Examples of Adding Predictors to Multivariate Models
Uses of the ESTIMATE Statement with the CLASS Statement
Comparisons of Multivariate Models with Classical MANOVA

The data for this example come from http://www.ats.ucla.edu/stat/sas/library/repeated_ut.htm

SAS Syntax for Data Manipulation – Original Wide Format:

```
*VALUE LABELS FOR EACH LEVEL OF OUR CATEGORICAL VARIABLES;
PROC FORMAT;
VALUE exercises
    1 = "Aerobic Stair Climbing"
    2 = "Racquetball"
    3 = "Weight Training";

VALUE diets
    1 = "Meat Eaters"
    2 = "Vegetarians";

VALUE intensities
      1 = "Pulse: Warm Up"
      2 = "Pulse: Jogging"
      3 = "Pulse: Running";
RUN;


DATA work.dietwide;
INPUT exertype pulse1 pulse2 pulse3 diet personID;

*LABELING VARIABLES;
LABEL exertype = "Exercise Type"
      pulse1 = "Pulse After Warmup"
      pulse2 = "Pulse After Jogging"
      pulse3 = "Pulse After Running"
      diet = "Diet Type";

*ADDING VALUE LABELS TO VARIABLES;
FORMAT exertype exercises. diet diets.;

*CREATING DUMMY CODED VARIABLES FOR EXERCISE TYPE:;
IF exertype = 1 THEN DO; dEXERCISE_ASC = 1; dEXERCISE_R = 0; dEXERCISE_WT = 0; END;
IF exertype = 2 THEN DO; dEXERCISE_ASC = 0; dEXERCISE_R = 1; dEXERCISE_WT = 0; END;
IF exertype = 3 THEN DO; dEXERCISE_ASC = 0; dEXERCISE_R = 0; dEXERCISE_WT = 1; END;


*CREATING DUMMY CODED VARIABLES FOR DIET TYPE;
IF diet = 1 THEN DO; dDIET_M = 1; dDIET_V = 0; END;
IF diet = 2 THEN DO; dDIET_M = 0; dDIET_V = 1; END;

*IMPORTING DATA;
DATALINES;
1         112        166        215        1       1
...

3         78         110        164        2      18
;
RUN;
```

SAS Syntax for Data Manipulation – Converting from Wide Format to Long/Stacked Format:

```
*CONVERTING DATA TO STACKED FORM FOR PROC MIXED;
DATA WORK.dietstack;
    SET WORK.dietwide;

       FORMAT intensity intensities.; *ADDING A FORMAT STATEMENT FOR INTENSITY VARIABLE;

       *FIRST OUTCOME: PULSE 1 (AFTER WARM UP);
       pulse = pulse1;
    intensity = 1;
       dINTENSITY_W = 1; dINTENSITY_J = 0; dINTENSITY_R = 0; *DUMMY CODED VARIABLES FOR ANALYSIS;
    OUTPUT; *OUTPUT MAKES THE LINE OF DATA GET WRITTEN TO THE NEW DATA SET;

       *SECOND OUTCOME: PULSE 2 (AFTER JOGGING);
    pulse = pulse2;
    intensity = 2;
       dINTENSITY_W = 0; dINTENSITY_J = 1; dINTENSITY_R = 0;
    OUTPUT;

       *THIRD OUTCOME: PULSE 3 (AFTER RUNNING);
    pulse = pulse3;
    intensity = 3;
       dINTENSITY_W = 0; dINTENSITY_J = 0; dINTENSITY_R = 1;
    OUTPUT;

RUN;
```

Resulting SAS Data Set (personID = 1 and personID =18 shown)

Obs	personID	pulse	intensity	dINTENSITY_W	dINTENSITY_J	dINTENSITY_R	diet	dDIET_V	dDIET_M
1	1	112	Pulse: Warm Up	1	0	0	Meat Eaters	0	1
2	1	166	Pulse: Jogging	0	1	0	Meat Eaters	0	1
3	1	215	Pulse: Running	0	0	1	Meat Eaters	0	1
52	18	78	Pulse: Warm Up	1	0	0	Vegetarians	1	0
53	18	110	Pulse: Jogging	0	1	0	Vegetarians	1	0
54	18	164	Pulse: Running	0	0	1	Vegetarians	1	0

PROC MIXED is modeling $\boldsymbol{y}_p = \begin{bmatrix} y_{p1} \\ y_{p2} \\ y_{p3} \end{bmatrix}$, where $\boldsymbol{y}$ is the pulse rate. For personID = 1 this would be: $\boldsymbol{y}_1 = \begin{bmatrix} 112 \\ 166 \\ 215 \end{bmatrix}$

Empty multivariate model predicting the mean pulse rate for each intensity level:

```
*MODEL FROM LAST CLASS: MULTIVARIATE EMPTY MODEL/UN R MATRIX STRUCTURE;
TITLE "EMPTY MULTIVARIATE MODEL, FULL UNSTRUCTURED ERROR VARIANCE MODEL:
       (PREDICTORS ARE INDICATORS OF WHICH VARIABLE) - MULTIVARIATE ANOVA ASSUMPTION";
PROC MIXED DATA=WORK.dietstack METHOD=ML COVTEST NOPROFILE ITDETAILS IC NAMELEN=50;
MODEL pulse = dINTENSITY_W dINTENSITY_R / S DDFM=KENWARDROGER;
REPEATED / SUBJECT=personID TYPE=UN R RCORR;
ODS OUTPUT INFOCRIT=info_model1 R=r_model1;
RUN;
TITLE;
```

Multivariate model becomes (simultaneously): $\boldsymbol{y}_p = \mathbf{X}_p\boldsymbol{\beta} + \boldsymbol{e}_p$

$$\boldsymbol{y}_p = \begin{bmatrix} y_{p1} \\ y_{p2} \\ y_{p3} \end{bmatrix}; \mathbf{X}_p = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}; \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_{I1} \\ \beta_{I3} \end{bmatrix}; \boldsymbol{e}_p = \begin{bmatrix} e_{p1} \\ e_{p2} \\ e_{p3} \end{bmatrix}$$

Where $\boldsymbol{e}_p \sim N_3(\mathbf{0}, \mathbf{R})$, and $\mathbf{R} = \begin{bmatrix} \sigma^2_{e1} & \sigma_{e_1,e_2} & \sigma_{e_1,e_3} \\ \sigma_{e_1,e_2} & \sigma^2_{e2} & \sigma_{e_2,e_3} \\ \sigma_{e_1,e_3} & \sigma_{e_2,e_3} & \sigma^2_{e3} \end{bmatrix}$ (the unstructured error covariance matrix model).

