Sharita Underwood

Geospatial Database Design

8/21/2021

Introduction

Many of the public natural amenities in Cobb County Georgia are hidden behind and within subdivisions which make them hard to find and result in the amenity functioning as a private feature despite being maintained using taxpayers' dollars. To demystify these amenities and provide more transparency to the public, I have created a database, "Park_Division", that can not only locate these features but be used to explore their relationship with subdivisions in more depth. To do this I will first build my database in ArcGIS Pro then run queries and create maps as necessary to solve three questions which will prove the functionality of my database. The questions are as follows: how many subdivisions are within each city, which subdivisions are within one-quarter of a mile from a park, and which subdivisions within one-quarter mile from a park were built before in or before 2003. Before I proceed to answer these questions however, I will first discuss my database in depth.

Database Description

Obtaining Data

While Park_Division is a database that can analyze the relationship between subdivisions and public amenities, it did not start this way. My original goal was to analyze the relationship between parcels and public amenities however, parcel data in Cobb County is not publicly available and hence I had to rethink my strategy. I considered switching to a nearby county, Fulton County, who did have their parcel data freely available but that wouldn't have attacked the root of Cobb County's problem which I really wanted to target. While looking at Cobb County's Open Data Hub (CCODH) I was able to find subdivision data which was expansive and included 8,564 records so I opted to use this data instead. I also from CCODH downloaded the following shapefiles: parks, boundaries, trails, and bodies of water.

