LAB 8 – RASTER ANALYSIS WITH QGIS

What you'll Learn: In this lab you'll learn the basics of raster analysis and be introduced to tools that include **Reclassify, Slope, Aspect, Value, Raster calculator**, and more.

What You'll Produce: You'll produce and submit 1) a word doc with screenshots and answers to questions posed throughout the lab, and 2) a map that shows suitable terrain for property development in a study area in southern South Africa.

Data: You'll find necessary data posted throughout the instructions.

Background: Raster data is quite different from vector data. Vector data has discrete features constructed out of vertices, and perhaps connected with lines and/or areas. Raster data, however, is like any image. Although it may portray various properties of objects in the real world, these objects don't exist as separate objects; rather, they are represented using pixels of various different color values.

Note that large sections of this lab are derived from or copied directly from the documentation activities at qgis.org. Follow the link below to read more:

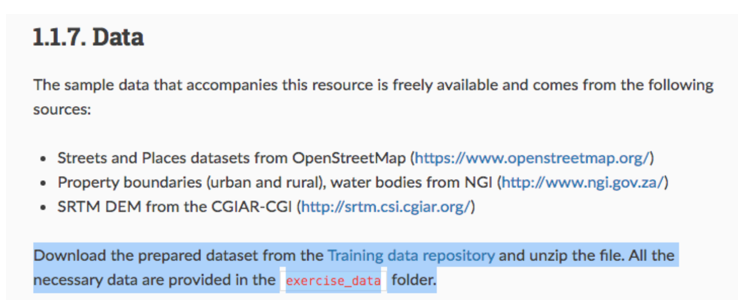
https://docs.qgis.org/3.4/en/docs/training_manual/rasters/data_manipulation.html

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.**

STEP 1: WORKING WITH RASTER DATA

Begin by setting up a workspace. Once you have that set up, we'll continue by simply downloading raster data and bringing it into QGIS. Navigate to the link below:

https://docs.qgis.org/3.4/en/docs/training_manual/foreword/foreword.html#



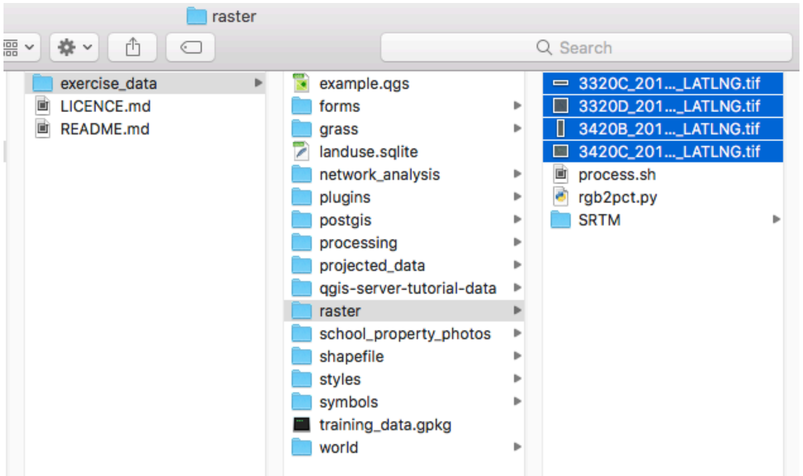
1.1.7. Data

The sample data that accompanies this resource is freely available and comes from the following sources:

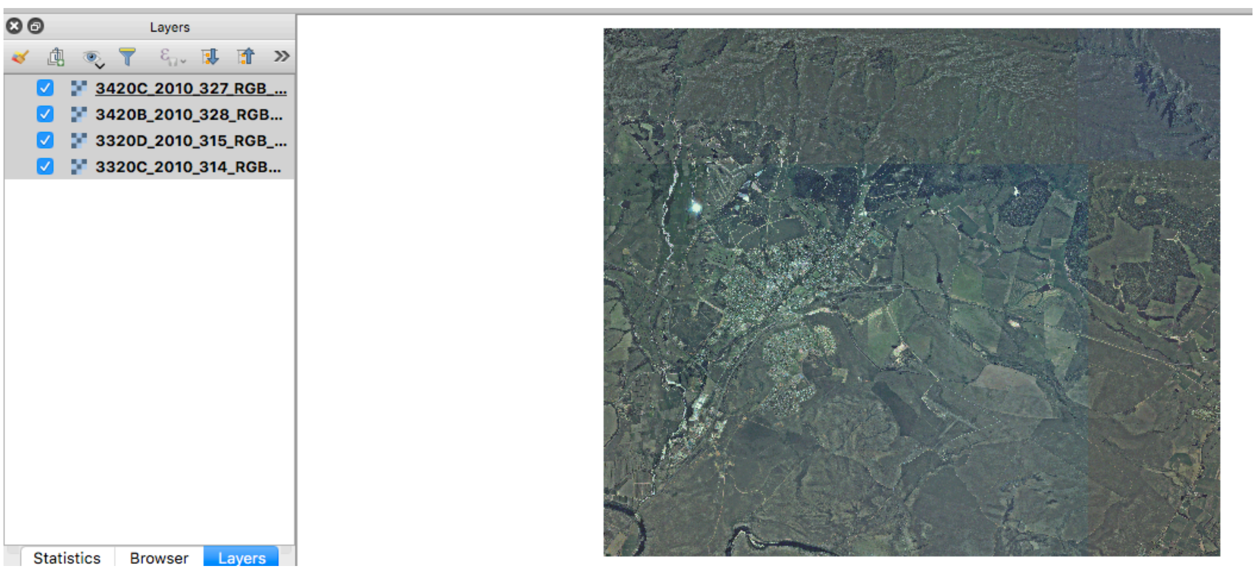
- Streets and Places datasets from OpenStreetMap (https://www.openstreetmap.org/)
- Property boundaries (urban and rural), water bodies from NGI (http://www.ngi.gov.za/)
- SRTM DEM from the CGIAR-CGI (http://srtm.csi.cgiar.org/)

Download the prepared dataset from the Training data repository and unzip the file. All the necessary data are provided in the `exercise_data` folder.

Download the sample data from section 1.1.7, as shown in the screenshot to the right. You should click the "Training data repository" and unzip the downloaded file. Move that data into an appropriate location in your workspace, and add the following files to a new QGIS project:



You should see the following map:

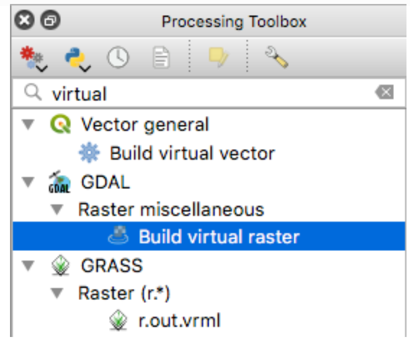


These four rasters provide partial coverage of the Marloth Nature Reserve in South Africa. Go ahead and click the four layers on and off to get a sense of their individual extents. Feel free to add a basemap (e.g., Google Satellite, OSM) to situate yourself as well.

