

# Figures for HIV Profile

*James Hunter*

*April 1, 2015*

- CD4+ T cells X-axis bases
- Create x-axis labels with n
- bring in n's for each year from luciano\_NAIVE spreadsheet

```
cdbasen <- c(2196, 21217, 17017, 20366, 24753, 23480, 25354, 25347, 26507)
cdyears <- seq(2001, 2009)
tickcd <- rep(NA, length(cdyears))
for (i in 1:length(cdyears)) {
  tickcd[i] <- paste(cdyears[i], "\n(", cdbasen[i], ")", sep = "")
}
```

- vl initial base n for x-axis

```
cvbasen <- c(855, 10079, 12997, 18516, 22671, 19872, 20684, 24690, 25613)
cvyears <- seq(2001, 2009)
tickcv <- rep(NA, length(cvyears))
for (i in 1:length(cvyears)) {
  tickcv[i] <- paste(cvyears[i], "\n(", cvbasen[i], ")", sep = "")
}
```

- vl base n for x-axis on vl < 400 graph

```
cvyear2 <- seq(2001, 2011)
load("cvn.RData")
tickcv2 <- rep(NA, length(cvyear2))
for (i in 1:length(cvyear2)) {
  tickcv2[i] <- paste(cvyear2[i], "\n(", cvn[i], ")", sep = "")
}
```

- resistance base n for x-axis on class resistance graph

```
load("resist.RData")
reyear <- seq(2003, 2009)
ren <- resist$n
tickre <- rep(NA, length(reyear))
for (i in 1:length(reyear)) {
  tickre[i] <- paste(reyear[i], "\n(", ren[i], ")", sep = "")
}
```

**Figure 1. CD4+ T cell counts among antiretroviral naïve individuals at first presentation from 2001 to 2009**

```
## Stacked Bar graph of CD4 levels
```

```
require(ggplot2)
```

```
## Loading required package: ggplot2
```

```
## Color version of CD4 stacked bar graph
```

```
cd4barc <- ggplot(cd4melt, aes(x = as.character(year), y = val, fill = var)) +  
  geom_bar(stat = "identity")
```

```
cd4barc <- cd4barc + theme_classic()
```

```
cd4barc <- cd4barc + scale_fill_brewer(palette="Set1",  
                                     labels = c("< 200",  
                                                "200-349",  
                                                "> 349"))
```

```
cd4barc <- cd4barc + ggtitle("CD4+ T Cell Count 2001 - 2009")
```

```
cd4barc <- cd4barc + theme(plot.title = element_text(size = rel(1.5),  
                                                    lineheight = .9,  
                                                    face = "bold"))
```

```
cd4barc <- cd4barc + ylab("Percent") + xlab("Year (n)")
```

```
cd4barc <- cd4barc + scale_x_discrete(breaks = seq(2001, 2009),  
                                     labels = tickcd)
```

```
cd4barc <- cd4barc + guides(fill = guide_legend(reverse = TRUE))
```

```
cd4barc <- cd4barc + labs(fill = "Count")
```

```
#cd4barc <- cd4barc + scale_fill_grey(labels = c("< 200", "200-349", "> 349"))  
cd4barc
```

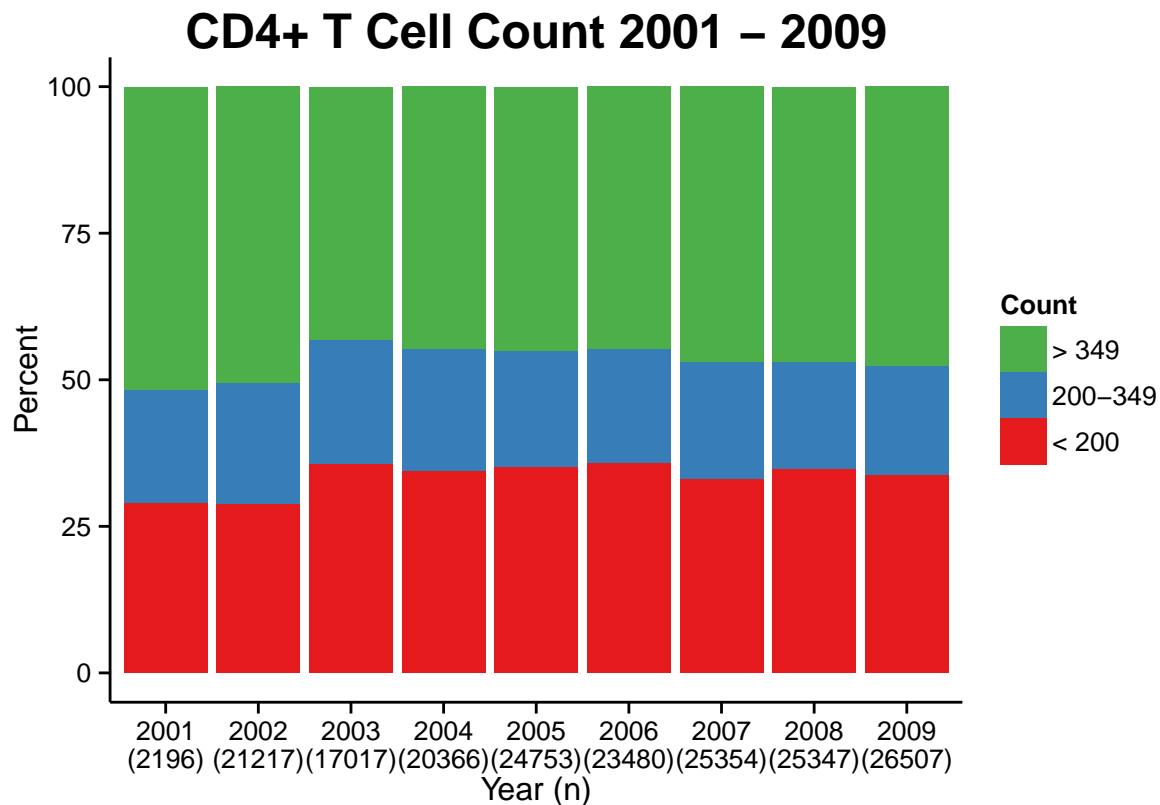


Figure 2A. Mean CD4+ T cell counts by gender according to the first measurement for each patient from 2001 to 2009.

```
## CD4 gender plot color version
load("cd4means.RData")
plot(cd4means$all ~ cd4means$year, type = "l",
     lwd = 2, xaxt = "n", ylim = c(300, 460), yaxp = c(300, 460, 8),
     main = "Mean CD4+ T Cell Count by Gender\n2001 - 2009",
     xlab = "Year",
     ylab = "Cell Count per mm3")
points(cd4means$all ~ cd4means$year, pch = 20)
lines(cd4means$men ~ cd4means$year,
     lwd = 2, lty = 2, col = "blue")
points(cd4means$men ~ cd4means$year,
     pch = 18, col = "blue")
lines(cd4means$women ~ cd4means$year,
     lwd = 2, lty = 3, col = "darkgreen")
points(cd4means$women ~ cd4means$year,
     pch = 4, col = "darkgreen")
grid(col = "darkgrey")
axis(1, at = cd4reg$year, labels = tickcd)
legend("top",
     legend = c("All patients", "Men", "Women"),
     lwd = 2,
     col = c("black", "blue", "darkgreen"),
     lty = seq(1,3), cex = 0.8,
     ncol = 3)
```

