Module: Classifying Vector Data

Classifying vector data allows you to assign different symbols to features (different objects in the same layer), depending on their attributes. This allows someone who uses the map to easily see the attributes of various features.

4.1 Lesson: Attribute Data

Up to now, none of the changes we have made to the map have been influenced by the objects that are being shown. In other words, all the farms look alike, and all the roads look alike. When looking at the map, the viewers don't know anything about the roads they are seeing; only that there is a road of a certain shape in a certain area.

But the whole strength of GIS is that all the objects that are visible on the map, also have attributes. Maps in a GIS aren't just pictures. They represent not only objects in locations, but also information about those objects.

The goal of this lesson: To explore the attribute data of an object and understand what the various data can be useful for.

4.1.1 ![tag icon] Follow along: Attribute data

Open the attribute table for the *places* layer (refer back to the section *"Working with Vector Data"* if necessary). Which field would be the most useful to represent in label form, and why?

Check your results

4.1.2 In conclusion

You now know how to use the attribute table to see what is actually in the data you're using. Any dataset will only be useful to you if it has the attributes that you care about. If you know which attributes you need, you can quickly decide if you're able to use a given dataset, or if you need to look for another one that has the required attribute data.

4.1.3 What's next?

Different attributes are useful for different purposes. Some of them can be represented directly as text for the map user to see. You'll learn how to do this in the next lesson.

4.2 Lesson: The Label Tool

Labels can be added to a map to show any information about an object. Any vector layer can have labels associated with it. These labels rely on the attribute data of a layer for their content.

Note: The *Layer Properties* dialog does have a *Labels* tab, but for this example we will not be using it. It has too few options and may be difficult to work with, which is why it will probably be taken out of new versions of QGIS (after 1.8). Instead, for this lesson we'll use the separate *Label tool*, which has more options and is more user-friendly.

The goal for this lesson: To apply useful and good-looking labels to a layer.

4.2.1 Follow along: Using labels

Before being able to access the Label tool, you will need to ensure that it has been activated.

- Go to the menu item *View → Toolbars*.

- Ensure that the *Label* item has a check mark next to it. If it doesn't, click on the *Label* item, and it will be activated.

- Click on the *places* layer in the *Layers list*, so that it is highlighted.

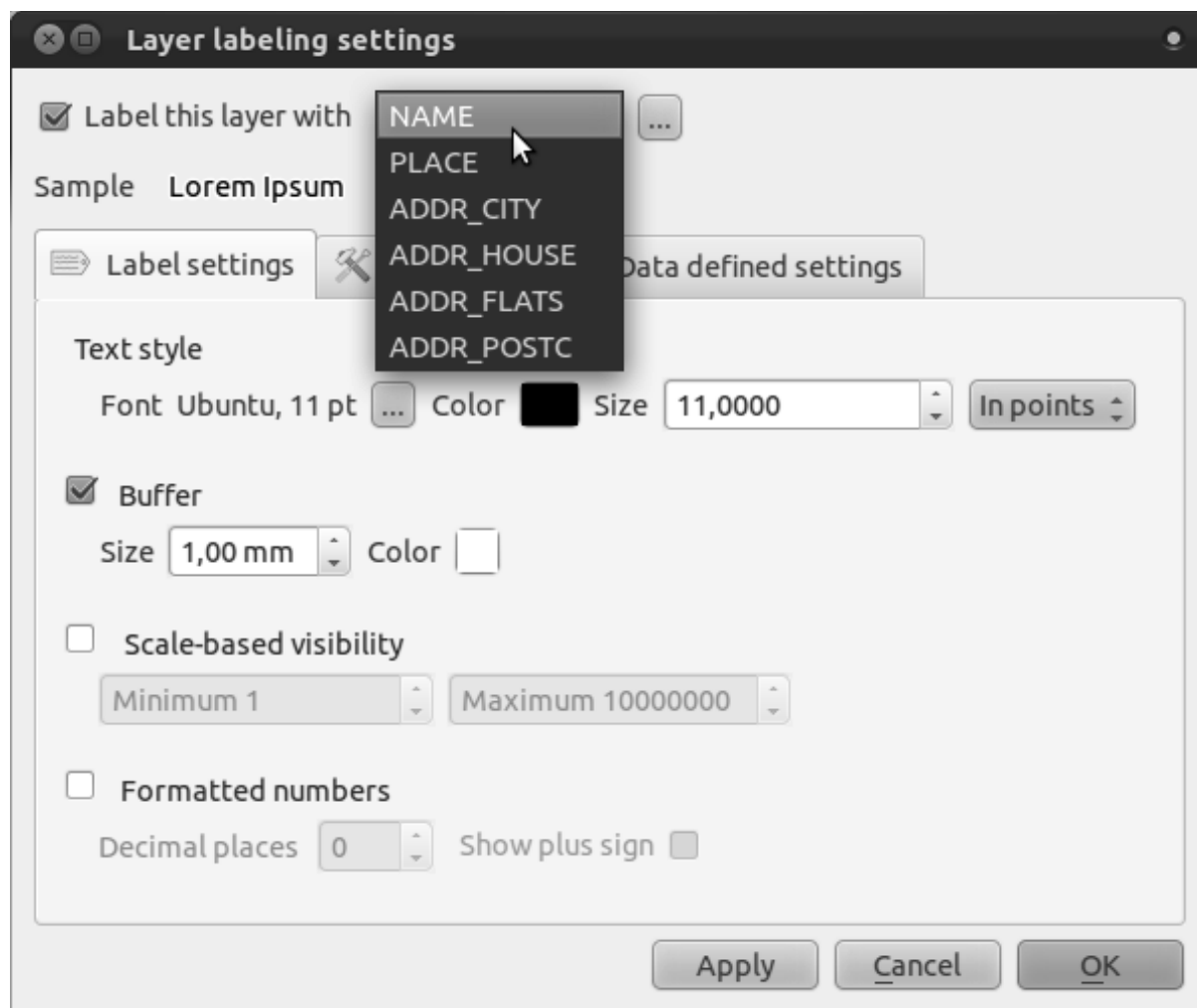- Click on the following toolbar button:

This gives you the *Layer labeling settings* dialog.

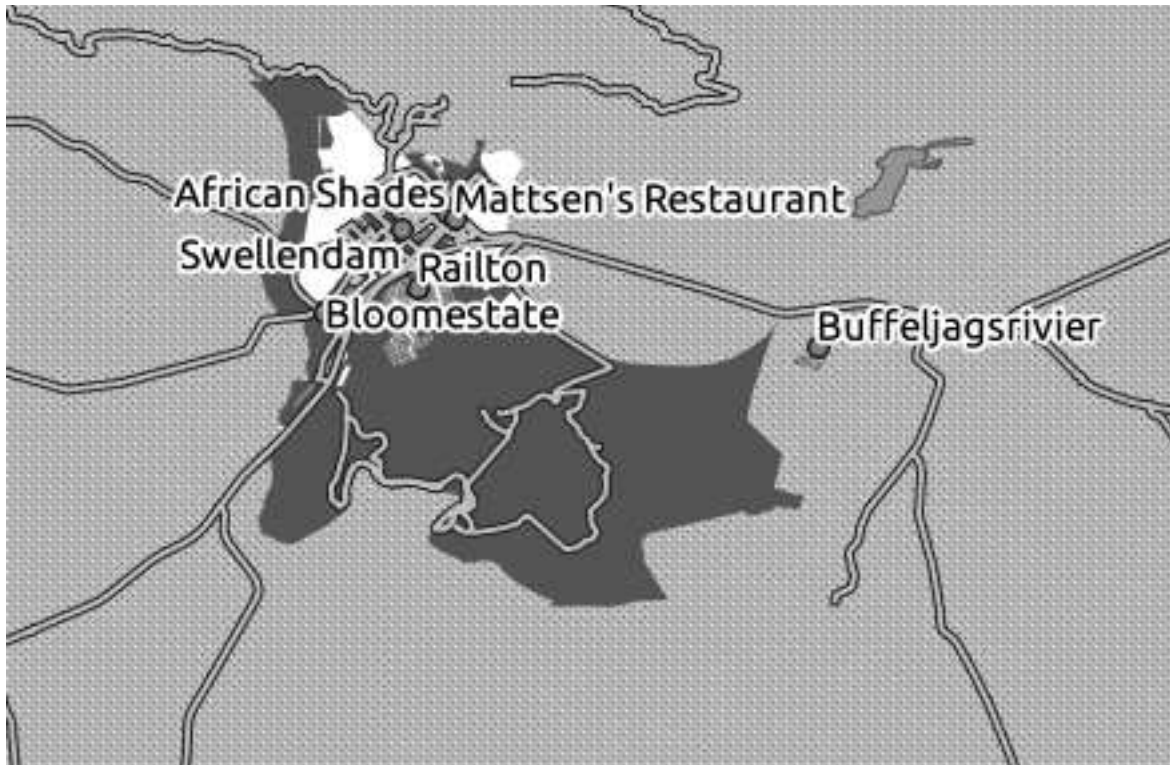- Check the box next to *Label this layer with....*
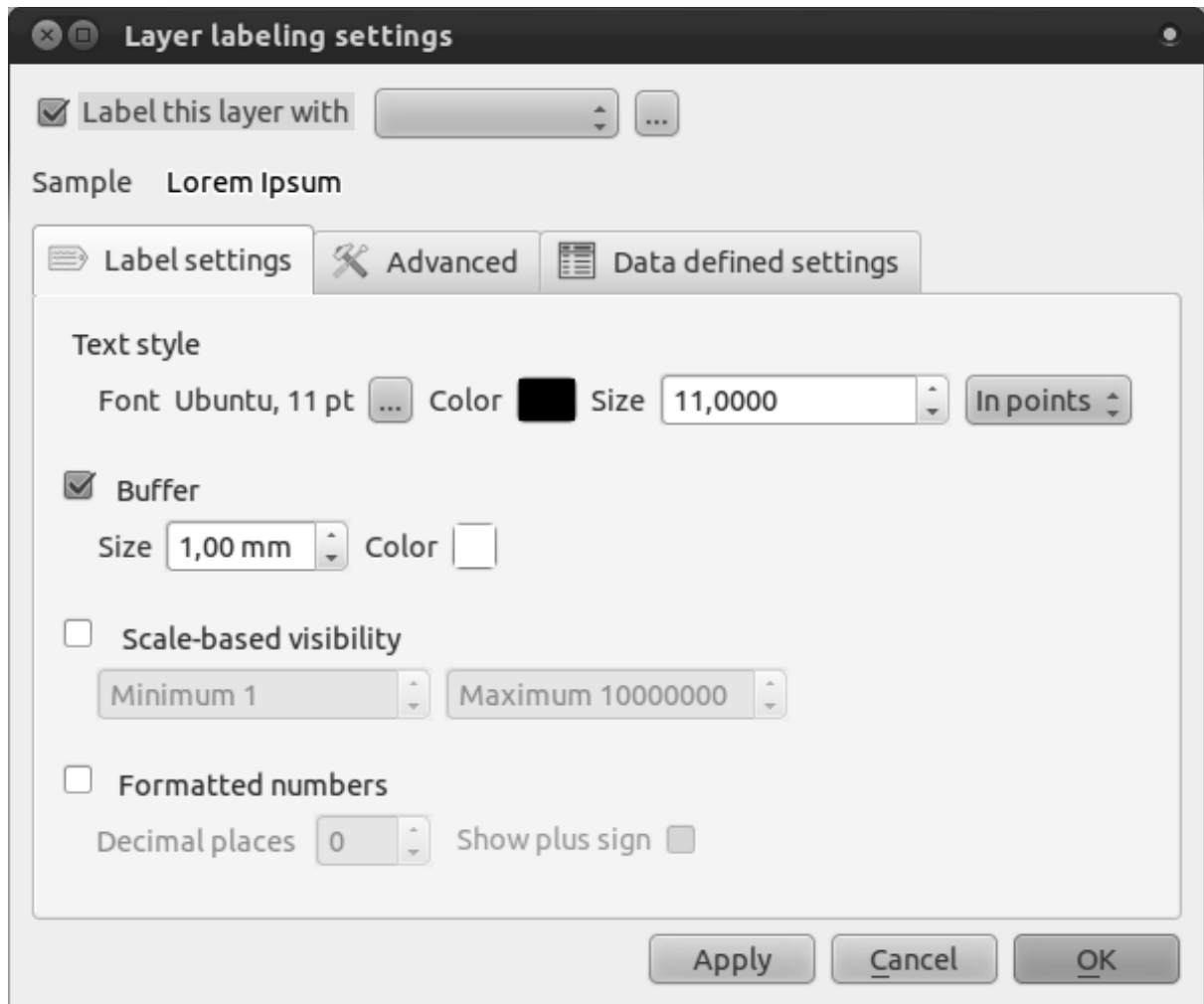
You'll need to choose which field in the attributes will be used for the labels. In the previous lesson, you decided that the NAME field was the most suitable one for this purpose.

- Select *NAME* from the list:

- Click *OK*.

The map should now have labels like this:

4.2.2 🏷 Follow along: Changing label options

What we have so far is good, but as you can see, the labels are overlapping the points that they are associated with. That doesn't look very nice. The text is also a bit larger than it needs to be. Let's fix these problems!

- Open the *Label tool* again by clicking on its button as before.

