## **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

List of Table	28	2
List of Figur	es	3
Chapter 1: N	Iulti-criteria Decision Analysis	4
1.1 Intr	roduction	4
1.2 Lite	erature Review	5
1.2.1	Multi-Attribute Decision Making (MADM)	5
1.2.2	Prevalent MADM Models	7
1.2.3	Weighting the Criteria	14
1.3 Me	thodology	15
1.3.1	Weighting of Criteria	15
1.3.2	Scores of the Criteria	16
1.3.3	Scoring and Ranking the Projects	16
1.4 Ana	alysis and Results	16
1.4.1	Weighting of Criteria	16
1.4.2	Value of the Criteria (MOEs)	17
1.4.3	Overall Scoring	
1.5 Cor	nclusion	
Chapter 2: K	Inowledge Transfer	
References		
APPENDIX	A: Tutorial for Measures of Effectiveness Estimation	
APPENDIX	B. Tutorial for Social Media Data Analysis	79

# **Table of Content**

# List of Tables

Table 1: Reviewed Literature in MADM	7
Table 2: Weights of different criteria	8
Table 3: Value of criteria obtained from detailed evaluation	8
Table 4: Scoring of alternatives	8
Table 5: Rank of the alternatives	9
Table 6: Weights and value of different criteria	9
Table 7: Table of preference by WPM	9
Table 8: Priority table by AHP	10
Table 9: Ranking by AHP	10
Table 10: Estimation of S <sub>j</sub> and R <sub>j</sub>	11
Table 11: Ranking by VIKOR method	12
Table 12: Preference matrix	13
Table 13: Overall Index Calculation	13
Table 14: Outranking Flow Calculation	13
Table 15: Net Flow and Ranking by PROMETHEE	14
Table 16: Preference rating point table	14
Table 17: Pairwise comparison Matrix	14
Table 18: Determining priorities from Pairwise comparison matrix	15
Table 19: Preference Rating Point Table	15
Table 20: Pairwise Comparison Matrix	15
Table 21: Determining Priorities from Pairwise Comparison Matrix	16
Table 22: Scoring Example of the Projects	16
Table 23: Pairwise Comparison of Criteria (MOEs) Based on Survey	17
Table 24: Weights of the Criteria	17
Table 25: Property Value Evaluation for SunRail Project	18
Table 26: Property Value Evaluation for I-4 Expansion Project	19
Table 27: Property Value Evaluation for Juice Bikeshare Project	20
Table 28: Job Accessibility Evaluation Change for SunRail Project	20
Table 29: Job Accessibility Change Evaluation for I-4 Expansion Project	21
Table 30: Job Accessibility Change Evaluation for Juice Bikeshare Project	21
Table 31: Commuting Travel Time Change Evaluation of SunRail Project	21
Table 32: Commuting Travel Time Change Evaluation of I-4 Expansion Project	22
Table 33: Commuting Travel Time Change Evaluation of Juice Bikeshare Project	22
Table 34: Land Use Change for SunRail Project	22
Table 35: Land Use Change for I-4 Expansion Project	23
Table 36: Land Use Change for Juice Bikeshare Project	23
Table 37: Travel Pattern Change Evaluation of SunRail Project	23
Table 38: Travel Pattern Change Evaluation of I-4 Expansion Project	24
Table 39: Travel Pattern Change Evaluation of Juice Bikeshare Project	24
Table 40: Scoring and Ranking of the Projects	24

# List of Figures

Figure 1: Classification of multicriteria decision making	5
Figure 2: Hierarchical process of multi-attribute decision making	6
Figure 3: The analytical hierarchy process	10
Figure 4: Steps of PROMETHEE method	12

## **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 <u>first step of the</u> <u>analysis</u>, 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 <u>second step</u>, 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 <u>third step</u>, 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 <u>fourth step</u>, 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 <u>first step</u>, each alternative is scored based on their criteria values and criteria specific weights. In the <u>final step</u>, 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.

Study	Model	Weighting method	Applied for a transportation project?
Bottero et al., 2018	Bottero et al., 2018 PROMETHEE		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 <sup>°</sup> 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	Review of available		No
Sanchez, 1997	method		110
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

 Table 1: Reviewed Literature in MADM

### 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):

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

 $A_{wsm}$  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*<sup>th</sup> alternative in terms of the *j*<sup>th</sup> criterion and  $\omega_j$  is the weight of importance of the *j*<sup>th</sup> criterion. For each criterion, the decision maker has to determine its importance, or weight,  $\omega_i$ . It is also assumed that the following relationship is always true:

 $\sum \omega_i = 1$ 

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,  $A_3 = Project C$ 

Alternatives Project A

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  $\sum \omega_i = 1$ . Assume that the assigned weights of the criteria are as follows:

Criteria	C1	C2	C3	C4	C5
Weight, $\omega_j$	0.3	0.2	0.2	0.2	0.1

 Table 2: Weights of different criteria

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:

			a monn actailea	eralaalon	
Alternatives	C1	C2	C3	C4	C5
Project A	10%	20%	10%	30%	5%
Project B	20%	30%	10%	10%	15%
Project C	30%	10%	20%	10%	20%

Table 3. Value of criteria obtained from detailed evaluation

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

Table 4. Scoring of alternatives		
	Scores	
	15.5	

Table 4:	Scoring	of alter	matives
----------	---------	----------	---------

I lojeet A	15.5
Project B	17.5
Project C	19.0
On the basis of scores, alternatives are ranked a	and the alternative, having the highest score, is

ranked 1st.

Table 5: Kank 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  $\omega_j$  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:

Criteria Alternatives  $C_1$  $\mathbf{C}_2$ **C**<sub>4</sub> **C**<sub>3</sub> C<sub>5</sub> Weight, w 0.2 0.3 0.2 0.2 0.1 Project A 10% 20% 30% 5% 10% 10% 15% Project B 20% 30% 10% Project C 30% 10% 20% 10% 20%

**Table 6:** Weights and value of different criteria

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:

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

#### Table 7: Table of preference by WPM

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: Phonty table by AHP						
Alternatives	Project A	Project B	Priority			
Project A	1	4	4/5			
Project B	1/4	1	1/5			

Table 8:	Priority	table by	AHP
----------	----------	----------	-----

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.

Criteria	Land value change	Travel commuting time	Job accessibility	<b>Overall</b> <b>Priority</b>	Rank
Weights	0.2	0.4	0.4	1 1101 Ity	
Project A	0.8	0.5	0.4	0.52	1
Project B	0.2	0.5	0.6	0.48	2

Table 9: Ranking by AHP

#### 1.2.2.4 VIKOR Method

The Compromise Ranking method, also known as the VIKOR method, introduces the Multicriteria 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*<sup>th</sup> function represents a benefit then  $f_i^* = \max f_{ij}$  and  $f_i^- = \min f_{ij}$ . In contrast, If the *i*<sup>th</sup> 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  $j = 1, 2, 3, \dots, J$  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,  $S_i$  and  $R_j$  for the three alternatives are presented in following table:

Altomotivos	Criteria						
Alternatives	C1	C2	С3	<b>C</b> 4	C5		
Weight, ω <sub>i</sub>	0.3	0.2	0.2	0.2	0.1	$S_i$	$R_i$
<b>f</b> *	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

**Table 10:** Estimation of  $S_j$  and  $R_j$ 

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

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

 $\nu$  is introduced as a weight for the strategy of maximum group utility and value of  $\nu$  ranges from 0 to 1 (Assumed 0.5 for this example). Here,  $s^* = min_jS_j = 0.40$ ,  $s^- = maxS_j = 0.70$ ,  $R^* = min_jR_j = 0.20$  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:

Alternatives	$Q_{j}$	Rank		
Project A	1.00	3		
Project B	0.30	2		
Project C	0.00	1		

Table 11: Ranking by VIKOR method

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.

$$P_j(a, b) = 1$$
 if  $f_{aj} - f_{bj} \ge \theta_j$ 

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

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

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

 Table 12: Preference matrix

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:

				eure ururen		
Overall Preference Index	П (1, 2)	П (1,3)	П (2, 1)	П (2, 3)	П (3, 1)	П (3, 2)
Estimated value	0.2	0.4	0.6	0.2	0.6	0.5

Table 13: Overall Index Calculation

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):

$$\phi_{(a)}^{+} = \frac{1}{n-1} \sum_{b \in A} \Pi(a, b)$$

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

$$\phi_{(a)}^{-} = \frac{1}{n-1} \sum_{b \in A} \Pi(b, a)$$

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

Alternatives	Positive flow	Negative flow			
Project A	$\frac{1}{2} (0.2+0.4) = 0.30$	$\frac{1}{2} (0.6+0.6) = 0.60$			
Project B	0.40	0.35			
Project C	0.65	0.30			

Table 14: Outranking Flow Calculation

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:

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

### **Table 15:** Net Flow and Ranking by PROMETHEE

## 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:

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	

### Table 16: Preference rating point table

\* T.L. Saaty (2008)

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

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

Table 17: Pairwise comparison Matrix

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).

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

**Table 18:** Determining priorities from Pairwise comparison matrix

### 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:

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

**Table 19:** Preference Rating Point Table

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.

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						

**Table 20:** Pairwise Comparison Matrix

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).

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

 Table 21: Determining Priorities from Pairwise Comparison Matrix

## 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.

Criteria	1	2	3	4	Casara	Daula					
Weights	0.53	0.27	0.13	0.07	Scores	капк					
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					

Table 22: Scoring Example of the Projects

## 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:

Oritaria	Property value	Job	Commuting	Land Use	Travel
Criteria	change	accessibility	time	Change	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

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

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).

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			

Table 24: Weights of the Criteria

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.

			~				3			
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)
	Single Family	456,558	315,503	59.13	40.87	18.27	53.89			
-	Multi- Family	1915,108	381,978	83.37	16.63	66.74	5.67	<b>27</b> (0)	176 471	
Downtown	Retail	570,925	163,763	77.71	22.29	55.42	21.20	35.40	176.471	
	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			1.99
	Single Family	221,968	201,879	52.37	47.63	4.74	53.89			
Outside	Multi- Family	767,571	369,499	67.50	32.50	35.01	5.67	14.15	520 41	
Downtown Phase I	Retail	192,578	144,824	57.08	42.92	14.15	21.20	14.15	529.41	
	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			
	Single Family	309,987	322,017	49.05	50.95	-1.90	53.89			
	Multi- Family	1459,670	163,480	89.93	10.07	79.86	5.67	•••••		
Phase II	Retail	955,465	98,549	90.65	9.35	81.30	21.20	20.99	294.12	
	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			

Table 25: Property Value Evaluation for SunRail Project

 $^{1}$  = 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.

								2		
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)
	Single Family	500,104	269,953	64.94	35.06	29.89	49.94			
	Multi- Family	771,574	1038,100	42.64	57.36	-14.73	7.82		1	
Attraction	Retail	869,466	492,299	63.85	36.15	27.70	31.34	30.49	773.531	
	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			
	Single Family	220,217	230,254	48.89	51.11	-2.23	49.94			
<b>D</b>	Multi- Family	836,234	269,488	75.63	24.37	51.26	7.82	20.02	569.97	569.97
Downtown	Retail	519,365	72,903	87.69	12.31	75.38	31.34	29.93		
	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			
	Single Family	354,816	143,717	71.17	28.83	42.34	49.94			0.88
	Multi- Family	1509,037	449,386	77.05	22.95	54.11	7.82		<<10 <b>7</b>	
Ivanhoe	Retail	270,338	30,586	89.84	10.16	79.67	31.34	55.66	664.97	
	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			
	Single Family	141,806	148,119	48.91	51.09	-2.18	49.94			
A.1	Multi- Family	588,380	699,996	45.67	54.33	-8.66	7.82	5.04	0.00 52	
Altamonte	Retail	68,099	102,032	40.03	59.97	-19.95	31.34	-5.86	868.53	
I	Industrial	142,913	231,700	38.15	61.85	-23.70	6.18	.8		
Institutio	Institutional	43,924	5,763	88.40	11.60	76.80	4.71			

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

 $^{1}$  = 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.

