Crime in London

Valeria Cortez

07/10/2017

Visualising crime in London by borough in R

Accessing the data

```
#Reads csv
crime_data <- read.csv("./data/crime_data.csv")
#Changes the name for the headings
names(crime_data) <- c("borough",2009:2013)
#Shows the first 5 entries
head(crime_data,5)
```

```
##                borough  2009  2010  2011 2012 2013
## 1       City of London  18.4  17.7  17.0 14.3 14.0
## 2 Barking and Dagenham 123.1 115.6 111.3 99.6 90.6
## 3               Barnet  86.1  82.8  83.6 78.6 68.7
## 4               Bexley  80.2  70.5  61.9 62.2 59.5
## 5                Brent 107.2 107.4 113.3 96.8 85.9
```

```
#Reads boroughs data as SpatialPolygonsDataFrame
library(rgdal)
```

```
## Warning: package 'rgdal' was built under R version 3.4.2
```

```
## Loading required package: sp
```

```
## rgdal: version: 1.2-11, (SVN revision 676)
##  Geospatial Data Abstraction Library extensions to R successfully loaded
##  Loaded GDAL runtime: GDAL 2.1.3, released 2017/20/01
##  Path to GDAL shared files: /Library/Frameworks/R.framework/Versions/3.4/Resources/library/rgdal/gdal
##  Loaded PROJ.4 runtime: Rel. 4.9.3, 15 August 2016, [PJ_VERSION: 493]
##  Path to PROJ.4 shared files: /Library/Frameworks/R.framework/Versions/3.4/Resources/library/rgdal/pr
##  Linking to sp version: 1.2-5
```

```
london_boroughs <- readOGR(dsn = "data/Map_data/LondonBoroughs.shp")
```

```
## OGR data source with driver: ESRI Shapefile
## Source: "data/Map_data/LondonBoroughs.shp", layer: "LondonBoroughs"
## with 33 features
## It has 8 fields
## Integer64 fields read as strings:  Pop_2001 PopDensity PopDen
```

```
head(london_boroughs@data,2)
```

```
##   ons_label                 name Partic_Per Pop_2001 PopDensity      AREA
## 0      00AF              Bromley       21.7   295535       1961 150664950
## 1      00BD Richmond upon Thames       26.6   172330       2945  58509871
##   PERIMETER PopDen
## 0  67554.41   1962
## 1  47941.02   2945
```

```
#joins london_boroughs@data with crime_data based on the borough names
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
london_boroughs@data <- left_join(london_boroughs@data, crime_data, by = c('name' = 'borough'))
head(london_boroughs@data)
```

```
##   ons_label                 name Partic_Per Pop_2001 PopDensity      AREA
## 1      00AF              Bromley       21.7   295535       1961 150664950
## 2      00BD Richmond upon Thames       26.6   172330       2945  58509871
## 3      00AS           Hillingdon       21.5   243006       2091 116202296
## 4      00AR             Havering       17.9   224262       1950 114999312
## 5      00AX Kingston upon Thames       24.4   147271       3968  37108272
## 6      00BF               Sutton       19.3   179767       4080  44056244
##   PERIMETER PopDen 2009 2010 2011 2012 2013
## 1  67554.41   1962 89.9 82.7 81.3 74.3 70.9
## 2  47941.02   2945 70.9 71.2 72.3 70.2 62.4
## 3  57255.86   2091 80.8 77.5 77.5 65.3 60.4
## 4  58830.85   1950 83.8 85.1 82.8 77.6 70.3
## 5  34461.94   3969 75.4 75.1 74.6 66.3 59.7
## 6  35161.95   4080 80.8 76.3 71.3 66.9 62.0
```

```
#passes the spatial data as a data.frame rather than a spatial object
library(ggplot2)
london_boroughs_f <- fortify(london_boroughs)
```

```
## Regions defined for each Polygons
```

```
# allocate an id variable to the sp data
london_boroughs$id <- row.names(london_boroughs)

# joins the data
london_boroughs_f <- left_join(london_boroughs_f, london_boroughs@data)
```