Cleaning Data/Attributes

Once I downloaded these five shapefiles, I cleaned each one and only kept relevant attributes. Each of them had five attributes in common which I kept. Those five attributes were *object_ID*, which is a primary key for each of the tables, *shape* which defines what kind of features the feature class holds, *name*, which is the name of the record, and *length and area*, which are both derived from shape. Separately the boundaries feature class has *full_name* which is a composite field composed of *name and state* and the parks feature class has *parkType*, which defines if the park is a county owned park, a federal park, or a park that the county is leasing from the federal government. The park feature class also has *TotalAcres*, which is the amount of acres each park covers. The subdivision feature class had *taxyr*, which is the year the subdivision first appeared on tax records and the bodies of water feature class has *water_type* which describes what kind of body of water the water is. In figure one, you will find a picture of this display which is color coded to indicate what kind of attribute each field is.

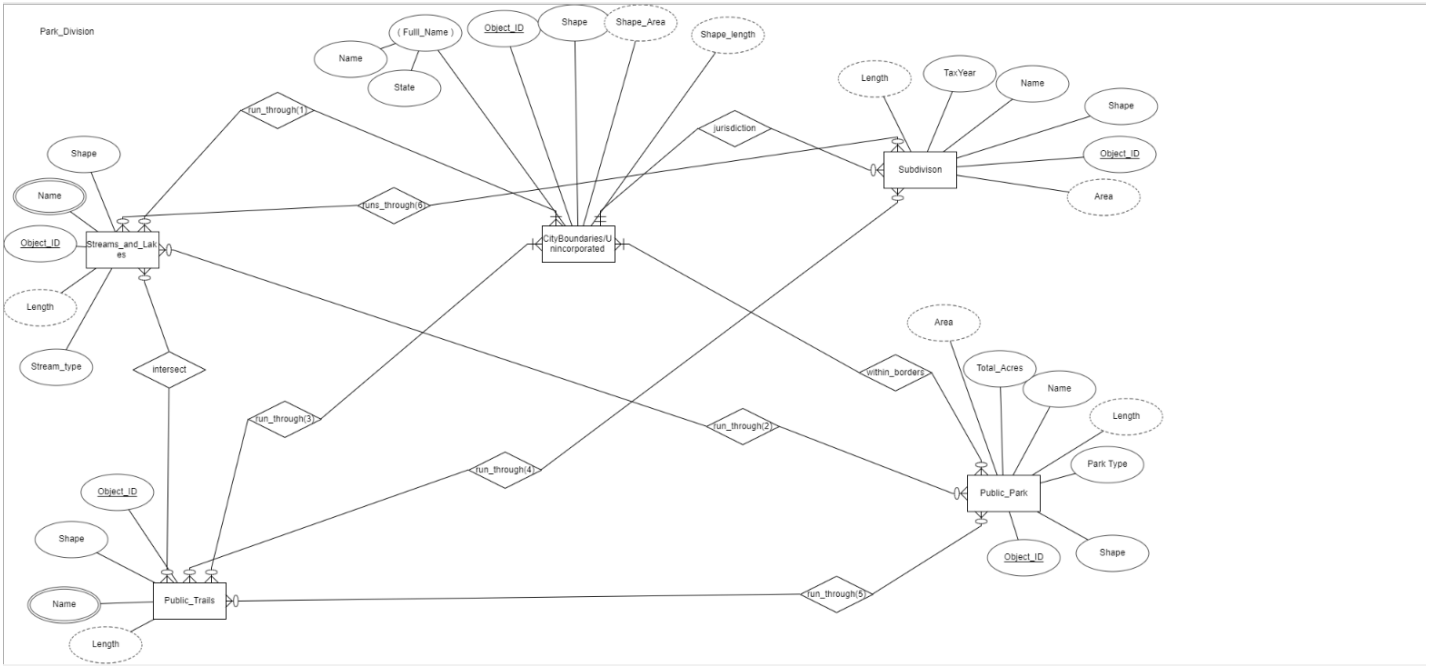
Feature Class Name	Attribute 1	Attribute 2	Attribute 3	Attribute 4	Attribute 5	Attribute 6	Attribute 7
Boundaries	Object_ID	Shape	Full_Name	Name	State	Length	Area
Parks	Object_ID	Shape	ParkType	Name	TotalAcres	Shape	Area
Subdivisions	Object_ID	Shape	Name	TaxYr	Shape	Area	
Trails	Object_ID	Shape	Name	Length			
Body of Water	Object_ID	Shape	Name	Water_Type	Shape		

