Stat123\_lab3

#1  
  
#a  
library(readxl)  
sfo <- read\_excel("SFO\_Recovery\_Onsite\_December\_2020.xlsx", sheet = 'Data')  
  
#b  
length(sfo$Q4)

[1] 1086

dim(sfo)[1]

[1] 1086

print("there are 1086 observations")

[1] "there are 1086 observations"

#c  
length(sfo)

[1] 42

dim(sfo)[2]

[1] 42

print("there are 42 columns")

[1] "there are 42 columns"

head(sfo)

# A tibble: 6 × 42  
 Respnum CCGID Source LANG TERM BAREA Q1 Q2 Q2a Q2b Q3\_1 Q3\_2  
 <dbl> <dbl> <dbl> <dbl> <dbl> <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
1 6 1 1 1 2 D 2 7 2 NA 1 2  
2 7 2 1 1 2 D 3 4 2 NA 4 NA  
3 8 3 1 1 2 D 2 1 1 1 0 NA  
4 9 4 1 1 2 D 6 1 2 NA 2 NA  
5 10 5 1 1 2 D 2 2 2 NA 0 NA  
6 11 6 1 1 2 D 3 2 2 NA 0 NA  
# ℹ 30 more variables: Q3\_3 <lgl>, Q4 <dbl>, Q5a <dbl>, Q5b <dbl>, Q5c <dbl>,  
# Q5d <dbl>, Q5e <dbl>, Q5f <dbl>, Q6\_1 <dbl>, Q6\_2 <dbl>, Q7 <dbl>,  
# Q7a <chr>, Q7a\_1 <dbl>, Q7a\_2 <dbl>, Q7a\_3 <dbl>, Q8 <dbl>, Q9 <chr>,  
# Q10 <chr>, Q11 <dbl>, Q12 <dbl>, Q12\_Country <chr>, CountryCode <dbl>,  
# Q13 <dbl>, Q14 <dbl>, Q14\_3\_other <lgl>, Q15 <dbl>, Q16 <chr>, Q16\_1 <dbl>,  
# Q16\_2 <dbl>, Q16\_3 <dbl>

print("after investigating the excel, we noted that the two columns can be considered individuals, so there are 40 variables in the dataset.")

[1] "after investigating the excel, we noted that the two columns can be considered individuals, so there are 40 variables in the dataset."

#d  
q4 <- sfo$Q4  
q4 <- factor( q4, levels = c(1,2,3,4,5,0))  
(tab <- table(q4))

q4  
 1 2 3 4 5 0   
159 186 291 253 188 9

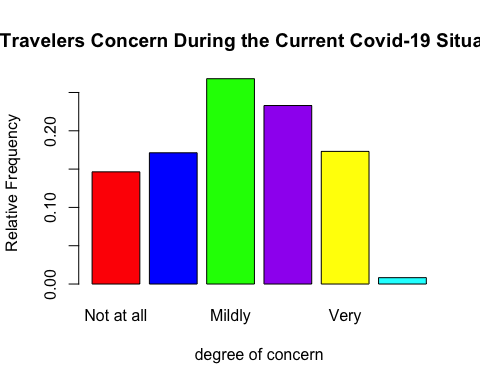
(tabrf = tab / sum(tab))

q4  
 1 2 3 4 5 0   
0.146408840 0.171270718 0.267955801 0.232965009 0.173112339 0.008287293

tabnames <- c("Not at all", "Slightly", "Mildly", "Somewhat", "Very", "No response")  
  
names(tabrf) <- tabnames  
  
tabrf

Not at all Slightly Mildly Somewhat Very No response   
0.146408840 0.171270718 0.267955801 0.232965009 0.173112339 0.008287293

barplot(tabrf, main = "Travelers Concern During the Current Covid-19 Situation", xlab = "degree of concern", ylab = 'Relative Frequency', col = c("red", "blue", "green", "purple", "yellow", "cyan"))



#e  
print("Most people feel a mild degree of concern about the covid situation, with a few more people somewhat or very concerned. The least people did not respond, or are not at all worried.")

[1] "Most people feel a mild degree of concern about the covid situation, with a few more people somewhat or very concerned. The least people did not respond, or are not at all worried."