Example 2a: Reviewing Main Effects in General Linear Models (as estimated using restricted maximum likelihood in SAS MIXED and STATA MIXED)

The models for this example come from Hoffman (2015) chapter 2. We will be examining the extent to which cognition (as measured by an information test outcome) can be predicted from age (centered at 85 years) grip strength (centered at 9 pounds), sex (with men as the reference group) and subsequent dementia status (none = 1, future = 2, and current = 3) in a sample of 550 older adults.

SAS Syntax and Output for Data Manipulation and Data Description:

```
* Defining global variable for file location to be replaced in code below;
* \\Client\ precedes actual path when using UIowa Virtual Desktop
%LET filesave=\\Client\C:\Dropbox\19_PSQF7375_Clustered\PSQF7375_Clustered_Example2;
* Location for SAS files for these models (uses macro variable filesave);
LIBNAME filesave "&filesave.";

* Import chapter 2 example data into work library;
DATA work.Chapter2; SET filesave.SAS_Chapter2;
* Centering continuous predictors;
age85 = age - 85;
grip9 = grip - 9;
* Creating manual contrasts for dementia groups (to be treated as continuous);
     IF demgroup=1 THEN DO; demNF=0; demNC=0; END; * None group is reference;
ELSE IF demgroup=2 THEN DO; demNF=1; demNC=0; END; * Future group difference;
ELSE IF demgroup=3 THEN DO; demNF=0; demNC=1; END; * Current group difference;
* Labeling all variables – note semi-colon is only at the end of ALL labels;
LABEL
age85=  "age85: Age in Years (0=85)"
grip9=  "grip9: Grip Strength in Pounds (0=9)"
sexMW=  "sexMW: Sex (0=Men, 1=Women)"
demNF=  "demNF: Dementia Contrast for None=0 vs Future=1"
demNC=  "demNC: Dementia Contrast for None=0 vs Current=1"
cognition= "Cognition Outcome"
demgroup=  "Dementia Group 1N 2F 3C";
RUN;

* Creating value labels to use as needed (not stored in data);
PROC FORMAT; VALUE FDemGroup 1="1None" 2="2Future" 3="3Current";
             VALUE FSex 0="0Men" 1="1Women"; RUN;

TITLE1 "Chapter 2: Descriptive Statistics for Example Variables";
PROC MEANS NOLABEL NDEC=2 NONOBS DATA=work.Chapter2; VAR age grip cognition; RUN;
PROC FREQ  DATA=work.Chapter2;
     TABLE sexMW*demgroup; RUN;
PROC CORR  DATA=work.Chapter2; VAR age grip sexMW cognition; RUN;
TITLE1;
```

STATA Syntax and Output for Data Manipulation and Data Description:

```
* Defining global variable for file location to be replaced in code below
global filesave "C:\Dropbox\19_PSQF7375_Clustered\PSQF7375_Clustered_Example2"

* Import chapter 2 data into temporary file and center predictors
use "$filesave\STATA_Chapter2.dta", clear
* Centering continuous predictors
gen age85 = age - 85
gen grip9 = grip - 9
* Creating manual contrasts for dementia groups
gen demnf=0
gen demnc=0
* Demgroup = none
replace demnf=0 if demgroup==1
replace demnc=0 if demgroup==1
```

```
* Demgroup = future
replace demnf=1 if demgroup==2
replace demnc=0 if demgroup==2
* Demgroup = current
replace demnf=0 if demgroup==3
replace demnc=1 if demgroup==3
* Adding value labels
label define fdemgoup 1 "1None" 2 "2Future" 3 "3Current"
label values demgroup fdemgroup
label define fsex 0 "0Men" 1 "1Women"
label values sexmw fsex
* Labeling all variables
label variable age85  "age85: Age in Years (0=85)"
label variable grip9  "grip9: Grip Strength in Pounds (0=9)"
label variable sexmw  "sexmw: Sex (0=Men, 1=Women)"
label variable demnf  "demnf: Dementia Contrast for None=0 vs Future=1"
label variable demnc  "demnc: Dementia Contrast for None=0 vs Current=1"
label variable cognition  "Cognition Outcome"
label variable demgroup   "Dementia Group 1N 2F 3C"

display as result "Chapter 2: Descriptive Statistics for Example Variables"
format age grip cognition %4.2f
summarize age grip cognition, format
tabulate sexmw demgroup, cell
```

SAS Old-School Listing Output (which I still use because it's easier to paste and annotate):

Variable	N	Mean	Std Dev	Minimum	Maximum
age	550	84.93	3.43	80.02	96.97
grip	550	9.11	2.98	0.00	19.00
cognition	550	24.82	10.99	0.00	44.00

sexMW(sexMW: Sex (0=M, 1=W))
demgroup(Dementia Group 1N 2F 3C)

