Urban MODs GIS Workshop – 2019

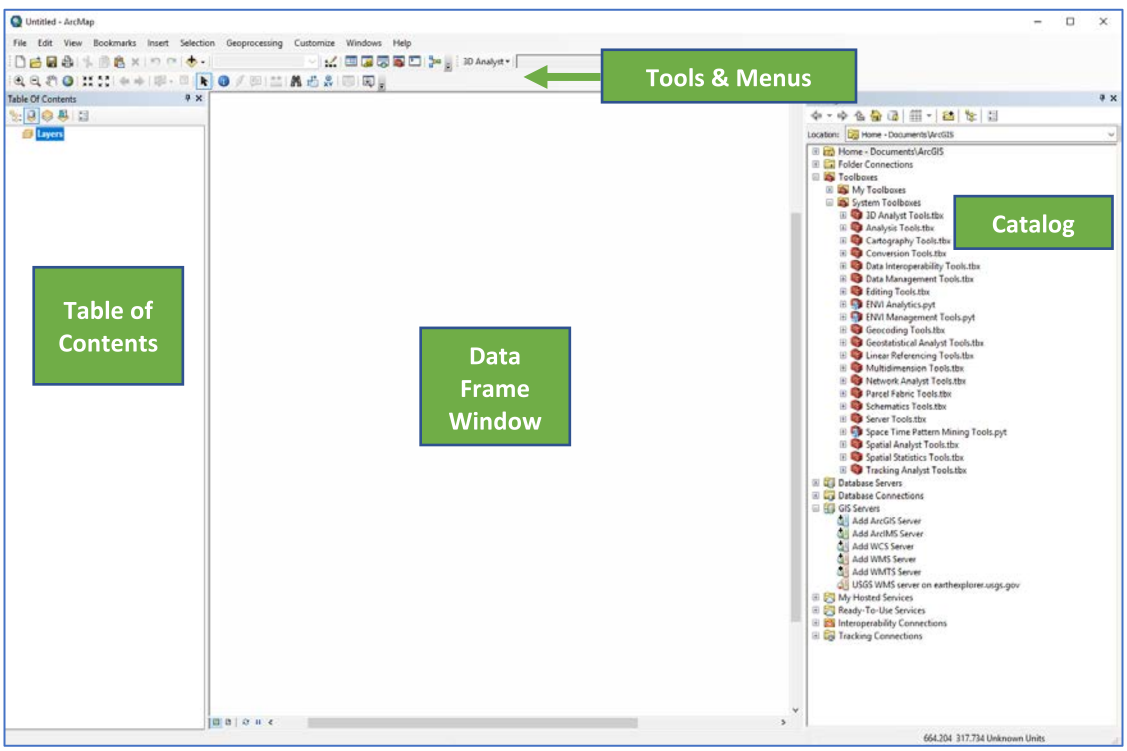
MODULE 1 – Temperature & Trees

1. Accessing and downloading data

- Go to https://tinyurl.com/2019modsGISweek1 and click 'Download' and then 'Save File.'
- Go to your Downloads folder. Right-click the "Urban MODS 2019 GIS Workshop" zip file and then click on "Extract All." Next, click "Extract."
- Open the file "MODs_GIS_2019.pdf". This pdf document provides all the instructions you'll need for today's workshop.

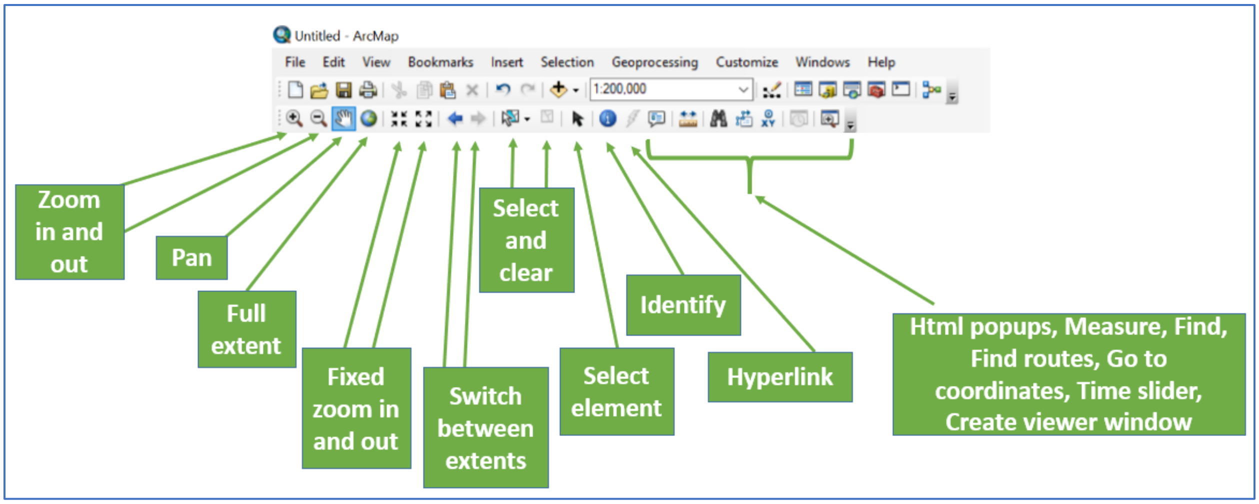
2.1. Getting started with ArcMap

- Start ArcMap
- Close the **ArcMap – Getting Started** window



Layers of data that are added to ArcMap will be listed in the **Table of Contents**. The spatial data will be displayed in the **Data Frame Window**. Tools can be accessed and searched for in the top menu bar or the **Toolbox Catalog**.

