hpv_imm_coverage.R

joebrew

Mon May 18 17:07:18 2015

```r
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
##
## The following object is masked from 'package:stats':
##
##     filter
##
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(RColorBrewer)

wd <- '~/Documents/fdoh/public/hpv'
setwd(wd)

#####
# Census data for county-specific population
#####
pop <- read.csv("florida_details.csv", skip = 1)
names(pop) <- gsub("[.]", "", names(pop))
pop <- tbl_df(pop)

pop$AreaName <- as.character(toupper(sub(" County", "", pop$AreaName)))
state <- pop %>%
  filter(RaceEthnicity == "Total" ) %>%
  select(AreaName, TotalUnder1Year:Total110YearsandOlder)
names(state)[c(2:3,
               102:104)] <- c("Total0Years", "Total1Years",
                                "Total100-104Years", "Total105-109Years", "Total110+Years")
rm(pop)


# Assign vector of region 3, florida, etc.
region3 <- toupper(c("Alachua",
                     "Baker",
                     "Bradford",
                     "Clay",
                     "Duval",
                     "Flagler",
                     "Gilchrist",
                     "Levy",
                     "Marion",
                     "Nassau",
```

```r
                        "Putnam",
                        "St. Johns",
                        "Union"))

# In state, assign a vector of region
# As before, each region is EXCLUSIVE (ie, Alachua is NOT part of region 3 is NOT part of florida)
state$region <- NA
state$region <- ifelse(state$AreaName == "ALACHUA", "ALACHUA",
                       ifelse(state$AreaName %in% region3, "REGION3",
                              "FLORIDA"))
rm(region3)

