

Mengyang Zheng HW3 4/26/2021**Exercise 2: Data Manipulation**

```
rm(list=ls())
#install.packages("lattice")
setwd("C:/Users/bobme/Desktop/ECON 613/Homework/A3/Data")

crime <- read.csv("C:/Users/bobme/Desktop/ECON 613/Homework/A3/Data/crime_long.csv", stringsAsFactors=TRUE)
officer <- read.csv("C:/Users/bobme/Desktop/ECON 613/Homework/A3/Data/officers.csv", stringsAsFactors=TRUE)
pop <- read.csv("C:/Users/bobme/Desktop/ECON 613/Homework/A3/Data/population.csv", stringsAsFactors=TRUE)
```

```
#Calculate total crime per month and plot the time series of crime
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
```

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.3      v purrr   0.3.4
## v tibble  3.0.6      v stringr 1.4.0
## v tidyr   1.1.3      v forcats 0.5.1
## v readr   1.4.0
```

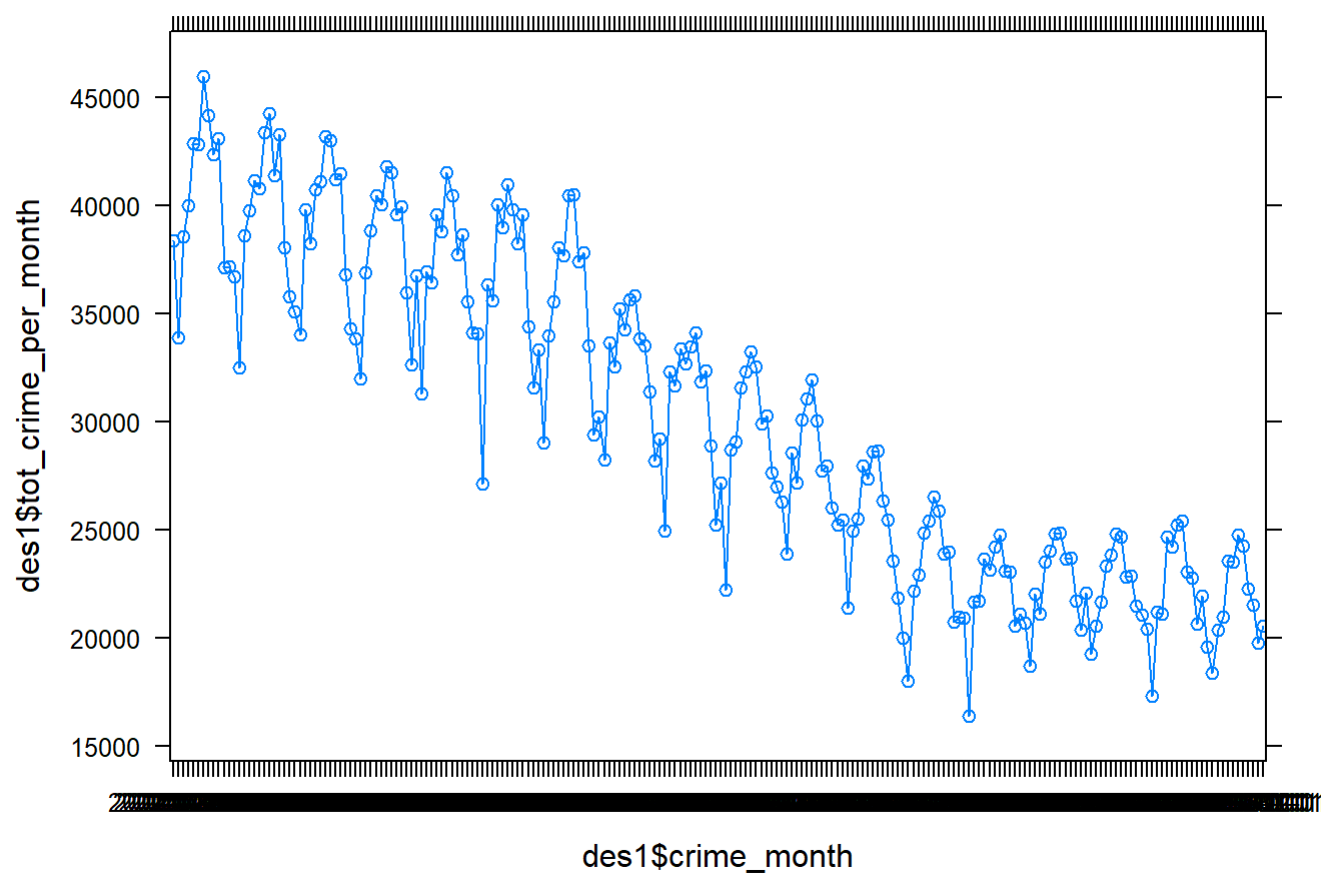
```
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(ggplot2)
library(lattice)
```

