HW1

Marjan Rezvani

9/14/2020

3:

Dice Rolls with sample

We can simulate one such die roll in R as follows:

```
sample(1:6, size = 1, replace = TRUE)
```

```
## [1] 1
```

now we want to roll two fair, six-sided dice and compute their sum Here's one way to do it:

```
sample(1:6, size = 1, replace = TRUE) + sample(1:6, size = 1, replace = TRUE)
```

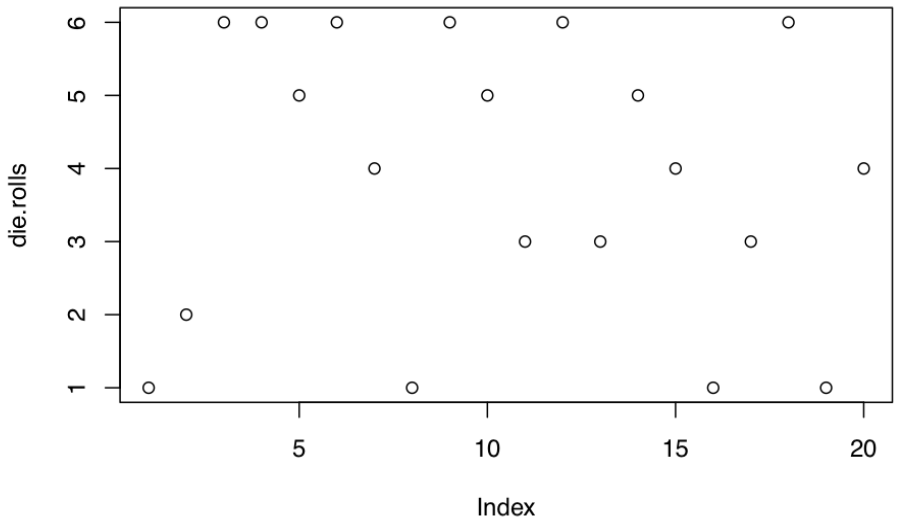
```
## [1] 2
```

I could repeat the experiment of rolling a single fair die 20 times as follows:

```
die.rolls <- sample(x = 1:6, size = 20, replace = TRUE)
die.rolls
```

```
##  [1] 1 2 6 6 5 6 4 1 6 5 3 6 3 5 4 1 3 6 1 4
```

```
#tmp <- dice(10,4)
plot(die.rolls)
```



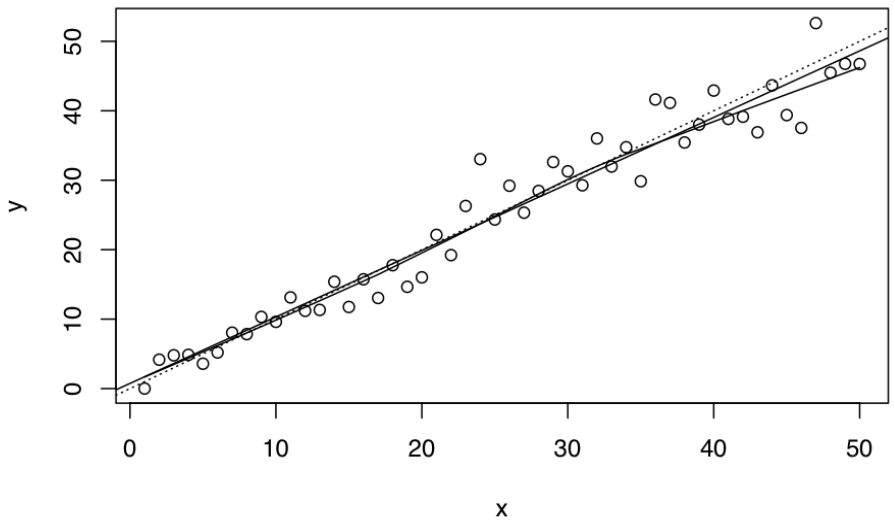
4:

```
x <- 1:50
w <- 1 + sqrt(x)/2
example1 <- data.frame(x=x, y= x + rnorm(x)*w)
attach(example1)
```

```
fm <- lm(y ~ x)
summary(fm)
```

```
##
## Call:
## lm(formula = y ~ x)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.2707 -1.7944 -0.3619  1.3977  9.3060
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.74308    0.93714   0.793    0.432
## x            0.95779    0.03198  29.946   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.264 on 48 degrees of freedom
## Multiple R-squared:  0.9492, Adjusted R-squared:  0.9481
## F-statistic: 896.7 on 1 and 48 DF,  p-value: < 2.2e-16
```

```
lrf <- lowess(x, y)
plot(x, y)
lines(x, lrf$y)
abline(0, 1, lty=3)
abline(coef(fm))
```




```
detach()
```

```
rm(list = ls(all = TRUE)) # clear workspace

load("/Users/marjanrezvani/Documents/Fall2020/eco_stat/data/acs2017_ny/acs2017_ny_data.RData")
acs2017_ny[1:10,1:7]
```

```
##    AGE female educ_nohs educ_hs educ_somecoll educ_college educ_advdeg
## 1   72      1         0       0             0            0           1
## 2   72      0         0       0             0            0           1
## 3   31      0         0       0             0            1           0
## 4   28      1         0       0             0            1           0
## 5   54      0         0       0             0            0           1
## 6   45      1         0       1             0            0           0
## 7   84      1         0       0             1            0           0
## 8   71      0         0       0             0            1           0
## 9   68      1         0       0             1            0           0
## 10  37      1         1       0             0            0           0
```

```
attach(acs2017_ny)
```

```
summary(acs2017_ny)
```

```
##       AGE            female         educ_nohs        educ_hs
##  Min.   : 0.00   Min.   :0.0000   Min.   :0.000   Min.   :0.0000
##  1st Qu.:22.00   1st Qu.:0.0000   1st Qu.:0.000   1st Qu.:0.0000
##  Median :42.00   Median :1.0000   Median :0.000   Median :0.0000
##  Mean   :41.57   Mean   :0.5156   Mean   :0.271   Mean   :0.2804
##  3rd Qu.:60.00   3rd Qu.:1.0000   3rd Qu.:1.000   3rd Qu.:1.0000
##  Max.   :95.00   Max.   :1.0000   Max.   :1.000   Max.   :1.0000
##
##  educ_somecoll    educ_college     educ_advdeg
##  Min.   :0.000   Min.   :0.0000   Min.   :0.000
##  1st Qu.:0.000   1st Qu.:0.0000   1st Qu.:0.000
```

```
##  Median :0.000   Median :0.0000   Median :0.000
##  Mean   :0.173   Mean   :0.1567   Mean   :0.119
##  3rd Qu.:0.000   3rd Qu.:0.0000   3rd Qu.:0.000
##  Max.   :1.000   Max.   :1.0000   Max.   :1.000
##
##                SCHOOL                                EDUC
##  N/A              :  5569   Grade 12                 :55119
##  No, not in school:144968   4 years of college       :30802
##  Yes, in school   : 46048   5+ years of college      :23385
##  Missing          :     0   1 year of college        :19947
##                             Nursery school to grade 4:14240
##                             2 years of college       :14065
##                             (Other)                  :39027
##                                          EDUCD
##  Regular high school diploma                 :35689
##  Bachelor's degree                           :30802
##  1 or more years of college credit, no degree:19947
##  Master's degree                             :17010
##  Associate's degree, type not specified      :14065
##  Some college, but less than 1 year          : 9086
##  (Other)                                     :69986
##                                       DEGFIELD
##  N/A                                     :142398
##  Business                                :  9802
##  Education Administration and Teaching   :  6708
##  Social Sciences                         :  4836
##  Medical and Health Sciences and Services:  3919
##  Fine Arts                               :  3491
##  (Other)                                 : 25431
##                                     DEGFIELDD
##  N/A                                    :142398
##  Psychology                             :  2926
##  Business Management and Administration:  2501
##  Accounting                             :  2284
##  General Education                      :  2238
##  English Language and Literature        :  2202
##  (Other)                                : 42036
##                                    DEGFIELD2
##  N/A                                   :190425
##  Business                              :   972
##  Social Sciences                       :   853
##  Education Administration and Teaching:   611
##  Fine Arts                             :   465
##  Communications                        :   352
##  (Other)                               :  2907
##                                                                DEGFIELD2D
##  N/A                                                                :190425
##  Psychology                                                         :   284
##  Economics                                                          :   260
##  Political Science and Government                                   :   243
##  Business Management and Administration                             :   217
##  French, German, Latin and Other Common Foreign Language Studies:   205
##  (Other)                                                            :  4951
##       PUMA            GQ           OWNERSHP       OWNERSHPD
```

```
##  Min.   : 100   Min.   :1.000   Min.   :0.000   Min.   : 0.00
##  1st Qu.:1500   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:12.00
##  Median :3201   Median :1.000   Median :1.000   Median :13.00
##  Mean   :2713   Mean   :1.148   Mean   :1.266   Mean   :14.95
##  3rd Qu.:3902   3rd Qu.:1.000   3rd Qu.:2.000   3rd Qu.:22.00
##  Max.   :4114   Max.   :5.000   Max.   :2.000   Max.   :22.00
##
##     MORTGAGE        OWNCOST          RENT         COSTELEC
##  Min.   :0.000   Min.   :    0   Min.   :   0   Min.   :   0
##  1st Qu.:0.000   1st Qu.: 1208   1st Qu.:   0   1st Qu.: 960
##  Median :1.000   Median : 2891   Median :   0   Median :1560
##  Mean   :1.453   Mean   :38582   Mean   : 393   Mean   :2311
##  3rd Qu.:3.000   3rd Qu.:99999   3rd Qu.: 630   3rd Qu.:2520
##  Max.   :4.000   Max.   :99999   Max.   :3800   Max.   :9997
##
##     COSTGAS        COSTWATR       COSTFUEL       HHINCOME
##  Min.   :   0   Min.   :   0   Min.   :   0   Min.   : -11800
##  1st Qu.: 840   1st Qu.: 320   1st Qu.:9993   1st Qu.:  41600
##  Median :2400   Median :1400   Median :9993   Median :  81700
##  Mean   :5032   Mean   :4836   Mean   :7935   Mean   : 114902
##  3rd Qu.:9993   3rd Qu.:9993   3rd Qu.:9993   3rd Qu.: 140900
##  Max.   :9997   Max.   :9997   Max.   :9997   Max.   :2030000
##                                               NA's   :10630
##     FOODSTMP        LINGISOL         ROOMS           BUILTYR2
##  Min.   :1.000   Min.   :0.000   Min.   : 0.000   Min.   : 0.000
##  1st Qu.:1.000   1st Qu.:1.000   1st Qu.: 4.000   1st Qu.: 1.000
##  Median :1.000   Median :1.000   Median : 6.000   Median : 3.000
##  Mean   :1.147   Mean   :1.002   Mean   : 5.887   Mean   : 3.711
##  3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.: 8.000   3rd Qu.: 5.000
##  Max.   :2.000   Max.   :2.000   Max.   :16.000   Max.   :22.000
##
##     UNITSSTR        FUELHEAT          SSMC            FAMSIZE
##  Min.   : 0.00   Min.   :0.000   Min.   :0.00000   Min.   : 1.000
##  1st Qu.: 3.00   1st Qu.:2.000   1st Qu.:0.00000   1st Qu.: 2.000
##  Median : 3.00   Median :2.000   Median :0.00000   Median : 3.000
##  Mean   : 4.39   Mean   :2.959   Mean   :0.01102   Mean   : 3.087
##  3rd Qu.: 6.00   3rd Qu.:4.000   3rd Qu.:0.00000   3rd Qu.: 4.000
##  Max.   :10.00   Max.   :9.000   Max.   :2.00000   Max.   :19.000
##
##      NCHILD           NCHLT5            RELATE          RELATED
##  Min.   :0.0000   Min.   :0.00000   Min.   : 1.000   Min.   : 101.0
##  1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.: 1.000   1st Qu.: 101.0
##  Median :0.0000   Median :0.00000   Median : 2.000   Median : 201.0
##  Mean   :0.5009   Mean   :0.08441   Mean   : 3.307   Mean   : 335.6
##  3rd Qu.:1.0000   3rd Qu.:0.00000   3rd Qu.: 3.000   3rd Qu.: 301.0
##  Max.   :9.0000   Max.   :5.00000   Max.   :13.000   Max.   :1301.0
##
##      MARST            RACE          RACED         HISPAN
##  Min.   :1.000   Min.   :1.00   Min.   :100   Min.   :0.0000
##  1st Qu.:1.000   1st Qu.:1.00   1st Qu.:100   1st Qu.:0.0000
##  Median :5.000   Median :1.00   Median :100   Median :0.0000
##  Mean   :3.742   Mean   :2.03   Mean   :205   Mean   :0.4153
##  3rd Qu.:6.000   3rd Qu.:2.00   3rd Qu.:200   3rd Qu.:0.0000
##  Max.   :6.000   Max.   :9.00   Max.   :990   Max.   :4.0000
```

```
##
##     HISPAND                  BPL                          BPLD
##  Min.   :  0.00   New York     :128517   New York          :128517
##  1st Qu.:  0.00   West Indies  :  8481   China             :  4116
##  Median :  0.00   China        :  4964   Dominican Republic:  3517
##  Mean   : 44.75   SOUTH AMERICA:  4957   Pennsylvania      :  3303
##  3rd Qu.:  0.00   India        :  3476   New Jersey        :  3127
##  Max.   :498.00   Pennsylvania :  3303   Puerto Rico       :  2272
##                   (Other)      : 42887   (Other)           : 51733
##                      ANCESTR1
##  Not Reported            :32021
##  Italian                 :20577
##  Irish, various subheads,:16388
##  German                  :12781
##  African-American        : 9559
##  United States           : 8209
##  (Other)                 :97050
##                                    ANCESTR1D             ANCESTR2
##  Not Reported                           :32021   Not Reported:141487
##  Italian (1990-2000, ACS, PRCS)         :20577   German      :  9476
##  Irish                                  :15651   Irish       :  9238
##  German (1990-2000, ACS/PRCS)           :12605   English     :  4895
##  African-American (1990-2000, ACS, PRCS): 9559   Italian     :  4531
##  United States                          : 8209   Polish      :  3113
##  (Other)                                :97963   (Other)     : 23845
##                            ANCESTR2D         CITIZEN          YRSUSA1
##  Not Reported                  :141487   Min.   :0.0000   Min.   : 0.000
##  German (1990-2000, ACS, PRCS) :  9441   1st Qu.:0.0000   1st Qu.: 0.000
##  Irish                         :  8809   Median :0.0000   Median : 0.000
##  English                       :  4895   Mean   :0.4793   Mean   : 5.377
##  Italian (1990-2000, ACS, PRCS):  4531   3rd Qu.:0.0000   3rd Qu.: 0.000
##  Polish                        :  3113   Max.   :3.0000   Max.   :92.000
##  (Other)                       : 24309
##     HCOVANY         HCOVPRIV         SEX            EMPSTAT
##  Min.   :1.000   Min.   :1.000   Male  : 95222   Min.   :0.000
##  1st Qu.:2.000   1st Qu.:1.000   Female:101363   1st Qu.:1.000
##  Median :2.000   Median :2.000                   Median :1.000
##  Mean   :1.951   Mean   :1.691                   Mean   :1.514
##  3rd Qu.:2.000   3rd Qu.:2.000                   3rd Qu.:3.000
##  Max.   :2.000   Max.   :2.000                   Max.   :3.000
##
##     EMPSTATD        LABFORCE          OCC              IND
##  Min.   : 0.00   Min.   :0.000   0      : 79987   0      :79987
##  1st Qu.:10.00   1st Qu.:1.000   2310   :  3494   7860   : 9025
##  Median :10.00   Median :2.000   5700   :  3235   8680   : 6354
##  Mean   :15.16   Mean   :1.331   430    :  3025   770    : 6279
##  3rd Qu.:30.00   3rd Qu.:2.000   4720   :  2666   8190   : 5873
##  Max.   :30.00   Max.   :2.000   4760   :  2563   7870   : 4041
##                                  (Other):101615   (Other):85026
##     CLASSWKR       CLASSWKRD        WKSWORK2        UHRSWORK
##  Min.   :0.000   Min.   : 0.00   Min.   :0.000   Min.   : 0.00
##  1st Qu.:0.000   1st Qu.: 0.00   1st Qu.:0.000   1st Qu.: 0.00
##  Median :2.000   Median :22.00   Median :1.000   Median :12.00
##  Mean   :1.116   Mean   :13.03   Mean   :2.701   Mean   :19.77
```

```
##  3rd Qu.:2.000   3rd Qu.:22.00   3rd Qu.:6.000   3rd Qu.:40.00
##  Max.   :2.000   Max.   :29.00   Max.   :6.000   Max.   :99.00
##
##      INCTOT           FTOTINC           INCWAGE          POVERTY
##  Min.   :  -7300   Min.   : -11800   Min.   :     0   Min.   :  0.0
##  1st Qu.:   8000   1st Qu.:  35550   1st Qu.:     0   1st Qu.:159.0
##  Median :  25000   Median :  74000   Median : 10000   Median :351.0
##  Mean   :  45245   Mean   : 107110   Mean   : 33796   Mean   :318.7
##  3rd Qu.:  56500   3rd Qu.: 132438   3rd Qu.: 47000   3rd Qu.:501.0
##  Max.   :1563000   Max.   :2030000   Max.   :638000   Max.   :501.0
##  NA's   :31129     NA's   :10817     NA's   :33427
##     MIGRATE1       MIGRATE1D        MIGPLAC1         MIGCOUNTY1
##  Min.   :0.000   Min.   : 0.00   Min.   :  0.000   Min.   :  0.000
##  1st Qu.:1.000   1st Qu.:10.00   1st Qu.:  0.000   1st Qu.:  0.000
##  Median :1.000   Median :10.00   Median :  0.000   Median :  0.000
##  Mean   :1.122   Mean   :11.51   Mean   :  6.184   Mean   :  4.117
##  3rd Qu.:1.000   3rd Qu.:10.00   3rd Qu.:  0.000   3rd Qu.:  0.000
##  Max.   :4.000   Max.   :40.00   Max.   :900.000   Max.   :810.000
##
##     MIGPUMA1        VETSTAT          VETSTATD         PWPUMA00
##  Min.   :    0   Min.   :0.0000   Min.   : 0.000   Min.   :    0
##  1st Qu.:    0   1st Qu.:1.0000   1st Qu.:11.000   1st Qu.:    0
##  Median :    0   Median :1.0000   Median :11.000   Median :    0
##  Mean   :  277   Mean   :0.8621   Mean   : 9.412   Mean   : 1255
##  3rd Qu.:    0   3rd Qu.:1.0000   3rd Qu.:11.000   3rd Qu.: 3100
##  Max.   :70100   Max.   :2.0000   Max.   :20.000   Max.   :59300
##
##     TRANWORK         TRANTIME         DEPARTS           in_NYC
##  Min.   : 0.000   Min.   :  0.00   Min.   :   0.0   Min.   :0.0000
##  1st Qu.: 0.000   1st Qu.:  0.00   1st Qu.:   0.0   1st Qu.:0.0000
##  Median : 0.000   Median :  0.00   Median :   0.0   Median :0.0000
##  Mean   : 9.725   Mean   : 14.75   Mean   : 373.3   Mean   :0.3615
##  3rd Qu.:10.000   3rd Qu.: 20.00   3rd Qu.: 732.0   3rd Qu.:1.0000
##  Max.   :70.000   Max.   :138.00   Max.   :2345.0   Max.   :1.0000
##
##     in_Bronx       in_Manhattan       in_StatenI       in_Brooklyn
##  Min.   :0.0000   Min.   :0.00000   Min.   :0.00000   Min.   :0.000
##  1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.000
##  Median :0.0000   Median :0.00000   Median :0.00000   Median :0.000
##  Mean   :0.0538   Mean   :0.04981   Mean   :0.02084   Mean   :0.126
##  3rd Qu.:0.0000   3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:0.000
##  Max.   :1.0000   Max.   :1.00000   Max.   :1.00000   Max.   :1.000
##
##    in_Queens      in_Westchester      in_Nassau          Hispanic
##  Min.   :0.0000   Min.   :0.00000   Min.   :0.00000   Min.   :0.0000
##  1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.0000
##  Median :0.0000   Median :0.00000   Median :0.00000   Median :0.0000
##  Mean   :0.1111   Mean   :0.04413   Mean   :0.07032   Mean   :0.1387
##  3rd Qu.:0.0000   3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:0.0000
##  Max.   :1.0000   Max.   :1.00000   Max.   :1.00000   Max.   :1.0000
##
##     Hisp_Mex          Hisp_PR         Hisp_Cuban         Hisp_DomR
##  Min.   :0.00000   Min.   :0.0000   Min.   :0.000000   Min.   :0.00000
##  1st Qu.:0.00000   1st Qu.:0.0000   1st Qu.:0.000000   1st Qu.:0.00000
```

```
##  Median :0.00000   Median :0.0000   Median :0.000000   Median :0.00000
##  Mean   :0.01626   Mean   :0.0436   Mean   :0.003403   Mean   :0.02827
##  3rd Qu.:0.00000   3rd Qu.:0.0000   3rd Qu.:0.000000   3rd Qu.:0.00000
##  Max.   :1.00000   Max.   :1.0000   Max.   :1.000000   Max.   :1.00000
##
##      white             AfAm           Amindian             Asian
##  Min.   :0.0000   Min.   :0.000   Min.   :0.000000   Min.   :0.00000
##  1st Qu.:0.0000   1st Qu.:0.000   1st Qu.:0.000000   1st Qu.:0.00000
##  Median :1.0000   Median :0.000   Median :0.000000   Median :0.00000
##  Mean   :0.6997   Mean   :0.125   Mean   :0.003779   Mean   :0.08656
##  3rd Qu.:1.0000   3rd Qu.:0.000   3rd Qu.:0.000000   3rd Qu.:0.00000
##  Max.   :1.0000   Max.   :1.000   Max.   :1.000000   Max.   :1.00000
##
##     race_oth        unmarried       veteran        has_AnyHealthIns
##  Min.   :0.0000   Min.   :0.00   Min.   :0.00000   Min.   :0.0000
##  1st Qu.:0.0000   1st Qu.:0.00   1st Qu.:0.00000   1st Qu.:1.0000
##  Median :0.0000   Median :0.00   Median :0.00000   Median :1.0000
##  Mean   :0.1324   Mean   :0.45   Mean   :0.04443   Mean   :0.9513
##  3rd Qu.:0.0000   3rd Qu.:1.00   3rd Qu.:0.00000   3rd Qu.:1.0000
##  Max.   :1.0000   Max.   :1.00   Max.   :1.00000   Max.   :1.0000
##
##  has_PvtHealthIns  Commute_car      Commute_bus      Commute_subway
##  Min.   :0.0000   Min.   :0.0000   Min.   :0.00000   Min.   :0.00000
##  1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.:0.00000
##  Median :1.0000   Median :0.0000   Median :0.00000   Median :0.00000
##  Mean   :0.6906   Mean   :0.2997   Mean   :0.02162   Mean   :0.07468
##  3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:0.00000   3rd Qu.:0.00000
##  Max.   :1.0000   Max.   :1.0000   Max.   :1.00000   Max.   :1.00000
##
##   Commute_rail     Commute_other     below_povertyline below_150poverty
##  Min.   :0.00000   Min.   :0.00000   Min.   :0.000     Min.   :0.0000
##  1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.000     1st Qu.:0.0000
##  Median :0.00000   Median :0.00000   Median :0.000     Median :0.0000
##  Mean   :0.01332   Mean   :0.05506   Mean   :0.122     Mean   :0.1965
##  3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:0.000     3rd Qu.:0.0000
##  Max.   :1.00000   Max.   :1.00000   Max.   :1.000     Max.   :1.0000
##
##  below_200poverty   foodstamps
##  Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.0000   1st Qu.:0.0000
##  Median :0.0000   Median :0.0000
##  Mean   :0.2676   Mean   :0.1465
##  3rd Qu.:1.0000   3rd Qu.:0.0000
##  Max.   :1.0000   Max.   :1.0000
##
```

```
print(NN_obs <- length(AGE))
```

```
## [1] 196585
```

```
summary(AGE[female == 1])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.00   23.00   44.00   42.72   61.00   95.00
```

```
summary(AGE[!female])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.00   21.00   40.00   40.35   59.00   95.00
```

```
mean(AGE[female == 1])
```

```
## [1] 42.71629
```

```
sd(AGE[female == 1])
```

```
## [1] 23.72012
```

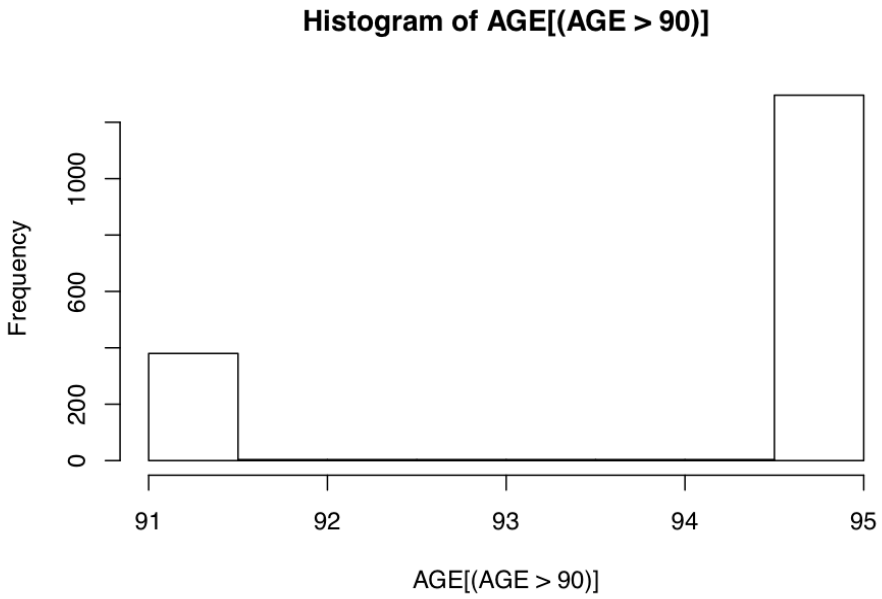
```
mean(AGE[!female])
```

```
## [1] 40.35398
```

```
sd(AGE[!female])
```

```
## [1] 23.1098
```

```
hist(AGE[(AGE > 90)])
```




```
mean(AGE[ (female == 1) & (AGE<90) ])
```

```
## [1] 41.98866
```

```
str(as.numeric(PUMA))
```

```
##  num [1:196585] 902 902 4002 4002 3803 ...
```

```
PUMA <- as.factor(PUMA)
female <- as.factor(female)
```

First the quick and dirty way.

