Asmaa Abdelbaki, Max Hess, and Sofia Viotto

Lab Assignment #1 – Cartography, Colors, and Displaying Remotely-Sensed Data

1. Logistics

Date assigned:	Monday, Oct 14, 2024
Date due:	Monday November 4, 2024 (via Moodle)
Points:	20 points

Please submit all homework as a single PDF file via Moodle.

In this lab, you will work with two types of geographic data (vector and raster) in a desktop environment. We will show how to make maps using common geographic data, and what elements should be included on all assignments. We will also discuss software needs for the class.

2. Software Installation and Account Creation

To make the class go smoothly, take a look at the 'Software Installation' document on Moodle. Please install the required software (**QGIS**, **GDAL**, and **SNAP**) to make sure that you will not fall behind in any of the labs. If you have a slow internet connection/computer, it is enough to install only **QGIS** today, but please make sure all other software is installed!

If you already have **Anaconda/miniconda** installed, that is the easiest way to install **GDAL**. Otherwise, you can install it from a stand-alone installer.

Next, please request access to **Google Earth Engine** (https://earthengine.google.com/), which we will use in a later lab. Sometimes it takes a few days for access to be granted, so please do this now!

3. Introduction to QGIS

You will continue to use QGIS over the next weeks' labs, so today we will focus on getting familiar with the software and making a simple map.

(1) Adding raster and vector data

The geospatial data you will work with comes in two primary forms: ***vector*** data are non-gridded shapes or lines (e.g., streets, city boundaries) that may or may not have directional or magnitude information (e.g., GPS geodetic data). In this lab, we will first load simple polygons of continent borders.

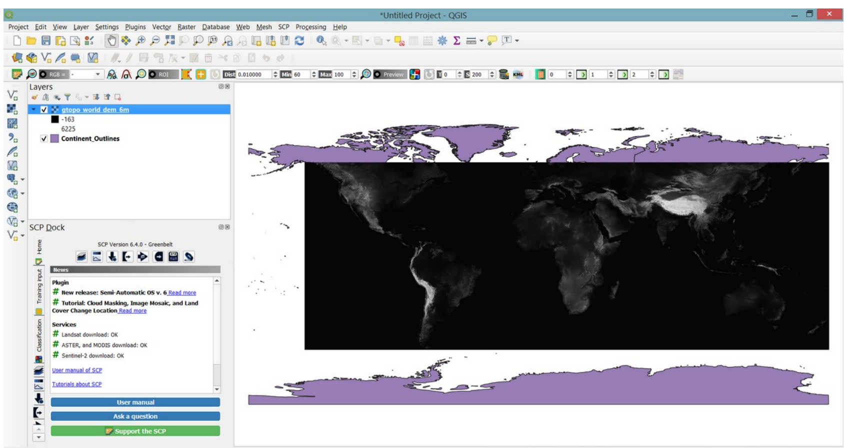
- In QGIS, navigate to the left-hand toolbar. The top button should be ‘Add Vector Layer.’ Click on this and, in the dialog box choose to ***Browse*** for datasets and open ***‘Continent_Outlines.shp’***
- When you try to open the shapefile, you will be prompted by the dialog box ***Coordinate Reference System Selector***. The data should be loaded in ‘WGS 84 / EPSG: 4326’. (This will probably already be selected). Select this coordinate system and load the data.

You should now have a map showing polygons of each continent.

Raster data are gridded datasets that provide pixel-based information, like a digital photograph. Rasters may be single-band (e.g., black and white image) or multiband (e.g., RGB image). In this lab, we will load a single-band digital elevation model (DEM), in which every pixel represents an elevation on Earth’s surface.

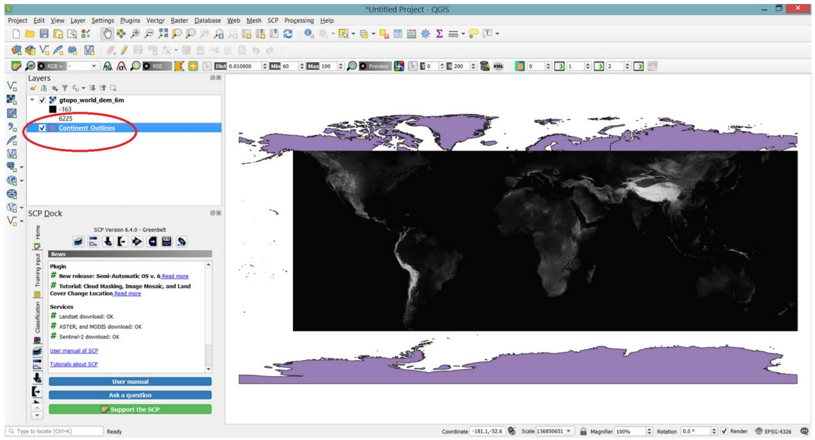
- Beneath the ‘Add Vector Data’ button is the ‘Add Raster Layer’ button. Select this and follow the same procedure to load the raster file, ***‘gtopo_dem_6m.tif’***.
- Again, be sure to load the data in the ‘WGS 84 / EPSG: 4326’ coordinate system.

You should now have a map that looks something like this:

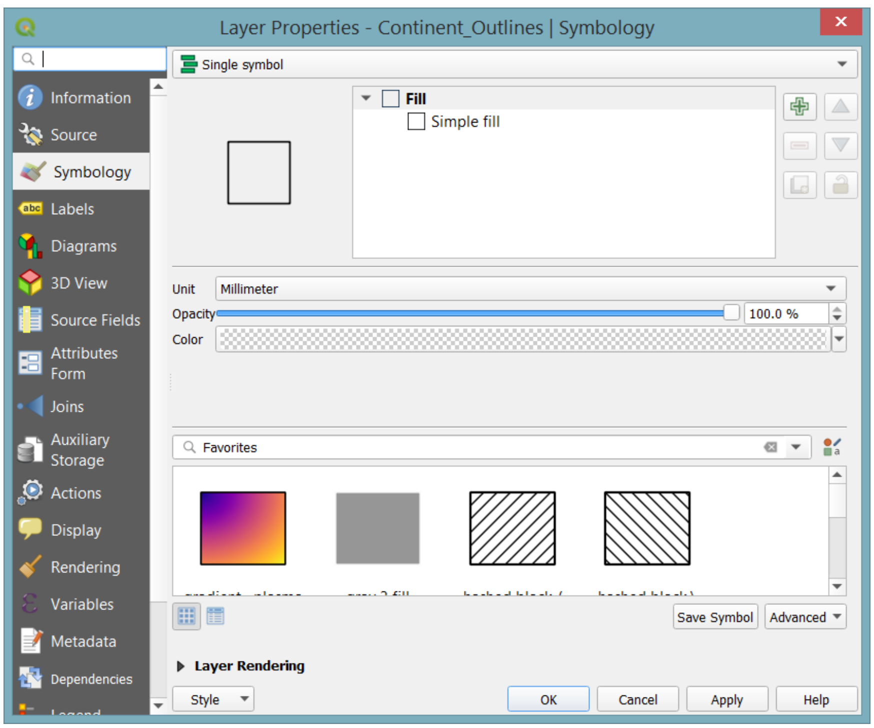


(2) <u>Editing data appearance</u>

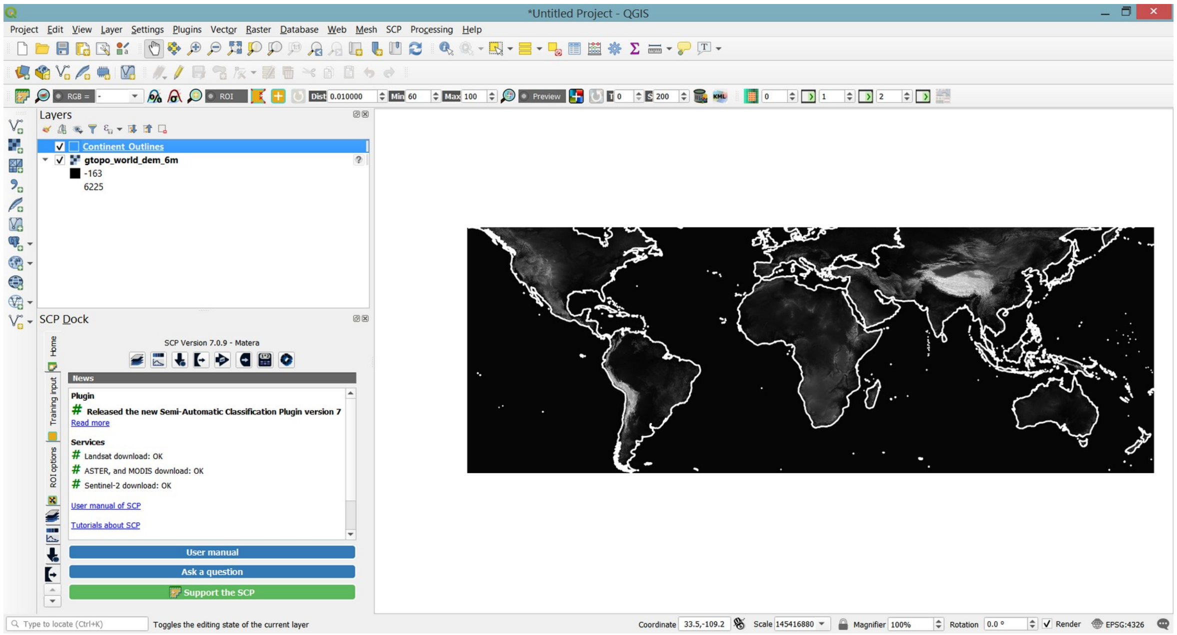
You can change the appearance of the data currently displayed in your map by the layer properties, accessible by right-clicking each layer in the Layers Panel.



