Washington State Medicaid Dental Expenses

Brian High

05/21/2015

Geographical Heatmap

This serves as an example of making a geographical "heatmap" (choropleth). In particular, a US state map with counties colored by the value of a variable. We will use the average annual expenses of Medicaid dental care per Medicaid user averaged by county. this project was inspired by a related example.

Data Sources

Medicaid data are from Washington Health Care Authority

Spatial data are from: GADM

Some Recent Research

If you are interested in more information about research on Medicaid dental expenses, see:

- Churchill SS, Williams BJ, Villareale NL. Characteristics of Publicly Insured Children with High Dental Expenses. Journal of Public Health Dentistry. 2007 Fall;67(4):199-207. http://www.ncbi.nlm.nih.gov/pubmed/18087990
- Bussma Ahmed Bugis, "Early Childhood Caries and the Impact of Current U.S. Medicaid Program: An Overview," International Journal of Dentistry, vol. 2012, Article ID 348237, 7 pages, 2012. doi:10.1155/2012/348237 http://www.hindawi.com/journals/ijd/2012/348237/

Setup

Java is required for package "XLConnect". Make sure Java is installed.

```
if (system2("java","-version")) {
    stop("Java not found. Install Java first. https://java.com/en/download/")
}
```

Load the required R packages.

```
for (pkg in c("knitr", "RCurl", "XLConnect", "rgeos", "maptools", "ggplot2", "scales",
              "gridExtra")) {
    if (! suppressWarnings(require(pkg, character.only=TRUE)) ) {
        install.packages(pkg, repos="http://cran.fhcrc.org", dependencies=TRUE)
        if (! suppressWarnings(require(pkg, character.only=TRUE)) ) {
            stop(paste0(c("Can't load package: ", pkg, "!"), collapse = ""))
        }
    }
}
```

Configure `knitr` options.

```
opts_chunk$set(tidy=FALSE, cache=TRUE)
```

Create the data folder, if necessary.

```
datadir <- "data"
dir.create(file.path(datadir), showWarnings=FALSE, recursive=TRUE)
```

Import data

Get the shapefile zip

```
shapefile.zip <- "USA_adm.zip"
shapefile.zip.path <- paste0(c(datadir, "/", shapefile.zip), collapse='')
shapefile.zip.url <- "http://biogeo.ucdavis.edu/data/gadm2/shp/USA_adm.zip"
if (! file.exists(shapefile.zip.path)) {
            print("Downloading data file...")
            download.file(url=shapefile.zip.url, destfile=shapefile.zip.path)
}
```

Extract the shapefile zip

```
shapefile <- "USA_adm2"
shapefile.path <- paste0(c(datadir, "/", shapefile, ".shp"), collapse='')
if (file.exists(shapefile.zip.path)) {
    if (! file.exists(shapefile.path)) {
            print("Unzipping data file...")
            unzip(zipfile=shapefile.zip.path, overwrite=TRUE, exdir=datadir)
    }
} else {
    stop(paste("Can't find", shapefile.zip.path, "!", sep=" "))
}

if (! file.exists(shapefile.path)) {
    stop(paste("Can't find", shapefile.path, "!", sep=" "))
}
```

Select WA state map data

```
usa <- readShapeSpatial(paste0(c(datadir, "/", shapefile), collapse=''))
wa <- usa[usa$NAME_1=="Washington", ]
```

Get county names and locations

These are the names and coordinates for the county name labels.

```
cnames.path <- paste0(c(datadir, "/cnames.csv"), collapse='')
cnames.url <- "https://github.com/brianhigh/wa-water-quality/blob/master/data/cnames.csv"
if (! file.exists(cnames.path)) {
    print("Downloading data file...")
    download.file(url=cnames.url, destfile=cnames.path,  mode="wb", method="curl")
}

if (file.exists(cnames.path)) {
    cnames <- read.csv(file = cnames.path, header = TRUE)
} else {
    stop(paste("Can't read", cnames.path, "!", sep=" "))
}
```

Get dental expenses

```
dentfile.path <- paste0(c(datadir, "/wa_hca_dental_summary.xls"), collapse='')
dent.url <- "http://www.hca.wa.gov/medicaid/dentalproviders/documents/999cntysumall.XLS"
if (! file.exists(dentfile.path)) {
    print("Downloading data file...")
    download.file(url=dent.url, destfile=dentfile.path,  mode="wb")
}

if (! file.exists(dentfile.path)) {
    stop(paste("Can't find", dentfile.path, "!", sep=" "))
}

# Read worksheet from Excel workbook twice - once for each column of interest
dent.cnty <- readWorksheetFromFile(dentfile.path, sheet=1, header=FALSE,
                                   startRow=5, endRow=43, startCol=1, endCol=1)
dent.exp <- readWorksheetFromFile(dentfile.path, sheet=1, header=FALSE,
                                  startRow=5, endRow=43, startCol=25, endCol=25)
expenses <- data.frame(county=dent.cnty$Col1, FY2014=dent.exp$Col1,
                       stringsAsFactors = FALSE)
```