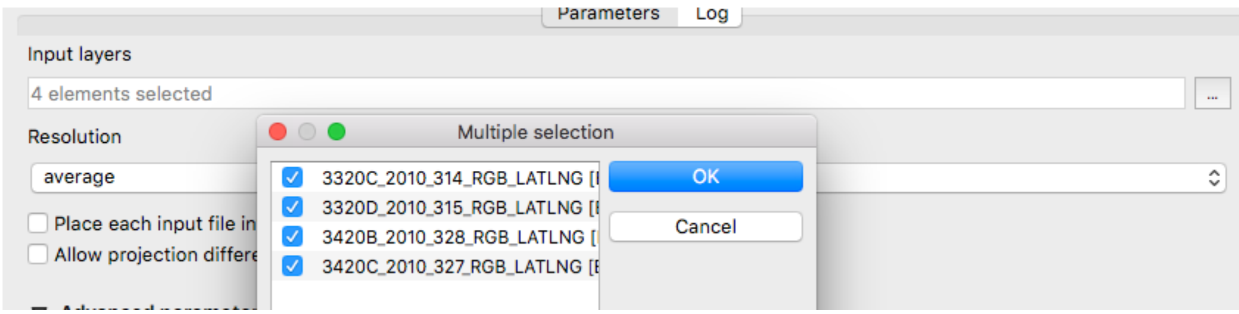
As you can see, your layer lies across all four photographs. What this means is that you're going to have to work with four rasters all the time. That's not ideal; it would be better to have one file for one (composite) image, right?

Luckily, QGIS allows you to do exactly this, and without needing to actually create a new raster file, which could take up a lot of space. Instead, you can create a **Virtual Raster**. This is also often called a **Catalog**, which explains its function. It's not really a new raster. Rather, it's a way to organize your existing rasters into one file for easy access: or, in other words, a **Catalog**.

First, navigate to the **Processing Toolbox** and open the **Build virtual raster** tool from **GDAL > Raster miscellaneous** (see screenshot to the right – you can also just search "virtual").



As shown below, select all 4 rasters as your input layers. Also be sure to uncheck "Place each input file into a separate band". Leave the "Advanced parameters" as they appear, and click **Run**.



Your new layer, *Virtual*, is a composite image, a catalog of the 4 rasters we previously had. Nice! This will be much easier for the rest of the analysis. You can now remove the original 4 rasters from the Layers Panel and leave only the output *Virtual*.

The above methods allow you to merge datasets using a catalog, and to reproject them "on the fly". However, if you are setting up data that you'll be using for quite a while, it may be more efficient to create new rasters that are already merged and reprojected. This improves performance while using the rasters in a map.

To reproject your raster, open the **Warp (reproject)** tool from **Processing toolbox > GDAL > Raster projection**, and as usual you can simply search "warp" as well. Selecting the *Virtual* layer as your input, set the output CRS to EPSG:32733. Just save it as a temporary output and examine the small changes between *Reprojected* and *Virtual*.

Let's export this newest file, *Reprojected*, to an appropriate spot in our workspace. Name it "lab8_aerial" and add it to the project when you're done. It may load with a slightly different coloring. If so, you can change this by navigating to the layer styling panel and setting the contrast enhancement to "No enhancement."

In the next section, we'll look at some raster data that isn't aerial imagery, and see how it can be manipulated with various symbology.

STEP 2: CHANGING RASTER SYMBOLOGY

Not all raster data consists of aerial photographs. There are many other forms of raster data, and in many of those cases, it’s essential to symbolize the data properly so that it becomes properly visible and useful.

Navigate to the *exercise_data* folder and in the *SRTM* folder add the layer “srtm_41_19_4326.tif”. Rename it “DEM”, and be sure to zoom to layer after loading it.

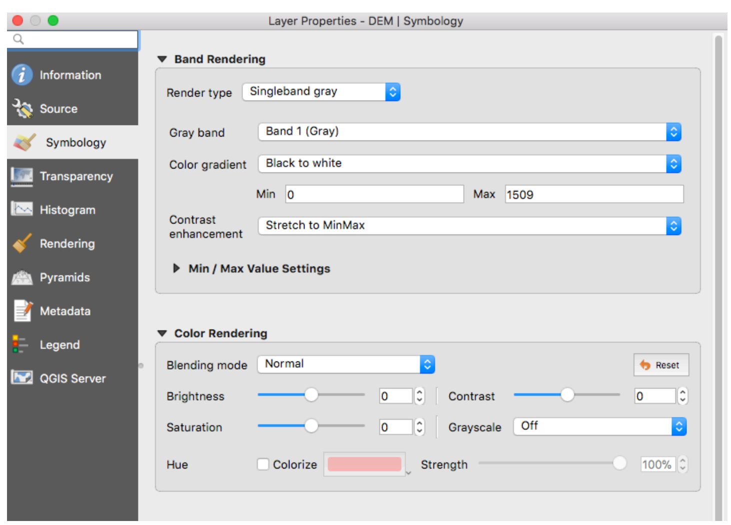
As we’ve discussed in previous labs, DEM stands for “digital elevation model”. These files show elevation (altitude) of the terrain, allowing us to see where the mountains and valleys are, for example. While each pixel of dataset of the previous section contained color information, in a *DEM* file, each pixel contains elevation values. Once it’s loaded, you’ll notice that it’s a basic stretched grayscale representation of the DEM. QGIS has automatically applied a stretch to the image for visualization purposes, and we will learn more about how this works as we continue.

> **QUESTION 1:** Is the aerial imagery we loaded in step 1 of categorical or quantitative type? Why? (See the lecture slides for 3/30/20 if you need help.)
>
> **QUESTION 2:** Likewise, is the DEM we just loaded of categorical or quantitative type? Why?

You have basically two different options to change the raster symbology:

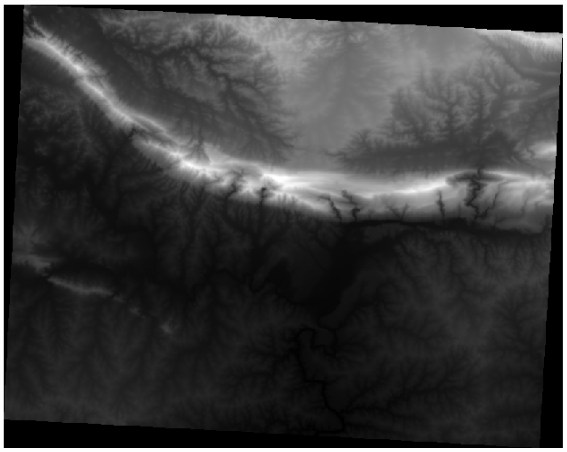
1. The properties tab;
2. The layer styling panel.

Choose the method you prefer to work with, and let’s change the symbology of the DEM we’ve just loaded. When you load a raster file, if it is not a photo image like the ones of the previous section, the default style is set to a grayscale gradient. Let’s explore some of the features of this renderer.



The default Color gradient is set to **Black to white**, meaning that low pixel values are black and while high values are white. Try to invert this setting to **White to black** and see the results.