Frequency Percent	1None	2Future	3Current	Total
0Men	168	40	19	227
	30.55	7.27	3.45	41.27
1Women	231	69	23	323
	42.00	12.55	4.18	58.73
Total	399	109	42	550
	72.55	19.82	7.64	100.00

STATA Old-School Listing Output (which I still use because it's easier to paste and annotate):

Variable	Obs	Mean	Std. Dev.	Min	Max
age	550	84.93	3.43	80.02	96.97
grip	550	9.11	2.98	0.00	19.00
cognition	550	24.82	10.99	0.00	44.00

sexmw: Sex (0=Men, 1=Women)	Dementia Group 1N 2F 3C 1None	2Future	3Current	Total
0Men	168	40	19	227
	30.55	7.27	3.45	41.27
1Women	231	69	23	323
	42.00	12.55	4.18	58.73
Total	399	109	42	550
	72.55	19.82	7.64	100.00

SAS MIXED Syntax and Output for Empty Means Model in Equation 2.3

$$\text{Cognition}_i = \beta_0 + e_i$$

```
TITLE1 'Eq 2.3: Empty Means Model';

PROC MIXED DATA=work.Chapter2 COVTEST NOCLPRINT NAMELEN=100 IMETHOD=REML;

     MODEL cognition =  / SOLUTION DDFM=BW;

RUN; TITLE1;
```

METHOD = REML → Least Squares (= GLM)
MODEL y = fixed effects of predictors
Options after / ; add "CL" for fixed effect Cis (not used here to save space for annotation)
DDFM indicates denominator DF method (doesn't matter here, but it will later)

Number of Observations

Number of Observations Read	550
Number of Observations Used	550
Number of Observations Not Used	0

This table tells you how many cases were removed due to incomplete data—make sure you pay attention to this if you are doing any model comparisons (which will need to be based on the exact same cases to be valid).

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error	Z Value	Pr > Z
Residual	120.76	7.2887	16.57	<.0001 **is σ_e^2**

This table will list all estimated parameters within the model for the variance. Right now all we have is residual variance, the variance of the e_i residuals. Because this is an empty means model with no predictors, this is ALL the variance to be predicted in cognition.

Fit Statistics

-2 Res Log Likelihood	4196.1
AIC (Smaller is Better)	4198.1
AICC (Smaller is Better)	4198.1
BIC (Smaller is Better)	4202.4

This first "Fit Statistics" table will index relative model fit in terms of **−2*LL**, in which **smaller is better** (stay tuned).

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	24.8218	0.4686	549	52.97	<.0001 is B0

This "Solution for Fixed Effects" table will list all estimated parameters in the model for the means.

It is not printed by default and must be requested by "SOLUTION" after the / on the MODEL line above.

Interpret each fixed effect:

Intercept β_0 =

STATA MIXED Syntax and Output for Empty Means Model in Equation 2.3

$$\text{Cognition}_i = \beta_0 + e_i$$

```
display as result "Eq 2.3: Empty Means Model"
mixed cognition , ///
      variance reml dfmethod(residual)
```

MODEL y fixed effects of predictors
/// used to continue single command across lines; second line are options:

- METHOD = REML → Least Squares (= GLM)
- Stata used to provide SD instead of variances by default, so "variance" ensures this version
- dfmethod invokes denominator DF (→ t and F)

```
Mixed-effects REML regression                 Number of obs     =        550
DF method: Residual                           DF:           min =     549.00
                                                            avg =     549.00
                                                            max =     549.00
                                              F(0,   549.00)    =          .
Log restricted-likelihood = -2098.0488        Prob > F          =          .
```

F (0,549) is where multivariate Wald test for model R^2 will go once we have model predictors.

Note that STATA provides log-likelihood (**LL**), in which **bigger is better**, whereas SAS provides **−2*LL** instead (in which **smaller is better**). The latter is more useful to simplify the math in variance model comparisons (stay tuned).

```
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   24.82182   .4685737    52.97   0.000      23.9014    25.74223
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
               var(Residual) |   120.7587   7.282025      107.2974    135.9089
------------------------------------------------------------------------------
```

Fixed effects are listed first, then the intercept (Beta0) which is labeled as "_cons" instead.

This table will list all estimated parameters in the model for the variance.

SAS MIXED Syntax and Output for Age, Grip, and Sex (0=M, 1=W) Model in Equation 2.7

$$\text{Cognition}_i = \beta_0 + \beta_1(\text{Age}_i - 85) + \beta_2(\text{Grip}_i - 9) + \beta_3(\text{SexMW}_i) + e_i$$

```
TITLE1 'Eq 2.7: Age + Grip + Sex (0=M 1=W, as continuous predictor)';
PROC MIXED DATA=work.Chapter2 COVTEST NOCLPRINT NAMELEN=100 METHOD=REML;
     MODEL cognition = age85 grip9 sexMW
                     / SOLUTION DDFM=BW
                       OUTPM=PredAgeGripSex;
CONTRAST "Model R2 F-Test" age85 1, grip9 1, sexMW 1;
RUN; TITLE1;

PROC CORR NOSIMPLE DATA=work.PredAgeGripSex;
     VAR cognition; WITH pred; RUN;
```

OUTPM requests a new dataset (after =) that contains predicted outcomes from the fixed effects. The new column is called "pred" by default.

