A3

Yifei.Peng

2021/4/13

```r
ppl_ori = read.csv('population.csv')
crm_ori = read.csv('crime_long.csv')
off_ori = read.csv('officers.csv')
```

```r
summary(ppl_ori)
```

```
##     month              period         district        tot_pop
##  Length:4230        Min.   :1.000   Min.   : 1.0   Min.   : 37992
##  Class :character   1st Qu.:1.000   1st Qu.: 6.0   1st Qu.: 72356
##  Mode  :character   Median :1.000   Median :12.0   Median :101879
##                     Mean   :1.923   Mean   :12.6   Mean   :113605
##                     3rd Qu.:3.000   3rd Qu.:19.0   3rd Qu.:139915
##                     Max.   :3.000   Max.   :25.0   Max.   :247373
##    tot_white         tot_black         tot_hisp          p50_inc
##  Min.   :   265   Min.   : 2028    Min.   :   628   Min.   :23961
##  1st Qu.:  3217   1st Qu.:10470    1st Qu.:  2543   1st Qu.:33147
##  Median : 23750   Median :32850    Median : 14123   Median :44900
##  Mean   : 35740   Mean   :36854    Mean   : 33125   Mean   :48803
##  3rd Qu.: 51491   3rd Qu.:61918    3rd Qu.: 55765   3rd Qu.:61849
##  Max.   :150564   Max.   :87288    Max.   :139854   Max.   :93098
```

```r
head(ppl_ori)
```

```
##        month period district tot_pop tot_white tot_black tot_hisp  p50_inc
## 1 2005-01-01      1        1   38472     22608      4953     2543 91084.91
## 2 2005-01-01      1        2   37992       630     35966      628 29890.17
## 3 2005-01-01      1        3   78629      3217     71792     1371 28047.56
## 4 2005-01-01      1        4  124519      9926     77314    35387 39010.22
## 5 2005-01-01      1        5   74384       843     70053     2524 33146.90
## 6 2005-01-01      1        6   89570       309     87288      906 34672.25
```

```r
summary(crm_ori)
```

```
##  crime_month          district        crime_type            crimes
##  Length:36979       Min.   : 1.00   Length:36979       Min.   :    1.0
##  Class :character   1st Qu.: 6.00   Class :character   1st Qu.:   51.0
##  Mode  :character   Median :12.00   Mode  :character   Median :   97.0
##                     Mean   :12.61                      Mean   :  177.5
##                     3rd Qu.:19.00                      3rd Qu.:  206.0
##                     Max.   :25.00                      Max.   : 1500.0
```

```r
#Ex2 data manipulate
