ECON 370 Final Project

SD:2, 5, 17, 19, 28

2022-11-8

# ============================================

# Set-up

# ============================================

library(readr)  
library(mosaic)

## Registered S3 method overwritten by 'mosaic':  
## method from   
## fortify.SpatialPolygonsDataFrame ggplot2

##   
## The 'mosaic' package masks several functions from core packages in order to add   
## additional features. The original behavior of these functions should not be affected by this.

##   
## Attaching package: 'mosaic'

## The following objects are masked from 'package:dplyr':  
##   
## count, do, tally

## The following object is masked from 'package:Matrix':  
##   
## mean

## The following object is masked from 'package:ggplot2':  
##   
## stat

## The following objects are masked from 'package:stats':  
##   
## binom.test, cor, cor.test, cov, fivenum, IQR, median, prop.test,  
## quantile, sd, t.test, var

## The following objects are masked from 'package:base':  
##   
## max, mean, min, prod, range, sample, sum

library(data.table)

##   
## Attaching package: 'data.table'

## The following objects are masked from 'package:dplyr':  
##   
## between, first, last

library("xlsx")  
library(openxlsx)

##   
## Attaching package: 'openxlsx'

## The following objects are masked from 'package:xlsx':  
##   
## createWorkbook, loadWorkbook, read.xlsx, saveWorkbook, write.xlsx

library(mvtnorm)  
library(tidyverse)

## ── Attaching packages  
## ───────────────────────────────────────  
## tidyverse 1.3.2 ──

## ✔ tibble 3.1.8 ✔ stringr 1.4.0  
## ✔ tidyr 1.2.0 ✔ forcats 0.5.1  
## ✔ purrr 0.3.4   
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ data.table::between() masks dplyr::between()  
## ✖ mosaic::count() masks dplyr::count()  
## ✖ purrr::cross() masks mosaic::cross()  
## ✖ mosaic::do() masks dplyr::do()  
## ✖ tidyr::expand() masks Matrix::expand()  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ data.table::first() masks dplyr::first()  
## ✖ ggstance::geom\_errorbarh() masks ggplot2::geom\_errorbarh()  
## ✖ dplyr::lag() masks stats::lag()  
## ✖ data.table::last() masks dplyr::last()  
## ✖ tidyr::pack() masks Matrix::pack()  
## ✖ mosaic::stat() masks ggplot2::stat()  
## ✖ mosaic::tally() masks dplyr::tally()  
## ✖ purrr::transpose() masks data.table::transpose()  
## ✖ tidyr::unpack() masks Matrix::unpack()

# ============================================

# Question 1 - Download Data

# ============================================

Note: r eval = FALSE prevents code chunk from running when data is already downloaded (the “switch” in our code). Delete eval = FALSE for first run of code to download files.

for (i in 1998:2010){  
   
url1 <- paste("http://www.nber.org/hcris/265-94/rnl\_rpt265\_94\_", i , ".csv",sep = '')  
url2 <- paste("http://www.nber.org/hcris/265-94/rnl\_nmrc265\_94\_", i, "\_long.csv", sep = '')  
url3 <- paste("http://www.nber.org/hcris/265-94/rnl\_alpha265\_94\_", i, "\_long.csv", sep = '')  
  
destfile <- paste("./hcris\_raw/rnl\_rpt265\_94\_", i , ".csv",sep = '')  
download.file(url1,destfile)  
destfile <- paste("./hcris\_raw/rnl\_nmrc265\_94\_", i, "\_long.csv", sep = '')  
download.file(url2,destfile)  
destfile <- paste("./hcris\_raw/rnl\_alpha265\_94\_", i, "\_long.csv", sep = '')  
download.file(url2,destfile)  
}

# ============================================

# Question 2 - Cleaning The Data

# ============================================

# Read in all CSVs  
for (i in 1998:2010){  
   
rpt\_name = paste("rpt", i, sep = '')  
nmrc\_name = paste("nmrc", i, sep = '')  
alpha\_name = paste("alpha", i, sep = '')  
  
rpt <- read\_csv(paste("hcris\_raw/rnl\_rpt265\_94\_", i , ".csv",sep = ''), show\_col\_types = F)  
nmrc <- read\_csv(paste("./hcris\_raw/rnl\_nmrc265\_94\_", i, "\_long.csv", sep = ''), show\_col\_types = F)  
alpha <- read\_csv(paste("./hcris\_raw/rnl\_alpha265\_94\_", i, "\_long.csv", sep = ''), show\_col\_types = F)  
  
assign(rpt\_name, rpt)  
assign(nmrc\_name, nmrc)  
assign(alpha\_name, alpha)  
}  
  
# clean up environment  
rm(rpt, nmrc, alpha, rpt\_name, nmrc\_name, alpha\_name)

# ============================================

# Question 3 - Analysis

# ============================================