Very important is the **Contrast enhancement** parameter: by default it is set to Stretch to MinMax meaning that the grayscale is stretched to the minimum and maximum values. Look at the difference with the enhancement (left) and without (right):



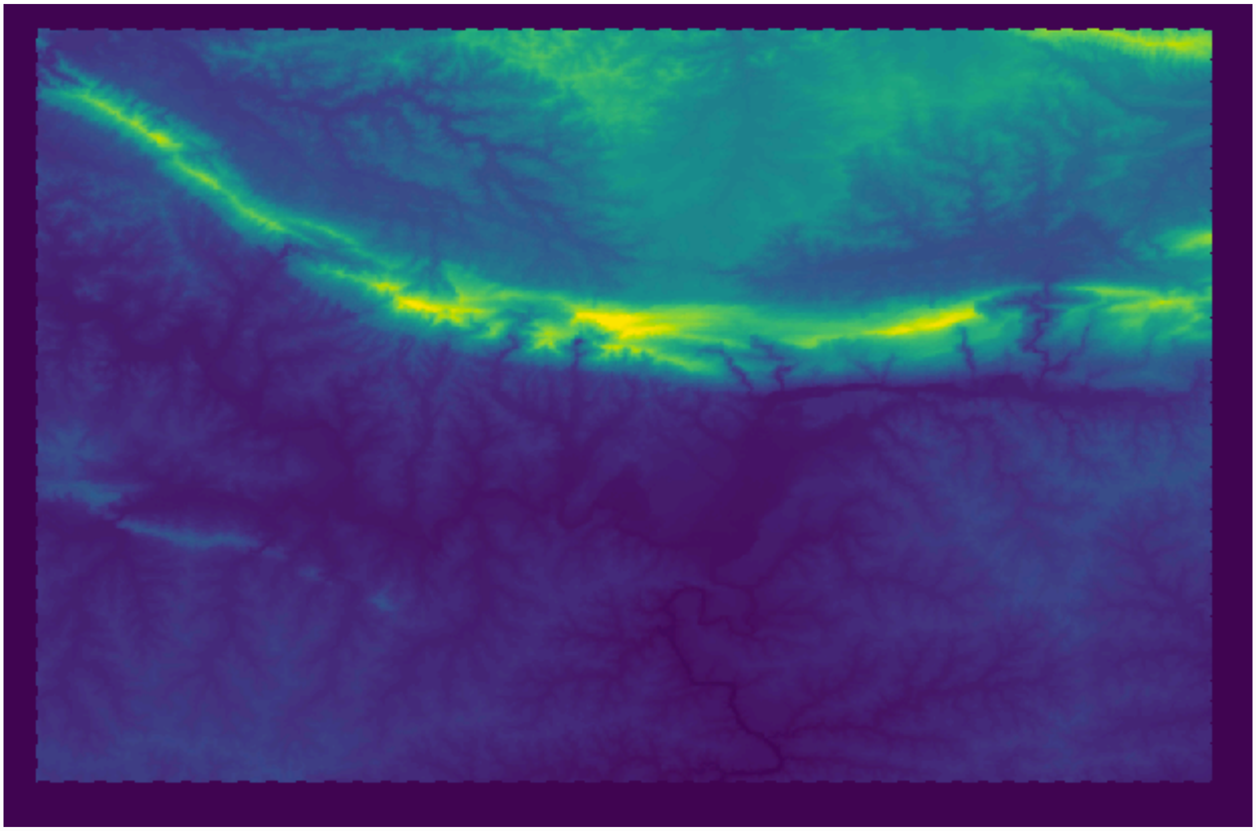
But what are the minimum and maximum values that should be used for the stretch? The ones that are currently under Min / Max Value Settings. There are many ways that you can use to calculate the minimum and maximum values and use them for the stretch:

1. **User Defined**: you choose both minimum and maximum values manually;
2. **Cumulative count cut**: this is useful when you have few extreme low or high values. It *cuts* the 2% (or the value you choose) of these values;
3. **Min / max**: the *real* minimum and maximum values of the raster;
4. **Mean +/- standard deviation**: the values will be calculated according to the mean value and the standard deviation.

Grayscales are not always great styles for raster layers. Let's try to make the DEM layer more colorful. First, change the **Render** type to **Singleband pseudocolor.**

If you don't like the default colors loaded, click on **Color ramp** and change them. Click the Classify button to generate a new color classification. If it is not generated automatically, click on the **OK** button to apply this classification to the DEM.

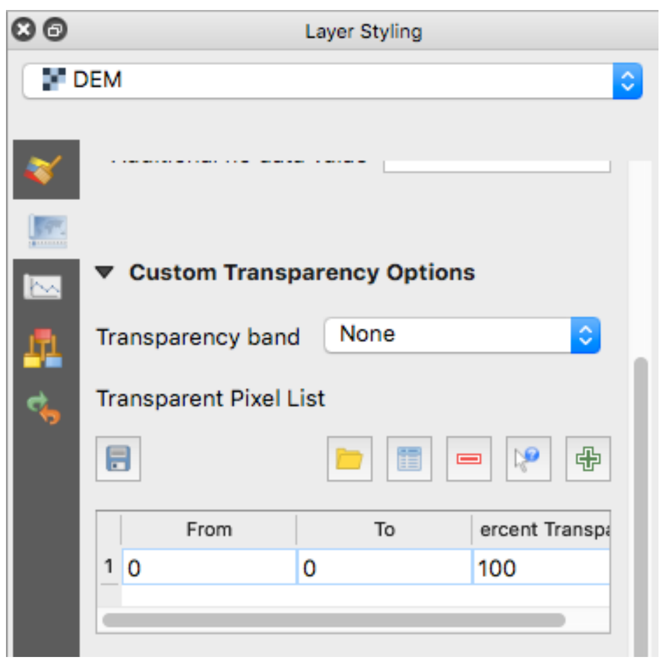
You should see an image that resembles the following, depending on what ramp you chose:

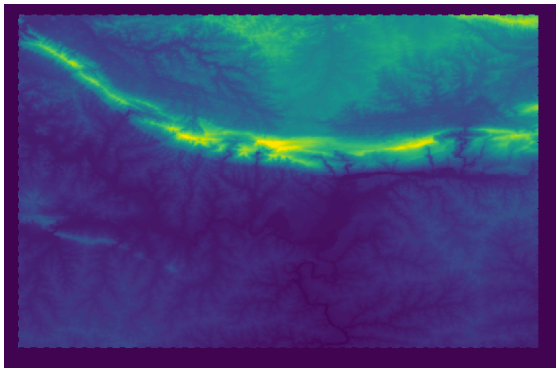
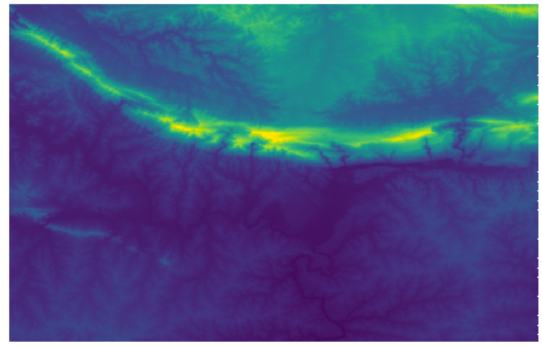


This is a potentially helpful way of looking at the DEM. You'll now see that the values of the raster are again properly displayed, with the darker colors representing valleys and the lighter ones, mountains.