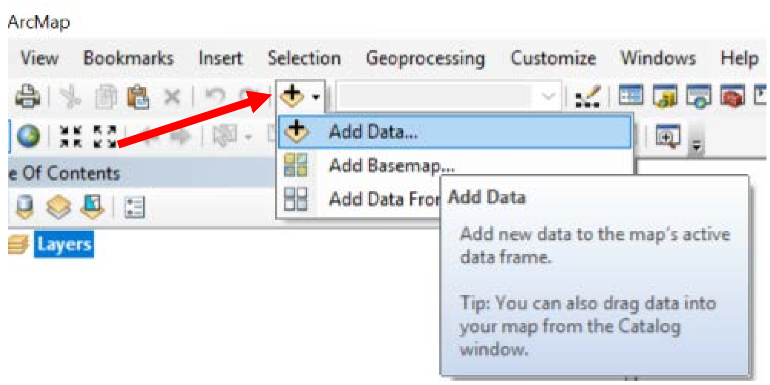
- **Important:** Activate the Spatial Analyst Extension by going to **Customize** (on the top menu bar), **Extensions**, and then make sure the **box next to Spatial Analyst is checked**. We'll need the Spatial Analyst tools for this workshop.



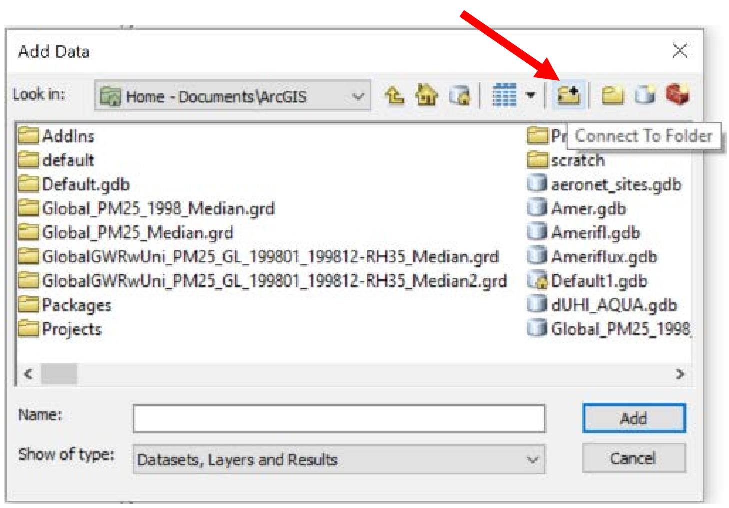
A number of tools are shown in the above figure to help you navigate the scene, zoom in and out, and explore and inspect data.

2.2. Adding Street Tree Data

- Click on **Add Data**



- Choose **Connect to Folder**
- Choose the folder where you downloaded and unzipped all the material



- Select the **Trees_new.shp** file
  You should see a number of points appear in your **Data Frame Window**. Each point represents the location of a tree, identified by URI staff and students. You should also note the layer appears in the **Table of Contents**. You may turn on and off layers by checking or unchecking the box next to that layer's name.
- To help us understand where these trees are located in New Haven, let's add a basemap to the workspace. Click the small triangle next to the **Add Data** icon and select **Add Basemap**. Feel free to take a look at all the provided basemaps, but for now, let's select the **Topographic Basemap**. Note, this may take a couple seconds to load.
- Spend a few minutes using the various exploration tools to explore this dataset. Zoom in and out in various places in the city and use the Identify tool to click on different trees (points) to see the associated attributes. You should see information such as the tree's address, common name, DBH, and others.

2.3 Visualizing Data

- Currently, the visualization of the street trees just shows us where each tree is located. It doesn't give us any information about DBH or Species. Let's change our visualization so that we can see the DBH of the trees.
- Right-click on **Trees_new** layer in the **Table of Contents**. Select **Properties**, and then go to the **Symbology** tab.
- Go to Quantities → Graduated Colors. Under the Value Field, select the variable DBH. *You may get a warming that the Maximum sample size was reached. To increase the sample size, close this pop-up and go through the following step:*
  - Still in the Graduated Colors dialog page, click the Classify button. Once the new window pops up, click the Sampling button and change the Maximum Sample Size to 50,000. Click OK twice to return to the Graduated Colors page. By increasing our maximum sample size, we make sure ArcMap plots all our data on the map.
- Select whichever Color Ramp and number of Classes you like best and hit OK. Now we have more information on our map. Not only do we know where the trees are located, but we are now able to visualize the general size of each tree as well. Take a few minutes to explore different methods to visualize the data. You may want to try the Graduated Symbols for DBH or go to the Categories tab on the left to try to visualize trees by their common name or species.

3.1 Mobile Meteorological Measurements

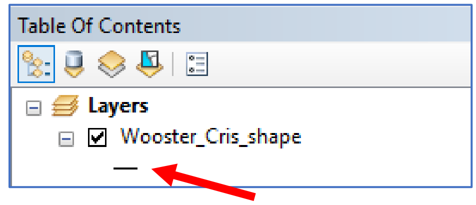
We're now going to bring in the mobile meteorological measurements that you collected on Monday on the bike trips to the parks and compare the temperature to the number of street trees around the route. A reference temperature from a nearby station was subtracted from the bike sensors, so we're looking at temperature anomalies, relative to the reference sensor.

