

4.11 Exercise: Advanced scatterplots for deeper analysis *(iNZight Lite version)*

This exercise will enable you to use iNZight Lite to explore the change in the rate of infant mortality and the average number of children per woman for different countries around the world over time.

The skills addressed are:

1. Create a scatterplot of two numeric variables, subset by a 3rd variable.
2. Explore the effect of a third and fourth variable using colour and size.
3. Extract and explore changes in individual groups within a categorical variable over time.

INSTRUCTIONS

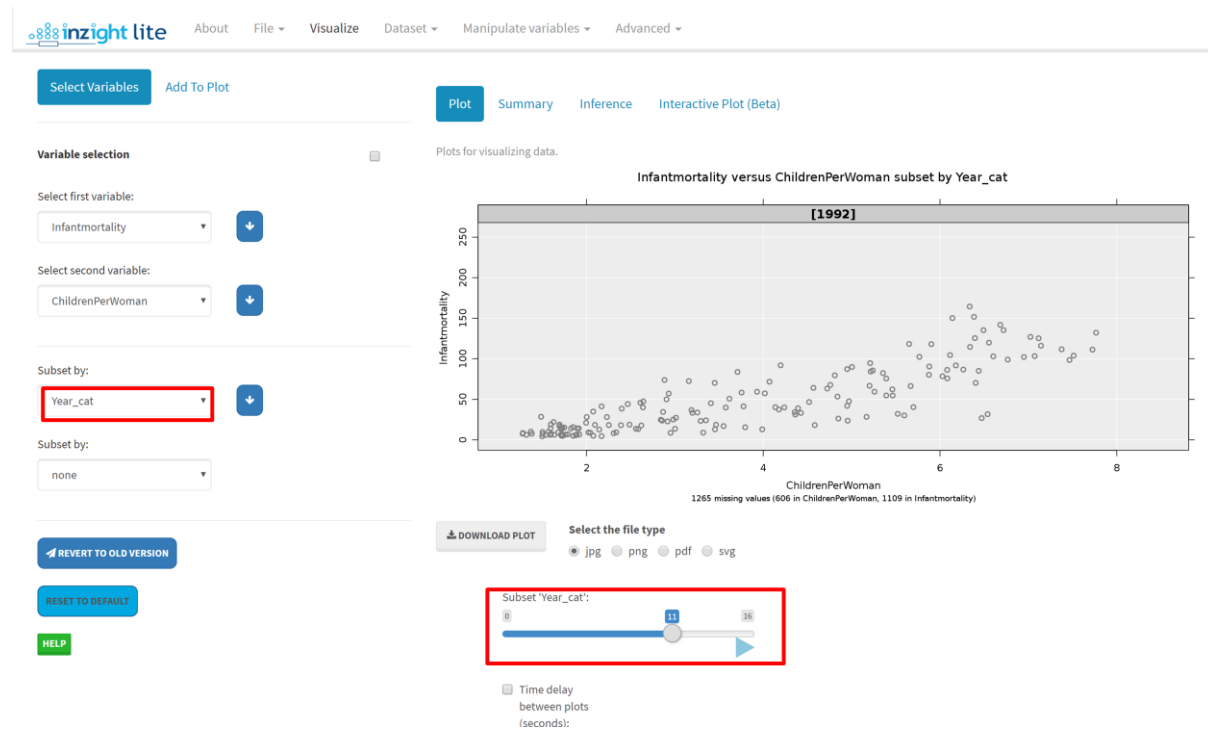
Import the **Gapminder** dataset into iNZight Lite:

- Select **File > Dataset Examples**
- Select Data set category: **Future-Learn**
- Select **Gapminder**
- Click on **Select Set**
- Click on **Visualize**

Create a scatterplot of two numeric variables, subset by a 3rd variable

We are going to explore the relationship between the variables **Infantmortality** and **ChildrenPerWoman** of countries in the **Gapminder** dataset.

Construct a scatterplot of **Infantmortality** (outcome variable) and **ChildrenPerWoman** (predictor variable, **subset by** the variable **Year_cat**). **Year_cat** is a categorical version of Year. [This way we get separate graphs for every subset formed by a value of the variable whereas iNZight cuts a numeric variable like Year up into just 4 class intervals.]



