

Skill lab R session 5

Bing-Jie Yen

February 19, 2017

In this module, I would introduce how to read survey data and basic graphs in R.

Download the following files

```
#BRFSS_MN2014.sas7bdat
#CountyHealth.csv
#NYC_HearingArguments.csv
#stateRPAS_MPH5962.csv
#TCPspending_long.csv
#SAMHSA_youthSmokNDWY.csv
#SAMHSA_youthOTPNDWY.csv
#YRBS_youthSmokNDWY.csv
#YRBS_youthOTPNDWY.csv
```

Survey data analysis

Inspect smoking status variable

```
table(data$variables$X_SMOKER3)

##
##      1      2      3      4      9
## 1623   688 4651 8880   577

#recode smoking status: 0=nonsmoker; 1=smoker
library(car)

data$variables$smoke <- recode(data$variables$X_SMOKER3,
                              recodes = "1:2 = 1; 3:4=0; 9=NA")

#check recode
with(data$variables, table(smoke, X_SMOKER3, useNA = "always"))

##           X_SMOKER3
## smoke      1      2      3      4      9 <NA>
##    0         0      0 4651 8880      0      0
##    1    1623   688      0      0      0      0
##   <NA>      0      0      0      0 577      0

#get adult smoking rate for MN in 2014

sm<- svymean(x = ~smoke, design=data, na.rm=T)
sm
```

```
##           mean      SE
## smoke 0.16276 0.0037

#calculate a confidence interval

attr(sm, "var")
str(attr(sm, "var"))
smse <- attr(sm, "var")
myci <- c(lb=sm-1.96*smse, rate = sm, ub = sm+1.96*smse)

#find smoking rates by sex

svyby(~smoke, by= ~SEX, design=data, svymean)

#test for difference

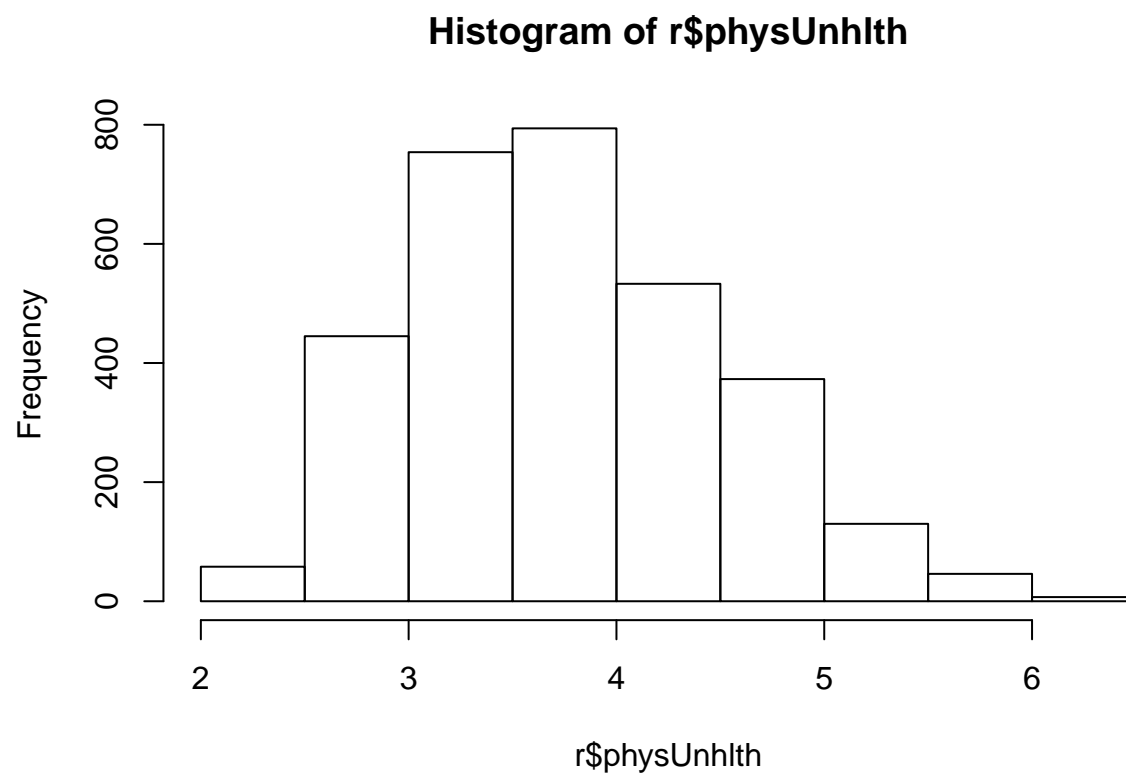
svyttest(smoke~SEX, data)

#remove the large dataset from memory
rm(data)
```

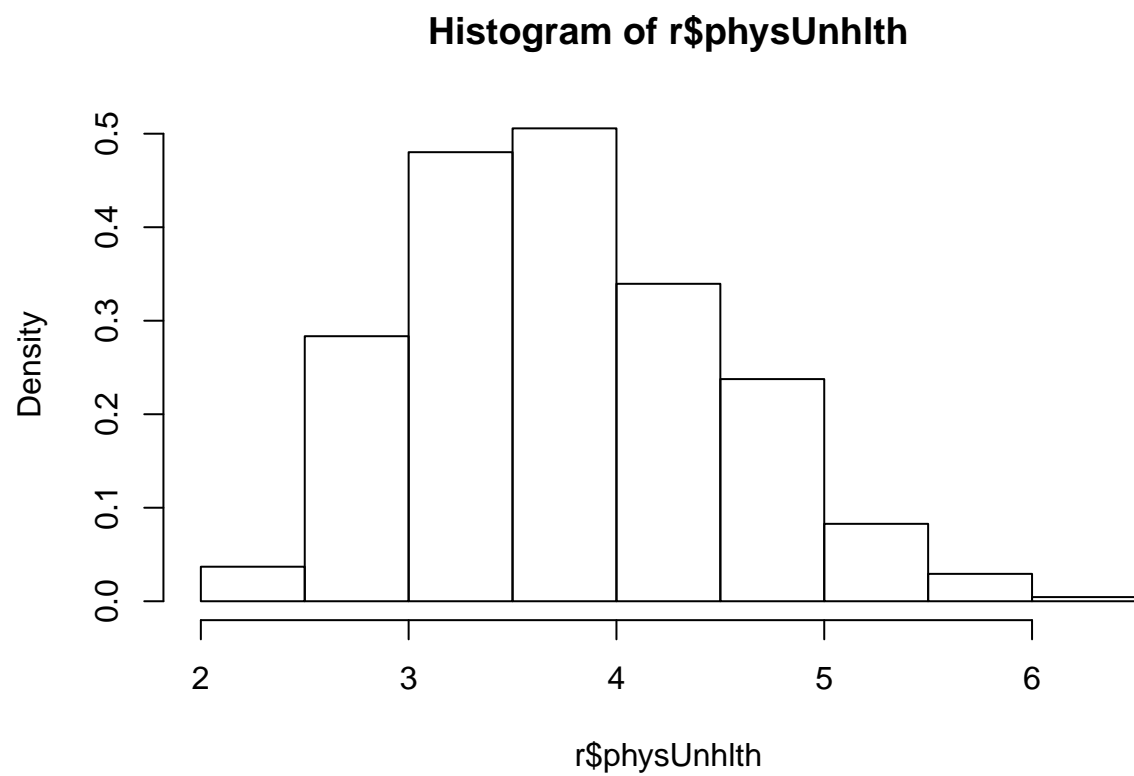
Graphics

Histograms

```
r <- read.csv("F:/Dropbox/2017 Spring/Skill Lab R/Data/CountyHealth.csv")
#histograms: default is frequency graph
hist(r$physUnhlth)
```

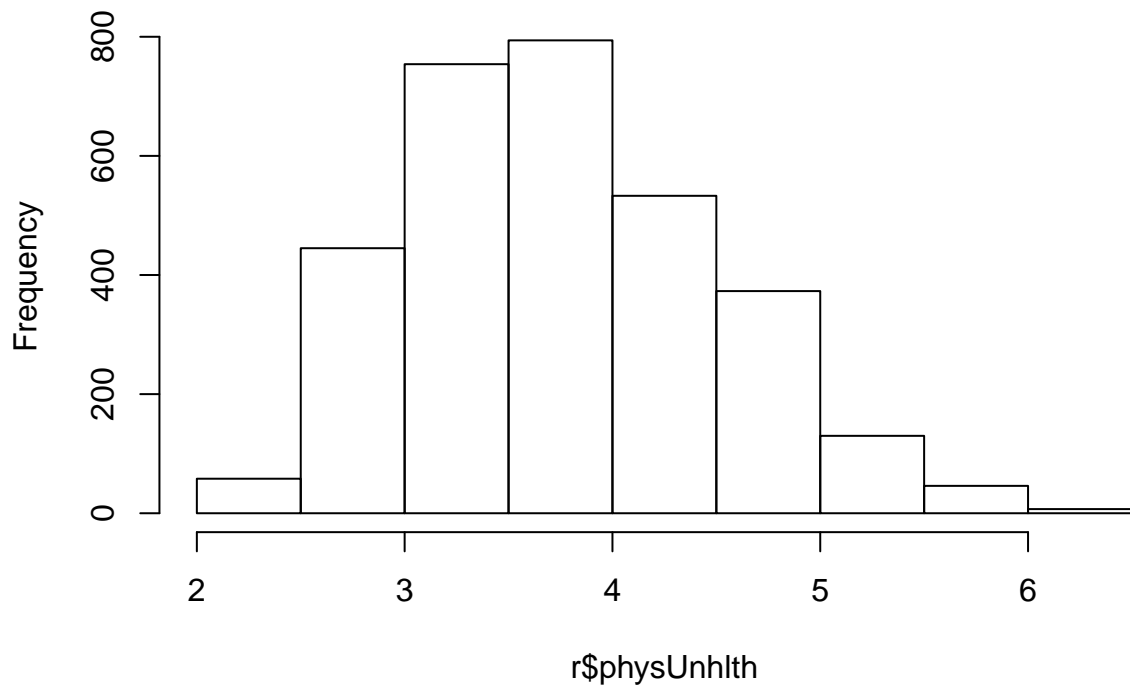


```
##hist  
# change histogram to density graph  
hist(r$physUnhlth, probability = T)
```



