GroupProject

LAUREN

11/23/2021

```
varinterst = c("Age","Gender","Race1","PhysActive","Depressed","Diabetes","BPSysAve", "BPDiaAve", "TotCh

nh1 = nh[,varinterst]
nh1 <- nh1[ which(nh1$Age >=18 & nh1$Age <=59), ]

nh1 <- nh1 %>%
  mutate(HypT = if_else(nh1$BPSysAve >= 130 | nh1$BPDiaAve >=80, 1, 0)) #Add hypertension variable
nh1 <- nh1 %>%
  mutate(Sex = ifelse(nh1$Gender == "female", 1, 0))
nh1 <- nh1 %>%
  mutate(PhysAc = ifelse(nh1$PhysActive == "No", 0, 1))
nh1 <- nh1 %>%
  mutate(Diabete = ifelse(nh1$Diabetes == "No", 0, 1))
nh1 <- nh1 %>%
  mutate(Depress = ifelse(nh1$Depressed == "None", 0, 1))
nh1 <- nh1 %>%
  mutate(RegularMarij2 = ifelse(nh1$RegularMarij == "No", 0, 1))
nh1$TotChol = log(nh1$TotChol)
nh1$Incomplete = complete.cases(nh1)
nh1$Race1 = as.factor(nh1$Race1)

drops = c("Sex")
nhF= nh1[ which(nh1$Sex == 1), ]
nhM = nh1[ which(nh1$Sex == 0), ]
nhF = nhF[,!(names(nhF) %in% drops)]
nhM = nhM[,!(names(nhM) %in% drops)]
```

```
## Shiny app
nhF %>%
 tbl_summary(by = Incomplete,
             missing = "no",
             statistic = list(all_continuous() ~ "{mean} ({sd})",
                              all_categorical() ~ "{n} ({p}%)"),
             ) %>%
  add_n %>%
  add_p(test = list(Age ~ "t.test",
                    Race1 ~ "chisq.test",
                    PhysAc ~ "chisq.test",
                    Depress ~ "chisq.test",
                    Diabete ~ "chisq.test",
                    HypT ~ "chisq.test",
```

```
                    TotChol ~ "t.test",
                    Testosterone ~ "t.test",
                    RegularMarij2 ~ "chisq.test"),
       test.args = all_tests("t.test") ~ list(var.equal = TRUE), ## Important argument!
       pvalue_fun = function(x) style_pvalue(x, digits = 2)) %>%
  bold_p(t = 0.05) %>%
  bold_labels
```

```
## Table printed with `knitr::kable()`, not {gt}. Learn why at
## http://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
## To suppress this message, include `message = FALSE` in code chunk header.
```

Characteristic	**N**	**FALSE**, N = 1,723	**TRUE**, N = 1,051	**p-value**
Age	2,774	38 (12)	39 (12)	0.50
Gender	2,774			>0.99
female		1,723 (100%)	1,051 (100%)	
male		0 (0%)	0 (0%)	
Race1	2,774			0.19
Black		224 (13%)	133 (13%)	
Hispanic		118 (6.8%)	68 (6.5%)	
Mexican		165 (9.6%)	86 (8.2%)	
White		1,048 (61%)	682 (65%)	
Other		168 (9.8%)	82 (7.8%)	
PhysActive	2,774	887 (51%)	642 (61%)	**<0.001**
Depressed	2,367			**0.003**
None		941 (72%)	807 (77%)	
Several		266 (20%)	156 (15%)	
Most		109 (8.3%)	88 (8.4%)	
Diabetes	2,774	83 (4.8%)	53 (5.0%)	0.79
BPSysAve	2,644	113 (14)	115 (15)	**0.001**
BPDiaAve	2,644	68 (12)	70 (11)	**<0.001**
TotChol	2,626	1.59 (0.20)	1.60 (0.21)	0.31