- Click on the ellipsis (...) button to change the text properties:
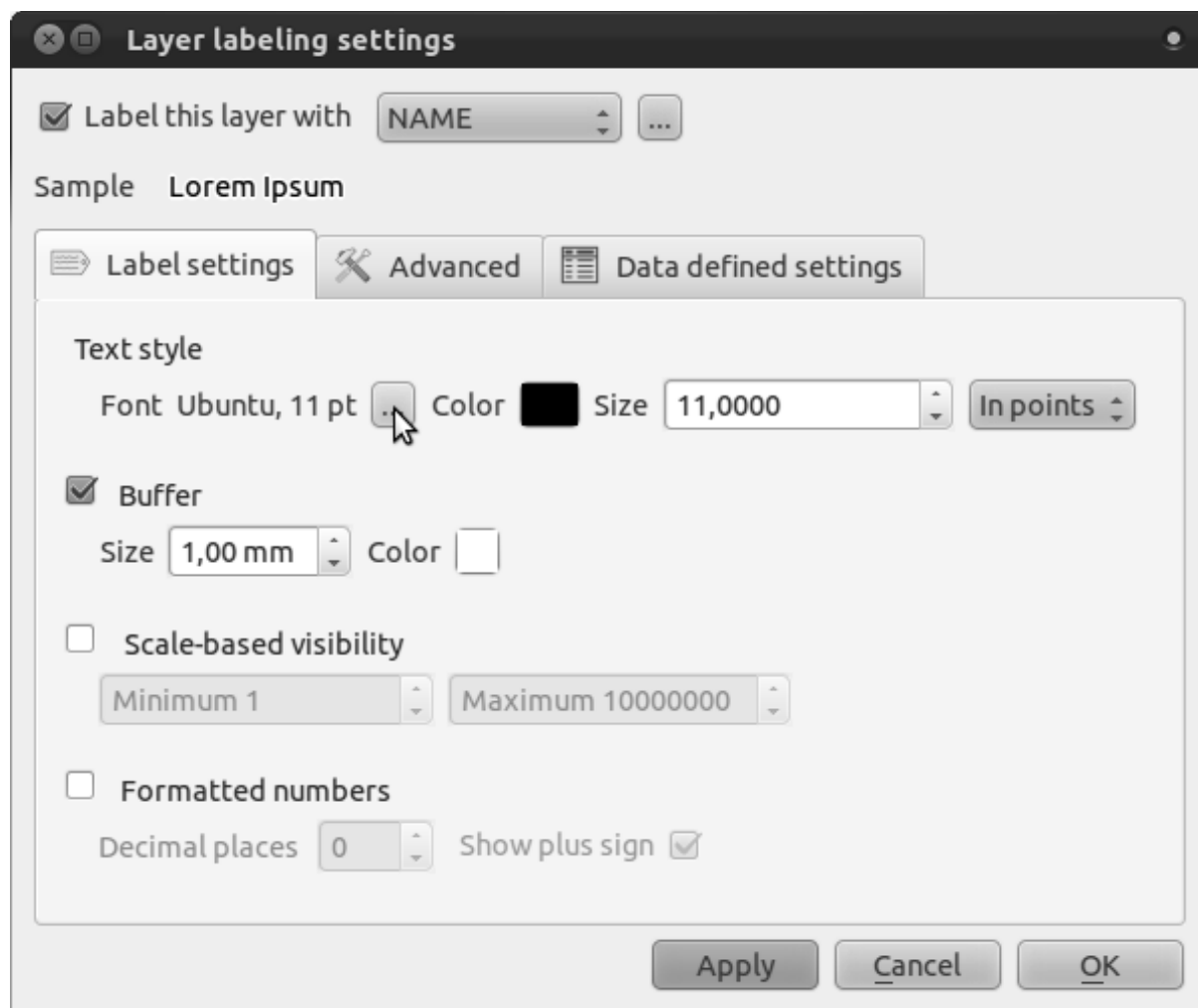
A standard text change dialog appears, similar to those in many other programs. Change the font to `Arial` size `9`. Your labels will now look like this:

That's the font problem solved! Now let's look at the problem of the labels overlapping the points, but before we do that, let's take a look at the *Buffer* option.

- Open the *Label tool* dialog.

- Deactivate the label buffer by clicking on the checkbox next to the text that says *Buffer*.

- Click *Apply*.

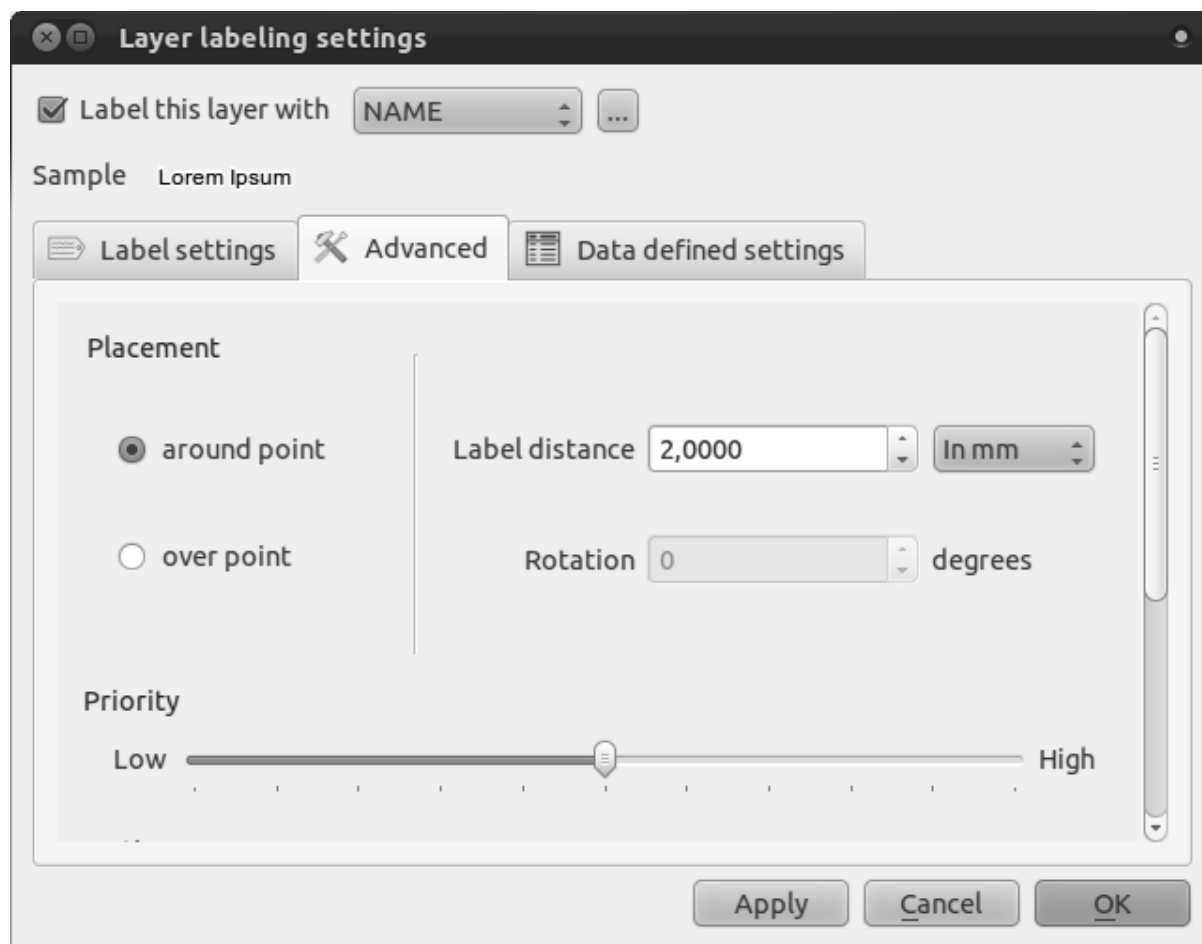Note the effects in the map:

Now you can see why we usually need label buffers!

- Reactivate the buffers by clicking in the same checkbox as before, and then clicking *Apply*.

Back to the problem of the labels that overlap points.

- In the *Label tool* dialog, go to the *Advanced* tab.

- Change the value of *Label distance* to 2.

- Click *Apply*.

This is the result:

4.2.3 ⬛ Follow along: Using labels instead of layer symbology

In many cases, the location of a point doesn't need to be very specific. For example, most of the points in the *places* layer refer to entire towns or suburbs, and the specific point associated with such features is not that specific on a large scale. In fact, giving a point that is too specific is often confusing for someone reading a map.

To name an example: on a map of the world, the point given for the European Union may be somewhere in Poland, for instance. To someone reading the map, seeing a point labeled *European Union* in Poland, it may seem that the capital of the European Union is therefore in Poland.

So, to prevent this kind of misunderstanding, it's often useful to deactivate the point symbols and replace them completely with labels.
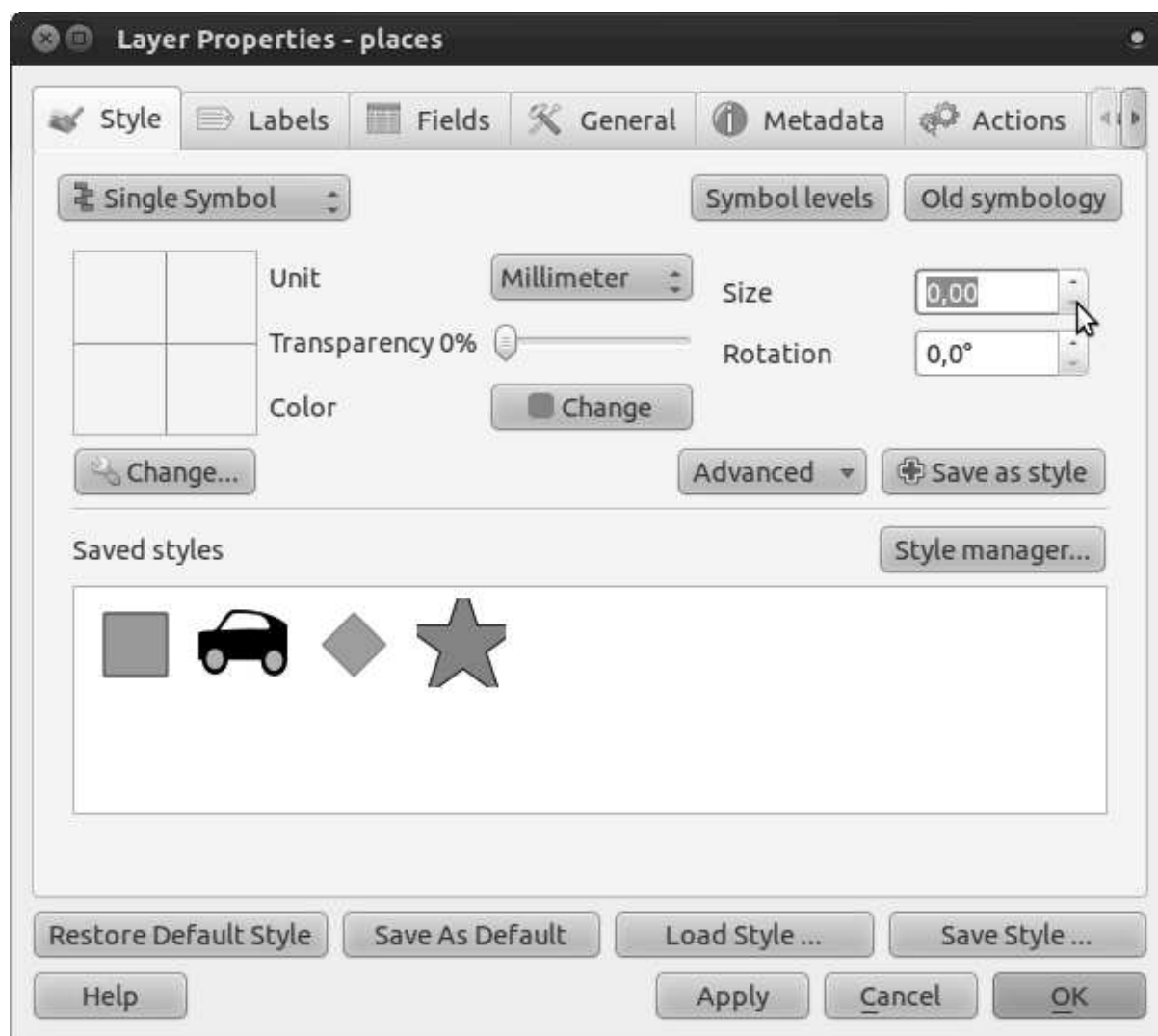
In QGIS, you can do this by changing the position of the labels to be rendered directly over the points they refer to.

- Open the *Layer labeling settings* dialog for the *places* layer.

- Go to the *Advanced* tab.

- Click on the *over point* button.
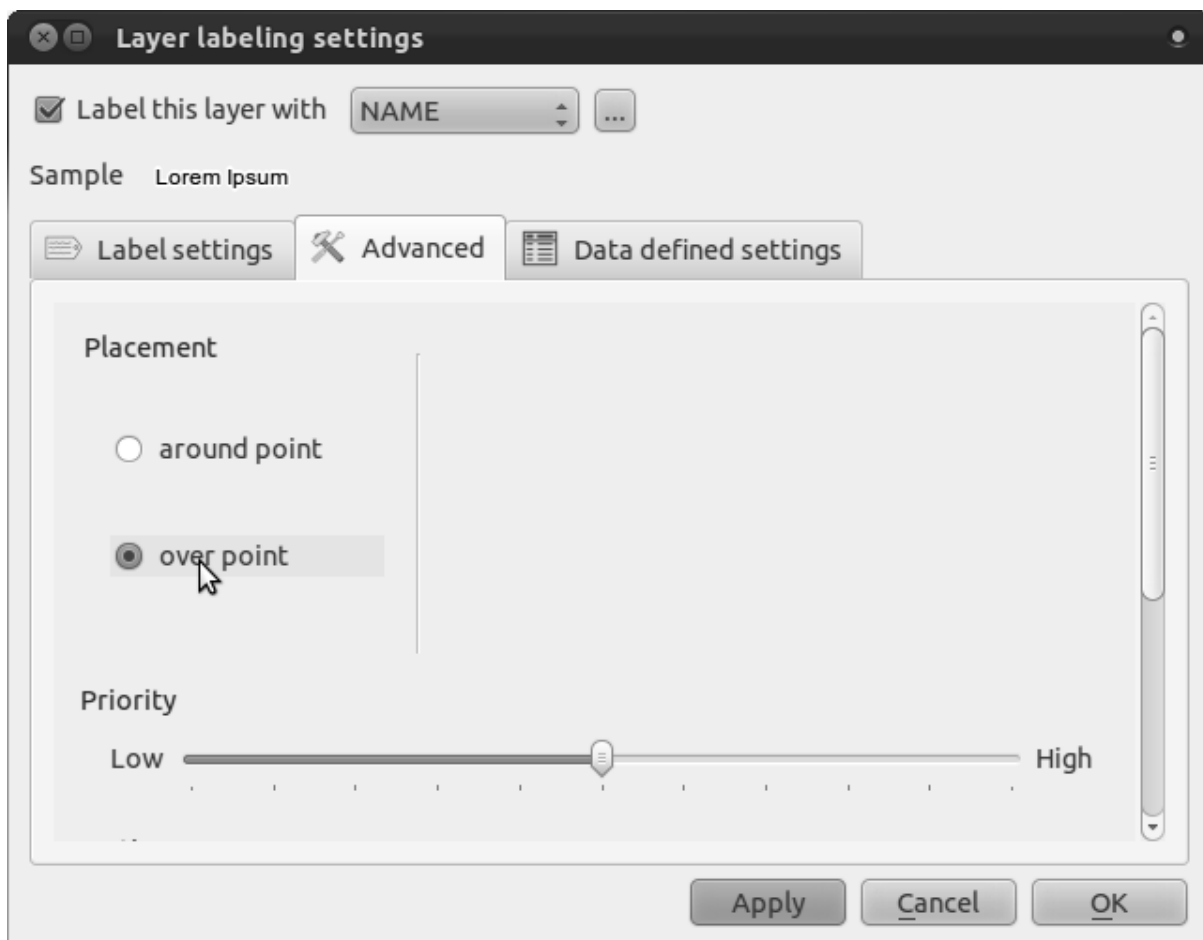
- Hide the point symbols by editing the layer style as usual, and setting the size of points

---

to 0:

• Click *OK* and you'll see this result:

As you can see, some of the labels are now missing to prevent overlap (at this scale). Sometimes this is what you want when dealing with datasets that have many points, but at other times you will lose useful information this way. There is another possibility for handling cases like this, which we'll cover in a later exercise in this lesson.