CONTRAST gives a multivariate Wald test for the significance of the model for the means (reduction in error variance by adding three fixed effects).

The CORR command requests the correlation between the predicted and original outcome. Squaring that correlation creates model R^2.

```
            Covariance Parameter Estimates
                        Standard         Z
Cov Parm     Estimate       Error     Value      Pr > Z
Residual       109.38      6.6200     16.52      <.0001 is new σ²e
```

```
                  Solution for Fixed Effects
                        Standard
Effect       Estimate       Error      DF    t Value    Pr > |t|
Intercept     26.9594      0.7389     546      36.49      <.0001  is B0
age85         -0.4338      0.1325     546      -3.27      0.0011  is B1
grip9          0.5460      0.1663     546       3.28      0.0011  is B2
sexMW         -3.7988      0.9904     546      -3.84      0.0001  is B3
```

```
                    Contrasts
                   Num      Den
Label               DF       DF     F Value     Pr > F
Model R2 F-Test      3      546       20.04     <.0001

Pearson Correlation Coefficients, N = 550
        Prob > |r| under H0: Rho=0

                    cognition
Pred                  0.31491 → squared = .0992
Predicted Mean         <.0001
```

FYI: In practice, the R^2 calculated as a reduction in residual variance will differ slightly from the R^2 calculated as the square of the correlation between the predicted and original outcome—the latter exactly matches the GLM output for model R^2.

We will use both of these techniques for different reasons in the multilevel models to come.

Calculate model R^2 as proportion reduction in residual (error) variance (PRE)
= (empty σ^2_e – current σ^2_e) / (empty σ^2_e) = (120.76 − 109.38) / (120.76) = **.0942**
The df=3 CONTRAST above says that this R^2 is significantly > 0, $F(3,546) = 20.04$, $p < .0001$.

STATA MIXED Syntax and Output for Age, Grip, and Sex (0=M, 1=W) Model in Equation 2.7

$$\text{Cognition}_i = \beta_0 + \beta_1(\text{Age}_i - 85) + \beta_2(\text{Grip}_i - 9) + \beta_3(\text{SexMW}_i) + e_i$$

```
display as result "Eq 2.7: Age + Grip + Sex (0=M 1=W, as continuous predictor)"
mixed cognition c.age85 c.grip9 c.sexmw, ///
                variance reml dfmethod(residual),
      predict predagegripsex, xb
      corr cognition predagegripsex
      display as result r(rho)^2
```

PREDICT creates a new column called "predagegripsex" in the original data that contains predicted outcomes from the fixed effects.

The CORR command requests the correlation between the predicted and original outcome. Squaring that correlation creates model R^2.

```
Mixed-effects REML regression                   Number of obs     =        550
DF method: Residual                             DF:           min =     546.00
                                                              avg =     546.00
                                                              max =     546.00

                                                F(3,   546.00)    =      20.04
Log restricted-likelihood = -2070.5586          Prob > F          =     0.0000
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       age85 |  -.4337719   .1324638    -3.27   0.001    -.6939729   -.1735709  is B1
       grip9 |   .5460019   .1662766     3.28   0.001     .2193818    .8726221  is B2
       sexmw |   -3.79878   .9903591    -3.84   0.000    -5.744161   -1.853399  is B3
       _cons |   26.95943   .7388729    36.49   0.000     25.50805    28.41081  is B0
------------------------------------------------------------------------------
```

F (3,546) = 20.04 is the multivariate Wald test for the model R^2 given three predictors in the model for the means.

Interpret each fixed effect:

Intercept β_0 =

Main effect of Age β_1 =

Main effect of Grip Strength β_2 =

Main effect of Sex β_3 =

```
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.      [95% Conf. Interval]
-----------------------------+------------------------------------------------
               var(Residual) |   109.3807   6.595904       97.1877    123.1034  is new σ²e
------------------------------------------------------------------------------
```

```
             | cognit~n     pred
-------------+------------------
   cognition |   1.0000
        pred |   0.3149   1.0000

.       display as result r(rho)^2
.09917075
```

.3149 is the correlation between the model-predicted and original outcome → R^2 = **.0992**

SAS MIXED Syntax and Output for the Dementia Group Model in Equation 2.8

$$\text{Cognition}_i = \beta_0 + \beta_1(\text{Age}_i - 85) + \beta_2(\text{Grip}_i - 9) + \beta_3(\text{SexMW}_i) + \beta_4(\text{DemNF}_i) + \beta_5(\text{DemNC}_i) + e_i$$

We can use the model equation to calculate the **dementia group means** for predicted cognition:

Cognition for None = β_0

Cognition for Future = $\beta_0 + \beta_4$

Cognition for Current = $\beta_0 + \beta_5$

Cognition for *Ever* Dementia (Mean of Future and Current) = $\beta_0 + (\beta_4 + \beta_5)/2$

We can determine the **differences between the dementia group means** as follows:

None vs. Future = Future – None = $(\beta_0 + \beta_4) - (\beta_0) = \beta_4$

None vs. Current = Current – None = $(\beta_0 + \beta_5) - (\beta_0) = \beta_5$

Future vs. Current = Current – Future = $(\beta_0 + \beta_5) - (\beta_0 + \beta_4) = \beta_5 - \beta_4 = -\beta_4 + \beta_5$

None vs. Ever = Ever – None = $\beta_0 + (\beta_4 + \beta_5)/2 - (\beta_0) = (\beta_4 + \beta_5)/2$

These values are then requested via the SAS ESTIMATE statements below…

```
TITLE1 'Eq 2.8: Adding Dementia Group';
TITLE2 'Using Manual Group Contrasts so Reference=None';
TITLE3 'sexMW, demNF, and demNC are all treated as continuous predictors';
PROC MIXED DATA=work.Chapter2 COVTEST NOCLPRINT NAMELEN=100 METHOD=REML;
     MODEL cognition = age85 grip9 sexMW demNF demNC
                     / SOLUTION DDFM=BW OUTPM=PredAgeGripSexDem;
```

The first CONTRAST below includes all fixed effects, and thus tests the full model R^2.
The second CONTRAST below includes only the new fixed effects, and thus tests the increment to the model R2 from adding dementia group. In this case this is also an "omnibus" ANOVA test for group.

```
     CONTRAST "Model R2 F-Test df=5" age85 1, grip9 1, sexmw 1, demNF 1, demNC 1;
     CONTRAST "Omnibus F-Test for Dementia Group df=2"          demNF 1, demNC 1;
```

The first 4 ESTIMATEs request predicted outcomes, so they must include the intercept.
The last 4 ESTIMATES request slopes for group differences, so they do NOT include the intercept.

```
* Request conditional (adjusted) group means (hold age=85, grip=9, men because not there);
  ESTIMATE "Intercept for None Group"    intercept 1 demNF  0 demNC 0; * Given as B0;
  ESTIMATE "Intercept for Future Group"  intercept 1 demNF  1 demNC 0; * Not given (B0+B4);
  ESTIMATE "Intercept for Current Group" intercept 1 demNF  0 demNC 1; * Not given (B0+B5);
  ESTIMATE "Intercept for Ever Group"    intercept 1 demNF .5 demNC .5; * B0 + (B4+B5)/2;
```

```
* Request group differences (unconditional because there are no interactions);
  ESTIMATE "None   vs. Future Group"   demNF  1 demNC  0; * Given as B4;
  ESTIMATE "None   vs. Current Group"  demNF  0 demNC  1; * Given as B5;
  ESTIMATE "Future vs. Current Group"  demNF -1 demNC  1; * Not given (B5-B4);
  ESTIMATE "None vs. Ever Group"       demNF .5 demNC .5; * Not given (B5+B4)/2;
RUN; TITLE1; TITLE2; TITLE3;
PROC CORR NOSIMPLE DATA=work.PredAgeGripSexDem;
     VAR cognition; WITH pred; RUN;
```

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error	Z Value	Pr > Z
Residual	88.0709	5.3401	16.49	<.0001 **is new σ_e^2**

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|	
Intercept	29.2643	0.6985	544	41.90	<.0001	is B0
age85	-0.4057	0.1189	544	-3.41	0.0007	is B1
grip9	0.6042	0.1498	544	4.03	<.0001	is B2
sexMW	-3.6574	0.8914	544	-4.10	<.0001	is B3
demNF	-5.7220	1.0191	544	-5.61	<.0001	is B4
demNC	-16.4798	1.5228	544	-10.82	<.0001	is B5

Interpret each fixed effect below:

Intercept β_0 =

Main effect of Age β_1 =

Main effect of Grip Strength β_2 =

Main effect of Sex β_3 =

Main effect of DemNF β_4 =

Main effect of DemNC β_5 =

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|	
Intercept for None Group	29.2643	0.6985	544	41.90	<.0001	B0
Intercept for Future Group	23.5424	1.0785	544	21.83	<.0001	B0+B4
Intercept for Current Group	12.7845	1.5302	544	8.35	<.0001	B0+B5
Intercept for Ever Group	18.1634	1.0115	544	17.96	<.0001	B0+(B4+B5)/2
None vs. Future Group	-5.7220	1.0191	544	-5.61	<.0001	B4
None vs. Current Group	-16.4798	1.5228	544	-10.82	<.0001	B5
Future vs. Current Group	-10.7578	1.7080	544	-6.30	<.0001	B5-B4
None vs. Ever Group	-11.1009	0.9744	544	-11.39	<.0001	(B4+B5)/2