Testosterone	1,253	23 (12)	26 (21)	0.053
RegularMarij	2,347	269 (21%)	238 (23%)	0.27
HypT	2,644	318 (20%)	230 (22%)	0.25
PhysAc	2,774	887 (51%)	642 (61%)	**<0.001**
Diabete	2,774	83 (4.8%)	53 (5.0%)	0.86
Depress	2,367	375 (28%)	244 (23%)	**0.004**
RegularMarij2	2,347	269 (21%)	238 (23%)	0.29

```
nhM %>%
 tbl_summary(by = Incomplete,
             missing = "no",
             statistic = list(all_continuous() ~ "{mean} ({sd})",
                              all_categorical() ~ "{n} ({p}%)"),
             ) %>%
  add_n %>%
  add_p(test = list(Age ~ "t.test",
                    Race1 ~ "chisq.test",
                    PhysAc ~ "chisq.test",
                    Depress ~ "chisq.test",
```

```
                    Diabete ~ "chisq.test",
                    HypT ~ "chisq.test",
                    TotChol ~ "t.test",
                    Testosterone ~ "t.test",
                    RegularMarij2 ~ "chisq.test"),
      test.args = all_tests("t.test") ~ list(var.equal = TRUE),## Important argument!
      pvalue_fun = function(x) style_pvalue(x, digits = 2)) %>%
  bold_p(t = 0.05) %>%
  bold_labels
```

```
## Table printed with `knitr::kable()`, not {gt}. Learn why at
## http://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
## To suppress this message, include `message = FALSE` in code chunk header.
```

Characteristic	**N**	**FALSE**, N = 1,664	**TRUE**, N = 1,204	**p-value**
Age	2,868	38 (12)	39 (12)	0.10
Gender	2,868			>0.99
female		0 (0%)	0 (0%)	
male		1,664 (100%)	1,204 (100%)	
Race1	2,868			0.070
Black		201 (12%)	114 (9.5%)	
Hispanic		93 (5.6%)	76 (6.3%)	
Mexican		202 (12%)	124 (10%)	
White		1,039 (62%)	785 (65%)	
Other		129 (7.8%)	105 (8.7%)	
PhysActive	2,868	962 (58%)	709 (59%)	0.56
Depressed	2,617			0.47
None		1,129 (80%)	985 (82%)	
Several		203 (14%)	157 (13%)	
Most		81 (5.7%)	62 (5.1%)	
Diabetes	2,868	111 (6.7%)	82 (6.8%)	0.88
BPSysAve	2,784	121 (15)	120 (14)	0.74
BPDiaAve	2,784	72 (12)	73 (11)	**0.019**
TotChol	2,723	1.60 (0.21)	1.57 (0.21)	**<0.001**
Testosterone	1,333	397 (165)	415 (170)	0.27
RegularMarij	2,594	462 (33%)	397 (33%)	0.89
HypT	2,784	566 (36%)	419 (35%)	0.60
PhysAc	2,868	962 (58%)	709 (59%)	0.59
Diabete	2,868	111 (6.7%)	82 (6.8%)	0.94
Depress	2,617	284 (20%)	219 (18%)	0.24
RegularMarij2	2,594	462 (33%)	397 (33%)	0.92

```
nh2 <- nh1 %>% drop_na() #drop anything with NA
m1 =  lm(log(Testosterone)~Age + Sex + Race1 + PhysAc + Depress + Diabete + HypT+ TotChol + RegularMari
summary(m1)
```

```
##
## Call:
## lm(formula = log(Testosterone) ~ Age + Sex + Race1 + PhysAc +
##     Depress + Diabete + HypT + TotChol + RegularMarij2, data = nh2)
```

```
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.10497 -0.27713  0.01199  0.31695  2.60208
