LAB 3 – PROJECTING PINE MOUNTAIN

What you'll Learn: Coordinate Reference Systems (CRS) and area calculations in QGIS.

Data: Zipped file, downloaded from Canvas, containing 5 shapefiles (Public Areas, Land cover, and administrative units for Pine Mountain, Kentucky, plus US states and counties) and 1 TIFF file (NLCD 2011 landcover).

What You'll Submit: Via Canvas, a Word document containing answers to the questions listed throughout.

Background: The Earth's surface complexly curved. We introduce unavoidable distortion when we flatten this curved surface onto a map, typically changing areas, lengths, and the shapes of features. Different map projections introduce different types of distortion, and we choose the projection that is best suited for our project. **We cannot mix map projections in an analysis, so we often have to reproject our data layers.** Using the example of Pine Mountain in Kentucky, we'll walk through the process of data (re)projection in different coordinate systems.

Lab naming conventions: Tools that you click will be bolded, e.g., QGIS Menu > File > New to create a new QGIS project file. Text that you'll type will have quotes around it, such as "MyNewProject.qgs" and names of existing datasets and directories will be *italicized*, e.g., *DataToUse.zip*. Key terms will be underlined. **Important tips and key instructions will be in bold red font.**

Note: Questions are posted throughout this lab as opposed to all located at the end.

STEP 1: GETTING STARTED

Begin by creating your workspace. As always the file structure should resemble:

- GEO309 folder
    - Lab 3 folder
        - Data downloaded folder

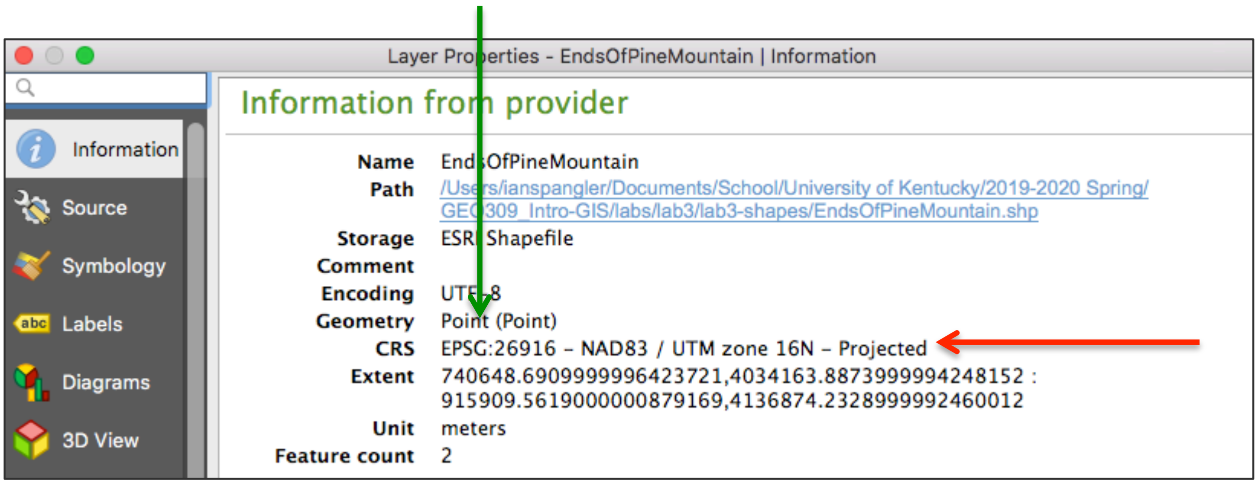
As we get further into the semester, take liberties with this structure and make it your own, if you wish (for example, you may want to create separate folders for downloaded raster data and downloaded vector data).

Add the 6 files. Ask yourself: how should these be layered in QGIS? Do you want the point or polygon layers on top? And which polygons do you want to be ordered first? Should some be given a degree of transparency?