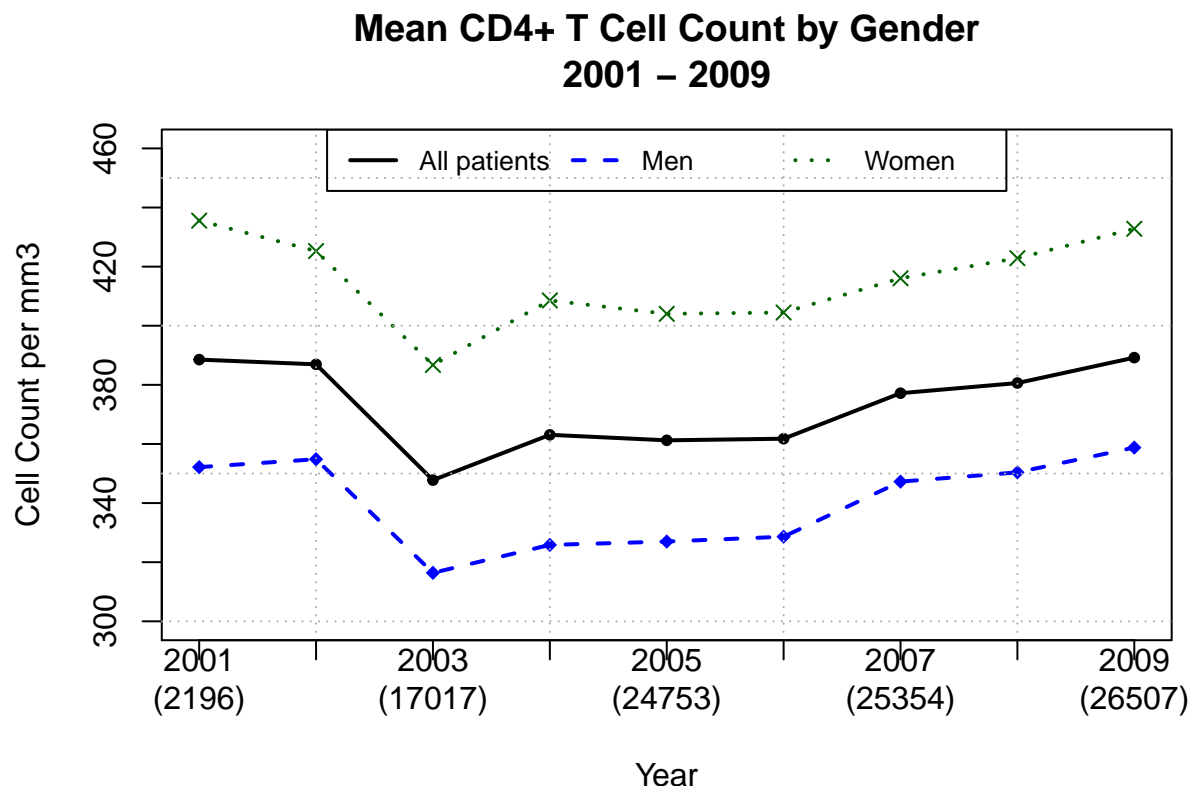


Figure 2B. Mean CD4+ T cell counts by age according to the first measurement for each patient from 2001 to 2009.

```
plot(cd4means$age1321 ~ cd4means$year, type = "l",
     lwd = 2, xaxt = "n", ylim = c(280, 580), yaxp = c(280, 580, 20),
     main = "Mean CD4+ T Cell Count by Age Group\n2001 - 2009",
     xlab = "Year",
     ylab = "Cell Count per mm3")
points(cd4means$age1321 ~ cd4means$year, pch = 20)
lines(cd4means$age2250 ~ cd4means$year,
      lwd = 2, lty = 2, col = "blue")
points(cd4means$age2250 ~ cd4means$year, pch = 18, col = "blue")
lines(cd4means$age50p ~ cd4means$year,
      lwd = 2, lty = 3, col = "darkgreen")
points(cd4means$age50p ~ cd4means$year, pch = 4, col = "darkgreen")
grid(col = "darkgrey")
axis(1, at = cd4reg$year, labels = tickcd)
legend("top",
      legend = c("13-21", "22-50", ">50"),
      lwd = 2,
      col = c("black", "blue", "darkgreen"),
      lty = seq(1,3), cex = 0.8,
      ncol = 3)
```