# Remove the aggregate florida county row
state <- state[which(state$AreaName != "FLORIDA"),]

# Alacua denom girls 13 to 17 years old
denom <- sum(state[which(state$AreaName == 'ALACHUA'),paste0('Total', 13:17, 'Years')]) / 2

#####
# CDC DATA HERE: http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6329a3.htm
#####
cdc <- data.frame(year = 2007:2013,
                  girls1 = c(25.1, 37.2, 44.3, 48.7, 53.0, 53.8, 57.3),
                  girls1lwr = c(22.3, 35.2, 42.4, 46.9, 51.4, 52, 55.4),
                  girls1upr = c(28.1, 39.3, 46.1, 50.5, 54.7, 55.7, 59.2),
                  girls2 = c(16.9, 28.3, 35.8, 40.7, 43.9, 43.4, 47.7),
                  girls2lwr = c(14.6, 26.4, 34.1, 38.9, 42.3, 41.5, 45.7),
                  girls2upr = c(19.6, 30.3, 37.6, 42.5, 45.6, 45.2, 49.6),
                  girls3 = c(5.9, 17.9, 26.7, 32.0, 34.8, 33.4, 37.6),
                  girls3lwr = c(4.4, 16.3, 25.2, 30.3, 33.2, 31.7, 35.7),
                  girls3upr = c(7.8, 19.6, 28.3, 33.6, 36.4, 35.2, 39.6),
                  boys1 = c(0, 0, 0, 0, 8.3, 20.8, 34.6),
                  boys1lwr = c(0, 0, 0, 0, 7.4, 19.4, 32.7),
                  boys1upr = c(0, 0, 0, 0, 9.3, 22.4, 36.5),
                  boys2 = c(0, 0, 0, 0, 3.8, 12.7, 23.5),
                  boys2lwr = c(0, 0, 0, 0, 3.2, 11.5, 21.8),
                  boys2upr = c(0, 0, 0, 0, 4.5, 14, 25.3),
                  boys3 = c(0, 0, 0, 0, 1.3, 6.8, 13.9),
                  boys3lwr = c(0, 0, 0, 0, 1, 5.9, 12.5),
                  boys3upr = c(0, 0, 0, 0, 1.7, 7.8, 15.3))

# MAKE AN ALACHUA df INTERPOLATING FROM CDC
alachua <- cdc
for (j in 2:ncol(cdc)){
  column <- cdc[,j]
  column <- column * denom / 100
  alachua[,j] <- round(column)
}

# Make charts
library(Hmisc)