Land Use	Property Value Change in Downtow n	Property Value Change in Outside Downtow n	% 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	53.8 9			
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	21.2 0	4.27	0.421	1016.7
Industrial	402773.13	448302.90	47.33	52.67	-5.35	13.0 3			
Institutional	1396720.68	1242913.07	52.91	47.09	5.83	6.21			

 Table 27: Property Value Evaluation for Juice Bikeshare Project

<sup>1</sup> = 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.

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	
Outside Downtown Phase I	12733.35	38325.90	24.94	75.06	-50.12	529.41	-2.38
Phase II	68855.06	50796.25	57.55	42.45	15.09	294.12	

Table 28: Job Accessibility Evaluation Change for SunRail Project

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.

			2 0			2	
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	
Downtown	57355	17113.78	77.02	22.98	54.04	569.97	1.02
Ivanhoe	42301	11921.67	78.01	21.99	56.03	664.97	1.23
Altamonte	5830	14611.18	28.52	71.48	-42.96	868.53	

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

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

=						J
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

Table 30: Job Accessibility Change Evaluation for Juice Bikeshare Project

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.

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	
Outside Downtown Phase I	-11.76	-4.62	-7.14	529.41	-0.52
Phase II	-6.86	-2.95	-3.91	294.12	

Table 31: Commuting Travel Time Change Evaluation of SunRail Project

## (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.

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	
Downtown	-6.55	-2.50	-4.05	569.97	0.07
Ivanhoe	-8.31	-7.18	-1.13	664.97	-0.07
Altamonte	-7.07	-3.99	-3.08	868.53	

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

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

Tuble eet e	ommuting Huver Hink		of suree Bikebi	ure i roject
% Reduction in	% Reduction in Outside	% Reduction (DT -	Investment in	% Change/ \$100
Downtown	Downtown	Outside DT)	\$million	million
-1.00	-5.87	4.88	0.42	1161.9

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

## 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.

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

**Table 34:** Land Use Change for SunRail Project

## (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.

	Change of	Change of	0/2	%				
Region	Vacant Area in	Vacant Area in	Change	Change	%	Investment	% Change/	
Region	Case	Control	in Case	in	Change	in \$million	\$100 million	
(Acre/sqmile)Attraction23.62	(Acre/sqmile)	in Case	Control					
Attraction	23.62	21.49	52.36	47.64	4.72	773.53		
Downtown	31.66	20.96	60.17	39.83	20.34	569.97	0.07	
Ivanhoe	21.70	24.47	47.00	53.00	-6.01	664.97	0.97	
Altamonte	81.31	9.23	89.81	10.19	79.62	868.53		

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

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

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

**Table 36:** Land Use Change for Juice Bikeshare Project

## 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.

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	
Outside Downtown Phase I	1.01	-0.66	1.68	529.41	0.47
Phase II	11.24	0.87	10.37	294.12	

Table 37: Travel Pattern Change Evaluation of SunRail Project

## (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.

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	
Downtown	-4.61	-4.8	0.19	569.97	0.11
Ivanhoe	-5.56	-3.84	-1.72	664.97	-0.11
Altamonte	-7.38	3.44	-10.82	868.53	

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

1 abic 57.	Traver I attern Change Lve	iluation of Jul	ee Dikeshare 1 loj	
% Change of Bike/Walk Share in	% Change of Bike/Walk	% Change	Investment Differential in	% Change/ \$100
Downtown	Share in Outside Downtown	Control)	\$million	millions
6.19	-14.44	20.63	0.42	4911.9

**Table 39:** Travel Pattern Change Evaluation of Juice Bikeshare Project

## 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.

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

Table 40: Scoring and Ranking of the Projects

## 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 1<sup>st</sup> 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 <u>http://www.people.cecs.ucf.edu/neluru/CommunityBuilding.html</u>. 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<sup>\*</sup>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. <u>https://altaplanning.com/wp-content/uploads/Redmond-Final-Report.pdf</u> 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 Nonhierarchy 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 1<sup>st</sup> 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:
    - ✓ 1<sup>st</sup> 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))







Q Workshop - ArcMap		- 🗆 X
1 40 40 71 52 121 121 121 121 121 121 121 121 121		
File Edit View Bookmarks Insert Selection Geoprocessin	g Customize Windows Help	
🗋 🛃 🖨 🐇 🛞 🖄 🗙 🔊 🗠 🚸 • [11.380.017	Noject – – ×	
i 🔍 🔍 🏹 🥥 💥 🎎 🚸 🕸 - 💷 🖢 🕥 🥖 🐼 🚢	Intro & Dataset or East on Class	
Table Of Contents		~
Sta 🔁 \ominus 🚑 🖾	Total Conditional System System System 2	
😑 😅 Layers	NAD_1983_UTH_Zone_17N	
C/Users\bi869991\Documents\ArcGIS\Dv	Output Dataset or Feature Class	
Orange_Merge	C: Users (bid6999 1) Documents (WcG25) Default.gdb) Drange_Project	
CAliforn' Public (Bibbar) Community Built	Output Coordinate System	
Velusia		
	Vetical Intional	
	Connection Transformation (and see	
	serge que la maison maior aporte ay	
	1.02	
Grange	•	
	×	
C:Users:Public:Bibhasi/Community Build     Enur County	*	
	and and another and a	
😑 🧱 C/(Users\Public\Bibhas\Community Build	OK Cancel Environments Show hep >>	
Orange County	II So Package	
CALINATED Dashiel Ritchard Community Rolls	🗉 🗞 Photos	
Osceola County	Separation and Transformations	
	🗉 🎭 Raster	
C/(Users\Public\Bibhas\Community Build	Stelch Project	
Seminole County	Control Control Control Control Control	
Contract Robid Richard Community Robi	Create Spatial Reference	
Volusia County	5, Define Projection	
	Project	
	< >> >>	
		~
< > D000000		>
Geoprocessing tool that projects spatial data from one coordinate syste	m to another.	1015415.911 1626494.698 Feet

(c)



(d)

e Edit View Bookmarks Insert Selection Geoproc	essing Customize Windows Help			Spatial Reference Properties	×
2 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	< Project		E 3	KY Coordinate System 2 Coordinate System	
R. R. D. IIIIII + + + R-     R. A. M.       Table Of Contents     + ×       S: B. O. B. S.       If <b>Depen</b> If <b>Depen</b> If Contents 04800011 Documents ArcG00.04	Trout Dataset or Peeture Clean Oxenge Brout Coerdinate System (astronal) 1940_23823_JTM_2are171 Oxfoud Dataset or Peeture Clean		•	The Transformer in security of the	1
Orange,Marge     O	C ( Users (billions ) ( Dollar entry Write Output: Coordinate System	25'pefault.gdbildrange_Project1		Europe indonesia Melaysia	
Semicale     Semicale     Collower Public (Stehas) Community Bulk	Vehical spitrowals		•	Control C	*
		Of Care Benerations	Sheer Help > +	Unear Unit: Neter (1.0)	v OC Careat
Vetosis County	• •	C Project	*	-138	221.642 1678201.679 Feet

CALLED AN ALL ALL ALL ALL ALL ALL ALL ALL ALL				Soutial Reference Properties	×
e Edit View Bookmarks Insert Selection Geoproc	cessing Cuitomice Windows Hel	P		agent restrict regimes	· ^
	- Propert			< KY Coordinate System Z Coordinate System	
. S. C. O II II + + II - II k O / S	D the Strengt Catagories of Feature Class				
e Of Contents 🛛 🕈 🗶	Orange			D   Tuberan ermehr	
0.4 1	Srput Coordinate System (apti	onal)		Q NAD 1903 UTM Zone 12N	· •
J Layers	HAD_1003_UTH_2are_17H		14	NAD 1983 UTM Zone 13N	
Choenabiliteren Decumenta Avcisis De	Output Dataset or Peature Cla	***		A MAD 1981 UTM Zone 198	
	CI Weers Wildening V Document	ta Wroll Stefault gdb Krange "Projects	6	NAD 1983 UTM Zone 16N	
E C/Uters/Public/Bibhas/Community Buik	Output Coordinate System		1.000	O NAD 1981 UTM Zone 11N	
E C Volusia	1		E.		
in 171 Seminale	Vehcal Isshoral)			NAD 1983 UTM Zone 30N	
	Geographic Transformation (or	ntorial)		Current coordinate system:	
🗄 门 Osceola	L			NAD_1983_UTM_Zone_17N	A
			+	WXD): 26917 Authenty: (PS0	
			100	Projection Transverse Mercator False Fastivo: 500000.0	
Chibern Public Bibhan Community Build			8	False_Northing: 0.0	
III D Four_County	171.		· •	Scale_Factori 0.9996	
Childrent Dabled Balt and Community Daily		OK Cantel	Environments	Latitude_OF_Origin: 0.0 Linear Unit: Heter (1.0)	
III Crosers Public Beneric Community Bun				-	*
		() SPackage			
Cr(Durs\Public\Bibhac\Community Build		Projections and Te	and omations		
G Osceole County		···· Se Ratter			
C/User/Public/Bibbe/Community Build		S Batch Project	10000000000	~	Const
C1 Seminole County		Convert Coord	inate Notation		
		Create Spatial	leference		
Columnity Busic		S Define Project	on .		
		S Project			
		C Re Rather			

(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 (<u>ftp://sdrftp03.dor.state.fl.us/Tax%20Roll%20Data%20Files/</u>).
- 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.

Crange.s	av [DataSet1] -	IBM SPSS Statistic	s Data Editor													- 0	×
Eile Edit	⊻iew Data	Transform	Analyze Graphs	s <u>U</u> tilities E <u>x</u> te	nsions Wind	low Help											
🔁 🗄			• 🖹 📥	= 1	H M		4	1		6							
												_			Vis	ible: 126 of 126	Variables
	& CO_NO	da F	PARCEL_ID	FILE_T	ASMNT_ YR	& BAS_ST RT	& ATV_STR T	🗞 GRP_NC	I DOR_UC	PA_UC	SPASS_ CD	<i>₿</i> ₽ JV	JV_CHN G	JV_CHN G_CD	NV_SD	& AV_NSD	🖉 Т.
1	58	272001000000	139	R	2012	1	1	3	· · · · ·	0		403310			403310	403310	3 *
2	58	272001000000	140	R	2012	1	1	2	2	0		128618			128618	128618	1
3	58	272001000000	141	R	2012	1	1	1	1	1		114279			114279	114279	
4	58	272001000000	142	R	2012	5	8	5	99	0		41595			41595	41595	
5	58	272001000000	143	R	2012	1	1	2	1	1		134149			134149	134149	1
6	58	272001000000	146	R	2012	1	1	2	1	1		163818			162905	162905	1
7	58	272001000000	147	R	2012	5	8	1	99	0		117365			117365	117365	1
8	58	272001000000	148	R	2012	3	8	1	69	30		270346			215180	215180	2
9	58	272001000000	149	R	2012	1	1	3	1	1		239681			188871	188871	1
10	58	272001000000	150	R	2012	1	1	2	1	1		161051			143240	143240	1
11	58	272001000000	151	R	2012	1	1	2	2	0		141961			141961	141961	1
12	58	272001000000	152	R	2012	1	1	1	2	0		63643			50553	50553	
13	58	272001000000	153	R	2012	5	8	2	99	0		570172			570172	570172	5
14	58	272001000000	54	R	2012	1	1	3	1	0		268653			130216	130216	1
15	58	272001000000	155	R	2012	4	8	2	0	1		98464			98464	98464	
16	58	272001000000	156	R	2012	5	8	5	99	0		51343			51343	51343	
17	58	272001000000	57	R	2012	3	8	5	68	1		61036			11706	11706	
18	58	272001000000	58	R	2012	3	8	1	69	30		64027			33629	33629	
19	58	272002000000	)14	R	2012	3	8	1	61	0		605459			207274	207274	2
20	58	272002000000	15	R	2012	3	8	5	61	0		226494			3330	3330	
21	58	272002000000	18	R	2012	1	1	3	1	2		468595			468595	468595	4
22	58	272002000000	19	R	2012	1	1	1	2	0		91615			91615	91615	-
	4																F
Data View	Variable View																

