**Breakout Session Data Sets**

There are four problems given below. We will split into two breakout rooms and work on these data sets. The goal is to have SAS output that your group will explain to the class.

**Problem 1: One Between One Within**

**/\* Data for One Between One Within Analysis \*/**

**data** onebonew;

input group day1 day2 day3 day4 day5;

datalines;

1 26 20 18 11 10

1 34 35 29 22 23

1 41 37 25 18 15

1 29 28 22 15 13

1 35 34 27 21 17

1 28 22 17 14 10

1 38 34 28 25 22

1 43 37 30 27 25

2 42 38 26 20 15

2 31 27 21 18 13

2 45 40 33 25 18

2 29 25 17 13 8

2 29 32 28 22 18

2 33 30 24 18 7

2 34 30 25 24 23

2 37 31 25 22 20

**run**;**proc print data=onebonew; run;**

**Problem 2: Repeated Measures with Contrasts**

Can you will yourself to fend off a cold? Can your mental state contribute to your getting sick? Recent studies in the field of psychoimmunology suggest a link exists between behavioral events and the functioning of ones’ immune system. Some have even suggested that the immune function might be conditioned to produce antibodies when challenged by antigens. In a study by Dr. William Keppel, 12 subjects, aged 20-50, were monitored during three distinct activities. The first activity, a baseline period, Phase A, consisted of neutral activity such as reporting tasks and answering questions. During the second activity, Phase B, subjects listened through headphones to a cassette tape of exercises including images of heaviness and warmth in the body, relaxation-deepening images, and suggestions to remember recovering from an illness. The third activity, Phase C. was a nonaudiotape follow-up stimulus consisting of continued relaxation as in Phase B and a verbal discussion of the positive aspects of the audiotape.

During each phase, the investigator measured interleukin-1 (IL-1) activity from blood samples. IL-1 plays a central role in immune system regulation, and it has a short half-life. The data are given in the file ex1614.csv.

The principal issues are (1) whether and how much difference there is between Phase B and the baseline Phase A levels of IL-1, and (2) whether and how much difference there is between the Phase C and Phase B responses. After answering the questions, comment on what inferences can be drawn from this study, in which the subjects were volunteers and the treatments were always given in the order A, B, C.

Here is the SAS code to enter the data. The repeated measures analysis is up to you.

title 'Person-Level Immune Data';

**data** immune;

infile '\\Client\H$\ex1614.csv' dlm=',' firstobs=**2**;

input Subject Phase1 Phase2 Phase3;

**run**;

**proc** **print** data=immune; **run**;

**Problem 3: Galton Height Data Accounting for Dependence in Family Memembership**

The data file ex0726 contains the heights of 453 adult females and 480 adult males with the heights of their mothers and fathers, collected by Francis Galton in 1885. Exercise 10.32 called for finding a repression model for predicting a person’s height from their gender and from their parents’ height, but suggested ignoring the possible dependence of heights of children from the same family. The variable Family in the data set is a code for family, from 1 to 205. Repeat the analysis – find and equation for predicting height from a person’s gender and their parents’ heights, but account for the possible dependence. SAS proc genmod is used to model correlated data.  We can use the class statement and the repeated statement to indicate that the observations are clustered into families (based on family) and that the observations may be correlated within families, but would be independent between families.

Here’s the code (assuming the SAS work file is called “galton”:

**proc** **genmod** data=galton;

class family gender;

model height = father mother gender ;

repeated subject=family / type=ind ;

**run**; **quit**;

Run the OLS code, too, and compare your results.

**Problem 4: Clever Hans Randomized Block Experiment**

This example uses statements for the analysis of a randomized block with two treatment factors occurring in a factorial structure. The data, from Neter, Wasserman, and Kutner ([1990](http://support.sas.com/documentation/cdl/en/statug/68162/HTML/default/statug_anova_references.htm#statug_anovanete_j90), p. 941), are from an experiment examining the effects of codeine and acupuncture on post-operative dental pain in male subjects. Both treatment factors have two levels. The codeine levels are a codeine capsule or a sugar capsule. The acupuncture levels are two inactive acupuncture points or two active acupuncture points. There are four distinct treatment combinations due to the factorial treatment structure. The 32 subjects are assigned to eight blocks of four subjects each based on an assessment of pain tolerance.

The data for the analysis are balanced, so PROC ANOVA is used. The data are as follows:

title1 'Randomized Complete Block With Two Factors';

**data** PainRelief;

input PainLevel Codeine Acupuncture Relief @@;

datalines;

1 1 1 0.0 1 2 1 0.5 1 1 2 0.6 1 2 2 1.2

2 1 1 0.3 2 2 1 0.6 2 1 2 0.7 2 2 2 1.3

3 1 1 0.4 3 2 1 0.8 3 1 2 0.8 3 2 2 1.6

4 1 1 0.4 4 2 1 0.7 4 1 2 0.9 4 2 2 1.5

5 1 1 0.6 5 2 1 1.0 5 1 2 1.5 5 2 2 1.9

6 1 1 0.9 6 2 1 1.4 6 1 2 1.6 6 2 2 2.3

7 1 1 1.0 7 2 1 1.8 7 1 2 1.7 7 2 2 2.1

8 1 1 1.2 8 2 1 1.7 8 1 2 1.6 8 2 2 2.4

;

Questions:

1. What is the blocking variable?
   1. Pain Level
2. What are the factors?
   1. Codeine Acupuncture
3. What is the response?
   1. Relief
4. Interpret the output.
   1. The interaction is not necessary as a two way ANOVA proves out the lack of significance (p = 0.0923)
5. Is the blocking variable necessary? Hint – consider the discussion on relative efficiency from the asynchronous material.
   1. Yes, it’s necessary because it’s significant in the analysis. Can’t judge on statistical significance. 13.66 times more efficient.

**Once we have explained results, we will come back together and examine panel plots for repeated measures data, as was done in the asynchronous material.**