ppl = ppl_ori
crm = crm_ori
tot_crm_mon = aggregate(crimes~crime_month,crm,sum)
```

```
head(tot_crm_mon)
```

```
##   crime_month crimes
## 1  2002-01-01  38405
## 2  2002-02-01  33909
## 3  2002-03-01  38583
## 4  2002-04-01  40032
## 5  2002-05-01  42913
## 6  2002-06-01  42834
```

```
tot_crm_mon$crime_month = ymd(tot_crm_mon$crime_month)
plot1 = ggplot(tot_crm_mon, aes(x = crime_month, y = crimes)) + geom_line()+ scale_x_date(date_breaks =
plot1
```
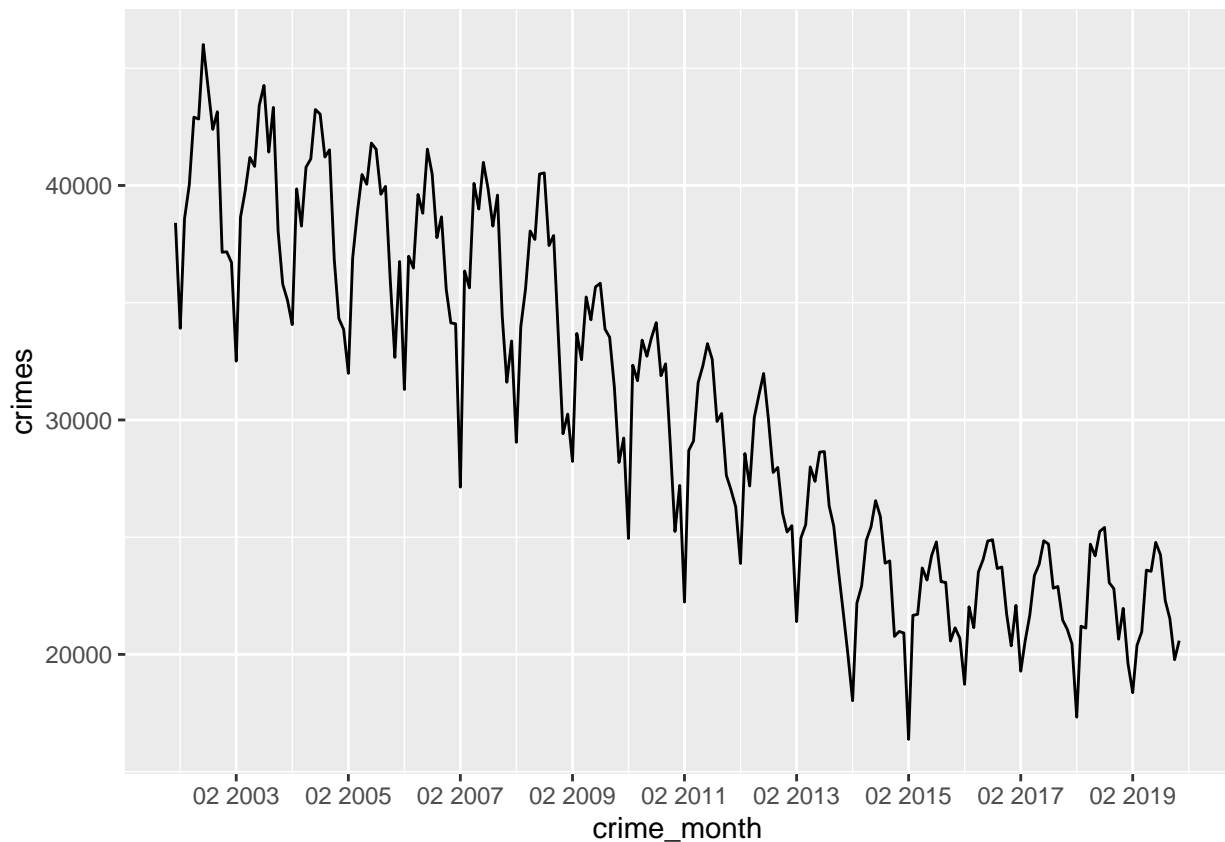


```
# merge
ppl_crm = left_join(ppl, crm, by = c('district'='district', 'month'='crime_month'))
head(ppl_crm,10)
```

```
##          month period district tot_pop tot_white tot_black tot_hisp  p50_inc
## 1   2005-01-01      1        1   38472     22608      4953     2543 91084.91
## 2   2005-01-01      1        1   38472     22608      4953     2543 91084.91
## 3   2005-01-01      1        1   38472     22608      4953     2543 91084.91
## 4   2005-01-01      1        1   38472     22608      4953     2543 91084.91
## 5   2005-01-01      1        1   38472     22608      4953     2543 91084.91
## 6   2005-01-01      1        1   38472     22608      4953     2543 91084.91
## 7   2005-01-01      1        1   38472     22608      4953     2543 91084.91
## 8   2005-01-01      1        1   38472     22608      4953     2543 91084.91
```

```
## 9  2005-01-01       1       2   37992       630      35966       628 29890.17
## 10 2005-01-01       1       2   37992       630      35966       628 29890.17
##     crime_type crimes
## 1         drug      1
## 2         drug    188
## 3        other     62
## 4        other    302
## 5     property    624
## 6     property    160
## 7      violent    150
## 8      violent     62
## 9         drug      1
## 10        drug    208
```

```r
# panel data of unit

# sum up the crime_type
a <- ppl_crm %>%group_by( , month, period, district, tot_pop, tot_white, tot_black, tot_hisp, p50_inc, 
```

```
## `summarise()` has grouped output by 'month', 'period', 'district', 'tot_pop', 'tot_white', 'tot_blac
```

```r
b <- spread(a, crime_type,crimes) %>% mutate(, tot_crimes = drug + other + property + violent, tot_crm_
panel5 = b[,c(1:3,15:17,8, 18:20)]#district,tot_crm_p_res,vio_crm_p_res, pro_crm_p_res, p50_inc,ro_blac
head(panel5,15)
```

