Shapefile Modification in R as the Basis for Linked Micromap Plots for New Geographic Regions

Jürgen Symanzik
Utah State University, Logan, UT

symanzik@math.usu.edu
http://www.math.usu.edu/~symanzik

--- with contributions from ---

Jeong Yong Ahn, Se-jin Park, Marc Weber, Michael McManus, Quinn Payton, XiaoTian Dai, Shuming Bao, Miao Shui & Bing She

Interface 2015, Morgantown, WV – Friday, June 12, 2015
Contents

- Linked Micromap (LM) Plots
  - Overview
  - Regional Micromaps

- Shapefile Modifications
  - Examples
  - Limitations

- Outlook and Future Work
Micromaps

- Link of row-labeled univariate (or multivariate) statistical summaries to corresponding geographical region
- Focus on statistical display and not on maps
- Useful for
  - environmental data
  - agricultural data
  - medical data
  - public health data
  - economical data

History of Micromaps

- First presented at 1996 American Statistical Association’s annual meeting (Olsen, Carr, Courbois, Pierson)

- Initial references:
  - Carr, Olsen, Courbois, Pierson, Carr (1998) Linked Micromap Plots …, SCSG, Vol. 9, No.1

Micromaps in R

- Two recently developed R Packages:
  - “micromap” (by Payton, Olsen, Weber, McManus, Kincaid): general purpose
  - “micromapST” (by Carr, Pearson, Pickle): focus on the 50 U.S. states
Regional Micromaps

- Micromaps that are related to subregions within a particular country

- Previously:
  - National level: United States and Korea
  - Provincial level: United States and France

- Recent developments:
  - National level: most countries of South America, Korea (advanced), and China
  - Provincial level: most provinces of Korea
Regional Micromap Construction

- Modification of existing shapefiles
  - Simplification of boundaries (via Douglas-Peucker 1973 algorithm) [previously solved]
  - Removal of tiny islands [previously solved]
  - Enlarge small areas such as capital regions, e.g., Washington DC in the United States
  - Shift and resize (enlarge or shrink) regions that are far away from the main area, e.g., Alaska or Hawaii in the United States
  - New R code created for the remaining tasks
Shapefile Sources

GADM database of Global Administrative Areas

GADM is a spatial database of the location of the world's administrative areas (or administrative boundaries) for use in GIS and similar software. Administrative areas in this database are countries and lower level subdivisions such as provinces, departments, bibhag, bundeslander, daerah istimewa, fivondronana, krong, landsvæðun, opština, sous-préfectures, counties, and thana. GADM describes where these administrative areas are (the "spatial features"), and for each area it provides some attributes, such as the name and variant names.

The current version is 2.0 (January 2012)

Source: http://www.gadm.org
New R function `EnlargeReplaceMapArea` heavily builds on R functions from `rgeos`, `rgdal`, `maptools`:

- `unionSpatialPolygons`: Aggregates polygons in a `SpatialPolygons` object
- `gBuffer`: Expands the given geometry
- `gIntersection`: Determines the intersection between two given geometries
- `gDifference`: Determines the difference between two given geometries

**Call:**
```
ChinaShapeRefined <- EnlargeReplaceMapArea("Beijing", "Hebei", ChinaShapefileThin, "ename", "gbcode", 50000)
```
Shapefile Modification (2)

- New R function **ShiftArea** modifies the slots in the spatial object that contain latitude and longitude, based on
  - slot: Returns or sets information about the individual slots in an object

- Call:

```r
KorShape1Sub28 <- ShiftArea(KorShape1Sub28, "Ongjin-gun", 1, "SIG_ENG_NM", c(30000, 30000, 120000, 120000))
KorShape1Sub28 <- ShiftArea(KorShape1Sub28, "Ongjin-gun", 2, "SIG_ENG_NM", c(0, 0, -20000, -20000))
```
Brazil
Simplification

- UL: Raw
- UR: Thinned boundaries
- LL: Capital & other regions enlarged
- LR: Capital region enlarged & shifted

Brazil Example

### Korea Example (raw)

<table>
<thead>
<tr>
<th>Province</th>
<th>1997 Population</th>
<th>2013 Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seoul</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gyeonggi-do</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Busan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gyeongsangnam-do</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gyeongsangbuk-do</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daegu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incheon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jeollanam-do</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jeollabuk-do</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chungcheongnam-do</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gangwon-do</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chungcheongbuk-do</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gwangju</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daejeon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ulsan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jeju-do</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sejong-si</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Maps**

- Median

- Maps
<table>
<thead>
<tr>
<th>Province</th>
<th>1997 Population</th>
<th>2013 Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seoul</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gyeonggi-do</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Busan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gyeongsangnam-do</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gyeongsangbuk-do</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daegu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incheon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jeollanam-do</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jeollabuk-do</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chungcheongnam-do</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gangwon-do</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chungcheongbuk-do</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gwangju</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daejeon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ulsan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jeju-do</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sejong-si</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Maps

Median
Korea / Incheon Example (template - raw)
Korea / Incheon Example (template - refined)
China Simplification

Raw

Thinned boundaries

Enlarged regions
China Example (1)
China Example (2)
Limitation (1): Japan
Limitation (1): Japan
Limitation (2): Korea / Gyeonggi
Limitation

(2) / Solution

Outline
Current Work:
Interactive Link: Micromaps / Religion Explorer
Software at the China Data Center (CDC),
University of Michigan
Future Work (1)

- Reintroduce 2-column identifier layout to micromap/micromapST R packages, based on Carr (2001)

Figure Source:
Future Work (2)

- Further extend new R functions to enlarge an area into several neighboring areas simultaneously and into the open sea.
- Further enhance, test, and debug new R functions for the creation of modified shapefiles.
- Add functionality to micromap R package once code is stable or release as stand-alone R package.
Questions ???
— or —
send e-mail to:

symanzik@math.usu.edu