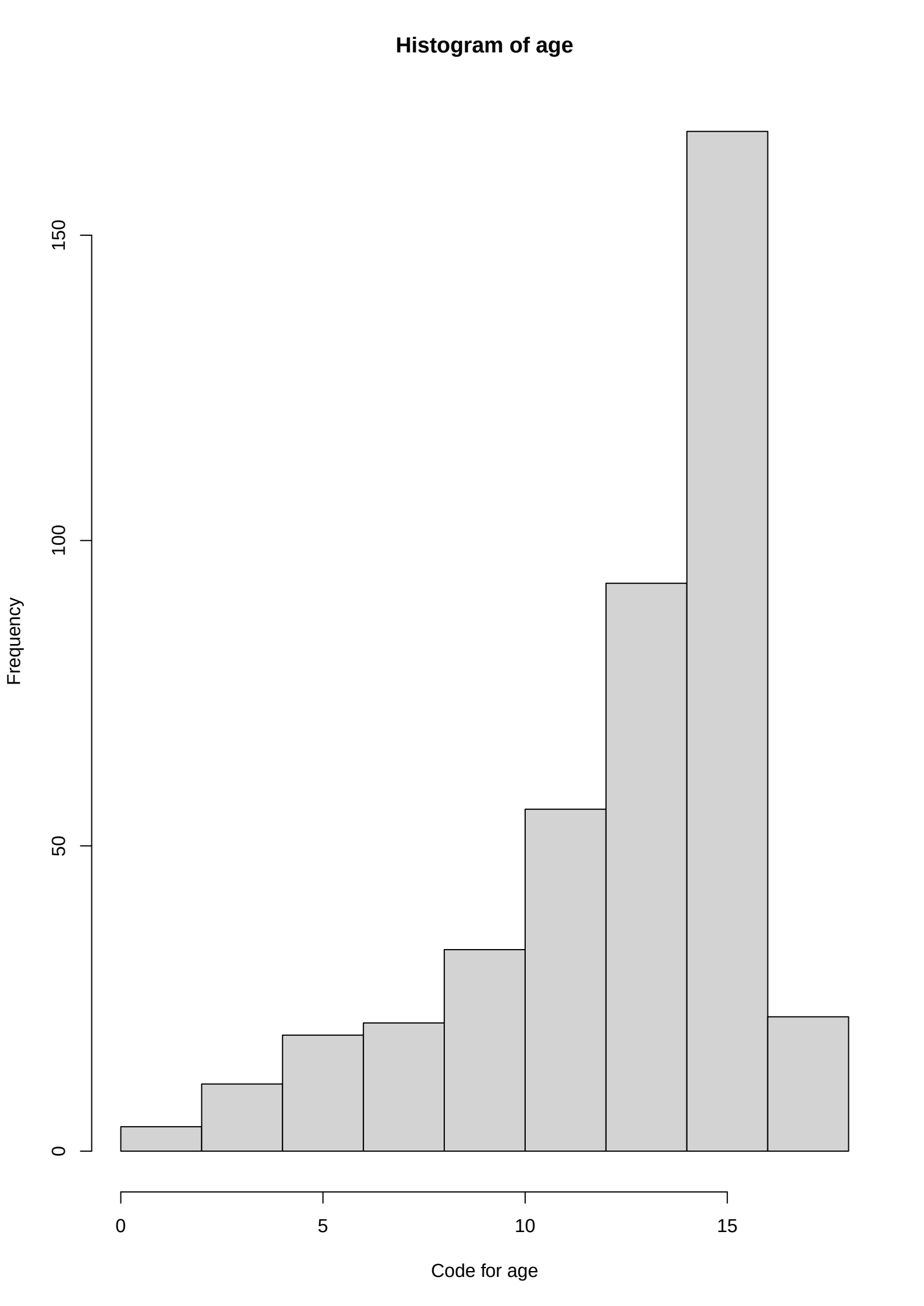
```
setwd("/Users/carolinaherrera/Documents/Assignment 2 - CRIM 250")
dat <- read.csv(file = 'dat.nsduh.small.csv')

summary(dat)
head(dat)

names(dat)

# Variables: mjage, cigever, alcever, AGE2, sexatract, speakengl, irsex. //// Coded age, have they ever smoked part of a cigarette, have they ever had a drink of alcohol, have they ever used mariijuana oor hashish, what best describes their sexual attractment.
# These are categorical variables

dim(dat)
# I got two numbers 426 and 7, is 426 how many individual are in the data, and 7 the number of categories?


hist(dat$AGE2, main="Histogram of age", xlab="Code for age", ylab="Frequency")

# The numbers/bins on the x-axis represent the code associated with the age of respondents. For example, 5 doesn't mean that the respondent is 5 years old - when we look at the age associated with that code in the codebook we would see that the respondent is 16 years old. Same for 15 on the x-axis, the codebook tells us that the respondennt is between 35-49 years old. The numbers on the y-axis represent the frequency of indviduals that fit into each of those groups.
```


Histogram of age
Frequency
0
50
100
150
0
5
10
15
Code for age