```

```
## Loading required package: grid
```

```
## Loading required package: lattice
## Loading required package: survival
## Loading required package: Formula
## Loading required package: ggplot2
##
## Attaching package: 'Hmisc'
##
## The following objects are masked from 'package:dplyr':
##
##     combine, src, summarize
##
## The following objects are masked from 'package:base':
##
##     format.pval, round.POSIXt, trunc.POSIXt, units
```

```r
# ALACHUA GIRLS
cols <- brewer.pal(5, 'Blues')[3:5]
columns <- c('girls1', 'girls2', 'girls3')
sub_data <- alachua[,columns]
obj <- t(as.matrix(sub_data))
girls_max <- max(obj)

lwrs <- paste0(columns, 'lwr')
uprs <- paste0(columns, 'upr')
lwrs <- t(as.matrix(alachua[,lwrs]))
uprs <- t(as.matrix(alachua[,uprs]))

bp <- barplot(obj,
              beside = TRUE,
              ylim = c(0, girls_max * 1.2),
              col = cols,
              border = NA,
              las = 1,
              names.arg = alachua$year,
              xlab = 'Year',
              ylab = 'Residents',
              main = 'HPV vaccine coverage among 13-17 Alachua resident females (estimated)',
              cex.main = 0.85)
legend('topleft',
       fill = cols,
       border = NA,
       legend = paste(1:3, 'dose'))
legend('topright',
       lty = 1,
       col = adjustcolor('darkred', alpha.f = 0.6),
       legend = '95% confidence interval',
       bty = 'n',
       cex = 0.8)

errbar(x = as.numeric(bp),
       y = as.numeric(obj),
       yplus = as.numeric(uprs),
       yminus = as.numeric(lwrs),
       pch = NA,
```
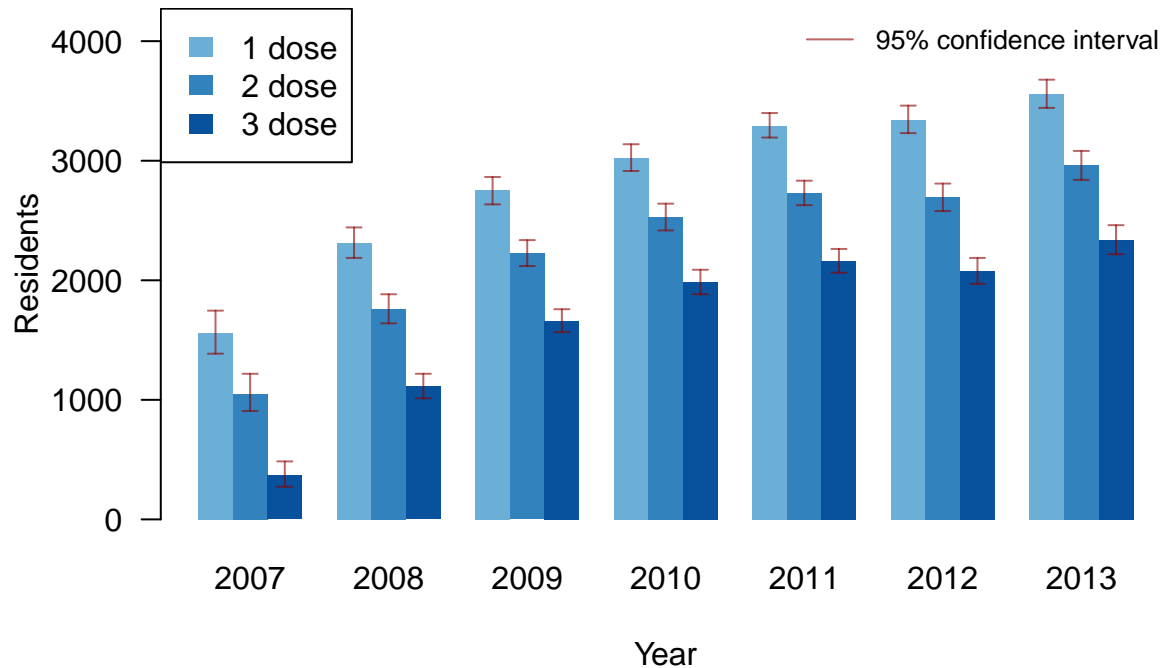
```
        add = TRUE,
        errbar.col = adjustcolor('darkred', alpha.f = 0.6))