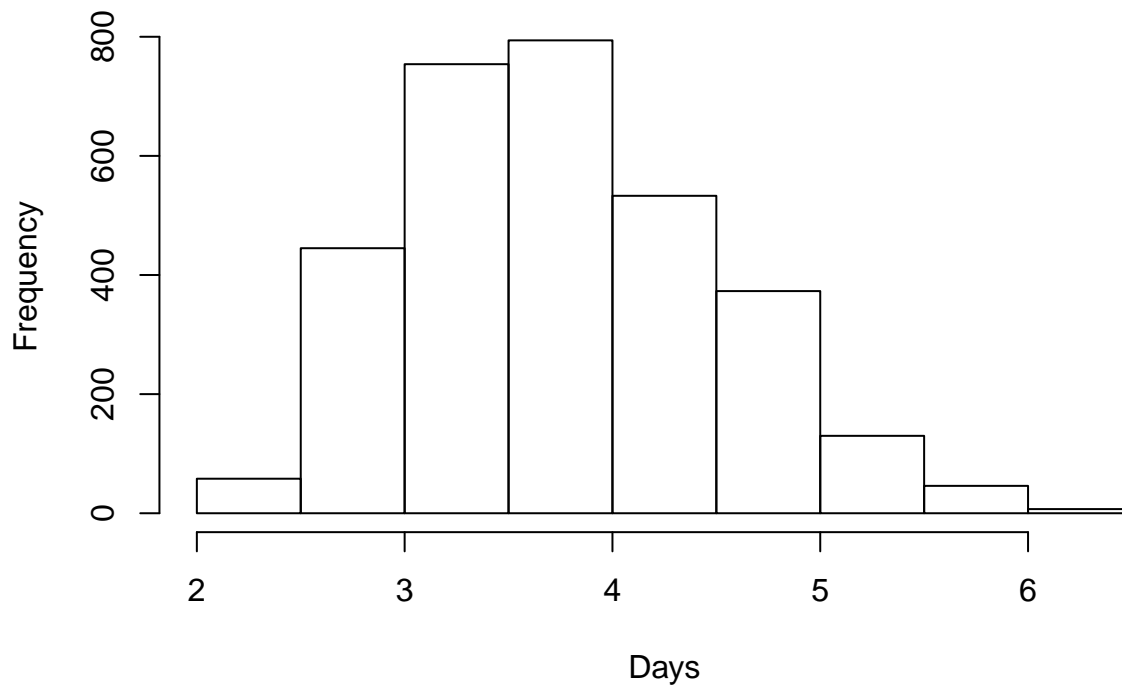
```
hist(r$physUnhlth, main="Average days per month of poor physical health for adults")
```

Average days per month of poor physical health for adults



```
hist(r$physUnhlth, main="Average days per month of poor physical health for adults",  
     xlab="Days")
```

Average days per month of poor physical health for adults

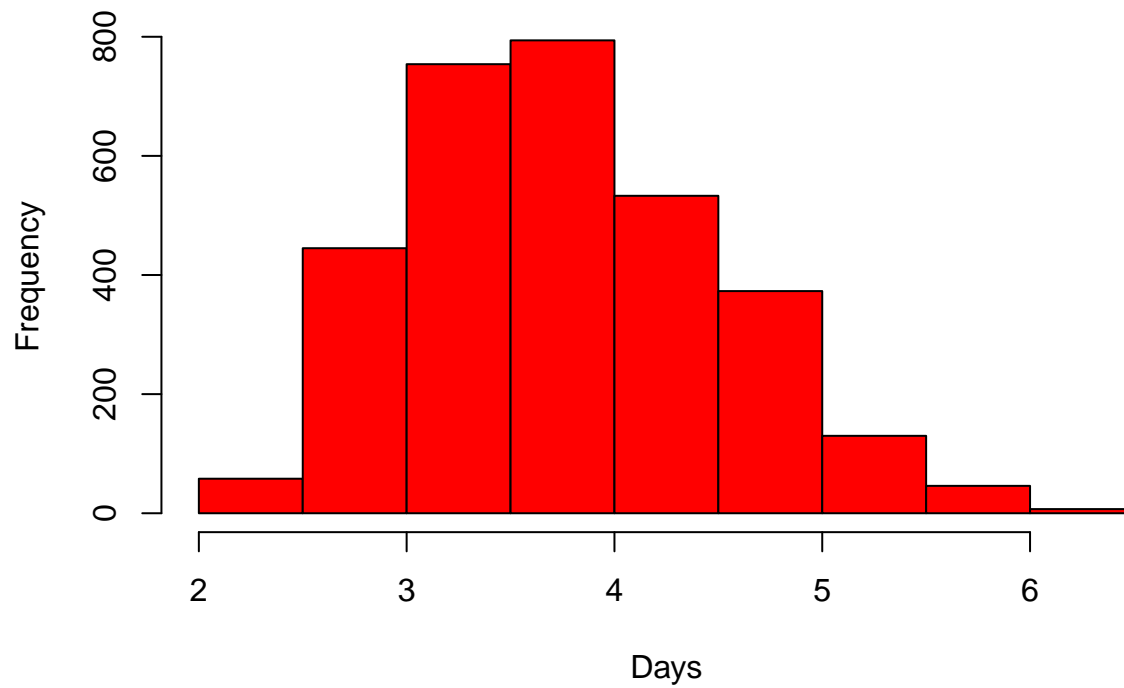


```
xx <- hist(r$physUnhlth, main="Average days per month of poor physical health for adults",  
          xlab="Days")
```

```
xx
```

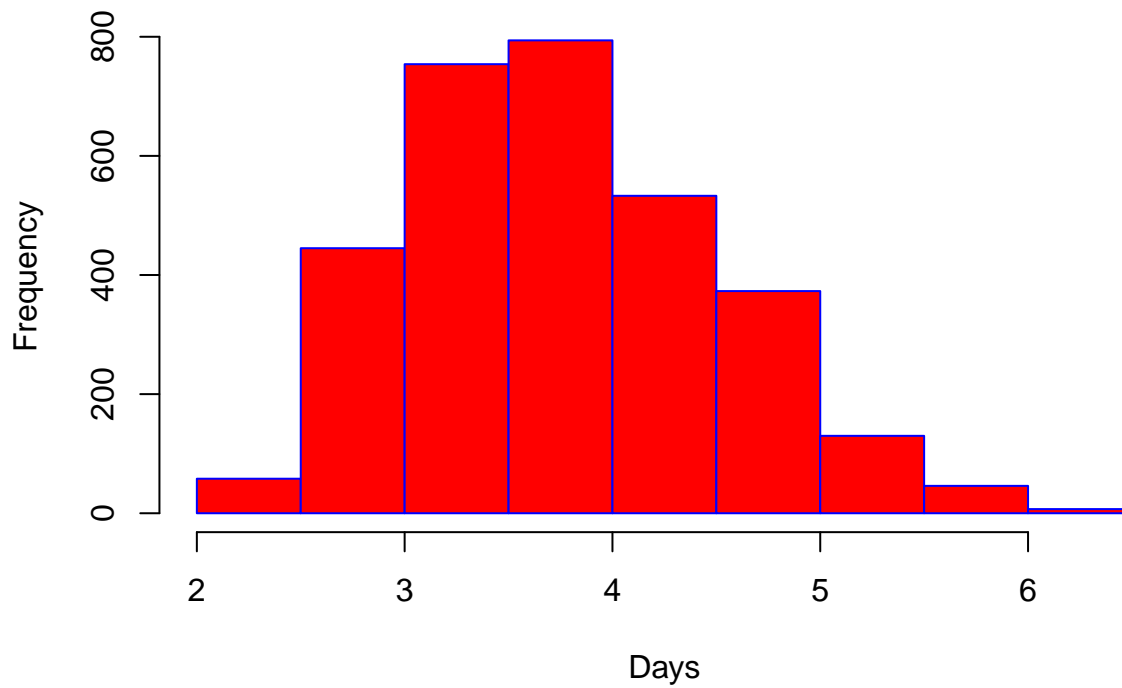
```
hist(r$physUnhlth, main="Average days per month of poor physical health for adults",  
     xlab="Days", col="red")
```

Average days per month of poor physical health for adults



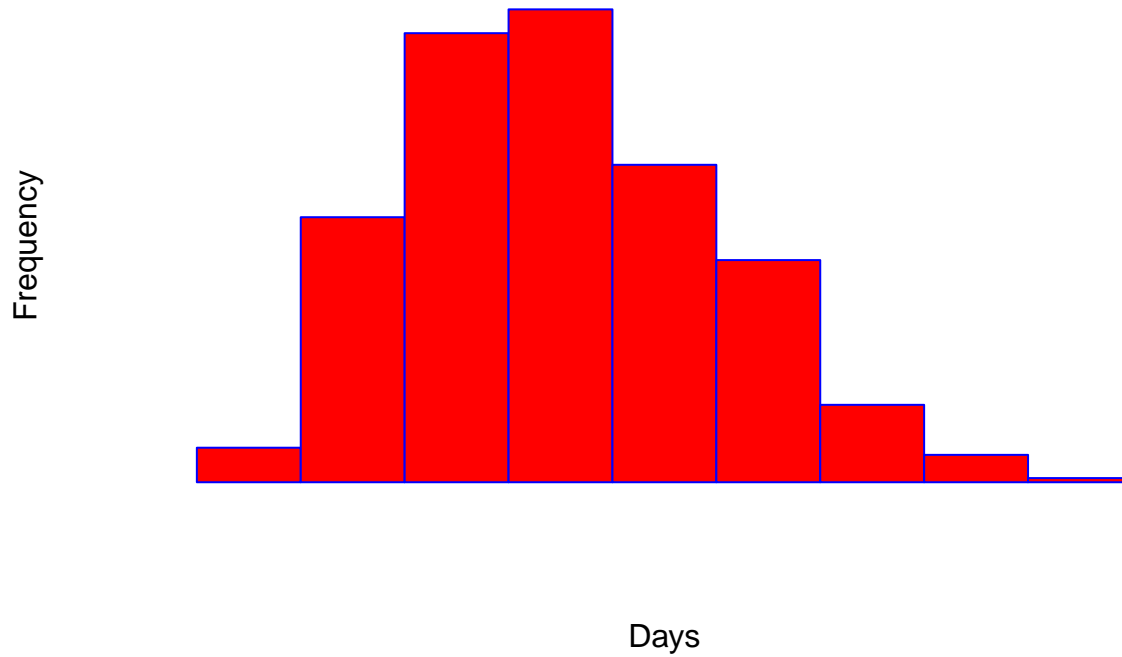
```
hist(r$physUnhlth, main="Average days per month of poor physical health for adults",  
     xlab="Days", col="red", border="blue")
```

Average days per month of poor physical health for adults



```
# axes= F means delete the axes
hist(r$physUnhlth, main="Average days per month of poor physical health for adults",
     xlab="Days", col="red", border="blue", axes=F)
```