Primary key/Requried *Composite* *Derived* *Multivalued* *(figure 1)*

**Note the Name for the Trails and Body of water feature classes are multivalued due to both having the ability to change names at various spots*

Database Conceptual Design

Figure two details the relationships between my feature classes and the Park_Division



(figure 2)

Database Conceptual Design Descriptions:

The following is a list of the spatial relationships that exist between entities as shown in the above image.

Boundaries can have jurisdiction over zero or many subdivisions.

Subdivisions can only be in the jurisdiction of one boundary.

Boundaries can have zero or many public parks within their borders

Parks can be within one or many boundaries.

A boundary can have zero or many public trails.

A trail can run through one or many boundaries.

A boundary can have zero or many bodies of water.

A body of water can run through one or many boundaries.

A body of water can run through zero or many subdivisions.

A subdivision can have zero or many subdivisions running through it.

A body of water can run through zero or many parks.

A park can have zero or many subdivisions running through them.

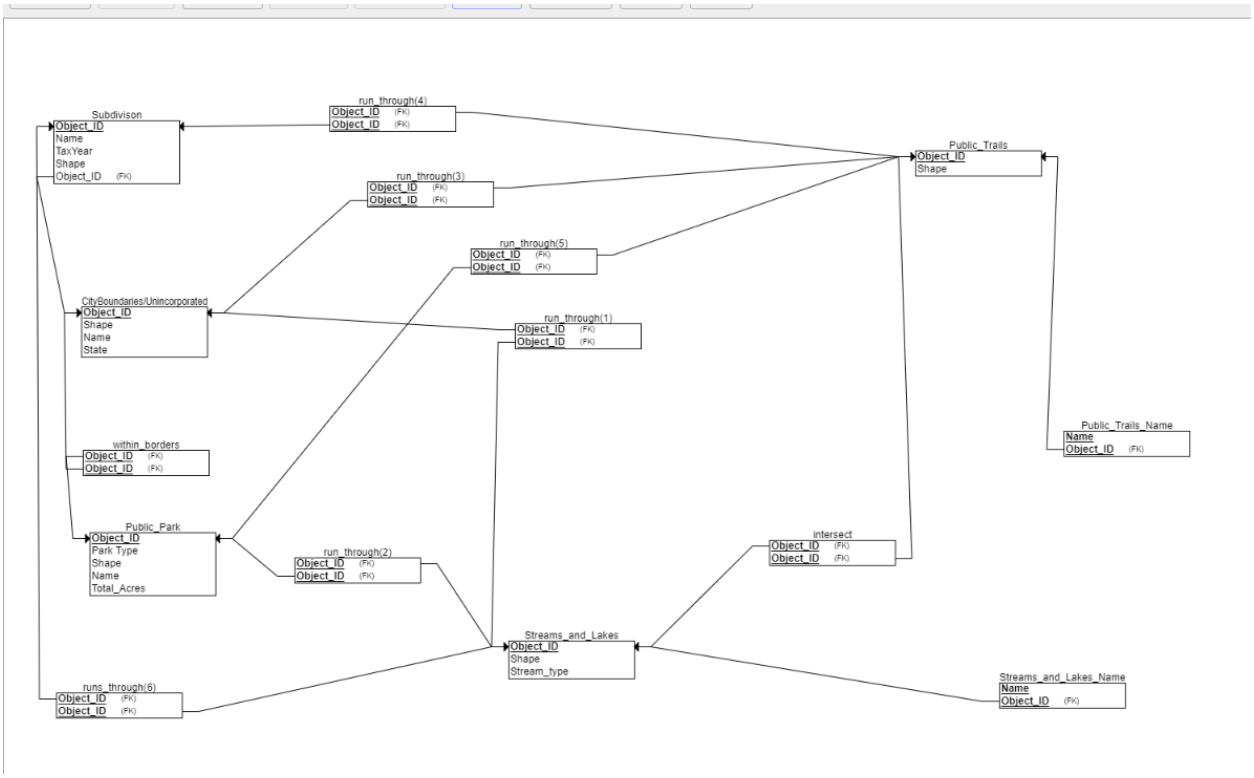
A trail can run zero or many subdivisions.