##
## Coefficients:
##                 Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    6.3613175  0.0859337   74.026  < 2e-16 ***
## Age           -0.0096663  0.0009369  -10.318  < 2e-16 ***
## Sex           -2.8848479  0.0210884 -136.798  < 2e-16 ***
## Race1Hispanic -0.0271273  0.0511246   -0.531  0.59574
## Race1Mexican  -0.0768069  0.0460189   -1.669  0.09525 .
## Race1White    -0.0659921  0.0336581   -1.961  0.05004 .
## Race1Other    -0.1241461  0.0473209   -2.623  0.00876 **
## PhysAc         0.0132567  0.0214976    0.617  0.53752
## Depress       -0.0128492  0.0257615   -0.499  0.61799
## Diabete       -0.1106606  0.0440763   -2.511  0.01212 *
## HypT          -0.0354984  0.0238415   -1.489  0.13665
## TotChol        0.0103321  0.0523056    0.198  0.84343
## RegularMarij2  0.0489403  0.0232149    2.108  0.03513 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4856 on 2242 degrees of freedom
## Multiple R-squared:  0.8985, Adjusted R-squared:  0.898
## F-statistic:  1654 on 12 and 2242 DF,  p-value: < 2.2e-16
```

```
quantile(nh2$Age, probs = seq(0, 1, 0.25))
```

```
##   0%  25%  50%  75% 100%
##   18   28   39   49   59
```