4.2.4 Try yourself... Customize the labels

- Return the label and symbol settings to the way they were before.

Check your results

- Set the map to the scale `1:50000`. You can do this by typing it into the *Scale* box in the *Status Bar*.

- Modify your labels to be suitable for viewing at this scale.

Check your results

4.2.5 Follow along: Labeling lines

Now that you know how labeling works, there's an additional problem. Points and polygons are easy to label, but what about lines? If you label them the same way as the points, your results would look like this:

This is not very useful! To make lines behave, we'll need to edit some options.

- Hide the *places* layer so that it doesn't distract you.

- Activate labels for the *streets* layer as before. (Remember to use the *Label tool* on the toolbar, not the one in *Label Properties*!)

- Set the font *Size* to 9 so that you can see more labels.

- Zoom in on the Swellendam town area.

- In the *Label tool* dialog's *Advanced* tab, choose the following settings:
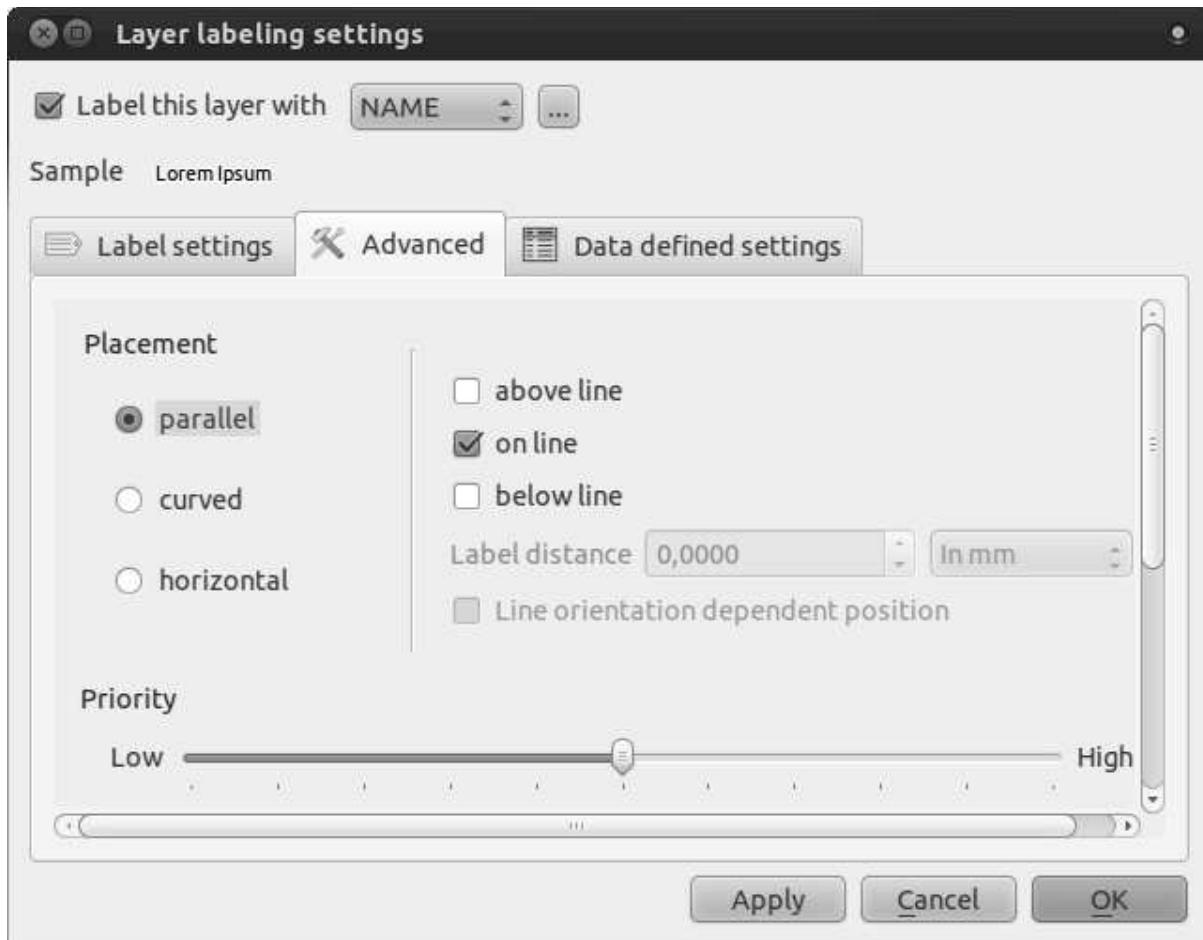
The map will look somewhat like this, depending on scale:

It's better than before, but still not ideal. For starters, some of the names appear more than once, and that's not always necessary. To prevent that from happening:

- Enable the option *Merge connected lines to avoid duplicate labels* (also under the *Advanced* tab - you may need to scroll down to see it).

Another useful function is to prevent labels being drawn for features too short to be of notice.

- Set the value of *Suppress labeling of features smaller than ...* to 5mm and note the results when you click *Apply*.

Try out different *Placement* settings as well (also under the *Advanced* tab). As we've seen before, the *horizontal* option is not a good idea in this case, so let's try the *curved* option instead!

- Select the *curved* option under the *Advanced* tab of the *Layer labeling settings* dialog.

Here's the result:

As you can see, this hides a lot of the labels that were previously visible, because of the difficulty of making some of them follow twisting street lines and still be legible. You can decide which of these options to use, depending on what you think seems more useful or what looks better.

4.2.6 Follow along: Data defined settings

- Deactivate labeling for the *streets* layer.

- Reactivate labeling for the *places* layer.

- Open the attribute table for *places*.

It has two fields that are of interest to us now: ADDR_CITY and IN_SWD. ADDR_CITY is the city that the feature is in. IN_SWD is derived from it, and tells you whether or not that feature is in Swellendam (1 if it is, 0 if not). We can use this data to influence the label styles.

- Navigate to the *Data defined settings* tab.

- In the *Italic* dropdown, select IN_SWD and click *Apply*.

Notice its effects:

4.2.7   Try yourself... Using data defined settings

Note: We're jumping ahead a bit here to demonstrate some advanced labeling settings. At the advanced level, it's assumed that you'll know what the following means. If you don't, feel free to leave out this section and come back later when you've covered the requisite materials.

- Open the attribute table for *places*.

- Enter edit mode by clicking this button (at the lower edge of the attribute table):

- Add a new column:

- Configure it like this:
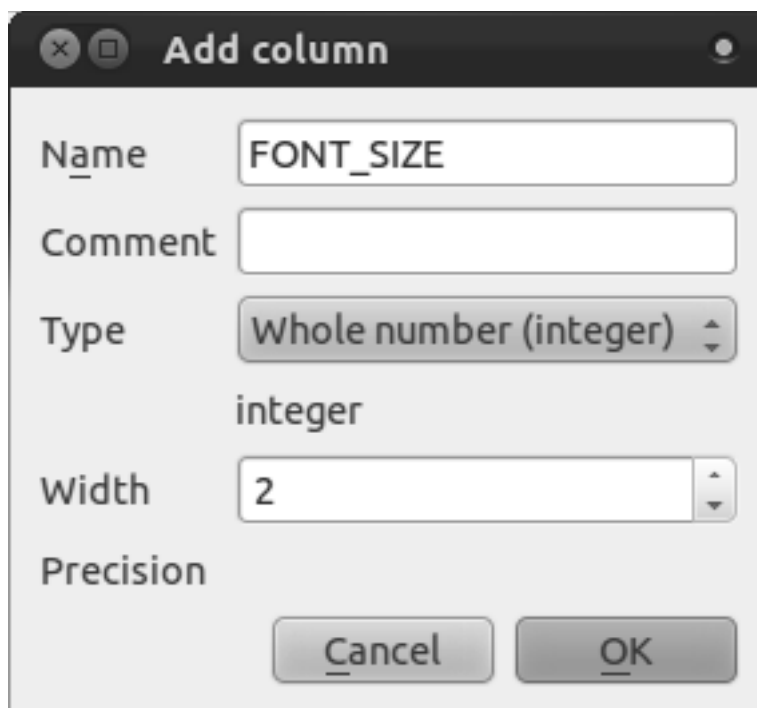
- Use this to set custom font sizes for each different type of place (i.e., each key in the `PLACE` field).

Check your results

4.2.8 Further possibilities with labeling

We can't cover every option in this course, but be aware that the *Label tool* has many other useful functions. You can set scale-based rendering, alter the rendering priority for labels in a layer, and set every label option using layer attributes. You can even set the rotation, XY position, and other properties of a label (if you have attribute fields allocated for the purpose), then edit these properties using the tools adjacent to the main *Label tool*:

(These tools will be active if the required attribute fields exist and you are in edit mode.)

Feel free to explore more possibilities of the labeling system.

4.2.9 In conclusion

You've learned how to use layer attributes to create dynamic labels. This can make your map a lot more informative and stylish!

4.2.10 What's next?

Now that you know how attributes can make a visual difference for your map, how about using them to change the symbology of objects themselves? That's the topic for the next lesson!

4.3 Lesson: Classification

Labels are a good way to communicate information such as the names of individual places, but they can't be used for everything. For example, let's say that someone wants to know which area the farms are in. Using labels, you'd get this:
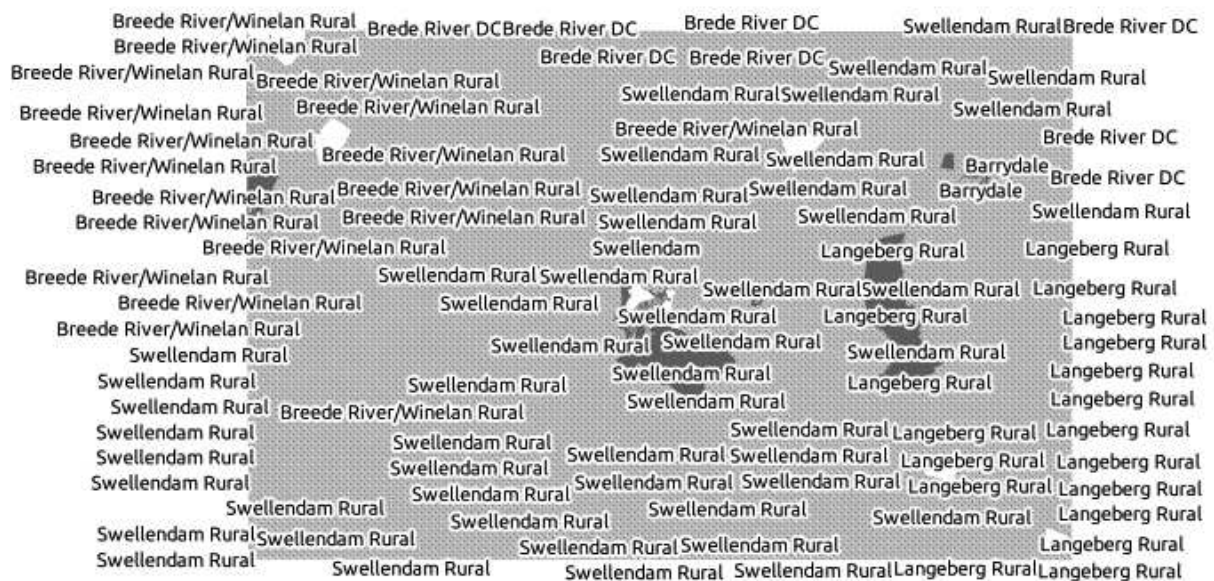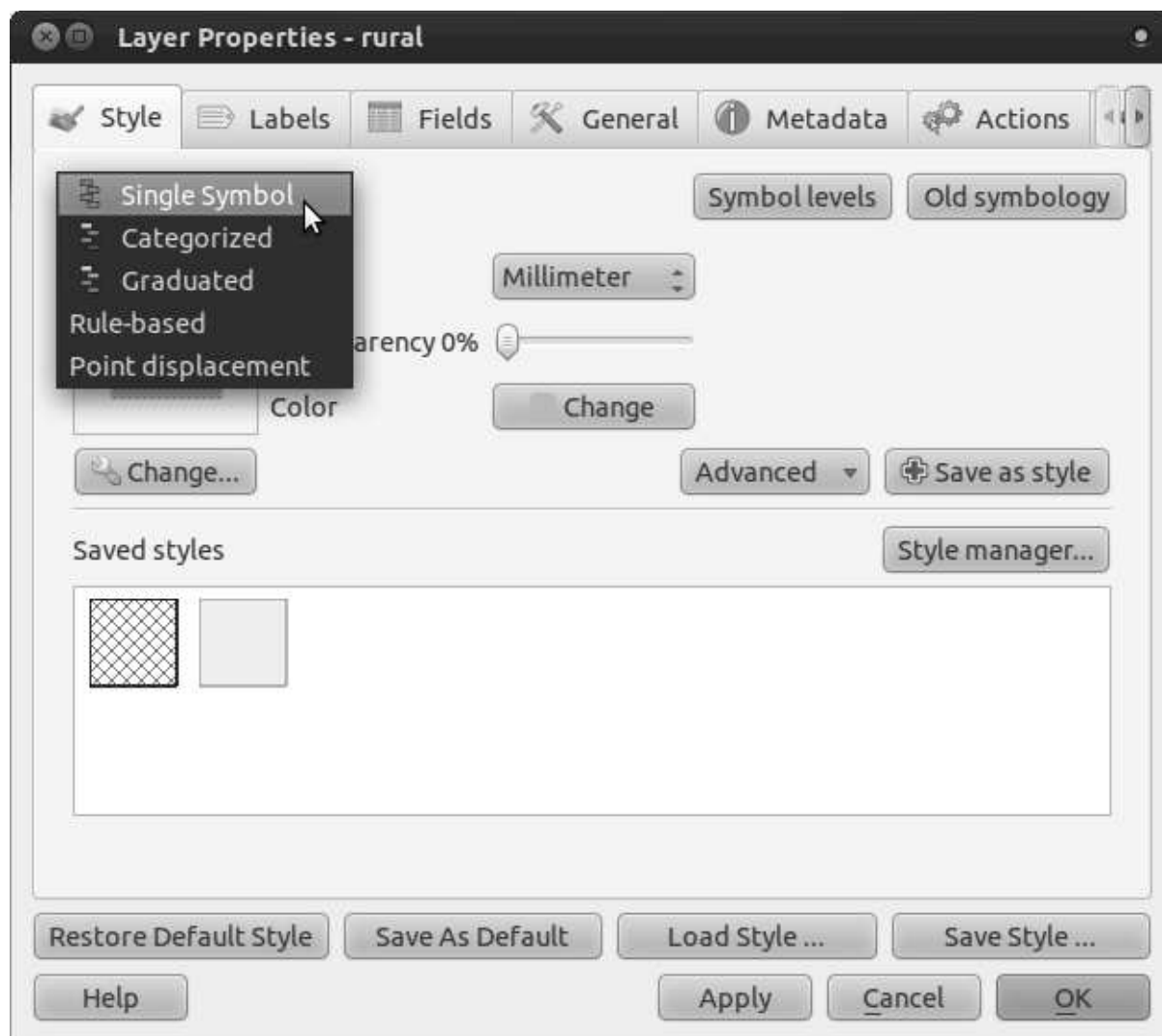