A subdivision can have zero or many trails running through it.

A body of water can intersect zero or many trails.

A trail can intersect zero or many bodies of water.

Database Logical Design



(figure 3)

<u>Park_Division Design Description Implementation</u>

I implemented my database in ArcGIS Pro. I had to change some of the data types, lengths, and alias' of the original data to fit my needs which meant I used the geoprocessing tool

Feature Class to Feature to customize the data exactly how I wanted. The schema I followed is below.

Geodatabase design forms

Geodatabase Parks_Division

Feature Dataset	FeatureClassType	Name	Alias
Amenities	**POLY**	**Cobb_County_Boundaries**	**Boundaries**
	POLY	**Cobb_County_Parks**	**Parks**
	L	**Body_of_Water**	**Water**
	L	**Trails**	**Trails**
Residents_Lots	**POLY**	**Cobb_County_Subdivisons**	**Subdivisons**

(figure 4)

Datasets worksheet: Descriptions of all datasets

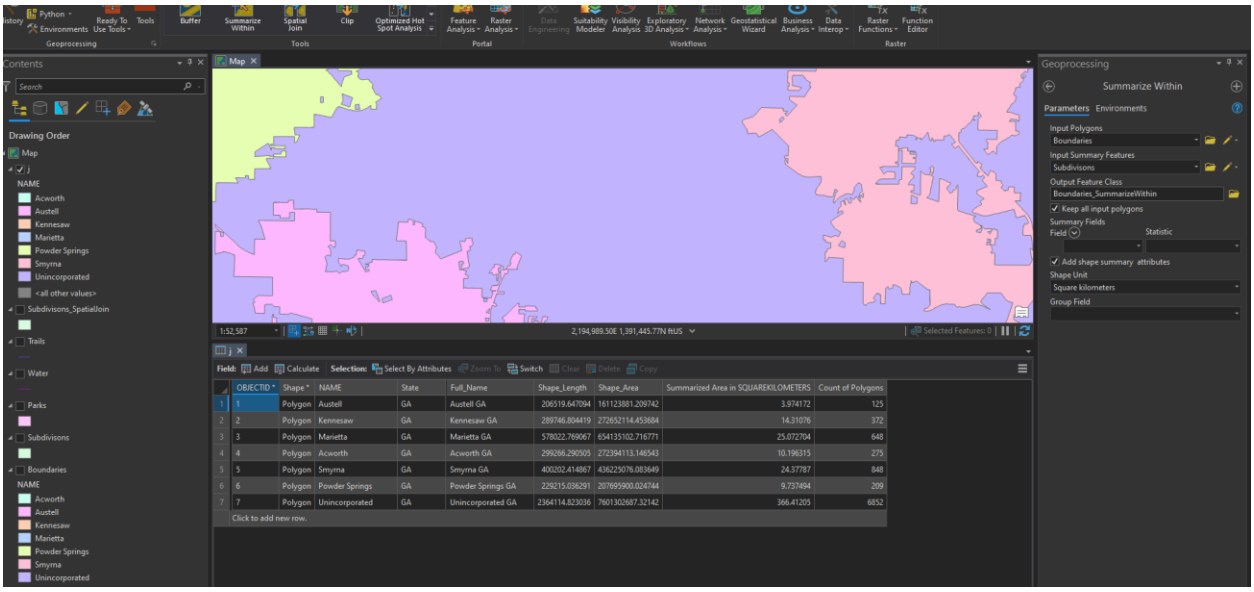
Featureclass or table name	Field name	Field type	Length	AliasName	Description	DomainName	DefaultValue	IsNullable(Y/	Subtype
Boundaries	Object_ID_1	SI	Auto	Object_ID	Primary Key			N	
	Shape	Geom	Auto	Shape	Geometry of record			N	
	Name	Text	14	Name	City/County Name			N	
	Shape_Length	Double	Auto	Length	length of polygon		0	N	
	Shape_Area	Double	Auto	Area	area of polygon		0	N	
	State	Text	7	State	State polygon is within			N	
	Full_Name	Text	50	Full_Name	Name + State			N	
Parks	Object_ID_1	SI	4	Object_ID	Primary key			N	
	Shape	Geom	25	Shape	Geometry of record			N	
	ParkType	Text	9	Type	Type of Park			N	
	Name	Text	30	Names	Name of Park			Y	
	Total_Acres	Double	50	Total_Aces	Total Acres of Shape			N	
	Shape_Length	Double	50	Length	Calculated length of shape		0	N	
	Shape_Area	Double	50	Area	Calculated area of shape		0	N	
Trails	Object_ID_1	SI	68	Object_ID	Primary Key			N	
	Shape	Geom	25	Shape	Geometry of record			N	
	Name	ST	14	Name	Name of Trail			Y	
	Shape_Length	Double	50	Length	Calculated length of shape		0	Y	
Water	Object_ID_1	SI	4	Object_ID	Primary Key			N	
	Shape	Geom	25	Shape	Geometry of record			N	
	Name	ST	30	Name	Name of body of water			Y	
	Shape_Length	Double	50	Length	Calculated length of shape		0	N	
	Type	Text	25	Type	Type of Water			N	
Subdivisions	Object_ID_1	SI	4	Object_ID	Primary Key			N	
	Shape	Geom	25	Shape	Geometry of record			Y	
	SUBDIVNAME	ST	50	Name	City/County Name			Y	
	TAXYR	SI	4	TaxYr	Year subdivision first appeared on tax records			Y	

(figure 5)

Database Implementation

Question 1: How many subdivisions are within each city?

To find how many subdivisions are within each boundary I used the summarize within tool. The summarize within tool stacks two feature layers, the input polygon and the input summary features, and counts the number of input summary features that fall within each record of the input polygon. In my case, the summarize within tool counted the number of subdivisions that were within each boundary and outputted the results in a table I named 'J'. Hence, my input polygon was boundaries and my input summary feature was subdivisons. The results of my query are below.