```
## Warning: package 'lattice' was built under R version 4.0.5
```

```
#crime$crime_month=as.Date(crime$crime_month,format="%y/%m/%d")
des1=crime %>%
  group_by(crime_month)%>%
  summarise(
    tot_crime_per_month=sum(crimes)
  )

xyplot(des1$tot_crime_per_month~des1$crime_month,type=c("l","p"))
```



```

#Merge the two datasets
pop=rename(pop,crime_month=month)

crime$merger=paste0(crime$crime_month,crime$district)
crime0= crime %>%
  group_by(merger,crime_type) %>%
  summarise(
    tot_crime=sum(crimes),.groups="drop"
  )
crime$merger1=paste0(crime$merger,crime$crime_type)
crime0$merger1=paste0(crime0$merger,crime0$crime_type)
crime0$merger=NULL
crime0$crime_type=NULL
crime=merge(crime,crime0,by="merger1")
crime0=NULL
crime=crime[!duplicated(crime[,c("merger1")]),]
crime$merger1=NULL
crime$crimes=NULL
pop$merger=paste0(pop$crime_month,pop$district)
crime1=crime
crime1$crime_month=NULL
crime1$district=NULL
panel=merge(pop,crime1,by="merger")

```

```

#Construct a panel data
panel1=panel
des2=panel1 %>%
  group_by(merger) %>%
  summarise(
    tot_crime_per_res=sum(tot_crime)/tot_pop,
    vio_crime_per_res=sum(tot_crime[crime_type=="violent"])/tot_pop,
    prop_crime_per_res=sum(tot_crime[crime_type=="property"])/tot_pop,
    share_black=tot_black/tot_pop,
    share_white=tot_white/tot_pop,
    share_hisp=tot_hisp/tot_pop,
    .groups="drop"
  )
panel1=merge(panel1,des2,by="merger")
panel1=panel1[,c(1:4,9,12:17)]
panel1=panel1[!duplicated(panel1[,c("merger")]),]
panel1$merger=NULL
panel1[,c(5:10)]=round(panel1[,c(5:10)],digits=4)
panel1=panel1[order(panel1$crime_month,panel1$district),]

head(panel1,20)

```

##	crime_month	period	district	p50_inc	tot_crime_per_res	vio_crime_per_res
## 1	2005-01-01	1	1	91084.91	0.0403	0.0055
## 177	2005-01-01	1	2	29890.17	0.0356	0.0096
## 289	2005-01-01	1	3	28047.56	0.0217	0.0068
## 305	2005-01-01	1	4	39010.22	0.0148	0.0049
## 321	2005-01-01	1	5	33146.90	0.0190	0.0058
## 337	2005-01-01	1	6	34672.25	0.0194	0.0066
## 353	2005-01-01	1	7	23960.59	0.0275	0.0099
## 369	2005-01-01	1	8	49069.15	0.0096	0.0025
## 385	2005-01-01	1	9	36323.73	0.0099	0.0029
## 17	2005-01-01	1	10	29147.49	0.0119	0.0038
## 33	2005-01-01	1	11	26337.89	0.0319	0.0087
## 49	2005-01-01	1	12	52560.59	0.0188	0.0049
## 65	2005-01-01	1	13	61949.53	0.0075	0.0019
## 81	2005-01-01	1	14	61849.46	0.0103	0.0025
## 97	2005-01-01	1	15	27940.01	0.0269	0.0081
## 113	2005-01-01	1	16	64130.01	0.0056	0.0011
## 129	2005-01-01	1	17	55060.06	0.0068	0.0014
## 145	2005-01-01	1	18	87939.24	0.0115	0.0021
## 161	2005-01-01	1	19	93097.57	0.0114	0.0021
## 193	2005-01-01	1	20	52332.72	0.0066	0.0016
##	prop_crime_per_res	share_black	share_white	share_hisp		
## 1	0.0204	0.1287	0.5876	0.0661		
## 177	0.0134	0.9467	0.0166	0.0165		
## 289	0.0077	0.9130	0.0409	0.0174		
## 305	0.0057	0.6209	0.0797	0.2842		
## 321	0.0063	0.9418	0.0113	0.0339		
## 337	0.0082	0.9745	0.0034	0.0101		
## 353	0.0089	0.9680	0.0037	0.0164		
## 369	0.0047	0.2111	0.2082	0.5654		
## 385	0.0044	0.1216	0.1517	0.6108		
## 17	0.0039	0.3374	0.0359	0.6191		
## 33	0.0077	0.8434	0.0248	0.1174		
## 49	0.0086	0.1742	0.3114	0.4125		
## 65	0.0041	0.1839	0.5159	0.2480		
## 81	0.0056	0.0722	0.4174	0.4609		
## 97	0.0065	0.9335	0.0163	0.0370		
## 113	0.0032	0.0101	0.6882	0.2319		
## 129	0.0037	0.0333	0.3863	0.4326		
## 145	0.0058	0.0928	0.7530	0.0498		
## 161	0.0068	0.0251	0.7976	0.1069		
## 193	0.0034	0.1066	0.5403	0.1834		

Exercise 3: Panel Data: Introduction