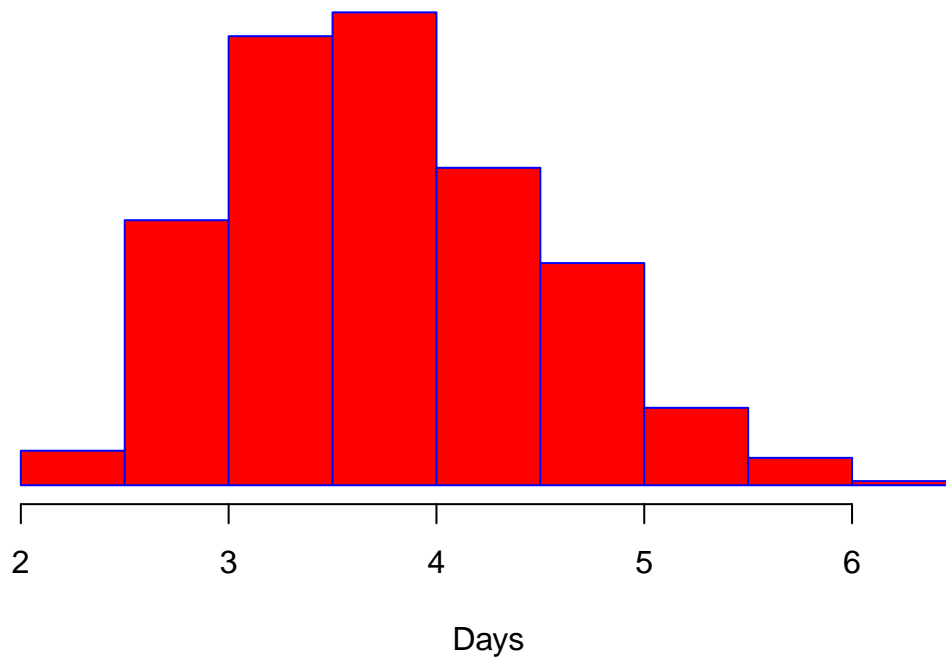

Average days per month of poor physical health for adults



```
# Delete default y label: frequency
hist(r$physUnhlth, main="Average days per month of poor physical health for adults",
     xlab="Days", col="red", border="blue", axes=F, ylab="")

# count per 1 unit
axis(1)
```

Average days per month of poor physical health for adults



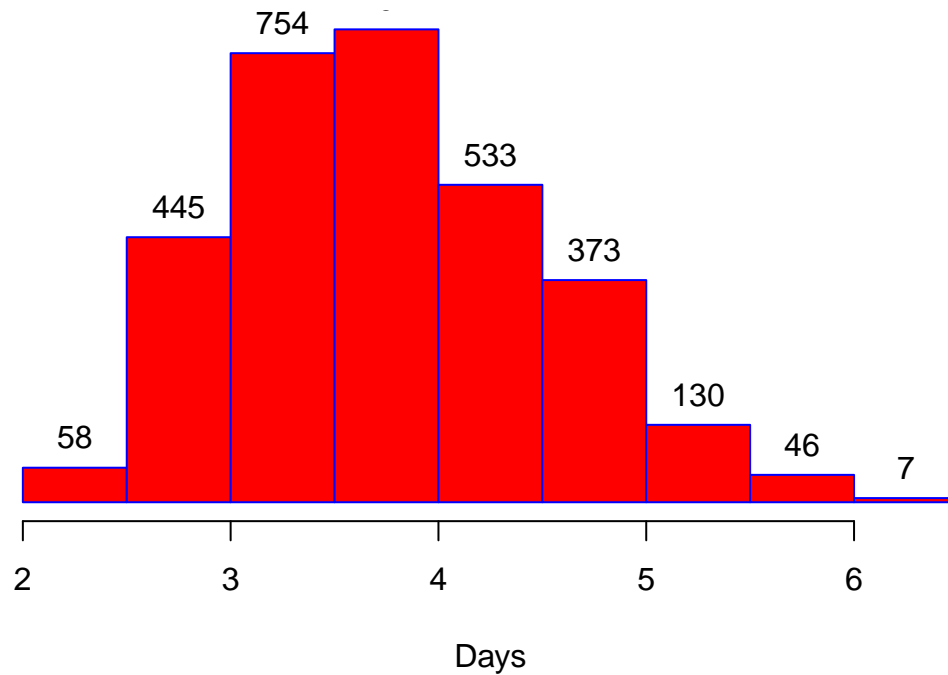
```
# Adjust the cut off (adjust the margin)
hist(r$physUnhlth, main="Average days per month of poor physical health for adults",
     xlab="Days", col="red", border="blue", axes=F, ylab="")

axis(1)

#?text

text(x=xx$mids, y=xx$counts, labels = xx$counts, pos=3)
```

Average days per month of poor physical health for adults



#Graph parameters

xpd takes two other values, TRUE and NA, which limit plotting to the figure and device region, respectively

```
par(xpd=T)
```

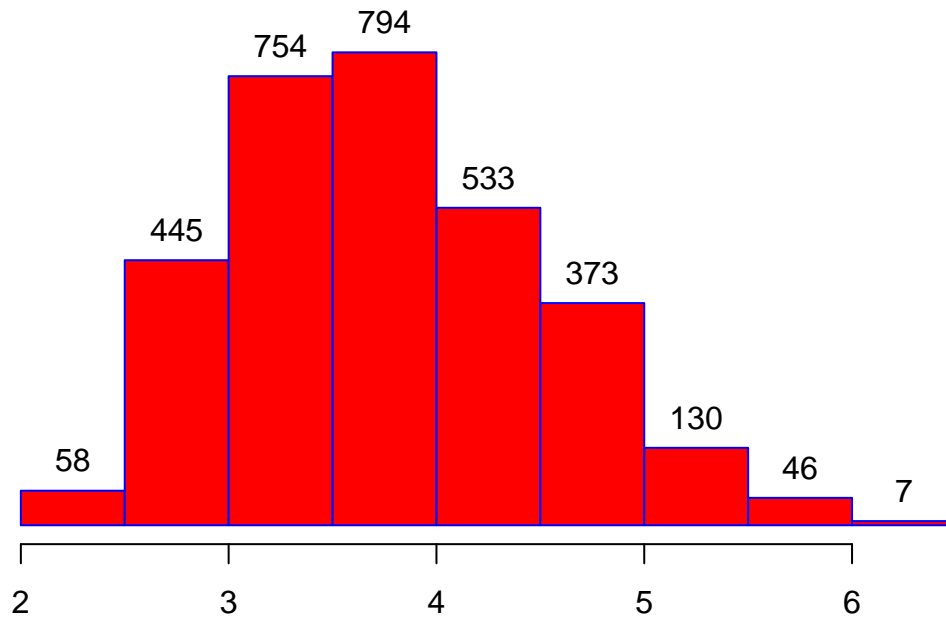
```
hist(r$physUnhlth, main="Average days per month of poor physical health for adults",  
     xlab="Days", col="red", border="blue", axes=F, ylab="")
```

```
axis(1)
```

```
text(x=xx$mids, y=xx$counts, labels = xx$counts, pos=3)
```

```
text(x= 1.5, y=-300, "Source: County Health Rankings and Roadmap 2016",pos=4, cex=0.8)
```

Average days per month of poor physical health for adults



Source: County Health Rankings and Roadmap 2016

Comparing the following two graphs

```
# Without text(x=xx$mids, y=xx$counts, labels = xx$counts, pos=3)
par(mai= c(1,1,1,1), xpd=T)

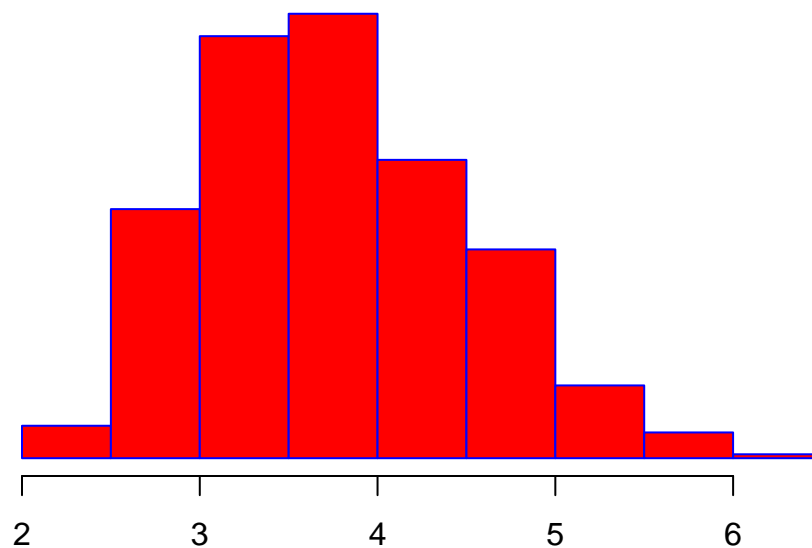
hist(r$physUnhlth, main="Average days per month of poor physical health for adults",
     xlab="Days", col="red", border="blue", axes=F, ylab="")

axis(1)

#text(x=xx$mids, y=xx$counts, labels = xx$counts, pos=3)

text(x= 1.5, y=-300, "Source: County Health Rankings and Roadmap 2016",pos=4, cex=0.8)
```

Average days per month of poor physical health for adults



Source: County Health Rankings and Roadmap 2016

```
#margins in inches
# With text(x=xx$mids, y=xx$counts, labels = xx$counts, pos=3)
par(mai= c(1,1,1,1), xpd=T)

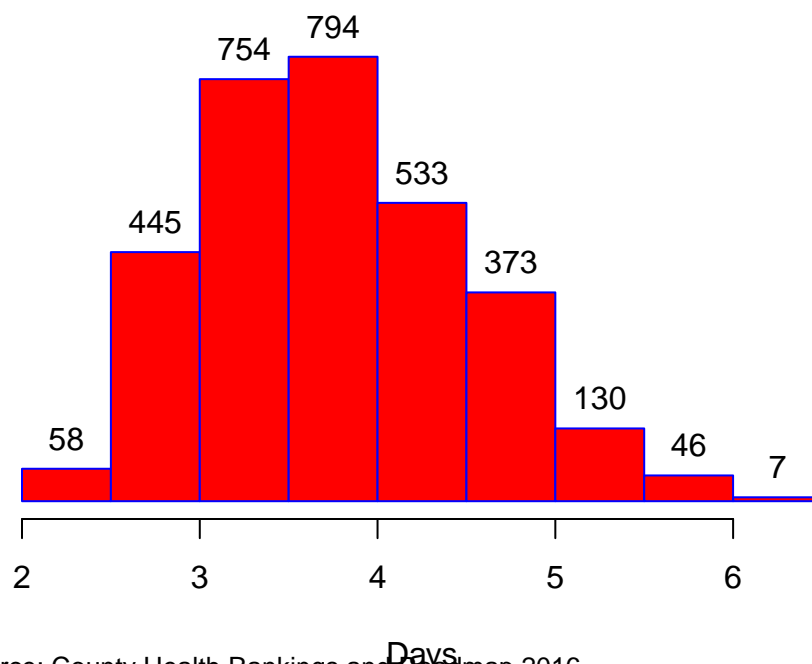
hist(r$physUnhlth, main="Average days per month of poor physical health for adults",
     xlab="Days", col="red", border="blue", axes=F, ylab="")

axis(1)

text(x=xx$mids, y=xx$counts, labels = xx$counts, pos=3)

text(x= 1.5, y=-300, "Source: County Health Rankings and Roadmap 2016",pos=4, cex=0.8)
```