Which leads to:

$$\begin{aligned} y_{p1} &= \beta_0 + \beta_{I1} + e_{p1} \\ y_{p2} &= \beta_0 + e_{p2} \\ y_{p3} &= \beta_0 + \beta_{I3} + e_{p3} \end{aligned}$$

Dimensions	
Covariance Parameters	6
Columns in X	3
Columns in Z	0
Subjects	18
Max Obs Per Subject	3

Useful output to make sure you are modeling what you think you are:
subjects should be your sample size
Max Obs Per Subject should match the # of DVs
Number of Observations Read/Used should be total N

Number of Observations	
Number of Observations Read	54
Number of Observations Used	54
Number of Observations Not Used	0

Estimated R Matrix for Subject 1

Row	Col1	Col2	Col3
1	264.36	315.00	373.72
2	315.00	446.77	539.49
3	373.72	539.49	727.25

The estimated R matrix and the COVTEST output that gives the estimated covariance matrix elements and their standard errors.

Note: disregard Z values (Wald test) as some covariance parameters are tested on their boundary.

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	personID	264.36	88.1204	3.00	0.0013
UN(2,1)	personID	315.00	109.88	2.87	0.0041
UN(2,2)	personID	446.77	148.92	3.00	0.0013
UN(3,1)	personID	373.72	135.79	2.75	0.0059
UN(3,2)	personID	539.49	184.99	2.92	0.0035
UN(3,3)	personID	727.25	242.42	3.00	0.0013

Information Criteria						
Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
408.1	9	426.1	430.2	427.2	434.1	443.1

For comparison with subsequent models.

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	134.11	4.9820	18	26.92	<.0001
dINTENSITY_W	-46.6111	2.1230	18	-21.96	<.0001
dINTENSITY_R	55.4444	2.2976	18	24.13	<.0001

The estimates for each parameter (the fixed effects)

Interpret each effect…

Intercept:

dINTENSITY_W:

dINTENSITY_R:

Adding a Predictor to a Multivariate Model

We will add diet (vegetarian or meat eater) as a predictor to the empty model:

```
TITLE "DIET TYPE MULTIVARIATE MODEL, FULL UNSTRUCTURED ERROR VARIANCE MODEL:
      (PREDICTORS ARE INDICATORS OF WHICH VARIABLE) - MULTIVARIATE ANOVA ASSUMPTION";
PROC MIXED DATA=WORK.dietstack METHOD=ML COVTEST NOPROFILE ITDETAILS IC NAMELEN=50;
MODEL pulse = dINTENSITY_W dINTENSITY_R dDIET_V dDIET_V*dINTENSITY_W dDIET_V*dINTENSITY_R / S
DDFM=KENWARDROGER;
REPEATED / SUBJECT=personID TYPE=UN R RCORR;
ODS OUTPUT INFOCRIT=info_model2 R=r_model2;
RUN;
TITLE;
```

dDIET_V = 1 if person is a vegetarian; = 0 if person is a meat eater

Here, we are adding what looks like a main effect of diet and an interaction of diet with intensity – all must be present to add variable to the analysis

Multivariate model becomes : $\boldsymbol{y}_p = \mathbf{X}_p\boldsymbol{\beta} + \boldsymbol{e}_p$

For meat eaters:

$$\boldsymbol{y}_p = \begin{bmatrix} y_{p1} \\ y_{p2} \\ y_{p3} \end{bmatrix}; \mathbf{X}_p = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}; \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_{I1} \\ \beta_{I3} \\ \beta_V \\ \beta_{V*I1} \\ \beta_{V*I3} \end{bmatrix}; \boldsymbol{e}_p = \begin{bmatrix} e_{p1} \\ e_{p2} \\ e_{p3} \end{bmatrix}$$

For vegetarians:

$$\boldsymbol{y}_p = \begin{bmatrix} y_{p1} \\ y_{p2} \\ y_{p3} \end{bmatrix}; \mathbf{X}_p = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}; \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_{I1} \\ \beta_{I3} \\ \beta_V \\ \beta_{V*I1} \\ \beta_{V*I3} \end{bmatrix}; \boldsymbol{e}_p = \begin{bmatrix} e_{p1} \\ e_{p2} \\ e_{p3} \end{bmatrix}$$

Where $\boldsymbol{e}_p \sim N_3(\mathbf{0}, \mathbf{R})$, and $\mathbf{R} = \begin{bmatrix} \sigma^2_{e1} & \sigma_{e_1,e_2} & \sigma_{e_1,e_3} \\ \sigma_{e_1,e_2} & \sigma^2_{e2} & \sigma_{e_2,e_3} \\ \sigma_{e_1,e_3} & \sigma_{e_2,e_3} & \sigma^2_{e3} \end{bmatrix}$ (the unstructured error covariance matrix model).

Which leads to the following simultaneous linear models:

For meat eaters:

$$y_{p1} = \beta_0 + \beta_{I1} + e_{p1}$$
$$y_{p2} = \beta_0 + e_{p2}$$
$$y_{p3} = \beta_0 + \beta_{I3} + e_{p3}$$

For vegetarians:

$$y_{p1} = \beta_0 + \beta_{I1} + \beta_V + \beta_{V*I1} + e_{p1}$$
$$y_{p2} = \beta_0 + \beta_V + e_{p2}$$
$$y_{p3} = \beta_0 + \beta_{I3} + \beta_V + \beta_{V*I3} + e_{p3}$$

Estimated R Matrix for Subject 1

Row	Col1	Col2	Col3
1	167.67	192.63	216.39
2	192.63	291.90	340.38
3	216.39	340.38	471.25

The **R** error covariance matrix has changed by adding the predictor – this is analogous to what happens in a univariate general linear model – the predictor shrinks the error variance:

For intensity #1 (pulse after warm up):