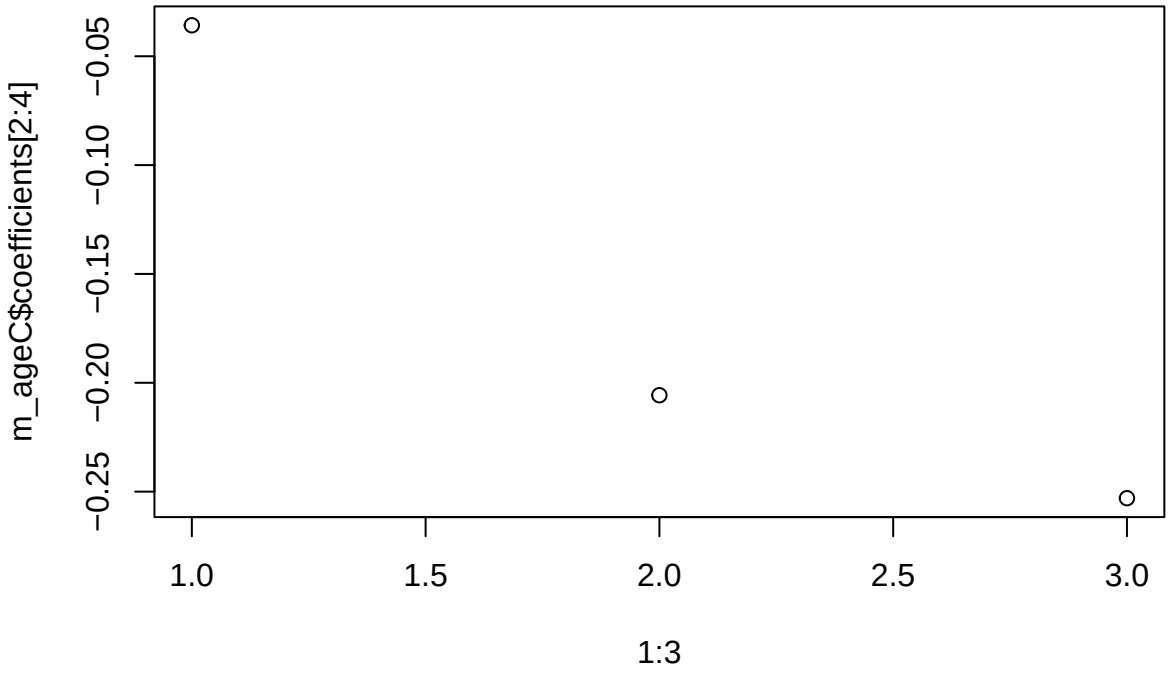
```
AgeQ1 <- nh2$Age<28
AgeQ2 <- nh2$Age>=28 & nh2$Age<39
AgeQ3 <- nh2$Age>=39 & nh2$Age<49
AgeQ4 <- nh2$Age>=49 & nh2$Age<59

m_ageC <- lm(log(Testosterone)~AgeQ2+AgeQ3+AgeQ4 + Sex + Race1 + PhysAc + Depress + Diabete + HypT+ TotC
summary(m_ageC)
```

```
##
## Call:
## lm(formula = log(Testosterone) ~ AgeQ2 + AgeQ3 + AgeQ4 + Sex +
##     Race1 + PhysAc + Depress + Diabete + HypT + TotChol + RegularMarij2,
##     data = nh2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.1181 -0.2700  0.0174  0.3190  2.5819
##
## Coefficients:
##               Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    6.17519    0.08624   71.605  < 2e-16 ***
```

```
## AgeQ2TRUE     -0.03579    0.02900   -1.234  0.21723
## AgeQ3TRUE     -0.20570    0.02983   -6.895 6.99e-12 ***
## AgeQ4TRUE     -0.25300    0.03058   -8.274  < 2e-16 ***
## Sex           -2.88415    0.02115 -136.374  < 2e-16 ***
## Race1Hispanic -0.02009    0.05124   -0.392  0.69501
## Race1Mexican  -0.07766    0.04616   -1.682  0.09266 .