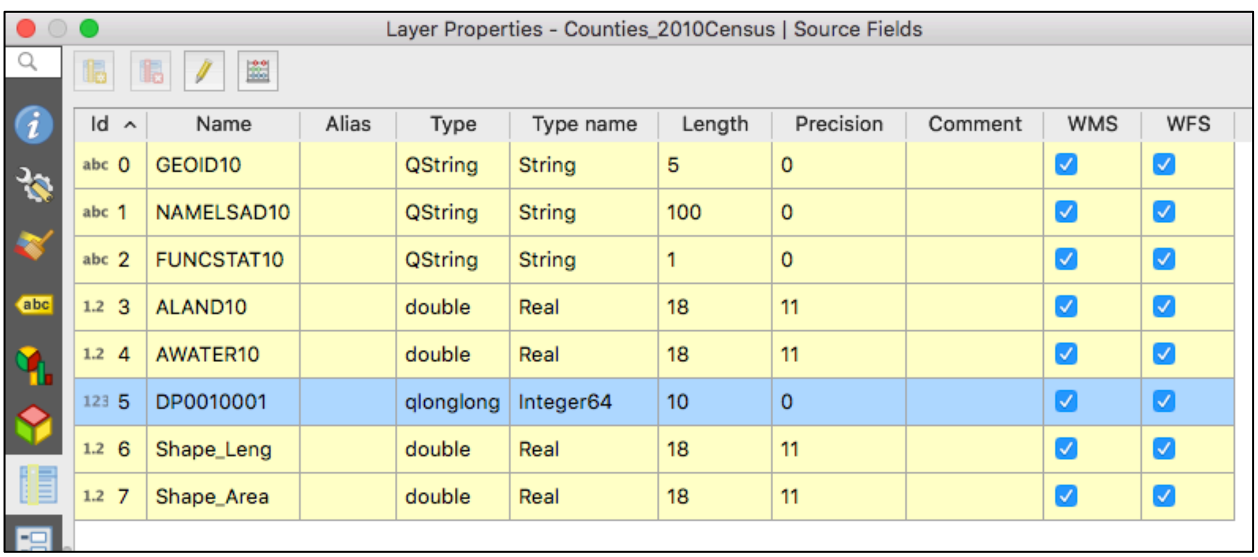
Inspect the layers by **right+clicking** one of them. Let's go with the *EndsOfPineMountain* layer. Select **Properties > Information** and identify the following:

1. Data model type?
2. Coordinate system?
3. Attributes?

The green arrow locates our data model – it is a vector point file – and the red arrow locates our coordinate system (NAD83), as well as how it's projected (UTM zone 16N).

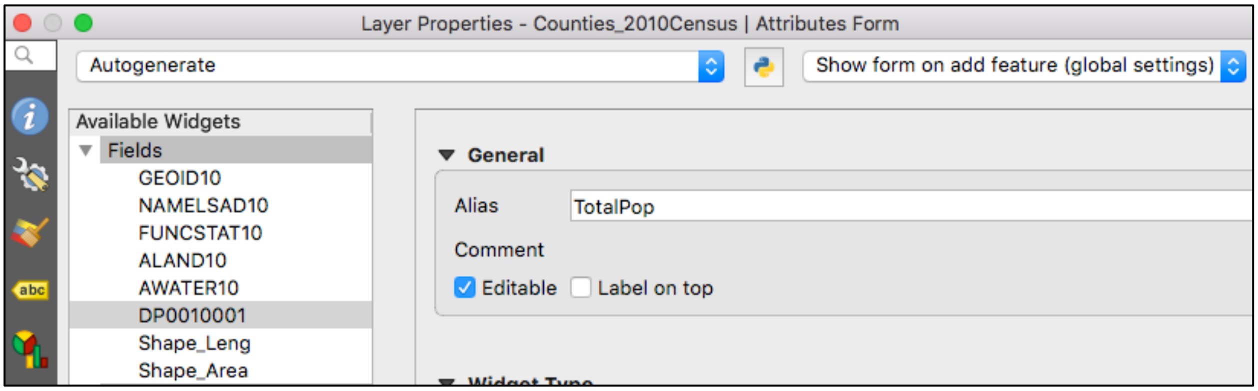


The *EndsOfPineMountain* layer only has 2 records and 1 field. Let's select another one: *Counties_2010Census* has much more for us to examine. We can view its attributes by selecting **Properties > Source Fields** (these are the same fields we'd see by opening up the attribute table):



