Blanca Linares
CIS 320 Project

Purpose of this project is to view data for US Arrests made
1.We will see data for four categories

2. We will see data from all 50 states

3. We will see a statistical data of for the four categories

4.We will see a selection of Histograms for every category

5. followed by scatterplots which we will see relationship between the groups

Based on the observations made, we will observe a statistics contained by 50 observations for 100,000 residents for assault, murder and rape for all 50 states in 1973 including a percentage of urban population. Based on the map we can determine that which states show the highest crime rates. The darker shades represent less murders while the lighter states represent higher rates of murder. We can determine that for the amount of crimes such as Murder, Rape and Assaults, per 100,000 people, we can see there is a relationship where there is a majority of people that live in these urban areas that are affected by this. We are able to determine that since the correlation of arrests vs. murder and rape vs. an urban population are close to 0 we can determine there is a constant relationship. The scatter plots represent information that states that if you are located in a more urban area there are higher chances of you coming across a murder, assault or rape.

```
mUSArrests<-read.table("https://vincentarelbundock.github.io/Rdatasets/csv/datasets/USArrests.csv",sep = ",", header = T)
class(USArrests)
```

```
## [1] "data.frame"
```

```
USArrests
```

```
##                Murder Assault UrbanPop Rape
## Alabama          13.2     236       58 21.2
## Alaska           10.0     263       48 44.5
## Arizona           8.1     294       80 31.0
## Arkansas          8.8     190       50 19.5
## California        9.0     276       91 40.6
## Colorado          7.9     204       78 38.7
## Connecticut       3.3     110       77 11.1
## Delaware          5.9     238       72 15.8
## Florida          15.4     335       80 31.9
## Georgia          17.4     211       60 25.8
## Hawaii            5.3      46       83 20.2
## Idaho             2.6     120       54 14.2
## Illinois         10.4     249       83 24.0
## Indiana           7.2     113       65 21.0
## Iowa              2.2      56       57 11.3
## Kansas            6.0     115       66 18.0
## Kentucky          9.7     109       52 16.3
## Louisiana        15.4     249       66 22.2
## Maine             2.1      83       51  7.8
## Maryland         11.3     300       67 27.8
## Massachusetts     4.4     149       85 16.3
## Michigan         12.1     255       74 35.1
## Minnesota         2.7      72       66 14.9
## Mississippi      16.1     259       44 17.1
## Missouri          9.0     178       70 28.2
## Montana           6.0     109       53 16.4
## Nebraska          4.3     102       62 16.5
## Nevada           12.2     252       81 46.0
## New Hampshire     2.1      57       56  9.5
## New Jersey        7.4     159       89 18.8
## New Mexico       11.4     285       70 32.1
## New York         11.1     254       86 26.1
## North Carolina   13.0     337       45 16.1
## North Dakota      0.8      45       44  7.3
## Ohio              7.3     120       75 21.4
## Oklahoma          6.6     151       68 20.0
## Oregon            4.9     159       67 29.3
## Pennsylvania      6.3     106       72 14.9
## Rhode Island      3.4     174       87  8.3
## South Carolina   14.4     279       48 22.5
## South Dakota      3.8      86       45 12.8
## Tennessee        13.2     188       59 26.9
## Texas            12.7     201       80 25.5
## Utah              3.2     120       80 22.9
## Vermont           2.2      48       32 11.2
## Virginia          8.5     156       63 20.7
## Washington        4.0     145       73 26.2
## West Virginia     5.7      81       39  9.3
## Wisconsin         2.6      53       66 10.8
## Wyoming           6.8     161       60 15.6
```

```
summary(USArrests) #summary of statistics for data set
```

```
##      Murder          Assault         UrbanPop          Rape
##  Min.   : 0.800   Min.   : 45.0   Min.   :32.00   Min.   : 7.30
##  1st Qu.: 4.075   1st Qu.:109.0   1st Qu.:54.50   1st Qu.:15.07
##  Median : 7.250   Median :159.0   Median :66.00   Median :20.10
##  Mean   : 7.788   Mean   :170.8   Mean   :65.54   Mean   :21.23
##  3rd Qu.:11.250   3rd Qu.:249.0   3rd Qu.:77.75   3rd Qu.:26.18
##  Max.   :17.400   Max.   :337.0   Max.   :91.00   Max.   :46.00
```

```r
summary(USArrests$Murder) #here we can check for any invalid outliers that are negative
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.800   4.075   7.250   7.788  11.250  17.400
```

```r
summary(USArrests$Assault)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    45.0   109.0   159.0   170.8   249.0   337.0
```

```r
summary(USArrests$UrbanPop)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   32.00   54.50   66.00   65.54   77.75   91.00
```

```r
summary(USArrests$Rape)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    7.30   15.08   20.10   21.23   26.17   46.00
```

```r
dim(USArrests) #50 Observations, 4 variables
```

```
## [1] 50  4
```

```r
USArrests [1,1] #shows the first row and first value
```

```
## [1] 13.2
```

```r
USArrests [1] #Show murders for all states
