Carlos Acuña
Benjamin Johnson
Sabrina Nelson
Sarina Weiss

URP 4273 Final Project: Where Should I Move?

Main Goal

The main goal of this project and report is to demonstrate how ArcGIS can be used in a daily life decision. This report will explain how anyone with access to the GIS program can find data for free online, put it into GIS, and use it to find the answer to a difficult question that might include many factors, requirements, or just too much data for someone to sort alone. Also, more specifically, this report aims to find the best place to live within Florida and Georgia for a recent college graduate making $75,000.

Background

In this project, we will take a look at physical, demographic, and socio-economic data in the United States on a zip code and county level in order to quantitatively determine the best place to live by using GIS mapping software. Although this software and data can be used to find the best place to live for any person, for this project specifically, we will be assuming the role of a young person in their 20's who aims to make roughly $75,000 per year. Of course, there are an infinite number of possible dimensions or "indicators" that could be used to define the best place to live. However, for the sake of being brief, we have chosen twelve indicators that will be represented by sets of data for this project. These are the factors that we have determined are most likely the most important in determining what makes up a good area to live (down to a zip code level). Our twelve indicators are broken up into four groups: Economics, Health, Community, and Recreation. More details on why we chose these categories and data sets within them will be explained in depth later in this report.

Scope

Although this process can be done using ArcGIS given data encompassing the entire United States, we have chosen to only focus this project on the area occupied by Florida and Georgia. Because we are students of the University of Florida and the person most likely to use this process is someone who lives in Florida, choosing Florida and Georgia as our two states was the clear choice because of their proximity to the University of Florida. Also, because the person we are using in this report as an example grew up in Florida, it seemed natural to pick two states with the most similar culture and values. Our group members also desired a location in the southeast due to the warmer climate and vicinity to family members.

We have done our best to collect data at the smallest level possible (zip code) because it can give data on a smaller geographical area which makes the process of selecting a place to live much easier. However, not all data is available at the zip code level, therefore some data was collected on a county-wide basis, however we integrated this data to the best of our ability.

Criteria (in no particular order)

Part A: Economics

1. Median Household Income by Zip Code (Dollars)
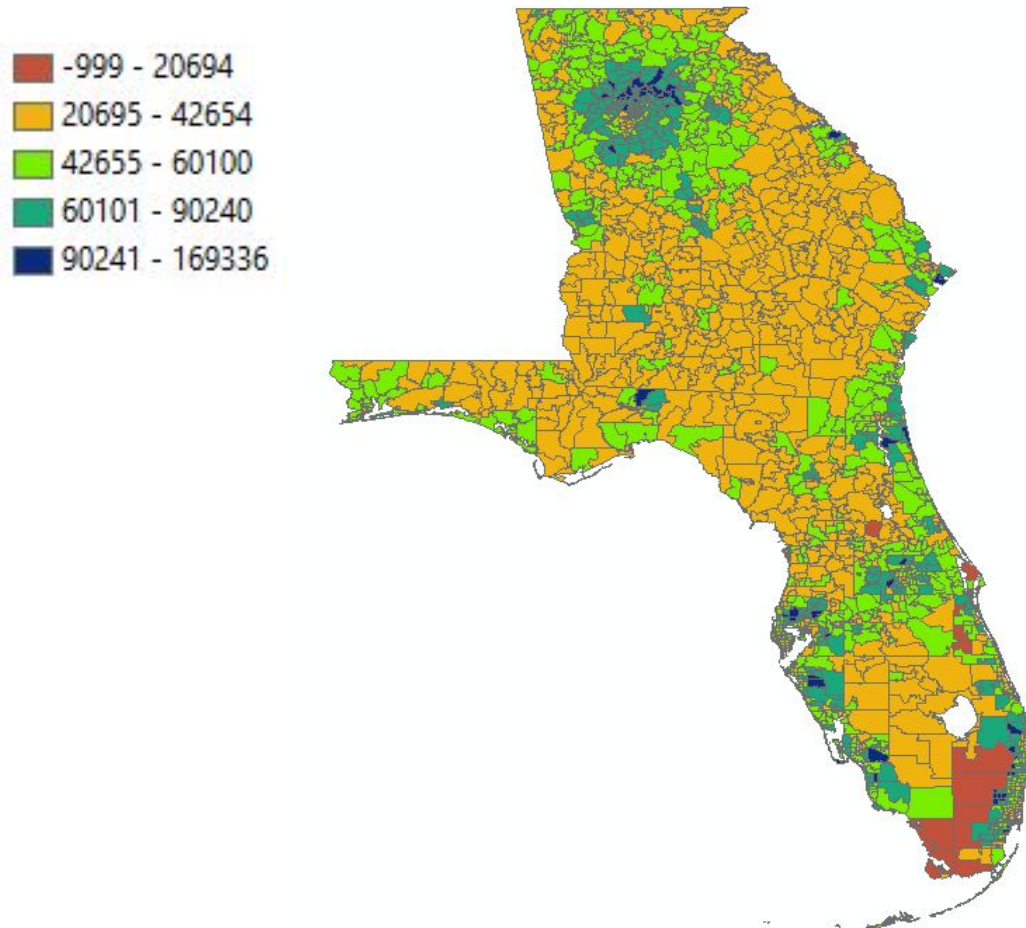
Median household income was chosen first because money is an important factor in living. Household income data is important because it gives an idea of what areas are best to live in to make the most money possible. More specifically, median household income was chosen because it better represents what the average person earns. If average household income was chosen, the values would be skewed by the incredibly rich people at the very top who would bring up the average household income amount. The effects of the very top 1% of earners has no real effect on the median household income data, making median the optimal choice for data use.

Of course, median household income does not take into account the cost of living. Areas with higher median income tend to have higher rent expenses. For that reason, the next indicator we've chosen is median disposable income, which takes into account the cost of living. Median household income can still be very useful nonetheless, especially in scenarios where housing might be provided for

you or if you can find a good deal on housing, which would make the effects of housing on your choice smaller.

We used the ArcGIS program to create five natural breaks in our data so that each class of median household income could be as close to actual classes as possible. Also, since our goal is to find the best place to live in these two states, it is especially helpful that we were able to find this data by zip code, which gives us the most accurate information.

2. Median Disposable Household Income by Zip Code (Dollars)
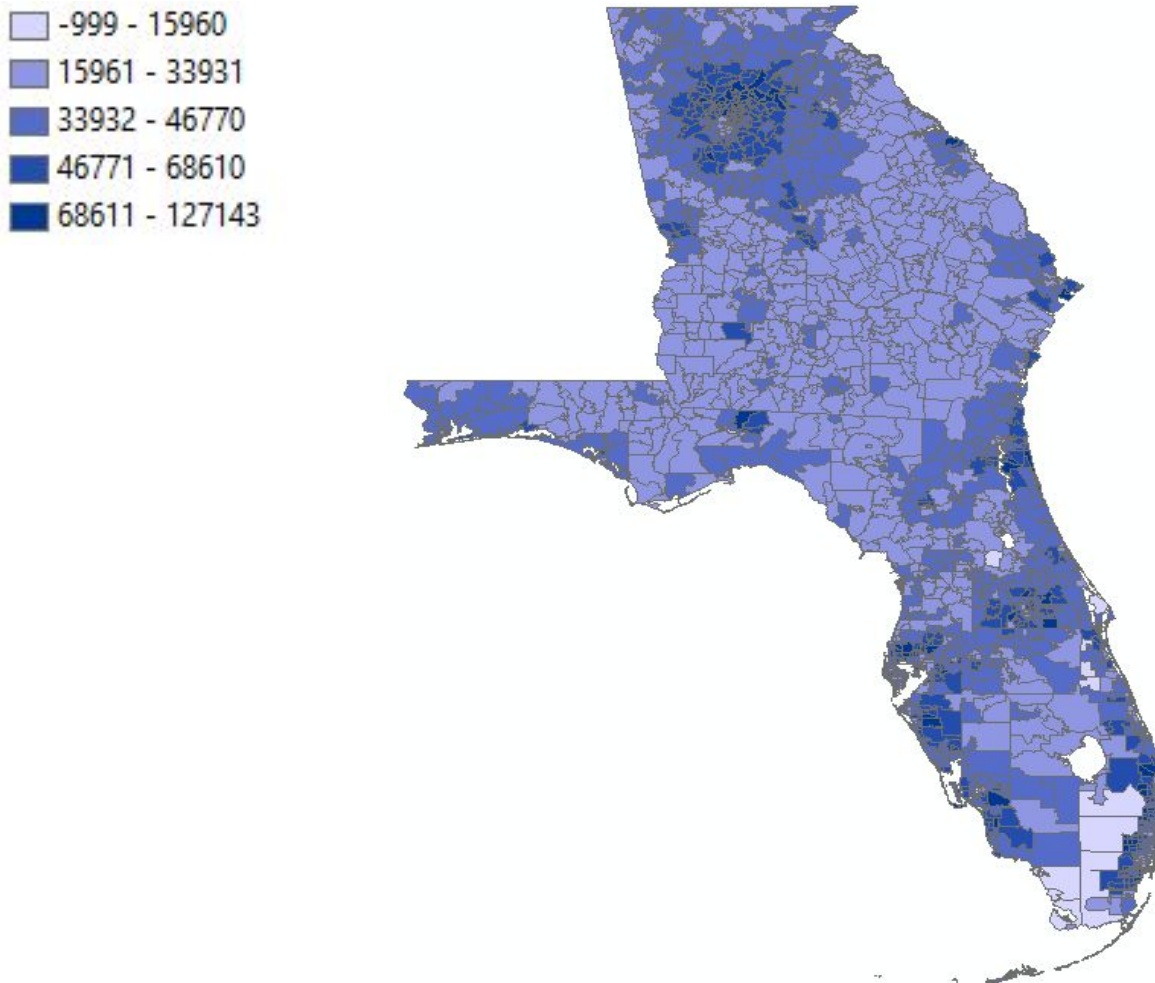
As stated in the paragraph above, median disposable household income is useful because it takes into account the effects of cost of living on a person's income. If a person is not confident in their ability to find affordable housing or isn't supplied housing assistance by their work, the effects of cost of living become dramatic, making median disposable income a very important indicator of where the best place to live is. Also as stated before, median data is best because it is unaffected by outlying data points We also used ArCGIS again to create five natural breaks in the data to create five classes that are as

distinct as possible. Data was also available at the zip code level which is very useful for finding the most precise location to live.