Contrasts

Label	Num DF	Den DF	F Value	Pr > F
Model R2 F-Test with df=5	5	544	41.75	<.0001
Omnibus F-Test for Dementia Group with df=2	2	544	67.06	<.0001

Pearson Correlation Coefficients, N = 550

	cognition
Pred	0.52662 → squared = .2774

Current model R^2 = (empty σ_e^2 – current σ_e^2) / (empty σ_e^2) = (120.76 – 88.07) / (120.76) = .2707
The df=5 CONTRAST above says that current R^2 is significantly > 0, $F(5,544) = 41.75$, $p < .0001$.
Change in model R^2 = (current R^2) − (previous R^2) = .2707 − .0942 = .1765
The df=2 CONTRAST above says that change in R^2 is significantly > 0, $F(2,544) = 67.06$, $p < .0001$.

STATA MIXED Syntax and Output for Dementia Group Model in Equation 2.8

$$\text{Cognition}_i = \beta_0 + \beta_1(\text{Age}_i - 85) + \beta_2(\text{Grip}_i - 9) + \beta_3(\text{SexMW}_i) + \beta_4(\text{DemNF}_i) + \beta_5(\text{DemNC}_i) + e_i$$

```
display as result "Eq 2.8: Adding Dementia Group"
display as result "Using Manual Group Contrasts so Reference=None"
mixed cognition c.age85 c.grip9 c.sexmw c.demnf c.demnc, ///
                variance reml dfmethod(residual),
      test (c.demnf=0) (c.demnc=0), small // Omnibus Dementia Group Test - change in R2
      predict predagegripsexdem, xb
      corr cognition predagegripsexdem
      display as result r(rho)^2
 lincom _cons*1 + c.demnf*0 + c.demnc*0, small // Intercept for None Group = B0
 lincom _cons*1 + c.demnf*1 + c.demnc*0, small // Intercept for Future Group = B0+B4
 lincom _cons*1 + c.demnf*0 + c.demnc*1, small // Intercept for Current Group = B0+B5
 lincom _cons*1 + c.demnf*.5 + c.demnc*.5, small // Intercept for Ever Group = B0+(B4+B5)/2
 lincom c.demnf*1  + c.demnc*0,  small // None vs Future = B4
 lincom c.demnf*0  + c.demnc*1,  small // None vs Current = B5
 lincom c.demnf*-1 + c.demnc*1,  small // Future vs Current = B5-B4
 lincom c.demnf*.5 + c.demnc*.5, small // None vs Ever = (B4+B5)/2
```

The option SMALL requests t-tests instead of z-tests.

```
Mixed-effects REML regression                   Number of obs     =        550
DF method: Residual                             DF:           min =     544.00
                                                              avg =     544.00
                                                              max =     544.00
                                                F(5,   544.00)    =      41.75
Log restricted-likelihood = -2008.1345          Prob > F          =     0.0000
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       age85 |   -.405734   .1188972    -3.41   0.001    -.6392878   -.1721802 is B1
       grip9 |   .6042256   .1497757     4.03   0.000      .310016    .8984351 is B2
       sexmw |  -3.657374   .8914326    -4.10   0.000    -5.408446   -1.906303 is B3
       demnf |  -5.721971   1.019078    -5.61   0.000    -7.723782    -3.72016 is B4
       demnc |  -16.47981   1.522754   -10.82   0.000    -19.47101   -13.48862 is B5
       _cons |   29.26433   .6985079    41.90   0.000     27.89222    30.63643 is B0
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
               var(Residual) |   88.07088   5.310874      78.25335     99.1201 is new σ_e^2
------------------------------------------------------------------------------
```

F (5,544) = 41.75 is the multivariate Wald test for the model R^2 given five model predictors.

```
.       * Omnibus Dementia Group Test (change in R2)
.       test (c.demnf=0) (c.demnc=0), small
 ( 1)  [cognition]demnf = 0
 ( 2)  [cognition]demnc = 0

       F(  2,544.00) =   67.06
            Prob > F =    0.0000

.       corr cognition pred  (obs=550)
             | cognit~n predag~m
-------------+------------------
predagegri~m |   0.5266   1.0000 → .2773294 R2 for model
```

F (2,544) = 67.06 is the multivariate Wald test for the model CHANGE in R^2 given two new model predictors. This is also known as the "omnibus" main effect of group in ANOVA.

