

Final Project DSC 520

Benjamin Bartek

6/1/2019

Summary of Problem Statement

The problem I aim to address is to identify the total nationwide prison population over time, as well as identify the potential correlation between demographic variables, sentencing, categories of crime, and other variables in order to glean insights into incarceration practices in the United States. I will also likely briefly identify how many U.S. citizens cannot vote due to their status as a convicted felon.

Install Packages

```
install.packages("ggplot2")
install.packages("pastecs")
install.packages("rcompanion")
install.packages("rmarkdown")
install.packages("corrplot")
install.packages("RColorBrewer")
install.packages("PerformanceAnalytics")
install.packages("ggm")
install.packages("dplyr")
library(dplyr)
library(pastecs)
library(ggplot2)
library(rcompanion)
library(rmarkdown)
library(corrplot)
library(RColorBrewer)
library(PerformanceAnalytics)
library(ggm)
```

Main Data Cleaning Narrative

Though we have been learning how to do data cleaning in R, most of my data cleaning did not occur within R; it occurred within Excel. And there are some very good reasons primarily relating to ensuring that the data could even be imported into R correctly. Generally in class, we start with mostly clean datasets. However, I obtained my datasets from disparate sources, with disparate formats. There was a lot of extraneous information that needed to be removed from the tops and bottoms of the spreadsheets that would not import into R properly.

I worked with numerous spreadsheets. I had to select which columns and rows I wanted from those spreadsheets, delete other columns and rows, and change the orientation of the columns and rows to fit my dataframe. I also had to create new variable names to fit the scheme of my dataset.

There were numerous extraneous characters that needed to be removed in order for R to read the data correctly.

I changed the spreadsheet cell formats where necessary to ensure that numbers were reading as numbers in R instead of strings.

At one point while examining the dataset for outliers, I discovered that a large chunk of cells was incorrectly displaying as negative values in R. I went back to the source spreadsheet and replaced them in my working data frame with the proper format and value.

There are many missing values in this dataset. That is a result of merging many datasets that measure some variables, but not others. But all of the numeric data relates to the entire US corrections population, the federal corrections population, the entire state corrections population, or each individual state's corrections population. I intend to deal with skipping over NA values when I begin running analyses with the individual variables and regression models.

Michigan and Georgia are also missing some prominent 2016 data that was presumably not reported to the BJS agency. I may need to try to find that data from the states themselves.

I may also need to find other data that relates to demographics.

Main Variable - Opening the File with the Correct Method

View Number of Rows and Columns

```
nrow(maindata)
```

```
## [1] 53
```

```
ncol(maindata)
```

```
## [1] 53
```

Structure of DataFrame

```
str(maindata)
```

```
## 'data.frame': 53 obs. of 53 variables:
```

```
## $ jurisdiction                                : Factor w/ 53 levels "Al
```

```
abama","Alaska",...: 45 46 3 1 2 4 5 6 7 8 ...
```

```
## $ prison_or_jail_2000                        : int 1938500 140100 117
```

```

7200 NA NA NA NA NA NA NA ...
## $ prison_or_jail_2010 : int 1517800 207000 131
0800 NA NA NA NA NA NA NA ...
## $ prison_or_jail_2015 : int 2166600 195600 124
3500 NA NA NA NA NA NA NA ...
## $ total_corr_pop_2016 : int 6582100 320000 626
2000 99800 12900 137500 72100 536100 121900 59600 ...
## $ corr_sup_rate_per_100k_residents_all_ages_2016: int 2030 100 1930 2050
1730 1970 2410 1360 2180 1670 ...
## $ prison_or_jail_2016 : int 2131000 188400 194
2600 40900 4400 55000 24000 202700 32100 15000 ...
## $ prob_or_parole_2016 : int 4537100 131700 440
5400 60700 8400 84800 51500 333300 90900 44700 ...
## $ Incarc_rate_per_100k_US_residents_18_or_older : int 850 80 780 1080 80
0 1030 1050 670 740 530 ...
## $ Incarc_rate_per_100k_US_residents_all_ages : int 660 60 600 840 600
790 800 510 580 420 ...
## $ state_pop_2016 : int NA NA NA 4860545 7
41522 6908642 2988231 39296476 5530105 3587685 ...
## $ violent_crime_total_2016 : int NA NA NA 25878 596
6 32542 16563 174796 19030 8169 ...
## $ murder_manslaughter_2016 : int NA NA NA 407 52 38
9 217 1930 189 79 ...
## $ rape_revised_2016 : int NA NA NA 1915 1053
3304 2214 13702 3635 791 ...
## $ robbery_2016 : int NA NA NA 4687 850
7045 2125 54789 3525 2711 ...
## $ agg_assault_2016 : int NA NA NA 18869 401
1 21804 12007 104375 11681 4588 ...
## $ property_crime_total_2016 : int NA NA NA 143259 24
876 207317 98092 1002070 152146 64875 ...
## $ burglary_2016 : int NA NA NA 34045 405
3 38216 23814 188304 23825 10107 ...
## $ larceny_2016 : int NA NA NA 97498 177
66 150618 67091 637010 108680 47642 ...
## $ vehicle_theft_2016 : int NA NA NA 11716 305
7 18483 7187 176756 19641 7126 ...
## $ state_population_2006 : int NA NA NA 4599030 6
70053 6166318 2810872 36457549 4753377 3504809 ...
## $ violent_crime_total_2006 : int NA NA NA 19553 461
0 33456 15538 194483 18794 10505 ...
## $ murder_manslaughter_2006 : int NA NA NA 382 36 53
3 205 2486 171 137 ...
## $ rape_revised_2006 : int NA NA NA 1646 512
2449 1325 9235 2158 696 ...
## $ robbery_2006 : int NA NA NA 7062 600
9491 2749 71142 3824 4460 ...
## $ agg_assault_2006 : int NA NA NA 10463 346
2 20983 11259 111620 12641 5212 ...
## $ property_crime_total_2006 : int NA NA NA 181249 23

```

975 294389 112025 1159988 163439 90638 ...	
## \$ burglary_2006	: int NA NA NA 44780 415
5 59418 31864 247230 32411 15583 ...	
## \$ larceny_2006	: int NA NA NA 121451 17
284 178321 72646 669103 110372 64765 ...	
## \$ vehicle_theft_2006	: int NA NA NA 15018 253
6 56650 7515 243655 20656 10290 ...	
## \$ vap2016	: int 247219588 NA NA 37
55483 552166 5205215 2272904 30023902 4199509 2826827 ...	
## \$ percentage_disenfranchised_2016	: num 0.02 NA NA 0.08 0.
03 0.04 0.03 0.01 0.01 0.01 ...	
## \$ prison_2001	: int NA 149852 NA 24741
4570 27710 11489 157142 17278 17507 ...	
## \$ prison_2002	: int NA 158216 NA 25100
4351 29359 11849 159695 18636 18716 ...	
## \$ prison_2003	: int NA 168144 NA 27614
4472 31084 12068 161785 19450 18023 ...	
## \$ prison_2004	: int NA 177600 NA 25635
4534 32384 12577 163939 19655 17993 ...	
## \$ prison_2005	: int NA 186364 NA 24315
4798 33345 12455 168035 21063 17928 ...	
## \$ prison_2006	: int NA 190844 NA 24103
5052 35752 12854 172298 22051 18902 ...	
## \$ prison_2007	: int NA 197285 NA 25253
5151 37700 13275 171444 22666 19438 ...	
## \$ prison_2008	: int NA 198414 NA 25363
4997 39455 13135 171085 23211 19905 ...	
## \$ prison_2009	: int NA 205087 NA 27241
5472 40544 13338 168830 22725 18938 ...	
## \$ prison_2010	: int NA 206968 NA 27345
5369 40130 14192 162821 22752 18629 ...	
## \$ prison_2011	: int NA 214774 NA 26813
6216 39949 14090 147578 21862 17877 ...	
## \$ prison_2012	: int NA 216915 NA 26768
6308 40013 14043 132935 20328 17164 ...	
## \$ prison_2013	: int NA 214989 NA 26825
5081 41031 14295 134339 20184 17319 ...	
## \$ prison_2014	: int NA 209561 NA 26145
6323 42136 15250 134430 20469 16814 ...	
## \$ prison_2015	: int NA 195622 NA 25212
5247 42204 15784 127815 19959 16024 ...	
## \$ prison_2016	: int NA 188311 NA 23745
4378 42248 15833 129416 19486 15040 ...	
## \$ White_2016_fed_sent	: int NA 46951 NA NA NA
NA NA NA NA NA ...	
## \$ Black_2016_fed_sent	: int NA 59405 NA NA NA
NA NA NA NA NA ...	
## \$ Hisp_2016_fed_sent	: int NA 57432 NA NA NA
NA NA NA NA NA ...	
## \$ Native_2016_fed_sent	: int NA 3417 NA NA NA N

```
A NA NA NA NA ...
## $ Asian_Pac_2016_fed_sent : int NA 2293 NA NA NA N
A NA NA NA NA ...
```

ViewDataFrame

```
#View(maindata)
```

Summary - Identifying Ranges, Averages, Missing Values

```
summary(maindata)

##      jurisdiction prison_or_jail_2000 prison_or_jail_2010
## Alabama      : 1      Min.      : 140100      Min.      : 207000
## Alaska       : 1      1st Qu.: 658650      1st Qu.: 758900
## All_States: 1      Median :1177200      Median :1310800
## Arizona      : 1      Mean      :1085267      Mean      :1011867
## Arkansas     : 1      3rd Qu.:1557850      3rd Qu.:1414300
## California: 1      Max.      :1938500      Max.      :1517800
## (Other)      :47      NA's      :50          NA's      :50
## prison_or_jail_2015 total_corr_pop_2016
## Min.      : 195600      Min.      : 7400
## 1st Qu.: 719550      1st Qu.: 26450
## Median :1243500      Median : 82900
## Mean      :1201900      Mean      : 365814
## 3rd Qu.:1705050      3rd Qu.: 143400
## Max.      :2166600      Max.      :6582100
## NA's      :50          NA's      :2
## corr_sup_rate_per_100k_residents_all_ages_2016 prison_or_jail_2016
## Min.      : 100          Min.      : 1700
## 1st Qu.:1370          1st Qu.: 11300
## Median :1730          Median : 30400
## Mean      :1733          Mean      : 117030
## 3rd Qu.:2120          3rd Qu.: 55000
## Max.      :2880          Max.      :2131000
## NA's      :2
## prob_or_parole_2016 Incarc_rate_per_100k_US_residents_18_or_older
## Min.      : 5500      Min.      : 80.0
## 1st Qu.: 17450      1st Qu.: 540.0
## Median : 61400      Median : 780.0
## Mean      : 252025      Mean      : 747.9
## 3rd Qu.: 102250      3rd Qu.: 900.0
## Max.      :4537100      Max.      :1310.0
## NA's      :2
## Incarc_rate_per_100k_US_residents_all_ages state_pop_2016
## Min.      : 60.0          Min.      : 584910
## 1st Qu.:420.0          1st Qu.: 1848378
## Median :600.0          Median : 4561135
## Mean      :576.2          Mean      : 6454432
```

```

## 3rd Qu.:690.0                                3rd Qu.: 7187861
## Max.      :990.0                                Max.      :39296476
##                                                  NA's      :3
## violent_crime_total_2016 murder_manslaughter_2016 rape_revised_2016
## Min.      : 851                                Min.      : 14.00                                Min.      : 170.0
## 1st Qu.: 6133                                1st Qu.: 61.75                                1st Qu.: 856.5
## Median : 17786                                Median : 222.50                                Median : 1823.0
## Mean      : 25547                                Mean      : 345.54                                Mean      : 2637.6
## 3rd Qu.: 31045                                3rd Qu.: 525.75                                3rd Qu.: 2842.8
## Max.      :174796                                Max.      :1930.00                                Max.      :13702.0
## NA's      :3                                NA's      :3                                NA's      :3
## robbery_2016 agg_assault_2016 property_crime_total_2016
## Min.      : 59                                Min.      : 589                                Min.      : 9705
## 1st Qu.: 1026                                1st Qu.: 4069                                1st Qu.: 42564
## Median : 3628                                Median : 10930                                Median : 114817
## Mean      : 6586                                Mean      : 15978                                Mean      : 157923
## 3rd Qu.: 7249                                3rd Qu.: 18784                                3rd Qu.: 185352
## Max.      :54789                                Max.      :104375                                Max.      :1002070
## NA's      :3                                NA's      :3                                NA's      :3
## burglary_2016 larceny_2016 vehicle_theft_2016
## Min.      : 1771                                Min.      : 7637                                Min.      : 184
## 1st Qu.: 7126                                1st Qu.: 31156                                1st Qu.: 4412
## Median : 20546                                Median : 80805                                Median : 10144
## Mean      : 30281                                Mean      :112355                                Mean      : 15287
## 3rd Qu.: 35056                                3rd Qu.:129508                                3rd Qu.: 15039
## Max.      :188304                                Max.      :637010                                Max.      :176756
## NA's      :3                                NA's      :3                                NA's      :3
## state_population_2006 violent_crime_total_2006 murder_manslaughter_2006
## Min.      : 515004                                Min.      : 905                                Min.      : 13.0
## 1st Qu.: 1780866                                1st Qu.: 5340                                1st Qu.: 51.0
## Median : 4246921                                Median : 16974                                Median : 192.5
## Mean      : 5976339                                Mean      : 28527                                Mean      : 342.8
## 3rd Qu.: 6426844                                3rd Qu.: 33422                                3rd Qu.: 506.5
## Max.      :36457549                                Max.      :194483                                Max.      :2486.0
## NA's      :3                                NA's      :3                                NA's      :3
## rape_revised_2006 robbery_2006 agg_assault_2006
## Min.      : 151.0                                Min.      : 72                                Min.      : 604
## 1st Qu.: 585.5                                1st Qu.: 1168                                1st Qu.: 3624
## Median :1295.5                                Median : 4950                                Median : 10397
## Mean      :1885.7                                Mean      : 8908                                Mean      : 17390
## 3rd Qu.:2166.2                                3rd Qu.:10730                                3rd Qu.: 22175
## Max.      :9235.0                                Max.      :71142                                Max.      :111620
## NA's      :3                                NA's      :3                                NA's      :3
## property_crime_total_2006 burglary_2006 larceny_2006
## Min.      : 13359                                Min.      : 2322                                Min.      : 9672
## 1st Qu.: 55538                                1st Qu.: 10046                                1st Qu.: 39516
## Median : 145900                                Median : 31085                                Median : 99811
## Mean      : 199851                                Mean      : 43823                                Mean      :132209
## 3rd Qu.: 243397                                3rd Qu.: 54980                                3rd Qu.:161606
## Max.      :1159988                                Max.      :247230                                Max.      :669103

```

```

## NA's :3          NA's :3          NA's :3
## vehicle_theft_2006  vap2016          percentage_disenfranchised_2016
## Min. : 620      Min. : 447212      Min. :0.00000
## 1st Qu.: 5074    1st Qu.: 1445192    1st Qu.:0.01000
## Median : 14332    Median : 3555911    Median :0.01000
## Mean : 23818      Mean : 9694886      Mean :0.02294
## 3rd Qu.: 29033    3rd Qu.: 6035540    3rd Qu.:0.02500
## Max. :243655      Max. :247219588      Max. :0.10000
## NA's :3          NA's :2          NA's :2
## prison_2001      prison_2002      prison_2003      prison_2004
## Min. : 1088      Min. : 1117      Min. : 1100      Min. : 1130
## 1st Qu.: 5193     1st Qu.: 5260     1st Qu.: 5528     1st Qu.: 5656
## Median : 15226     Median : 16263     Median : 16227     Median : 16503
## Mean : 26075       Mean : 26811       Mean : 27336       Mean : 27879
## 3rd Qu.: 28142     3rd Qu.: 29577     3rd Qu.: 30390     3rd Qu.: 30877
## Max. :157142       Max. :159695       Max. :168144       Max. :177600
## NA's :2          NA's :2          NA's :2          NA's :2
## prison_2005      prison_2006      prison_2007      prison_2008
## Min. : 1318      Min. : 1337      Min. : 1392      Min. : 1379
## 1st Qu.: 5994     1st Qu.: 6074     1st Qu.: 5943     1st Qu.: 5888
## Median : 16781     Median : 16827     Median : 17478     Median : 17896
## Mean : 28381       Mean : 29184       Mean : 29674       Mean : 29862
## 3rd Qu.: 30568     3rd Qu.: 30470     3rd Qu.: 31400     3rd Qu.: 31361
## Max. :186364       Max. :190844       Max. :197285       Max. :198414
## NA's :2          NA's :2          NA's :2          NA's :2
## prison_2009      prison_2010      prison_2011      prison_2012
## Min. : 1436      Min. : 1416      Min. : 1385      Min. : 1413
## 1st Qu.: 6026     1st Qu.: 5910     1st Qu.: 6381     1st Qu.: 6474
## Median : 17255     Median : 17028     Median : 17109     Median : 16919
## Mean : 29895       Mean : 29832       Mean : 29510       Mean : 29096
## 3rd Qu.: 30757     3rd Qu.: 31244     3rd Qu.: 30750     3rd Qu.: 30456
## Max. :205087       Max. :206968       Max. :214774       Max. :216915
## NA's :2          NA's :2          NA's :2          NA's :2
## prison_2013      prison_2014      prison_2015      prison_2016
## Min. : 1576      Min. : 1696      Min. : 1750      Min. : 1639
## 1st Qu.: 6238     1st Qu.: 6526     1st Qu.: 6181     1st Qu.: 6116
## Median : 17319     Median : 16814     Median : 16024     Median : 15059
## Mean : 29123       Mean : 29006       Mean : 28249       Mean : 27785
## 3rd Qu.: 30742     3rd Qu.: 32198     3rd Qu.: 31362     3rd Qu.: 31154
## Max. :214989       Max. :209561       Max. :195622       Max. :188311
## NA's :2          NA's :2          NA's :2          NA's :2
## White_2016_fed_sent Black_2016_fed_sent Hisp_2016_fed_sent
## Min. :46951      Min. :59405      Min. :57432
## 1st Qu.:46951     1st Qu.:59405     1st Qu.:57432
## Median :46951     Median :59405     Median :57432
## Mean :46951       Mean :59405       Mean :57432
## 3rd Qu.:46951     3rd Qu.:59405     3rd Qu.:57432
## Max. :46951       Max. :59405       Max. :57432
## NA's :52         NA's :52         NA's :52
## Native_2016_fed_sent Asian_Pac_2016_fed_sent

```

```
## Min. :3417      Min. :2293
## 1st Qu.:3417    1st Qu.:2293
## Median :3417    Median :2293
## Mean   :3417    Mean   :2293
## 3rd Qu.:3417    3rd Qu.:2293
## Max.   :3417    Max.   :2293
## NA's   :52      NA's   :52
```

Head

```
head(maindata)
```

```
## jurisdiction prison_or_jail_2000 prison_or_jail_2010 prison_or_jail_2015
## 1 US_ALL 1938500 1517800 2166600
## 2 US_Federal 140100 207000 195600
## 3 All_States 1177200 1310800 1243500
## 4 Alabama NA NA NA
## 5 Alaska NA NA NA
## 6 Arizona NA NA NA
## total_corr_pop_2016 corr_sup_rate_per_100k_residents_all_ages_2016
## 1 6582100 2030
## 2 320000 100
## 3 6262000 1930
## 4 99800 2050
## 5 12900 1730
## 6 137500 1970
## prison_or_jail_2016 prob_or_parole_2016
## 1 2131000 4537100
## 2 188400 131700
## 3 1942600 4405400
## 4 40900 60700
## 5 4400 8400
## 6 55000 84800
## Incarc_rate_per_100k_US_residents_18_or_older
## 1 850
## 2 80
## 3 780
## 4 1080
## 5 800
## 6 1030
## Incarc_rate_per_100k_US_residents_all_ages state_pop_2016
## 1 660 NA
## 2 60 NA
## 3 600 NA
## 4 840 4860545
## 5 600 741522
## 6 790 6908642
## violent_crime_total_2016 murder_manslaughter_2016 rape_revised_2016
## 1 NA NA NA
## 2 NA NA NA
```


## 3		NA		NA		NA
## 4		25878		407		1915
## 5		5966		52		1053
## 6		32542		389		3304
##	robbery_2016	agg_assault_2016	property_crime_total_2016	burglary_2016		
## 1	NA	NA	NA	NA		NA
## 2	NA	NA	NA	NA		NA
## 3	NA	NA	NA	NA		NA
## 4	4687	18869	143259	34045		
## 5	850	4011	24876	4053		
## 6	7045	21804	207317	38216		
##	larceny_2016	vehicle_theft_2016	state_population_2006			
## 1	NA	NA	NA			
## 2	NA	NA	NA			
## 3	NA	NA	NA			
## 4	97498	11716	4599030			
## 5	17766	3057	670053			
## 6	150618	18483	6166318			
##	violent_crime_total_2006	murder_manslaughter_2006	rape_revised_2006			
## 1	NA	NA	NA			
## 2	NA	NA	NA			
## 3	NA	NA	NA			
## 4	19553	382	1646			
## 5	4610	36	512			
## 6	33456	533	2449			
##	robbery_2006	agg_assault_2006	property_crime_total_2006	burglary_2006		
## 1	NA	NA	NA	NA		
## 2	NA	NA	NA	NA		
## 3	NA	NA	NA	NA		
## 4	7062	10463	181249	44780		
## 5	600	3462	23975	4155		
## 6	9491	20983	294389	59418		
##	larceny_2006	vehicle_theft_2006	vap2016			
## 1	NA	NA	247219588			
## 2	NA	NA	NA			
## 3	NA	NA	NA			
## 4	121451	15018	3755483			
## 5	17284	2536	552166			
## 6	178321	56650	5205215			
##	percentage_disenfranchised_2016	prison_2001	prison_2002	prison_2003		
## 1	0.02	NA	NA	NA		
## 2	NA	149852	158216	168144		
## 3	NA	NA	NA	NA		
## 4	0.08	24741	25100	27614		
## 5	0.03	4570	4351	4472		
## 6	0.04	27710	29359	31084		
##	prison_2004	prison_2005	prison_2006	prison_2007	prison_2008	prison_2009
## 1	NA	NA	NA	NA	NA	NA
## 2	177600	186364	190844	197285	198414	205087
## 3	NA	NA	NA	NA	NA	NA

```

## 4      25635      24315      24103      25253      25363      27241
## 5      4534      4798      5052      5151      4997      5472
## 6      32384     33345     35752     37700     39455     40544
##  prison_2010 prison_2011 prison_2012 prison_2013 prison_2014 prison_2015
## 1      NA      NA      NA      NA      NA      NA
## 2     206968     214774     216915     214989     209561     195622
## 3      NA      NA      NA      NA      NA      NA
## 4     27345     26813     26768     26825     26145     25212
## 5      5369     6216     6308     5081     6323     5247
## 6     40130     39949     40013     41031     42136     42204
##  prison_2016 White_2016_fed_sent Black_2016_fed_sent Hisp_2016_fed_sent
## 1      NA      NA      NA      NA
## 2     188311      46951      59405      57432
## 3      NA      NA      NA      NA
## 4     23745      NA      NA      NA
## 5      4378      NA      NA      NA
## 6     42248      NA      NA      NA
##  Native_2016_fed_sent Asian_Pac_2016_fed_sent
## 1      NA      NA
## 2      3417      2293
## 3      NA      NA
## 4      NA      NA
## 5      NA      NA
## 6      NA      NA

```

Tail

```
tail(maindata)
```

```

##      jurisdiction prison_or_jail_2000 prison_or_jail_2010
## 48      Vermont      NA      NA
## 49      Virginia      NA      NA
## 50      Washington      NA      NA
## 51 West_Virginia      NA      NA
## 52      Wisconsin      NA      NA
## 53      Wyoming      NA      NA
##  prison_or_jail_2015 total_corr_pop_2016
## 48      NA      7400
## 49      NA     120000
## 50      NA     127000
## 51      NA      20100
## 52      NA     100500
## 53      NA      9400
##  corr_sup_rate_per_100k_residents_all_ages_2016 prison_or_jail_2016
## 48      1180      1700
## 49      1420     57500
## 50      1730     30400
## 51      1100     10100
## 52      1740     35600
## 53      1600      3900

```

##	prob_or_parole_2016	Incarc_rate_per_100k_US_residents_18_or_older	
## 48	5800		340
## 49	62500		880
## 50	100600		530
## 51	10100		690
## 52	64900		790
## 53	5500		870
##	Incarc_rate_per_100k_US_residents_all_ages	state_pop_2016	
## 48	280	623354	
## 49	680	8414380	
## 50	410	7280934	
## 51	550	1828637	
## 52	620	5772917	
## 53	660	584910	
##	violent_crime_total_2016	murder_manslaughter_2016	rape_revised_2016
## 48	851	14	170
## 49	18495	482	2830
## 50	22101	195	3133
## 51	6633	85	684
## 52	17716	232	2005
## 53	1431	20	206
##	robbery_2016	agg_assault_2016	property_crime_total_2016
## 48	78	589	9705
## 49	4826	10357	157292
## 50	5649	13124	254994
## 51	720	5144	37282
## 52	4707	10772	111911
## 53	59	1146	11460
##	burglary_2016		
## 48			1884
## 49			20159
## 50			49249
## 51			9127
## 52			19498
## 53			1771
##	larceny_2016	vehicle_theft_2016	state_population_2006
## 48	7637	184	623908
## 49	127285	9848	7642884
## 50	173423	32322	6395798
## 51	25657	2498	1818470
## 52	82455	9958	5556506
## 53	8889	800	515004
##	violent_crime_total_2006	murder_manslaughter_2006	rape_revised_2006
## 48	905	14	173
## 49	21683	403	1829
## 50	22303	197	2748
## 51	5113	80	404
## 52	15899	165	1233
## 53	1307	13	151
##	robbery_2006	agg_assault_2006	property_crime_total_2006
## 48	114	604	14920
## 49	7779	11672	190257
## 50	6424	12934	288456
## 51	853	3776	47941
## 52	5564	8937	156748
## 53	72	1071	15381
##	burglary_2006		
## 48			3443
## 49			32166
## 50			58843
## 51			11579
## 52			27038
## 53			2322
##	larceny_2006	vehicle_theft_2006	vap2016
##			percentage_disenfranchised_2016

##	48	10857		620	506119		0.00
##	49	143292		14799	6512571		0.08
##	50	183179		46434	5558509		0.01
##	51	32305		4057	1464532		0.01
##	52	115682		14028	4476711		0.01
##	53	12255		804	447212		0.05
##		prison_2001	prison_2002	prison_2003	prison_2004	prison_2005	prison_2006
##	48	1400	1361	1491	1968	2078	2215
##	49	31472	30937	30915	31072	30722	30823
##	50	15226	16263	16227	16503	16781	16827
##	51	3403	3575	3790	3990	4036	4367
##	52	20766	21666	22153	22152	21921	22564
##	53	1487	1641	1630	1765	2096	2035
##		prison_2007	prison_2008	prison_2009	prison_2010	prison_2011	prison_2012
##	48	2116	2116	2220	2079	2053	2034
##	49	32977	32585	30995	31911	30531	29708
##	50	16911	17188	17255	17028	17109	16919
##	51	4907	4896	5062	5072	5149	5335
##	52	22839	22282	22262	22033	22388	22419
##	53	1927	1704	1743	2092	2162	2187
##		prison_2013	prison_2014	prison_2015	prison_2016	White_2016_fed_sent	
##	48	2078	1979	1750	1735		NA
##	49	29985	30050	30430	29882		NA
##	50	17760	17180	17222	17228		NA
##	51	5708	5867	5925	5899		NA
##	52	22443	22572	22914	23163		NA
##	53	2288	2369	2400	2352		NA
##		Black_2016_fed_sent	Hispanic_2016_fed_sent	Native_2016_fed_sent			
##	48		NA	NA		NA	
##	49		NA	NA		NA	
##	50		NA	NA		NA	
##	51		NA	NA		NA	
##	52		NA	NA		NA	
##	53		NA	NA		NA	
##		Asian_Pac_2016_fed_sent					
##	48		NA				
##	49		NA				
##	50		NA				
##	51		NA				
##	52		NA				
##	53		NA				

Missing Values

##		jurisdiction
##		0
##		prison_or_jail_2000
##		50
##		prison_or_jail_2010

```

##                                     50
##                                prison_or_jail_2015
##                                     50
##                                total_corr_pop_2016
##                                     2
## corr_sup_rate_per_100k_residents_all_ages_2016
##                                     2
##                                prison_or_jail_2016
##                                     0
##                                prob_or_parole_2016
##                                     2
## Incarc_rate_per_100k_US_residents_18_or_older
##                                     0
##      Incarc_rate_per_100k_US_residents_all_ages
##                                     0
##                                state_pop_2016
##                                     3
##                                violent_crime_total_2016
##                                     3
##                                murder_manslaughter_2016
##                                     3
##                                rape_revised_2016
##                                     3
##                                robbery_2016
##                                     3
##                                agg_assault_2016
##                                     3
##                                property_crime_total_2016
##                                     3
##                                burglary_2016
##                                     3
##                                larceny_2016
##                                     3
##                                vehicle_theft_2016
##                                     3
##                                state_population_2006
##                                     3
##                                violent_crime_total_2006
##                                     3
##                                murder_manslaughter_2006
##                                     3
##                                rape_revised_2006
##                                     3
##                                robbery_2006
##                                     3
##                                agg_assault_2006
##                                     3
##                                property_crime_total_2006
##                                     3
##                                burglary_2006

```

```

##                                     3
##                                larceny_2006
##                                     3
##                                vehicle_theft_2006
##                                     3
##                                vap2016
##                                     2
##    percentage_disenfranchised_2016
##                                     2
##                                prison_2001
##                                     2
##                                prison_2002
##                                     2
##                                prison_2003
##                                     2
##                                prison_2004
##                                     2
##                                prison_2005
##                                     2
##                                prison_2006
##                                     2
##                                prison_2007
##                                     2
##                                prison_2008
##                                     2
##                                prison_2009
##                                     2
##                                prison_2010
##                                     2
##                                prison_2011
##                                     2
##                                prison_2012
##                                     2
##                                prison_2013
##                                     2
##                                prison_2014
##                                     2
##                                prison_2015
##                                     2
##                                prison_2016
##                                     2
##                                White_2016_fed_sent
##                                     52
##                                Black_2016_fed_sent
##                                     52
##                                Hisp_2016_fed_sent
##                                     52
##                                Native_2016_fed_sent
##                                     52

```

```
## Asian_Pac_2016_fed_sent
## 52
```

1. What is the total prison population in relation to the entire United States population? (There are some variables that I can use to infer the overall population or to sum and see the overall population, but I may want to find a population-specific dataset.)

The incarceration rate for the entire US population aged 18 and older is the first value in rate: 850 per 100,000 residents.

```
rate <- maindata$Incarc_rate_per_100k_US_residents_18_or_older
rate
## [1] 850 80 780 1080 800 1030 1050 670 740 530 880 900 1160 500
## [15] 900 620 850 560 780 1010 1270 380 610 360 730 380 1260 940
## [29] 700 610 890 410 460 930 480 680 540 790 1310 640 810 370
## [43] 820 880 930 1050 540 340 880 530 690 790 870
```

Finding a percentage value for 850 / 100,000

```
popprct <- 850 / 100000
popprct
## [1] 0.0085
```

Inferring 18+ population of the US based on the rate of imprisonment and the number of reported prisoners (from prison_or_jail_2016)

```
prison_pop_over_18 <- 2131000
us_pop_over_18 <- prison_pop_over_18 / popprct
us_pop_over_18
## [1] 250705882
```

Based on the figures in the dataset, the US population of people 18 and over is 250,705,882. As previously determined, the incarceration rate is 850 people per 100,000 people 18+.

```
prison_pct <- prison_pop_over_18 / us_pop_over_18
prison_pct
## [1] 0.0085
```

The percentage of adults 18 and older in the US in 2016 that were imprisoned can be stated as a raw percentage as well: 0.85% of the US population, or less than 1%.

2. How has the U.S. prison population changed over time?

Prison or Jail Population 2000, 2010, 2015, 2016 Subset

```
##
## Attaching package: 'dplyr'

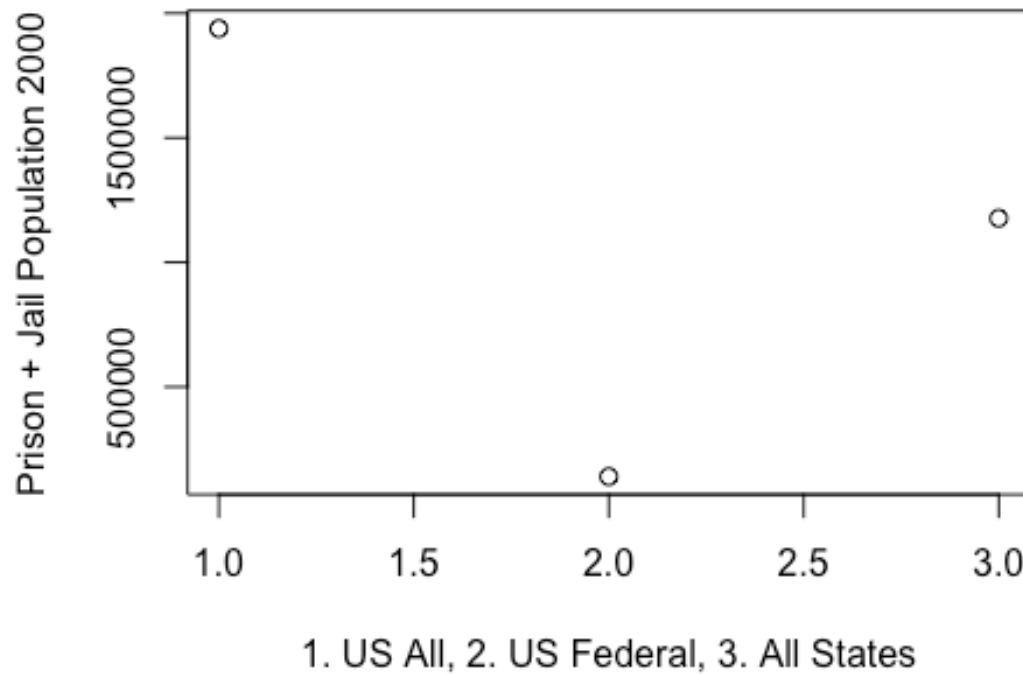
## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

## Observations: 3
## Variables: 5
## $ jurisdiction      <fct> US_ALL, US_Federal, All_States
## $ prison_or_jail_2000 <int> 1938500, 140100, 1177200
## $ prison_or_jail_2010 <int> 1517800, 207000, 1310800
## $ prison_or_jail_2015 <int> 2166600, 195600, 1243500
## $ prison_or_jail_2016 <int> 2131000, 188400, 1942600

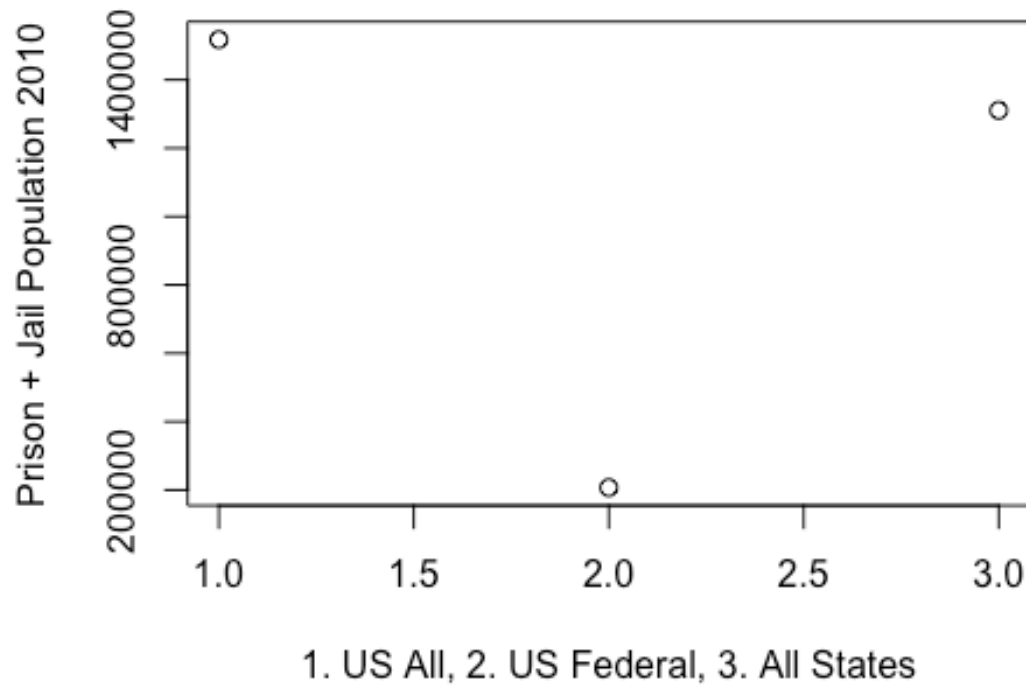
plot(prison_or_jail$prison_or_jail_2000, main = "US Prison + Jail Population
2000", xlab = "1. US All, 2. US Federal, 3. All States", ylab="Prison + Jail
Population 2000")
```


US Prison + Jail Population 2000



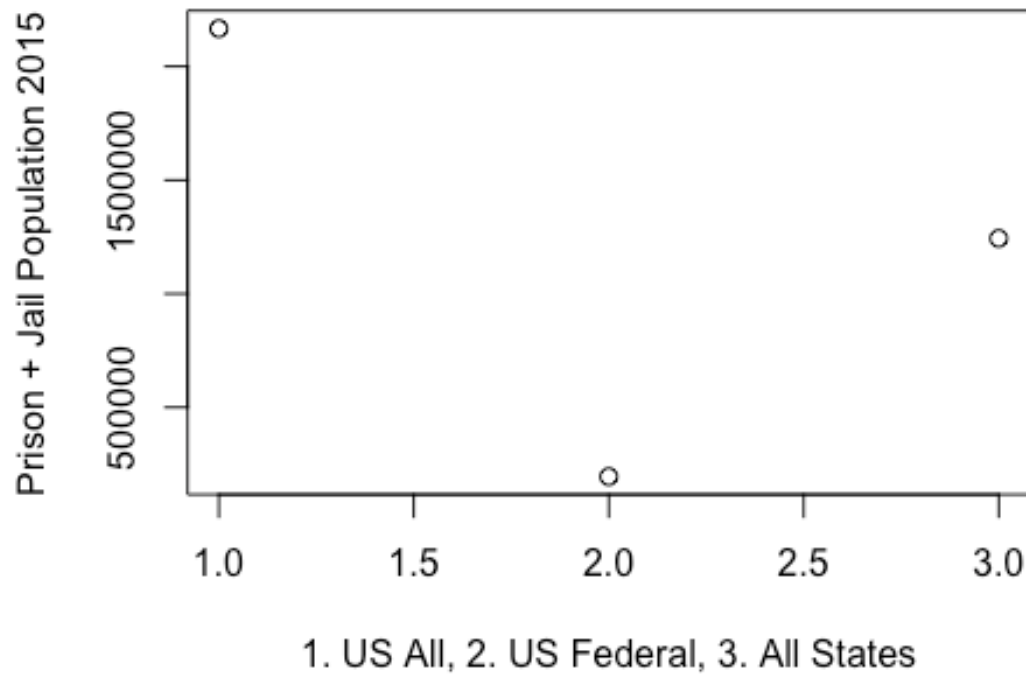
```
plot(prison_or_jail$prison_or_jail_2010, main = "US Prison + Jail Population  
2010", xlab = "1. US All, 2. US Federal, 3. All States", ylab="Prison + Jail  
Population 2010")
```

US Prison + Jail Population 2010



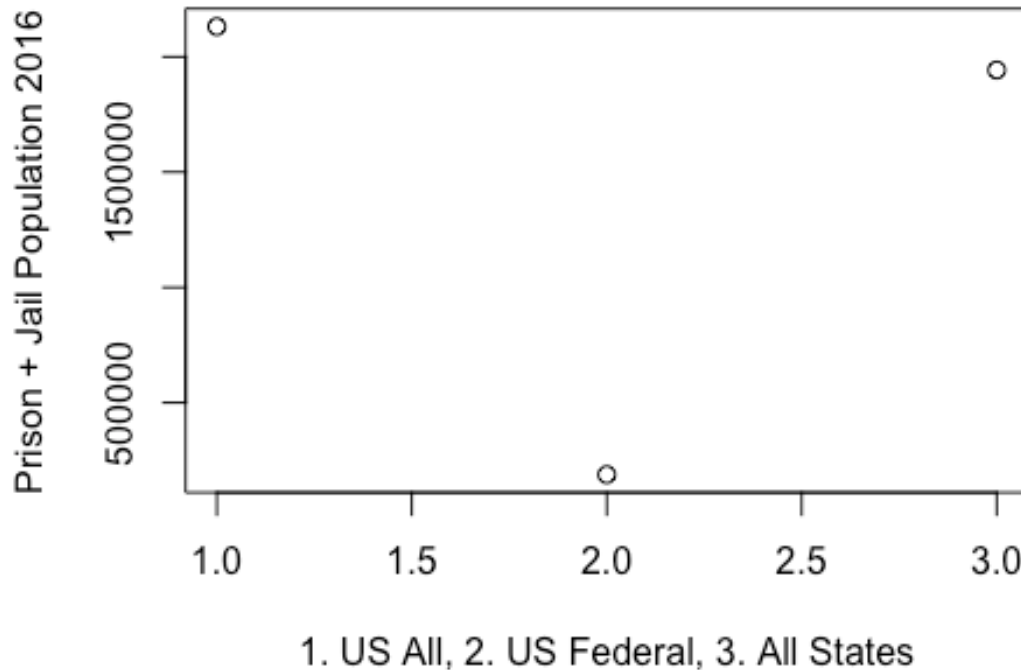
```
plot(prison_or_jail$prison_or_jail_2015, main = "US Prison + Jail Population 2015", xlab = "1. US All, 2. US Federal, 3. All States", ylab="Prison + Jail Population 2015")
```

US Prison + Jail Population 2015



```
plot(prison_or_jail$prison_or_jail_2016, main = "US Prison + Jail Population 2016", xlab = "1. US All, 2. US Federal, 3. All States", ylab="Prison + Jail Population 2016")
```

US Prison + Jail Population 2016



#50

State + Federal Prison Population 2001 - 2016 Subset

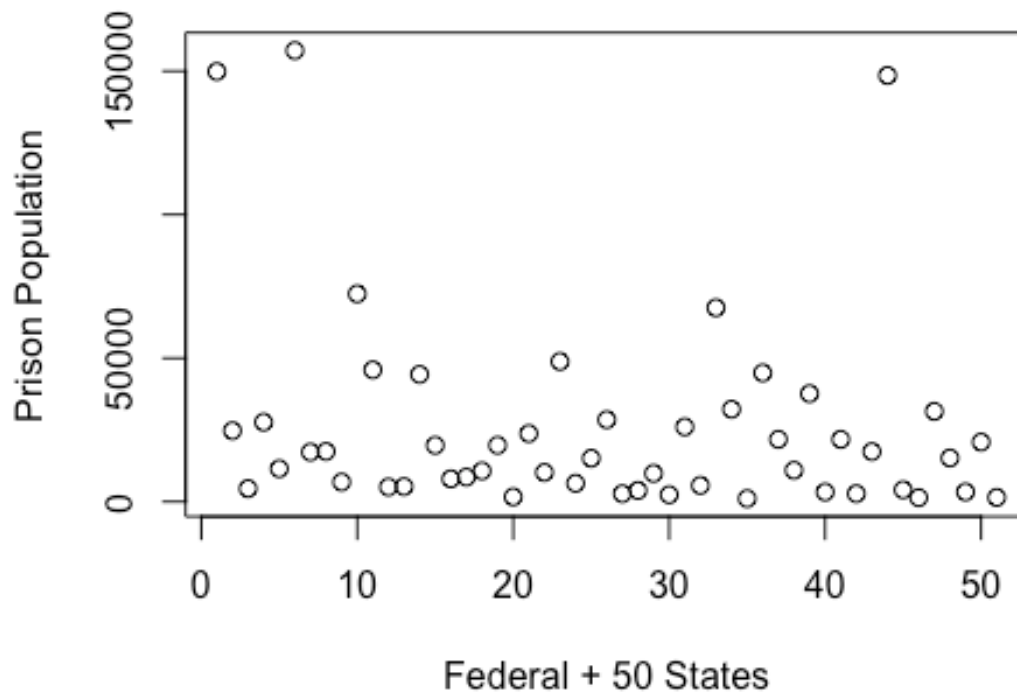
```
## Observations: 51
## Variables: 17
## $ jurisdiction <fct> US_Federal, Alabama, Alaska, Arizona, Arkansas, Cal...
## $ prison_2001 <int> 149852, 24741, 4570, 27710, 11489, 157142, 17278, 1...
## $ prison_2002 <int> 158216, 25100, 4351, 29359, 11849, 159695, 18636, 1...
## $ prison_2003 <int> 168144, 27614, 4472, 31084, 12068, 161785, 19450, 1...
## $ prison_2004 <int> 177600, 25635, 4534, 32384, 12577, 163939, 19655, 1...
## $ prison_2005 <int> 186364, 24315, 4798, 33345, 12455, 168035, 21063, 1...
## $ prison_2006 <int> 190844, 24103, 5052, 35752, 12854, 172298, 22051, 1...
## $ prison_2007 <int> 197285, 25253, 5151, 37700, 13275, 171444, 22666, 1...
## $ prison_2008 <int> 198414, 25363, 4997, 39455, 13135, 171085, 23211, 1...
## $ prison_2009 <int> 205087, 27241, 5472, 40544, 13338, 168830, 22725, 1...
## $ prison_2010 <int> 206968, 27345, 5369, 40130, 14192, 162821, 22752, 1...
## $ prison_2011 <int> 214774, 26813, 6216, 39949, 14090, 147578, 21862, 1...
## $ prison_2012 <int> 216915, 26768, 6308, 40013, 14043, 132935, 20328, 1...
## $ prison_2013 <int> 214989, 26825, 5081, 41031, 14295, 134339, 20184, 1...
## $ prison_2014 <int> 209561, 26145, 6323, 42136, 15250, 134430, 20469, 1...
## $ prison_2015 <int> 195622, 25212, 5247, 42204, 15784, 127815, 19959, 1...
## $ prison_2016 <int> 188311, 23745, 4378, 42248, 15833, 129416, 19486, 1...

##      jurisdiction  prison_2001      prison_2002      prison_2003
## Alabama      : 1      Min.      : 1088      Min.      : 1117      Min.      : 1100
```

```
## Alaska      : 1      1st Qu.: 5193      1st Qu.: 5260      1st Qu.: 5528
## Arizona     : 1      Median : 15226     Median : 16263     Median : 16227
## Arkansas    : 1      Mean      : 26075     Mean      : 26811     Mean      : 27336
## California: 1      3rd Qu.: 28142     3rd Qu.: 29577     3rd Qu.: 30390
## Colorado    : 1      Max.      :157142     Max.      :159695     Max.      :168144
## (Other)     :45
## prison_2004      prison_2005      prison_2006      prison_2007
## Min.      : 1130      Min.      : 1318      Min.      : 1337      Min.      : 1392
## 1st Qu.: 5656      1st Qu.: 5994      1st Qu.: 6074      1st Qu.: 5943
## Median : 16503     Median : 16781     Median : 16827     Median : 17478
## Mean      : 27879     Mean      : 28381     Mean      : 29184     Mean      : 29674
## 3rd Qu.: 30877     3rd Qu.: 30568     3rd Qu.: 30470     3rd Qu.: 31400
## Max.      :177600     Max.      :186364     Max.      :190844     Max.      :197285
##
## prison_2008      prison_2009      prison_2010      prison_2011
## Min.      : 1379     Min.      : 1436     Min.      : 1416     Min.      : 1385
## 1st Qu.: 5888      1st Qu.: 6026      1st Qu.: 5910      1st Qu.: 6381
## Median : 17896     Median : 17255     Median : 17028     Median : 17109
## Mean      : 29862     Mean      : 29895     Mean      : 29832     Mean      : 29510
## 3rd Qu.: 31361     3rd Qu.: 30757     3rd Qu.: 31244     3rd Qu.: 30750
## Max.      :198414     Max.      :205087     Max.      :206968     Max.      :214774
##
## prison_2012      prison_2013      prison_2014      prison_2015
## Min.      : 1413     Min.      : 1576     Min.      : 1696     Min.      : 1750
## 1st Qu.: 6474      1st Qu.: 6238      1st Qu.: 6526      1st Qu.: 6181
## Median : 16919     Median : 17319     Median : 16814     Median : 16024
## Mean      : 29096     Mean      : 29123     Mean      : 29006     Mean      : 28249
## 3rd Qu.: 30456     3rd Qu.: 30742     3rd Qu.: 32198     3rd Qu.: 31362
## Max.      :216915     Max.      :214989     Max.      :209561     Max.      :195622
##
## prison_2016
## Min.      : 1639
## 1st Qu.: 6116
## Median : 15059
## Mean      : 27785
## 3rd Qu.: 31154
## Max.      :188311
##
```

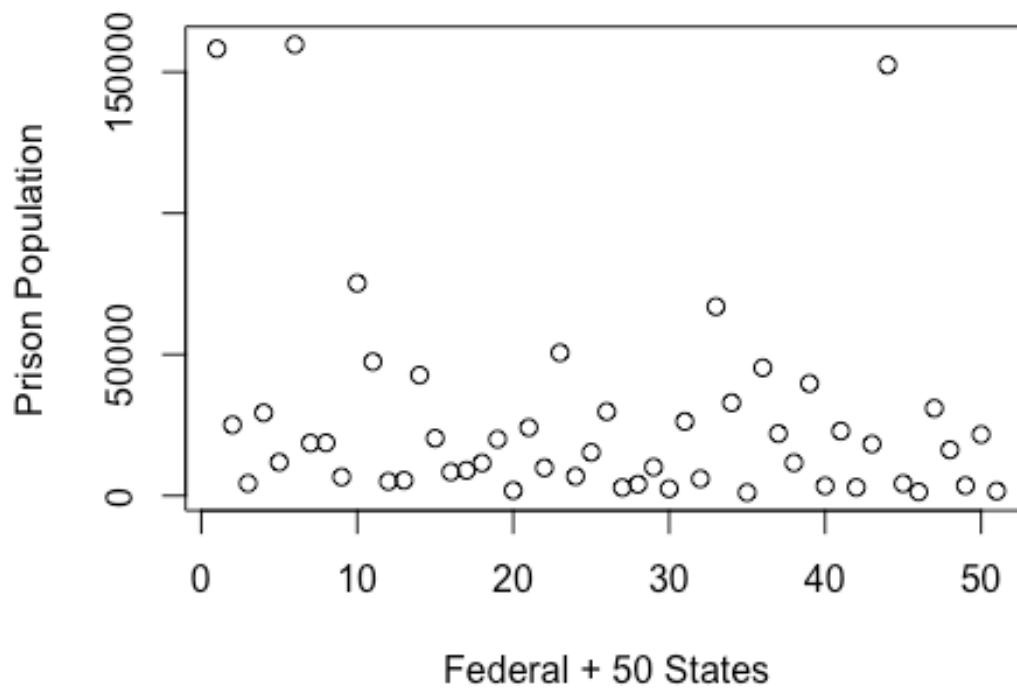
```
plot(allstates$prison_2001, main = "2001 Prison Population by Federal and Individual State", xlab = "Federal + 50 States", ylab="Prison Population")
```

2001 Prison Population by Federal and Individual St



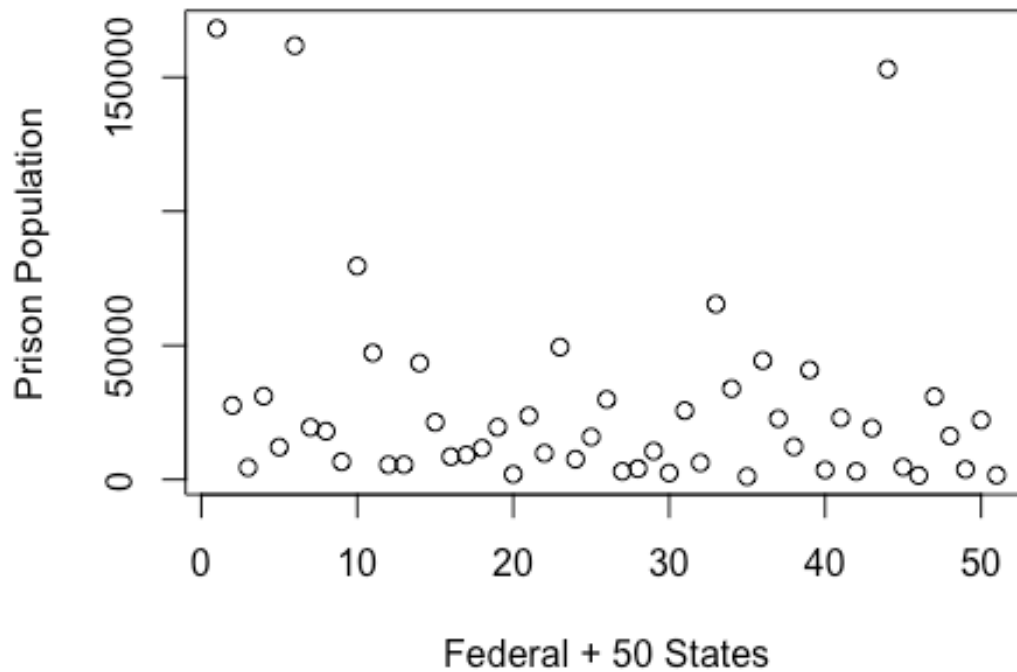
```
plot(allstates$prison_2002, main = "2002 Prison Population by Federal and Individual State", xlab = "Federal + 50 States", ylab="Prison Population")
```

2002 Prison Population by Federal and Individual St



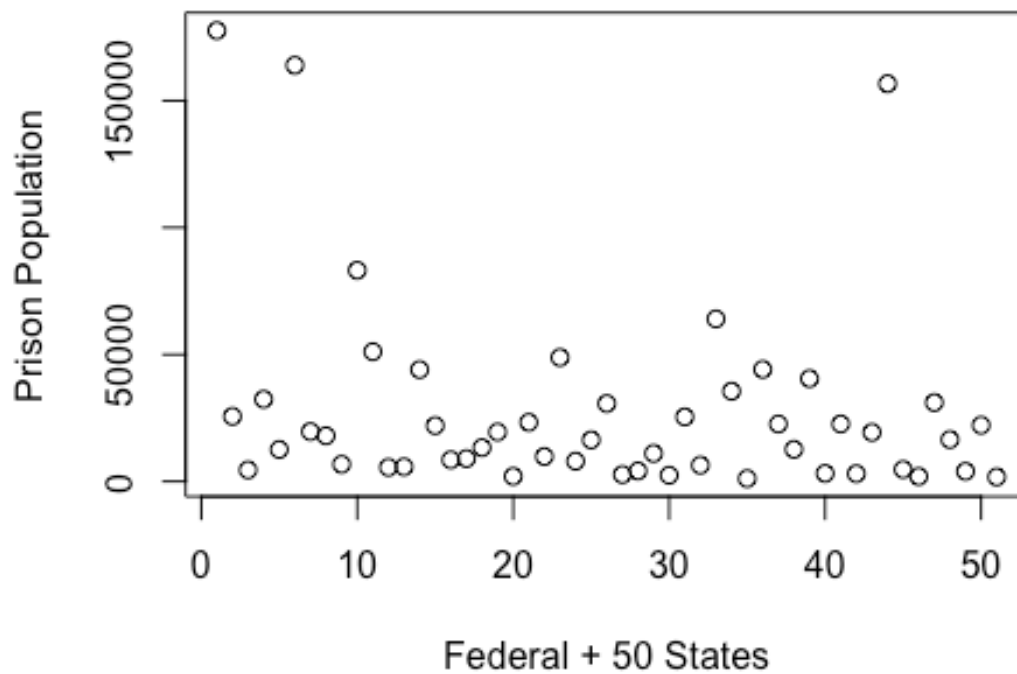
```
plot(allstates$prison_2003, main = "2003 Prison Population by Federal and Individual State", xlab = "Federal + 50 States", ylab="Prison Population")
```

2003 Prison Population by Federal and Individual St



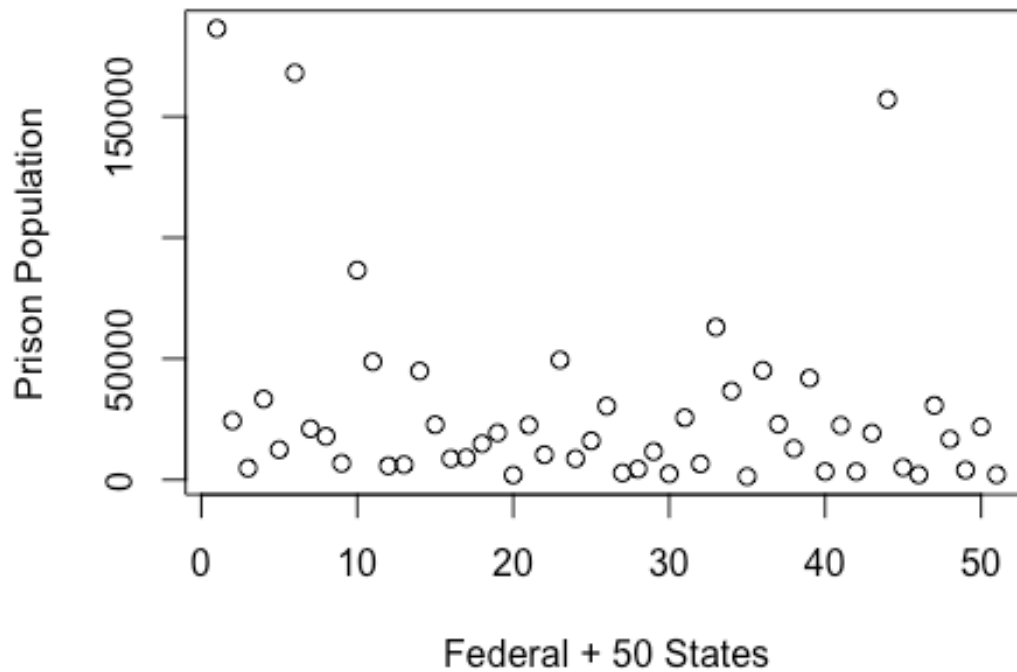
```
plot(allstates$prison_2004, main = "2004 Prison Population by Federal and Individual State", xlab = "Federal + 50 States", ylab="Prison Population")
```


2004 Prison Population by Federal and Individual St



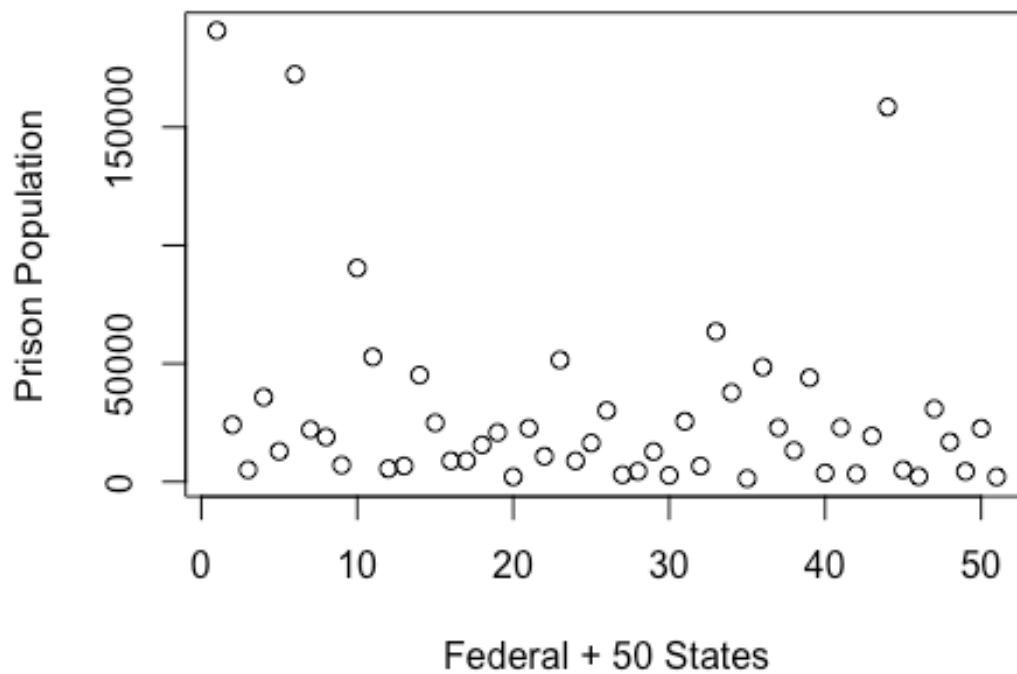
```
plot(allstates$prison_2005, main = "2005 Prison Population by Federal and Individual State", xlab = "Federal + 50 States", ylab="Prison Population")
```

2005 Prison Population by Federal and Individual St



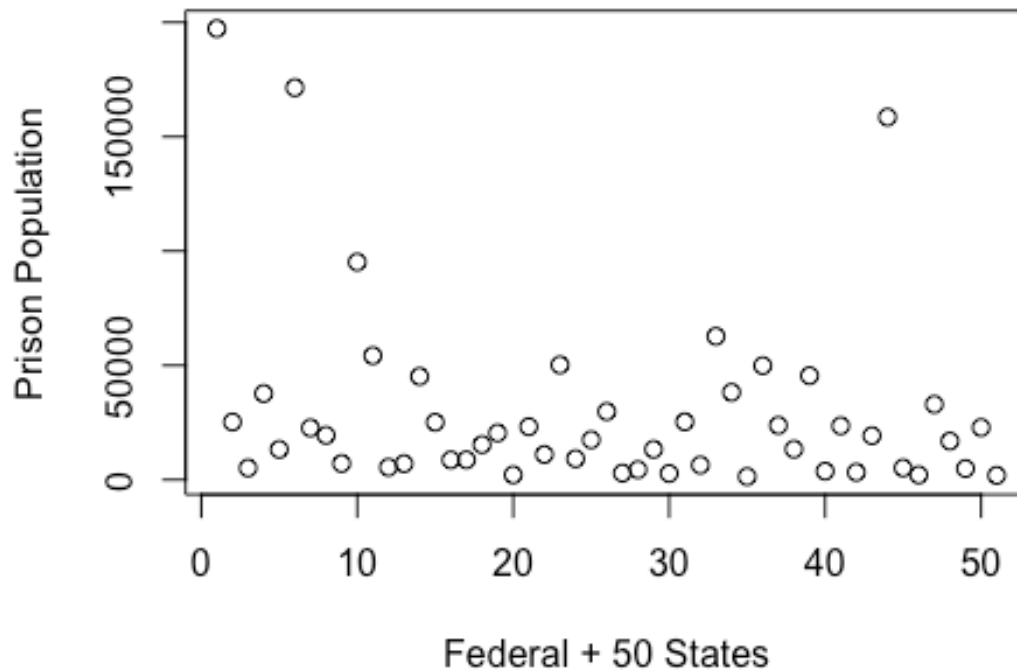
```
plot(allstates$prison_2006, main = "2006 Prison Population by Federal and Individual State", xlab = "Federal + 50 States", ylab="Prison Population")
```

2006 Prison Population by Federal and Individual St



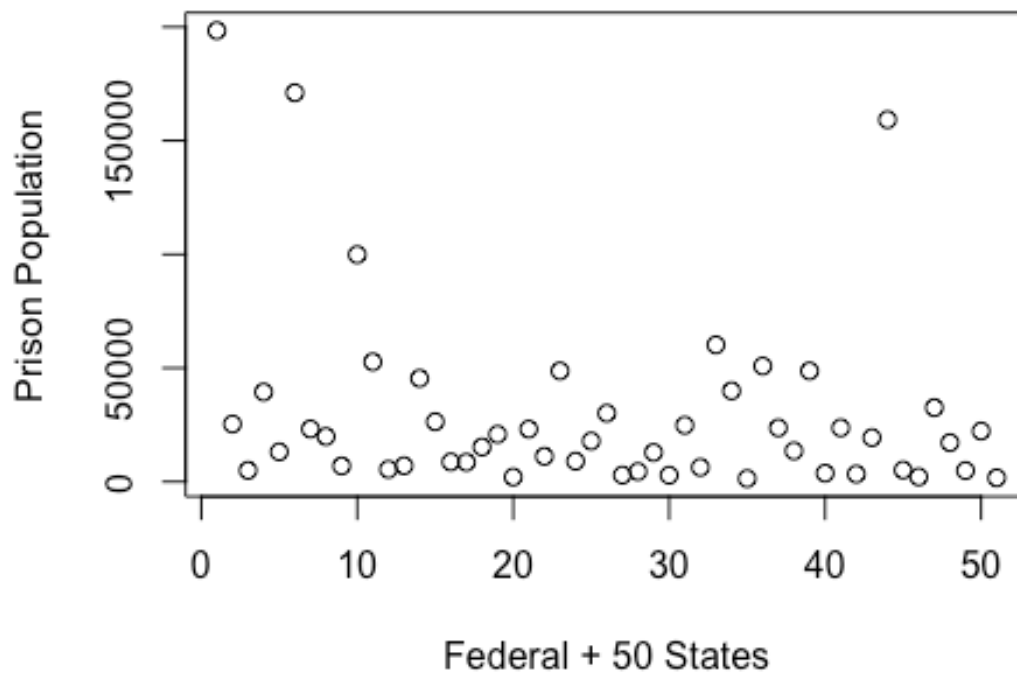
```
plot(allstates$prison_2007, main = "2007 Prison Population by Federal and Individual State", xlab = "Federal + 50 States", ylab="Prison Population")
```

2007 Prison Population by Federal and Individual St



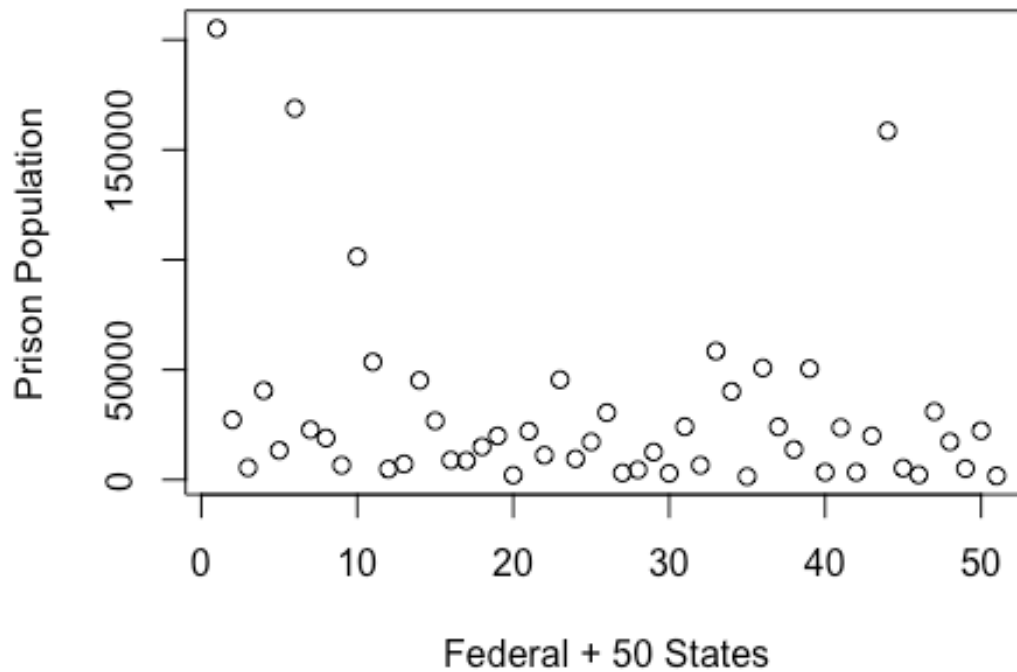
```
plot(allstates$prison_2008, main = "2008 Prison Population by Federal and Individual State", xlab = "Federal + 50 States", ylab="Prison Population")
```

2008 Prison Population by Federal and Individual St



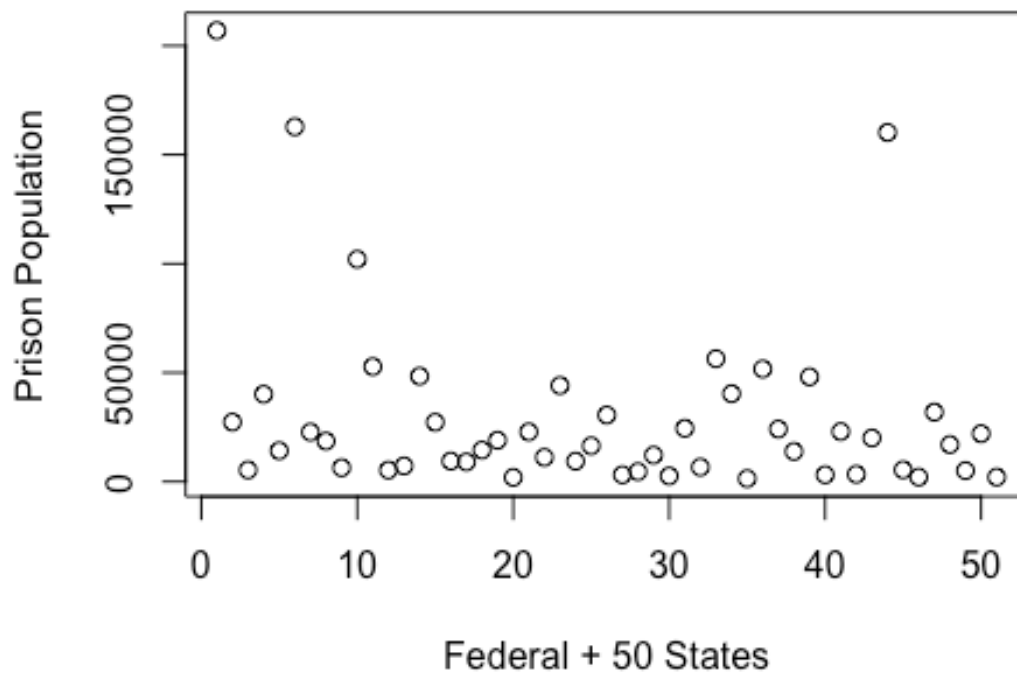
```
plot(allstates$prison_2009, main = "2009 Prison Population by Federal and Individual State", xlab = "Federal + 50 States", ylab="Prison Population")
```

2009 Prison Population by Federal and Individual St



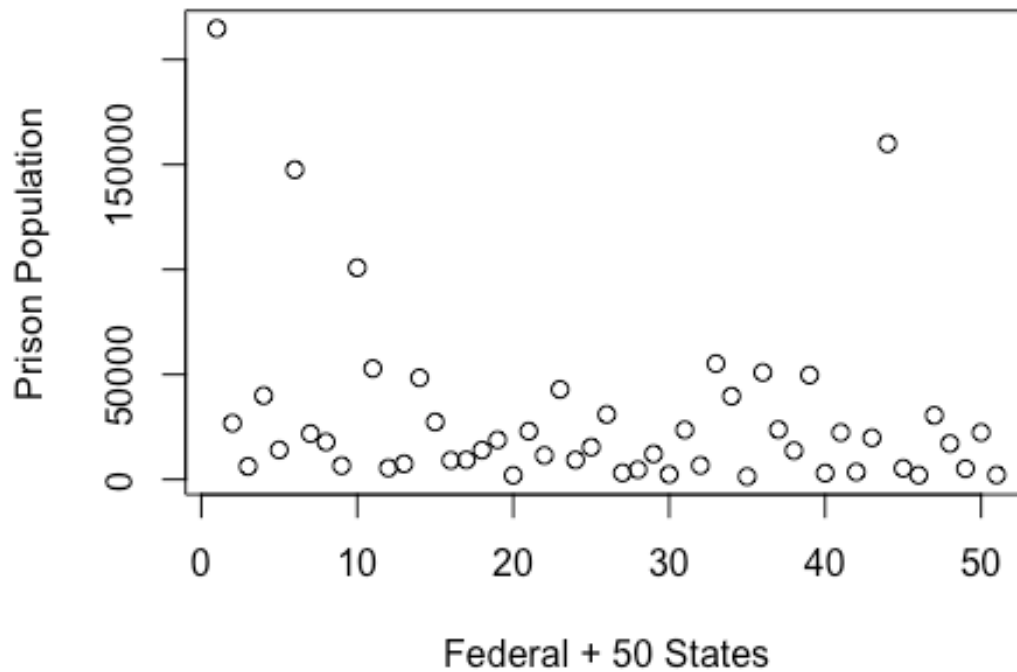
```
plot(allstates$prison_2010, main = "2010 Prison Population by Federal and Individual State", xlab = "Federal + 50 States", ylab="Prison Population")
```

2010 Prison Population by Federal and Individual St



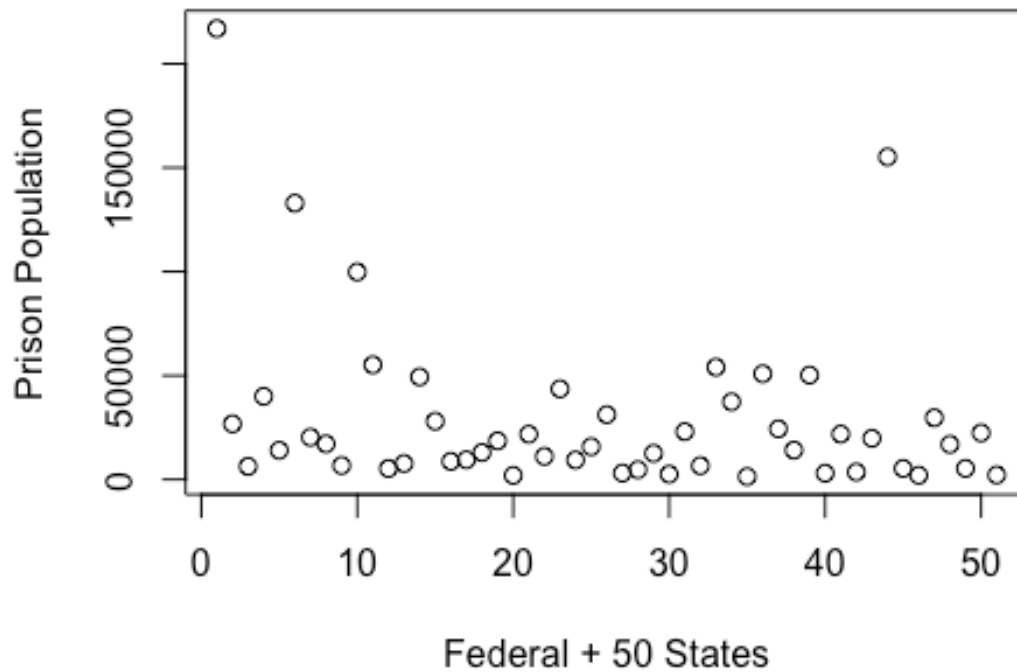
```
plot(allstates$prison_2011, main = "2011 Prison Population by Federal and Individual State", xlab = "Federal + 50 States", ylab="Prison Population")
```

2011 Prison Population by Federal and Individual St



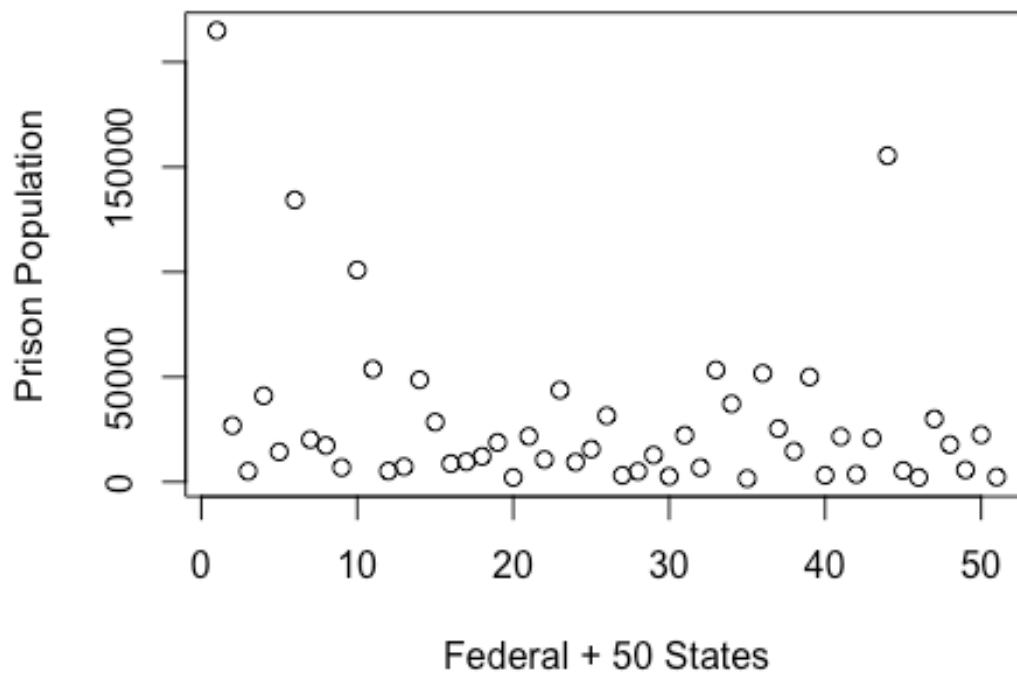
```
plot(allstates$prison_2012, main = "2012 Prison Population by Federal and Individual State", xlab = "Federal + 50 States", ylab="Prison Population")
```


2012 Prison Population by Federal and Individual St



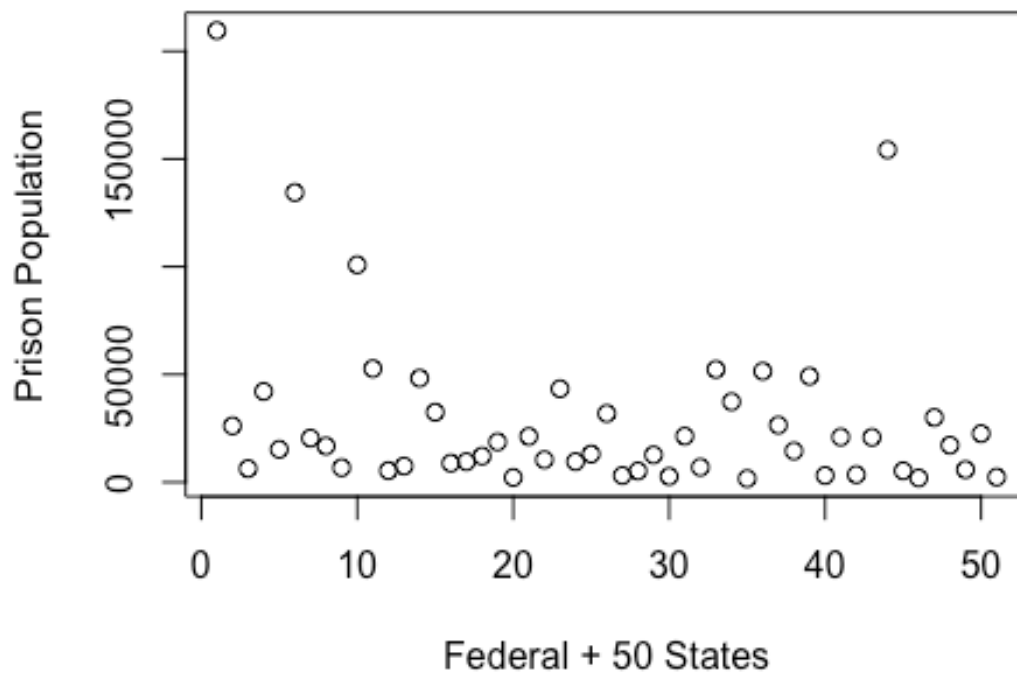
```
plot(allstates$prison_2013, main = "2013 Prison Population by Federal and Individual State", xlab = "Federal + 50 States", ylab="Prison Population")
```

2013 Prison Population by Federal and Individual St



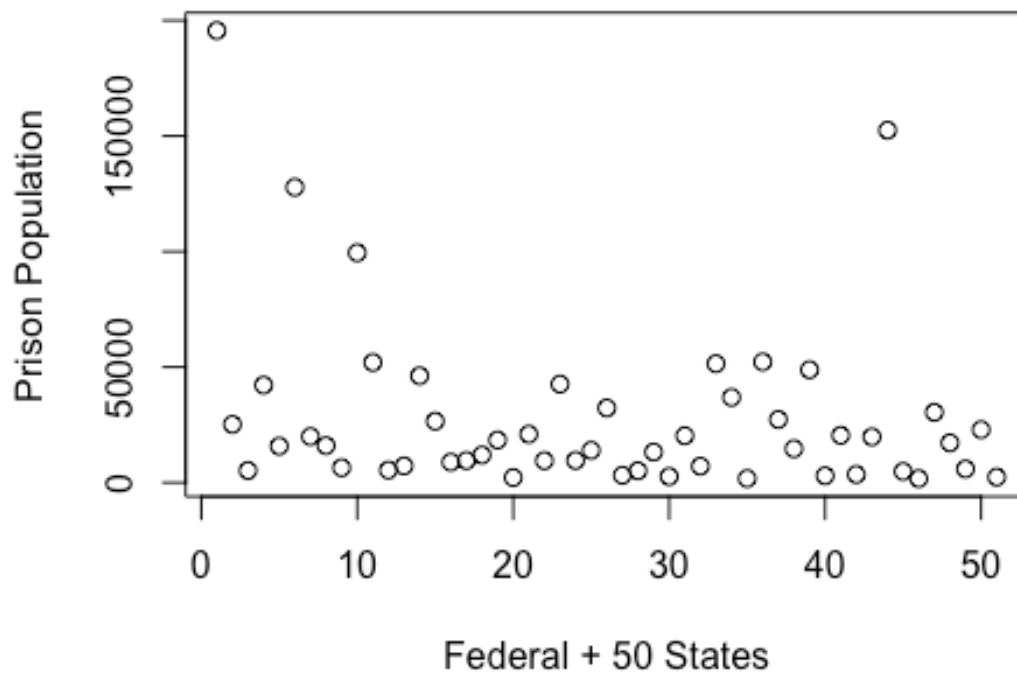
```
plot(allstates$prison_2014, main = "2014 Prison Population by Federal and Individual State", xlab = "Federal + 50 States", ylab="Prison Population")
```

2014 Prison Population by Federal and Individual St



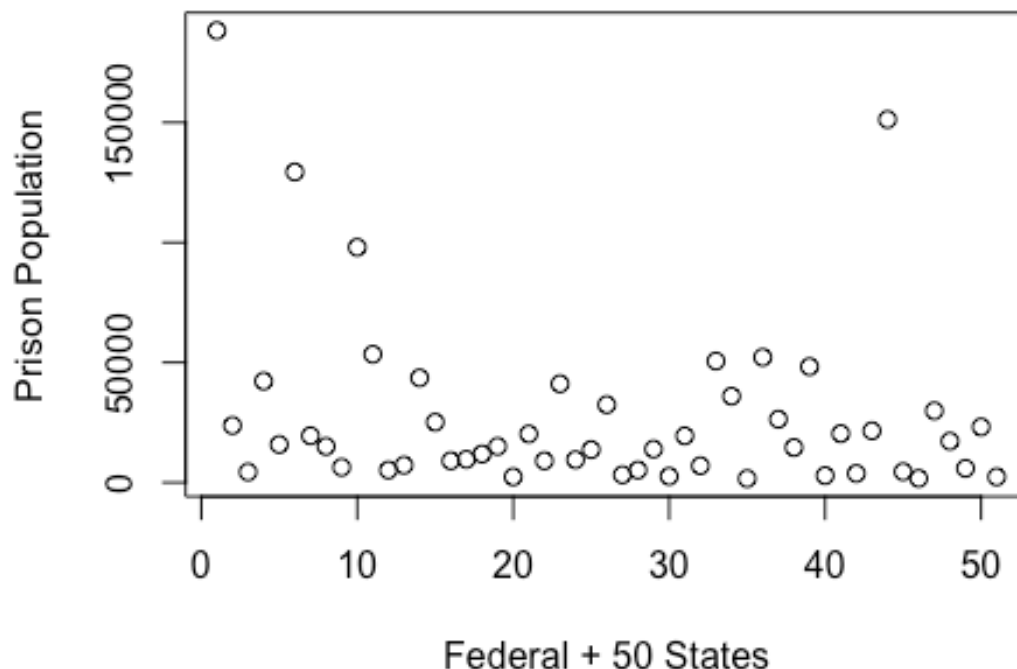
```
plot(allstates$prison_2015, main = "2015 Prison Population by Federal and Individual State", xlab = "Federal + 50 States", ylab="Prison Population")
```

2015 Prison Population by Federal and Individual St



```
plot(allstates$prison_2016, main = "2016 Prison Population by Federal and Individual State", xlab = "Federal + 50 States", ylab="Prison Population")
```

2016 Prison Population by Federal and Individual St



3. What demographic factors are most closely correlated with being incarcerated?

Demographic Subset

```
library(dplyr)
dem_subset <- subset(maindata, select = c(jurisdiction, prison_2016, White_2016_fed_sent, Black_2016_fed_sent, Hisp_2016_fed_sent, Native_2016_fed_sent, Asian_Pac_2016_fed_sent))
demographic <- dem_subset[complete.cases(dem_subset),]
```

```
glimpse(demographic)
```

```
## Observations: 1
## Variables: 7
## $ jurisdiction      <fct> US_Federal
## $ prison_2016       <int> 188311
## $ White_2016_fed_sent <int> 46951
## $ Black_2016_fed_sent <int> 59405
## $ Hisp_2016_fed_sent <int> 57432
## $ Native_2016_fed_sent <int> 3417
## $ Asian_Pac_2016_fed_sent <int> 2293
```

```
dem_prison_pop_all <- demographic$prison_2016
white <- demographic$White_2016_fed_sent
black <- demographic$Black_2016_fed_sent
hispanic <- demographic$Hispanic_2016_fed_sent
native <- demographic$Native_2016_fed_sent
asian <- demographic$Asian_Pac_2016_fed_sent
```

Percentage White: 24.9%

```
white / dem_prison_pop_all
## [1] 0.2493269
```

Percentage Black: 31.5%

```
black / dem_prison_pop_all
## [1] 0.3154622
```

Percentage Hispanic: 30.4%

```
hispanic / dem_prison_pop_all
## [1] 0.3049848
```

Percentage Native American: 1.8%

```
native / dem_prison_pop_all
## [1] 0.01814551
```

Percentage Asian & Pacific Islander: 1.2%

```
asian / dem_prison_pop_all
## [1] 0.01217667
```

4. How many Americans cannot vote due to their status as a convicted felon?

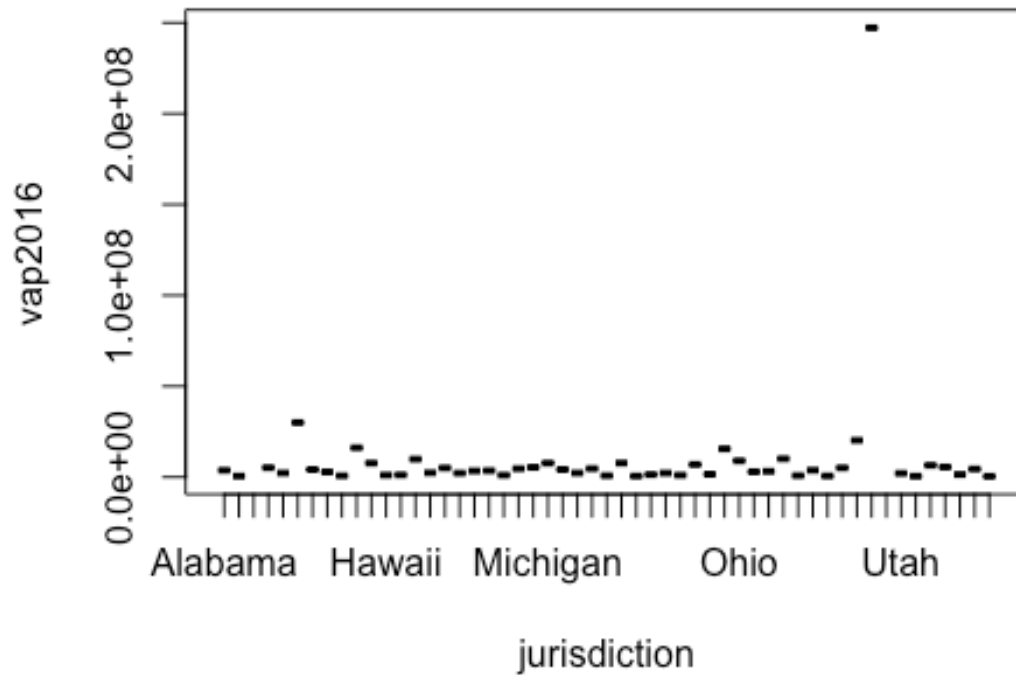
Population Disenfranchised from Voting Due to Imprisonment

Voting Subset

```
##      jurisdiction  vap2016 percentage_disenfranchised_2016
## 1      US_ALL 247219588                                0.02
```

## 4	Alabama	3755483	0.08
## 5	Alaska	552166	0.03
## 6	Arizona	5205215	0.04
## 7	Arkansas	2272904	0.03
## 8	California	30023902	0.01
## 9	Colorado	4199509	0.01
## 10	Connecticut	2826827	0.01
## 11	Delaware	741548	0.02
## 12	Florida	16166143	0.10
## 13	Georgia	7710688	0.03
## 14	Hawaii	1120770	0.01
## 15	Idaho	1222093	0.02
## 16	Illinois	9901322	0.01
## 17	Indiana	2395103	0.02
## 18	Iowa	5040224	0.01
## 19	Kansas	2192084	0.01
## 20	Kentucky	3413425	0.09
## 21	Louisiana	3555911	0.03
## 22	Maine	1072948	0.00
## 23	Maryland	4658175	0.00
## 24	Massachusetts	5407335	0.00
## 25	Michigan	7715272	0.01
## 26	Minnesota	4205207	0.02
## 27	Mississippi	2265485	0.10
## 28	Missouri	4692196	0.02
## 29	Montana	806529	0.01
## 30	Nebraska	7752234	0.01
## 31	Nevada	583001	0.00
## 32	New_Hampshire	1425853	0.01
## 33	New_Jersey	2221681	0.04
## 34	New_Mexico	1066610	0.00
## 35	New_York	6959192	0.01
## 36	North_Carolina	1588201	0.02
## 37	North_Dakota	15584974	0.01
## 38	Ohio	8984946	0.01
## 39	Oklahoma	2950017	0.02
## 40	Oregon	3166121	0.00
## 41	Pennsylvania	10112229	0.01
## 42	Rhode_Island	845254	0.00
## 43	South_Carolina	3804558	0.01
## 44	South_Dakota	647145	0.02
## 45	Tennessee	5102688	0.08
## 46	Texas	20257343	0.02
## 47	Utah	2083423	0.00
## 48	Vermont	506119	0.00
## 49	Virginia	6512571	0.08
## 50	Washington	5558509	0.01
## 51	West_Virginia	1464532	0.01
## 52	Wisconsin	4476711	0.01
## 53	Wyoming	447212	0.05

```
plot(vap2016 ~ jurisdiction, data = disenfranchised)
```



```
dis_calc <- disenfranchised$percentage_disenfranchised_2016 * disenfranchised
$vap2016
```

```
dis_calc
```

```
## [1] 4944391.76 300438.64 16564.98 208208.60 68187.12 300239.02
## [7] 41995.09 28268.27 14830.96 1616614.30 231320.64 11207.70
## [13] 24441.86 99013.22 47902.06 50402.24 21920.84 307208.25
## [19] 106677.33 0.00 0.00 0.00 77152.72 84104.14
## [25] 226548.50 93843.92 8065.29 77522.34 0.00 14258.53
## [31] 88867.24 0.00 69591.92 31764.02 155849.74 89849.46
## [37] 59000.34 0.00 101122.29 0.00 38045.58 12942.90
## [43] 408215.04 405146.86 0.00 0.00 521005.68 55585.09
## [49] 14645.32 44767.11 22360.60
```

```
disenfranchised$tot_disenfranchised <- dis_calc
disenfranchised
```

```
## jurisdiction vap2016 percentage_disenfranchised_2016
## 1 US_ALL 247219588 0.02
## 4 Alabama 3755483 0.08
## 5 Alaska 552166 0.03
```


## 6	Arizona	5205215	0.04
## 7	Arkansas	2272904	0.03
## 8	California	30023902	0.01
## 9	Colorado	4199509	0.01
## 10	Connecticut	2826827	0.01
## 11	Delaware	741548	0.02
## 12	Florida	16166143	0.10
## 13	Georgia	7710688	0.03
## 14	Hawaii	1120770	0.01
## 15	Idaho	1222093	0.02
## 16	Illinois	9901322	0.01
## 17	Indiana	2395103	0.02
## 18	Iowa	5040224	0.01
## 19	Kansas	2192084	0.01
## 20	Kentucky	3413425	0.09
## 21	Louisiana	3555911	0.03
## 22	Maine	1072948	0.00
## 23	Maryland	4658175	0.00
## 24	Massachusetts	5407335	0.00
## 25	Michigan	7715272	0.01
## 26	Minnesota	4205207	0.02
## 27	Mississippi	2265485	0.10
## 28	Missouri	4692196	0.02
## 29	Montana	806529	0.01
## 30	Nebraska	7752234	0.01
## 31	Nevada	583001	0.00
## 32	New_Hampshire	1425853	0.01
## 33	New_Jersey	2221681	0.04
## 34	New_Mexico	1066610	0.00
## 35	New_York	6959192	0.01
## 36	North_Carolina	1588201	0.02
## 37	North_Dakota	15584974	0.01
## 38	Ohio	8984946	0.01
## 39	Oklahoma	2950017	0.02
## 40	Oregon	3166121	0.00
## 41	Pennsylvania	10112229	0.01
## 42	Rhode_Island	845254	0.00
## 43	South_Carolina	3804558	0.01
## 44	South_Dakota	647145	0.02
## 45	Tennessee	5102688	0.08
## 46	Texas	20257343	0.02
## 47	Utah	2083423	0.00
## 48	Vermont	506119	0.00
## 49	Virginia	6512571	0.08
## 50	Washington	5558509	0.01
## 51	West_Virginia	1464532	0.01
## 52	Wisconsin	4476711	0.01
## 53	Wyoming	447212	0.05
##	tot_disenfranchised		
## 1		4944391.76	

## 4	300438.64
## 5	16564.98
## 6	208208.60
## 7	68187.12
## 8	300239.02
## 9	41995.09
## 10	28268.27
## 11	14830.96
## 12	1616614.30
## 13	231320.64
## 14	11207.70
## 15	24441.86
## 16	99013.22
## 17	47902.06
## 18	50402.24
## 19	21920.84
## 20	307208.25
## 21	106677.33
## 22	0.00
## 23	0.00
## 24	0.00
## 25	77152.72
## 26	84104.14
## 27	226548.50
## 28	93843.92
## 29	8065.29
## 30	77522.34
## 31	0.00
## 32	14258.53
## 33	88867.24
## 34	0.00
## 35	69591.92
## 36	31764.02
## 37	155849.74
## 38	89849.46
## 39	59000.34
## 40	0.00
## 41	101122.29
## 42	0.00
## 43	38045.58
## 44	12942.90
## 45	408215.04
## 46	405146.86
## 47	0.00
## 48	0.00
## 49	521005.68
## 50	55585.09
## 51	14645.32
## 52	44767.11
## 53	22360.60

The first value is the entire population disenfranchised from voting based on imprisonment throughout the country. That number is 4,944,391.

Entire disenfranchised population by adding states with values, then subtracting the US value.

```
dis_calc2 <- (sum(dis_calc)) - 4944391.76
dis_calc2

## [1] 6195696
```

Missing values are problematic here, but it gives a higher overall number: 6,195,696

Averaging the two Disenfranchisement Totals

```
(sum(dis_calc))/2

## [1] 5570044
```

5,570,044

5. What kinds of crime are most closely associated with incarceration?

This was problematic in the end, as I cannot use the different crime categories as predictors when my data is solely about those who have been imprisoned. As with the demographic categories, I looked at percentages.

```
library(dplyr)
crime_subset <- subset(maindata, select = c(jurisdiction, prison_or_jail_2016,
, violent_crime_total_2016, murder_manslaughter_2016, rape_revised_2016,
robbery_2016, agg_assault_2016, property_crime_total_2016, burglary_2016,
, larceny_2016, vehicle_theft_2016))
crimes <- crime_subset[complete.cases(crime_subset),]
```

```
glimpse(crimes)
```

```
## Observations: 50
## Variables: 11
## $ jurisdiction      <fct> Alabama, Alaska, Arizona, Arkansas, Ca...
## $ prison_or_jail_2016 <int> 40900, 4400, 55000, 24000, 202700, 321...
## $ violent_crime_total_2016 <int> 25878, 5966, 32542, 16563, 174796, 190...
## $ murder_manslaughter_2016 <int> 407, 52, 389, 217, 1930, 189, 79, 58, ...
## $ rape_revised_2016 <int> 1915, 1053, 3304, 2214, 13702, 3635, 7...
## $ robbery_2016 <int> 4687, 850, 7045, 2125, 54789, 3525, 27...
## $ agg_assault_2016 <int> 18869, 4011, 21804, 12007, 104375, 116...
## $ property_crime_total_2016 <int> 143259, 24876, 207317, 98092, 1002070,...
## $ burglary_2016 <int> 34045, 4053, 38216, 23814, 188304, 238...
## $ larceny_2016 <int> 97498, 17766, 150618, 67091, 637010, 1...
## $ vehicle_theft_2016 <int> 11716, 3057, 18483, 7187, 176756, 1964...
```

```
summary(crimes)
```

```
##      jurisdiction prison_or_jail_2016 violent_crime_total_2016
## Alabama      : 1      Min.      : 1700      Min.      : 851
## Alaska       : 1      1st Qu.: 10400      1st Qu.: 6133
## Arizona      : 1      Median : 28550      Median : 17786
## Arkansas     : 1      Mean     : 38812      Mean     : 25547
## California   : 1      3rd Qu.: 47650      3rd Qu.: 31045
## Colorado     : 1      Max.     :218500      Max.     :174796
## (Other)      :44
## murder_manslaughter_2016 rape_revised_2016 robbery_2016
## Min.      : 14.00      Min.      : 170.0      Min.      : 59
## 1st Qu.: 61.75      1st Qu.: 856.5      1st Qu.: 1026
## Median : 222.50      Median : 1823.0      Median : 3628
## Mean     : 345.54      Mean     : 2637.6      Mean     : 6586
## 3rd Qu.: 525.75      3rd Qu.: 2842.8      3rd Qu.: 7249
## Max.     :1930.00      Max.     :13702.0      Max.     :54789
##
## agg_assault_2016 property_crime_total_2016 burglary_2016
## Min.      : 589      Min.      : 9705      Min.      : 1771
## 1st Qu.: 4069      1st Qu.: 42564      1st Qu.: 7126
## Median : 10930      Median : 114817      Median : 20546
## Mean     : 15978      Mean     : 157923      Mean     : 30281
## 3rd Qu.: 18784      3rd Qu.: 185352      3rd Qu.: 35056
## Max.     :104375      Max.     :1002070      Max.     :188304
##
## larceny_2016      vehicle_theft_2016
## Min.      : 7637      Min.      : 184
## 1st Qu.: 31156      1st Qu.: 4412
## Median : 80805      Median : 10144
## Mean     :112355      Mean     : 15287
## 3rd Qu.:129508      3rd Qu.: 15039
## Max.     :637010      Max.     :176756
##
```

```
jurisdiction <- crimes$jurisdiction
incarcerated <- crimes$prison_or_jail_2016
violent <- crimes$violent_crime_total_2016 / incarcerated
murder <- crimes$murder_manslaughter_2016 / incarcerated
rape <- crimes$rape_revised_2016 / incarcerated
robbery <- crimes$robbery_2016 / incarcerated
assault <- crimes$agg_assault_2016 / incarcerated
property <- crimes$property_crime_total_2016 / incarcerated
burglary <- crimes$burglary_2016 / incarcerated
larceny <- crimes$larceny_2016 / incarcerated
vehicletheft <- crimes$vehicle_theft_2016 / incarcerated
```

Violent Crime Percentages by State

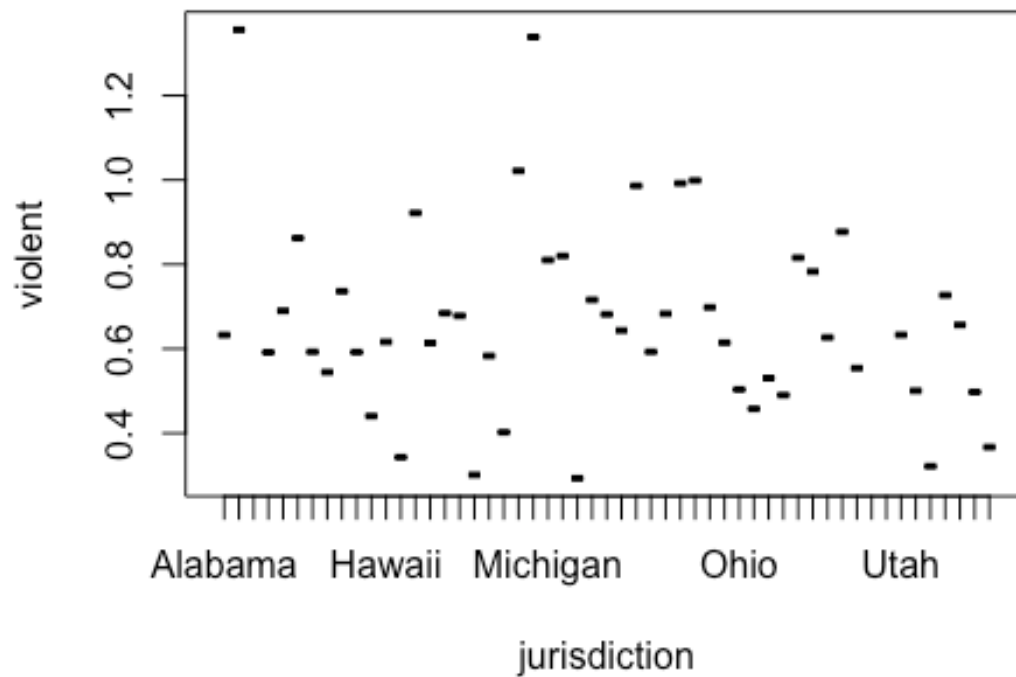
```
violent_subset <- subset(crimes, select = c(jurisdiction))  
violent_subset$violence_rate <- violent
```

```
violent_sorted <- violent_subset[order(-violent_subset$violence_rate),]  
violent_sorted
```

##	jurisdiction	violence_rate
## 5	Alaska	1.3559091
## 24	Massachusetts	1.3389175
## 23	Maryland	1.0217958
## 35	New_York	0.9988575
## 34	New_Mexico	0.9921769
## 31	Nevada	0.9863366
## 16	Illinois	0.9219408
## 45	Tennessee	0.8772521
## 8	California	0.8623384
## 26	Minnesota	0.8199387
## 42	Rhode_Island	0.8158065
## 25	Michigan	0.8103009
## 43	South_Carolina	0.7830841
## 11	Delaware	0.7362121
## 50	Washington	0.7270066
## 28	Missouri	0.7160271
## 36	North_Carolina	0.6980961
## 7	Arkansas	0.6901250
## 18	Iowa	0.6843284
## 33	New_Jersey	0.6831563
## 29	Montana	0.6817544
## 19	Kansas	0.6781977
## 51	West_Virginia	0.6567327
## 30	Nebraska	0.6432955
## 47	Utah	0.6329915
## 4	Alabama	0.6327139
## 44	South_Dakota	0.6268966
## 14	Hawaii	0.6164286
## 37	North_Dakota	0.6145161
## 17	Indiana	0.6137963
## 32	New_Hampshire	0.5928889
## 9	Colorado	0.5928349
## 12	Florida	0.5921228
## 6	Arizona	0.5916727
## 21	Louisiana	0.5831938
## 46	Texas	0.5540686
## 10	Connecticut	0.5446000
## 40	Oregon	0.5305797
## 38	Ohio	0.5036479
## 48	Vermont	0.5005882
## 52	Wisconsin	0.4976404

```
## 41    Pennsylvania    0.4901578
## 39      Oklahoma    0.4578205
## 13       Georgia    0.4405689
## 22       Maine     0.4021951
## 53      Wyoming     0.3669231
## 15       Idaho     0.3430088
## 49      Virginia    0.3216522
## 20      Kentucky    0.3012104
## 27      Mississippi 0.2930662
```

```
plot(violent ~ jurisdiction, data = crimes)
```



```
murder_subset <- subset(crimes, select = c(jurisdiction))
murder_subset$murder_rate <- murder
murder_sorted <- murder_subset[order(-murder_subset$murder_rate),]
murder_sorted
```

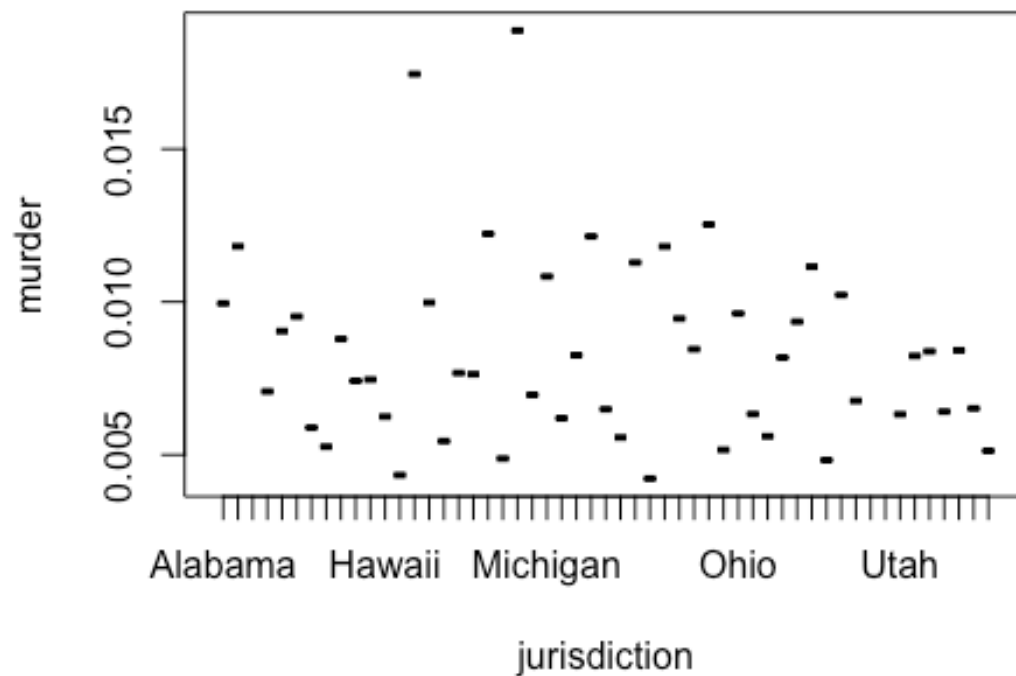
```
##      jurisdiction murder_rate
## 23      Maryland 0.018873239
## 16      Illinois 0.017450658
## 36 North_Carolina 0.012532348
## 21      Louisiana 0.012224670
## 28      Missouri 0.012144470
## 5       Alaska 0.011818182
```

```

## 33      New_Jersey 0.011812500
## 31      Nevada 0.011287129
## 43 South_Carolina 0.011152648
## 25      Michigan 0.010831858
## 45      Tennessee 0.010227273
## 17      Indiana 0.009976852
## 4       Alabama 0.009951100
## 38      Ohio 0.009619718
## 8       California 0.009521460
## 34      New_Mexico 0.009455782
## 42      Rhode_Island 0.009354839
## 7       Arkansas 0.009041667
## 11      Delaware 0.008787879
## 35      New_York 0.008454301
## 51 West_Virginia 0.008415842
## 49      Virginia 0.008382609
## 27      Mississippi 0.008257840
## 48      Vermont 0.008235294
## 41      Pennsylvania 0.008179612
## 19      Kansas 0.007674419
## 20      Kentucky 0.007636888
## 13      Georgia 0.007461707
## 12      Florida 0.007416555
## 6       Arizona 0.007072727
## 24      Massachusetts 0.006958763
## 46      Texas 0.006764302
## 52      Wisconsin 0.006516854
## 29      Montana 0.006491228
## 50      Washington 0.006414474
## 39      Oklahoma 0.006333333
## 47      Utah 0.006324786
## 14      Hawaii 0.006250000
## 26      Minnesota 0.006196319
## 9       Colorado 0.005887850
## 40      Oregon 0.005603865
## 30      Nebraska 0.005568182
## 18      Iowa 0.005447761
## 10      Connecticut 0.005266667
## 37      North_Dakota 0.005161290
## 53      Wyoming 0.005128205
## 22      Maine 0.004878049
## 44      South_Dakota 0.004827586
## 15      Idaho 0.004336283
## 32      New_Hampshire 0.004222222

```

```
plot(murder ~ jurisdiction, data = crimes)
```



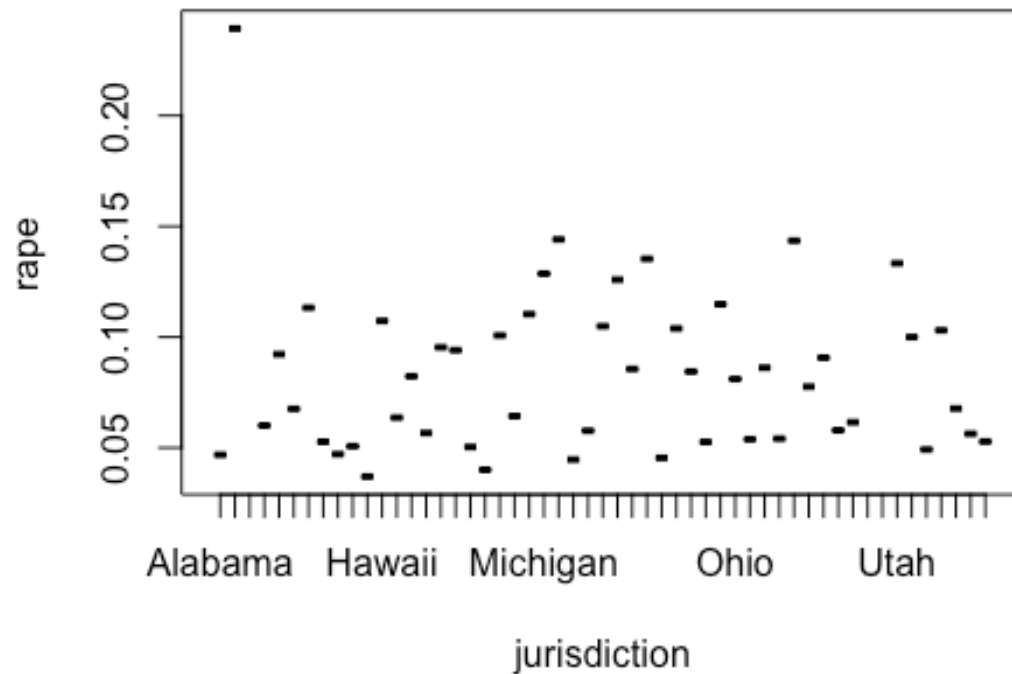
```
rape_subset <- subset(crimes, select = c(jurisdiction))
rape_subset$rape_rate <- rape
rape_sorted <- rape_subset[order(-rape_subset$rape_rate),]
rape_sorted
```

```
##      jurisdiction  rape_rate
## 5      Alaska 0.23931818
## 26    Minnesota 0.14411043
## 42  Rhode_Island 0.14354839
## 32 New_Hampshire 0.13533333
## 47      Utah 0.13333333
## 25    Michigan 0.12863717
## 30    Nebraska 0.12590909
## 37 North_Dakota 0.11483871
## 9      Colorado 0.11323988
## 24 Massachusetts 0.11036082
## 14      Hawaii 0.10732143
## 29    Montana 0.10491228
## 34    New_Mexico 0.10387755
## 50    Washington 0.10305921
## 22      Maine 0.10073171
## 48    Vermont 0.10000000
## 18      Iowa 0.09537313
```



```
## 19      Kansas 0.09406977
## 7       Arkansas 0.09225000
## 44    South_Dakota 0.09068966
## 40      Oregon 0.08618357
## 31      Nevada 0.08559406
## 35     New_York 0.08442204
## 16     Illinois 0.08228618
## 38      Ohio 0.08104225
## 43    South_Carolina 0.07760125
## 51    West_Virginia 0.06772277
## 8      California 0.06759743
## 23     Maryland 0.06433099
## 15      Idaho 0.06362832
## 46      Texas 0.06153776
## 6      Arizona 0.06007273
## 45     Tennessee 0.05791322
## 28     Missouri 0.05772009
## 17     Indiana 0.05668981
## 52     Wisconsin 0.05632022
## 41     Pennsylvania 0.05411408
## 39     Oklahoma 0.05384615
## 53     Wyoming 0.05282051
## 10     Connecticut 0.05273333
## 36    North_Carolina 0.05262477
## 12      Florida 0.05072096
## 20     Kentucky 0.05037464
## 49     Virginia 0.04921739
## 11     Delaware 0.04712121
## 4      Alabama 0.04682152
## 33     New_Jersey 0.04543750
## 27     Mississippi 0.04466899
## 21     Louisiana 0.04006608
## 13      Georgia 0.03699125
```

```
plot(rape ~ jurisdiction, data = crimes)
```

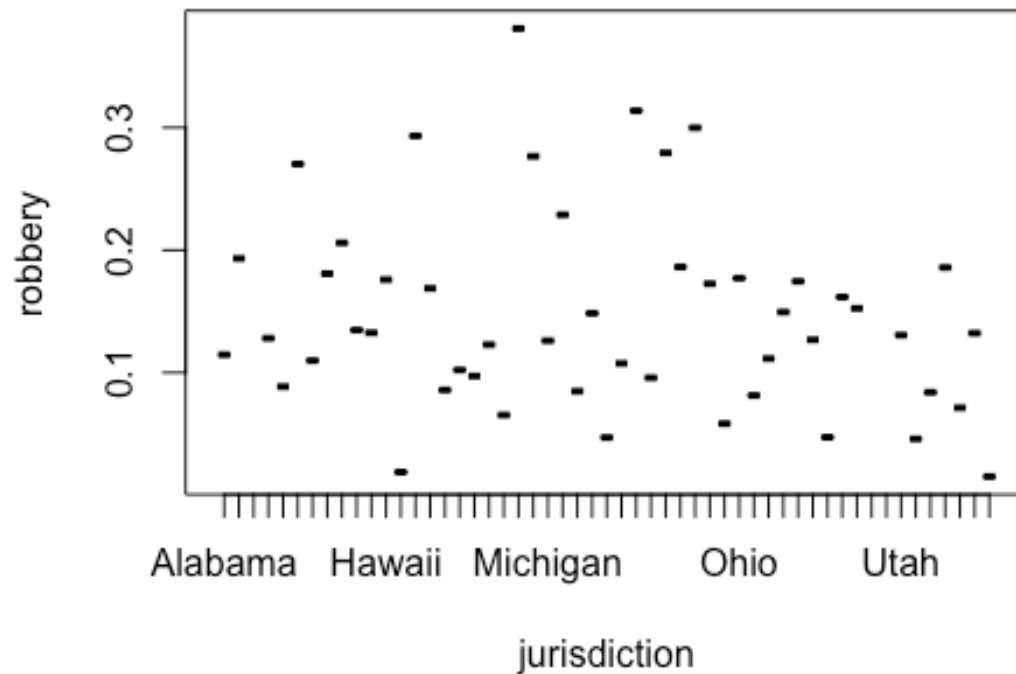


```
robbery_subset <- subset(crimes, select = c(jurisdiction))
robbery_subset$robbery_rate <- robbery
robbery_sorted <- robbery_subset[order(-robbery_subset$robbery_rate),]
robbery_sorted
```

```
##      jurisdiction robbery_rate
## 23      Maryland    0.38080986
## 31       Nevada    0.31376238
## 35    New_York    0.29986559
## 16     Illinois    0.29324013
## 33   New_Jersey    0.27937500
## 24 Massachusetts    0.27654639
## 8      California    0.27029600
## 26     Minnesota    0.22883436
## 11     Delaware    0.20590909
## 5       Alaska    0.19318182
## 34   New_Mexico    0.18619048
## 50    Washington    0.18582237
## 10   Connecticut    0.18073333
## 38        Ohio    0.17709859
## 14        Hawaii    0.17589286
## 42   Rhode_Island    0.17451613
## 36 North_Carolina    0.17256932
```

## 17	Indiana	0.16877315
## 45	Tennessee	0.16175620
## 46	Texas	0.15239359
## 41	Pennsylvania	0.14944175
## 28	Missouri	0.14844244
## 12	Florida	0.13467957
## 13	Georgia	0.13253829
## 52	Wisconsin	0.13221910
## 47	Utah	0.13059829
## 6	Arizona	0.12809091
## 43	South_Carolina	0.12682243
## 25	Michigan	0.12607080
## 21	Louisiana	0.12279736
## 4	Alabama	0.11459658
## 40	Oregon	0.11144928
## 9	Colorado	0.10981308
## 30	Nebraska	0.10750000
## 19	Kansas	0.10215116
## 20	Kentucky	0.09714697
## 32	New_Hampshire	0.09577778
## 7	Arkansas	0.08854167
## 18	Iowa	0.08582090
## 27	Mississippi	0.08477352
## 49	Virginia	0.08393043
## 39	Oklahoma	0.08138462
## 51	West_Virginia	0.07128713
## 22	Maine	0.06536585
## 37	North_Dakota	0.05838710
## 44	South_Dakota	0.04724138
## 29	Montana	0.04701754
## 48	Vermont	0.04588235
## 15	Idaho	0.01884956
## 53	Wyoming	0.01512821

```
plot(robbery ~ jurisdiction, data = crimes)
```

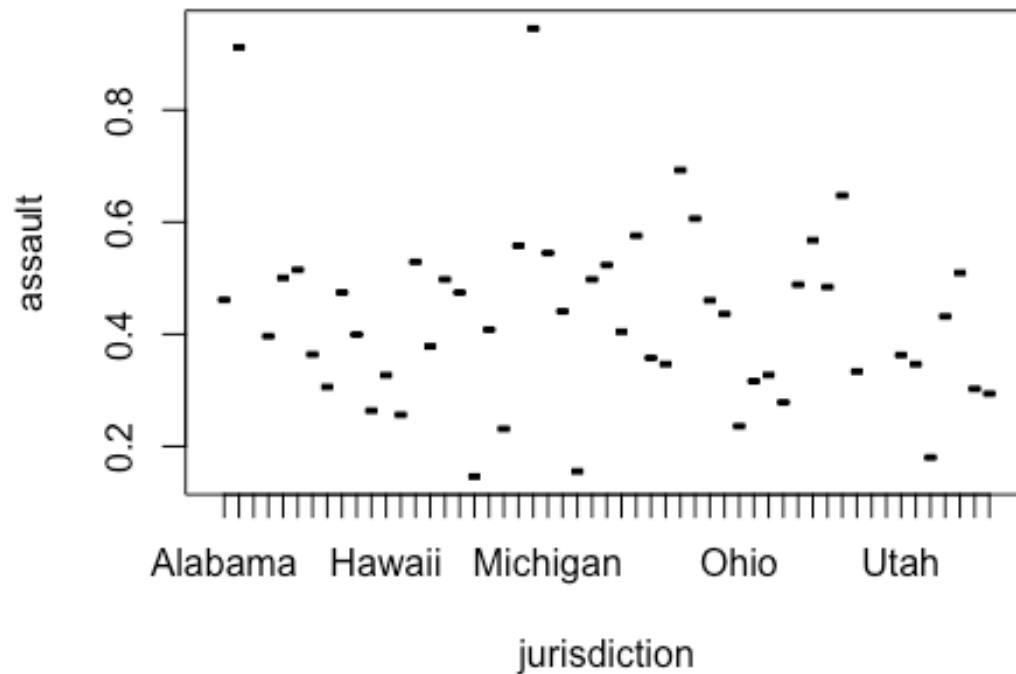


```
assault_subset <- subset(crimes, select = c(jurisdiction))
assault_subset$assault_rate <- assault
assault_sorted <- assault_subset[order(-assault_subset$assault_rate),]
assault_sorted
```

```
##      jurisdiction assault_rate
## 24  Massachusetts    0.9450515
## 5    Alaska          0.9115909
## 34   New_Mexico      0.6926531
## 45   Tennessee      0.6473554
## 35   New_York        0.6061156
## 31   Nevada          0.5756931
## 43  South_Carolina   0.5675078
## 23   Maryland       0.5577817
## 25   Michigan       0.5447611
## 16   Illinois       0.5289638
## 29   Montana        0.5233333
## 8    California     0.5149235
## 51  West_Virginia    0.5093069
## 7    Arkansas       0.5002917
## 28   Missouri       0.4977201
## 18   Iowa           0.4976866
## 42   Rhode_Island   0.4883871
```

## 44	South_Dakota	0.4841379
## 11	Delaware	0.4743939
## 19	Kansas	0.4743023
## 4	Alabama	0.4613447
## 36	North_Carolina	0.4603697
## 26	Minnesota	0.4407975
## 37	North_Dakota	0.4361290
## 50	Washington	0.4317105
## 21	Louisiana	0.4081057
## 30	Nebraska	0.4043182
## 12	Florida	0.3993057
## 6	Arizona	0.3964364
## 17	Indiana	0.3783565
## 9	Colorado	0.3638941
## 47	Utah	0.3627350
## 32	New_Hampshire	0.3575556
## 33	New_Jersey	0.3465313
## 48	Vermont	0.3464706
## 46	Texas	0.3333730
## 40	Oregon	0.3273430
## 14	Hawaii	0.3269643
## 39	Oklahoma	0.3162564
## 10	Connecticut	0.3058667
## 52	Wisconsin	0.3025843
## 53	Wyoming	0.2938462
## 41	Pennsylvania	0.2784223
## 13	Georgia	0.2635777
## 15	Idaho	0.2561947
## 38	Ohio	0.2358873
## 22	Maine	0.2312195
## 49	Virginia	0.1801217
## 27	Mississippi	0.1553659
## 20	Kentucky	0.1460519

```
plot(assault ~ jurisdiction, data = crimes)
```

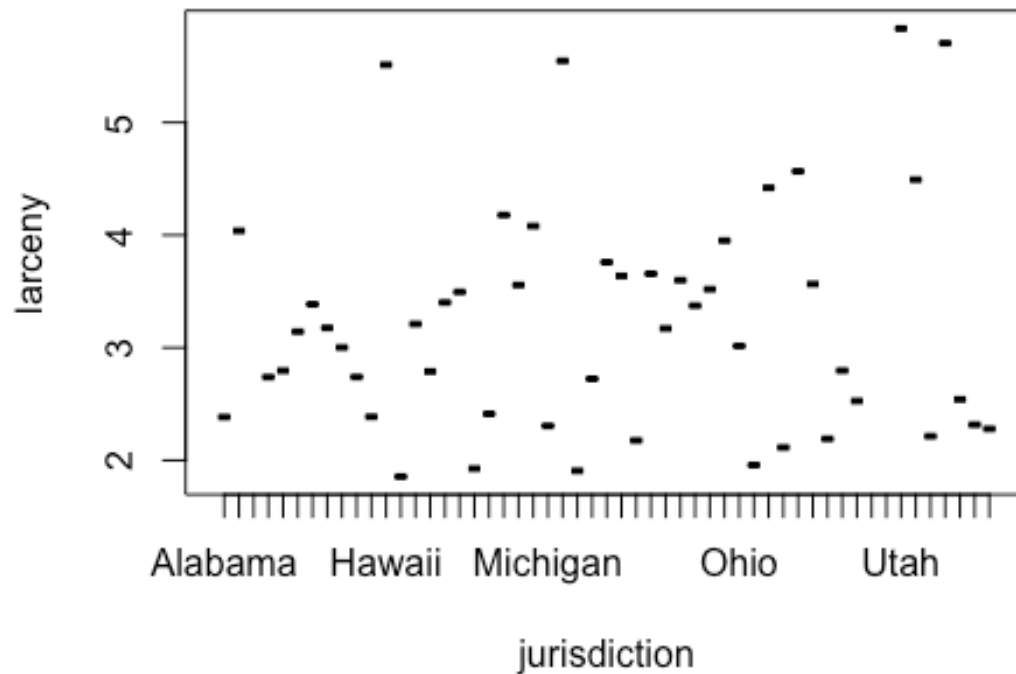


```
larceny_subset <- subset(crimes, select = c(jurisdiction))
larceny_subset$larceny_rate <- larceny
larceny_sorted <- larceny_subset[order(-larceny_subset$larceny_rate),]
larceny_sorted
```

```
##      jurisdiction larceny_rate
## 47      Utah      5.834274
## 50 Washington      5.704704
## 26  Minnesota      5.547301
## 14    Hawaii      5.512679
## 42 Rhode_Island      4.566774
## 48    Vermont      4.492353
## 40    Oregon      4.420628
## 22    Maine      4.178293
## 24 Massachusetts      4.080155
## 5      Alaska      4.037727
## 37 North_Dakota      3.951290
## 29    Montana      3.759123
## 32 New_Hampshire      3.656444
## 30    Nebraska      3.637386
## 34    New_Mexico      3.599116
## 43 South_Carolina      3.565576
## 23    Maryland      3.556162
```

```
## 36 North_Carolina      3.518983
## 19      Kansas        3.495291
## 18      Iowa          3.401493
## 9       Colorado       3.385670
## 35      New_York       3.373522
## 16      Illinois      3.210461
## 10      Connecticut    3.176133
## 33      New_Jersey     3.170594
## 8       California     3.142625
## 38      Ohio           3.015521
## 11      Delaware       3.001818
## 45      Tennessee      2.796550
## 7       Arkansas       2.795458
## 17      Indiana        2.788287
## 12      Florida        2.739332
## 6       Arizona        2.738509
## 28      Missouri       2.723454
## 51 West_Virginia      2.540297
## 46      Texas          2.526801
## 21      Louisiana      2.411718
## 13      Georgia        2.386510
## 4       Alabama        2.383814
## 52      Wisconsin     2.316152
## 25      Michigan       2.305292
## 53      Wyoming        2.279231
## 49      Virginia       2.213652
## 44      South_Dakota   2.190517
## 31      Nevada         2.175693
## 41      Pennsylvania   2.113434
## 39      Oklahoma       1.957897
## 20      Kentucky       1.926023
## 27      Mississippi    1.906794
## 15      Idaho          1.855044
```

```
plot(larceny ~ jurisdiction, data = crimes)
```

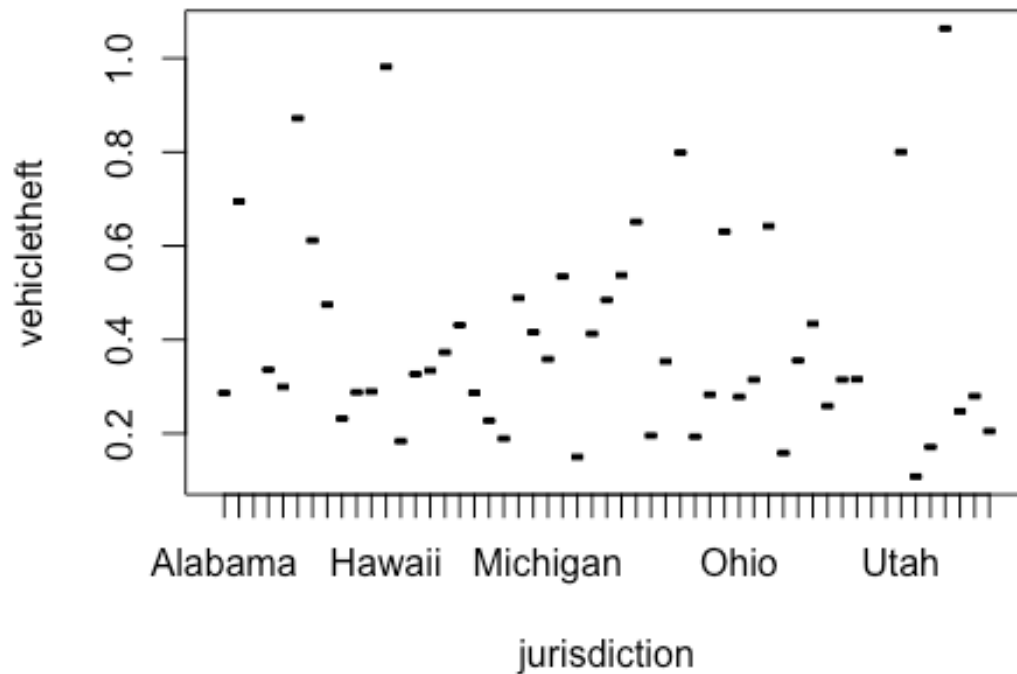


```
vehicletheft_subset <- subset(crimes, select = c(jurisdiction))
vehicletheft_subset$vehicletheft_rate <- vehicletheft
vehicletheft_sorted <- vehicletheft_subset[order(-vehicletheft_subset$vehicle
theft_rate),]
vehicletheft_sorted
```

##	jurisdiction	vehicletheft_rate
## 50	Washington	1.0632237
## 14	Hawaii	0.9819643
## 8	California	0.8720079
## 47	Utah	0.8001709
## 34	New_Mexico	0.7988435
## 5	Alaska	0.6947727
## 31	Nevada	0.6513366
## 40	Oregon	0.6420290
## 37	North_Dakota	0.6303226
## 9	Colorado	0.6118692
## 30	Nebraska	0.5373864
## 26	Minnesota	0.5349693
## 23	Maryland	0.4890493
## 29	Montana	0.4847368
## 10	Connecticut	0.4750667
## 43	South_Carolina	0.4339564

## 19	Kansas	0.4311047
## 24	Massachusetts	0.4160309
## 28	Missouri	0.4130023
## 18	Iowa	0.3732836
## 25	Michigan	0.3586195
## 42	Rhode_Island	0.3558065
## 33	New_Jersey	0.3538750
## 6	Arizona	0.3360545
## 17	Indiana	0.3342361
## 16	Illinois	0.3267599
## 46	Texas	0.3157437
## 45	Tennessee	0.3148554
## 39	Oklahoma	0.3147692
## 7	Arkansas	0.2994583
## 13	Georgia	0.2895295
## 12	Florida	0.2879506
## 20	Kentucky	0.2866859
## 4	Alabama	0.2864548
## 36	North_Carolina	0.2829205
## 52	Wisconsin	0.2797191
## 38	Ohio	0.2783380
## 44	South_Dakota	0.2586207
## 51	West_Virginia	0.2473267
## 11	Delaware	0.2318182
## 21	Louisiana	0.2275330
## 53	Wyoming	0.2051282
## 32	New_Hampshire	0.1955556
## 35	New_York	0.1933333
## 22	Maine	0.1892683
## 15	Idaho	0.1838053
## 49	Virginia	0.1712696
## 41	Pennsylvania	0.1586044
## 27	Mississippi	0.1500697
## 48	Vermont	0.1082353

```
plot(vehicletheft ~ jurisdiction, data = crimes)
```

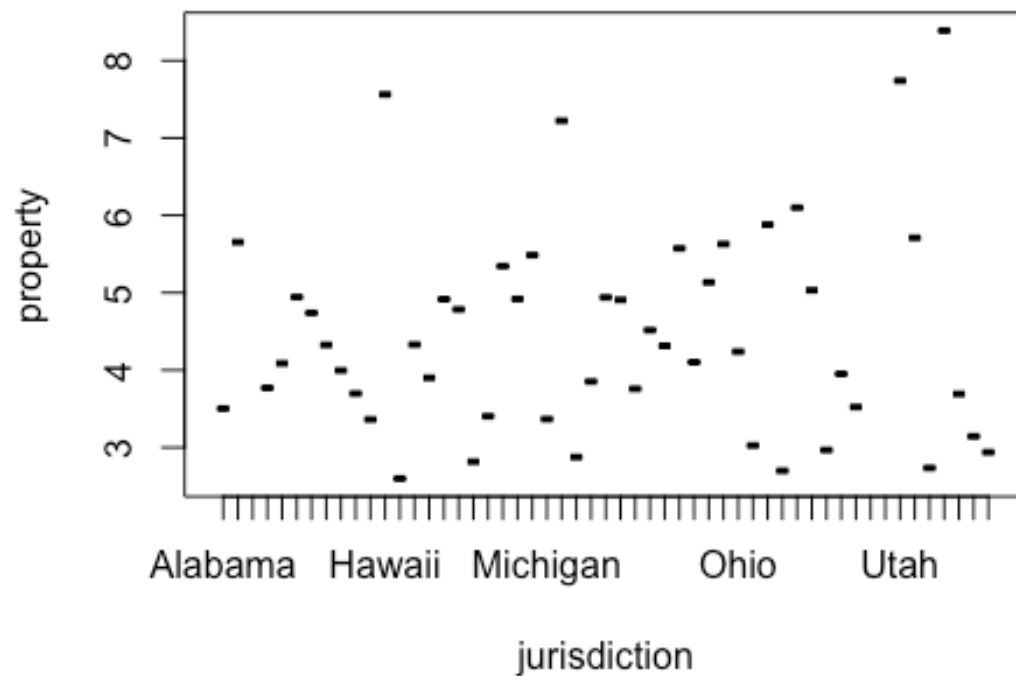


```
property_subset <- subset(crimes, select = c(jurisdiction))
property_subset$property_rate <- property
property_sorted <- property_subset[order(-property_subset$property_rate),]
property_sorted
```

```
##      jurisdiction property_rate
## 50      Washington      8.387961
## 47         Utah      7.740256
## 14         Hawaii      7.563036
## 26      Minnesota      7.222270
## 42   Rhode_Island      6.100000
## 40         Oregon      5.880000
## 48         Vermont      5.708824
## 5         Alaska      5.653636
## 37   North_Dakota      5.629355
## 34    New_Mexico      5.573469
## 24  Massachusetts      5.486082
## 22           Maine      5.343415
## 36 North_Carolina      5.134288
## 43 South_Carolina      5.032212
## 8         California      4.943611
## 29         Montana      4.939649
## 23         Maryland      4.919577
```

## 18	Iowa	4.917015
## 30	Nebraska	4.908750
## 19	Kansas	4.784477
## 9	Colorado	4.739751
## 32	New_Hampshire	4.516222
## 16	Illinois	4.329868
## 10	Connecticut	4.325000
## 33	New_Jersey	4.312875
## 38	Ohio	4.238662
## 35	New_York	4.102473
## 7	Arkansas	4.087167
## 11	Delaware	3.995455
## 45	Tennessee	3.950393
## 17	Indiana	3.899537
## 28	Missouri	3.852393
## 6	Arizona	3.769400
## 31	Nevada	3.758515
## 12	Florida	3.697009
## 51	West_Virginia	3.691287
## 46	Texas	3.523533
## 4	Alabama	3.502665
## 21	Louisiana	3.403326
## 25	Michigan	3.367239
## 13	Georgia	3.362199
## 52	Wisconsin	3.143567
## 39	Oklahoma	3.025897
## 44	South_Dakota	2.966724
## 53	Wyoming	2.938462
## 27	Mississippi	2.875296
## 20	Kentucky	2.815937
## 49	Virginia	2.735513
## 41	Pennsylvania	2.698956
## 15	Idaho	2.597965

```
plot(property ~ jurisdiction, data = crimes)
```

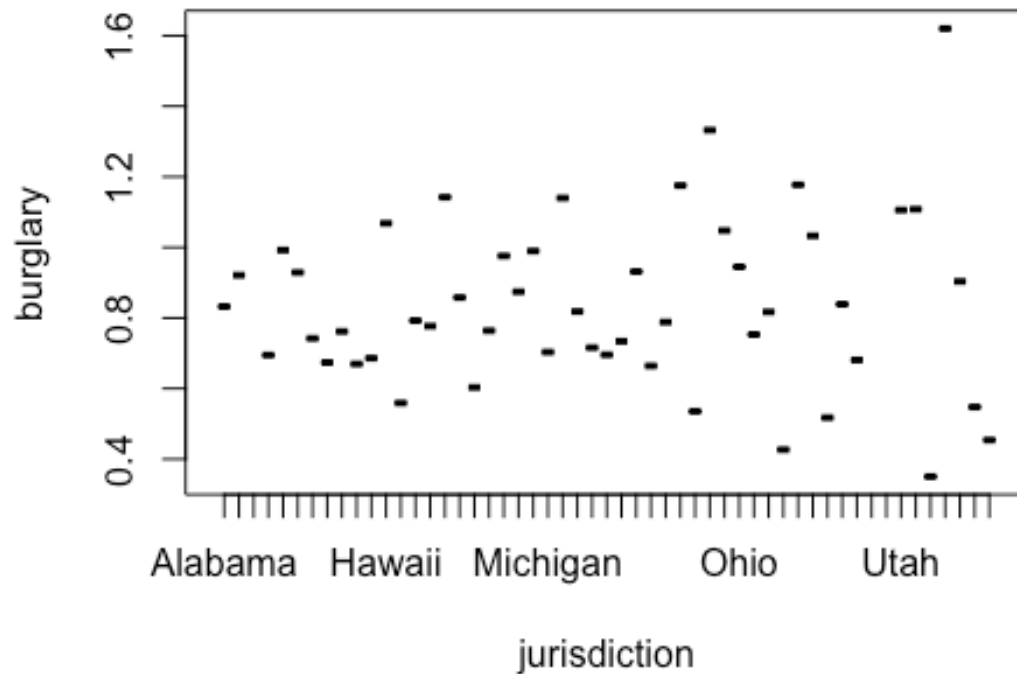


```
burglary_subset <- subset(crimes, select = c(jurisdiction))
burglary_subset$burglary_rate <- burglary
burglary_sorted <- burglary_subset[order(-burglary_subset$burglary_rate),]
burglary_sorted
```

```
##      jurisdiction burglary_rate
## 50      Washington      1.6200329
## 36 North_Carolina      1.3323845
## 42   Rhode_Island      1.1774194
## 34    New_Mexico      1.1755102
## 18         Iowa      1.1422388
## 26    Minnesota      1.1400000
## 48     Vermont      1.1082353
## 47         Utah      1.1058120
## 14         Hawaii      1.0683929
## 37  North_Dakota      1.0477419
## 43 South_Carolina      1.0326791
## 7      Arkansas      0.9922500
## 24 Massachusetts      0.9898969
## 22         Maine      0.9758537
## 38         Ohio      0.9448028
## 31         Nevada      0.9314851
## 8      California      0.9289788
```

## 5	Alaska	0.9211364
## 51	West_Virginia	0.9036634
## 23	Maryland	0.8743662
## 19	Kansas	0.8580814
## 45	Tennessee	0.8389876
## 4	Alabama	0.8323961
## 27	Mississippi	0.8184321
## 40	Oregon	0.8173430
## 16	Illinois	0.7926480
## 33	New_Jersey	0.7884062
## 17	Indiana	0.7770139
## 21	Louisiana	0.7640749
## 11	Delaware	0.7618182
## 39	Oklahoma	0.7532308
## 9	Colorado	0.7422118
## 30	Nebraska	0.7339773
## 28	Missouri	0.7159368
## 25	Michigan	0.7033274
## 29	Montana	0.6957895
## 6	Arizona	0.6948364
## 13	Georgia	0.6861597
## 46	Texas	0.6809886
## 10	Connecticut	0.6738000
## 12	Florida	0.6697263
## 32	New_Hampshire	0.6642222
## 20	Kentucky	0.6032277
## 15	Idaho	0.5591150
## 52	Wisconsin	0.5476966
## 35	New_York	0.5356183
## 44	South_Dakota	0.5175862
## 53	Wyoming	0.4541026
## 41	Pennsylvania	0.4269175
## 49	Virginia	0.3505913

```
plot(burglary ~ jurisdiction, data = crimes)
```



6. Which states have the highest incarceration rates?

```
state_incarc_subset <- subset(maindata, select = c(jurisdiction, Incarc_rate_
per_100k_US_residents_18_or_older))
states <- state_incarc_subset[complete.cases(state_incarc_subset),]
```

states

##	jurisdiction	Incarc_rate_per_100k_US_residents_18_or_older
## 1	US_ALL	850
## 2	US_Federal	80
## 3	All_States	780
## 4	Alabama	1080
## 5	Alaska	800
## 6	Arizona	1030
## 7	Arkansas	1050
## 8	California	670
## 9	Colorado	740
## 10	Connecticut	530
## 11	Delaware	880
## 12	Florida	900
## 13	Georgia	1160
## 14	Hawaii	500
## 15	Idaho	900

## 16	Illinois	620
## 17	Indiana	850
## 18	Iowa	560
## 19	Kansas	780
## 20	Kentucky	1010
## 21	Louisiana	1270
## 22	Maine	380
## 23	Maryland	610
## 24	Massachusetts	360
## 25	Michigan	730
## 26	Minnesota	380
## 27	Mississippi	1260
## 28	Missouri	940
## 29	Montana	700
## 30	Nebraska	610
## 31	Nevada	890
## 32	New_Hampshire	410
## 33	New_Jersey	460
## 34	New_Mexico	930
## 35	New_York	480
## 36	North_Carolina	680
## 37	North_Dakota	540
## 38	Ohio	790
## 39	Oklahoma	1310
## 40	Oregon	640
## 41	Pennsylvania	810
## 42	Rhode_Island	370
## 43	South_Carolina	820
## 44	South_Dakota	880
## 45	Tennessee	930
## 46	Texas	1050
## 47	Utah	540
## 48	Vermont	340
## 49	Virginia	880
## 50	Washington	530
## 51	West_Virginia	690
## 52	Wisconsin	790
## 53	Wyoming	870

Highest Incarceration Rates by State

Top 5: Oklahoma, Louisiana, Mississippi, Georgia, Alabama Bottom 5: Maine, Minnesota, Rhode Island, Vermont, Massachusetts

```
states_sorted <- states[order(-states$Incarc_rate_per_100k_US_residents_18_or_older),]
states_sorted
```

##	jurisdiction	Incarc_rate_per_100k_US_residents_18_or_older
## 39	Oklahoma	1310
## 21	Louisiana	1270
## 27	Mississippi	1260
## 13	Georgia	1160
## 4	Alabama	1080
## 7	Arkansas	1050
## 46	Texas	1050
## 6	Arizona	1030
## 20	Kentucky	1010
## 28	Missouri	940
## 34	New_Mexico	930
## 45	Tennessee	930
## 12	Florida	900
## 15	Idaho	900
## 31	Nevada	890
## 11	Delaware	880
## 44	South_Dakota	880
## 49	Virginia	880
## 53	Wyoming	870
## 1	US_ALL	850
## 17	Indiana	850
## 43	South_Carolina	820
## 41	Pennsylvania	810
## 5	Alaska	800
## 38	Ohio	790
## 52	Wisconsin	790
## 3	All_States	780
## 19	Kansas	780
## 9	Colorado	740
## 25	Michigan	730
## 29	Montana	700
## 51	West_Virginia	690
## 36	North_Carolina	680
## 8	California	670
## 40	Oregon	640
## 16	Illinois	620
## 23	Maryland	610
## 30	Nebraska	610
## 18	Iowa	560
## 37	North_Dakota	540
## 47	Utah	540
## 10	Connecticut	530
## 50	Washington	530
## 14	Hawaii	500
## 35	New_York	480
## 33	New_Jersey	460
## 32	New_Hampshire	410
## 22	Maine	380
## 26	Minnesota	380

##	42	Rhode_Island	370
##	24	Massachusetts	360
##	48	Vermont	340
##	2	US_Federal	80