I'm so sorry that here I use the new data to try dplyr. Because I want to keep the result, I also put this code on github. And the try of bridges data is in the downloadBridges.R

Here is the data description of this dataset:http://archive.ics.uci.edu/ml/datasets/Communities+and+Crime+Unnormalized# (http://archive.ics.uci.edu/ml/datasets/Communities+and+Crime+Unnormalized#)

Data Import

```
library(plyr)
```

```
## Warning: package 'plyr' was built under R version 3.3.2
```

```
library(choroplethr)
```

```
## Warning: package 'choroplethr' was built under R version 3.3.2
```

```
## Loading required package: acs
```

```
## Warning: package 'acs' was built under R version 3.3.2
```

```
## Loading required package: stringr
```

```
## Loading required package: XML
```

```
##
## Attaching package: 'acs'
```

```
## The following object is masked from 'package:base':
##
##     apply
```

```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 3.3.2
```

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

```
## The following object is masked from 'package:acs':
##
##     combine
```

```
## The following objects are masked from 'package:plyr':
##
##     arrange, count, desc, failwith, id, mutate, rename, summarise,
##     summarize
```

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

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

```
library(readr)
```

```
## Warning: package 'readr' was built under R version 3.3.2
```

```
library(data.table)
```

```
## Warning: package 'data.table' was built under R version 3.3.2
```

```
## ------------------------------------------------------------------------
```

```
## data.table + dplyr code now lives in dtplyr.
## Please library(dtplyr)!
```

```
## ------------------------------------------------------------------------
```

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

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

```
dest="http://archive.ics.uci.edu/ml/machine-learning-databases/00211/CommViolPredUnnormalizedData.txt"
tmp=fread(dest,na.strings = '?')
```

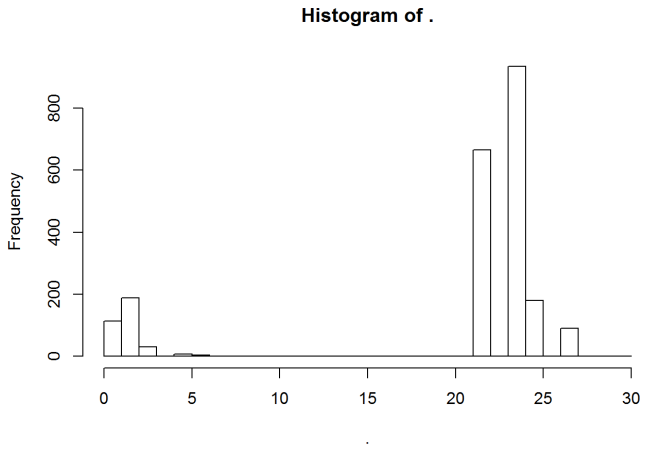
```
tmp=as.tbl(tmp)
CC=tmp
page="http://archive.ics.uci.edu/ml/datasets/Communities+and+Crime+Unnormalized#"
lines=read_lines(page)
needed.lines=lines[566:712]
(variables.names=sub(pattern="<br>@attribute\\s(.*?)\\s(.*)",replacement="\\1",x=needed.lines))
```

```
##   [1] "communityname"       "State"               "countyCode"
##   [4] "communityCode"       "fold"                "pop"
##   [7] "perHoush"            "pctBlack"            "pctWhite"
##  [10] "pctAsian"            "pctHisp"             "pct12-21"
##  [13] "pct12-29"            "pct16-24"            "pct65up"
##  [16] "persUrban"           "pctUrban"            "medIncome"
##  [19] "pctWwage"            "pctWfarm"            "pctWdiv"
##  [22] "pctWsocsec"          "pctPubAsst"          "pctRetire"
##  [25] "medFamIncome"        "perCapInc"           "whitePerCap"
##  [28] "blackPerCap"         "NAperCap"            "asianPerCap"
##  [31] "otherPerCap"         "hispPerCap"          "persPoverty"
##  [34] "pctPoverty"          "pctLowEdu"           "pctNotHSgrad"
##  [37] "pctCollGrad"         "pctUnemploy"         "pctEmploy"
##  [40] "pctEmployMfg"        "pctEmployProfServ"   "pctOccupManu"
##  [43] "pctOccupMgmt"        "pctMaleDivorc"       "pctMaleNevMar"
##  [46] "pctFemDivorc"        "pctAllDivorc"        "persPerFam"
##  [49] "pct2Par"             "pctKids2Par"         "pctKids-4w2Par"
##  [52] "pct12-17w2Par"       "pctWorkMom-6"        "pctWorkMom-18"
##  [55] "kidsBornNevrMarr"    "pctKidsBornNevrMarr" "numForeignBorn"
##  [58] "pctFgnImmig-3"       "pctFgnImmig-5"       "pctFgnImmig-8"
##  [61] "pctFgnImmig-10"      "pctImmig-3"          "pctImmig-5"
##  [64] "pctImmig-8"          "pctImmig-10"         "pctSpeakOnlyEng"
##  [67] "pctNotSpeakEng"      "pctLargHousFam"      "pctLargHous"
##  [70] "persPerOccupHous"    "persPerOwnOccup"     "persPerRenterOccup"
##  [73] "pctPersOwnOccup"     "pctPopDenseHous"     "pctSmallHousUnits"
##  [76] "medNumBedrm"         "houseVacant"         "pctHousOccup"
##  [79] "pctHousOwnerOccup"   "pctVacantBoarded"    "pctVacant6up"
##  [82] "medYrHousBuilt"      "pctHousWOphone"      "pctHousWOplumb"
##  [85] "ownHousLowQ"         "ownHousMed"          "ownHousUperQ"
##  [88] "ownHousQrange"       "rentLowQ"            "rentMed"
##  [91] "rentUpperQ"          "rentQrange"          "medGrossRent"
##  [94] "medRentpctHousInc"   "medOwnCostpct"       "medOwnCostPctWO"
##  [97] "persEmergShelt"      "persHomeless"        "pctForeignBorn"
## [100] "pctBornStateResid"   "pctSameHouse-5"      "pctSameCounty-5"
## [103] "pctSameState-5"      "numPolice"           "policePerPop"
## [106] "policeField"         "policeFieldPerPop"   "policeCalls"
## [109] "policCallPerPop"     "policCallPerOffic"   "policePerPop2"
## [112] "racialMatch"         "pctPolicWhite"       "pctPolicBlack"
## [115] "pctPolicHisp"        "pctPolicAsian"       "pctPolicMinority"
## [118] "officDrugUnits"      "numDiffDrugsSeiz"    "policAveOT"
## [121] "landArea"            "popDensity"          "pctUsePubTrans"
## [124] "policCarsAvail"      "policOperBudget"     "pctPolicPatrol"
## [127] "gangUnit"            "pctOfficDrugUnit"    "policBudgetPerPop"
## [130] "murders"             "murdPerPop"          "rapes"
## [133] "rapesPerPop"         "robberies"           "robbbPerPop"
## [136] "assaults"            "assaultPerPop"       "burglaries"
## [139] "burglPerPop"         "larcenies"           "larcPerPop"
## [142] "autoTheft"           "autoTheftPerPop"     "arsons"
## [145] "arsonsPerPop"        "violentPerPop"       "nonViolPerPop"
```

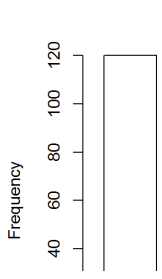
```
variables.names=as.vector(variables.names)
names(CC)=variables.names
```

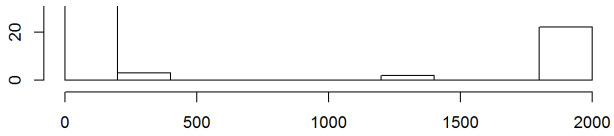
Test NULL

```
is.na(CC) %>% rowSums %>% hist(breaks=30)
```




```
is.na(CC) %>% colSums %>% hist
```





.

Select needed variables

```
needed=c("communityname","State","countyCode","communityCode","pop","perHoush","medIncome","medFamIncome","perCapInc","persPoverty","pctPoverty","pctAllDivorc","murdPerPop","rapesPerPop","robbbPerPop","assaultPerPop","burglPerPop","larcPerPop","autoTheftPerPop","arsonsPerPop","violentPerPop","nonViolPerPop")
needed.data=select(CC,one_of(needed))
```

ggplot of total crime vs some variables I am interested in

```
library(ggplot2)
```

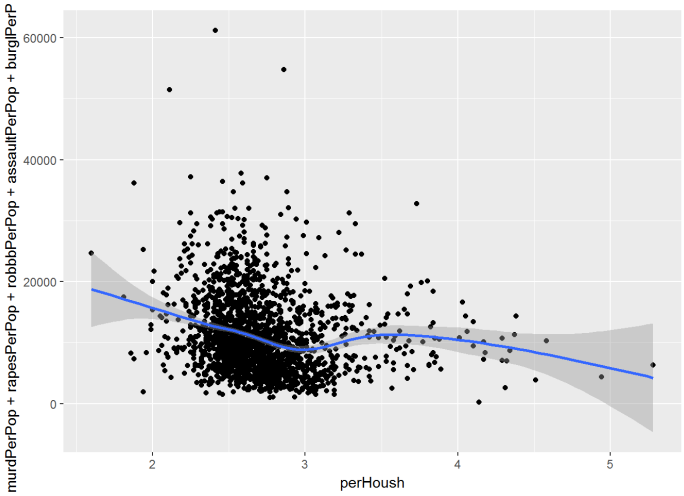
```
## Warning: package 'ggplot2' was built under R version 3.3.2
```

```
ggplot(data=needed.data,mapping = aes(x=perHoush,y=murdPerPop+rapesPerPop+robbbPerPop+assaultPerPop+burglPerPop+larcPerPop+autoTheftPerPop+arsonsPerPop+violentPerPop+nonViolPerPop))+geom_point()+geom_smooth()
```

```
## `geom_smooth()` using method = 'gam'
```

```
## Warning: Removed 313 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 313 rows containing missing values (geom_point).
```

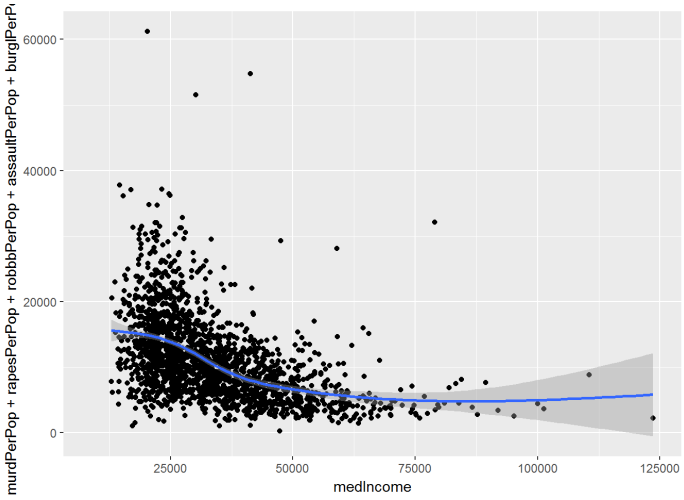



```
ggplot(data=needed.data,mapping = aes(x=medIncome,y=murdPerPop+rapesPerPop+robbbPerPop+assaultPerPop+burglPerPop+larcPerPop+autoTheftPerPop+arsonsPerPop+violentPerPop+nonViolPerPop))+geom_point()+geom_smooth()
```

```
## `geom_smooth()` using method = 'gam'
```

```
## Warning: Removed 313 rows containing non-finite values (stat_smooth).

## Warning: Removed 313 rows containing missing values (geom_point).
```

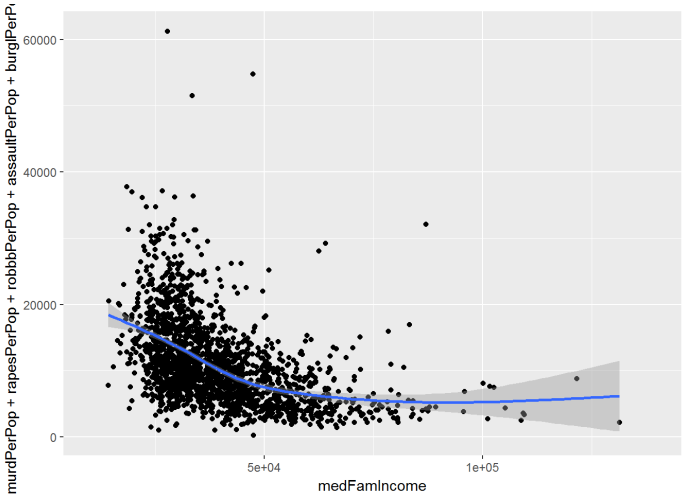



```
ggplot(data=needed.data,mapping = aes(x=medFamIncome,y=murdPerPop+rapesPerPop+robbbPerPop+assaultPerPop+burglPerPop+larcPerPop+autoTheftPerPop+arsonsPerPop+violentPerPop+nonViolPerPop))+geom_point()+geom_smooth()
```

```
## `geom_smooth()` using method = 'gam'
```

```
## Warning: Removed 313 rows containing non-finite values (stat_smooth).

## Warning: Removed 313 rows containing missing values (geom_point).
```

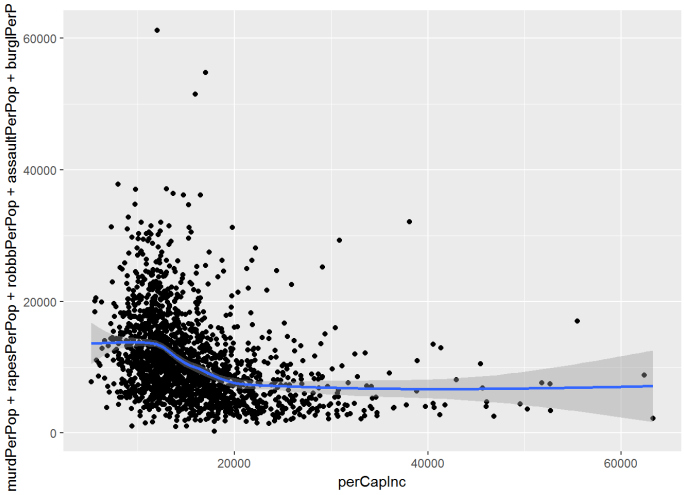



```
ggplot(data=needed.data, mapping = aes(x=perCapInc, y=murdPerPop+rapesPerPop+robbbPerPop+assaultPerPop+burglPerPop+larcPerPop+autoTheftPerPop+arsonsPerPop+violentPerPop+nonViolPerPop))+geom_point()+geom_smooth()
```

```
## `geom_smooth()` using method = 'gam'
```

```
## Warning: Removed 313 rows containing non-finite values (stat_smooth).

## Warning: Removed 313 rows containing missing values (geom_point).
```

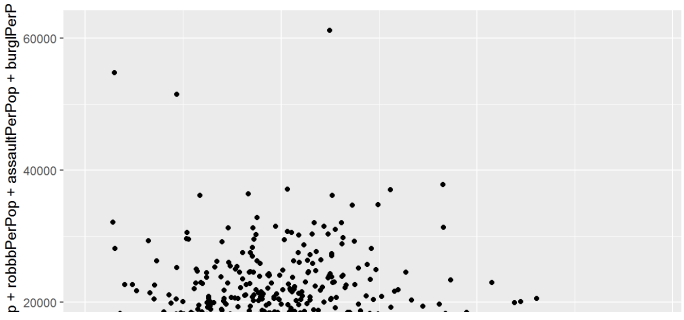


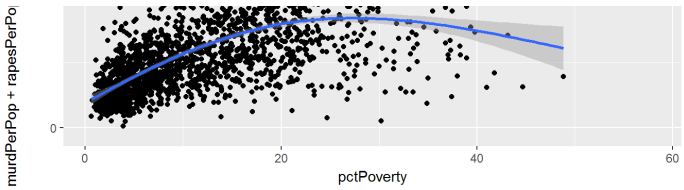

```
ggplot(data=needed.data, mapping = aes(x=pctPoverty, y=murdPerPop+rapesPerPop+robbbPerPop+assaultPerPop+burglPerPop+larcPerPop+autoTheftPerPop+arsonsPerPop+violentPerPop+nonViolPerPop))+geom_point()+geom_smooth()
```

```
## `geom_smooth()` using method = 'gam'
```

```
## Warning: Removed 313 rows containing non-finite values (stat_smooth).

## Warning: Removed 313 rows containing missing values (geom_point).
```



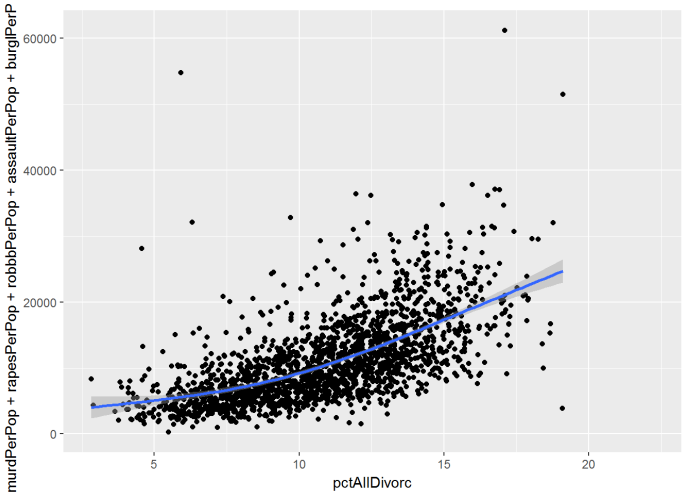



```
ggplot(data=needed.data,mapping = aes(x=pctAllDivorc,y=murdPerPop+rapesPerPop+robbbPerPop+assaultPerPop+burglPerPop+larcPerPop+autoTheft
PerPop+arsonsPerPop+violentPerPop+nonViolPerPop))+geom_point()+geom_smooth()
```

```
## `geom_smooth()` using method = 'gam'
```

```
## Warning: Removed 313 rows containing non-finite values (stat_smooth).

## Warning: Removed 313 rows containing missing values (geom_point).
```



Total crime per 100k population with wisconsin map

```
library(choroplethrMaps)
```

```
## Warning: package 'choroplethrMaps' was built under R version 3.3.2
```

```
wi=filter(needed.data,State=="WI")
wi = mutate(wi, fips = 55000+countyCode)
by_fips=group_by(wi,fips)
sumcrime=summarise(by_fips,sum(murdPerPop+rapesPerPop+robbbPerPop+assaultPerPop+burglPerPop+larcPerPop+autoTheftPerPop+arsonsPerPop+viol
entPerPop+nonViolPerPop))
sumcrime=data.frame(sumcrime)
sumcrime%>% transmute(region = fips, value = sumcrime[,2]) %>% county_choropleth(state_zoom = "wisconsin") #total crime
```

```
## Warning in super$initialize(map.df, user.df): Your data.frame contains the
## following regions which are not mappable: NA
```

```
## Warning in self$bind(): The following regions were missing and are being
## set to NA: 55107, 55111, 55113, 55119, 55123, 55125, 55011, 55013, 55019,
## 55021, 55023, 55015, 55053, 55067, 55078, 55099, 55109, 55121, 55135,
## 55037, 55041, 55043, 55047, 55049, 55057, 55065, 55077, 55083, 55085,
## 55091, 55093, 55061, 55051, 55079, 55001, 55137, 55029, 55069, 55081,
## 55095, 55103, 55115, 55129, 55003, 55005, 55007
```