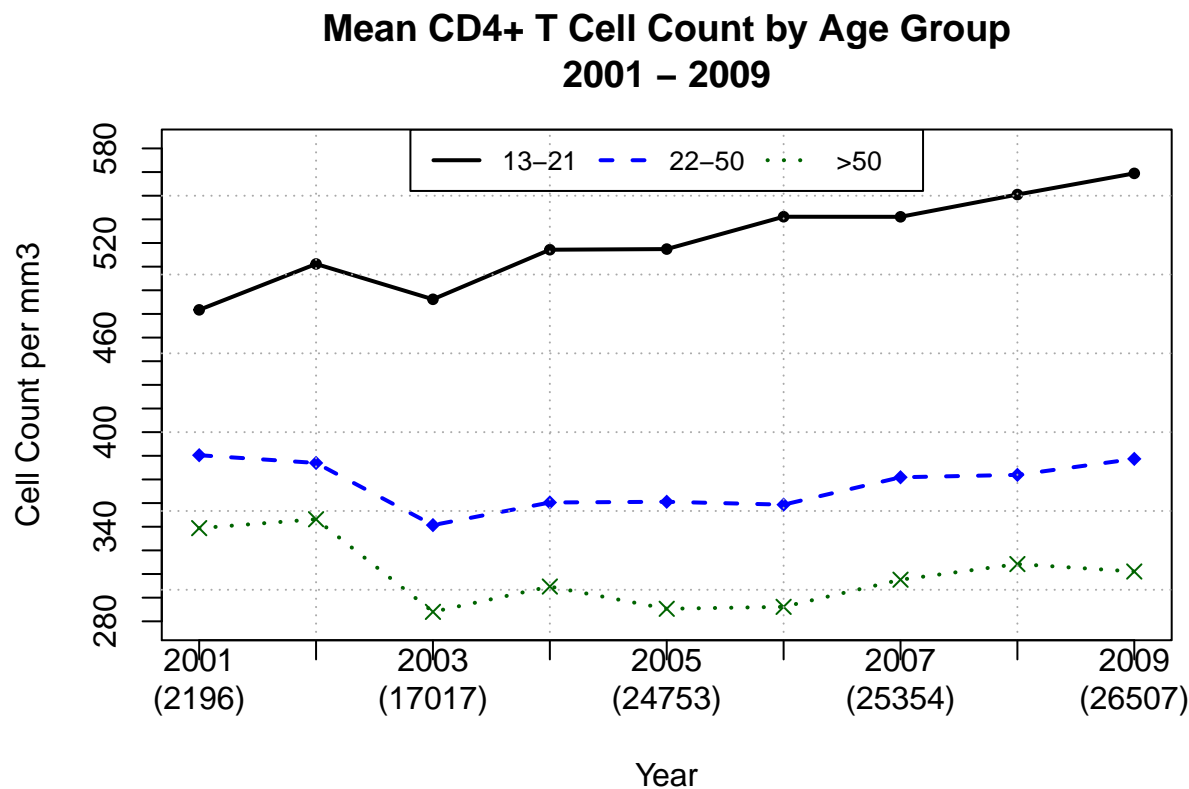


Figure 2C. Mean CD4+ T cell counts by region according to the first measurement for each patient from 2001 to 2009.

```
plot(cd4reg$se ~ cd4reg$year, type = "l",
     lwd = 2, xaxt = "n", ylim = c(280, 420),
     yaxp = c(280, 420, 7),
     main = "Mean CD4+ T Cell Count by Region\n2001 - 2009",
     xlab = "Year",
     ylab = "Cell Count per mm3")
lines(cd4reg$s ~ cd4reg$year, lwd = 2, lty = 2, col = "blue")
lines(cd4reg$co ~ cd4reg$year, lwd = 2, lty = 3, col = "darkgreen")
lines(cd4reg$ne ~ cd4reg$year, lwd = 2, lty = 4, col = "darkred")
lines(cd4reg$n ~ cd4reg$year, lwd = 2, lty = 5, col = "orange")
grid(col = "darkgrey")
axis(1, at = cd4reg$year, labels = tickcd)
legend("top",
      legend = c("South East", "South", "Central West", "Northeast", "North"),
      lwd = 2,
      col = c("black", "blue", "darkgreen", "darkred", "orange"),
      lty = seq(1,5), cex = 0.8,
      ncol = 5)
```

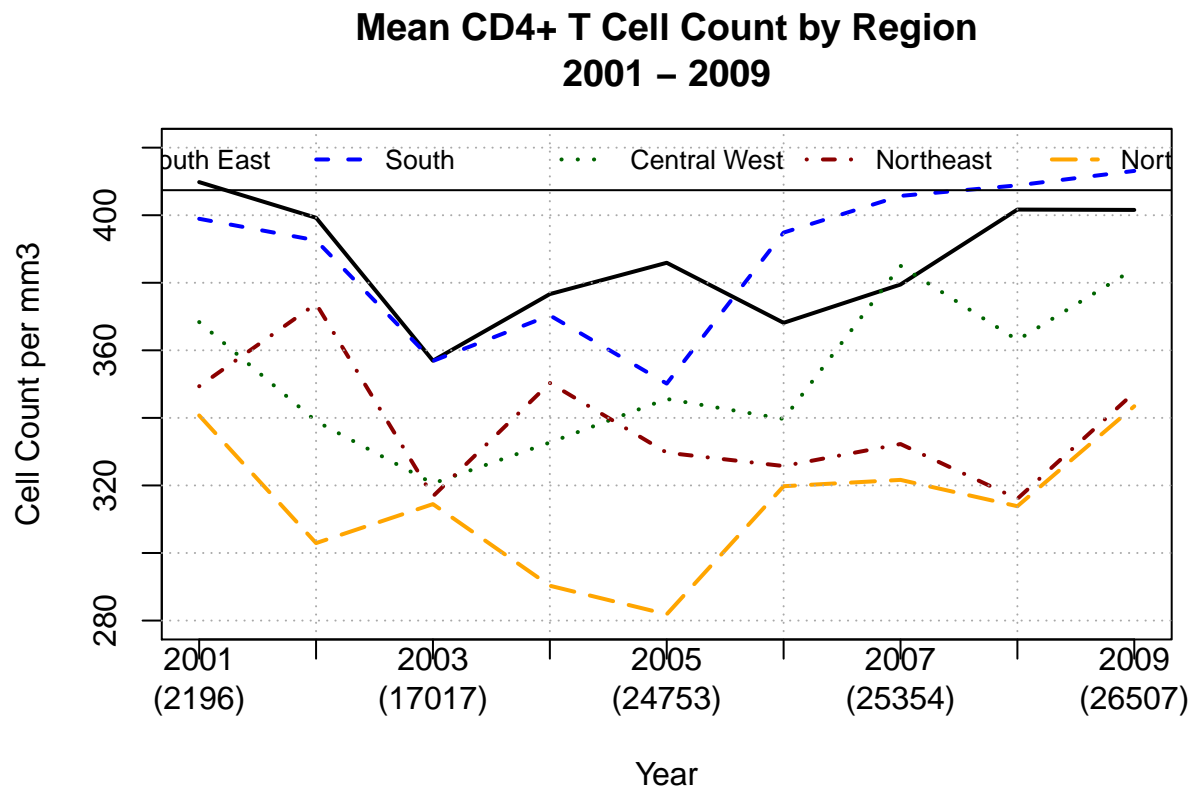


Figure 3. Mean Basal viral load by age and gender according to the first measurement for each patient from 2001 to 2011.

```
plot(log10(cvmeans$age1321) ~ cvmeans$year, type = "l",
     lwd = 2, xaxt = "n", ylim = c(4.5, 5.5),
     main = "Mean Basal Viral Load by Age Group/Gender\n2001 - 2009",
     xlab = "Year",
     ylab = "Log10 Copies per ml of plasma ")
# points(log10(cvmeans$age50p) ~ cvmeans$year, pch = 20)
lines(log10(cvmeans$age2250) ~ cvmeans$year, lwd = 2, lty = 2,
      col = "blue")
# points(log10(cvmeans$age2250) ~ cvmeans$year, pch = 18)
lines(log10(cvmeans$age50p) ~ cd4means$year, lwd = 2, lty = 3,
      col = "darkgreen")
# points(log10(cvmeans$age1321) ~ cd4means$year, pch = 4)
lines(log10(cvmeans$men) ~ cvmeans$year, lwd = 2, lty = 4,
      col = "darkred")
lines(log10(cvmeans$women) ~ cvmeans$year, lwd = 2, lty = 5,
      col = "orange")
grid(col = "darkgrey")
axis(1, at = cvmeans$year, labels = tickcv)
legend("top",
      legend = c("13 - 21", "22 - 50", "50 +", "Men", "Women"),
      lwd = 2,
      lty = seq(1,5), cex = 0.8,
      col = c("black", "blue", "green", "darkred", "orange"),
      ncol = 5)
```