```
## # A tibble: 15 x 10
## # Groups:   month, period, district, p50_inc [15]
##     month       period district tot_crm_p_res vio_crm_p_res pro_crm_p_res p50_inc
##     <chr>        <int>   <int>         <dbl>         <dbl>         <dbl>   <dbl>
## 1 2005-01-01        1       1        0.0403       0.00551        0.0204  91085.
## 2 2005-01-01        1       2        0.0356       0.00961        0.0134  29890.
## 3 2005-01-01        1       3        0.0217       0.00678        0.00771 28048.
## 4 2005-01-01        1       4        0.0148       0.00495        0.00573 39010.
## 5 2005-01-01        1       5        0.0190       0.00585        0.00628 33147.
## 6 2005-01-01        1       6        0.0194       0.00660        0.00819 34672.
## 7 2005-01-01        1       7        0.0275       0.00994        0.00891 23961.
## 8 2005-01-01        1       8        0.00963      0.00255        0.00474 49069.
## 9 2005-01-01        1       9        0.00993      0.00294        0.00437 36324.
## 10 2005-01-01       1      10        0.0119       0.00376        0.00392 29147.
## 11 2005-01-01       1      11        0.0319       0.00865        0.00771 26338.
## 12 2005-01-01       1      12        0.0188       0.00491        0.00864 52561.
## 13 2005-01-01       1      13        0.00746      0.00193        0.00406 61950.
## 14 2005-01-01       1      14        0.0103       0.00250        0.00558 61849.
## 15 2005-01-01       1      15        0.0269       0.00806        0.00649 27940.
## # ... with 3 more variables: ro_black <dbl>, ro_hisp <dbl>, ro_white <dbl>
```

```r
#Ex3

#pool estimation
off = off_ori
panel5_3 = cbind(panel5, b$tot_crimes) %>% rename(,c("tot_crimes" = "...11"))
```

```
## New names:
## * NA -> ...11
```

```r
offcri= left_join(off, panel5_3, by=c("unit"="district" , "month"="month"))
tot_crimes = unlist(offcri$tot_crimes)
```

```
offcri = data.frame(offcri[,1:13], tot_crimes)

reg1 <- plm(arrest ~ tenure + p50_inc + tot_crimes + ro_black + ro_hisp + ro_white, data = offcri, model
summary(reg1)
```

```
## Pooling Model
##
## Call:
## plm(formula = arrest ~ tenure + p50_inc + tot_crimes + ro_black +
##      ro_hisp + ro_white, data = offcri, model = "pooling")
##
## Unbalanced Panel: n = 13028, T = 1-132, N = 1077905
##
## Residuals:
##     Min.  1st Qu.   Median  3rd Qu.      Max.
## -0.50168 -0.49929 -0.49808  0.50082  5.50249
##
## Coefficients:
##                 Estimate  Std. Error t-value Pr(>|t|)
## (Intercept)   5.0671e-01  1.2781e-02 39.6438   <2e-16 ***
## tenure       -4.1607e-06  8.3541e-06 -0.4980   0.6185
## p50_inc       1.6175e-08  9.1862e-08  0.1761   0.8602
## tot_crimes    2.2296e-07  1.8046e-06  0.1236   0.9017
## ro_black     -8.1016e-03  1.3403e-02 -0.6045   0.5455
## ro_hisp      -5.3626e-03  1.3911e-02 -0.3855   0.6999
## ro_white     -1.2073e-02  1.6323e-02 -0.7397   0.4595
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    538530
## Residual Sum of Squares: 538530
## R-Squared:      2.0317e-06
## Adj. R-Squared: -3.5347e-06
## F-statistic: 0.364999 on 6 and 1077898 DF, p-value: 0.90142
```

As the result shown above, when using the pooling method and does not use the fixed effect or random effect
model, The result is not significant.