Here are the results of the separate LINCOM statements:

```
.           lincom _cons*1 + c.demnf*0 + c.demnc*0, small // Intercept for None Group = B0
 ( 1)  [cognition]_cons = 0
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |   29.26433   .6985079    41.90   0.000     27.89222    30.63643
------------------------------------------------------------------------------

.           lincom _cons*1 + c.demnf*1 + c.demnc*0, small // Intercept for Future Group = B0+B4
 ( 1)  [cognition]demnf + [cognition]_cons = 0
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |   23.54235    1.07853    21.83   0.000     21.42376    25.66095
------------------------------------------------------------------------------

.       lincom _cons*1 + c.demnf*0 + c.demnc*1, small // Intercept for Current Group = B0+B5
 ( 1)  [cognition]demnc + [cognition]_cons = 0
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |   12.78451   1.530193     8.35   0.000     9.778701    15.79032
------------------------------------------------------------------------------

.  lincom _cons*1 + c.demnf*.5 + c.demnc*.5, small // Intercept for Ever Group = B0+(B4+B5)/2
 ( 1)  .5*[cognition]demnf + .5*[cognition]demnc + [cognition]_cons = 0
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |   18.16343   1.011474    17.96   0.000     16.17656    20.15031
------------------------------------------------------------------------------

.           lincom c.demnf*1  + c.demnc*0, small // None vs Future = B4
 ( 1)  [cognition]demnf = 0
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |  -5.721971   1.019078    -5.61   0.000    -7.723782    -3.72016
------------------------------------------------------------------------------

.           lincom c.demnf*0  + c.demnc*1, small // None vs Current = B5
 ( 1)  [cognition]demnc = 0
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |  -16.47981   1.522754   -10.82   0.000    -19.47101   -13.48862
------------------------------------------------------------------------------

.           lincom c.demnf*-1 + c.demnc*1, small // Future vs Current = B5-B4
 ( 1)  - [cognition]demnf + [cognition]demnc = 0
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |  -10.75784   1.707957    -6.30   0.000    -14.11284   -7.402844
------------------------------------------------------------------------------

.  lincom c.demnf*.5 + c.demnc*.5, small // None vs Ever = (B4+B5)/2
 ( 1)  .5*[cognition]demnf + .5*[cognition]demnc = 0
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |  -11.10089   .9743564   -11.39   0.000    -13.01485    -9.18693
------------------------------------------------------------------------------
```

SAS MIXED Syntax and Output for Dementia Group Model in Equation 2.8
Using CLASS statement (SAS-coded contrasts instead of manually created contrasts)

Because the default reference group is the HIGHEST group numerically or last alphabetically, I have changed the model to reflect "Current" (group=3) as the reference:

$$\text{Cognition}_i = \beta_0 + \beta_1(\text{Age}_i - 85) + \beta_2(\text{Grip}_i - 9) + \beta_3(\text{SexMW}_i) + \beta_4(\text{DemCN}_i) + \beta_5(\text{DemCF}_i) + e_i$$

```
TITLE1 'Eq 2.8: Adding Dementia Group';
TITLE2 'Categorical Predictor for Dementia Group on CLASS statement';
PROC MIXED DATA=work.Chapter2 COVTEST NOCLPRINT NAMELEN=100 METHOD=REML;
     CLASS demgroup; * CLASS statement demgroup replaces previous dem contrasts;
     FORMAT demgroup Fdemgroup.; * Use value labels defined earlier in output;
     MODEL cognition = age85 grip9 sexMW demgroup / SOLUTION DDFM=BW;
     CONTRAST "Model R2 F-Test with df=5" age85 1, grip9 1, sexmw 1,
                                          demgroup -1 1 0, demgroup -1 0 1;
* Request conditional (adjusted) group means (hold age=85, grip=9, men) and all diffs;
     LSMEANS demgroup / DIFF=ALL AT(age85 grip9 sexMW) = (0 0 0);
* Request conditional (adjusted) group means and all differences for demonstration;
     LSMEANS demgroup / DIFF=ALL;
*** All of the code below is redundant with LSMEANS, but here is how you get all the info;
*** The exceptions are the lines that refer to the "ever" group, which is not a default;
  CONTRAST "Omnibus F-Test for Dementia Group with df=2"  demgroup -1 1 0, demgroup -1 0 1;
* Request conditional (adjusted) group means (hold age=85, grip=9, men);
  ESTIMATE "Intercept for None Group"    intercept 1 demgroup 1 0 0; * Not given (B0+B4);
  ESTIMATE "Intercept for Future Group"  intercept 1 demgroup 0 1 0; * Not given (B0+B5);
  ESTIMATE "Intercept for Current Group" intercept 1 demgroup 0 0 1; * Given as B0;
  ESTIMATE "Intercept for Ever Group"    intercept 1 demgroup 0 .5 .5; * (B0+B5)/2;
* Request group differences (unconditional because there are no interactions);
  ESTIMATE "None   vs. Future Group"   demgroup -1  1  0; * Not given (B5-B4);
  ESTIMATE "None   vs. Current Group"  demgroup -1  0  1; * Given as B5;
  ESTIMATE "Future vs. Current Group"  demgroup  0 -1  1; * Given as B4;
  ESTIMATE "None   vs. Ever Group"     demgroup -1 .5 .5; * (B0+B5)/2 - B4;
RUN; TITLE1; TITLE2;
```

New output after using the CLASS statement for demgroup:

The row with the dot indicates which group is the reference. The other rows then indicate group mean differences relative to the reference group.