Empty model $\sigma^2_{e1} = 264.36$; After adding DIET as a predictor: $\sigma^2_{e1} = 167.67$
Therefore $R^2_{I1} = \frac{264.36-167.67}{264.36} = 0.366$

For intensity #2 (pulse after jogging):

Empty model $\sigma^2_{e2} = 446.77$; After adding DIET as a predictor: $\sigma^2_{e2} = 291.90$
Therefore $R^2_{I2} = \frac{446.77-291.90}{446.77} = 0.347$

For intensity #3 (pulse after running):

Empty model $\sigma^2_{e3} = 727.25$; After adding DIET as a predictor: $\sigma^2_{e3} = 471.25$
Therefore $R^2_{I3} = \frac{727.25-471.25}{727.25} = 0.352$

These R^2 reflect the proportion of variance accounted for **marginally** for each variable. What about the covariances? Notice, adding a predictor reduces those, too. What is needed is a **joint** description of how the **combined** variation of all three variables is reduced by the predictor. To get that, we can use the **generalized sample variance** (the determinant of **R)**

```
*MACRO FOR ANALYSIS OF THE R MATRIX – syntax run after PROC MIXED;
%MACRO R_Matrix_Determinant(rmatrix_library,rmatrix_dataset,output_library,output_dataset,modelnum);
PROC IML;

	USE &rmatrix_library..&rmatrix_dataset.;
	READ ALL VAR _ALL_ INTO FullX;

	K = NCOL(FullX);

	COV = FullX[,3:K];
	DET_R = DET(COV);

	OUTPUTMAT = &modelnum. || DET_R;

	CNAME = {"modelnumber","Gen Variance R"};
	CREATE &output_library..&output_dataset. FROM OUTPUTMAT [COLNAME = cname];
	APPEND FROM OUTPUTMAT;

QUIT;

%MEND;

*FOR MODEL COMPARISON;
DATA work.info_model1;
	SET work.info_model1;
	modelname = 'Empty Means; Unstructured Covariances (All Estimated)';
	modelnumber = 1;
RUN;

*CALCULATION OF GENERALIZED VARIANCE OF RMATRIX;
%R_Matrix_Determinant(work,R_model1,work,GV_R_model1,1);

*MERGE WITH INFORMATION CRITERIA FROM MODEL OUTPUT;
DATA work.info_model1;
	MERGE work.info_model1 work.GV_R_model1;
RUN;

*FOR MODEL COMPARISON;
DATA work.info_model2;
	SET work.info_model2;
	modelname = 'Diet; Unstructured Covariances (All Estimated)';
	modelnumber = 2;
RUN;

*CALCULATION OF GENERALIZED VARIANCE OF RMATRIX;
%R_Matrix_Determinant(work,R_model2,work,GV_R_model2,2);

*MERGE WITH INFORMATION CRITERIA FROM MODEL OUTPUT;
DATA work.info_model2;
	MERGE work.info_model2 work.GV_R_model2;
RUN;
```

<u>For all three variables jointly:</u>

Empty model generalized variance: $|\mathbf{R}|$ =1,411,077.1972; After adding DIET as a predictor: $|\mathbf{R}|$ = 859,978.5574

Therefore the overall R^2 is: $R^2 = \frac{1,411,077.1972 - 859,978.5574}{1,411,077.1972} = .391$

This value is bigger than each of the marginal R^2 values because of the contributions of the covariances.

Information Criteria						
Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
399.2	12	423.2	430.8	424.7	433.9	445.9

Because we are using Maximum Likelihood as our estimator, we can use the information criteria to determine if DIET significantly improved model fit. This is essentially testing whether the multivariate R^2 change is zero. More formally, we are testing the hypothesis:

$$H_0: \beta_V = \beta_{V*I1} = \beta_{V*I3} = 0$$
$$H_A: \text{At least one not equal to zero}$$

Note: if you are using REML (the default estimator in SAS PROC MIXED), you cannot do this likelihood ratio test as the model for the means (the fixed effects) are different between the empty model and the diet-predictor model.

Empty model $-2LL = 408.9$; 9 parameters. DIET model: $-2LL = 399.2$; 12 parameters

$$\chi^2 = 408.9 - 399.2 = 9.7, df = 12 - 9 = 3, p = 0.021$$

We can conclude that DIET significantly improved the fit of the **multivariate** model – but, what that does not tell us is if it improved the fit for each dependent variable **marginally**.

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	146.56	5.6950	18	25.73	<.0001
dINTENSITY_W	-49.2222	2.8734	18	-17.13	<.0001
dINTENSITY_R	59.0000	3.0255	18	19.50	<.0001
dDIET_V	-24.8889	8.0540	18	-3.09	0.0063
dINTENSITY_W*dDIET_V	5.2222	4.0636	18	1.29	0.2150
dINTENSITY_R*dDIET_V	-7.1111	4.2787	18	-1.66	0.1138

Interpret each effect…

Intercept:

dINTENSITY_W:

dINTENSITY_R:

dDIET_V:

dINTENSITY_W*dDIET_V:

dINTENSITY_R*dDIET_V:

What we cannot find from our **direct model output** (when coded this way – and without a CONTRAST statement) are some key hypothesis tests that show up for this type of analysis:

- Within subjects main effect of intensity (do mean pulse rates differ when pulse is taken after different intensities?)
    - Should be 2 degrees of freedom (3 intensity levels - 1)
- Between subjects main effect of diet type (do mean pulse rates differ for vegetarians and meat eaters?)
    - Should be 1 degree of freedom (2 diet types - 1)
- Interaction of within subjects intensity and between subjects diet type (are there differences in pulse rate for varying combinations of intensity and diet?)
    - Should be 2 degrees of freedom (3 intensity levels - 1)* (2 diet types - 1)

NOTE: using the CLASS statement for diet and intensity will make these steps unnecessary (see last example in handout)