```
#Ex4
#using within estimator

reg2 = plm(arrest ~ tenure + p50_inc + tot_crimes + ro_black + ro_hisp + ro_white, data = offcri, model

summary(reg2)
```

```
## Oneway (individual) effect Within Model
##
## Call:
## plm(formula = arrest ~ tenure + p50_inc + tot_crimes + ro_black +
##      ro_hisp + ro_white, data = offcri, model = "within", index = c("month"))
##
## Unbalanced Panel: n = 132, T = 7656-8829, N = 1077905
##
## Residuals:
##     Min.  1st Qu.   Median  3rd Qu.      Max.
```

```
## -0.52161 -0.49999 -0.49260  0.50107  5.51664
##
## Coefficients:
##              Estimate  Std. Error t-value Pr(>|t|)
## tenure     -5.0030e-06  8.4332e-06 -0.5933   0.5530
## p50_inc     2.5222e-08  9.3348e-08  0.2702   0.7870
## tot_crimes -4.5860e-07  2.1593e-06 -0.2124   0.8318
## ro_black   -6.5442e-03  1.3802e-02 -0.4741   0.6354
## ro_hisp    -3.6415e-03  1.4363e-02 -0.2535   0.7999
## ro_white   -1.1387e-02  1.6382e-02 -0.6951   0.4870
##
## Total Sum of Squares:     538470
## Residual Sum of Squares: 538470
## R-Squared:      2.099e-06
## Adj. R-Squared: -0.00012502
## F-statistic: 0.377042 on 6 and 1077767 DF, p-value: 0.89407
```

```
#Ex5
#Q1
reg5_1 = plm(arrest ~ tenure + p50_inc + tot_crimes + ro_black + ro_hisp + ro_white, data = offcri, mode
reg5_2 = plm(arrest ~ tenure + p50_inc + tot_crimes + ro_black + ro_hisp + ro_white, data = offcri, mode
reg5_3 = plm(arrest ~ tenure + p50_inc + tot_crimes + ro_black + ro_hisp + ro_white, data = offcri, mode
```

```
summary(reg5_1)
```

```
## Oneway (individual) effect Within Model
##
## Call:
## plm(formula = arrest ~ tenure + p50_inc + tot_crimes + ro_black +
##     ro_hisp + ro_white, data = offcri, model = "within", index = c("NUID",
##     "month", "unit"))
##
## Unbalanced Panel: n = 13028, T = 1-132, N = 1077905
##
## Residuals:
##     Min.  1st Qu.   Median  3rd Qu.     Max.
## -1.75002 -0.50739 -0.42829  0.49287  5.51161
##
## Coefficients:
##              Estimate  Std. Error t-value Pr(>|t|)
## tenure      2.3791e-05  2.4074e-05  0.9883   0.3230
## p50_inc    -2.7719e-07  2.4407e-07 -1.1357   0.2561
## tot_crimes  3.3373e-06  2.8896e-06  1.1549   0.2481
## ro_black   -4.6214e-02  3.1990e-02 -1.4446   0.1486
## ro_hisp    -5.2359e-02  3.4092e-02 -1.5358   0.1246
## ro_white   -3.2526e-02  4.2254e-02 -0.7698   0.4414
##
## Total Sum of Squares:     531920
## Residual Sum of Squares: 531920
## R-Squared:      4.275e-06
## Adj. R-Squared: -0.012235
## F-statistic: 0.75872 on 6 and 1064871 DF, p-value: 0.60237
```

```
summary(reg5_2)
```

```
## Oneway (individual) effect Between Model
```

