logitudinal_final_proj

Lynette Pan

12/3/2018

Final Project: Smoking in the Framingham Heart Study

Descriptive statistics

```
##     cursmoke          randid        sex         totchol
##  Min.   :0.0000   Min.   :   2448   1:5022   Min.   :107.0
##  1st Qu.:0.0000   1st Qu.:2474378   2:6605   1st Qu.:210.0
##  Median :0.0000   Median :5006008            Median :238.0
##  Mean   :0.4325   Mean   :5004741            Mean   :241.2
##  3rd Qu.:1.0000   3rd Qu.:7472730            3rd Qu.:268.0
##  Max.   :1.0000   Max.   :9999312            Max.   :696.0
##                                              NA's   :409
##       age             sysbp           diabp            cigpday
##  Min.   :32.00   Min.   : 83.5   Min.   : 30.00   Min.   : 0.00
##  1st Qu.:48.00   1st Qu.:120.0   1st Qu.: 75.00   1st Qu.: 0.00
##  Median :54.00   Median :132.0   Median : 82.00   Median : 0.00
##  Mean   :54.79   Mean   :136.3   Mean   : 83.04   Mean   : 8.25
##  3rd Qu.:62.00   3rd Qu.:149.0   3rd Qu.: 90.00   3rd Qu.:20.00
##  Max.   :81.00   Max.   :295.0   Max.   :150.00   Max.   :90.00
##                                                   NA's   :79
##       bmi         diabetes   bpmeds          heartrte         glucose
##  Min.   :14.43   0:11097   0   :10090   Min.   : 37.00   Min.   : 39.00
##  1st Qu.:23.09   1:  530   1   :  944   1st Qu.: 69.00   1st Qu.: 72.00
##  Median :25.48             NA's:  593   Median : 75.00   Median : 80.00
##  Mean   :25.88                          Mean   : 76.78   Mean   : 84.12
##  3rd Qu.:28.07                          3rd Qu.: 85.00   3rd Qu.: 89.00
##  Max.   :56.80                          Max.   :220.00   Max.   :478.00
##  NA's   :52                             NA's   :6        NA's   :1440
##    educ       prevchd   prevap    prevmi    prevstrk  prevhyp
##  1   :4690   0:10785   0:11000   0:11253   0:11475   0:6283
##  2   :3410   1:  842   1:  627   1:  374   1:  152   1:5344
##  3   :1885
##  4   :1347
##  NA's: 295
##
##
##       time           period           hdlc             ldlc
##  Min.   :   0   Min.   :1.000   Min.   : 10.00   Min.   : 20.0
##  1st Qu.:   0   1st Qu.:1.000   1st Qu.: 39.00   1st Qu.:145.0
##  Median :2156   Median :2.000   Median : 48.00   Median :173.0
##  Mean   :1957   Mean   :1.899   Mean   : 49.37   Mean   :176.5
##  3rd Qu.:4252   3rd Qu.:3.000   3rd Qu.: 58.00   3rd Qu.:205.0
##  Max.   :4854   Max.   :3.000   Max.   :189.00   Max.   :565.0
##                                 NA's   :8600     NA's   :8601
##      death            angina          hospmi          mi_fchd
```

```
##  Min.   :0.0000   Min.   :0.0000   Min.   :0.00000   Min.   :0.0000
##  1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.:0.0000
##  Median :0.0000   Median :0.0000   Median :0.00000   Median :0.0000
##  Mean   :0.3033   Mean   :0.1636   Mean   :0.09925   Mean   :0.1538
##  3rd Qu.:1.0000   3rd Qu.:0.0000   3rd Qu.:0.00000   3rd Qu.:0.0000
##  Max.   :1.0000   Max.   :1.0000   Max.   :1.00000   Max.   :1.0000
##
##      anychd           stroke             cvd            hyperten
##  Min.   :0.0000   Min.   :0.00000   Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.:0.0000   1st Qu.:0.0000
##  Median :0.0000   Median :0.00000   Median :0.0000   Median :1.0000
##  Mean   :0.2716   Mean   :0.09125   Mean   :0.2493   Mean   :0.7433
##  3rd Qu.:1.0000   3rd Qu.:0.00000   3rd Qu.:0.0000   3rd Qu.:1.0000
##  Max.   :1.0000   Max.   :1.00000   Max.   :1.0000   Max.   :1.0000
##
##      timeap         timemi        timemifc       timechd
##  Min.   :   0   Min.   :   0   Min.   :   0   Min.   :   0
##  1st Qu.:6224   1st Qu.:7212   1st Qu.:7050   1st Qu.:5598
##  Median :8766   Median :8766   Median :8766   Median :8766
##  Mean   :7242   Mean   :7594   Mean   :7543   Mean   :7008
##  3rd Qu.:8766   3rd Qu.:8766   3rd Qu.:8766   3rd Qu.:8766
##  Max.   :8766   Max.   :8766   Max.   :8766   Max.   :8766
##
##     timestrk       timecvd        timedth        timehyp      age_ctg
##  Min.   :   0   Min.   :   0   Min.   :  26   Min.   :   0   35-  :  24
##  1st Qu.:7295   1st Qu.:6004   1st Qu.:7798   1st Qu.:   0   35-44:1785
##  Median :8766   Median :8766   Median :8766   Median :2429   45-54:4095
##  Mean   :7661   Mean   :7166   Mean   :7854   Mean   :3599   55-64:3701
##  3rd Qu.:8766   3rd Qu.:8766   3rd Qu.:8766   3rd Qu.:7329   65-74:1819
##  Max.   :8766   Max.   :8766   Max.   :8766   Max.   :8766   75+  : 203
##
##     sysbp_ctg                  diabp_ctg
##  normal  :2746   normal and elevated:4367
##  crisis  : 529   crisis             :  54
##  elevated:2351   stage1             :4180
##  stage1  :2066   stage2             :3026
##  stage2  :3935
##
##
```

missing values

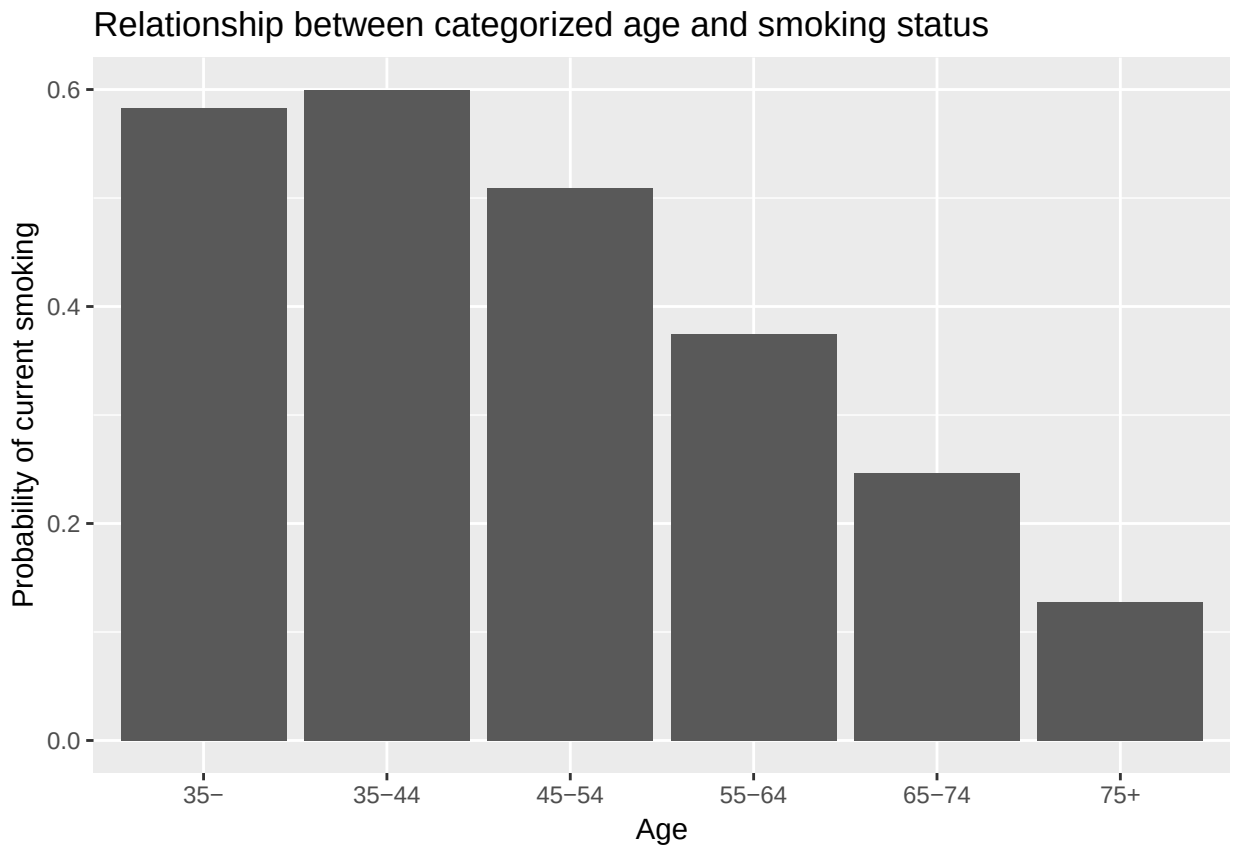
Impute missing values using MICE package

- The output tells us that 2243 samples are complete, 7077 samples miss both hdlc and ldlc, 4 samples miss only the glucose value and so on.
- there are 447 patients who have only one obs; there are 781 patients who have only two obs and 3206 patients have all three obs. In total, there are 4434 patients in the study.

1. Target question:

1) Is there a relationship between age and smoking status? Does this relationship differ by sex?

cursmoke vs age_ctg without adjustment



Finding confounders

age and cursmoke – confounder: hdlc, ldlc

```
## $totchol
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.875  FALSE
##
## $sysbp
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.879  FALSE
##
```

```
## $diabp
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.875  FALSE
##
## $cigpday
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.875  FALSE
##
## $bmi
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.877  FALSE
##
## $diabetes
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.875  FALSE
##
## $bpmeds
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.877  FALSE
##
## $heartrte
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.872  FALSE
##
## $glucose
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.882  FALSE
##
## $educ
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
```