You can also opt to change the transparency of your raster, just as you would any old vector layer. Navigate to the "Transparency" tab and set the **Global opacity** a bit lower.

Finally, notice how this does not isolate the boundary outside of the DEM. We can get rid of that pesky boundary with the use of custom transparency settings. Navigate to the "Custom transparency options" in the layer styling panel, and click the green plus button. Set the "From" and "To" values at 0, and be sure the "Percent Transparency" is set to 100. See the old vs. new result below:



In the final section, we'll get a sense of how to actually perform some analysis on these raster layers – specifically, we'll be developing a hillshade in order to derive some new information about the selected terrain.

STEP 3: HILLSHADE, SLOPE, AND ASPECT

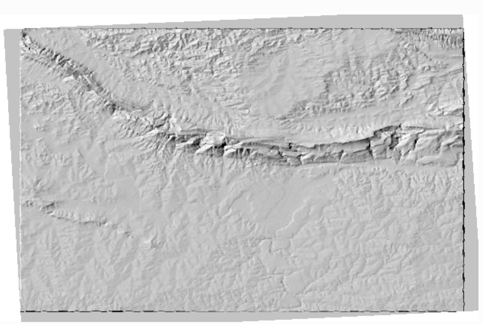
Where the previous section used "srtm_41_19_4326.tif", this time we're going to begin by adding "srtm_41_19.tif" to the data frame. Remove the previous DEM from your project, and rename "srtm_41_19.tif" to "DEM". As in the previous step, set the *DEM* layer to singleband pseudocolor. I recommend the Viridis color ramp.

DEM shows you the elevation of the terrain, but it can sometimes seem a little abstract. It contains all the 3D information about the terrain that you need, but it doesn't look like a 3D object. To get a better look at the terrain, it is possible to calculate a **hillshade**, which (as we've discussed in previous labs) is a raster that maps the terrain using light and shadow to create a 3D-looking image.

Navigate to the **Hillshade** tool, located within **Processing > Raster > Raster terrain analysis**. The algorithm allows you to specify the position of the light source.

The Azimuth parameter has values from 0 (North) through 90 (East), 180 (South) and 270 (West) while the Vertical angle sets how high the light is. We will leave the default values. Leave the default settings, and **Run** the tool.

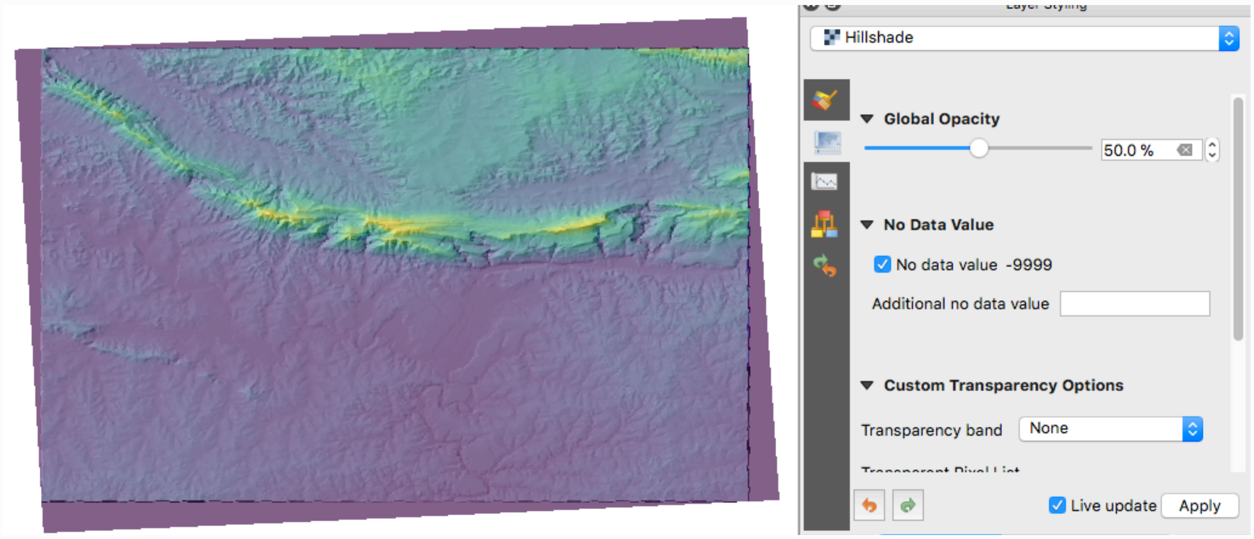
Your output should resemble the following:



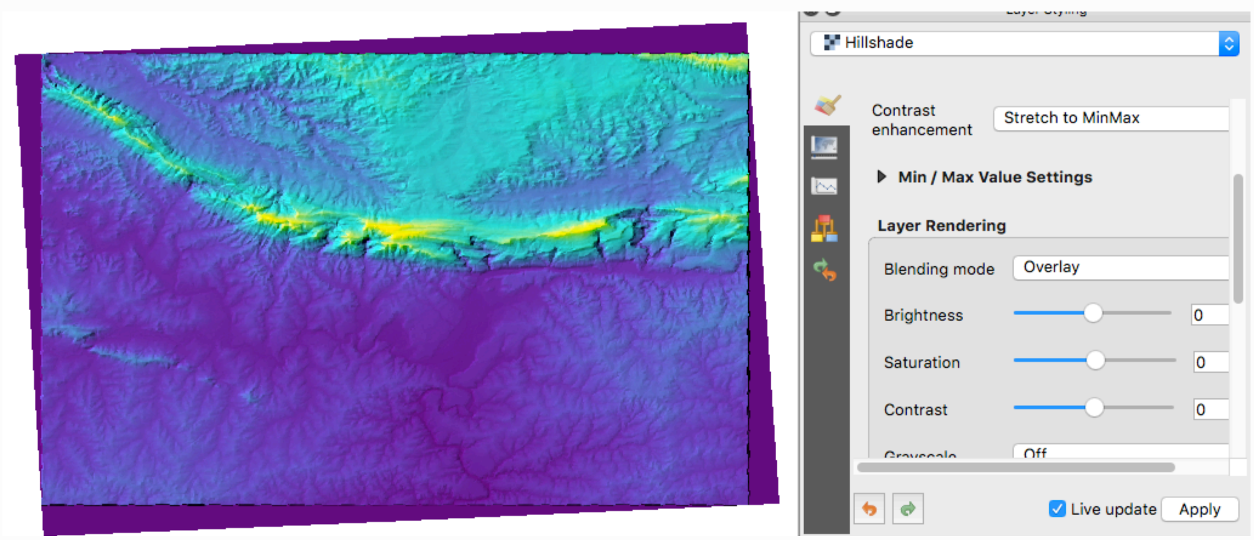
This is fine, and gives us a sense of elevation that is more visually intuitive than before. However, we can improve it by **overlaying** it with the original DEM.

