```
> library(haven)          # load the package that contain functions to read
SAS files into R environment
> library(plm)            # package that can process panel data in R
> library(margins)        # package that can compute the marginal effect of
 model
>
> file<-read_sas("college_admission.sas7bdat") # read SAS file into R
>
> file$COLLEGE<-file$PSECHOICE!=1
> file$COLLEGE<-as.numeric(file$COLLEGE)        # set up COLLEGE variable
>
> probit<-glm(COLLEGE~GRADES+FAMINC+FAMSIZ+PARCOLL+FEMALE+BLACK,family=b
inomial(link="probit"),data=file)
> summary(probit)                               # probit model

Call:
glm(formula = COLLEGE ~ GRADES + FAMINC + FAMSIZ + PARCOLL +
    FEMALE + BLACK, family = binomial(link = "probit"), data = file)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-3.6547   0.0891   0.3615   0.6878   1.7145

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  2.693704   0.285643   9.430  < 2e-16 ***
GRADES      -0.294552   0.027452 -10.730  < 2e-16 ***
FAMINC       0.005392   0.001976   2.728 0.006366 **
FAMSIZ      -0.053108   0.037349  -1.422 0.155047
PARCOLL      0.476540   0.141154   3.376 0.000735 ***
FEMALE       0.023788   0.101626   0.234 0.814926
BLACK        0.610896   0.215740   2.832 0.004631 **
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1058.86 on 999 degrees of freedom
Residual deviance: 832.44 on 993 degrees of freedom
AIC: 846.44

Number of Fisher Scoring iterations: 6
```

```
>
> logit<-glm(COLLEGE~GRADES+FAMINC+FAMSIZ+PARCOLL+FEMALE+BLACK,family=bi
nomial(link="logit"),data=file)
> summary(logit)                          # logit model

Call:
glm(formula = COLLEGE ~ GRADES + FAMINC + FAMSIZ + PARCOLL +
    FEMALE + BLACK, family = binomial(link = "logit"), data = file)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-3.3697   0.1252   0.3579   0.6656   1.7735

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) 4.541835   0.519120   8.749 < 2e-16 ***
GRADES     -0.517438   0.050464 -10.254 < 2e-16 ***
FAMINC      0.013042   0.003897   3.346 0.000819 ***
FAMSIZ     -0.088256   0.064466  -1.369 0.170987
PARCOLL     0.835371   0.262977   3.177 0.001490 **
FEMALE      0.034279   0.176918   0.194 0.846365
BLACK       1.073344   0.372292   2.883 0.003938 **
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1058.86 on 999 degrees of freedom
Residual deviance: 829.95 on 993 degrees of freedom
AIC: 843.95

Number of Fisher Scoring iterations: 6

>
> sum(file$COLLEGE)/length(file$COLLEGE)       # the rate that attend coll
ege
[1] 0.778
> mean(file$FAMINC)                            # sample average of family income
variable
[1] 51.3935
```

```
> x<-data.frame("GRADES"=10,"FAMINC"=51.3935,"FAMSIZ"=4,"PARCOLL"=1,"FEM
ALE"=1,"BLACK"=1)
> predict.glm(probit,newdata=x)          # predict the result of a new obs
ervation with probit model
      1
0.924101
> marginal_effects(probit,x)            # see the marginal effect of the
variables in this observation
  dydx_GRADES dydx_FAMINC dydx_FAMSIZ dydx_PARCOLL dydx_FEMALE dydx_BLACK
1 -0.07667187 0.00140358 -0.01382398    0.1240433 0.006192081 0.1590163
>
>
> data<-read_sas("college2.sas7bdat")     # read the other SAS file to R
>
> pdata<-plm.data(data,index=c("hh","t"))      # set up panel data
>
> ols<-lm(l~x,data=pdata)                        # ordinary least squar
e model
> pool<-plm(l~x,data=pdata,model="pooling")
> fixed<-plm(l~x,data=pdata,model="within")
> summary(fixed)                         # fixed effect model
Oneway (individual) effect Within Model

Call:
plm(formula = l ~ x, data = pdata, model = "within")

Balanced Panel: n=40, T=3, N=120

Residuals :
   Min. 1st Qu. Median 3rd Qu.    Max.
-1.6700 -0.5490 -0.0494 0.5540 2.0500

Coefficients :
  Estimate Std. Error t-value Pr(>|t|)
x 0.020742   0.020908 0.9921   0.3242

Total Sum of Squares:    77.107
Residual Sum of Squares: 76.159
R-Squared      : 0.012305
     Adj. R-Squared : 0.0081005
F-statistic: 0.984166 on 1 and 79 DF, p-value: 0.3242
>
```

```
> random<-plm(l~x,data=pdata,model="random")
> summary(random)                            # random effect model
Oneway (individual) effect Random Effect Model
   (Swamy-Arora's transformation)

Call:
plm(formula = l ~ x, data = pdata, model = "random")

Balanced Panel: n=40, T=3, N=120

Effects:
              var std.dev share
idiosyncratic 0.9640 0.9819 0.571
individual    0.7251 0.8515 0.429
theta:  0.4459

Residuals :
   Min. 1st Qu. Median 3rd Qu.    Max.
-2.2600 -0.6970 0.0787 0.5530 2.2300

Coefficients :
            Estimate Std. Error t-value Pr(>|t|)
(Intercept) 0.9690324 0.5210052 1.8599 0.0653859 .
x           0.0265755 0.0070126 3.7897 0.0002391 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:   126.61
Residual Sum of Squares: 112.88
R-Squared      : 0.1085
     Adj. R-Squared : 0.1067
F-statistic: 14.3618 on 1 and 118 DF, p-value: 0.00023906
>
> pFtest(fixed,pool)                        # test the hypothesis that individu
al intercepts all equal


        F test for individual effects

data: l ~ x
F = 3.1755, df1 = 39, df2 = 79, p-value = 6.647e-06
alternative hypothesis: significant effects
```