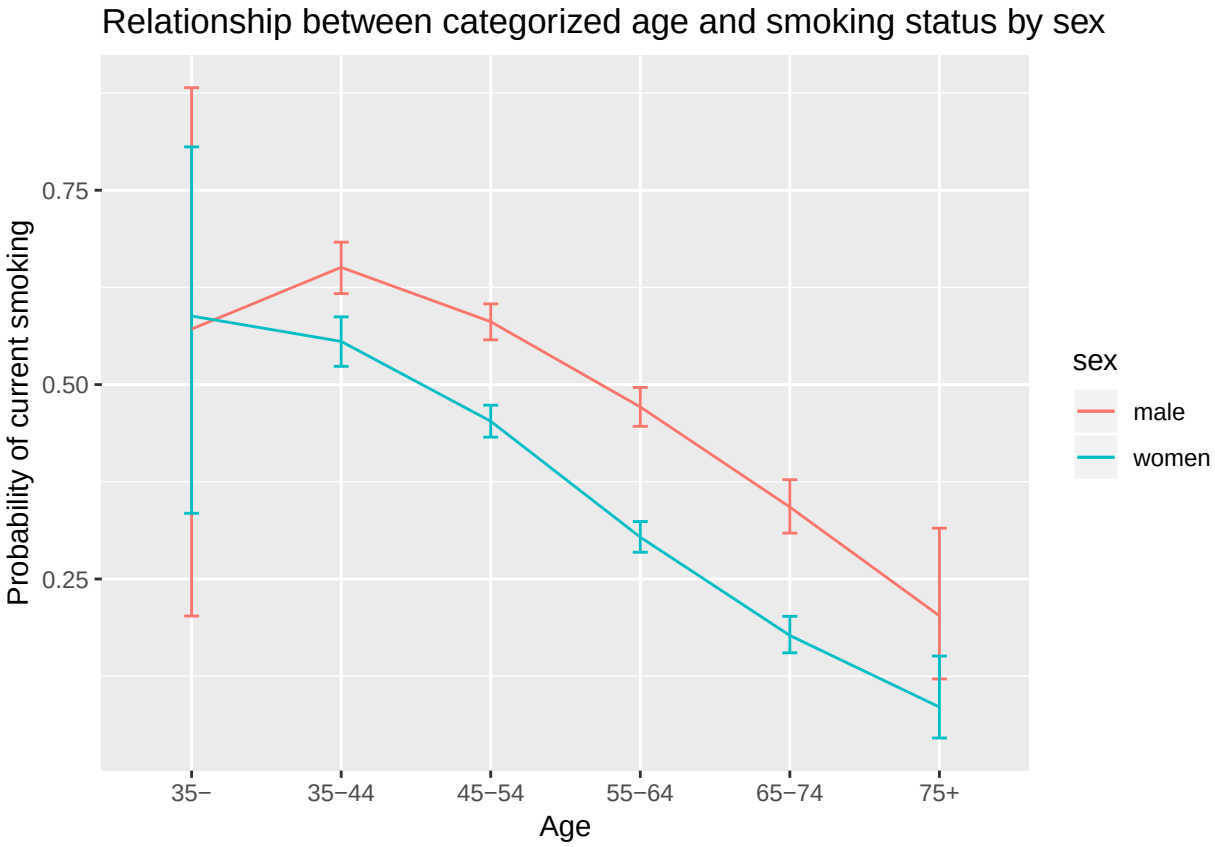
```
## age         0.874      0.864      0.885        0.873  FALSE
##
## $prevchd
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.876  FALSE
##
## $prevap
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.876  FALSE
##
## $prevmi
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.875  FALSE
##
## $prevstrk
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.874  FALSE
##
## $prevhyp
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885         0.88  FALSE
##
## $period
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.903  TRUE
##
## $hdlc
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.935  TRUE
##
## $ldlc
##
##
```

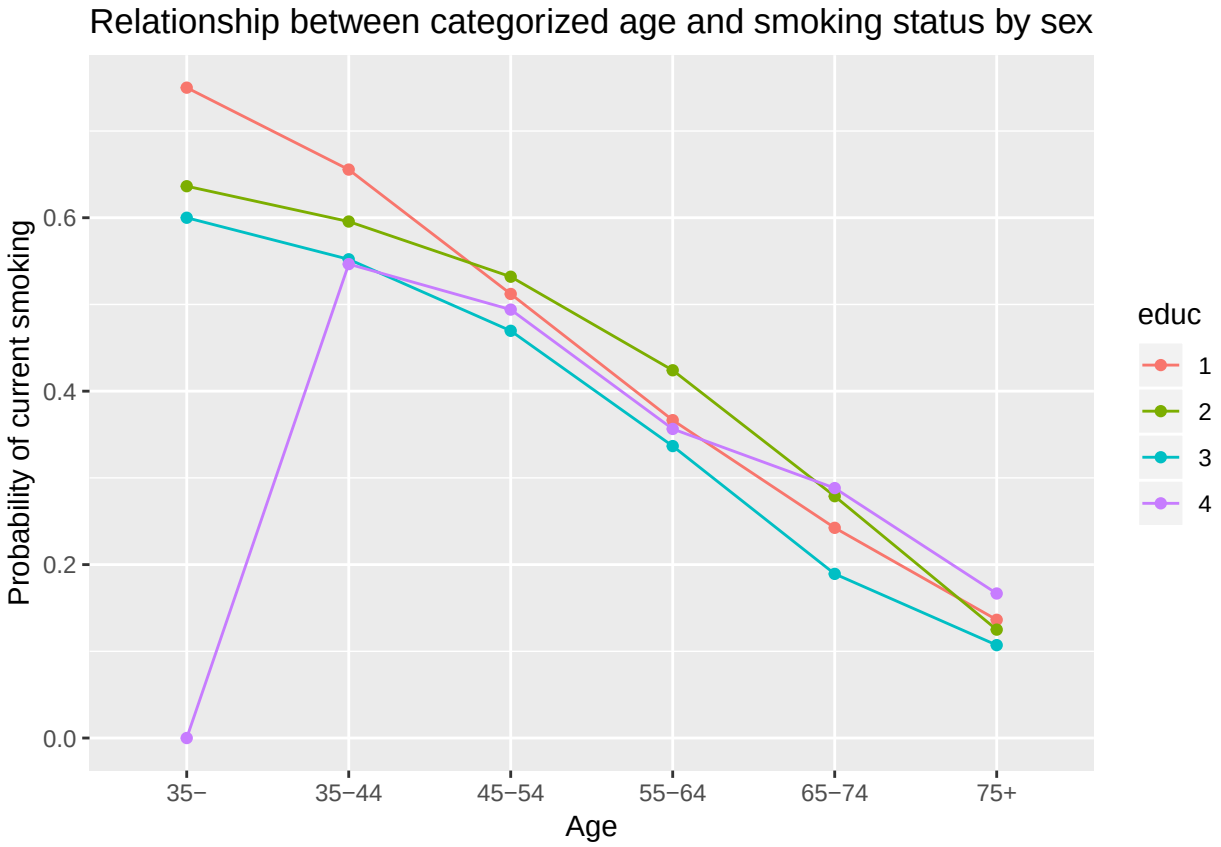
```
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.935  TRUE
##
## $death
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.867  FALSE
##
## $angina
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.874  FALSE
##
## $hospmi
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.874  FALSE
##
## $mi_fchd
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.873  FALSE
##
## $anychd
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.873  FALSE
##
## $stroke
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.873  FALSE
##
## $cvd
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.871  FALSE
##
## $hyperten
```

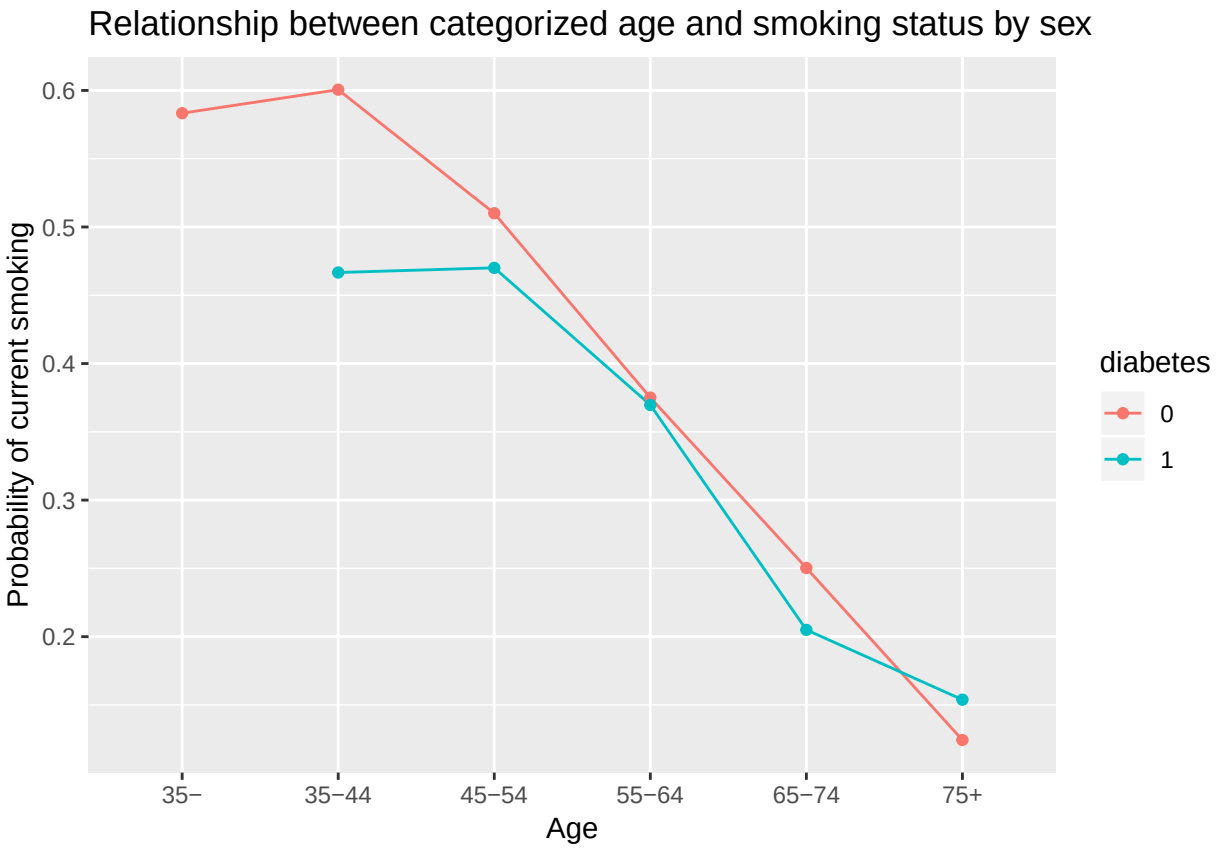
```
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.876  FALSE
##
## $age_ctg
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.861  TRUE
##
## $sysbp_ctg
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.879  FALSE
##
## $diabp_ctg
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.874      0.864      0.885        0.875  FALSE
```