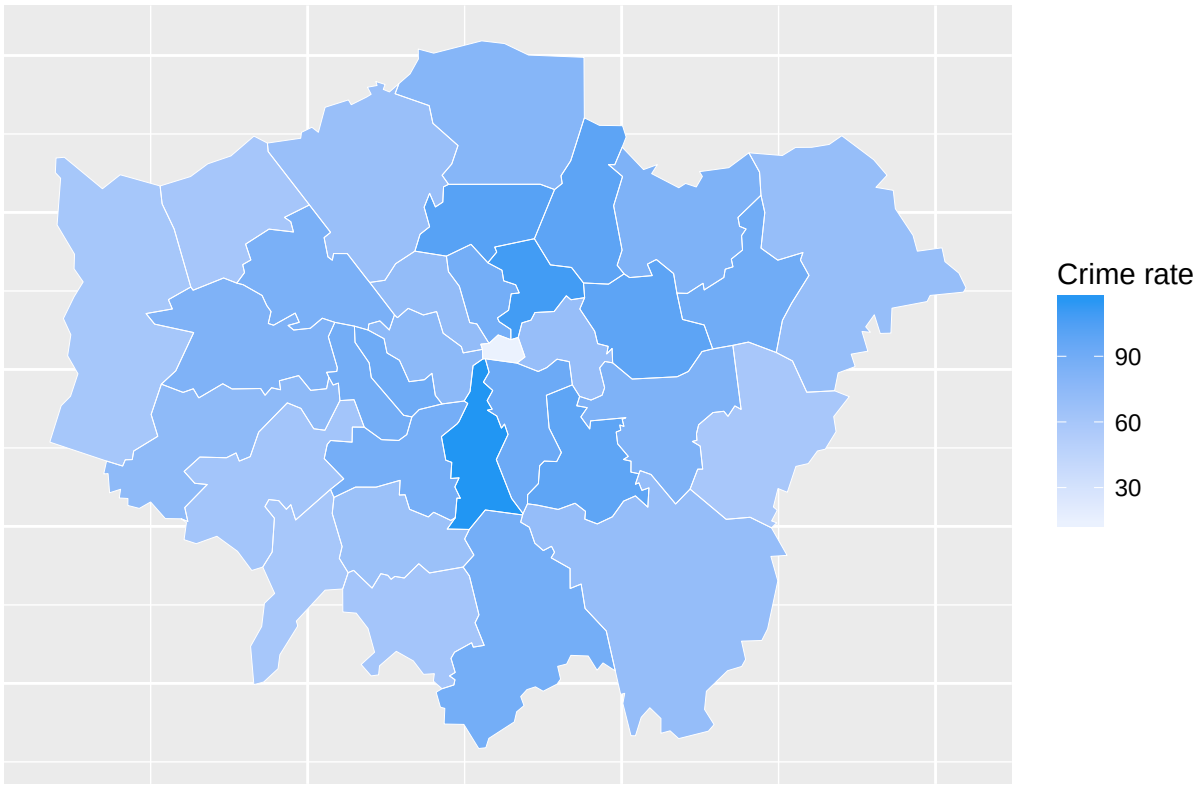
```
## Joining, by = "id"
```

```
head(london_boroughs_f,5)
```

```
##       long      lat order  hole piece id group ons_label    name
## 1 541177.7 173555.7     1 FALSE     1  0   0.1      00AF Bromley
## 2 541872.2 173305.8     2 FALSE     1  0   0.1      00AF Bromley
## 3 543441.5 171429.9     3 FALSE     1  0   0.1      00AF Bromley
## 4 544361.6 172379.2     4 FALSE     1  0   0.1      00AF Bromley
## 5 546662.4 170451.9     5 FALSE     1  0   0.1      00AF Bromley
##   Partic_Per Pop_2001 PopDensity      AREA PERIMETER PopDen 2009 2010 2011
## 1       21.7   295535       1961 150664950  67554.41   1962 89.9 82.7 81.3
## 2       21.7   295535       1961 150664950  67554.41   1962 89.9 82.7 81.3
## 3       21.7   295535       1961 150664950  67554.41   1962 89.9 82.7 81.3
```

```
## 4       21.7   295535       1961 150664950  67554.41   1962 89.9 82.7 81.3
## 5       21.7   295535       1961 150664950  67554.41   1962 89.9 82.7 81.3
##   2012 2013
## 1 74.3 70.9
## 2 74.3 70.9
## 3 74.3 70.9
## 4 74.3 70.9
## 5 74.3 70.9
```