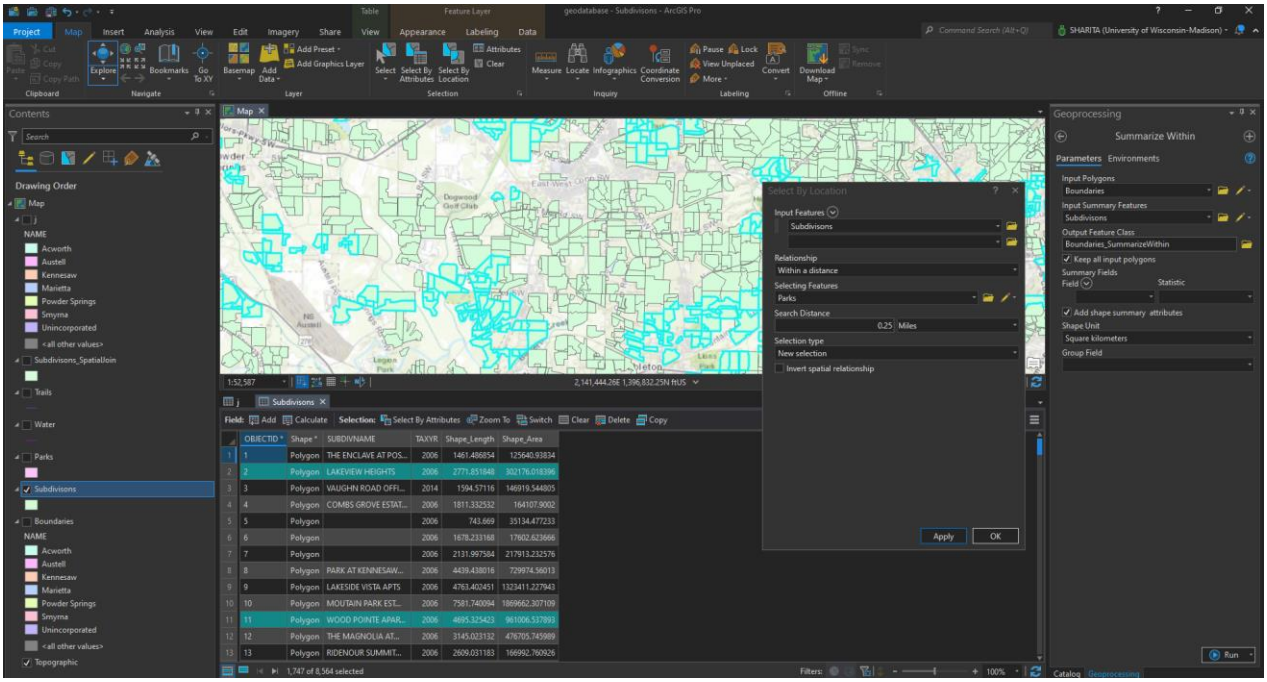


(figure 6)

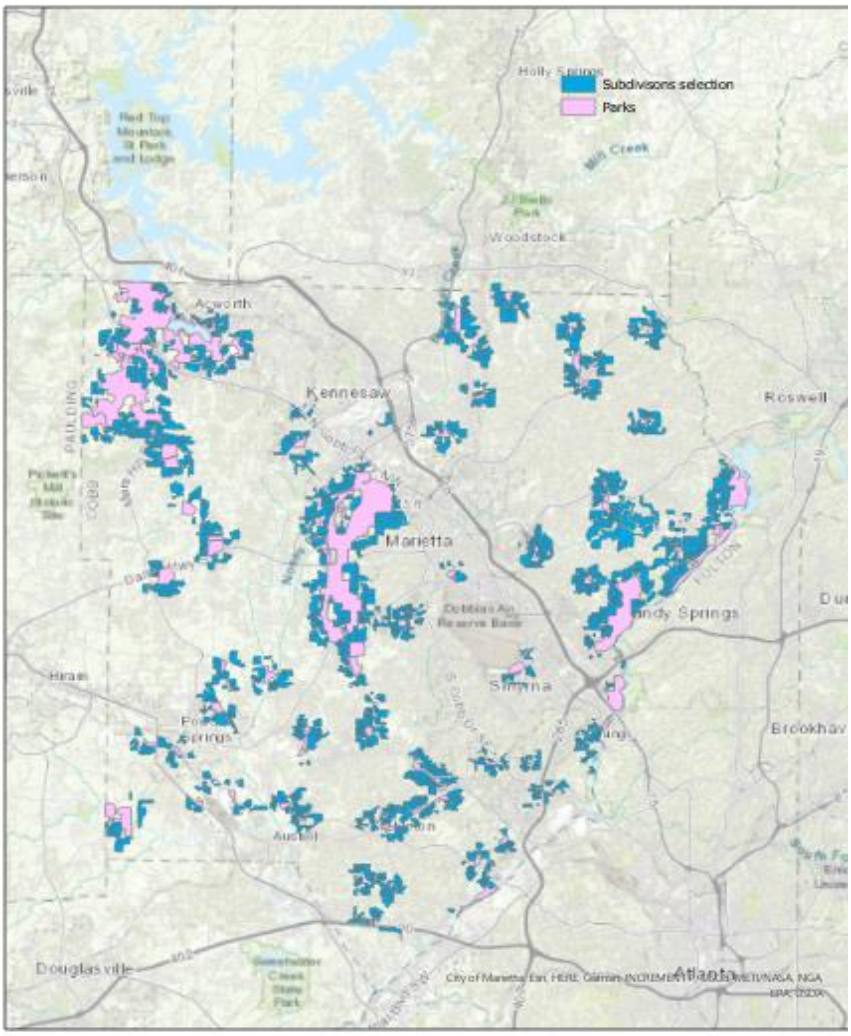
Question 2: Which subdivisions are within one-quarter of a mile from a park?