```

HPV vaccine coverage among 13–17 Alachua resident females (estimated)


```
# ALACHUA BOYS
cols <- brewer.pal(5, 'Greens')[3:5]
columns <- c('boys1', 'boys2', 'boys3')
sub_data <- alachua[,columns]
obj <- t(as.matrix(sub_data))

lwrs <- paste0(columns, 'lwr')
uprs <- paste0(columns, 'upr')
lwrs <- t(as.matrix(alachua[,lwrs]))
uprs <- t(as.matrix(alachua[,uprs]))


bp <- barplot(obj,
              beside = TRUE,
              ylim = c(0, girls_max * 1.2),
              col = cols,
              border = NA,
              las = 1,
              names.arg = alachua$year,
              xlab = 'Year',
              ylab = 'Residents',
              main = 'HPV vaccine coverage among 13-17 Alachua resident males (estimated)',
              cex.main = 0.85)
legend('topleft',
       fill = cols,
       border = NA,
```

```
            legend = paste(1:3, 'dose'))

legend('topright',
       lty = 1,
       col = adjustcolor('darkred', alpha.f = 0.6),
       legend = '95% confidence interval',
       bty = 'n',
       cex = 0.8)

errbar(x = as.numeric(bp),
       y = as.numeric(obj),
       yplus = as.numeric(uprs),
       yminus = as.numeric(lwrs),
       pch = NA,
       add = TRUE,
       errbar.col = adjustcolor('darkred', alpha.f = 0.6))
```
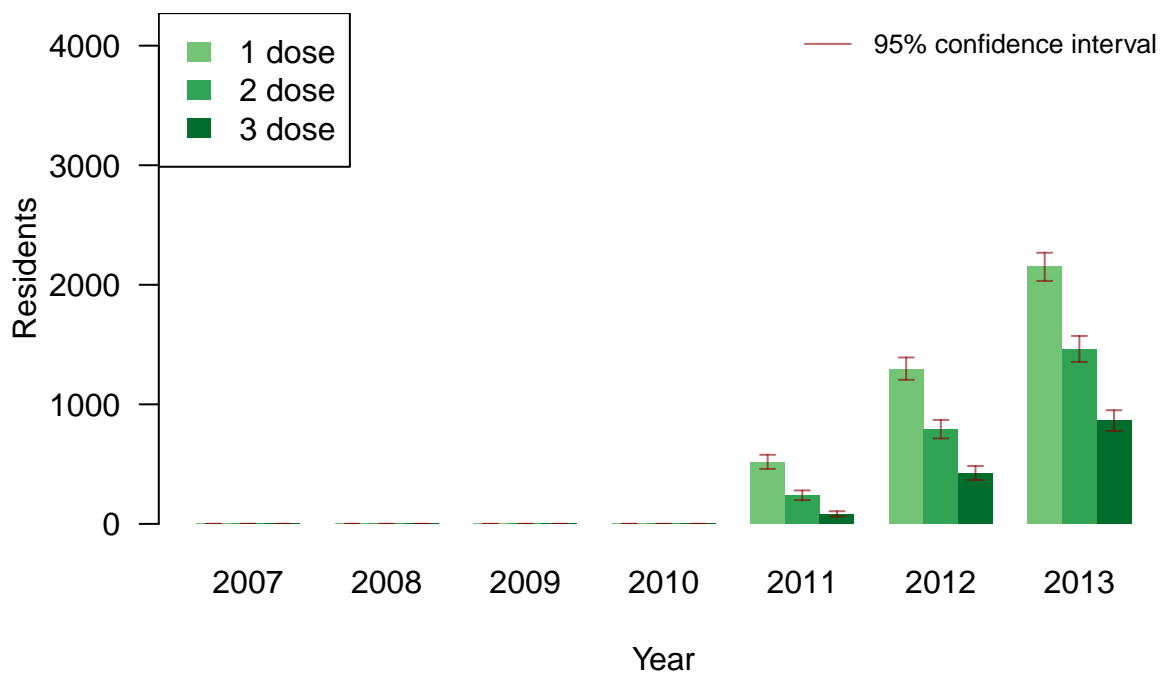
HPV vaccine coverage among 13–17 Alachua resident males (estimated)


```
# PROPORTIONAL GIRLS
cols <- brewer.pal(5, 'Blues')[3:5]
columns <- c('girls1', 'girls2', 'girls3')
sub_data <- cdc[,columns]
obj <- t(as.matrix(sub_data))
girls_max <- max(obj)

lwrs <- paste0(columns, 'lwr')
uprs <- paste0(columns, 'upr')
lwrs <- t(as.matrix(cdc[,lwrs]))
uprs <- t(as.matrix(cdc[,uprs]))
```

```r