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:
  - ✓ 1<sup>st</sup> 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





Q Workshop - ArcMap		- 0 X
前 前 門 茂 白 田 致   前 南 🚥   ~		
Eile Edit View Bookmarks Insert Celestia	Geographics Customics Windows Hele	Information Addad
The cur view bookmans inser selection	deoprocessing customize windows help	
	1:1.009,342 🗸 🔛 🗿 🔂 🎦 🎥	
0, 0, 2 0 11 12 + + 12 - 13 🖡	0/==:::::::::::::::::::::::::::::::::::	
Table Of Contents # x	Table	stfer loining NAL Hile
	12、13、14110111111	E E
E El Layers	Orange	×
C:\Users\Public\Bibhas\Community uild	NAL2012_PA NAL2012_JV NAL2012_DO Shape_Area	
□ Sunrail_Station_17	29223537720218 19646 1 6768.684851	Dancel Filet
		rarcer rue.
	29223637720215 20063 1 6769.881494	1) Dancel ID
Li Sunrail_Koute	29223537720213 27295 8 6770 6931	1) rarcel ID
-+	29223491681806 22265 1 3279.199883	2) Shana Anga
C:\Users\bi869991\Documents\ArcGIS\Delta	29223491681807 30285 8 6347.446709	2) Shape Area
Merge Counties	29223491681805 24308 8 3198.903255	
	29223487380001 2601573 87 314957.49268	+
C:\Users\Public\Bibhas\Community Built	29223491681804 20390 1 3195.974742	
🖹 🗌 Volusia	29223491081704 8898 0 3139.002703	
	29223491681701 100 80 399.952812	ŧ.
😑 🔲 Seminole	29223491081700 8898 0 2999.401418	Added Information from NAL Files
	29223491681809 5598 1 519.790822	Added Information from NAL File:
🖂 🔲 Osceola	29223491681810 11211 0 3511.024344	1) IV Bronster Value
	29223491681811 24480 1 3506.012507	1) JV - Froprty Value
Clip_Buffer	29223491081812 11211 0 3500.96/751	2) DOD UC Land Use Code
	29223491681814 2000 0 3490.903997	2) DOK UC - Land Use Code
C:\Users\Public\Bibhas\Community Build	29223491681815 43783 1 6966.685082	
□ Sunrail 1mile	29223491681709 33219 1 3466.983907	
	29223491681713 26678 1 3877.299103	
C:\Users\Public\Bibbas\Community Build	29223491681809 8898 0 2996.284997	
	29223491081508 14948 0 7312.439095	
	29223491691606 21932 1 3652.501668	
	29223491681504 11211 0 3647.562492	
	29223491681503 20885 1 3645.074062	
	29223491681502 11211 0 3642.597454	
	29223491681501 11211 80 3590.142752	
	29223491681607 79296 1 3686.815393	
	H28223491001009 11211 0 3007.008314	t •
	c	
	14 4 0 F FI 🔚 💻 (0 out of 46175 Select	ed)
	Orange	
(		
	haid to it the	
		-29757601.382 10478016.728 Feet

(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))









(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.

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

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

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))



- 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)



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] \ge 1 and [DOR\_UC] \le 9 Then
x="Residential"
elseif [DOR UC] >= 11 and [DOR UC] \leq= 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] \geq 85 and [DOR UC] \leq 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))







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).

🕲 Workshop - Archiap	- 🗆 ×
Ne Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help	
and of Contents	
times to the times	
Mege	
C/User/bill69991/Document/ArcGS/D	4
Coange, Merge     Search      Search      Search	(a)
AcTos Actos Geopticeumg todasets into a	
E Crubes/Vales/Behan/Community Balt 🛠 Environ single services output dataset.	(B)
T Revits	
🗉 🗆 Seminole 🛛 🔰 ModelBuilden	
Python	
III Oucesta Geoprocessing Options	
R Orana	
🔤 🛄 C/Uhen/Public/Bibhal/Community Buik	
E Four County	
- Children Publich Babbah Community Bab	
II Orange County	
🗉 😂 C/Wars/Public/Bibhar/Community Build	
I Dicess County	
🗉 🧰 C. Waters Public (Biblio) Community Build	
iii 🗋 Seminole County	
Critierri/Judic/Bellse/Community Built     Community Built	
C remark county	
3 00 0 H 4	
	722398.05 1906092.704 Feet

(a)

Q Workshop - ArcMap			- 0 X
File Edit View Backmade Incast Selection Generation	Customine Windows Male		
Det Data and the standard states	Costomet windows mep		
	- Merge -		
R R C O 11 11 + + H F O / D 2	Input Datasets	~	
Table Of Contents 7 ×		- 🛃 💷	â 📮
2: U 🗢 🕂 🖾	Crance .		
E California (1999) Decomposition (1999) Decompo	Oscesia		2
II CI Orange Merge	Seminole	×	
	Volutia	+	24
C:\Users\Public\Bibhas\Community Buik			14 J
		•	
Seminole			
Osceola	Culture Million methods (2006 all officiance Mene)	-	
	Endel Man (antional)		
	→ CETOTERO		
😑 🚞 C:\Usen\Public\Bibhas\Community Build	PARCELNO (Text)	· ·	
Four_County	8 Shape_Leng (Double)		
California Dabled Babard Community, Built	OK Cancel Environments	Show Help >>	
Croses (Patric Biords) Commany Ban     Diange County			
Cr\Users\Public\Bibhas\Community Build			
Osceola County			
CAUsers/Public/Bibbar).Community Built			
Seminole County			
C:\Users\Public\Bibhas\Community Build			
U Venus County			
_			
< > D00000 c			2
			876058.494 1199495.481 Feet

(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<sup>1</sup> 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)).

<sup>&</sup>lt;sup>1</sup>Downtown Stations are fixed based on the downtown area projected at 'I-4 Ultimate Project' construction map at <u>https://i4ultimate.com/construction-info/construction-map/#constructionAlerts</u>









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))





(c) Figure A.19: Case Area Selection

## A.2.4.2 Overlapping Problem

B 🗍 G C:\Users\bi869991\ C: ninole 

> inge C:\. C:\. Sun Us PI -

rnotos Projections and Transfu Batch Project Convert Coordinate Create Custom Geog Create Spatial Refere Define Projection Project

Relation 66

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.

29734888.511 10365914.567 Fee



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:

- 1<sup>st</sup> 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)







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)).

tClip_12_S	R_Works							-	- 0 X
Eile Edit	View Data Transform Analyze Graphs	s Utilities Expinsions Window	Help						
<b>a b</b>	Set Measurement Level for Linknown		-Z	1	A46				
	Copy Data Properties							Vis	sible: 7 of 7 Variables
	New Custom Attribute	a_Acris 🧳	near_fid 🔗 DOR_U	ic 🔗 JV 💰	🖉 LND_VAL 🛛 🚜 LAND_U	SE var	var	var var	var
1	29 2261	7.74 17393621	13 8	89 284841316	11553200 Public				-
3	29 2132 B Define Multiple Response Sets	57 2500971690	5 1	85 172416139	5749818 Public				
4	29 2356 🔡 Identify Duplicate Cases	6.77 47850783	7 8	85 135449956	5018842 Public				
5	29 2350 📆 Compare Datasets	4.40 85696933	7 8	85 121616537	3785502 Public				
6	29 226( 🗃 Sort Cases	4.5 55983124	13 1	18 84923372	12195214 Retail/Office		-		
8	29 2266 Sort Variables	8.04026546505	13 0	RA 72727738	20595881 Institutional				
9	29 302( Transpose	6.92252820451	7 8	68594899	2766210 Public				
10	29 226 Adjust String Widths Across Files	1.0862732396	13	3 59526955	3770100 MultiFamilyRes				
11	29 226: Merge Files	9.20056026986	6 8	87 59494492	8074740 Public				
12	21 529 Per Desiruciure	25.89 55115650	11	73 55474200	2823700 Institutional				
13	29 2251 Propensity Score Matching	3.04 (19850314	13	18 49038829	5699865 Retail/Office				
15	29 2234 Case Control Matching	5.65 94992805	6	43556616	7022571 Retail/Office				
16	29 2264 Aggregate	.67 123373077	6	18 40065637	1716046 Retail/Office				
17	29 235 R Copy Dataset	.8 137348701	13	18 39058089	2190212 Retail/Office				
18	29 201. 29 226 🗄 Split into Files	7 72 96023757	6 1	16 39031884 R3 34991215	6300156 Public				
20	29 213: 🚟 Split Eile	26.00121804510	5 1	89 32872982	2293280 Public				
21	29 226 🔛 Select Cases	1.29 21336672	13	18 32062186	10711250 Retail/Office				
22	29 225 1 Weight Cases	2.24 18404442	13 1	74 30233388	3381345 Institutional				-
23	29 235630602000	2.8-035547902	13 8	89 30160367 20100676	4964680 Public				
24	292235135402010	.56609974879	13	18 30047564	1528920 Retail/Office				
26	292226735222009	1.50042018718	13	18 29971515	4052196 Retail/Office				
27	292226735222010	1.60948164902	13	18 29427399	4346758 Retail/Office				-
	1	0.04700240440	12		CTOREDO Data NOTas				<b>N</b>
Aggregate	Variable View					IBM SPSS Statistics	Processor is read	y Unicode:0	DN
				(a)					
Eile Edit	R_Workshop.sav [DataSet2] - IBM SPSS Statistics Data View Data Transform Analyze Graphs	Editor Utilities Extensions Window	Help						o x
					ABC				
				1.0					
		Aggrega	le Data		× .		(	Visi	ible: 7 of 7 Variables
1	292226630501000	7 7471739		Break Variable(s):	AND_U	Estimat	a Tot	var var	var
2	292226621900010	7.5104543 Area_A	cres	A LAND_USE		Lounau	101	ai	1
3	292213278001000	57.2550097 🗞 near_fi				Land He	o Va	lue	
4	292235640300010	6.7774785 DOR_0	C Agg	uladii/asiablaa			n va	iuc	
5	292235019700010	4.4078569 A E166609	AL	Summaries of Var	lable(s):	and Acr	es hv	Land	
7	292226631700010	8.0482654		Area_Acres_sum	= SUM(Area_Acres)	and Acr	cs by	Lanu	
8	302205940063010	50.6368232				LIGE Tyn	ρ		
9	292302030000010	6.9225282				Use Typ	L		
10	292226133500020	1.0856273							
12	212529187400010010	25.8975511		Function	Mome 2 Label				
13	292226226300020	2.9214536	Mur	mber of <u>c</u> ases 1	lame: N_BREAK				
14	292225814600021	3.0471985 Save	6 <b>1</b> 175				1		
15	292223564001001	5.6559499 O Add ag	gregated variables to activ	e dataset					
16	292226509101000	.6792337 @ Create	a new dataset containing	only the aggregated	variables				
18	292201366401001	30.5418610 Dat	iset name: a						
19	292226160605011	7.7209602 O Write a	new data file containing o	nly the aggregated	rariables				
20	292213513200001	26.0642180 File	C:\Users\Public	Bibhas\Communit	y Building\Propert_\aggr.sav				
21	292226002700030	1.2942133	Nen Large Datasets						
22	292225630602000	2.2471840 Options fo	already sorted on break va	riable(s)					
24	292423000000004	96.3532069 Sort fil	e before aggregating						
25	292235135402010	5660997							
26	292226735222009	1.5004201	OK Paste	Reset Cancel	Help				
		1 887 1 1 1							
27	292226735222010	1.60948164902	13 1	8 29427399	4346758 Retail/Office				-
27	292226735222010 11	1.60948164902	13 1	8 29427399	4346758 Retail/Office				*
27 Data View	292226735222010 202225735226043 Variable View	1.60948164902	13 1	8 29427399	4346758 Retail/Office				×