Effect	Dementia Group 1N 2F 3C	Estimate	Standard Error	DF	t Value	Pr > |t|	
Intercept		12.7845	1.5302	544	8.35	<.0001	new B0
age85		-0.4057	0.1189	544	-3.41	0.0007	same B1
grip9		0.6042	0.1498	544	4.03	<.0001	same B2
sexMW		-3.6574	0.8914	544	-4.10	<.0001	same B3
demgroup	1None	16.4798	1.5228	544	10.82	<.0001	new B4
demgroup	2Future	10.7578	1.7080	544	6.30	<.0001	new B5
demgroup	3Current	0	.	.	.	.	new ref group

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
age85	1	544	11.65	0.0007
grip9	1	544	16.27	<.0001
sexMW	1	544	16.83	<.0001
demgroup	**2**	**544**	**67.06**	**<.0001**

The multivariate Wald test for the omnibus main effect of demgroup is now given by default when using CLASS for demgroup.

Contrasts

Label	Num DF	Den DF	F Value	Pr > F
Model R2 F-Test with df=5	5	544	41.75	<.0001
Omnibus F-Test for Dementia Group with df=2	**2**	**544**	**67.06**	**<.0001 now given by default**

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept for None Group	29.2643	0.6985	544	41.90	<.0001
Intercept for Future Group	23.5424	1.0785	544	21.83	<.0001
Intercept for Current Group	12.7845	1.5302	544	8.35	<.0001
None vs. Future Group	-5.7220	1.0191	544	-5.61	<.0001
None vs. Current Group	-16.4798	1.5228	544	-10.82	<.0001
Future vs. Current Group	-10.7578	1.7080	544	-6.30	<.0001

These group means and mean differences, which we requested, are given through LSMEANS below (with less typing).

Least Squares Means

Effect	Dementia Group 1N 2F 3C	age85	grip9	sexMW	Estimate	Standard Error	DF	t Value	Pr > |t|
demgroup	1None	0.00	0.00	0.00	29.2643	0.6985	544	41.90	<.0001
demgroup	2Future	0.00	0.00	0.00	23.5424	1.0785	544	21.83	<.0001
demgroup	3Current	0.00	0.00	0.00	12.7845	1.5302	544	8.35	<.0001
demgroup	1None	-0.07	0.11	0.59	27.2143	0.4702	544	57.88	<.0001
demgroup	2Future	-0.07	0.11	0.59	21.4923	0.9024	544	23.82	<.0001
demgroup	3Current	-0.07	0.11	0.59	10.7345	1.4486	544	7.41	<.0001

The first three rows are adjusted group means at the specified levels of the other predictors.
The second three rows are group means at the sample mean values of the other predictors instead.

Differences of Least Squares Means

Dementia Group 1N 2F 3C	Dementia Group 1N 2F 3C	age85	grip9	sexMW	Estimate	Standard Error	DF	t Value	Pr > |t|
1None	2Future	0.00	0.00	0.00	5.7220	1.0191	544	5.61	<.0001
1None	3Current	0.00	0.00	0.00	16.4798	1.5228	544	10.82	<.0001
2Future	3Current	0.00	0.00	0.00	10.7578	1.7080	544	6.30	<.0001

The first three rows are adjusted group mean differences at the specified levels of the other predictors.
The second three rows are group mean differences at the sample mean values of the other predictors instead.
Because demgroup does not interact with these other predictors, the group mean differences are the same.

1None	2Future	-0.07	0.11	0.59	5.7220	1.0191	544	5.61	<.0001
1None	3Current	-0.07	0.11	0.59	16.4798	1.5228	544	10.82	<.0001
2Future	3Current	-0.07	0.11	0.59	10.7578	1.7080	544	6.30	<.0001

So to CLASS or not to CLASS? Either can work in every circumstance. The use of CLASS for categorical predictors can be more convenient in models with more than one categorical predictor (e.g., to get marginal and cell means for factorial designs), whereas manual group contrasts can be more convenient when most other predictors are continuous, or when some of your effects pertain to only some levels of the grouping variable (i.e., nested effects; stay tuned).

See example results section in chapter 2 of Hoffman (2015) for text describing the results.