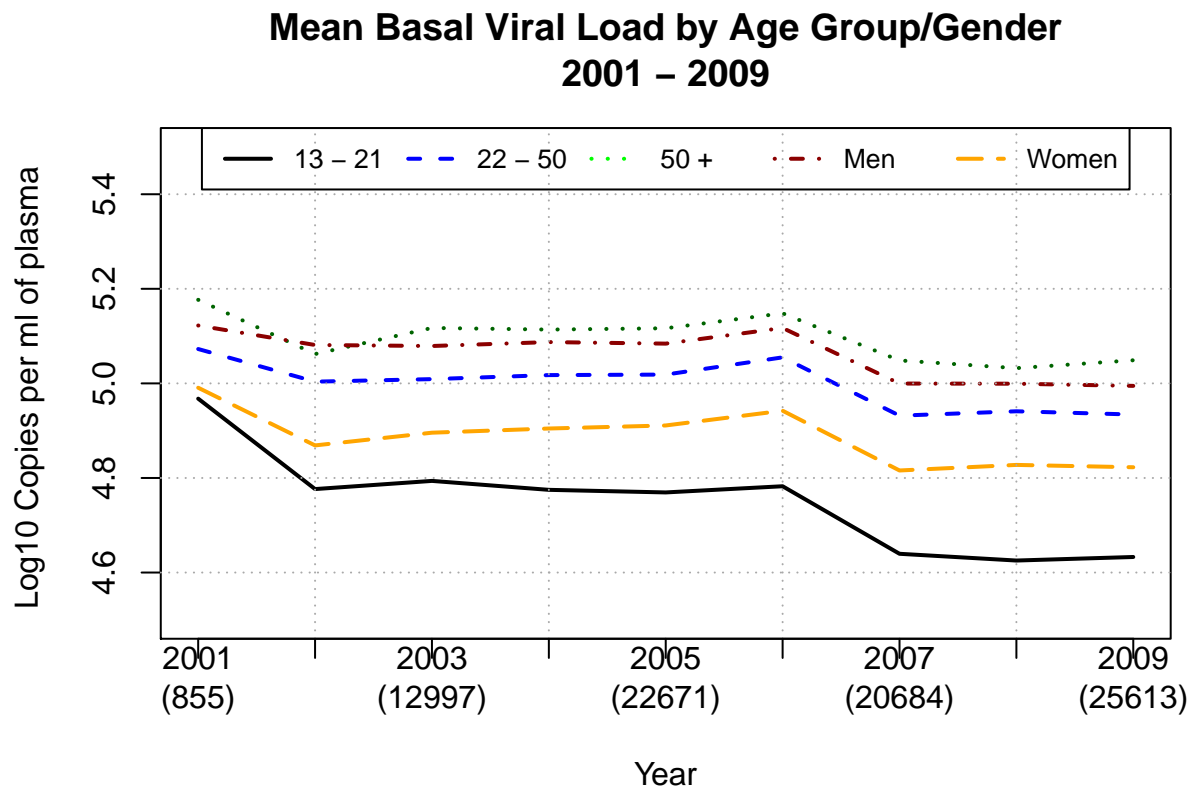


Figure 4. Number of treatments containing non-nucleoside reverse transcriptase inhibitors (NNRTIs), unboosted protease inhibitors (PIs) and ritonavir-boosted PI from 2004 to 2009

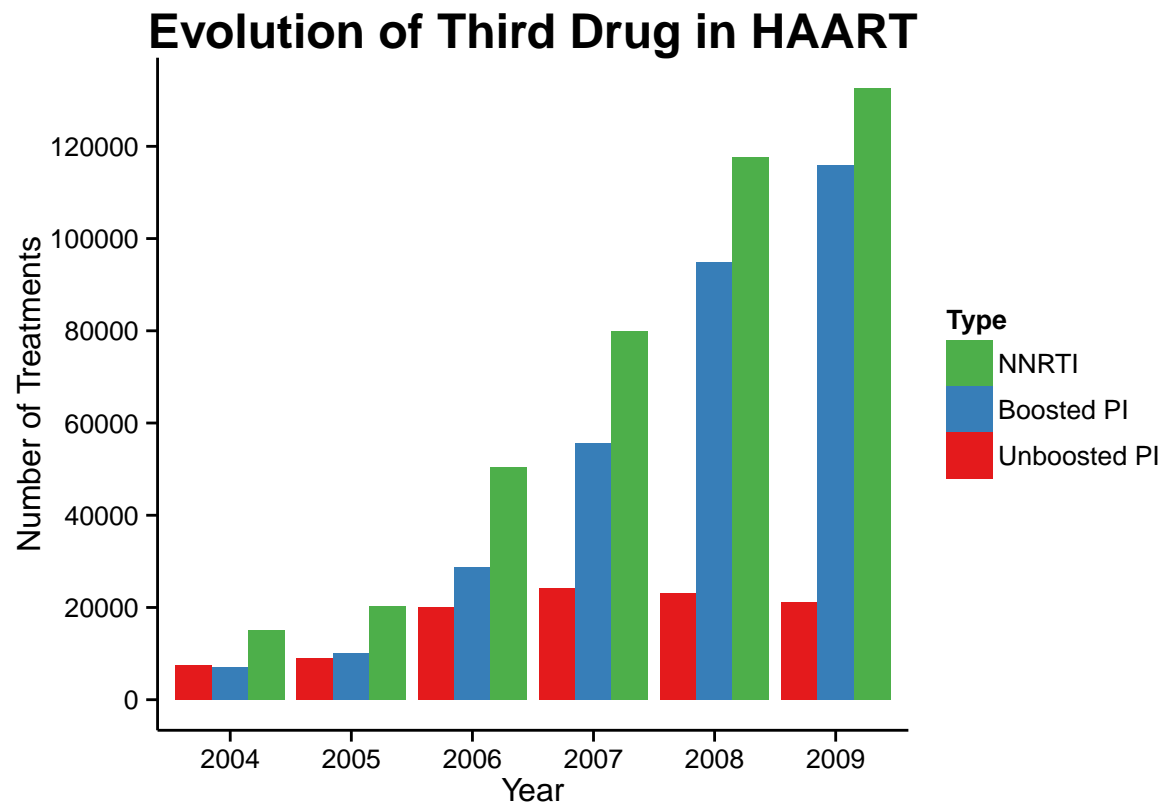
```
meds <- read.csv("meds.csv", header = TRUE)
require(reshape2)
```

```
## Loading required package: reshape2
```

```
medsmelt <- melt(meds, id.vars = "year",
  measure.vars = c("unboostpi", "boostpi", "nnrti"),
  variable.name = "type",
  value.name = "treat")
#medsyear <- as.character(seq(min(meds$year), max(meds$year)))
```