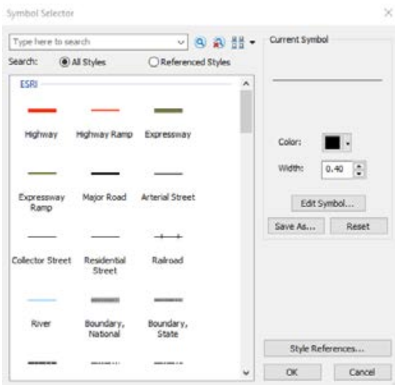
We've already converted the bike routes to shapefiles for use in this workshop. Ask an instructor or TA if you're interested in how this was done.

- Click on **Add Data** and select the merged bike routes file (BikeRoute_week1.shp).

Now you are viewing your bike route shapefile on the map as well as the street trees. By default, the bike route may be displayed as a thin line that is difficult to see on the map.

- For better visualization, left-click directly on the line below the layer's name in the **Table of Contents**.





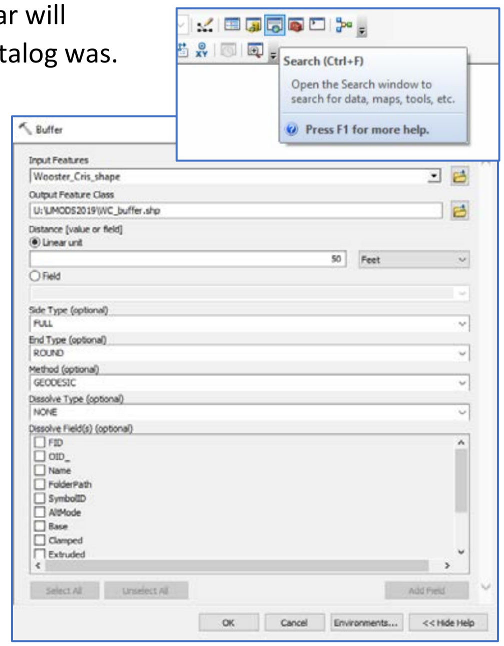
You can now edit the style of the line through preselected options or by changing the color and width yourself.

- Edit the style of the line so that you can better see it on the map and click **OK** when you're done.

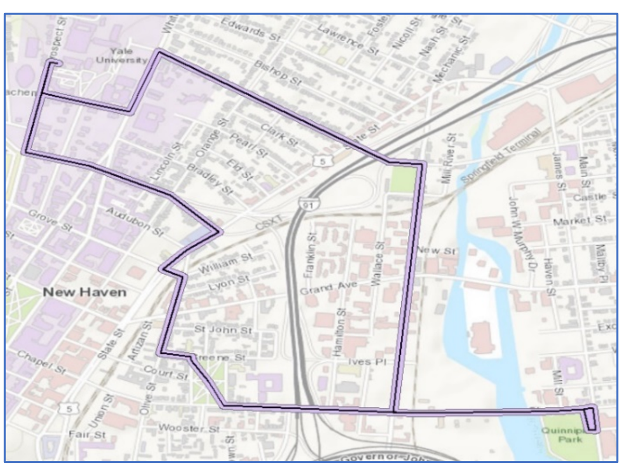
Note, you can change the drawing order by dragging layers in the Table of Contents above or below other layers.

3.2. Create a buffer around the route shape file

- Open the **Search** tool in the upper toolbar. A search bar will appear on the right side of the interface, where the Catalog was. Search for "Buffer"
- Select the **Buffer (Analysis) Tool**
- Use your route shapefile (BikeRoute_week1.shp) as the **Input Feature**.
- Navigate to your data folder and provide a name for the new shapefile in the **Output Feature Class** field.
- Change the Distance to **50 Feet**
- Change the Method to **GEODESIC**
- Click **OK** to run the tool



Your buffer should look something like the image below (although you may have a different route). The black line here is the route, and the purple area is the 50-foot buffer around this feature.



Because we want to directly compare the effect of street trees on bike temperature, we will use this buffer to select only the trees that are right next to the road along the bike route. We call this buffered region our "area of interest".

3.3. Mobile meteorological data

- Use the **Add Data** button to load the bike temperature shapefile for the routes (Bike_temp.shp). You can see that this data is point data, with each point corresponding to a measurement (1 measurement per second). You can go into this layer's Symbology properties as you did earlier with DBH to assign colors to the points based on their temperature. However, for better visualization, we will convert this vector file to a raster.
- Use the **Search** function to search for "IDW (Spatial Analyst)". This tool interpolates a raster from point data using an inverse distance weighted (IDW) technique.
- Select your bike temperature file for the **Input point features**
- For the **Z value field**, select air temperature (delta_Ta)
- Save the **Output raster** to your data folder
- Click **OK** to run the tool

Note, we will have large biases away from the sampling points, but we will use our buffer to clip out areas of the raster that are outside our area of interest.

3.4. Clipping data to area of interest

Our final step in comparing the bike temperature and street tree data is to clip both of those layers to our area of interest. There are separate clip tools for raster and vector data.