bp <- barplot(obj,
              beside = TRUE,
              ylim = c(0, girls_max * 1.2),
              col = cols,
              border = NA,
              las = 1,
              names.arg = cdc$year,
              xlab = 'Year',
              ylab = 'Percent',
              main = 'HPV vaccine coverage among 13-17 cdc resident females (estimated)',
              cex.main = 0.85)
legend('topleft',
       fill = cols,
       border = NA,
       legend = paste(1:3, 'dose'))

legend('topright',
       lty = 1,
       col = adjustcolor('darkred', alpha.f = 0.6),
       legend = '95% confidence interval',
       bty = 'n',
       cex = 0.8)

errbar(x = as.numeric(bp),
       y = as.numeric(obj),
       yplus = as.numeric(uprs),
       yminus = as.numeric(lwrs),
       pch = NA,
       add = TRUE,
       errbar.col = adjustcolor('darkred', alpha.f = 0.6))
```
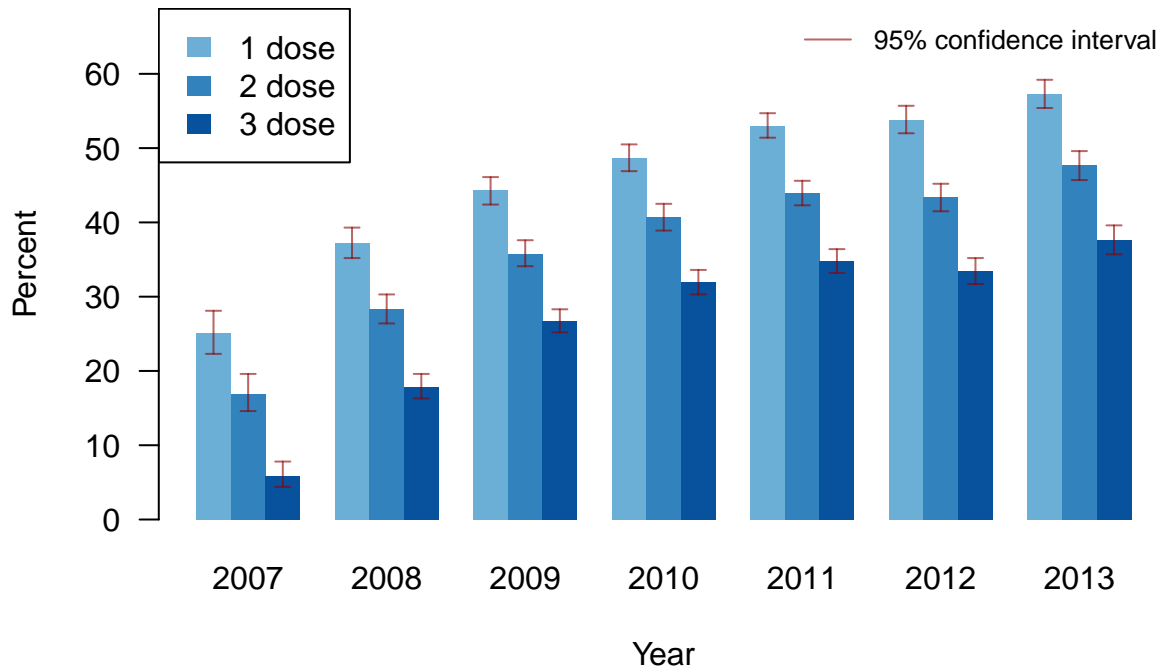
HPV vaccine coverage among 13–17 cdc resident females (estimated)


```
# PROPORTIONAL BOYS
cols <- brewer.pal(5, 'Greens')[3:5]
columns <- c('boys1', 'boys2', 'boys3')
sub_data <- cdc[,columns]
obj <- t(as.matrix(sub_data))


# CURRENT

lwrs <- paste0(columns, 'lwr')
uprs <- paste0(columns, 'upr')
lwrs <- t(as.matrix(cdc[,lwrs]))
uprs <- t(as.matrix(cdc[,uprs]))

bp <- barplot(obj,
              beside = TRUE,
              ylim = c(0, girls_max * 1.2),
              col = cols,
              border = NA,
              las = 1,
              names.arg = cdc$year,
              xlab = 'Year',
              ylab = 'Percent',
              main = 'HPV vaccine coverage among 13-17 cdc resident males (estimated)',
              cex.main = 0.85)
legend('topleft',
       fill = cols,
       border = NA,
       legend = paste(1:3, 'dose'))
```
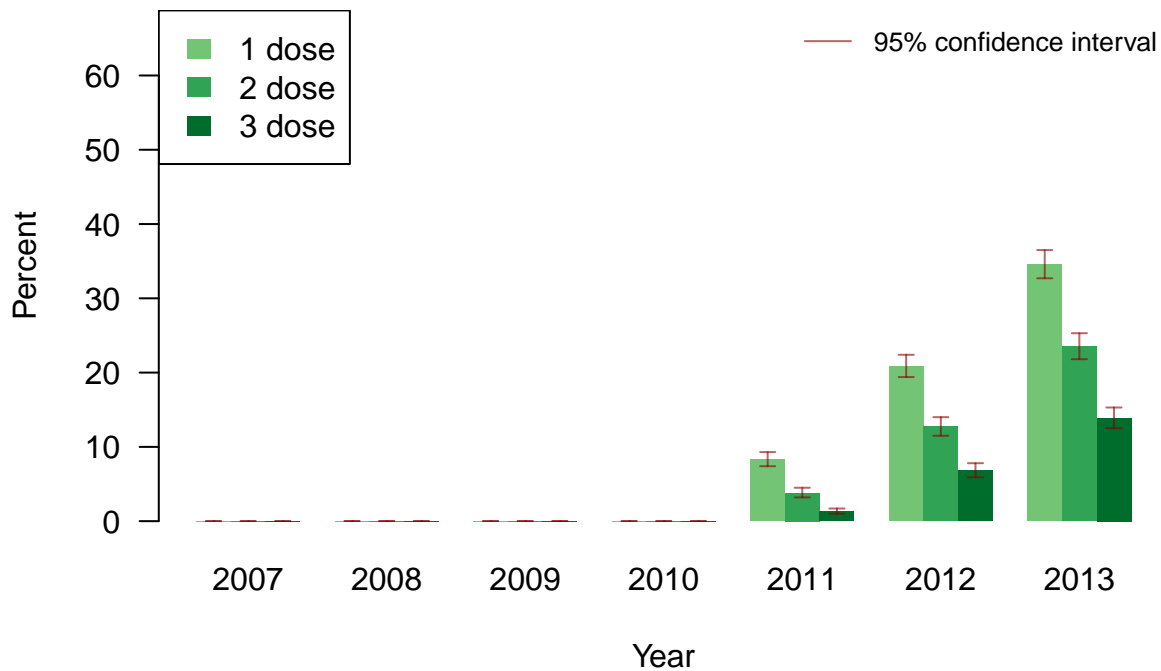
```
legend('topright',
       lty = 1,
       col = adjustcolor('darkred', alpha.f = 0.6),
       legend = '95% confidence interval',
       bty = 'n',
       cex = 0.8)

errbar(x = as.numeric(bp),
       y = as.numeric(obj),
       yplus = as.numeric(uprs),
       yminus = as.numeric(lwrs),
       pch = NA,
       add = TRUE,
       errbar.col = adjustcolor('darkred', alpha.f = 0.6))
```

HPV vaccine coverage among 13–17 cdc resident males (estimated)


```
#####
# READ IN FLORIDA SHOTS HPV DATA
#####
flshots <- read.csv('florida_shots_hpv.csv')

# GET TOTAL STATE POPULATION
state$region <- NULL
pop <- data.frame(AreaName = state$AreaName,
                  pop = apply(state[,-1], 1, sum ))

# JOIN TOGETHER POPULATION AND HPV IMMUNIZATION DATA
pop <- left_join(pop, flshots)
```

```
## Joining by: "AreaName"
```

```
# GET PERCENTAGE
pop$p <- pop$total / pop$pop * 100

# ORDER BY PERCENTAGE
pop <- pop[order(pop$p),]

# Barplot
par(mar = c(7,4,4,2))
cols <- colorRampPalette(brewer.pal(9, 'Spectral'))(nrow(pop))
bp <- barplot(pop$p,
              names.arg = pop$AreaName,
              col = cols,
              border = NA,
              cex.names = 0.5,
              las = 3,
              ylab = 'Percentage',
              main = 'HPV vaccination coverage\n(Doses / Total Population)')
```
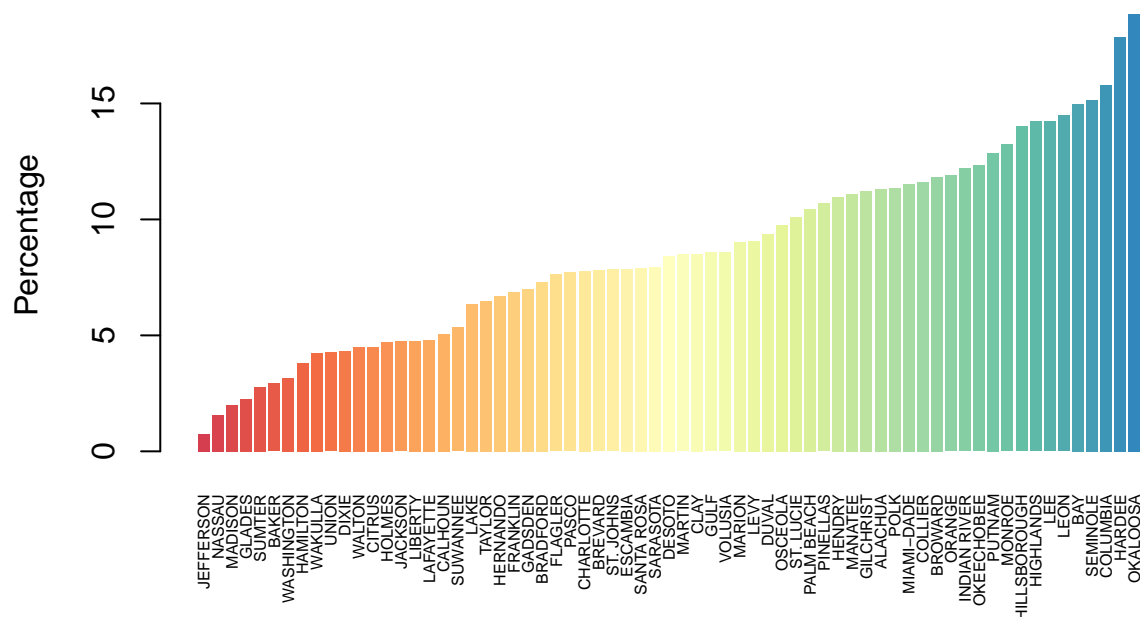

HPV vaccination coverage
(Doses / Total Population)

```
# Histogram
h <- hist(pop$p, breaks = 20, plot = FALSE)
cols <- colorRampPalette(brewer.pal(9, 'Spectral'))(length(h$mids))
plot(h,
     col = cols,
     border = NA,
     xlab = 'Percentage',
