

The objective of this assignment is continuing practice with the SAS software. For this assignment the SAS programming will require;

- 1) input of an external data set, (ex0524.csv). **Store this dataset in C:\TEMP\.**
- 2) Start the data input on the second line of the dataset. The first line reads “ZINC GROUP” and is followed by data on the second line. These are the input variables. The first variable (ZINC) is quantitative and the second is an alphanumeric character variable.
- 3) Run an Analysis of Variance using both PROC GLM and PROC MIXED.
- 4) Output of residuals with the option “/ outp=*somename*” on the MODEL statement of the PROC MIXED. Print this dataset and note that it contains the predicted values (PRED) and residuals (RESID).
- 5) Use PROC UNIVARIATE to process the variable RESID from the output dataset.
- 6) Create a scatter plot the residuals on both the variable BONEGRP and on the predicted value (PRED). Include the vref option on the PROC PLOT.

In addition to the above, include the following from our recent discussion of SAS programs. You may have included these previously.

```
dm'log;clear;output;clear';
options nodate nocenter nonumber ps=512 ls=132;
ODS HTML style=minimal rs=none body='C:\TEMP\anyname.html' ;
```

Also, sort the data by the variable GROUP, output selected statistics from a MEANS statement and print the output. The statements you need will resemble the following.

```
Title2 'PROC Means BY DIET';
proc sort data=MouseFeed; by diet lifetime; run;
Title3 'Means statement with a noprint option and output';
Proc means data=mousefeed noprint; by diet;
  var lifetime;
  output out=dietmean n=n mean=ybar stddev=std stderr=stderr median=med;
run;
Title3 'Listing of means from an output statement';
proc print data=dietmean; run;
```

You may use the same output variables I used in the mouse diet analysis. Note that other output variables could be requested by using any of the following options as keywords in the output statement: n, mean, stddev, min and max, CLM, RANGE, CSS, SKEW, CV, KURTOSIS, STDERR, LCLM, SUM, SUMWGT, UCLM, USS, VAR, NMISS, the quantiles MEDIAN|P50, Q3|P75, P1, P90, P5, P95, P10, P99, Q1|P25, QRANGE.

The dataset is given below. The datasets from the text come as CSV files (comma separated values).

"ZINC", "GROUP"	174, "PREGVEG"	174, "PREGVEG"
185, "PREGNON"	202, "PREGVEG"	184, "PREGVEG"
189, "PREGNON"	171, "PREGVEG"	186, "PREGVEG"
187, "PREGNON"	207, "PREGVEG"	210, "NONPREGVEG"
181, "PREGNON"	125, "PREGVEG"	139, "NONPREGVEG"
150, "PREGNON"	189, "PREGVEG"	172, "NONPREGVEG"
176, "PREGNON"	179, "PREGVEG"	198, "NONPREGVEG"
171, "PREGVEG"	163, "PREGVEG"	177, "NONPREGVEG"