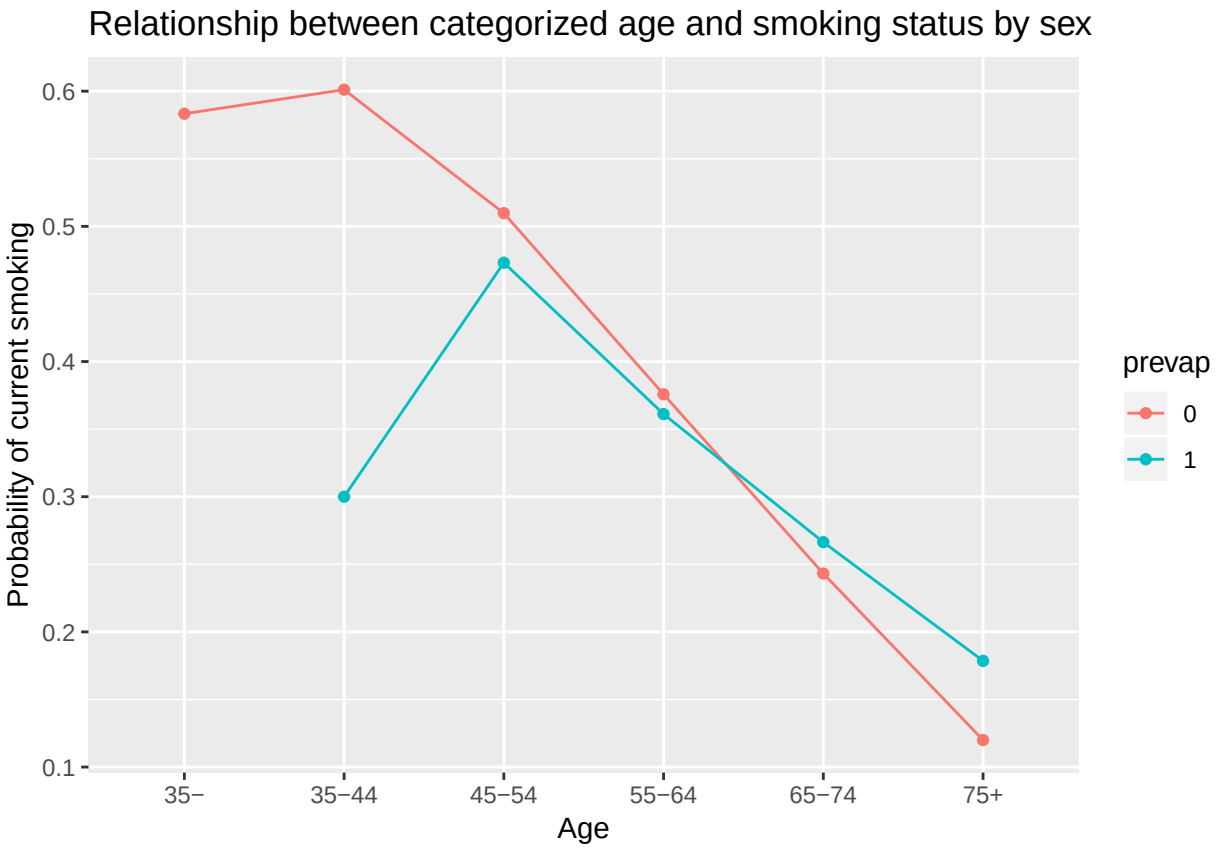
Finding modifier

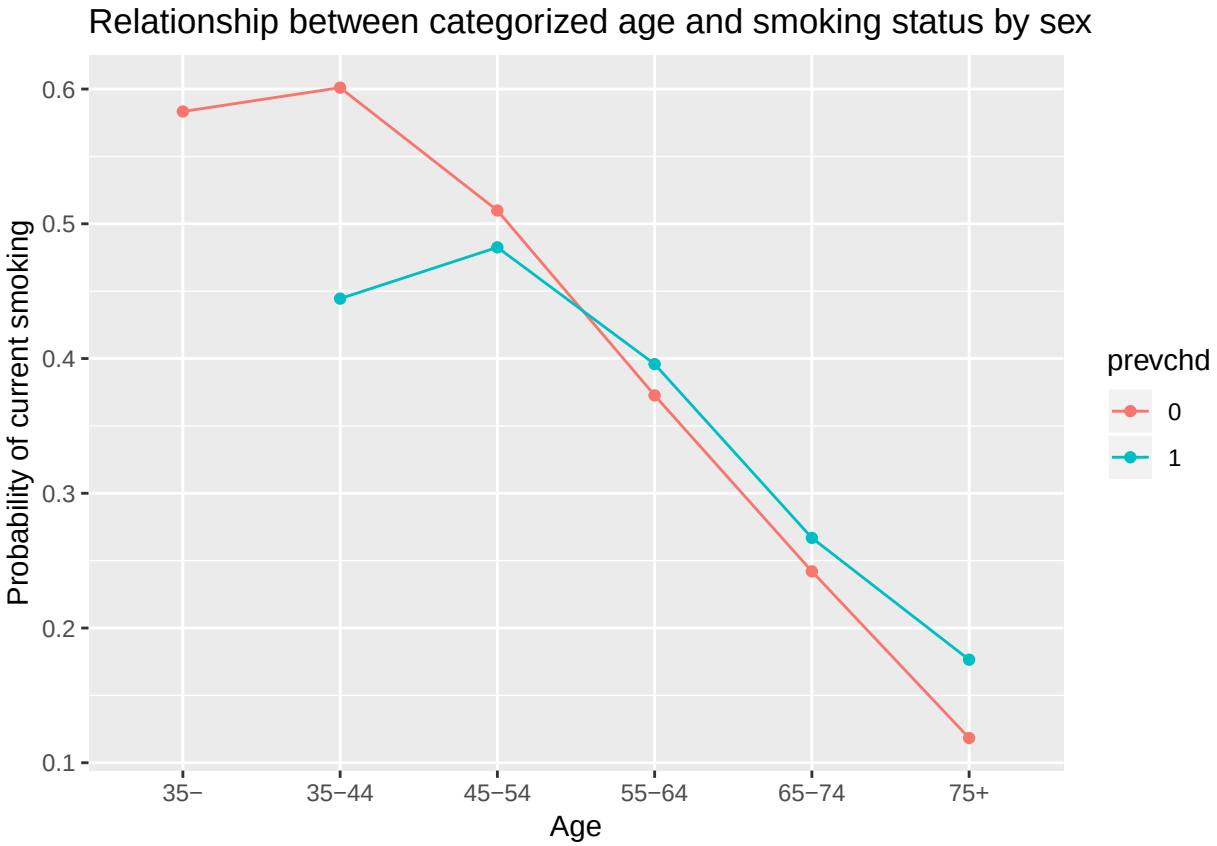
```
## Warning in prop.test(4L, 7L): Chi-squared approximation may be incorrect

## Warning in prop.test(4L, 7L): Chi-squared approximation may be incorrect
```

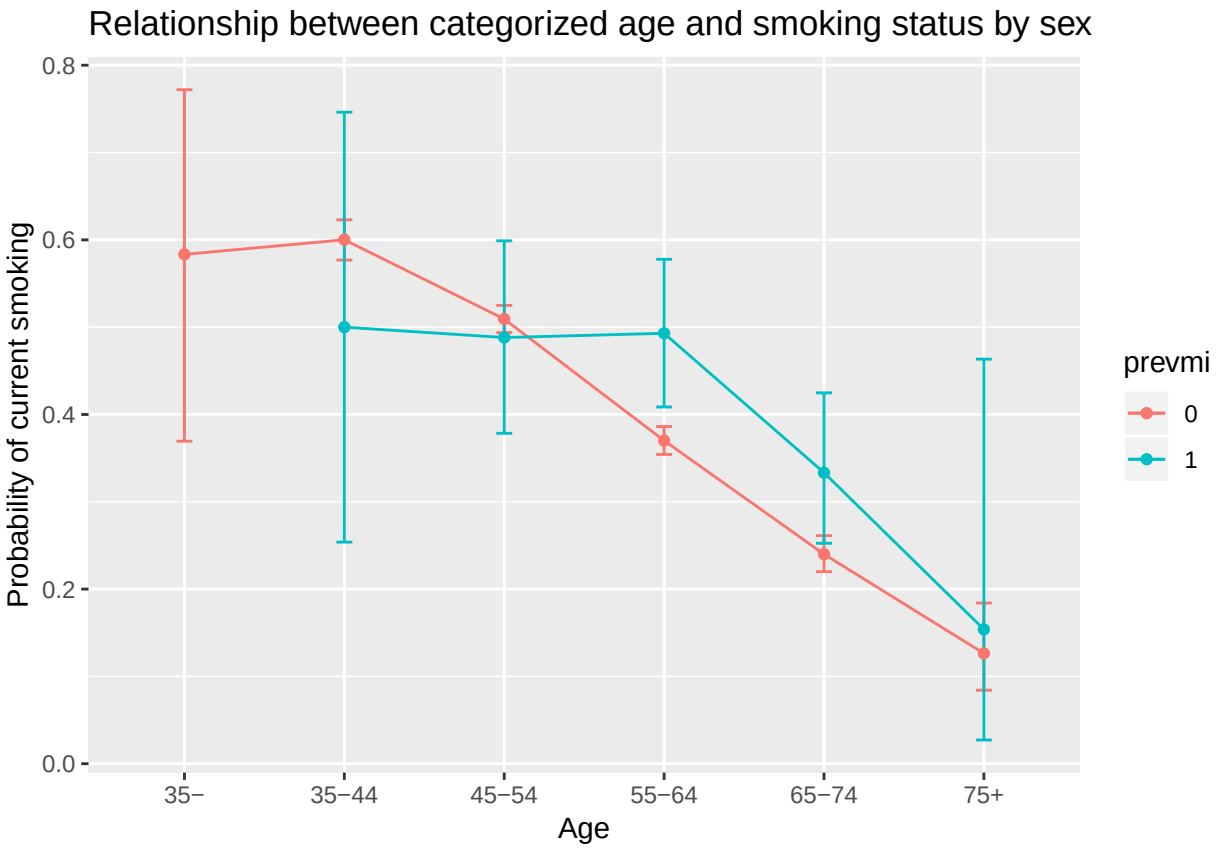

Relationship between categorized age and smoking status by sex
Probability of current smoking
0.75
0.50
0.25
35−
35−44
45−54
55−64
65−74
75+
Age
sex
male
women


Relationship between categorized age and smoking status by sex
Probability of current smoking
0.0
0.2
0.4
0.6
35−
35−44
45−54
55−64
65−74
75+
Age
educ
1
2
3
4

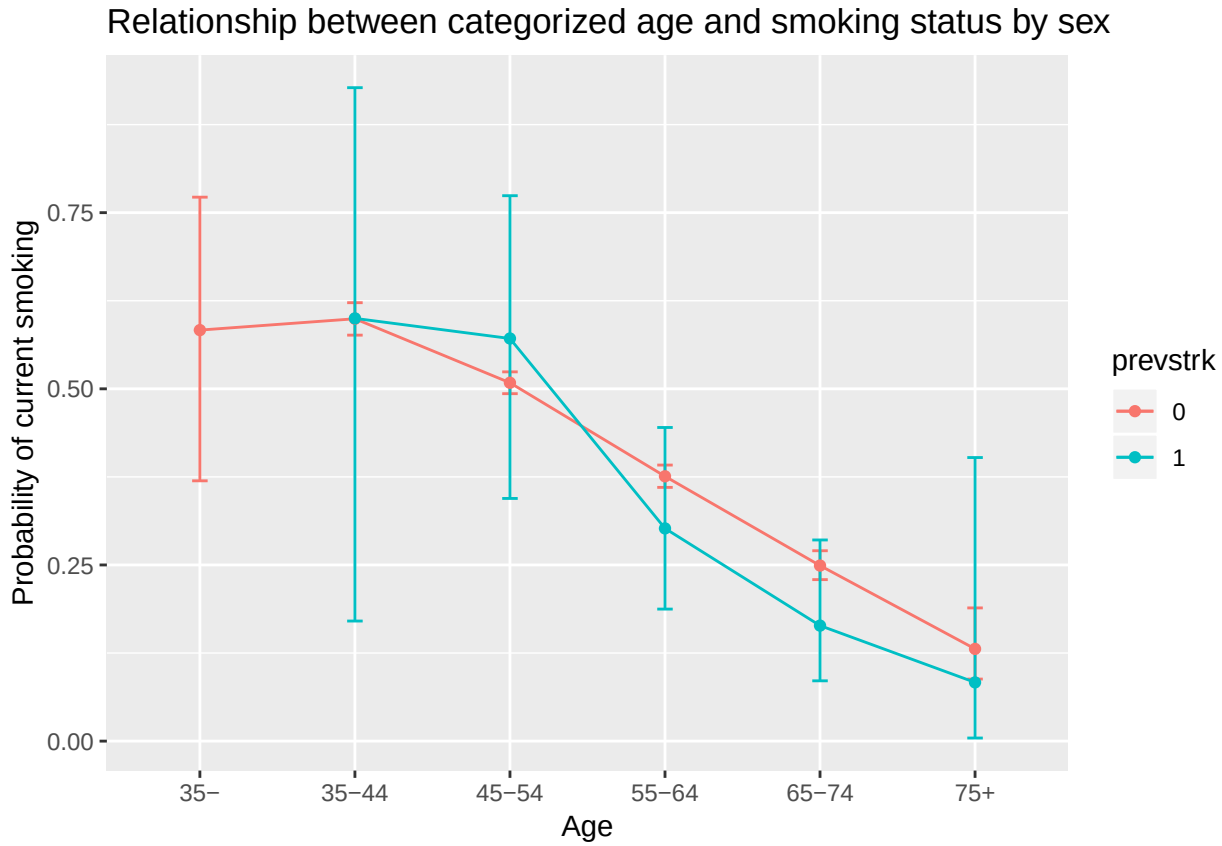

Relationship between categorized age and smoking status by sex
Probability of current smoking
0.6
0.5
0.4
0.3
0.2
35–
35–44
45–54
55–64
65–74
75+
Age
diabetes
0
1