```
panel2=left_join(panel1,officer,by=c("crime_month" = "month", "district" = "unit"))
panel2=drop_na(panel2)
```

```
ex3model=lm(arrest~tenure+p50_inc+tot_crime_per_res+share_black+share_white+share_hisp-1,data=panel2)
summary(ex3model)
```

```
##
## Call:
## lm(formula = arrest ~ tenure + p50_inc + tot_crime_per_res +
##      share_black + share_white + share_hisp - 1, data = panel2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.5526 -0.5089 -0.4678  0.4938  5.5270
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## tenure           4.025e-05  8.268e-06   4.868 1.13e-06 ***
## p50_inc           2.846e-07  9.513e-08   2.992  0.00277 **
## tot_crime_per_res  8.477e-01  1.248e-01   6.795 1.08e-11 ***
## share_black       4.761e-01  3.423e-03 139.091 < 2e-16 ***
## share_white       5.382e-01  9.790e-03  54.976 < 2e-16 ***
## share_hisp       5.010e-01  3.715e-03 134.872 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7074 on 1077899 degrees of freedom
## Multiple R-squared:  0.3317, Adjusted R-squared:  0.3317
## F-statistic: 8.917e+04 on 6 and 1077899 DF,  p-value: < 2.2e-16
```

Exercise 4: Panel Data: More Controls

```
ex4model=lm(arrest~tenure+p50_inc+tot_crime_per_res+share_black+share_white+share_hisp+factor(di
strict)+factor(crime_month)-1,data=panel2)
summary(ex4model)
```

```
##
## Call:
## lm(formula = arrest ~ tenure + p50_inc + tot_crime_per_res +
##      share_black + share_white + share_hisp + factor(district) +
##      factor(crime_month) - 1, data = panel2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.5280 -0.5003 -0.4920  0.5008  5.5155
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## tenure          -3.666e-06  8.526e-06  -0.430   0.6672
## p50_inc          -4.559e-08  6.523e-07  -0.070   0.9443
## tot_crime_per_res -6.023e-01  3.940e-01  -1.529   0.1263
## share_black      -7.807e-02  1.044e-01  -0.748   0.4544
## share_white     -1.729e-01  1.845e-01  -0.937   0.3487
## share_hisp      -1.432e-01  2.025e-01  -0.707   0.4795
## factor(district)1  6.448e-01  1.065e-01   6.053 1.42e-09 ***
## factor(district)2  5.968e-01  9.716e-02   6.142 8.16e-10 ***
## factor(district)3  5.959e-01  9.915e-02   6.010 1.85e-09 ***
## factor(district)4  6.145e-01  1.231e-01   4.994 5.93e-07 ***
## factor(district)5  5.958e-01  9.942e-02   5.992 2.07e-09 ***
## factor(district)6  5.943e-01  9.711e-02   6.120 9.33e-10 ***
## factor(district)7  5.948e-01  9.941e-02   5.984 2.18e-09 ***
## factor(district)8  6.443e-01  1.561e-01   4.128 3.66e-05 ***
## factor(district)9  6.311e-01  1.472e-01   4.287 1.81e-05 ***
## factor(district)10 6.295e-01  1.559e-01   4.037 5.41e-05 ***
## factor(district)11 6.073e-01  1.080e-01   5.622 1.89e-08 ***
## factor(district)12 6.374e-01  1.389e-01   4.589 4.45e-06 ***
## factor(district)13 6.427e-01  1.399e-01   4.595 4.33e-06 ***
## factor(district)14 6.621e-01  1.535e-01   4.313 1.61e-05 ***
## factor(district)15 5.969e-01  1.008e-01   5.919 3.23e-09 ***
## factor(district)16 6.618e-01  1.488e-01   4.446 8.74e-06 ***
## factor(district)17 6.397e-01  1.405e-01   4.552 5.31e-06 ***
## factor(district)18 6.527e-01  1.281e-01   5.095 3.48e-07 ***
## factor(district)19 6.603e-01  1.373e-01   4.808 1.52e-06 ***
## factor(district)20 6.356e-01  1.266e-01   5.019 5.21e-07 ***
## factor(district)21 5.913e-01  8.645e-02   6.839 7.98e-12 ***
## factor(district)22 6.225e-01  1.129e-01   5.514 3.52e-08 ***
## factor(district)23 6.478e-01  1.351e-01   4.794 1.64e-06 ***
## factor(district)24 6.281e-01  1.218e-01   5.154 2.54e-07 ***
## factor(district)25 6.452e-01  1.631e-01   3.956 7.64e-05 ***
## factor(crime_month)2007-02-01 2.687e-03  1.108e-02   0.242  0.8084
## factor(crime_month)2007-03-01 5.463e-03  1.098e-02   0.498  0.6188
## factor(crime_month)2007-04-01 -4.474e-03  1.099e-02  -0.407  0.6840
## factor(crime_month)2007-05-01 9.719e-03  1.103e-02   0.881  0.3781
## factor(crime_month)2007-06-01 -1.513e-02  1.098e-02  -1.378  0.1683
## factor(crime_month)2007-07-01 -2.419e-03  1.103e-02  -0.219  0.8264
## factor(crime_month)2007-08-01 -9.274e-03  1.098e-02  -0.844  0.3984
## factor(crime_month)2007-09-01 2.832e-03  1.095e-02   0.259  0.7959
## factor(crime_month)2007-10-01 5.450e-03  1.095e-02   0.498  0.6188
## factor(crime_month)2007-11-01 3.153e-03  1.088e-02   0.290  0.7721
```

