

# Enriching and Visualising Data

This exercise looks at taking some simple data and enriching it in order to do some visualisation. For the purposes of this exercise we are going to look at world population data published on Wikipedia.

## Required Software

In this exercise we are going to make use of Google Fusion tables, an experimental google service that allows for quick merging and visualising of data. Although the service can be run from any browser by visiting [www.google.com/fusiontables](http://www.google.com/fusiontables), it is recommended that you installed the add-on to Google Chrome. To do this add a new tab to Google Chrome, click “App Store” and search for Fusion Tables.



### Google Chrome + Add-ons

Googles Web Browser



### Fusion Tables

Bust your data out of its silo! Combine it with other data on the web. Collaborate, visualize and share.

**Note:** This exercise cannot be performed if you are logged into google using an organisational account. You will need to ensure that you are logged into google using a private account, e.g. @gmail.com

## Importing Data

Once loaded, search for a public dataset about population by typing this in the web search box:

Or search public data tables

From the results list, pick the population dataset from Wikipedia that contains over 200 rows of data and import it by clicking **import data**.

Once it is imported you will need to click the **see table** link to view your first fusion table.

## A Quick Map

To perform a simple mapping visualisation of population data, first click the **file menu** and select **geocode**. Ensure **country** is selected and click start. This will enable the map tab!

Clicking on a marker will show the record for that country.



Having all the markers the same colour is not very useful, however we can change this. To do this we first need to check that the data schema is correct and that the population data is numeric (0-9 only) and not a string value. Sorting a string is different from numeric values, additionally numeric sorting operations are not available if fusion tables thinks your data is a string.



By browsing back to our table data, select **change** from the dropdown list against the population column. This will pop up a box where we can change the data in the column to be numeric in type. Be careful to select the correct number template.

**Population**

Change column

Name

Type **Number**

Format **1,234.56**

**Save changes** **Cancel**

When done click **save changes**, this should not have changed the data. If you wish, do the same for the other columns to ensure they are all of the correct type.

Going back to our map, from the **tools menu** select **change map styles**. Under the **marker icons** section, select **buckets** and specify some new ranges and marker colours for the population data.

[Fixed](#) [Column](#) **Buckets**

Updating the map will now show a number of colours representing the varying populations.

**Divide into** **Custom** buckets

Column **Population**

**use this range**

+	0.0	up to 1,000,000.0		
+	-	1,000,000.0	up to 10,000,000.0	
+	-	10,000,000.0	up to 100,000,000.0	
+	-	100,000,000.0	up to 1,000,000,000.0	
+	-	1,000,000,000.0	up to 2,000,000,000.0	
		2,000,000,000.0		



## Polygons

Although markers are one way to display the data, it would probably be clearer to fill the whole country in with a colour dependant on the size of the population.

To do this from the **tools menu** select **change map styles** and this time select a **gradient fill** from the polygons section. Select your colour and fill in the range values to map those from the data.

When done, save the changes and you will notice that nothing changes. Even though geocoding the dataset allows you to place a marker in a country, the geocode operation only provides a centre pointy, and not the polygon data.

Polygon background colors

[Fixed](#) [Column](#) **Buckets** **Gradient**

**Show a gradient**

Column **Population**

**use this range**

From **0**

+	
+	
+	

Through **1354040000**

In order to complete this exercise, we need to merge our data with another dataset that contains this information.

## Enriching Data

To enrich our data, select **find a table to merge with** from the file menu. As we are seeking polygon data we need to ensure that we are looking for tables that match on our country data.

From the results we can then choose one that looks most appropriate to the data we currently have and has the largest number of matches.

Once you have found an appropriate one import it by clicking **next**.

Inconveniently, this will create you a new table. As a result you will need to set up all your **map styles** again.

Merge: Select a table

Suggest tables matching on

Country (or dependent territory) ▾



Each match adds about 8 rows.

<input type="checkbox"/>	Birth Rate 2009	<a href="#">view table</a>
	68% of rows have a match.	
<input checked="" type="checkbox"/>	World Country Boundaries.kml	<a href="#">view table</a>
	52% of rows have a match.	
<input type="checkbox"/>	world_adm0.kml	<a href="#">view table</a>
	66% of rows have a match.	
<input type="checkbox"/>	World Country Boundaries	<a href="#">view table</a>
	52% of rows have a match.	

Or [select a table from Google Drive](#)

Cancel

Next



## Extension Exercises

1. Can you find another dataset to merge with this one to display some point data alongside the population data to show correlation?
2. Taking one of the other datasets used in this course can you think of a way to enrich it and visualise it in fusion tables?