Prepare bar plot:

```
require(ggplot2)
medsg <- ggplot(medsmelt, aes(x = as.character(year),
  y = treat, fill = type)) + geom_bar(stat = "identity", position = "dodge")
medsg <- medsg + theme_classic()
medsg <- medsg + scale_fill_brewer(palette = "Set1", labels = c("Unboosted PI", "Boosted PI", "NNRTI"))
medsg <- medsg + scale_y_continuous(breaks =
  seq(0, 140000, by = 20000))
medsg <- medsg + ggtitle("Evolution of Third Drug in HAART")
medsg <- medsg + theme(plot.title = element_text(size = rel(1.5),
  lineheight = .9,
  face = "bold"))
medsg <- medsg + ylab("Number of Treatments") + xlab("Year")
medsg <- medsg + guides(fill = guide_legend(reverse = TRUE))
medsg <- medsg + labs(fill = "Type")
medsg
```



Revised Figure 4 to show y axis in Percentages.

```

meds <- read.csv("meds.csv", header = TRUE)
require(reshape2)
medsmelt <- melt(meds, id.vars = "year",
  measure.vars = c("unboostpi", "boostpi", "nnrti"),
  variable.name = "type",
  value.name = "treat")

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

yeartot <- summarise(group_by(medsmelt, year), treatsum = sum(treat))
for (i in 1:nrow(medsmelt)) {
  medsmelt$treatpct[i] <- 100 * medsmelt$treat[i] /
    yeartot$treatsum[yeartot$year == medsmelt$year[i]]
}
#medsyear <- as.character(seq(min(meds$year), max(meds$year)))

```



Prepare bar plot:

```
require(ggplot2)
medsg <- ggplot(medsmelt, aes(x = as.character(year),
                             y = treatpct, fill = type)) + geom_bar(stat = "identity", position = "dodge")
medsg <- medsg + theme_classic()
medsg <- medsg + scale_fill_brewer(palette = "Set1", labels = c("Unboosted PI", "Boosted PI", "NNRTI"))
#medsg <- medsg + scale_y_continuous(breaks =
#                                   seq(0, 140000, by = 20000))
medsg <- medsg + ggtitle("Evolution of Third Drug in HAART")
medsg <- medsg + theme(plot.title = element_text(size = rel(1.5),
                                                  lineheight = .9,
                                                  face = "bold"))
medsg <- medsg + ylab("Percent - Number of Treatments") + xlab("Year")
medsg <- medsg + guides(fill = guide_legend(reverse = TRUE))
medsg <- medsg + labs(fill = "Type")
medsg
```

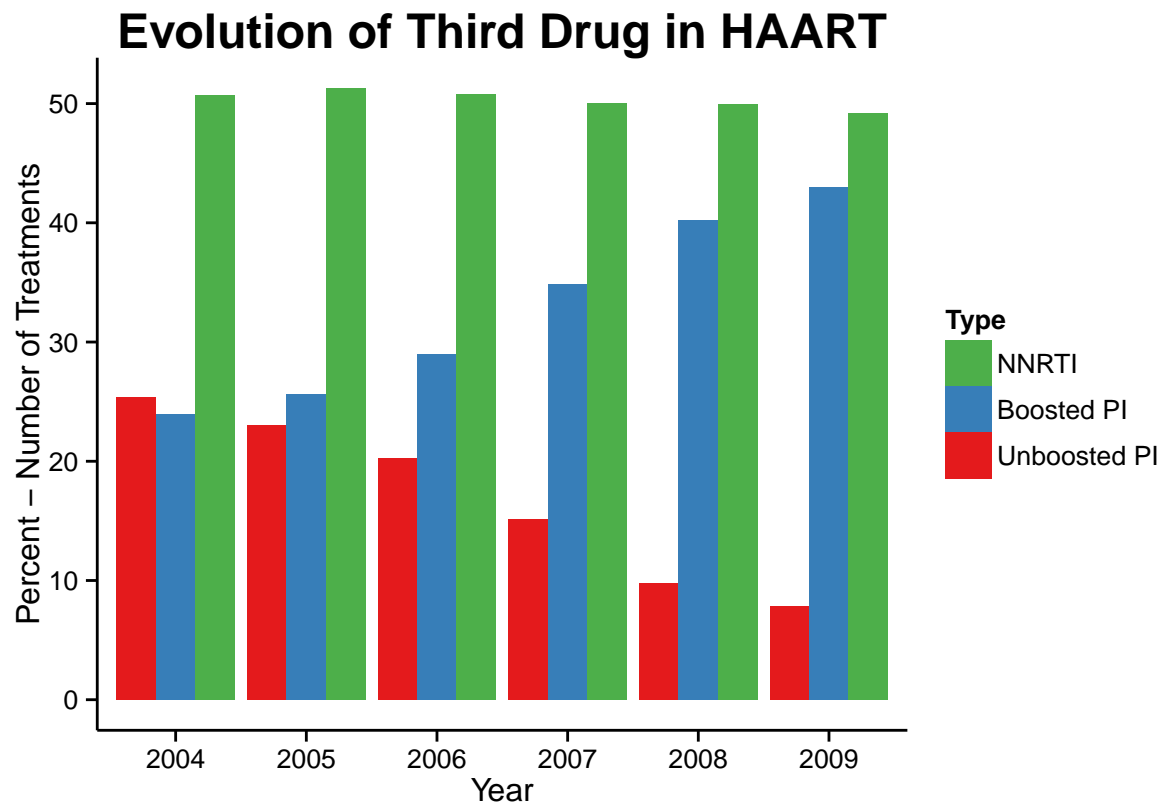


Figure 5. Percentage of viral loads below detection among all individuals on antiretroviral treatment by age from 2002 to 2011.

```
load("vlage.RData")
plot(vlage$pct[vlage$age == "13 - 21"] * 100 ~
     vlage$year[vlage$age == "13 - 21"],
     type = "l", lwd = 2, xaxt = "n", ylim = c(30, 100),
```

```

main = "Percentage of Patients by Age\nwith Viral Load < 400  2002 - 2011",
xlab = "Year",
ylab = "% of Patients with VL < 400/ncopies RNA/mm3 plasma")
lines(vlage$pct[vlage$age == "22 - 50"] * 100 ~
      vlage$year[vlage$age == "22 - 50"],
      lwd = 2, lty = 2, col = "blue")
lines(vlage$pct[vlage$age == "50 plus"] * 100 ~
      vlage$year[vlage$age == "50 plus"],
      lwd = 2, lty = 3, col = "darkgreen")
lines(locvall$pct * 100 ~ locvall$year,
      lwd = 2, lty = 4, col = "darkred")
grid(col = "darkgrey")
axis(1, at = cvyear2, labels = tickcv2)
legend("top",
      legend = c("13-21", "22-50", ">50", "all"),
      lwd = 2,
      col = c("black", "blue", "darkgreen", "darkred"),
      lty = seq(1,4), cex = 0.8,
      ncol = 4)

```

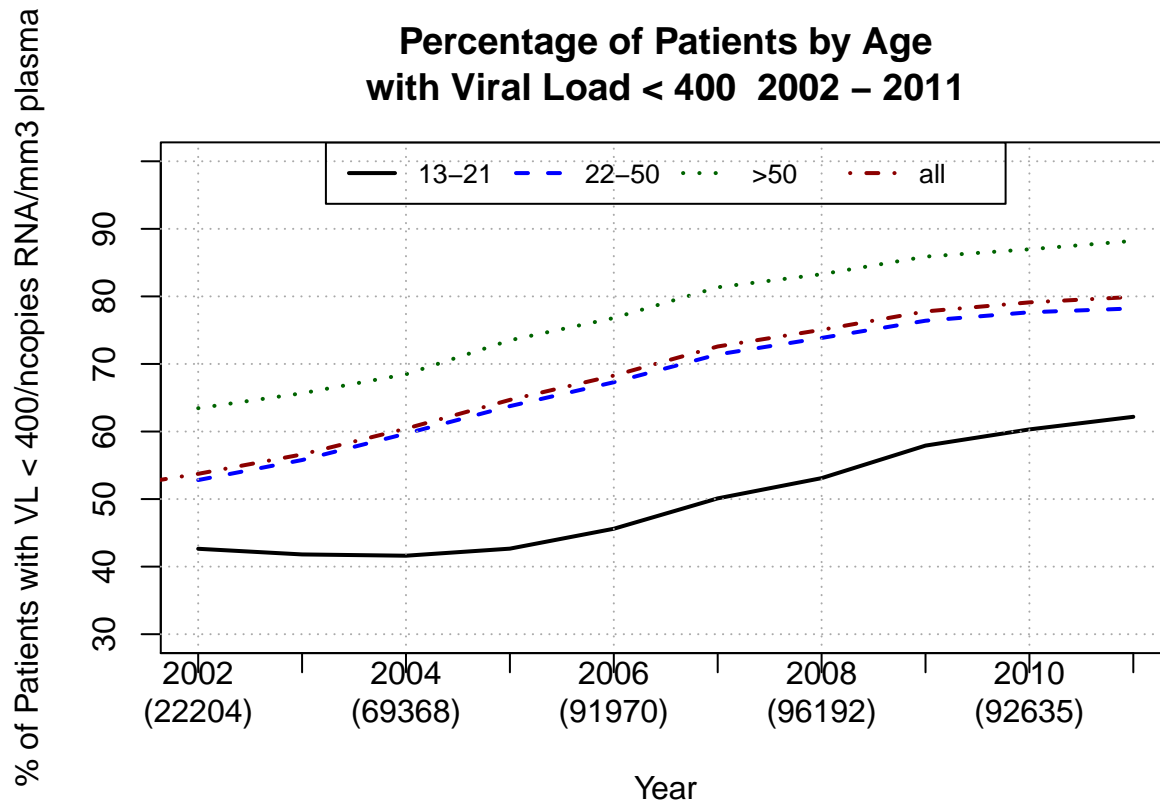


Figure 6. Percentage of antiretroviral resistance to different antiretroviral classes and number of 3 class resistances among individuals in which the genotype test have been performed in Brazil from 2003 to 2009.

```
plot(resist$NRTI ~ resist$year, type = "l",
     lwd = 2, xaxt = "n", ylim = c(10, 100),
     main = "Percentage of Drug Resistance by Year\n2003 - 2009",
     xlab = "Year(n)",
     ylab = "Percentage ")
lines(resist$year, resist$NNRTI, lwd = 2, lty = 2, col = "blue")
lines(resist$year, resist$PI, lwd = 2, lty = 3, col = "darkgreen")
lines(resist$year, resist$class3, lwd = 2, lty = 4, col = "darkred")
legend("top",
      legend = c("NRTI", "NNRTI", "PI", "3 class"),
      lwd = 2,
      col = c("black", "blue", "darkgreen", "darkred"),
      lty = seq(1,4), cex = 0.8,
      ncol = 4)
axis(1, at = resist$year, labels = tickre)
grid(col = "darkgrey")
```