```
print(levels(female))
```

```
## [1] "0" "1"
```

```
levels(female) <- c("male","female")
```

```
educ_indx <- factor((educ_nohs + 2*educ_hs + 3*educ_somecoll + 4*educ_college + 5*educ_advdeg), levels=
```

```
summary(female)
```

```
##   male female
##  95222 101363
```

```
summary(PUMA)
```

```
##    2500     200    4112    4009    4103    4017     500     403    2402
##    3051    2624    2438    2422    2290    2234    2214    2202    2172
##    4105     704    4110    2600    2002    2203     703    1500    1400
##    2148    2043    2040    2039    2030    1950    1945    1914    1909
##    4016    1300     800     100    1207     402    3107     906    4101
##    1903    1885    1871    1819    1787    1774    1762    1752    1748
##     300    1204    4014    1900    4111     701    3105     600    3903
##    1724    1701    1698    1694    1694    1692    1679    1670    1621
##    2802     401    3201    2801    1000    2702    4013    3106    1600
##    1620    1597    1543    1542    1536    1533    1519    1518    1510
##    4015    1700    2100    4106    3202    2202    1802    4104    3102
##    1479    1407    1404    1393    1376    1358    1356    1344    1341
##    1205    3305    2903    4008    4004    4102    2201    3705    3709
##    1333    1326    1323    1321    1320    1316    1311    1311    1307
##    4113    3901    4001    4012     901    3103    3303    3304    3204
##    1304    1303    1293    1291    1288    1283    1279    1268    1267
##     702    4010    3801    3003    3205    3206    2902    3208    3902
##    1265    1250    1238    1232    1215    1207    1195    1188    1173
##    4005    3805    3710    4109    1102     904    3810    3308     902
##    1168    1167    1166    1158    1153    1141    1136    1131    1126
##    1202    2701    3704    1206    3707    2401    2300     905    3306
##    1122    1118    1115    1112    1112    1109    1105    1099    1097
## (Other)
##   44328
```

```
summary(educ_indx)
```

```
##  No HS     HS SmColl   Bach    Adv
##  53267  55119  34012  30802  23385
```

```
#ddply(acs2017_ny, .(PUMA), summarize, mean = round(mean(AGE), 2), sd = round(sd(AGE), 2))
```

Here's the 90th and 10th percentiles of wages by neighborhood,

```
#dat_use1 <- subset(acs2017_ny,((INCWAGE > 0) & in_NYC))
#ddply(dat_use1, .(PUMA), summarize, inc90 = quantile(INCWAGE,probs = 0.9), inc10 = quantile(INCWAGE,pr
summary(educ_indx)
```

```
##  No HS     HS SmColl   Bach    Adv
##  53267  55119  34012  30802  23385
```

```r
table(educ_indx,female)
```

```
##          female
## educ_indx  male female
##    No HS  27180  26087
##    HS     27309  27810
##    SmColl 15847  18165
##    Bach   14632  16170
##    Adv    10254  13131
```

```r
xtabs(~educ_indx + female)
```

```
##          female
## educ_indx  male female
##    No HS  27180  26087
##    HS     27309  27810
##    SmColl 15847  18165
##    Bach   14632  16170
##    Adv    10254  13131
```

Want proportions instead of counts?

```r
prop.table(table(educ_indx,female))
```

```
##          female
## educ_indx       male     female
##    No HS  0.13826080 0.13270087
##    HS     0.13891701 0.14146552
##    SmColl 0.08061144 0.09240278
##    Bach   0.07443091 0.08225450
##    Adv    0.05216064 0.06679553
```

```r
mean(educ_nohs[(AGE >= 25)&(AGE <= 55)])
```

```
## [1] 0.08354656
```

```r
mean(educ_hs[(AGE >= 25)&(AGE <= 55)])
```

```
## [1] 0.2974594
```

```r
mean(educ_somecoll[(AGE >= 25)&(AGE <= 55)])
```

```
## [1] 0.2057843
```

```r
mean(educ_college[(AGE >= 25)&(AGE <= 55)])
```

```
## [1] 0.2383112
```

```r
mean(educ_advdeg[(AGE >= 25)&(AGE <= 55)])
```

```
## [1] 0.1748986
```

```r
# alternatively
restrict1 <- as.logical((AGE >= 25)&(AGE <= 55))
dat_age_primeage <- subset(acs2017_ny, restrict1)

detach()
```

```r
attach(dat_age_primeage)

mean(educ_nohs)
```

```
## [1] 0.08354656
```

```r
mean(educ_hs)
```

```
## [1] 0.2974594
```

```r
mean(educ_somecoll)
```

```
## [1] 0.2057843
```

```r
mean(educ_college)
```

```
## [1] 0.2383112
```

```r
mean(educ_advdeg)
```

```
## [1] 0.1748986
```

```r
detach()
```

the education level between female and male seems to be equally distributed.