     main = 'Distribution of HPV vaccination coverage\n(Doses / Total Population)')
val <- pop$p[which(pop$AreaName == 'ALACHUA')]
abline(v = val,
```
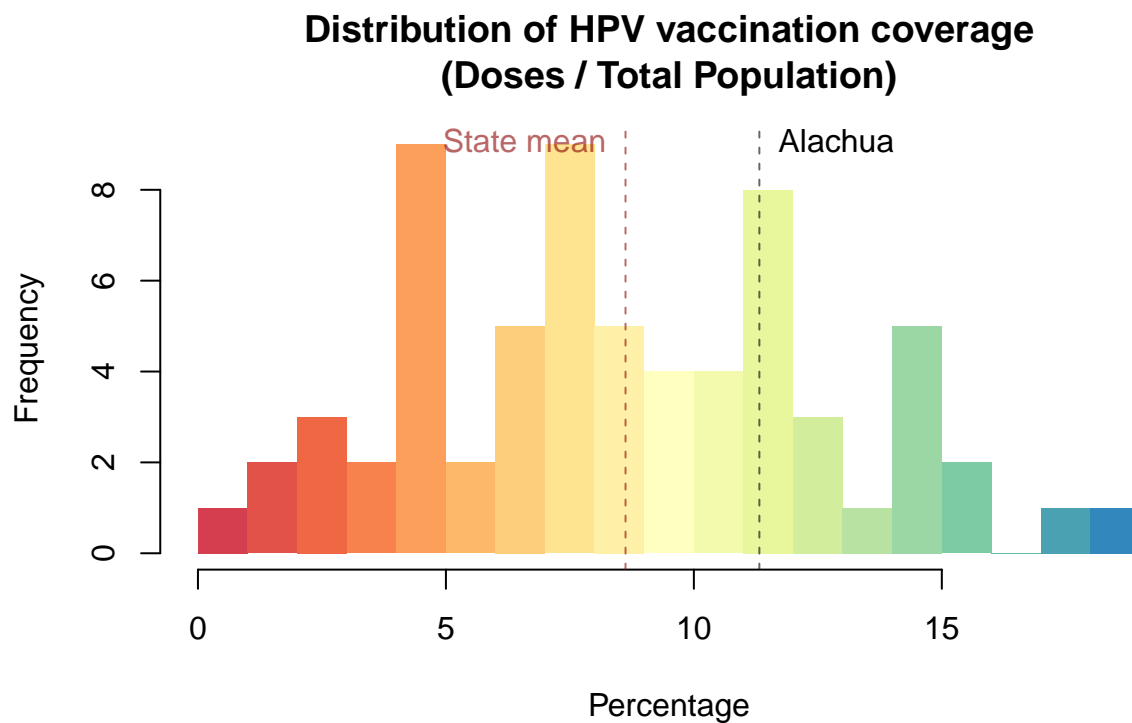
```
          lty = 2,
          col = adjustcolor('black', alpha.f = 0.6))
text(x = val,
     y = round(0.95 * max(h$counts)),
     pos = 4,
     labels = 'Alachua')

val <- mean(pop$p)
abline(v = val,
       lty = 2,
       col = adjustcolor('darkred', alpha.f = 0.6))
text(x = val,
     y = round(0.95 * max(h$counts)),
     pos = 2,
     col = adjustcolor('darkred', alpha.f = 0.6),
     labels = 'State mean')
```
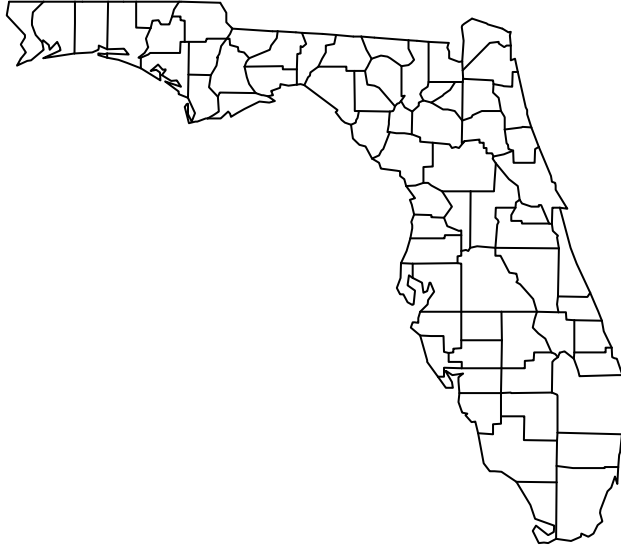


```
#####
# MAP
#####
library(maps)
fl <- map('county', 'florida')
```

```
fl$names <- toupper(gsub('florida,', '', fl$names))
fl$names[46:47] <- 'OKALOOSA'
fl$names[56:57] <- c('ST. JOHNS', 'ST. LUCIE')
fl$names[14] <- 'DESOTO'

fl$p <- NA
for (i in 1:length(fl$names)){
  fl$p[i] <-
    pop$p[which(pop$AreaName == fl$names[i])]
}

fl$p2 <- round(fl$p * 100)
cols <- colorRampPalette(brewer.pal(9, 'Spectral'))(2000)
fl$col <- cols[fl$p2]
map('county',
    'fl',
    fill = TRUE,
    col = fl$col,
    border = adjustcolor('black', alpha.f = 0.4))
title(main = 'HPV vaccination coverage\n(Doses / Total Population)')

legend_seq <- seq(200, 2000, length = 10)
legend('bottomleft',
       fill = cols[legend_seq],
       legend = paste0(legend_seq/100, '%'))
map.text('county', 'fl',
         labels = as.character(round(fl$p, digits = 1)),
         add = TRUE,
         col = adjustcolor('black', alpha.f = 0.6),
         cex = 0.35)
```

HPV vaccination coverage
(Doses / Total Population)

Legend:
- 2%
- 4%
- 6%
- 8%
- 10%
- 12%
- 14%
- 16%
- 18%
- 20%