					×				Visible	7 of 7 Variab
	ParcelNo	Area Acres	Aggregate Data		A	var	Var	var	var	var
1	292226630501000	7.7471739362		Break Variable(s):		Î.				
2	292226621900010	7.5104543558	ParcelNo	ALAND_USE						
3	292213278001000	57.2550097169	DOR UC	a hear_tid						
4	292235640300010	6.7774785078	JV V	- Appregated variables						
5	292235019700010	4.4078569693	S LND_VAL	Summariae of Variania/e1						
6	292226002700050	4.5155598312		JV sum = SUM(JV)						
7	292226631700010	8.0482654650		Area_Acres_sum = SUM(Area_Acres)						
8	302205940063010	50.6368232965								
9	292302030000010	6.9225282045								
10	292226133500020	1.0856273239								
11	292226308404010	9.2685602698								
12	212529187400010010	25.8975511565		EunctionName & Laber						
13	292226226300020	2.9214536312		Number of cases Name: N_BREAK						
14	292225814600021	3.0471985031								
15	292223564001001	5.6559499280	Save							
16	292226509101000	.6792337307	O Add aggregated var	ables to active dataset						
17	292235135402030	.8113734870	Orgate a new datas	et containing only the aggregated variables						
18	292201366401001	30.5418610658	Dataset name:	b						
19	292226160605011	7.7209602375	O Write a new data file	containing only the aggregated variables						
20	292213513200001	26.0642180451	THE C	1/ Isers\Public\Bibbas\CommunityBuilding\Propert_\an	orsav					
21	292226002700030	1.2942133667		and a second sec						
22	292225617101000	2.2471840444	- Options for Very Large I	Datasets						
23	292235630602000	2.8493554790	File is already sorte	d on break variable(s)						
24	29242300000004	96.3532069568	Sort file before aggr	egating						
25	292235135402010	.5660997487								
26	292226735222009	1.5004201871	OH	Paste Reset Cancel Help						
27	292226735222010	1.6094816490								
00	000000705000040	0.0450.051.0440	10	40 000000 070000 D 1 1000						18

(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)

	Single Family	Multi- Family	Retail/Offi	Industria	Institution
Station	Residential (USD)	Residential (USD)	ce (USD)	l (USD)	al (USD)
	Do	wntown 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	625 /09	474 380	1 150 111	110 080	1 /02 057
Blvd Station	023,405	474,380	1,133,111	415,005	1,492,097
	Phase-I Ou	tside Downtown Sta	tions		
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	295 864	373 609	820 133	120 185	653 5/8
Station	255,804	373,005	025,155	425,105	055,548
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	918 072	626 616	1 208 935	724 904	1 083 417
Health Village Station	510,072	020,020	1,200,000	/ _ 1,501	1,000,117
Sand Lake Road	456,825	363,302	405,738	256,050	280,571
Station					
	P	hase-II Stations			
DeLand Station	111,661	86,914	56,488	71,328	108,124
Meadow Woods	534,753	351,368	75,014	387,552	159,837
Station					
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

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

## 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 1<sup>st</sup> 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
  - 1<sup>st</sup> 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.





(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.

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)).<sup>2</sup>

- 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'

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



(a)







(d)



(f)

File Folle Many Backwords Jacob Calendary Conserva-		
File Edit View Bookmarks Insert Selection Geoproci	essing Customize Windows Help	
□ 🖆 🖬 👶 🛸 🖹 🏝 × 🔹 < ♦ • 14,310,898	✓ ビ 回 課 第 回 第 。 Network Analyst・ 配 → が、 居 回 2000 第 第 。	
역 역 전 @ 11111 수 수 1월 ~ 의 k @ 7 등 26 A : Shie Of Contents	New Network Dataset ×	
	The default connectivity settings for network datasets establish connectivity only at	
🗏 😅 Layers	coincident endpoints of line features during the build process.	
C\Users\Bibhas\Desktop\Community Buildin     maxspeed	If you want to use differencementarity settings, clock the connected than below. You can be due to connectivity settings now, or you can change them allot many	
- Children Bilder Deskton Community Buildin	Connectivity X	
Crosers/bionas/besktop/community buildin     Sunrail Route	Connectivity: Connectivity Constraints	
-	Source Connectivity Policy 1	
Suntair stations	maxspeed End Point V	
C:\Users\Bibhas\Documents\ArcGIS\Default.g	Any Vertex	
c2012_Com_Intersect_Statisti c2012_Com_Intersect_Statisti		
C:\Users\Bibhas\Desktop\Community Buildin		
□ 2012_Com		
Access 30 stat		
	Group Columns: 1 + Subtypes OK Cancel	
	<back next=""> Cancel</back>	
· · · · · · · · · · · · · · · · · · ·	- All and a second s	
		767353.974 3048749.489 Meters
	(g)	
	(8)	
Access Time New.mxd - ArcMap		- a x
Access_Time_New.mxd - ArcMap File Edit View Bookmarks Insert Selection Geoproc	ssing Customize Windows Help	- a x
Access_Time_New.mid - ArcMap     File Edit View Bookmarks Insert Selection Geoproce     De	tsing Customize Windows Help 」 近 回回回回 デッ Network Analyst - 回 点気留回 massored ND 、 同時間。	- a ×
Access_Time_Newmid - ArdMap File Edit View Bookmarks Insert Selection Geoproce D = = = = = = = = = = = = = = = = = = =	tssing Customize Windows Help 」::: 回口意図「テ」 Network Analyst・ほ・ため留口 maxsored ND 、密稿。 さぶ New Network Dataset X	- a x
Access_Time_New.mad - ArcMap File Edit View Bookmarks Insert Selection Geoproce B B B B S B S X P ∩ ( ⊕ - [14310398 Q Q D S III + + 18 + 0   ⊕ - [14310398 B C Contents + 18 + 0   € 0 / 0 III + 1 B C Contents + 18	ssing Customize Windows Help ン :: (回口目目目 : 「 」 「 」 「 」 「 」 「 」 「 」 「 」 「 」 「 」 「	Catalog 0 x
Access_Time_Newmud - ArcMap File Edit View Bookmarks Inset Selection Geoproce Selection (1+310.888 Selection (1+310.888) Selection (1+310.	ssing Customize Windows Help 기 :: : : :::::::::::::::::::::::::::::	Catalog = × $\phi \to \phi_{0} \oplus \phi_{0} \oplus \phi_{1} = +   \oplus   \psi_{0}   = -$
Access_Time_New.mid - ArcMap File Edit View Bookmarks Insert Selection Geoproc Company Company Compa	tsing Customize Windows Help : :::::::::::::::::::::::::::::::::::	Catalog • × • • • • • • • • • • • • • • • • •
Access_Time_New.mxd - ArcMap         File       Edit       View       Bookmarks       Insert       Selection       Geoproce         Image: Selection       Image: Selection<	tsding Customize Windows Help : :::::::::::::::::::::::::::::::::::	Catalog P × P + P & D (2) (2) (2) (2) (2) (2) (2) (2) (2) (2)
Access_Time_New.msd - ArcMap         File       Edit View Bookmarks Inset Selection Geoproce         Image: Second Secon	stong Customize Windows Help : :::::::::::::::::::::::::::::::::::	Catalog * * * * * * * * * * * * * * * * * * *
Access_Time_New.mad - ArcMap File Edit View Bookmarks Inset Selection Geoproce Comparison of the Selection	ssing Customize Windows Help St. The Rev Network Analyst · R · L · L · L · L · L · L · L · L · L	Catalog 3 × 4 + + + + + + + + + + + + + + + + + + +
Access_Time_Newmod - ArcMap File Edit View Bookmarks Inset Selection Geoproce Comparison of the Inset Selection Geoproce Inset Selection Geoproce	ssing Customize Windows Help ✓	Catalog • × × • • • • • • • • • • • • • • • • •
Access_Time_New.msd - ArcMap File Edit View Bookmarks Insert Selection Geoproc Edit View Bookmarks Insert Select	tsing Customize Windows Help : :: : : : : : : : : : : : : : : : : :	Catalog • × × • • • • • • • • • • • • • • • • • • •
Access_Time_Newsmid - ArcMap File Edit View Bookmarks Inset Selection Geoproc Community Community Community Selection Community Community Selection Community Selection Service Area Community Selection Service Area Community Selection Community Selectio	essing Customize Windows Help 1: I I I I I I I I I I I I I I I I I I I	Catalog • × × • • • • • • • • • • • • • • • • • • •
Access_Time_New.mxd - ArcMap File Edit View Bookmarks Inset Selection Geoproc Comparison of the Selection Geoproc Compa	essing Customize Windows Help State To reveal you like to model the elevation of your network features? Now Network Dataset Very Now Dataset Very State To model the elevation of your network features? None Using 2 Coordinate Values from Geometry Using 2 Coordinate Values from Geometry Coordinate Values from Geometry Co	Catalog * * * * * * * * * * * * * * * * * * *
Access_Time_Nev.mud - ArcMap File Edit View Bookmarks Inset Selection Geoproc Comparison of the selection Geoproc Comparison of the selection Geoproc Comparison of the selection of the sele	essing Customize Windows Help  Submit Retwork Analyst	Catalog * X Catalog * X * * * * * * * * * * * * * Loadie: @ maugesd-du Ladie: Codect * CAUsers/Bibhas Cutsers/Bibhas Cu
Access_Time_Nev.m.rd - ArcMap File Edit View Bookmarks Insert Selection Geoproc File Edit View Bookmarks Insert Selection Geoproc File Selection Geoproc File Selection Geoproc File Selection Geoproc File Selection Geoproc File Service Area Conserts Bibhas/Desktop/Community Buil File Children Bibhas/Desktop/Community Buil File Geoper Street Network File Street Network File Street Network File Street Network File Street Network File Street Network	essing Customize Windows Help : :: Retwork Analyst* :: +5 +2, # :: maxseed ND : New Network Dataset New Networ	Catalog *** ********************************
Access_Time_New.mad - ArcMap  File Edit View Bookmarks Inset Selection Geoproc  Comparison of the selection of the selection Geoproc  Comparison of the selection of t	essing Customize Windows Help State Contract Analyst: R+& R maxapeed ND Network Analyst: R+& R maxapeed ND New Network Dataset For would you like to model the elevation of your network features? None Contract Values from Geometry Using Coordinate Values from Geometry Coordinate Values from Geometry Coordinate Values from Geometry Coordinate Values from Geome	Catalog 3 × Catalog 3 × Ca
Access_Time_Nev.rmsd - ArcMap File Edit View Bookmarks Inset Selection Geoproc Selection Control (1310.898 Control (1310.898) Control (1310.89	essing Customize Windows Help Mit The Windows Help Mit The Wethork Analyst - I + + + I m masseed ND New Network Dataset New Network	Catalog
Access_Time_Nexcmod - ArcMap File Edit View Bookmarks Insert Selection Geoproc Edit View Insert View Insert Selection Geoproc Edit View Insert Selection Geoproc E	essing Customize Windows Help  set Retwork Analyst	Catalog • × × • • • • • • • • • • • • • • • • •
Access_Time_New.mud - ArcMap File Edit View Bookmarks Insert Selection Geoproc Comparison of the Selection Geoproc Selection Geoproc Comparison of the Selection Geoproc Comparison of the Sele	essing Customize Windows Help  string Customize Windows Help  with Provide Network Analyst* # +% # # maxageed ND  with New Network Dataset  New Network Dataset New Network Dataset  New Network Datas	Catalog *** **** ***************************
Access_Time_New.mxd - ArcMap  File Edit View Bookmarks Inset Selection Geoproc	essing Customize Windows Help State Control C	Catalog
Access_Time_Next.mid - ArtMap  File Edit View Bookmarks Insert Selection Geoproc	essing Customize Windows Help Mit The Wethork Analyst	Catalog
Access_Time_New.cm.d - ArcMap File Edit View Bookmarks Insert Selection Geoproc Characteristic (Community File) (Community File) Characteristic (Community File) (Community File) Community File) (Community File) (Community File) Community File) (Community File)	essing Customize Windows Help  Set Network Analyst: I + I + I = M B B A  Network Analyst: I + I + I = M B B A  Network Dataset  New Net	Catalog • * * • • • • • • • • • • • • • • • • • • •
Access_Time_New.msd - ArcMap  File Edit View Bookmarks Inset Selection Geoproc  Comparison of the selection of the selection Geoproc  Comparison of the selection of t	esing Customize Windows Help State The Network Analyst: R.4.5.F. maxageed ND Network Dataset Very Weldy Dataset Very Sector finate Values from Geometry Using 2C coordinate Values from Geometry Using Coordinate Values fr	Catalog ***
Access_Time_Neverned - ArcMap  File Edit View Bookmarks Inset Selection Geoproc	essing Customize Windows Help St. New Network Analyst: I I I I I I I I I I I I I I I I I I I	Catalog * * * * * * * * * * * * * * * * * * *
Access_Time_Nexcmad - ArcMap File Edit View Bookmarks Insert Selection Geoproc Selection Geoproc Selectio	esing Customize Windows Help Mit The Wethork Analyst	Catalog
Access_Time_New.cm.d - ArcMap File Edit View Bookmarks Inset Selection Geoproc Selection Geoproc Selectio	essing Customize Windows Help Metwork Analyst: I I I I I I I I I I I I I I I I I I I	Catalog *** Catalog **** Catalog **** Catalog **** Catalog **** Catalog **** Catalog **** Catalog ***** Catalog ***** Catalog ***** Catalog ************************************
Access, Time, New.mod - ArcMap  File Edit View Bookmarks Inset Selection Geoproc  Select	essing Customize Windows Help State Control C	Catalog       > ×         Ca
Access_Time_Neverned - ArcMap  File Edit View Bookmarks Inset Selection Geoproc  Selecti	esing Customize Windows Help Metwork Analyst: I I I I I I I I I I I I I I I I I I I	Catalog       ×         Image of the second seco
Access_Time_Next.mad - ArtMap  File Edit View Bookmarks Insert Selection Geoproc  Selecti	esing Customize Windows Help M To Default Network Analyst: I + 1 + 2 = 1 masseed ND Network Analyst: I + 1 + 2 = 1 masseed ND New Network Analyst: I + 1 + 2 = 1 masseed ND New Network Analyst: I + 1 + 2 = 1 masseed ND New Network Analyst: I + 1 + 2 = 1 masseed ND New Network Analyst: I + 1 + 2 = 1 masseed ND New Network Analyst: I + 1 + 2 = 1 masseed ND New Network Analyst: I + 1 + 2 = 1 masseed ND New Network Analyst: I + 1 + 2 = 1 masseed ND New Network Analyst: I + 1 + 2 = 1 masseed ND New Network Analyst: I + 1 + 2 = 1 masseed ND New Network Analyst: I + 1 + 2 = 1 +	Catalog       ***         Catalog       ****         Catalog       ************************************
Access, Time, Newmad - ArcMap  Ric Edit View, Bookmarks Inset Selection Geoproc  Solution  Solu	story esting Customize Windows Help Wetwork Analyst:	Catalog       ***         Candidag       **** <t< td=""></t<>

(h)



(j)



(1)



(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<sup>3</sup>.

- 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

<sup>&</sup>lt;sup>3</sup> Create drive time areas in ArcGIS (<u>https://doc.arcgis.com/en/arcgis-online/analyze/create-drive-time-areas.htm</u>)


(a)



(b)



(d)





(f)



(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).

🗧 HP 11.8" Chromebook Intel Cele: 😠   🎽 Getting started — Twitter Devel: 🗴   😄 project/social media	tutorial/Tuti 🗴 🧧 Social_Media_Data_Analysis - Jut 🗴 🔘 Anaconda   Individual E	idion x +	– a ×
← → C i anaconda.com/products/individual			🖈 🚷 i
📰 Apps 📒 order 📃 CASH BACK 📃 python 🌘 University of Centra 💶 YouTube 🍡 Google	Translate 💁 Mail - Jiechao Zhen 📃 social 📒 bank 📙 logit model 📒 Ti	A material 🐧 https://www.sprinta 🖪 mobility 📙 vissim 📒 map 📒 mosqi 📃 sumo	Cther bookmarks
	Anaconda Installe	rs	
Windows 🕊	MacOS 🗯	Linux 🛆	
Python 3.7	Python 3.7	Python 3.7	
64-Bit Graphical Installer (466 M	8) 64-Bit Graphical Installer (442 MB)	64-Bit (x86) Installer (522 MB)	
32-Bit Graphical Installer (423 MI	<ol> <li>64-Bit Command Line Installer (430 MB)</li> </ol>	64-Bit (Power8 and Power9) Installer (276 MB)	
Python 2.7	Python 2.7		
64-Bit Graphical Installer (413 MI	8) 64-Bit Graphical Installer (637 MB)	Python 2.7	
32-Bit Graphical Installer (356 MI	8) 64-Bit Command Line Installer (409 MB)	64-Bit (x86) Installer (477 MB)	
		64-Bit (Power8 and Power9) Installer (295 MB)	
This website uses cookies to ensure	you get the best experience on our website. Privacy Policy	Accept	
● ■ ② 第 申 ○ へ ■			10:51 PM 7/21/2020

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**.

	Applications on tase (root)	* Chaineis		1		
ments	CADE use Prompt END and the Prompt Burk of these detrimined with your current	An extended according of the provided and the the the provided according to the provided and the the the provided according to the provided accordin	Natabook 613 web-base intercolor computing material encorporation of an or on	Powenheil Prompt Est 1 Run & Powenheil Cerminal with your corpor enconcence from Instrument	Clearle Glavel 1932 Muttermensional data visualization extra	Crange 3 335 Component based dos maring framework. Then subsequence for All Advances for
ity	Land	Jupyter Notebook and Architecture.	human-readable docs while describing the data analysis.	activated	emorg related datasets.	novice and expert. Interactive workflows with a large toolbox.
	CC Concele 4:15 Pyde Contrace topocots With refluence registry of the Conceleration of the Conceleration registry of the Conceleration of the Conceleration registry of the Conceleration of the Conceleration of the Conceleration registry of the Conceleration of	Estudio 1186 Ast of Integrated Iosis designed to help plue be more enabled waith R instalates R exemutia and notabooks.	Spyder 4.3 Solentifik Prinhon Development bruitenment, Flowerful Python 108 with advenad editing, interactive taxing, debugging advinceopacition features	-		
Ration	Instit	man	Integri			

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**.

	💭 jupyter			Quit Logo
s	Files Running Clusters			Upload New -
	0 • h /			Name 🍁 cast Modified 🛛 File st
	Contacts			3 days ago 3 days ago
Open ← → × ↑	« social media tutorial » Tutorial » Tutorial	v ŏ Search Tutoria		23 days ago 3 days ago
Organize 💌 N	lew folder	0:1	• • •	3 days ago
<ul> <li>This PC</li> <li>3D Objects</li> <li>Desktop</li> <li>Documents</li> </ul>	Name     Jpynb_checkpoints     teoic_anabyis_sec     Social_Media_Data_Analysis     Social_Media_Data_Analysis_instruction	Date modified 7/15/2020 10:54 PM 6/12/2020 2:22 PM 7/21/2020 5:56 PM 2/13/2020 12:16 AM	Type File folder File folder IPYNB File Microsoft W	3 days ago 3 days ago 23 days ago 3 days ago
<ul> <li>Downloads</li> <li>Music</li> <li>Pictures</li> <li>Videos</li> </ul>	topic_analysis_src.rar	IPYNB fi	lle	3 days ago 3 days ago 23 days ago 55
<ul> <li>Local Disk (C</li> <li>New Volume</li> <li></li></ul>	2.) (D:) )		,	13 days ago 16.7
240010	File name:	✓ All Files	~ Cancel	

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.

🔲 HP 11.8" Chromebook Intel Celo: 🗙 🎽 Getting started — Twitter Develo: 🛪 🍐	😋 project/social media tutorial/Tuti: 🗙 📔 🤌 Social_Media_Data_Analysis - Ju	: 🗙 🎽 Apply for access – Twitter Develo 🗙 🛨			– a ×
← → C @ developer.twitter.com/en/apply-for-access					\$ 💮 i
🔠 Apps 🦲 order 📙 CASHBACK 🦲 python 🐞 University of Centra 🕊	🛢 YouTube 🧕 Google Translate 🧕 Mail - Jechao Zhan 📒 social	🛿 bank 📙 logit model 📒 14 material 🐧 https:	//www.sprinta 📙 mobility 📙 vissim 📕	map 🧧 masqi 📙 sumo	Cher bookmarks
🈏 Developer Use cases Products				Apply Apps Q 🧕	
Get starte	d with Twitter APIs and tools				
Δη	nly for access				
All new develo	opers must apply for a developer account to access	Twitter APIs.			
Apply for a de	Restricted use cases >				
Standard	APIs Premium APIs	Enterprise APIs	Ads APIs		
Our free, standa	rd APIs are great for Our premium APIs offer scalable	Our enterprise APIs offer the highest	The Ads API gives partners a		
getting started, t	testing an integration, access to Twitter data for those	level of access and reliability to those	programmatic way to integrate with		
or validating a or	oncept. looking to grow, experiment, and innovate.	who depend on Twitter data.	the Twitter Ads platform.		
Apply for a deve	eloper account >	Apply for enterprise access >	Apply for Ads APIs access >		
	Apply for premium access >				
🔳 🔎 O 🖽 🐂 🌍 🔳 🥥 🔿					12:41 AM 7/22/2020

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.