The vector data should be ‘hollow’ so we can see the elevation data underneath. For example, choosing an empty fill:



Will look something like this if you choose a ‘white’ outline:


*Untitled Project - QGIS
Project Edit View Layer Settings Plugins Vector Raster Database Web Mesh SCP Processing Help
Layers
Continent Outlines
gtopo_world_dem_6m
-163
6225
SCP Dock
SCP Version 7.0.9 - Matera
News
Plugin
Released the new Semi-Automatic Classification Plugin version 7
Read more
Services
Landsat download: OK
ASTER, and MODIS download: OK
Sentinel-2 download: OK
User manual of SCP
Tutorials about SCP
User manual
Ask a question
Support the SCP
Type to locate (Ctrl+K)
Toggles the editing state of the current layer
Coordinate 33.5,-109.2
Scale 145416880
Magnifier 100%
Rotation 0.0 °
Render
EPSG:4326

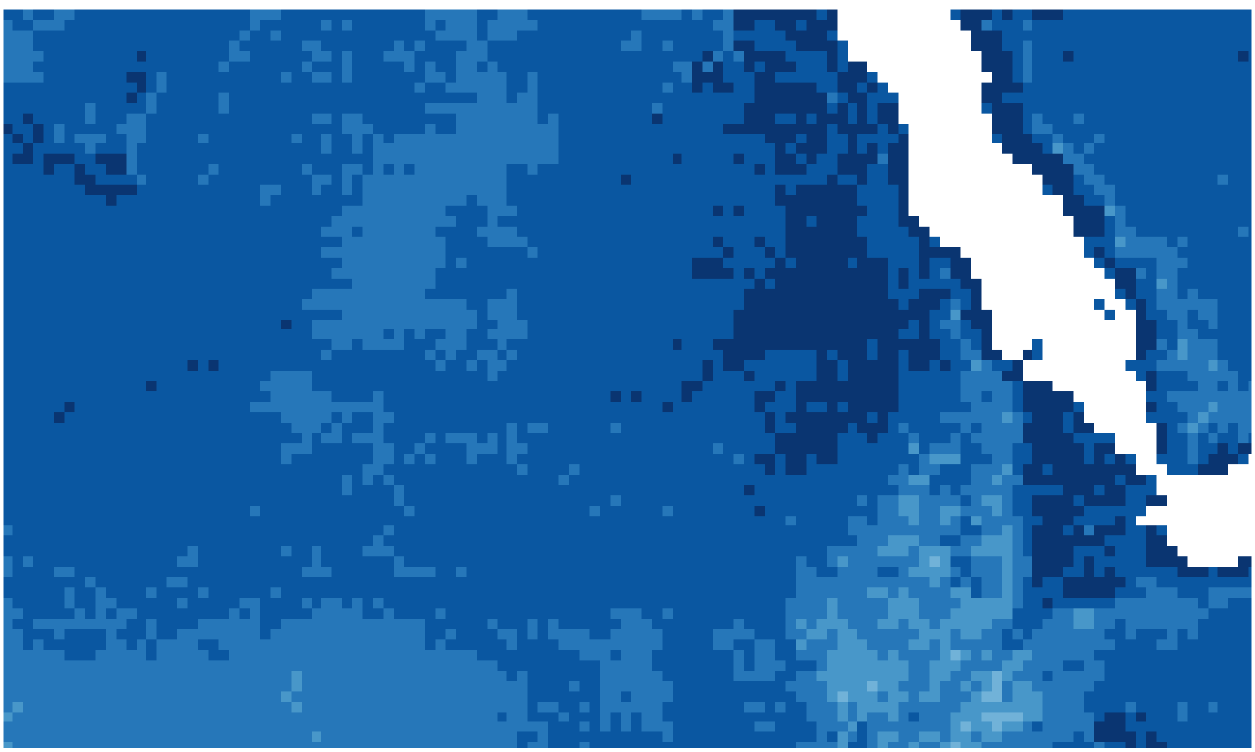
4. Color Schemes and Displaying Data

Many books have been written about how to best use colors when displaying data (e.g., https://colorbrewer2.org/, paper reference on Moodle, many others). There are three main ways to think about separating colors, and which one you choose will be highly dependent on your dataset. The three choices are generally:

1. Categorical – Colors are widely different, often used to separate labelled data (for example, giving each country a different random color on the map).
2. Simple Ramp – Colors blend into each other, often used for continuous data (for example, countries are colored from red to green based on how many people live there).
3. Divergent – Colors converge towards a value, and then diverge away from it (for example, countries are colored based on how close they are to the equator).

There are of course variations these options – many of which are built into QGIS and other software platforms. It is important to keep in mind what your data represents, how it is scaled, and how people will interpret your color scheme.

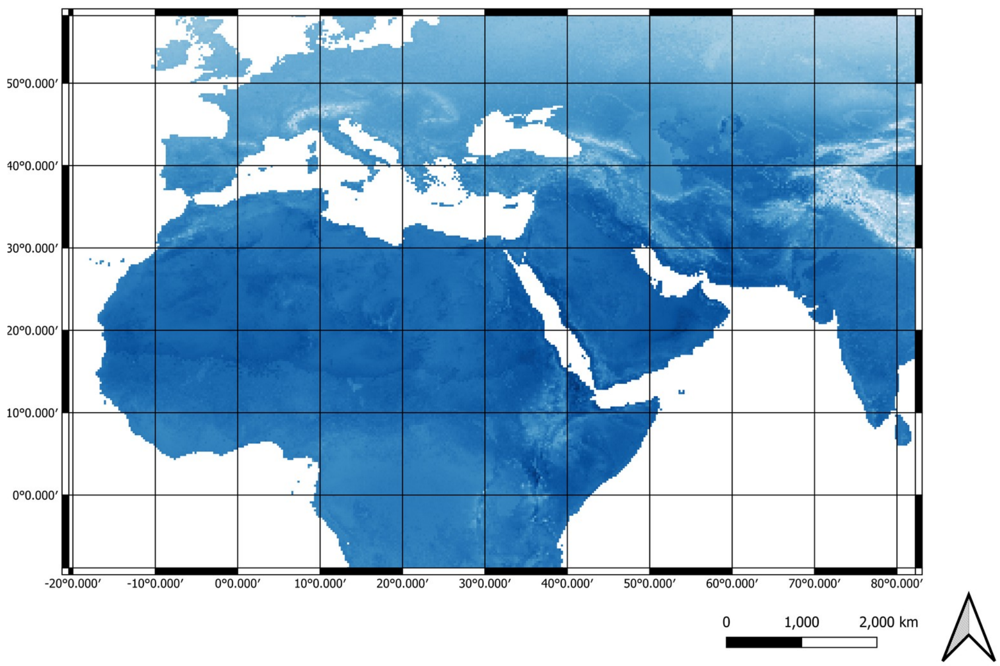
Let’s take an example:



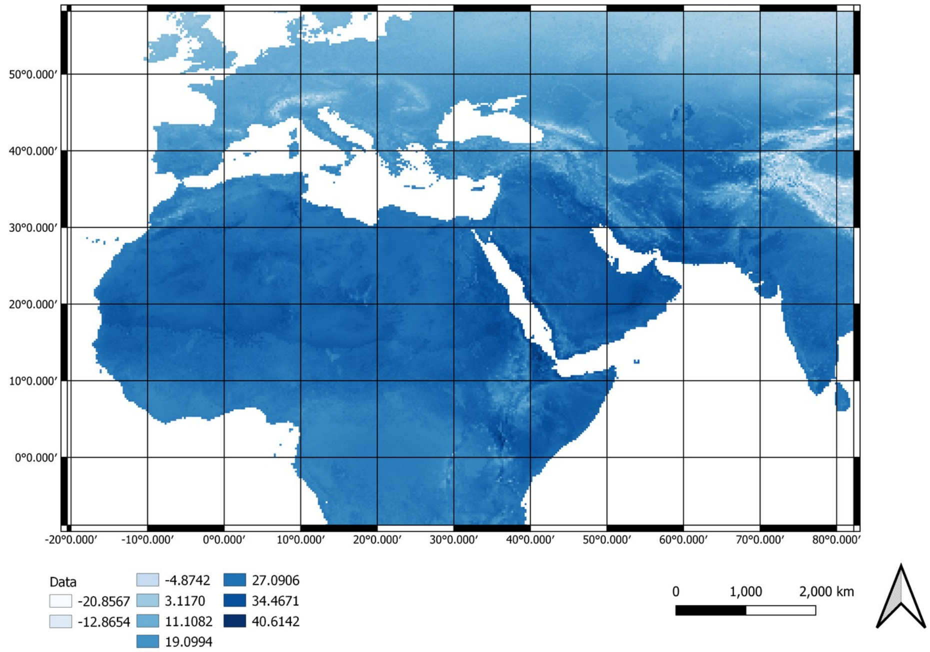
What can you say from this map? What do you think it represents?

It is very hard to say anything from this map because I don’t know where I am looking, what data is being displayed, and what the colors mean!

Let’s take the same data and add some of these markers and try to make some sense of this map.

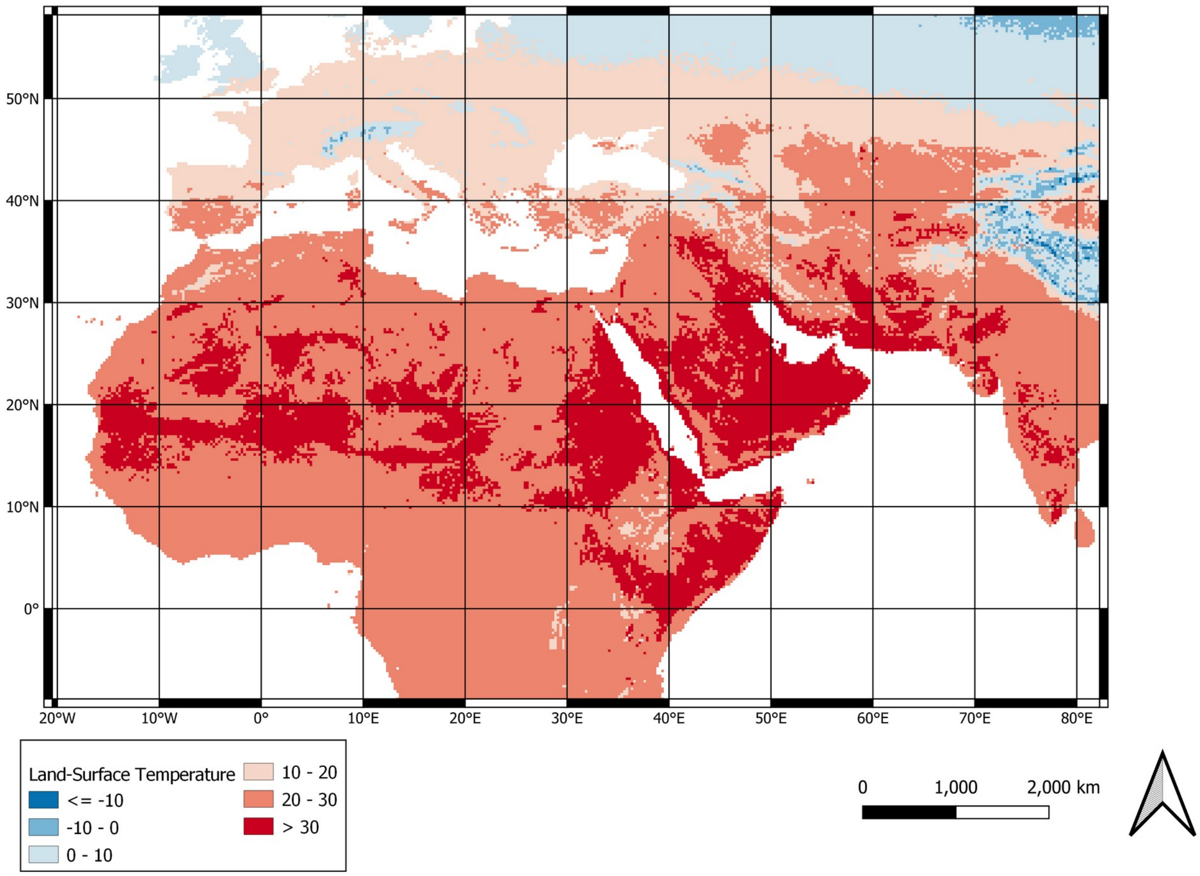


Ok, a bit better! This map already does a few things very well – it has a North Arrow, a Scale, and grid lines. These three things tell us where on the world we are looking. Can you guess what the colors represent? Which are the high ones and which are the low ones?



