tmap: An R Package for thematic maps

by Martijn Tennekes

Abstract A thematic map is a geographical map in which statistical data are visualized. The theme refers to the statistical phenomena that is shown, such as the unemployment rate at municipal level. The best known thematic map type is the choropleth, where regions are coloured according to a statistical variable, for instance unemployment rate. Another popular thematic map type is the bubble map, in which the sizes of the bubbles are defined by a statistical variable, for instance metropolitan population. With the **tmap** package, thematic maps can be generated with great flexibility. A thematic map is created by stacking layers, for instance one for colouring municipalities, one for thick borders of higher level regions, and one for text labels. The standard work flow that is needed to create a thematic map is embedded in tmap by several convenient functions for reading, appending, and transforming spatial data.

Introduction

Visualization is key in data science. Without looking at data, it is difficult to know the data, to unveil anomalies, and, moreover, to extract valuable knowledge (Tufte, 1983). Software tools to visually explore, analyse, and present data should therefore belong to any data scientist's toolkit. The R language and its packages contain many functions to craft elegant and clarifying graphics, most notably **ggplot2** (Wickham, 2009). Although it is also possible to achieve great looking geospatial visualizations, it is often more complicated and requires the usage of several R pacakges together. Therefore, I introduce the **tmap** package by which thematic maps can be created in a flexible and layer-based way.

Thematic maps are geographic maps in which spatial data distributions are shown, such as population density. Although they are widely used for publication purposes, which is no surprise due to their visual appeal and recognizability, they also proved to be succesful for discovery and exploration of spatial data (Friendly, 1995). One of the most illustrative examples is the dot map that physician John Show used to locate the source of the cholera outbreak in London in 1854 (Snow, 1855). Roughly, there are five types of thematic maps:

choropleth Cultural areas are colored according to a variable

proportional symbol map Symbols are scaled in proportion to a variable. The best know variant is the bubble map

dot map Individual data points are positioned on the map, like all cholera occurences in John Snow's map.

isopleth Regions are defined by homogeneity of a quantitative variable. For instance altitude or wheather maps.

cartogram The geographic maps is distorted. Regions are sized according to a quantitative variable.

The aim of the **tmap** package is to provide R users an elegant and flexible way of making thematic maps, with a minimum of code required. The syntax resembles the syntax of **ggplot2**, and follows the layered grammar of graphics (Wickham, 2010) broadly.

Cartograms are not supported yet. Also, isopleths cannot be made in a straightforward manner with **tmap**. Typically the **raster** (Hijmans, 2015) package provides excellent options for this.

An example

A thematic map of the world that has been created with **tmap** is depicted in Figure 1. It shows the relation between level of income per country and the distribution of emerging metropolitan areas. While the metropolitan areas in high income countries have a annual growth rate that is less than two percent (notice that the world population increases with 2.7 percent), metropolitan areas in lower income countries grow very fast, especially in Asia but also western and middle Africa.

The code required to create this map is the following.

```
#load spatial objects contained in tmap
data(World)
data(metro)
```
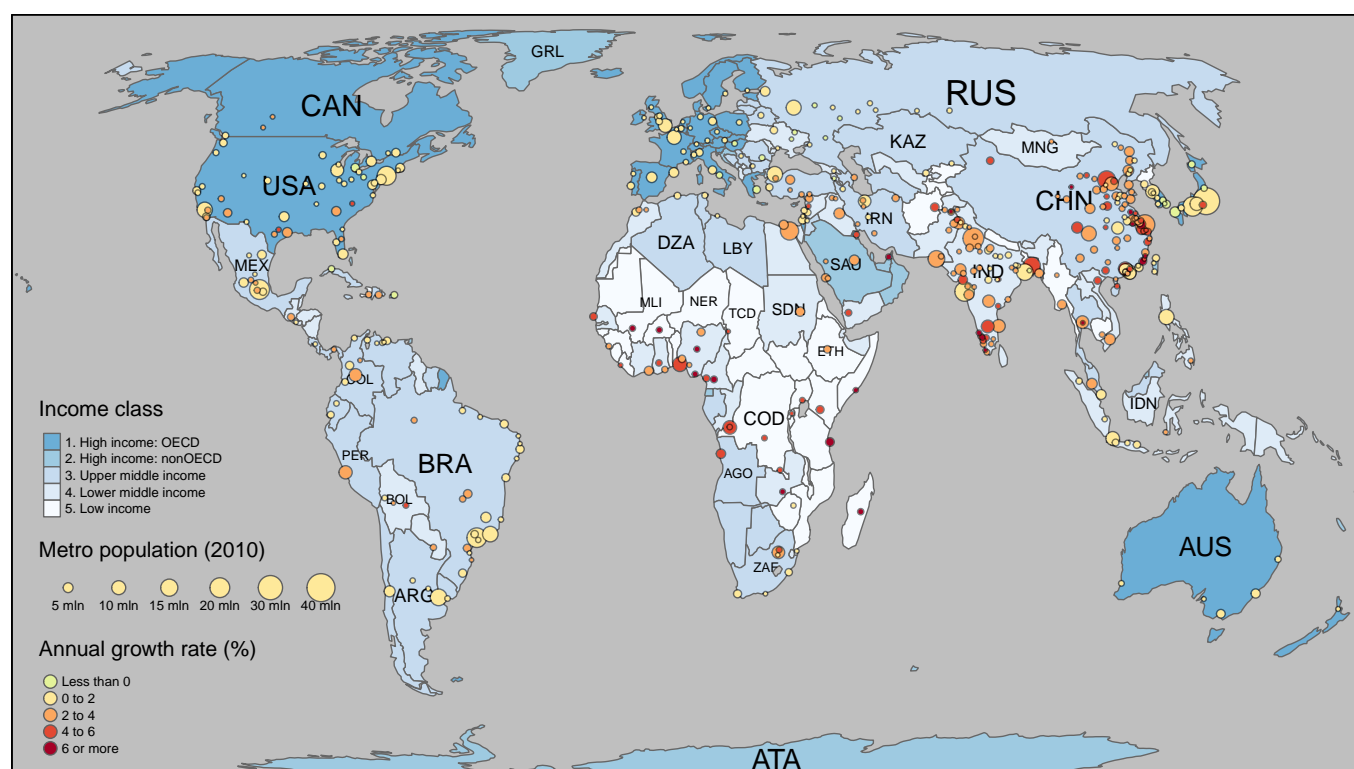
Figure 1: World map about income and urbanization.

```
#derive new variable
metro$growth <- (metro$X2020 - metro$X2010) / (metro$X2010 * 10) * 100

#plot
tm_shape(World) +
  tm_fill("income_grp", style="kmeans", palette="-Blues") +
  tm_borders() +
  tm_text("iso_a3", size="AREA", scale=1.5) +
tm_shape(metro) +
  tm_bubbles("X2010", col = "growth", border.col = "black",
    border.alpha = .5, style="fixed", breaks=c(-Inf, 0, 2, 4, 6, Inf) ,
    palette="-RdYlGn", contrast=1) +
tm_layout_World(title="", legend.titles=c(fill="Income class",
  bubble.size="Metro population (2010)", bubble.col="Annual growth rate (%)"))
```