```

```
##                Murder
## Alabama          13.2
## Alaska           10.0
## Arizona           8.1
## Arkansas          8.8
## California        9.0
## Colorado          7.9
## Connecticut       3.3
## Delaware          5.9
## Florida          15.4
## Georgia          17.4
## Hawaii            5.3
## Idaho             2.6
## Illinois         10.4
## Indiana           7.2
## Iowa              2.2
## Kansas            6.0
## Kentucky          9.7
## Louisiana        15.4
## Maine             2.1
## Maryland         11.3
## Massachusetts     4.4
## Michigan         12.1
## Minnesota         2.7
## Mississippi      16.1
## Missouri          9.0
## Montana           6.0
## Nebraska          4.3
## Nevada           12.2
## New Hampshire     2.1
## New Jersey        7.4
## New Mexico       11.4
## New York         11.1
## North Carolina   13.0
## North Dakota      0.8
## Ohio              7.3
## Oklahoma          6.6
## Oregon            4.9
## Pennsylvania      6.3
## Rhode Island      3.4
## South Carolina   14.4
## South Dakota      3.8
## Tennessee        13.2
## Texas            12.7
## Utah              3.2
## Vermont           2.2
## Virginia          8.5
## Washington        4.0
## West Virginia     5.7
## Wisconsin         2.6
## Wyoming           6.8
```

```r
USArrests[2] #shows Assaults for all states
```

```
##                Assault
## Alabama            236
```

```
## Alaska            263
## Arizona           294
## Arkansas          190
## California        276
## Colorado          204
## Connecticut       110
## Delaware          238
## Florida           335
## Georgia           211
## Hawaii             46
## Idaho             120
## Illinois          249
## Indiana           113
## Iowa               56
## Kansas            115
## Kentucky          109
## Louisiana         249
## Maine              83
## Maryland          300
## Massachusetts     149
## Michigan          255
## Minnesota          72
## Mississippi       259
## Missouri          178
## Montana           109
## Nebraska          102
## Nevada            252
## New Hampshire      57
## New Jersey        159
## New Mexico        285
## New York          254
## North Carolina    337
## North Dakota       45
## Ohio              120
## Oklahoma          151
## Oregon            159
## Pennsylvania      106
## Rhode Island      174
## South Carolina    279
## South Dakota       86
## Tennessee         188
## Texas             201
## Utah              120
## Vermont            48
## Virginia          156
## Washington        145
## West Virginia      81
## Wisconsin          53
## Wyoming           161
```

```
USArrests [3] #shows Urbanpop for all states
```

```
##               UrbanPop
## Alabama             58
## Alaska              48
## Arizona             80
## Arkansas            50
## California          91
## Colorado            78
## Connecticut         77
## Delaware            72
## Florida             80
## Georgia             60
## Hawaii              83
## Idaho               54
## Illinois            83
## Indiana             65
## Iowa                57
## Kansas              66
## Kentucky            52
## Louisiana           66
## Maine               51
## Maryland            67
## Massachusetts       85
## Michigan            74
## Minnesota           66
## Mississippi         44
## Missouri            70
## Montana             53
## Nebraska            62
## Nevada              81
## New Hampshire       56
## New Jersey          89
## New Mexico          70
## New York            86
## North Carolina      45
## North Dakota        44
## Ohio                75
## Oklahoma            68
## Oregon              67
## Pennsylvania        72
## Rhode Island        87
## South Carolina      48
## South Dakota        45
## Tennessee           59
## Texas               80
## Utah                80
## Vermont             32
```

```
## Virginia             63
## Washington           73
## West Virginia        39
## Wisconsin            66
## Wyoming              60
```

```
USArrests [4] #shows Rape for all states
```

```
##                Rape
## Alabama        21.2
## Alaska         44.5
## Arizona        31.0
## Arkansas       19.5
## California     40.6
## Colorado       38.7
## Connecticut    11.1
## Delaware       15.8
## Florida        31.9
## Georgia        25.8
## Hawaii         20.2
## Idaho          14.2
## Illinois       24.0
## Indiana        21.0
## Iowa           11.3
## Kansas         18.0
## Kentucky       16.3
## Louisiana      22.2
## Maine           7.8
## Maryland       27.8
## Massachusetts  16.3
## Michigan       35.1
## Minnesota      14.9
## Mississippi    17.1
## Missouri       28.2
## Montana        16.4
## Nebraska       16.5
## Nevada         46.0
## New Hampshire   9.5
## New Jersey     18.8
## New Mexico     32.1
## New York       26.1
## North Carolina 16.1
## North Dakota    7.3
## Ohio           21.4
## Oklahoma       20.0
## Oregon         29.3
## Pennsylvania   14.9
## Rhode Island    8.3
## South Carolina 22.5
## South Dakota   12.8
## Tennessee      26.9
## Texas          25.5
## Utah           22.9
## Vermont        11.2
## Virginia       20.7
## Washington     26.2
## West Virginia   9.3
## Wisconsin      10.8
## Wyoming        15.6
```

```
library(ggplot2)
```