Ok – so we have some values on our map, but still no idea what they are. Any guesses? Does the number -4.8742 help at all?

It is important to add useful labels to your colorbar, and to choose numbers that are intuitive to the data you want to display! Finally, it is important to choose colors that help explain your map. In this case, this data is **Land-Surface Temperature**. Does a high ‘blue’ value really make sense to represent ‘hot’ places on the earth? How about if we change the color scaling and color bar:

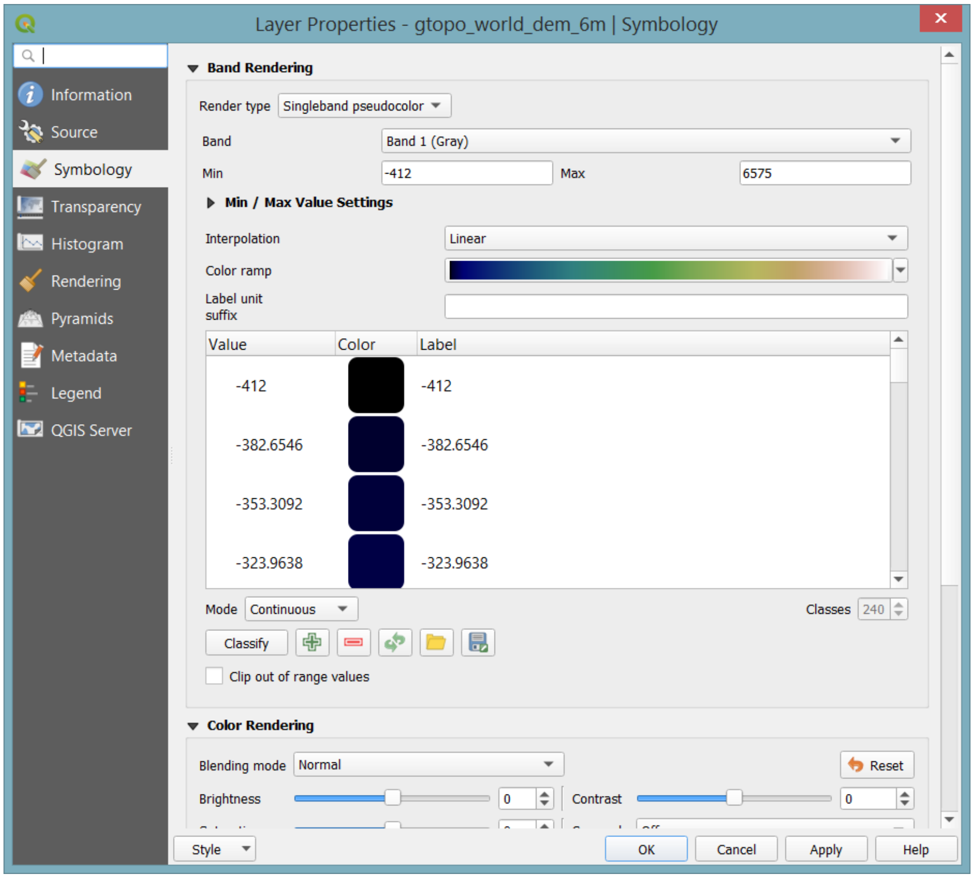


What do you think? Easier to understand?

This map does a much better job of explaining itself: (1) ‘cold’ areas are blue, ‘warm’ areas are red, (2) there is a useful legend with reasonable numbers to interpret the data, (3) you can locate the data on the earth using the scale, arrow, and grid.

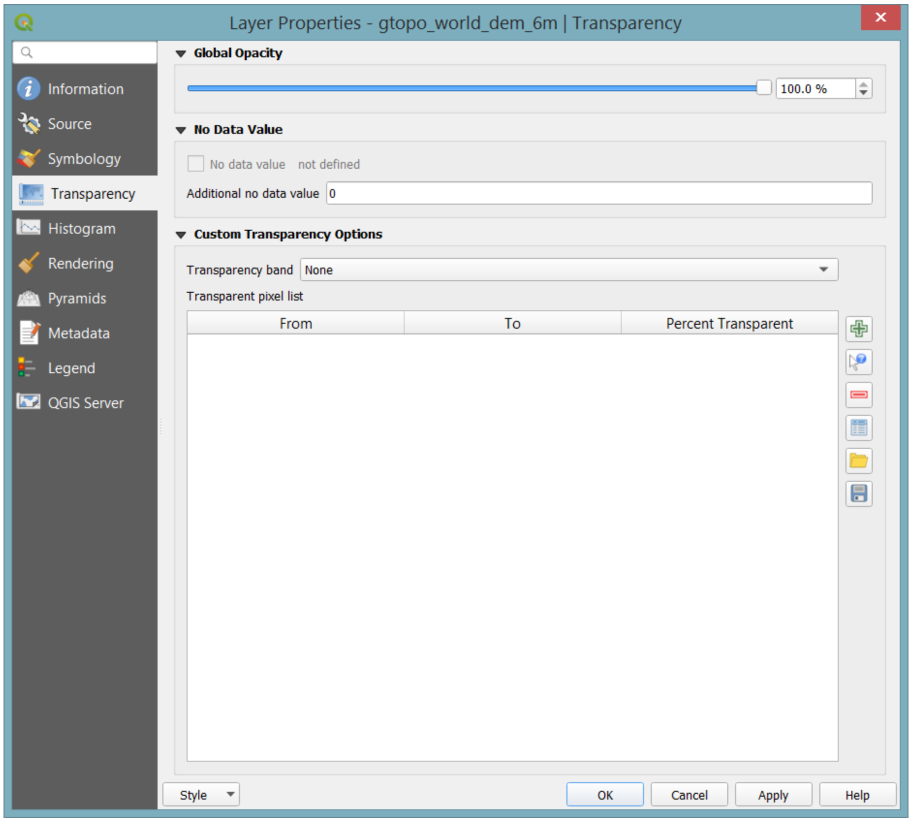
5. Symbolizing our Elevation Dataset

The elevation data should be symbolized with a color scheme that provides a reasonable differentiation between low elevations, high elevations, and oceans, such as that above. You can manually edit the bin values by clicking the numbers in the 'value' column.

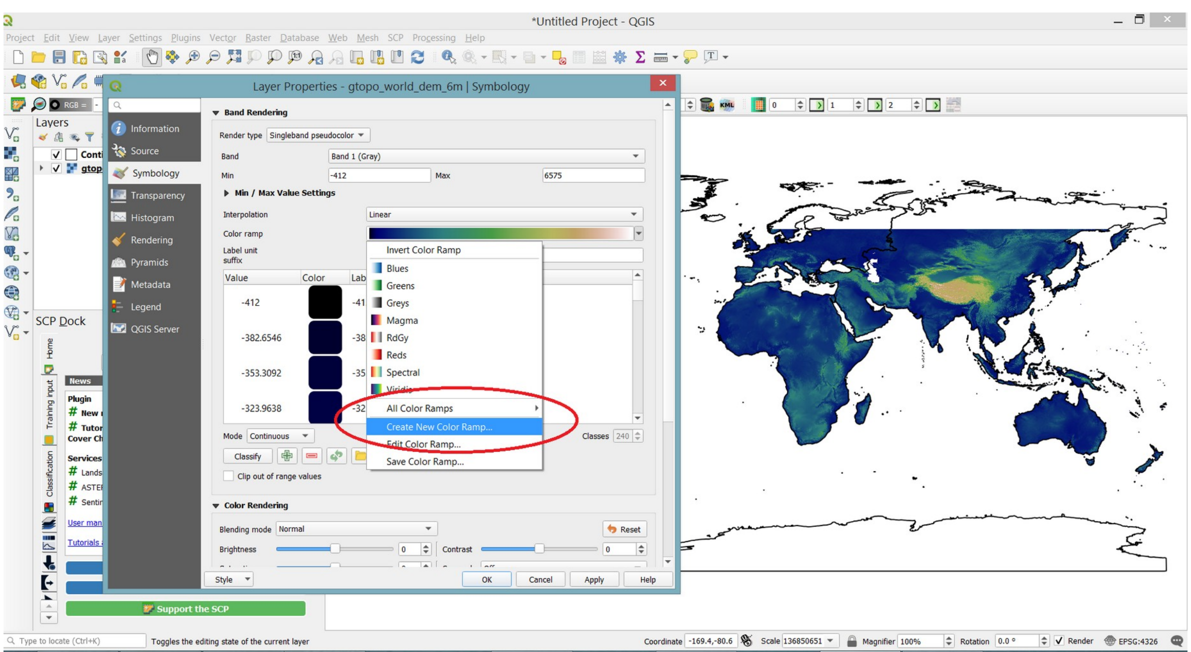


Now, since we have elevation represented by color in the DEM, we can place the raster data on top of the raster data by dragging the vector up the list – layers are drawn from the top down in QGIS (and most GIS systems).

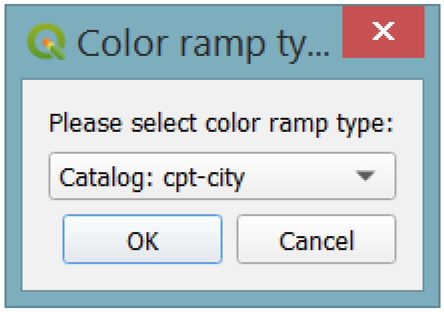
Note that the oceans are likely a weird color – this is because they are defined as 'zero' in our dataset, and there are also areas of the world with an elevation below zero. To change this, we need to define a 'NoData' value for the dataset. Some datasets have this already defined, and some do not. We can check this under properties – transparency:



