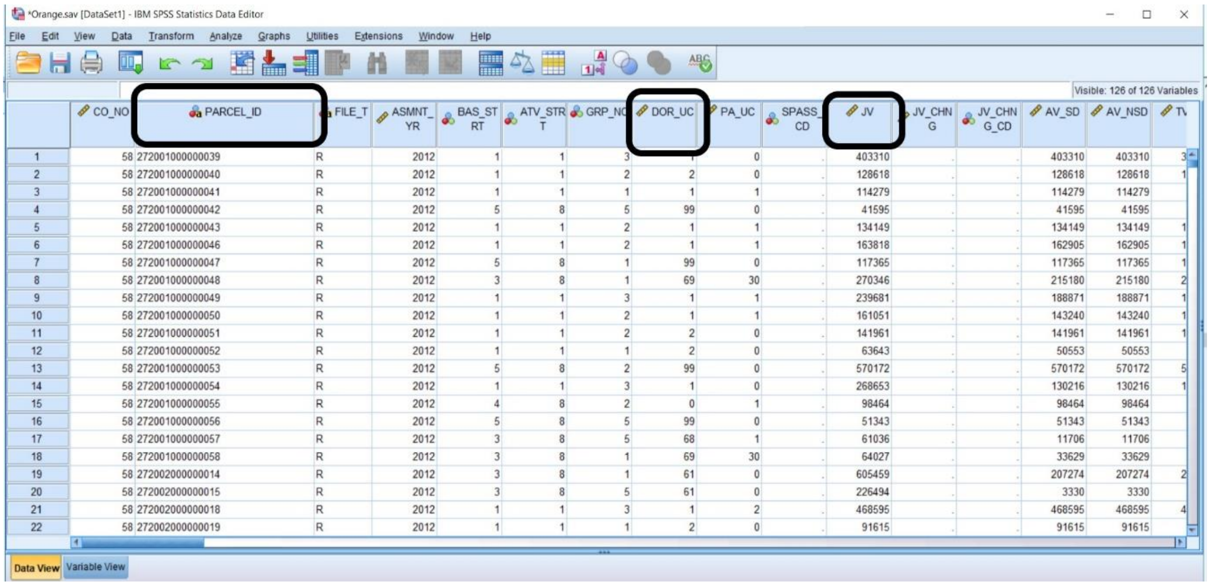


Figure A.9: Name-Address-Legal (NAL) File

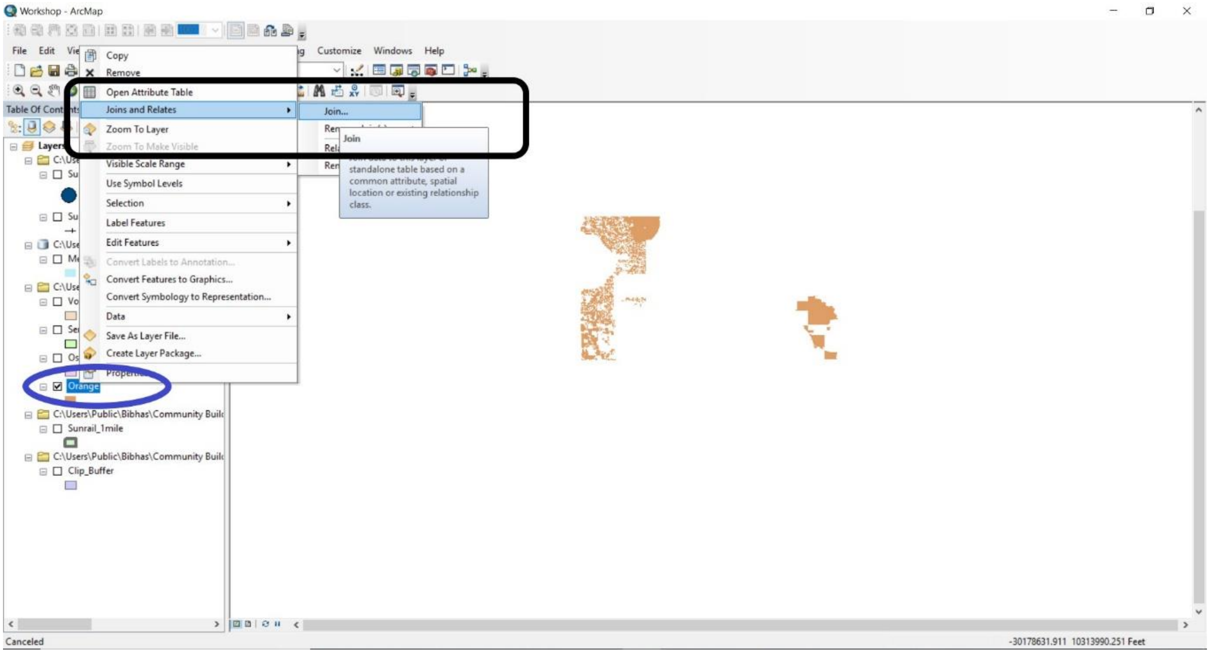
A.2.1.4 Merging NAL File Information

- Parcel No from county parcel shapefile contains unique parcels within each county file linking it with equivalent parcel level attribute information contained in the Name-Address-Legal (NAL) file.
- Following are the steps to linking parcel shapefile with NAL file information:
  - ✓ 1st click right cursor
  - ✓ Select ‘Join and Relates’ button
  - ✓ Double click on ‘Join’ bar (See Figure A.6(a))

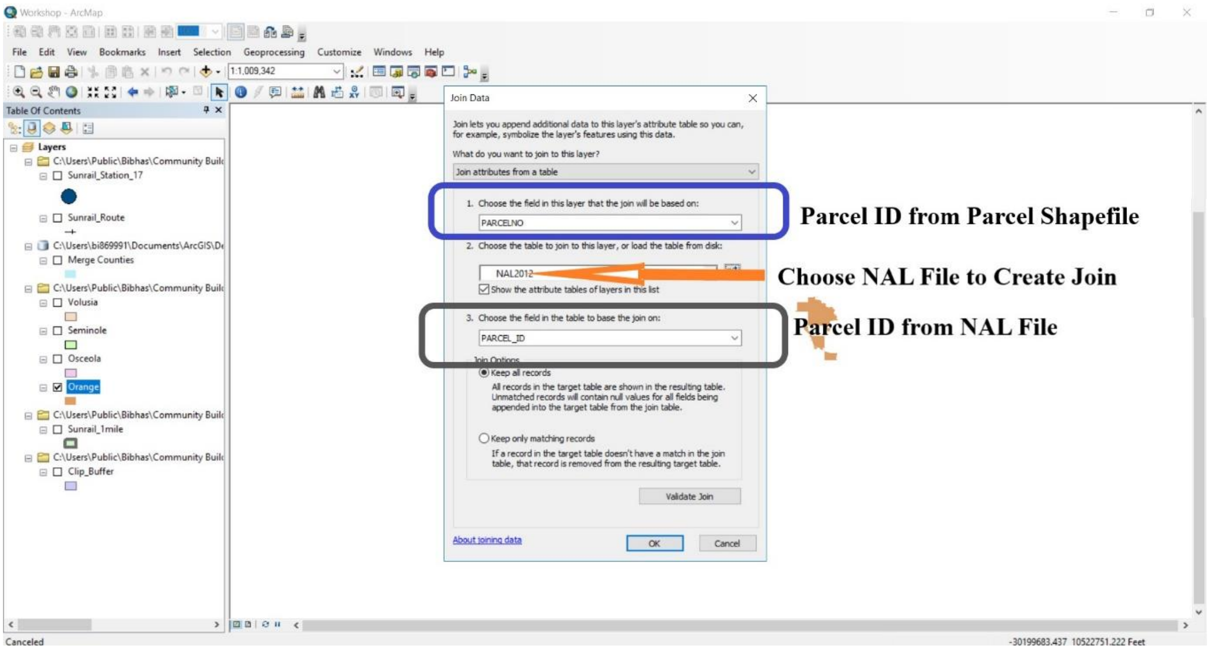
- There are three options in Join tool - 1 and 3 is the joining field option based on these options NAL file information were added to Parcel shapefile while 2 is the option where NAL file must be included (See Figure 6(b))
  - ✓ Select ‘Parcel No’ from Parcel shapefile in option 1
  - ✓ Choose ‘Parcel ID’ from NAL file in option 3

- ✓ Select NAL file for corresponding year for option 2
- ✓ Finally click Ok

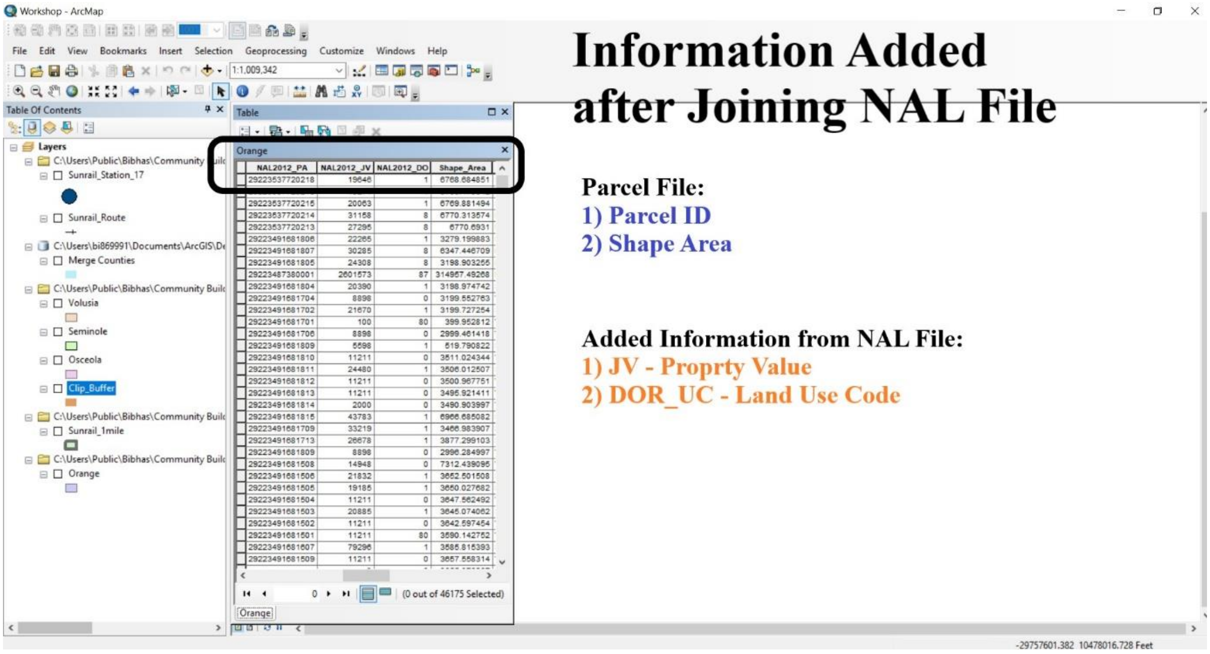
- From Figure 6(c), it is clearly seen that property value (JV) and land use type value (DOR_UC) was added to shape file attributes table.
  - ✓ Select shapefile and click on right side of mouse
  - ✓ Select 'Open Attribute Table' to see the new variables



(a)



(b)



(c)

Figure A.10: Adding NAL Information to Parcel Shapefile

A.2.2 Create Appropriate Information

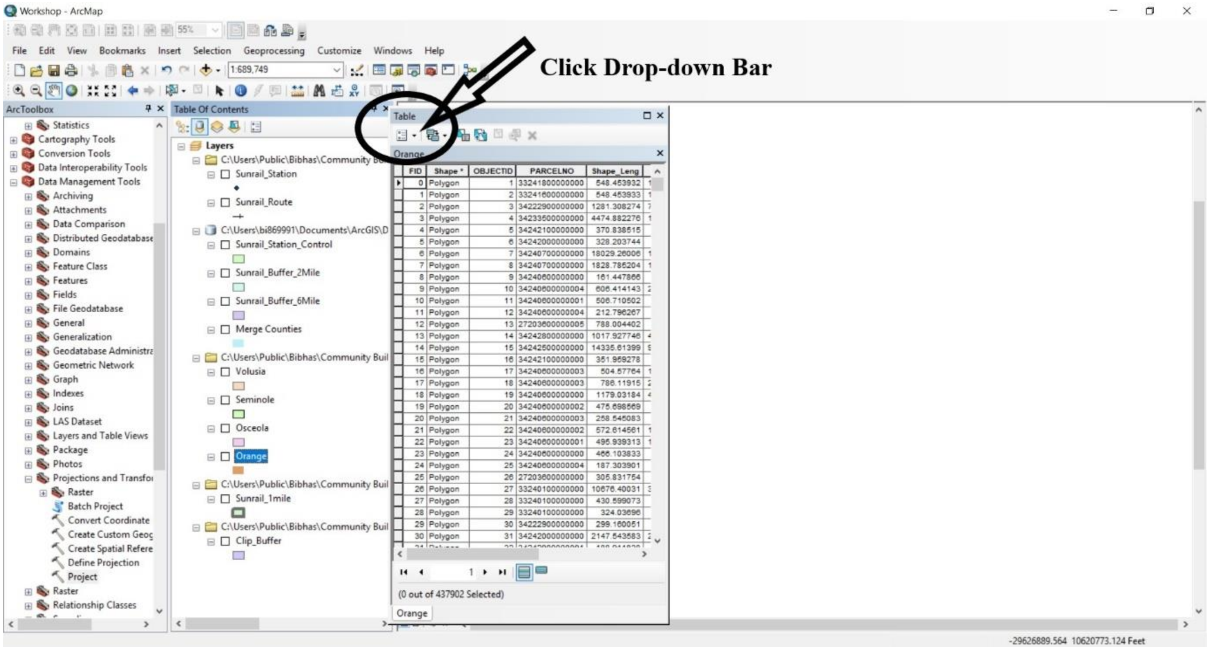
To continue further estimation, we categorized few land use types from DOR_UC (Land use value) information and area was transformed into acres unit also.

A.2.2.1 Land Use Type

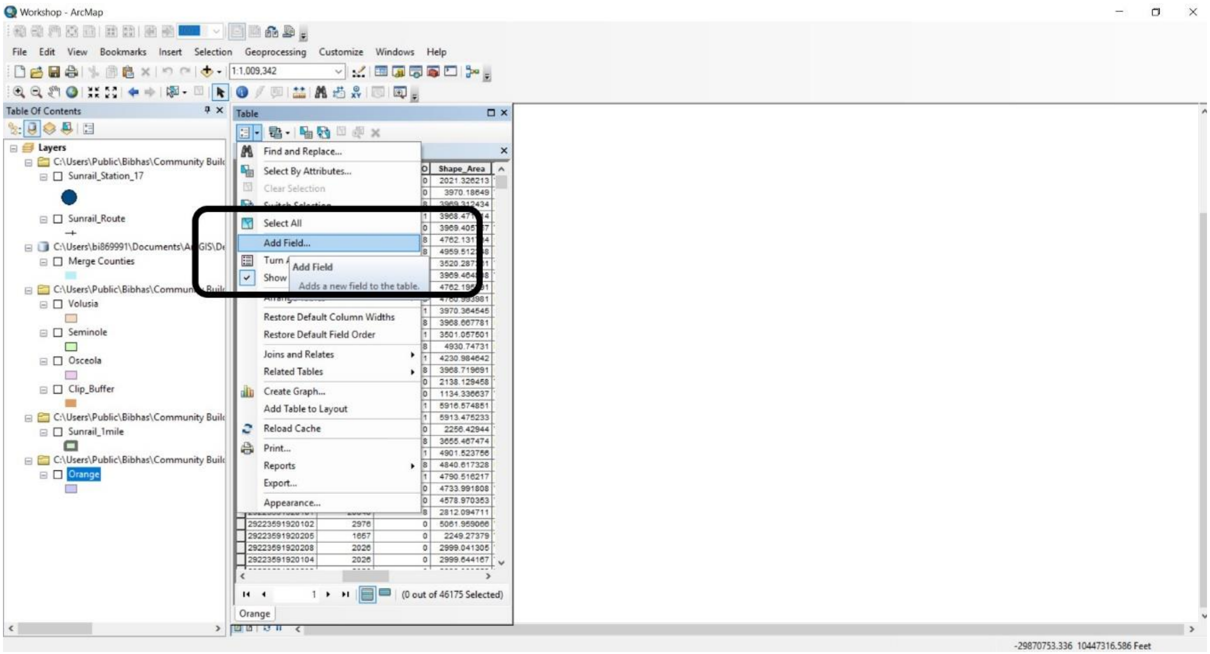
A.2.2.1.1 Adding New Field

Several steps were followed to categorize DOR_UC to various land use types as follows:

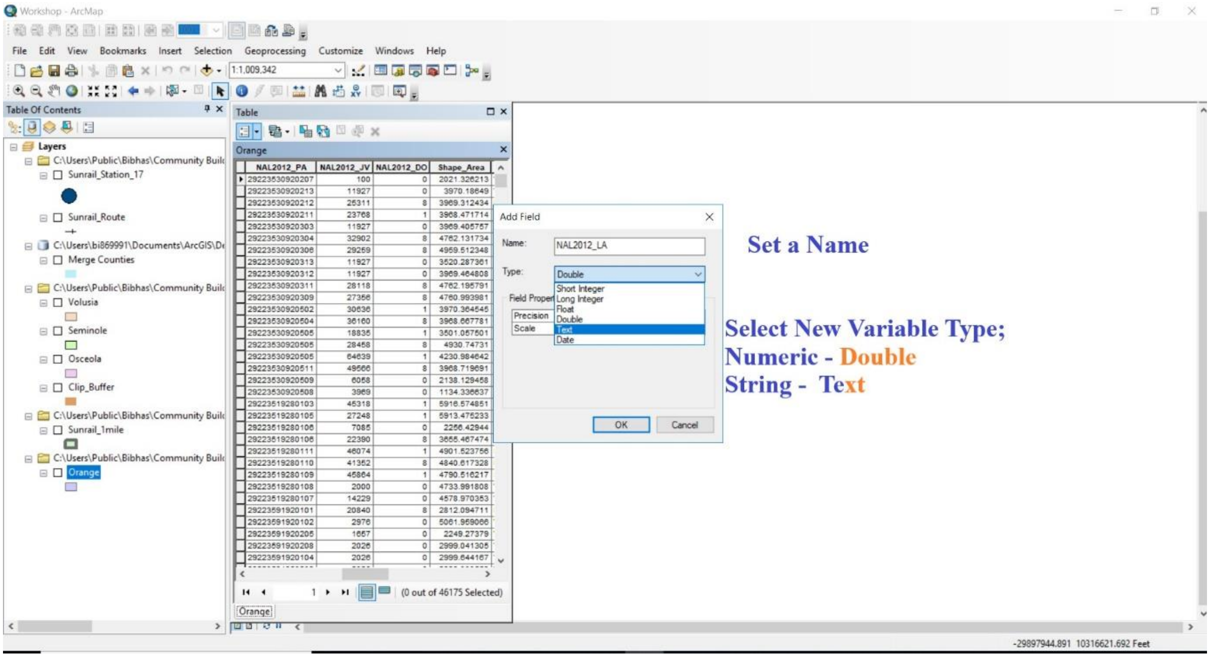
- Select shapefile and click on right cursor
- Select 'Open Attribute Table'
- Click drop-down bar as shown in Figure A.7(a)
- Click on 'Add Field' (See Figure A.7(b))
- Choose a new name as 'NAL2012_LA'
- Select variable type (See Figure A.7(c))
- Select 'String' as a variable type since land use type is a string variable (See Figure A.7(c))


Workshop - ArcMap
Click Drop-down Bar
ArcToolbox
Table Of Contents
Table
(0 out of 437902 Selected)
Orange

(a)


Workshop - ArcMap
Table Of Contents
Table
Find and Replace...
Select By Attributes...
Clear Selection
Select All
Add Field...
Add Field
Adds a new field to the table.
Restore Default Column Widths
Restore Default Field Order
Joins and Relates
Related Tables
Create Graph...
Add Table to Layout
Reload Cache
Print...
Reports
Export...
Appearance...
(0 out of 46175 Selected)
Orange

(b)



(c)

Figure A.11: New Field Adding Procedure

A.2.2.1.2 Selection of Land Use Category

- For our analysis purpose, we consolidated the land use categories reported by DOR into 12 land use categories. These are Single Family Residential, Multi-Family Residential, Retail/Office, Industrial/Manufacturing, Agriculture, Institutional/Infrastructure, Public, Recreational, Water, Vacant, and Others (see Table A.1). See Appendix A for DOR land use code.
- However, we will be reporting values for the following 5 out of the 12 categories: (1) Single family residential, (2) Multiple family residential, (3) Retail/Office area, (4) Institutional, and (5) Industrial.

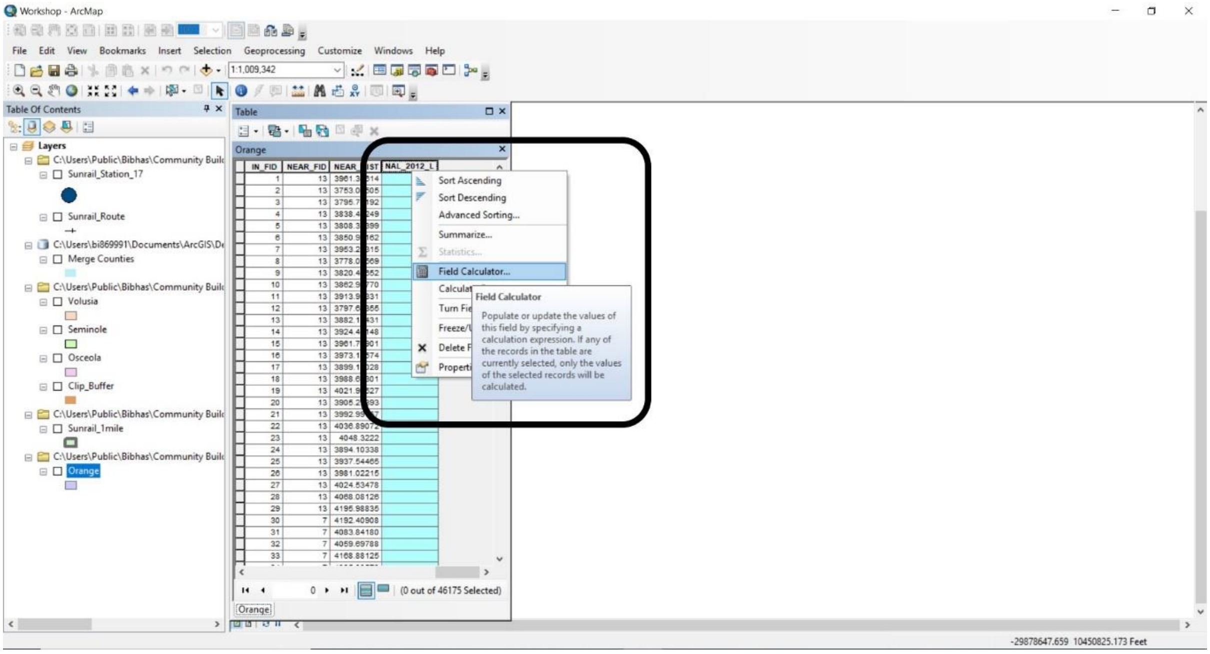
Table A.41: Land Use Category Based on DOR Land Use Codes

Land Use Category	DOR Land Use Code
Single Family Residential	1
Multi-Family Residential	3,8
Other Residential	2,4-7,9
Retail/Office	11-39
Industrial	41-49
Agricultural	50-69
Institutional	71-79, 81, 84
Public	83, 85-91
Recreational	82, 97
Water	95
Vacant	0, 10, 40, 70, 80
Others	92-96, 98, 99, 100, 995, 999

A.2.2.1.3 Land Use Type Conversion Technique from DOR Land Value

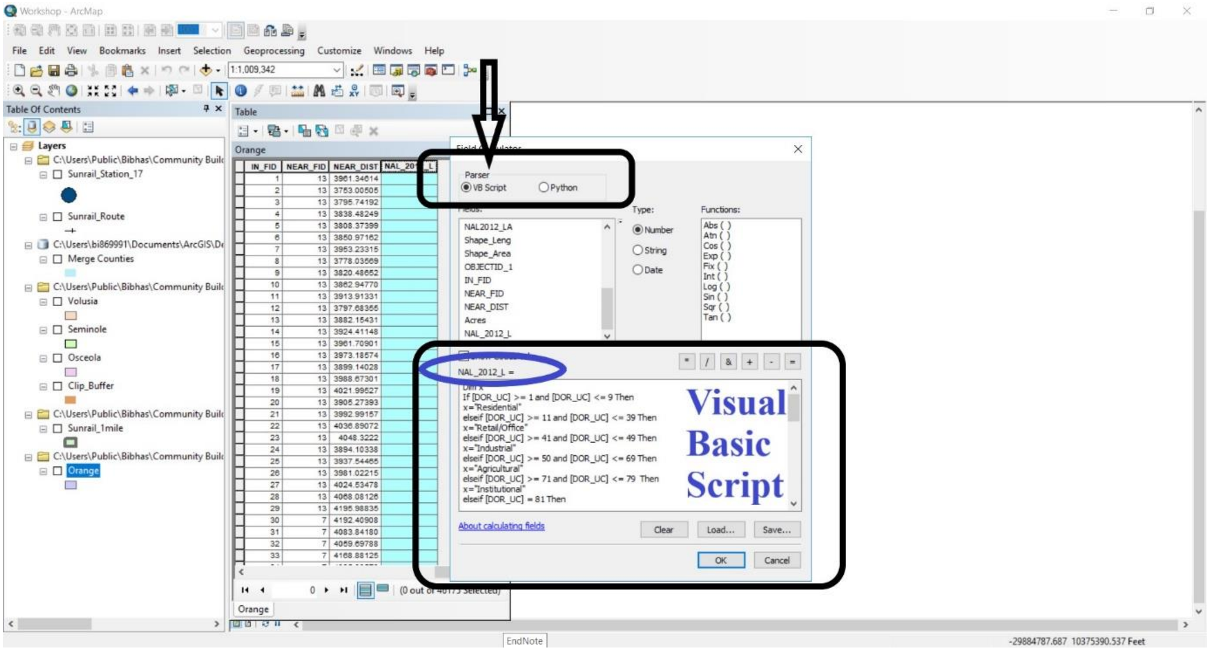
Following steps were used to convert DOR_UC land value to selected 12 land use categories.

- Select new added variable 'NAL2012_LA' and click on right cursor
- Select 'Field Calculator' (See Figure A.8(a))



(a)

- One can select either Visual Basic (VB) Script or Python option
- VB Script was selected for conversion of land use type category (See Figure A.8(b))
- A code was written within the box to convert DOR land value (See Table A.2)



(b)

Figure A.12: Land Use Type Conversion Technique from DOR Land Value

Table A.42: VB Script for Land Use Type Conversion

```
Dim x
If [DOR_UC] >= 1 and [DOR_UC] <= 9 Then
x="Residential"
elseif [DOR_UC] >= 11 and [DOR_UC] <= 39 Then
x="Retail/Office"
elseif [DOR_UC] >= 41 and [DOR_UC] <= 49 Then
x="Industrial"
elseif [DOR_UC] >= 50 and [DOR_UC] <= 69 Then
x="Agricultural"
elseif [DOR_UC] >= 71 and [DOR_UC] <= 79  Then
x="Institutional"
elseif [DOR_UC] = 81 Then
x="Institutional"
elseif [DOR_UC] = 84 Then
x="Institutional"
elseif [DOR_UC] >= 85 and [DOR_UC] <= 91  Then
x="Public"
elseif [DOR_UC] = 83 Then
x="Public"
elseif [DOR_UC] >= 92 and [DOR_UC] <= 96  Then
x="Other"
elseif [DOR_UC] = 98 Then
x="Other"
elseif [DOR_UC] = 99 Then
x="Other"
elseif [DOR_UC] = 82 Then
x="Recreational"
elseif [DOR_UC] = 97 Then
x="Recreational"
elseif [DOR_UC] = 95 Then
x="Water"
elseif [DOR_UC] = 0 Then
x="Vacant"
elseif [DOR_UC] = 10 Then
x="Vacant"
elseif [DOR_UC] = 40 Then
x="Vacant"
elseif [DOR_UC] = 70 Then
x="Vacant"
elseif [DOR_UC] = 80 Then
x="Vacant"

else x=0

end if
```

A.2.2.2 Area Unit Conversion

A.2.2.2.1 Adding New Field

Several steps were followed to categorize DOR_UC to various land use types as follows (See Figure A.9).

