

W07-1: Publication quality graphics

This worksheet revisits the visualization examples from last session but this time we will try to come up with publication quality plots. After completing this worksheet you should be able to plot enhanced visualizations which can be used for publications.

Things you need for this worksheet

- R — the interpreter can be installed on any operation system. For Linux, you should use the r-cran packages supplied for your Linux distribution. If you use Ubuntu, [this](#) is one of many starting points. If you use windows, you could install R from the official [CRAN](#) web page.
- R Studio — we recommend to use R Studio for (interactive) programming with R. You can download R Studio from the official [web page](#).
- your combined world bank data sets computed within [W05-1: Melting a data set](#)
- [this](#) csv data set which matches world bank country acronyms to regions according to this [overview](#).

Learning log assignments

😄 As always, please add these entries to your today's learning log at [teachwiki](#):

- Favorite aspect of the session (if any)
- Superfluous aspect of the session (if any)
- Eureka effect (if any)
- Links to what I've learned so far (if any)
- Questions (if any)

For more information see this short [howto](#).

As today's special, please complete the following assignment:

As part of [W06-1: Visualizing data](#) you have already combined your world bank data with a country-to-region list and produced the following visualizations:

1. Histogram for the original CO₂ and GNI values
2. Histogram for the transformed CO₂ and GNI values
3. Box whisker plots to summarize the CO₂ and GNI value distributions for each world bank region
4. Scatter plot between all CO₂ and GNI values

Let's pimp them a little bit so they become to a publication quality level. Please rework your code from [W06-1](#) so that only the following graphics are computed (instead of the ones which are in your code from last session right now):

😄 Please combine all four histogram plots (original, transformed values) into one visualization page and make sure that the reader knows which kind of transformation has been performed although the original data values should still be directly extractable from all four histograms (i.e. does not have transformed values on the x axis of the transformed data set).

😬 Please combine both box whisker plots (which should be computed based on the transformed values) in one plot and also use different colors for the individual variables. Make sure that the original data values are printed on the y axis. The box whisker plots should not be differentiated by the regions this time (i.e. just one box for GNI, one for CO2).

For the next task, the plot from last session should not only be pimped but it should be extended by a linear regression line. You can use the `lm()` function to compute a linear regression between two (or more) variables. Make sure you use the same variables as used in the plot because then you can use the `regLine()` function of the “car” library to add the regression line like this:

```
linear_model <- lm(...) # your linear model
plot(...) # your scatter plot
regLine(linear_model)
```

😬 Please make a scatter plot which shows the two transformed variables and add a colored linear regression line to it. Again, make sure that the original data values are printed on each of the axis.

😬 After you have finished your code, please upload it to your learning log along with the produced graphics.

To store graphics, you can use some code like this:

```
my_plot <- plot(...)
png() # open a device for writing to PNG format; file is written to
working directory
plot(my_plot) # write the plot to the PNG file
dev.off() # close the device
```

If you want to save another visualization, you have to open, plot and close the device again!

Although it is not noted explicitly above, publication quality plots must have meaningful titles, axis labels etc. so the reader knows what he is looking at (without looking into figure captions).

From:
<http://moc.environmentalinformatics-marburg.de/> - Marburg Open Courseware

Permanent link:
<http://moc.environmentalinformatics-marburg.de/doku.php?id=courses:msc:data-management:worksheets:dm-ws-07-1>

Last update: 2014/12/11 06:52