Obviously this is not ideal, so we need another solution. That's what this lesson is about!

The goal for this lesson: To learn how to classify vector data effectively.

4.3.1 Follow along: Classifying nominal data

- Open *Layer Properties*.
- Go to the *Style* tab.
- Click on the dropdown that says *Single Symbol*:

• Change it to *Categorized* and the interface will change:
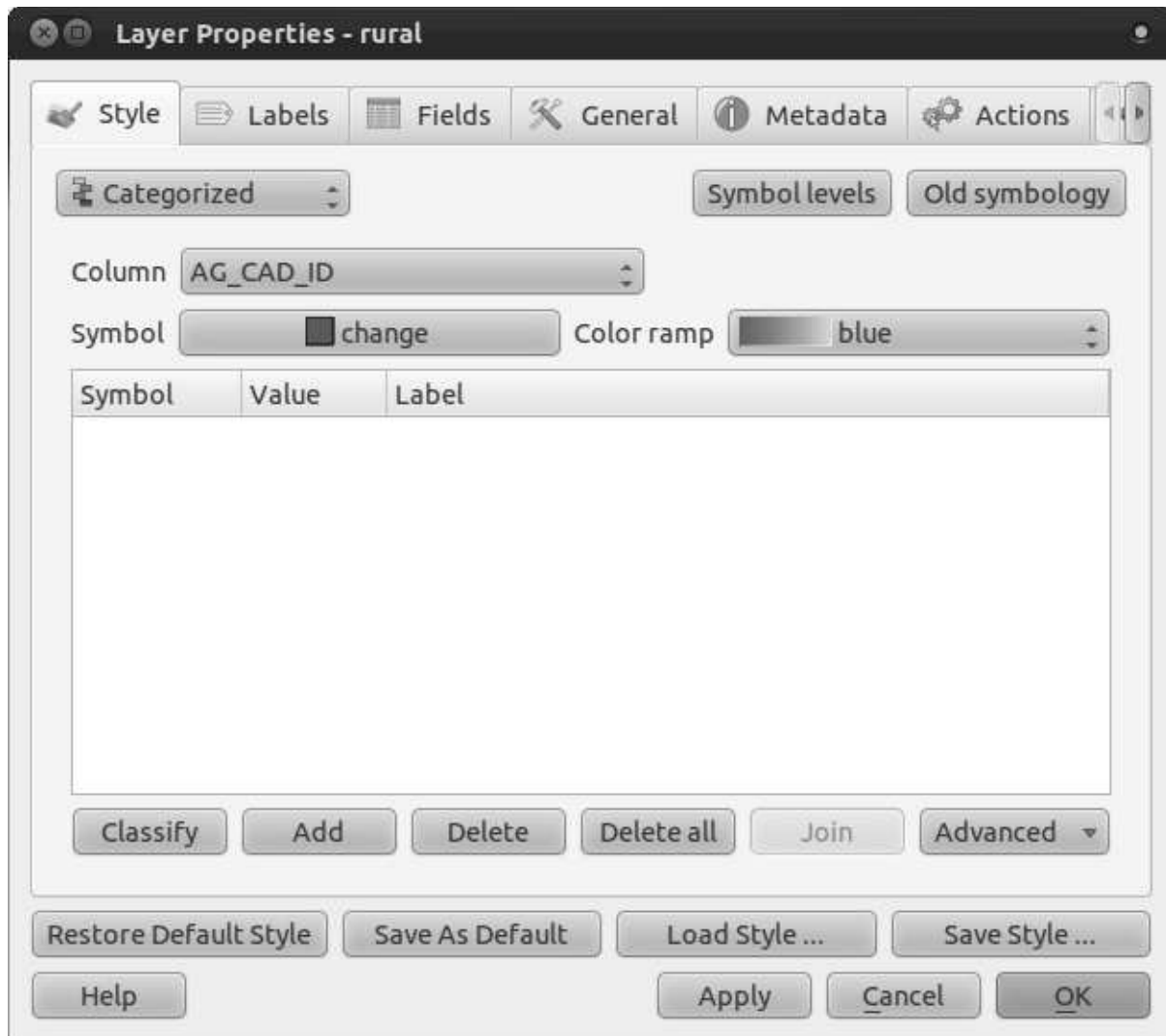
- Change the *Column* to *TOWN* and the *Color ramp* to *random*:

• Click the button labeled *Classify* and then click *OK*.
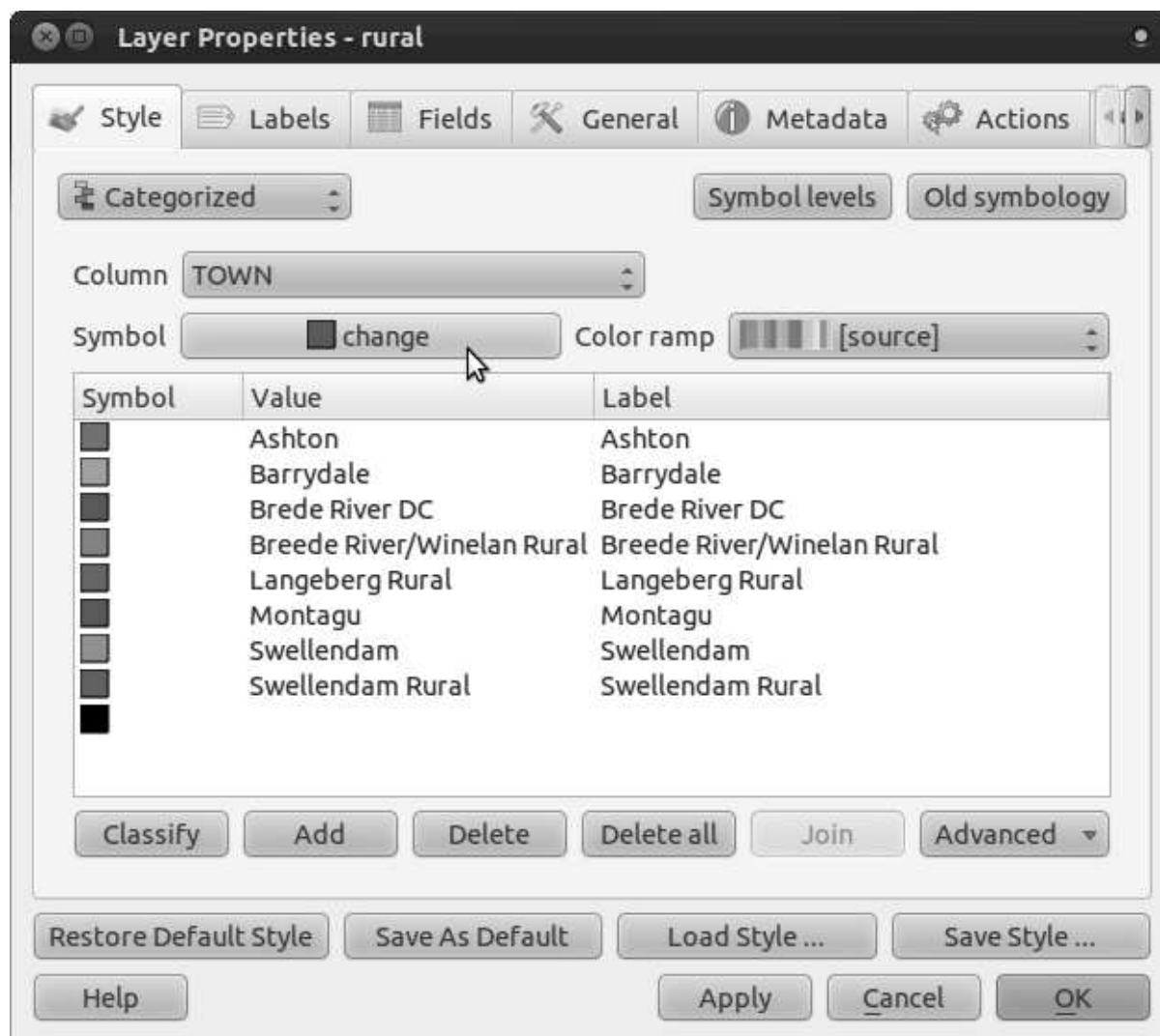
You'll see something like this:
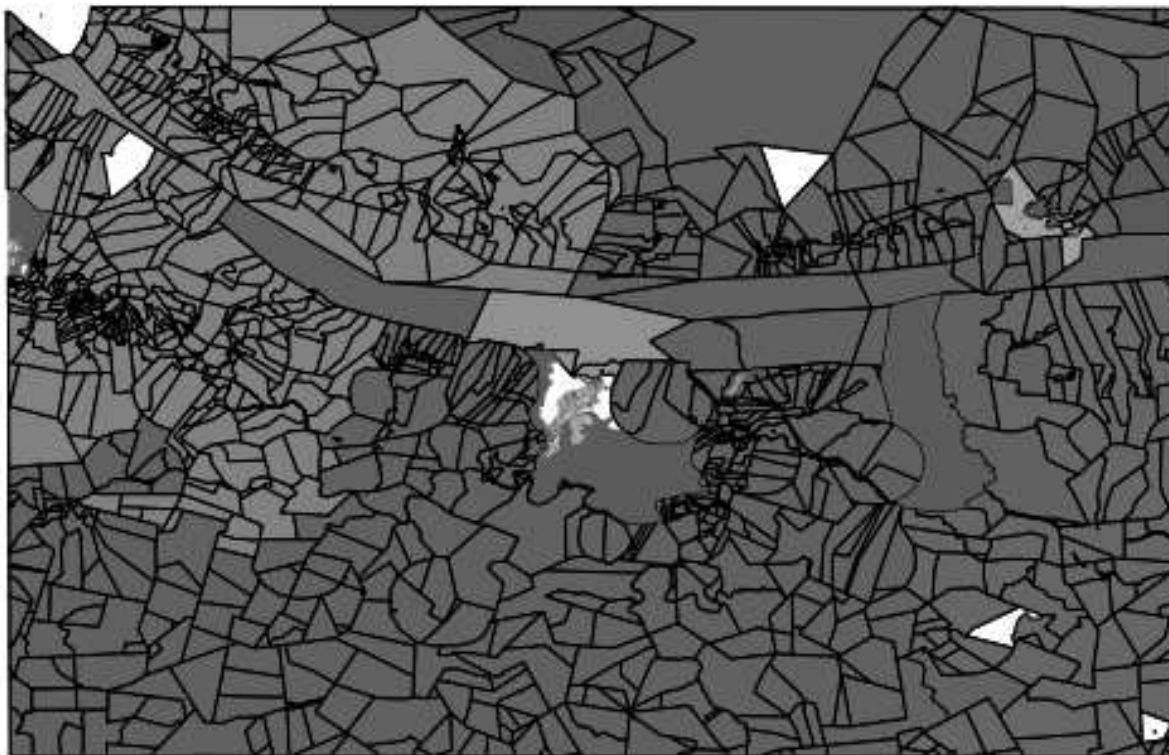
- Click the arrow (or plus sign) next to *rural* in the *Layer list*, you'll see the categories explained:
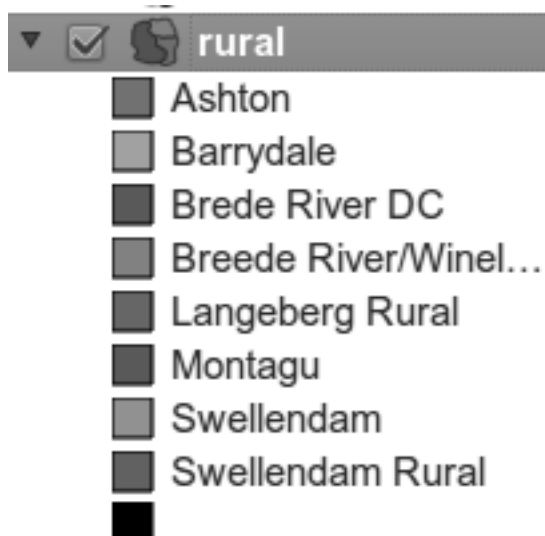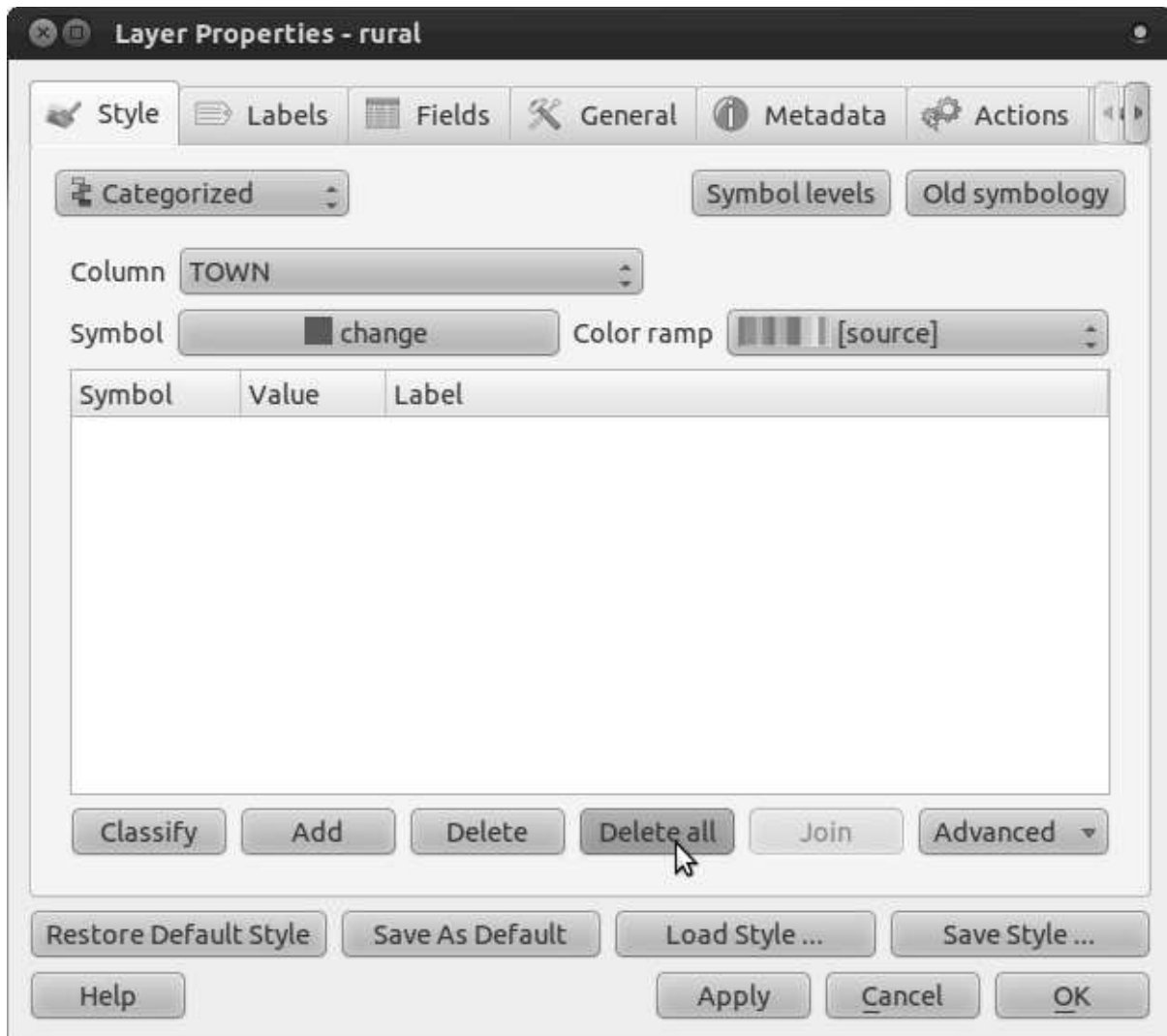

