US GUN VIOLENCE

Chirantan Ganguly

22/05/2020

Lets first Load the data into our system.

```
getwd()
```

```
## [1] "F:/Education/Data_Science"
```

```
setwd("F:/Education/Data_Science")
murders<-read.csv("murders.csv")
```

Let us have a look at the structure of the data we have.

```
dim(murders)
```

```
## [1] 51   5
```

```
str(murders)
```

```
## 'data.frame':    51 obs. of  5 variables:
##  $ state     : Factor w/ 51 levels "Alabama","Alaska",..: 1 2 3 4 5 6 7 8 9 10 ...
##  $ abb       : Factor w/ 51 levels "AK","AL","AR",..: 2 1 4 3 5 6 7 9 8 10 ...
##  $ region    : Factor w/ 4 levels "North Central",..: 3 4 4 3 4 4 2 3 3 3 ...
##  $ population: int  4779736 710231 6392017 2915918 37253956 5029196 3574097 897934 601723 19687653 ...
##  $ total     : int  135 19 232 93 1257 65 97 38 99 669 ...
```

Lets create a new variable violence.rate in murders to find the rate of violence per 100,000 people

```
murders$violence.rate<-(murders$total/murders$population*100000)
murders$violence.rate
```

```
##  [1]  2.8244238  2.6751860  3.6295273  3.1893901  3.3741383  1.2924531
##  [7]  2.7139722  4.2319369 16.4527532  3.3980688  3.7903226  0.5145920
## [13]  0.7655102  2.8369608  2.1900730  0.6893484  2.2081106  2.6732010
## [19]  7.7425810  0.8280881  5.0748655  1.8021791  4.1786225  0.9992600
## [25]  4.0440846  5.3598917  1.2128379  1.7521372  3.1104763  0.3798036
## [31]  2.7980319  3.2537239  2.6679599  2.9993237  0.5947151  2.6871225
## [37]  2.9589340  0.9396843  3.5977513  1.5200933  4.4753235  0.9825837
## [43]  3.4509357  3.2013603  0.7959810  0.3196211  3.1246001  1.3829942
## [49]  1.4571013  1.7056487  0.8871131
```

Lets visualise the data we just obtained.

```
argnames<-murders$abb[order(murders$violence.rate)]
barplot(height=sort(murders$violence.rate),names.arg = argnames ,space =1,
        horiz=T,las=2,xlim=c(0,20))
```