Average days per month of poor physical health for adults



Source: County Health Rankings and Roadmap 2016

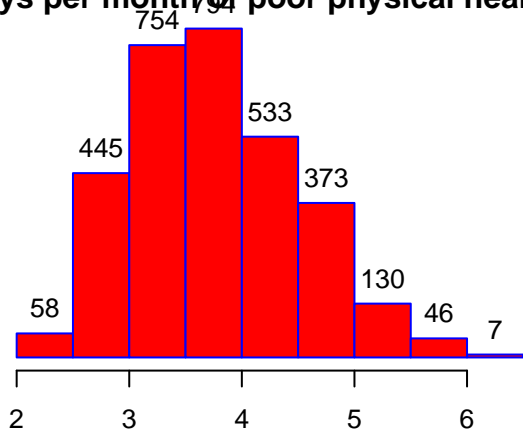
Create panels for multiple graphs

```
par(mai= c(0.2,0.2,0.2,0.2), mfrow= c(2,2), xpd=T)

hist(r$physUnhlth, main="Average days per month of poor physical health for adults",
     xlab="Days", col="red", border="blue", axes=F, ylab="")
axis(1)
text(x=xx$mids, y=xx$counts, labels = xx$counts, pos=3)

text(x= 1.5, y=-300, "Source: County Health Rankings and Roadmap 2016", pos=4, cex=0.8)
```

days per month of poor physical health f

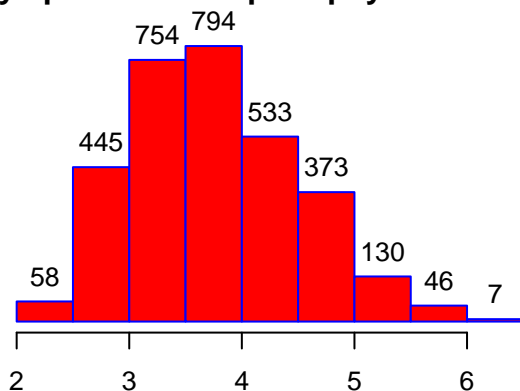


```
#increase top margin
par(mai= c(0.2,0.2,0.5,0.2), mfrow= c(2,2), xpd=T)

hist(r$physUnhlth, main="Average days per month of poor physical health for adults",
     xlab="Days", col="red", border="blue", axes=F, ylab="")
axis(1)
text(x=xx$mids, y=xx$counts, labels = xx$counts, pos=3)

text(x= 1.5, y=-300, "Source: County Health Rankings and Roadmap 2016",
     pos=4, cex=0.8)
```

days per month of poor physical health f



```
#increase bottom margin
par(mai= c(0.5,0.2,0.5,0.2), mfrow= c(2,2), xpd=T)

hist(r$physUnhlth, main="Average days per month of poor physical health for adults",
     xlab="Days", col="red", border="blue", axes=F, ylab="")
axis(1)
text(x=xx$mids, y=xx$counts, labels = xx$counts, pos=3)

text(x= 1.5, y=-300, "Source: County Health Rankings and Roadmap 2016",
     pos=4, cex=0.8)

hist(r$physUnhlth, main="Average days per month of poor physical health for adults",
     xlab="", col="red", border="blue", axes=F, ylab="")

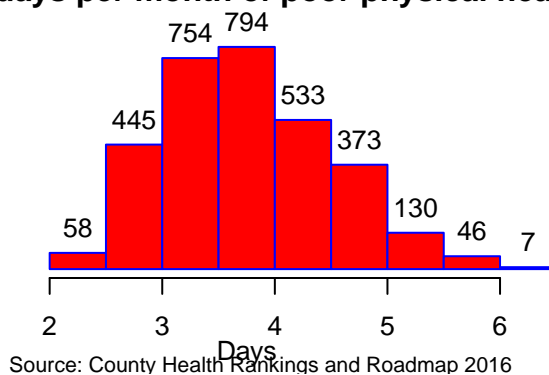
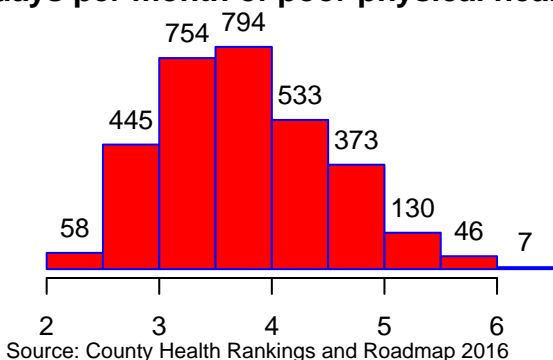
axis(1)
text(x=xx$mids, y=xx$counts, labels = xx$counts, pos=3)

text(x= 1.5, y=-350, "Source: County Health Rankings and Roadmap 2016",
     pos=4, cex=0.8)

text(x=xx$mids[round(length(xx$mids)/2)], y=-300, "Days")

par(mai= c(0.75,0.5,0.5,0.5), mfrow= c(2,2), xpd=T)
```


days per month of poor physical health f



```
hist(r$physUnhlth, main="", xlab="", col="red", border="blue", axes=F, ylab="")
hist(r$physUnhlth, ann=F, axes=F, col="purple", border="blue")

axis(1)
text(x=xx$mids, y=xx$counts, labels = xx$counts, pos=3)

text(x= 1.5, y=-500, "Source: County Health Rankings and Roadmap 2016",
      pos=4, cex=0.8)

text(x=xx$mids[round(length(xx$mids)/2)], y=-350, "Days")

text(x= xx$mids[round(length(xx$mids)/2)], y=950,
      "Average days per month of poor physical health for adults", pos=3, font=2)

text(x=1.5, y=794/2, "No. of counties", srt=90)

dev.off()

## null device
##          1
```

Graph parameters 'las'

```
par(mai= c(1,0.75,0.5,1), mfrow= c(1,2), xpd=T, las=1)

#figure c(1,1)
hist(r$physUnhlth, ann=F, axes=F, col="purple", border="blue")

axis(1)

axis(2, cex.axis=0.8)

text(x= 1.5, y=-250, "Source: County Health Rankings and Roadmap 2016",
      pos=4, cex=0.8)

text(x=xx$mids[round(length(xx$mids)/2)], y=-150, "Days")

text(x= xx$mids[round(length(xx$mids)/2)], y=850,
      "Average days per month of poor physical health for adults", pos=3, font=2, cex=0.9)

text(x=1, y=794/2, "No. of counties", srt=90)

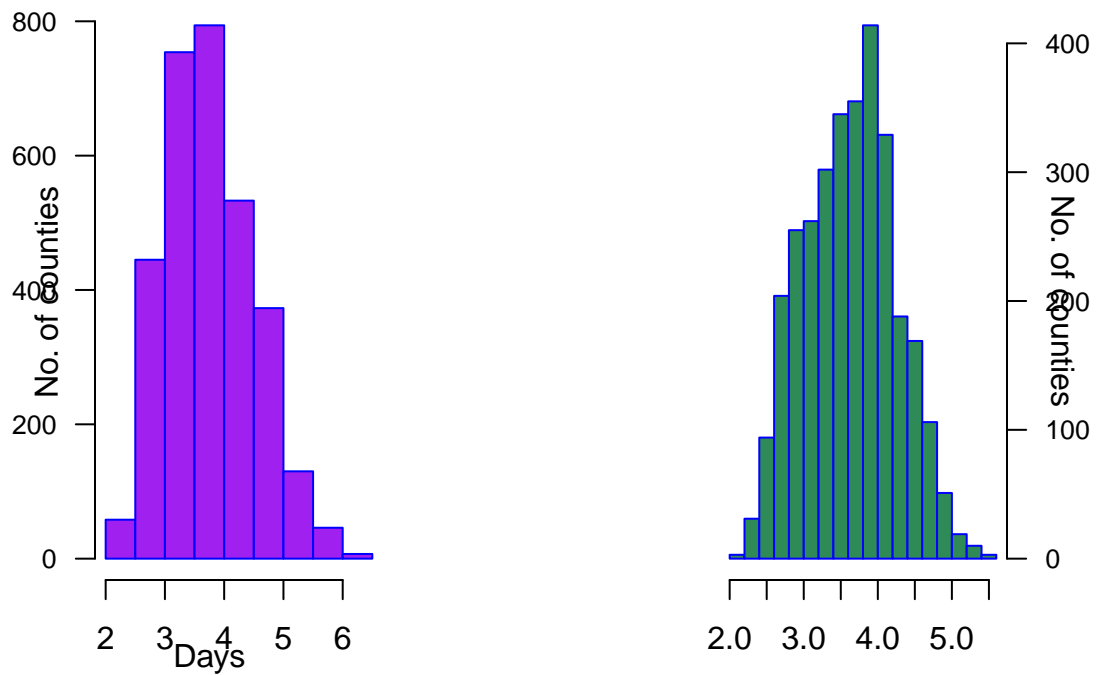
# figure c(1,2)
hist(r$mentUnhlth, ann=F, axes=F, col="seagreen", border="blue")

axis(1)

axis(4, cex.axis=0.8)

text(x=6.5, y=200, "No. of counties", srt=270)
```