So, this is useful! But it hurts your eyes to look at it, so let's see what we can do about that.

- Open *Layer Properties* and go to the *Style* tab again.

- Change the symbol:

- Get rid of the outline the same way you did before, and click *OK*. (If you need to, go back to the lesson where we covered this to remember how it's done.)

- Click the *Delete all* button:

This gets rid of the ugly classes so you can try again.

• Click *Classify* again, and the new symbols will appear.

You'll notice they don't have outlines. This is because you just removed the outlines!

• Change the color for each town by double-clicking on the colored block representing its symbol:
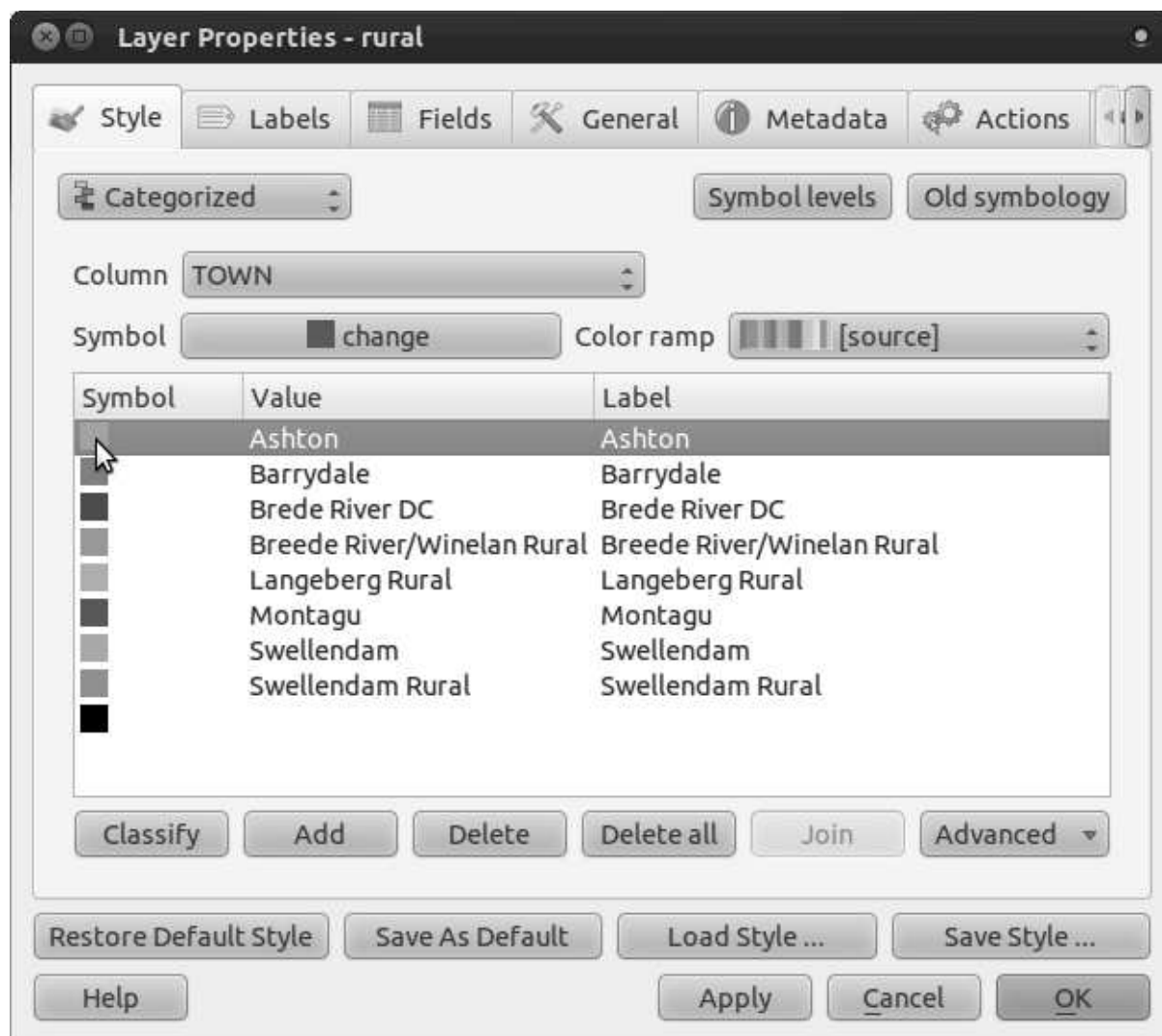
Use your own colors, but make sure they're not too bright, so that the resulting map isn't as ugly as the old one. In the example, we'll use these colors:
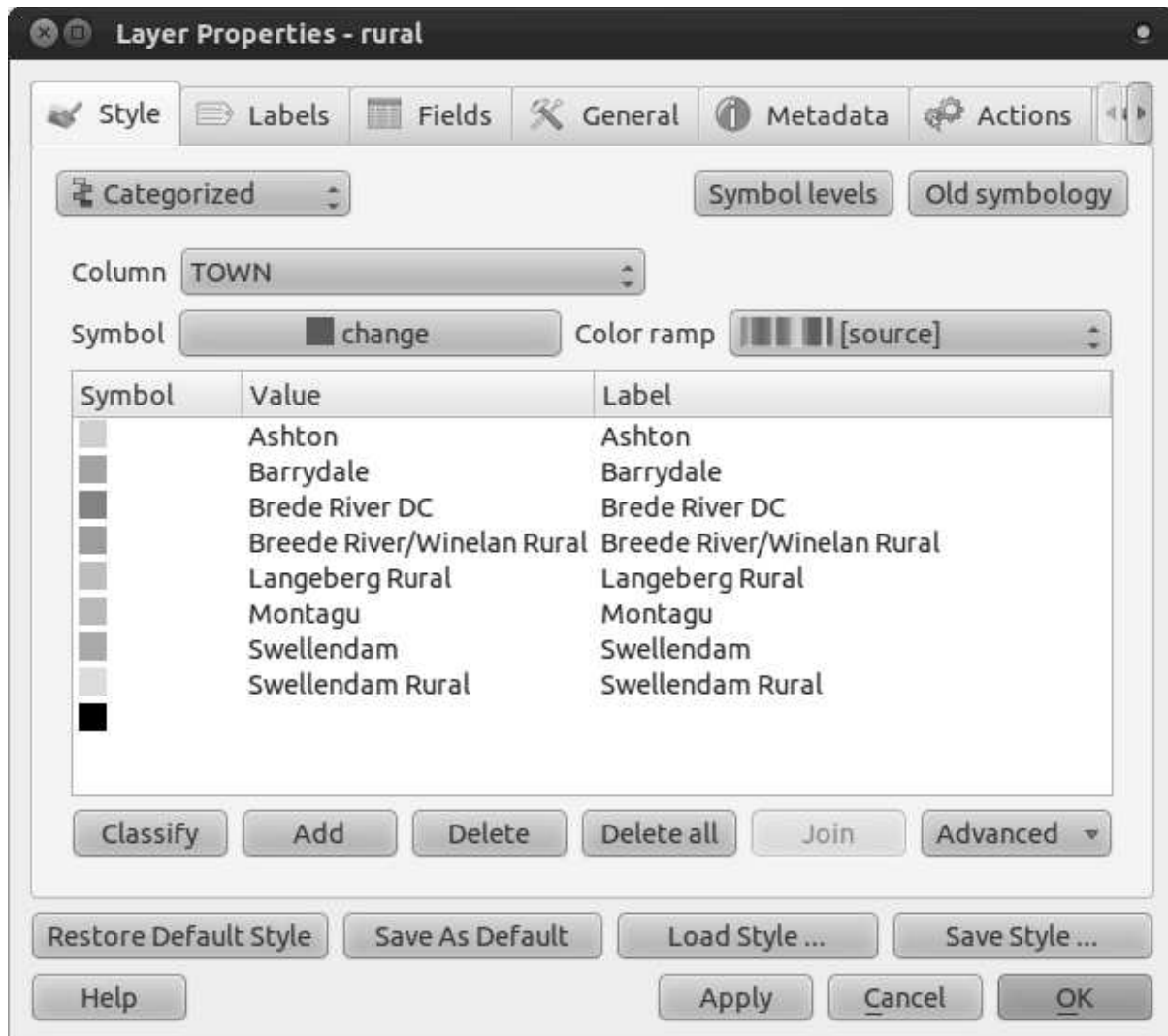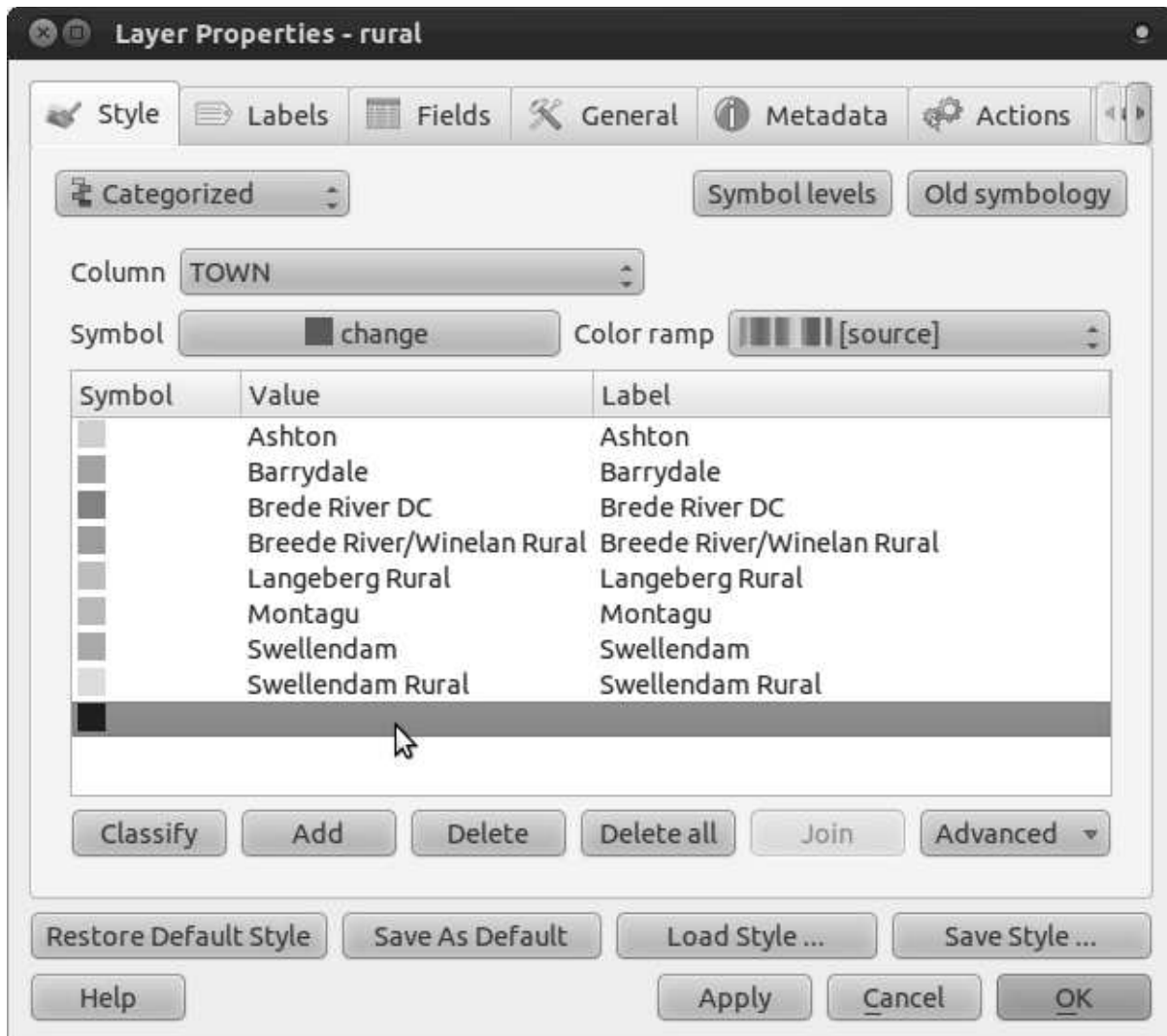
This gives us a nice map:

(Only *rural* and *water* are switched on here.)

By the way, there's one category that's empty:

- Select the empty category.
- Click the *Delete* button.

This only gets rid of the symbol, not the data, so don't worry about messing up; you're not actually deleting anything that you can't recover.

Remember to save your map now so that you don't lose all your hard-earned changes!

4.3.2 ⬟ Try yourself... More classification

If you're only following the basic-level content, use the knowledge you gained above to classify the *urban* areas. Use darker colors to set them apart from the farms.

4.3.3 ⬟ Follow along: Ratio classification

There are four types of classification: *nominal*, *ordinal*, *interval* and *ratio*.

In nominal classification, the categories that objects are classified into are name-based; they have no order. For example: town names, district codes, etc.

In ordinal classification, the categories are arranged in a certain order. For example, world cities are given a rank depending on their importance for world trade, travel, culture, etc.

In interval classification, the numbers are on a scale with positive, negative and zero values. For example: height above/below sea level, temperature above/below freezing (0 degrees Celsius), etc.

In ratio classification, the numbers are on a scale with only positive and zero values. For example: temparature above absolute zero (0 degrees Kelvin), distance from a point, the average amount of traffic on a given street per month, etc.

In the example above, we used nominal classification to assign each farm to the town that it is administered by. Now we will use ratio classification to classify the farms by area.

- Save your rural symbology (if you want to keep it) by clicking on the *Save Style ...* button in the *Style* dialog.

We're going to reclassify the layer, so existing classes will be lost if not saved.

- Close the *Style* dialog.

- Open the layer attributes for the *rural* layer.

We want to classify these farms by area, but there's a problem: they don't have an area field! We'll have to make one.
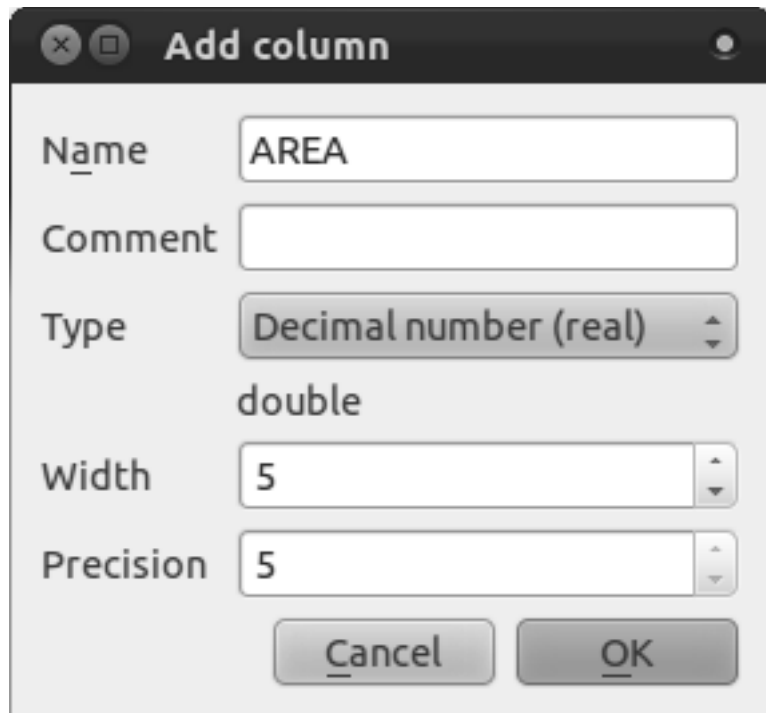
- Enter edit mode by clicking this button:

- Add a new column with this button:

- Set up the dialog that appears, like this:
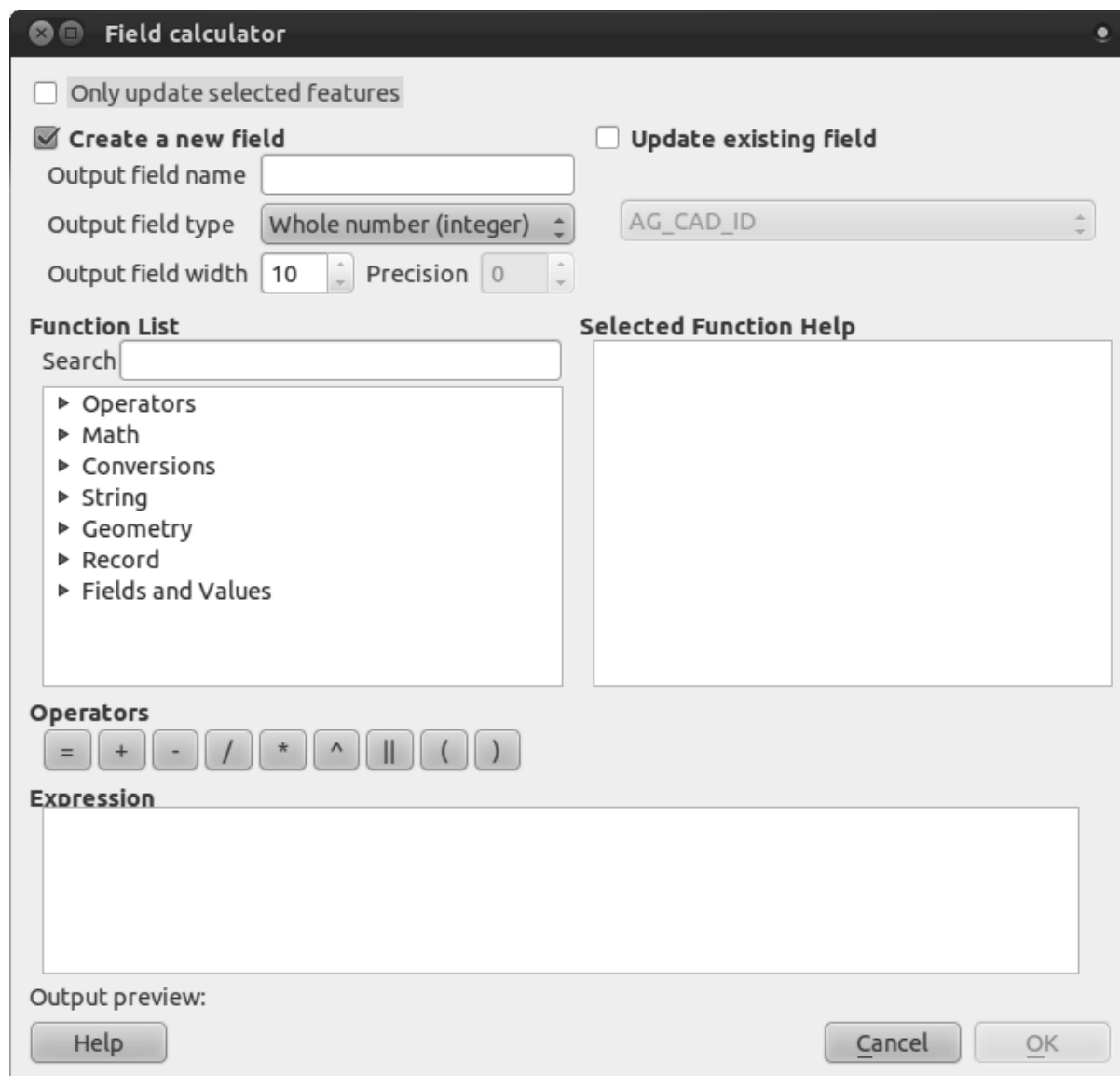
- Click *OK*.

The new field will be added (at the far right of the table; you may need to scroll horizontally to see it). However, at the moment it is not populated, it just has a lot of NULL values.

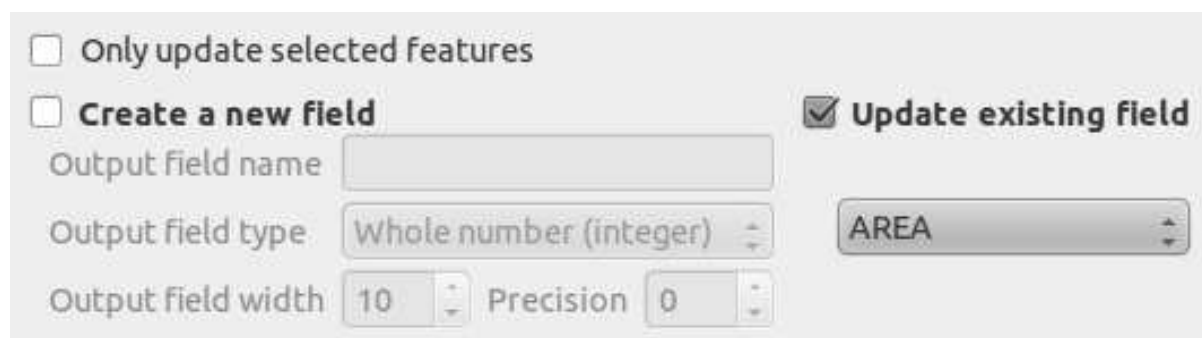To solve this problem, we'll need to calculate the areas.
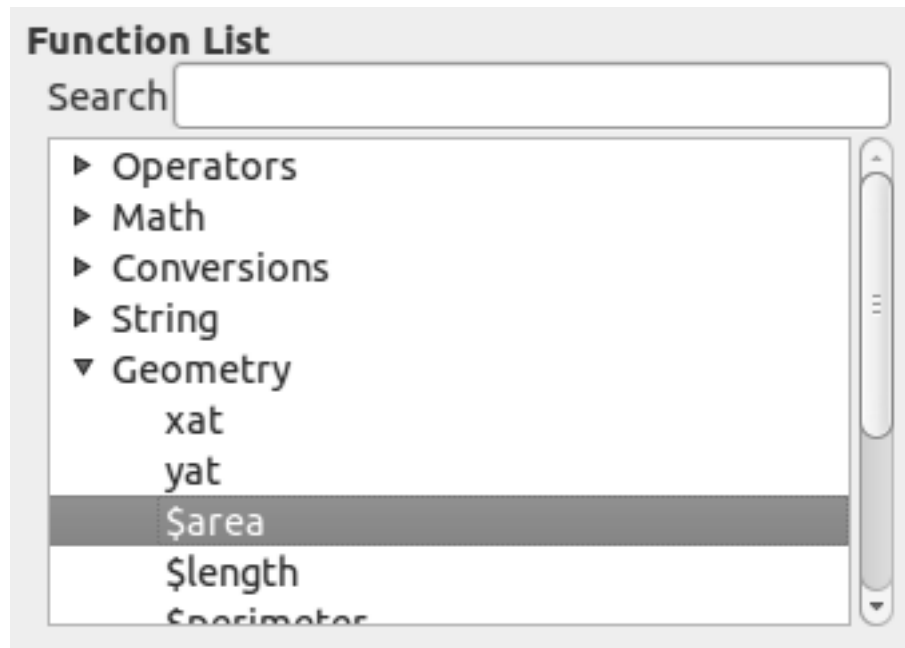
- Open the field calculator:

You'll get this dialog:

- Change the values at the top of the dialog to look like this:


- In the *Function List*, select *Geometry* → *$area*:

- Double-click on it so that it appears in the *Expression* field.
- Click *OK*.

Now your `AREA` field is populated with values! Admire them, then close the attribute table.

