Project.7

1.

Use the Batting.csv file (inside the "core" folder) to discover who is a member of the 40-40 club, namely, who has hit 40 home runs and also has (simultaneously) stolen 40 bases in the same season.
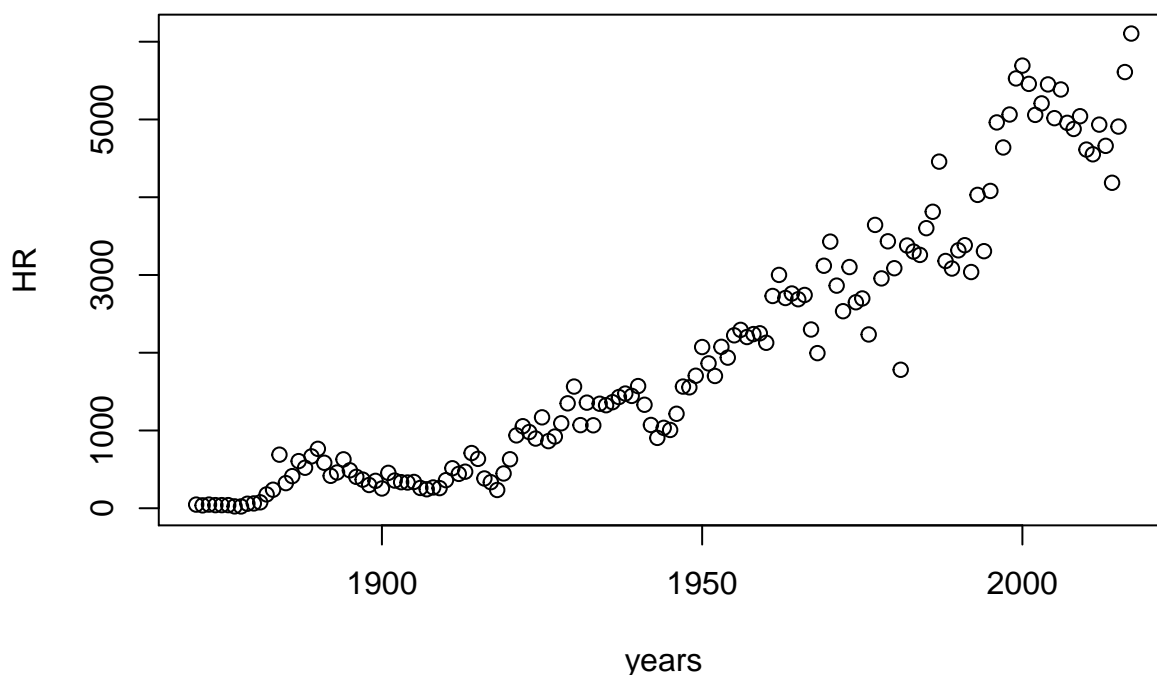
```
myDF <- read.csv ("Downloads/baseballdatabank-master/core/Batting.csv")
myDF[myDF$HR>=40 & myDF$SB>=40,]
```

```
##          playerID yearID stint teamID lgID   G  AB   R   H X2B X3B HR RBI SB
## 65469 cansejo01   1988     1    OAK   AL 158 610 120 187  34   0 42 124 40
## 74270 bondsba01   1996     1    SFN   NL 158 517 122 159  27   3 42 129 40
## 77656 rodrial01   1998     1    SEA   AL 161 686 123 213  35   5 42 124 46
## 88497 soriaal01   2006     1    WAS   NL 159 647 119 179  41   2 46  95 41
##       CS  BB  SO IBB HBP SH SF GIDP
## 65469 16  78 128  10  10  1  6   15
## 74270  7 151  76  30   1  0  6   11
## 77656 13  45 121   0  10  3  4   12
## 88497 17  67 160  16   9  2  3    3
```

2.

Make a plot that depicts the total number of home runs per year (across all players on all teams). The plot should have the years as the labels for the x-axis, and should have the number of home runs as the labels for the y-axis.

```
v <- tapply(myDF$HR,myDF$yearID,sum)
plot(names(v),v,xlab = "years",ylab = "HR")
```

3

```r
#a.
myBatting <- read.csv ("Downloads/baseballdatabank-master/core/Batting.csv")
myPeople <- read.csv ("Downloads/baseballdatabank-master/core/People.csv")
myDF <- merge(myBatting, myPeople, by="playerID")
#b.
Fullname <- paste(myDF$nameFirst,myDF$nameLast)
myDF$fullname <- Fullname
#c
myDF[myDF$HR>=40 & myDF$SB>=40,]
```

```
##          playerID yearID stint teamID lgID   G  AB   R   H X2B X3B HR RBI SB
## 8520   bondsba01   1996     1    SFN   NL 158 517 122 159  27   3 42 129 40
## 14008  cansejo01   1988     1    OAK   AL 158 610 120 187  34   0 42 124 40
## 80212  rodrial01   1998     1    SEA   AL 161 686 123 213  35   5 42 124 46
## 88502  soriaal01   2006     1    WAS   NL 159 647 119 179  41   2 46  95 41
##        CS  BB  SO IBB HBP SH SF GIDP birthYear birthMonth birthDay
## 8520    7 151  76  30   1  0  6   11      1964          7       24
## 14008  16  78 128  10  10  1  6   15      1964          7        2
## 80212  13  45 121   0  10  3  4   12      1975          7       27
## 88502  17  67 160  16   9  2  3    3      1976          1        7
##        birthCountry         birthState          birthCity deathYear
## 8520            USA                 CA          Riverside        NA
## 14008          Cuba          La Habana          La Habana        NA
## 80212           USA                 NY           New York        NA
## 88502          D.R. San Pedro de Macoris San Pedro de Macoris      NA
##        deathMonth deathDay deathCountry deathState deathCity nameFirst
## 8520           NA       NA                                      Barry
## 14008          NA       NA                                       Jose
## 80212          NA       NA                                       Alex
## 88502          NA       NA                                    Alfonso
##         nameLast          nameGiven weight height bats throws      debut
## 8520       Bonds        Barry Lamar    185     73    L      L 1986-05-30
## 14008    Canseco               Jose    240     76    R      R 1985-09-02
## 80212  Rodriguez Alexander Enmanuel    230     75    R      R 1994-07-08
## 88502    Soriano  Alfonso Guilleard    195     73    R      R 1999-09-14
##         finalGame   retroID   bbrefID         fullname
## 8520   2007-09-26 bondb001 bondsba01     Barry Bonds
## 14008  2001-10-06 cansj001 cansejo01     Jose Canseco
## 80212  2016-08-12 rodra001 rodrial01   Alex Rodriguez
## 88502  2014-07-05 soria001 soriaal01 Alfonso Soriano
```