Dataset Simulation

Humphrey

2022-07-17

Section 1. Simulate SV 1849 and 1854

```
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
```

```
#SIMULATING SV1849 SV1854
set.seed(1)
SV_func<-function(mean_deaths, sd_deaths, n, year, mean_pop, sd_pop){
  SVonly<-round(rnorm(mean = mean_deaths, sd= sd_deaths, n = n), digits = 0)
  SV_district<-paste("SV", seq(1:n),sep = "_")
  year<-rep(year, length(SVonly))
  population<-round(rnorm(mean = mean_pop, sd = sd_pop, n = n ), digits = 0) + SVonly*10
  SV<-data.frame(cbind(deaths = SVonly, district_code = SV_district,
                       year = year, population))
  SV<-SV %>% mutate(deaths = as.numeric(as.character(deaths)),
                    population = as.numeric(as.character(population)))
  return(SV)
}
```

SV1849

```
SV1849<-SV_func(mean_deaths = 120, sd_deaths = 50, n = 20, year = 1849, mean_pop = 8045,sd_pop = 1000)
write.csv(SV1849, paste(here::here(), "SV1849.csv", sep = "/"), row.names = FALSE)
head(SV1849)
```

```
##   deaths district_code year population
## 1     89          SV_1 1849       9854
## 2    129          SV_2 1849      10117
## 3     78          SV_3 1849       8900
## 4    200          SV_4 1849       8056
## 5    136          SV_5 1849      10025
## 6     79          SV_6 1849       8779
```

SV1854

```
SV1854<-SV_func(mean_deaths = 60, sd_deaths = 50, n = 20, year = 1854, mean_pop = 8045,sd_pop = 1000)
write.csv(SV1854, paste(here::here(), "SV1854.csv", sep = "/"), row.names = FALSE)
head(SV1854)
```

```
##   deaths district_code year population
## 1     52          SV_1 1854      10967
## 2     47          SV_2 1854       8476
## 3     95          SV_3 1854       9685
## 4     88          SV_4 1854       8953
## 5     26          SV_5 1854       7562
## 6     25          SV_6 1854       8484
```

Section 2. Simulate Lam 1849 and 1854

```
set.seed(1)
Lam_func<-function(mean_deaths, sd_deaths, n, year, mean_pop, sd_pop) {
  Lamonly<-round(rnorm(mean = mean_deaths, sd= sd_deaths, n = n), digits = 0)
  Lam_district<-paste("Lambeth", seq(1:n), sep = "_")
  year<- rep(year, length(Lamonly))
  population<-round(rnorm(mean = mean_pop, sd = sd_pop, n = n ), digits = 0) + Lamonly*15
  Lam<-data.frame(cbind(deaths = Lamonly, district_code = Lam_district, year = year, population))
  Lam<-Lam %>% mutate(deaths = as.numeric(as.character(deaths)),
                      population = as.numeric(as.character(population)))
  return(Lam)
}
```

Lam1849

```
Lam1849<-Lam_func(mean_deaths = 200, sd_deaths = 50, n = 20, year = 1849, mean_pop = 9045,sd_pop = 1200)
write.csv(Lam1849, paste(here::here(), "Lam1849.csv", sep = "/"), row.names = FALSE)
head(Lam1849)
```

```
##   deaths district_code year population
## 1    169     Lambeth_1 1849      12683
## 2    209     Lambeth_2 1849      13119
## 3    158     Lambeth_3 1849      11504
## 4    280     Lambeth_4 1849      10858
## 5    216     Lambeth_5 1849      13029
## 6    159     Lambeth_6 1849      11363
```

Lam1854

```
Lam1854<-Lam_func(mean_deaths = 40, sd_deaths = 50, n = 20, year = 1854, mean_pop = 9045,sd_pop = 1200)
write.csv(Lam1854, paste(here::here(), "Lam1854.csv", sep = "/"), row.names = FALSE)
head(Lam1854)
```

```
##   deaths district_code year population
## 1     32     Lambeth_1 1854      12407
## 2     27     Lambeth_2 1854       9403
## 3     75     Lambeth_3 1854      10998
## 4     68     Lambeth_4 1854      10099
```

```
## 5      6     Lambeth_5 1854       8243
## 6      5     Lambeth_6 1854       9347
```