```

## factor(crime_month)2007-12-01 -8.609e-03 1.089e-02 -0.791 0.4291
## factor(crime_month)2008-01-01 -1.592e-02 1.090e-02 -1.461 0.1441
## factor(crime_month)2008-02-01 -8.108e-03 1.095e-02 -0.741 0.4590
## factor(crime_month)2008-03-01 -7.140e-03 1.090e-02 -0.655 0.5124
## factor(crime_month)2008-04-01 7.854e-03 1.089e-02 0.721 0.4709
## factor(crime_month)2008-05-01 1.881e-02 1.090e-02 1.727 0.0842
## factor(crime_month)2008-06-01 9.747e-04 1.087e-02 0.090 0.9285
## factor(crime_month)2008-07-01 5.324e-05 1.088e-02 0.005 0.9961
## factor(crime_month)2008-08-01 1.109e-03 1.090e-02 0.102 0.9189
## factor(crime_month)2008-09-01 9.334e-03 1.085e-02 0.861 0.3895
## factor(crime_month)2008-10-01 -3.419e-03 1.084e-02 -0.315 0.7524
## factor(crime_month)2008-11-01 -1.711e-03 1.083e-02 -0.158 0.8745
## factor(crime_month)2008-12-01 1.068e-02 1.087e-02 0.982 0.3259
## factor(crime_month)2009-01-01 -6.246e-03 1.085e-02 -0.575 0.5650
## factor(crime_month)2009-02-01 -6.056e-03 1.091e-02 -0.555 0.5788
## factor(crime_month)2009-03-01 -9.077e-03 1.085e-02 -0.836 0.4029
## factor(crime_month)2009-04-01 -5.104e-03 1.087e-02 -0.469 0.6388
## factor(crime_month)2009-05-01 -1.688e-03 1.086e-02 -0.155 0.8765
## factor(crime_month)2009-06-01 4.015e-03 1.086e-02 0.370 0.7116
## factor(crime_month)2009-07-01 5.287e-03 1.086e-02 0.487 0.6265
## factor(crime_month)2009-08-01 -4.864e-03 1.087e-02 -0.448 0.6545
## factor(crime_month)2009-09-01 -5.000e-03 1.088e-02 -0.460 0.6457
## factor(crime_month)2009-10-01 -1.741e-03 1.088e-02 -0.160 0.8728
## factor(crime_month)2009-11-01 -5.459e-03 1.089e-02 -0.501 0.6161
## factor(crime_month)2009-12-01 -1.016e-02 1.094e-02 -0.928 0.3533
## factor(crime_month)2010-01-01 1.579e-03 1.093e-02 0.144 0.8851
## factor(crime_month)2010-02-01 3.326e-03 1.112e-02 0.299 0.7649
## factor(crime_month)2010-03-01 -9.071e-03 1.099e-02 -0.825 0.4093
## factor(crime_month)2010-04-01 6.832e-03 1.101e-02 0.621 0.5348
## factor(crime_month)2010-05-01 -3.665e-03 1.101e-02 -0.333 0.7393
## factor(crime_month)2010-06-01 -6.041e-03 1.102e-02 -0.548 0.5836
## factor(crime_month)2010-07-01 -1.367e-02 1.102e-02 -1.240 0.2152
## factor(crime_month)2010-08-01 9.653e-03 1.103e-02 0.875 0.3817
## factor(crime_month)2010-09-01 2.029e-03 1.103e-02 0.184 0.8541
## factor(crime_month)2010-10-01 -2.748e-03 1.103e-02 -0.249 0.8032
## factor(crime_month)2010-11-01 -9.368e-03 1.109e-02 -0.845 0.3981
## factor(crime_month)2010-12-01 -5.083e-03 1.119e-02 -0.454 0.6496
## factor(crime_month)2011-01-01 -8.493e-03 1.117e-02 -0.760 0.4470
## factor(crime_month)2011-02-01 -5.756e-03 1.137e-02 -0.506 0.6128
## factor(crime_month)2011-03-01 -4.797e-03 1.120e-02 -0.428 0.6685
## factor(crime_month)2011-04-01 -1.480e-03 1.120e-02 -0.132 0.8948
## factor(crime_month)2011-05-01 -7.996e-04 1.117e-02 -0.072 0.9429
## factor(crime_month)2011-06-01 -5.562e-05 1.115e-02 -0.005 0.9960
## factor(crime_month)2011-07-01 3.844e-03 1.113e-02 0.345 0.7299
## factor(crime_month)2011-08-01 -2.402e-04 1.114e-02 -0.022 0.9828
## factor(crime_month)2011-09-01 -1.158e-02 1.100e-02 -1.052 0.2927
## factor(crime_month)2011-10-01 1.026e-02 1.099e-02 0.933 0.3509
## factor(crime_month)2011-11-01 -7.209e-03 1.104e-02 -0.653 0.5137
## factor(crime_month)2011-12-01 4.909e-05 1.106e-02 0.004 0.9965
## factor(crime_month)2012-01-01 -1.070e-02 1.110e-02 -0.964 0.3350
## factor(crime_month)2012-02-01 2.720e-03 1.115e-02 0.244 0.8073
## factor(crime_month)2012-03-01 1.794e-03 1.103e-02 0.163 0.8708
## factor(crime_month)2012-04-01 4.315e-03 1.111e-02 0.388 0.6979
## factor(crime_month)2012-05-01 4.066e-03 1.107e-02 0.367 0.7134

```