Days per month of poor physical health for adults



Source: County Health Rankings and Ro

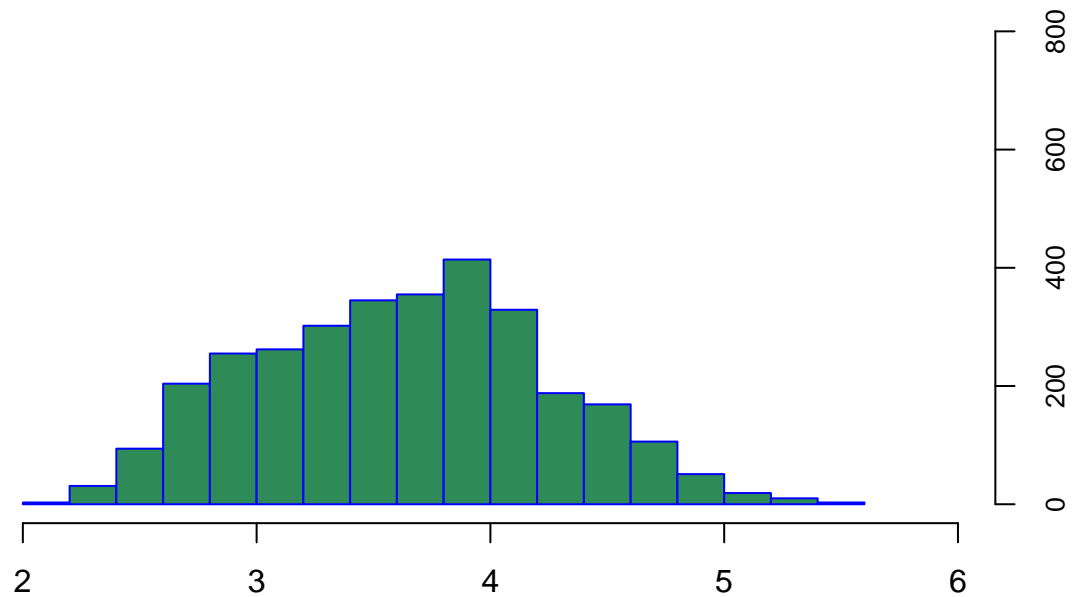
#graph parameters axis limits

```
hist(r$mentUnhlth, ann=F, axes=F, col="seagreen", border="blue",
     xlim=c(2,6), ylim=c(0,800))
```

```
axis(1)
```

```
axis(4, cex.axis=0.8)
```

```
text(x=7, y=794/2, "No. of counties", srt=270)
```



```

par(mai= c(1,0.75,0.5,1), mfrow= c(1,2), xpd=T, las=1)

hist(r$physUnhlth, ann=F, axes=F, col="purple", border="blue")

axis(1)

axis(2, cex.axis=0.8)

text(x= 1.5, y=-250, "Source: County Health Rankings and Roadmap 2016",
      pos=4, cex=0.8)

text(x=xx$mids[round(length(xx$mids)/2)], y=-150, "Days")

text(x= xx$mids[round(length(xx$mids)/2)], y=850,
      "Average days per month of poor physical health for adults", pos=3, font=2, cex=0.9)

text(x=1, y=794/2, "No. of counties", srt=90)

hist(r$mentUnhlth, ann=F, axes=F, col="seagreen", border="blue",
      xlim=c(2,6), ylim=c(0,800))

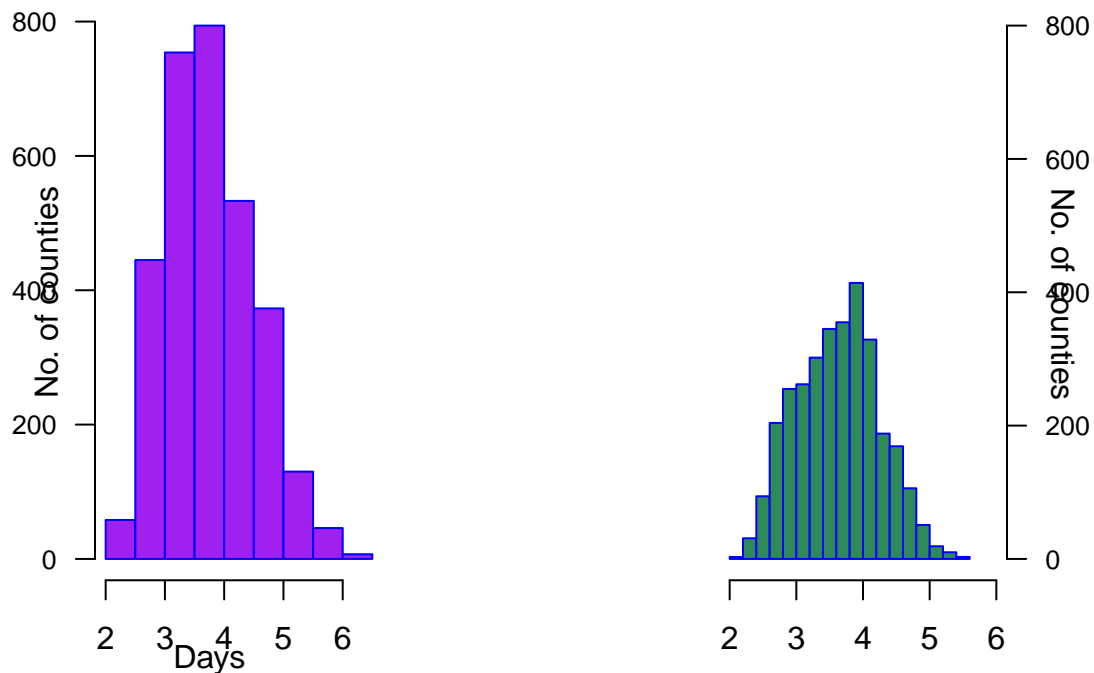
axis(1)

```

```
axis(4, cex.axis=0.8)

text(x=7, y=794/2, "No. of counties", srt=270)
```

Days per month of poor physical health for adults



Source: County Health Rankings and Roadmap 2016

```
#match breaks

par(mai= c(1,0.75,0.5,1), mfrow= c(1,2), xpd=T, las=1)

hist(r$physUnhlt, ann=F, axes=F, col="purple", border="blue")

axis(1)

axis(2, cex.axis=0.8)

text(x= 1.5, y=-250, "Source: County Health Rankings and Roadmap 2016",
      pos=4, cex=0.8)

text(x=xx$mids[round(length(xx$mids)/2)], y=-150, "Days")

text(x= xx$mids[round(length(xx$mids)/2)], y=850,
      "Average days per month of poor physical health for adults", pos=3, font=2, cex=0.9)

text(x=1, y=794/2, "No. of counties", srt=90)
```

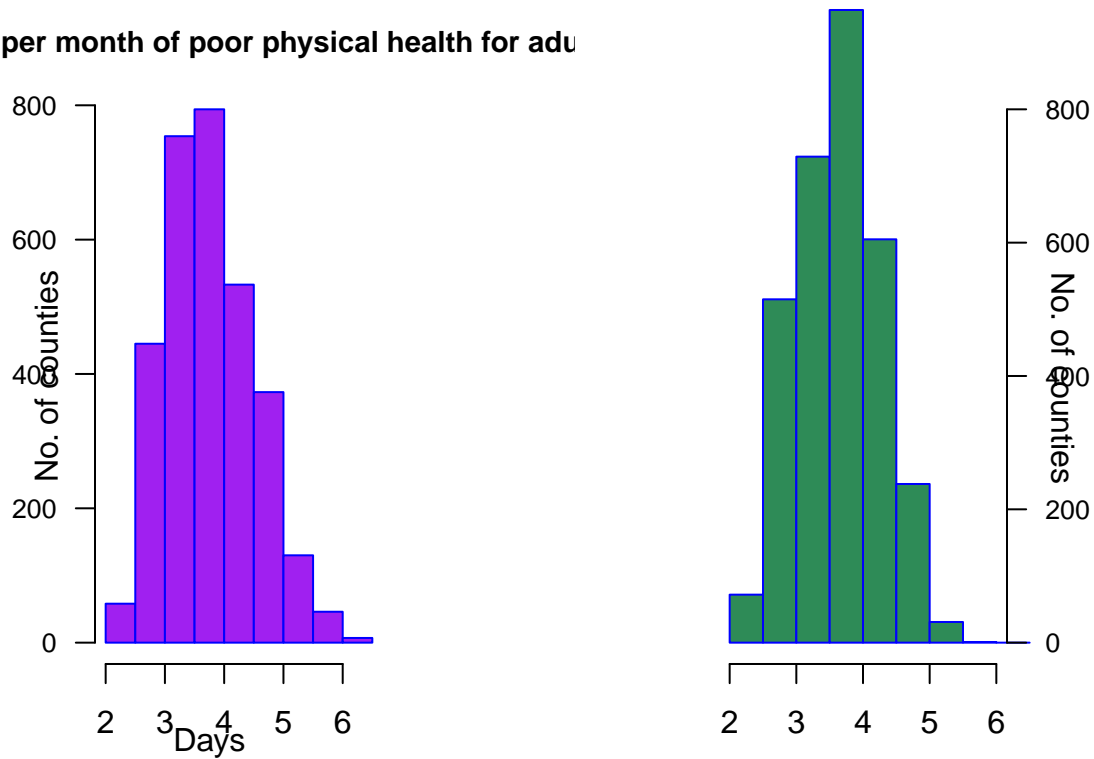
```
hist(r$mentUnhlth, ann=F, axes=F, col="seagreen", border="blue",
     xlim=c(2,6), ylim=c(0,800), breaks = xx$breaks)

axis(1)

axis(4, cex.axis=0.8)

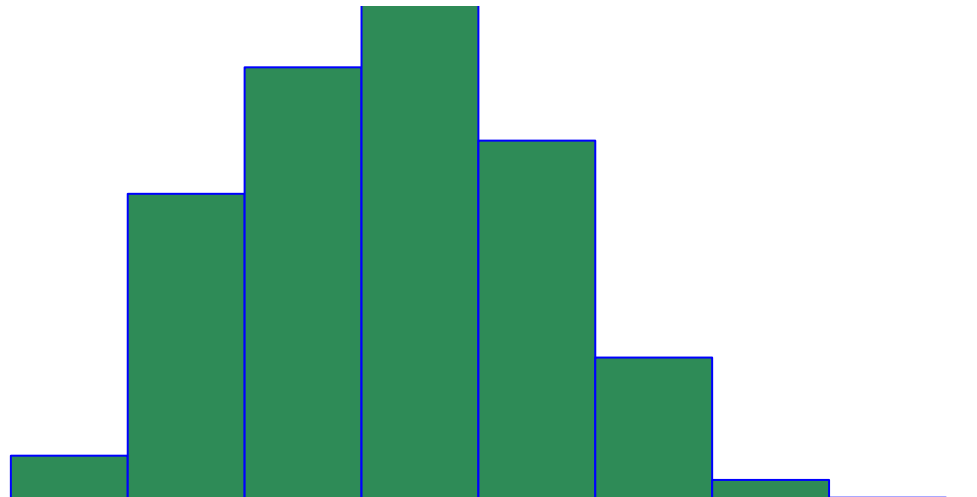
text(x=7, y=794/2, "No. of counties", srt=270)
```

ays per month of poor physical health for adu



Source: County Health Rankings and Ro

```
xm<- hist(r$mentUnhlth, ann=F, axes=F, col="seagreen", border="blue",
          xlim=c(2,6), ylim=c(0,800), breaks = xx$breaks)
```



```
xm #look at totals (counts)
```

```
## $breaks
## [1] 2.0 2.5 3.0 3.5 4.0 4.5 5.0 5.5 6.0 6.5
##
## $counts
## [1] 72 515 729 949 605 238 31 1 0
##
## $density
## [1] 0.0458598726 0.3280254777 0.4643312102 0.6044585987 0.3853503185
## [6] 0.1515923567 0.0197452229 0.0006369427 0.0000000000
##
## $mids
## [1] 2.25 2.75 3.25 3.75 4.25 4.75 5.25 5.75 6.25
##
## $xname
## [1] "r$mentUnhhlth"
##
## $equidist
## [1] TRUE
##
## attr("class")
## [1] "histogram"
```

```
par(mai= c(1,0.75,0.5,1), mfrow= c(2,1), xpd=T, las=1)
```

```
hist(r$physUnhhlth, ann=F, axes=F, col="purple", border="blue",
```

```

xlim=c(2,6.5), ylim=c(0,1000))

axis(1)

axis(2, cex.axis=0.8)

text(x= 1.5, y=-300, "Source: County Health Rankings and Roadmap 2016",
     pos=4, cex=0.8)

text(x=xx$mids[round(length(xx$mids)/2)], y=-200, "Days")

text(x= xx$mids[round(length(xx$mids)/2)], y=1050,
     "Average days per month of poor physical health for adults", pos=3, font=2, cex=0.9)

text(x=1, y=500, "No. of counties", srt=90)

hist(r$mentUnhlth, ann=F, axes=F, col="seagreen", border="blue",
     xlim=c(2,6.5), ylim=c(0,1000), breaks = xx$breaks)

axis(1)

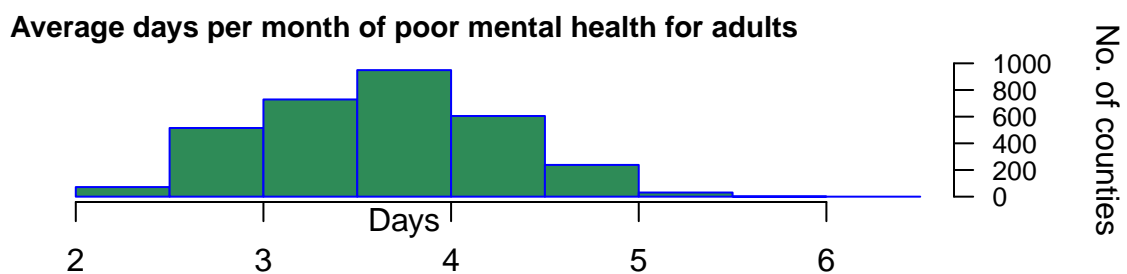
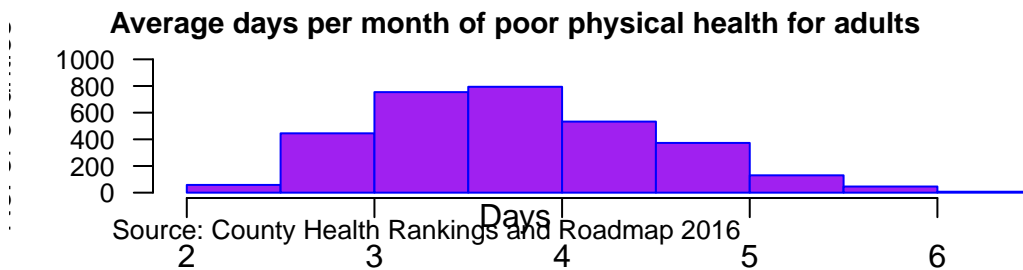
axis(4, cex.axis=0.8)

text(x=7.5, y=500, "No. of counties", srt=270)

text(x= xm$mids[round(length(xm$mids)/2)], y=1050,
     "Average days per month of poor mental health for adults", pos=3, font=2, cex=0.9)

text(x=xx$mids[round(length(xx$mids)/2)], y=-200, "Days")