Note that the different fields have different "Type" values. <u>Type</u> refers to what kind of characters an attribute field can record: for instance, a "Real" type can only record numbers and not letters. A "String" field can encode anything, but it will only be encoded as a text character and not a numerical character, which means that you can't perform mathematical analysis on it. **When you create fields, you always must select a "Type," so be careful to select the correct one.**

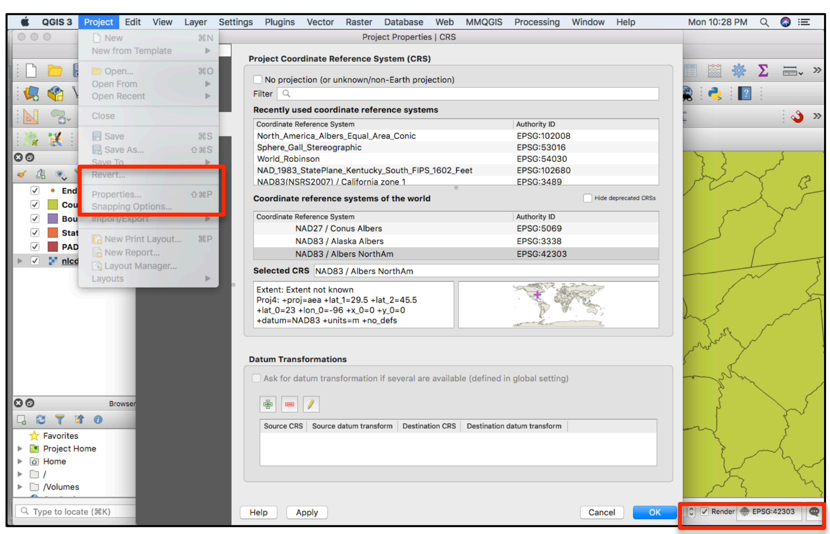
There's a field called *DP0010001* that doesn't make a lot of sense. This is actually a field that refers to total population in each county. Navigate to the **Attributes Form** tab and you can designate the field an **Alias**. Name it "TotalPop":



These have both been vector data properties, but what do raster ones look like? Open the **Properties** for the layer *nlcd_2011_landcover* and note that there are new options (e.g., **Histogram**, **Pyramids**, etc.). For now, let's check the **Information** panel and examine the coordinate system and projection.

Yikes – it shares a coordinate system, but the projection is markedly different. In fact, if you hover your cursor over each layer, you'll see a popup that shows at least three different projection sources. So… why isn't our data showing up in completely wacky ways – why does it render properly???

QGIS cleverly uses On-the-fly projection (or OTFP) to automatically transform data with different sources (e.g., different projections) into the same projection. On-the-fly projection is automatically enabled in QGIS 3 and you can't turn it off. QGIS will read the projection of whichever data you load first, and use that as the OTFP; for instance, I loaded the *nlcd_2011_landcover* layer first, and the map is projecting my data in EPSG:42303 – the same projection as that layer. But you can change this in the **Project Properties** panel, accessible via the buttons pictured here:



We've established so far that our layers share different projection, but can be reprojected automatically in QGIS using its built-in OTFP function.

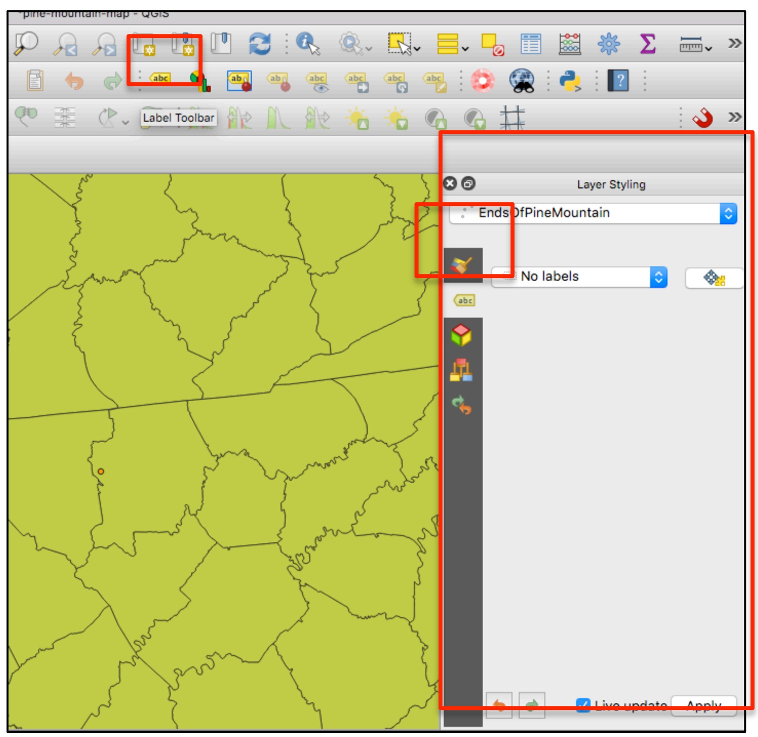
Unfortunately this isn't suitable for when we want to actually do analysis between layers – in those cases, everything needs to be in the same projection, or else we'll have some problematic outputs from the geoprocessing tools we run.