Relationship between categorized age and smoking status by sex
Probability of current smoking
0.6
0.5
0.4
0.3
0.2
0.1
35−
35−44
45−54
55−64
65−74
75+
Age
prevap
0
1

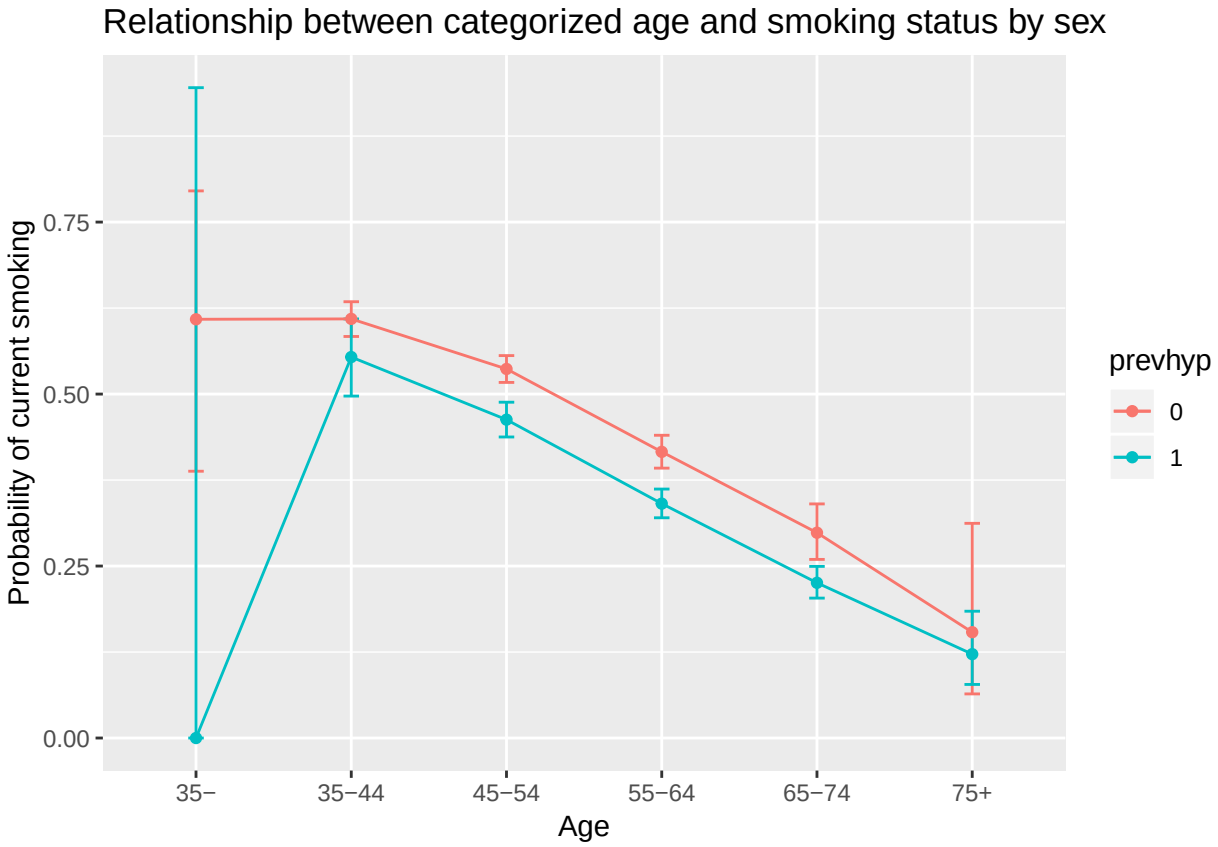

Relationship between categorized age and smoking status by sex
Probability of current smoking
0.6
0.5
0.4
0.3
0.2
0.1
35−
35−44
45−54
55−64
65−74
75+
Age
prevchd
0
1




```
## Warning in prop.test(3L, 5L): Chi-squared approximation may be incorrect
## Warning in prop.test(3L, 5L): Chi-squared approximation may be incorrect
```




```
## Warning in prop.test(0L, 1L): Chi-squared approximation may be incorrect
## Warning in prop.test(0L, 1L): Chi-squared approximation may be incorrect
```



Fit model using glmer

```
## Generalized linear mixed model fit by maximum likelihood (Adaptive
##   Gauss-Hermite Quadrature, nAGQ = 0) [glmerMod]
##  Family: binomial  ( logit )
## Formula: cursmoke ~ age * sex + educ + hdlc + ldlc + age * prevhyp + (1 |
##     randid)
##    Data: frmgham
##
##      AIC      BIC   logLik deviance df.resid
##   3581.6   3653.5  -1778.8   3557.6     2939
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.3319 -0.7483 -0.5126  1.0475  3.1773
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  randid (Intercept) 0.007196 0.08483
## Number of obs: 2951, groups:  randid, 2951
##
## Fixed effects:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)   2.1089335  0.6408587   3.291 0.000999 ***
## age          -0.0410951  0.0100028  -4.108 3.98e-05 ***
```

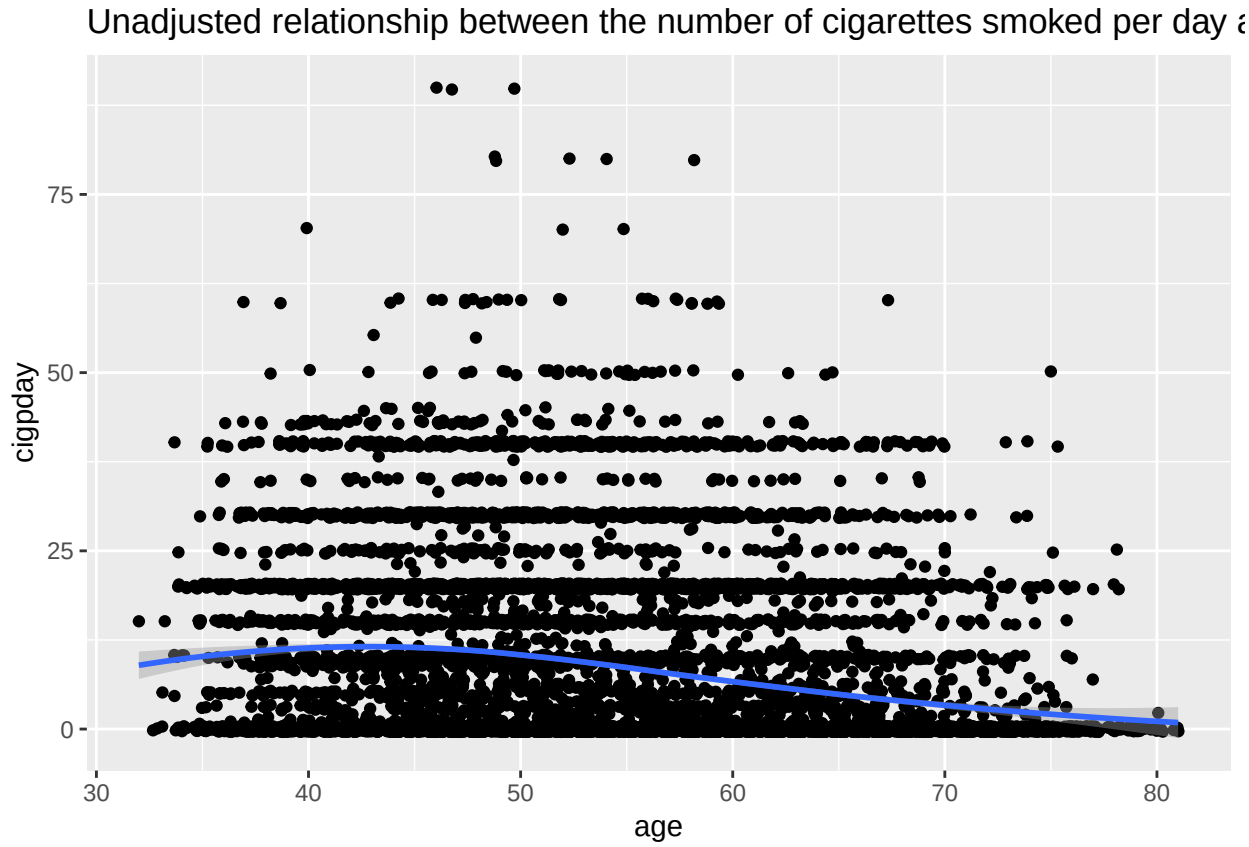
```
## sex2          1.5708819  0.6289735   2.498 0.012506 *
## educ2         0.1879275  0.0983733   1.910 0.056088 .
## educ3        -0.1768304  0.1215055  -1.455 0.145579
## educ4        -0.1589561  0.1331928  -1.193 0.232701
## hdlc         -0.0004737  0.0028068  -0.169 0.865979
## ldlc          0.0005599  0.0008889   0.630 0.528808
## prevhyp1     -0.1490198  0.6473765  -0.230 0.817944
## age:sex2     -0.0321176  0.0105458  -3.046 0.002323 **
## age:prevhyp1 -0.0031512  0.0109453  -0.288 0.773418
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) age    sex2   educ2  educ3  educ4  hdlc   ldlc   prvhy1
## age         -0.934
## sex2        -0.529  0.536
## educ2       -0.159  0.107  0.012
## educ3       -0.054  0.014 -0.031  0.385
## educ4       -0.089  0.043  0.021  0.348  0.273
## hdlc        -0.255  0.022 -0.020 -0.023 -0.043 -0.049
## ldlc        -0.301  0.035  0.072 -0.016 -0.016 -0.013  0.189
## prevhyp1    -0.583  0.643  0.042 -0.028 -0.046 -0.020  0.013 -0.032
## age:sex2     0.529 -0.530 -0.990 -0.016  0.022 -0.008 -0.028 -0.095 -0.028
## age:prvhyp1  0.586 -0.656 -0.027  0.027  0.048  0.027 -0.010  0.025 -0.991
##             ag:sx2
## age
## sex2
## educ2
## educ3
## educ4
## hdlc
## ldlc
## prevhyp1
## age:sex2
## age:prvhyp1  0.013
```

(2) Is there a relationship between the number of cigarettes smoked per day and age? Does this relationship differ by sex?

While answering these questions, please account for any confounders that you have evidence may impact the relationship between age and sex with smoking.

