M448 Course Project

Gabrielle Salamanca

3/9/2022

## R Markdown

#calling libraries  
library(caret)

## Loading required package: ggplot2

## Loading required package: lattice

library(ggplot2)  
  
#calling dataset  
Rnames = read.csv("C:/Users/knigh/OneDrive/Documents/Resume Names.csv", header=T, na.strings="?", stringsAsFactors = T)  
  
  
#checking out the specs  
dim(Rnames)

## [1] 4870 28

names(Rnames)

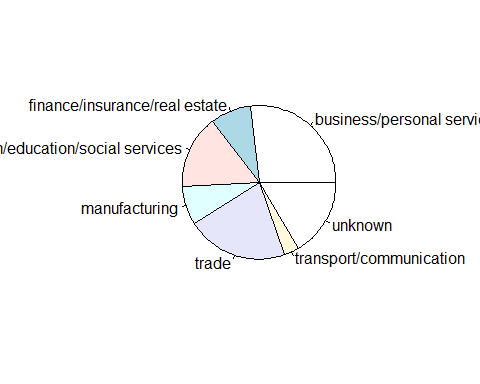
## [1] "X" "name" "gender" "ethnicity" "quality"   
## [6] "call" "city" "jobs" "experience" "honors"   
## [11] "volunteer" "military" "holes" "school" "email"   
## [16] "computer" "special" "college" "minimum" "equal"   
## [21] "wanted" "requirements" "reqexp" "reqcomm" "reqeduc"   
## [26] "reqcomp" "reqorg" "industry"

summary(Rnames)

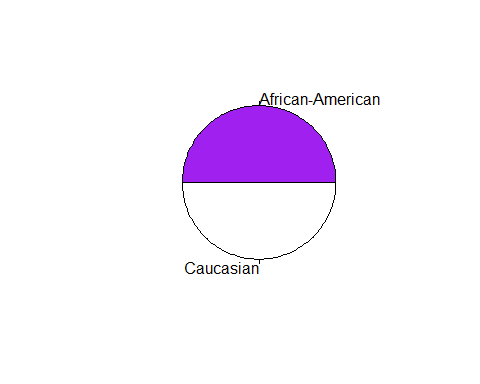
## X name gender ethnicity quality call   
## Min. : 1 Tamika : 256 female:3746 afam:2435 high:2446 no :4478   
## 1st Qu.:1218 Anne : 242 male :1124 cauc:2435 low :2424 yes: 392   
## Median :2436 Allison: 232   
## Mean :2436 Latonya: 230   
## 3rd Qu.:3653 Emily : 227   
## Max. :4870 Latoya : 226   
## (Other):3457   
## city jobs experience honors volunteer   
## boston :2166 Min. :1.000 Min. : 1.000 no :4613 no :2866   
## chicago:2704 1st Qu.:3.000 1st Qu.: 5.000 yes: 257 yes:2004   
## Median :4.000 Median : 6.000   
## Mean :3.661 Mean : 7.843   
## 3rd Qu.:4.000 3rd Qu.: 9.000   
## Max. :7.000 Max. :44.000   
##   
## military holes school email computer special college   
## no :4397 no :2688 no :2145 no :2536 no : 874 no :3269 no :1366   
## yes: 473 yes:2182 yes:2725 yes:2334 yes:3996 yes:1601 yes:3504   
##   
##   
##   
##   
##   
## minimum equal wanted requirements reqexp   
## none :2746 no :3452 manager : 741 no :1036 no :2750   
## some :1064 yes:1418 office support: 578 yes:3834 yes:2120   
## 2 : 356 other : 736   
## 3 : 331 retail sales : 818   
## 5 : 163 secretary :1621   
## 1 : 142 supervisor : 376   
## (Other): 68   
## reqcomm reqeduc reqcomp reqorg   
## no :4262 no :4350 no :2741 no :4516   
## yes: 608 yes: 520 yes:2129 yes: 354   
##   
##   
##   
##   
##   
## industry   
## business/personal services :1304   
## finance/insurance/real estate : 414   
## health/education/social services: 754   
## manufacturing : 404   
## trade :1042   
## transport/communication : 148   
## unknown : 804

## Including Plots

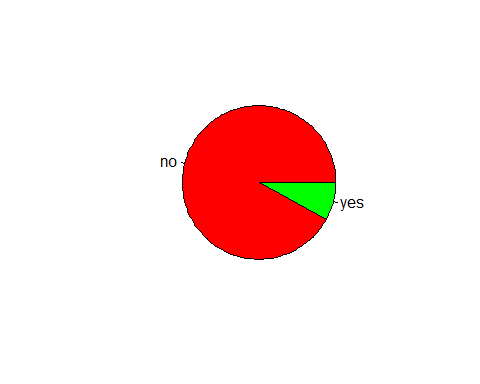
#industry variable  
pie(table(Rnames$industry))



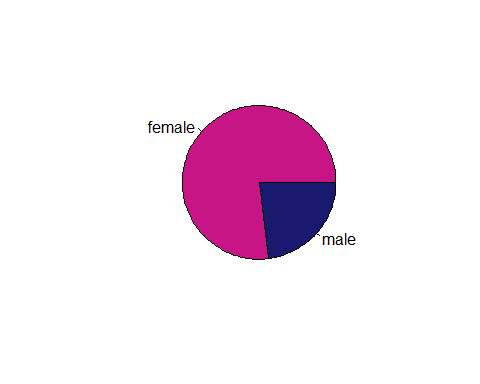
#ethnicity variable  
colors = c("purple", "white")  
ethnic <- c("African-American", "Caucasian")  
pie(table(Rnames$ethnicity), col = colors, labels = ethnic)



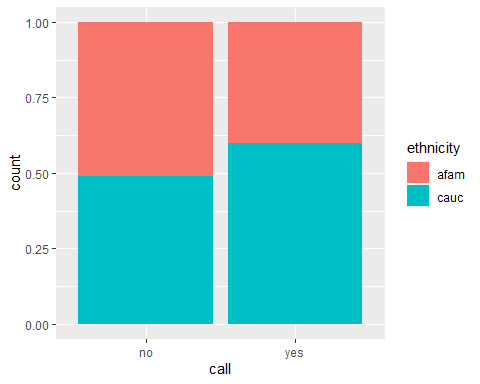
#call variable  
colors = c("red", "green")  
pie(table(Rnames$call),col = colors)



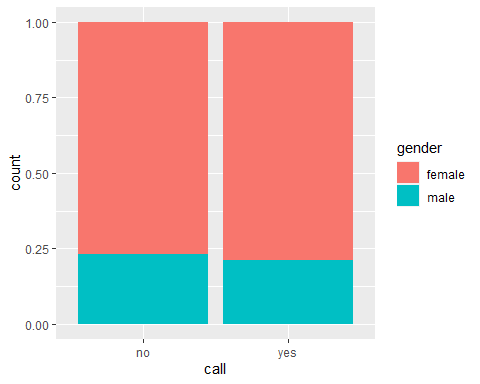
#gender variable  
colors = c("mediumvioletred", "midnightblue")  
pie(table(Rnames$gender),col = colors)



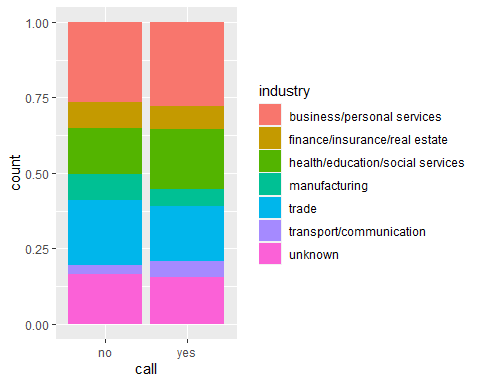
#plots w/ x = call  
ggplot(Rnames, aes(x = call, fill = ethnicity)) + geom\_bar(position = "fill")



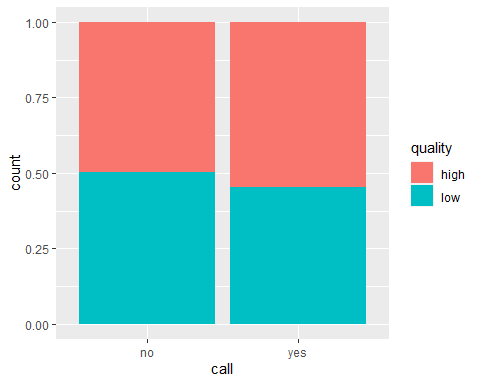
ggplot(Rnames, aes(x = call, fill = gender)) + geom\_bar(position = "fill")



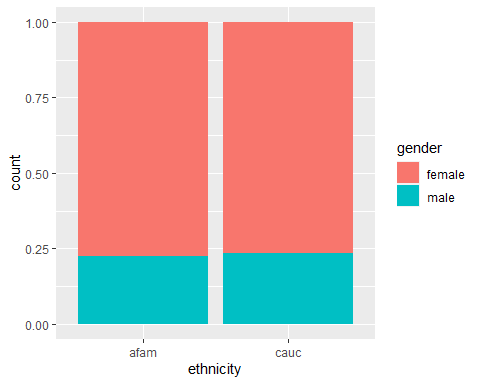
ggplot(Rnames, aes(x = call, fill = industry)) + geom\_bar(position = "fill")



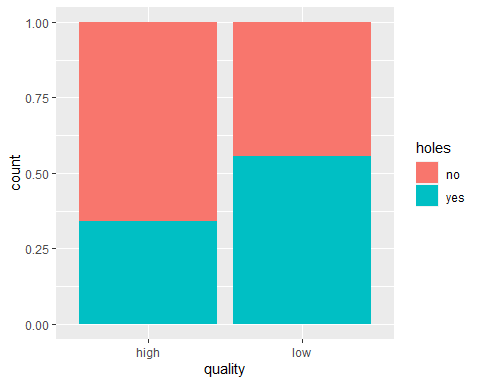
ggplot(Rnames, aes(x = call, fill = quality)) + geom\_bar(position = "fill")



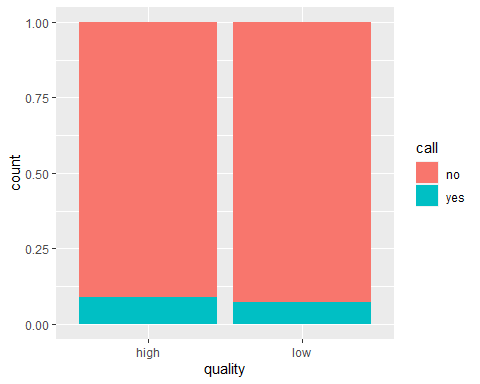
#plots w/ other x  
ggplot(Rnames, aes(x = ethnicity, fill = gender)) + geom\_bar(position = "fill")



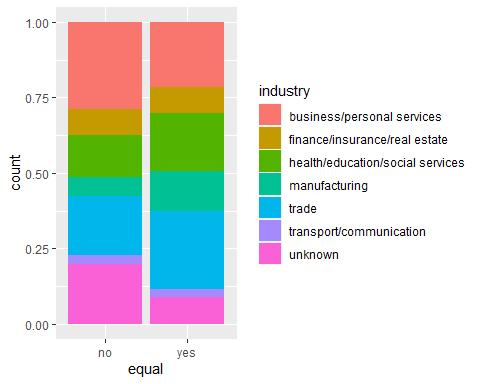
ggplot(Rnames, aes(x = quality, fill = holes)) + geom\_bar(position = "fill")



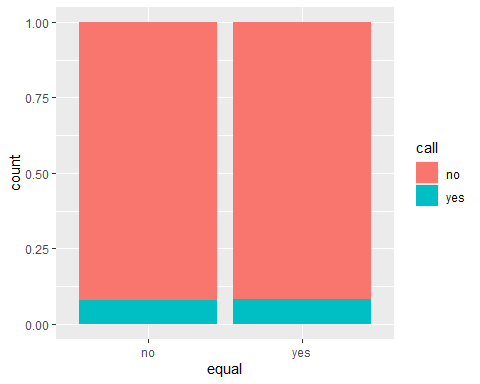
ggplot(Rnames, aes(x = quality, fill = call)) + geom\_bar(position = "fill")



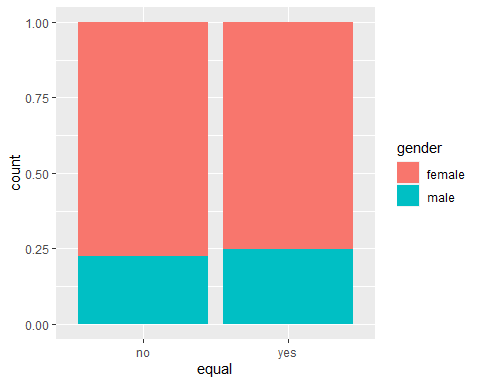
#plots w/ x = equal  
ggplot(Rnames, aes(x = equal, fill = industry)) + geom\_bar(position = "fill")



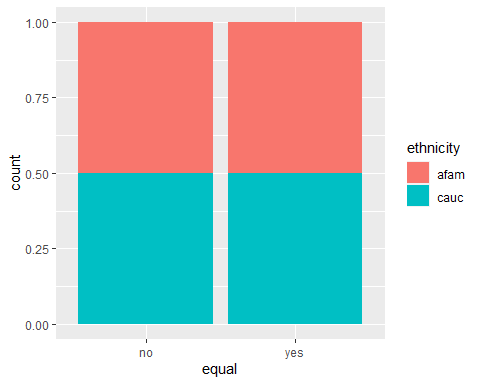
ggplot(Rnames, aes(x = equal, fill = call)) + geom\_bar(position = "fill")



ggplot(Rnames, aes(x = equal, fill = gender)) + geom\_bar(position = "fill")



ggplot(Rnames, aes(x = equal, fill = ethnicity)) + geom\_bar(position = "fill")



#pairs  
Names = subset(Rnames, select = -c(X, name, city, holes, email, equal, wanted, industry))  
pairs(Names)