In the next few steps, you'll learn how to reproject data – not just OTF, but for good, too.

STEP 2: IS THIS COORDINATE SYSTEM THE BEST? (RE)PROJECTING DATA IN QGIS

Before moving on, let's style the layers to make the map more readable. You learned last week how to edit data from the **Layer Properties** panel, but there's also a button that opens up a separate tab for **Layer Styling** that can make the process easier.

As demonstrated in the figure to the right, click the **Layer Labeling Options** button, which should open the **Layer Styling** window. Navigate to the **Symbology** tab on the left-hand side. Voila! Now you can make changes to your data in real time.



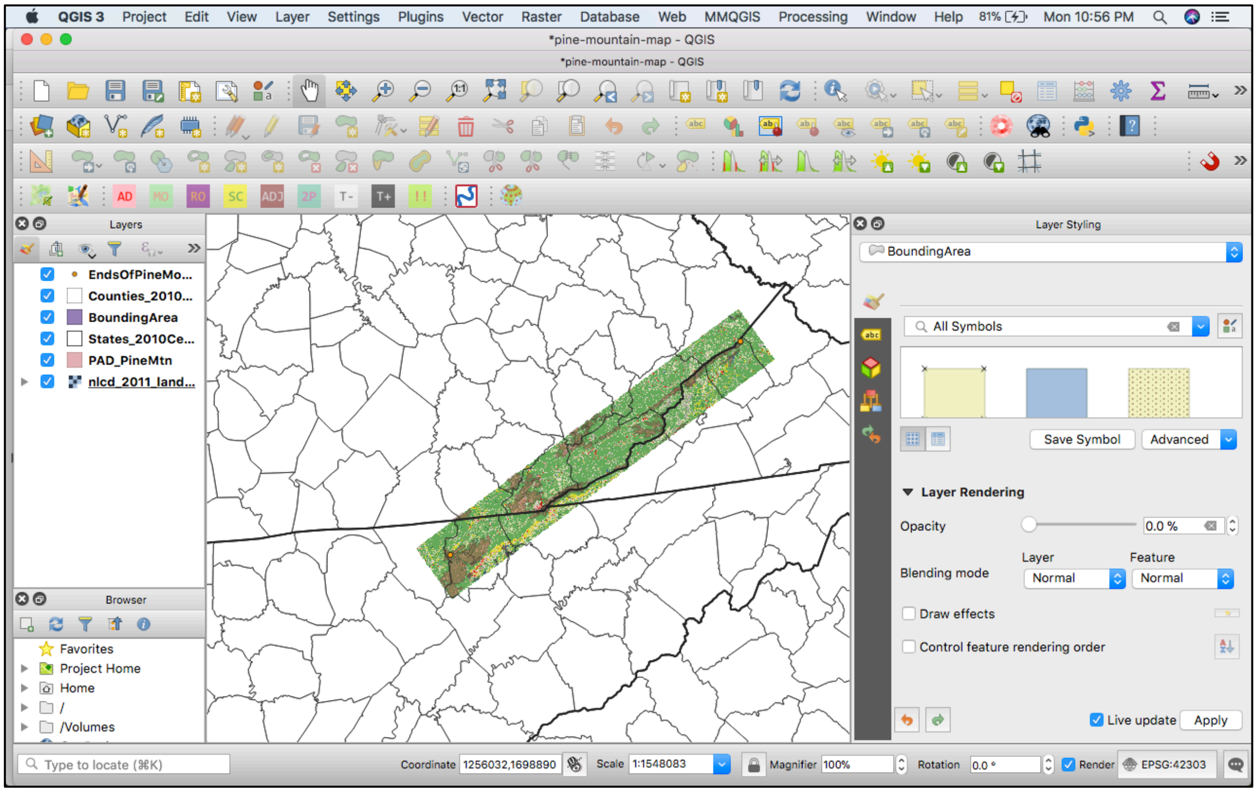
Make the following changes:

1. Make the counties and states hollow-filled (e.g., just an outline). You may want to make states a thicker line than counties.
2. Make the PAD layer slightly transparent.
3. Uncheck "Bounding Layer."

Note you will have to select "Simple Fill" to make most of these changes.

Also note that the interface for Layer Styling can be tricky to navigate if you're not already familiar with QGIS, so if you struggle, feel free to make these changes in the Layer Properties > Symbology tab.

Once you're done, you should see a map resembling the following:

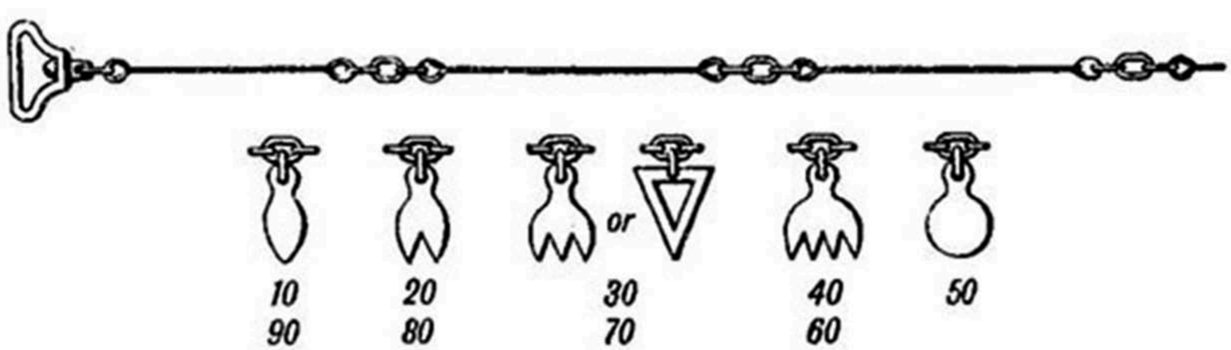


Check and uncheck layers once again to get a feel for which is which.

The symbology is already defined in the raster layer. Please… don't change it.

STEP 3: MEASURING SPACE, FROM GUNTER TO GIS

In 1620, a man named Edmund Gunter stunned the European world with his new invention, Gunter's chain: a device that introduced standards for how plots of land would be measured, both legally and commercially. It was literally just a 100-link chain.



Not only did Gunter's chain allow the market and the state to see land in a new way – that is, they now had a standard by which they could survey the earth – but it also created new ways of knowing, trading, and interacting with space and its inhabitants.

You might ask, *"Why is this man talking about Gunter and his chain?"* Well, rightfully so! You're here to learn GIS – not about Gunter. I promise I'll shut up, but not before insisting that Gunter's chain is an important technology in the development of surveying and measuring land, and just as Gunter's chain introduced new ways of knowing and interacting with contemporary real estate markets, so too do GIS technologies. These tools do not just *enter into* existing worlds – they actively help to *create them*.

Now, that said, let's once again break out the **Measure Tool** in QGIS.

Techniques of measuring the earth have improved greatly since the time of Gunter and QGIS' Measure Tool is no exception. As you learned last week, you can use it to capture measurements between different points in space. However, what those measurements end up being will vary depending on the system of surveying and the techniques used to flatten the globe on a flat plane – in other words, depending on the coordinate system and map projection.

Using the **Measure Tool**, select **Miles** as units. Click on the two *EndsOfPineMountain* points to record the distance between them.