```

## factor(crime_month)2012-06-01 -2.423e-03 1.108e-02 -0.219 0.8268
## factor(crime_month)2012-07-01 2.962e-04 1.108e-02 0.027 0.9787
## factor(crime_month)2012-08-01 -5.480e-03 1.111e-02 -0.493 0.6219
## factor(crime_month)2012-09-01 2.011e-03 1.116e-02 0.180 0.8570
## factor(crime_month)2012-10-01 -1.481e-02 1.115e-02 -1.328 0.1842
## factor(crime_month)2012-11-01 3.206e-03 1.120e-02 0.286 0.7748
## factor(crime_month)2012-12-01 -1.713e-02 1.124e-02 -1.524 0.1276
## factor(crime_month)2013-01-01 3.522e-03 1.126e-02 0.313 0.7545
## factor(crime_month)2013-02-01 -8.281e-03 1.142e-02 -0.725 0.4685
## factor(crime_month)2013-03-01 -1.605e-02 1.130e-02 -1.420 0.1556
## factor(crime_month)2013-04-01 3.339e-03 1.127e-02 0.296 0.7670
## factor(crime_month)2013-05-01 1.395e-02 1.120e-02 1.246 0.2129
## factor(crime_month)2013-06-01 6.782e-03 1.123e-02 0.604 0.5458
## factor(crime_month)2013-07-01 1.735e-02 1.120e-02 1.549 0.1213
## factor(crime_month)2013-08-01 -3.001e-03 1.121e-02 -0.268 0.7889
## factor(crime_month)2013-09-01 2.823e-03 1.126e-02 0.251 0.8020
## factor(crime_month)2013-10-01 -8.861e-03 1.128e-02 -0.785 0.4322
## factor(crime_month)2013-11-01 -5.925e-03 1.134e-02 -0.522 0.6015
## factor(crime_month)2013-12-01 3.212e-03 1.141e-02 0.282 0.7783
## factor(crime_month)2014-01-01 -1.506e-02 1.152e-02 -1.307 0.1912
## factor(crime_month)2014-02-01 2.439e-03 1.160e-02 0.210 0.8335
## factor(crime_month)2014-03-01 -4.560e-03 1.142e-02 -0.399 0.6897
## factor(crime_month)2014-04-01 -1.876e-02 1.138e-02 -1.650 0.0990
## factor(crime_month)2014-05-01 -6.046e-03 1.130e-02 -0.535 0.5926
## factor(crime_month)2014-06-01 -1.930e-03 1.124e-02 -0.172 0.8636
## factor(crime_month)2014-07-01 5.418e-03 1.122e-02 0.483 0.6291
## factor(crime_month)2014-08-01 -1.241e-03 1.124e-02 -0.110 0.9121
## factor(crime_month)2014-09-01 9.119e-04 1.125e-02 0.081 0.9354
## factor(crime_month)2014-10-01 -1.873e-03 1.125e-02 -0.167 0.8677
## factor(crime_month)2014-11-01 -8.672e-03 1.137e-02 -0.763 0.4457
## factor(crime_month)2014-12-01 -7.118e-03 1.137e-02 -0.626 0.5313
## factor(crime_month)2015-01-01 -5.023e-03 1.140e-02 -0.440 0.6596
## factor(crime_month)2015-02-01 -9.394e-03 1.165e-02 -0.806 0.4201
## factor(crime_month)2015-03-01 4.066e-03 1.139e-02 0.357 0.7210
## factor(crime_month)2015-04-01 -3.556e-03 1.141e-02 -0.312 0.7552
## factor(crime_month)2015-05-01 -6.859e-03 