STATA MIXED Syntax and Output for Dementia Group Model in Equation 2.8

Using factor variable statement (STATA-coded contrasts instead of manually created contrasts)

To match the SAS results, I changed the reference group to the HIGHEST group numerically or last alphabetically, I have changed the model to reflect "Current" (group=3) as the reference:

$$\text{Cognition}_i = \beta_0 + \beta_1(\text{Age}_i - 85) + \beta_2(\text{Grip}_i - 9) + \beta_3(\text{SexMW}_i) + \beta_4(\text{DemCN}_i) + \beta_5(\text{DemCF}_i) + e_i$$

```
display as result "Eq 2.8: Adding Dementia Group"
display as result "Using Manual Group Contrasts so Reference=None"
mixed cognition c.age85 c.grip9 c.sexmw ib(last).demgroup, ///
                variance reml dfmethod(residual),

      contrast i.demgroup, small
      margins  i.demgroup, at(c.age85=0 c.grip9=0 c.sexmw=0) df(544)
      margins  i.demgroup, df(544)
      margins  i.demgroup, pwcompare(pveffects) df(544)
lincom _cons*1 + 1.demgroup*0  + 2.demgroup*.5 + 3.demgroup*.5, small // Ever = (B0+B5)/2
lincom           1.demgroup*-1 + 2.demgroup*.5 + 3.demgroup*.5, small // NvE=(B0+B5)/2 -B4
```

The option SMALL cannot be used on MARGINS. You must provide the exact denominator DF instead or use LINCOM to get them instead.

```
Mixed-effects REML regression                   Number of obs     =        550
DF method: Residual                             DF:           min =     544.00
                                                              avg =     544.00
                                                              max =     544.00

                                                F(5,   544.00)    =      41.75
Log restricted-likelihood = -2008.1345          Prob > F          =     0.0000
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       age85 |   -.405734   .1188972    -3.41   0.001    -.6392878   -.1721802 same B1
       grip9 |   .6042256   .1497757     4.03   0.000      .310016    .8984351 same B2
       sexmw |  -3.657374   .8914326    -4.10   0.000    -5.408446   -1.906303 same B3
             |
    demgroup |
      1None  |   16.47981   1.522754    10.82   0.000     13.48862    19.47101  new B4
    2Future  |   10.75784   1.707957     6.30   0.000     7.402844    14.11284  new B5
             |
       _cons |   12.78451   1.530193     8.35   0.000     9.778701    15.79032  new B0
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
               var(Residual) |   88.07088   5.310874      78.25335     99.1201 same σ²e
------------------------------------------------------------------------------
```

```
.       contrast i.demgroup, small
Contrasts of marginal linear predictions
Margins      : asbalanced
------------------------------------------------------------
             |         df        ddf           F        P>F
-------------+----------------------------------------------
cognition    |
    demgroup |          2     544.00       67.06     0.0000
------------------------------------------------------------
```

F (2,544) = 67.06 is the multivariate Wald test for the model CHANGE in R^2 given two new model predictors. This is also known as the "omnibus" main effect of group in ANOVA.

```
.       margins  i.demgroup, at(c.age85=0 c.grip9=0 c.sexmw=0) df(544)
Adjusted predictions                            Number of obs     =        550
Expression   : Linear prediction, fixed portion, predict()
at           : age85           =           0
               grip9           =           0
               sexmw           =           0
------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    demgroup |
      1None  |   29.26433   .6985079    41.90   0.000     27.89222    30.63643
    2Future  |   23.54235    1.07853    21.83   0.000     21.42376    25.66095
   3Current  |   12.78451   1.530193     8.35   0.000     9.778701    15.79032
------------------------------------------------------------------------------

.           margins  i.demgroup, df(544)
Predictive margins                              Number of obs     =        550
Expression   : Linear prediction, fixed portion, predict()
------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    demgroup |
      1None  |   27.21427   .4702141    57.88   0.000     26.29061    28.13792
    2Future  |    21.4923    .902428    23.82   0.000     19.71963    23.26497
   3Current  |   10.73445   1.448621     7.41   0.000     7.888878    13.58003
------------------------------------------------------------------------------

.       margins  i.demgroup, pwcompare(pveffects) df(544)
Pairwise comparisons of predictive margins
Expression   : Linear prediction, fixed portion, predict()
------------------------------------------------------------
                     |            Delta-method    Unadjusted
                     |   Contrast   Std. Err.      t    P>|t|
---------------------+--------------------------------------
            demgroup |
   2Future vs 1None  |  -5.721971   1.019078    -5.61   0.000
  3Current vs 1None  |  -16.47981   1.522754   -10.82   0.000
3Current vs 2Future  |  -10.75784   1.707957    -6.30   0.000
------------------------------------------------------------

. lincom _cons*1 + 1.demgroup*0  + 2.demgroup*.5 + 3.demgroup*.5, small // Intercept Ever = (B0+B5)/2
 ( 1)  .5*[cognition]2.demgroup + .5*[cognition]3b.demgroup + [cognition]_cons = 0
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |   18.16343   1.011474    17.96   0.000     16.17656    20.15031
------------------------------------------------------------------------------

. lincom           1.demgroup*-1 + 2.demgroup*.5 + 3.demgroup*.5, small // None vs Ever = (B0+B5)/2 -B4
 ( 1)  - [cognition]1.demgroup + .5*[cognition]2.demgroup + .5*[cognition]3b.demgroup = 0
------------------------------------------------------------------------------
   cognition |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |  -11.10089   .9743564   -11.39   0.000    -13.01485    -9.18693
------------------------------------------------------------------------------
```