To clip raster data (bike temperature):

- **Search** for "Clip"
- Select the **Clip (Data Management)** tool
- Select your bike temperature raster as the **Input Raster**
- For the **Output Extent**, select your buffer
- Save the **Output Raster Dataset** to your data folder
- Check the **Use Input Features for Clipping Geometry (optional)** box
- Check the **Maintain Clipping Extent** box
- Click **OK** to run the tool

To clip vector data (street trees):

- Having already searched for "Clip", select the **Clip (Analysis)** tool
- Select the street tree data (Trees_new) as your **Input Features**
- For **Clip Features**, select your buffer
- Save the **Output Feature Class** to your data folder
- Click **OK** to run the tool

Now you should have the bike temperature raster and the street tree data clipped to your bike route. Turn off the other layers (uncheck the boxes in the **Table of Contents**), and toggle between the bike temperature layer and the street trees layer. You may want to update the symbology of the temperature data to make the variations more prominent. Do you see any spatial relationship between these two layers? Discuss with those around you.

We'll pause here and discuss these patterns as a group before moving on to the next section.

4. Scale up the analysis using city-wide data

Now that we've examined the data we collected on the ground, we're going to scale up the analysis to the city-scale. Since our street tree dataset does not include all vegetation within New Haven (e.g. trees in people's yards) and our mobile met measurements are limited, we will use satellite data for this task. We will also relate these relationships to socio-economic data of New Haven residents with census data.

4.1. Load data

- Before loading the new data, let's remove all existing layers from the previous section, except for the basemap, by right-clicking on them in the **Table of Contents** and selecting **Remove**.

Census data

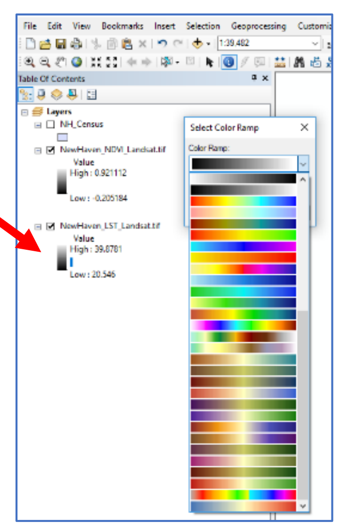
- Use the **Add Data** button to add the NH_census_data.shp layer from your data folder. This is a vector file that contains the outlines of each census tract in New Haven. Right-click on this layer in your **Table of Contents** and select **Open Attribute Table**. You can see the attributes associated with each polygon in this file, including the census tract name, median income (Med_Inc), crime rate (Crime_Rate), and poverty rate (Pov_Rate).
- Let's visualize the spatial variation of poverty rate in New Haven. Right-click on the New Haven census layer, select **Properties**, and go to the **Symbology** tab. Under Quantities (on the left side), select **Graduated colors**. In the Value field, select Pov_Rate. Choose a color ramp of your liking and hit **OK**. Now we can see which census tracts of New Haven have higher or lower poverty rates. Note that the range of values associated with the different colors now appears under this layer's name in the **Table of Contents**.

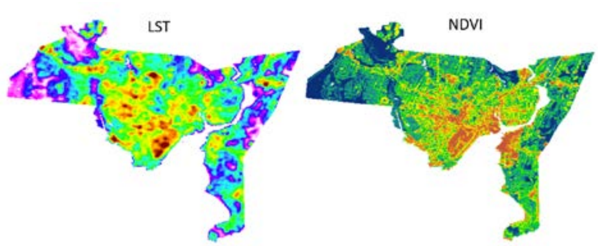
Satellite data

- Use the **Add Data** button again to load NewHaven_LST_Landsat.tif and NewHaven_NDVI_Landsat.tif. As you may have guessed, the first raster file is the Land Surface Temperature (LST) of New Haven derived from the Landsat satellite. The second file is the Normalized Difference Vegetation Index (NDVI) derived from Landsat.

Note: the spatial resolution (pixel size) of each of these two raster layers is 30m. You may notice that there is more detail in the NDVI layer than the LST layer. The actual resolution of the thermal bands of Landsat is 100m; however, the USGS resamples these bands to 30m for easier analysis with the other 30m bands.

- By default, both layers will be displayed as a black-white gradient. You can improve the visualization using the Color Ramp options. Click on the color bar in the associated with the files in the **Table of Contents** to change the color ramp of each layer.





4.2. Zonal statistics

Because we'll be comparing the LST and NDVI patterns with census data, we first need to calculate the mean value of these rasters for each census tract. We do this using the **Zonal Statistics** tool.

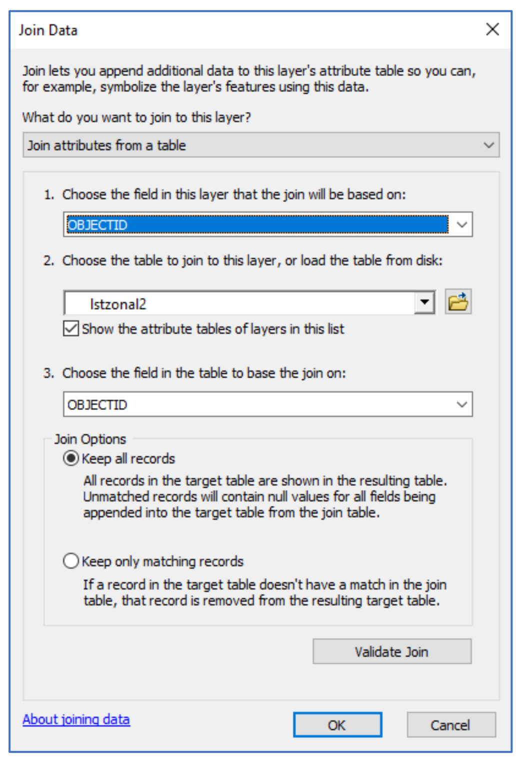
- **Search** for "Zonal Statistics as Table" tool
- Select the NH_census_data.shp file as the **Input raster for feature zone data**
- For **Zone Field**, select 'OBJECTID'
- Select the LST raster (NewHaven_LST_Landsat.tif) for the **Input value raster**
- Select your data folder for saving the **Output table**
- For **Statistics type**, select ***Mean***
- Click OK to run the tool
- Repeat this process for the NDVI raster (NewHaven_NDVI_Landsat.tif)

We've just created two new tables with the mean LST and NDVI for each of New Haven's census tracts. For visualization, we'll need to join these tables with the New Haven census shapefile. But first, to avoid confusion, we'll do a quick renaming of the "MEAN" fields in each of the tables, so we don't get them confused with each other.