We can also use our friend, the ESTIMATE statement to provide us with cell means that we may find useful…

```
*INVESTIGATING INTERACTION EFFECT OF INTENSITY*DIET;
ESTIMATE 'Predicted Pulse After Warmup for Vegetarians'  intercept 1 dINTENSITY_W 1 dDIET_V 1
                                                          dDIET_V*dINTENSITY_W 1;
ESTIMATE 'Predicted Pulse After Warmup for Meat Eaters'  intercept 1 dINTENSITY_W 1;
ESTIMATE 'Predicted Pulse After Jogging for Vegetarians' intercept 1 dDIET_V 1;
ESTIMATE 'Predicted Pulse After Jogging for Meat Eaters' intercept 1;
ESTIMATE 'Predicted Pulse After Running for Vegetarians' intercept 1 dINTENSITY_R 1 dDIET_V 1
                                                          dDIET_V*dINTENSITY_R 1;
ESTIMATE 'Predicted Pulse After Running for Meat Eaters' intercept 1 dINTENSITY_R 1;
```

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
Predicted Pulse After Warmup for Vegetarians	77.6667	4.3162	18	17.99	<.0001
Predicted Pulse After Warmup for Meat Eaters	97.3333	4.3162	18	22.55	<.0001
Predicted Pulse After Jogging for Vegetarians	121.67	5.6950	18	21.36	<.0001
Predicted Pulse After Jogging for Meat Eaters	146.56	5.6950	18	25.73	<.0001
Predicted Pulse After Running for Vegetarians	173.56	7.2361	18	23.98	<.0001
Predicted Pulse After Running for Meat Eaters	205.56	7.2361	18	28.41	<.0001

The two-way interaction between intensity and diet type can be found with a CONTRAST statement. Contrast statements are for multiple degree of freedom test. Here, the interaction is testing $H_0: \beta_{V*I1} = \beta_{V*I3} = 0$:

```
* THE CONTRAST STATEMENT PROVIDES MULTIPLE DEGREE OF FREEDOM TESTS;
CONTRAST 'Within Subjects Test of Intensity*Diet' dDIET_V*dINTENSITY_W 1, dDIET_V*dINTENSITY_R 1;
```

Contrasts

Label	Num DF	Den DF	F Value	Pr > F
Within Subjects Test of Intensity*Diet	2	17	1.61	0.2294

We find the between diet and intensity interaction is not significant – but, we can now look at the cell mean differences between each combination of diet and intensity to see if any are significantly different. Here, each estimate statement comes from subtracting one predicted value (from above) from another:

```
*  POST HOC DIFFERENCES IN INTERACTION CELL MEANS  -----------------------------------------------------;
ESTIMATE 'Difference In Pulse After Warmup for Diet' dDIET_V 1 dDIET_V*dINTENSITY_W 1;
ESTIMATE 'Difference In Pulse After Jogging for Diet' dDIET_V 1;
ESTIMATE 'Difference in Pulse After Running for Diet' dDIET_V 1 dDIET_V*dINTENSITY_R 1;

ESTIMATE 'Difference in Pulse After Warmup and Pulse After Jogging for Vegetarians'
                                                            dINTENSITY_W 1 dDIET_V*dINTENSITY_W 1;
ESTIMATE 'Difference in Pulse After Warmup and Pulse After Running for Vegetarians'
                    dINTENSITY_W 1 dINTENSITY_R -1 dDIET_V*dINTENSITY_W 1 dDIET_V*dINTENSITY_R -1;
ESTIMATE 'Difference in Pulse After Jogging and Pulse After Running for Vegetarians'
                                                            dINTENSITY_R -1 dDIET_V*dINTENSITY_R -1;

ESTIMATE 'Difference in Pulse After Warmup and Pulse After Jogging for Meat Eaters' dINTENSITY_W 1;
ESTIMATE 'Difference in Pulse After Warmup and Pulse After Running for Meat Eaters'
                                                                     dINTENSITY_W 1 dINTENSITY_R -1;
ESTIMATE 'Difference in Pulse After Jogging and Pulse After Running for Meat Eaters' dINTENSITY_R -1;

ESTIMATE 'Difference in Pulse After Warmup for Vegetarians and After Jogging for Meat Eaters'
                                                   dDIET_V 1 dINTENSITY_W 1 dDIET_V*dINTENSITY_W 1;
ESTIMATE 'Difference in Pulse After Warmup for Vegetarians and After Running for Meat Eaters'
                                   dINTENSITY_W 1 dDIET_V 1 dDIET_V*dINTENSITY_W 1 dINTENSITY_R -1;

ESTIMATE 'Difference in Pulse After Warmup for Meat Eaters and After Jogging for Vegetarians'
                                                                         dINTENSITY_W 1 dDIET_V -1;
ESTIMATE 'Difference in Pulse After Warmup for Meat Eaters and After Running for Vegetarians'
                                 dINTENSITY_W 1 dINTENSITY_R -1 dDIET_V -1 dDIET_V*dINTENSITY_R -1;

ESTIMATE 'Difference in Pulse After Jogging for Vegetarians and Pulse After Running for Meat Eaters'
                                                                         dDIET_V 1 dINTENSITY_R -1;
ESTIMATE 'Difference in Pulse After Jogging for Meat Eaters and Pulse After Running for Vegetarians'
                                              dINTENSITY_R -1 dDIET_V -1 dDIET_V*dINTENSITY_R -1;
```

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
Difference in Pulse After Running for Diet	-32.0000	10.2334	18	-3.13	0.0058
Difference in Pulse After Warmup and Pulse After Jogging for Vegetarians	-44.0000	2.8734	18	-15.31	<.0001
Difference in Pulse After Warmup and Pulse After Running for Vegetarians	-95.8889	4.7858	18	-20.04	<.0001
Difference in Pulse After Jogging and Pulse After Running for Vegetarians	-51.8889	3.0255	18	-17.15	<.0001
Difference in Pulse After Warmup and Pulse After Jogging for Meat Eaters	-49.2222	2.8734	18	-17.13	<.0001
Difference in Pulse After Warmup and Pulse After Running for Meat Eaters	-108.22	4.7858	18	-22.61	<.0001
Difference in Pulse After Jogging and Pulse After Running for Meat Eaters	-59.0000	3.0255	18	-19.50	<.0001
Difference in Pulse After Warmup for Vegetarians and After Jogging for Meat Eaters	-68.8889	7.1458	20.3	-9.64	<.0001
Difference in Pulse After Warmup for Vegetarians and After Running for Meat Eaters	-127.89	8.4256	21.4	-15.18	<.0001
Difference in Pulse After Warmup for Meat Eaters and After Jogging for Vegetarians	-24.3333	7.1458	20.3	-3.41	0.0028
Difference in Pulse After Warmup for Meat Eaters and After Running for Vegetarians	-76.2222	8.4256	21.4	-9.05	<.0001
Difference in Pulse After Jogging for Vegetarians and Pulse After Running for Meat Eaters	-83.8889	9.2084	19.4	-9.11	<.0001
Difference in Pulse After Jogging for Meat Eaters and Pulse After Running for Vegetarians	-27.0000	9.2084	19.4	-2.93	0.0084

Because of the DIET*INTENSITY interaction, the main effects for DIET and INTENSITY now are conditional – and the actual hypothesis tests are for the **marginal values** (the mean difference in levels aggregated across all levels of the other variable). We can come up with the **marginal means** for each effect using the ESTIMATE statement. Here, the DIVISOR = 2 divides each of the coefficients in the estimate statement by 2 (there are two levels of DIET)

```
*INVESTIGATING MARGINAL EFFECT OF INTENSITY;
ESTIMATE 'Marginal Mean of Pulse After Warmup'
                           intercept 2 dINTENSITY_W 2 dDIET_V 1 dDIET_V*dINTENSITY_W 1 / DIVISOR = 2;

ESTIMATE 'Marginal Mean of Pulse After Jogging' intercept 2 dDIET_V 1 / DIVISOR = 2;

ESTIMATE 'Marginal Mean of Pulse After Running'
                           intercept 2 dINTENSITY_R 2 dDIET_V 1 dDIET_V*dINTENSITY_R 1 / DIVISOR = 2;
```

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
Marginal Mean of Pulse After Warmup	87.5000	3.0520	18	28.67	<.0001
Marginal Mean of Pulse After Jogging	134.11	4.0270	18	33.30	<.0001
Marginal Mean of Pulse After Running	189.56	5.1167	18	37.05	<.0001
Difference in Marginal Mean of Pulse After Warmup and Jogging	-46.6111	2.0318	18	-22.94	<.0001
Difference in Marginal Mean of Pulse After Warmup and Running	-110.83	6.1673	18.4	-17.97	<.0001
Difference in Marginal Mean of Pulse After Jogging and Running	-55.4444	2.1393	18	-25.92	<.0001

The associated contrast functions by forming a pair of differences: (1) between the marginal mean for pulse after warm up and marginal mean for pulse after Jogging [showing up before the comma in the CONTRAST statement] and (2) between the marginal mean pulse after jogging and the marginal mean pulse after running [showing up after the comma in the CONTRAST statement]

```
CONTRAST 'Within Subjects Test of Intensity' dINTENSITY_W 1 dDIET_V*dINTENSITY_W .5,
                                             dINTENSITY_R -1 dDIET_V*dINTENSITY_R -.5;
```

Contrasts				
Label	Num DF	Den DF	F Value	Pr > F
Within Subjects Test of Intensity	2	17	431.33	<.0001

We can also ask for post-hoc mean differences between each of the marginal means of intensity:

```
* POST HOC DIFFERENCES IN MARGINAL CELL MEANS OF INTENSITY;
ESTIMATE 'Difference in Marginal Mean of Pulse After Warmup and Jogging'
                                        dINTENSITY_W 2 dDIET_V*dINTENSITY_W 1 / DIVISOR = 2;

ESTIMATE 'Difference in Marginal Mean of Pulse After Warmup and Running'
 dINTENSITY_W 2 dINTENSITY_R -2 dDIET_V*dINTENSITY_W 1 dDIET_V*dINTENSITY_W -1/ DIVISOR = 2;

ESTIMATE 'Difference in Marginal Mean of Pulse After Jogging and Running'
                                     dINTENSITY_R -2 dDIET_V*dINTENSITY_R -1 / DIVISOR = 2;
```

Estimates					
Label	Estimate	Standard Error	DF	t Value	Pr > |t|
Difference in Marginal Mean of Pulse After Warmup and Jogging	-46.6111	2.0318	18	-22.94	<.0001
Difference in Marginal Mean of Pulse After Warmup and Running	-110.83	6.1673	18.4	-17.97	<.0001
Difference in Marginal Mean of Pulse After Jogging and Running	-55.4444	2.1393	18	-25.92	<.0001

Finally, we can do the same for the other marginal variable: Diet:

```
*INVESTIGATING MARGINAL EFFECT OF DIET;
ESTIMATE 'Marginal Mean of Vegetarians'
intercept 3 dINTENSITY_W 1 dDIET_V 3 dDIET_V*dINTENSITY_W 1 dINTENSITY_R 1 dDIET_V*dINTENSITY_R 1
                                        /DIVISOR=3;

ESTIMATE 'Marginal Mean of Meat Eaters' intercept 3 dINTENSITY_W 1 dINTENSITY_R 1 /DIVISOR=3;

CONTRAST 'Between Subjects Test of Diet' dDIET_V .33 dDIET_V*dINTENSITY_W .33 dDIET_V*dINTENSITY_R .33;

* POST HOC DIFFERENCES IN MARGINAL CELL MEANS OF DIET;
ESTIMATE 'Difference Marginal Mean of Vegetarians and Meat Eaters'
                          dDIET_V 3 dDIET_V*dINTENSITY_W 1 dDIET_V*dINTENSITY_R 1/DIVISOR=3;
```

Estimates					
Label	Estimate	Standard Error	DF	t Value	Pr > |t|
Marginal Mean of Vegetarians	124.30	5.4768	18	22.70	<.0001
Marginal Mean of Meat Eaters	149.81	5.4768	18	27.35	<.0001
Difference Marginal Mean of Vegetarians and Meat Eaters	-25.5185	7.7454	18	-3.29	0.0040

Contrasts				
Label	Num DF	Den DF	F Value	Pr > F
Between Subjects Test of Diet	1	18	10.84	0.0040

Using the CLASS statement with LSMEANS, ESTIMATE, and CONTRAST

We take a detour from the multivariate analysis to show how the CLASS statement could simplify the post-hoc analyses. The CLASS statement is useful for categorical IVs in that it makes SAS provide the coding system for the variables on the line:

```
TITLE "MULTIVARIATE MODEL WITH DIET PREDICTOR ONLY - CLASS STATEMENT";
PROC MIXED DATA=WORK.dietstack METHOD=ML COVTEST NOPROFILE ITDETAILS IC NAMELEN=50;
CLASS intensity diet;
MODEL pulse = intensity diet intensity*diet/ S DDFM=KENWARDROGER;
REPEATED / SUBJECT=personID TYPE=UN R RCORR;
```

Here, we just need to put the name of the variables on the CLASS statement – SAS provides all effects implied by the model automatically…

Class Level Information

Class	Levels	Values
intensity	3	Pulse: Jogging Pulse: Running Pulse: Warm Up
diet	2	Meat Eaters Vegetarians

NOTE: this is where you can find the reference groups for the analysis: Pulse after Warm Up / Vegetarians. All fixed effects will reflect this coding.