## Race1White    -0.07090    0.03374   -2.102  0.03570 *
## Race1Other    -0.12752    0.04749   -2.685  0.00730 **
## PhysAc         0.02563    0.02143    1.196  0.23197
## Depress       -0.01337    0.02582   -0.518  0.60467
## Diabete       -0.12108    0.04423   -2.738  0.00623 **
## HypT          -0.04978    0.02370   -2.100  0.03581 *
## TotChol       -0.03140    0.05168   -0.608  0.54349
## RegularMarij2  0.04880    0.02328    2.096  0.03621 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4867 on 2240 degrees of freedom
## Multiple R-squared:  0.8981, Adjusted R-squared:  0.8975
## F-statistic:  1411 on 14 and 2240 DF,  p-value: < 2.2e-16
```

```
plot(1:3, m_ageC$coefficients[2:4])
```




```
m_ageC <- lm(log(Testosterone)~AgeQ2+AgeQ3+AgeQ4 + Sex + Race1 + PhysAc + Depress + Diabete + HypT+ TotC
```

```
cor(nh2[c("Diabete", "TotChol", "HypT", "PhysAc","Depress")])
```

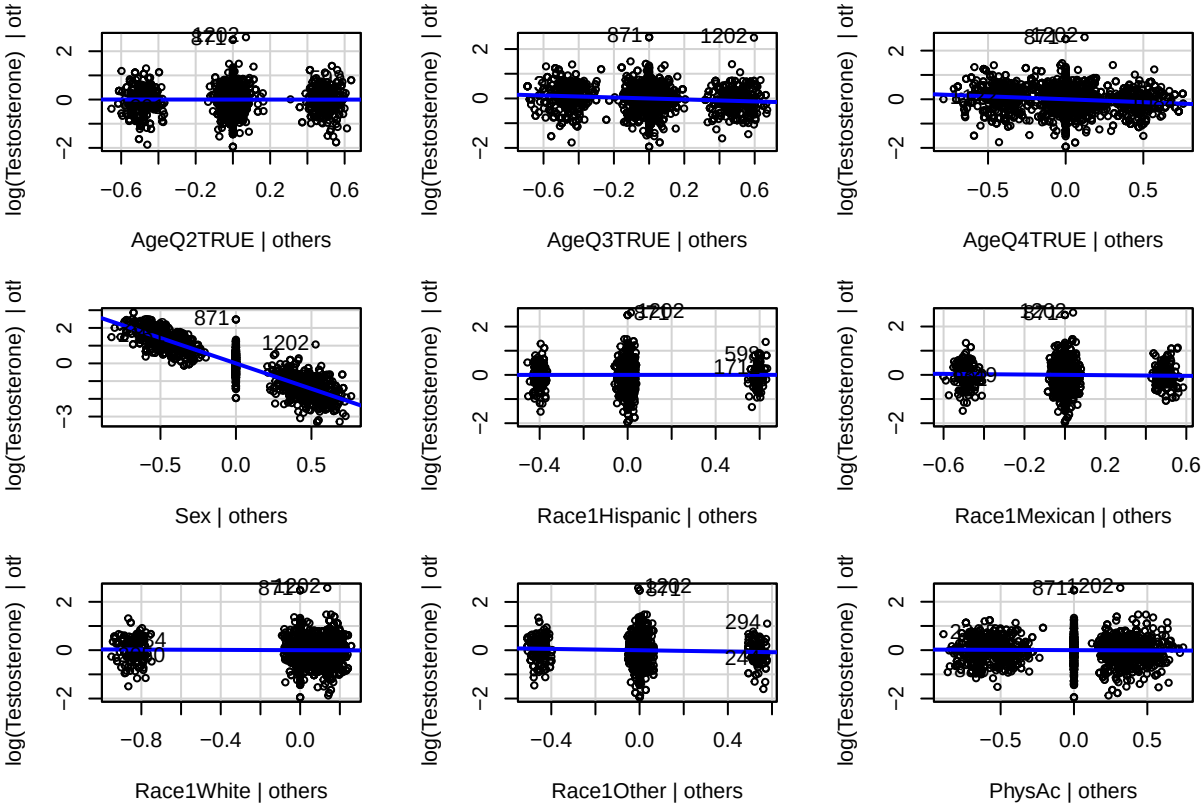
```
##             Diabete     TotChol        HypT      PhysAc     Depress
## Diabete  1.00000000  0.01684888  0.08317883 -0.05294167  0.03832360
## TotChol  0.01684888  1.00000000  0.18259500 -0.06464567  0.06684462
## HypT     0.08317883  0.18259500  1.00000000 -0.01563832  0.05758225
## PhysAc  -0.05294167 -0.06464567 -0.01563832  1.00000000 -0.11063907
## Depress  0.03832360  0.06684462  0.05758225 -0.11063907  1.00000000
```