3. Unemployment Rate by County (%)
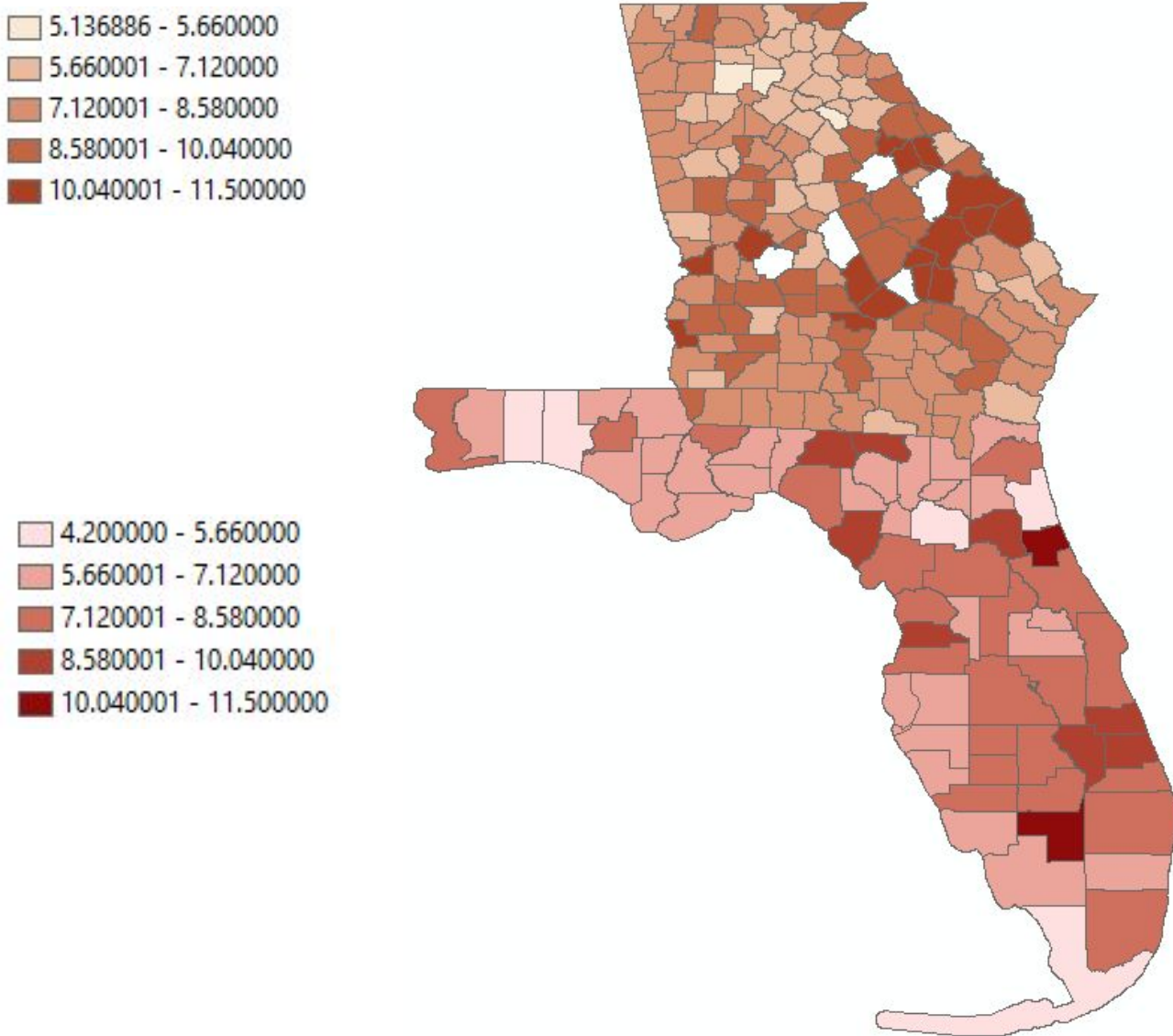
The third economic indicator we chose for finding the best place to live is the county unemployment rate. Unfortunately, unemployment rates are not taken at the zip code level, which is why this data set is done by county. Being able to find a job is perhaps the most important part of deciding where to live. Because unemployment rate is generally a good statistic to describe the local job market and local economy, a person might want to choose to live in an area with the lowest unemployment rate possible.

For this data, once again, natural data breaks were used to sort the data into five distinct classes on the map. Although there are two different legends shown above for the data classes of each state,

only the first class is different because some counties in Florida have an unemployment rate of 4.2 %, which is lower than any rate in Georgia. However, we wanted the employment rate to be within the lower 40%, therefore making the range for both states the same.

Part B: Health

4a . Average Male Obesity Rate by County (%)
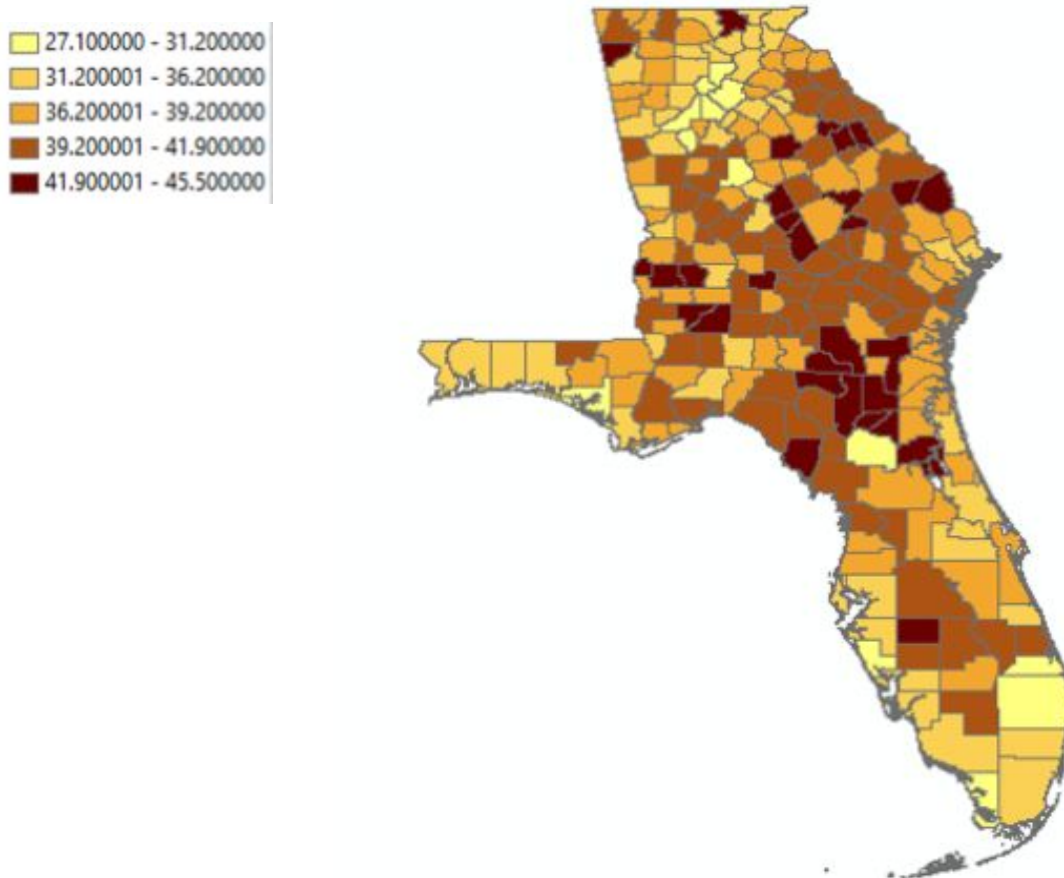
Obesity Rates in general can be an incredibly useful tool for deciding which area to live. If you are the type of person who likes to exercise, you might prefer to live in an area where the obesity rates are low because you would be more likely to find people with a similar lifestyle. They can also give some insight into the types of food that are most popular in certain areas, which is a big aspect of an area's culture.
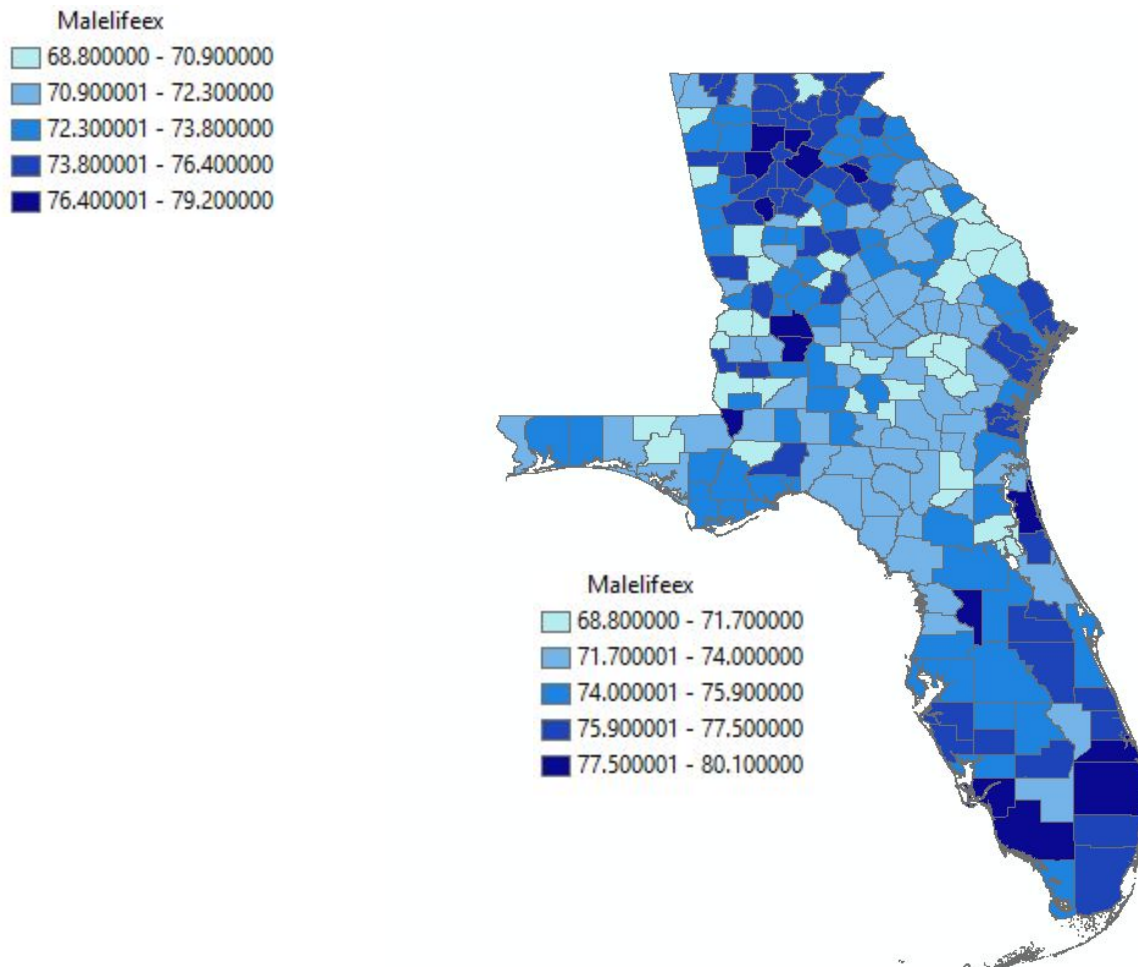
Unfortunately, this data is only available at the county level. It is also important to point out that the way the five classes were arranged were by using natural breaks in the data calculated by the ArcGIS program. By putting data into distinct colors on the map, it becomes easier to analyze. Moreover, since obesity is a characteristic where a person is either obese or not obese, an average obesity rate is the only possible way to quantify the data. Female Obesity rates will be included as our next data set.