To answer my second question, I used the Select By Location Tool. The Select By Location tool looks at the spatial relationships between and amongst various features. The input feature is the feature you want to query, and the selecting feature is the feature which sets the parameters for your input feature. Since I wanted to observe the relationship between subdivisions and parks my input feature was subdivisions and my selecting feature was parks. Next, I set my relationship to within a distance and my search distance to 0.25 miles which operates like a temporary buffer around the parks and searches for subdivisions inside of that

buffer. As figure 7 shows, once this task was completed my attribute table for subdivisions showed that 1,747 subdivisions met this criterion. In figure 8 this same data is shown expect in map form which I created by making a layer from my selected features and creating a layout featuring my selected features and the parks layer.



(figure 6)

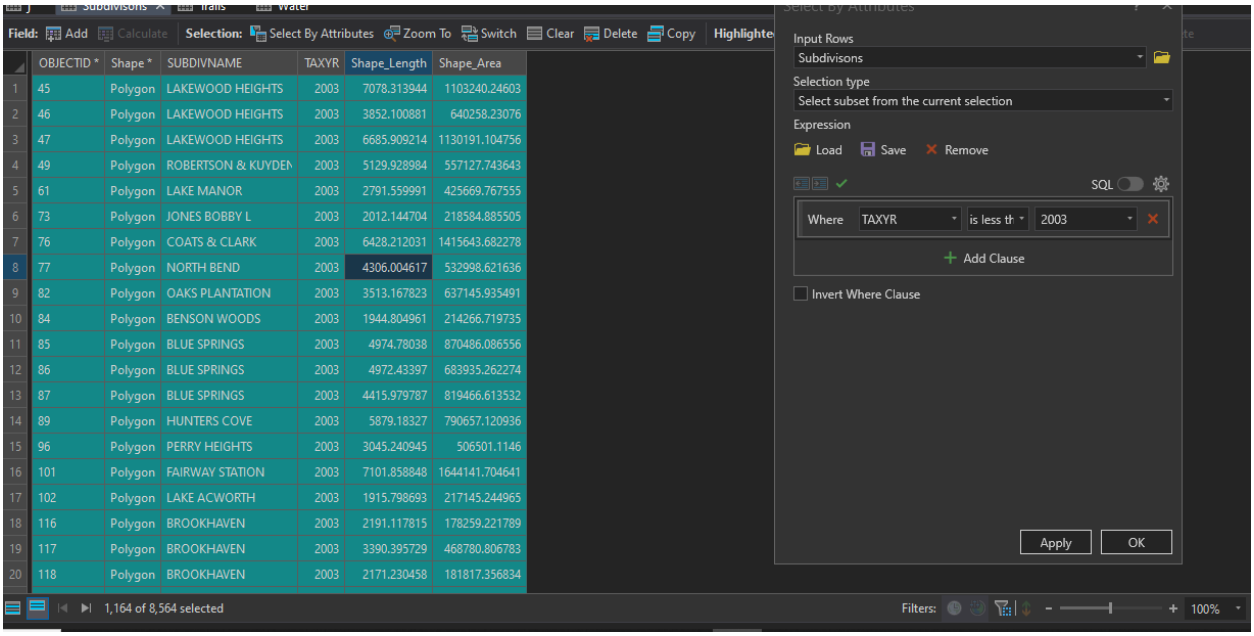


(figure 7)