Note: These areas are in degrees area. Later, we will compute them in square meters.
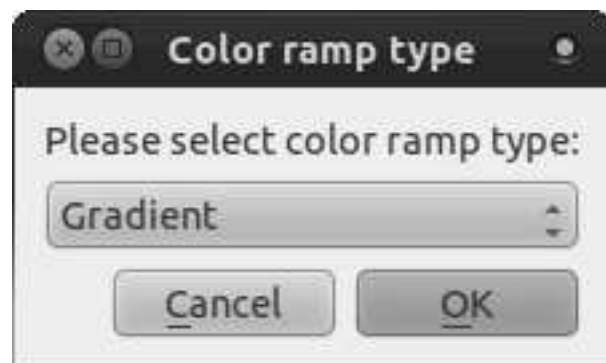
- Open the *Layer properties* dialog's *Style* tab.
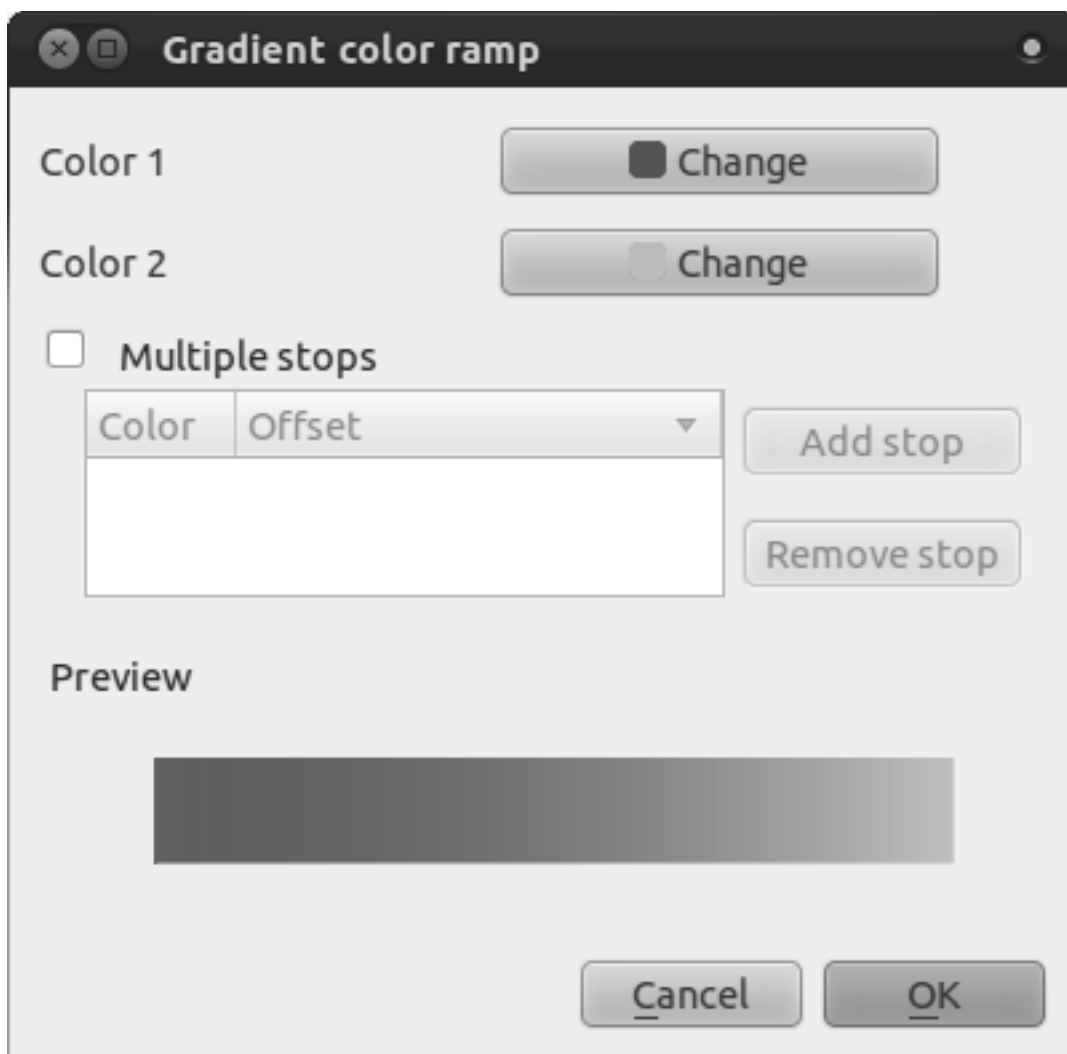- Change the classification style from *Classified* to *Graduated*:
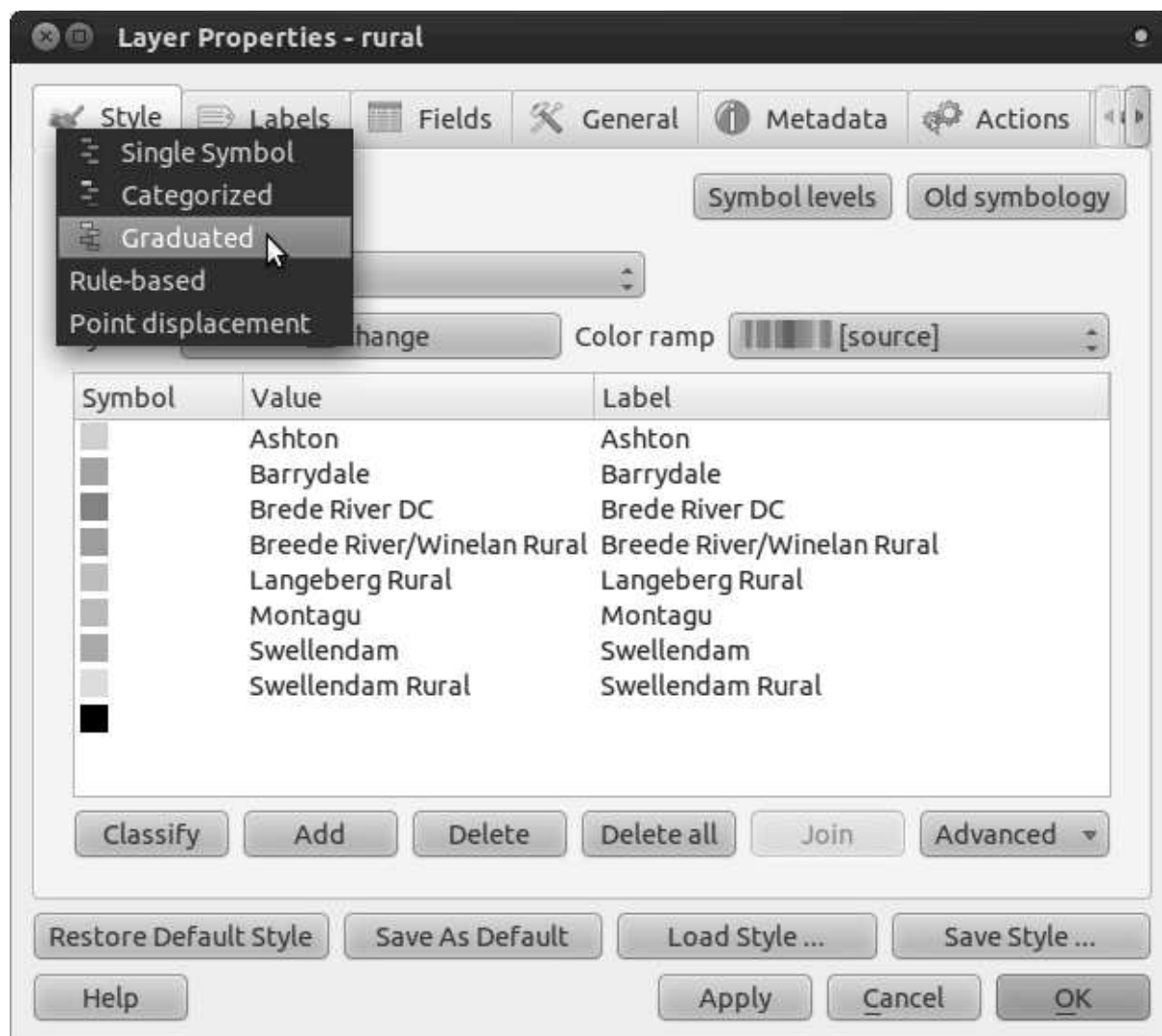
• Change the *Column* to *AREA*:
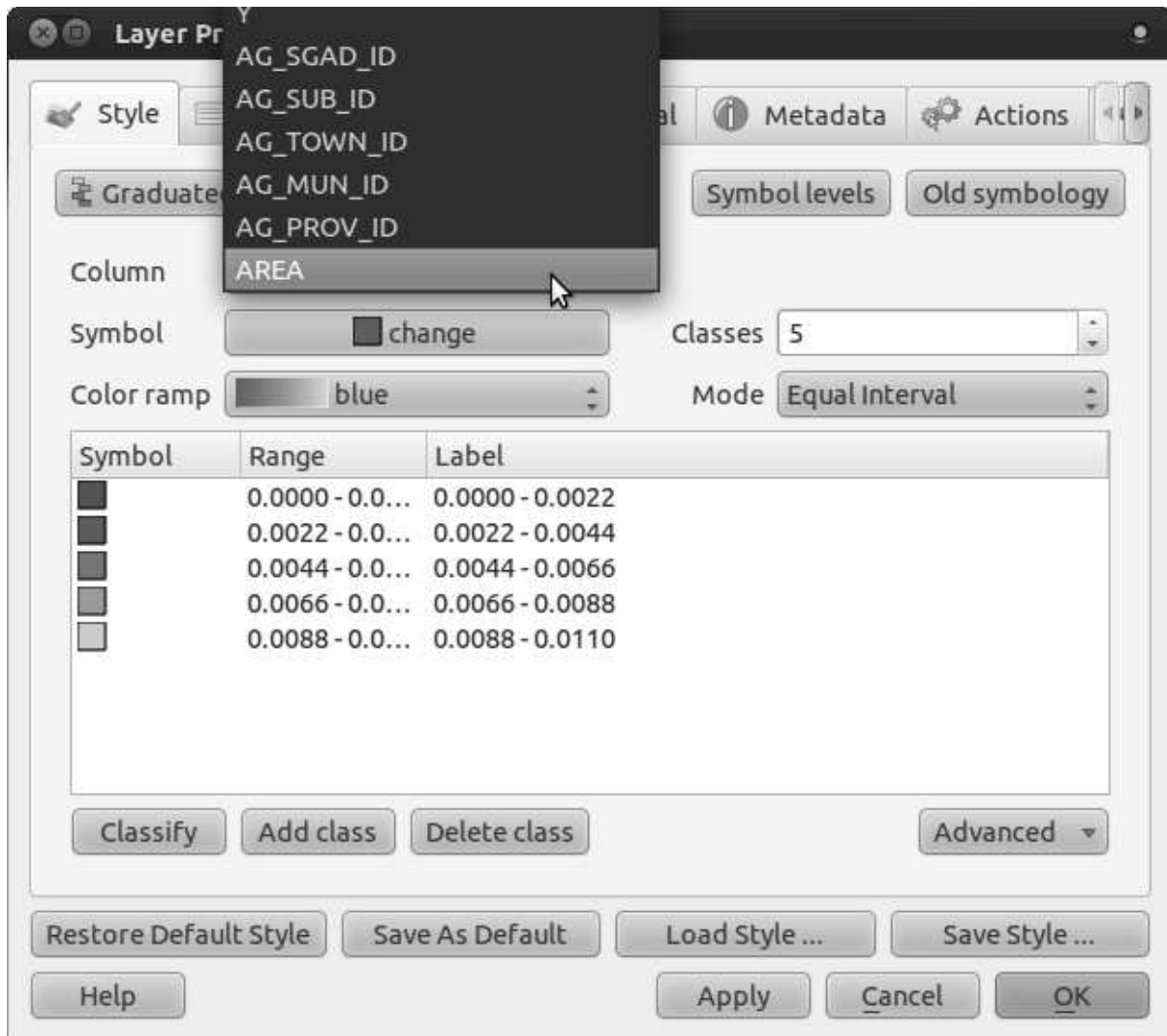
- Under *Color ramp*, choose the option *New color ramp...* to get this dialog:
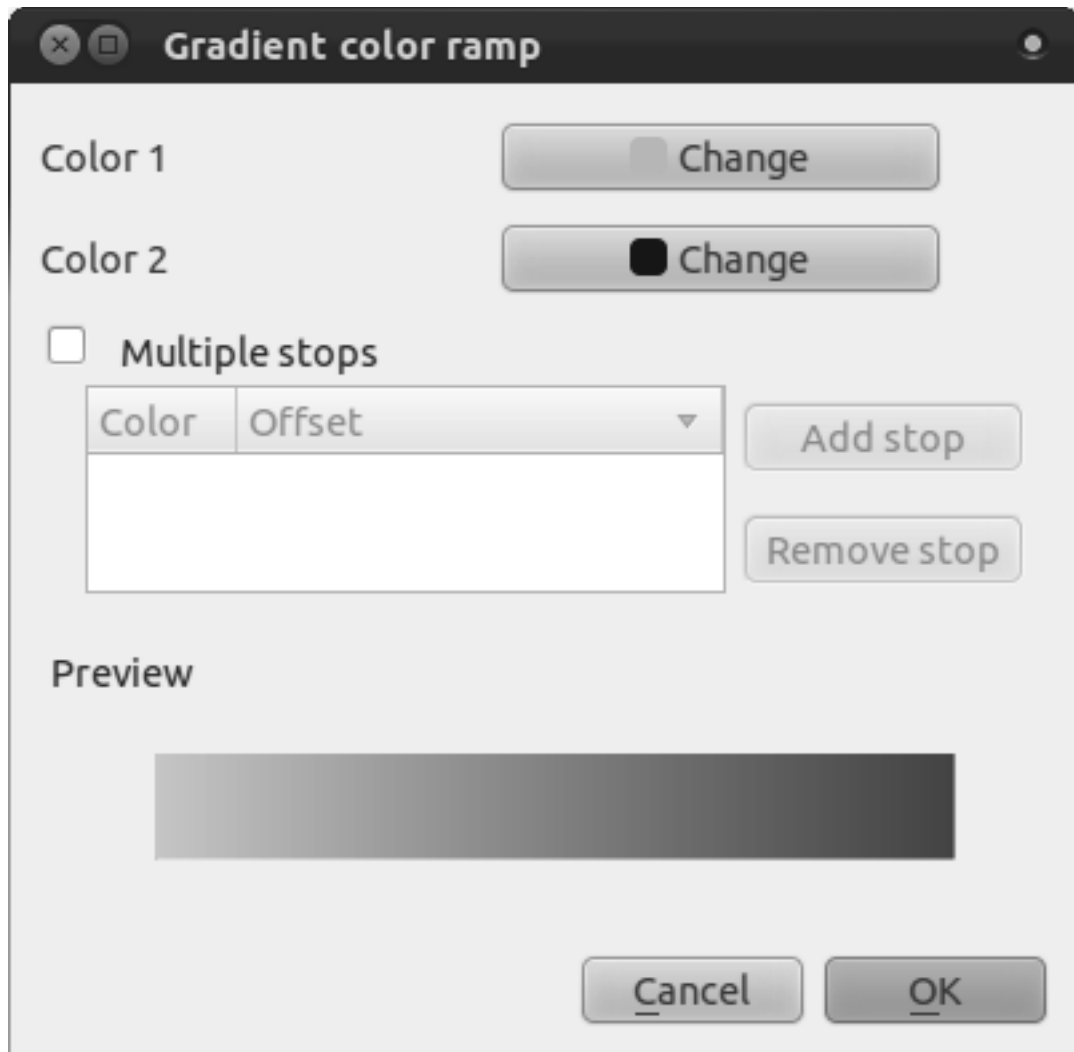