Edit	View Insert Cell Kernel Widgets Help	Trusted Python 3
+ * 4	I h + H Run I C H Code	
	1 Data Collection (user accounts) ¶	Necessary Packages
In [10]:	#l/usr/bin/env python # encoding: utf-8	
	<pre>import tweepy # https://github.com/tweepy/tweepy import csv, pdb import time</pre>	
	<pre># Twitter API credentiols twitter_app_auth = {     consumer_key':         'consumer_secret         'access_token':         'access_token':         'access_token':         'access_token':         ////         ///         ///</pre>	Path - List_User
	List_User = r'D:\project\social media tutorial\example data\List_User.csv	Path – Save File
	<pre>def get_all_tweets(screen_name):</pre>	
	# Twitter only allows access to a users most recent 3240 tweets with	this method

Figure B.5 Example of User Account Data Collection Code

AutoSave 💽 🛱 🦻	G		List_User - Excel	🛕 złyjsl123@gmail.com 😰
File Home Insert Pa	ige Layout Formulas Da	ta Review View Help 🔎 Search		6
Paste S Format Painter	I U - ⊞ - A A ·	= = = ↔ . ﷺ Wrap Text = = = = = = = ■ Merge & Center -	General - \$ - % 9 1% 48 Conditional Format as Cell Formatting - Table - Styles	↓     ↓ </td
Clipboard Fs	Font G	Alignment G	Number 15 Styles	Cells Editing Ideas
110 • I × ~	fx fx	L	ist_User File Exa	ample
A B	C D	F C	7 11 6 11 11	R S T
1 FL511_Estatal 2 321Transit				
3 965traffic				
4 BikeWalkCFL				
5 fl_511_i4 6 Fl 511_95Express				
7 fl511_central				
8 fl511_i10				
AutoSave Core Insart Page L File Home Insart Page L Page Copy - Copeand Painter Copeand Fil	yout Formulas Data Revi $-11 - A^* A^* = \equiv   $ $l -    ⊕ +    ⊕ + A - r_{r_0}  ≡ ≡$	Alexander Spring, Deco       ew     View       Help     Ø Search       Image: Spring Deco     Image: Spring Deco       Image: Spring Deco     Image: Spring Deco	enter 66,11 Jarets - Doct	A dept323/geneticen
A1 * I X V J	fr 6/10/2020 6:49:32 PM			
A B C	D E F	G H I J K BrestRass 1.27Fa18 1.27Fa18 3.09Fa09 F	L M N O P Q 4155 F4155 Twitter Wittker( ania TRUE F4155	R S T U V W A
2				
3 ######### b'RT @Hic spring_art. 4	0 10 1.27E+18		ALSE FALSE Twitter for User[_api+_FALSE _FALSE	######## IKUE 333333 ['hashtags' b'\xe2\x8C 2.
5 ######## b'Tailorbyr BestShopp	0 0 1.27E+18	TT_ A	Det E	(hashtags'b" 1.
7 ######## b'.@AmCa AmCaritas	0 0 1.27E+18	User A	ccounts Data Fil	e Example (hashtags' b'Care is tf 1.
8 9 88888888 5'97 @caf 84137688	0 9 1 276+18			Phashtam NyfOygP 1
10	5 9 1.2/(+16			maxings o woodshirt.
11 ######## b'RT @wa Spring_Soc 12	0 540 1.27E+18			['hashtags' b"\xf0\x9f 1.
13 ######## b'@Alexar AmiesPhili	0 0 1 275+18	AmiesPhili 1.27E		Thashtags b'Naturalis 8.
	0 0 21676-20			
15 ######## b'RT @dio lau_lfrchi	0 418 1.27E+18		ALSE FALSE Twitter for User(_api= FALSE FALSE	auanaaa FALSE 333333 (hashtags b'Plus vif q 2.
14 15 ####################################	0 418 1.27E+18		ALSE FALSE Twitter for User(_api+ FALSE FALSE ALSE FALSE Twitter for User(_api+ FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE TWITTER for User(_api+ FALSE FALS	REINERER FALSE 333333 (hashtagi b'Plus vif q 2,

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 'O.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**.

<pre>tupport tweetyy import civ from tweetyy import cov from tweetyy import cov from tweetyy import Stream is consumer_key = c</pre>
<pre>consumer_key = consumer_key = consumer_key.consumer_key, consumer_secret) access_token_secret = access_token_access_token_access_token_secret) auth=steepy.OAttMindler(consumer_key, consumer_secret) auth=steepy.OAttMindler(cons</pre>
<pre>auth = tweepy.OAuthHandler(consumer_key, consumer_secret) auth.set_access_token(access_token, access_token_secret) auth.set_access_token(access_token, access_token_secret) auth.set_access_token(access_token, access_token_secret) auth.set_access_token(access_token, access_token_secret) auth.set_access_token(access_token, access_token_secret) auth.set_access_token(access_token, access_token, access_tok</pre>
<pre># Open/Create a file to append data Resystile = open(Xs_tweets.csv' X == q, 'w', encoding='utf-8') Rule csv writer Resystile = csv.writer(csvFile) List_User = r'D:\project\social media tutorial\example data\0. List_KW.csv' #define the path for line in open(List_User, 'r', encoding='utf-8'):</pre>
<pre>List_User = r'D:\project\social media tutorial\example data\0. List_KX.csv' #define the path for line in open(List_User, 'r', encoding='utf-8'):</pre>
<pre>for line in open((ist_User, 'r', encoding='utf-8'):     keyword = line.strip()     save_nath = r'D:\project\social media tutorial\example data\key word\%s_December_06_11_tweets.csv' % keyword #define the</pre>
(a) rta_Analysis-As: X
(a) ata_Analysis-Ar; X Mail-HechaoZhang-Outlook X + tria_Data_Analysis:pynD wersky of Centra. VorTube & Google Translate & Mail-JacchaoZhan is social is bark is logit model is Ta material is https://www.spinta is mobility is vision jupyter Social_Media_Data_Analysis Last Checkpoint a day ago (autosaved) Ele Edit View Inset Call Kernel Widgets Help Turtled > Python3 HazeTrafficOrl completed total success- 25 Data Collection (keywords) In [8]: Import tweey: from tweey: Import Outlinhandler from tweey: Streaming Import Streamilistener # consumer_scret = auth =: tweey.Outlinhandler(consumer_key, consumer_secret) auth.set_access_token_access_token_secret) auth.set_access_token(access_token_access_token_secret) auth.set_access_token(access_token_access_token_secret) auth.set_access_token(access_token_access_token_secret) auth.set_access_token_access_token_access_token_secret) auth.set_access_token_access_token_access_token_secret) auth.set_access_token_access_token_secret) auth.set_access_token_access_token_secret
(d) tra_Analysis _h; x  Mail-Jierbao Zhang-Outlook x + dia_Data_Analysis.pynob hereity of Certs • Voulle  Google Translate  Mail-Jierbao Zhan • social  bark • logt model • TA material • https://www.spirita • mobility • vision j jupyter Social_Media_Data_Analysis Last Checkpoint e day ago (autosaved)
ta_Analysis-Ja; x       Mail-Jiecha Zhang-Outook       x       +         fai_Data_Analysis/pynb <ul> <li>writy of Centra.</li> <li>foul/use</li> <li>Google Tanslate</li> <li>Mail-Jiecha Zhan,</li> <li>social</li> <li>bank</li> <li>logit model</li> <li>Ta material</li> <li>http://www.sprinta.</li> <li>mobility</li> <li>vision</li> <li>jupyter</li> <li>Social_Media_Data_Analysis Last Checkpoint a day age (autosaved)</li> <li>Logo</li> <li>Logo</li> <li>Logo</li> <li>Logo</li> <li>NazeTrafficOr1</li> <li>completed</li> <li>total success- 25</li> <li><b>2 Data Collection (keywords)</b></li> <li>in [8]:</li> <li>inpert Corr</li> <li>from tweepy inpert Stream</li> <li>from tweepy.streaming import Stream</li> <li>from tweepy.streaming import Stream</li> <li>from tweepy.streaming import Stream</li> <li>g cocss_token_secret = ''</li> <li>auth = tweepy.OuthHandler(consumer_key, consumer_secret)</li> <li>auth = tweepy.Markandler(consumer_key, consumer_secret)</li></ul>
Jal_Data_Analysis/pynb versity of Centra. ● YouTube ● Google Tanslate ● Mal-Jeckao Zhan. ● ocial ● bark ● logit model ● Ta material ● https://www.sprinta. ● mobiliy ● vision ● Jupyter Social_Media_Data_Analysis Last Checkpoint a day ago (autosaved) ● Logo ■ ● ● ■ ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ●
<pre>werking of Centra.</pre>
> jupyter Social_Media_Data_Analysis Last Checkpoint a day ago (autosaved)         Logo
File       Edit View Inset Cell Karnel Widgets Help       Trusted Python3         D + B: O C       + HRun B C H Code < ID         NazeTraffich-1 completed total success- 25         Data Collection (keywords)         In [8]:       import tweepy import cw from tweepy import Stream from tweepy.streaming import Stream.         # consumer_key = '' # access_token_secret = '' auth = tweepy.oluthHandler(consumer_key, consumer_secret) auth.set_access_token(access_token_access_token_secret) auth.set_access_token(access_token_access_token_secret)
Pie       Ext       View       Instant       View       Net       View       Python 3         B       ★       ★       W Run       ■       C       W Code       ✓       W         MazeTraffiCO1       completed       total success-25       2       Data Collection (keywords)         In [8]:       import tweepy       import Stream       From tweepy import OutHandler         from tweepy import Couthwardler       from tweepy import Stream       from tweepy import Stream         from tweepy import Couthwardler       #       occess_token = -''       #         #       access_token_secret = -''       #       access_token_secret;       auth = tweepy.olduthHandler(consumer_key, consumer_secret)         auth = tweepy.olduthHandler(consumer_key, consumer_secret)       auth = tweepy.olduthHandler(consumer_key, consumer_secret)       auth = tweepy.olduthHandler(consumer_key, consumer_secret)         auth = tweepy.olduthHandler(consumer_key, consumer_secret)       auth = tweepy.olduthHandler(consumer_key, consumer_secret)       auth = tweepy.olduthHandler(consumer_key = from tweepy)
<pre>Water and the secret = '' auth = treey.OuthHandler(consumer_key, consumer_secret) auth = set access_token_access_token_access_token_secret) auth = set access_token_access_token_access_token_secret auth = set access_token_access_token_access_token_secret) auth = set access_token_access_token_access_token_secret auth = set access_token_access_token_access_token_secret auth = set access_token_access_token_secret auth = set access_token_access_token_secret auth = set access_token_access_token_secret auth = set access_token_secret auth = set access_token_secret auth = set access_token_access_token_secret auth = set access_token_secret auth = set acc</pre>
<pre>Data Collection (keywords)  In [8]: import twepy import csv from tweey import OdutHandler from tweey.inport OdutHandler from tweey.inport Streamlistener # consumer_key = "' # access_token = -" auth = tweey.OdutHandler(consumer_key, consumer_secret) auth.set_access_token(access_token, access_token_secret) auth.set_access_token(access_token, access_token_secret) auth.set_access_token(access_token, access_token_secret)</pre>
<pre>2 Data Collection (keywords)  In [8]: import twepy import csv from twepy import OutHMandler from twepy.streaming import StreamListener</pre>
<pre>In [8]: import twepy import csv from twepy import OutHHandler from twepy.streaming import StreamListener # consume_key - '' # consume_secret = '' # access_token_secret = '' auth = twepy.OutHHandler(consume_key, consume_secret) auth.set_access_token(access_token, access_token_secret) auth.set_access_token(access_token, access_token_secret) auth.set_access_token(access_token, access_token_secret)</pre>
<pre>in [s]: import tweepy import tweepy import Stream from tweepy import OutHHandler from tweepy.streaming import StreamListener # consume_key = '' # consume_secret = '' # access_token_secret = '' auth = tweepy.OutHHandler(consume_key, consume_secret) auth.set_access_token(access_token,access_token_secret) auth.set_access_token(access_token,access_token_secret) auth.set_access_token(access_token,access_token_secret)</pre>
<pre>from tweepy import Stream from tweepy.streamIng import StreamListener  # consume_key = '' # consume_key = '' # access_token = c'' auth = tweepy.OAutHMandler(consumer_key, consumer_secret) auth.set_access_token(access_token, access_token_secret) auth.set_access_token(access_token, access_token, access_token_secret) auth.set_access_token(access_token, access_token, acces</pre>
<pre>from tweepy.streamIng import StreamListener  # consumer_key = '' # consumer_scret = '' # access_token_scret = '' # access_token_scret = '' auth = tweepy.OAutHMandler(consumer_key, consumer_secret) auth.set_access_token(access_token, access_token_secret) auth.set_access_token(access_token, access_token_secret)</pre>
<pre># consumer_key = '' # consumer_secret = '' # access_token = c'' # access_token_secret = '' auth = tweepy.OAutHMandler(consumer_key, consumer_secret) auth.set_access_token(access_token, access_token_secret) apl = tweepy.API(auth)</pre>
<pre># access_token = '' # access_token_secret = '' auth = tweepy.OdutHiandler(consumer_key, consumer_secret) auth.set_access_token(access_token, access_token_secret) ap1 = tweepy.API(auth)</pre>
<pre>auth = tweepy.OAuthHandler(consumer_key, consumer_secret) auth.set_access_token(access_token, access_token_secret) api = tweepy.API(auth)</pre>
# Open/Create a file to append data #csvfile = open(%z_g_ueets.csv % = q, `w', encoding='utf-8') #Use csv writer
List_User = r'D:\project\social media tutorial\example data\0.List_K0.csv'
for line in open(List_User,'r', encoding-'utf-8'): Time Period
<pre>save_path = n'D:\project\social media tutorial\example data(to)lection_data\%s_August_19_18_tweets.csv' % keyword <pre>save_path = n'D:\project\social media tutorial\example data(to)lection_data\%s_August_19_18_tweets.csv' % keyword </pre></pre>
csvile - opensate party, w, encourse (resp) csviviter = csv.writer(csvFile) try:
for twee in tweey.Cursor(spi.search.q.(keywdro], count-100,\ lang="en", since="2020-06-00", until="2020-06-13").items(): #can be used for upto 11 Day
<pre>print ("Rufaing</pre>
<pre>tweet.in_reply_to_screen_name, tweet.in_reply_to_status[d, tweet.in_reply_to_status[d]str, tweet.in_reply tweet.retweeted, tweet.source, tweet.user, tweet.truncated,tweet.user.contributors_mabled,tweet.user.crt tweet.user.time_rone, tweet.user.profile_text_color,tweet.entities,tweet.user.description.encode('utf-8')</pre>