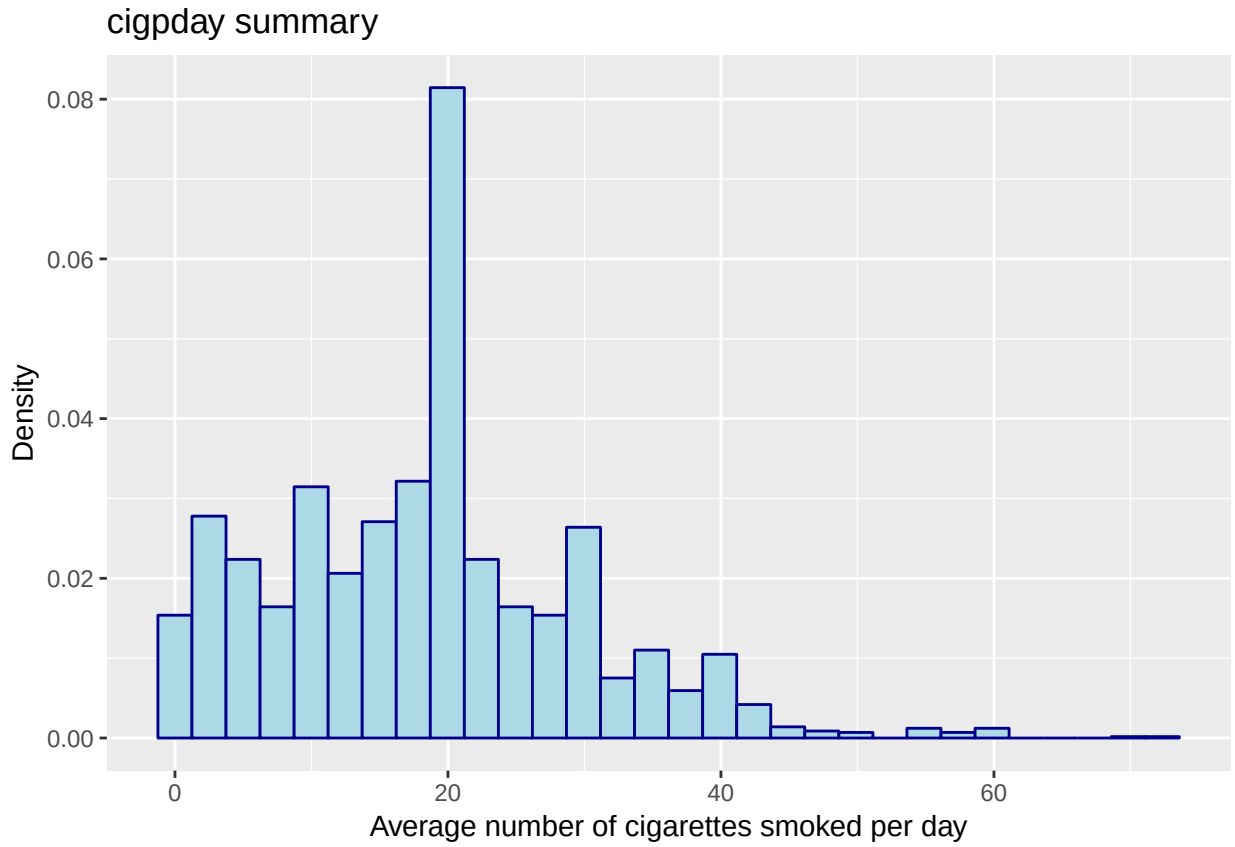
```
## Warning: Removed 79 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 79 rows containing missing values (geom_point).
```




```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.00   10.00   20.00   19.25   25.00   90.00
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27

## running glm to get initial regression estimate

## (Intercept)         age
##  4.05032331 -0.03649173

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

## (Intercept)         age
##  0.82801390 -0.03120521

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##   (Intercept)           age       totchol
##  8.205811e-01 -3.137216e-02  5.639067e-05

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##  (Intercept)          age        sysbp
##  1.066322565 -0.028655440 -0.002785127

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

## (Intercept)         age       diabp
##  1.23996287 -0.03046680 -0.00546731

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
```

```
## running glm to get initial regression estimate

## (Intercept)         age         bmi
##  1.91200020 -0.02985014 -0.04534089

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

## (Intercept)         age   diabetes1
##  0.82084112 -0.03102093 -0.06646395

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

## (Intercept)         age     bpmeds1
##  0.76865830 -0.02980929 -0.20231430

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##  (Intercept)          age     heartrte
##  0.363609848 -0.031496341  0.006206298

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##  (Intercept)          age      glucose
##  0.927633265 -0.030768877 -0.001514277

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

## (Intercept)         age       educ2       educ3       educ4
##  0.82982420 -0.03097717  0.03249024 -0.11339801 -0.04966555

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

## (Intercept)         age    prevchd1
##  0.83742613 -0.03143404  0.04274108

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

## (Intercept)         age     prevap1
##  0.82587661 -0.03115428 -0.01231164

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

## (Intercept)         age     prevmi1
##  0.84424516 -0.03162447  0.19199742

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

## (Intercept)         age   prevstrk1
##  0.82082170 -0.03103646 -0.17419051

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

## (Intercept)         age    prevhyp1
##  0.73776289 -0.02834957 -0.14914763
```

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##  (Intercept)          age       period
##  0.827036125 -0.031078124 -0.003147334

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##  (Intercept)          age         hdlc
##  1.650730480 -0.044188754 -0.002226944

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##   (Intercept)           age          ldlc
##  1.5620275288 -0.0441417089 -0.0001318165

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

## (Intercept)         age       death
##  0.96470363 -0.03544637  0.28952416

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

## (Intercept)         age      angina
##  0.82626015 -0.03109184 -0.02757864

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

## (Intercept)         age      hospmi
##  0.82847349 -0.03169515  0.23878512

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

## (Intercept)         age     mi_fchd
##  0.84568736 -0.03223356  0.22859760

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

## (Intercept)         age      anychd
##  0.84328875 -0.03204868  0.10966852

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

## (Intercept)         age      stroke
##  0.86181547 -0.03211417  0.16410163

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

## (Intercept)         age         cvd
##  0.88338256 -0.03325683  0.21247220

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate
```

```
## (Intercept)         age    hyperten
##  0.86000759 -0.02965192 -0.16074355

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##  (Intercept)          age age_ctg35-44 age_ctg45-54 age_ctg55-64
##   0.48208732  -0.03029572   0.23914998   0.34671138   0.32978877
## age_ctg65-74   age_ctg75+
##   0.18781581  -0.22088289

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##       (Intercept)               age   sysbp_ctgcrisis sysbp_ctgelevated
##        0.80170439       -0.02876858       -0.21640033       -0.13036911
##   sysbp_ctgstage1   sysbp_ctgstage2
##       -0.10525143       -0.15981715

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##     (Intercept)             age diabp_ctgcrisis diabp_ctgstage1
##      0.86310716     -0.03044677     -0.16707868     -0.10021979
## diabp_ctgstage2
##     -0.15973730

## $cursmoke
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $totchol
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $sysbp
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.974  FALSE
##
## $diabp
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.974  FALSE
##
## $bmi
##
```

```
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.974  FALSE
##
## $diabetes
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $bpmeds
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.974  FALSE
##
## $heartrte
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  TRUE
##
## $glucose
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $educ
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $prevchd
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.974  FALSE
##
## $prevap
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.974  FALSE
##
```

```
## $prevmi
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $prevstrk
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $prevhyp
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.975  FALSE
##
## $period
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.972  TRUE
##
## $hdlc
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.957  TRUE
##
## $ldlc
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.957  TRUE
##
## $death
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.971  TRUE
##
## $angina
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
```

```
## age         0.975      0.973      0.978        0.973  FALSE
##
## $hospmi
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $mi_fchd
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $anychd
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $stroke
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $cvd
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.972  TRUE
##
## $hyperten
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.974  FALSE
##
## $age_ctg
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.972  TRUE
##
## $sysbp_ctg
##
##
```

```
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.974  FALSE
##
## $diabp_ctg
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.974  FALSE
## $cursmoke
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $totchol
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $sysbp
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.974  FALSE
##
## $diabp
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.974  FALSE
##
## $bmi
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.974  FALSE
##
## $diabetes
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $bpmeds
```

```
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.974  FALSE
##
## $heartrte
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  TRUE
##
## $glucose
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $educ
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $prevchd
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.974  FALSE
##
## $prevap
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.974  FALSE
##
## $prevmi
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $prevstrk
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
```

```
##
## $prevhyp
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.975  FALSE
##
## $period
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.972  TRUE
##
## $hdlc
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.957  TRUE
##
## $ldlc
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.957  TRUE
##
## $death
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.971  TRUE
##
## $angina
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $hospmi
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $mi_fchd
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
```

```
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $anychd
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $stroke
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.973  FALSE
##
## $cvd
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.972  TRUE
##
## $hyperten
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.974  FALSE
##
## $age_ctg
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.972  TRUE
##
## $sysbp_ctg
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.974  FALSE
##
## $diabp_ctg
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## age         0.975      0.973      0.978        0.974  FALSE
## Generalized linear mixed model fit by maximum likelihood (Adaptive
##   Gauss-Hermite Quadrature, nAGQ = 0) [glmerMod]
```

```
##  Family: poisson  ( log )
## Formula: cigpday ~ age + sex + heartrte + hdlc + ldlc + totchol + (1 |
##     randid)
##    Data: frmgham_smoker
##
##      AIC      BIC   logLik deviance df.resid
##   7876.2   7915.7  -3930.1   7860.2     1013
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.83291 -0.26160  0.05622  0.18987  0.52723
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  randid (Intercept) 0.3593   0.5994
## Number of obs: 1021, groups:  randid, 1021
##
## Fixed effects:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  3.6529114  0.2412169  15.144  < 2e-16 ***
## age         -0.0167559  0.0028804  -5.817 5.99e-09 ***
## sex2        -0.2544544  0.0438405  -5.804 6.47e-09 ***
## heartrte     0.0041263  0.0016404   2.515   0.0119 *
## hdlc        -0.0027549  0.0017678  -1.558   0.1191
## ldlc        -0.0004538  0.0011361  -0.399   0.6896
## totchol      0.0007457  0.0011999   0.621   0.5343
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) age    sex2   hertrt hdlc   ldlc
## age      -0.705
## sex2      0.033  0.074
## heartrte -0.494  0.002 -0.030
## hdlc     -0.225  0.073 -0.241  0.017
## ldlc     -0.048  0.084  0.006  0.031  0.579
## totchol  -0.110 -0.074 -0.068 -0.060 -0.561 -0.917
## Generalized linear mixed model fit by maximum likelihood (Adaptive
##   Gauss-Hermite Quadrature, nAGQ = 0) [glmerMod]
##  Family: poisson  ( log )
## Formula: cigpday ~ age + sex + heartrte + hdlc + totchol + (1 | randid)
##    Data: frmgham_smoker
##
##      AIC      BIC   logLik deviance df.resid
##   7874.4   7908.9  -3930.2   7860.4     1014
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.83512 -0.26504  0.05623  0.18987  0.52476
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  randid (Intercept) 0.3593   0.5995
```

```
## Number of obs: 1021, groups:  randid, 1021
##
## Fixed effects:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  3.6482799  0.2409483  15.141  < 2e-16 ***
## age         -0.0166593  0.0028704  -5.804 6.48e-09 ***
## sex2        -0.2543480  0.0438424  -5.801 6.58e-09 ***
## heartrte     0.0041465  0.0016398   2.529   0.0114 *
## hdlc        -0.0023462  0.0014415  -1.628   0.1036
## totchol      0.0003063  0.0004795   0.639   0.5229
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) age    sex2   hertrt hdlc
## age      -0.705
## sex2      0.033  0.074
## heartrte -0.494 -0.001 -0.030
## hdlc     -0.242  0.030 -0.300 -0.001
## totchol  -0.386  0.008 -0.157 -0.079 -0.092
```