```

#google 'colors in R' or 'hex color red'

Bar plots

```
r <- read.csv("F:/Dropbox/2017 Spring/Skill Lab R/Data/stateRPAS_MPH5962.csv")
```

Reshape the variables RPAS12, RPAS14, RPAS 15 to variables time and RPAS. Every time means variable+2000(years)

```
r2 <- reshape(r, varying=3:5, idvar="State", times=12:14, sep="", direction="long", drop="abbr")
r2$time <- r2$time + 2000
```

```
rt <- by(r2$RPAS, r2$time, table)
rt
```

```
## r2$time: 2012
```

```
##
```

```
## 0 1 3 4 5 6 8 9 10 12 13 15 16 21 23 25
```

```
## 7 3 1 8 1 2 6 4 3 6 1 1 2 1 1 1
```

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

```
## r2$time: 2014
```

```
##
```

```
## 1 2 3 4 6 7 8 9 10 11 12 13 14 15 17 18 22 23 24 28
```

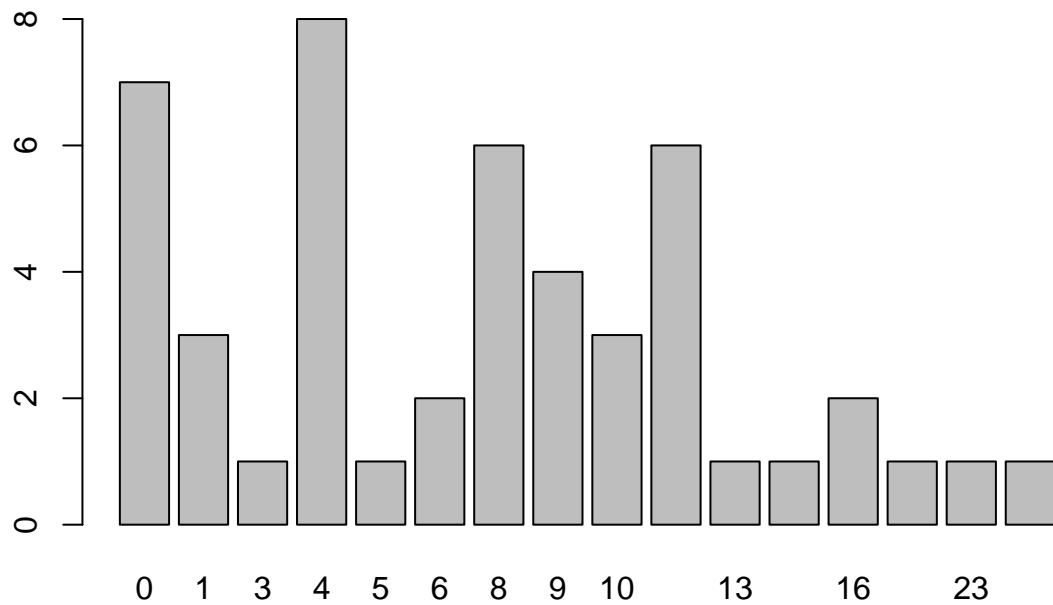
```
## 1 1 2 1 3 2 6 1 1 4 2 3 2 7 2 3 1 4 1 1
```

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

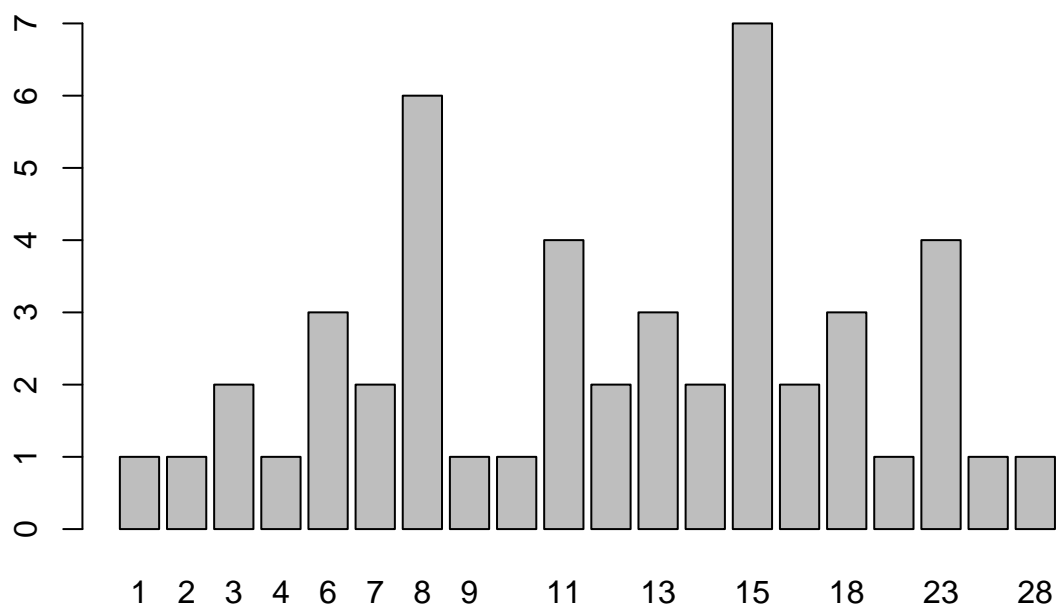
```
## r2$time: 2015
##
##  3  5  6  7  9 10 11 12 13 14 16 17 18 19 20 21 23 25 26 27 28 33
##  4  1  5  1  5  4  4  2  1  3  1  3  3  1  1  4  1  2  1  1  1  1
```

Since “rt” includes 3 tables which include 2012, 2014 and 2015, we need to specify which table we want to show

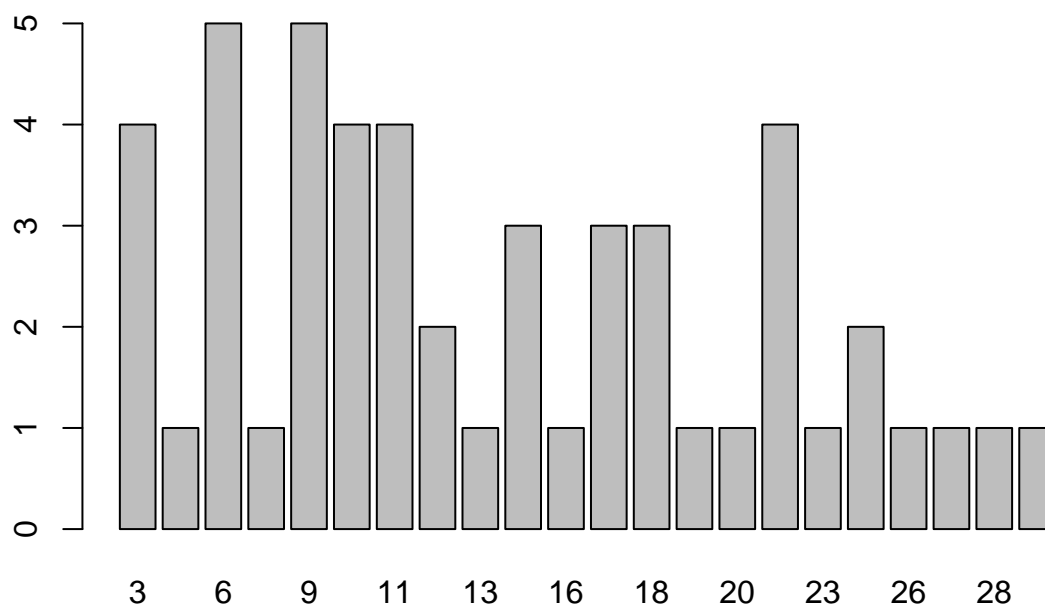
```
barplot(rt[[1]]) # 2012
```



```
barplot(rt[[2]]) # 2014
```



```
barplot(rt[[3]]) # 2015
```



```
library(car)
r2$scorecat <- recode(r2$RPAS, recodes=
  " 0:5='0-5';
  6:10='6-10';
  11:15='11-15';
  16:20= '16-20';
  21:25='21-25';
  26:30='26-30';
  31:33='31-33'", as.factor.result=T)

rt <- by(r2$scorecat, r2$time, table)

# levels: order the categories
r2$scorecat <- recode(r2$RPAS, recodes=
  " 0:5='0-5';
  6:10='6-10';
  11:15='11-15';
  16:20= '16-20';
  21:25='21-25';
  26:30='26-30';
  31:33='31-33'", as.factor.result=T,
  levels = c('0-5', '6-10', '11-15', '16-20',
    '21-25', '26-30', '31-33'))

rt <- by(r2$scorecat, r2$time, table)
rt
```

```

## r2$time: 2012
##
##   0-5  6-10 11-15 16-20 21-25 26-30 31-33
##   20   15   8     2     3     0     0
## -----
## r2$time: 2014
##
##   0-5  6-10 11-15 16-20 21-25 26-30 31-33
##   5    13   18    5     6     1     0
## -----
## r2$time: 2015
##
##   0-5  6-10 11-15 16-20 21-25 26-30 31-33
##   5    15   10    9     7     3     1

str(rt)

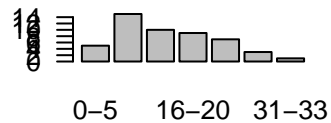
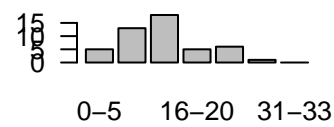
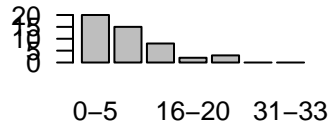
## List of 3
## $ 2012: 'table' int [1:7(1d)] 20 15 8 2 3 0 0
##   ..- attr(*, "dimnames")=List of 1
##     .. ..$ : chr [1:7] "0-5" "6-10" "11-15" "16-20" ...
## $ 2014: 'table' int [1:7(1d)] 5 13 18 5 6 1 0
##   ..- attr(*, "dimnames")=List of 1
##     .. ..$ : chr [1:7] "0-5" "6-10" "11-15" "16-20" ...
## $ 2015: 'table' int [1:7(1d)] 5 15 10 9 7 3 1
##   ..- attr(*, "dimnames")=List of 1
##     .. ..$ : chr [1:7] "0-5" "6-10" "11-15" "16-20" ...
## - attr(*, "dim")= int 3
## - attr(*, "dimnames")=List of 1
##   ..$ r2$time: chr [1:3] "2012" "2014" "2015"
## - attr(*, "call")= language by.default(data = r2$scorecat, INDICES = r2$time, FUN = table)
## - attr(*, "class")= chr "by"

par(mai=rep(1,4), xpd=T, las=1, mfrow=c(2,2))

for(i in 1:3) {barplot(rt[[i]])}

#clean up (hor, names, colors, borders, labels, axis limits, margins)