1.133e-02 -0.606 0.5448
## factor(crime_month)2015-06-01 -7.279e-03 1.133e-02 -0.642 0.5207
## factor(crime_month)2015-07-01 -8.698e-03 1.129e-02 -0.770 0.4411
## factor(crime_month)2015-08-01 -5.392e-03 1.128e-02 -0.478 0.6326
## factor(crime_month)2015-09-01 -1.122e-03 1.136e-02 -0.099 0.9213
## factor(crime_month)2015-10-01 -8.472e-03 1.136e-02 -0.746 0.4557
## factor(crime_month)2015-11-01 2.225e-03 1.145e-02 0.194 0.8459
## factor(crime_month)2015-12-01 -7.488e-03 1.145e-02 -0.654 0.5131
## factor(crime_month)2016-01-01 -5.450e-03 1.149e-02 -0.474 0.6353
## factor(crime_month)2016-02-01 -7.161e-03 1.157e-02 -0.619 0.5359
## factor(crime_month)2016-03-01 -1.460e-02 1.141e-02 -1.279 0.2009
## factor(crime_month)2016-04-01 9.733e-03 1.144e-02 0.851 0.3949
## factor(crime_month)2016-05-01 9.769e-03 1.134e-02 0.861 0.3891
## factor(crime_month)2016-06-01 -5.276e-03 1.132e-02 -0.466 0.6412
## factor(crime_month)2016-07-01 -1.090e-02 1.130e-02 -0.964 0.3349
## factor(crime_month)2016-08-01 -1.938e-02 1.127e-02 -1.719 0.0856
## factor(crime_month)2016-09-01 -1.698e-03 1.131e-02 -0.150 0.8806
## factor(crime_month)2016-10-01 2.934e-03 1.132e-02 0.259 0.7955
## factor(crime_month)2016-11-01 -1.413e-02 1.141e-02 -1.239 0.2154

```



```
## factor(crime_month)2016-12-01 -1.303e-02 1.145e-02 -1.138 0.2549
## factor(crime_month)2017-01-01 -8.919e-05 1.147e-02 -0.008 0.9938
## factor(crime_month)2017-02-01 -5.932e-03 1.158e-02 -0.512 0.6085
## factor(crime_month)2017-03-01 6.001e-03 1.152e-02 0.521 0.6025
## factor(crime_month)2017-04-01 -5.744e-03 1.149e-02 -0.500 0.6170
## factor(crime_month)2017-05-01 8.357e-03 1.142e-02 0.732 0.4645
## factor(crime_month)2017-06-01 -1.222e-02 1.140e-02 -1.072 0.2837
## factor(crime_month)2017-07-01 -6.198e-03 1.138e-02 -0.545 0.5858
## factor(crime_month)2017-08-01 -2.657e-03 1.139e-02 -0.233 0.8156
## factor(crime_month)2017-09-01 -8.604e-03 1.145e-02 -0.752 0.4523
## factor(crime_month)2017-10-01 -9.814e-03 1.143e-02 -0.859 0.3905
## factor(crime_month)2017-11-01 -1.686e-02 1.150e-02 -1.467 0.1423
## factor(crime_month)2017-12-01 -8.989e-03 1.149e-02 -0.782 0.4342
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7068 on 1077743 degrees of freedom
## Multiple R-squared:  0.3328, Adjusted R-squared:  0.3327
## F-statistic: 3318 on 162 and 1077743 DF, p-value: < 2.2e-16
```

Exercise 5: Panel Data: Individual Fixed Effects