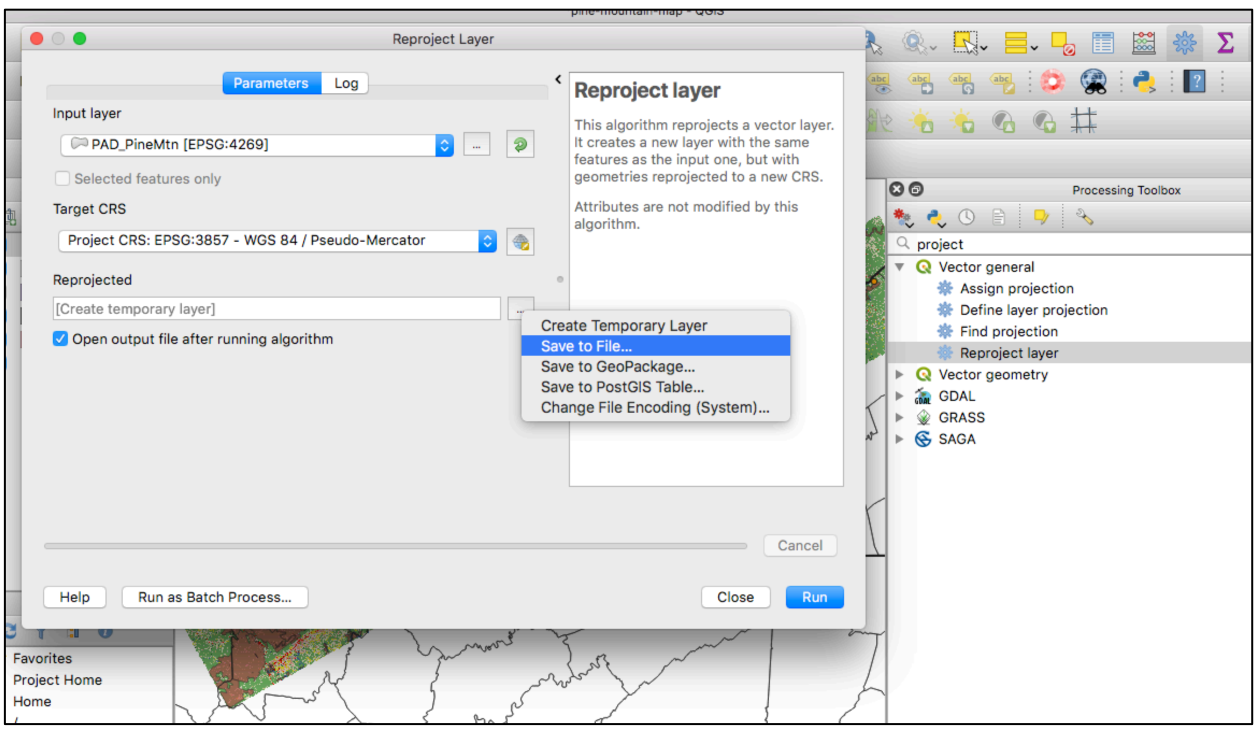
> **QUESTION 1:** What is the distance in miles that you measured?
>
> **QUESTION 2:** What coordinate system are you measuring in?
>
> **QUESTION 3:** What projection?

It's one thing to measure the distance between two points, but it's another thing entirely to measure the area of a polygon within different projections. The next section will walk you through that process.

STEP 4: REPROJECTION IN QGIS

In the **Processing Toolbox,** navigate to the **Reproject Layer** tool. You can do this by searching for it in the search bar or finding it under **Vector General**.

Select the *PAD_PineMtn* layer. Set the target CRS as "EPSG:3857 – WGS 84 / Pseudo-Mercator" (note you may have to **click** the globe in order to locate this). Then, **click** the ellipsis and select "Save as file…"



Once you've done this, a window that allows you to navigate through your file finder should appear. Create a **New Folder** titled "reprojected-layers" and title the output file

"PAD_wgs84". Click **Save** – **be sure to save it as a shapefile** – and when you're back to the **Reproject Layer** tool, click **Run**.

If the file didn't add to your **Layers Panel**, add it manually. Open its Properties and inspect the coordinate system: it should now be projected in EPSG:3857.