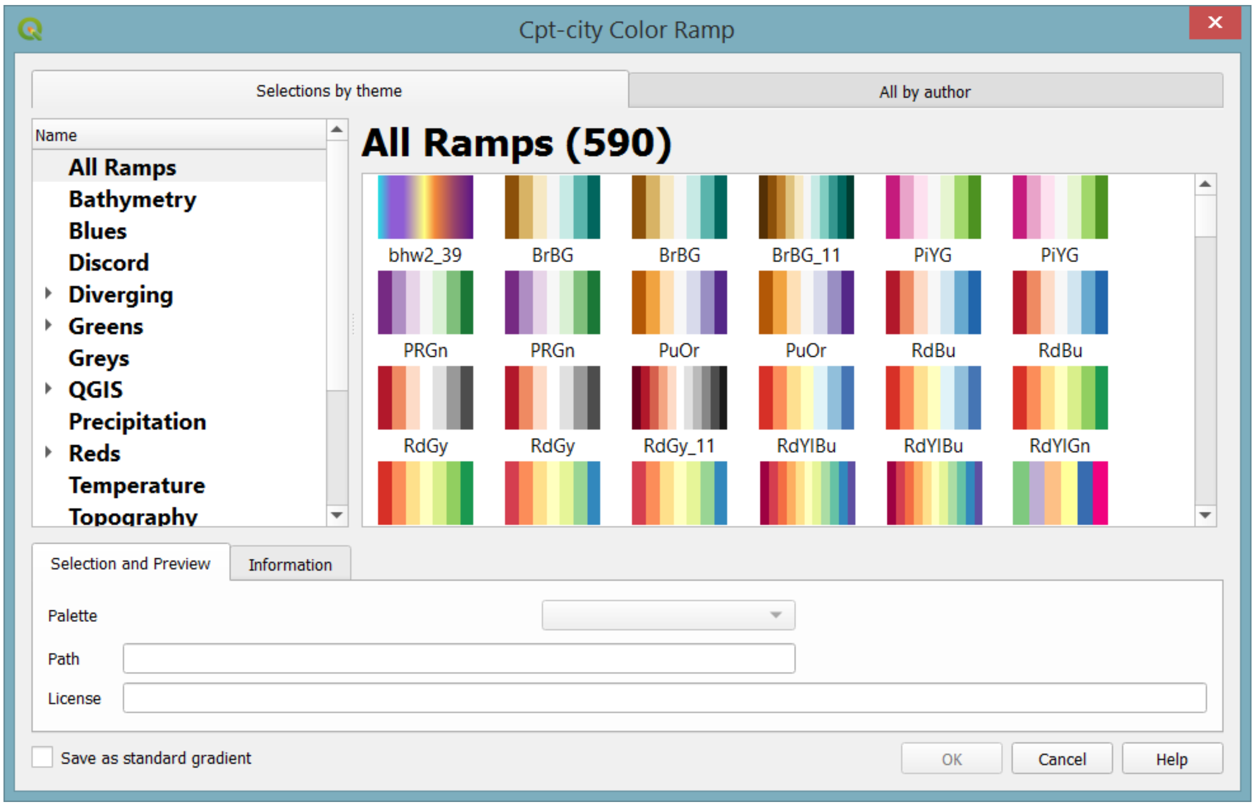
If we then rescale our color map, we should have a nice map that excludes the oceans. We can look for additional colormaps under ‘create new color ramp’:



Choose ‘cpt- city’:



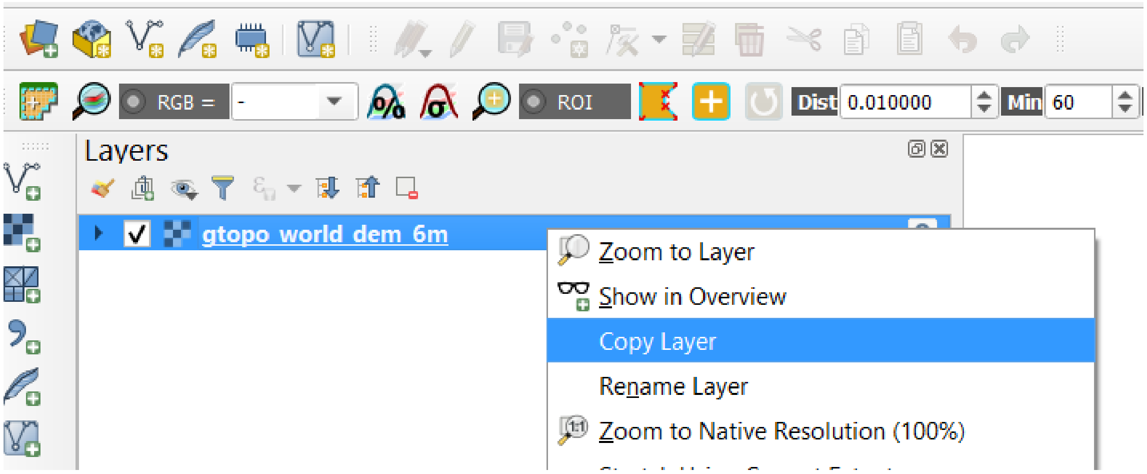
And you will find a list of possible additional color ramps:



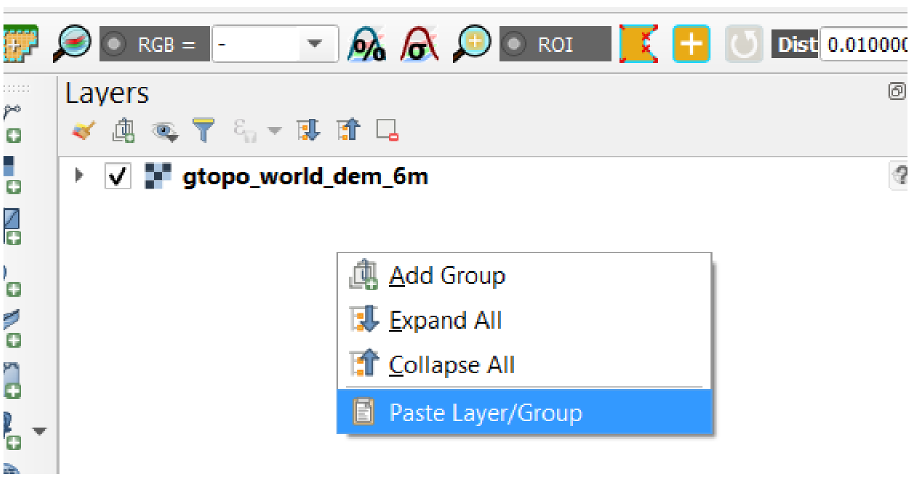
6. Adding a Simple Hillshade

Hillshades are a very common way to give 'depth' to topographic data, and allow you to symbolize terrain underneath whatever variable you want to include in your map. Fortunately, this is very quick and easy to do in QGIS.

First, make a copy of your DEM layer:

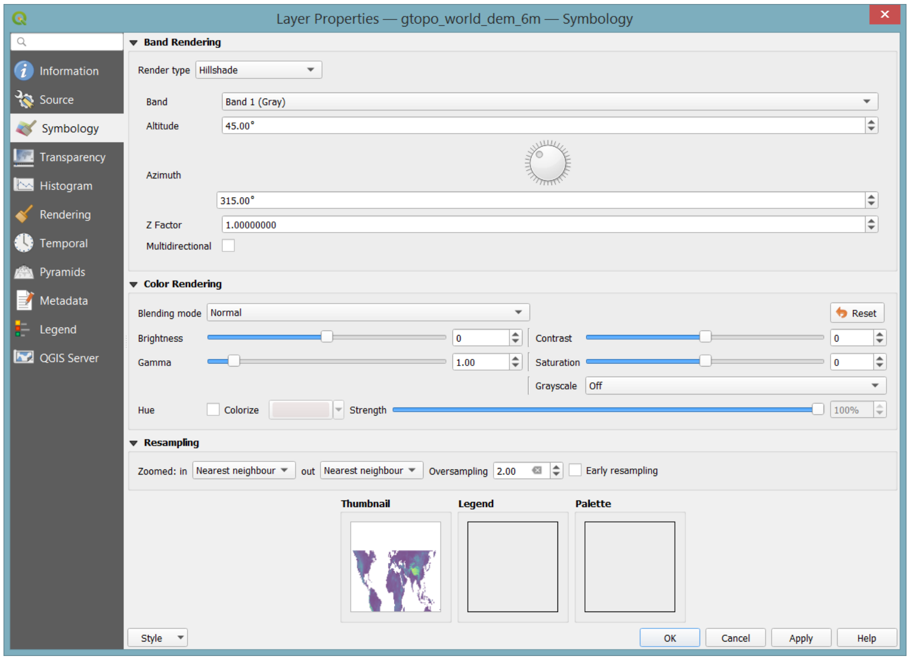


Next, paste that layer:



You should now have two copies of the DEM layer to work with.

Choose whichever one is on the bottom, and change the symbology to ‘Hillshade’:



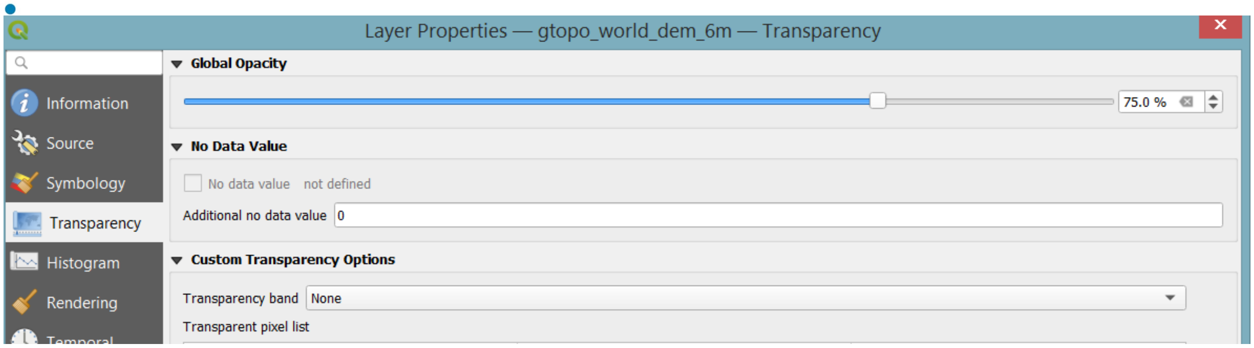
This panel lets you control the relative sun azimuth and altitude (e.g., which part of the landscape is ‘illuminated’ and which part is ‘dark’). You can also set a ‘z-factor’, which is useful for controlling how much you want to exaggerate topography.

In our case, the z-factor also stands in as a translation from *horizontal* to *vertical* resolution – QGIS assumes that pixels are scaled in the same way in both width and height. For our data, which lacks a projection, the vertical unit is *meters* but the horizontal unit is *decimal degrees*, meaning the hillshade will look very odd unless we set a z-factor.