4b. Average Female Obesity Rate by County (%)
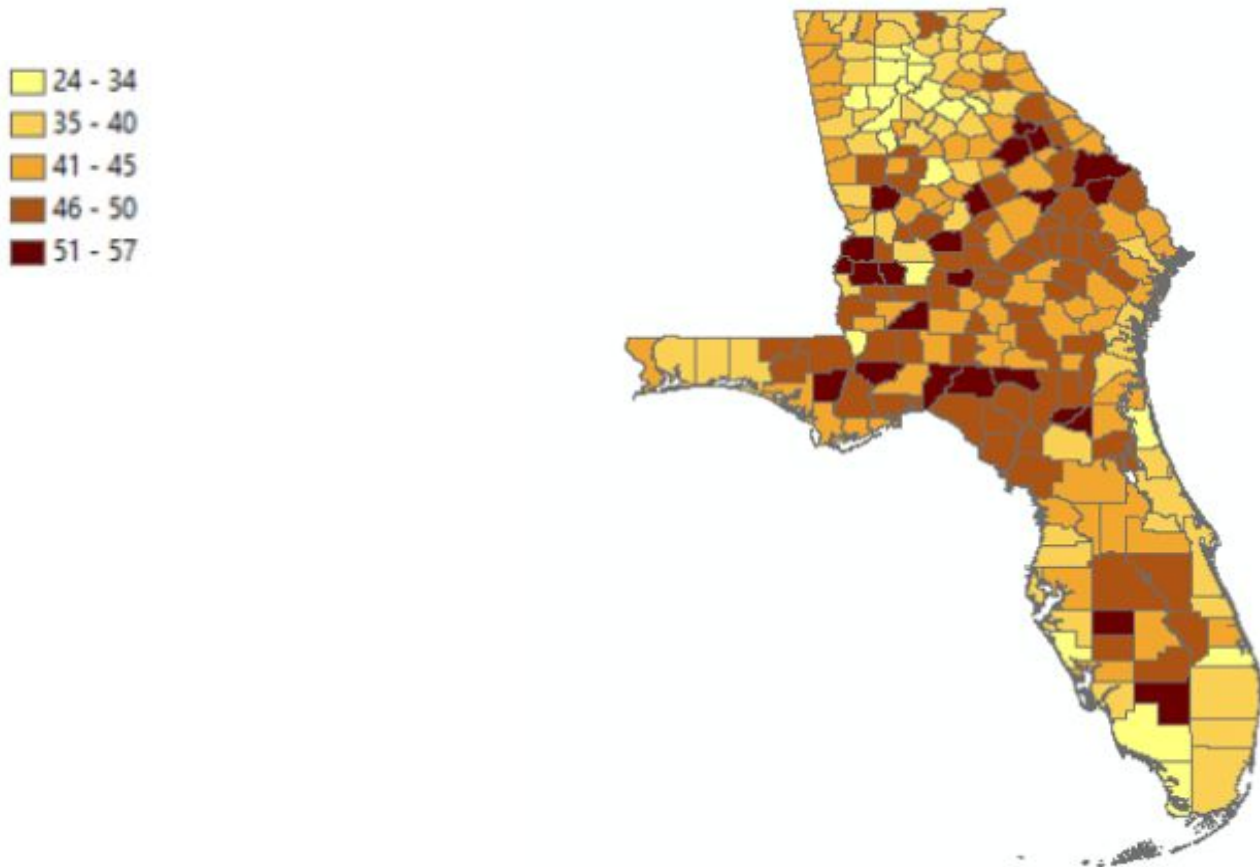
For the sake of inclusiveness, we have also included average female obesity rates. As stated above, knowing the obesity rate of an area can give some insight into the health, exercise habits, and food culture of an area. Details about why we used an average obesity rate and how we sorted our data into five classes are outlined in the previous paragraph.
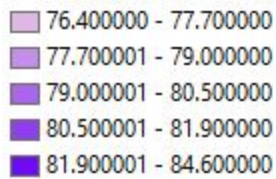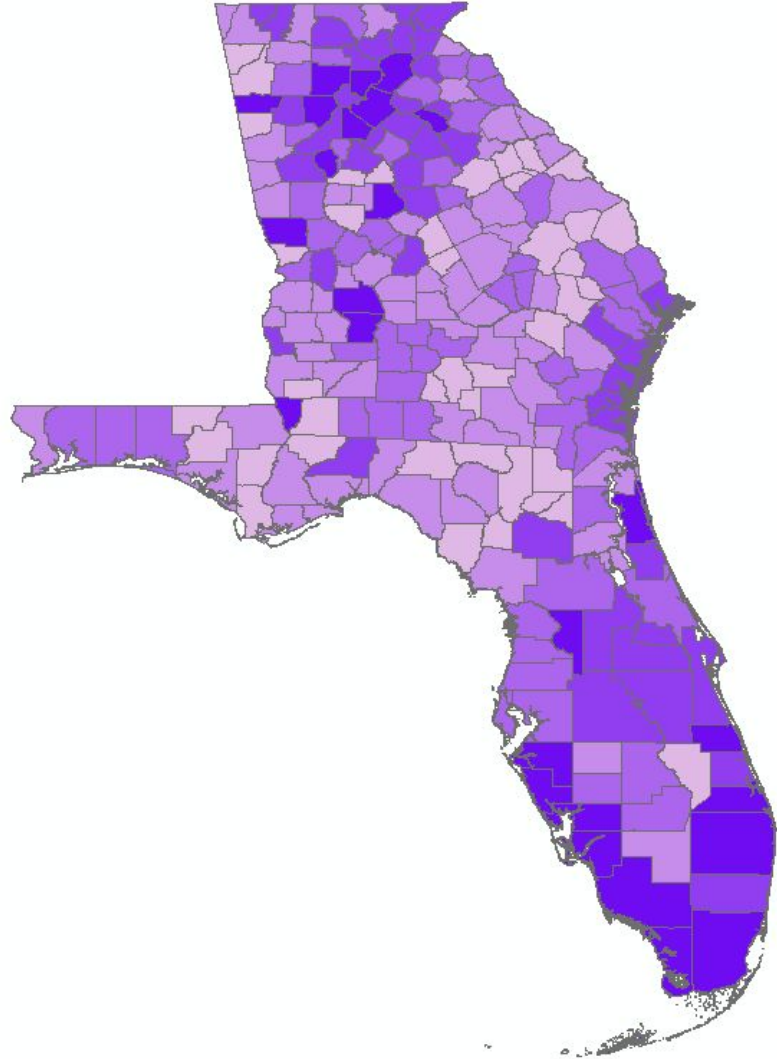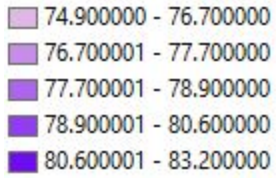
5a. Average Male Life Expectancy by County (years)

Malelifeex
- 68.800000 - 70.900000
- 70.900001 - 72.300000
- 72.300001 - 73.800000
- 73.800001 - 76.400000
- 76.400001 - 79.200000

Malelifeex
- 68.800000 - 71.700000
- 71.700001 - 74.000000
- 74.000001 - 75.900000
- 75.900001 - 77.500000
- 77.500001 - 80.100000

Life expectancy is the best overall indicator of the health standards of an area. This statistic can give insight into the sanitation, physical activity, rates of diseases, and quality of healthcare in a given area. Arguably, this is one of the most important statistics alongside disposable income in determining where the best place to live is. Unfortunately, this data was only available by county, which can only help us determine the best county to live in. However, we were able to find separate data for male life expectancy and female life expectancy so that someone other than our chosen persona could use the data if they wanted. Average female life expectancy data will be shown in the next section.