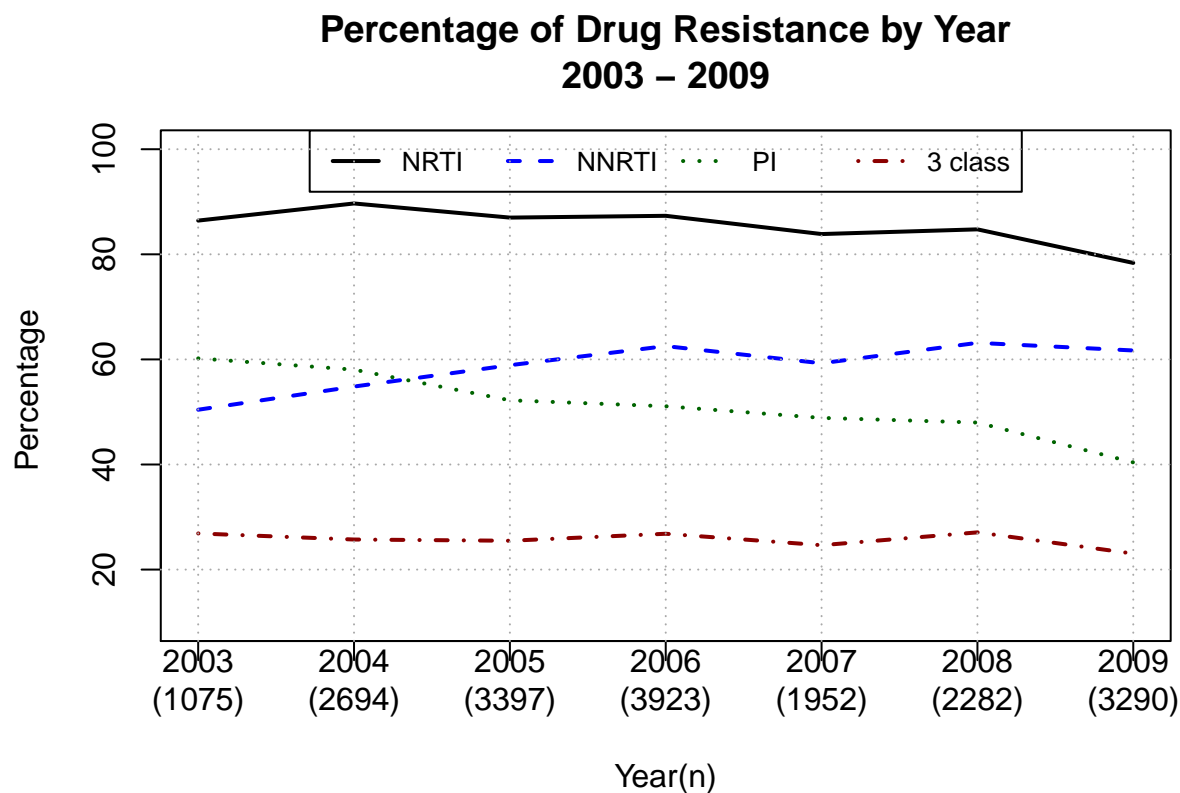


Figure 7. Distribution of Clades for Study Period by Year

```
library(lattice)
library(RColorBrewer)
barchart(data = cladeLong, x = pct ~ clade | year,
  main = "Distribution of Clades for Study Period\nby Year (Total N)",
  xlab = "Clade", ylab = "Percent",
  col = brewer.pal(5, "Set1"),
  index.cond = list(c(4,5,6,7,1,2,3)))
```

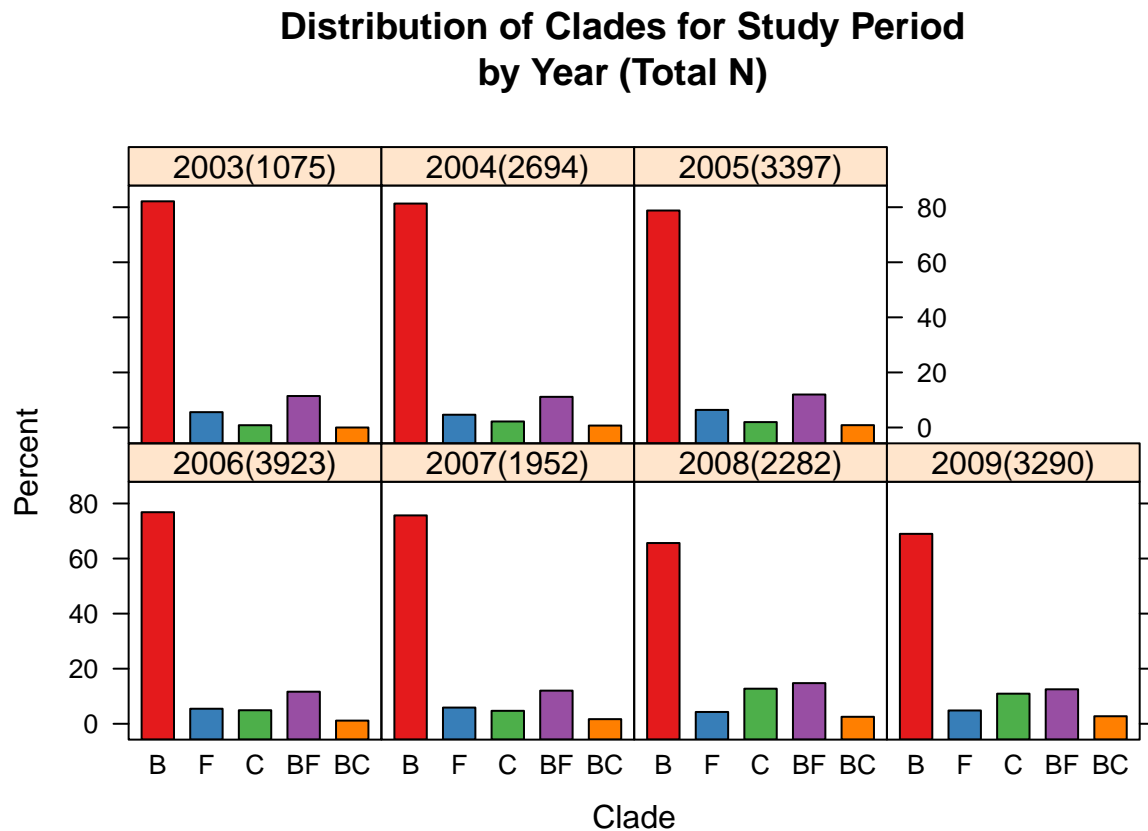


Figure 8. Distribution of Clades for by Region

```
library(lattice)
library(RColorBrewer)
barchart(data = claderegpc[claderegpc$ano == 2007,],
  x = pct*100 ~ cl | reg,
  main = "Distribution of Clades for Study Period\nby Region for 2007 (Percent)",
  xlab = "Clade", ylab = "Percent",
  col = brewer.pal(5, "Set1"))
```