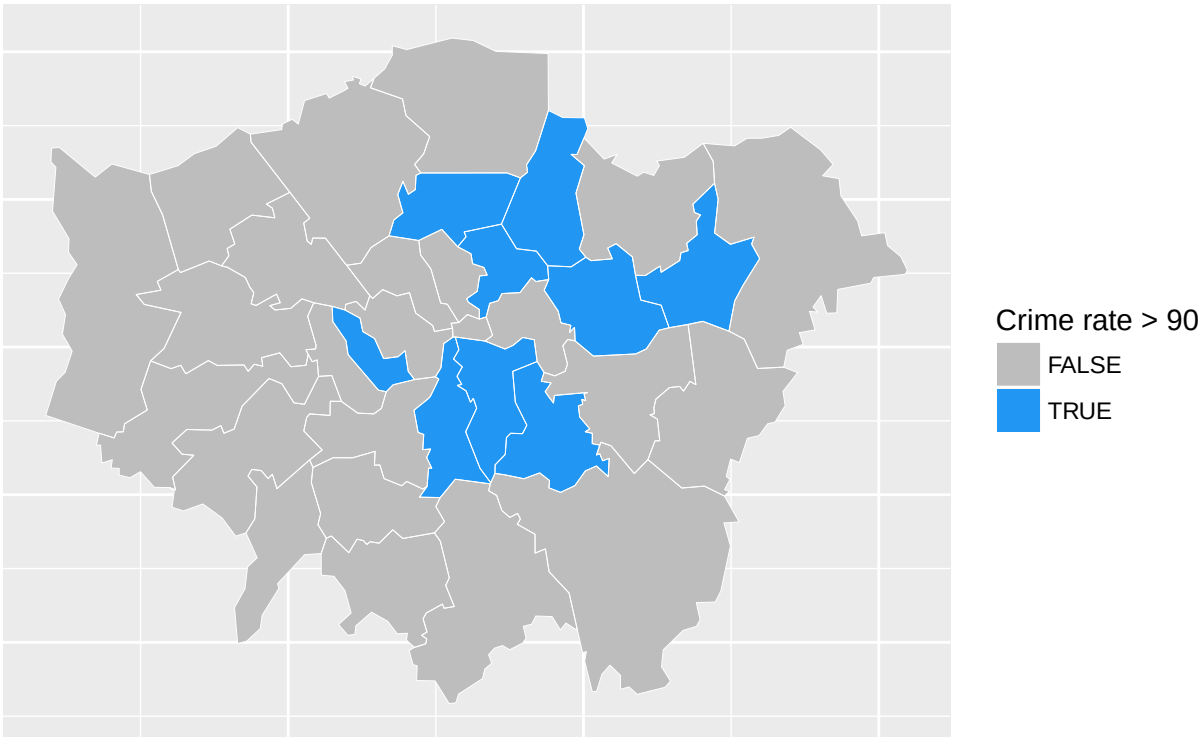
```
#plots the boroughs with highest crime rate of 100
ggplot(london_boroughs_f, aes(long, lat, group = group, fill=`2013`)) +
  geom_polygon() + geom_path(colour="white", lwd=0.05) + coord_equal() +
  labs(x = "lat", y = "lon",
       fill = "Crime rate") +
  scale_fill_gradient2(low = "#BDBDBD",high = "#2196F3", # colors
                       name = "Crime rate") + # legend options
  ggtitle("Crime rate (2013) ") +
  theme(axis.text = element_blank(), # change the theme options
        axis.title = element_blank(), # remove axis titles
        axis.ticks = element_blank()) # remove axis ticks
```




```
#plots the boroughs with highest crime rate of 100
ggplot(london_boroughs_f, aes(long, lat, group = group, fill=`2013`>90)) +
  geom_polygon() + geom_path(colour="white", lwd=0.05) + coord_equal() +
  labs(x = "lat", y = "lon",
       fill = "Crime rate > 90") +
  scale_fill_manual(values=c("#BDBDBD","#2196F3")) +
  ggtitle("Crime rate > 90 (2013) ") +
```

```
  theme(axis.text = element_blank(), # change the theme options
        axis.title = element_blank(), # remove axis titles
        axis.ticks = element_blank()) # remove axis ticks
```