- Right-click and open your LST table
- Now right-click on the 'MEAN' field and go to Properties
- Change the Alias to 'LST_MEAN'
- Repeat for the NDVI table to rename 'MEAN' to 'NDVI_MEAN'

4.3. Joining data to the New Haven census file

For visualization and comparison to the socioeconomic data from the census, we now need to join the LST and NDVI tables to the New Haven census shapefile.



<u>To join the LST & NDVI data:</u>

- Right-click on the NH_census_data.shp layer, go to **Join and Relates** and select **Join**
- In the drop-down box **What do you want to join to this layer?** Select **Join attributes from a table**
- 1. Select the 'OBJECTID' field
- 2. Select the LST zonal statistics table
- 3. Select the 'OBJECTID' field
- Click OK to perform the Join
- Repeat this process for the NDVI table

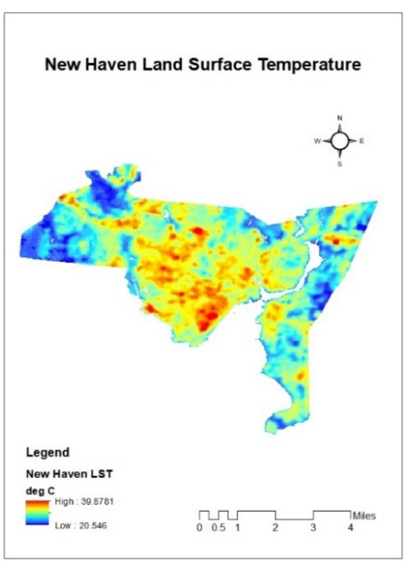
Now we're done with joining all our data to the census file. If you open the attribute table for the New Haven census data layer, you will find additional columns for 'LST_MEAN' and 'NDVI_MEAN'. If you're interested, take a minute to visualize LST or NDVI by census tract.

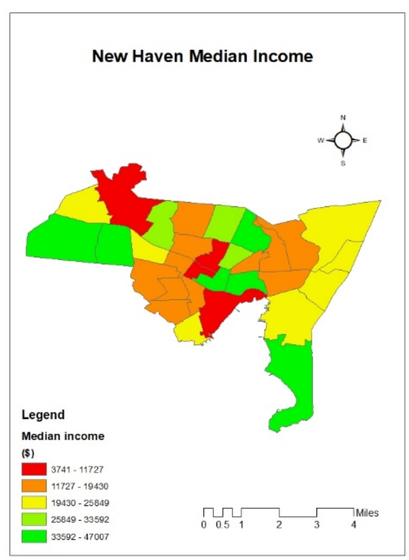
5. Creating a publishable map

In this section, we'll go through the process of creating a publishable map. The finished map should include a scale bar, legend, title, among others, so that other people can easily interpret your results.

- Let's start with the New Haven LST raster file. Uncheck all the other layers in the **Table of Contents** so that the LST file is the only layer displaced. If you haven't already done so, change the color ramp to a range of colors appropriate for visualizing temperature.
- The first thing we need to do is to change from **Data View** to **Layout View**. On the top menu, go to **View** and select **Layout View**. Now we are viewing our LST layer on a piece of paper. You can use the **Pan**  and **Fixed Zoom In/Out**  buttons to position your layer on the page.
- Now let's add some information to our map. In the top menu bar, go to **Insert** to select these options:
    - *Title* – give your map a main title
    - *Legend* – show the color bar for LST
        - Note: In your Legend, by default, the name of your layer will be shown. To change that, left-click on the name of the layer in the **Table of Contents** and change 'NewHaven_LST_Landsat.tif' to 'New Haven LST'. This will update your legend. You can do the same with 'Value'. Change this to 'deg C'.
    - *North Arrow* – Select whichever one you like
    - *Scale Bar* – Select your favorite, go to Properties to change units if you want.
    - *Text* – you may also want to add additional text to add your name or provide more information about the map (optional here).

You can repeat this process with any of the data we've loaded thus far. Try to create at least one additional map using the census data. Examples of the LST map and a census map shown below. You can export your map by going to **File → Export Map**.





Let's pause here to discuss our maps as a group.

6. Examine relationships between satellite and census data

We've already exported the census-tract averaged data from ArcMap to a csv file. To examine the relationship between vegetation, temperature, and the socioeconomic data, we'll move to R.

- Close ArcMap (do not need to save your map before exiting the program)
- Open RStudio
- In the top menu bar, go to File→Open, and navigate to your data folder and open the file **GIS_MODS2019.R**

Comments, which describe what each section of code does, are preceded by a #.

```
# This is a comment
```

- The first thing we need to do is load the packages we need for all functions in the script to work. Put your cursor anywhere on line 10 and hit **CTRL-ENTER** to run this line of code. In the Console, you should see that R is installing this package.

```
9  ##Install the packages we need for this script to run
10 install.packages("openxlsx") # To open Excel files
11 install.packages("ggplot2")  # For better plots & visualizations
12 install.packages("hash")     # To create hash table
```

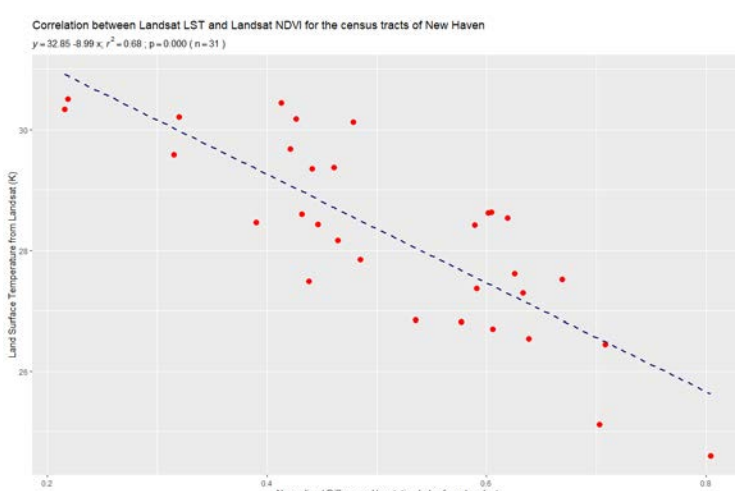
- Your cursor should have automatically moved to line 11. Hit **CTRL-ENTER** again to run this line, and again for line 12.
- After these packages are installed, let's comment them out (add a # before each line) so that R doesn't try to repeatedly install the packages each time we run the script.

```
9  ##Install the packages we need for this script to run
10 #install.packages("openxlsx") # To open Excel files
11 #install.packages("ggplot2")  # For better plots & visualizations
12 #install.packages("hash")     # To create hash table
```

- Set the working directory on line 23 to your unzipped folder location (use forward slashes)

```
21 ##Set working directory (change to where you saved the downloaded material)
22 #-----------------------------------------------------------------------------
23 setwd("N:/UMODS_2019/Final cleaned")
24 #-----------------------------------------------------------------------------
```

- Now the script is ready to be run. Use **CTRL-A** to highlight the entire script and then hit **CTRL-ENTER** to run. You can also click the **Run** button at the upper right of the script.



The variables that are initially set up to be plotted are NDVI (x-axis) and LST (y-axis) (these are set on lines 33-34). After running the script, you should see a figure that looks like the figure to the right.

- Why do we see this relationship? Does this make sense?

- Now let’s look at the relationship between LST and poverty rate. We can change the variables in lines 33-34.

```
31 ##Change the variables here and the rest of the script should respond automatically
32 #--------------------------------------------------------------------------------
33 y_nam="Pov_Rate"    # y-axis
34 x_nam="LST_Landsat"   # x-axis
35 #--------------------------------------------------------------------------------
```

- Highlight the entire script (**CTRL-A**) and run again (**CTRL-ENTER**)

Examine your new figure. What is the relationship between LST and poverty rate?

- Continue to edit lines 33-34 to look at the relationships between
    - NDVI (‘NDVI_Landsat’) (x-axis) & median income (‘Med_Inc’) (y-axis)
    - Poverty rate (‘Pov_Rate’) (x-axis) & crime rate (‘Crime_Rate’) (y-axis)
- If time allows, feel free to explore other variables and relationships

We’ll pause here to discuss as a group before moving on to the stormwater module.

MODULE 2 – Urban Stormwater Flow

We're switching gears a bit and moving onto urban stormwater flow. Our goals for this module are to delineate the Beaver Ponds Park watershed (above-ground drainage) and sewershed (below-ground drainage through New Haven's sewer system).

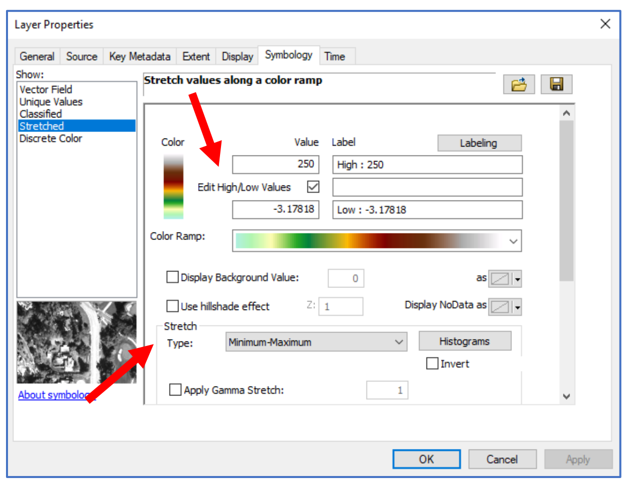
7. Watershed delineation

Let's start with the watershed. To determine the urban area that naturally drains into the Beaver Ponds Park, we'll apply GIS tools to elevation data to map the flow of water over the landscape. Note, we've prepared some intermediate files to simplify the process for this workshop. There are some additional steps if you were starting this process from scratch. An instructor or TA can provide more information if you're interested.

- Reopen ArcMap
- Go to **Add Data** and select **NewHaven_elevation.tif**. If ArcMap asks to create pyramids for faster display, select **Yes**.