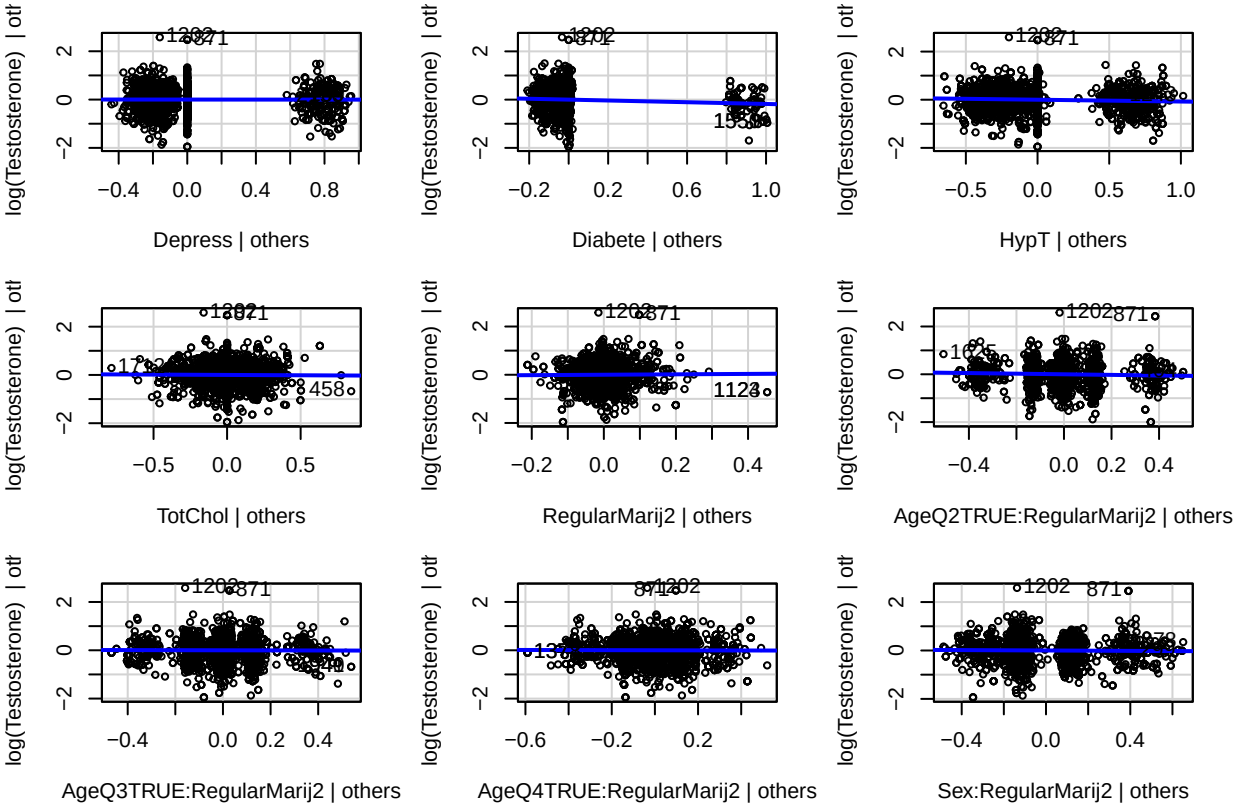
```
car::vif(m_ageC)
```

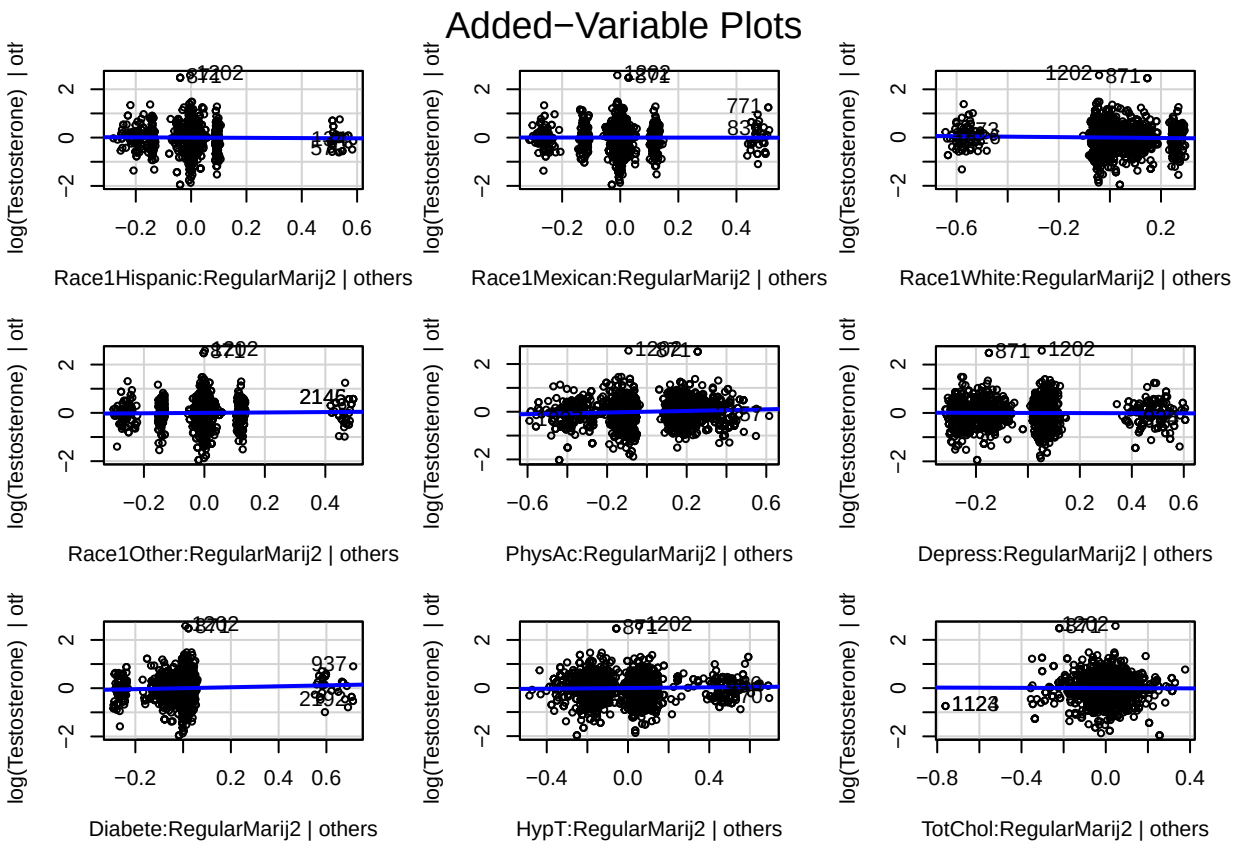
```
##                             GVIF Df GVIF^(1/(2*Df))
## AgeQ2                   2.040759  1        1.428551
## AgeQ3                   2.114436  1        1.454110
## AgeQ4                   2.398694  1        1.548772
## Sex                     1.449860  1        1.204101
## Race1                   3.394524  4        1.165058
## PhysAc                  1.512573  1        1.229867
## Depress                 1.558318  1        1.248326
## Diabete                 1.490743  1        1.220960
## HypT                    1.518555  1        1.232297
## TotChol                 1.602995  1        1.266095
## RegularMarij2          64.969309  1        8.060354
## AgeQ2:RegularMarij2     2.514017  1        1.585565
## AgeQ3:RegularMarij2     2.678259  1        1.636539
## AgeQ4:RegularMarij2     3.243785  1        1.801051
## Sex:RegularMarij2       2.110015  1        1.452589