As a point of order, you should have two layers in your Layers Panel: 1) a file called *DEM* that is styled with a singleband pseudocolor, and 2) a file called *Hillshade* that emulates a 3D landscape. *Hillshade* should be above *DEM* in the layer ordering.

There are two ways to overlay your hillshade layer with the original DEM. First, you can simply adjust transparency. In the layer styling panel for *Hillshade*, set the global opacity to 50%. The output should resemble below:



Alternatively, you can use QGIS' built-in **Blending** techniques. Set the global opacity back to 100%. Then, in the **Symbology** tab for *Hillshade*, scroll down to "Layer Rendering" and reset "Normal" to "Overlay". The result should resemble the following:

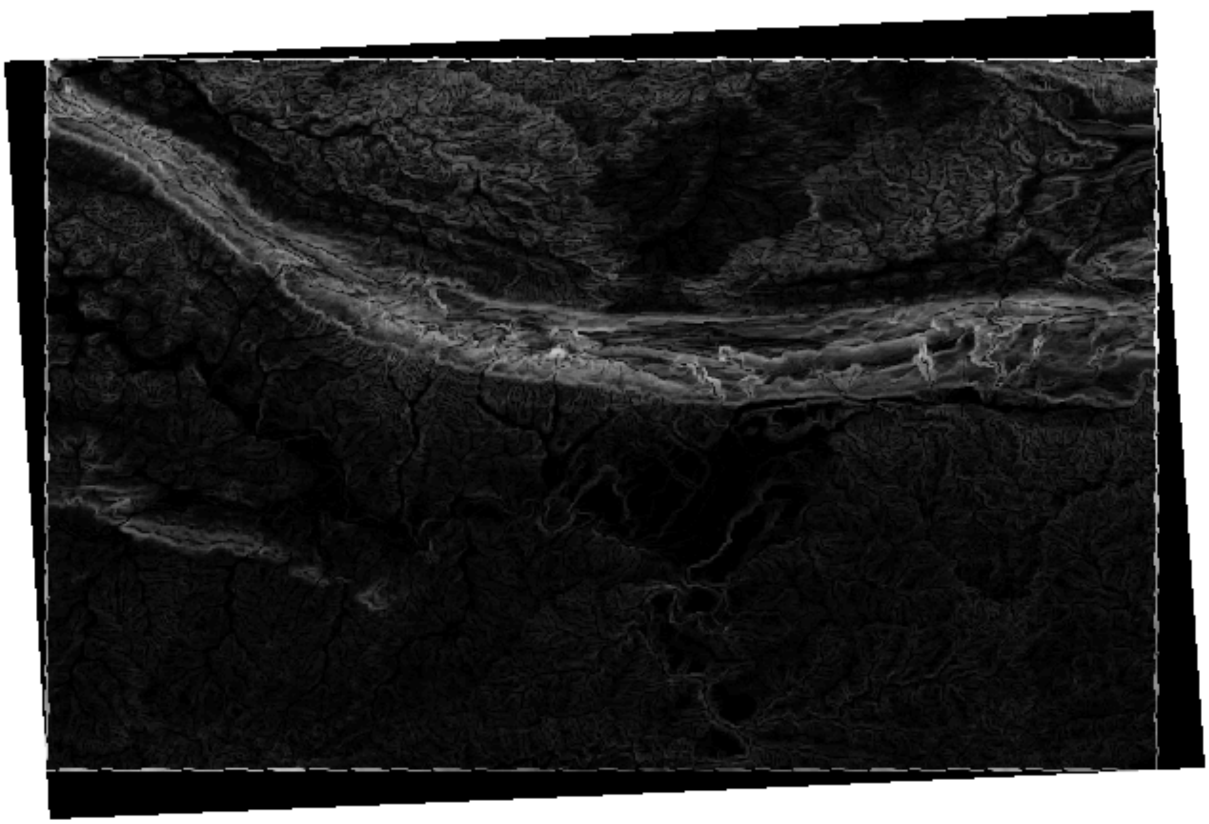


Feel free to play around with the rendering techniques. Slide the brightness and contrast around, or maybe play with the opacity. You can also try putting *DEM* on top of *Hillshade*, and using the “Darken,” “Multiply,” or “Burn” blending techniques – I think all of those combinations look pretty cool. There is no “right” answer here; rather, it’s up to you to decide on what the output should look like.

> **QUESTION 3:** Take a screenshot of your QGIS data frame with the layers blended and include it in your word doc for submission. In a sentence, note what parameters you selected, including the opacity, the blending mode, and any changes to brightness, saturation, or contrast.

Now we’ve got a good-looking overlay. Next, we’ll next be calculating *slope*, or the how steep the land actually is. This is important, for example, if you want to build houses in an area and need land that is relatively flat.

To do this, navigate to the **Slope** tool in the processing toolbox. Choose “DEM” as your input layer, and click **Run**. Make sure that both *Hillshade* and *DEM* are unchecked in the layers panel. The output should resemble the following:



In this output, black pixels are flat terrain, while white pixels represent steep terrain.

Sticking with the example of selecting an area for residential development, we should determine the study area’s *Aspect* in addition to its slope. *Aspect* is the compass direction that the slope of the terrain faces. An aspect of 0 means that the slope is North-facing, 90 East-facing, 180 South-facing, and 270 West-facing.

Since this study is taking place in the Southern Hemisphere, properties should ideally be built on a north-facing slope so that they can remain in the sunlight.

Select the **Aspect** tool in **Processing** > **Raster terrain analysis** to produce the layer.

> **QUESTION 4:** Take a screenshot of the **Aspect** output and copy-paste it into your word doc.

In the next section, we'll be imagining a scenario that expands on the current example of selecting an area for residential development.

STEP 4: RASTER CALCULATION AND RECLASSIFICATION

Let's imagine that residential development has begun in this area. Now, the Zook family wants to purchase some land, and build a small house on the property. In the Southern Hemisphere, an ideal plot for development needs to have the following criteria:

- An aspect that is north-facing
- A slope of less than five degrees
- **BUT,** if the slope is less than 2 degrees, then the aspect doesn't matter

Fortunately, you already have rasters showing you the slope as well as the aspect, but you have no way of knowing where both conditions are satisfied at once. How could this analysis be done?

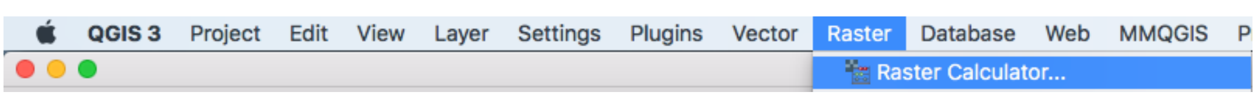
The answer lies with the Raster calculator.

