Intro to Social Science Data Analysis

Seminar 6: Data Visualisation in R

Christopher Gandrud

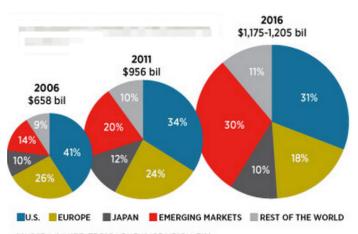
October 11, 2012

Principles of Graphical Excellence

What does it mean to "show the data"?

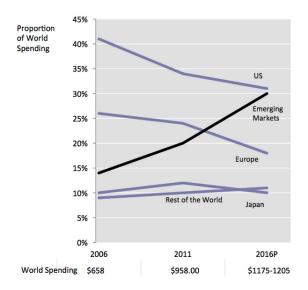
What is a high data-ink ratio?

What's the story? How well does the graph help you see it? How could it be better?



SOURCE: IMS INSTITUTE FOR HEALTHCARE INFORMATICS.

Better



Source: http://junkcharts.typepad.com/junk_charts/2012/10/expanding-circles-of-error.html

Base R Graphics

Make a graph in base R (using any data).

Make sure the figure is labeled well.

With ggplot2 make:

- A scatter plot,
- A density plot
- ► A line graph
- ► A bar graph

With ggplot2 make:

- ► A scatter plot,
- A density plot
- ► A line graph
- ► A bar graph

With ggplot2 make:

- A scatter plot,
- A density plot
- ► A line graph
- ► A bar graph

With ggplot2 make:

- A scatter plot,
- ► A density plot
- ► A line graph
- ► A bar graph

Maps

Make a map with googleVis.

Assignment 2

Due: Friday 19 October

Describe at least **3** variables in a data set.

You need to select a **range of descriptive statistical tools**. The tools should include both **numerical descriptive statistics** and **graphics**.

These tools should describe the variables':

- central tendency,
- variation,
- their relationships with the other variables.

The descriptions need to be discussed **in paragraph form**.

The description must be **reproducible**. So you should email me the link to a Dropbox folder with:

- the .csv data set,
- ▶ the .Rmd R markdown file,
- ▶ the final .html file.