Solution for Fixed Effects

Effect	intensity	Diet Type	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept			77.6667	4.3162	18	17.99	<.0001
intensity	Pulse: Jogging		44.0000	2.8734	18	15.31	<.0001
intensity	Pulse: Running		95.8889	4.7858	18	20.04	<.0001
intensity	Pulse: Warm Up		0	.	.	.	.
diet		Meat Eaters	19.6667	6.1040	18	3.22	0.0047
diet		Vegetarians	0	.	.	.	.
intensity*diet	Pulse: Jogging	Meat Eaters	5.2222	4.0636	18	1.29	0.2150
intensity*diet	Pulse: Jogging	Vegetarians	0	.	.	.	.
intensity*diet	Pulse: Running	Meat Eaters	12.3333	6.7682	18	1.82	0.0851
intensity*diet	Pulse: Running	Vegetarians	0	.	.	.	.
intensity*diet	Pulse: Warm Up	Meat Eaters	0	.	.	.	.
intensity*diet	Pulse: Warm Up	Vegetarians	0	.	.	.	.

Interpret all of these parameters:

Effect	intensity	Diet Type
Intercept		
intensity	Pulse: Jogging	
intensity	Pulse: Running	
diet		Meat Eaters
intensity*diet	Pulse: Jogging	Meat Eaters
intensity*diet	Pulse: Running	Meat Eaters

Using the CLASS statement, we also receive an overall test for the three within and between subjects hypothesis tests:

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
intensity	2	17	431.33	<.0001
diet	1	18	10.85	0.0040
intensity*diet	2	17	1.61	0.2294

We found these using the CONTRAST statement in the previous analysis

With the CLASS statement, the categorical IVs can now be used in combination with the LSMEANS statement, which will give you the conditional and marginal means for all statements (all the estimate statements from the previous analysis!). The SLICE statement provides your tests of group differences at all levels of the "SLICING" variable – for investigating interactions.

```
LSMEANS intensity diet intensity*diet / PDIFF=ALL SLICE=diet SLICE=intensity;
```

Least Squares Means							
Effect	intensity	Diet Type	Estimate	Standard Error	DF	t Value	Pr > |t|
intensity*diet	Pulse: Jogging	Meat Eaters	146.56	5.6950	18	25.73	<.0001
intensity*diet	Pulse: Jogging	Vegetarians	121.67	5.6950	18	21.36	<.0001
intensity*diet	Pulse: Running	Meat Eaters	205.56	7.2361	18	28.41	<.0001
intensity*diet	Pulse: Running	Vegetarians	173.56	7.2361	18	23.98	<.0001
intensity*diet	Pulse: Warm Up	Meat Eaters	97.3333	4.3162	18	22.55	<.0001
intensity*diet	Pulse: Warm Up	Vegetarians	77.6667	4.3162	18	17.99	<.0001
intensity*diet	Pulse: Jogging	Meat Eaters	146.56	5.6950	18	25.73	<.0001
intensity*diet	Pulse: Jogging	Vegetarians	121.67	5.6950	18	21.36	<.0001
intensity*diet	Pulse: Running	Meat Eaters	205.56	7.2361	18	28.41	<.0001
intensity*diet	Pulse: Running	Vegetarians	173.56	7.2361	18	23.98	<.0001
intensity*diet	Pulse: Warm Up	Meat Eaters	97.3333	4.3162	18	22.55	<.0001
intensity*diet	Pulse: Warm Up	Vegetarians	77.6667	4.3162	18	17.99	<.0001
intensity	Pulse: Jogging		134.11	4.0270	18	33.30	<.0001
intensity	Pulse: Running		189.56	5.1167	18	37.05	<.0001
intensity	Pulse: Warm Up		87.5000	3.0520	18	28.67	<.0001
diet		Meat Eaters	149.81	5.4768	18	27.35	<.0001
diet		Vegetarians	124.30	5.4768	18	22.70	<.0001

These are the marginal means we found in the previous analysis

Differences of Least Squares Means									
Effect	intensity	Diet Type	_intensity	Diet Type	Estimate	Standard Error	DF	t Value	Pr > |t|
intensity*diet	Pulse: Jogging	Meat Eaters	Pulse: Jogging	Vegetarians	24.8889	8.0540	18	3.09	0.0063
intensity*diet	Pulse: Jogging	Meat Eaters	Pulse: Running	Meat Eaters	-59.0000	3.0255	18	-19.50	<.0001
intensity*diet	Pulse: Jogging	Meat Eaters	Pulse: Running	Vegetarians	-27.0000	9.2084	19.4	-2.93	0.0084