QGIS has different raster calculators available:

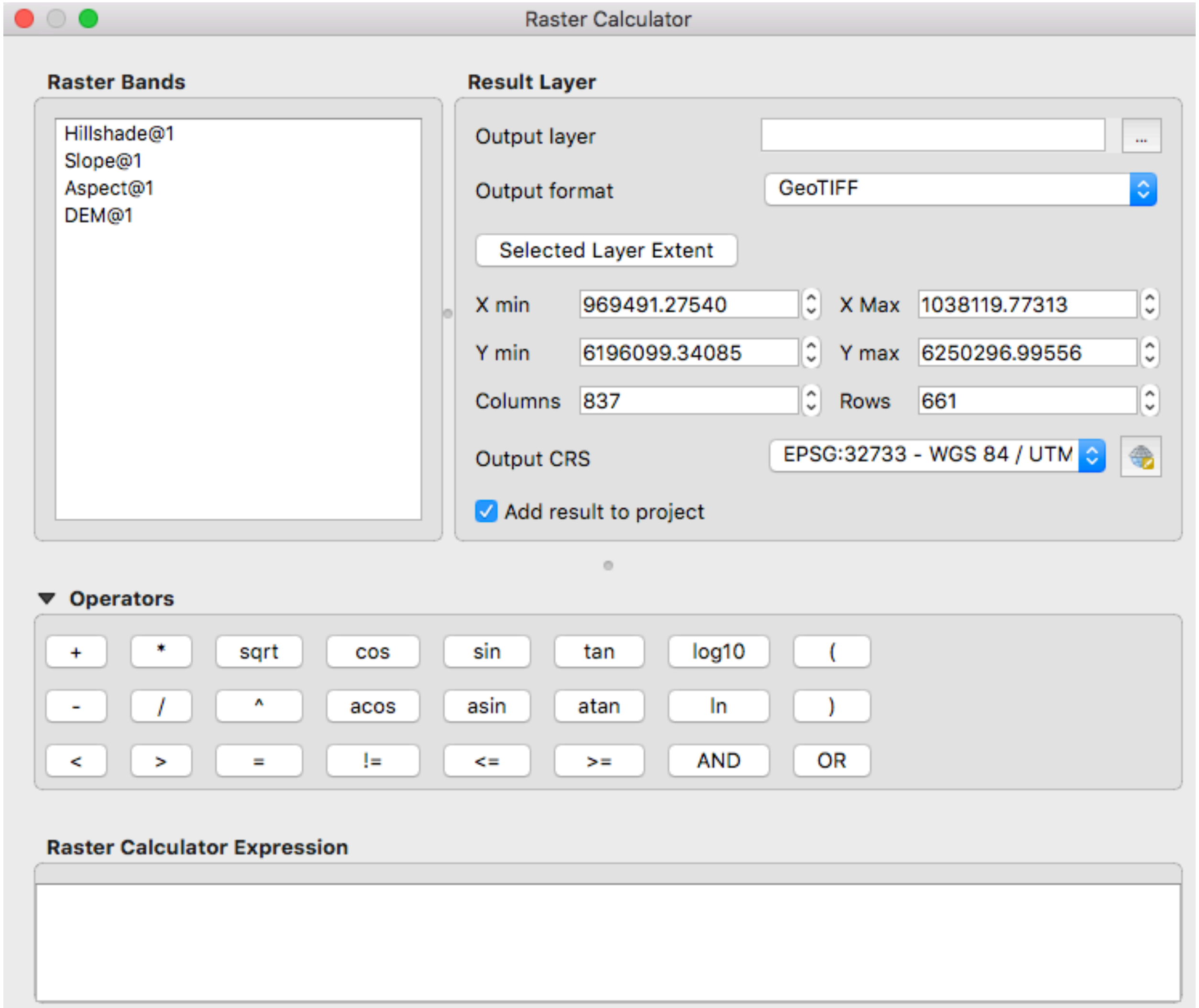
- **Raster > Raster Calculator**
- **Processing > Raster Analysis > Raster calculator**
- **Processing> GDAL > Raster miscellaneous > Raster calculator**
- **SAGA > Raster calculus > Raster calculator**

Each tool is leading to the same results, but the syntax may be slightly different and the availability of operators may vary. We will use **Raster > Raster calculator.**

Open the tool by navigating to **Raster > Raster calculator** in the main toolbar:



You should then see a pretty busy-looking dialog box:



The upper left part of the dialog lists all the raster layers loaded in the legend as “name@N”. “Name” is the name of the layer in your Layers panel, and “N” is the raster band used.

In the middle part you will see a lot of different **operators**. Consider that a raster is an image, so you can think of it as a 2D matrix filled with numbers, as in the image to the right. We can perform various mathematical operations on these values to get new ones.

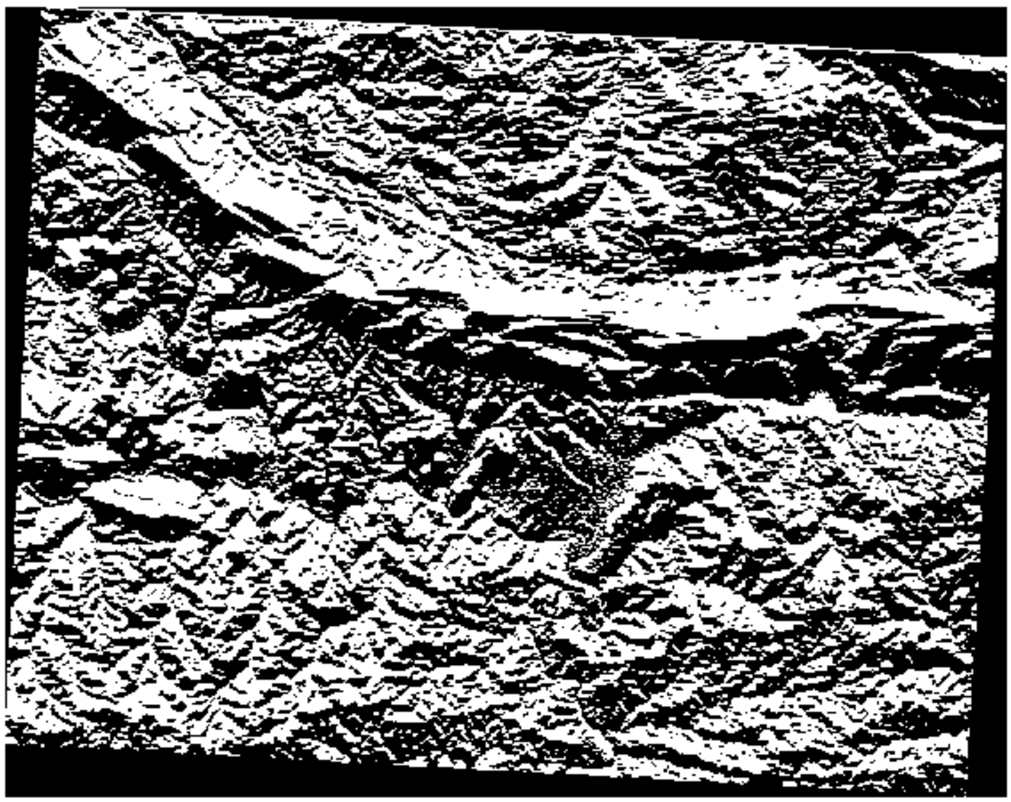
1	1	1	1	1	1	1	3	3	3
1	1	1	1	1	1	1	3	3	3
1	1	1	1	1	1	3	3	3	3
1	1	1	2	2	2	2	3	3	3
1	1	1	2	2	2	2	3	3	3
1	1	1	2	2	2	2	3	3	3
1	1	1	1	2	2	2	3	3	3
1	1	1	1	1	1	3	3	3	3
1	1	1	1	1	1	1	3	3	3
1	1	1	1	1	1	1	1	3	3

North is at 0 (zero) degrees, so for the terrain to face north, its aspect needs to be greater than 270 degrees and less than 90 degrees.