## Race1:RegularMarij2    21.745286  4        1.469504
## PhysAc:RegularMarij2    2.775730  1        1.666052
## Depress:RegularMarij2   1.940016  1        1.392845
## Diabete:RegularMarij2   1.592168  1        1.261811
## HypT:RegularMarij2      1.983350  1        1.408315
## TotChol:RegularMarij2  63.334636  1        7.958306
```

linearity check

```
car::avPlots(m_ageC)
```


log(Testosterone) | oth
AgeQ2TRUE | others
AgeQ3TRUE | others
AgeQ4TRUE | others
Sex | others
Race1Hispanic | others
Race1Mexican | others
Race1White | others
Race1Other | others
PhysAc | others


log(Testosterone) | oth
Depress | others
Diabete | others
HypT | others
TotChol | others
RegularMarij2 | others
AgeQ2TRUE:RegularMarij2 | others
AgeQ3TRUE:RegularMarij2 | others
AgeQ4TRUE:RegularMarij2 | others
Sex:RegularMarij2 | others


Added−Variable Plots
log(Testosterone) | oth
Race1Hispanic:RegularMarij2 | others
Race1Mexican:RegularMarij2 | others
Race1White:RegularMarij2 | others
Race1Other:RegularMarij2 | others
PhysAc:RegularMarij2 | others
Depress:RegularMarij2 | others
Diabete:RegularMarij2 | others
HypT:RegularMarij2 | others
TotChol:RegularMarij2 | others