Now run the tool two more times on the *PAD_PineMtn* layer according to these parameters:

1. Export as "EPSG:3089" and save as "PAD_ky-single"
2. Export as "EPSG:102004" and save as "PAD_lambert"

Add each of these new exports to the map. You should inspect each output to confirm it successfully exported to the desired coordinate system and projection.

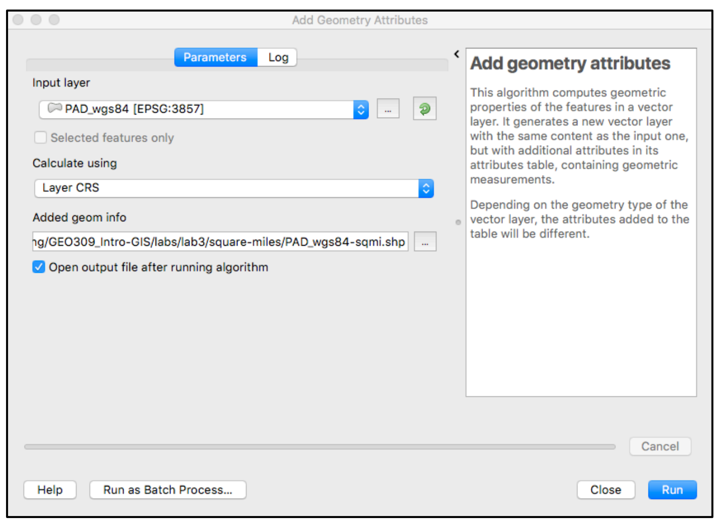
In the next step, we'll create new fields to measure each of these layers' geometries, comparing the differences that coordinate systems make.

STEP 5: MAKING GEOMETRY MEASUREMENTS AN ATTRIBUTE IN DIFFERENT COORDINATE SYSTEMS

Calculating geometry (area, perimeter, etc…) in the different coordinate systems will produce different and incorrect measurements. In this step, we'll make those measurements permanent, so we need to use the right coordinate system.

Using a similar workflow as the previous step, complete the following:

1. Using the **Processing Toolbox**, locate and open the **Add geometry attributes** tool
2. Select *PAD_wgs84* as the input
3. Calculate layer using "Layer CRS"
4. "Save as file…" to a new folder titled "geom-calcs"
5. Run the tool



Note that you may get an error message while the tool is running and that is okay.

If the output doesn't add, add it manually. Inspect the **Attribute Table** by scrolling all the way to the right. You should see two new fields titled "area" and "perimeter".

Run the tool two more times:

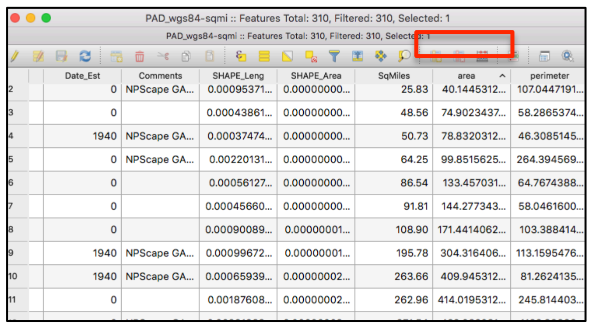
1. Once using *PAD_ky-single* as the input; save this file as “PAD_ky-single-geom”
2. Again using *PAD_lambert* as the input; save this file as “PAD_lambert-geom”

Add the two new files to QGIS if they haven’t automatically added. You’ve just created three shapefiles for Pine Mountain, each with uniquely calculated geometries and each in a different projection.

Using the attribute table to inspect your three newest files, answer the following questions:

> **QUESTION 4:** In the “area” field, what is the smallest value of *PAD_wgs84-geom*?
>
> **QUESTION 5:** In the “area” field, what is the smallest value *PAD_ky-single-geom*?
>
> **QUESTION 6:** In the “area” field, what is the smallest value of *PAD_lambert-geom*?