```
between=panel2 %>%
  group_by(NUID) %>%
  summarise(mean_arrest=mean(arrest),
            mean_tenure=mean(tenure),
            mean_tot_crime=mean(tot_crime_per_res),
            mean_median=mean(p50_inc),
            mean_share_black=mean(share_black),
            mean_share_white=mean(share_white),
            mean_share_hisp=mean(share_hisp))

between_est=lm(mean_arrest~mean_tenure+mean_tot_crime+mean_median+ mean_share_black+mean_share_w
hite+mean_share_hisp,data=between)
summary(between_est)
```

```
##
## Call:
## lm(formula = mean_arrest ~ mean_tenure + mean_tot_crime + mean_median +
##      mean_share_black + mean_share_white + mean_share_hisp, data = between)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.50861 -0.06195 -0.00366  0.05368  2.49784
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    4.873e-01  2.639e-02  18.466  <2e-16 ***
## mean_tenure     1.115e-05  1.436e-05   0.776   0.438
## mean_tot_crime  -4.475e-01  2.889e-01  -1.549   0.121
## mean_median     1.566e-07  1.970e-07   0.795   0.427
## mean_share_black 1.720e-02  2.605e-02   0.660   0.509
## mean_share_white -4.928e-03  3.525e-02  -0.140   0.889
## mean_share_hisp  1.669e-02  2.742e-02   0.609   0.543
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1466 on 13021 degrees of freedom
## Multiple R-squared:  0.0004767, Adjusted R-squared:  1.608e-05
## F-statistic: 1.035 on 6 and 13021 DF, p-value: 0.4002
```

```
within=panel2 %>%
  group_by(NUID)%>%
  summarise(meandiff_arrest=arrest-mean(arrest),
            meandiff_tenure=tenure-mean(tenure),
            meandiff_tot_crime=tot_crime_per_res-mean(tot_crime_per_res),
            meandiff_median=p50_inc-mean(p50_inc),
            meandiff_share_black=share_black-mean(share_black),
            meandiff_share_white=share_white-mean(share_white),
            meandiff_share_hisp=share_hisp-mean(share_hisp))
```

```
## `summarise()` has grouped output by 'NUID'. You can override using the `.groups` argument.
```

```
within_est=lm(meandiff_arrest~meandiff_tenure+meandiff_tot_crime+meandiff_median+ meandiff_share
_black+meandiff_share_white+meandiff_share_hisp,data=within)
summary(within_est)
```

```
##
## Call:
## lm(formula = meandiff_arrest ~ meandiff_tenure + meandiff_tot_crime +
##      meandiff_median + meandiff_share_black + meandiff_share_white +
##      meandiff_share_hisp, data = within)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7500 -0.5074 -0.4283  0.4928  5.5117
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -1.684e-15  6.766e-04   0.000    1.000
## meandiff_tenure    2.168e-05  2.461e-05   0.881    0.378
## meandiff_tot_crime  2.109e-01  2.332e-01   0.905    0.366
## meandiff_median   -2.797e-07  2.473e-07  -1.131    0.258
## meandiff_share_black -4.183e-02  3.130e-02  -1.337    0.181
## meandiff_share_white -2.639e-02  4.224e-02  -0.625    0.532
## meandiff_share_hisp -4.352e-02  3.302e-02  -1.318    0.188
##
## Residual standard error: 0.7025 on 1077898 degrees of freedom
## Multiple R-squared:  3.781e-06, Adjusted R-squared:  -1.785e-06
## F-statistic: 0.6793 on 6 and 1077898 DF,  p-value: 0.6665
```