As always, this data set was divided into five classes using natural breaks in the data via the ArcGIS program. Because we're forcing relatively normally distributed data into five distinct color categories, the most accurate way to describe the data on a map is to use natural breaks. However, it's important to remember that because we are using different class sizes for Florida and Georgia, the colors representing data for one state might be different for another state. Also, it's important to note that this data was only available in the form of averages, instead of median values which are usually more representative of the typical person.

5b. Average Female Life Expectancy by County (years)

74.900000 - 76.700000
76.700001 - 77.700000
77.700001 - 78.900000
78.900001 - 80.600000
80.600001 - 83.200000

76.400000 - 77.700000
77.700001 - 79.000000
79.000001 - 80.500000
80.500001 - 81.900000
81.900001 - 84.600000

As stated above, life expectancy is an important statistic because it allows the interpreter to gain insight into the physical activity, food culture, sanitation, and standard of healthcare in an area. We have included both male and female life expectancies so that a male or female could use the GIS program and data to help determine the best place to live. Also, women tend to live longer than men, which means that an area with a much higher ratio of women to men might have a higher life expectancy, which is a hidden factor. This way, the users can see the pure data before the amount of women and men in an area can affect the results of life expectancy. Details on why the data was separated into these five classes and why our data is at the county level are shared above in the previous section.
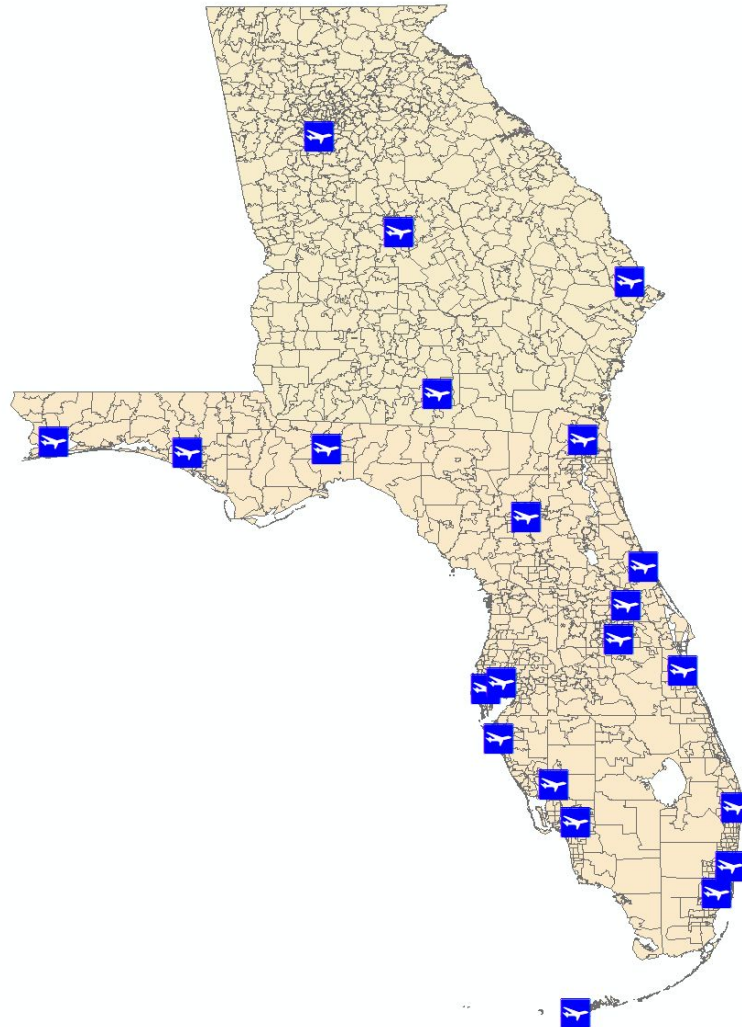
Part C: Proximity to Airports, National Parks, and Major Roads

6. Proximity to Airports (miles)

Being close in proximity to an airport can be a major factor in deciding where a person may want to live. Often times, people will have to travel for work, which means it might be optimal to live within a certain distance from an airport because not everybody has a car or can afford to take a taxi service to a far-away airport. Additionally, the more accessible an airport is, the more accomodating it is when travelling for recreational purposes.

For this data, airports are listed as points on the map rather than taking up an area like a county or zip code would. We did this because it might be misleading to show counties that possess a major airport because, as said before, the goal is to live within a certain specific achievable distance from an airport. You could live in a county with an airport but still be too far away from it. Also, we felt that showing zip codes that possess major airports wouldn't be very helpful either because that method

would make it seem like you have to live in the zip code that has an airport, which is not true. It's entirely possible to be able to get to an airport that's in a different zip code. For those two reasons, we are showing this data in the form of points and we will use each zip code's proximity to the points to determine where the best place to live is. The proximity range for airports that we will use for our analysis will be explained later in this report.

7. Proximity to National Parks (miles)
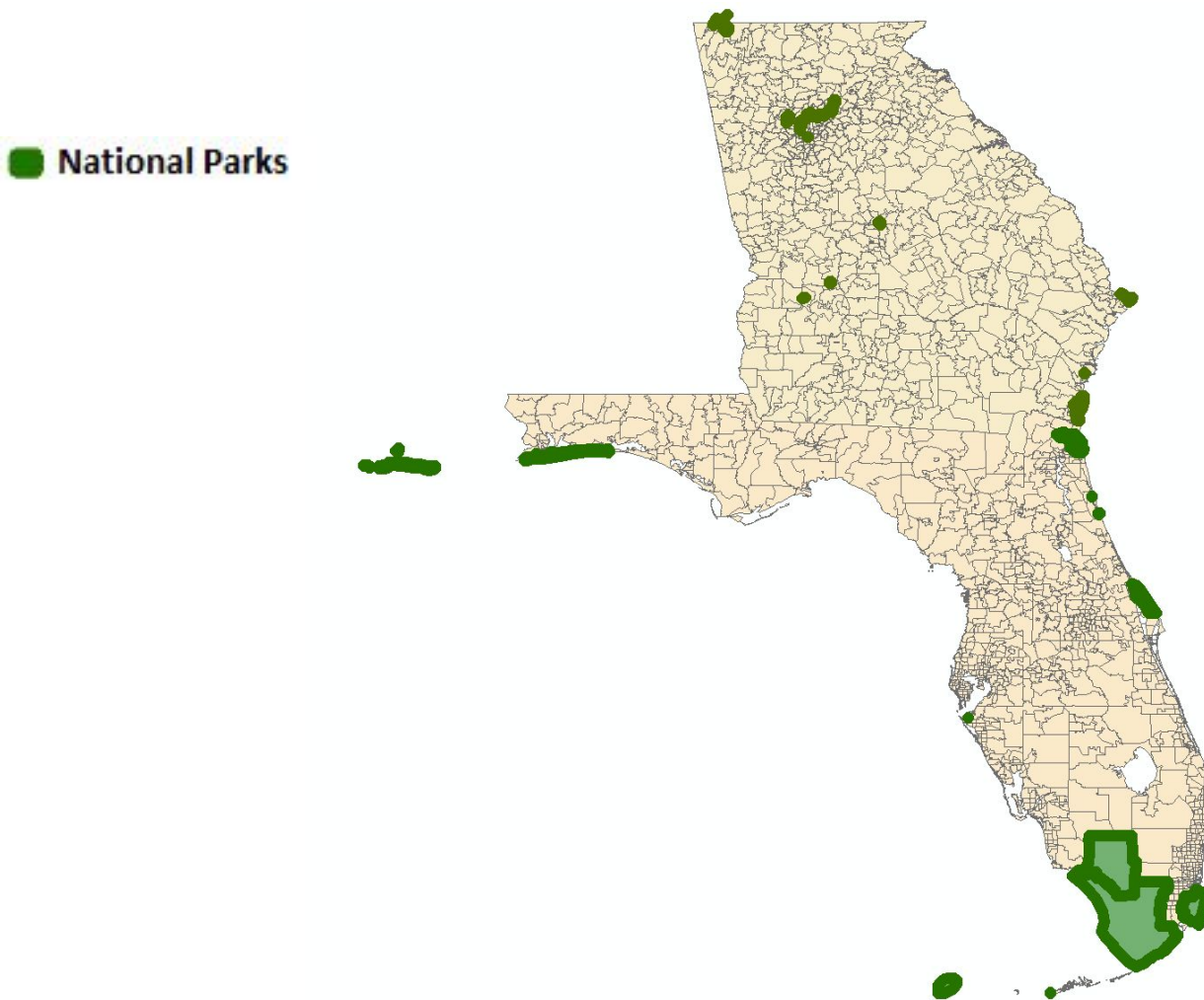
Another important aspect that can affect one's decision on where is best to live is the beauty of the area itself. More specifically, national parks can be a very good place to live near because they provide an activity to do in your free time and also introduce tourism to the area, which leads to a more positive economy in the area.

National parks are naturally defined as areas, so we will be narrowing down our decision on where to live by taking into account a county or zip code's proximity to one of these national parks. Of

course, this is not the most important of the data sets, but it can definitely be important to outdoors enthusiasts or act as a tiebreaker for areas that might be close in their scores from other data sets. The proximity range to national parks that we will use for our analysis will be explained later in this report.

