

Better R coding -- without attach

```
> options(scipen=999) # To suppress scientific notation
> math =
read.table("http://www.utstat.toronto.edu/brunner/data/legal/mathcat.data.txt")
> head(math)
  hsgpa hsenl1 hscalc  course passed outcome
1  78.0    80    Yes Mainstrm    No  Failed
2  66.0    75    Yes Mainstrm   Yes  Passed
3  80.2    70    Yes Mainstrm   Yes  Passed
4  81.7    67    Yes Mainstrm   Yes  Passed
5  86.8    80    Yes Mainstrm   Yes  Passed
6  76.7    75    Yes Mainstrm   Yes  Passed
> ls()
[1] "math"
> mean(hsgpa)
Error in mean(hsgpa) : object 'hsgpa' not found
> with(math, mean(hsgpa))
[1] 79.73756
> # Need numeric binary outcome
> math2 = within(math, {
+ pass = c1 = c2 = numeric(length(hsgpa))
+ pass[passed=='Yes'] = 1
+ c1[course=='Catch-up'] = 1
+ c2[course=='Elite'] = 1
+ }) # End within
> head(math2)
  hsgpa hsenl1 hscalc  course passed outcome pass c1 c2
1  78.0    80    Yes Mainstrm    No  Failed    0  0  0
2  66.0    75    Yes Mainstrm   Yes  Passed    1  0  0
3  80.2    70    Yes Mainstrm   Yes  Passed    1  0  0
4  81.7    67    Yes Mainstrm   Yes  Passed    1  0  0
5  86.8    80    Yes Mainstrm   Yes  Passed    1  0  0
6  76.7    75    Yes Mainstrm   Yes  Passed    1  0  0
> with(math2, {
+ print(table(pass,passed)) # Need print or only the last one shows
+ cat('\n')
+ print(table(course,c1))
+ cat('\n')
+ print(table(course,c2))
+ cat('\n')
+ }) # End with
      passed
pass  No Yes
    0 158  0
    1  0 236

      c1
course  0  1
Catch-up  0 35
Elite    31 0
Mainstrm 328 0

      c2
course  0  1
Catch-up 35 0
Elite    0 31
Mainstrm 328 0
```

```
> modell = glm(pass ~ hsgpa + hsengl + c1 + c2, family=binomial, data=math2)
> summary(modell)
```

Call:

```
glm(formula = pass ~ hsgpa + hsengl + c1 + c2, family = binomial,
     data = math2)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.5404	-0.9852	0.4110	0.8820	2.2109

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-14.18265	2.06382	-6.872	0.000000000000633	***
hsgpa	0.21939	0.02988	7.342	0.000000000000021	***
hsengl	-0.03534	0.01766	-2.001	0.04539	*
c1	-1.29137	0.45190	-2.858	0.00427	**
c2	0.75847	0.49308	1.538	0.12399	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 530.66 on 393 degrees of freedom
Residual deviance: 424.76 on 389 degrees of freedom
AIC: 434.76

Number of Fisher Scoring iterations: 4

```
> ls()
[1] "math" "math2" "modell"
```

This document was prepared by [Jerry Brunner](#), University of Toronto. It is licensed under a Creative Commons Attribution - ShareAlike 3.0 Unported License:

http://creativecommons.org/licenses/by-sa/3.0/deed.en_US. Use any part of it as you like and share the result freely. It is available in OpenOffice.org format from the course website: <http://www.utstat.toronto.edu/brunner/oldclass/312f22>