Create the Map

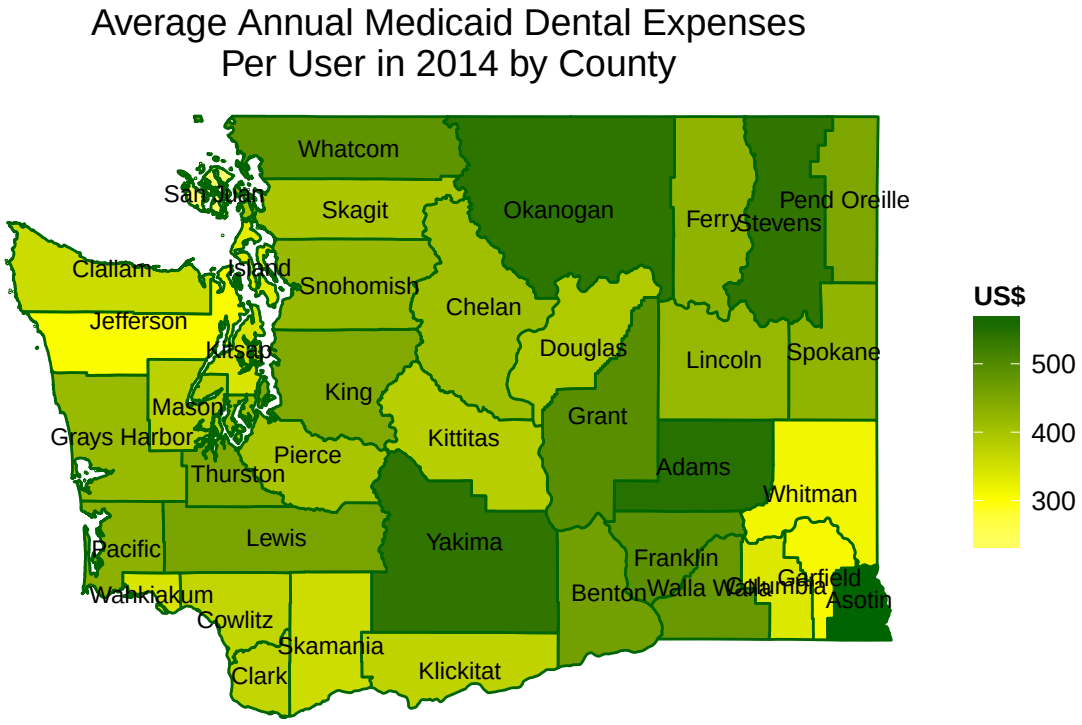
```
# Create a custom theme from theme_classic
theme_bare <- function(...) {
    theme_classic() +
    theme(axis.line=element_blank(),
          axis.text.x=element_blank(),
          axis.text.y=element_blank(),
          axis.ticks=element_blank(),
          axis.title.x=element_blank(),
          axis.title.y=element_blank())
}

# Prepare map data for ggplot
wa <- fortify(wa, region="NAME_2")
```

```
# Map Washington State Medicaid dental expenses by county
gmap <- ggplot() + geom_map(data=expenses, aes(map_id=county, fill=FY2014),
                     color="darkgreen", map=wa) +
    expand_limits(x=wa$long, y=wa$lat) + theme_bare() +
    geom_text(data=cnames, aes(long, lat, label = subregion), size=3) +
    labs(title = "Average Annual Medicaid Dental Expenses\nPer User in 2014 by County") +
    scale_fill_gradient2(space="Lab", high="darkgreen", mid="yellow",
                         low="white", midpoint=300, name="US$")

# Create source attribution string
data.src <- paste0(collapse = ' ', c('Data sources:',
                                     'WA Health Care Authority (www.hca.wa.gov),',
                                     'and GADM (gadm.org)'))

# Layout map with source attribution string at bottom of plot area
gmap <- arrangeGrob(gmap, sub = textGrob(data.src, x=0, hjust=-.1, vjust=0.1,
                                  gp = gpar(fontface="italic", fontsize=10)))
gmap
```



Data sources: WA Health Care Authority (www.hca.wa.gov), and GADM (gadm.org)

Figure 1: