Quick introduction to Plot.ly

This worksheet provides a brief introduction to using the ploy.ly library to create and share rich visualisations of data.

The objective of this exercise is to create a heatmap comparing population characteristics (e.g. education level) to employment rate.

The datasets used for this example are available openly from data.gov.uk.

http://data.gov.uk/dataset/employment_rates_of_working_age_by_qualification_level

Step 1 - Select your data

First, select from the many datasets that exist in the source spreadsheet. The source spreadsheet does not organise the data into nice rows and columns that we can simply plot. Nor are the categories very linear.

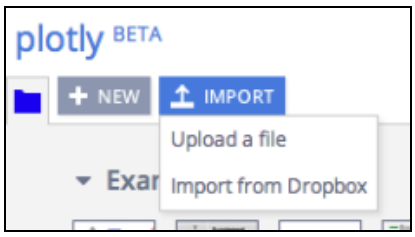
Plot.ly requires data organised in row and column based formats with only one set of row and column titles. Using your favorite spreadsheet tool, decide on a subset of the data you want to plot, and make a new worksheet that is in a plain format as below.

The example in this worksheet uses the following data pertaining to UK region and education level:

UK Region	**Level 7-8**	**Level 4-6**	**Level 3**	**Level 2**	**No qualifications**
North East	6.6	22.0	22.7	24.0	8.2
North West	6.8	24.0	20.8	22.7	8.5
Yorkshire & the Humber	7.3	23.4	21.1	22.0	7.6
East Midlands	5.7	23.7	21.5	21.5	8.6
West Midlands	7.0	23.0	20.8	22.2	9.6
East	7.1	21.7	21.0	22.2	8.7
London	14.2	32.2	14.6	16.1	7.2
South East	8.2	26.6	21.3	20.7	5.9
South West	6.7	24.9	21.9	22.0	5.7

Once you have selected your data, save it out as a CSV file.

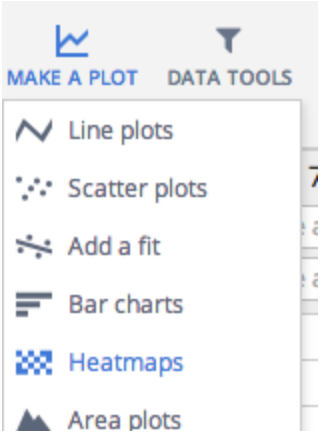
Step 2 - Import into plot.ly



Register yourself at account at **http://plot.ly**, create a

new project and **import your data**.

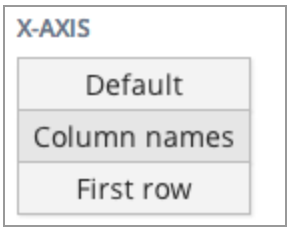
Step 3 - Make a plot



Ensure that you have the data table displayed on your screen. In plot.ly this is named the “grid layout” and can be accessed via the grid icon.

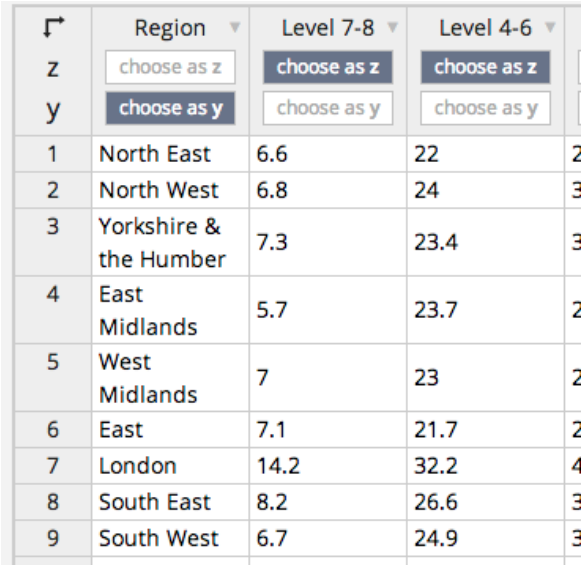
With the data displayed, it is very important to ensure that plot.ly has detected the correct column titles. If not you may need to re-import or manually change them.

At this point you can create a plot by clicking the **make a plot** button. For the purpose of this exercise we are going to create a heatmap to compare UK region to education level. We are comparing the relationships between all data points across 2 dimensions.



Both axis in our dataset are categorical so we need to ensure that our plot is aware of this. To do this select **column names** from the x-axis selection for the heatmap.

Lastly, it is necessary to select the data for the x (or sometimes z in ploy.ly) and y axis. In this example we have selected the **UK Region** as the **y axis** value and the **qualification levels** as the **x/z axis**.



	Region	Level 7-8	Level 4-6	
z	choose as z	choose as z	choose as z	
y	choose as y	choose as y	choose as y	
1	North East	6.6	22	2
2	North West	6.8	24	3
3	Yorkshire & the Humber	7.3	23.4	3
4	East Midlands	5.7	23.7	2
5	West Midlands	7	23	2
6	East	7.1	21.7	2
7	London	14.2	32.2	4
8	South East	8.2	26.6	3
9	South West	6.7	24.9	3

Once complete, click **make heatmap** to create your plot.

Step 4 - Sharing you plot

Before sharing your plot, use the plot.ly toolbar to add axis titles and customise the colours used.

Once done, click the share button, give your plot a name and make it publicly available.

You could now simply send people a link to your plot, however we are going to embed this page in the website created as part of the **Publishing Data in Github** worksheet. In order to embed your visualisation you will need to have completed steps 1-3 inclusive from this worksheet.

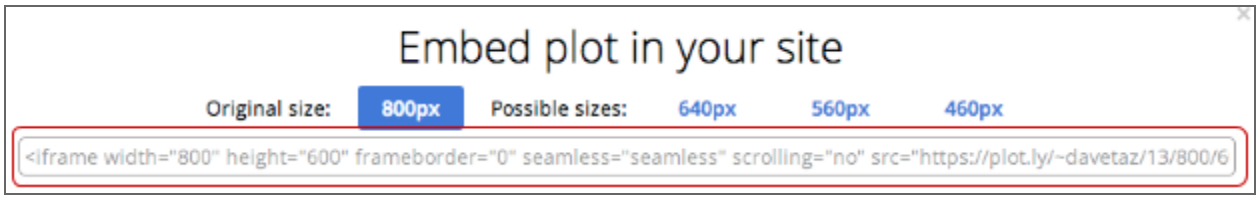
In a separate tab, log into GIthub and open your data publication project. From the list of files click on **index.html** to open it. From here click the **edit** icon (pencil) and scroll down to find the following line:

{% include data.html %}

Above this line add the following:

```
<h2>Visualisation exemplar from data</h2>
```

Keep this window/tab open and **return to plot.ly** and click the embed button in the sharing settings. In the window that appears, copy the code block and paste this below the `<h2>` element you just created in Github.



In Github, commit your change (with useful comment) and then browse to your website to view your new embedded visualisation. The link to your website is available via the **settings** menu of the repository.

Note that your visualisation is interactive, you can zoom in and select specific regions. There is also a direct link to the data used for the visualisation and the code used to generate it, making it fully open.