8. Proximity to Major Roads (miles)


— GA Major Roads
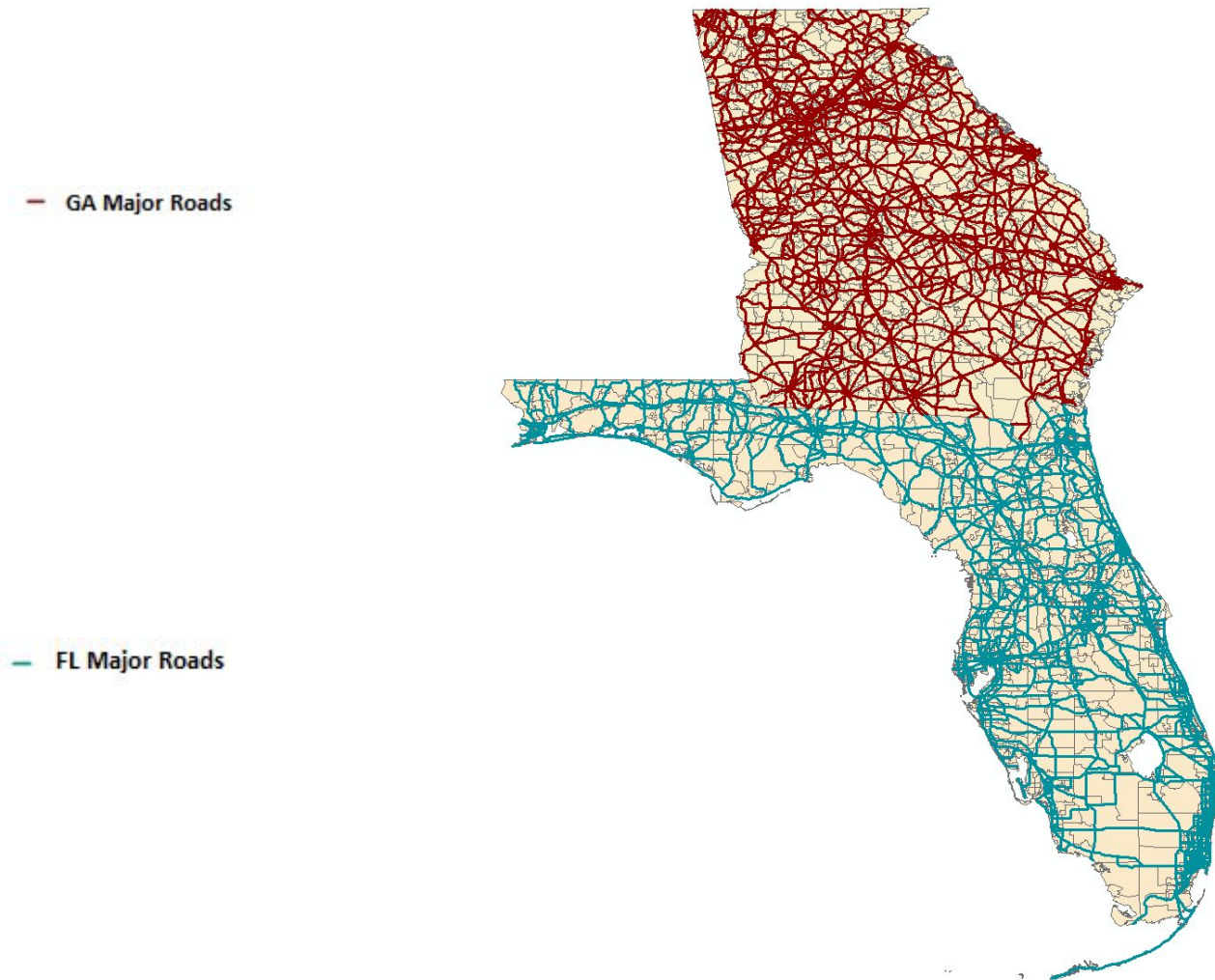
— FL Major Roads

       Like airports, it can be very important to live close to major roads. Depending on your job, you may be in a situation where you need to commute a long distance to work or travel for your job, which means that being near a major road would really help make your life more efficient. Although most counties have a major road within them, not all zip codes do, which we will use for our analysis. Instead of counting the major roads present in each county or zip code, we've decided the best way quantify our data is to show each zip code's proximity to the nearest point on the line that makes up a major road. All proximity values, just like with the other data sets, will be calculated using the ArcGIS program.

       Our determination of the desired proximity to a major road will be explained later in this report.
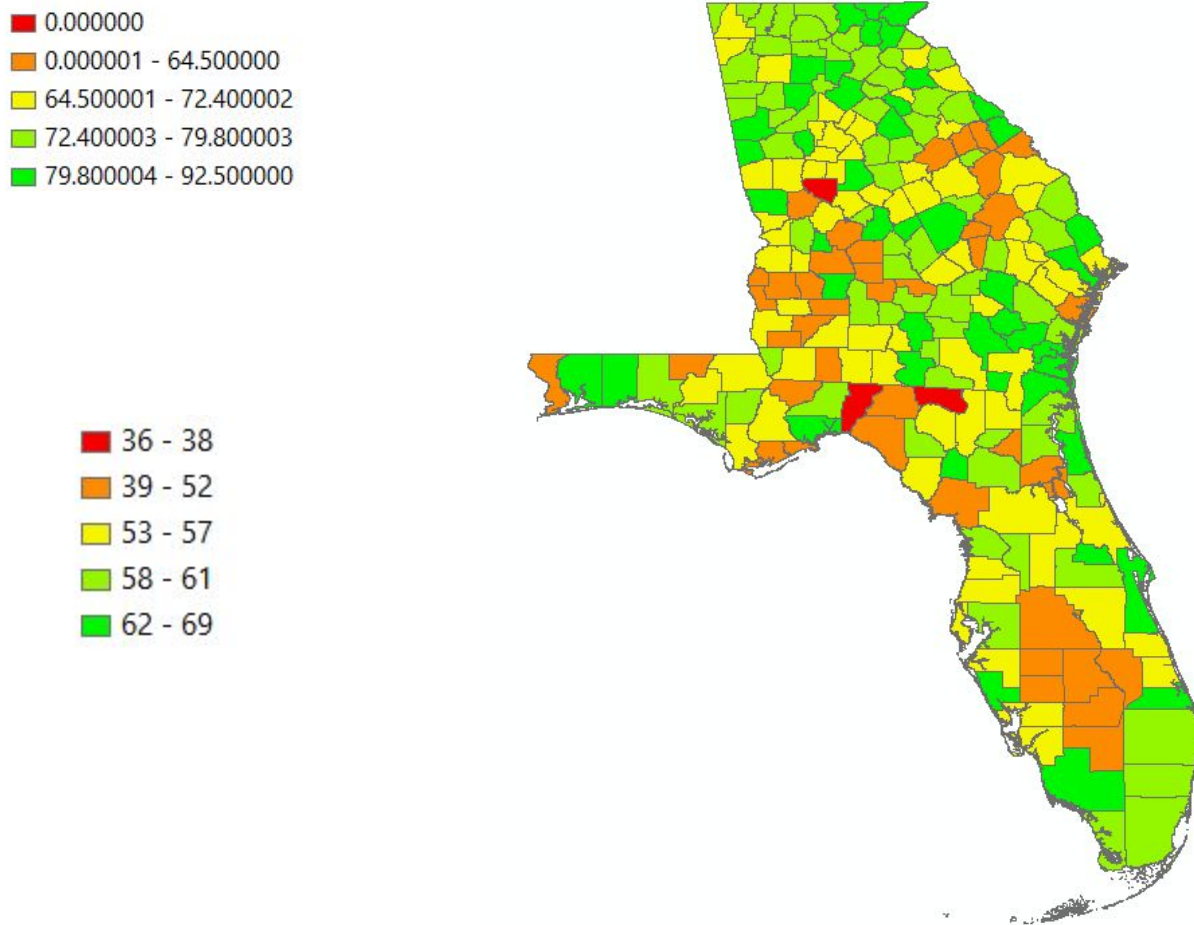
Part D: Community

9. Crime Index per County
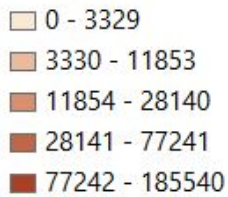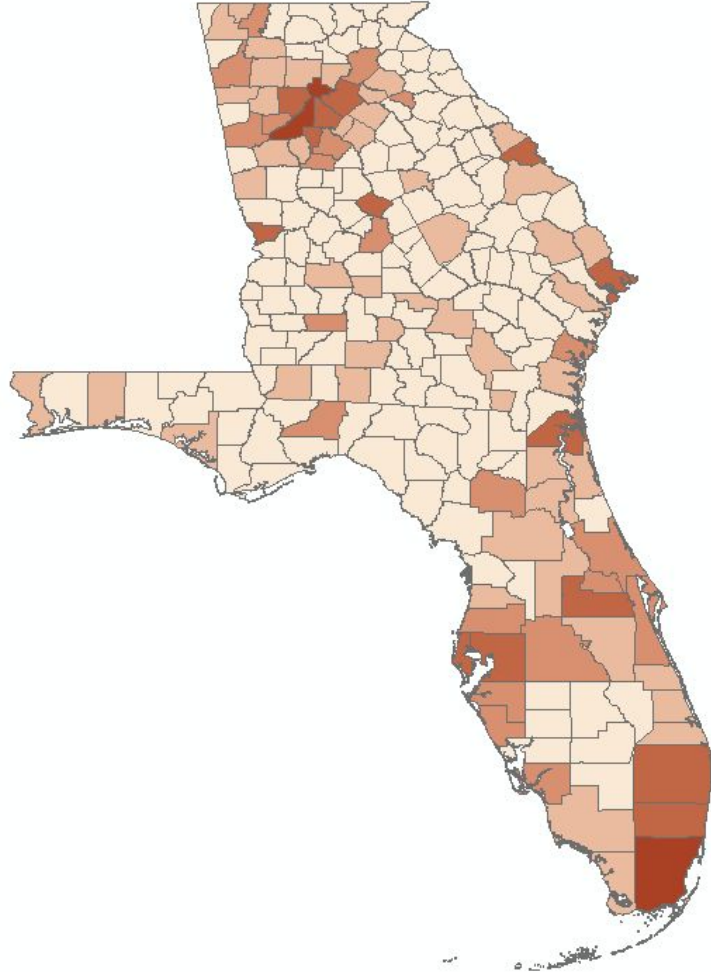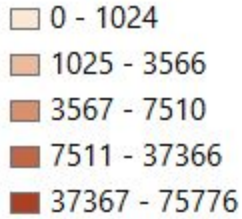
Living in a safe area is perhaps one of the most important factors in determining where you might want to live. An area with a low crime index allows a person peace of mind when they go to bed at night. Often times, there is not enough crime in a certain zip code to get a good enough comparison between zip codes, which is why crime is usually measured at the county level.