```
firstdiff=panel2 %>%
  group_by(NUID)%>%
  summarise(meandiff2_arrest=lag(arrest),
            meandiff2_tenure=lag(tenure),
            meandiff2_tot_crime=lag(tot_crime_per_res),
            meandiff2_median=lag(p50_inc),
            meandiff2_share_black=lag(share_black),
            meandiff2_share_white=lag(share_white),
            meandiff2_share_hisp=lag(share_hisp),
            .groups="drop",
            order_by=crime_month)
firstdiff=drop_na(firstdiff)
firstdiff_est=lm(meandiff2_arrest~meandiff2_tenure+meandiff2_tot_crime+meandiff2_median+ meandiff2_share_black+meandiff2_share_white+meandiff2_share_hisp,data=firstdiff)
summary(firstdiff_est)
```

```
##
## Call:
## lm(formula = meandiff2_arrest ~ meandiff2_tenure + meandiff2_tot_crime +
##      meandiff2_median + meandiff2_share_black + meandiff2_share_white +
##      meandiff2_share_hisp, data = firstdiff)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.5015 -0.4991 -0.4982  0.5010  5.5024
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    5.061e-01  1.280e-02  39.554  <2e-16 ***
## meandiff2_tenure -2.864e-06  8.420e-06  -0.340   0.734
## meandiff2_tot_crime -1.770e-02  1.274e-01  -0.139   0.890
## meandiff2_median    3.234e-08  9.581e-08   0.338   0.736
## meandiff2_share_black -7.674e-03  1.269e-02  -0.605   0.545
## meandiff2_share_white -1.278e-02  1.704e-02  -0.750   0.453
## meandiff2_share_hisp -4.810e-03  1.331e-02  -0.361   0.718
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7069 on 1064870 degrees of freedom
## Multiple R-squared:  1.897e-06, Adjusted R-squared:  -3.737e-06
## F-statistic: 0.3367 on 6 and 1064870 DF, p-value: 0.9178
```

#However I could not add in the individual fixed effects above since it is too large to run, so instead I use plm for calculation

```
between1= plm(arrest~tenure+p50_inc+tot_crime_per_res+share_black+share_white+share_hisp+factor(
crime_month)+factor(district),index="NUID",model="between",data=panel2)
```

```
## Error in plm(arrest ~ tenure + p50_inc + tot_crime_per_res + share_black + : could not find f
unction "plm"
```

```
summary(between1)
```

```
## Error in summary(between1): object 'between1' not found
```

```
within1= plm(arrest~tenure+p50_inc+tot_crime_per_res+share_black+share_white+share_hisp+factor(c
rime_month)+factor(district),index="NUID",model="within",data=panel2)
```

```
## Error in plm(arrest ~ tenure + p50_inc + tot_crime_per_res + share_black + : could not find f
unction "plm"
```

```
summary(within1)
```

```
## Error in summary(within1): object 'within1' not found
```

```
firstdiff1= plm(arrest~tenure+p50_inc+tot_crime_per_res+share_black+share_white+share_hisp+factor(crime_month)+factor(district),index="NUIID",model="fd",data=panel2)
```

```
## Error in plm(arrest ~ tenure + p50_inc + tot_crime_per_res + share_black + : could not find function "plm"
```

```
summary(firstdiff1)
```

```
## Error in summary(firstdiff1): object 'firstdiff1' not found
```

```
#Use a GMM approach to estimate all parameters in 1 step
# Do not know why it keeps giving me errors, do not know how to fix it
panel2=unique(panel2)
ex5model=pgmm(formula=arrest~tenure+p50_inc+tot_crime_per_res+share_black+share_white+share_hisp|crime_month+district+NUIID,data=panel2,effect="individual",model="onestep")
```

```
## Error in pgmm(formula = arrest ~ tenure + p50_inc + tot_crime_per_res + : could not find function "pgmm"
```

```
summary(ex5model)
```

```
## Error in summary(ex5model): object 'ex5model' not found
```

```
#betahat=(X'ZWZ'X)^(-1)(X'ZWZ'Y) for OLS solution by GMM
#I do not know how to debug this but just want to give an idea, plus the data is very large, it cannot be done through my computer
gmm=function(theta,x,momentFun){
  avg=colMeans(momentFun(theta,x))
  sum(avg^2)
}
init=0
optim(init,gmm,x=panel2,moment=sum(panel2)/nrow(panel2),control=list(reltol=1e-19,maxit=20000))$par
```

```
## Warning in optim(init, gmm, x = panel2, moment = sum(panel2)/nrow(panel2), : one-dimensional optimization by Nelder-Mead is unreliable:
## use "Brent" or optimize() directly
```

```
## Error in FUN(X[[i]], ...): only defined on a data frame with all numeric variables
```