Therefore the formula is:

"Aspect@1" <= 90 OR "Aspect@1" >= 270

Paste this formula into the Raster Calculator Expression. The output CRS should be EPSG:32733. Save the output in an appropriate location as “aspect_north” and click **Run.** You should see the following output:



What are we looking at here? The output values are either 0 or 1 (you can see this in the Layers panel, but you can also inspect it more closely by navigating to the **Histogram** in the Properties panel). The formula we had written contains the conditional operator OR; therefore, the output is either False (0) or True (1). These binary, 0/1 relationships are called *Boolean* variables – variables with only 2 possible values, true or false. In plain terms, the raster we’re looking at shows north-facing terrain in the white (True) cells, and not north-facing in the black (False) cells.

Recall the current scenario, in which the Zook family wants to build a new home on some land that meets three criteria:

- A north-facing aspect
- A slope of less than five degrees
- **BUT,** if the slope is less than 2 degrees, then the aspect doesn’t matter

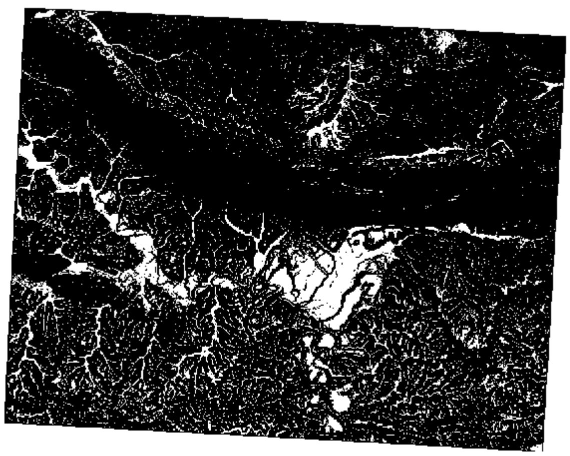
We've successfully identified the first thing (terrain with a north-facing aspect), but how can we figure out the next two? We'll simply run the **Raster calculator** again – only this time, we'll use the *Slope* layer as the input.

Open up the **Raster calculator** and enter the following in the expression dialog box:

"Slope@1" <= 2

Name your output "slope_lte2" and **Run**. Then, run the tool again, changing the 2 to a 5 and saving the output "slope_lte5". The results should resemble the following:

slope_lte2



slope_lte5

Now you should have 3 layers produced using the **Raster calculator**:

- *aspect_north*: the terrain faces north in white cells
- *slope_lte2*: the slope is at or below 2 degrees in white cells
- *slope_lte5*: the slope is at or below 5 degrees in white cells

What if we wanted to combine these three separate layers to encode all of this information in a single raster? How would we go about it?

Each of the cells where the conditions were met contains a 1 value; other cells contain 0. Therefore, if you multiply one of these rasters by another one, you will get areas where both of them are equal to 1.

The conditions to be met are: at or below 5 degrees of slope, the terrain must face north; but at or below 2 degrees of slope, the direction that the terrain faces in does not matter. Therefore, you need to find areas where the slope is at or below 5 degrees AND the terrain is facing north, OR the slope is at or below 2 degrees. Such terrain would be suitable for development.

Open the **Raster calculator** again. Use the Layer panel, the Operators buttons, and your keyboard to build this expression in the Expressions text area:

("aspect_north@1" = 1 AND "slope_lte5@1" = 1) OR "slope_lte2@1" = 1

Save the output as "all_conditions.tif" and click **Run.** Your results should resemble this:

QUESTION 5: Paste a screenshot of the output *all_conditions* from the most recent raster calculation. What do the black cells represent? What do the white cells represent?

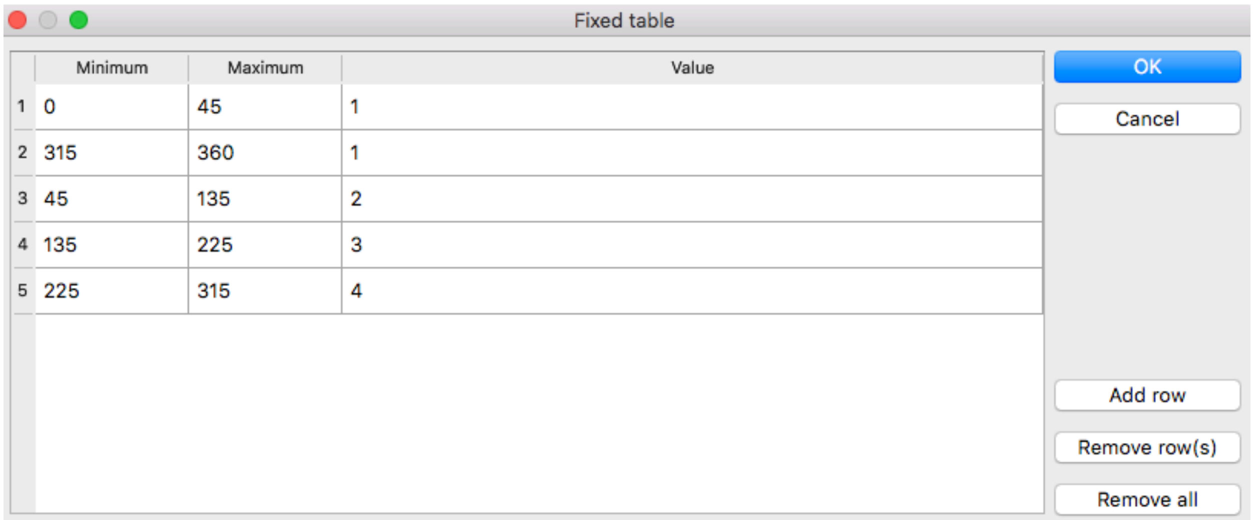
Finally, we are going to finish this analysis by using the **Reclassify** tool, which allows us to extract information from existing layers in a different way.