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 O.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**).



(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.

	5 Sentiment Analysis Visualization
]:	import pandas as pd import matplotlib.pyplot as plt import numpy as np
	<pre>path = n'f:\sentiment result\final_data\whole result\sunshine skymay.csv' df = pd_read_csv(path, header = 0, names =  'id', time', text', 'account', peotagned', sentiment', polarity', 'subjectivity']) output_path = n'f:\sentiment result\final_data\whole result\sunshine skyway.png'</pre>
	df.time = pd.to_datetime(df.time) #select data based on the time (half year) df_1 = df[(df.time.dt.year == 2017)&(df.time.dt.month>1)&(df.time.dt.month<8)] df_2 = df[(df.time.dt.year == 2017)&(df.time.dt.month>7)&(df.time.dt.month<1)] df_3 = df[(df.time.dt.year == 2018)&(df.time.dt.month>0)&(df.time.dt.month<9)]
	fig, axes = plt.subplots(3, 1, sharex=True, sharey=True) fig.set_size_inches(5,10) Set up the time periods
	axes[0].stt[d['Enpotability'] eensity' = 1, bins=40, color= r',) axes[0].stt[d['February 2017 - July 2017'] axes[1].stt[d['August 2017 - December 2017] axes[1].stt]d['August 2017 - December 2017] axes[1].stt]d['August 2017 - December 2017] axes[1].stt]d['August 2017 - December 2017]

# 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

Normal New Propage Control A					
Product         Date         Date <thdate< th="">         Date         Date         &lt;</thdate<>	AntoSale (8.09) 🔠 19 - (2 - 4		florida bus - Excel	🖉 🗿 florida bus - Notepad	- 🗆 🗙
A       C       F       C       F       C       F       C       F       C       F       C       C       F       C       C       F       C       C       F       C       C       F       C	File Home Insert Page Layout Form	ulas Data Review View Help 🔎 Searc	h	File Edit Format View Help	
Al C D C D C T T C H I J K L M 1 2 0 7 restance pathy subjectively 2 7 restance pathy subjectively 3 8 7 2 0 7 restance pathy subjectively 3 9 restance of C M I J K L M 3 0 7 restance pathy subjectively 4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	A Cut         Calibri         11           Basic         Grownat Plainter         B         I         U         -         H           Claboard         G         Font         Font         -         Set         -         <		General - \$ - % \$ \$ \$ Condition Formatti	1.0309664470679512e+18 "RT ChrisKin 1.0309663678377124e+18 "RT ChrisKin 1.030964885092528e+18 "RT ChrisKing 1.0309567547865496e+18 "RT joannefe	gFL Everywhere we go in ^ gFL Everywhere we go in FL Everywhere we go in a Awesome bus tour educ
A b c b r restrict of the second s	A1 - 1 × √ fr			1.0309502900049950e+18 KT hammelli	I navenxezx80x99t said
A         B         C         D         F         G         H         J         L         1000305302240660810=18<" "subtle hint from the Panama city;           1         3         2         7         return polenty subtle hint from the Panama city;           1         3         2         7         return polenty subtle hint from the Panama city;           1         3         2         7         return polenty subtle hint from the Panama city;           1         3         3         3         1         3				1.0309533215994470+18 "RT GwenGraha	= South Florida friends
A         B         C         D         E         F         G         H         J         K         L           1         1         2         0         7         reserved in the statement of 0         0         0         1         0 <td< td=""><td></td><td></td><td></td><td>1 03005303234066010+10 "Subtle hint</td><td>from the Danama CityEl</td></td<>				1 03005303234066010+10 "Subtle hint	from the Danama CityEl
1         2         6         7 externme platform independence         1.0399316378610502743499418 W1 Chr13k1.Bgt. EveryWhite's use (5) 1.03993163786104418 W1 Chr13k1.Bgt. EveryWhite's use (5) 1.039934393904384418 W1 Chr13k1.Bgt. EveryWhite's use (5) 1.0399343934042676418 W1 Chr13k1.Bgt. EveryWhite's use (5) 1.0399343934042676418 W1 Chr13k1.Bgt. EveryWhite's use (5) 1.039934394042618418 W1 Chr13k1.Bgt. EveryWhite's use (5) 1.03993439404218418 W1 Chr13k1.Bgt. EveryWhite's use (5) 1.03993443940418418 W1 Chr13k1.Bgt. EveryWhite's use (5) 1.03993443940418418 W1 Chr13k1.Bgt. EveryWhite's use (5) 1.03993443940418418 W1 Chr13k1.Bgt. EveryWhite's use (5) 1.0399344941848 W1 Chr13k1.Bgt. EveryWhite's use (5) 1.03993443940418418 W1 Chr13k1.Bgt. EveryWhite's use (5) 1.03993443940418418 W1 Chr13k1.Bgt. EveryWhite's use (5) 1.039934484184 W1 Chr13k1.Bgt. EveryWhite's use (5) 1.0399344848 W1 Chr13k1.Bgt. EveryWhite's use (	A B C D	E F G H I	J K L M	1.03093392324900810418 Subtle Hillt	Trom che Panama cicyri
2         0         ####### Wit Ches 1001-18         isotherm         0         0.005           1         ####### Wit Ches 1001-18         isotherm         0         0.005           4         ####### Wit Ches 1001-18         isotherm         0         0.005           5         ####### Wit Ches 1001-18         isotherm         0         0.005           7         ####### Wit Ches 1001-18         isotherm         0         0           7         ######## Wit Ches 1001-18         isotherm         0         0           7         ####### Wit Ches 1001-18         isotherm         0         0         0           8         ########## Wit Ches 1001-18         isotherm         0         0         0           11         ####################################	1 1 2 6	7 sentiment polarity subjectively		1.0309516505743809e+18 KT Chriskin	gFL Everywhere we go in
3         1         ressent with Conk 1.001-18         bestment         0         0.0025           2         2         bestment         0         0.0025         1.0309409909052670338841-18         "Wtr Christing": L Everywhere we go in 1.030940990909528441-18         "Wtr Christing": L Everywhere we go in 1.03094099090954141-18         "Wtr Christing": L Everywhere we go in 1.030940990909541418         "Wtr Christing": L Everywhere we go in 1.03094099090954118         "Wtr Christing": L Everywhere we go in 1.0309409909084118         "Wtr Christing": L Everywhere we go in 1.0309409909084118         "Wtr Christing": L Everywhere we go in 1.030940990909318         "Wtr Christing": L Everywhere we go in 1.030940990909318         "Wtr Christing": L Everywhere we go in 1.030940990919318         "Wtr Christing": L Everywherewe everywhere we go in 1.030940990919318         <	2 0 ######## bRT ChrisK 1.03E+18	Sentimenti 0 0.0625		1.0309511678561894e+18 "RT GwenGrah	am Our South Florida Ea
4         2         3	3 1 ######### bRT ChrisK 1.03E+18	Sentiment 0 0.0625		1.0309509267633888e+18 "RT ChrisKin	gFL Everywhere we go in
5         1         Sexesses Withows 101:18         Sexesses Withows	4 2 ######### bRT ChrisK 1.03E+18	Sentiment: 0 0.0625		1.0309498999935384e+18 "Everywhere	we go in Southwest Flor
0         1         4         4         4         4         4         4         4         1.0309361350823906841e+18         "RT CissydJudge Cissy is very prou           0         4         5         5         5         5         5         5         5         5         5         1.0309361350823906841e+18         "RT CissydJudge Cissy is very prou           0         6	5 3 ######### bRT joanni 1.03E+18	Sentiment 0.519048 0.590476		1 0309494540426076e+18 "BT joannefe	a Awesome bus tour educ
7         5         Ferressent Decidion 1. 101-18         Sectioner         0.23         0.333333           8         6         6         6         6         0	5 4 ######## bRT hamm 1.03E+18	Sentiment 0 0		1.03004533140600410-10 "0T Ciscuta	dae Ciesu is yeary proud
Image: Setting of Control 1001-10         Setting of Set	7 5 ######## brealDona 1.03E+18	Sentiment 0.25 0.333333		1.03094323140008410+18 KT C1559430	uge crossy is very proud
9 7 research Under N. 101-18 testiment 0 0 0005 1 8 research Under N. 101-18 testiment 0 0 0005 1 8 research Under N. 101-18 testiment 0 0 0005 1 1 8 research Under N. 101-18 testiment 0 0 0005 1 1 8 research Under N. 101-18 testiment 0 0 0005 1 1 8 research Under N. 101-18 testiment 0 0 0005 1 1 8 research Under N. 101-18 testiment 0 0 0005 1 1 8 research Under N. 101-18 testiment 0 0 0005 1 1 8 research Under N. 101-18 testiment 0 0 0005 1 1 8 research Under N. 101-18 testiment 0 0 0005 1 1 8 research Under N. 101-18 testiment 0 0 0005 1 1 8 research Under N. 101-18 testiment 0 0 0005 1 1 8 research Under N. 101-18 testiment 0 0 0005 1 2 research Under N. 101-18 testiment 0 0 0005 2 research Under N. 101-18 testime	8 6 ######### bRT Gwerv 1.03E+18	Sentiment 0.2 0.45			
0     0 <td>9 7 ######## bSubtle hir 1.03E+18</td> <td>Sentiment 0 0</td> <td></td> <td>Output File B</td> <td>xample</td>	9 7 ######## bSubtle hir 1.03E+18	Sentiment 0 0		Output File B	xample
1         1	0 8 ######## bRT ChrisK 1.03E+18	Sentimenti 0 0.0625		ouputint	Mampie
2         10 </td <td>1 9 ######## bRT Gwenv 1.03E+18</td> <td>Sentiment 0.5 0.6</td> <td></td> <td></td> <td></td>	1 9 ######## bRT Gwenv 1.03E+18	Sentiment 0.5 0.6			
11         imments         0         0.005           12         imments         0.005         1.039932392239204743         N1         0.001101 cm         0.001	2 10 ######## bRT ChrisK 1.03E+18	Sentiment 0 0.0625		1 030036350033000000110 "BT jooppofe	a Autocomo buo tour aduc
1         1         Sections: 0.01/01         Sections: 0.01/01         1         0.0000/027/24/25/22/26/18         Advectore         Dur of Unit         Durof Unit         Dur of Unit <thd< td=""><td>3 11 ######## bEverywhe 1.03E+18</td><td>Sentiment 0 0.0625</td><td></td><td>1.03033035382393086+18 KI Joannete</td><td>a Awesome bus cour educ</td></thd<>	3 11 ######## bEverywhe 1.03E+18	Sentiment 0 0.0625		1.03033035382393086+18 KI Joannete	a Awesome bus cour educ
1       1       1.030931863581184e18       "At GeenGraham Our South Florida F         1       1.030931863581184e18       "At GeenGraham Our South Florida Frien         1       1.030931863581184e18       "At GeenGraham South Florida Frien         1       1.03093129118       "At GeenGraham South Florida Frien         1       1.03093129118       "At GeenGraham South Florida Frien         1       1.0309322498707600392418       "At mamelil 1 havenxez24809951 said         1       1.030900714918277488       Thavenez24809951 said         1       1.030900714918277488       Thavenez24809951 said         1       1.030900714918277488       "At Avenez24809951 said         1       1.03090071491827488       "At Avenez24809951 said         1       1.030	4 12 ######### bRT joanni 1.03E+18	Sentiment 0.519048 0.590476		1.0309362734255226e+18 "Awesome bus	tour educating the vot
0         1           0         10           1         10           1         10           10         10           10         10           10         10           10         10           10         10           10         10           10         10           10         10           10         10           10         10           10         10           11         10           11         10           11         10           11         10           11         10           11         10           11         10           11         10           11         10           11         10           11         10           11         10           11         10           11         10           11         10           11         10           11         10           11         10           12         10           12         10	5 13 #			1.0309319836381184e+18 "RT GwenGrah	am Our South Florida Ea
10         10<	6 14 #	nut File Example	a	1.0309318630980198e+18 "RT hammel11	I havenxe2x80x99t said
0         14         300	7 15 #	put i ne Brampi		1,0309313480962662e+18 "RT GwenGrah	am South Florida friend
9         1) ####### 01100000000000000000000000000	8 16 #			1 02002522409704590119 "PT GuonGrah	am South Florida friend
21         ####### 01,2001:3         Settement 03:000         1.039962794132         N1 namme111         1.navenxe2x80x9915 3a10           21         ####### 01/0000         Dikitis         Settement 03:000         Dikitis         D	3 1/ Sharana orr journe 1000-10	Seminene 0.313040 0.330410		1.03032322430704380418 KT Gwellor an	T have a support and
10         10<	0 18 ######## bRT joanny 1.03E+18	Sentiment: 0.519048 0.590476		1.030918/0/630039e+18 "KI nammel11	1 havenxe2x80x99t said
2         20 </td <td>19 mmmmmm 6Awesomi 1.03E+18</td> <td>sentiment 0.278571 0.385714</td> <td></td> <td>1.0309062709482291e+18 "RT GwenGrah</td> <td>am South Florida friend</td>	19 mmmmmm 6Awesomi 1.03E+18	sentiment 0.278571 0.385714		1.0309062709482291e+18 "RT GwenGrah	am South Florida friend
Distant         Distant <thdistant< th=""> <th< td=""><td>2 20 demonder DRT Gwenv 1.03E+18</td><td>Sentimenti 0.5 0.6</td><td></td><td>1.0309047633936957e+18 "RT GwenGrah</td><td>am South Florida friend</td></th<></thdistant<>	2 20 demonder DRT Gwenv 1.03E+18	Sentimenti 0.5 0.6		1.0309047633936957e+18 "RT GwenGrah	am South Florida friend
2         semment on town 100+13         Settement out, U.O.         1.039900744942737+18         Th howersez2885x995         Said a damn this           5         2         semment on town 100+13         Settement o         0.0         1.0399000744942737+18         Th howersez2885x995         Said a damn this           5         2         semment on town 100+13         Settement o         0         1.039900018194816e+18         "MobuFoday Ray from Davenport Flori           7         2         semment on town 100+13         Settement o         0         1.039900018194816e+18         "MobuFoday Ray from Davenport Flori	3 21 ######## bill hamm 1.03E+18	Sentimenti 0 0		1.0309013396505025e+18 "RT hammel11	I havenxe2x80x99t said
C2         27         25         ######## OHT Owen 101/18         Sentiment         0.2         0.45           27         25         ######## UHIT Owen 101/18         Sentiment         0.2         0.45	4 22 ######## bRT Gwen 1.03E+18	Sentiment 0.2 0.45		1 03000074404947370+18 "T havenyo2y	Payoot said a damp thin
to a measure on instant Just-16 sectored o v I.1.8389060108194816e+18 "MDMTOday Ray from Davenport FJO11	23 ######### ORT Gween 1.03E+18	Sentimento 0.2 0.45		1.05050074454047370418 1 Havenxezx	boxssc saru a damii thin
C/ 23 MINIMUM DI LOUT LUSETED SECTIONES U.2 U.45	o 24 annual off hamm 1.031+18	Sentimenti 0 0		1.0309000108194816e+18 "WDWToday Ra	y from Davenport Florid v
Handa has	a 25 minimum ORI Gweny 1.03E+18	sentiment 0.2 0.45		<	>