```



Line charts

TCP spending

```
r <- read.csv("F:/Dropbox/2017 Spring/Skill Lab R/Data/TCPspending_long.csv")
```

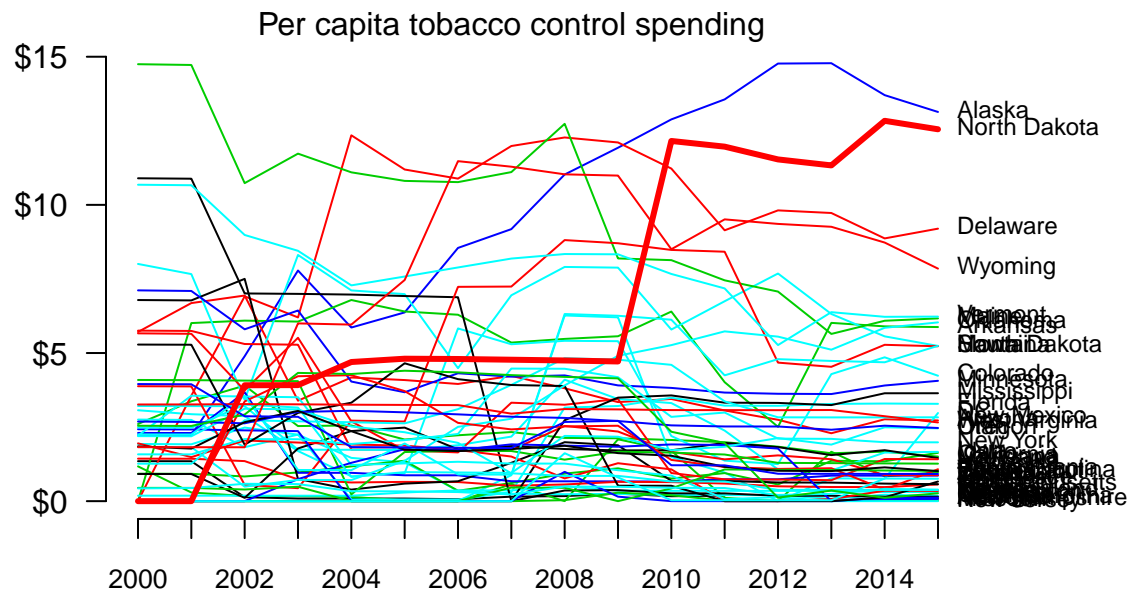
In the following figure, the North Dakota is emphasized with thicker line.

```
par(mai = c(1,1,1,1), family = "sans", xpd = T, las= 1)
plot(c(2000,2015), c(0,15), type = "n", ann = F, axes = F)
# plot(c(2000,2015), c(0,15), type = "p", ann = F, axes = F)

for (i in unique(r$state)){
  r2 <- r[r$state == i,]
  with(r2, lines(x = time, y = percap, col = sample(1:5)))
}
with(r[r$state == "North Dakota",], lines(x = time, y = percap, lwd = 3, col = "red"))
axis(2, at=seq(0,15,5), labels= paste("$",seq(0,15,5), sep=""))
axis(1, 2000:2015, cex.axis=0.8)

for (i in 1:51){
  r2 <- r[r$state == unique(r$state)[i]& r$time==2015,]
  text(2015, r2$percap, r2$state, pos=4, cex=0.75)
}
```

```
text(2007, 16, "Per capita tobacco control spending")
```



```
#SAMHSA & YRBS for ND & WY
```

```
r <- read.csv("F:/Dropbox/2017 Spring/Skill Lab R/Data/SAMHSA_youthSmokNDWY.csv")
r2 <- read.csv("F:/Dropbox/2017 Spring/Skill Lab R/Data/SAMHSA_youthOTPNDWY.csv")
r3 <- read.csv("F:/Dropbox/2017 Spring/Skill Lab R/Data/YRBS_youthSmokNDWY.csv")
r4 <- read.csv("F:/Dropbox/2017 Spring/Skill Lab R/Data/YRBS_youthOTPNDWY.csv")
```

```
layout(matrix(c(1,2,3,3,4,5), 3,2, byrow=T), heights = c(3,1,3))
```

```
mycol <- c("blue", "brown", "purple", "forestgreen")
```

```
myylim <- c(0, 45)
```

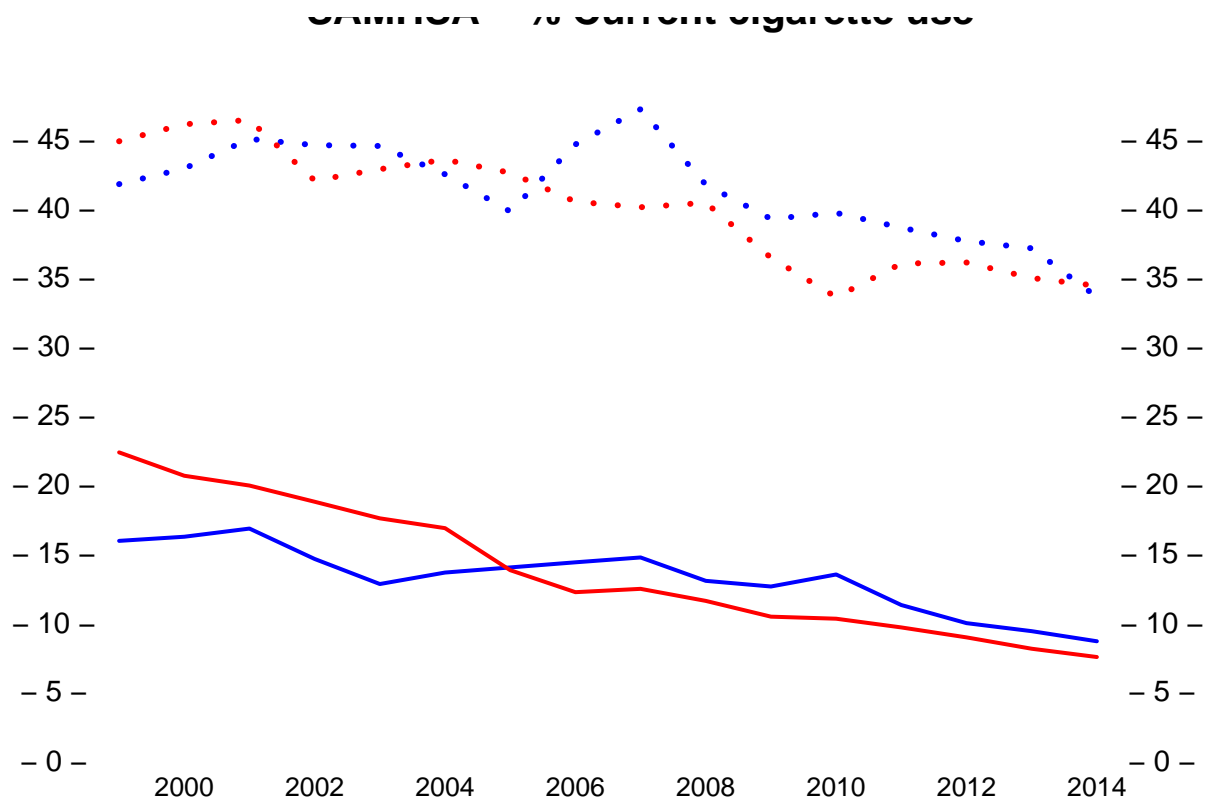
In the following, it shows four graphs with 2x2 order. In each sub-figure,

```
# PLOT 1 - adolescents 12-17 cigarette SAMHSA
```

```
par(mai= c(0.5,0.5,0.5,0.5), xpd= T, las=1)
plot(age1~year, r[r$state=="Wyoming",], ylim= myylim, type="l", col=mycol[1], lwd=2, axes=F, ann=F)
with( r[r$state=="North Dakota",], lines(x = year, y = age1, col = "red", lwd=2))
text(1998, seq(0,45,5), paste("-", seq(0,45,5), "-"), cex=0.95)
text(seq(2000,2014,2),0, seq(2000,2014,2), cex=0.85, pos=1)
text(2015, seq(0,45,5), paste("-", seq(0,45,5), "-"), cex=0.95)
text(2007, 54, "SAMHSA - % Current cigarette use", font=2, cex=1.25)
```

```
#young adults 18-25 cigarette
```