Like with some of the previous data datasets, we used natural breaks in the data determined by the ArcGIS program to sort the crime indices into five distinct classes in the most understandable and accurate way possible. However, it's important to remember that because we are using different class sizes for Florida and Georgia, the colors representing data for one state might be different for another state. The determination of the desired values for crime indices will be explained later in this report.

10. Average Grading of School Districts per County

- ■ 0.000000
- ■ 0.000001 - 64.500000
- □ 64.500001 - 72.400002
- ■ 72.400003 - 79.800003
- ■ 79.800004 - 92.500000

- ■ 36 - 38
- ■ 39 - 52
- □ 53 - 57
- ■ 58 - 61
- ■ 62 - 69

        The quality of school districts can be an important factor for determining where to live for people with children. Much of the time, parents will go incredible lengths to make sure their child has a good chance at a prosperous future, for that reason it was mandatory that we include this so that someone wanting to use our program can take this into account.

        As usual, the data was sorted into five distinct classes for each state using natural breaks in the data determined by ArcGIS. However, it's important to remember that because we are using different class sizes for Florida and Georgia, the colors representing data for one state might be different for another state. Average-type data was the only available type of data orientation available, which is also a good because median-type data would be misleading because there are not many school districts in every county. The exact grade that we will use to narrow down our selection of the best place to live will be explained later in this report.

Georgia legend:
- 0 – 900
- 901 – 2637
- 2638 – 5408
- 5409 – 9451
- 9452 – 19725

Florida legend:
- -999 – 1106
- 1107 – 2593
- 2594 – 4249
- 4250 – 6590
- 6591 – 11995
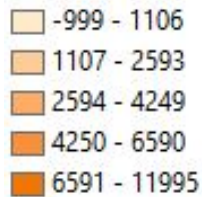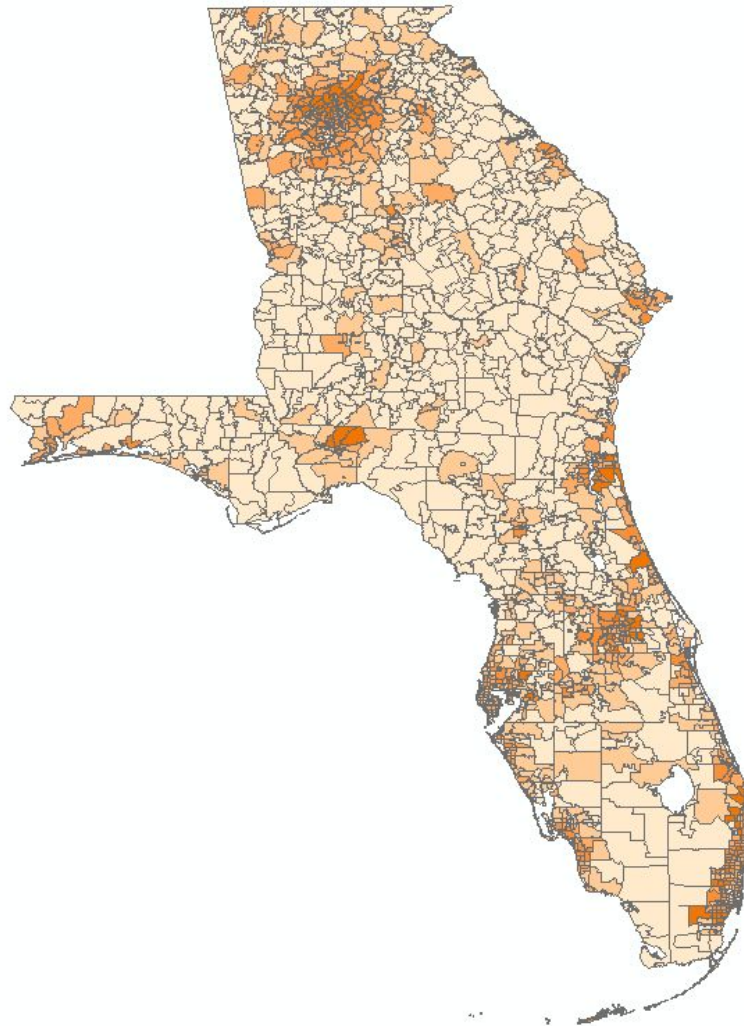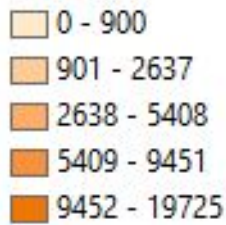
11. Amount of People with a Bachelor's Degree per Zip Code

More specifically to our person, who is in their 20's, finding an area with a large amount of other people who have a bachelor's degree is very important. First off, living in an area where a lot of people have bachelor's degrees means that maybe there are lots of employers who hire people straight out of college or just people with only bachelor's degrees. In an area with a high amount of bachelor's degrees, a person in their 20's might better be able to find a job or make friends.

Luckily, this data was collected on a zip code level, meaning that it helps is pinpoint the best place to live more accurately. For each of the two states, our data was divided into five classes using natural breaks in the data determined by the ArcGIS program. This way, the five different levels shown on the graph most closely corresponds to five different levels in the real data. However, it's important to remember that because we are using different class sizes for Florida and Georgia, the colors

representing data for one state might be different for another state.A more detailed explanation of what level we specifically desire will be given later in this report.

12a. Percentage of Males Completing Sufficient Physical Activity by County (%)
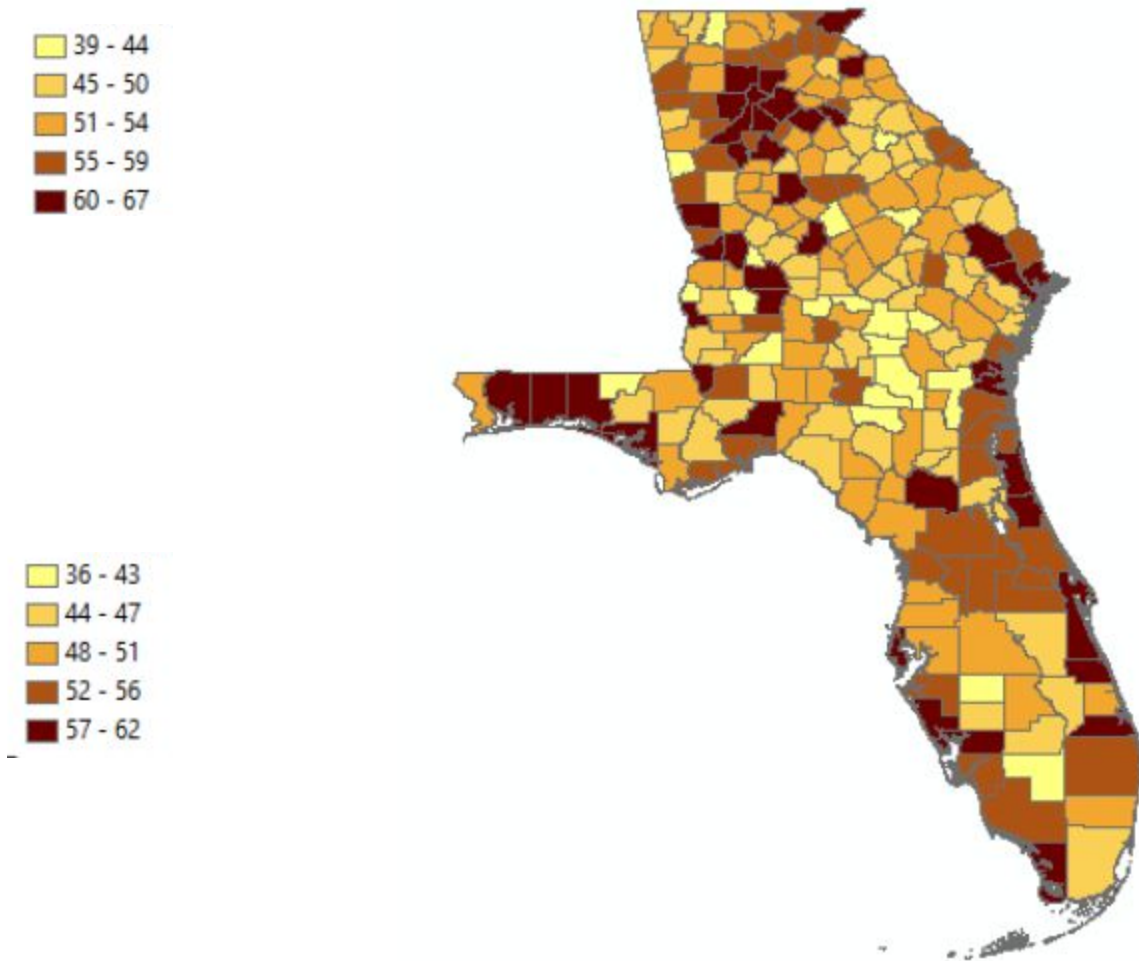
Physical activity is more than just a hobby for many people. It's a lifestyle. A person, such as ours, who enjoys physical activity might want to live in an area with a higher percentage of the population being active. This way, that person can meet more people who have similar hobbies, make more friends, and assimilate to the culture easier. As well, this statistic might give some insight into the quality of the local infrastructure to support outdoor activities such as running or hiking. For that reason, we're including physical activity as one of the indicators that will help us decide where the best place to live is.

Just like with most of our data, we divided the data for each of the two states into five classes using natural breaks in the data determined by the ArcGIS program. This way, the colors on the map can

best represent different levels in the actual data. However, it's important to remember that because we are using different class sizes for Florida and Georgia, the colors representing data for one state might be different for another state.