```
## Generalized linear mixed model fit by maximum likelihood (Adaptive
##   Gauss-Hermite Quadrature, nAGQ = 0) [glmerMod]
##  Family: poisson  ( log )
## Formula: cigpday ~ age + sex + heartrte + hdlc + (1 | randid)
##    Data: frmgham_smoker
##
##      AIC      BIC   logLik deviance df.resid
##   7872.8   7902.4  -3930.4   7860.8     1015
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.83222 -0.26666  0.05589  0.19140  0.52811
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  randid (Intercept) 0.3595   0.5996
## Number of obs: 1021, groups:  randid, 1021
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  3.707732   0.222277  16.681  < 2e-16 ***
## age         -0.016674   0.002871  -5.809 6.30e-09 ***
## sex2        -0.249947   0.043303  -5.772 7.83e-09 ***
## heartrte     0.004229   0.001635   2.587  0.00969 **
## hdlc        -0.002262   0.001436  -1.576  0.11513
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) age    sex2   hertrt
## age      -0.761
## sex2     -0.030  0.076
## heartrte -0.570  0.000 -0.043
```

```
## hdlc     -0.302  0.031 -0.320 -0.008
```

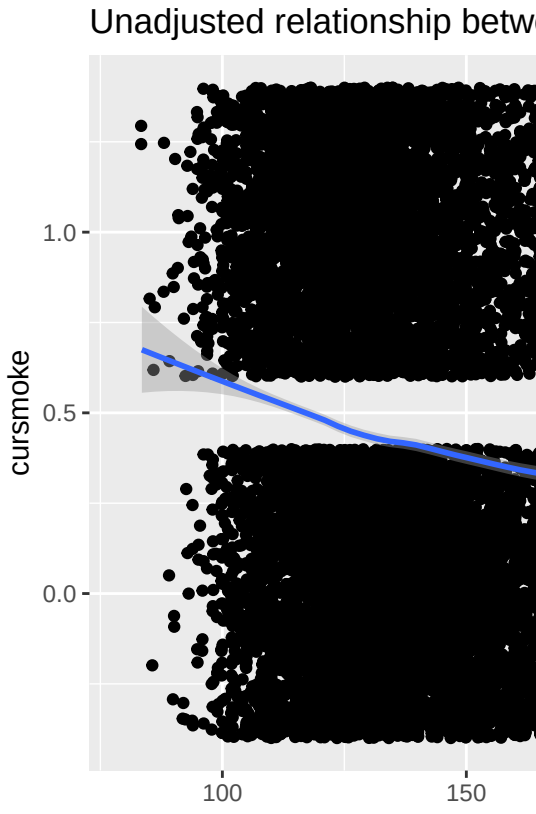
```
## Generalized linear mixed model fit by maximum likelihood (Adaptive
##   Gauss-Hermite Quadrature, nAGQ = 0) [glmerMod]
##  Family: poisson  ( log )
## Formula:
## cigpday ~ age + sex + heartrte + hdlc + age:heartrte + (1 | randid)
##    Data: frmgham_smoker
##
##      AIC      BIC   logLik deviance df.resid
##   7870.3   7904.8  -3928.1   7856.3     1014
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.84350 -0.26944  0.04809  0.19496  0.48497
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  randid (Intercept) 0.3581   0.5984
## Number of obs: 1021, groups:  randid, 1021
##
## Fixed effects:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)   1.6388170  1.0171175   1.611   0.1071
## age           0.0190057  0.0173552   1.095   0.2735
## sex2         -0.2466685  0.0432612  -5.702 1.19e-08 ***
## heartrte      0.0305622  0.0127412   2.399   0.0165 *
## hdlc         -0.0022279  0.0014334  -1.554   0.1201
## age:heartrte -0.0004549  0.0002183  -2.084   0.0372 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) age    sex2   hertrt hdlc
## age         -0.990
## sex2        -0.042  0.048
## heartrte    -0.984  0.978  0.030
## hdlc        -0.076  0.016 -0.319  0.010
## age:heartrt  0.976 -0.986 -0.036 -0.992 -0.011
```

```
## Generalized linear mixed model fit by maximum likelihood (Adaptive
##   Gauss-Hermite Quadrature, nAGQ = 0) [glmerMod]
##  Family: poisson  ( log )
## Formula: cigpday ~ age + sex + heartrte + hdlc + age * heartrte + (1 |
##     randid)
##    Data: frmgham_smoker
##
##      AIC      BIC   logLik deviance df.resid
##   7870.3   7904.8  -3928.1   7856.3     1014
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.84350 -0.26944  0.04809  0.19496  0.48497
##
## Random effects:
```

```
##  Groups Name        Variance Std.Dev.
##  randid (Intercept) 0.3581   0.5984
## Number of obs: 1021, groups:  randid, 1021
##
## Fixed effects:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)   1.6388170  1.0171175   1.611   0.1071
## age           0.0190057  0.0173552   1.095   0.2735
## sex2         -0.2466685  0.0432612  -5.702 1.19e-08 ***
## heartrte      0.0305622  0.0127412   2.399   0.0165 *
## hdlc         -0.0022279  0.0014334  -1.554   0.1201
## age:heartrte -0.0004549  0.0002183  -2.084   0.0372 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) age    sex2   hertrt hdlc
## age         -0.990
## sex2        -0.042  0.048
## heartrte    -0.984  0.978  0.030
## hdlc        -0.076  0.016 -0.319  0.010
## age:heartrt  0.976 -0.986 -0.036 -0.992 -0.011
```

2.Target question:

Next you are interested in the relationship between certain health outcomes and smoking status. In particular you are interested in :



(1)The relationship between current smoking status and systolic blood pressure.

```
## # A tibble: 3 x 6
##   variable    OR lower_CI upper_CI conf_OR confounder
##   <chr>    <dbl>    <dbl>    <dbl>   <dbl> <lgl>
## 1 diabp    0.966    0.966    0.966   0.953 TRUE
## 2 prevhyp  0.966    0.966    0.966   0.978 TRUE
## 3 hyperten 0.966    0.966    0.966   0.971 TRUE
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control
## $checkConv, : Model failed to converge with max|grad| = 0.546003 (tol =
## 0.001, component 1)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unide
##  - Rescale variables?
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: cursmoke ~ diabp_c + sysbp_c + prevhyp + hyperten + (1 | randid)
##    Data: frmgham
##
##      AIC      BIC   logLik deviance df.resid
##  11570.6  11614.8  -5779.3  11558.6    11621
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.5272 -0.2154 -0.1233  0.2498  3.2823
##
## Random effects:
```

```
##  Groups Name        Variance Std.Dev.
##  randid (Intercept) 21.23    4.607
## Number of obs: 11627, groups:  randid, 4434
##
## Fixed effects:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.2448667  0.0001298  1887.0   <2e-16 ***
## diabp_c      0.0403299  0.0001298   310.7   <2e-16 ***
## sysbp_c     -0.0360534  0.0001300  -277.3   <2e-16 ***
## prevhyp1    -0.9957484  0.0001298 -7673.7   <2e-16 ***
## hyperten    -1.2052589  0.0001298 -9288.1   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) dibp_c sysbp_ prvhy1
## diabp_c  0.000
## sysbp_c  0.000  0.000
## prevhyp1 0.000  0.000  0.000
## hyperten 0.000  0.000  0.000  0.000
## convergence code: 0
## Model failed to converge with max|grad| = 0.546003 (tol = 0.001, component 1)
## Model is nearly unidentifiable: very large eigenvalue
##  - Rescale variables?
```

(2)The relationship between current smoking status and diastolic blood pressure.

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27

## running glm to get initial regression estimate

## (Intercept)       diabp
##  0.83258871 -0.01331811

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##  (Intercept)        diabp        sysbp
##  1.201158028  0.009278866 -0.016517237

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##   (Intercept)         diabp      prevhyp1
## -0.0606238215  0.0007722528 -0.6125294096

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##  (Intercept)        diabp     hyperten
##  0.561729741 -0.005909097 -0.464765767

## # A tibble: 3 x 6
##   variable    OR lower_CI upper_CI conf_OR confounder
##   <chr>    <dbl>    <dbl>    <dbl>   <dbl> <lgl>
## 1 sysbp    0.994    0.992    0.997   1.01  TRUE
## 2 prevhyp  0.994    0.992    0.997   1.000 TRUE
## 3 hyperten 0.994    0.992    0.997   0.998 TRUE

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
```

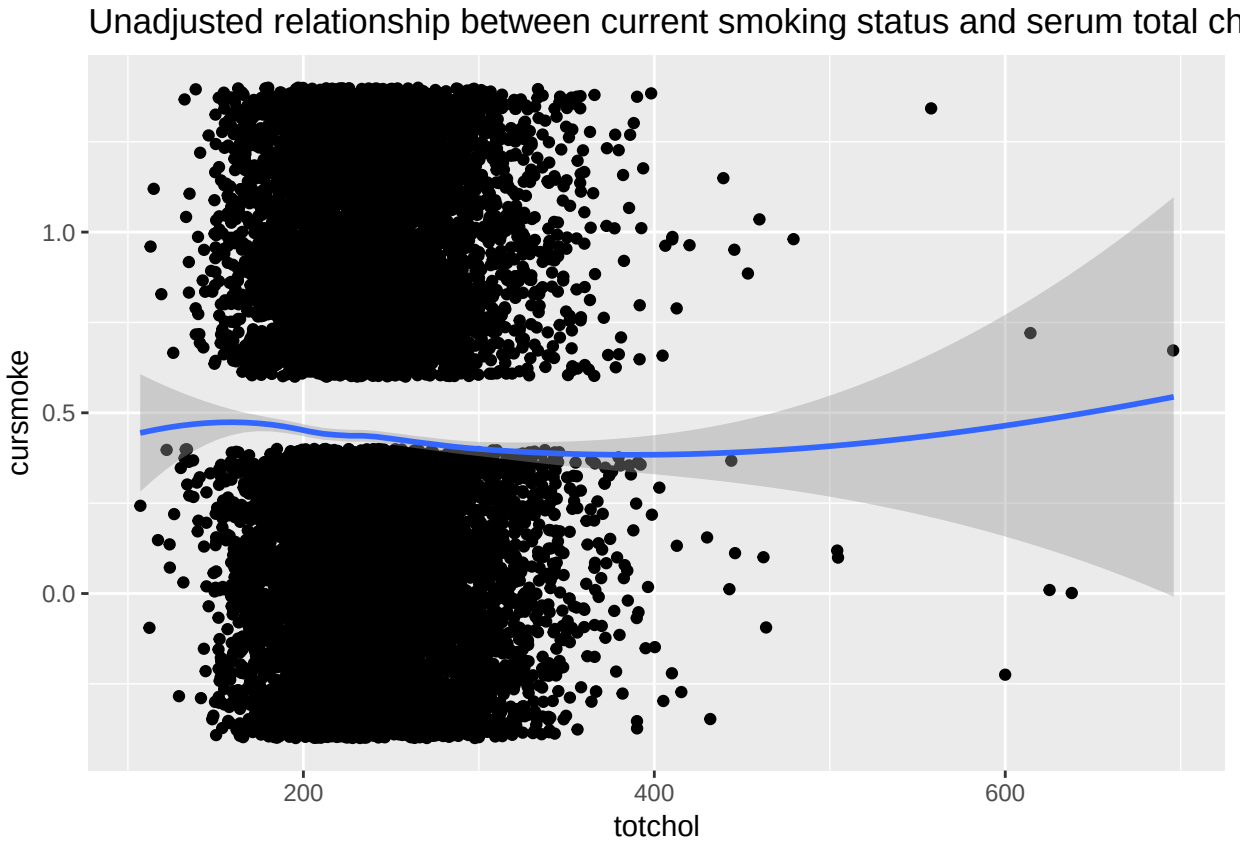
```
## running glm to get initial regression estimate
## (Intercept)       diabp       sysbp    prevhyp1    hyperten
##  0.42248755  0.01219174 -0.01035239 -0.35384039 -0.19034229
##
##  GEE:  GENERALIZED LINEAR MODELS FOR DEPENDENT DATA
##  gee S-function, version 4.13 modified 98/01/27 (1998)
##
## Model:
##  Link:                      Logit
##  Variance to Mean Relation: Binomial
##  Correlation Structure:     Unstructured
##
## Call:
## gee(formula = cursmoke ~ diabp + sysbp + prevhyp + hyperten,
##     id = randid, data = frmgham, family = binomial, corstr = ("unstructured"))
##
## Summary of Residuals:
##        Min         1Q     Median         3Q        Max
## -0.5772852 -0.4322399 -0.3434050  0.5378886  0.7834695
##
##
## Coefficients:
##                 Estimate  Naive S.E.   Naive z Robust S.E.  Robust z
## (Intercept)  0.358492204 0.140153972  2.557845 0.136318579  2.629812
## diabp        0.009015059 0.001987666  4.535501 0.001937856  4.652079
## sysbp       -0.007680584 0.001185816 -6.477047 0.001125213 -6.825895
## prevhyp1    -0.251511525 0.040964453 -6.139751 0.038933722 -6.459992
## hyperten    -0.288846931 0.066297876 -4.356805 0.068177985 -4.236660
##
## Estimated Scale Parameter:  0.9931821
## Number of Iterations:  3
##
## Working Correlation
##           [,1]      [,2]      [,3]
## [1,] 1.0000000 0.7731811 0.5108965
## [2,] 0.7731811 1.0000000 0.5728086
## [3,] 0.5108965 0.5728086 1.0000000
```

From the table above, we found "sysbp", "prevhyp" and "hyperten" are confounders. After adding the confounders, the final model is shown above. For a one-unit increase in the diastolic blood pressure, the expected change in log odds is 0.009, adjusting for systolic blood pressure, prevalent hypertensive and hypertensive.

---

(3)The relationship between current smoking status and serum total cholesterol. Again, while answering these questions, please account for any confounders that you have evidence may impact these relationships.

```
## Warning: Removed 409 rows containing non-finite values (stat_smooth).
## Warning: Removed 409 rows containing missing values (geom_point).
```




```
##
##     0     1
##   124 11503
```

Finding confounders

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27

## running glm to get initial regression estimate

##  (Intercept)      totchol
##  0.180474997 -0.001869103

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##  (Intercept)      totchol         sex2
##  0.304364508 -0.001008861 -0.599150529

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##   (Intercept)       totchol           age
##  2.8313902251 -0.0001034184 -0.0566406814

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##  (Intercept)      totchol        sysbp
```

```
##  1.725363617 -0.000925225 -0.013079212
```

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate
```

```
##  (Intercept)      totchol        diabp
##  1.122586351 -0.001456582 -0.012565835
```

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate
```

```
## Warning: glm.fit: algorithm did not converge
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
##   (Intercept)       totchol       cigpday
## -2.149461e+01 -1.846024e-04  3.959552e+01
```

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate
```

```
##  (Intercept)      totchol          bmi
##  2.267859530 -0.001381211 -0.085631423
```

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate
```

```
##  (Intercept)      totchol    diabetes1
##  0.199463411 -0.001870223 -0.436048131
```

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate
```

```
## (Intercept)     totchol     bpmeds1
##  0.22186119 -0.00176108 -0.66682312
```

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate
```

```
##  (Intercept)      totchol     heartrte
## -0.577000661 -0.002105534  0.010599667
```

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate
```

```
##  (Intercept)      totchol      glucose
##  0.731301297 -0.002075508 -0.005853906
```

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate
```

```
##  (Intercept)      totchol        educ2        educ3        educ4
##  0.061276975 -0.001769349  0.309068482 -0.057506010  0.101512857
```

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate
```

```
##  (Intercept)      totchol     prevchd1
##  0.188137717 -0.001799442 -0.352449578
```

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate
```

```
##  (Intercept)      totchol      prevap1
##  0.181597221 -0.001775954 -0.465446390
```

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##   (Intercept)       totchol       prevmi1
##  0.1804896950 -0.0018690669 -0.0007246913

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##  (Intercept)      totchol    prevstrk1
##  0.181481427 -0.001843588 -0.601436167

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##  (Intercept)      totchol     prevhyp1
##  0.254704355 -0.001093214 -0.587008909

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##  (Intercept)      totchol       period
##  0.727805611 -0.001822751 -0.299025251

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##  (Intercept)      totchol         hdlc
##  0.044882356 -0.002570154 -0.001898161

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##  (Intercept)      totchol         ldlc
##  0.250677947 -0.009187024  0.007152646

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##  (Intercept)      totchol       angina
##  0.175656048 -0.001729512 -0.177845483

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##  (Intercept)      totchol       hospmi
##  0.195417198 -0.002078423  0.351737296

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##  (Intercept)      totchol      mi_fchd
##  0.189787408 -0.002082404  0.270651764

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate

##  (Intercept)      totchol       anychd
##  0.180322574 -0.001891149  0.020107886

## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate
```

```
##  (Intercept)      totchol       stroke
##  0.184952930 -0.001865479 -0.058643127
```

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate
```

```
##  (Intercept)      totchol          cvd
##  0.171412976 -0.001961615  0.124512945
```

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate
```

```
##  (Intercept)      totchol     hyperten
##  0.448140968 -0.001369585 -0.525759376
```

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate
```

```
##   (Intercept)       totchol  age_ctg35-44  age_ctg45-54  age_ctg55-64
##  0.4237096503 -0.0004515521  0.0775985537 -0.2787440726 -0.8304547644
##  age_ctg65-74    age_ctg75+
## -1.4405447539 -2.2964765535
```

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate
```

```
##       (Intercept)           totchol   sysbp_ctgcrisis sysbp_ctgelevated
##      0.3823593209     -0.0009234769     -1.0191658985     -0.3670021499
##   sysbp_ctgstage1   sysbp_ctgstage2
##     -0.4383582160     -0.7076450514
```

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate
```

```
##     (Intercept)         totchol diabp_ctgcrisis diabp_ctgstage1
##     0.260736559    -0.001482552    -0.452776545    -0.202567676
## diabp_ctgstage2
##    -0.386936757
```

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate
```

```
##  (Intercept)      totchol      diabp_c
##  0.079147604 -0.001456582 -0.012565835
```

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate
```

```
##  (Intercept)      totchol      sysbp_c
## -0.057648336 -0.000925225 -0.013079212
```

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate
```

```
##  (Intercept)      totchol  totchol_ctg
##  0.303357304 -0.001839479 -0.131485645
```

```
## Beginning Cgee S-function, @(#) geeformula.q 4.13 98/01/27
## running glm to get initial regression estimate
```

```
##  (Intercept)      totchol      bmi_cat
##  0.768854509 -0.001887641 -0.197038129
```

```
## $sex
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
##
## $age
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        1.001  TRUE
##
## $sysbp
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1            1  FALSE
##
## $diabp
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
##
## $cigpday
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
##
## $bmi
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1            1  FALSE
##
## $diabetes
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
##
## $bpmeds
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
```

```
## totchol     0.999      0.998          1        0.999  FALSE
##
## $heartrte
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
##
## $glucose
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
##
## $educ
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
##
## $prevchd
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
##
## $prevap
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
##
## $prevmi
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
##
## $prevstrk
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
##
## $prevhyp
##
##
```

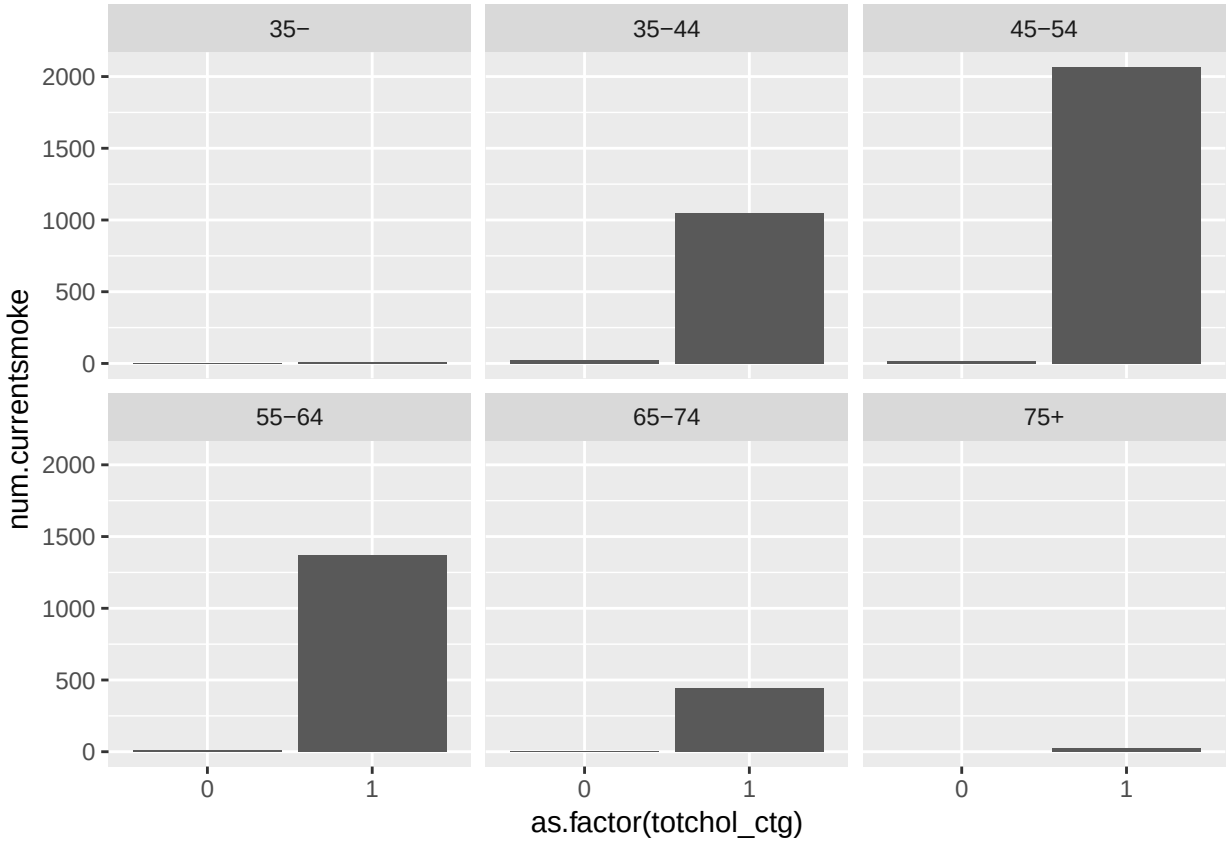
```
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1            1  FALSE
##
## $period
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1            1  FALSE
##
## $hdlc
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.997  TRUE
##
## $ldlc
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.991  TRUE
##
## $angina
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
##
## $hospmi
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
##
## $mi_fchd
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
##
## $anychd
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
##
## $stroke
```

```
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
##
## $cvd
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
##
## $hyperten
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
##
## $age_ctg
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1            1  TRUE
##
## $sysbp_ctg
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1            1  FALSE
##
## $diabp_ctg
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
##
## $diabp_c
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
##
## $sysbp_c
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1            1  FALSE
```

```
##
## $totchol_ctg
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
##
## $bmi_cat
##
##
## variable       OR   lower_CI   upper_CI   potconf_OR  confounder
## ---------  ------  ---------  ---------  -----------  -----------
## totchol     0.999      0.998          1        0.999  FALSE
```