The `#plot` part of the code is exampled in detail in section 2.4. In a nutshell, two groups of layers are drawn. The first group defines a choropleth that uses the `SpatialPolygonsDataFrame` object `World`. The drawing layers define the fill colors of polygons (`tm_fill`), their borders (`tm_borders`), and the supplementary text labels (`tm_text`). The second group defines a bubble map based on the `SpatialPointsDataFrame` `metro`. The only layer consist of the bubbles created by `tm_bubbles`.

Related R-packages

The **tmap** package stands on the shoulders of the following six great packages.

sp (Pebesma, 2005; Roger S. Bivand, 2013) This package contains classes and methods for spatial data.

rgdal (Bivand et al., 2014) Reading and writing ESRI shape files is done by this package, as well as map projections via the PROJ.4 library (Warmerdam and Evenden, 2012).

rgeos (Bivand and Rundel, 2014) All geometric processing is carried out by this package.

classInt (Bivand, 2015) This package facilitates class intervals, which are used in choropleths of quantitative variables.

RColorBrewer (Bivand, 2015) All used color schemes have been proposed by Brewer et al. (2003) and
implemented in this package.

grid This package provides all basic plotting functions.

Syntax

Another section

There will likely be several sections, perhaps including code snippets, such as:

Summary

This file is only a basic article template. For full details of *The R Journal* style and information on how
to prepare your article for submission, see the Instructions for Authors.

Martijn Tennekes
Statistics Netherlands
CBS-Weg 11, 6412 EX Heerlen
Netherlands
mtennekes@gmail.com

Bibliography

R. Bivand. *classInt: Choose univariate class intervals*, 2015. URL http://CRAN.R-project.org/package=
classInt. R package version 0.1-22. [p2, 3]

R. Bivand and C. Rundel. *rgeos: Interface to Geometry Engine - Open Source (GEOS)*, 2014. URL
http://CRAN.R-project.org/package=rgeos. R package version 0.3-8. [p2]

R. Bivand, T. Keitt, and B. Rowlingson. *rgdal: Bindings for the Geospatial Data Abstraction Library*, 2014.
URL http://CRAN.R-project.org/package=rgdal. R package version 0.9-1. [p2]

C. A. Brewer, G. W. Hatchard, and M. A. Harrower. ColorBrewer in print: A catalog of color schemes
for maps. *Cartography and Geographic Information Science*, 30(1):5–32, 2003. [p3]

M. Friendly. Milestones in the history of thematic cartography, statistical graphics, and data visu-
alization. In *13TH INTERNATIONAL CONFERENCE ON DATABASE AND EXPERT SYSTEMS
APPLICATIONS (DEXA 2002), AIX EN PROVENCE*, pages 59–66. Press, 1995. [p1]

R. J. Hijmans. *raster: Geographic Data Analysis and Modeling*, 2015. URL http://CRAN.R-project.org/
package=raster. R package version 2.3-33. [p1]

R. B. Pebesma, E.J. Classes and methods for spatial data in r. R News 5 (2), 2005. URL http://cran.r-
project.org/doc/Rnews/. [p2]

V. G.-R. Roger S. Bivand, Edzer Pebesma. *Applied spatial data analysis with R*. Springer, NY, second
edition edition, 2013. URL http://www.asdar-book.org/. [p2]

J. Snow. On the mode of communication of cholera. London, 1855. [p1]

E. R. Tufte. *The visual display of quantitative information*. Graphics Press, 1983. [p1]

F. Warmerdam and G. Evenden. Proj.4 - cartographic projections library, version 4.8.0, 2012. URL
http://trac.osgeo.org/proj/. [p2]

H. Wickham. *ggplot2: elegant graphics for data analysis*. Springer New York, 2009. ISBN 978-0-387-98140-6.
URL http://had.co.nz/ggplot2/book. [p1]

H. Wickham. A layered grammar of graphics. *Journal of Computational and Graphical Statistics*, 19(1):
3–28, 2010. [p1]