Back to the aspect layer: we know now that it has numeric values within a range from 0 through 360. What we want to do is to reclassify this layer with other discrete values (from 1 to 4) depending on the aspect:

1 = North (from 0 to 45 and from 315 to 360);
2 = East (from 45 to 135)
3 = South (from 135 to 225)
4 = West (from 225 to 315)

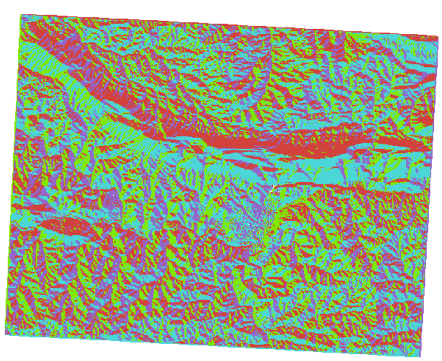
Recall how in previous labs, we "ranked" places based on socioeconomic distress. This process is similar. The desired operation could be achieved with the raster calculator but the formula would become very large. Instead, we'll use the **Reclassify by table** tool located in **Processing > Raster analysis**.

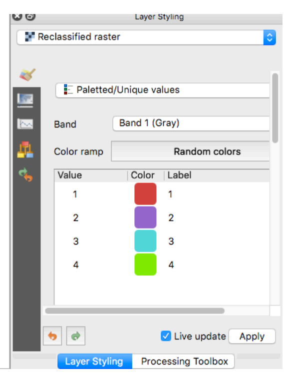
First, open the tool. Use *Aspect* as the Raster Layer – **not *aspect_north!*** Then, clicking the ellipsis next to "Reclassification table", and add the following rows:



	Minimum	Maximum	Value
1	0	45	1
2	315	360	1
3	45	135	2
4	135	225	3
5	225	315	4

You can leave all the Advanced Parameters as they are, and click **Run**. In the output, navigate to the Layer Styling panel and set the categorization to "Paletted/Unique values", and click **Classify**. The raster should automatically assign 4 random colors to each class, as shown below:



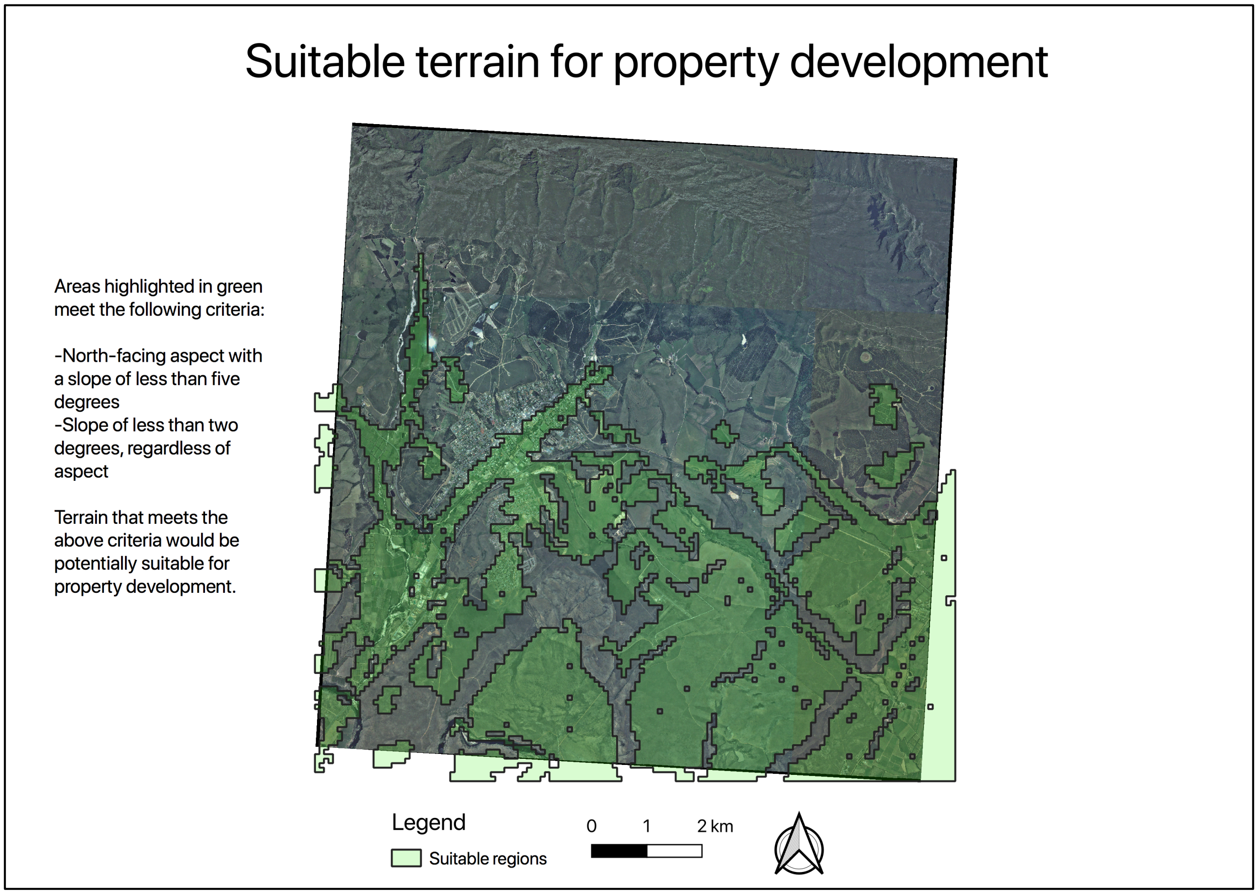


This reclassification makes it much easier to actually distinguish between aspect areas, rather than estimating based on a large color ramp with dozens (and sometimes even hundreds!) of discrete values.

> **QUESTION 6:** Copy and paste your reclassified raster into your word doc for submission. Open up the Properties tab of your reclassified raster and select **Histogram**. Which value – 1, 2, 3, or 4 – contains the most instances?

STEP 5: REVIEW & SUBMISSION OF FINAL MAP

In the final step, you will produce a map that shows suitable terrain for property development within the study area. You already have most of the data you need, but you'll have to run a couple of final tools in order to produce a map that resembles the following:



How can you make this map? You'll want to begin by loading the following layers into a QGIS project (and probably getting rid of the rest of them):

- *lab8_aerial*
- *all_conditions*

Once you've done this, complete the following workflow:

1. **Clip raster by extent** using *all_conditions* as the input and *lab8_aerial* as extent – save as temporary layer
2. **Polygonize** using the output *Clipped (extent)* as the input – save as temporary layer
3. Open the new *Vectorized* layer and remove all records where DN = 0

4. Create a new field called “sq_meters” and calculate area using the **Field calculator** – the expression should be “$area” (without quotes) – **note that your Project Properties should be using square meters as units for measurement**
5. Using **Filter** or **Select by expression**, select all records where “sq_meters” > 120000
6. Save the selected records as “suitable_development” in an appropriate location in your workspace
7. You should now have 1) an aerial image and 2) a vector layer that encompasses all regions with a north-facing aspect and a slope less than 5, OR a slope less than 2
8. Layout in Print Composer, including your name and date, and submit!

You should submit completed materials for Lab 8 via Canvas by Monday, 4/20 at 11:59pm. **Recall that I am not penalizing for late submissions anymore, but you must submit before Monday 5/4 in order to be graded.** The completed materials include:

- A Word document containing answers to the questions posed throughout the lab
- One map that shows suitable terrain for property development

As always, be in touch if you have any questions!