Move the slider from left to right or push **play** to watch the changes over years.

Do you see anything interesting in your plots that might be worth investigating further?

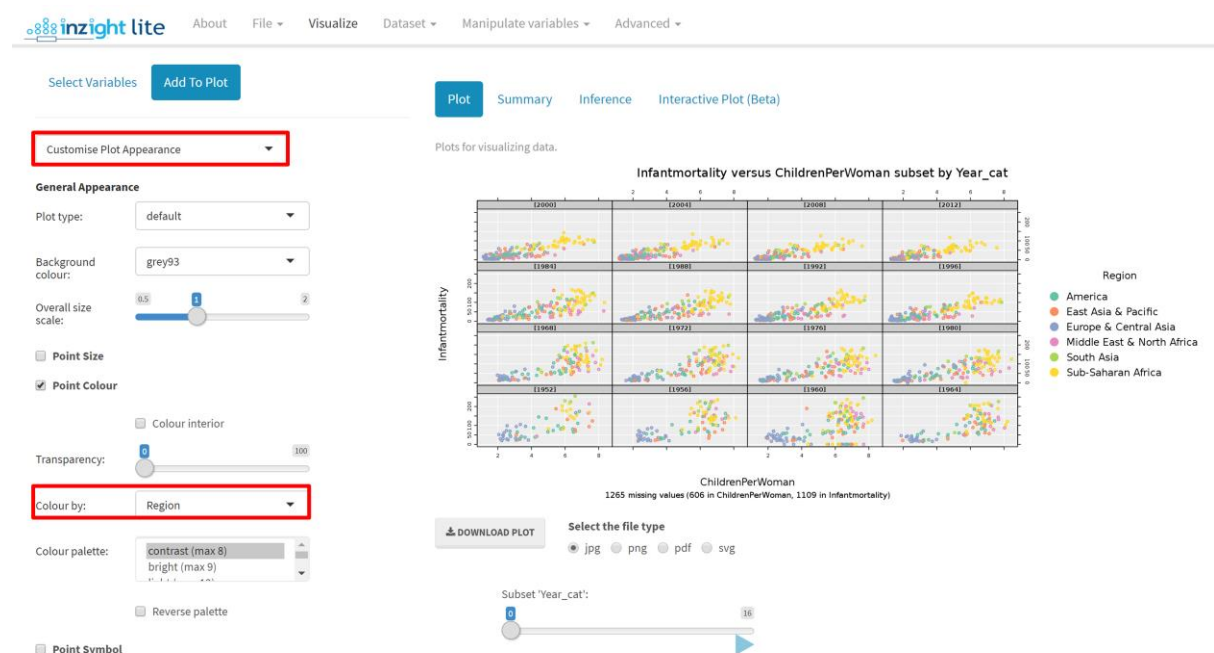
Explore the effect of a third variable using colour and size

We suspect that there is a lot more information that can be graphed to give us more insight into the rate of infant mortality and the number of children per woman.

- How does the graph change over time?
- What is happening in larger populations?
- Do different regions have similarities?

Explore the effect of **Region** by:

- Select **Add to Plot** tab and **Customise Plot Appearance**
- Select **Point Colour** to reveal options
- Select Colour by: **Region**
- Move the slider back to 0



What changes can you see over time? Use the slider and the **play** button to explore the changes over time.

Now adjust the size of the points according to the population:

- Select **Add to Plot** tab and **Customise Plot Appearance**
- Select **Point Size** to reveal options

- Select **Resize Points by: Populationtotal**
- Move the subset slider to 11 (1992)



Now change the appearance of the symbols so that it is easier to see:

- Select **Add to Plot** tab and **Customise Plot Appearance**
- Click the box for **Colour interior**
- Move the slider for **Transparency of symbols** to about 60
- Move the slider for **Size of symbols** to change the sizes of the plot symbols. Experiment with different size settings.



