PLZ Karten

Jan-Philipp Kolb

29 8 2018

PLZ Datensatz einlesen

- **Quelle** für PLZ Shapefiles

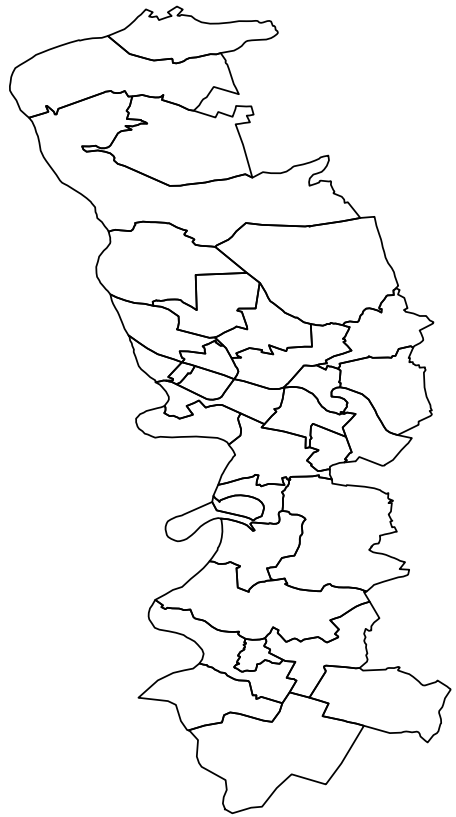
```
library(rgdal)
```

```
setwd(data_path)
plz <- readOGR ("post_pl.shp","post_pl")
```

```
## OGR data source with driver: ESRI Shapefile
## Source: "D:\Daten\Daten\GeoDaten\post_pl.shp", layer: "p
## with 8270 features
## It has 3 fields
```

Die Daten plotten

```
plzbereich <- substr(plz@data$PLZ99,1,2)
plot(plz[plzbereich=="68",])
```



Räumliche Stichprobe

```
set.seed(323)
n <- 100
sampplz <- "68239"
spatsamp <- spsample(plz[plz@data$PLZ99==sampplz,],
                     n,type="random")
```

Reverse Geokodierung

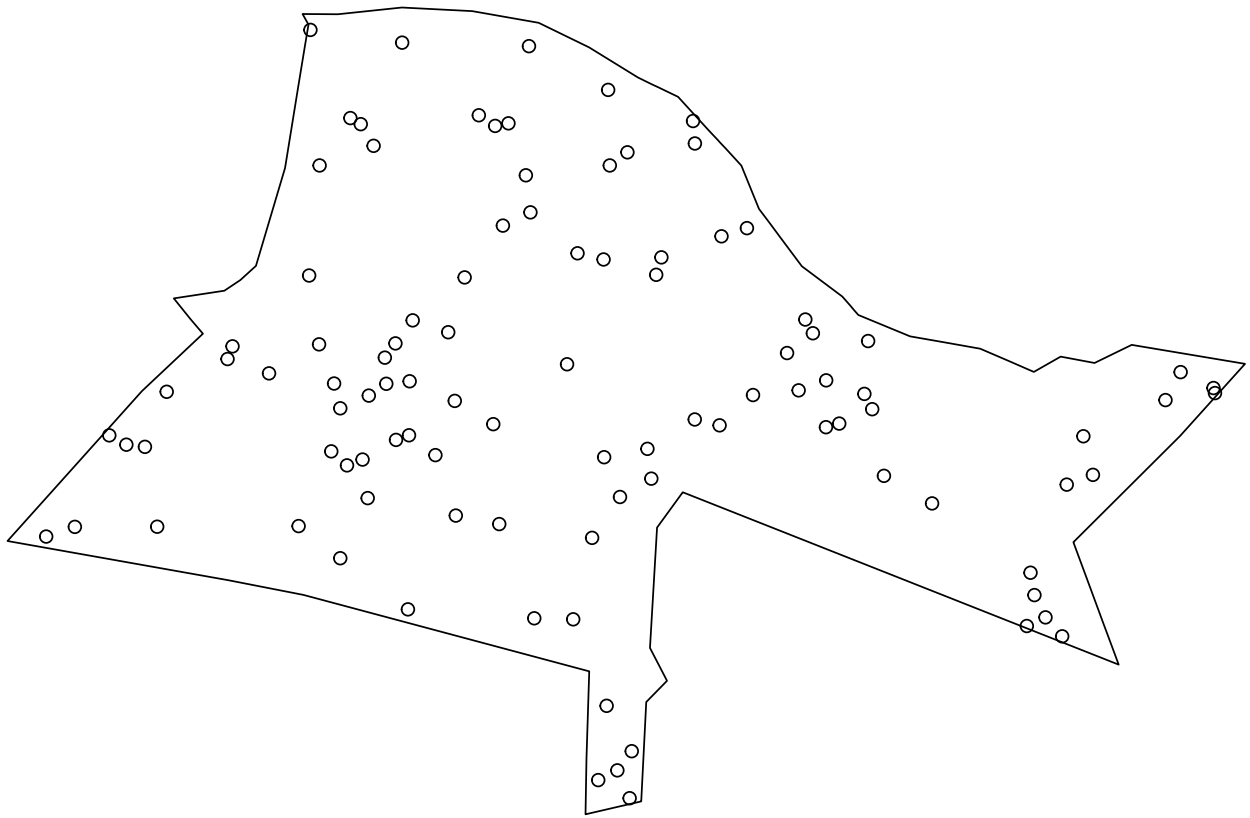
```
library(ggmap)
```

```
spatlist <- list()
for (i in 1:n){
  spatlist[[i]] <- ggmap::revgeocode(c(spatsamp[i,]$x,
                                       spatsamp[i,]$y))
}

spatvec <- unlist(spatlist)
spatsamp$adress <- spatvec
```

Die räumliche Stichprobe plotten

```
plot(plz[plz@data$PLZ99==sampplz,])
points(spatsamp)
```



Nur tatsächliche Adressen

- **Reguläre Ausdrücke** in R

```
addr_list <- spatsamp$adress
  # Adressen raus nehmen, die NA sind
indna <- which(is.na(addr_list))
addr_list <- as.character(addr_list)
addr_list2 <- strsplit(x = addr_list,split = " ")
addr_list2b <- unlist(lapply(addr_list2,length))
ind_ua <- which(addr_list2b<3)
addr_list3 <- unlist(lapply(addr_list2,function(x)x[1]))
  # Adressen rauß nehmen, die Landstraßen
  # oder Autobahnen sind
addr_list3 <- tolower(addr_list3)
ind_str <- grep("^[a-z][1-9]", addr_list3, value = F)

addr_list_t <- addr_list[-c(ind_str,ind_ua,indna)]
```

Das Ergebnis plotten

```
plot(plz[plz@data$PLZ99==sampplz,])
points(spatsamp,pch=20)
points(spatsamp[ind_str,],pch=20,col="green")
points(spatsamp[ind_ua,],pch=20,col="purple")
points(spatsamp[indna,],pch=20,col="red")
```