Note you can easily determine the above questions by opening the Attribute Table, scrolling to the “area” field, and clicking its header to sort by size:



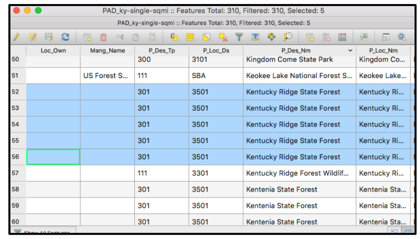
		Date_Est	Comments	SHAPE_Leng	SHAPE_Area	SqMiles	area	perimeter
2		0	NPScape GA...	0.00095371...	0.00000000...	25.83	40.1445312...	107.0447191...
3		0		0.00043861...	0.00000000...	48.56	74.9023437...	58.2865374...
4		1940	NPScape GA...	0.00037474...	0.00000000...	50.73	78.8320312...	46.3085145...
5		0	NPScape GA...	0.00220131...	0.00000000...	64.25	99.8515625...	264.394569...
6		0		0.00056127...	0.00000000...	86.54	133.457031...	64.7674388...
7		0		0.00045660...	0.00000000...	91.81	144.277343...	58.0461600...
8		0		0.00090089...	0.00000001...	108.90	171.4414062...	103.388414...
9		1940	NPScape GA...	0.00099672...	0.00000001...	195.78	304.316406...	113.1595476...
10		1940	NPScape GA...	0.00065939...	0.00000002...	263.66	409.945312...	81.2624135...
11		0		0.00187608...	0.00000002...	262.96	414.0195312...	245.814403...

So even though we ran the same tool on the three layers – each of which was derived from the same file – the calculated area for each record in Pine Mountain is different in each layer!

You’ll also note that the values for “area” in the *PAD_ky-single-geom* layer are quite lower (they’re small differences for small polygons but increasingly large as the polygons grow in size) than they are in corresponding polygons for the other two layers. If you check the **Information** tab for all three layers, you’ll see that the *PAD_ky-single-geom* layer uses “Feet” as its unit, while the other two use “Meters.” The **Add Geometry** tool calculated area in feet or meters accordingly.

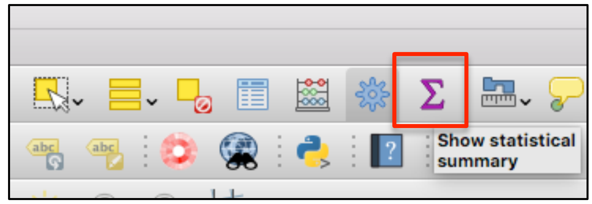
> **FOR EXTRA CREDIT:** Calculate the area in square feet for all records in both the *PAD_wgs84-geom* and *PAD_lambert-geom* layers and explain how you completed this process.

On to the final step: instead of First, in the Attribute table for one of the PAD layers, sort the records by clicking on the field “P_Des_Name”. Locate the five records for the “Kentucky Ridge State Forest” (note that you may have to expand the field width in order to see the full name). Once you’ve done so, **Shift+Click** to highlight all five. Then, on the bottom right-hand side of the attribute table, change **Show All Features** to **Show Selected Features**, as seen on the right:



You can also determine the total area of multiple records by using the **Show Statistical Summary** tool.

You can find the tool on the top right-hand side of the toolbar in your QGIS GUI. At this point, click the **Show Statistical Summary** tool.



You should see a toolbar pop up in the same place that the Layers and Browser tabs live. Be sure to select the layer you've selected the 5 features for. Set the field as "area," and check the "Selected features only" button. The result should be a table that summarizes your selected field and features by a variety of metrics.

Using the workflow elaborated above, answer the following questions:

> **QUESTION 7:** What is the sum value of *all records* for the Kentucky Ridge State Forest in *PAD_wgs84-geom?*
>
> **QUESTION 8:** What is the sum value of *all records* for the Kentucky Ridge State Forest *PAD_ky-single-geom*?
>
> **QUESTION 9:** What is the sum value of *all records* for the Kentucky Ridge State Forest in *PAD_lambert-geom*?
>
> **QUESTION 10:** In a few sentences, describe why values for the same records are different in each shapefile, but why they visually render the same in QGIS.
>
> **QUESTION 11**: Which do you think is closest to the real area of Kentucky Pine Ridge, and why?

STEP 6: REVIEW & SUBMIT

You should submit completed materials for Lab 2 via Canvas by Monday, 2/10 at 11:59pm. The completed materials only include the Word document containing answers to the questions posed throughout these lab instructions.

Be in touch if you have any questions!