```
##
## Call:
## plm(formula = arrest ~ tenure + p50_inc + tot_crimes + ro_black +
##     ro_hisp + ro_white, data = offcri, model = "between", index = c("NUID",
##     "month", "unit"))
##
## Unbalanced Panel: n = 13028, T = 1-132, N = 1077905
## Observations used in estimation: 13028
##
## Residuals:
##        Min.    1st Qu.     Median    3rd Qu.       Max.
## -0.5094400 -0.0621304 -0.0036527  0.0534082  2.4986345
##
## Coefficients:
##               Estimate  Std. Error t-value Pr(>|t|)
## (Intercept)  4.6690e-01  2.6859e-02 17.3835  < 2e-16 ***
## tenure       1.2027e-05  1.4395e-05  0.8355  0.40345
## p50_inc      1.4290e-07  1.9008e-07  0.7518  0.45219
## tot_crimes  -7.8368e-06  4.4254e-06 -1.7709  0.07661 .
## ro_black     4.0418e-02  2.8372e-02  1.4245  0.15431
## ro_hisp      4.9050e-02  2.9404e-02  1.6681  0.09531 .
## ro_white     2.2170e-02  3.3474e-02  0.6623  0.50779
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    280.08
## Residual Sum of Squares: 279.93
## R-Squared:      0.00053308
## Adj. R-Squared: 7.2527e-05
## F-statistic: 1.15748 on 6 and 13021 DF, p-value: 0.32605
```

```
summary(reg5_3)
```

```
## Oneway (individual) effect First-Difference Model
##
## Call:
## plm(formula = arrest ~ tenure + p50_inc + tot_crimes + ro_black +
##     ro_hisp + ro_white, data = offcri, model = "fd", index = c("NUID",
##     "month", "unit"))
##
## Unbalanced Panel: n = 13028, T = 1-132, N = 1077905
## Observations used in estimation: 1059380
##
## Residuals:
##         Min.     1st Qu.      Median     3rd Qu.        Max.
## -6.00070240 -0.99901908  0.00019606  0.99943717  6.00061040
##
## Coefficients:
##               Estimate  Std. Error t-value Pr(>|t|)
## (Intercept) -1.2211e-04  1.2958e-03 -0.0942   0.9249
## tenure      -9.5753e-05  8.4693e-04 -0.1131   0.9100
## p50_inc     -1.6188e-06  4.4146e-06 -0.3667   0.7139
## tot_crimes  -4.3614e-06  7.2379e-06 -0.6026   0.5468
## ro_black     5.0587e-01  8.1091e-01  0.6238   0.5327
## ro_hisp      9.8436e-03  1.5247e+00  0.0065   0.9948
```

```
## ro_white     1.1685e+00  1.4531e+00  0.8041    0.4213
##
## Total Sum of Squares:     1058500
## Residual Sum of Squares: 1058500
## R-Squared:      2.5775e-06
## Adj. R-Squared: -3.0863e-06
## F-statistic: 0.455082 on 6 and 1059373 DF, p-value: 0.84183
```

```r
rmarkdown::render('A3.Rmd', 'pdf_document')
```

```
## Warning in has_crop_tools():
## Tool(s) not installed or not in PATH: ghostscript
## -> As a result, figure cropping will be disabled.

##
##
## processing file: A3.Rmd

## Error in parse_block(g[-1], g[1], params.src, markdown_mode): Duplicate chunk label 'setup', which ha
## knitr::opts_chunk$set(echo = TRUE)
## chooseCRANmirror(graphics=FALSE, ind=1)
## knitr::opts_chunk$set(error = TRUE)
## library(pastecs)
## library(tidyverse)
## library(data.table)
## library(stringr)
## library(mfx)
## library(MASS)
## library(bayesm)
## library(rmarkdown)
## library(ggplot2)
## library(dplyr)
## library(lubridate)
## library(plm)
## setwd('C:/Users/sophi/Desktop/Econ613/A3')
## options(knitr.purl.inline = TRUE)
```