- Select shapefile and click on right cursor
- Select 'Open Attribute Table'
- Click drop-down bar
- Click on 'Add Field'
- Choose a new name as 'Area_Acres'
- Select variable type
- Select 'Double' as a variable type since area is a numeric variable

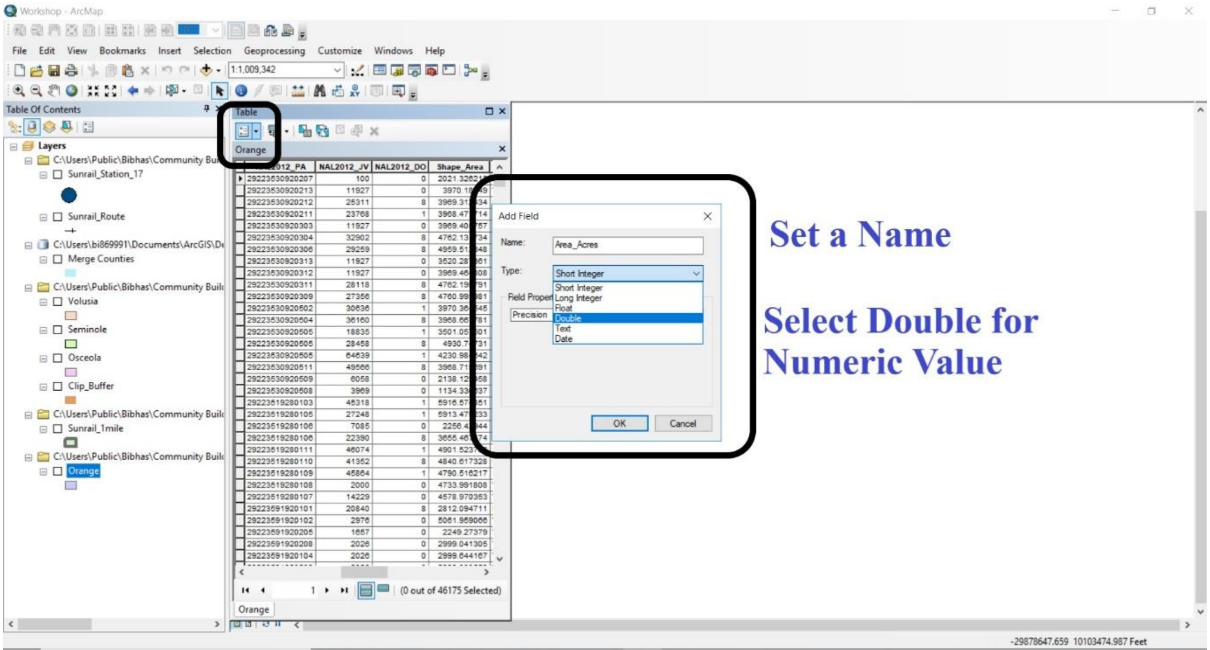
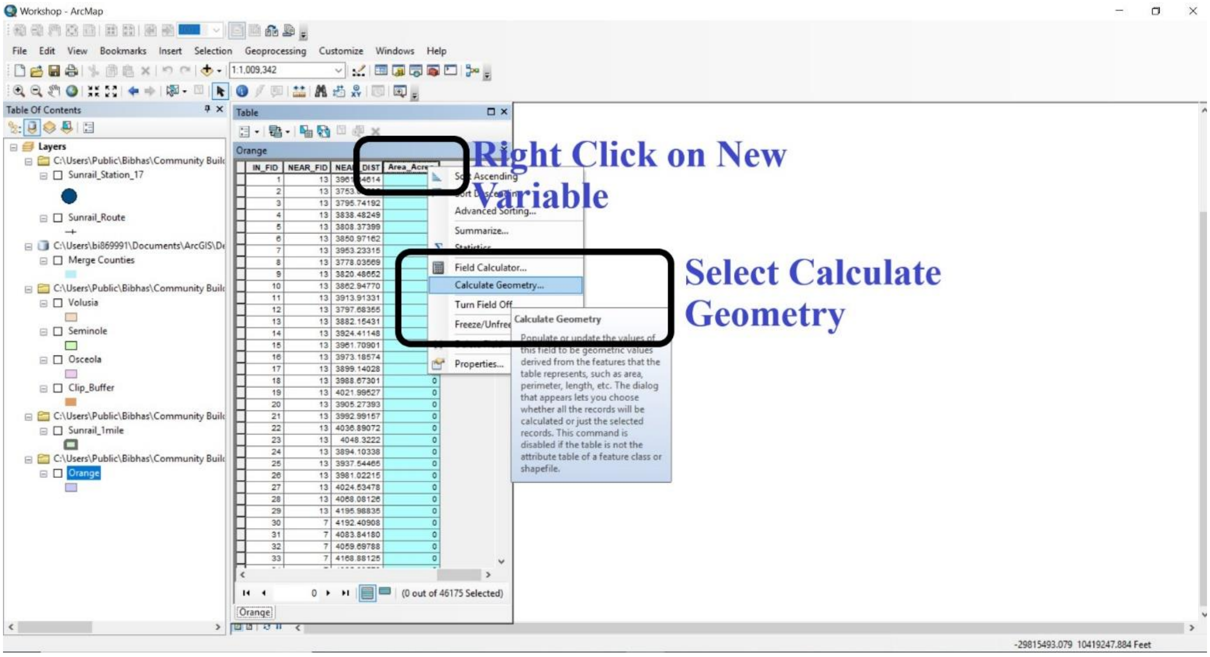


Figure A.13: Area Unit Conversion

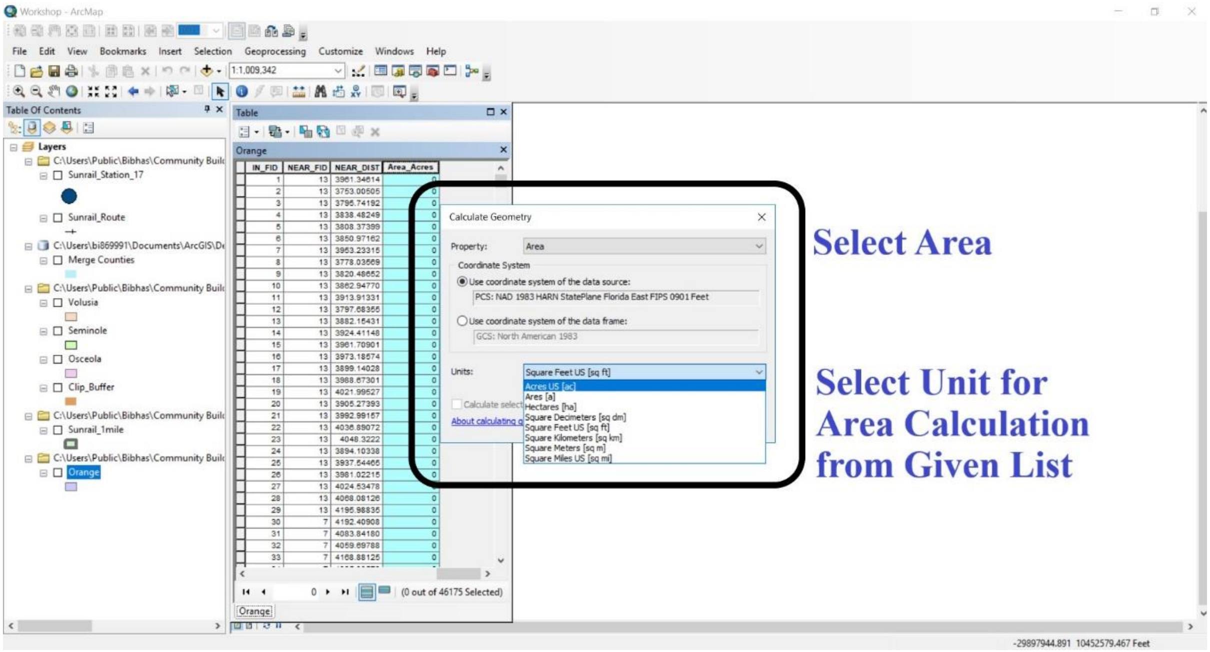
A.2.2.2.1 Area Unit Conversion

Following steps were used to convert shape area to new area unit such as acres.

- Select new added variable 'Area_Acres' and click on right cursor
- Select 'Calculate Geometry' (See Figure A.10(a))
- Then choose Area option on top and select area unit such as Acres (See Figure A.10(b))



(a)



(b)

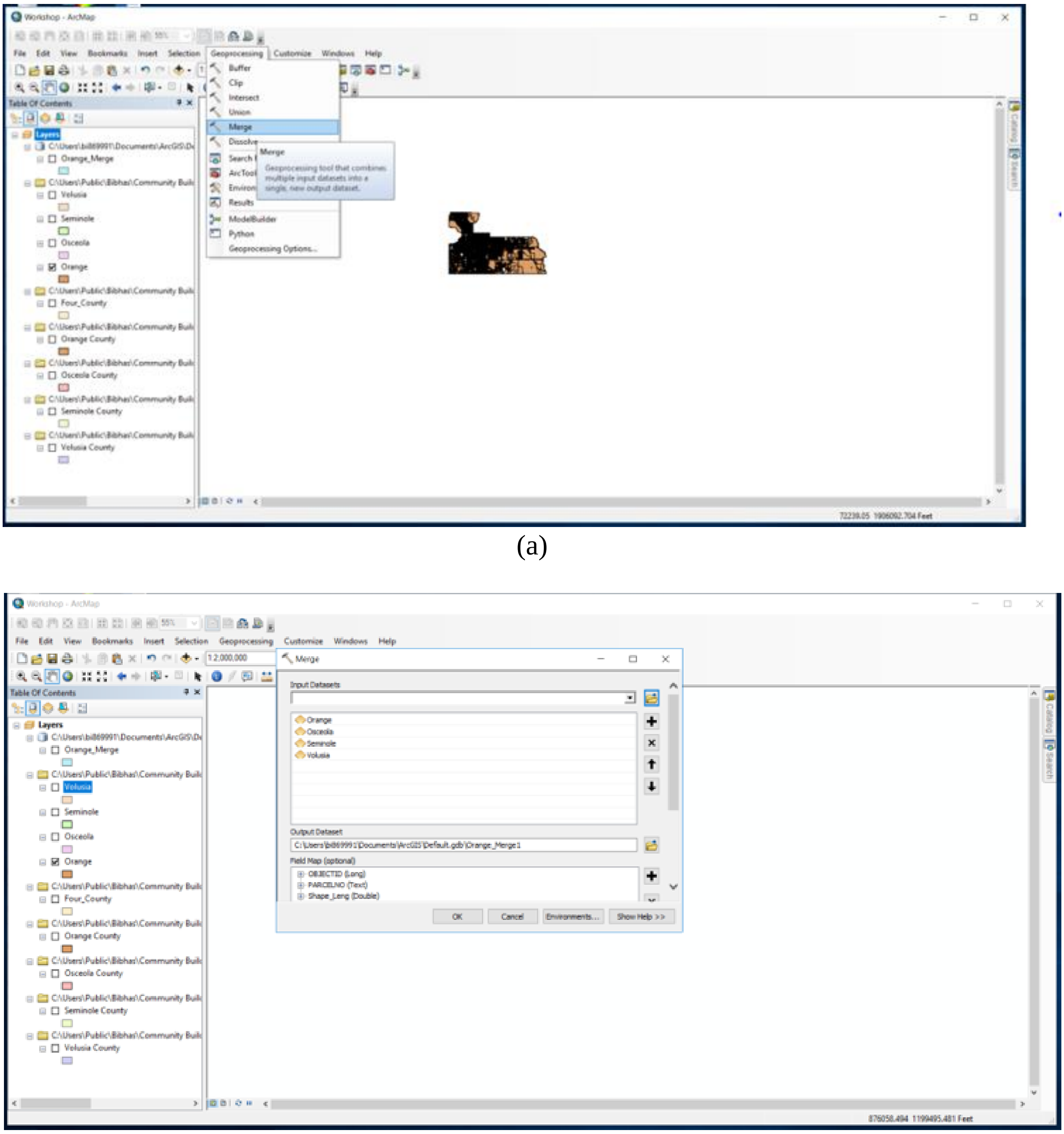
Figure A.14: Area Unit Conversion

A.2.3 Merge Counties

After preparing parcel data layer for all four counties (Orange, Seminole, Seminole and Volusia), a merged county shapefile was created. Following steps were followed to merge all four counties.

- At first, select 'Geoprocessing' toolbar
- Then click on 'Merge' option (see Figure A.11(a))
- Then put all of the counties within 'Merge' toolbar (see Figure A.11(b))
- Finally, click 'Ok'

After merge all counties a new shapefile was created (see Figure A.12).



(a)

(b)

Figure A.15: Merging Techniques

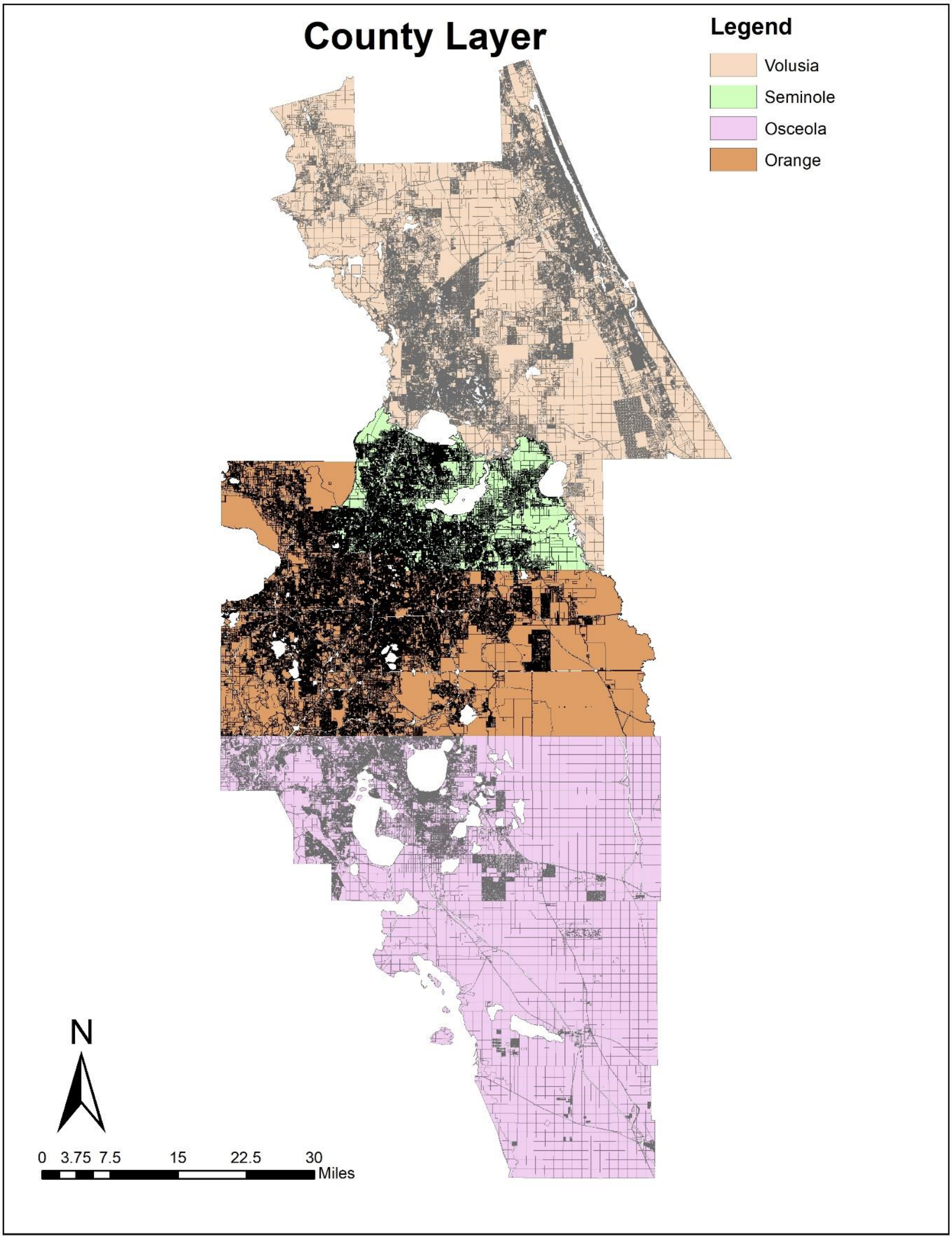


Figure A.16: Merged Counties Shapefile

A.2.4 SunRail Stations Layer Preparation

We divided the stations into three segments: (1) Downtown Stations[1] including Lynx Central station, Church Street station, and Orlando Health/Amtrak station; (2) Outside Downtown Stations comprised of DeBary, Sanford, Lake Mary, Longwood, Altamonte Springs, Maitland, Winter Park, Florida Hospital Health Village, and Sand Lake Road stations; (3) Phase-2 stations including northbound DeLand and Southbound Meadow Woods, Osceola Parkway, Kissimmee Amtrak, and Poinciana stations. Figure A.13 represents all 17 SunRail stations along with SnRail route.

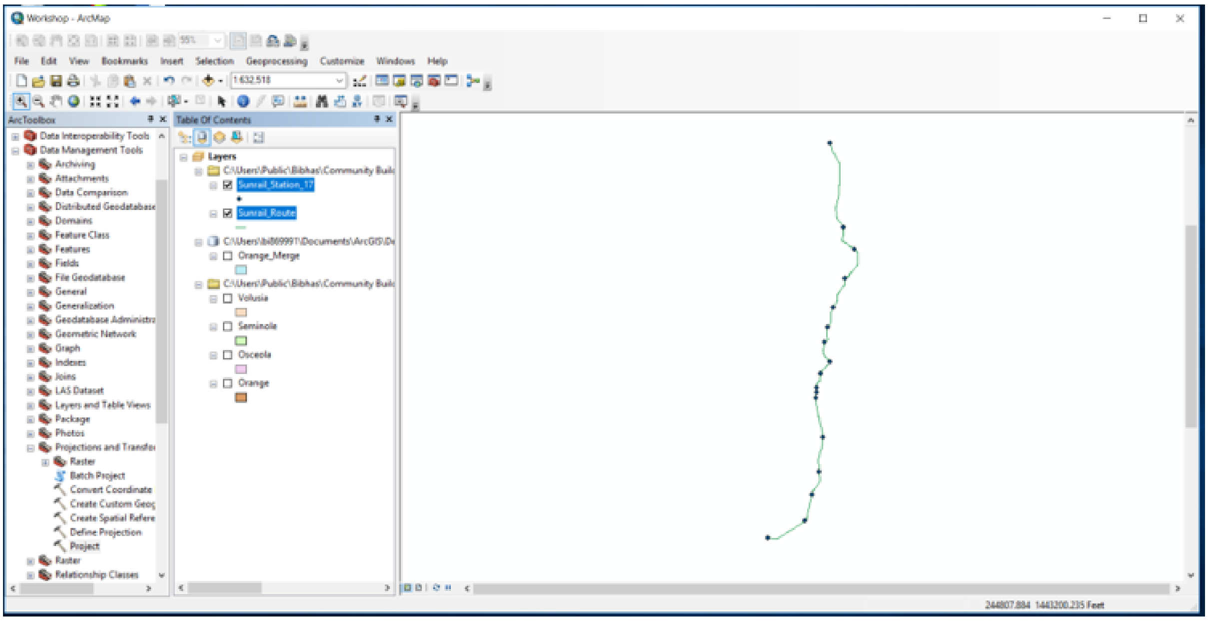


Figure A.17: SunRail Staions

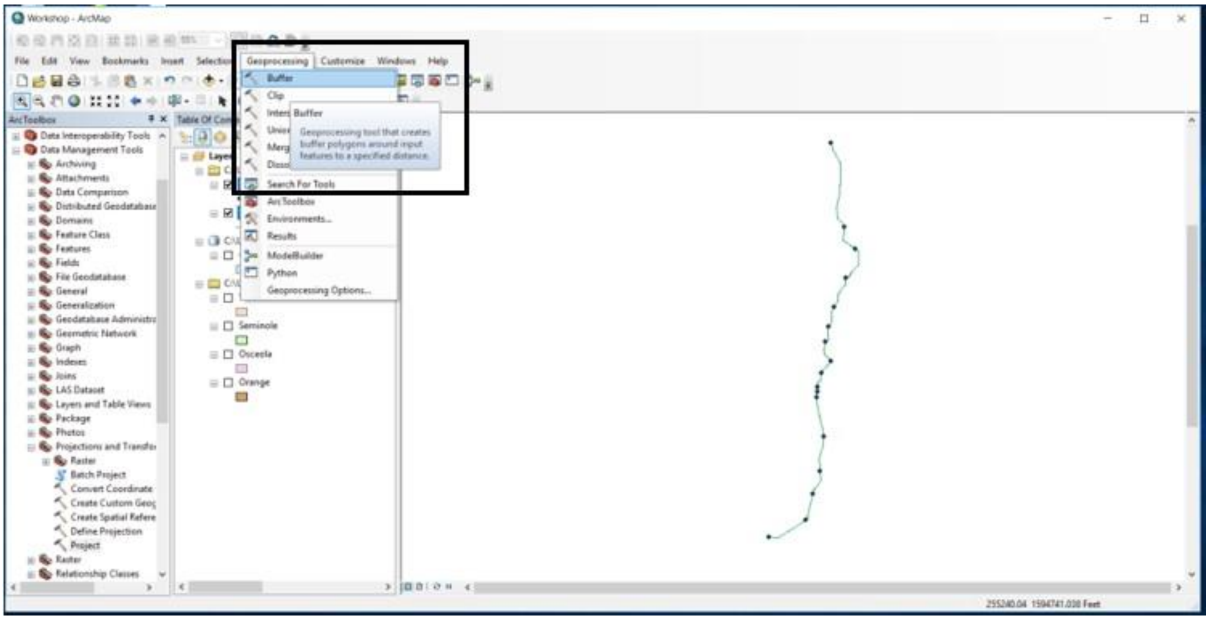
A.2.4.1 Case Area Selection

A.2.4.1.1 Creating Buffer

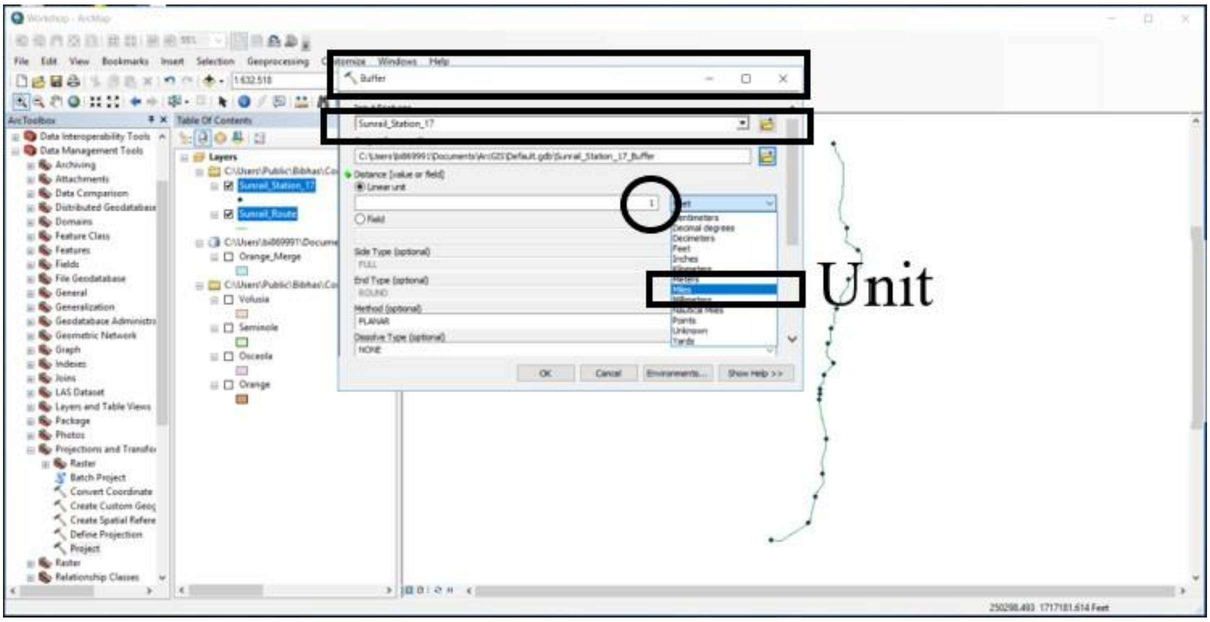
- At first, select ‘Geoprocessing’ toolbar
- Then click on ‘Buffer’ option (see Figure A.14(a))
- Then put SunRail station’s shapefile in ‘Buffer’ toolbar (see Figure A.14(b))
- In ‘Linear Unit’ option put the numeric value such as 1 and choose mile as unit
- Select default buffer type ‘Round’
- Finally, click ‘Ok’

A new buffer map was created around SunRail stations (see Figure A.14(c)).

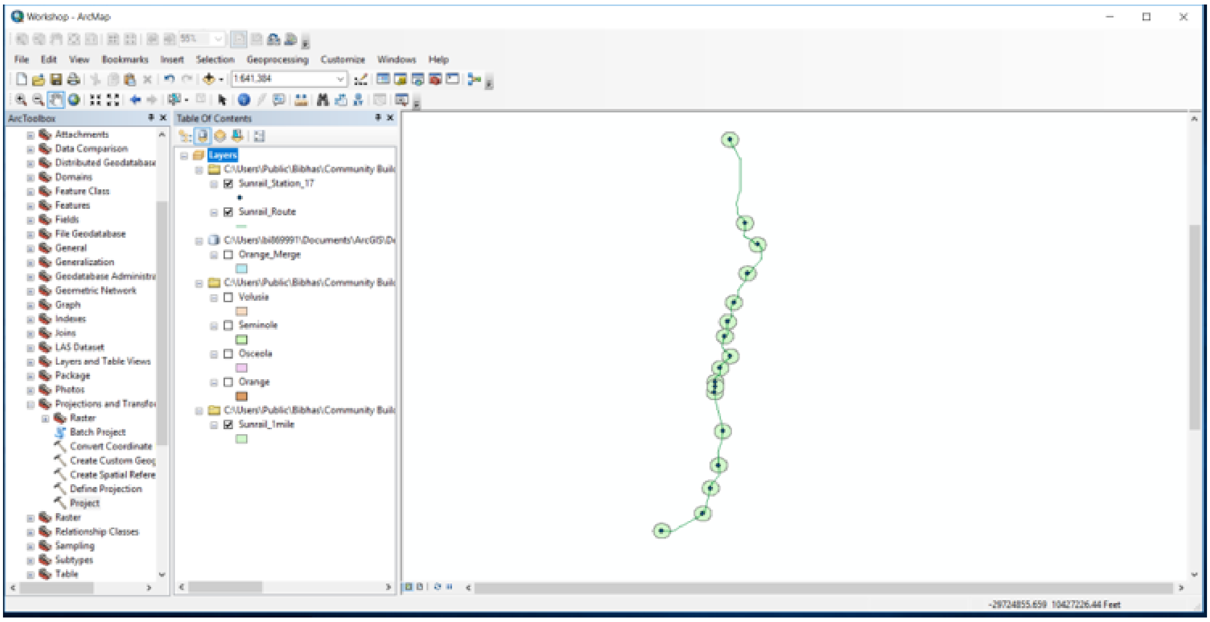
[1]Downtown Stations are fixed based on the downtown area projected at ‘I-4 Ultimate Project’ construction map at https://i4ultimate.com/construction-info/construction-map/#constructionAlerts


Workshop - ArcMap
File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help
Buffer
Clip
Geoprocessing tool that creates buffer polygons around input features to a specified distance.
Search For Tools
ArcToolbox
Environments...
Results
ModelBuilder
Python
Geoprocessing Options...

(a)


Buffer
Sunrail_Station_17
Unit

(b)

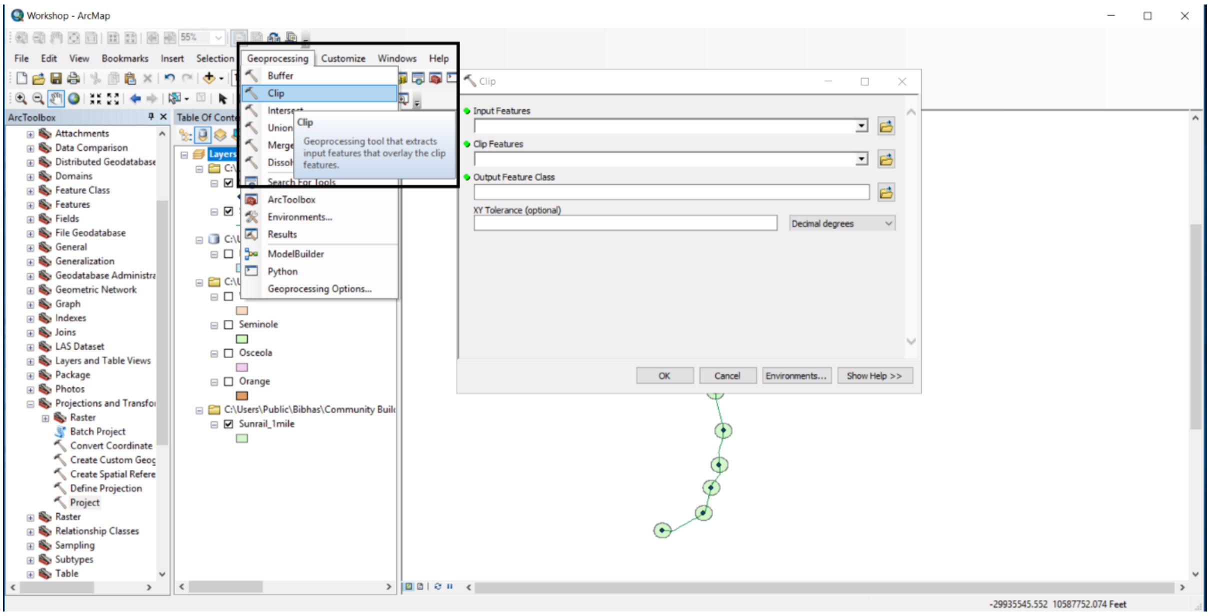


(c)

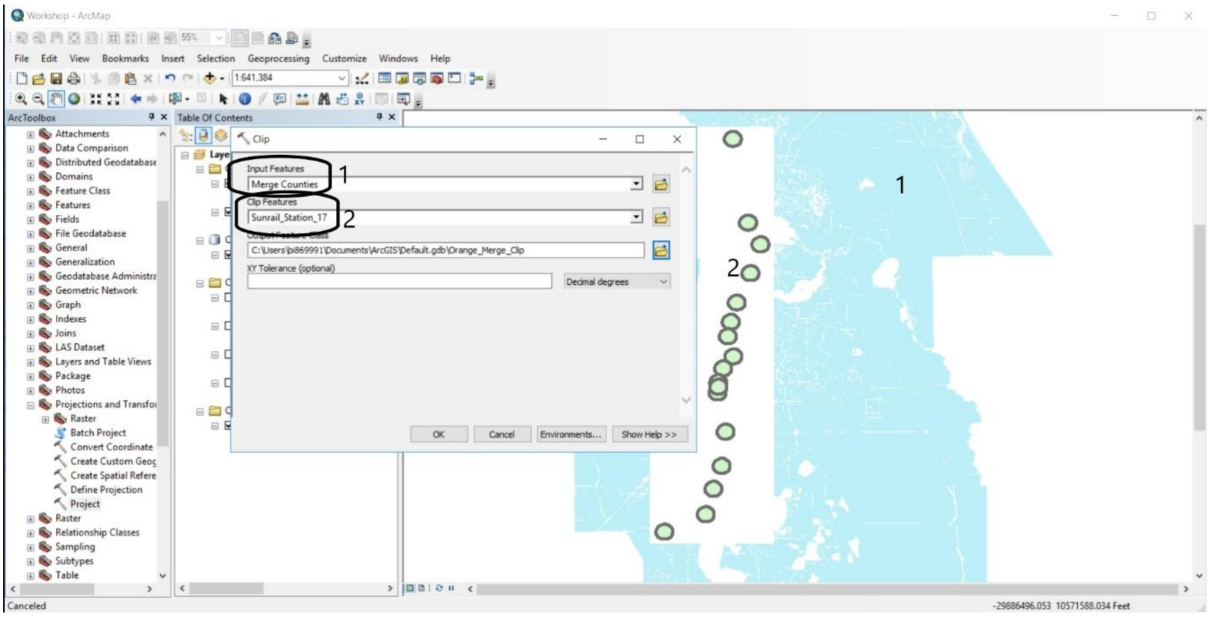
Figure A.18: 1 mile Buffer Around SunRail Stations

A.2.4.1.2 Clip from Merge Counties

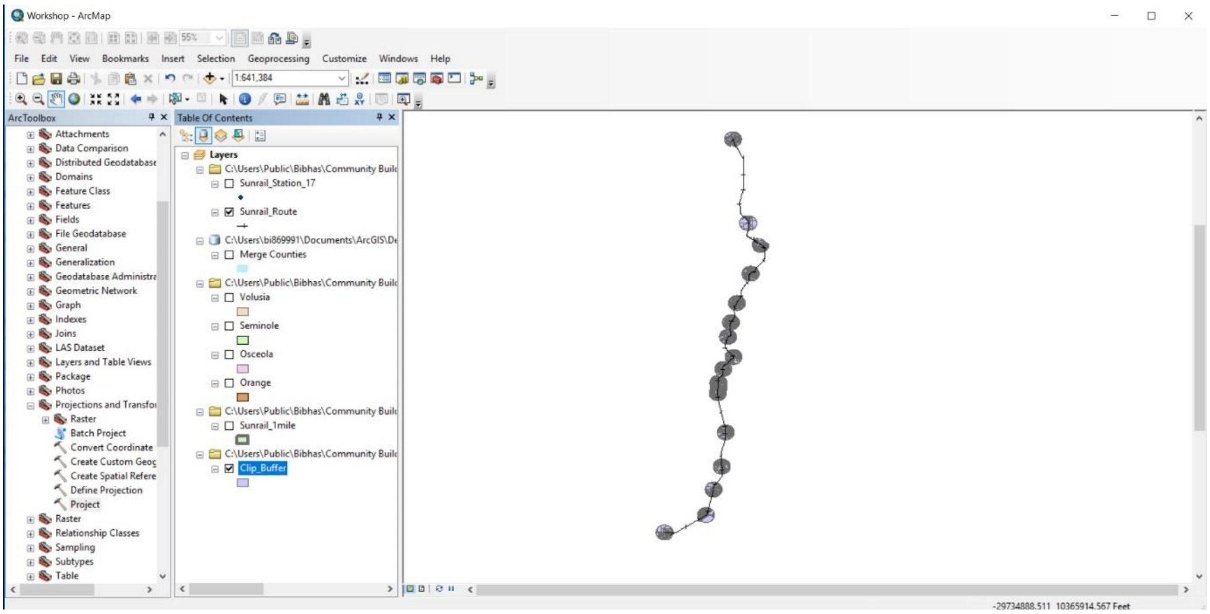
- At first, select 'Geoprocessing' toolbar
- Then click on 'Clip' option (see Figure A.15(a))
- In 'Input Features'section put Merge counties (see Figure A.15(b))
- Upload SunRail buffer created in previous step on 'Clip Features' section (see Figure A.15(b))
- Finally, a new buffer layer was created contains corresponding parcel level information such property value, land use type, area etc. (see Figure A.15(c))



(a)



(b)



(c)

Figure A.19: Case Area Selection

A.2.4.2 Overlapping Problem

A.2.4.2.1 Theoretical Approach

- A 1-mile buffer was created around each of the SunRail stations. Please note that the nearness of the stations, particularly in the downtown areas, cause overlapping problem.
- As a result of the overlapping, the same parcel might be part of two different stations.
- ArcGIS proximity tool (Near Generate Table operation) was used to assign a parcel to a unique station. More specifically, we computed the straight line distances from each parcel to the nearest station and the parcel was assigned to the station which was the nearest. Figure A.16 demonstrates an example of the station overlapping problem in the downtown area.