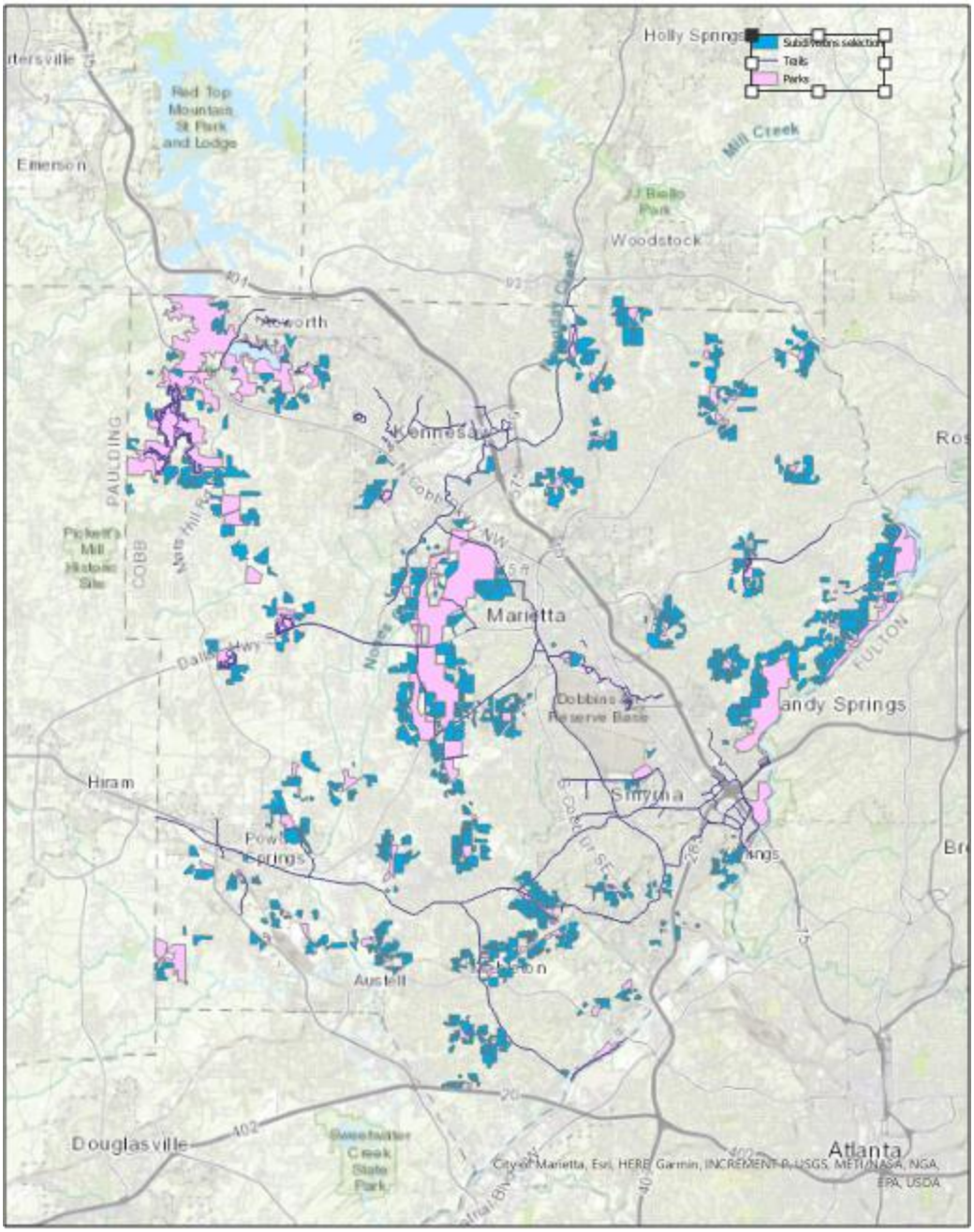
Question 3: Which subdivisions within one-quarter of a mile from a park were built in or before 2003?

To do this I reuse my Select By Location results and combine it with Select By Attributes Tool. Select By Attributes records based on the data contained in the attribute table. Similar to Select By Location, the input feature is the feature class which you would like to query. Once again in my case, this is Subdivisions. The selection type box answers the question of what to do if a selection already exists and expression is the SQL statement used to select a subset of your data. Hence, combing Select By Location with Select By Attributes acts as an and statement.

Figure 8 shows my results which reflect that 1,164 of the subdivisions built before 2003 are within one-quarter of a mile from a park and figure 9 shows these subdivisions in map form which I created using the same steps I used to create figure 7.



(figure 8)



(figure 9)

Conclusion

In conclusion, the database I built can be used to query relationships between subdivisions and public amenities. This could aid in the general mapping of some of the amenities that are hidden behind subdivisions and give a broad picture of the amenities and

subdivisions. This would provide transparency and help residents from feeling their taxes are being used in private goods.