Differences of Least Squares Means									
Effect	**intensity**	**Diet Type**	**_intensity**	**Diet Type**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
intensity*diet	Pulse: Jogging	Meat Eaters	Pulse: Warm Up	Meat Eaters	49.2222	2.8734	18	17.13	<.0001
intensity*diet	Pulse: Jogging	Meat Eaters	Pulse: Warm Up	Vegetarians	68.8889	7.1458	20.3	9.64	<.0001
intensity*diet	Pulse: Jogging	Vegetarians	Pulse: Running	Meat Eaters	-83.8889	9.2084	19.4	-9.11	<.0001
intensity*diet	Pulse: Jogging	Vegetarians	Pulse: Running	Vegetarians	-51.8889	3.0255	18	-17.15	<.0001
intensity*diet	Pulse: Jogging	Vegetarians	Pulse: Warm Up	Meat Eaters	24.3333	7.1458	20.3	3.41	0.0028
intensity*diet	Pulse: Jogging	Vegetarians	Pulse: Warm Up	Vegetarians	44.0000	2.8734	18	15.31	<.0001
intensity*diet	Pulse: Running	Meat Eaters	Pulse: Running	Vegetarians	32.0000	10.2334	18	3.13	0.0058
intensity*diet	Pulse: Running	Meat Eaters	Pulse: Warm Up	Meat Eaters	108.22	4.7858	18	22.61	<.0001
intensity*diet	Pulse: Running	Meat Eaters	Pulse: Warm Up	Vegetarians	127.89	8.4256	21.4	15.18	<.0001
intensity*diet	Pulse: Running	Vegetarians	Pulse: Warm Up	Meat Eaters	76.2222	8.4256	21.4	9.05	<.0001
intensity*diet	Pulse: Running	Vegetarians	Pulse: Warm Up	Vegetarians	95.8889	4.7858	18	20.04	<.0001
intensity*diet	Pulse: Warm Up	Meat Eaters	Pulse: Warm Up	Vegetarians	19.6667	6.1040	18	3.22	0.0047
intensity	Pulse: Jogging		Pulse: Running		-55.4444	2.1393	18	-25.92	<.0001
intensity	Pulse: Jogging		Pulse: Warm Up		46.6111	2.0318	18	22.94	<.0001
intensity	Pulse: Running		Pulse: Warm Up		102.06	3.3841	18	30.16	<.0001
diet		Meat Eaters		Vegetarians	25.5185	7.7454	18	3.29	0.0040

Tests of Effect Slices						
Effect	**intensity**	**Diet Type**	**Num DF**	**Den DF**	**F Value**	**Pr > F**
intensity*diet		Meat Eaters	2	17	242.69	<.0001
intensity*diet		Vegetarians	2	17	190.25	<.0001
intensity*diet	Pulse: Jogging		1	18	9.55	0.0063
intensity*diet	Pulse: Running		1	18	9.78	0.0058
intensity*diet	Pulse: Warm Up		1	18	10.38	0.0047

Finally, there are times you will need to use the ESTIMATE statement when you have the CLASS statement as part of your syntax. Most commonly, this is due to the LSMEANS statement not accepting continuous IVs – so if you need, for instance, a set of within-group slopes (or their differences), you'll need the ESTIMATE statement again.

The CLASS statement changes the behavior of the estimate statement – now each level of each effect on CLASS needs a number. Here, the / E prints out the list of effects where each number goes:

```
*INVESTIGATING INTERACTION EFFECT OF INTENSITY*DIET;
ESTIMATE 'Predicted Pulse After Warmup for Vegetarians'
                           intercept 1 diet 0 1 intensity 0 0 1 diet*intensity 0 0 0 0 0 1 / E;
```

Coefficients for Predicted Pulse After Warmup for Vegetarians			
Effect	**intensity**	**Diet Type**	**Row1**
Intercept			1
intensity	Pulse: Jogging		
intensity	Pulse: Running		
intensity	Pulse: Warm Up		1
diet		Meat Eaters	
diet		Vegetarians	1
intensity*diet	Pulse: Jogging	Meat Eaters	
intensity*diet	Pulse: Jogging	Vegetarians	
intensity*diet	Pulse: Running	Meat Eaters	
intensity*diet	Pulse: Running	Vegetarians	
intensity*diet	Pulse: Warm Up	Meat Eaters	
intensity*diet	Pulse: Warm Up	Vegetarians	1

Estimates					
Label	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Predicted Pulse After Warmup for Vegetarians	77.6667	4.3162	18	17.99	<.0001

Finally, the following syntax will replicate the cell means (given by LSMEANS in this analysis):

```
*INVESTIGATING INTERACTION EFFECT OF INTENSITY*DIET;
ESTIMATE 'Predicted Pulse After Warmup for Vegetarians'
                    intercept 1 diet 0 1 intensity 0 0 1 diet*intensity 0 0 0 0 0 1 / E;
ESTIMATE 'Predicted Pulse After Warmup for Meat Eaters'
                    intercept 1 diet 1 0 intensity 0 0 1 diet*intensity 0 0 0 0 1 0;
ESTIMATE 'Predicted Pulse After Jogging for Vegetarians'
                    intercept 1 diet 0 1 intensity 0 1 0 diet*intensity 0 0 0 1 0 0;
ESTIMATE 'Predicted Pulse After Jogging for Meat Eaters'
                    intercept 1 diet 1 0 intensity 0 1 0 diet*intensity 0 0 1 0 0 0;
ESTIMATE 'Predicted Pulse After Running for Vegetarians'
                    intercept 1 diet 0 1 intensity 1 0 0 diet*intensity 0 1 0 0 0 0;
ESTIMATE 'Predicted Pulse After Running for Meat Eaters'
                    intercept 1 diet 1 0 intensity 1 0 0 diet*intensity 1 0 0 0 0 0;
```

Estimates					
Label	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Predicted Pulse After Warmup for Vegetarians	77.6667	4.3162	18	17.99	<.0001
Predicted Pulse After Warmup for Meat Eaters	97.3333	4.3162	18	22.55	<.0001
Predicted Pulse After Jogging for Vegetarians	173.56	7.2361	18	23.98	<.0001
Predicted Pulse After Jogging for Meat Eaters	205.56	7.2361	18	28.41	<.0001
Predicted Pulse After Running for Vegetarians	121.67	5.6950	18	21.36	<.0001
Predicted Pulse After Running for Meat Eaters	146.56	5.6950	18	25.73	<.0001

Comparison of PROC MIXED Multivariate Analyses with Classical Multivariate Analysis of Variance (MANOVA)

A semi-frequent request when running multivariate analyses is for some type of "MANOVA" tests. Usually, this is due to the person requesting the classical tests having been trained in a customary multivariate analysis course. This section of the handout describes MANOVA and how what we have just done can provide a MANOVA-style test if you ever are asked for one.

