lab2.R

brintzb

Wed Oct 14 09:46:26 2015

library(ggplot2)

## Warning: package 'ggplot2' was built under R version 3.2.2

load(url("http://stat511.cwick.co.nz/data/cdc.rda"))
males <- subset(cdc, gender == 'm')
head(males)

## genhlth exerany hlthplan smoke100 height weight wtdesire age gender
## 1 good 0 1 0 70 175 175 77 m
## 7 very good 1 1 0 71 194 185 31 m
## 8 very good 0 1 0 67 170 160 45 m
## 10 good 1 1 0 70 180 170 44 m
## 11 excellent 1 1 1 69 186 175 46 m
## 12 fair 1 1 1 69 168 148 62 m

# 1 What is the average age of female respondents?
females = subset(cdc,gender=="f")
mean(females$age)

## [1] 45.79772

# 2 Find the median weight of respondents without health insurance. (Hint: create a new data.frame of people without health insurance, then find the median of the weight column in that new data.frame. There is a median function.)
nohealth=subset(cdc,hlthplan==0)
median(nohealth$weight)

## [1] 163

#or in one line
median(subset(cdc,hlthplan==0)$weight)

## [1] 163

# 3 Find the median weight of respondents without health insurance and who had not exercised in the last month.
nohealthnoex=subset(cdc,hlthplan==0 & exerany==0)
median(nohealthnoex$weight)

## [1] 160

# 4 How many females are younger than age 30?
dim(subset(females,age < 30))

## [1] 2168 9

#or
length(subset(females,age < 30)$age)

## [1] 2168

# 5 (Harder) Find the average difference between weight and wtdesire (the desired weight) for all respondents and for males and females separately.
mean(cdc$weight - cdc$wtdesire)

## [1] 14.5891

mean(males$weight - males$wtdesire)

## [1] 10.70613

mean(females$weight - females$wtdesire)

## [1] 18.15118

######### Plotting ######
#1
qplot(gender,age,data=cdc)



#2
qplot(age,data=cdc) + facet\_wrap(~gender,nrow=2)

## stat\_bin: binwidth defaulted to range/30. Use 'binwidth = x' to adjust this.
## stat\_bin: binwidth defaulted to range/30. Use 'binwidth = x' to adjust this.



#3
qplot(factor(exerany),weight,data=cdc,geom="boxplot")



#4
qplot(gender,weight,data=cdc,color=gender,geom="boxplot")



#5
qplot(factor(exerany),weight,geom="violin",data=cdc) + facet\_wrap(~gender)



#7
qplot(factor(exerany),weight,geom="violin",data=cdc) + facet\_wrap(~gender) + theme\_bw()



#8
qplot(factor(exerany),weight,geom="violin",data=cdc) + facet\_wrap(~gender) + theme\_bw() +
 ggtitle("Hello,I'm a title")