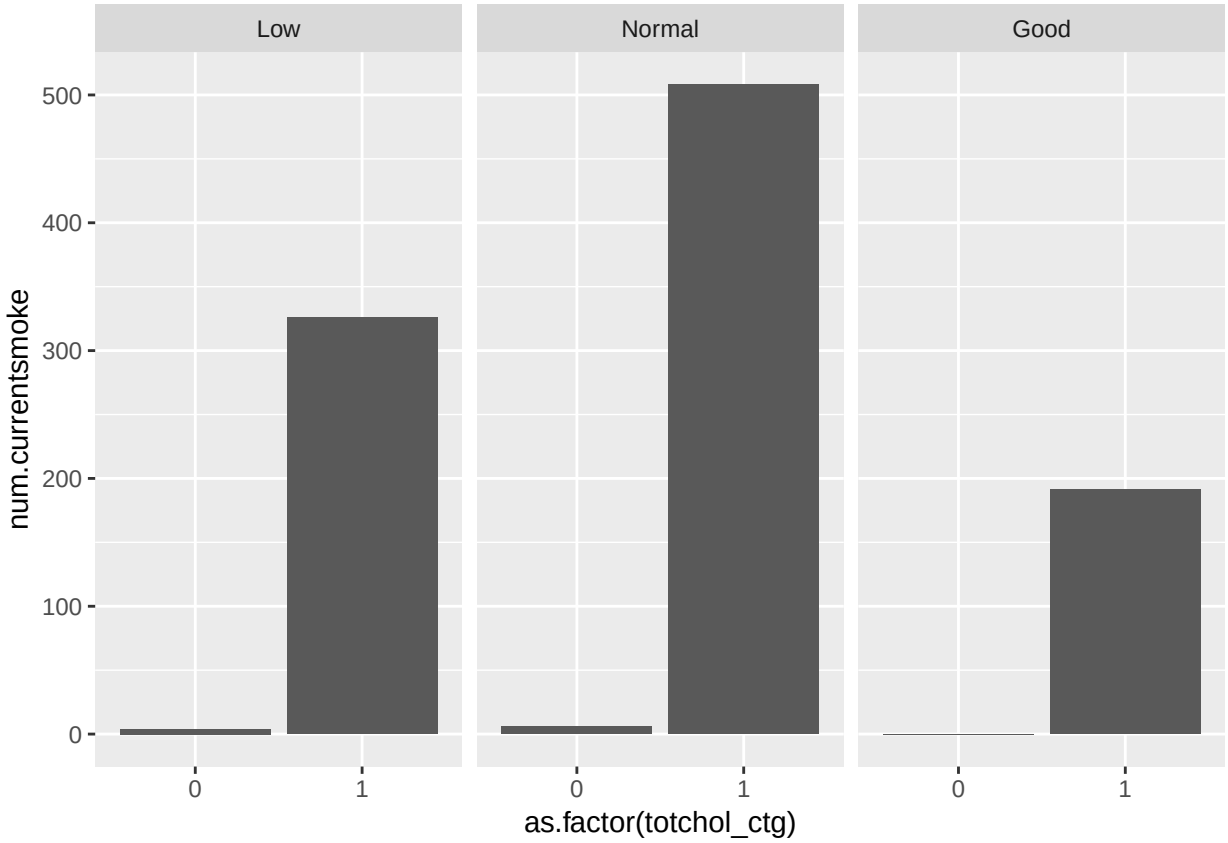
(1) age

```
## # A tibble: 11 x 4
## # Groups:   totchol_ctg [?]
##    totchol_ctg age_ctg percent.currentsmoke     n
##          <dbl> <fct>                   <dbl> <int>
##  1           0 35-                     1         1
##  2           0 35-44                   0.677    31
##  3           0 45-54                   0.439    41
##  4           0 55-64                   0.424    33
##  5           0 65-74                   0.333    18
##  6           1 35-                     0.565    23
##  7           1 35-44                   0.598  1754
##  8           1 45-54                   0.510  4054
##  9           1 55-64                   0.374  3668
## 10           1 65-74                   0.245  1801
## 11           1 75+                     0.128   203
```



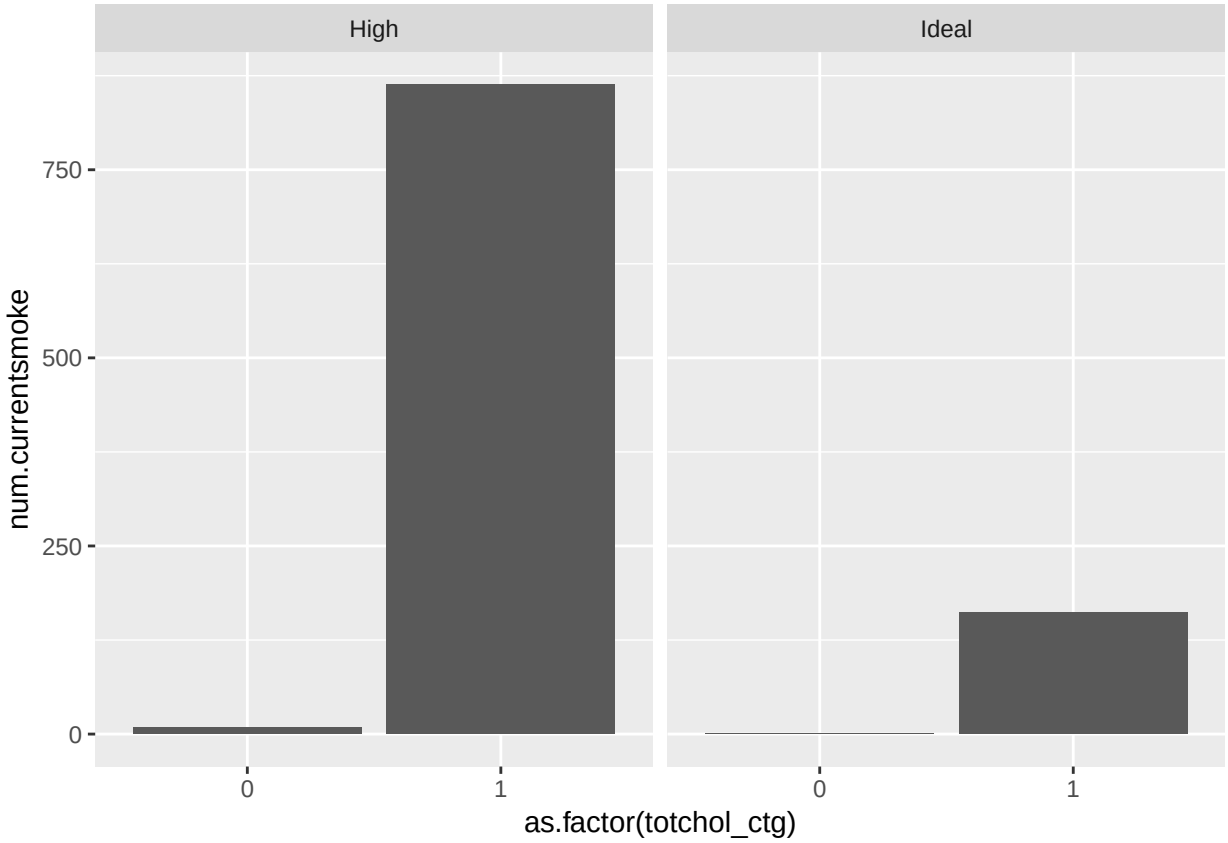
(2)hdlc

```
## # A tibble: 6 x 4
## # Groups:   totchol_ctg [?]
##   totchol_ctg hdlc_ctg percent.currentsmoke     n
##         <dbl> <fct>                   <dbl> <int>
## 1           0 Low                     0.571     7
## 2           0 Normal                  0.375    16
## 3           0 Good                    0         2
## 4           1 Low                     0.358   911
## 5           1 Normal                  0.339  1497
## 6           1 Good                    0.322   594
```



(3)ldlc

```
## # A tibble: 4 x 4
## # Groups:   totchol_ctg [?]
##   totchol_ctg ldlc_ctg percent.currentsmoke     n
##         <dbl> <chr>                   <dbl> <int>
## 1           0 High                    0.45     20
## 2           0 Ideal                   0.2       5
## 3           1 High                    0.336  2571
## 4           1 Ideal                   0.377   430
```



From the table above, we found age, ldlc and hdlc are confounders.

model fitting

```
##           totchol         age        hdlc        ldlc
## totchol 1.0000000  0.14940630  0.17649924  0.85723230
## age     0.1494063  1.00000000 -0.01070049  0.01044638
## hdlc    0.1764992 -0.01070049  1.00000000 -0.13746200
## ldlc    0.8572323  0.01044638 -0.13746200  1.00000000
```

```
## Generalized linear mixed model fit by maximum likelihood (Adaptive
##   Gauss-Hermite Quadrature, nAGQ = 0) [glmerMod]
##  Family: binomial  ( logit )
## Formula: cursmoke ~ totchol + age + hdlc + ldlc + (1 | randid)
##    Data: frmgham
##
##      AIC      BIC   logLik deviance df.resid
##   3691.2   3727.3  -1839.6   3679.2     3020
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.2509 -0.7394 -0.5252  1.0545  2.9054
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  randid (Intercept) 0.01273  0.1128
```

```
## Number of obs: 3026, groups:  randid, 3026
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  3.864097   0.376818  10.255  < 2e-16 ***
## totchol     -0.008788   0.002352  -3.737 0.000186 ***
## age         -0.064952   0.005127 -12.669  < 2e-16 ***
## hdlc         0.003762   0.003331   1.129 0.258740
## ldlc         0.007134   0.002214   3.222 0.001271 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##         (Intr) totchl age    hdlc
## totchol -0.155
## age     -0.780 -0.107
## hdlc    -0.226 -0.627  0.080
## ldlc    -0.034 -0.920  0.094  0.624
```

```
##                      Est           LL           UL
## (Intercept)  3.864097478  3.125534960  4.602659997
## totchol     -0.008788264 -0.013397730 -0.004178797
## age         -0.064951700 -0.075000390 -0.054903009
## hdlc         0.003762329 -0.002767096  0.010291754
## ldlc         0.007133770  0.002794694  0.011472846
```

```
##                    Est         LL         UL
## (Intercept) 47.6602386 22.7720740 99.7492958
## totchol      0.9912502  0.9866916  0.9958299
## age          0.9371127  0.9277431  0.9465770
## hdlc         1.0037694  0.9972367  1.0103449
## ldlc         1.0071593  1.0027986  1.0115389
```