We will begin our discussion of MANOVA by starting with the univariate ANOVA (a one-way ANOVA with one categorical IV). In univariate ANOVA, we develop the hypothesis test of group differences in the group mean of the outcome using the concept of sums of squares.

$$H_0: \mu_1 = \mu_2 = \cdots = \mu_G$$
$$H_A: \text{at least one } \mu \text{ not equal to the others}$$

Sums of squares between groups comes from the sum of squared differences from each group mean $\bar{y}_g$ from the overall grand mean $\bar{y}$ (where N_g is the within-group sample size for group g):

$$SS_B = \sum_{g=1}^{G} N_g \left(\bar{y}_g - \bar{y}\right)^2$$

The sums of squares within groups (here referred to as the sums of squares for error) comes from the squared difference of each person's outcome and their group's mean:

$$SS_E = \sum_{g=1}^{G} \sum_{n_g=1}^{N_g} \left(y_p - \bar{y}_g\right)^2$$

The F-test is the ratio of these two terms, each divided by their respective degrees of freedom:

$$F = \frac{SS_B/df_B}{SS_E/df_E}$$

A MANOVA extends a univariate ANOVA to test hypotheses about **mean vectors** across categorical independent variables. The hypothesis test is now:

$$H_0: \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \cdots = \boldsymbol{\mu}_G$$
$$H_A: \text{at least one } \boldsymbol{\mu} \text{ not equal to the others}$$

Where $\boldsymbol{\mu}_g = \begin{bmatrix} \mu_{g1} \\ \mu_{g2} \\ \vdots \\ \mu_{gV} \end{bmatrix}$, for Vobserved outcome variables (i.e., a multivariate analysis). The same concepts of univariate ANOVA still apply, just with vectors and matrices instead of individual means. The matrix analog to the sums of squares between groups is the Between Groups Sums of Squares and Cross Products matrix (using SAS' notation **H**, which stands for Hypothesis) is formed using each group's mean vector $\bar{\boldsymbol{y}}_g$ (size $V\ x\ 1$) and the overall grand mean vector $\bar{\boldsymbol{y}}$ (also size $V\ x\ 1$)

$$\mathbf{H}_{(V\ x\ V)} = \sum_{g=1}^{G} N_g \left(\bar{\boldsymbol{y}}_g - \bar{\boldsymbol{y}}\right)\left(\bar{\boldsymbol{y}}_g - \bar{\boldsymbol{y}}\right)^T$$

The matrix analog to the sums of squares within groups (here referred to as the sums of squares for error) is the Error Sums of Squares and Cross Products Matrix which comes from comparing each person's vector of outcomes $\boldsymbol{y}_p$ with their group mean vector $\bar{\boldsymbol{y}}_g$:

$$\mathbf{E}_{(V\,x\,V)} = \sum_{g=1}^{G}\sum_{n_g=1}^{N_g}(\boldsymbol{y}_p - \bar{\boldsymbol{y}})(\boldsymbol{y}_p - \bar{\boldsymbol{y}})^T$$

Note: for our multivariate models using maximum likelihood and an unstructured **R** matrix, $\mathbf{R} = \frac{\boldsymbol{E}}{\boldsymbol{N}}$

Now, instead of forming an F-ratio, we can form a matrix that incorporates this ratio:

$$\mathbf{F} = \mathbf{E}^{-1}\mathbf{H}$$

Where MANOVA differs from univariate ANOVA is that there is no one best test statistic for summarizing **F**. Instead, four popular statistics are formed:

- Wilk's Lambda = $\frac{\det(\mathbf{E})}{\det(\mathbf{H}+\mathbf{E})}$ (we will show how to get this using a modification of a likelihood ratio test)
- Pillai's trace = $trace(\mathbf{H}(\mathbf{H}+\mathbf{E})^{-1})$
- Hotelling-Lawley trace = $trace(\mathbf{E}^{-1}\mathbf{H})$
- Roy's greatest (largest) root = largest eigenvalue of $\mathbf{E}^{-1}\mathbf{H}$

What we seek to say is this: even though these statistics don't directly come out of PROC MIXED, you can still obtain (at least one of) them. Therefore, the analysis we have done subsumes classical MANOVA (or Multivariate Regression) into a more general framework for investigating multivariate hypotheses. We will now compare a MANOVA with PROC GLM and show how we can achieve the same result from PROC MIXED.

A Classical MANOVA Using PROC GLM

The following syntax is to conduct a MANOVA for the test of differences in the mean vectors of pulse rate across the two groups of the DIET variable:

```
TITLE "CLASSICAL MANOVA: MULTIVARIATE MODEL WITH DIET PREDICTOR ONLY";
*COMPARISON WITH MANOVA - RUNS CLASSICAL MANOVA ANALYSIS;
PROC GLM DATA=work.dietwide;
      MODEL pulse1 pulse2 pulse3 = DIET / SOLUTION NOUNI;
      MANOVA H=DIET / PRINTE;
RUN;
```

E = Error SSCP Matrix			
	pulse1	**pulse2**	**pulse3**
pulse1	3018	3467.3333333	3895
pulse2	3467.3333333	5254.2222222	6126.8888889
pulse3	3895	6126.8888889	8482.4444444

The **E** matrix is essentially what we found with **R – if we divide these numbers by N = 18**

167.67	192.63	216.39	$\frac{\mathbf{E}}{18}$
192.63	291.90	340.38	
216.39	340.38	471.27	

Estimated R Matrix for Subject 1			
Row	**Col1**	**Col2**	**Col3**
1	167.67	192.63	216.39
2	192.63	291.90	340.38
3	216.39	340.38	471.25

The four MANOVA test statistics are shown in the output below.

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall diet Effect H = Type III SSCP Matrix for diet E = Error SSCP Matrix S=1 M=0.5 N=6					
Statistic	**Value**	**F Value**	**Num DF**	**Den DF**	**Pr > F**
Wilks' Lambda	0.60944827	2.99	3	14	0.0669
Pillai's Trace	0.39055173	2.99	3	14	0.0669
Hotelling-Lawley Trace	0.64082835	2.99	3	14	0.0669
Roy's Greatest Root	0.64082835	2.99	3	14	0.0669

We can obtain Wilks Lambda from a model comparison of our previous two multivariate models in PROC MIXED:

Empty Model $-2LL_{EMPTY} = 408.1$
DIET Model $-2LL_{DIET} = 399.2$

Wilks Lambda =

$$\Lambda = \exp\left(\frac{-(-2LL_{EMPTY} - -2LL_{DIET})}{N}\right) = \exp\left(\frac{-(408.1 - 399.2)}{18}\right) = 0.609$$