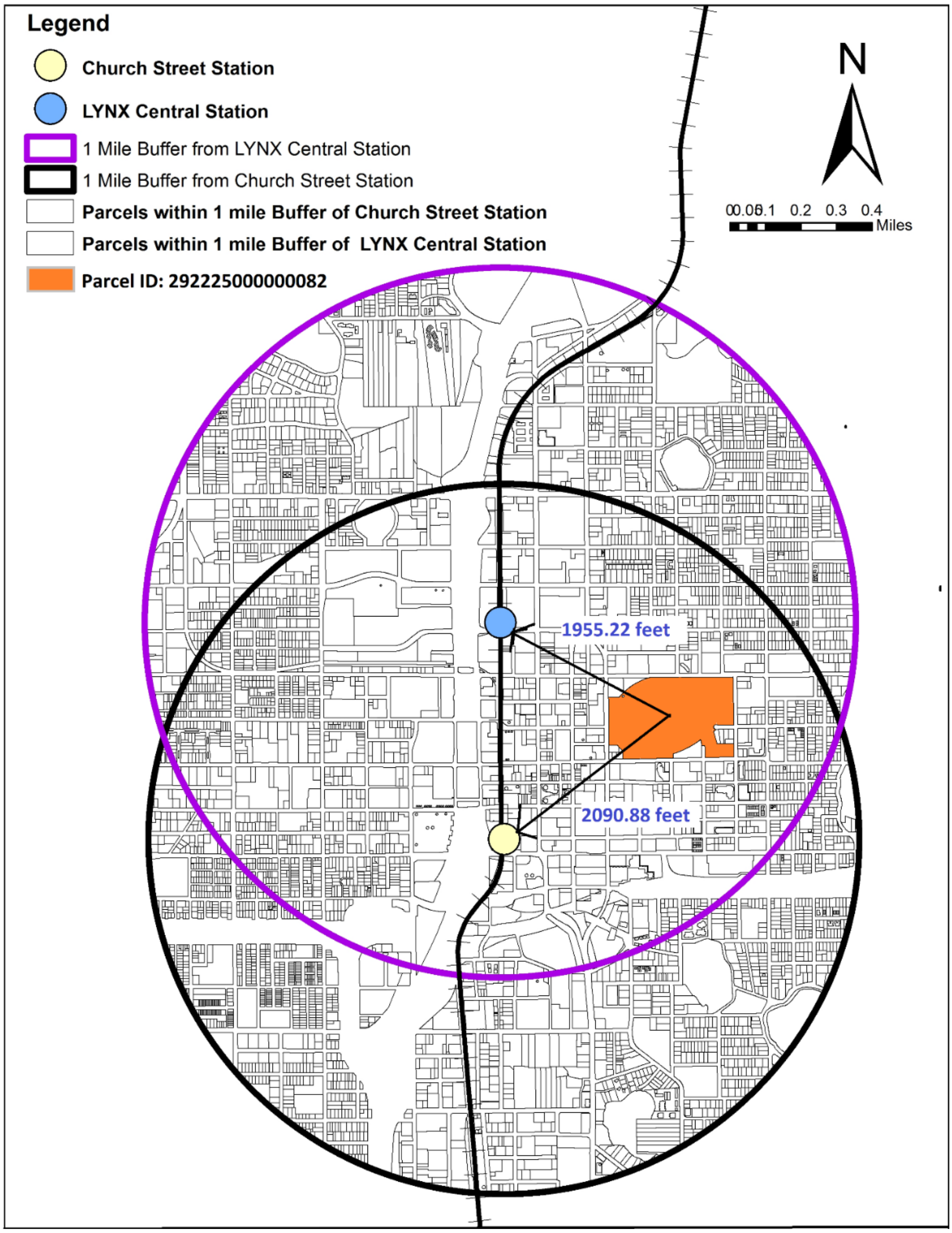


Figure A.20: Example of Overlapping Buffers and Proximity Analysis

A.2.4.2.2 Practical Overlapping Solution Technique

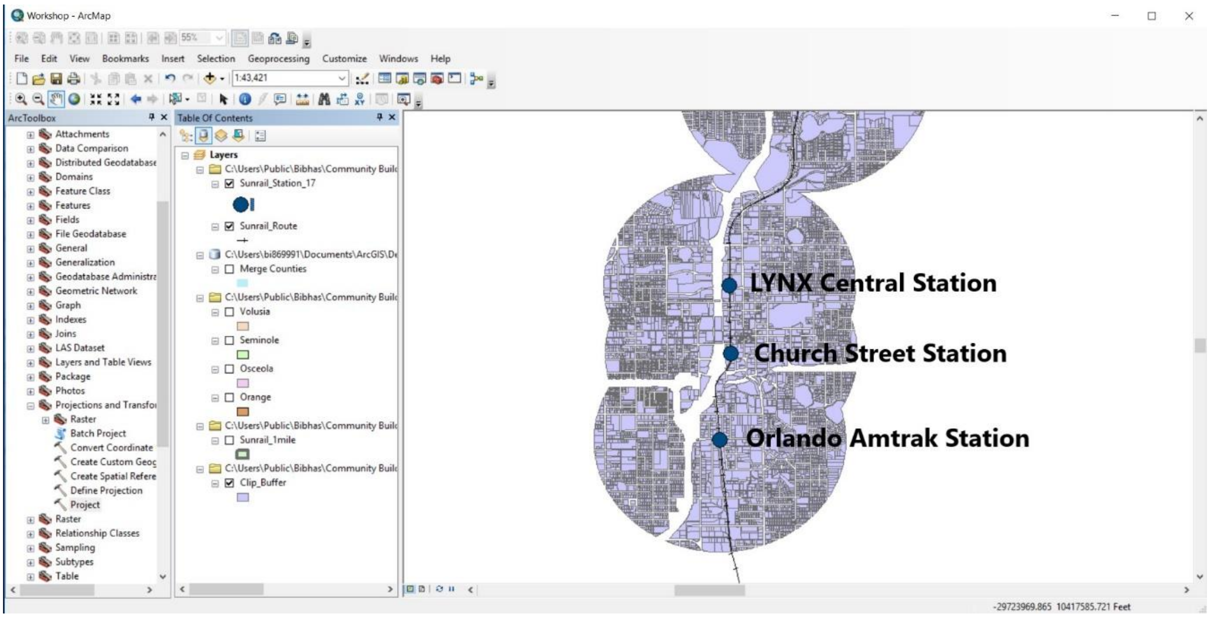
Three downtown stations are clear example of overlapping parcels (Figure A.17(a)). Following are the steps to solve the overlapping buffer problem:

- 1st select Geoprocessing
- Choose Arc Toolbox bar in Geoprocessing (See Figure A.17(b))
- Select 'Analysis Tools' (See Figure A.17(b))
- Select 'Proximity' from 'Analysis Tools'
- Then click to 'Near' for the overlapping solution
- Select parcel buffer layer shapefile in as input features (See Figure A.17(c))
- Then put SunRail stations as Near Features

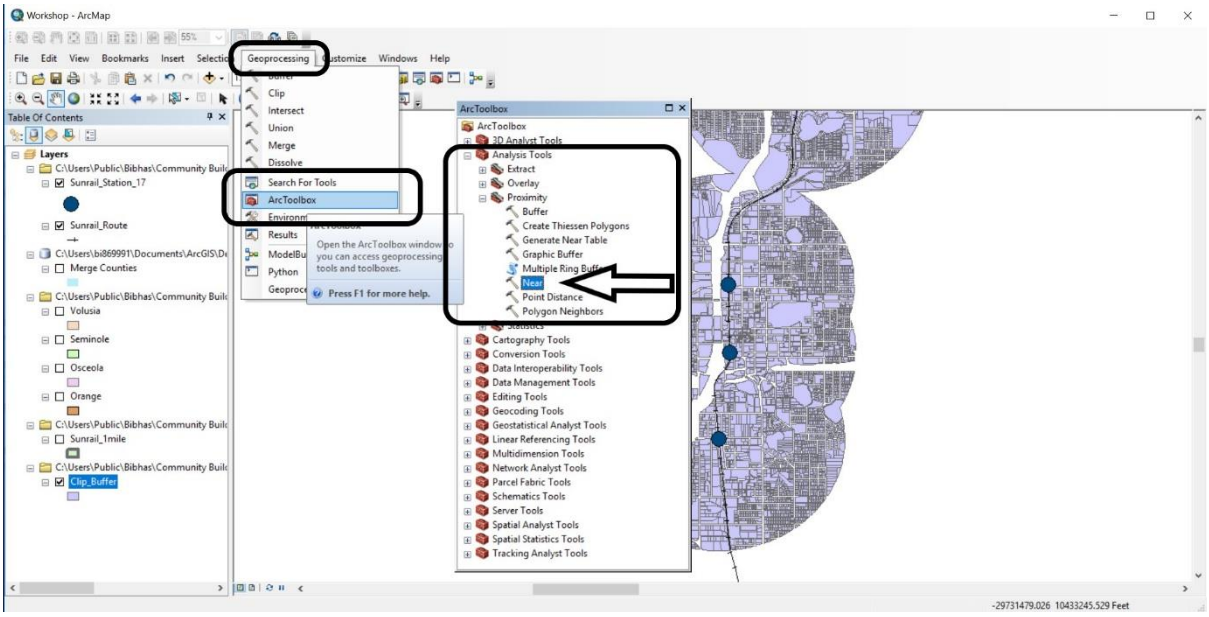
- Click ‘Ok’

After all the actions were taken, three new columns will be added in the attribute table (See Figure A.17(d)). They are:

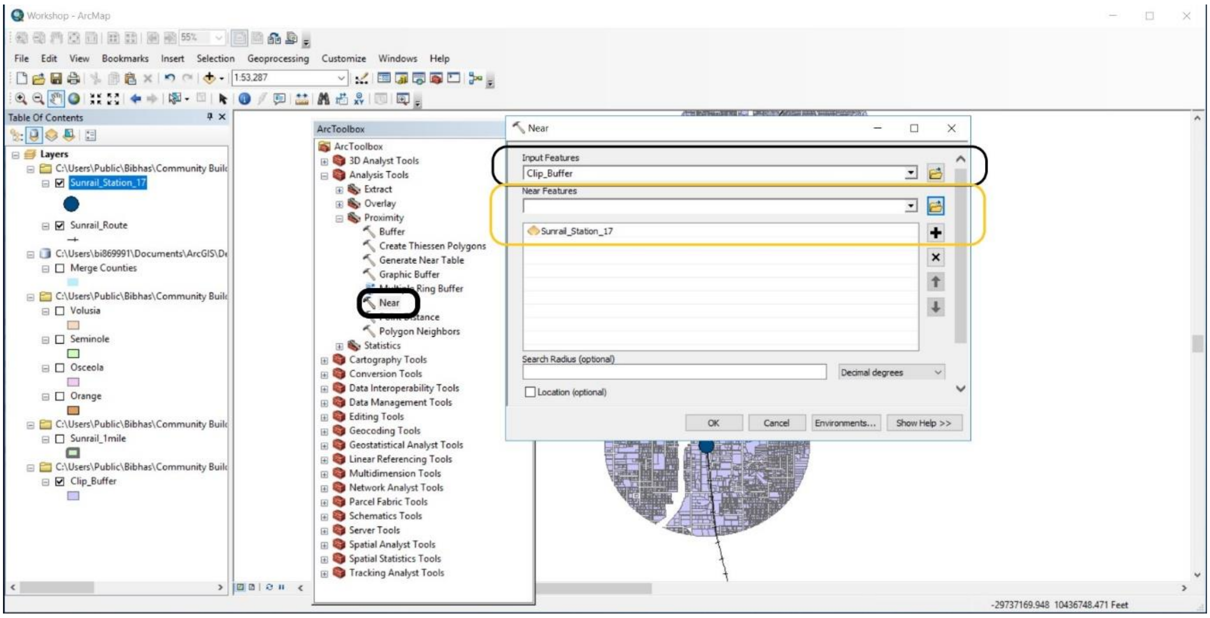
- In_FID = Parcel ID
- Near_FID = Station ID
- Near_Dist = Estimated nearest distance from parcel to each SunRail statins



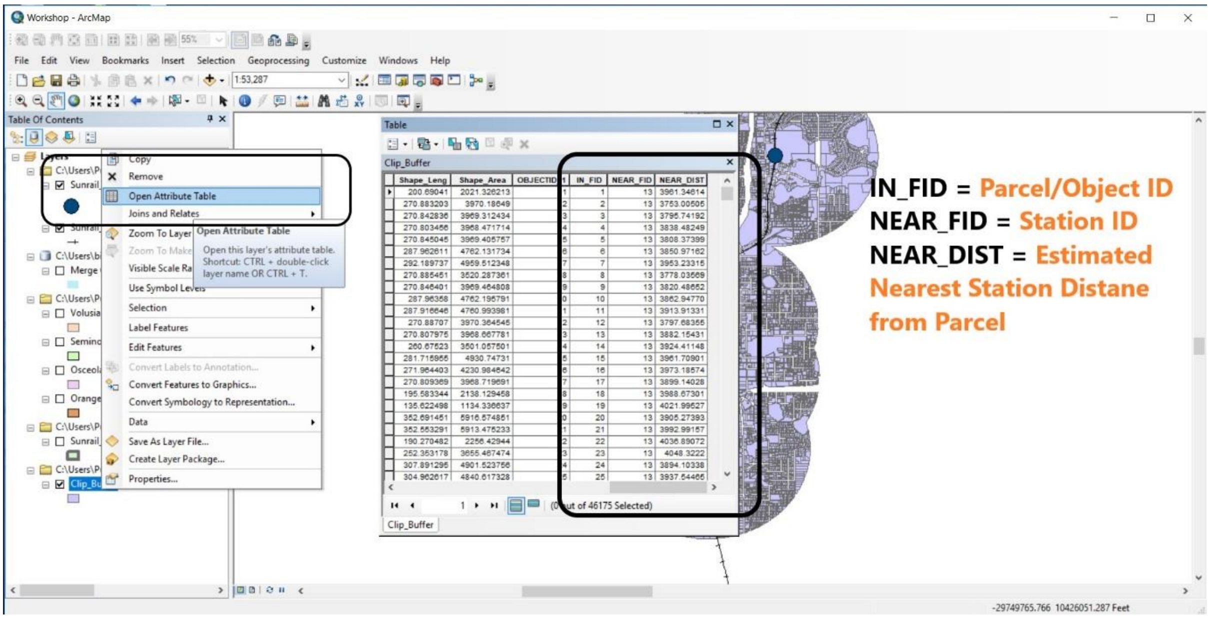
(a)



(b)



(c)



(d)

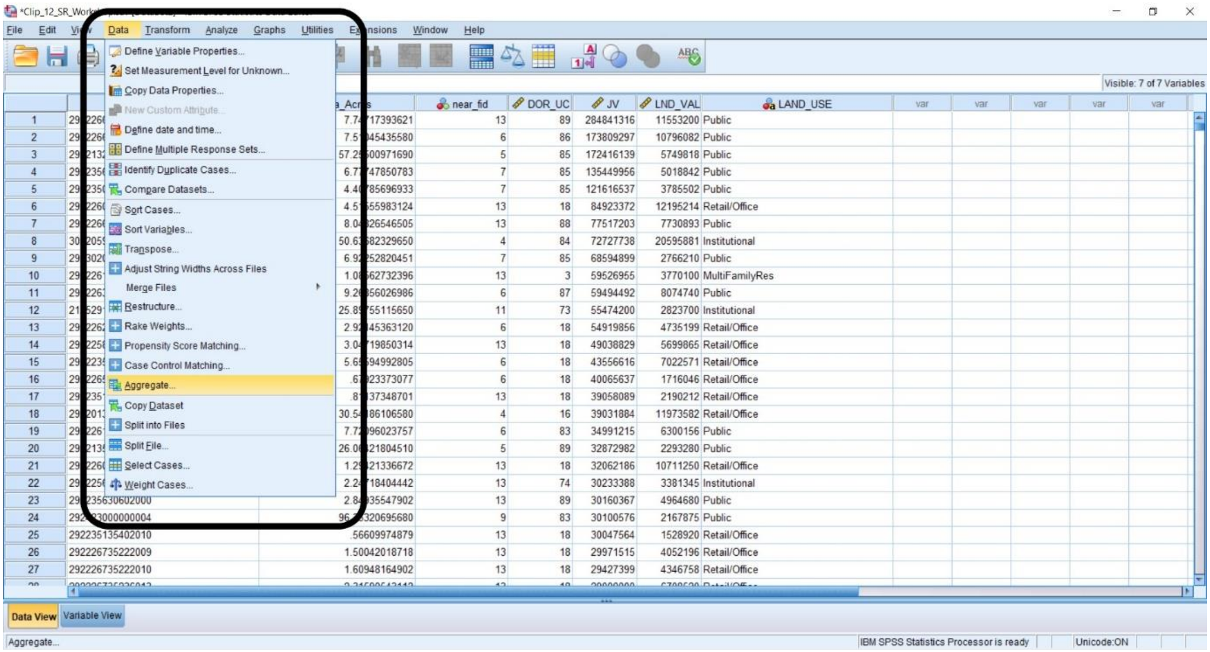
Figure A.21: Uses of Near Tool to Overcome Overlapping Problem on ArcGIS

A.2.4.3 Property Value Estimation

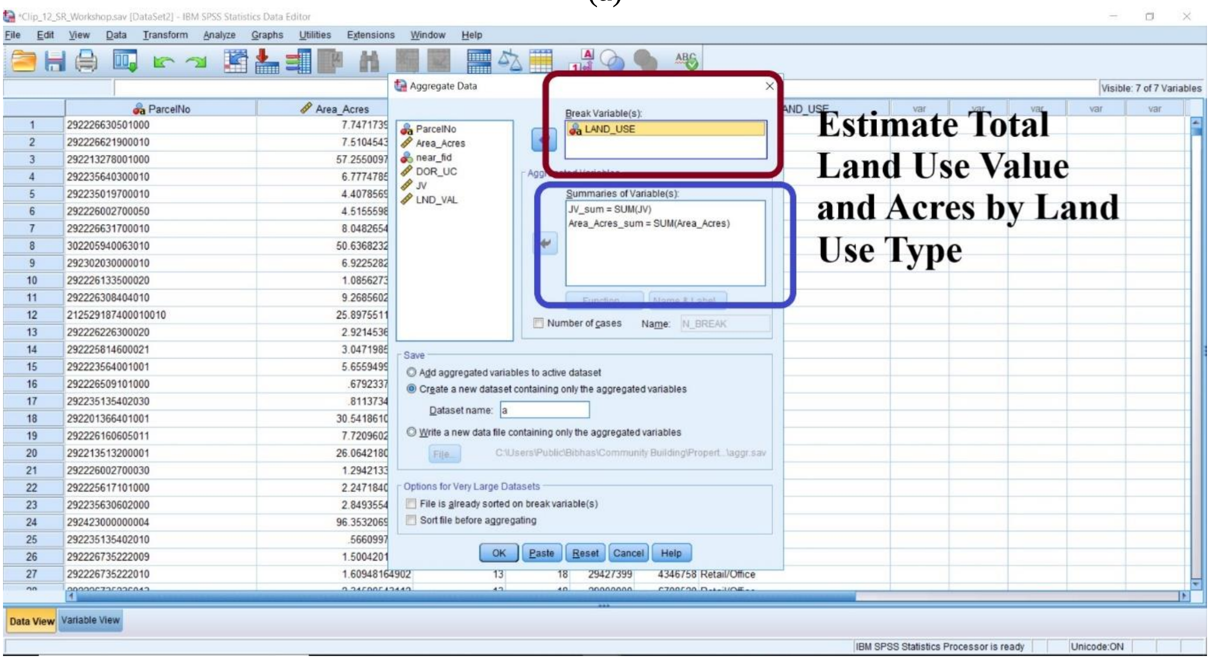
After allocating all parcels to their nearest stations, dbf file was converted to SPSS file for estimation of property value by land use type. Following steps were followed:

- Select 'Data' toolbar
- Click on 'Aggregate' option (See Figure A.18(a))
- Put 'Land Use Type' as break variable (See Figure A.18(b))
- For 'Summary of Variables' section choose JV as property value and Area (Acres)
- Also change the 'Function' option from default 'Mean' to 'Sum'

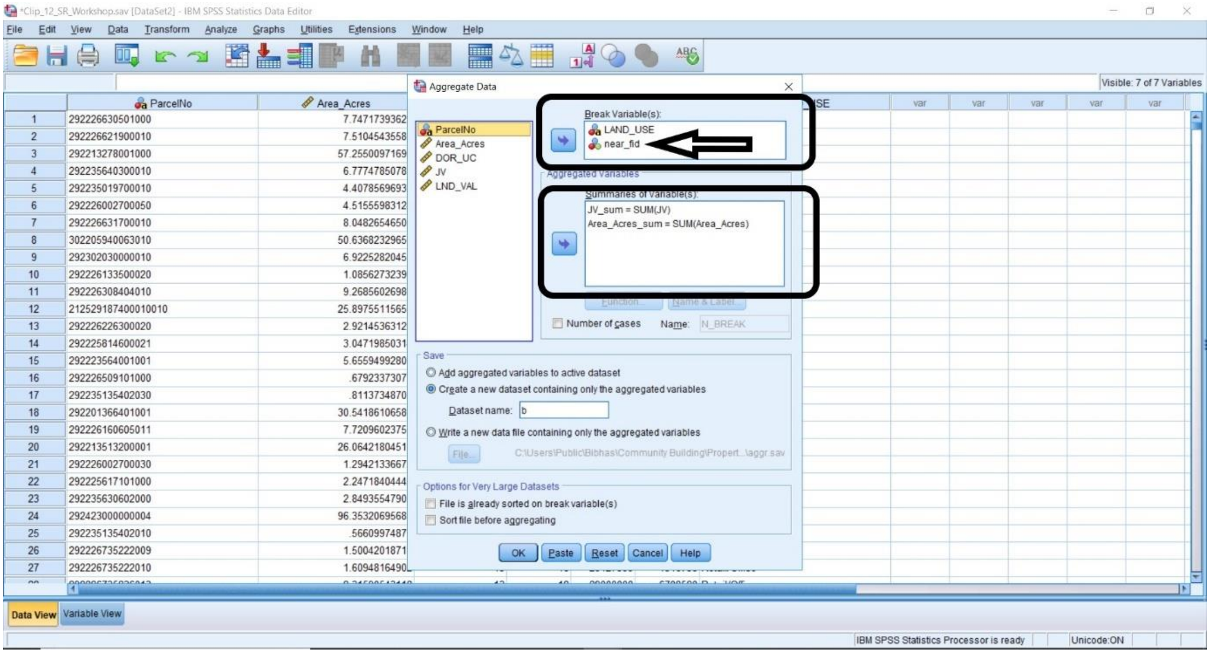
This action will give a new dataset of total property value and total area in acres for each land use type. Then average property value was estimated for each land use type by dividing the total property value by total area in acres. Please note that, property value by land use type around each SunRail station' buffer needed to be estimated, then put 'Land Use Type' and 'SunRail Station ID (Near_FID)' in 'Aggregate' section as break variables (See Figure 18(c)).



(a)



(b)



(c)

Figure A.22: Average Property Value Estimation

A.2.4.4 Average Property Value Estimation

- The property value evaluation was carried out for the parcels within the 1-mile buffer. These parcels are referred to as Case parcels. Figure A.19-A.22 presents the result.
- The average property value (per acre) for all parcels for each station by 5 land use types mentioned before (see Table 3).

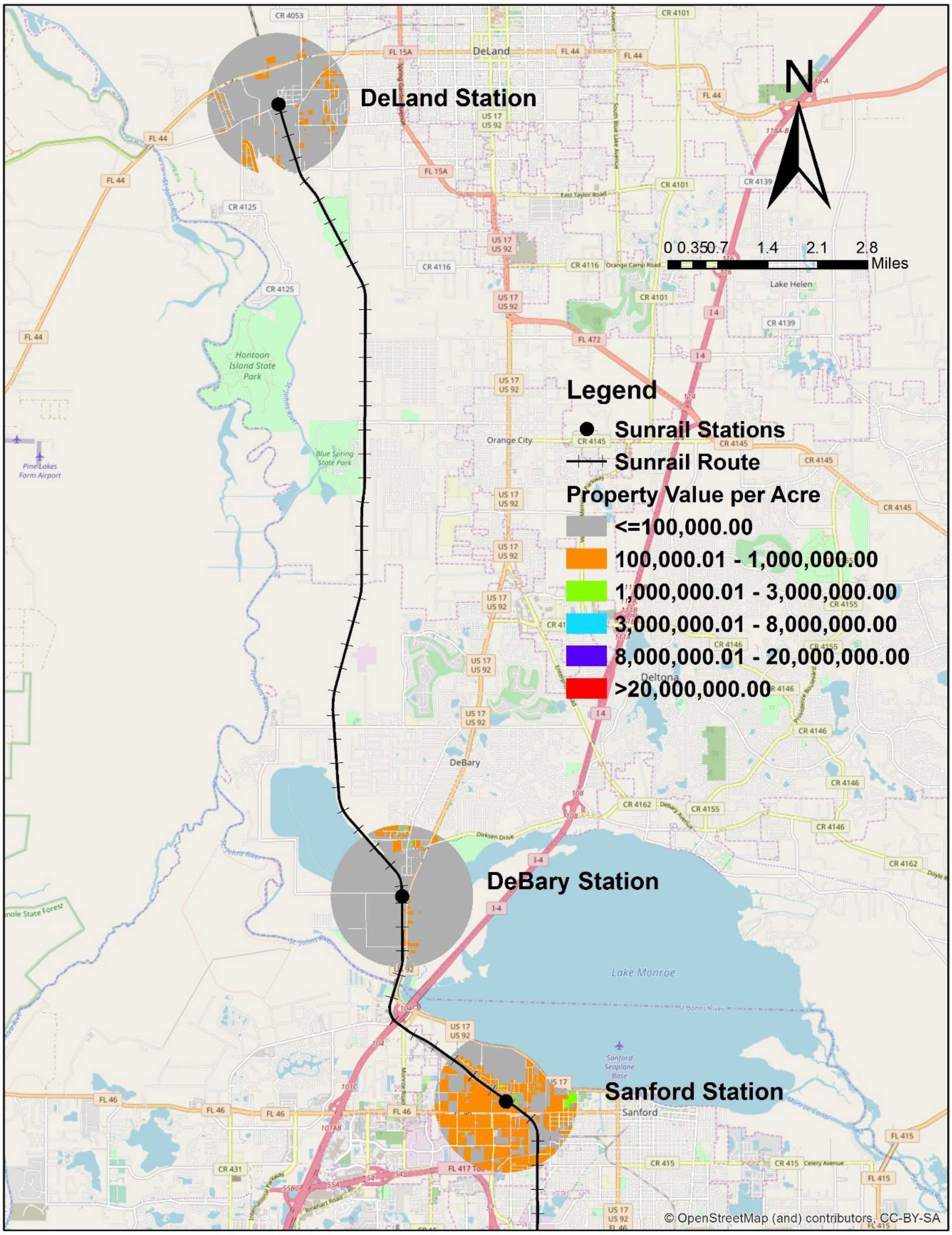


Figure A.23: Average Property Value (DeLand, DeBary and Sanford Station)

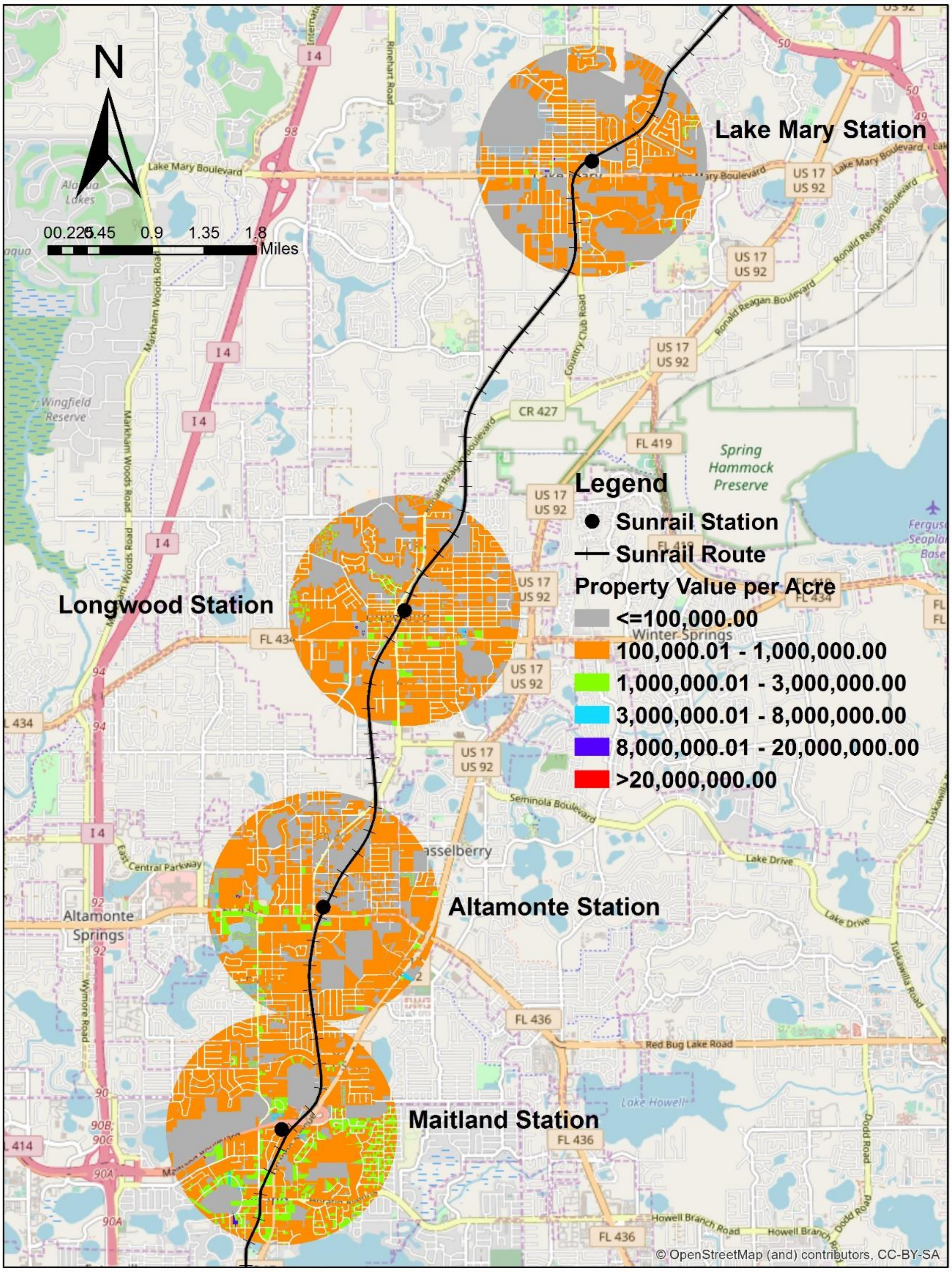


Figure A.24: Average Property Value (Lake Mary, Longwood, Altamonte Springs and Maitland Station)

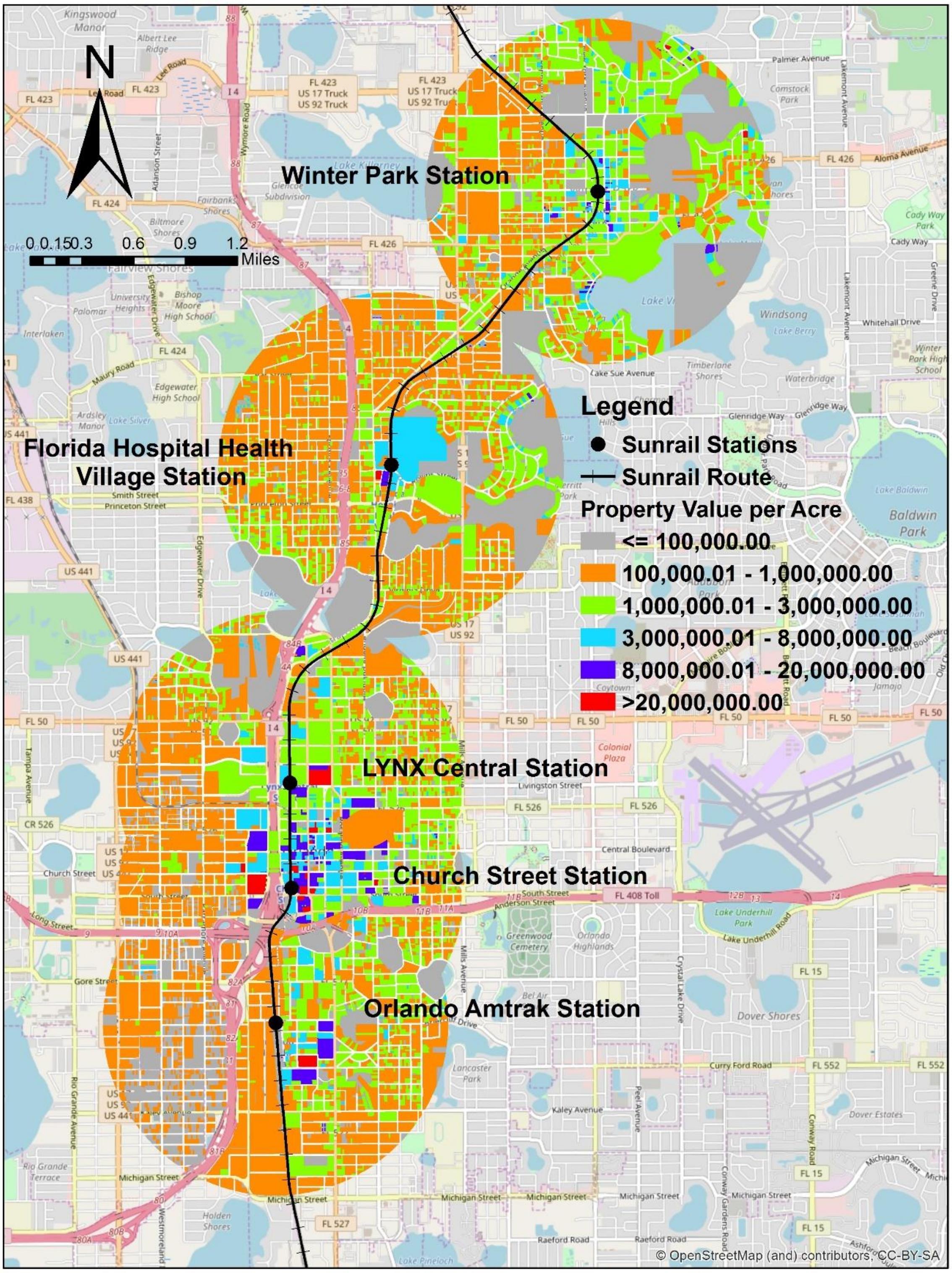


Figure A.25: Average Property Value (Winter Park, Florida Hospital Health Village, LYNX Central, Church Street and Orlando Amtrak Station)

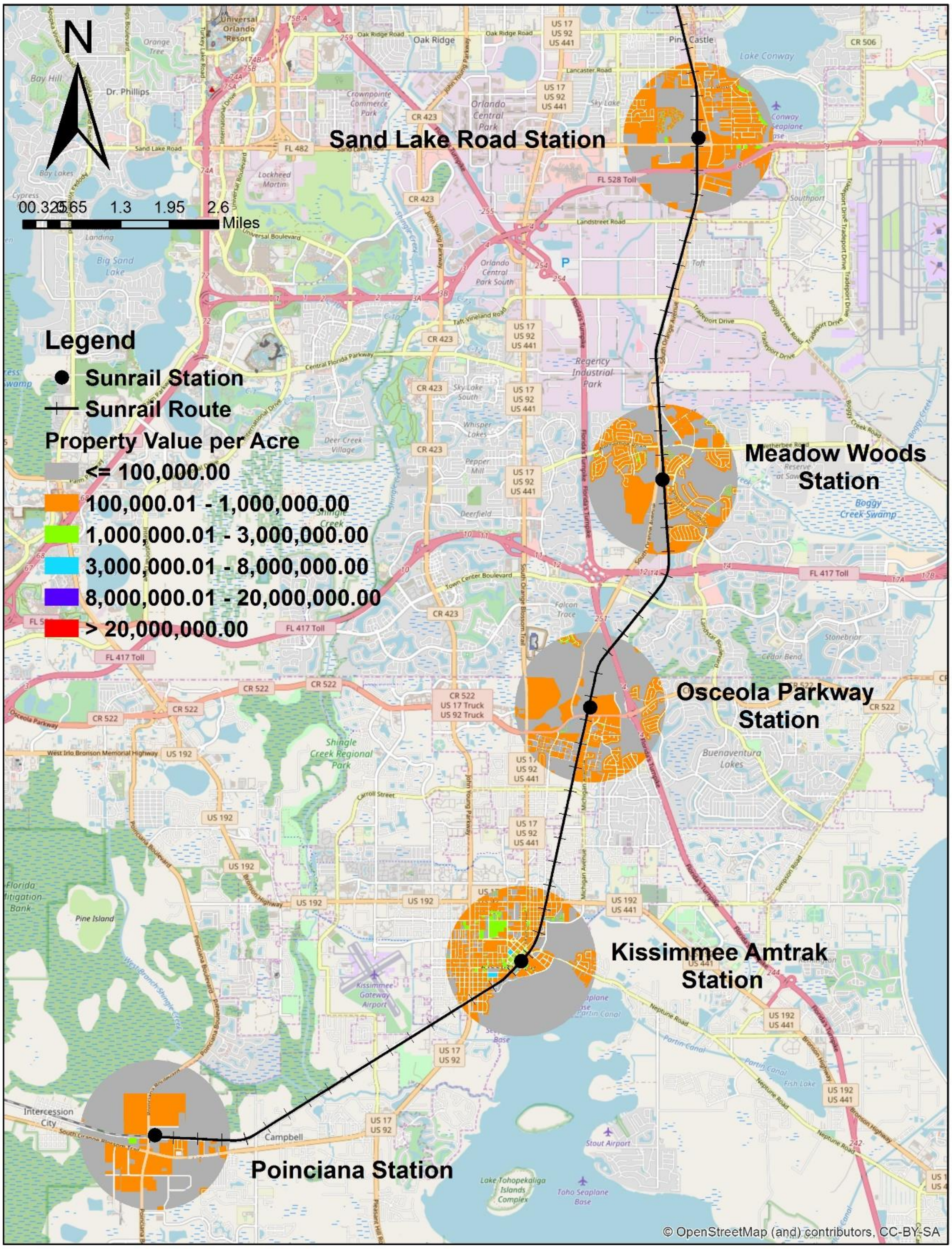


Figure A.26: Average Property Value (Sand Lake Road, Meadow Woods, Osceola Parkway, Kissimmee Amtrak and Poinciana Station)

Table A.43: Average Property Value per Station by Land Use Type for 2012

Station	Single Family Residential (USD)	Multi- Family Residential (USD)	Retail/Office (USD)	Industrial (USD)	Institutional (USD)
Downtown Stations					
LYNX Central Station	906,590	988,491	1,790,503	630,578	1,462,136
Church Street Station	981,280	2,401,727	5,214,377	281,022	4,683,842
Orlando Amtrak/Sligh Blvd Station	625,409	474,380	1,159,111	419,089	1,492,057
Phase-I Outside Downtown Stations					
DeBary Station	49,601	--	136,409	225,568	181,761
Sanford Station	401,223.	570,141	254,061	361,616	400,609
Lake Mary Station	288,673	337,571	673,920	--	81,433
Longwood Station	345,402	344,385	599,405	413,580	564,793
Altamonte Springs Station	295,864	373,609	829,133	429,185	653,548
Maitland Station	632,226	903,955	708,436	430,167	569,418
Winter Park Station	1,393,663	1,353,358	1,601,312	789,060	1,449,902
Florida Hospital Health Village Station	918,072	626,616	1,208,935	724,904	1,083,417
Sand Lake Road Station	456,825	363,302	405,738	256,050	280,571
Phase-II Stations					
DeLand Station	111,661	86,914	56,488	71,328	108,124
Meadow Woods Station	534,753	351,368	75,014	387,552	159,837
Osceola Parkway Station	414,276	245,964	272,880	204,007	161,955
Kissimmee Amtrak Station	255,253	406,806	693,784	317,913	1,034,599
Poinciana	173,863	--	129,603	379,231	175,979

A.2.5 Control Area Selection

While property values in the vicinity of the stations have substantially increased it is not possible to attribute all the increase to SunRail construction without examining the other parts of the urban region. To determine if the changes in property values is truly influenced by SunRail's development, control areas were systematically selected.

We adopted the following procedure for selecting the control areas.

- First, we created 2 and 8 mile buffer, respectively around the stations. The parcels located within that 6 mile buffer were selected to be the candidate control areas.
- Next, based on land use type and property value range (within 15% of the mean property value found for each land use type for case areas), control areas for analysis were identified.

The same number of control parcels were selected for each land use type. Second, the control parcels were assigned to a unique station by using the nearest distance analysis.

- Third, the same procedure as case area is followed to estimate average property price per land use category type for downtown, outside downtown, and Phase-2 stations.

A.2.1.1 1st Step Technique

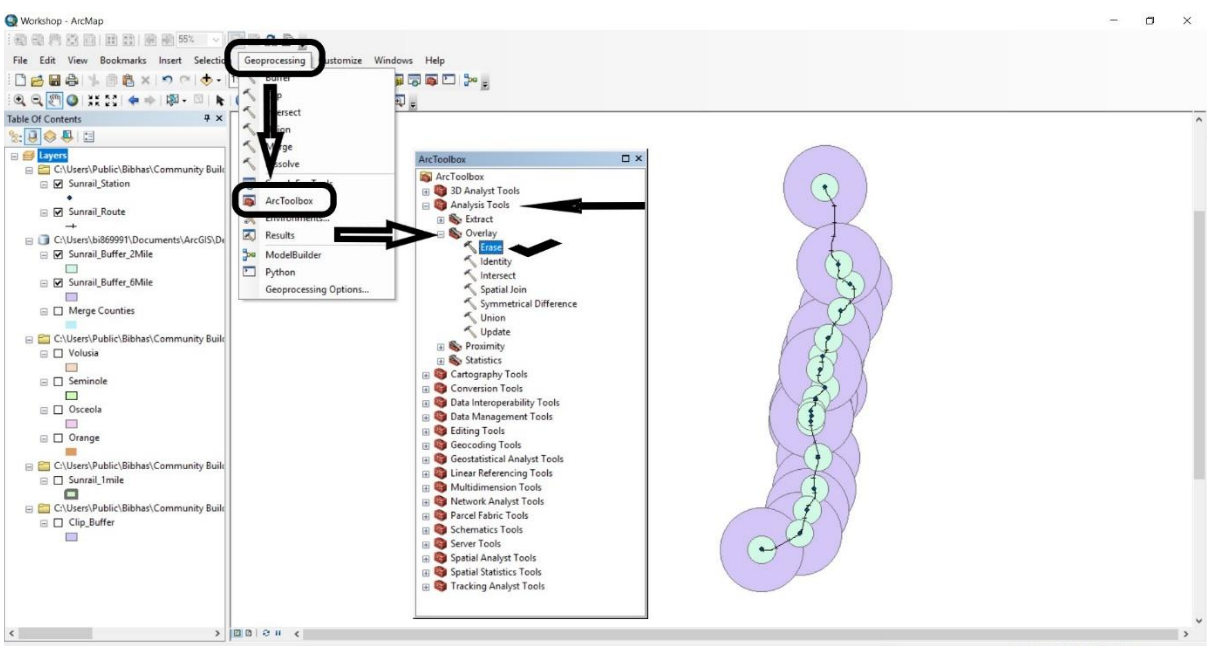
A.2.2.1.1 Draw 2 and 8 mile Buffer

- 2 and 8 mile buffer around each SunRail station was created using similar technique as case area selection

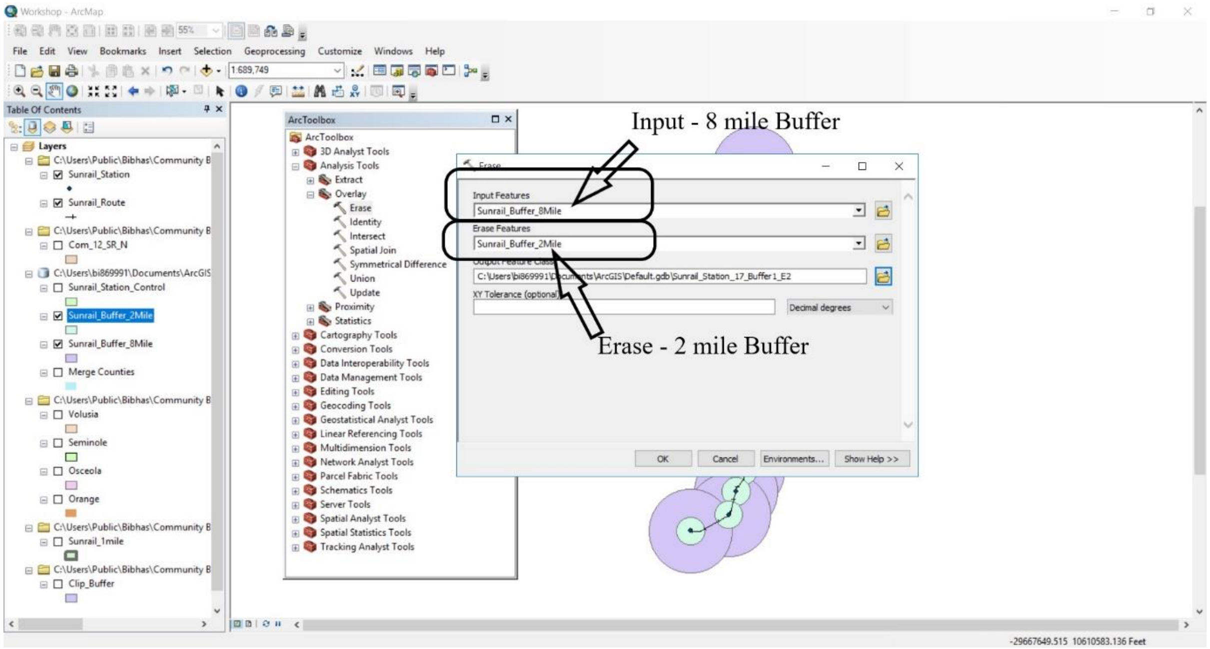
A.2.2.1.2 Erase Inner 2 Miles

- 1st select Geoprocessing toolbar
- Choose Arc Toolbox bar in Geoprocessing (See Figure A.23(a))
- Select 'Analysis Tools' (See Figure A.23(a))
- Select 'Overlay' from 'Analysis Tools'
- Then click to 'Erase' for the overlapping solution
- Select 8 mile buffer layer shapefile in as input features (See Figure A.23(b))
- Then put 2 mile buffer layer shapefile as Erase Features
- Click 'Ok'

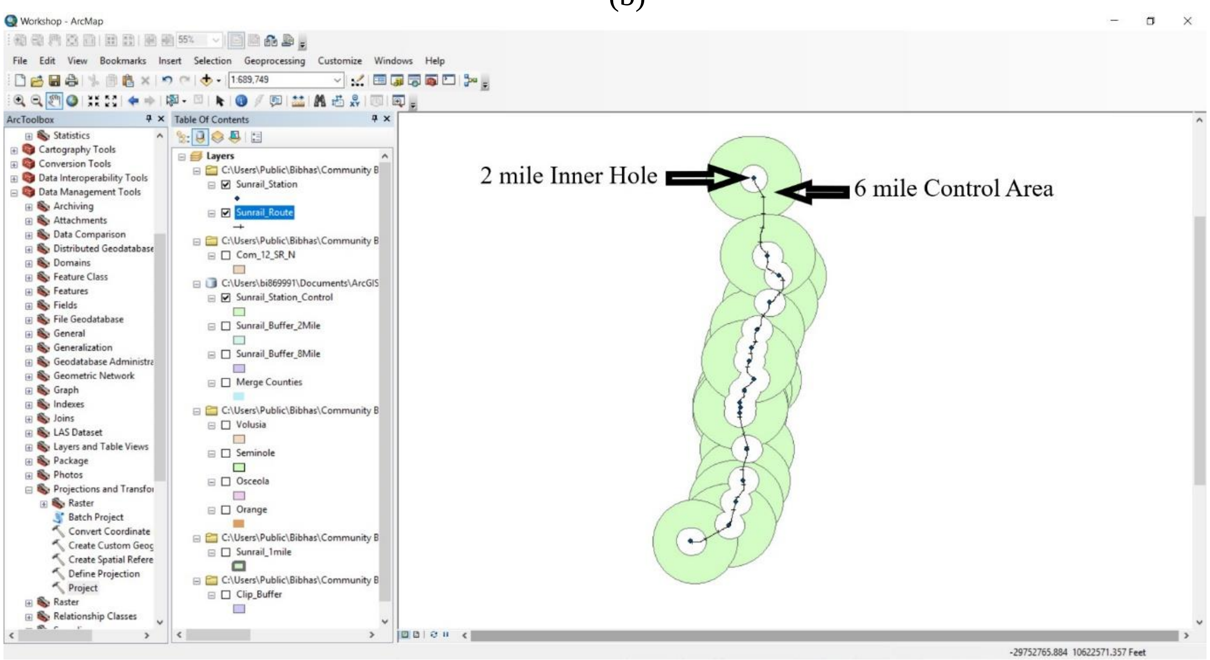
Figure A.23(c) represents the control area where 6 mile area was accounted after 2 mile inner radius from each SunRail station.



(a)



(b)



(c)

Figure A.27: Control Area Selection

A.2 Accessibility to Employment

Job accessibility can be defined as number of jobs accessible from a desirable point. To capture the change in number of jobs around the chosen investment projects, the employment (number of workers in the labor force) data for the years 2011-2016 was drawn from American Community Survey (ACS). This data contains information on total employment of individuals aged 20 through 64 years. These data were merged with the Florida census tract shapefile using the unique ID created by concatenating county and census tract IDs.

A.2.1 Case Area Selection

Job accessibility was computed using jobs accessible within a particular driving distance. Several travel time values are potentially used in literature to identify job. In our study, we used 10 minutes' drive time from our origin of interest as the appropriate threshold. The driving distance was computed using weekday peak period (8am on Tuesday). Street network of Florida has been used to draw driving area for both driving time and driving distance. 2011-2016 street network of 'NAVSTREET' data was used.

A.2.1.1 Driving Network Area

10 minutes driving network area around all SunRail stations was created to select case area for 'Job Accessibility' estimation. This procedure can be divided into two parts.

A.2.1.1.1 Road Network

- At first, a street network must need to be created to draw a driving area around SunRail stations.
- 'NAVSTREET' street network shapefile was used to create street network.
- Please note that, to estimate driving time, we need speed limit of the corresponding street. We define a fixed speed for a street from variable called 'Speed Category'. Conversion of speed from defined speed limit range is shown in Table A.4.

Table A.44: Speed Definition

Speed Category	Definition (MPH)	Speed, V (MPH)
1	Above 80	80
2	65-80	70
3	55-64	60
4	41-54	50
5	31-40	40
6	21-30	30
7	6-20	20
8	Below 6	6

- Since driving area was estimated based on time, so travel time need to be calculated on street network file.

- Travel time (in minutes) needed to travel the corresponding street was estimated by using equation, T = (L/V) *60 where T is travel time needed to travel the total length of street in minutes, L is total length in miles and V is speed in mph (as mentioned Table 4).
- Three new variables as Speed, length and minutes need to be created by using similar to A.2.1.2.1.

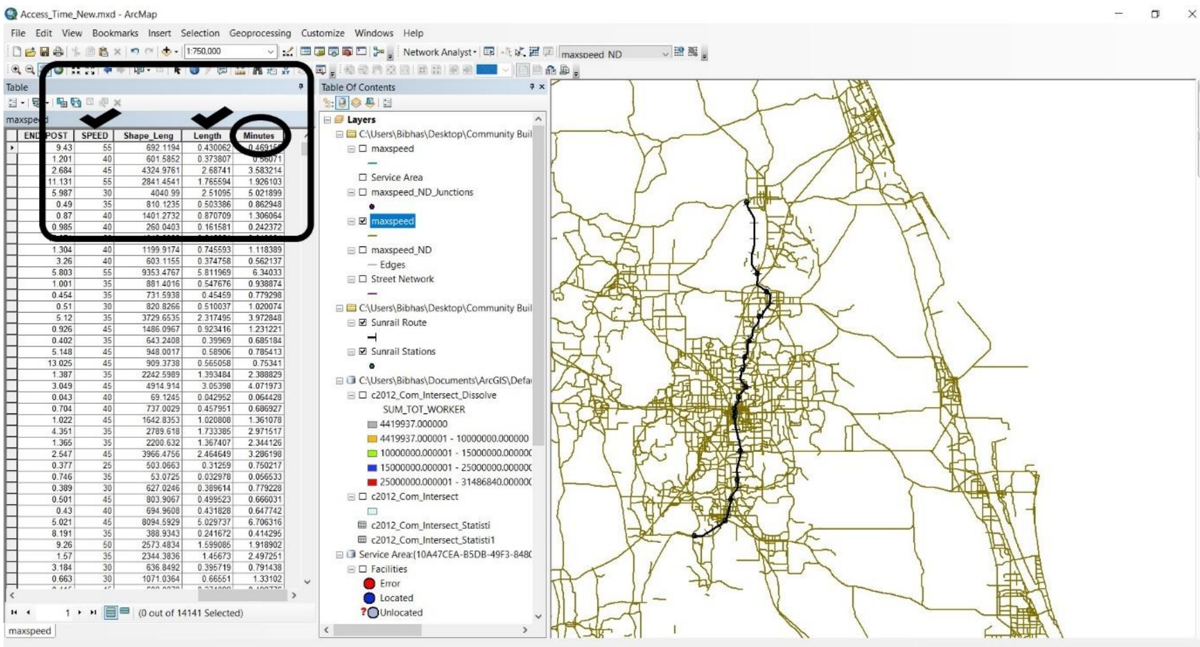
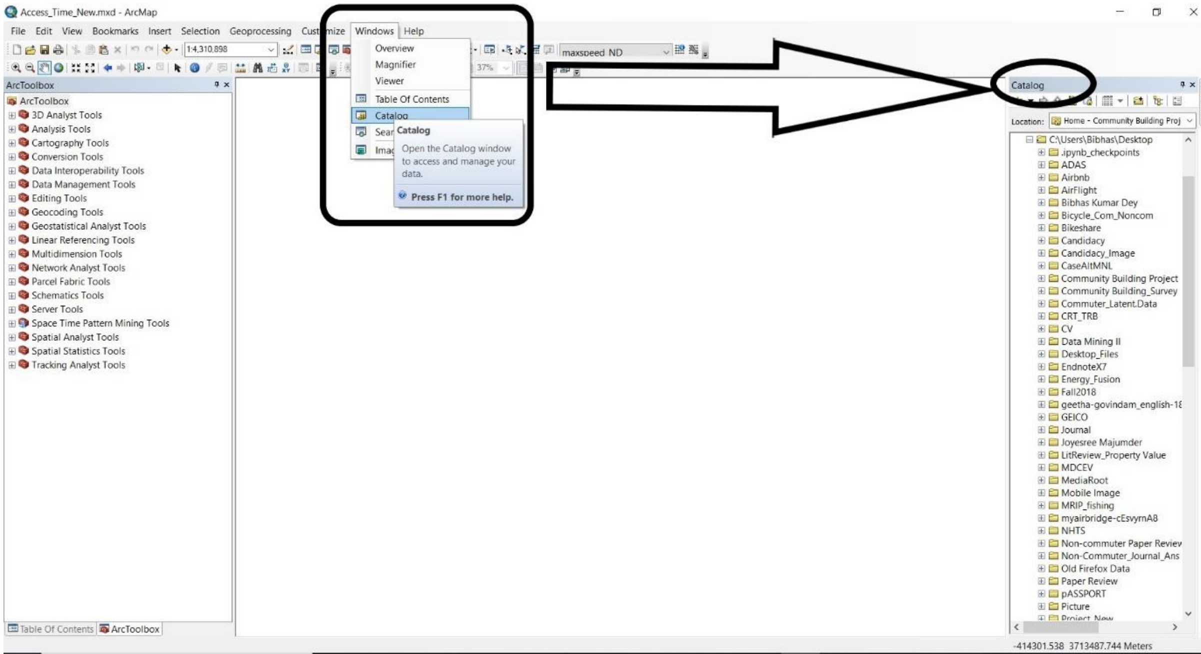


Figure A.28: Travel Time Added

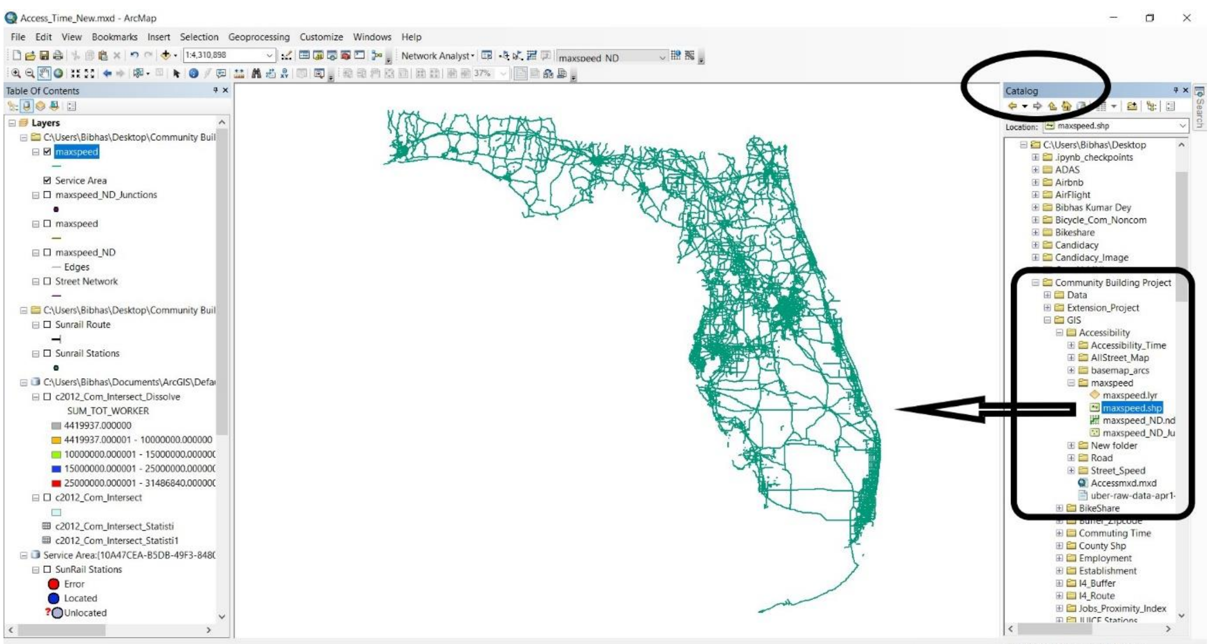
Several steps were followed to create 'Road Network' by using Network Analyst tool on ArcGIS (See Figure 25 (a) – 25 (m)).[2]

- Select 'Catalog' from 'Windows' toolbar (See Figure A.25 (a))
- Select Road Network shapefile by using 'Catalog' (See Figure A.25 (b))
- Click on the 'New Network Dataset' from 'Road Network' file
- Follow all the steps shown in Figure A.25 (d) to Figure A.25 (l)
- All the above steps will create a new road network with 'junction' and 'edges'

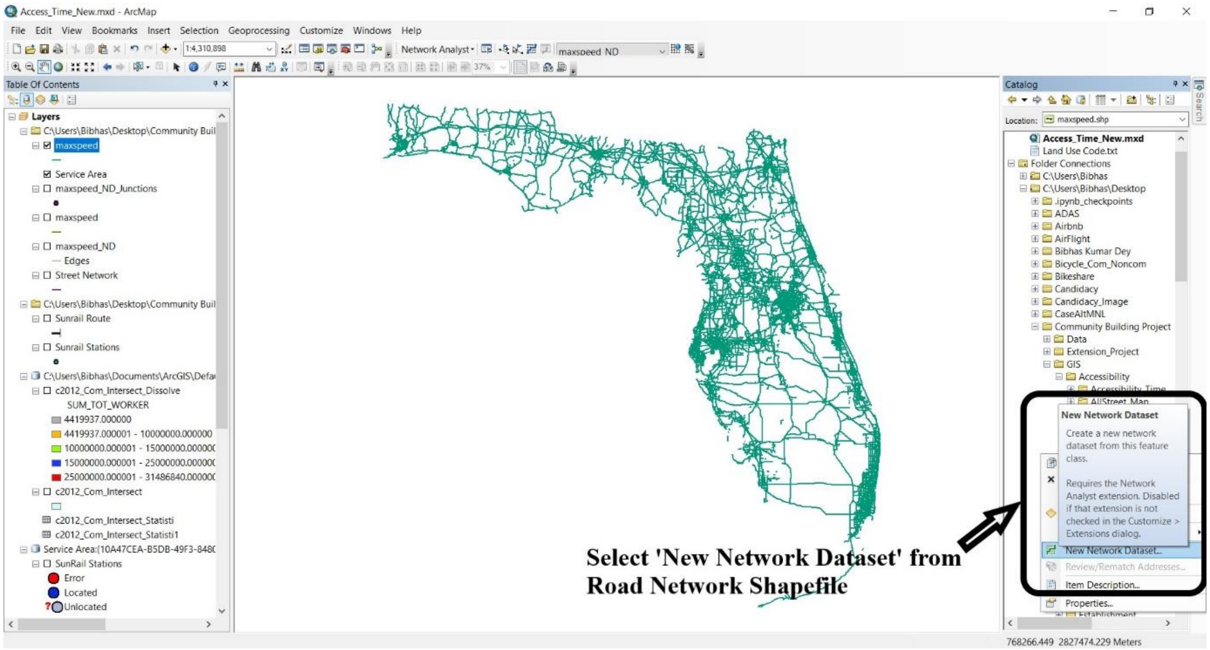
[2] How to create a road network in ArcGIS can be found on this YouTube link (https://www.youtube.com/watch?v=IcETd6oHZtQ)



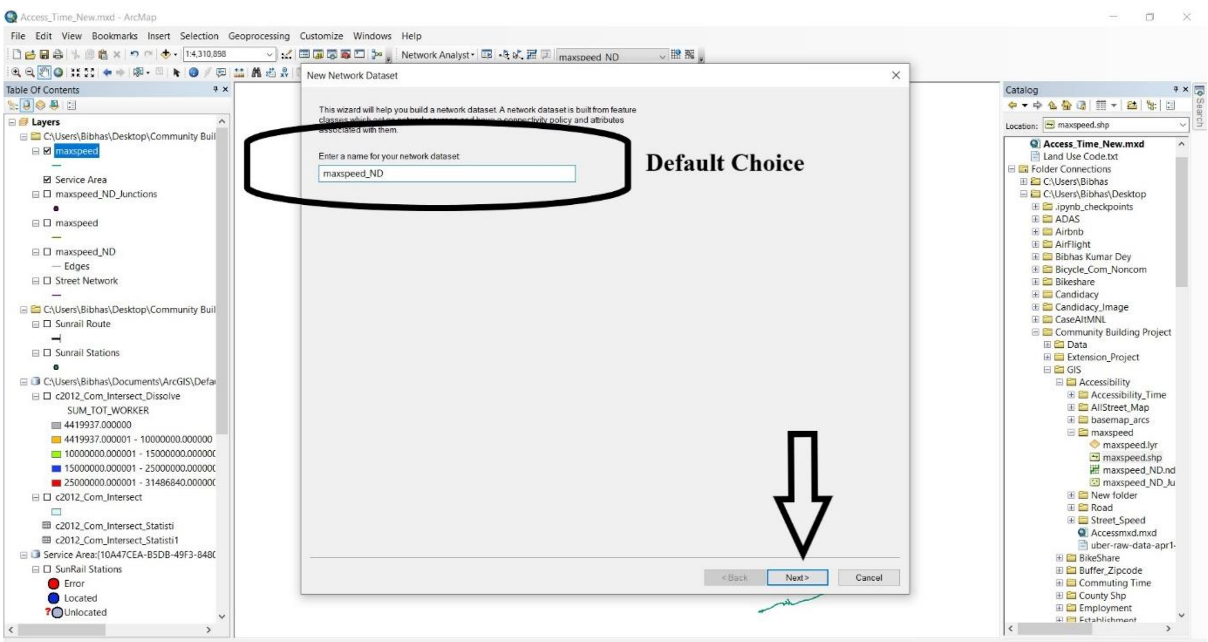
(a)



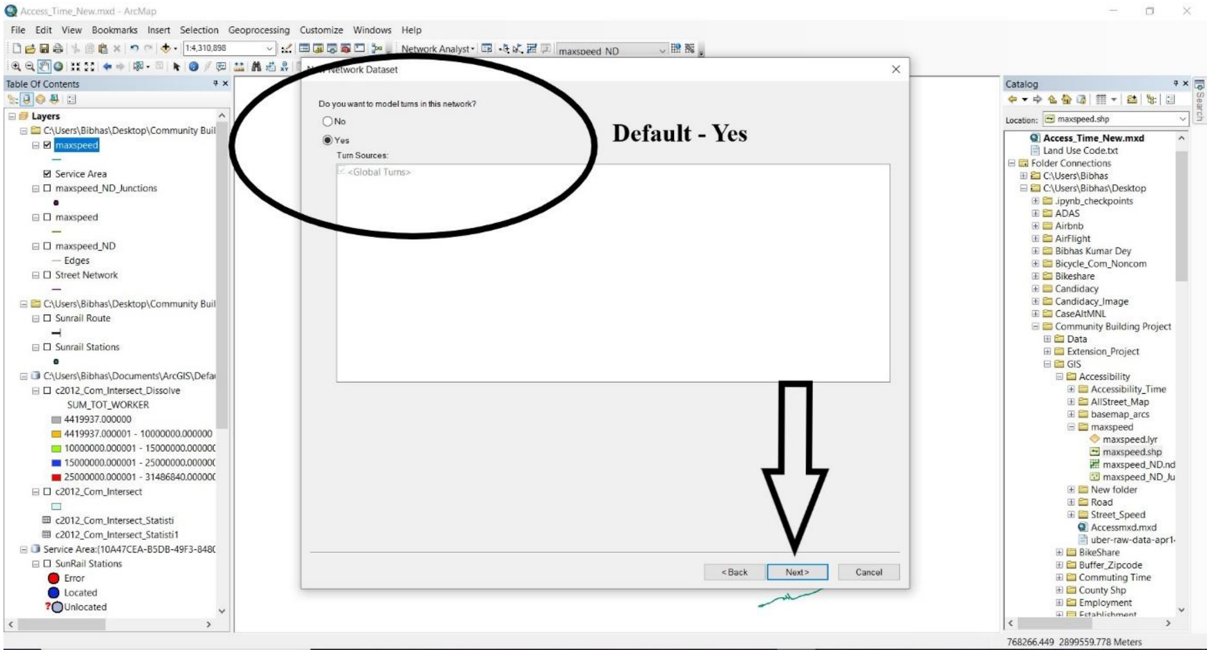
(b)


Select 'New Network Dataset' from
Road Network Shapefile
New Network Dataset
Create a new network dataset from this feature class.
Requires the Network Analyst extension. Disabled if that extension is not checked in the Customize > Extensions dialog.

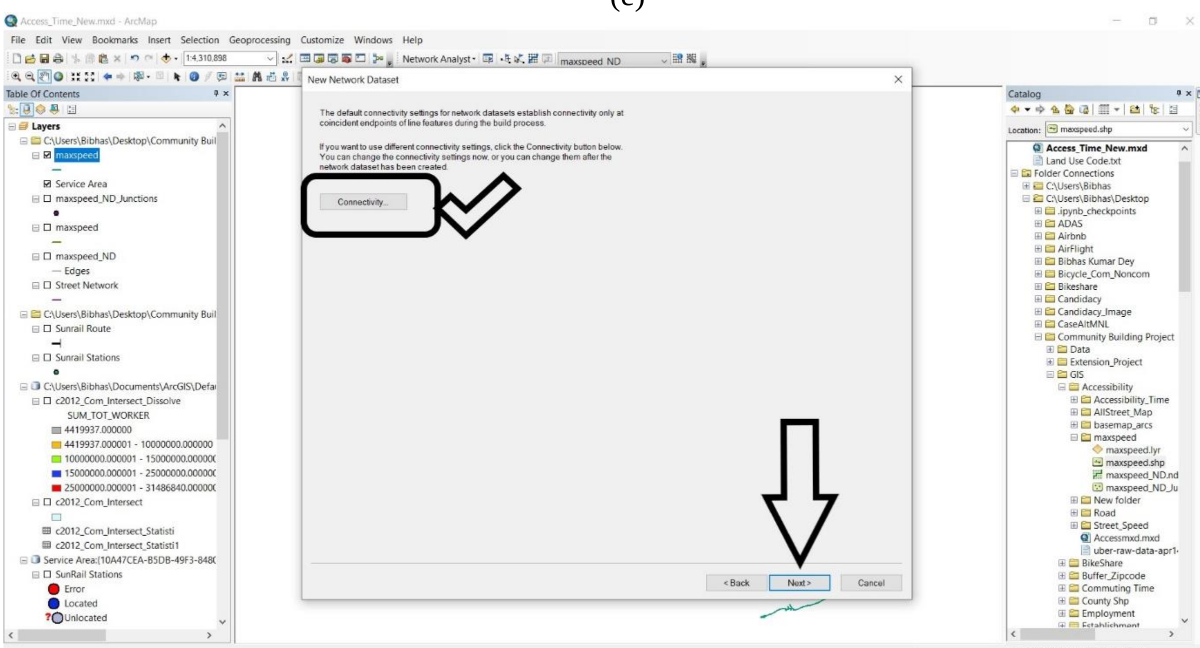
(c)


New Network Dataset
This wizard will help you build a network dataset. A network dataset is built from feature classes which act as network sources and have a connectivity policy and attributes associated with them.
Enter a name for your network dataset
maxspeed_ND
Default Choice
< Back
Next >
Cancel

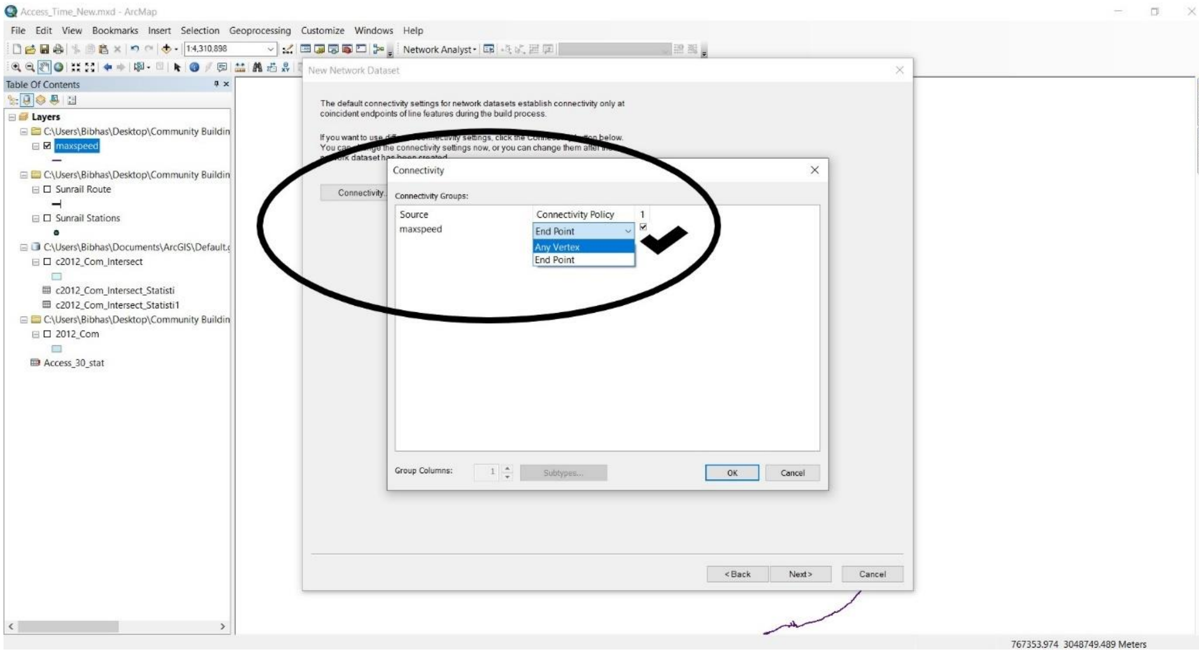
(d)


Access_Time_New.mxd - ArcMap
File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help
New Network Dataset
Do you want to model turns in this network?
No
Yes
Turn Sources:
<Global Turns>
Default - Yes
< Back
Next >
Cancel
Table Of Contents
Catalog

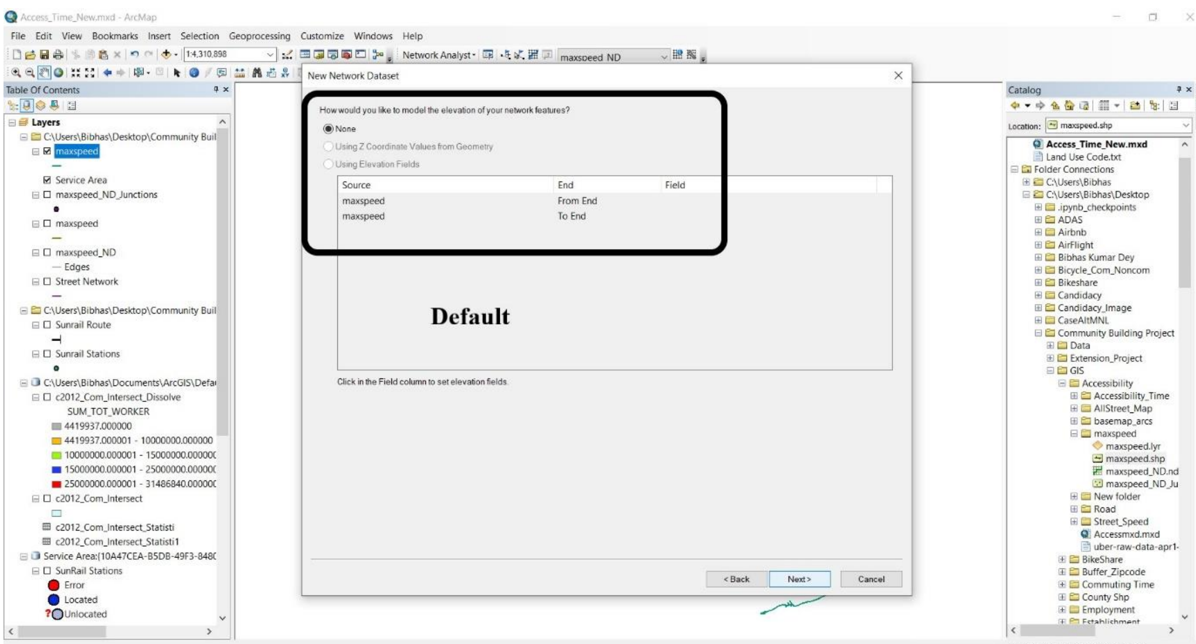
(e)


Access_Time_New.mxd - ArcMap
File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help
New Network Dataset
The default connectivity settings for network datasets establish connectivity only at coincident endpoints of line features during the build process.
If you want to use different connectivity settings, click the Connectivity button below. You can change the connectivity settings now, or you can change them after the network dataset has been created.
Connectivity...
< Back
Next >
Cancel
Table Of Contents
Catalog

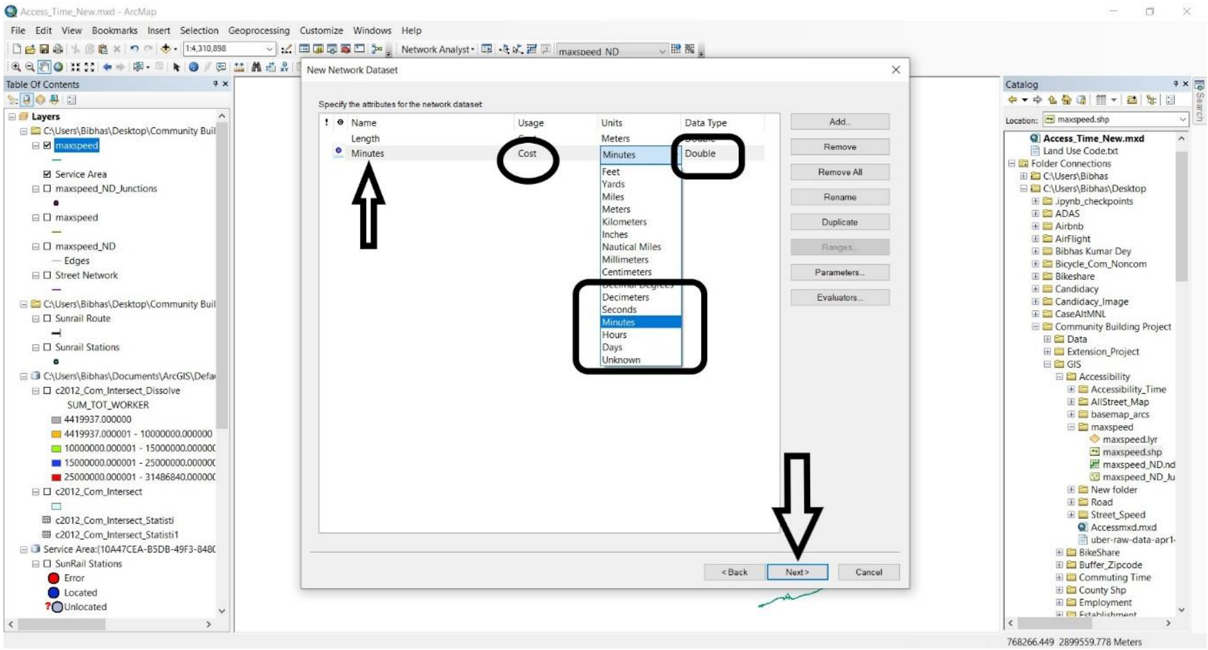
(f)


Access_Time_New.mxd - ArcMap
File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help
Network Analyst
Table Of Contents
Layers
maxspeed
Sunrail Route
Sunrail Stations
c2012_Com_Intersect
c2012_Com_Intersect_Statisti
c2012_Com_Intersect_Statisti1
2012_Com
Access_30_stat
New Network Dataset
The default connectivity settings for network datasets establish connectivity only at coincident endpoints of line features during the build process.
Connectivity
Connectivity Groups:
Source
Connectivity Policy
maxspeed
End Point
Any Vertex
End Point
Group Columns:
Subtypes...
OK
Cancel
< Back
Next >
Cancel
767353.974 3048749.489 Meters

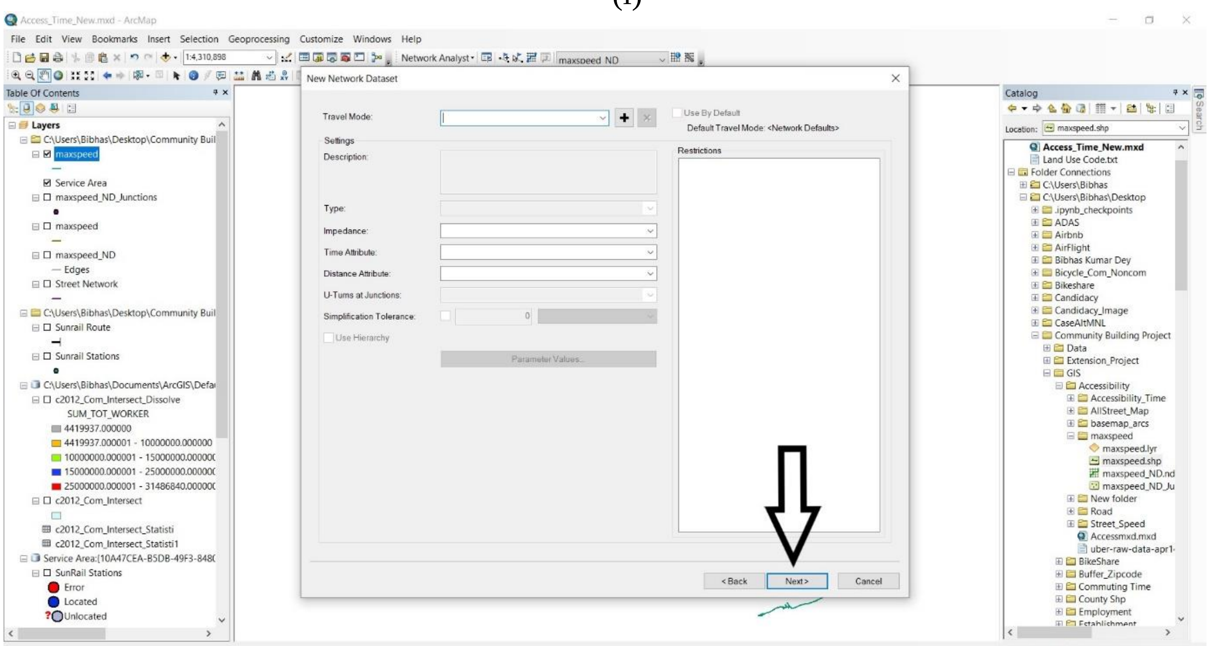
(g)


Access_Time_New.mxd - ArcMap
File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help
Network Analyst
maxspeed_ND
Table Of Contents
Layers
maxspeed
Service Area
maxspeed_ND_Junctions
maxspeed_ND
Edges
Street Network
Sunrail Route
Sunrail Stations
c2012_Com_Intersect_Dissolve
SUM_TOT_WORKER
c2012_Com_Intersect
SunRail Stations
Error
Located
Unlocated
New Network Dataset
How would you like to model the elevation of your network features?
None
Using Z Coordinate Values from Geometry
Using Elevation Fields
Source
End
Field
maxspeed
From End
maxspeed
To End
Default
Click in the Field column to set elevation fields.
< Back
Next >
Cancel
Catalog
Location: maxspeed.shp
Access_Time_New.mxd
Land Use Code.txt
Folder Connections
Community Building Project
GIS
Accessibility
maxspeed
maxspeed.lyr
maxspeed.shp
maxspeed_ND.nd
768266.449 2899559.778 Meters

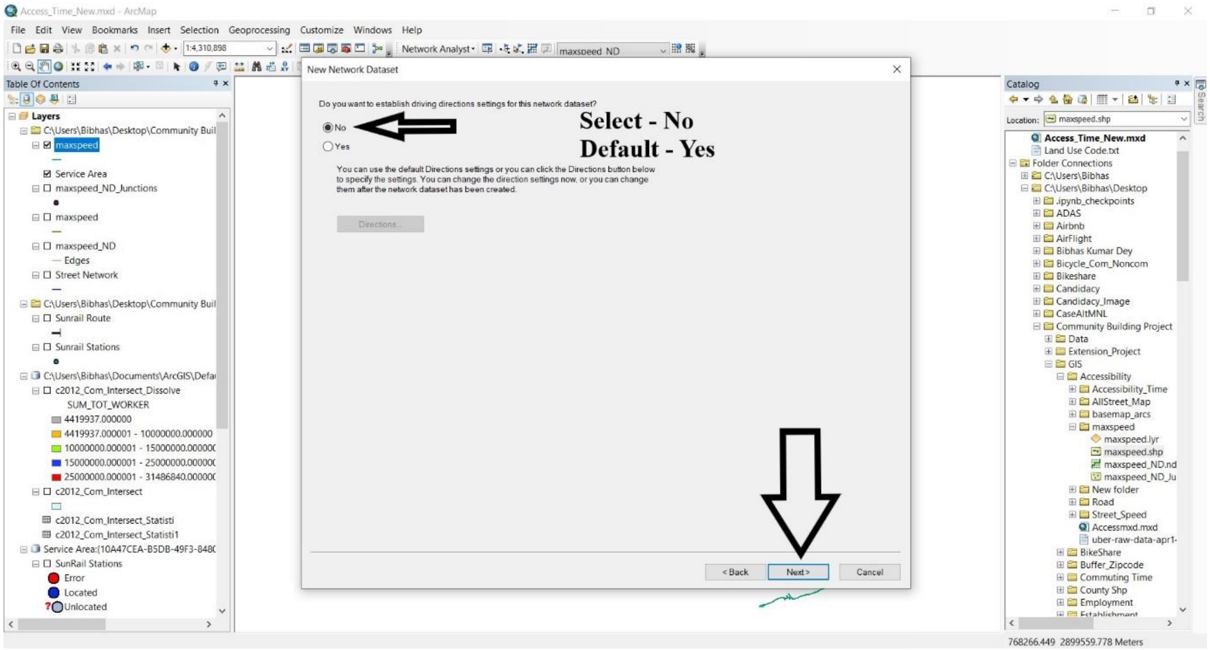
(h)



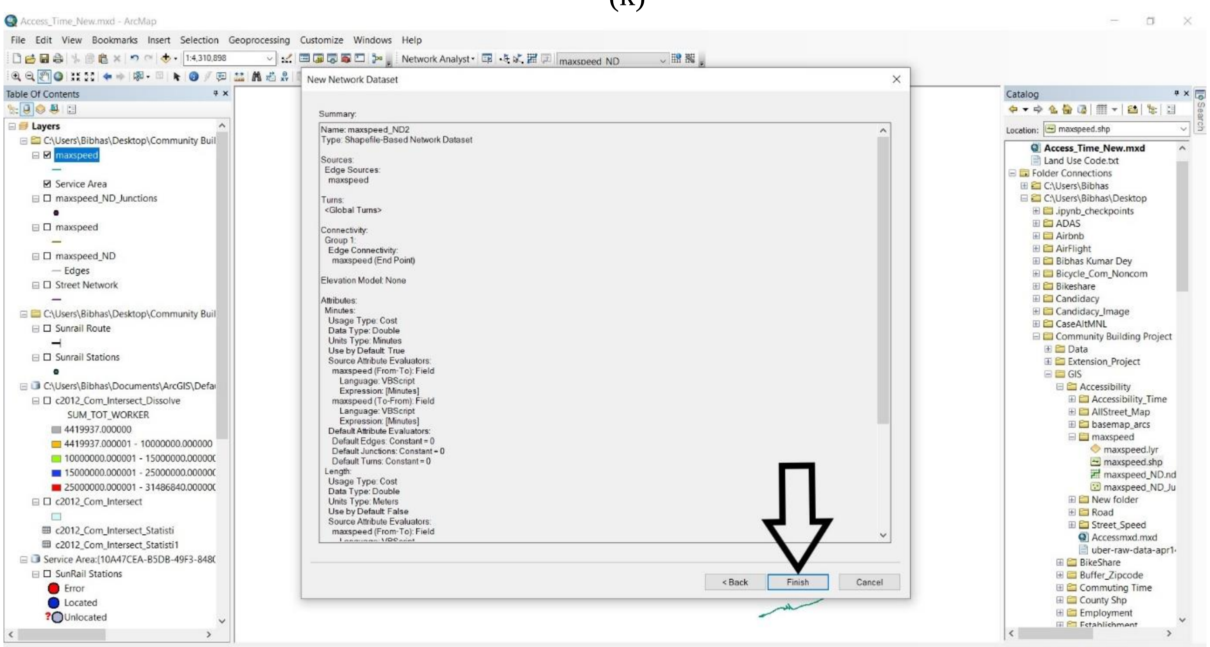
(i)



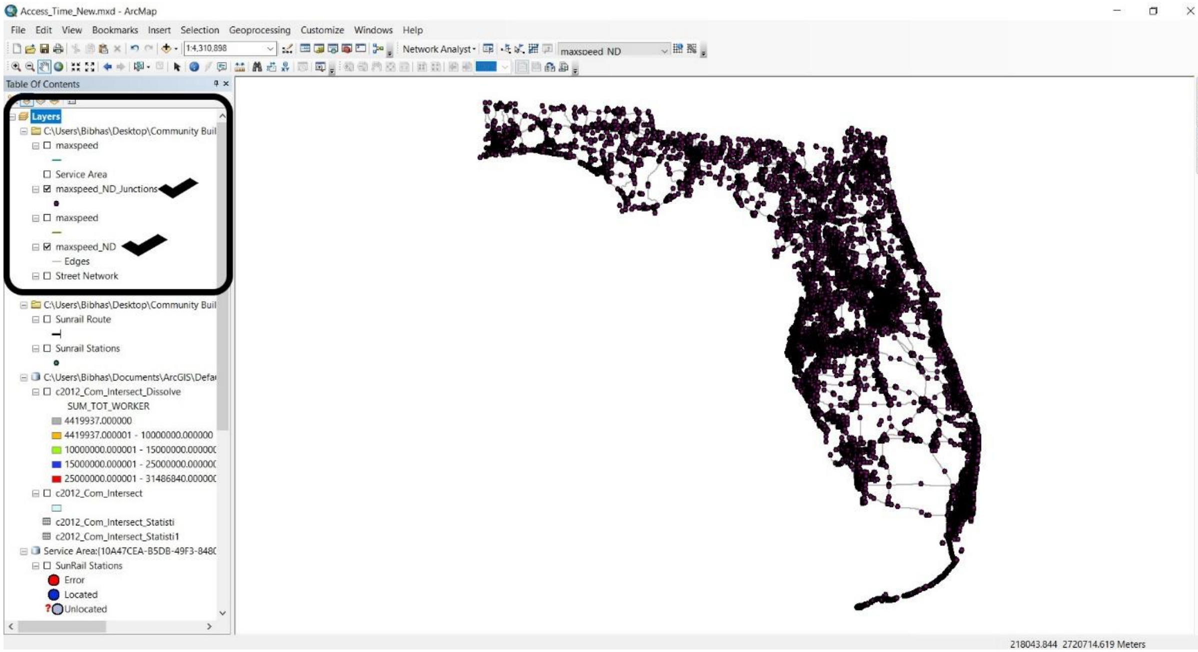
(j)


New Network Dataset
Do you want to establish driving directions settings for this network dataset?
No
Yes
Select - No
Default - Yes
You can use the default Directions settings or you can click the Directions button below to specify the settings. You can change the direction settings now, or you can change them after the network dataset has been created.
Directions...
< Back
Next >
Cancel

(k)


New Network Dataset
Summary:
Name: maxspeed_ND2
Type: Shapefile-Based Network Dataset
Sources:
Edge Sources:
maxspeed
Turns:
<Global Turns>
Connectivity:
Group 1:
Edge Connectivity:
maxspeed (End Point)
Elevation Model: None
Attributes:
Minutes:
Usage Type: Cost
Data Type: Double
Units Type: Minutes
Use by Default: True
Source Attribute Evaluators:
maxspeed (From-To): Field
Language: VBScript
Expression: [Minutes]
maxspeed (To-From): Field
Language: VBScript
Expression: [Minutes]
Default Attribute Evaluators:
Default Edges: Constant = 0
Default Junctions: Constant = 0
Default Turns: Constant = 0
Length:
Usage Type: Cost
Data Type: Double
Units Type: Meters
Use by Default: False
Source Attribute Evaluators:
maxspeed (From-To): Field
< Back
Finish
Cancel

(l)



(m)

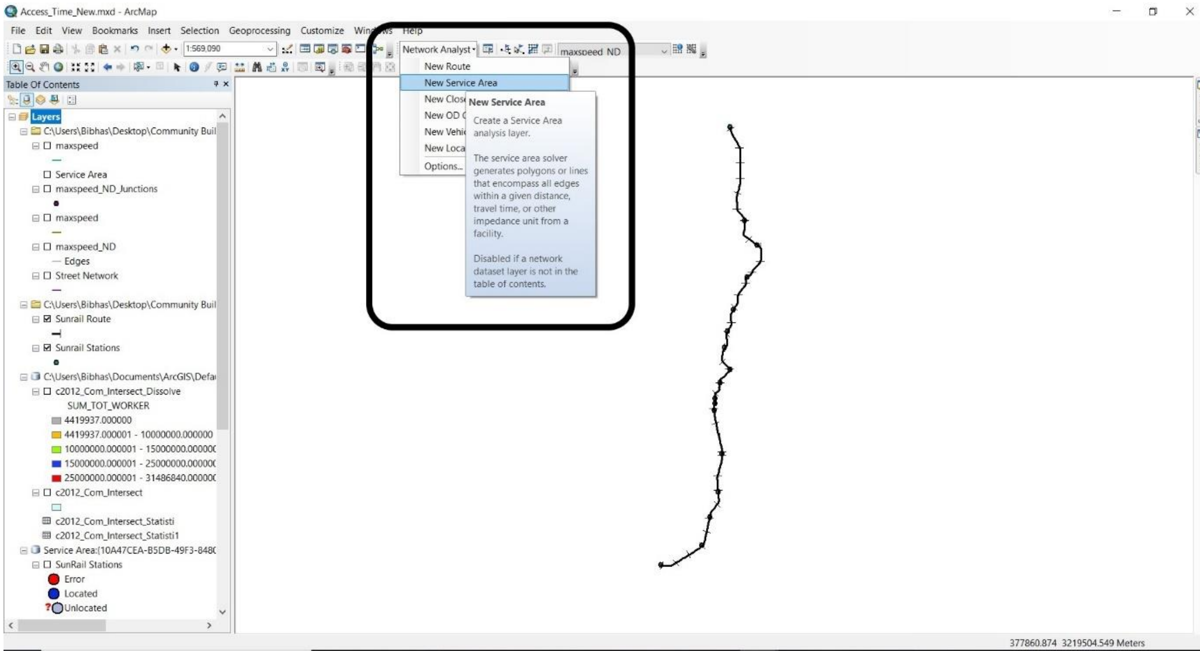
Figure A.29: Road Network Create in ArcGIS

A.2.1.1.2 Driving Area

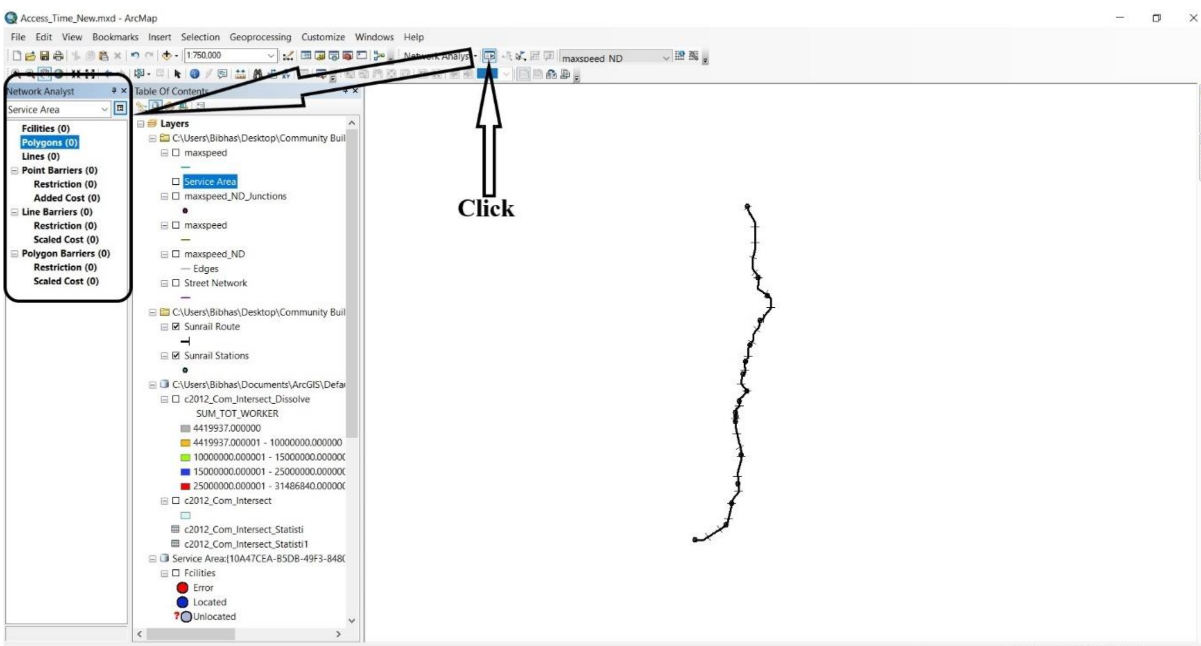
Several steps were followed to create a 10 minutes driving area by using road network in ArcGIS[3].

- Select 'Network Analysts'
- Click on 'New Service Area' (See Figure A.26 (a))
- Click Network Analyst Window (See Figure A.26 (b))
- Within Network Analyst toolbar, select 'Facilities'
- Click 'Load Locations' in Facilities (See Figure A.26 (c))
- Put 'SunRail Stations' in Load Locations (See Figure A.26 (d))
- Click 'Service Area' in Layers and select 'Properties' (See Figure A.26 (e))
- Click on 'Analysis Settings' in Properties menu bar (See Figure A.26 (f))
- In Impedance option, select 'Travel Time (Minutes)'
- In 'Default Breaks' option, put 10 for creating10 minutes driving area
- For the time of the day 8 am was selected
- Tuesday was selected as 'Day of Week'
- Click on 'Solve' as shown in Figure A.26 (g) and new 10 minute driving area was created
- For further estimation procedure, 10 minute driving area data need to be export (See Figure 26 (h))
- Figure A.27 represents the 10 minutes case area around all SunRail stations

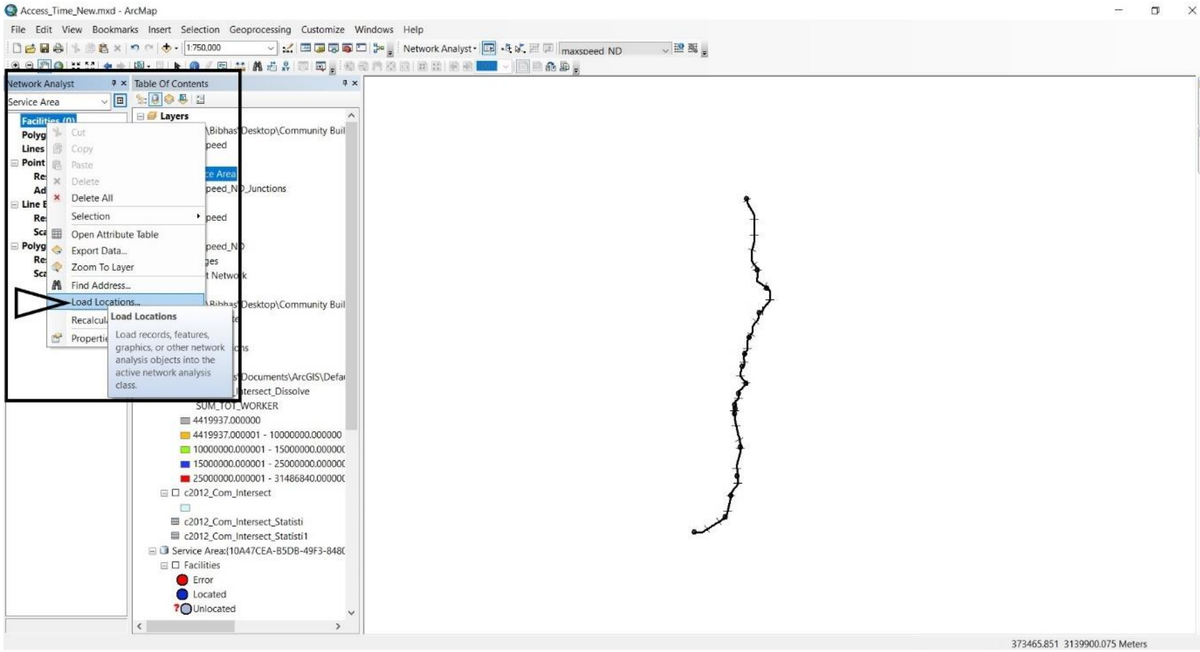
[3] Create drive time areas in ArcGIS (https://doc.arcgis.com/en/arcgis-online/analyze/create-drive-time-areas.htm)


Access_Time_New.mxd - ArcMap
File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help
Network Analyst
New Route
New Service Area
New Service Area
Create a Service Area analysis layer.
The service area solver generates polygons or lines that encompass all edges within a given distance, travel time, or other impedance unit from a facility.
Disabled if a network dataset layer is not in the table of contents.
Options...
Table Of Contents
Layers
maxspeed
Service Area
maxspeed_ND_Junctions
maxspeed_ND
Edges
Street Network
Sunrail Route
Sunrail Stations
c2012_Com_Intersect_Dissolve
SUM_TOT_WORKER
c2012_Com_Intersect
SunRail Stations
Error
Located
Unlocated
377860.874 3219504.549 Meters

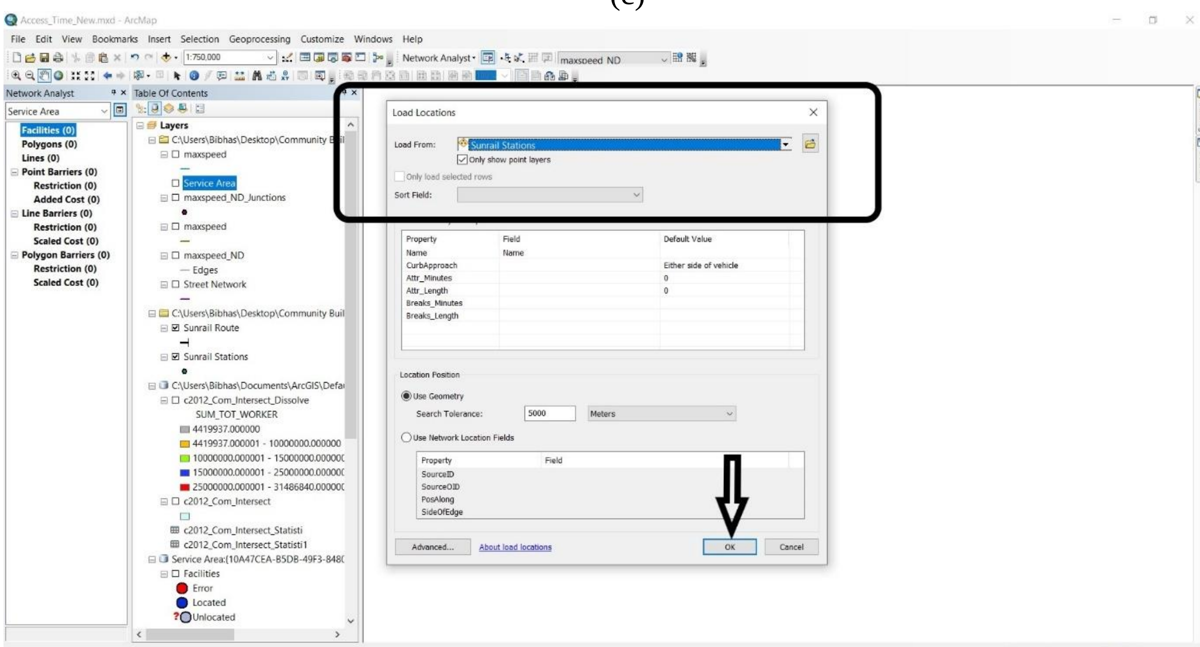
(a)


Access_Time_New.mxd - ArcMap
File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help
Network Analyst
Service Area
Fcilities (0)
Polygons (0)
Lines (0)
Point Barriers (0)
Restriction (0)
Added Cost (0)
Line Barriers (0)
Restriction (0)
Scaled Cost (0)
Polygon Barriers (0)
Restriction (0)
Scaled Cost (0)
Click
Table Of Contents
Layers
maxspeed
Service Area
maxspeed_ND_Junctions
maxspeed_ND
Edges
Street Network
Sunrail Route
Sunrail Stations
c2012_Com_Intersect_Dissolve
SUM_TOT_WORKER
c2012_Com_Intersect
Fcilities
Error
Located
Unlocated
370925.846 3156410.108 Meters

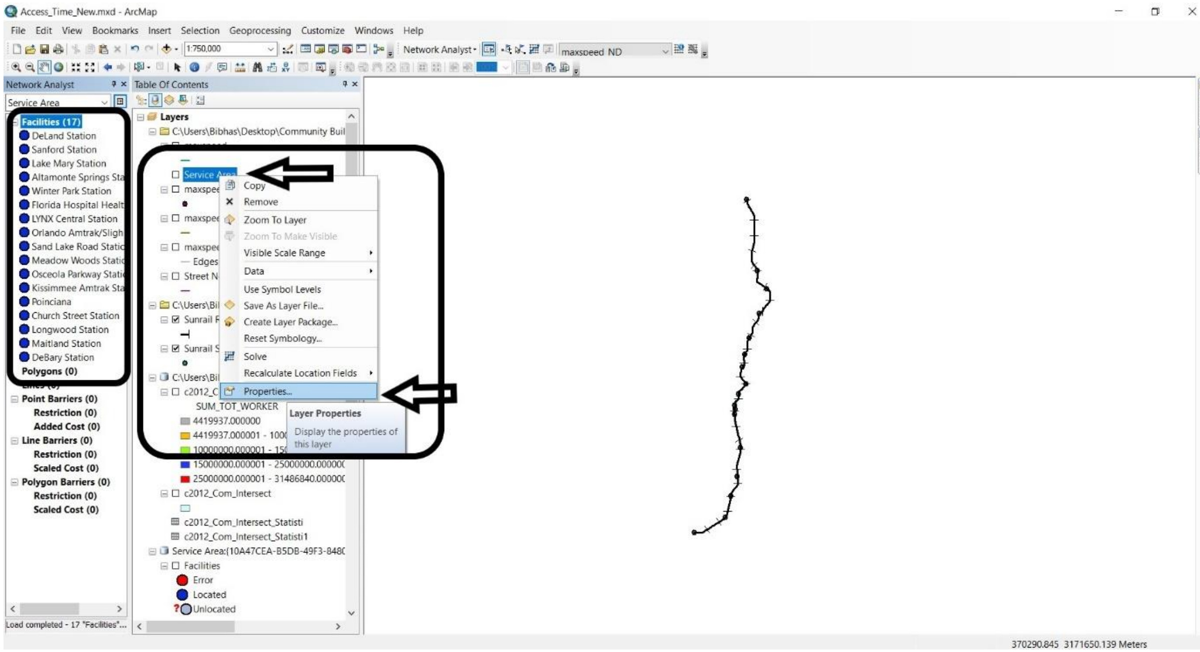
(b)


Access_Time_New.mxd - ArcMap
File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help
Network Analyst
maxspeed_ND
Table Of Contents
Service Area
Layers
Cut
Copy
Paste
Delete
Delete All
Selection
Open Attribute Table
Export Data...
Zoom To Layer
Find Address...
Load Locations...
Load Locations
Load records, features, graphics, or other network analysis objects into the active network analysis class.
SUM_TOT_WORKER
4419937.000000
4419937.000001 - 10000000.000000
10000000.000001 - 15000000.000000
15000000.000001 - 25000000.000000
25000000.000001 - 31486840.000000
c2012_Com_Intersect
c2012_Com_Intersect_Statisti
c2012_Com_Intersect_Statisti1
Service Area:{10A47CEA-B5DB-49F3-848
Facilities
Error
Located
Unlocated
373465.851 3139900.075 Meters

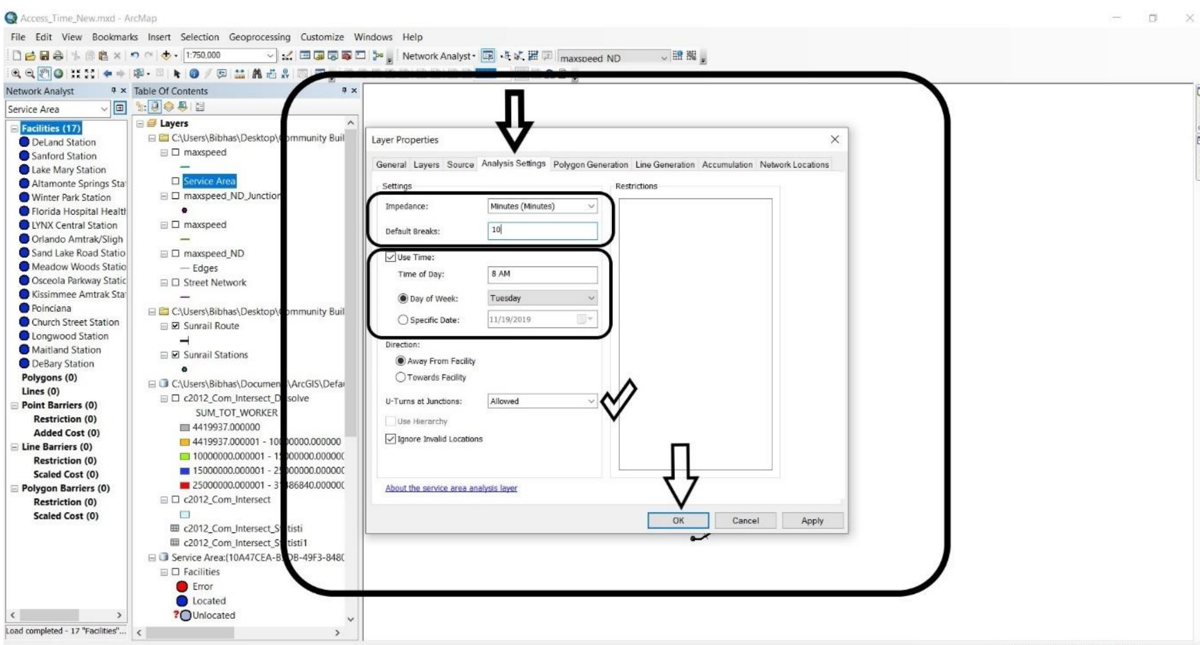
(c)


Access_Time_New.mxd - ArcMap
File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help
Network Analyst
Service Area
Facilities (0)
Polygons (0)
Lines (0)
Point Barriers (0)
Restriction (0)
Added Cost (0)
Line Barriers (0)
Restriction (0)
Scaled Cost (0)
Polygon Barriers (0)
Restriction (0)
Scaled Cost (0)
Table Of Contents
Layers
C:\Users\Bibhas\Desktop\Community B
maxspeed
Service Area
maxspeed_ND_Junctions
maxspeed
maxspeed_ND
Edges
Street Network
C:\Users\Bibhas\Desktop\Community Buil
Sunrail Route
Sunrail Stations
C:\Users\Bibhas\Documents\ArcGIS\Defa
c2012_Com_Intersect_Dissolve
SUM_TOT_WORKER
4419937.000000
4419937.000001 - 10000000.000000
10000000.000001 - 15000000.000000
15000000.000001 - 25000000.000000
25000000.000001 - 31486840.000000
c2012_Com_Intersect
c2012_Com_Intersect_Statisti
c2012_Com_Intersect_Statisti1
Service Area:{10A47CEA-B5DB-49F3-848
Facilities
Error
Located
Unlocated
Load Locations
Load From:
Sunrail Stations
Only show point layers
Only load selected rows
Sort Field:
Property
Field
Default Value
Name
Name
CurbApproach
Either side of vehicle
Attr_Minutes
0
Attr_Length
0
Breaks_Minutes
Breaks_Length
Location Position
Use Geometry
Search Tolerance:
5000
Meters
Use Network Location Fields
Property
Field
SourceID
SourceOID
PosAlong
SideOfEdge
Advanced...
About load locations
OK
Cancel
373465.851 3139900.075 Meters

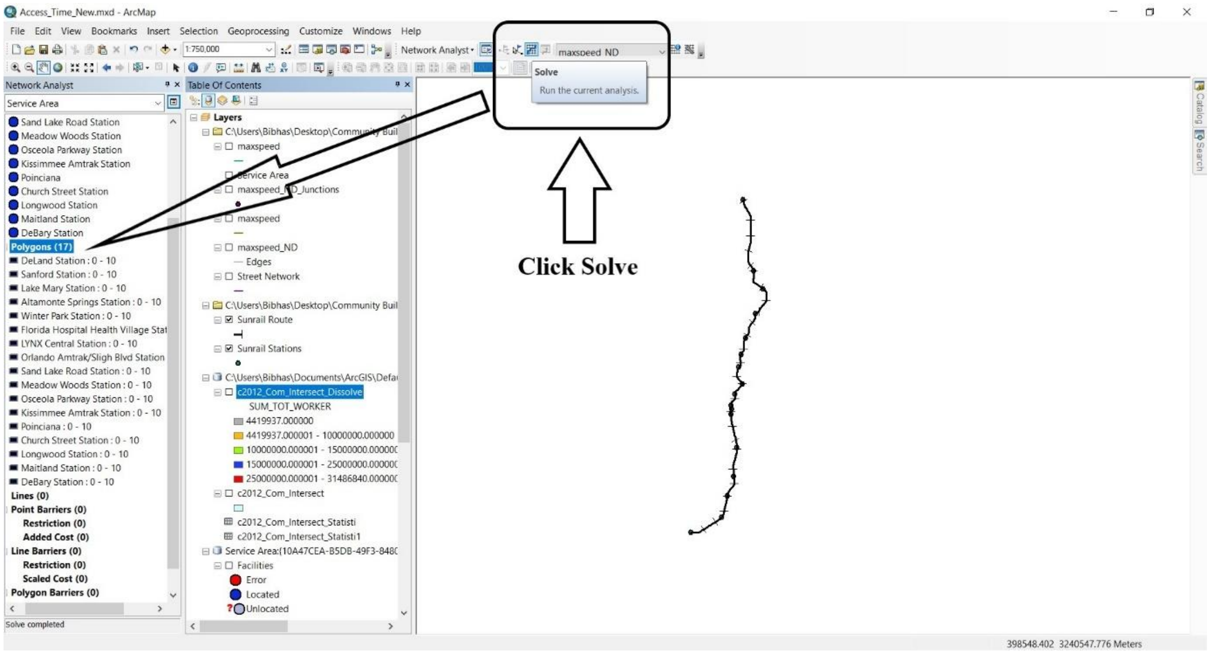
(d)


Access_Time_New.mxd - ArcMap
File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help
Network Analyst
Service Area
Facilities (17)
DeLand Station
Sanford Station
Lake Mary Station
Winter Park Station
LYNX Central Station
Poinciana
Church Street Station
Longwood Station
Maitland Station
DeBary Station
Polygons (0)
Point Barriers (0)
Restriction (0)
Added Cost (0)
Line Barriers (0)
Scaled Cost (0)
Polygon Barriers (0)
Table Of Contents
Layers
Service Area
Copy
Remove
Zoom To Layer
Zoom To Make Visible
Visible Scale Range
Data
Use Symbol Levels
Save As Layer File...
Create Layer Package...
Reset Symbology...
Solve
Recalculate Location Fields
Properties...
Layer Properties
Display the properties of this layer
SUM_TOT_WORKER
c2012_Com_Intersect
Facilities
Error
Located
Unlocated
370290.845 3171650.139 Meters

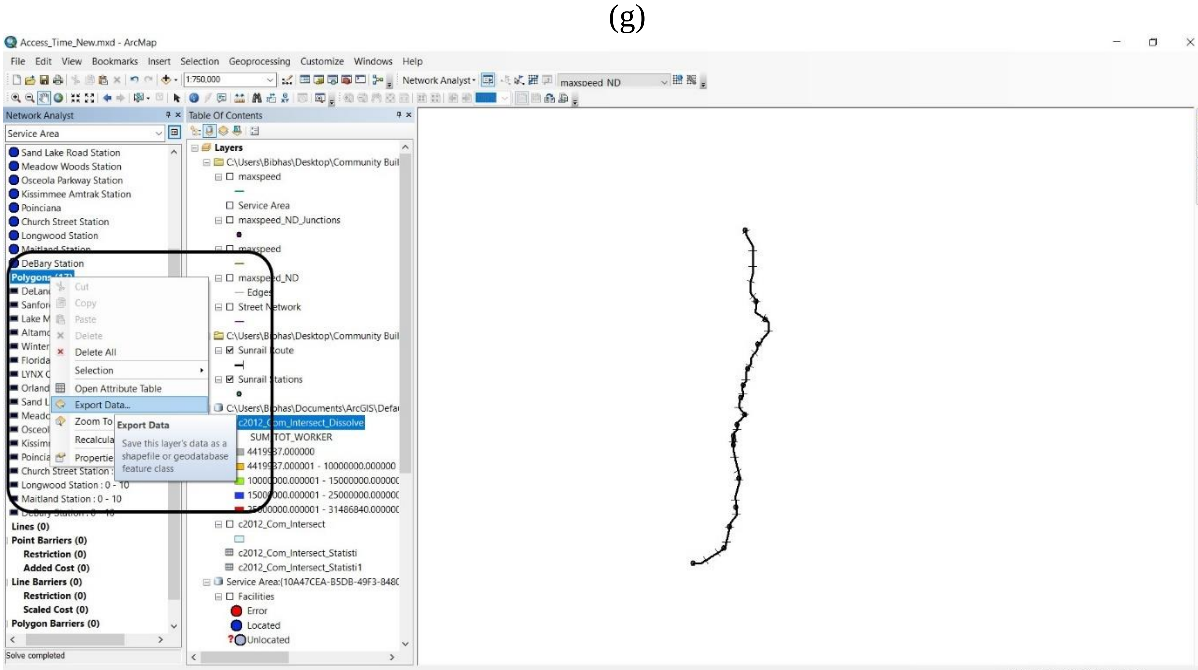
(e)


Access_Time_New.mxd - ArcMap
Layer Properties
General Layers Source Analysis Settings Polygon Generation Line Generation Accumulation Network Locations
Settings
Restrictions
Impedance: Minutes (Minutes)
Default Breaks: 10
Use Time:
Time of Day: 8 AM
Day of Week: Tuesday
Specific Date: 11/19/2019
Direction:
Away From Facility
Towards Facility
U-Turns at Junctions: Allowed
Use Hierarchy
Ignore Invalid Locations
About the service area analysis layer
OK
Cancel
Apply
369655.844 3174983.895 Meters

(f)



(g)



(h)

Figure A.30: Network Driving Area

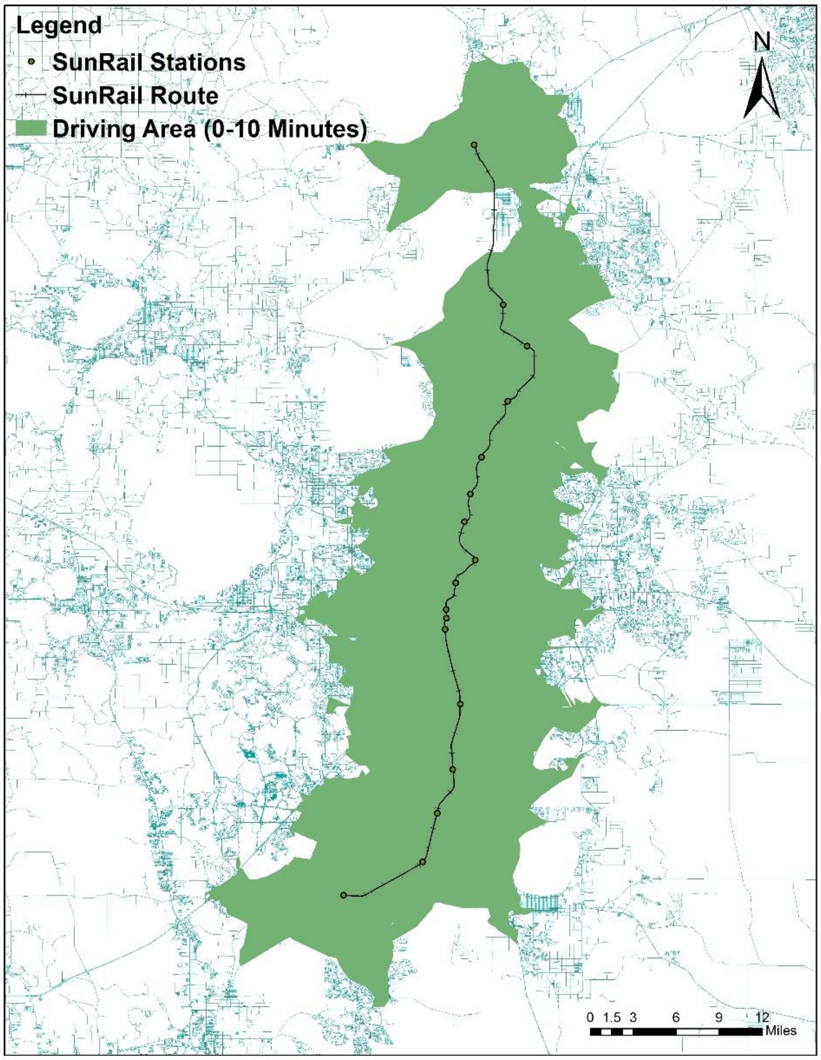


Figure A.31: Driving Network Area Across SunRail Stations

A.2.1.2 Control Area Selection

- To examine the economic impact of SunRail commuter system with respect to number of employed persons, control areas were selected using following procedure: First, we draw a 10 minutes car driving area around the stations.
- We select 20-30 minutes car driving time as our control threshold. Second, the census tracts located within this 10 minute threshold area (between 20 and 30 minutes) were selected to be the candidate control.
- Control area selection procedure is almost same as case area. We put 20 and 30 minutes in 'Default Breaks' option instead of 10 for case area (see Figure A.28).
- Figure A.29 represents the control area around all SunRail stations.

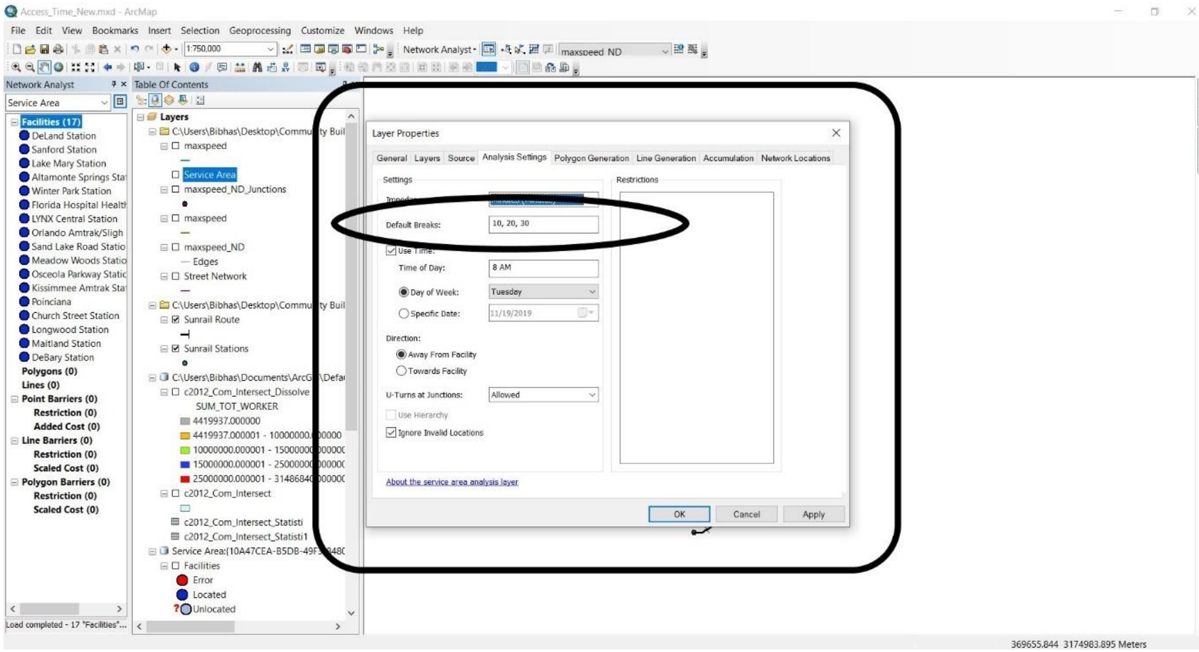


Figure 32: Control Area

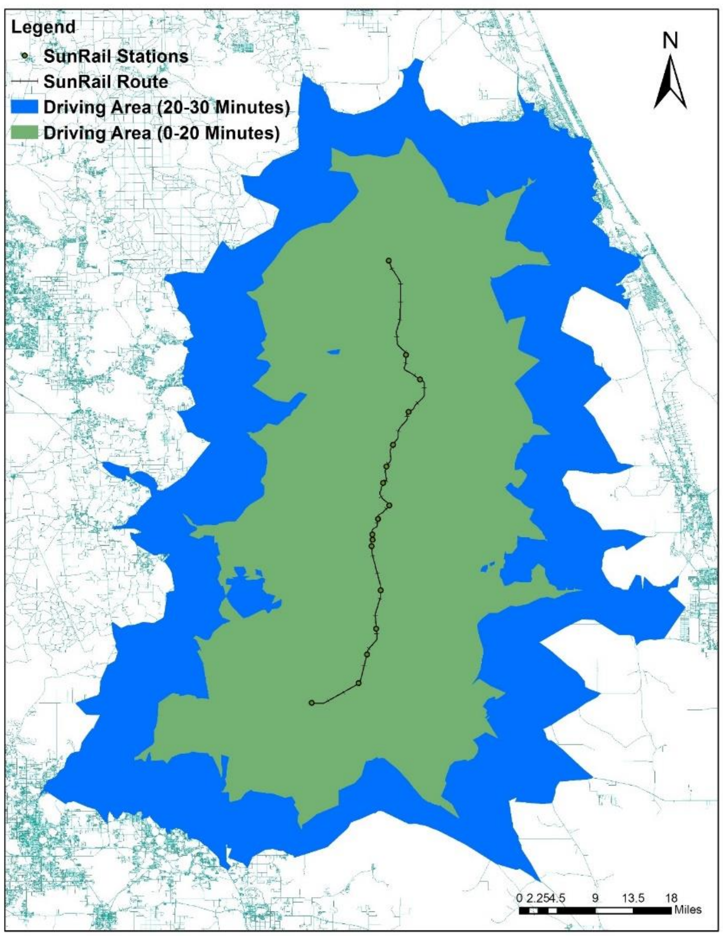


Figure A.33: Control Area Across SunRail Station

A.2.1.3 Accessible Job Estimation

After case and control area selection, all other procedure is quite similar to property value estimation. The employment (number of workers in the labor force) data was drawn from American Community Survey (ACS) was used instead of parcel data for property value.

A.3 Commuting Time

The whole procedure is similar to property value estimation except average commuting time data (journey to work in minutes) per census tract of Florida drawn from American Community Survey (ACS) were used as an alternative of parcel data.

A.4 Land Use Change

The same parcel data similar to property value was used to estimate land use change. After case and control area selection, SPSS file was used to estimate the total area change from vacant to various land use type.

A.5 Travel Pattern for Zero Car Households

The means of transportation to work by household vehicle fleet size data at the census tract level for 2011-2016 was extracted from American Community Survey (ACS) was used to estimate the percentage of trip number by various modes for zero car households. The estimation procedure is similar to property value estimation.

Appendix B. Tutorial for Social Media Data Analysis

This tutorial is an instruction for running the code of social media data analysis. It mainly contains four parts and the codes can be seen as the attached Jupyter Notebook – '**Social Media Data Analysis.ipynb**'. It has four sections: i) data collection, ii) sentiment analysis, iii) visualization of the sentiment analysis results, iv) visualization of the topic analysis results, and iv) topic analysis.

B.1 Software Installation and Data Downloading

B.1.1 Software Installation

In this project, we use Jupyter notebook as the integrated development environment for Python. To install the Jupyter notebook, we can download the anaconda from the official website, shown in **Figure B.1**. The link of the website is: https://www.anaconda.com/products/individual. We can choose the suitable version of anaconda from the website (e.g. 64-Bit Graphical Installer).

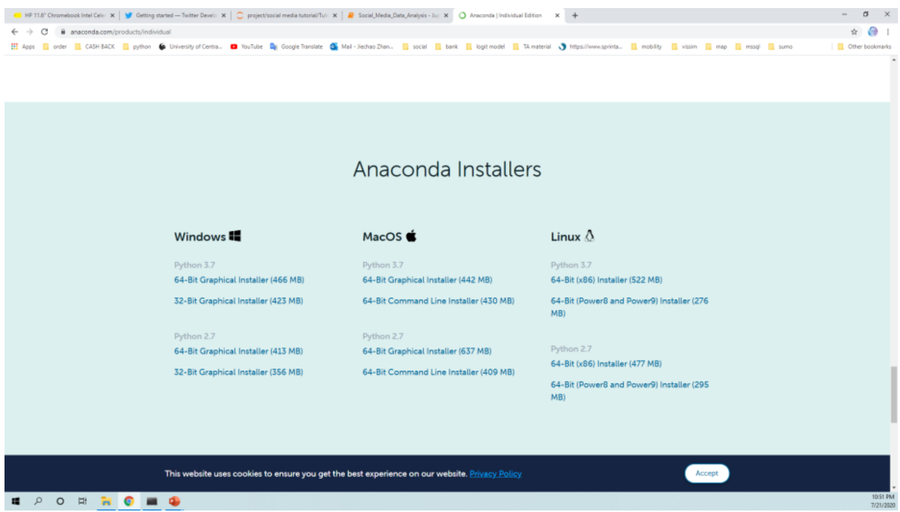


Figure B.1 The official website of Anaconda

After installing the anaconda, we can open the Jupyter notebook from anaconda navigator, shown in **Figure B.2**.

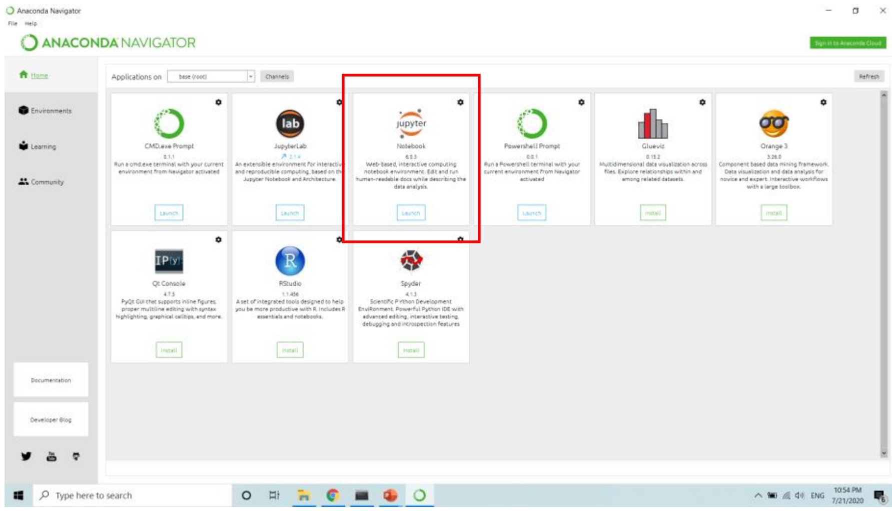


Figure B.2 The platform of anaconda navigator

In the Jupyter notebook, we can update the code file - Social Media Data Analysis.ipynb, shown in **Figure B.3**.

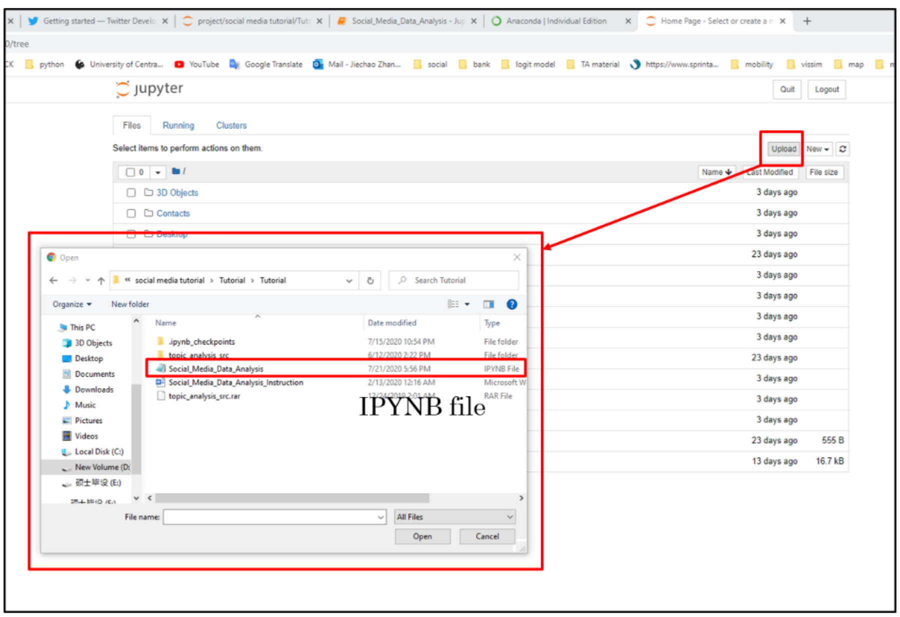


Figure B.3 The platform of Jupyter notebook

B.1.2 Data Collection

B.1.2.1 Apply for a Twitter Developer Account

To collect the Twitter data, one of the prerequisites is to have a Twitter developer account. Thus, we need to apply for a Twitter developer account online through the following link: https://developer.Twitter.com/en/apply-for-access. **Figure B.4** shows the website.

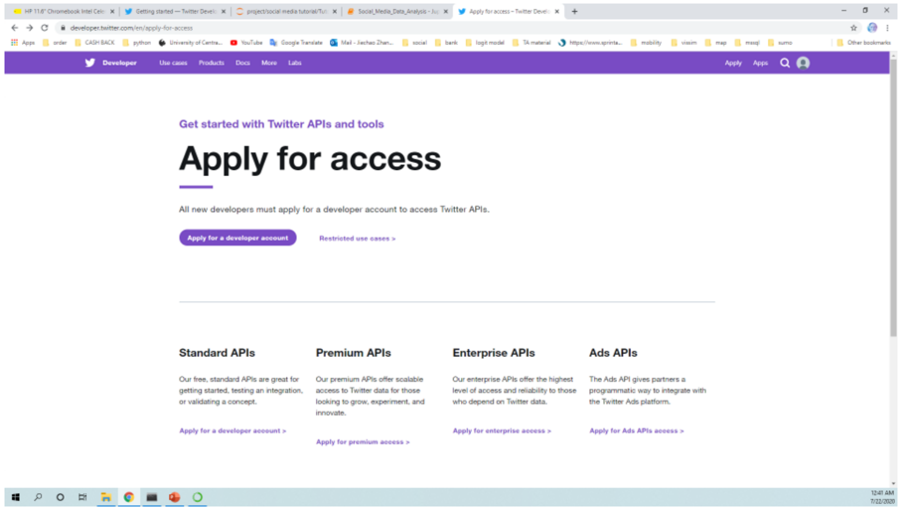


Figure B.4 The website for applying the Twitter developer account

B.1.2.2 Data Collection (User Accounts)

In the Jupyter Notebook, the code for data collection is shown under the heading '**1 Data Collection (User Accounts)**', **seen as Figure B.5**.

The **Twitter_app_auth** is the Twitter API credentials for which one needs to apply to the Twitter official website.

To collect Twitter data by user accounts, we need a 'csv' file called 'List_User.csv' (this file name should be fixed), seen in **Figure B.6**. The 'List_User.csv' file contains all the user accounts for which data need to be collected, and the format can be seen in **Figure B.6**. Put the 'List_User.csv' file and the code in the same folder.

For different collection time, we can create different folders to save the Twitter data. Each folder must contain both the 'List_User.csv' file and the data collection code. From the code, we can change the 'June_10_tweets' (shown in **Figure B.5**) to the expected date to save the Twitter data with a different file name.

After all the files are prepared, run the code and the Twitter data will be collected in the same folder where the code for data collection is kept.

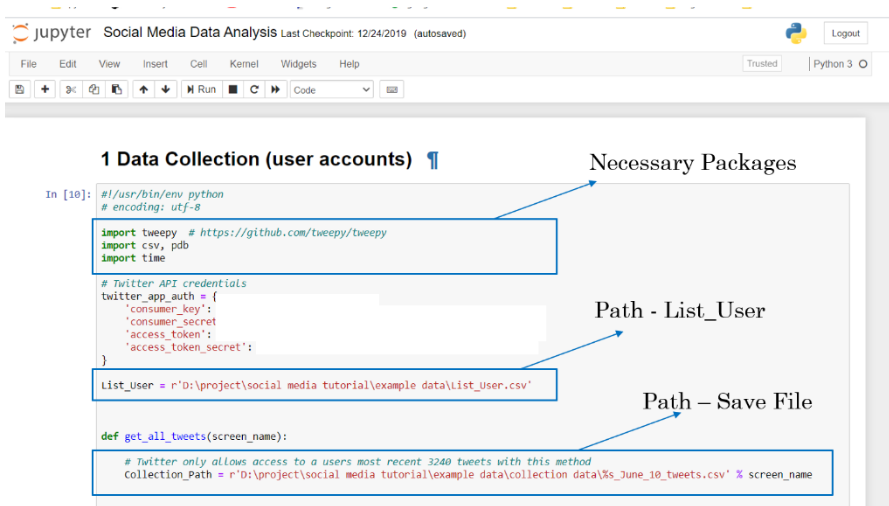


Figure B.5 Example of User Account Data Collection Code

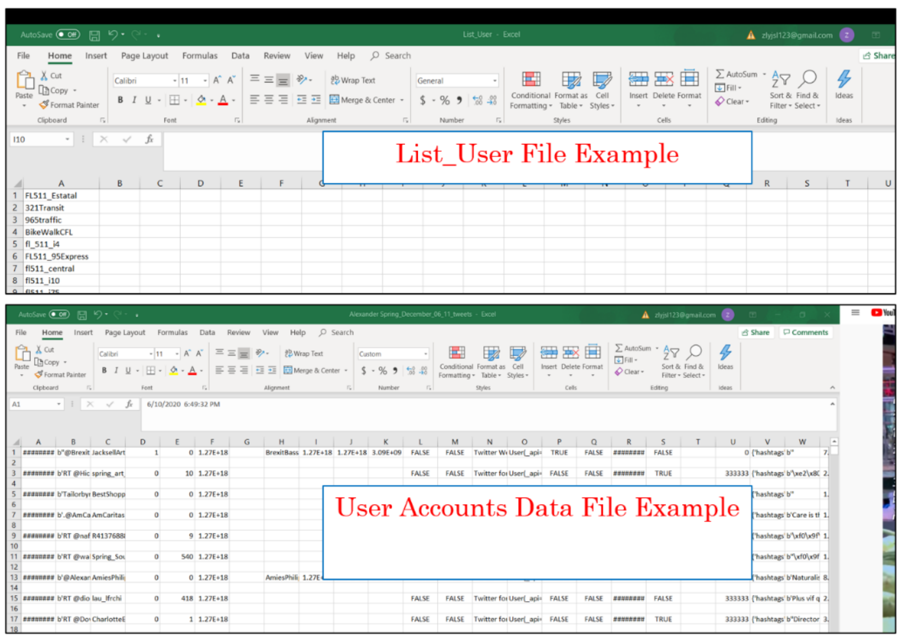


Figure B.6 The example of List_User.csv and output file

B.1.2.3 Data Collection (Keywords)

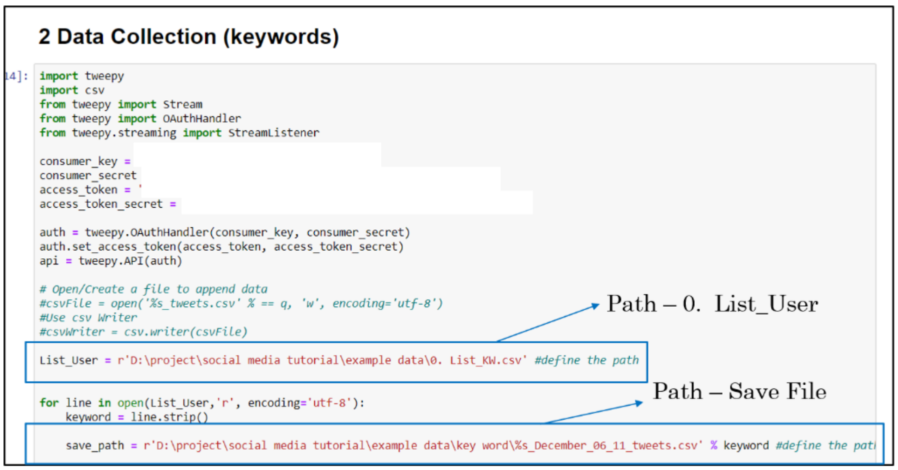
The code for data collection is shown under the heading ‘**2 Data Collection (keywords)**’, **seen as Figure B.7 (a)**.

The **Twitter_app_auth** is the Twitter API credentials for which one needs to apply to the Twitter official website.

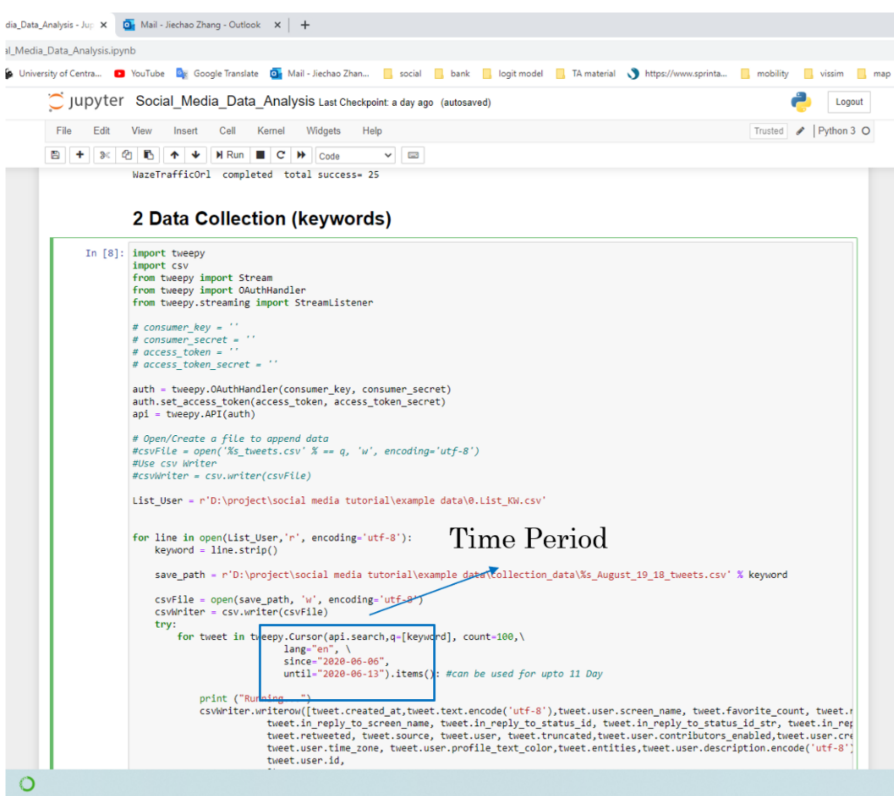
To collect Twitter data by user accounts, we need a ‘csv’ file called ‘0.List_KW.csv’ (this file name should be fixed), seen as **Figure B.8**. The ‘0.List_KW.csv’ file contains all the keywords which need to be collected, and the format can be seen as **Figure B.8**. Put the '0.List_KW.csv' file and the code in the same folder.

For different collection time, we can create different folders to save the Twitter data. Each folder must contain both the '0.List_KW.csv' file and the data collection code. From the code, we can change the time periods (shown in **Figure B.7 (b)**) to the expected date to save the Twitter data with different file name.

After all the files are prepared, run the code and the Twitter data will be collected in the same folder where code for data collection is kept seen as **Figure B.8**.



(a)



(b)

Figure B.7 Example of Keyword Data Collection Code: (a) the code for input and output files; (b) the code for changing the time period of collection

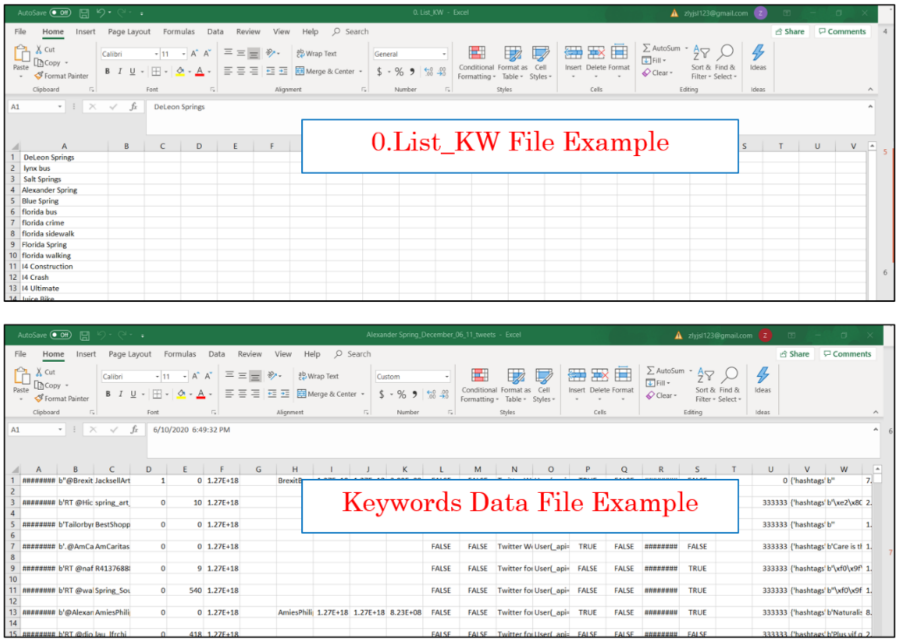


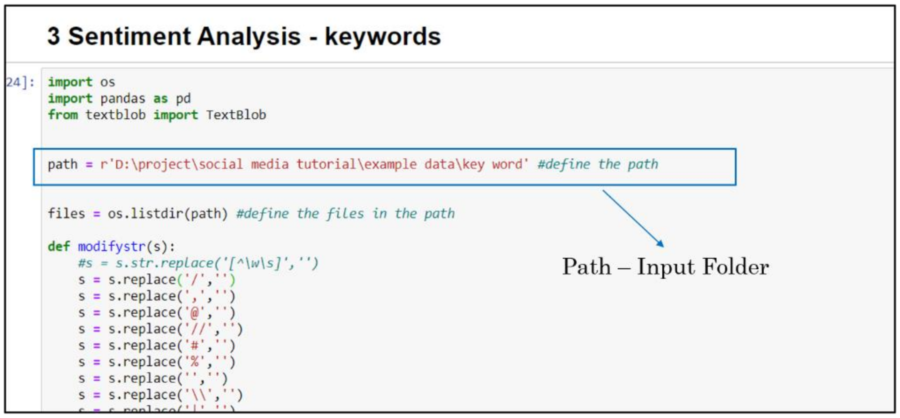
Figure B.8 The example of 0.List_KW.csv and output file

B.2 Sentiment Analysis and Visualization

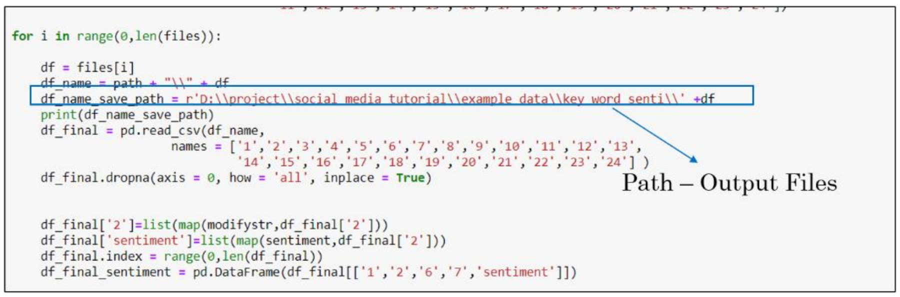
B.2.1 Sentiment Analysis

The second part in the Jupyter Notebook is the sentiment analysis which can be seen under the heading '**3 Sentiment Analysis**'. The input of the sentiment analysis is the data collected from the part 1. From the codes, the '**path**' defines the path of the input file folder and the 'files' is a list of the filename. Given the specific folder path, the output of this program is the sentiment analysis results. The sentiment analysis results contain the 'user_id', 'polarity', and 'subjectivity'. The examples can be seen as **Figure B.9**.

