DCP 2002: Intro to GIS II
Final Project Paper

Analyzing Gainesville's Bicycle Infrastructure

Abstract:

The main purpose of this project is to determine which areas within Gainesville's city limits are in need for new bicycle infrastructure expansion and/or improvements. This project aims to guide any future allocation of resources for the expansion of the city's bicycle network to areas that are in most need because of a supposed high ridership from the people that live and work in such areas, the system's poor current connectivity, and the concentration of bike-related accidents.

Background:

Gainesville, FL and the University of Florida campus have been both recognized as some of the most bike-friendly urban areas and campuses in the United States, yet, their bike systems lack connectivity in some areas and therefore their ease of use is compromised, which in turn limits ridership and encourages higher accidents leading to injuries or fatalities.

As a regular user of the bike system in the City of Gainesville and the University of Florida, I have seen some particular examples of well-planned areas and poor-planned areas. This has made me want to analyze the entire system to see if there are any patterns that can be found through the use of geographic information systems to improve those specific poor-planned areas.

Geographic data has been collected from Alachua County Department of Growth Management, the University of Florida Department of Planning, Design and Construction, previous research done by University of Florida students, census data from the National Historical Geographic Information System (NHGIS), and geo-referenced police reports for accident data from SignalFour Analytics, as well as some data already collected by the instructor Juna Papajorgji, PhD.

Scope and characteristics of the study area:

The area analyzed in this project is restricted to a buffer of 2 miles away from the city limits of Gainesville, FL. Gainesville, FL is a city of roughly 130,000 residents, from which around half are students or staff at the University of Florida. The City of Gainesville has a historic downtown district with relatively high density when compared to other cities in Alachua county and surrounding counties. The downtown area and the University of Florida campus are relatively close, separated by only 12 to 13 city blocks. Even though this historical central area hosts most of the activities performed in the city, residences and commercial areas have been built away from this core during the past decades, creating a small-town urban sprawl and expanding the city's limits as a consequence.

This project focuses on the entire city of Gainesville, at a local level, where data from the City, the County and UF's campus merge and create a complete version of the urban area's infrastructure.

Socio economic data was analyzed at the block group level. There are more than 110 census block groups in the city of Gainesville.

Objectives for accomplishing the main goal:

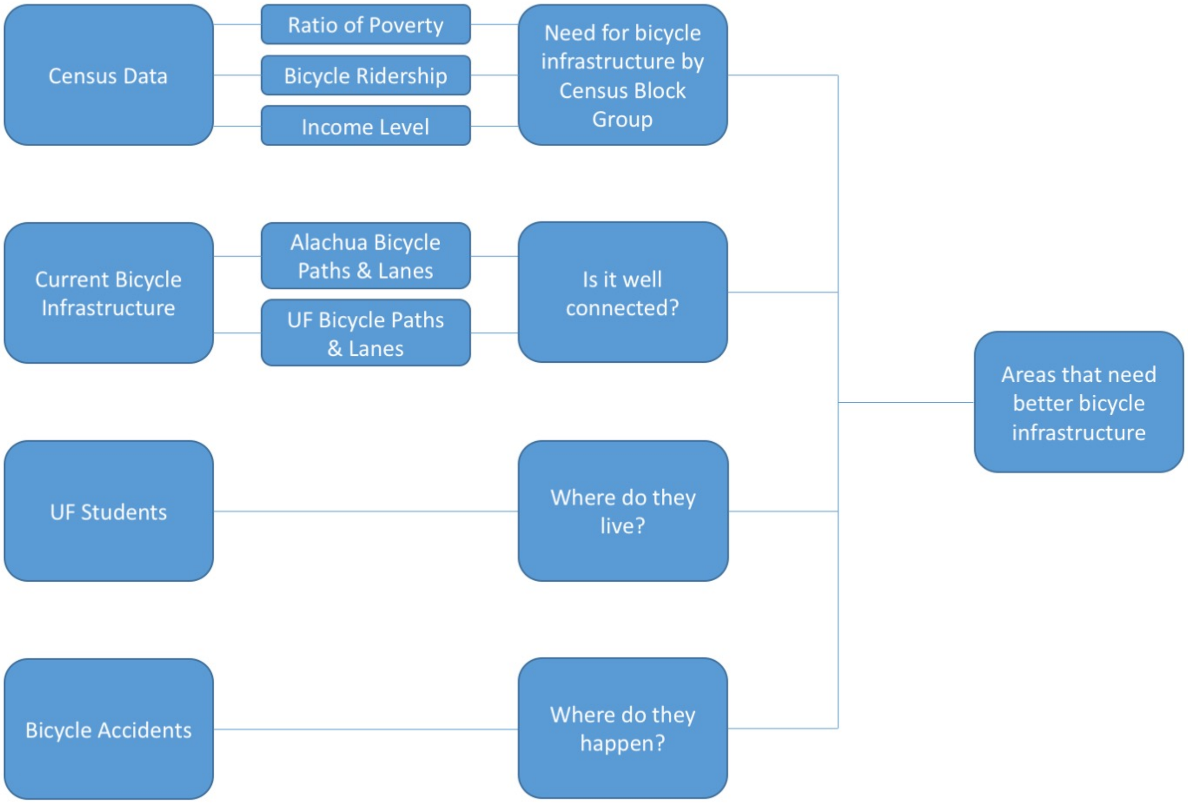
Create a map with all existing bike paths, shared roads, and bike lanes within the University of Florida campus and the city of Gainesville, FL. Analyze this bike infrastructure's overall ease of use utilizing spatial analyst tools and determine where future paths are needed.

Criteria:

Four main criteria elements will help determine which areas within the City of Gainesville need to upgrade their bike infrastructure. All these elements are necessary to understand why some people may feel threatened while biking or may not bike altogether because of a lack of adequate infrastructure:

1. Connectivity to other paths:
   Relative long distances between bike paths determine poor connectivity among them.
2. Socioeconomic data (2010 Census)
   A high amount of people that bike to work as main mode of transportation, per block group, means there is a higher need for adequate bike infrastructure.
   Poverty ratios per block group should mean more people rely on bikes than on personal vehicles for personal transportation.
   The income level per block group indicates which areas are more likely to not be able to afford a car and therefore bike more frequently.
3. Accident reports:
   Areas with the highest amount of bike accidents should indicate areas where bike infrastructure is mostly needed. This is based on police accident reports in which bicycles were involved.
4. Student addresses:
   Students are more likely to bike to school than to drive a car. A high concentration of UF students should likely mean a high need for adequate bike infrastructure.

The chart below shows this logical thinking:



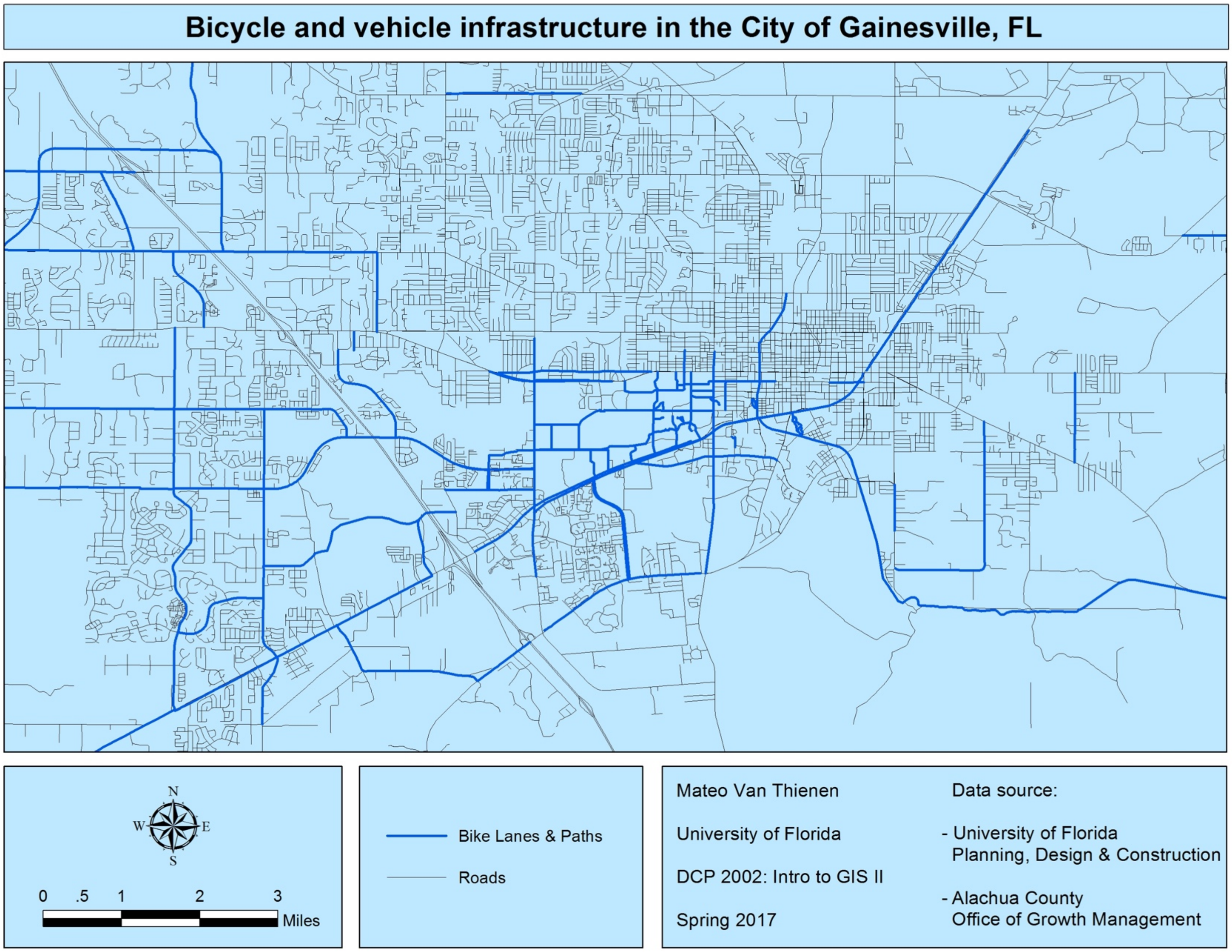
Methodology:

A. Using GIS step by step:

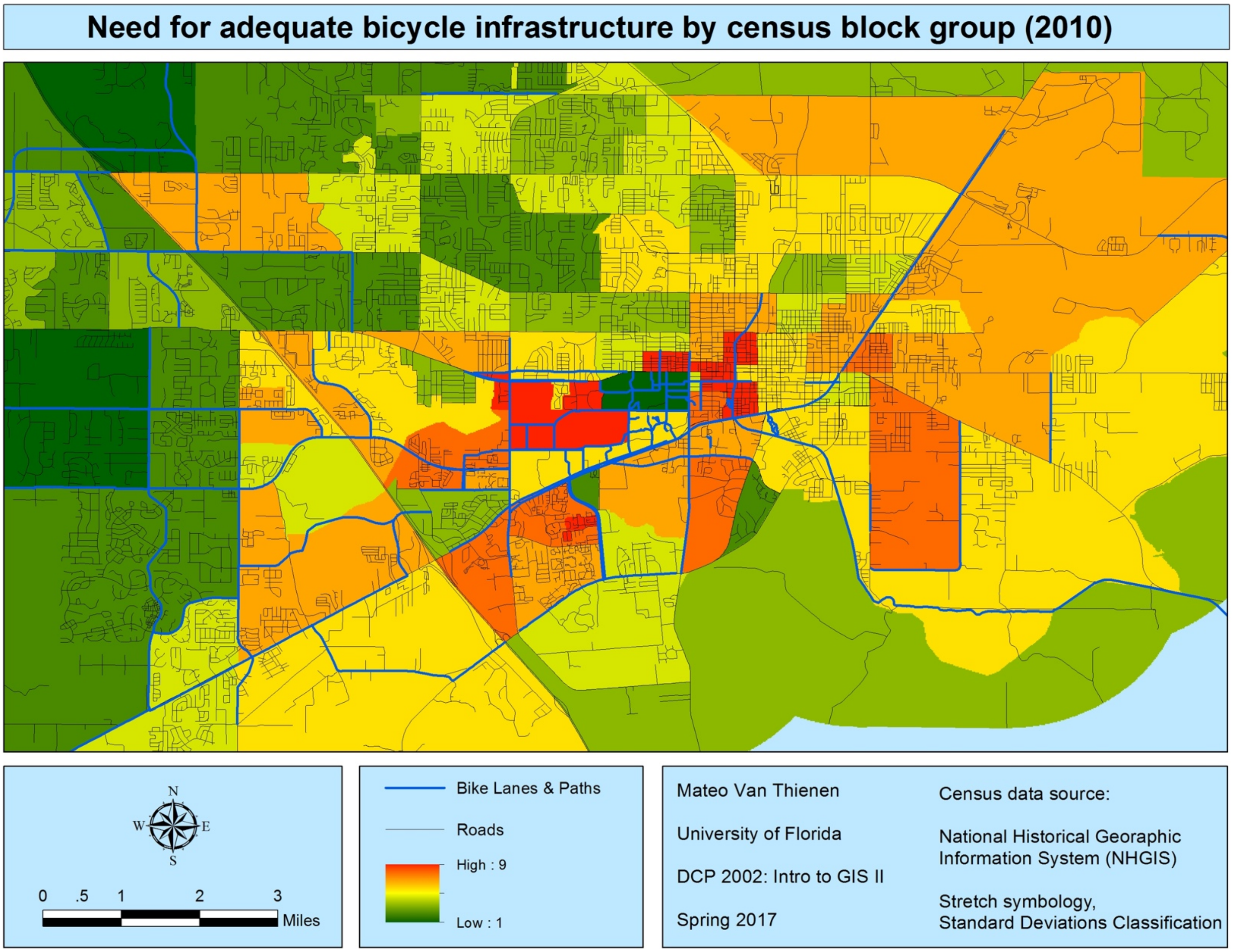
Census data in table format that contained information about modes of transportation, income levels and ratios of poverty per census block group were first synthesized and classified before adding them to ArcMap.

The first step performed in ArcMap was adding the 'Urban Areas' shapefile and then performing a *DEFINITION QUERY* to select 'Gainesville, FL'. Now that all other urban areas in the United States were hidden and only the study area was showing on the map, I performed a *BUFFER* of 2 miles around Gainesville, FL in order to include any relevant information to this study that may have gotten excluded for being located right outside sharp city limits. This buffer was used to clip most of the other layers later added along the project.