- Choose *Gradient* (if it's not selected already) and click *OK*. You'll see this:

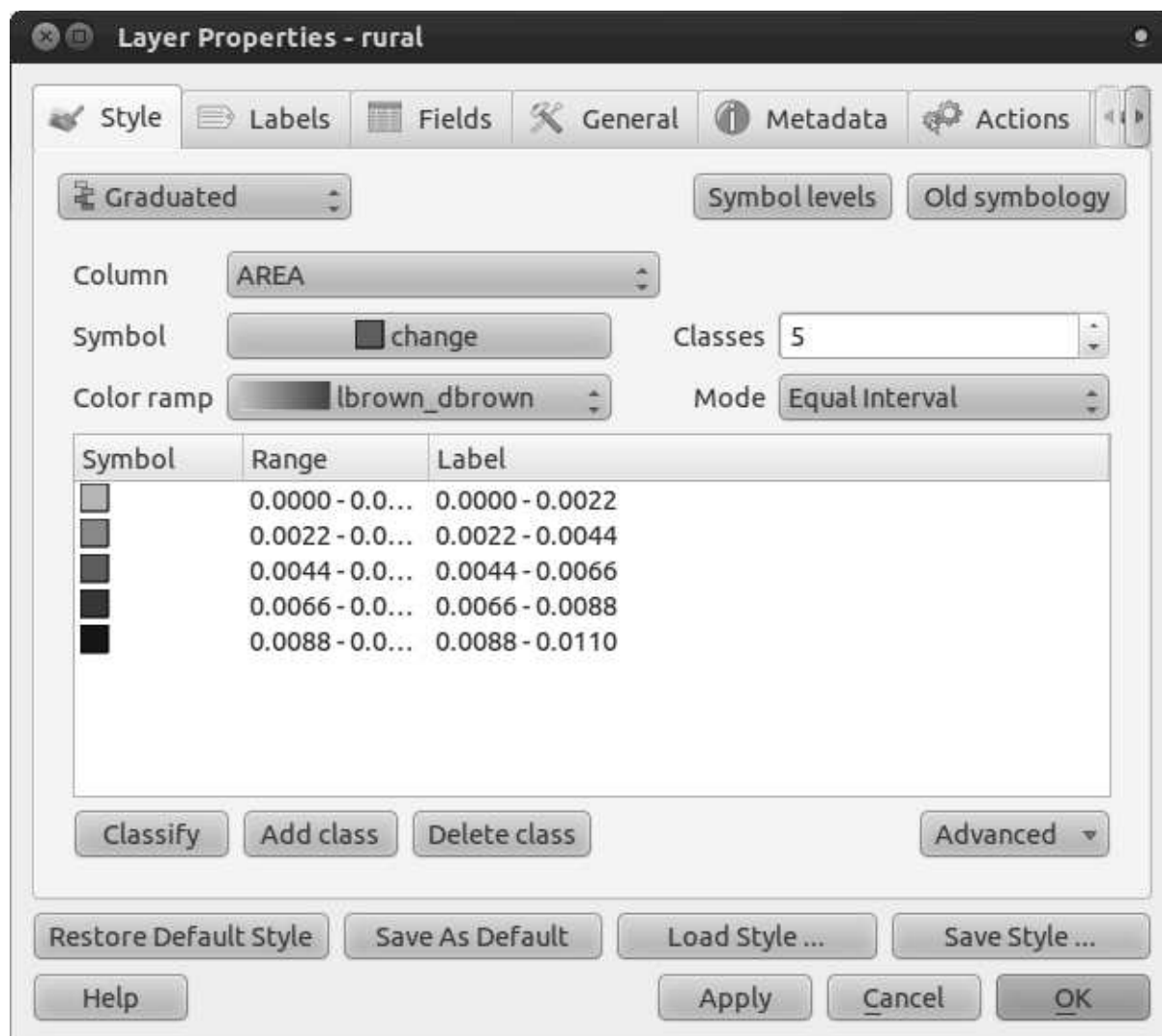You'll be using this to denote area, with small areas as *Color 1* and large areas as *Color 2*.

- Choose appropriate colors.

In the example, the result looks like this:

- Click *OK*.

- Choose a suitable name for the new color ramp.

- Click *OK* after filling in the name.

Now you'll have something like this:

Leave everything as-is.

- Click *Apply*:

### 4.3.4 ⬥ Try yourself... Refine the classification

- Get rid of the lines between the classes.

- Change the values of *Mode* and *Classes* until you get a classification that makes sense.

Check your results

### 4.3.5 ⬥ Follow along: Rule-based classification

It's often useful to combine multiple criteria for a classification, but unfortunately normal classification only takes one attribute into account. That's where rule-based classification comes in handy.

- Open the *Layer Properties* dialog for the *rural* layer.

- Switch to the *Style* tab.

- Switch the classification style to *Rule-based*. You'll get this:
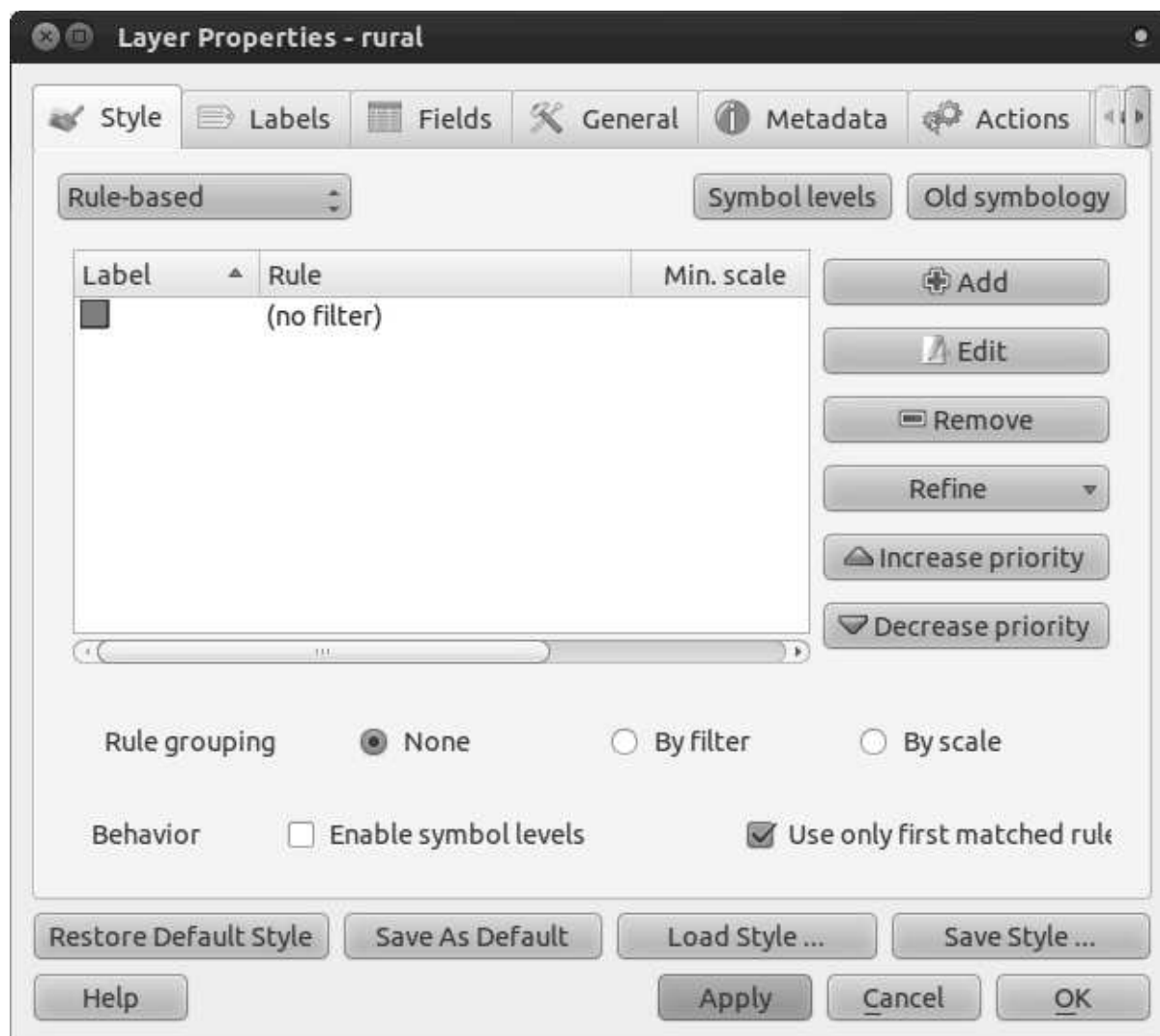
- Click the *Add* button.

- A new dialog that appears.

- Click the ellipsis ... button next to the *Filter* text area.

- Using the query builder that appears, enter the criterion `AREA >= 0.00085` and choose a dark color for it.

- Add the criterion `AREA <= 0.00085` and choose a light color.

- Add the criterion `TOWN != 'Swellendam Rural'` and assign it the color black, with transparency at `85%`.

- Click on the `TOWN ...` criterion in the list of rules, and then on *Increase priority*.

Your dialog should look like this:
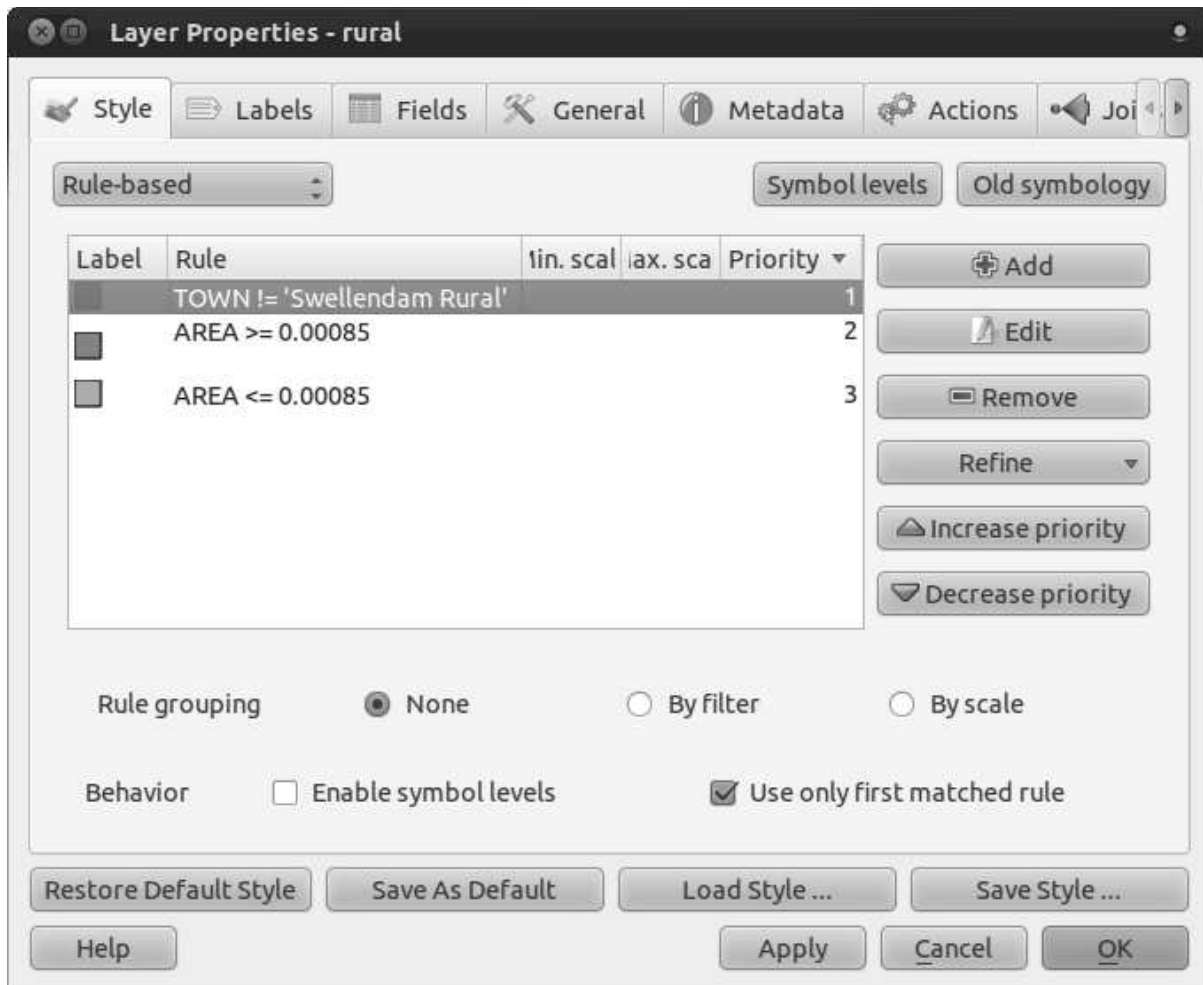
- Apply this symbology.

Your map will look like this:

Now you have two area classes, with the farms in the Swellendam Rural area emphasized.

4.3.6 In conclusion

Symbology allows us to represent the attributes of a layer in an easy-to-read way. It allows us as well as the map reader to understand the significance of features, using any relevat attributes that we choose. Depending on the problems you face, you'll apply different classification techniques to solve them.

4.3.7 What's next?

Now we have a nice-looking map, but how are we going to get it out of QGIS and into a format we can print out, or make into an image or PDF? That's the topic of the next lesson!