After setting the input path, run the codes and the '**df_final_sentiment**' is the output data frame. The example of input file path in the codes can be seen as **Figure B.9 (a)** and the output file path can be seen as **Figure B.9 (b)**.



(a)



(b)

Figure B.9 Example of Sentiment Analysis Code: (a) the code for input files; (b) the code for output files

B.2.2 Visualization of Sentiment Analysis Results

The visualization of sentiment analysis results can be seen as '**5 Sentiment Analysis Visualization**' in the code file. The example of the code for visualizing the sentiment analysis is shown as **Figure B.10**. From the code, we can change the path of input and output files as well as the time periods seen from **Figure B.10**.

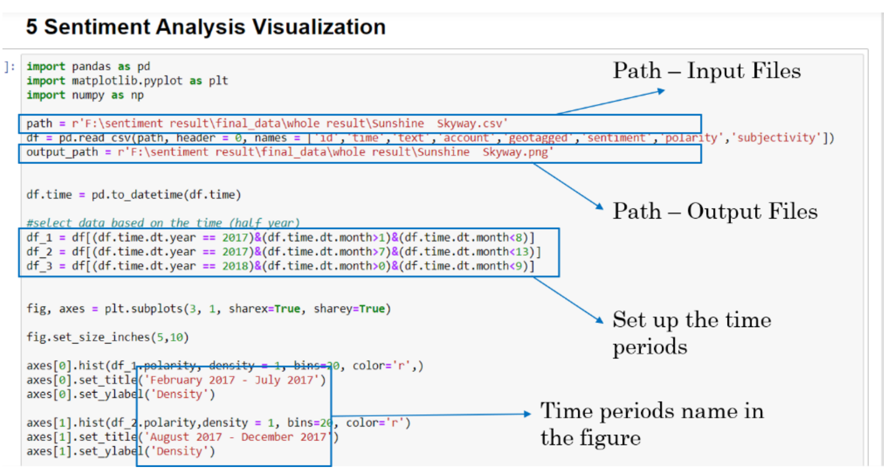


Figure B.10 Example of Sentiment Analysis Code: (a) the code for input files; (b) the code for output files

Run this code with the input file, the figure will be generated. **Figure B.11** shows an example figure.

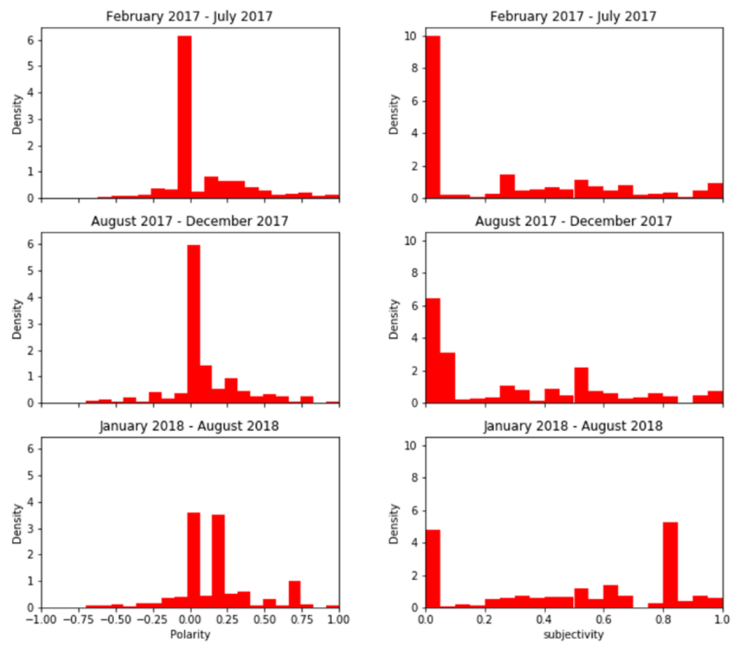


Figure B.11 Example of the visualization of sentiment analysis

B.3 Topic Analysis and Visualization

B.3.1 Data Processing for Topic Analysis

The third part in the Jupyter Notebook is the topic analysis. The first step of the topic analysis is to process the data which can be seen under the heading **‘7 Topic Model Data Processing’**, shown in Figure 3.1. From the code, we can change the path of input and output files to save the processed data for topic analysis. Some examples of input and output files are shown in **Figure B.12**.

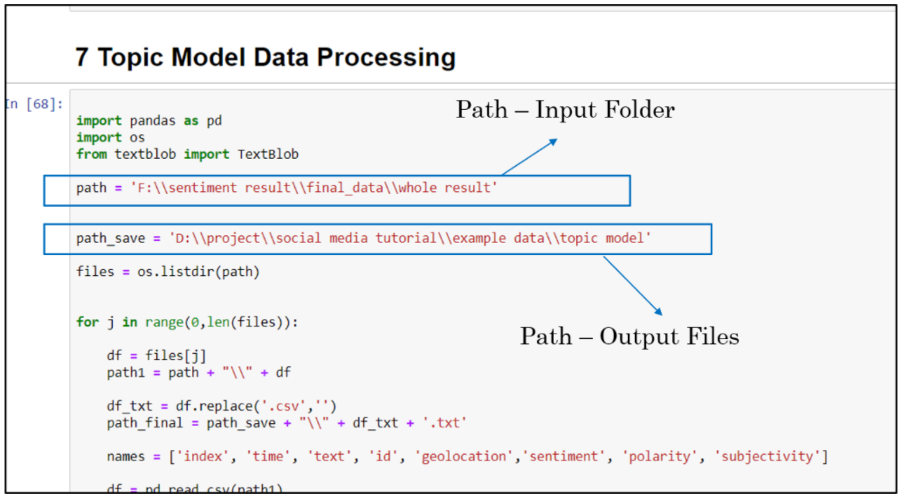


Figure B.12 Example of the topic analysis

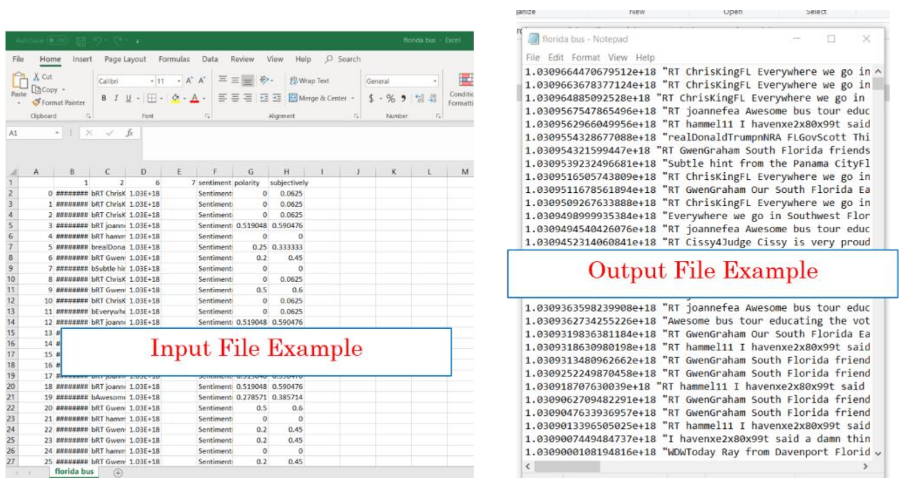


Figure B.12 Example of the input and output files of topic analysis

B.3.2 Prerequisite for Topic Analysis

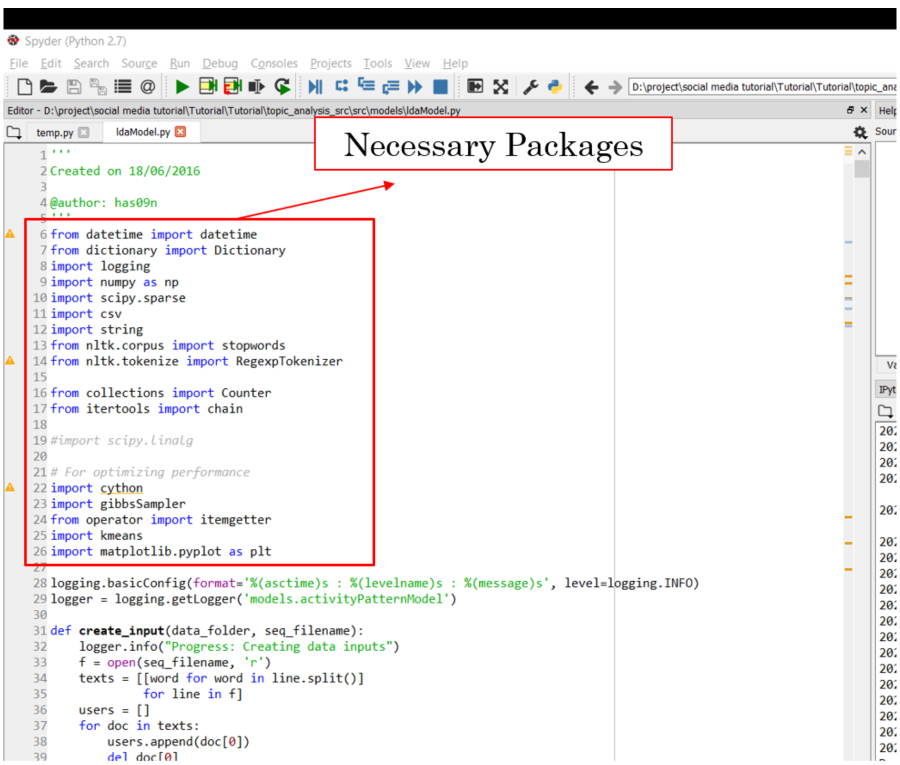
The code for topic analysis is written in Python 2 version which means that we need to use python 2 version to run the topic analysis. We use the Spyder IDE (in anaconda navigator) for python 2.7 version to apply the topic analysis. We can follow the instructions to install the required software and run the model.

- Download Anaconda (python 2.7) 32-bit Graphical Installer
- Open Spyder from the anaconda navigator (python 2.7) version
- Open the ldaModel.py
- Install all the necessary python packages
- Change the input path and file name
- Run the model

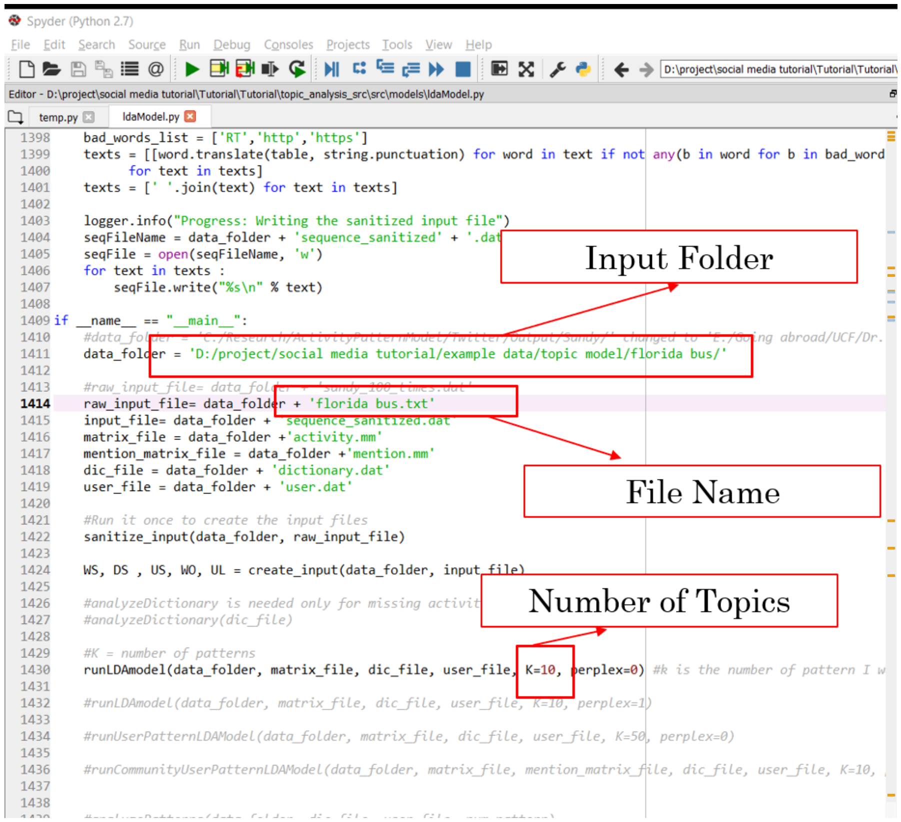
B.3.3 Topic Analysis

For the topic analysis, we use a tool based on Python 2.7 version environment. Thus, Python 2.7 is required in the topic analysis program. The procedure of topic analysis can be seen as follows:

- Download the 'topic_analysis_src.rar' archive and unzip the archive.
- Find the ldaModel.py file, which contains the code for topic analysis.
- Based on the ldaModel.py, install all the required python packages, seen as **Figure B.13 (a)**.
- Process the raw Twitter data into the input file of topic analysis. The input file of topic analysis contains two column – 'user_id' and 'tweets', which can be found in Figure 6.
- At the end of the codes (ldaModel.py), the '**data_folder**' (seen as **Figure B.13 (b))** should be changed into the path where the 'topic_analysis_src/model' is and the 'raw_input_file' is the path of the input file. The example of input file can be seen as **Figure B.14.**
- In the **runLDAmodel** function, ***k*** represents the number of topics. Set the number of topics by changing the value of '**k**'.
- After running all the above process, run the ldaModel.py and the results will be saved in the same path of the 'topic_analysis_src'. The output file name is 'RT_LDA_patterns'.



(a)



(b)

Figure B.13 Example of the topic analysis: (a) necessary packages; (b) input path

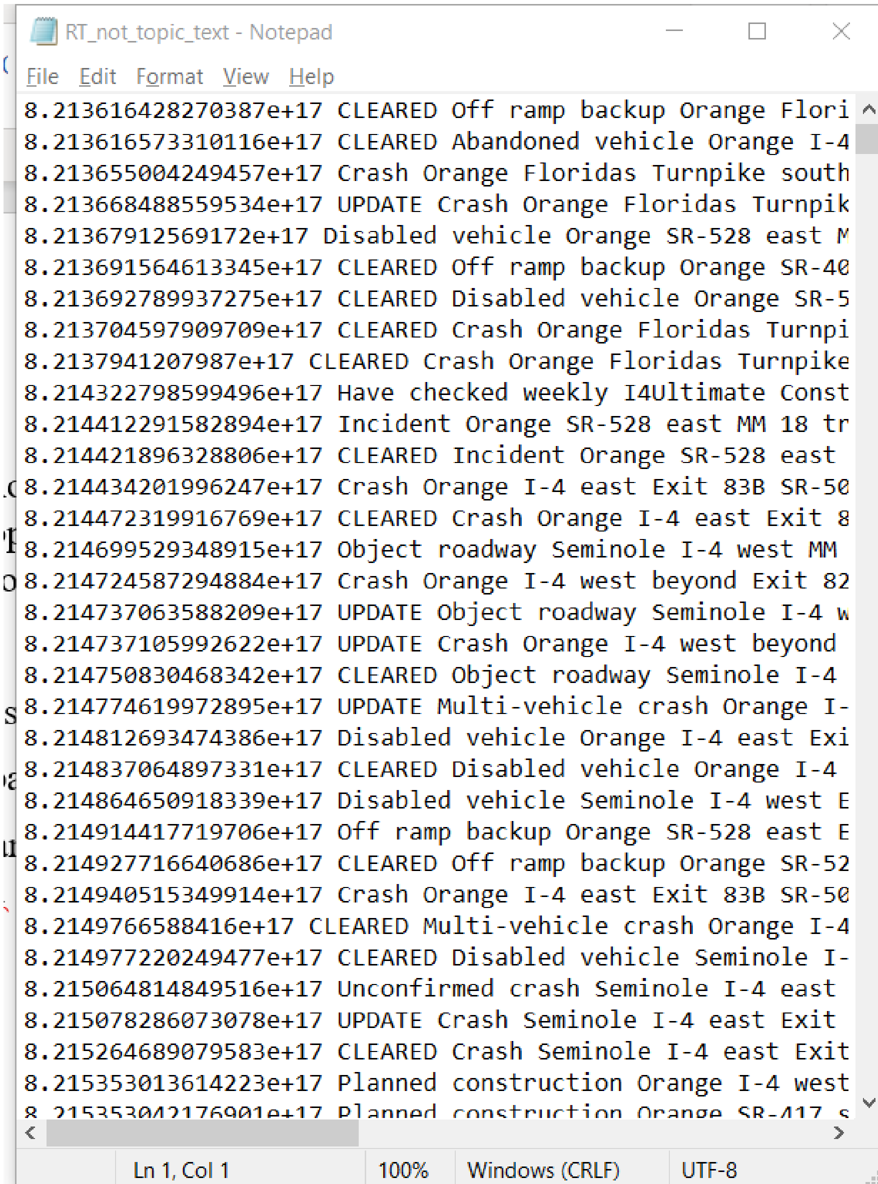

RT_not_topic_text - Notepad
File Edit Format View Help
```
8.213616428270387e+17 CLEARED Off ramp backup Orange Flori
8.213616573310116e+17 CLEARED Abandoned vehicle Orange I-4
8.21365504249457e+17 Crash Orange Floridas Turnpike south
8.213668488559534e+17 UPDATE Crash Orange Floridas Turnpik
8.2136791256917 2e+17 Disabled vehicle Orange SR-528 east M
8.213691564613345e+17 CLEARED Off ramp backup Orange SR-40
8.213692789937275e+17 CLEARED Disabled vehicle Orange SR-5
8.213704597909709e+17 CLEARED Crash Orange Floridas Turnpi
8.2137941207987e+17 CLEARED Crash Orange Floridas Turnpike
8.214322798599496e+17 Have checked weekly I4Ultimate Const
8.214412291582894e+17 Incident Orange SR-528 east MM 18 tr
8.214421896328806e+17 CLEARED Incident Orange SR-528 east
8.214434201996247e+17 Crash Orange I-4 east Exit 83B SR-50
8.214472319916769e+17 CLEARED Crash Orange I-4 east Exit 8
8.214699529348915e+17 Object roadway Seminole I-4 west MM
8.214724587294884e+17 Crash Orange I-4 west beyond Exit 82
8.214737063588209e+17 UPDATE Object roadway Seminole I-4 w
8.214737105992622e+17 UPDATE Crash Orange I-4 west beyond
8.214750830468342e+17 CLEARED Object roadway Seminole I-4
8.214774619972895e+17 UPDATE Multi-vehicle crash Orange I-
8.214812693474386e+17 Disabled vehicle Orange I-4 east Exi
8.214837064897331e+17 CLEARED Disabled vehicle Orange I-4
8.214864650918339e+17 Disabled vehicle Seminole I-4 west E
8.214914417719706e+17 Off ramp backup Orange SR-528 east E
8.214927716640686e+17 CLEARED Off ramp backup Orange SR-52
8.214940515349914e+17 Crash Orange I-4 east Exit 83B SR-50
8.2149766588416e+17 CLEARED Multi-vehicle crash Orange I-4
8.214977220249477e+17 CLEARED Disabled vehicle Seminole I-
8.215064814849516e+17 Unconfirmed crash Seminole I-4 east
8.215078286073078e+17 UPDATE Crash Seminole I-4 east Exit
8.215264689079583e+17 CLEARED Crash Seminole I-4 east Exit
8.215353013614223e+17 Planned construction Orange I-4 west
8.215353042176901e+17 Planned construction Orange SR-417 s
```
Ln 1, Col 1 | 100% | Windows (CRLF) | UTF-8

Figure B.14 Example of the input file

There are also three required files – dictionary.dat, sequence_sanitized.dat and user.dat, seen as **Figure B.15.**

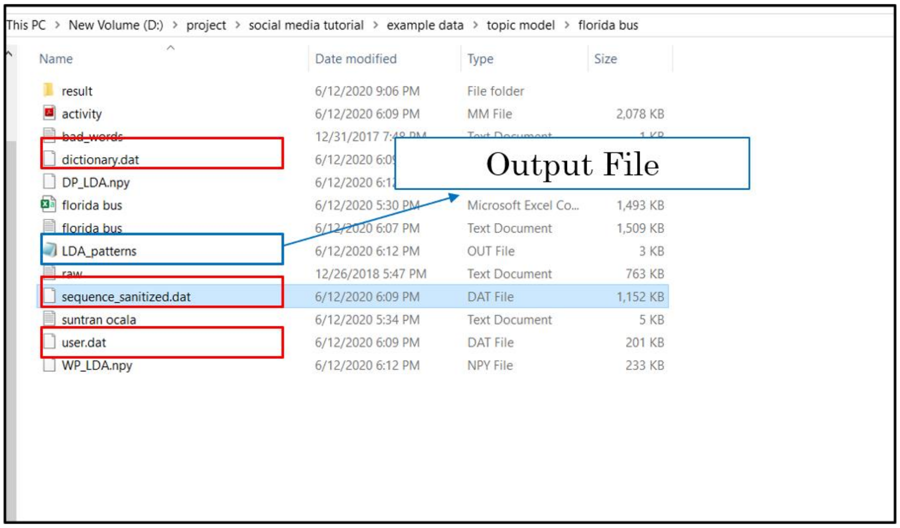


Figure B.15 Example of required files and output file for topic analysis

B.3.4 Topic Analysis Results Visualization

The first step for visualization of the topic analysis results is to process the data which can be seen under the heading '**8 Data Processing for Visualization**'. In the codes, the '**path_input**' is the path of the input file (format of the input file can be seen as **Figure 3.7**). The code for topic analysis visualization can be seen as **Figure B.16.** In the code, we should change the input file path to read the input file and output figure path to save the output figures.

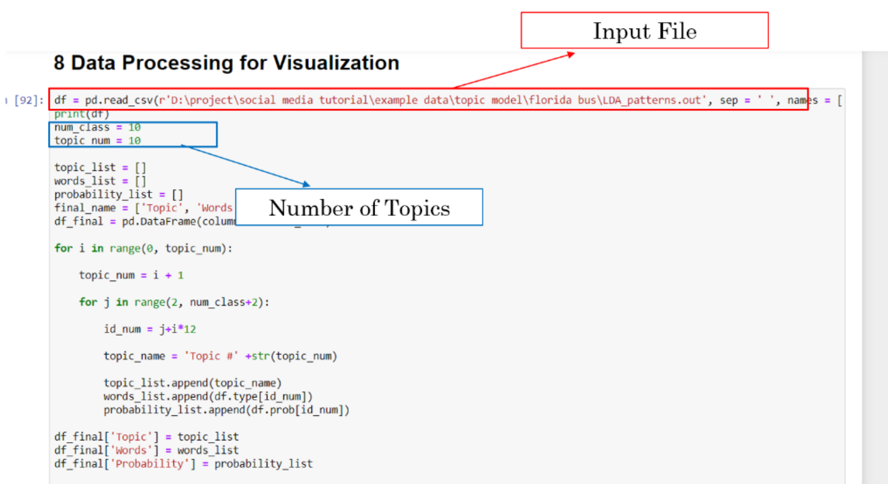


Figure B.16 Data Processing for Visualization of Topic Analysis

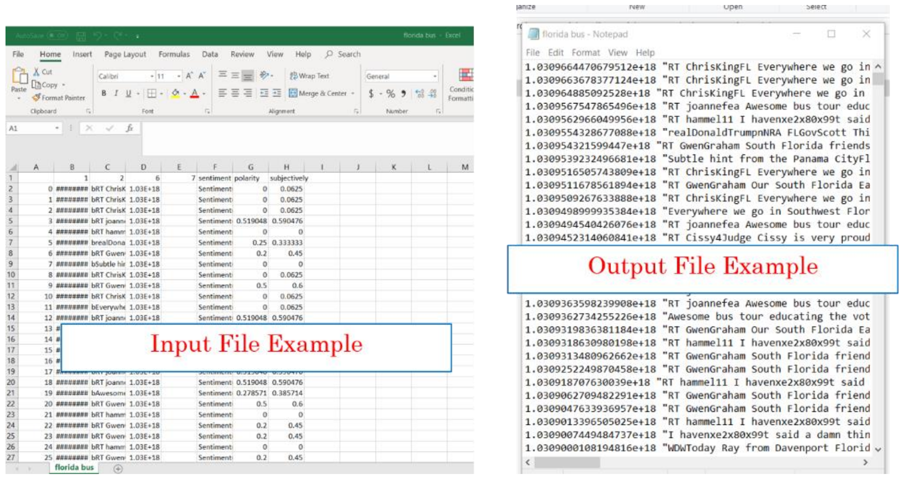


Figure B.17 Data Samples for Visualization of Topic Analysis

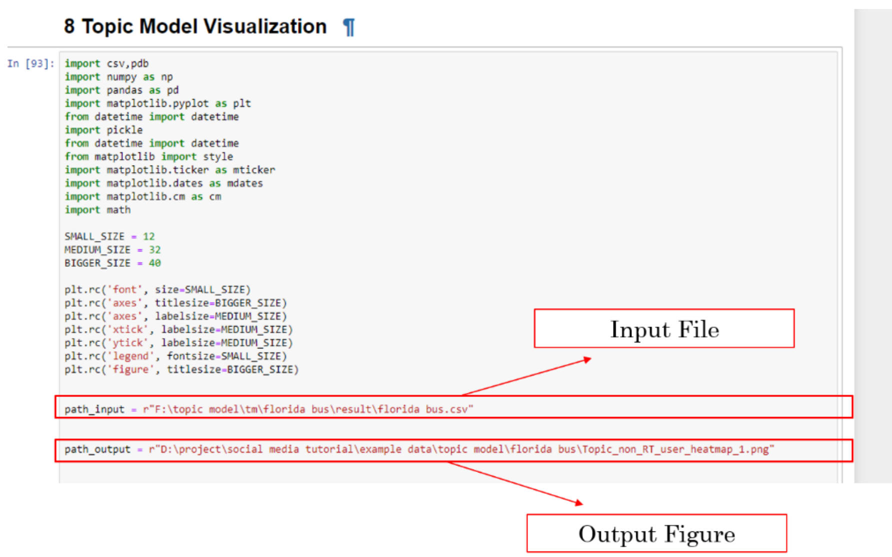


Figure B.18 Example of Visualization of Topic Analysis

Run this code with the input file, the figure can be shown. **Figure B.19** shows one of the examples.

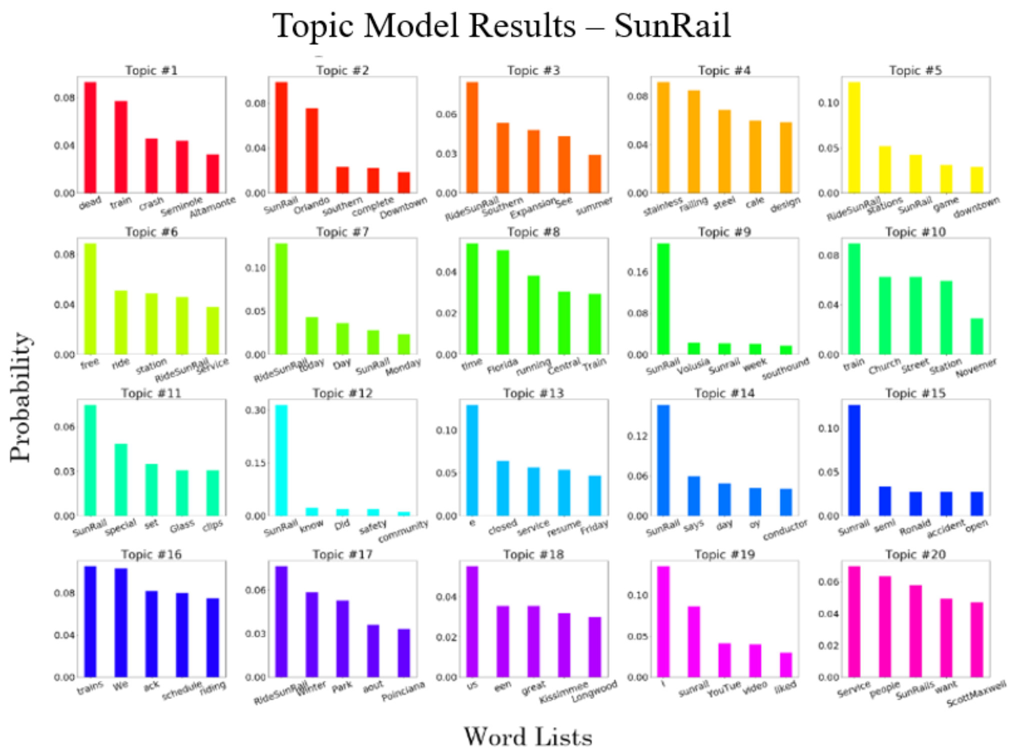


Figure B.19 Example of the visualization of topic analysis