Crime rate > 90 (2013)




```
#changes the structure of the dataset to be able to plot it to a graph
library(tidyr)
london_boroughs_f_long <- gather(london_boroughs_f, crime_rate_year, crime_rate, -long, -lat, -order, -
head(london_boroughs_f_long)
```

```
##       long      lat order  hole piece id group ons_label    name
## 1 541177.7 173555.7     1 FALSE     1  0   0.1      00AF Bromley
## 2 541872.2 173305.8     2 FALSE     1  0   0.1      00AF Bromley
## 3 543441.5 171429.9     3 FALSE     1  0   0.1      00AF Bromley
## 4 544361.6 172379.2     4 FALSE     1  0   0.1      00AF Bromley
## 5 546662.4 170451.9     5 FALSE     1  0   0.1      00AF Bromley
## 6 548187.1 170582.3     6 FALSE     1  0   0.1      00AF Bromley
##   Partic_Per Pop_2001 PopDensity      AREA PERIMETER PopDen
## 1       21.7   295535       1961 150664950  67554.41   1962
## 2       21.7   295535       1961 150664950  67554.41   1962
## 3       21.7   295535       1961 150664950  67554.41   1962
## 4       21.7   295535       1961 150664950  67554.41   1962
## 5       21.7   295535       1961 150664950  67554.41   1962
## 6       21.7   295535       1961 150664950  67554.41   1962
##   crime_rate_year crime_rate
## 1            2009       89.9
## 2            2009       89.9
## 3            2009       89.9
## 4            2009       89.9
## 5            2009       89.9
```

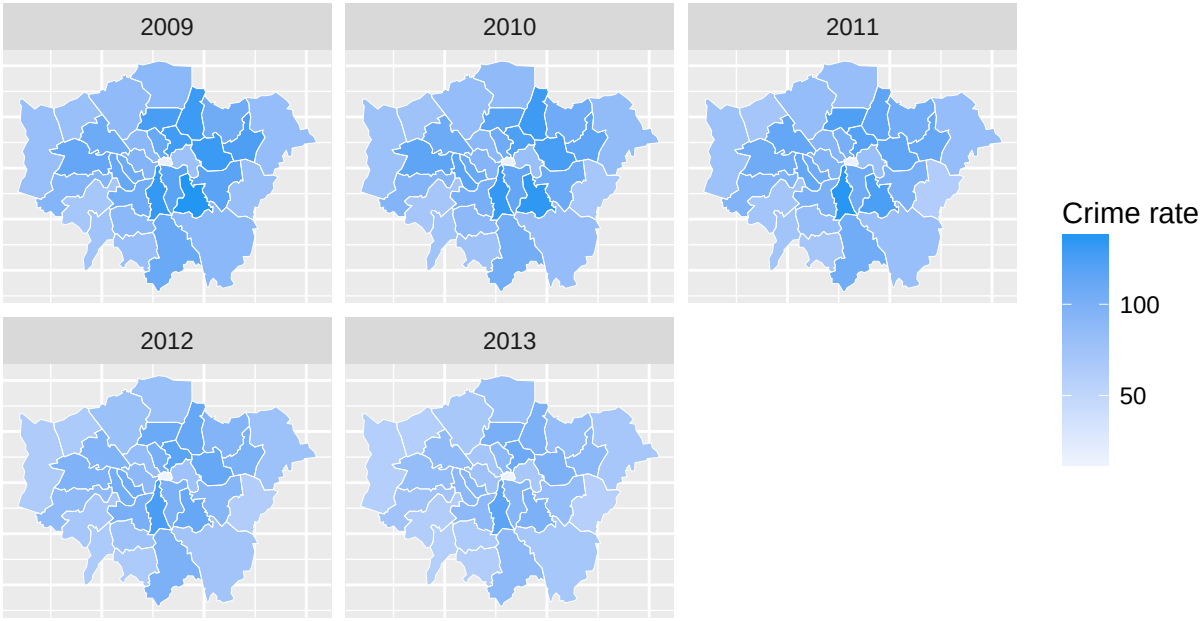


Figure 1: Crime rate by borough from 2009 to 2013

```
## 6            2009       89.9
```

```
ggplot(data = london_boroughs_f_long, # the input data
       aes(x = long, y = lat, fill = crime_rate, group = group)) + # define variables
  geom_polygon() + # plot the boroughs
  geom_path(colour="white", lwd=0.05) + # borough borders
  coord_equal() + # fixed x and y scales
  facet_wrap(~ crime_rate_year) + # one plot per year
  scale_fill_gradient2(low = "#64B5F6",high = "#2196F3", # colors
                       name = "Crime rate") + # legend options
  theme(axis.text = element_blank(), # change the theme options
        axis.title = element_blank(), # remove axis titles
        axis.ticks = element_blank()) # remove axis ticks
```