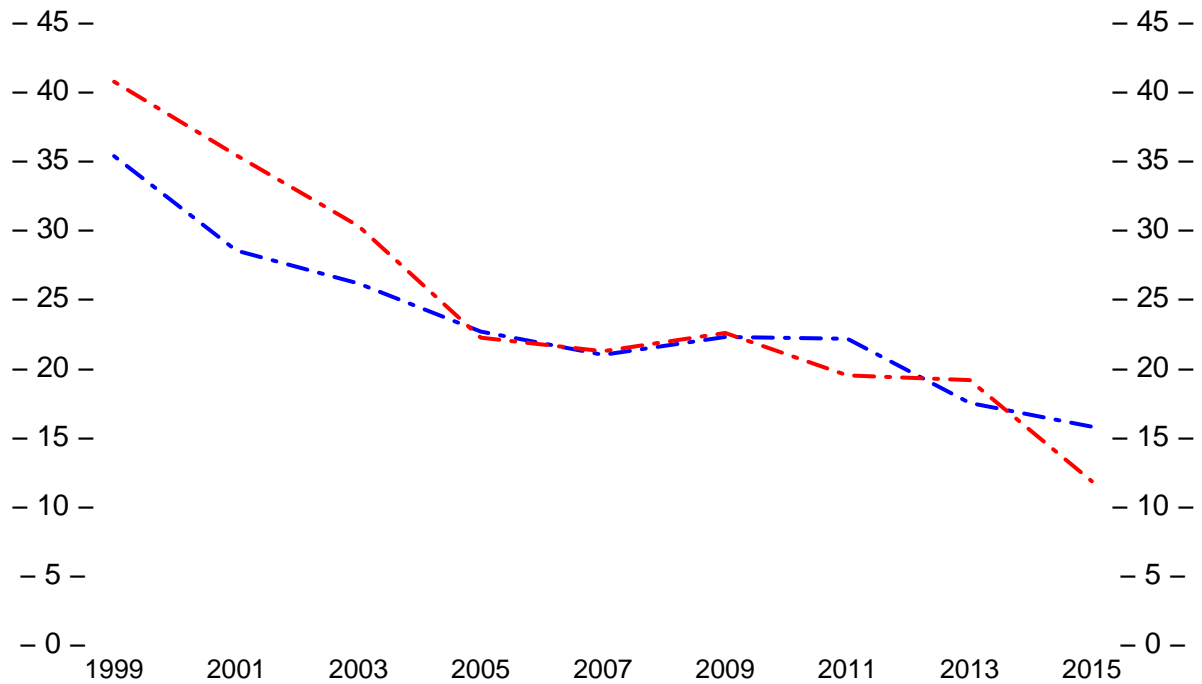
```
with( r[r$state=="Wyoming",], lines(x = year, y = age2, col = mycol[1], lwd=3, lty=3))
with( r[r$state=="North Dakota",], lines(x = year, y = age2, col = "red", lwd=3, lty=3))
```



```
#PLOT 2 - middle & high schoolers smoke YRBS
```

```
par(mai= c(0.5,0.5,0.5,0.5), xpd= T, las=1)
plot(smokany30~year, r3[r3$state=="Wyoming",], ylim= myylim, type="l", col=mycol[1], lwd=2, lty=6, axes=F)
with( r3[r3$state=="North Dakota",], lines(x = year, y = smokany30, col = "red", lwd=2, lty=6))
text(1998, seq(0,45,5), paste("-", seq(0,45,5), "-"), cex=0.95)
text(seq(1999,2015,2), 0, seq(1999,2015,2), cex=0.85, pos=1)
text(2016, seq(0,45,5), paste("-", seq(0,45,5), "-"), cex=0.95)
text(2007, 54, "YRBS - % Current cigarette use", font=2, cex=1.25)
```


TRDS - % Current cigarette use



```
#Legend 2
par(mai=c(0,0,0,0))
plot(c(0,10), c(0,10), type="n", ann=F, axes=F)
segments(x0=1,x1=2, y0=c(9,7,5,3), col=c("blue", "red", "blue", "red"), lwd=2, lty= c(3,3,1,1))
text(2,c(9,7,5,3), c("WY young adults (18-24)",
                    "ND young adults (18-24)",
                    "WY adolescents (12-17)",
                    "ND adolescents (12-17)"), col=c("blue", "red", "blue", "red"),pos=4, font=2)

segments(x0=6, x1=7, y0 = c(7,5), col = c("blue", "red"), lwd=2, lty=6)

text(7,c(7,5), c("WY middle & high school students",
                "ND middle & high school students"), col = c("blue", "red"), pos=4, font=2)
```

..... WY young adults (18–24)

..... ND young adults (18–24) - - - - - WY middle & high school

———— WY adolescents (12–17) - - - - - ND middle & high school

———— ND adolescents (12–17)

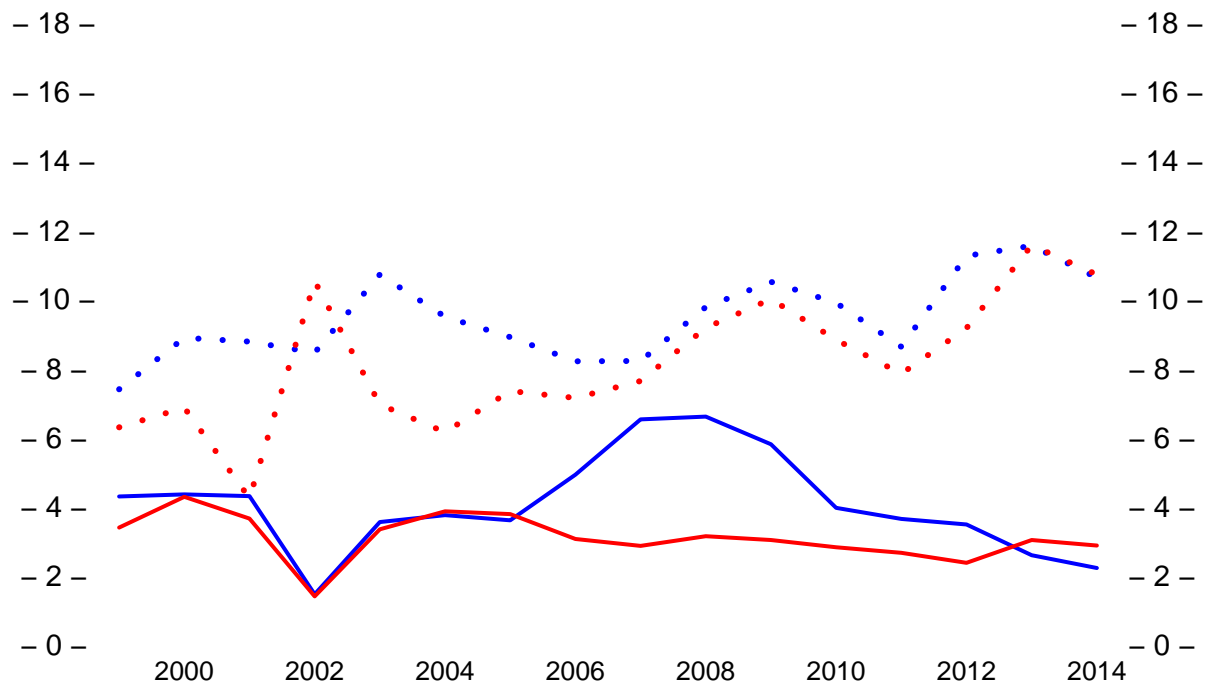
```
#PLOT 3 - adolescents 12-17 otp SAMHSA
```

```
myylim <- c(0, 18)
```

```
par(mai= c(0.5,0.5,0.5,0.5), xpd= T, las=1)
plot(age1~year, r2[r2$state=="Wyoming",], ylim= myylim, type="l", col=mycol[1], lwd=2, axes=F, ann=F)
with( r2[r2$state=="North Dakota",], lines(x = year, y = age1, col = "red", lwd=2))
text(1998, seq(0,18,2), paste("-",seq(0,18,2), "-"), cex=0.95)
text(seq(2000,2014,2),0, seq(2000,2014,2), cex=0.85, pos=1)
text(2015, seq(0,18,2), paste("-", seq(0,18,2), "-"), cex=0.95)
```

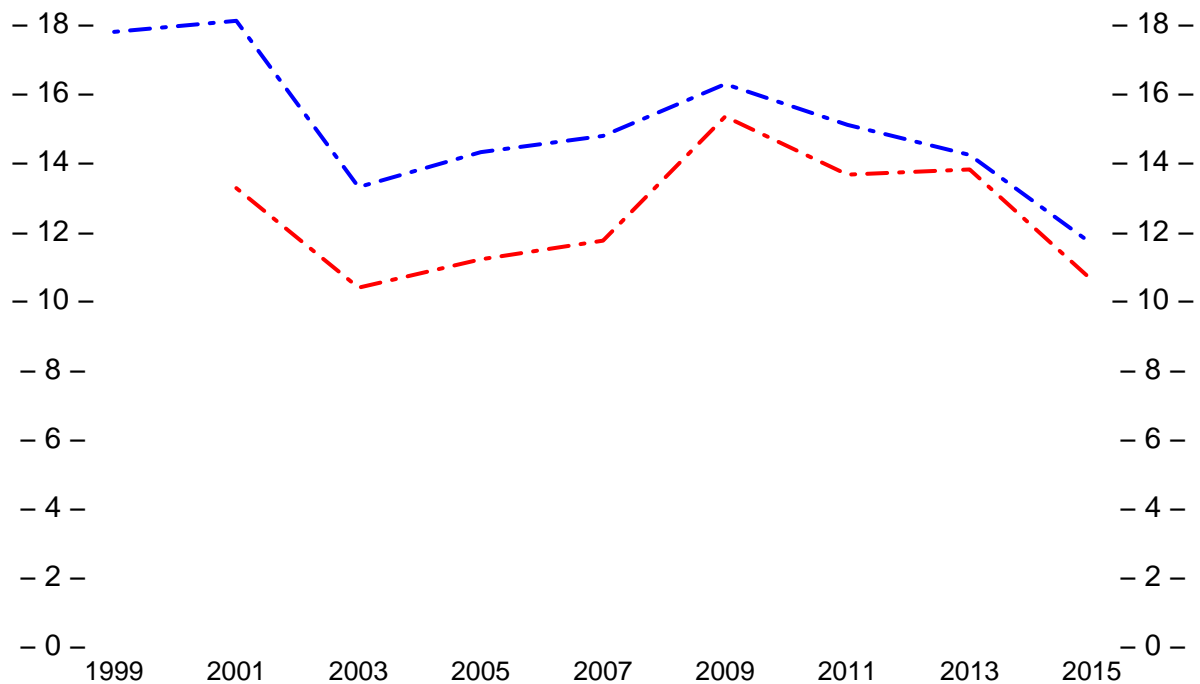
```
#young adults 18-25 otp
```

```
with( r2[r2$state=="Wyoming",], lines(x = year, y = age2, col = mycol[1], lwd=3, lty=3))
with( r2[r2$state=="North Dakota",], lines(x = year, y = age2, col = "red", lwd=3, lty=3))
text(2007,22,"SAMHSA - % Current non-cigarette tobacco use", font=2, cex=1.25 )
```



#PLOT 4 - middle & high schoolers smokeless YRBS

```
plot(smokeless~year, r4[r4$state=="Wyoming",], ylim= myylim, type="l", col=mycol[1], lwd=2, lty=6, axes=
with( r4[r4$state=="North Dakota",], lines(x = year, y = smokeless, col = "red", lwd=2, lty=6))
text(1998, seq(0,18,2), paste("-",seq(0,18,2), "-"), cex=0.95)
text(seq(1999,2015,2),0, seq(1999,2015,2), cex=0.85, pos=1)
text(2016, seq(0,18,2), paste("-", seq(0,18,2), "-"), cex=0.95)
text(2007,22,"YRBS - % Current smokeless tobacco use", font=2, cex=1.25 )
```



Treemaps

```
library(treemap)
#windowsFonts(A=windowsFont("Georgia"),B=windowsFont("Myriad Pro"))

#dev.off()
par(mar=rep(1,4), xpd=T, las=1, family="sans")
r<- read.csv("F:/Dropbox/2017 Spring/Skill Lab R/Data/NYC_HearingArguments.csv", stringsAsFactors = F)

treemap(r, index=c("stance", "argue"), vSize="num",
        type="color",vColor="mycol", border.col="white",
        fontfamily.labels="sans", fontsize.labels=c(25,15),fontcolor.labels="white",
        align.labels= list(c("center", "top"),c("center", "center")),bg.labels=0,
        title="")

## Warning in `[.data.table`(dtfDT, , `:=`("c", fact), with = FALSE):
## with=FALSE ignored, it isn't needed when using :=. See ?':=' for examples.
```