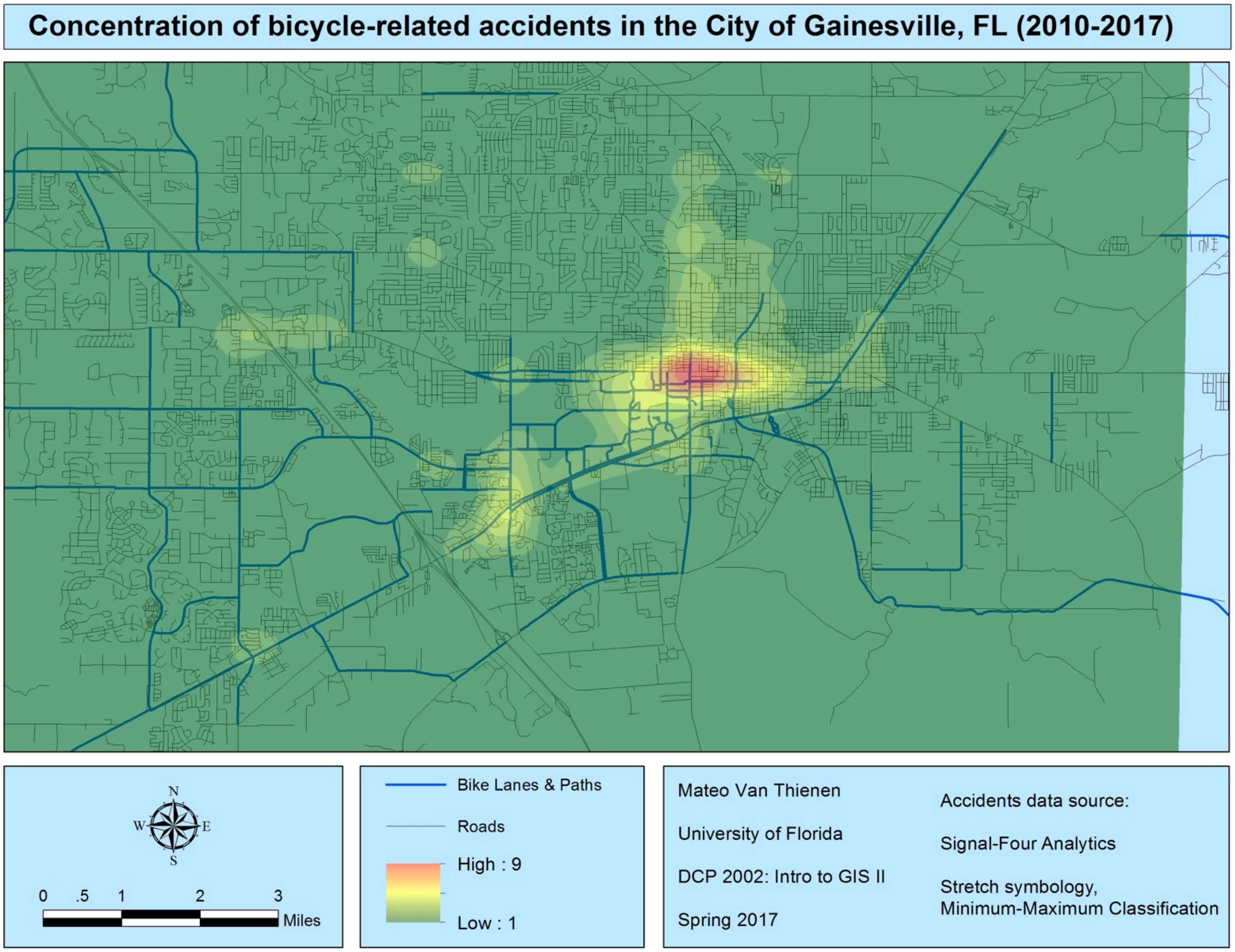
Roads and bicycle paths and lanes were then added to ArcMap and again clipped using the Gainesville, FL 2-mile buffer:



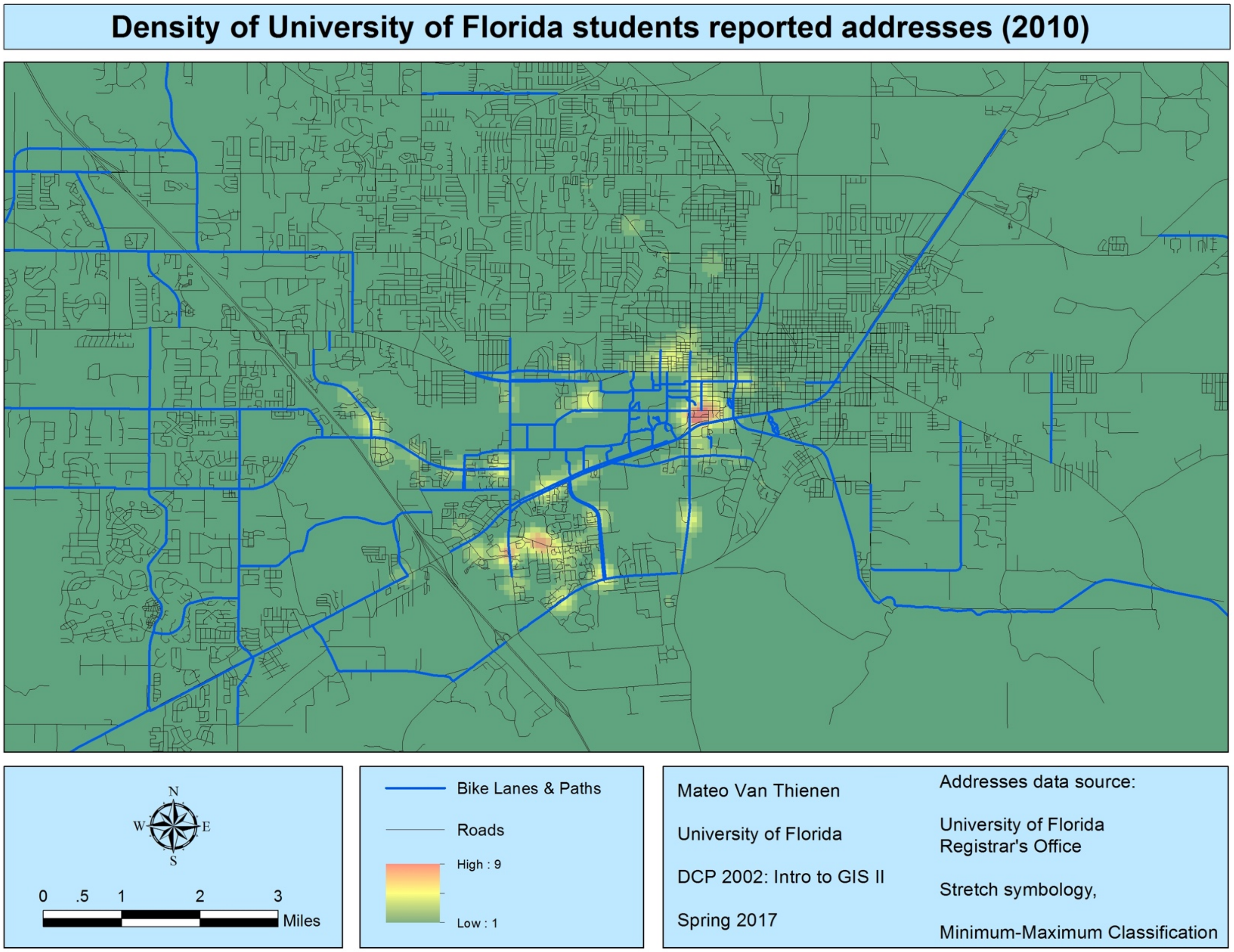
Census block groups for the State of Florida were then added to ArcMap. A *CLIP* operation was performed using the aforementioned 2-mile buffer around the Gainesville, FL urban area. Census data tables (percentage of people riding a bicycle to work, ratio of poverty and income level) were then added to the Table of Contents and a *JOIN* operation was performed to add the census data to the 'Florida census block groups' shapefile. After this, the symbology of the census block groups shapefile was changed to a gradient of colors showing the difference between census block groups among their bicycle ridership, their ratios of poverty and their income levels. The higher the bicycle ridership and ratios of poverty and the lower the income level, the higher the need for adequate bicycle infrastructure:



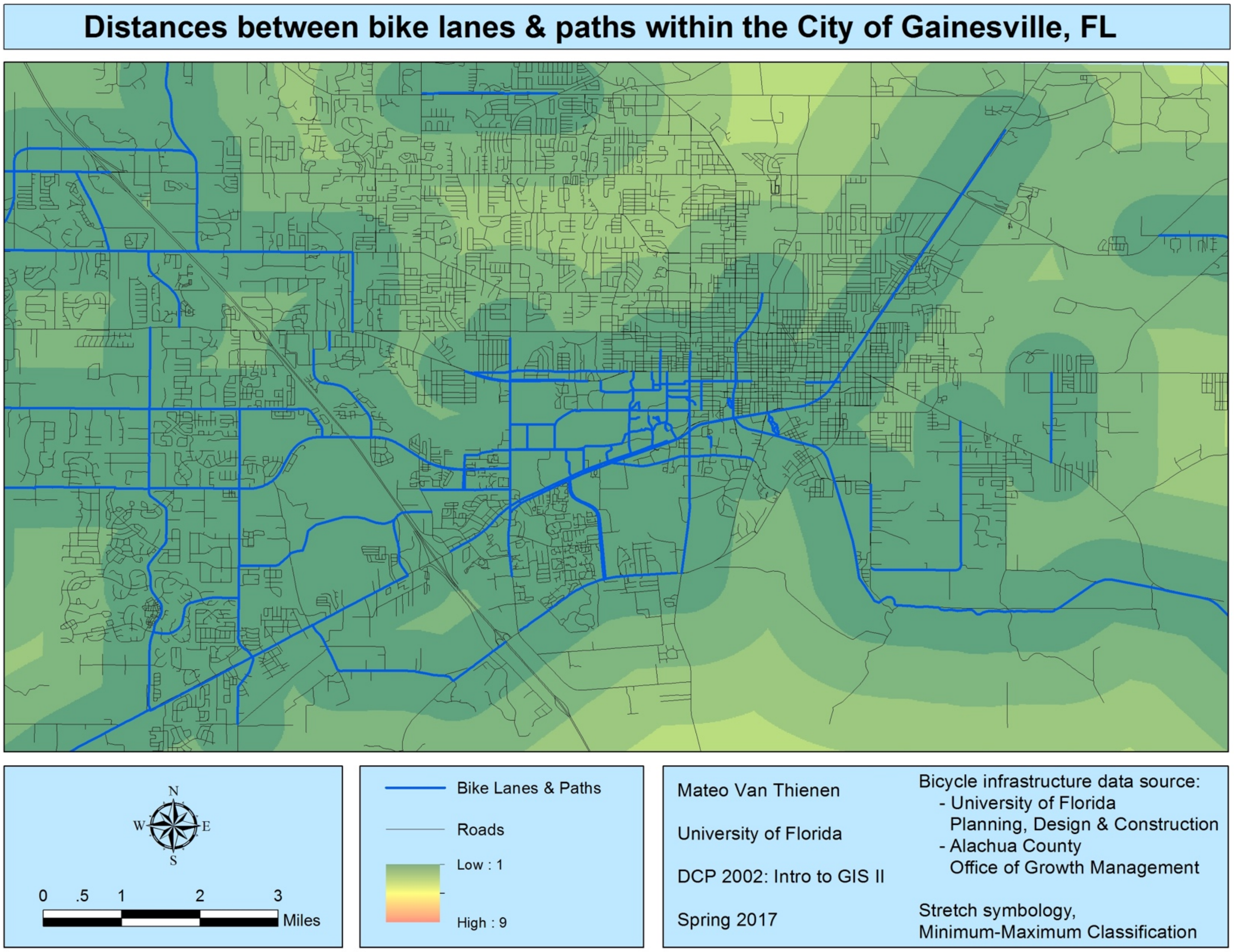
Bicycle accidents in Alachua County were then added to ArcMap and clipped using the Gainesville study area. This data source contained all accidents in Alachua County from 2010 to 2017 that reported at least one bicycle being involved, as primary involvement or collateral damage. These data source was provided in points so a *KERNEL DENSITY* was performed to see where most of the accidents happen. A direct correlation between the number of bicycle accidents and the lack of adequate infrastructure can be drawn by just looking at the following map:



University of Florida students' reported addresses were then added to ArcMap. University of Florida students were taken into account in this project because they are more likely to bike to class or work than the regular Gainesville resident. These students require adequate bicycle infrastructure to get to and from class on a daily basis, especially those students residing outside the university's campus. This is why all the address points were classified through a *KERNEL DENSITY* analysis, which is available in the following map:



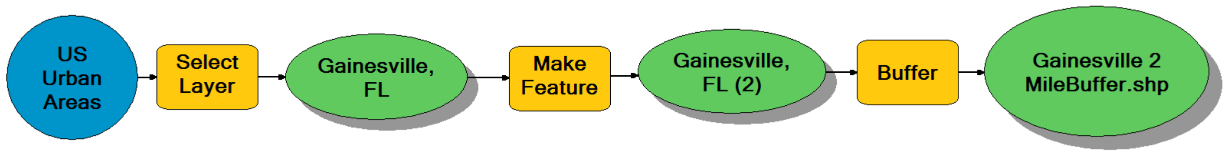
Lastly, a *EUCLIDEAN DISTANCE* operation was performed using the existing bicycle infrastructure in Gainesville in order to test its connectivity. In a scale from 1-9, these distances show how disconnected some bike lanes and paths are from the rest of the bicycle network. The more isolated the bicycle lanes, the higher the need for additional connecting lanes. Just as how vehicular infrastructure barely disconnects from the rest of the road network, bicycle lanes should not be disconnected from one another. Connectivity is essential to increase ridership, decrease the likelihood of accidents and increase the overall ease of use. The following map shows the distances among bicycle lanes and paths:



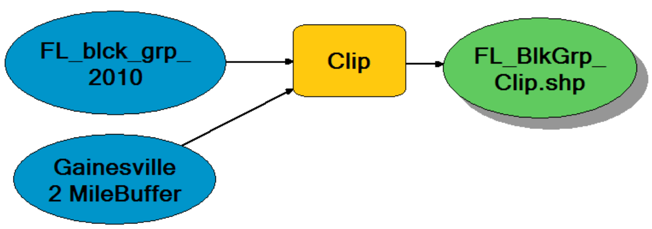
B. Automated GIS Models

10 different raw datasets were first added to the project, some of them as .dbf formatted tables and the rest as shapefiles. The .dbf tables containing census data were first classified and manipulated to narrow down only essential information for this project, before being added to ArcMap. These tables were then joined to the census block groups. Both the tables and the census block groups were downloaded from the same source, NHGIS, and were therefore easily joined because they both had a common field called GISJOIN.

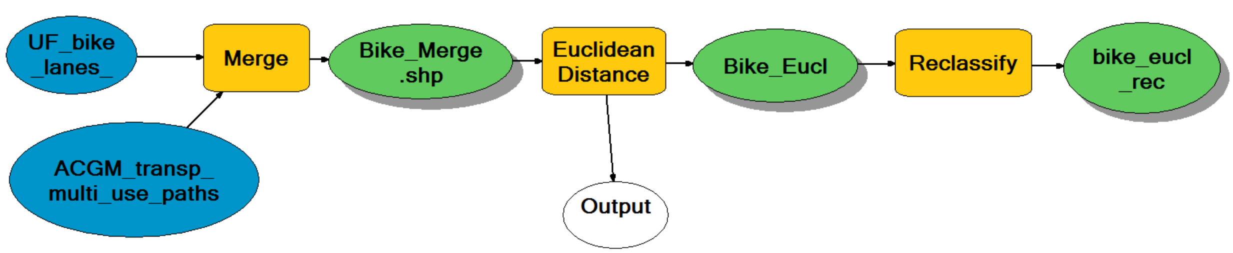
Starting with the Urban Areas shapefile, a “selection by attributes” was performed to select Gainesville, FL from all urban areas in the United States. After Gainesville, FL was the selected, a new layer was created in order to show Gainesville as the only urban area showing in the map. A 2-mile buffer was later produced using Gainesville’s urban area’s polygon. This was done in order to avoid cutting data out that may have been relevant to the study and that could have been left out due to sharp city limits. This buffer was used throughout the project in order to clip all other data sources and keep uniformity by having all data within the same study area. Model 1.1, in the chart below, shows this first step:



Model 2.1 shows how Florida 2010 Census block groups were clipped using Gainesville's 2-mile buffer previously created in Model 1.1. These census block groups were previously joined to tables with information about the percentage of people that bike to work, the ratio of poverty, and the level of income. Depending on each category's information, the census block groups were classified into 10 different classes, lower being the census block groups with the lower percentage of people biking to work, high income levels and low ratios of poverty, and higher being the opposite conditions. This was not automated in the Model because it was .dbf table manipulation, which cannot be automated. Model 2.1 can be replicated in the future when new census information becomes available or when additional demographic information is considered in the analysis:

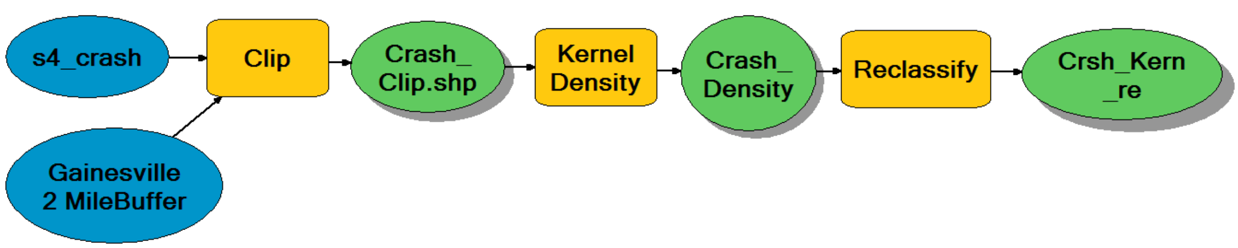


Model 2.2 shows the merge between University of Florida's and Gainesville's bike infrastructure. By calculating the Euclidean distance from each bike lane, a connectivity analysis could be done. 10 different distances were produced in a raster-format map. Classifications ranging from 0 to 9 were assigned to each one of the 10 distances calculated: 0 being the closest distances and 9 being the furthest distances, to show that bike lanes further apart need newer bike infrastructure to allow better connectivity among them:

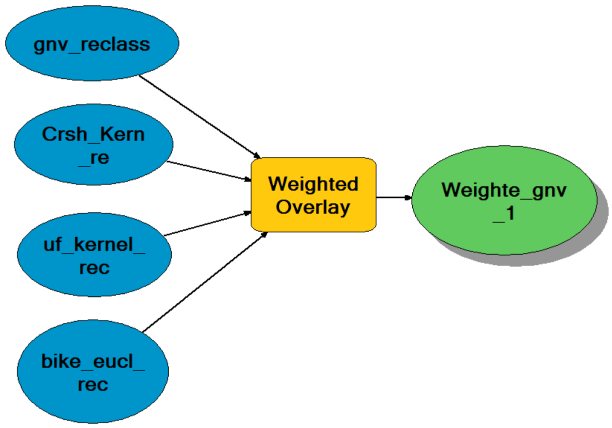


In model 2.3, UF students' addresses were used to determine where the most students live since this is a population that is highly likely to bike to class and therefore likely to be active users of the bike system in Gainesville. Since there were students living outside the city limits, these addresses were clipped using the 2-mile Gainesville buffer shapefile. A 'kernel density' analysis was performed which then showed where the highest concentrations of students are. These densities were also classified into 10 categories, ranging from 0 to 9, with 9 being the highest concentration of UF students:

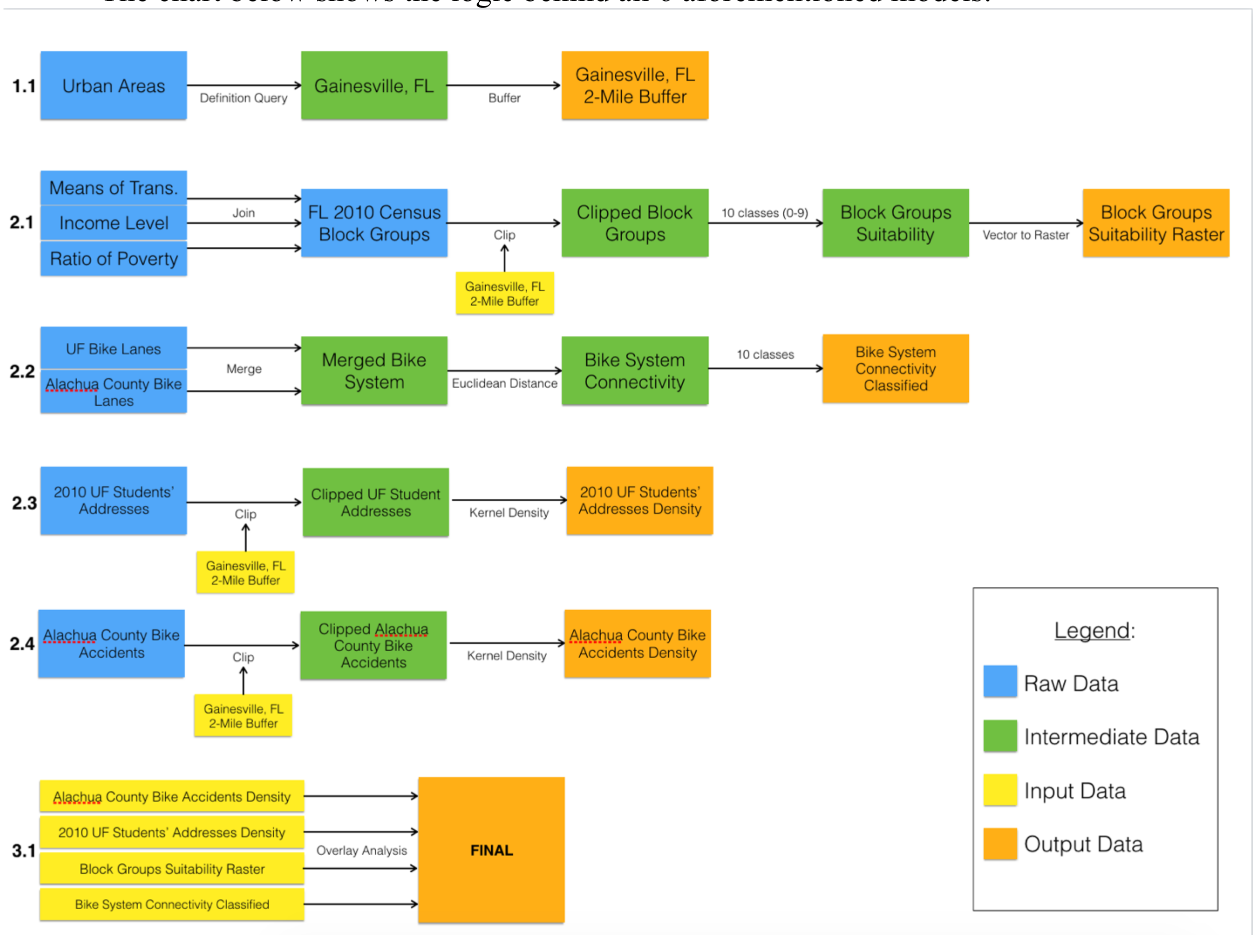
The data in Model 2.4 was derived from SignalFour Analytics, which provided data for all of Alachua County, which is why it first needed to be clipped using the same 2-mile Gainesville buffer in order to keep the data analysis within the topic area. This data showed points in the map in where traffic accidents involving bicycles occurred from 2010 until 2017. This data included both injuries and deaths, major and minor damages, nighttime and daytime accidents and accidents among bikes as well as as accidents involving bikes and other modes of transportation. Similar to what was done in model 2.3, these points were analyzed using a 'kernel density' analysis tool in order to see where the highest concentration of accidents occurred. Also, after these densities were calculated, they were classified into 10 categories, ranging from 0 to 9 as in all other models, 9 being the highest concentration of accidents:



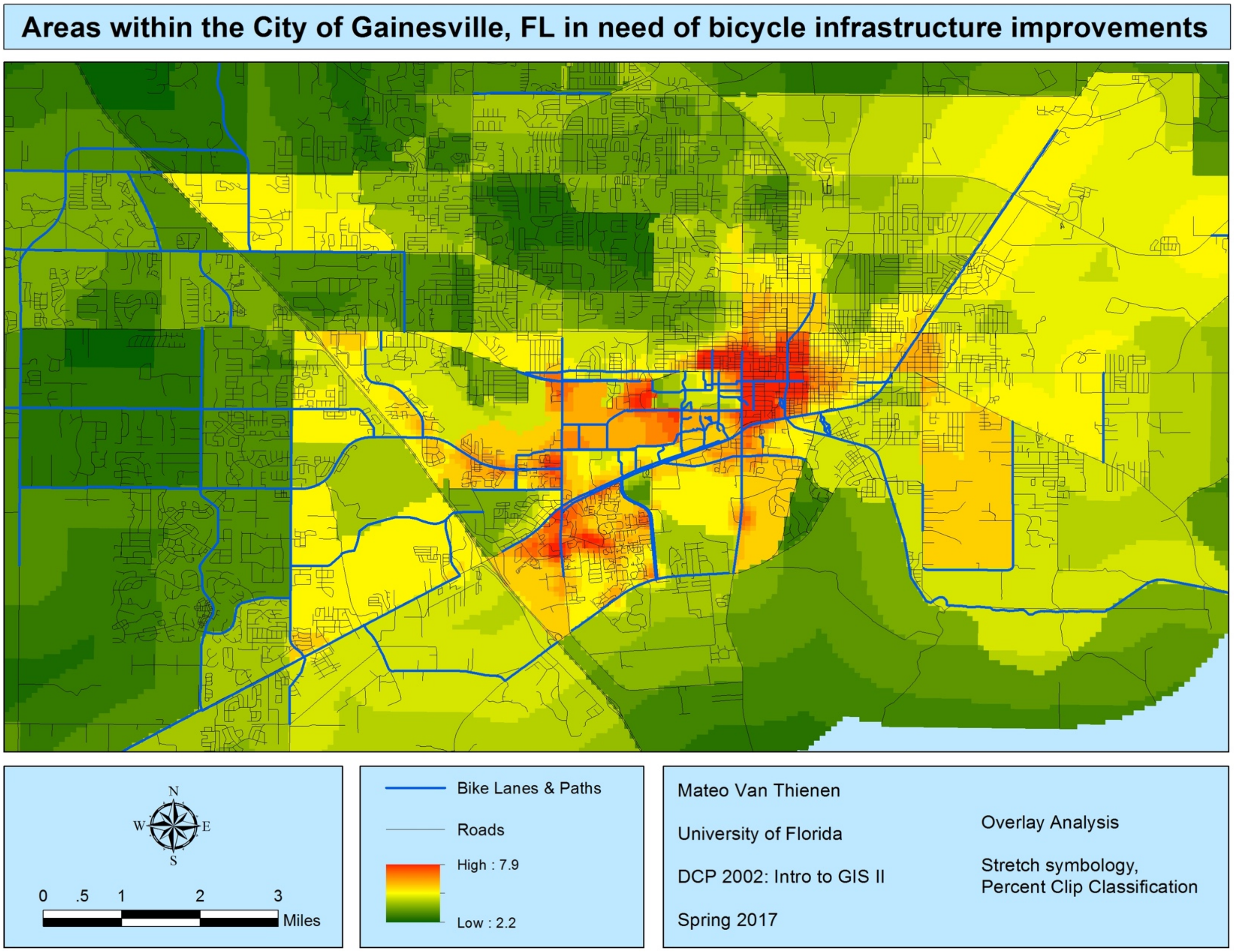
Lastly, in model 3.1, all four main raster datasets from Models 2.1, 2.2, 2.3 and 2.4 were combined through an 'weighted overlay analysis'. 30% of the weight was assigned to the accidents and the census block groups suitability respectively and 20% to the UF students addresses and the system's connectivity. Cells with an overlay analysis output closer to 9 are the places in where allocations for future bike lanes should most likely be built. According to the four factors analyzed in this project, cells with an output closer to 9 are cells in where people are probably low-income, frequently bike to work, are UF students, are far away from bike lanes and experience a high amount of traffic accidents:



The chart below shows the logic behind all 6 aforementioned models:



Results and Discussion:

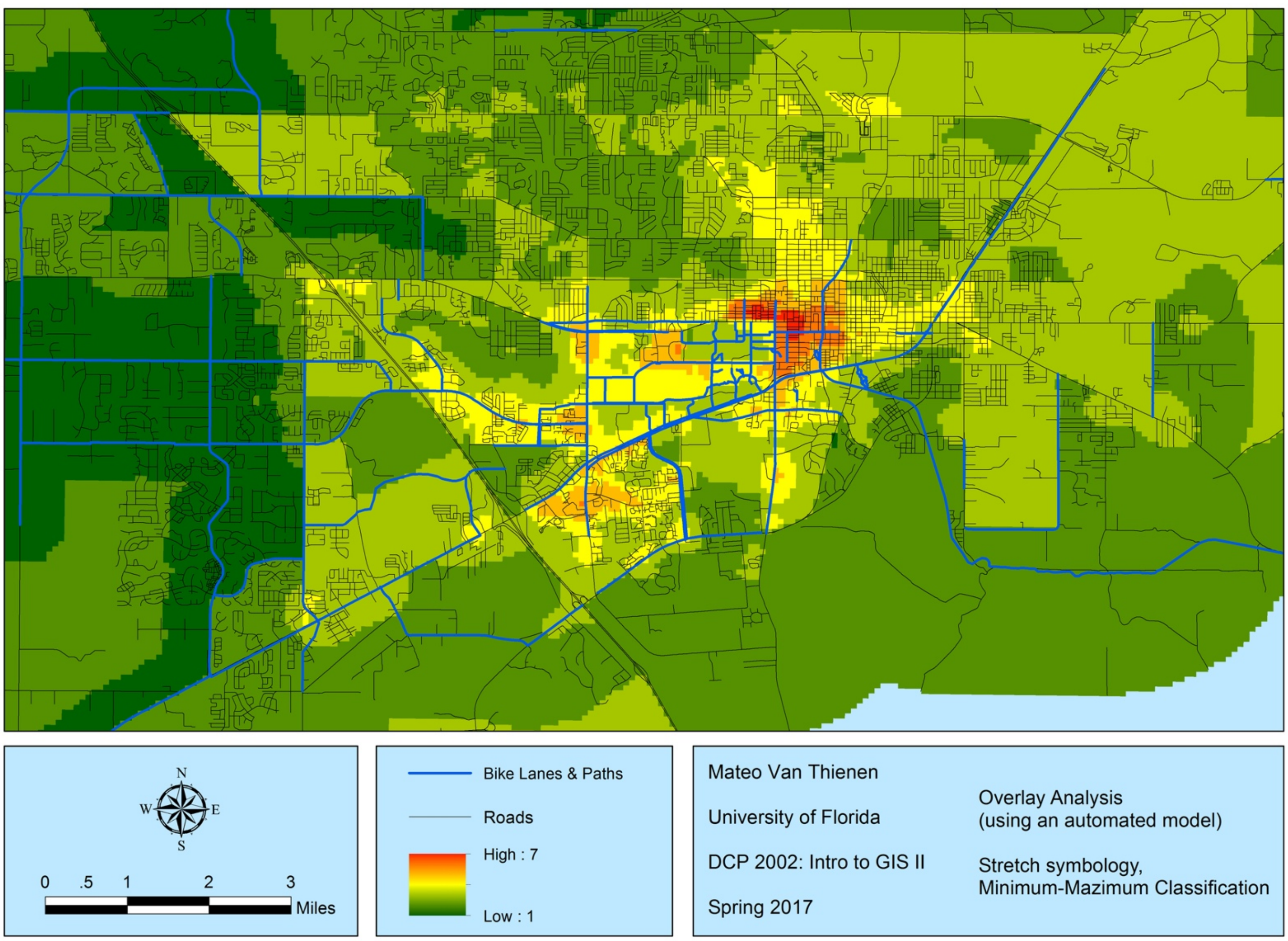


This map shows the results of the final overlay analysis using the step-by-step GIS methodology. The red areas in the map are the areas that this project has found are in the most need for new or upgraded bicycle infrastructure. These areas include downtown Gainesville all the way west to the University of Florida campus and north of campus in Midtown. This primarily due to the fact that a lack of bicycle infrastructure in University Avenue increases the likelihood of accidents among bicycles and vehicles. Even though SW 2nd Ave is considered a Bicycle Boulevard, it cannot be the only bicycle-accessible route in between Downtown Gainesville and the University of Florida campus. More connectors like this one are needed, and University Ave should be one of them. The midtown area is also poorly served, as 13th Street does not properly accommodate bicycles north of University Ave.

The historic University of Florida campus is almost entirely shown as a neutral place in either yellow or green colors, whereas sections in the southwest of campus appear in red. This may be due to the fact that the historical northeast corner of campus is more densely built, with more walkable areas than the southwest region.

Areas across Butler Plaza, near Interstate 75 also appear in red and these are commonly known to be low-density residential with a few apartment complexes that are apart from each other and where infrastructure is mostly designed for the private automobile and not precisely for bicycles. The high concentration of businesses in this area attracts residents from all of Alachua County, which is precisely why the area surrounding Archer Rd and 34th Street should be planned to accommodate alternative modes of transportation and not only private vehicles.

Also apparent by looking at the map, the northeast side of Gainesville shows yellowish colors predominantly, whereas the northwest side shows mostly green colors, with the exception of one census block group in yellow. This clearly shows a disparity between low-income areas in the northeast and high-income areas in the northwest. The current bicycle infrastructure mostly serves wealthy areas in the west, while the struggling census block groups in the east have little to no existing bicycle infrastructure to meet their needs.



Areas in the City of Gainesville, FL in need of bicycle infrastructure improvements

The map above shows the same overlay analysis, this time using the automated model strategy. The results are slightly different, however, very similar conclusions can be drawn from both maps.

Assumptions and Limitations

This project was based on several assumptions and does not take into account other planning disciplines that may shed light upon issues, problems or opportunities not described nor taken into account in this project. If this is to be continued, land use laws and regulations should be included, as well as other bicycle infrastructure influencers such as population density.

Some assumptions made in this project are:

- the fact that not all UF students bike to class or work and many walk and/or use public transportation instead.
- Not all low-income people necessarily bike more than drive a car and not all bike accidents are due to poor infrastructure design.

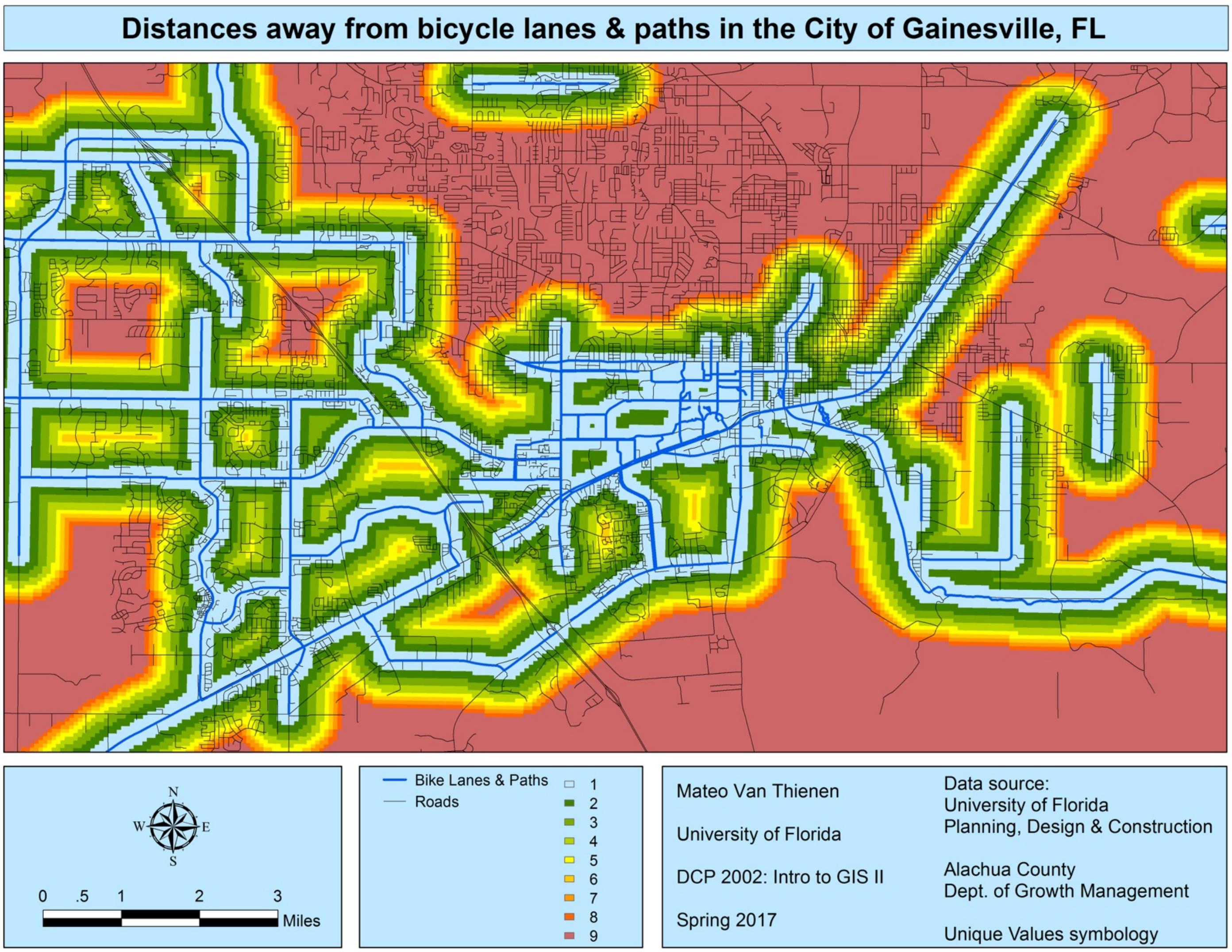
Apart from these restrictions, this project does a good job at locating potential zones that may need better bicycle infrastructure than others, especially as it takes several factors ranging from socio economic to demographic and combines them using spatial thinking techniques.

The automation of this project is limited because of the manipulation of census data that is necessary before being able to draw conclusions using geographic information systems.

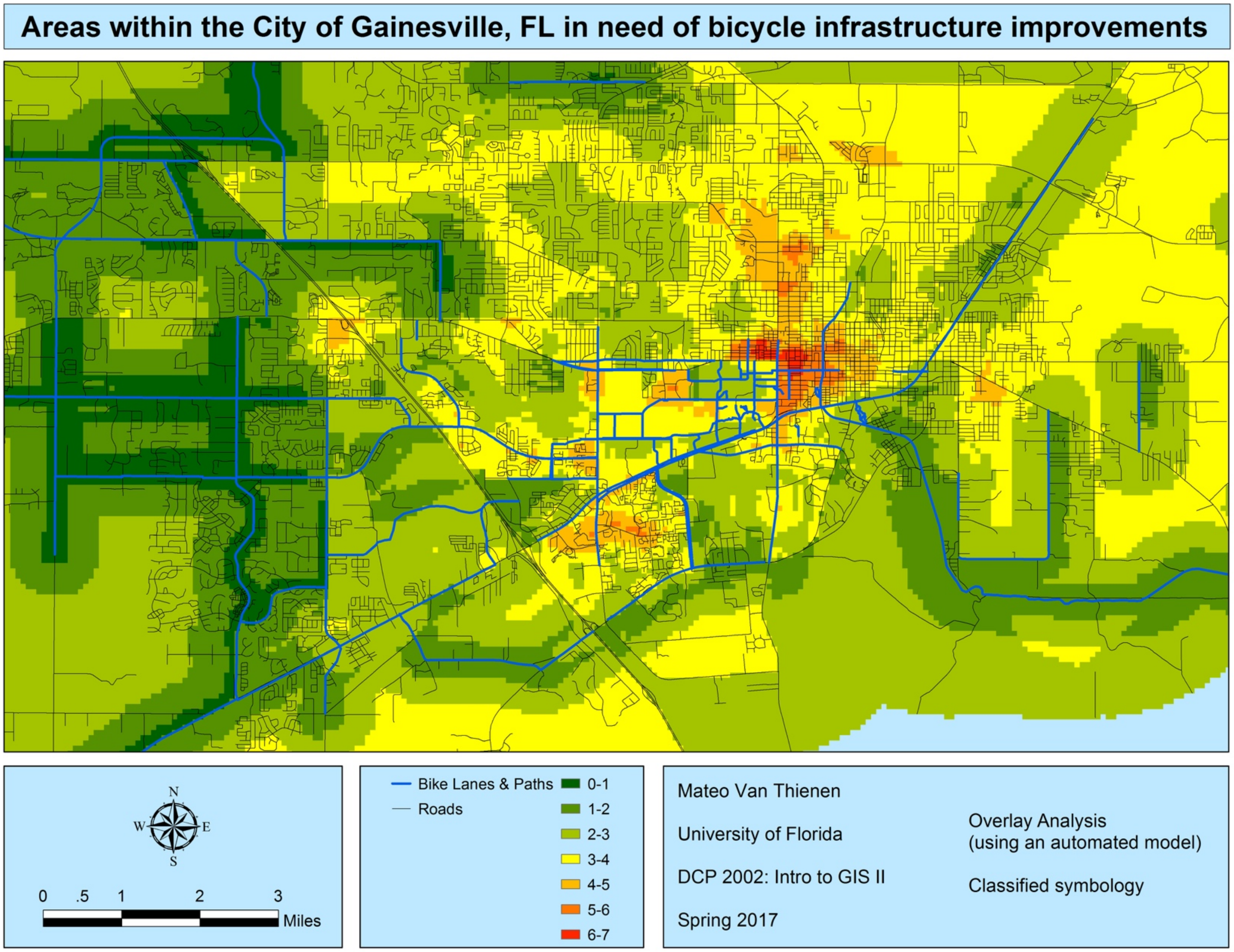
By using the automated models, different results will be readily available based on changes in criteria. It is usually best to try different methods and different variables in order to draw the most neutral and unbiased results possible. Some criteria that could be added to this project includes the population's age ranges and the presence of recreational sites.

Testing Alternative Criteria:

In order to narrow down the results and therefore produce more accurate deliverables for Gainesville city planners, a change in the Bicycle Network Connectivity's criteria was performed. The Euclidean distance function was set with a limit of 1000 units, with cells being 25 units in size. 9 classes were again produced, showing the relative distance among bicycle infrastructure in the city of Gainesville. The map below shows the result of this new criteria:



By changing the criteria in one of the four main components in the project, the overall results also suffered some changes. The map below shows a new overlay analysis, taking into account the new criteria mentioned above:



This new final map not only shows narrower green areas, meaning the need for bicycle infrastructure overall increased but also the northern area also increased in red, meaning a higher need in that specific area. The northern area along NW 13th Street lacks bicycle lanes, has a high concentration of bicycle accidents and has a large number of businesses and residences that make it a popular destination for various uses.

It is important to note how important different criteria may impact the results of a GIS analysis, as witnessed in this project. This simple change in Euclidean distance parameters produced a rather important change in results, hence the possibilities of further improving the results are still possible and necessary to thoroughly asses Gainesville's bicycle infrastructure.

Metadata/Documentation:

a. Urban Areas Shapefile – Juna Papajorgji, PhD.
The Bureau of the Census Urbanized Area Boundaries data set contains boundary information for urban areas with a population greater than 50,000. The data set includes boundaries for urban areas in all 50 states, the District of Columbia, and Puerto Rico. A UA may contain both place and non-place territory. The U.S. Census Bureau delineates UAs to provide a better separation of urban and rural territory, population, and housing in the vicinity of large places. At least 35,000 people in a UA must live in an area that is not part of a military reservation. The data provide users with information about the locations, names, and size of urban areas. The data are used primarily for national planning applications.

b. UF Bicycle Lanes – UF Department of Planning, Design and Construction

c. Alachua County Bicycle Lanes and Paths – Alachua County Department of Growth Management.
This dataset contains existing and proposed multi-use paths of Alachua County, Florida. Multi-use paths are paved paths or sidewalks that are typically 8 ft. or greater in width and accommodate both pedestrian and bicycle traffic. This dataset corresponds with the existing and future multi-use paths displayed on the adopted map "Bike/Pedestrian Existing & Future Network" which is part of the Transportation Mobility Element in the Alachua County Comprehensive Plan.

d. Alachua County Bicycle Accidents (2010-2017) – SignalFour Analytics

e. UF Student Addresses – Juna Papajorgji, PhD

f. Alachua County Roads – Alachua County Department of Growth Management
This data set was provided by the former GIS Division at the Department of Growth Management, Board of Commissioners, Alachua County. This data set shows the inventory of locally maintained roads with an allowed driving speed up to 35 miles per hour. The data set was prepared in 2008 by the GIS Division of Alachua County with input from the Department of Public Works at the City of Gainesville, in support of the policy analysis and recommendations provided by the Alachua County Energy Conservation Strategies Commission (ECSC) via its final report.

g. Florida 2010 Census Block Groups Shapefile – NHGIS
These data are based upon work supported by the National Institutes of Health and the National Science Foundation. Its purpose is twofold. First, the NHGIS created and freely distributes a database incorporating all available aggregate census information for the United States between 1790 and 2010. The database contains information for a wide variety of statistical (blocks, block groups, census tracts, metropolitan statistical areas, etc.) and administrative units (places, minor civil divisions, counties, states, etc.). Second, the NHGIS produced and freely distributes boundary files for small areas (census tracts and counties) in the United States. Boundary files for tracts are available for the 1910-

2010 decennial censuses and American Community Survey releases, and boundary files for counties and states/territories are available for the 1790-2010 decennial censuses and American Community Survey releases. The boundaries contained in this file do not necessarily represent current legal boundaries. Instead, they represent the boundaries over which the U.S. Census Bureau tabulated and published data at that time. In order for others to use the information in the Census MAF/TIGER database in a geographic information system (GIS) or for other geographic applications, the Census Bureau releases to the public extracts of the database in the form of TIGER/Line Shapefiles

h. Florida 2010 Modes of Transportation – NHGIS
Means of Transportation to Work
Universe: Workers 16 years and over. Source code: B08301. NHGIS code: JM0

i. Florida 2010 Median Income – NHGIS
Median Family Income in the Past 12 Months (in 2010 Inflation-Adjusted Dollars)
Universe: Families. Source code: B19113. NHGIS code: JPO

j. Florida 2010 Ratio of Poverty – NHGIS
Ratio of Income to Poverty Level in the Past 12 Months
Universe: Population for whom poverty status is determined. Source code: C17002. NHGIS code: JOC