### Distribution of Clades for Study Period by Region for 2007 (Percent)

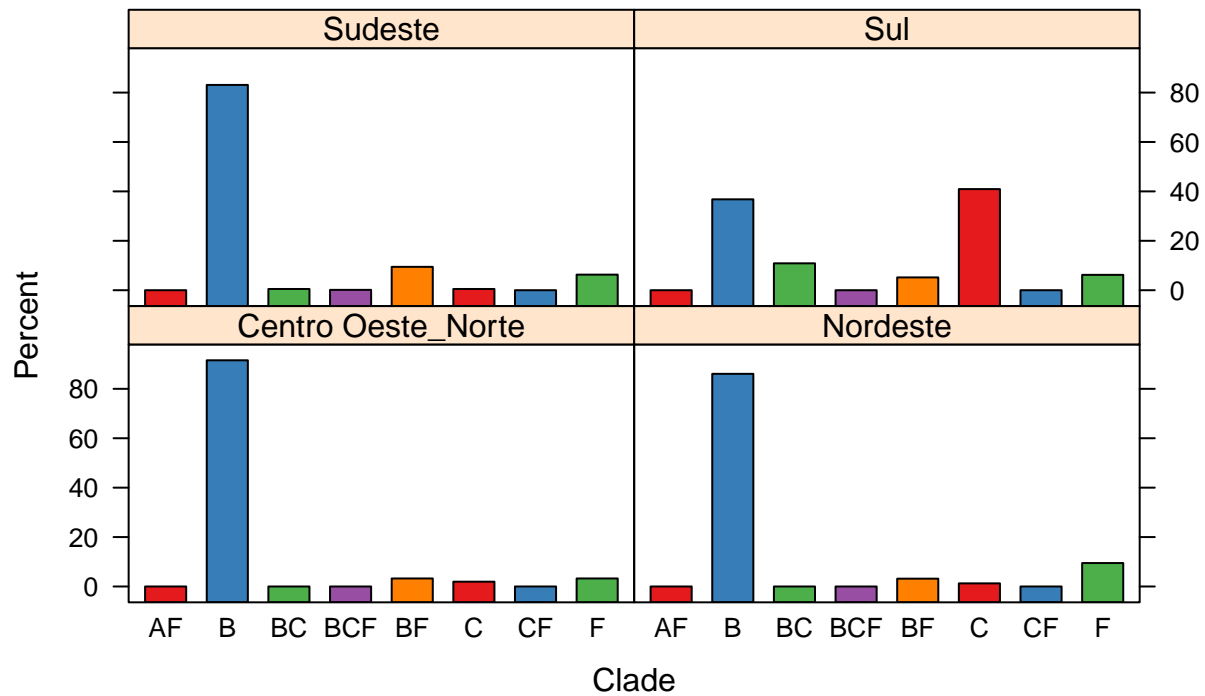
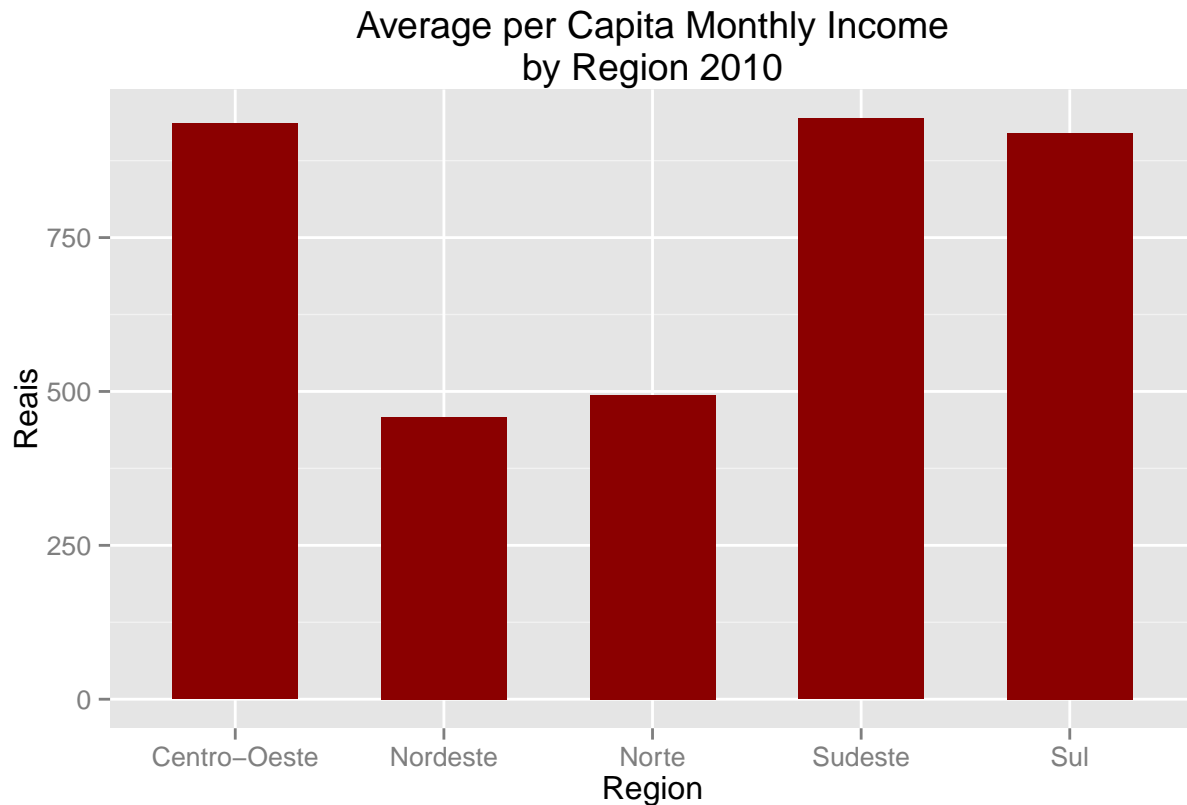


Figure 9. Average per Capita Monthly Income by Region, 2010.

```
require(ggplot2)
load("rendaregional.RData")
incPlot <- ggplot(renda.regional[-6,],
  aes(x = Region, y = renda_media))
incPlot <- incPlot + geom_bar(stat = "identity", width = 0.6,
  fill = "darkred")
incPlot <- incPlot + ggtitle("Average per Capita Monthly Income\nby Region 2010")
incPlot <- incPlot + xlab("Region") + ylab("Reais")
incPlot
```



**Figure 10. Percentage of Patients by Clade Showing 3-Class Antiretroviral Resistance.**

```
load("cl3resist.RData")
# create subset with just B,C,F
require(dplyr)
cl3bcf <- select(cl3resist, ano, totn, B, F, C, B3cpct, C3cpct, F3cpct)
# get rid of total row (last row)
cl3bcf <- cl3bcf[-8,]
# draw a graph of the 3 clades over the period
xyyears <- cl3bcf$ano
xn <- cl3bcf$totn
tickmark <- rep(NA, length(cl3bcf$ano))
for (i in 1:length(cl3bcf$ano)) {
  tickmark[i] <- paste(xyyears[i], "\n(",xn[i],")", sep = "")
}
cladeNames <- c("Clade B", "Clade C", "Clade F")
plot(1:7, cl3bcf$B3cpct, xaxt = "n", type = "l",
     lwd = 2, xlab = "Year", ylim = c(.1, .3),
     ylab = "Proportion Showing Resistance",
     main = "Percentage of Patients by Clade Showing\n3-Class Antiretroviral Resistance")
lines(1:7, cl3bcf$C3cpct, lwd = 2, col = "blue", lty = 2)
lines(1:7, cl3bcf$F3cpct, lwd = 2, col = "darkred", lty = 3)
grid(col = "darkgrey")
axis(1, at = 1:7, labels = tickmark)
```

```

legend("top",
      legend = cladeNames,
      lwd = 2,
      col = c("black", "blue", "darkred"),
      lty = seq(1,3), cex = 0.8,
      ncol = 3)

```