Choosing a z-factor can be a bit tricky (see this guide for more details: https://www.esri.com/arcgis-blog/products/product/imagery/setting-the-z-factor-parameter-correctly/)

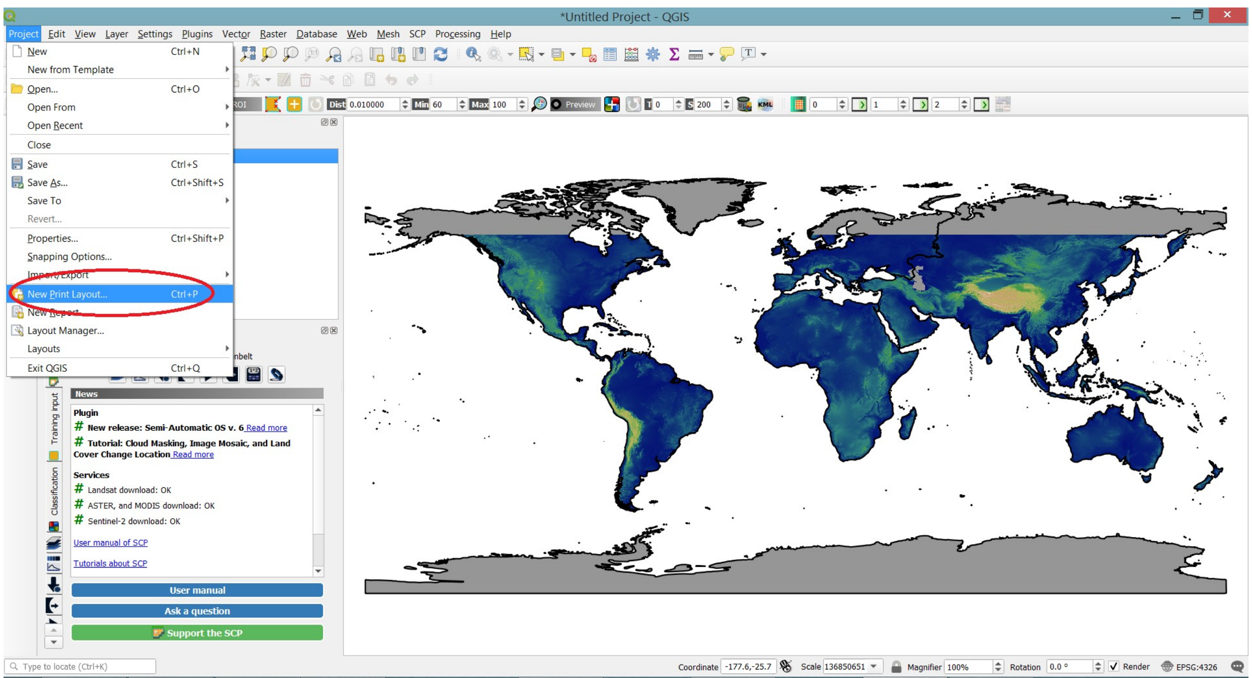
However, if it is simply for display purposes, you can be flexible. I chose a value of **0.0002**.

You can then change the transparency of your top layer and visualize both color and false-depth at once:

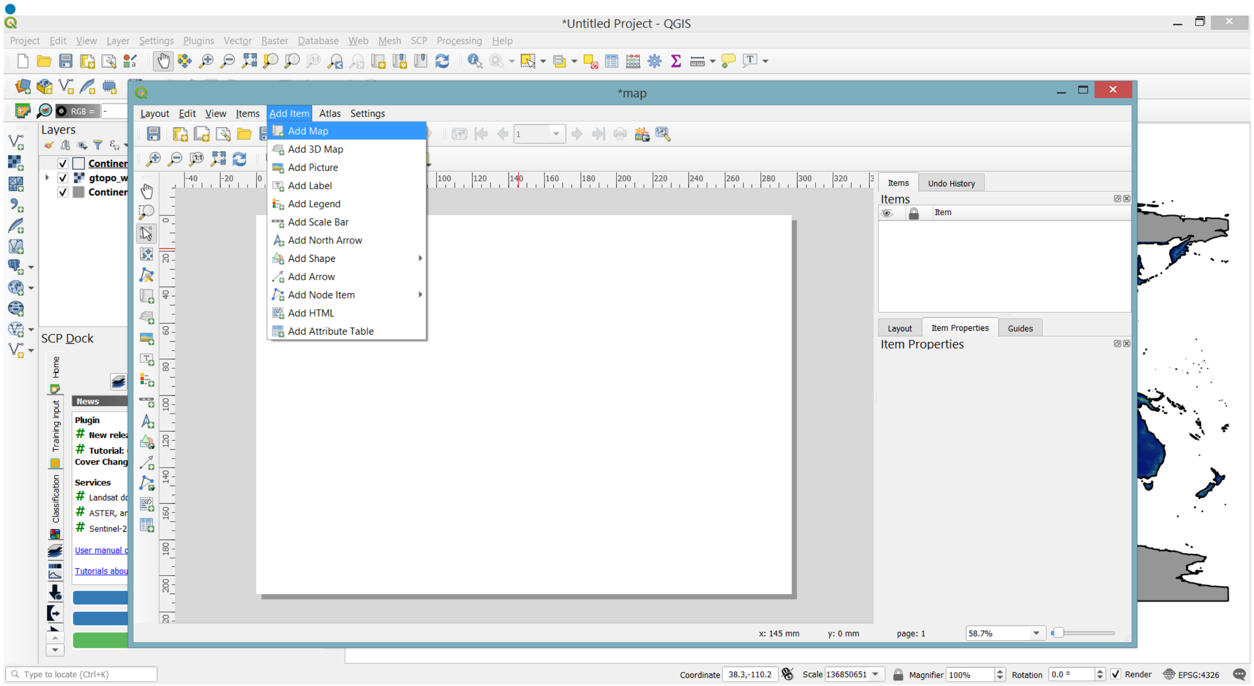


7. Making a Map for Export

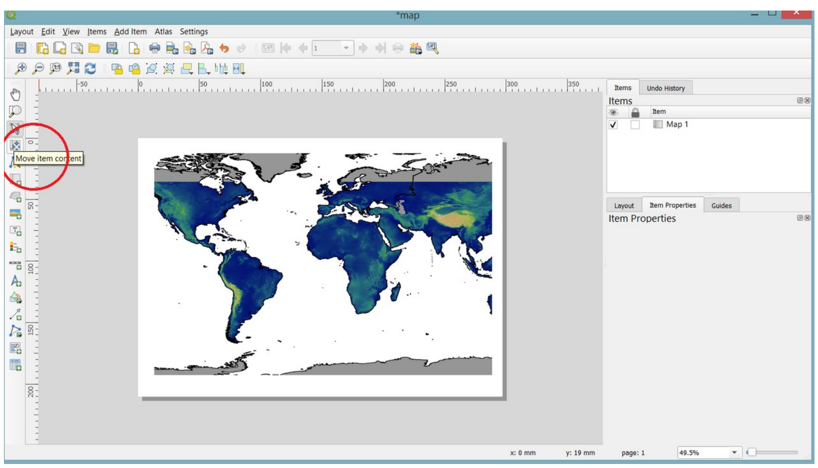
First, open a Print Composer using the 'Project' menu:



Switch the Print Composer window. Go to Layout ‣ Add Map. Use your cursor to draw a rectangle in the print composer. You will see that the rectangle window will be rendered with the map from the main QGIS canvas. The rendered map may not cover the full extent of our interest area.

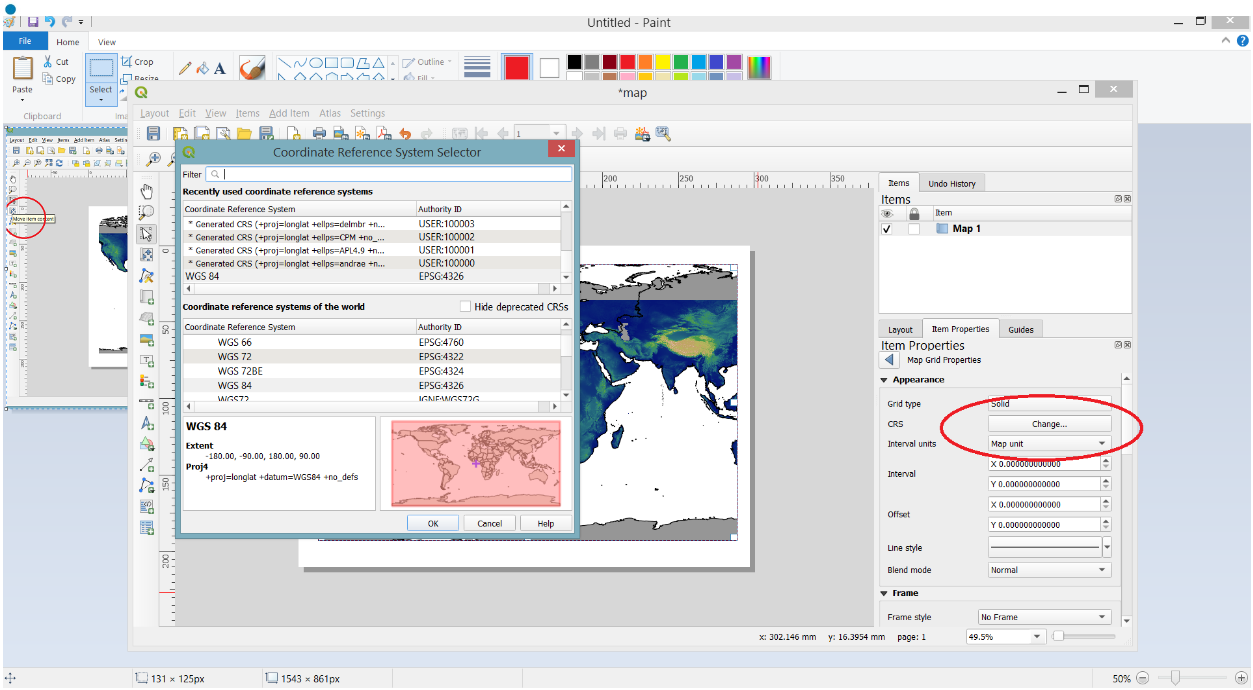


You can use the ‘Select/Move Item’ and ‘Move Item Content’ to fit the map to display only the regions with data.

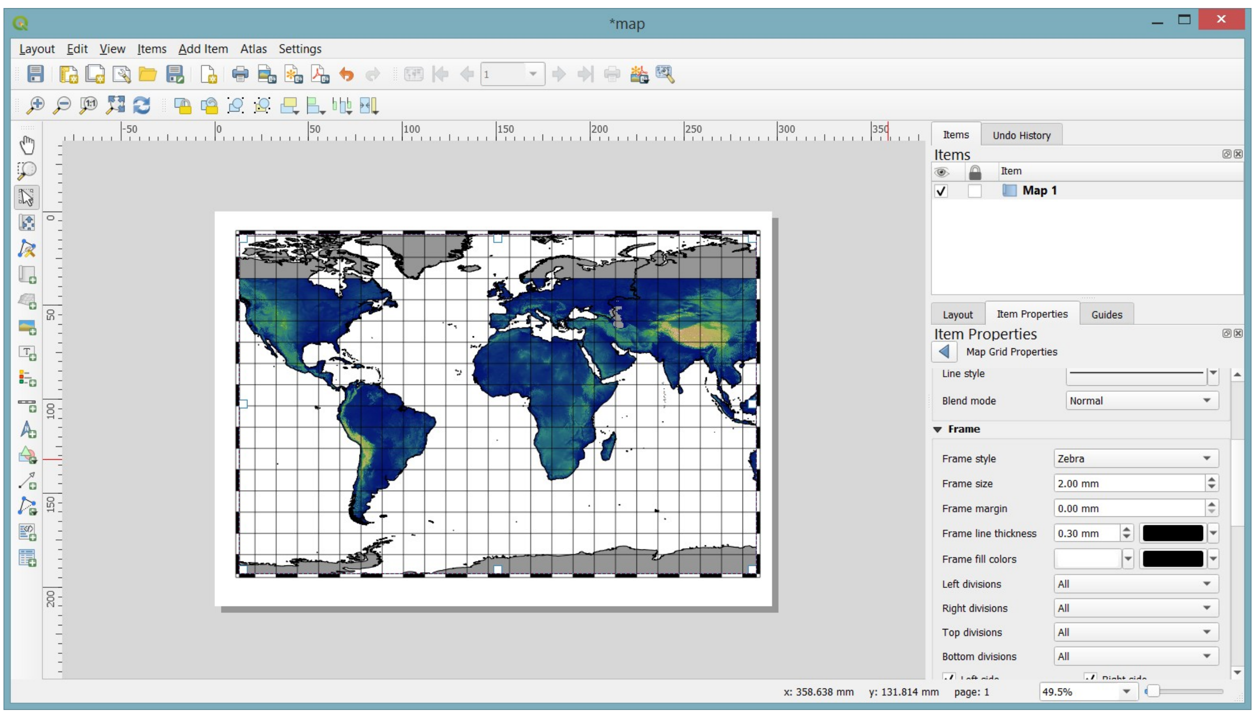


Once you’ve adjusted the project to display only the data you desire, we will add a grid and zebra border to the main map. Select the Map 0 object from the Items panel. In the Item properties tab, scroll down to the Grids section. Click the Add a new grid button. By default, the grid lines use the same units and projections as the currently selected map projections. However, it is more common and useful to display grid lines in degrees. We can select a different CRS for the grid. Click on the “change...” button next to CRS.

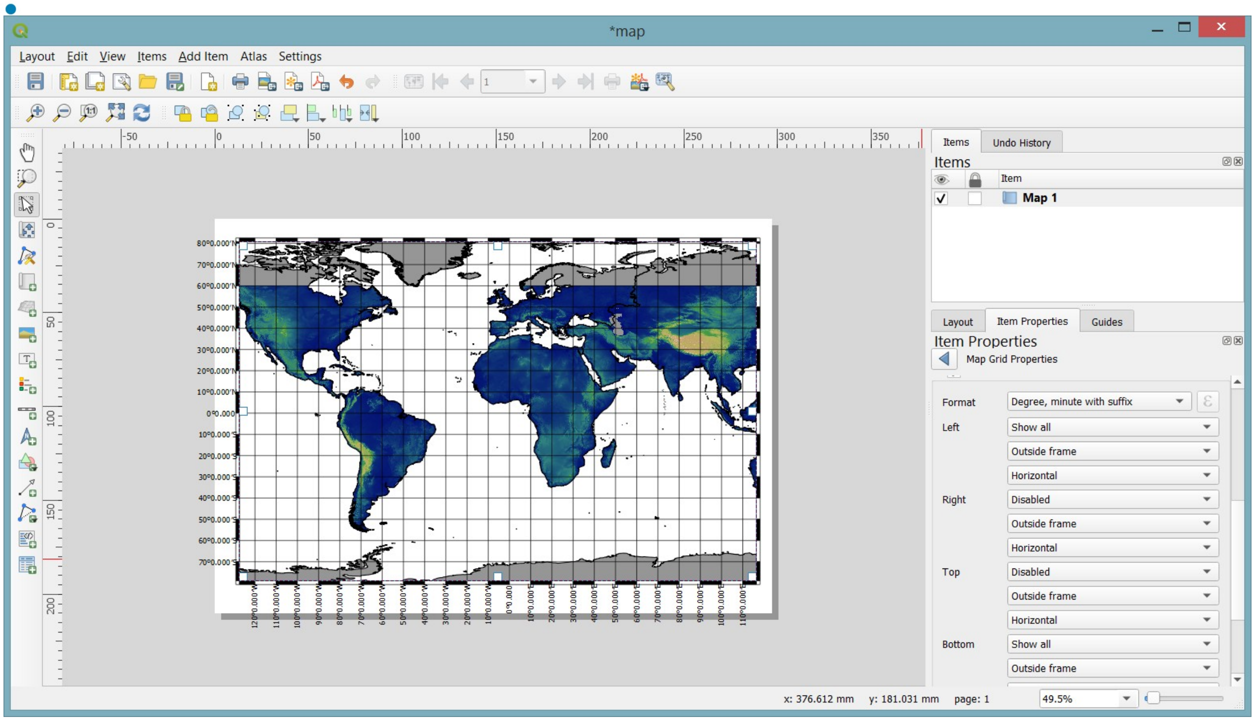
In the Coordinate Reference System Selector dialog, enter 4326 in the Filter box. From the results, select the WGS84 EPSG:4326 as the CRS. Click OK. Select the Interval values as 10 degrees in both X and Y direction. You can adjust the Offset to change where the grid lines appear.



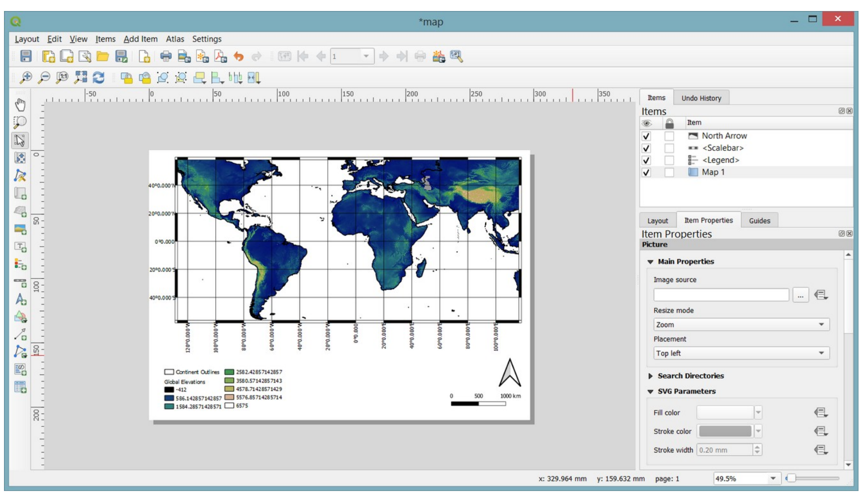
Scroll down to the Grid frame section and select the zebra frame.



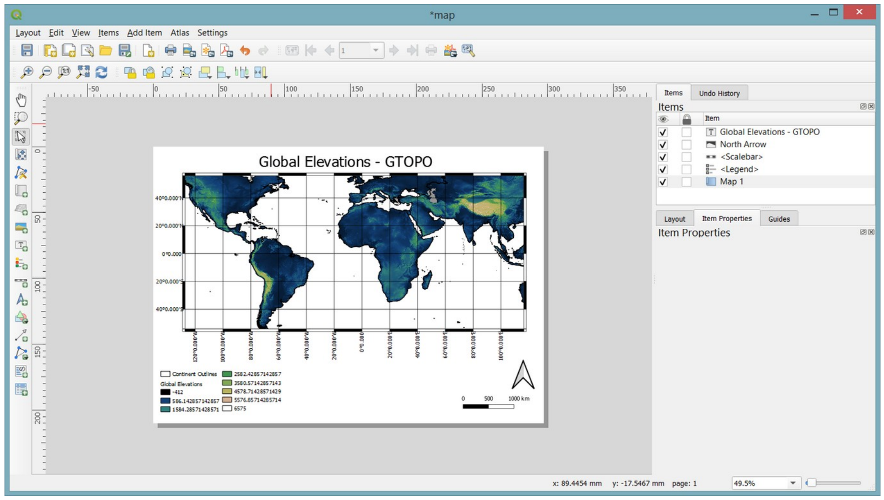
Next, click ‘Draw Coordinates’ further down to add labels.



To finish the map, include a Legend, Scale Bar, and North Arrow. Legends and scale bars are easily added with the ‘Add new legend’ and ‘Add new scalebar’ features. To insert a north arrow, you will need to use the ‘Add north arrow’ feature. Holding your left mouse button, draw a rectangle on the top-right corner of the map canvas.



Finish by adding a title to the map using the ‘Add new label’ feature.



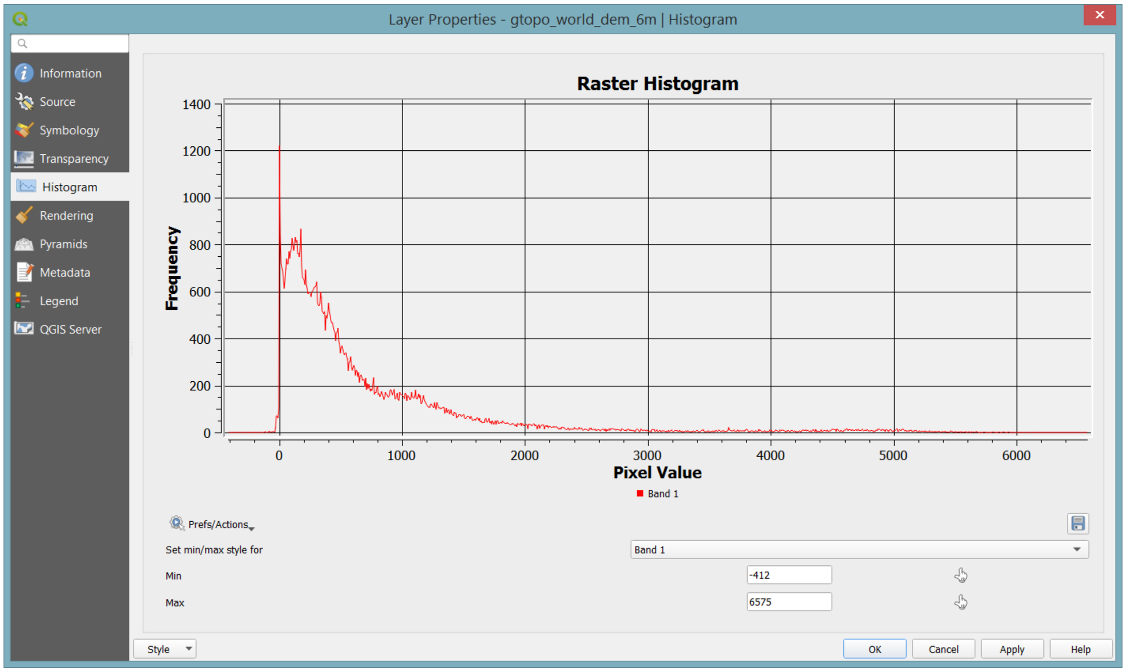
<u>Question 1:</u> Download your own raster data, explain where you got the data, and use it to create a simple map. The map should include a (1) scale bar, (2) color palette/legend, and (3) north arrow. There are many resources where you can download data and many datasets that might be interesting to you. Some good resources are the NASA EarthData (https://search.earthdata.nasa.gov/search) and USGS EarthExplorer ([https://earthexplorer.usgs.gov/)](https://earthexplorer.usgs.gov/) data stores. You will need to create an account, but registration is free. Feel free to download data for any region of interest to you. **Important – Do not simply turn in a map of the elevation data! It is important to find your own data to map for this assignment.** (10 points)

8. Basic GIS Tools and Analysis

Displaying data for visual analysis is only one key function of a desktop GIS. During this course, we will spend several lab sessions working within the QGIS toolbox. To familiarize ourselves with the interface, let's do some basic analysis on the elevation data we have.