Below are plots

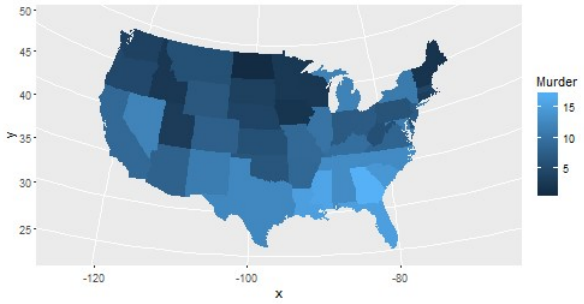
```
library(ggplot2)
arrestData<-data.frame(state=tolower(rownames(USArrests)),USArrests)
head(arrestData)
```

```
##                 state Murder Assault UrbanPop Rape
## Alabama       alabama   13.2     236       58 21.2
## Alaska         alaska   10.0     263       48 44.5
## Arizona       arizona    8.1     294       80 31.0
## Arkansas     arkansas    8.8     190       50 19.5
## California california    9.0     276       91 40.6
## Colorado     colorado    7.9     204       78 38.7
```

```
stMap<-map_data("state")
```

```
##
##  # maps v3.1: updated 'world': all lakes moved to separate new #
##  # 'lakes' database. Type '?world' or 'news(package="maps")'.  #
```

```
ggplot(arrestData, aes(map_id=state, fill=Murder))+geom_map(map=stMap)+expand_limits(x=stMap$long, y=stMap$lat)+coord_map("polyconic")
```



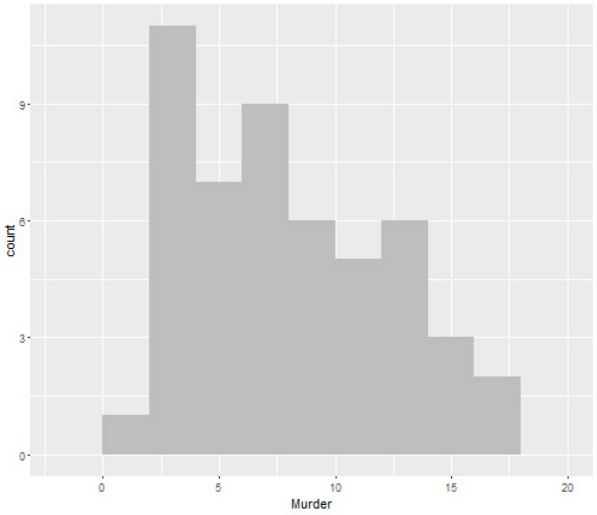

```
USArrests<- subset(USArrests, (USArrests$Murder>0 & USArrests$Murder<100 & USArrests$Assault>0 ))
cor(USArrests$Murder, USArrests$Assault) #This is the correlation between arrests and murder
```

```
## [1] 0.8018733
```

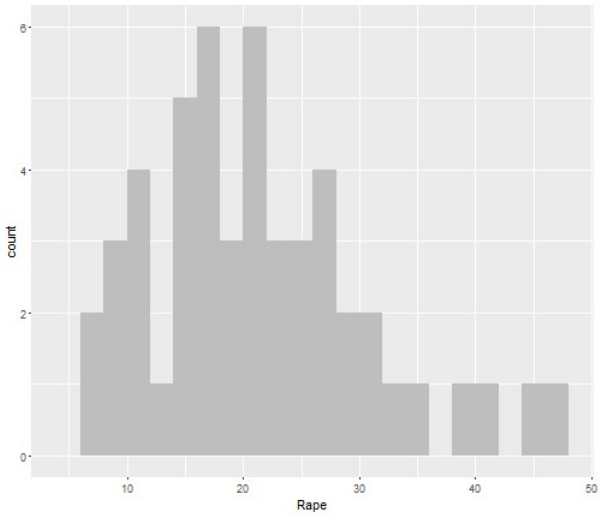
```
USArrests<- subset(USArrests, (USArrests$Rape>0 & USArrests$Rape<100 & USArrests$UrbanPop>0 ))
cor(USArrests$Rape, USArrests$UrbanPop) #this is the correlation between rape and urbanpopulation
```

```
## [1] 0.4113412
```

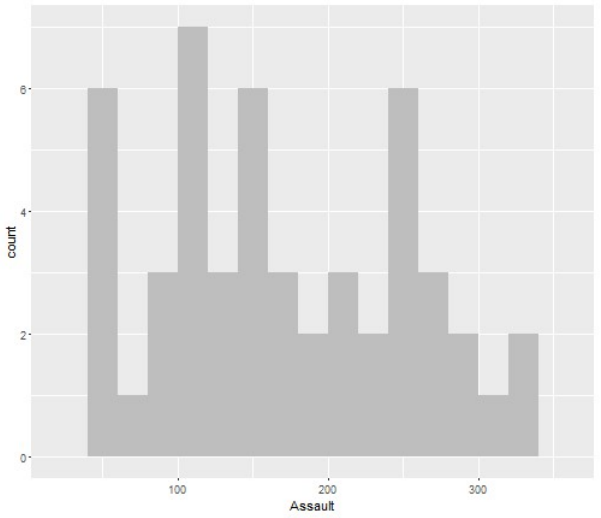
```
ggplot(USArrests)+ geom_histogram(aes(x=Murder), binwidth = 2, fill="gray") #distribution of Murders skewed to the left
```



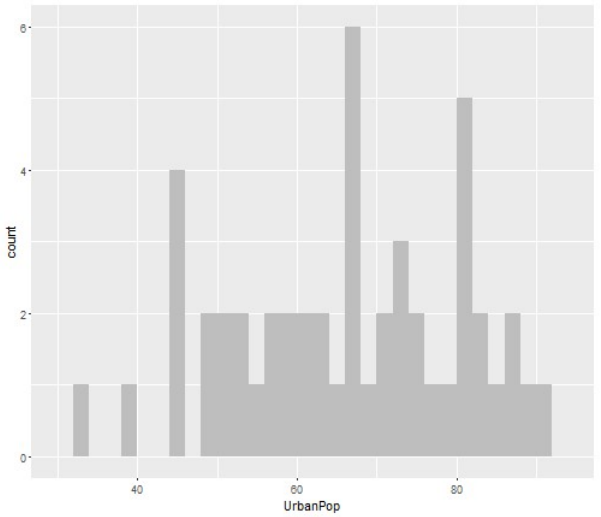

```
ggplot(USArrests)+ geom_histogram(aes(x=Rape), binwidth = 2, fill="gray") #bimodal distribution
```



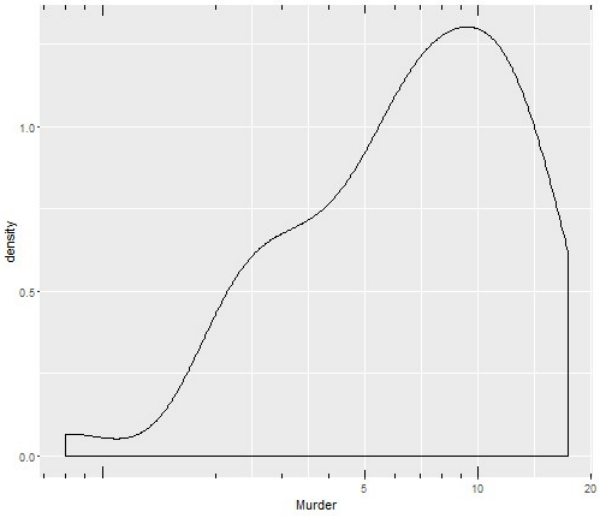

```
ggplot(USArrests)+ geom_histogram(aes(x=Assault), binwidth = 20, fill="gray")#multimodal distribution
```



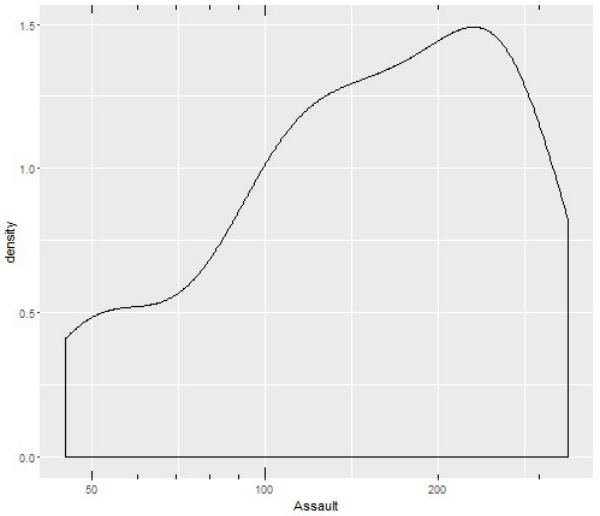

```
ggplot(USArrests)+ geom_histogram(aes(x=UrbanPop), binwidth = 2, fill="gray") #multimodal distribution
```



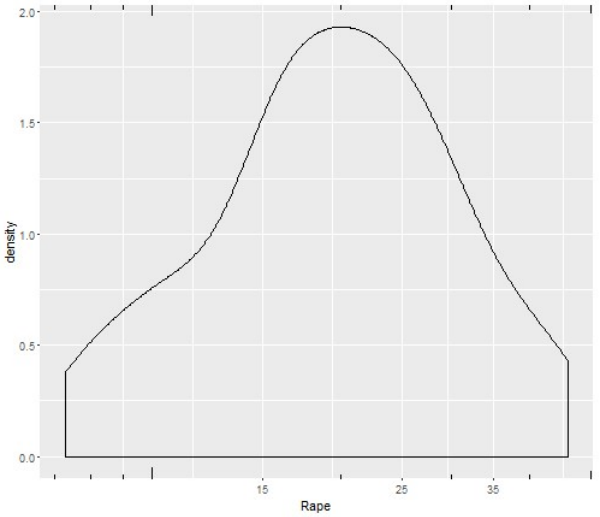

```
library(scales)

ggplot(USArrests)+ geom_density(aes(x=Murder))+ scale_x_log10(breaks=c(5, 10, 20))+ annotation_logticks(sides = "bt") #Most of the distri
```



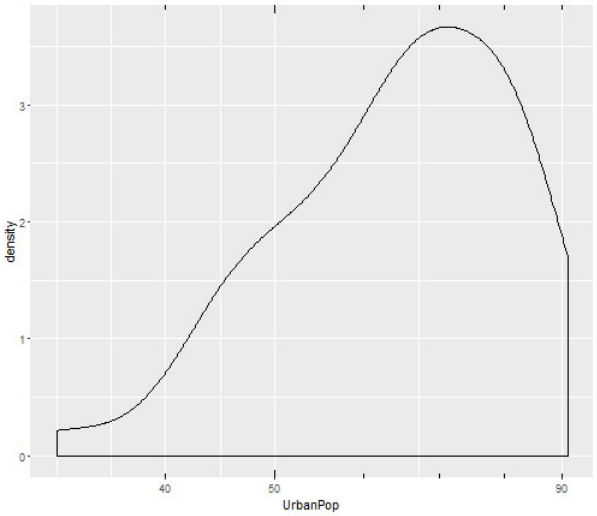

```
ggplot(USArrests)+ geom_density(aes(x=Assault))+ scale_x_log10(breaks=c(50, 100, 200))+ annotation_logticks(sides = "bt") #most of the di
```



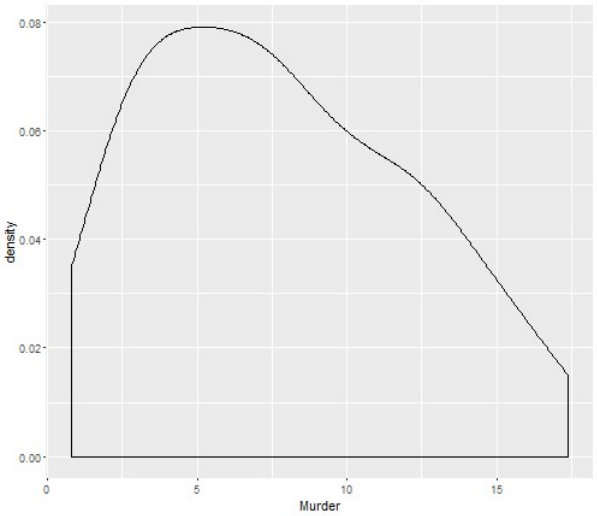

```
ggplot(USArrests)+ geom_density(aes(x=Rape))+ scale_x_log10(breaks=c(15, 25, 35))+ annotation_logticks(sides = "bt") #Most of the distrib
```



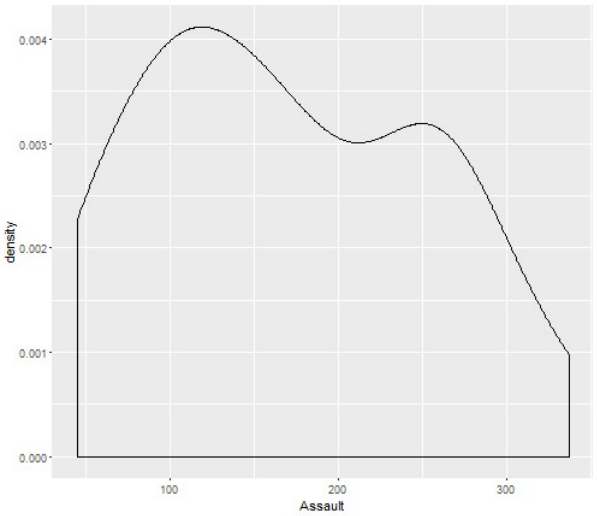

```
ggplot(USArrests)+ geom_density(aes(x=UrbanPop))+ scale_x_log10(breaks=c(40, 50, 90))+ annotation_logticks(sides = "bt") # most of the di
```



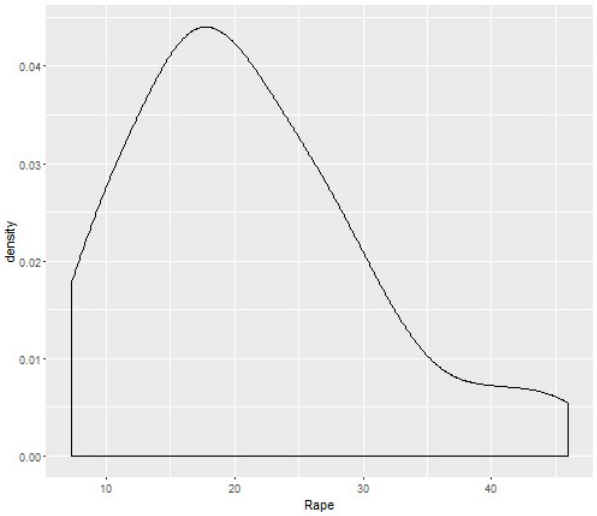

```
ggplot(USArrests)+ geom_density(aes(x=Murder))
```



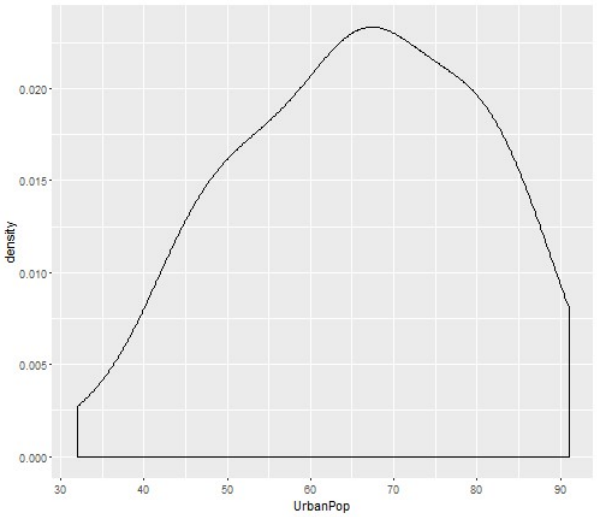

```
ggplot(USArrests)+ geom_density(aes(x=Assault))
```



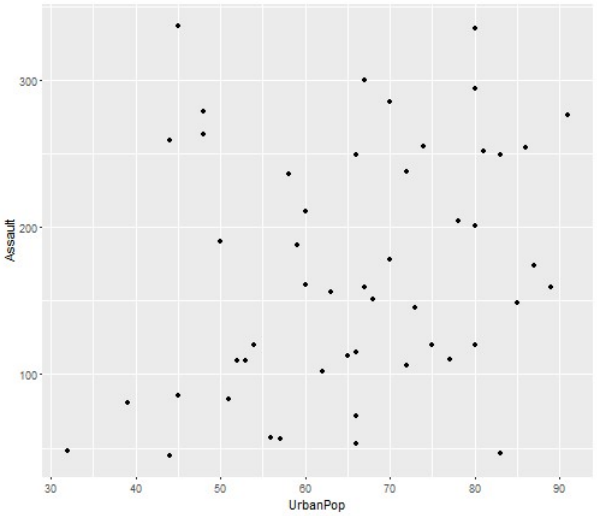

```
ggplot(USArrests)+ geom_density(aes(x=Rape))
```



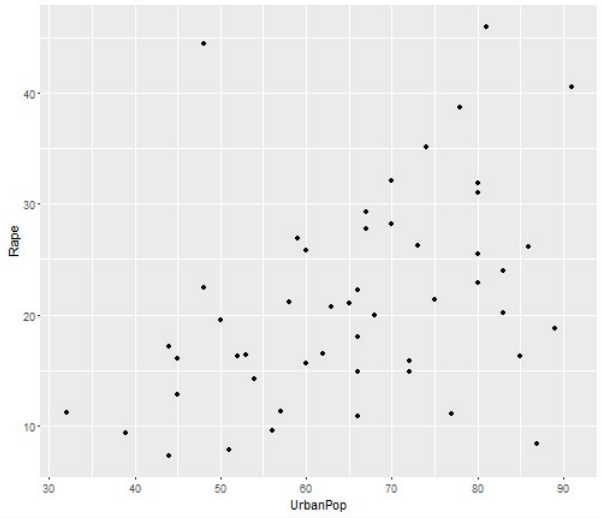

```
ggplot(USArrests)+ geom_density(aes(x=UrbanPop))
```




```
p<-ggplot(USArrests, aes(UrbanPop, Assault))
p+geom_point() #relationship between urbanpop and assault which shows positive corelation
```




```
p<-ggplot(USArrests, aes(UrbanPop, Rape))
p+geom_point() #relationship between rape and urbanpopulation which shows a positive corelation
```




```
p<-ggplot(USArrests, aes(UrbanPop, Murder))
p+geom_point() #relationship between murder and urbanpopulation which shows a positive corelation
```