This raster file is the elevation of the surface of New Haven at 1m resolution. The "surface" includes buildings and other structures on the land. This type of elevation data is referred to as a "Digital Surface Model" (DSM), which is slightly different than a "Digital Elevation Model" (DEM), which gives the elevation of bare ground.

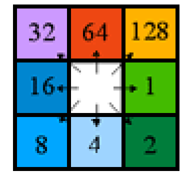
- By default, the elevation layer will be displayed in black and white. Click on the color bar in the **Table of Contents** and change the color ramp. It looks like the maximum and minimum elevation values are stretching out our color ramp. Let's edit the displayed values to better visualize elevation variations within the city.
    - Right click on the elevation layer, go to **Properties** and then the **Symbology** tab
    - Change **Stretch Type** to *Minimum-Maximum*
    - Check the **Edit High/Low Values** box
    - Change the maximum value to 250
    - Click **Apply** and **OK**



Use the **Identify** tool to examine the elevations of different areas around the city. Note the low elevation of Beaver Ponds Park.

- First, we need to calculate the flow direction of water. **Search** for "Flow Direction" and select the **Flow Direction (Spatial Analyst)** tool.
- Use the elevation file as the **Input surface raster** and save the **Output flow direction raster** to your data folder. Give it a descriptive name (e.g. FlowDir.tif) before saving.

This will produce a new raster showing the direction of flow out of each cell. The value of the cell is a number representing a direction as shown by the figure to the right. For example, a cell value of 64 means that the flow out of it is to the north.



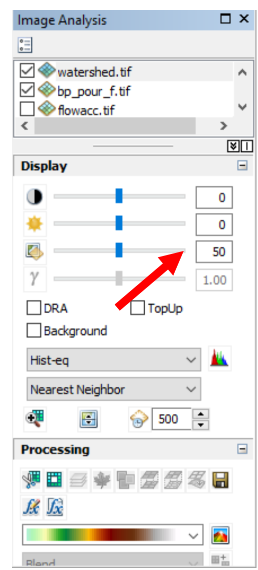
Direction coding

Our next step is to calculate the watershed area of Beaver Ponds using the flow direction raster we just created.

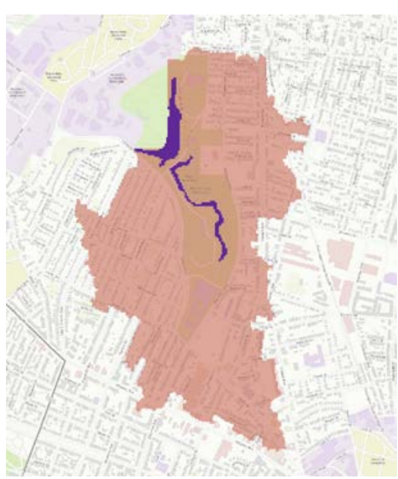
- Go to **Add Data** and select **BP_water.tif**. This is a raster file that has the extent of the Beaver Ponds Park water bodies, which we need for our delineation of the watershed. This raster is called a "pour point" or "pour region", representing the outlet through which water drains out of a region. It was created used the Snap Pour Point tool after determining where water was accumulating using the Flow Accumulation tool.
- **Search** for "Watershed" and select the **Watershed (Spatial Analyst)** tool.
- Select the flow direction raster file for the **Input flow direction raster** and the **BP_water.tif** file for the **Input raster or feature pour point data**.
- Save the **Output raster** as "BP_watershed.tif" to your data folder.

This final step produces a raster file (all zeros) of the land area that drains into the Beaver Ponds Park water bodies. Let's clean up the workspace for better visualization.

- Turn off all other layers except for the watershed layer and the BP_water.tif layer
- Change the color bar of the watershed layer (since the raster is all zeros, it will only display the lowest color)
- Add a basemap
- Adjust the transparency of the watershed raster:
  - Go to **Windows** on the top menu bar and select **Image Analysis**
  - Click on the watershed layer to highlight it
  - Change the transparency to 50
  - Close the **Image Analysis** box



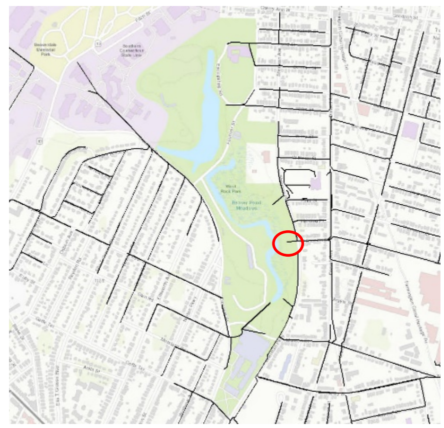
Now we have our watershed! At this point, you should have something that looks like the image below, that includes a basemap, the Beaver Ponds Park water bodies, and a 50% transparent layer of the Beaver Ponds Park watershed.



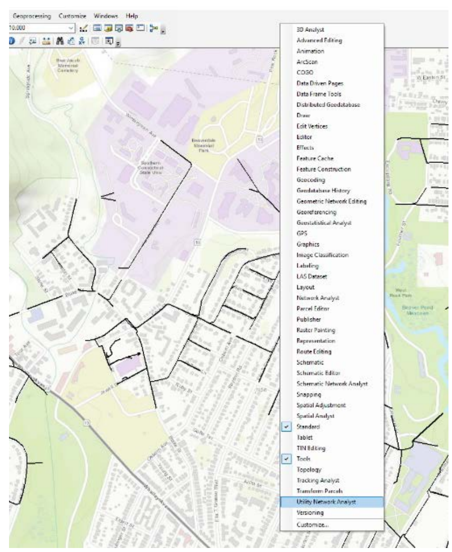
8. Sewershed delineation

Stormwater flows into Beaver Ponds Park not only by running over the land surface, but also through New Haven's stormwater sewer system. Our next (and last!) step in the workshop is to map the extent of the stormwater sewer network that flows into Beaver Ponds Park. ArcMap designs geospatial tools specifically for analysis of urban utility networks, which we will use here.