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.

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





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



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.

	Input File
8 Data Processing for Visualization	
<pre>]: df = pd.read_csv(r'D:\project\social media tutorial\example data\topic mode: print(dF) num_class = 10 topic_num = 10 topic_list = [] words_list = []</pre>	l\florida bus\LDA_patterns.out', sep = ' ', nam∳s = [
probability_list = [] final_name = ['Topic', 'Words df_final = pd.DataFrame(column	
<pre>tor i in range(o, copic_num): topic_num = i + 1</pre>	
<pre>tor j in range(2, num_class+2): id_num = j+i*12</pre>	
<pre>topic_name = 'Topic #' +str(topic_num) topic_list.append(topic_name) words_list.append(df.type[id_num]) probability_list.append(df.prob[id_num])</pre>	
<pre>df_final['Topic'] = topic_list df_final['Words'] = words_list df_final['Probability'] = probability_list</pre>	
df final	

Figure B.16 Data Processing for Visualization of Topic Analysis

	janize ivew Upen Select
AutoState (Bill) 📑 27 - (14 - a - Book	🖌 🎯 florida bus - Notepad — 🗆 🗙
File Home Insert Page Layout Formulas Data Review View Help O Search	File Edit Format View Help
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$	1.0309664470679512e+18 "RT ChrisKingFL Everywhere we go in ∧ 1.0309663678377124e+18 "RT ChrisKingFL Everywhere we go in 1.03096485092528e+18 "RT ChrisKingFL Everywhere we go in 1.0309567547865496e+18 "RT joannefea Awesome bus tour educ
<u>A</u>	1.0309562966049956e+18 "RT hammel11 I havenxe2x80x99t said 1.0309554228679088e+18 "realDonaldrimmpnRRA FLGovScott Thi 1.030954321599447e+18 "RT GwenGraham South Florida friends
A         B         C         D         E         F         G         H         I         J         K         L         M           3         1         2         6         7 sentiment polarity         subjectively         subjectivetively	1.839953923496881e+18 "Subtle hint from the Panama (119/FI 1.839951659743809e+18 "Rf thriskingFL Everywhere we go in 1.8309511678561884e+18 "Rf GwenGraham Our South Florida Ea 1.0309509267633888e+18 "RT ChriskingFL Everywhere we go in 1.83094099935384e+18 "Everywhere we go in Southwest Flor 1.8309409354866876e+18 "Rf igoannefea Awescome bus tour educ 1.830943548466876e+18 "Rf igoannefea Awescome bus tour educ
0         6         ####################################	Output File Example
12         10 ######## OKrisk 1.014-18         Sentiment         0         0.0625           13         11 ######## DKVorpych 1.031+18         Sentiment         0         0.0625           14         12 ####### DKT joann 1.038+18         Sentiment 0         0.0625	1.0309363598239908e+18 "RT joannefea Awesome bus tour educ 1.0309362734255226e+18 "Awesome bus tour educating the vot
Input File Example	<ol> <li>0309319836381184e+18 "RT GwenGraham Our South Florida Ea 1.0309318630980198e+18 "RT hammell1 I havenxezx808099t said 1.0309313480962662e+18 "RT GwenGraham South Florida friend 0.030053140208052e+18 "RT GwenGraham South Florida friend</li> </ol>
19 17 #hermann out gourne 1001-100 Sciences 0.020010 Construction	1.03092522498704580+18 KT Gwengranam South Florida Triend
10 4014000000     100100 L002100 Seminent 0.32008 0.030070     10 40100000     100101 1.00210     100100	1.030910/07030033etto NT Hammelli I havenxezx80x99t Salu
22 20 ####### bBT Govern 1.03F+18 Sentiment 0.5 0.6	1.0309002/094822910+18 KI GWenGraham South Florida friend
23 21 ######## b6T hamm 1.03E+18 Sentiment 0 0	1.0309047633936957e+18 "RT GwenGraham South Florida friend
24 22 ######### b6T Gwen 1.03E+18 Sentiment 0.2 0.45	1.0309013396505025e+18 "RT hammel11 I havenxe2x80x99t said
25 23 ######### bRT Gween 1.03E+18 Sentiment 0.2 0.45	1.0309007449484737e+18 "I havenxe2x80x99t said a damn thin
26 24 ######### bRT hamm 1.03E+18 Sentiment 0 0	1.0309000108194816e+18 "WDWToday Ray from Davenport Florid -
27 25 ######### bRT Gwen 1.03E+18 Sentiment 0.2 0.45	
florida bus	· · · · · · · · · · · · · · · · · · ·

Figure B.17 Data Samples for Visualization of Topic Analysis

]:	import csv,pdb				
	import numpy as np				
	import pandas as po				
	from datatime import datatime				
	import pickle				
	from datetime import datetime				
	from matplotlib import style				
	import matplotlib.ticker as mticker				
	import matplotlib.dates as mdates				
	import matplotlib.cm as cm				
	import mach				
	SMALL SIZE = 12				
	MEDIUM SIZE = 32				
	BIGGER_SIZE = 40				
	plt.rc('aves' titlesize_PIGGEP_SIZE)				
	plt.rc('axes', labelsize=MEDIUM_SIZE)				
	plt.rc('xtick', labelsize=MEDIUM SIZE)	Input File			
	plt.rc('ytick', labelsize=MEDIUM_SIZE)	input i ne			
	<pre>plt.rc('legend', fontsize=SMALL_SIZE)</pre>	*			
	<pre>plt.rc('figure', titlesize=BIGGER_SIZE)</pre>				
. J					
	path input = r"F:\topic model\tm\florida bus\result\florida bus.csv"				
1					
- E	math output = n"D:\nnoinst\corial modia tutonial\ovample data\te	anic model\flopida buc\Topic pop PT user bestman 1 ppg"			
۰.	pach_output = 1 b. (project (social media cucorial (example data(co	opic model(liorida bus(lopic_non_ki_user_neacmap_1.pmg			
		<u> </u>			
		Output Eigenne			

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