

# Overview of visualization with R

Amy F. Szczepański, Remote Data Analysis and Visualization Center, University of Tennessee  
aszczepa@utk.edu    <http://rdav.nics.tennessee.edu/>    [help@xsede.org](mailto:help@xsede.org)

## Overview of today's session

1. Introduction to visualization
2. Visualization in R with **ggplot2**. Example: `ggplot2-diamonds.R` with the built-in data set **diamonds**. Working with layers, such as geometries, aesthetics, and scales. Customizing plots. Saving plots.
3. Complementary tools for visualization. Sometimes even a Swiss army knife is not enough to solve your problems.
4. Map visualization with **ggmap**. Examples: `earthquake-map.R` and `uber-cars.R` with the data sets `quakes2.csv` and `all.tsv`. The latter can be downloaded from <http://www.infochimps.com/datasets/uber-anonymized-gps-logs>. Integrating the features of **ggplot2** with a mapping layer.
5. Specialized visualization tools. Making a motion chart with Google Vis. Example: `google-motion.R` with the data set `euData.tsv`.
6. Specialized visualization tools. Plotting graphs and networks with **igraph**. Example: `facebook.R` with the data set `3980.edges`.
7. Conclusions

## Other software mentioned

**Inkscape**: Free vector graphics editor.

**ImageMagick**: Free and highly scriptable tools for working with image formats and other tasks.

**ffmpeg**: Convert video formats; convert a sequence of frames into a video.

**Sweave**: Paradigm for integrating  $\text{\LaTeX}$  and R

## Exercises

1. Create a visualization that allows the viewer to make a hypothesis about the relationship between the cut and the clarity of a diamond.
2. Make a map with the locations of all of the sites for this workshop. Hint: the function `geocode("university of tennessee")` will return the latitude and longitude of UT.
3. Using the tool of your choice, parse the date/time field of the car service data. Create a visualization that shows the difference in car service usage in:
  - (a) Day versus night.
  - (b) Weekday versus weekend.
4. Can you find an example of a data set that is well-suited to a Motion Chart? Bonus points if it is unrelated to the application that Hans Roling used in his TED talk.
5. Download a Twitter data set from SNAP and create a directed graph of one of the ego networks.

## References

- [1] D. Borland and R.M. Taylor. Rainbow color map (still) considered harmful. *Computer Graphics and Applications, IEEE*, 27(2):14–17, 2007.
- [2] Gabor Csardi and Tamas Nepusz. The igraph software package for complex network research. *InterJournal, Complex Systems*:1695, 2006.
- [3] P. Dalgaard. *Introductory Statistics With R*. Statistics and Computing. Springer-Verlag, 2002.
- [4] Trafton Drew, Karla Evans, Melissa L. H. Võ, Francine L. Jacobson, and Jeremy M. Wolfe. Informatics in radiology: What can you see in a single glance and how might this guide visual search in medical images? *Radiographics*, 33(1):263–274, 2013.
- [5] Markus Gesmann and Diego de Castillo. googleVis: Interface between R and the Google Visualisation API. *The R Journal*, 3(2):40–44, December 2011.
- [6] David Kahle and Hadley Wickham. *ggmap: A package for spatial visualization with Google Maps and OpenStreetMap*, 2013. R package version 2.3.
- [7] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2013.
- [8] Matthew Ward, Georges Grinstein, and Daniel A. Keim. *Interactive Data Visualization: Foundations, Techniques, and Application*. A.K. Peters, Ltd, May 2010.
- [9] Hadley Wickham. *ggplot2: elegant graphics for data analysis*. Springer New York, 2009.
- [10] Hadley Wickham. *scales: Scale functions for graphics.*, 2012. R package version 0.2.3.
- [11] N. Yau. *Visualize This: The FlowingData Guide to Design, Visualization, and Statistics*. Wiley, 2011.