12b. Percentage of Females Completing Sufficient Physical Activity by County (%)

☐	31.300000 - 39.300000
☐	39.300001 - 44.000000
☐	44.000001 - 48.500000
■	48.500001 - 54.000000
■	54.000001 - 60.000000

|	
☐	40.100000 - 45.600000
☐	45.600001 - 49.300000
☐	49.300001 - 53.300000
■	53.300001 - 56.700000
■	56.700001 - 64.200000
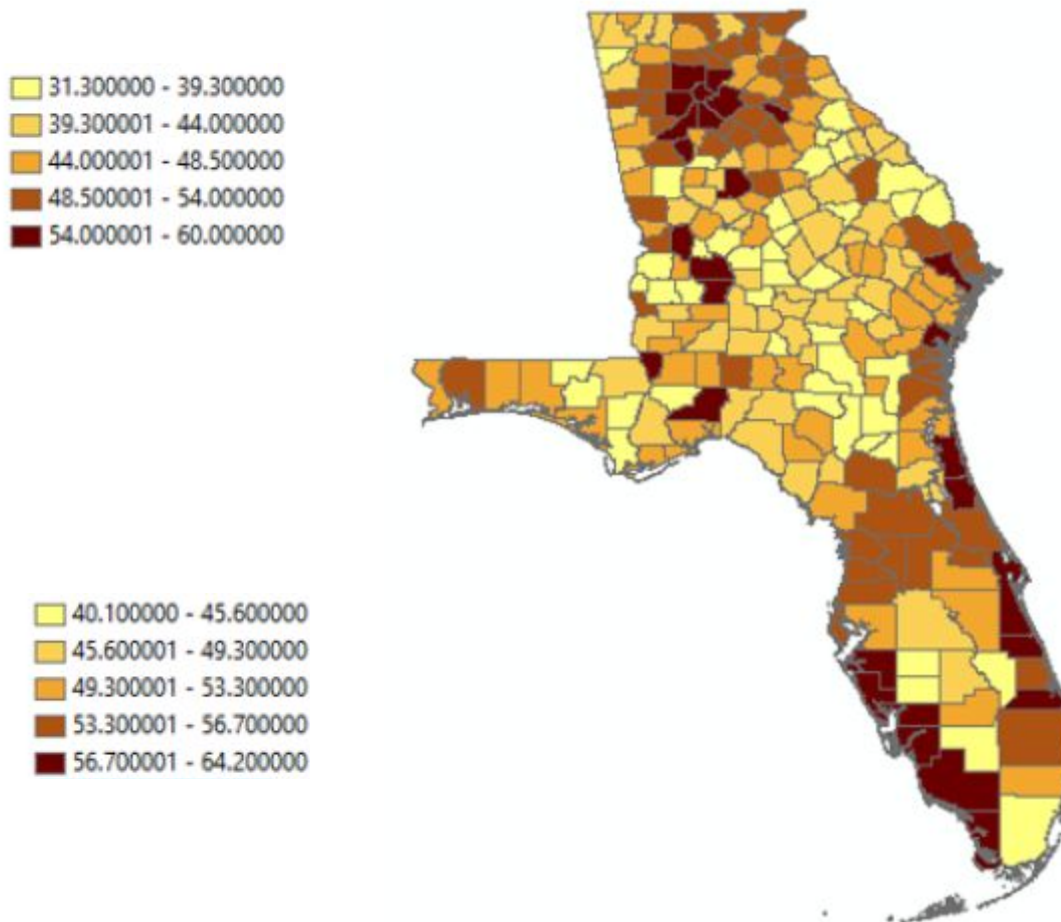
        For the sake of inclusiveness, we included both male and female physical activity percentage data so that anyone other than our chosen persona could use our data and GIS program to get the result specific to them. As well, a person might be more interested in finding one gender's physical activity amount relative to the other for whatever reason. It's always possible that a man might use this data to find where he could meet the most young active females or vice versa.

        The data for this map was divided and classified the exact same way as the corresponding map for males and can be explained in the paragraph above. And, as always, it's important to remember that having two different natural break sets of classes for Georgia and Florida means that you should not compare colors across states because the same values might have different colors in different states.

Methodology Using GIS flow charts:
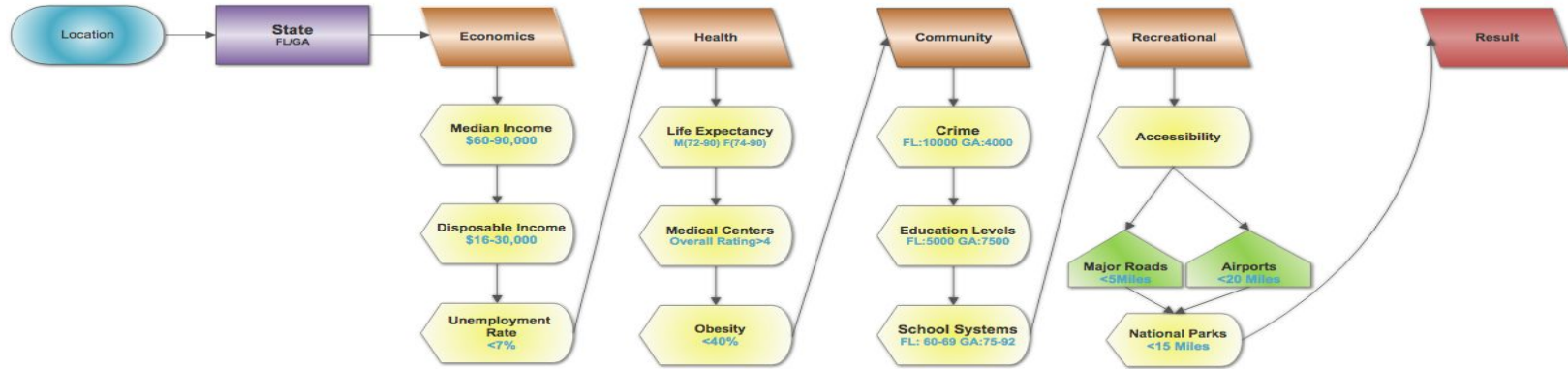
   We had data that was in the county form as well as the zipcode form. This mixture of data had provided us with additional challenges. We had to filter the data in the zip code format and in the county format, and once we had results for both, intersect the data to find the zip code data that had fallen within the county's that met our criteria.

   Once all of the data was collected and put into the proper format, we added it to the ArcMap, one state at a time. We used the feature of "Select by Attributes", creating a new selection from the Median Household Income. We then selected from the current selection the criteria we wanted for the Disposable Income, and then the Education status of the people in the community. These were all indicators that data had been found in zip code format. For the county data, we began with the Unemployment Rate of the county, creating a new selection and then proceeding with the health indicators. From there, we selected from the current selection, refining it with the crime rates and school systems grades. Furthermore, we used the feature of "Select by Location" since all of our recreation indicators were location based. We narrowed our results to counties within range of an airport, major roads, and national parks, then did the same for our zip code results up to that point. Once we had the data for all of our counties and zip codes separately, we intersected the data using the tool "Select by Location" finding both counties intersected with zip codes, and zip codes intersected with counties that were selected.

Indicator	Parameters
Median Income	Our desired value for median income was $75,000, which was about the average income that our four members could expect to earn. We decided to use a range of $60,000 to $90,000 to include enough of a variety. We also found that it was an appropriate value that would highlight zip codes that were above average but not too exclusive.
Disposable Income	Disposable income for Florida and Georgia are relatively low relative to the rest of the US. For that reason, we chose a value starting at $16,000 and ending at $30,000 which also fell within the range of above average but not too exclusive.

Unemployment Rate	Unemployment rates in Florida and Georgia are pretty high, so we aimed to find an employment rate that was slightly below the national average. For that reason, we chose 7% because it was below average and was a nice round number that was within the lower 40% of our data.
Life Expectancy	Life expectancy ran differently for both genders. We decided to implement the data separately in order to have a fair result. For females, the range is set a little higher and it starts at 74 and goes to 90 years old. For males, it starts at 72 and goes to 90 years old.
Medical Centers	The data found within this indicator included a variety of evaluations of the medical centers. Due to lack of correlation between them, we decided to stick to the overall rating to ensure the prime results. Additionally, the ratings were placed high due to the excess of medical centers meeting the rating of 3 out of 5. For this reason, a rating of 4 out of 5 was placed.
Obesity Rates	Obesity between Florida and Georgia is high. Initially, we wanted to set the population to have an obesity rate of less than 30%. After conducting a spatial analysis, we noticed that there were no results meeting this requirement. So this is when we decided to increase the parameter to 40%.
Crime Rates	We chose the crime rates based on the range that the state had. We desired an area that was within the 75th percentile of safety in the state.
Education Levels	We measured the amount of people that had a Bachelors degree and chose the numbers 5,000 (FL) and 7,500(GA) due to the data that was presented to us. These numbers were greater than 75% of the data.
School Systems	For Florida, the school rating index ranged from 0-69, therefore we wanted school districts rated 60 or higher. Georgia was based on a 0-100 scale, however the highest grade was 92.5, therefore we wanted schools rated greater than 75.
Airports	In this day and age we want to live somewhere with high accessibility, so we narrowed down our selection by selecting the counties that are within 20 miles of a public airport.
Major Roads	Since we want to live in the city where there is high accessibility, we decided we want to live within 5 miles of a major road.
National Parks	Though we value accessibility and the city, we are all very outdoorsy people who enjoy visiting national parks in our free time. We decided that we want to live in a county that is within 15 miles of a national

|	
	park.
Sufficient Activity	The sufficient activity indicator was added when we needed to narrow down results further and measured the percent of people who completed sufficient physical activity. We decided to make it more than 45% considering the range was 31%-67%.

Results and Discussions:

When observing the final results of the spatial analysis, the impact of our GIS analysis process can be truly cherished. The method of "funneling" the criteria through measurements and, furthermore, indicators narrowed down the two adjacent states to 21 counties and 32 cities within these counties.