What do you notice when you use the slider to step through the years? If you wish to identify any countries of interest, use:

- First of all check that iNZight has identified this as a scatter plot by selecting the **Add to Plot** tab and **Change Plot Appearance** and selecting Plot type: **scatter plot**.
- Select **Add to Plot** tab and : **Identify points**
- Select type of Labelling: **Text label**
- Method of selection: **Extremes**
- Use the slider to investigate the data

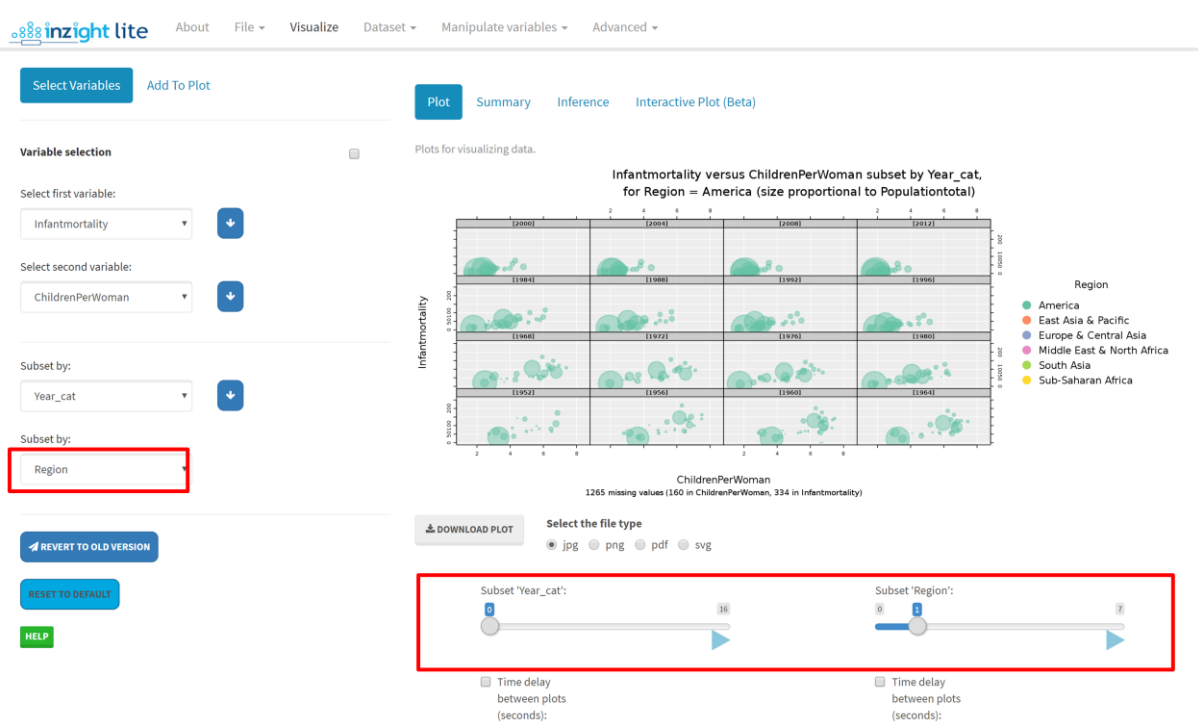
Note: you will need to view one year at a time to identify points.

Extract and explore changes in individual groups within a categorical variable over time

Sometimes we may just want to look at the changes in each **Region** separately. To do this:

- Select **Region** in the 2nd **Subset by** slot.
- Move the slider for **Region** onto the region you are interested in
- Move the slider for **Year_cat** to **0** or push **play**

Here is a tiled plot of the **Region** America:



What can you see on your series of plots about the changes in the rate of infant mortality (deaths per 100 live births) and the number of children per woman?

Look for something interesting in the data, post a comment and read what other people have discovered.

Move both sliders to **0**.

What happens?

Optional

If you want to play with a more recent version of the data download

[Gapminder_1952-2016_by_4yrs.csv](#) from

https://www.stat.auckland.ac.nz/~wild/data/gapminder_datasets/.

[Some variable names may differ somewhat from those in the Exercise above.]

Common questions

How can I edit iNZight graphics files saved from iNZight?

If you can use graphics software such as Adobe Illustrator (professional), Affinity Designer (inexpensive) or the free alternative Inkscape, you can use these to import pdf files saved from iNZight and edit the graphics in the graphics program.