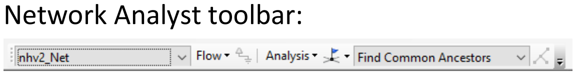
- Go to **Add Data**, navigate to your data folder and *double-click* on the file **UrbanMods.gdb**. Double-clicking will open the geodatabase revealing a number of file. Now *single-click* on **StormPipeDataset** and select Add to load our network of stormwater sewers. If you do accidently navigate into the **StormPipeDataset**, there will be three files listed (nhv2_Net, nhv2_Net_Junctions, and pipes2), and you need to load them all.
- After you load the storm pipe data set, two layers will appear in your **Table of Contents**: nhv2_Net_Junctions, and pipes2. You can uncheck the junctions layer as we will just be working with the pipes.
- Let's add a basemap at this point to see where we're at. It may also be helpful for visualization to change the style of the sewer lines. Remember, just click on the line itself in the Table of Contents to change its style.
- Navigate to Beaver Ponds Park. There is a storm pipe outlet on the southeastern edge of Beaver Ponds Park (this outlet is marked by the red circle in the image to the right. We're interested in determining which subsection of pipes in New Haven drain through to this outlet.



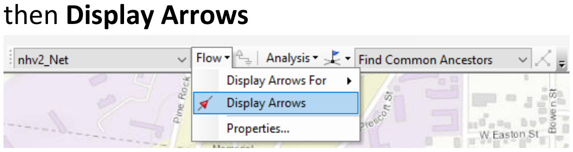
- Now we need to open the Utility Network Analyst tool to trace the upstream flow from this storm pipe outlet.



  - Right-click anywhere on the gray top menu bar area and select **Utility Network Analyst**.
  - If all goes to plan, a new tool bar should show up near the top of your screen. You may dock this alongside of the other tools by pulling the tool up to this region.
  - You should see the below fields in your Utility Network Analyst toolbar:

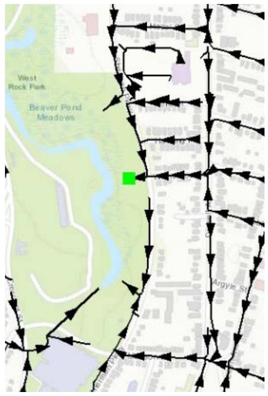


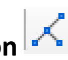
- First, let's look at the direction of flow within these pipes.
  - In the **Utility Network Analyst**, click on **Flow**, and then **Display Arrows**



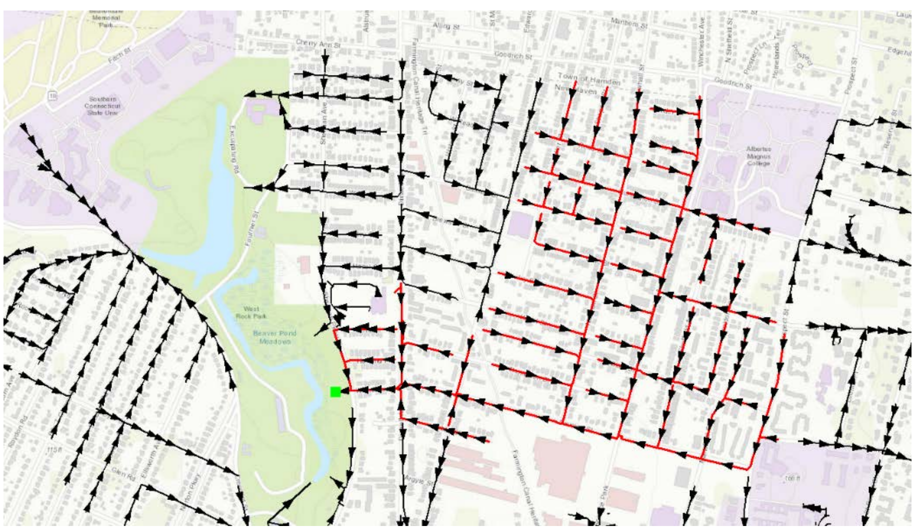
You should now see the direction of flow displayed as arrows on the pipe segments. Let's trace backwards from the Beaver Ponds Park outflow pipe to determine which storm pipes drain to this location.

- Click on the **flag icon**  in the Utility Network Analyst Tool.
- Click to place the flag at the Beaver Ponds Park outlet pipe:



- Now click the drop-down list that is currently set to **Find Common Ancestors**. Change that to **Trace Upstream**.
- After selecting **Trace Upstream**, hit the **solve button**  (to the right of **Trace Upstream**).

That's it! Now, the pipes that are highlighted in red are the pipes that are flowing towards that Beaver Ponds Park outlet.



Additional steps could include converting the Beaver Ponds Park watershed and sewershed into shapefiles (polygons) and overlaying them on the same map to visualize the entire area that drains into Beaver Ponds Park.

That's all for the GIS workshop!

Let's do a final wrap-up as a group.