Initially, there were significantly more results. We decided to tighten the initial ranges of our criteria. Firstly, the threshold for the accessibility indicator was narrowed down for the distances from major roads and airports. Additionally, hospital ratings were increased to meet a rating of 4 out of 5 instead of 3. In contrast, the parameters set for obesity had to be tweaked due to the lack of overlap with previous intersections. Originally, we started with an upper limit of 30 percent, but had to increase it to 40 percent in order to proceed since there were no results.

The zip codes that we ended up selecting were within the following counties and cities:

Florida		Georgia	
Counties	**Cities**	**Counties**	**Cities**
Collier	Bradenton	Cobb	Atlanta
Duval	Destin	DeKalb	Decatur
Escambia	Fernandina Beach	Douglas	Douglasville
Hillsborough	Goodland	Fayette	Ellenwood
Manatee	Gulf Breeze	Fulton	Fairburn
Miami- Dade	Homestead	Henry	Jonesboro
Monroe	Jacksonville	Houston	Mableton
Nassau	Jacksonville Beach	Rockdale	Macon
Pinellas	Marco Island		Rex
Santa Rosa	Miami		Savannah
Sarasota	Neptune Beach		Smyrna
Volusia	Niceville		Stockbridge
Walton	Ponte Vedra		Warner Robins
	Port Orange		
	Saint Augustine		
	Saint Petersburg		
	Sarasota		

	Shalimar Summerland Key		
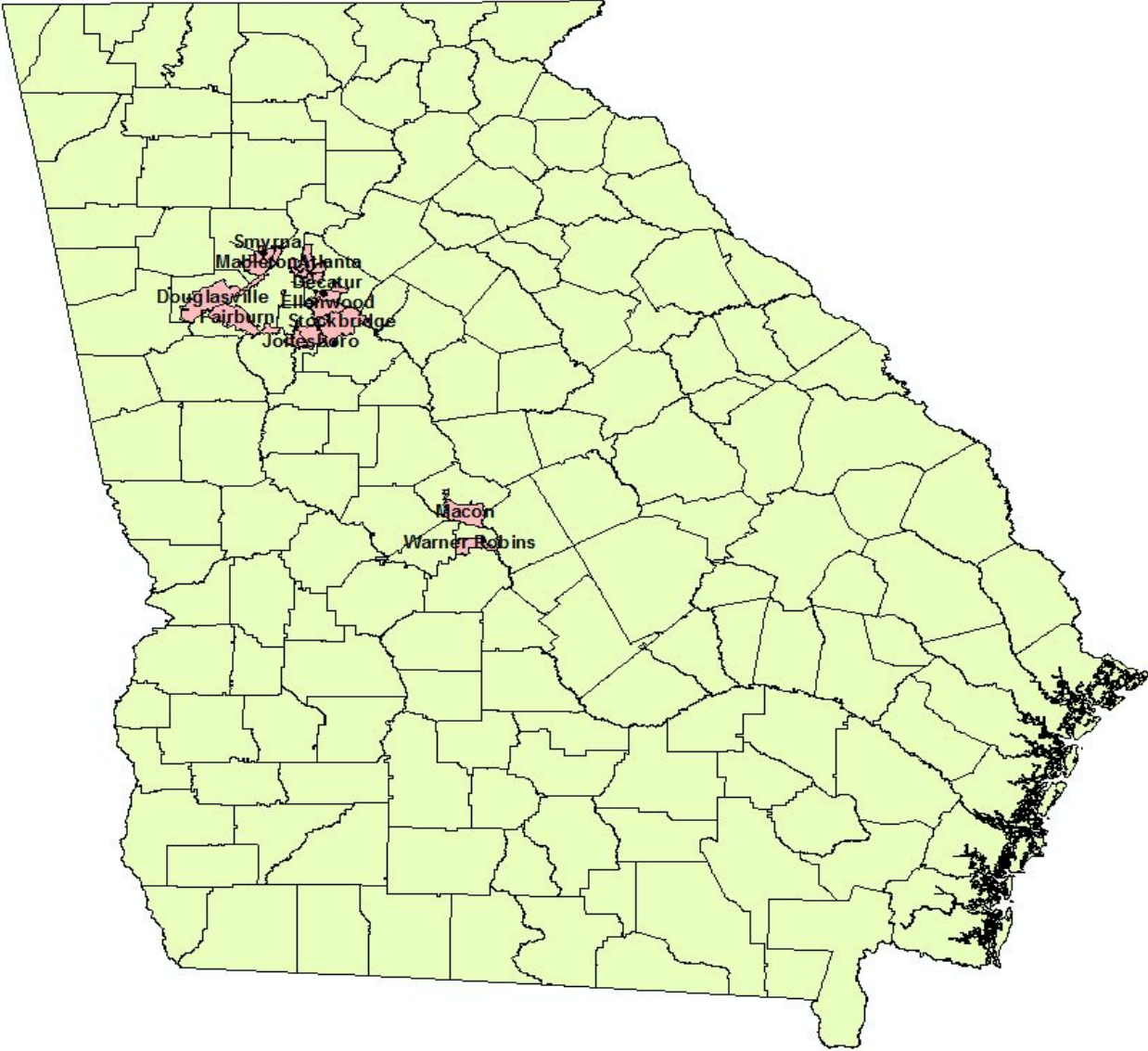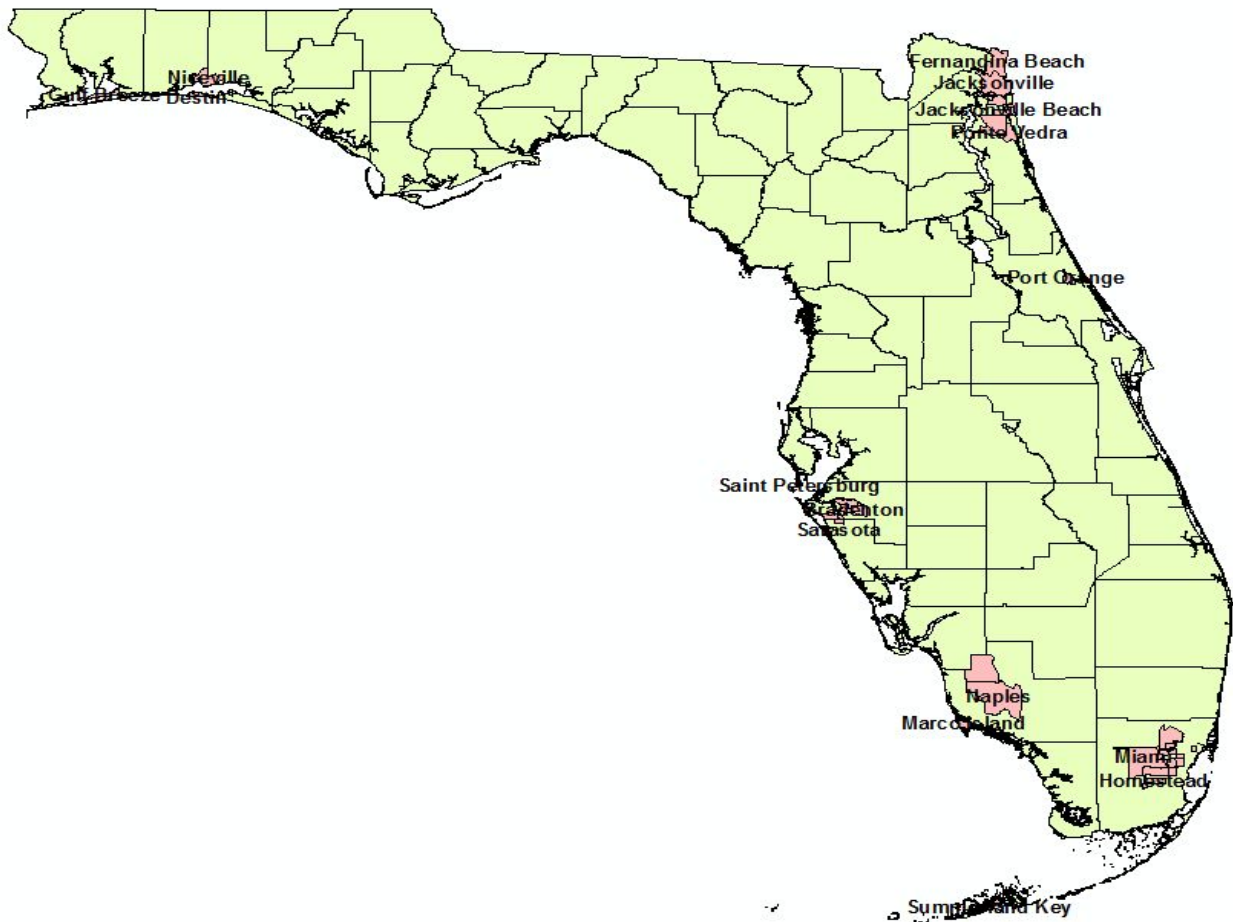
Conclusions

The focus of this project was to set out on an exploration for the best place to live within a combined total of 112,000 square miles of land. Through a straightforward analysis process and concise set of parameters, data was funneled down to result in a collection of strong locations that provide an ideal living situation. The balance of priorities within our measurements have lead us to end up with a strong set of results that leave us with comfortable and profitable parts of the respective states. Within the spatial analysis, the conflict between having data bounded by zip code and county led us to choose zip codes that were within a county that fit our criteria as well as satisfying the criteria of the zip codes.

The major issue with our data coming from both county and zip code is having information that may not be accurate when looking into specific zip codes. If within our data we had used an average for the entire county, it doesn't mean that the zip code is representative of the entire counties average. This could mean that there were zip codes that were within our criteria that were discontinued due to their respective counties not meeting our criteria. Because of other zip codes within that county that were further below our criteria, the average was decreased overall and the county didn't register as meeting

the criteria. All things considered, we believe this may have harmed our data, but feel as though our methodology was proficient in producing desired results.