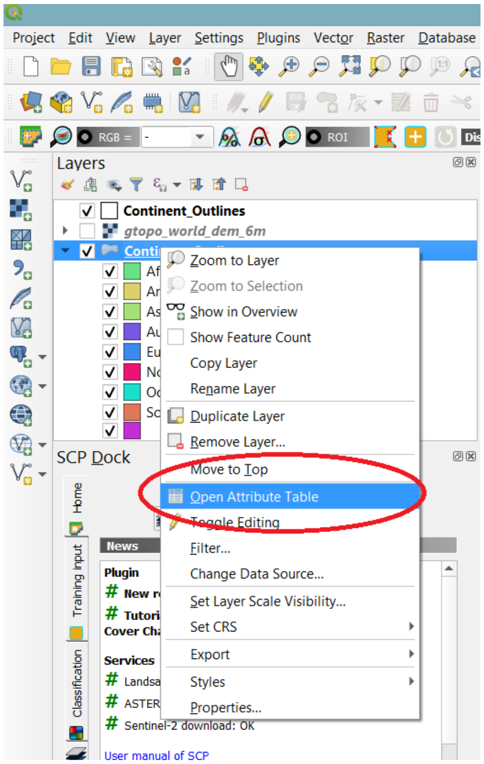
A **hypsometric curve** is an empirical cumulative distribution function of elevations in a given area. This can be small catchments, an entire drainage basin, a continent, or a rectangle of any area. A hypsometric curve shows the occurrences (or probability distribution) of topographic elevation.

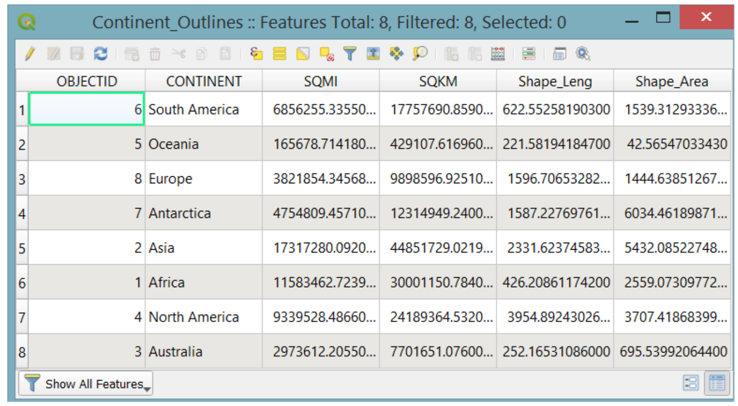
Comparing the elevation distributions of different areas is a simple way to compare the topography of different regions. We can create a hypsometric curve for the whole globe using QGIS:



Notice the large spike of values at zero, which is actually the coastlines. QGIS's simple histogram function still takes account of the NoData values we set in 'Transparency'. We can test this by turning off the NoData value and running the histogram again.

Let's take a second look at our 'continents' shapefile. There are several fields in the dataset that we can symbolize.

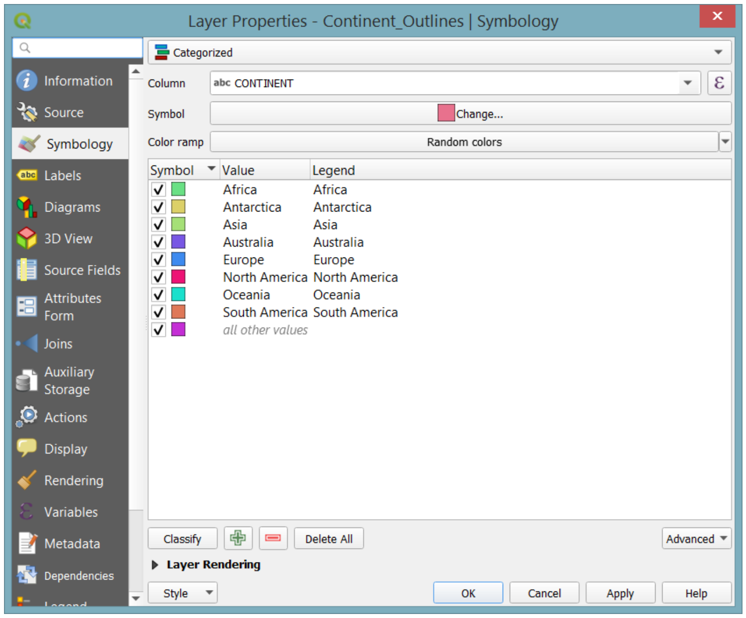



Continent_Outlines :: Features Total: 8, Filtered: 8, Selected: 0

	OBJECTID	CONTINENT	SQMI	SQKM	Shape_Leng	Shape_Area
1	6	South America	6856255.33550...	17757690.8590...	622.55258190300	1539.31293336...
2	5	Oceania	165678.714180...	429107.616960...	221.58194184700	42.56547033430
3	8	Europe	3821854.34568...	9898596.92510...	1596.70653282...	1444.63851267...
4	7	Antarctica	4754809.45710...	12314949.2400...	1587.22769761...	6034.46189871...
5	2	Asia	17317280.0920...	44851729.0219...	2331.62374583...	5432.08522748...
6	1	Africa	11583462.7239...	30001150.7840...	426.20861174200	2559.07309772...
7	4	North America	9339528.48660...	24189364.5320...	3954.89243026...	3707.41868399...
8	3	Australia	2973612.20550...	7701651.07600...	252.16531086000	695.53992064400

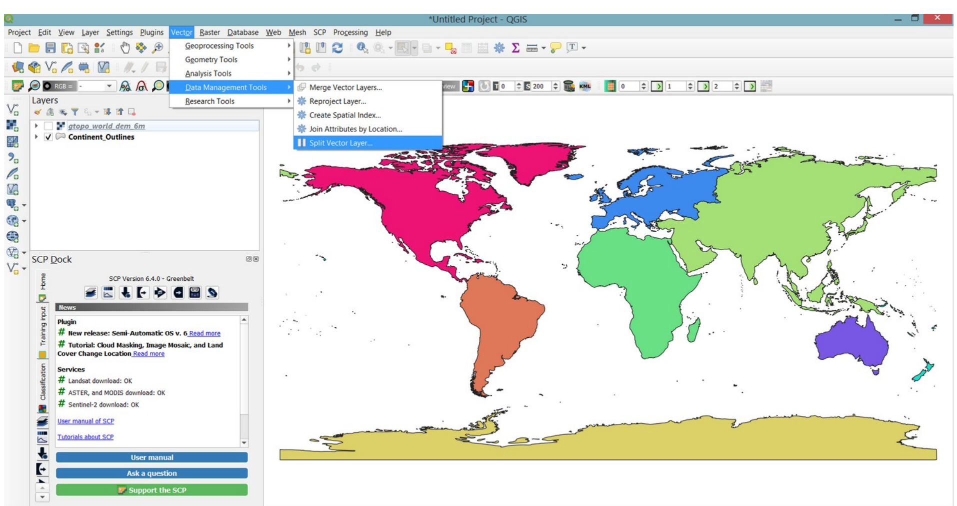
Show All Features

If we open this, we get a table describing each of the shapes in the layer. Notice that each continent is given a label, length, and area. We can symbolize each continent with a different color based on the names by changing the symbology to ‘categorized’:



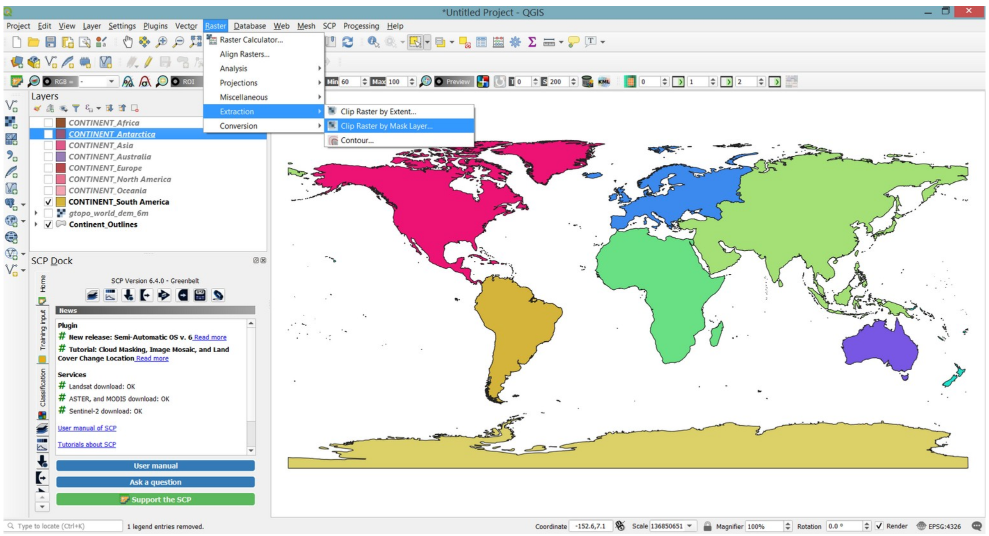
While the global hypsometry is interesting, a comparison of the histograms for each continent will be much more useful. To do this, we need to do two mains steps: (1) Select individual continent shapefiles, and (2) clip the elevation data to the extent of those continents. We can do all of this from within QGIS.

First, let’s clip out one of the continents. QGIS includes an easy split function for just such a case:

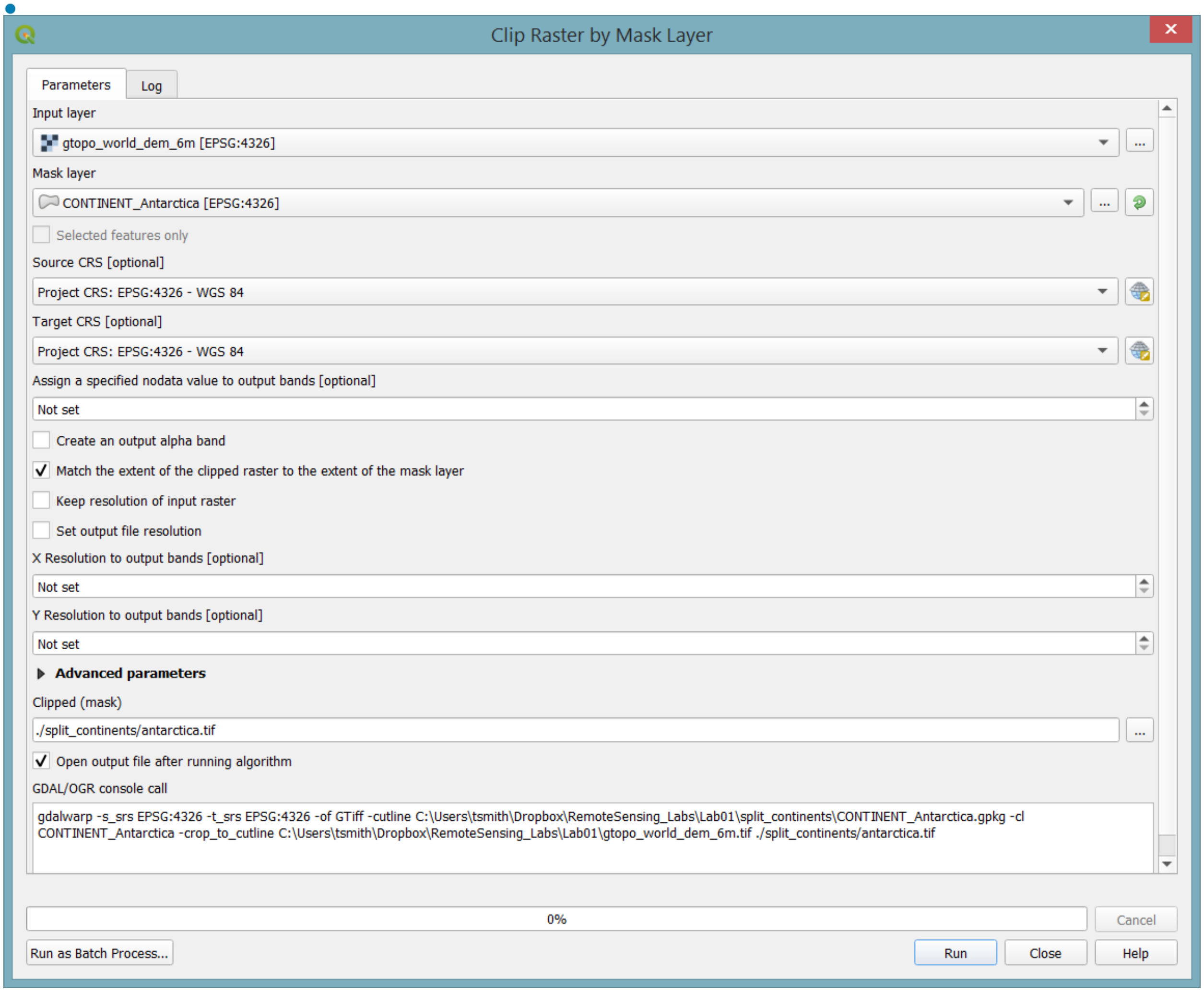


We can define a save location, and it exports several datasets in the gpkg format, which maintains color and other information as well.

Now that we have each of the continents as their own single file, we can go about clipping our elevation dataset one continent at a time. Load in one (or all) of the vectors and use the ‘Extract by Mask Layer’ tool:

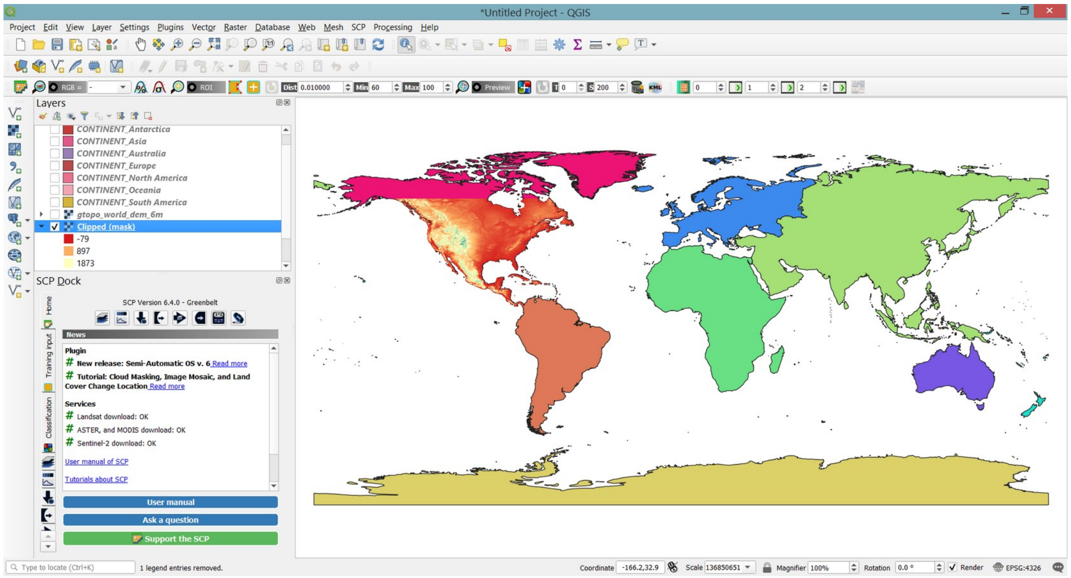


Note that we will change some settings, as seen here:

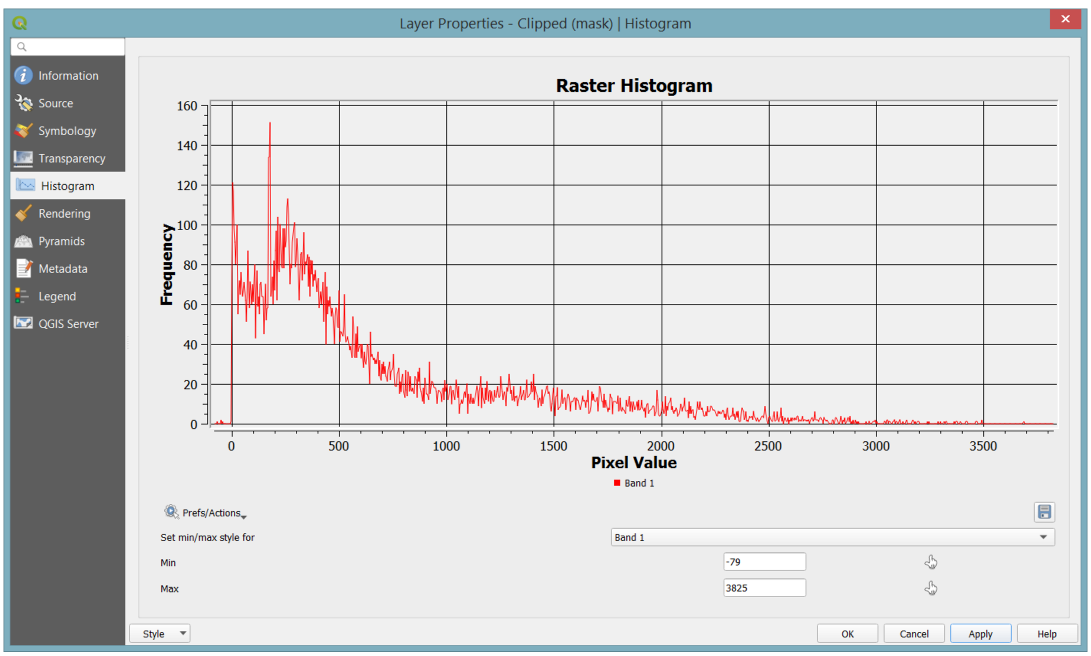


On the bottom you will see the GDAL command that QGIS is actually running, which we will return to later in this course. If you are running tools on the command line, you can also copy and paste that line of code to your terminal and create clipped files that way.

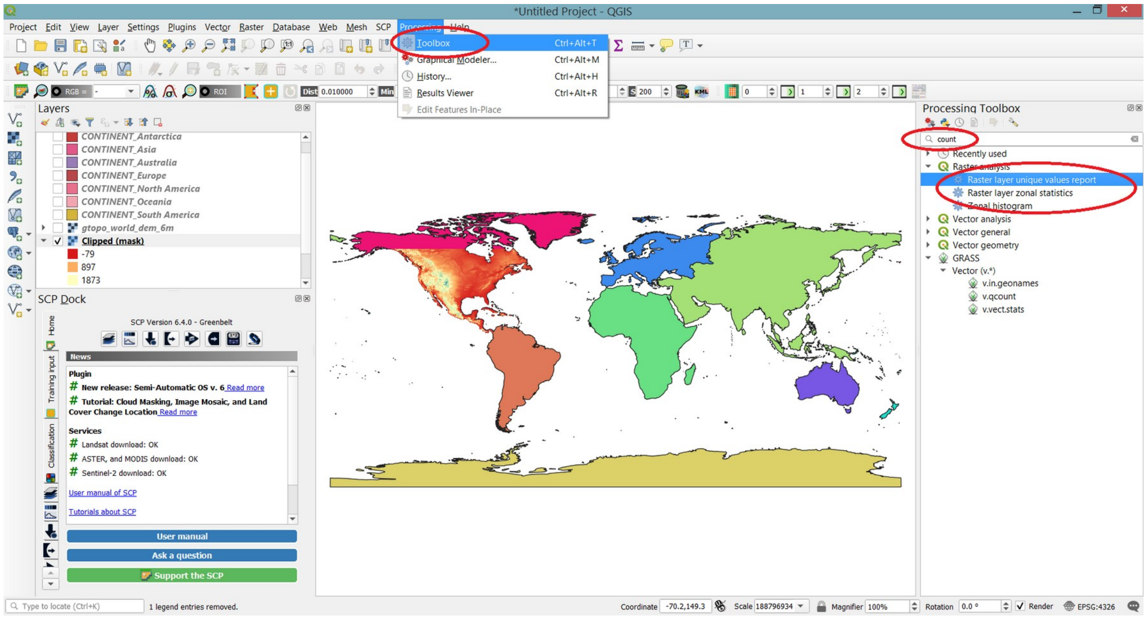
The result should be something like this:



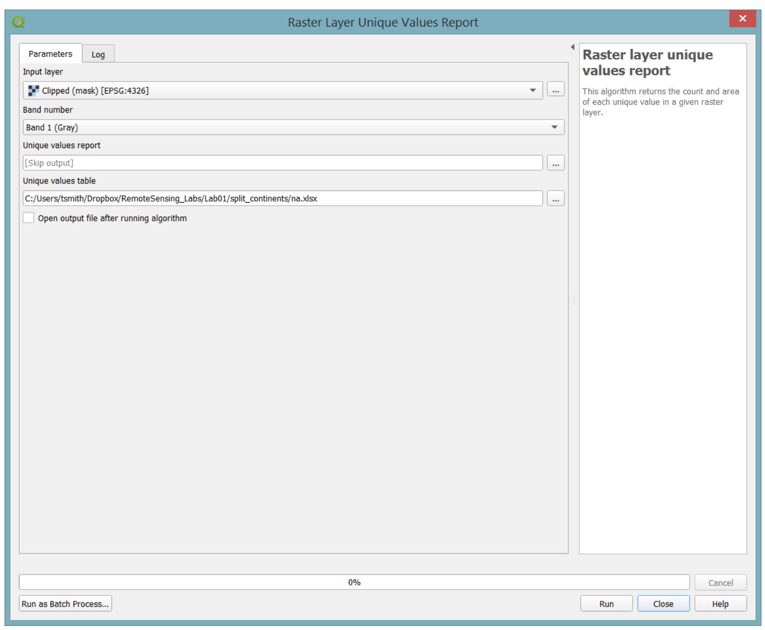
With a new histogram:



You can also export the histogram as raw data with the 'Raster Unique Values Report' tool:



This allows you to save the number of pixels with each value to a text file or excel spreadsheet.



<u>Question 2:</u> Using the tools found in this lab, create a single image showing the histogram of elevations for each continent. You are welcome to create the histogram in whatever program you choose (e.g., python, Excel, R, matlab, etc). <u>Hint: The easiest way to do this is likely to export each set of</u>

counts to a separate table and then combine them with another tool. See below for a possible acceptable output (10 points).

Sample output via Python/matplotlib: