This exercise uses the ***permanent*** IVH dataset that you saved in Homework 2 (the one that has missing values coded as '**.**') or use mine (g:\shared\bio113\ivh1.sas7bdat). This exercise may be done as a ***single long program (preferred)*** or as a ***series of small programs***. As always, turn in your program, SAS log and output.

Starting with your ***permanent*** SAS dataset, **ivh1, *add*** the following **8** new variables to it and make a ***new permanent*** dataset named **ivh2**.

Using the weight adjusted sum of colloid and crystalloid variables (cc1-cc4) that you created in homework 2 calculate the difference between cc1 and cc2, cc2 and cc3, cc3 and cc4 (**first variable minus the second**). Name the 3 resulting variables name them ccdif1, ccdif2, and ccdif3. Using the percent weight change variables you created in homework 2, calculate the difference between pctwt2 and pctwt1, pctwt3 and pctwt2, pctwt4 and pctwt3 (**second variable minus the first**). Name the resulting variables wtdif1, wtdif2 and wtdif3.Use ARRAYS to create these 6 variables.

Create a variable that divides gestational age into 3 categories: < 26 weeks, 26-28 weeks, and > 28 weeks.

Use a SAS function to abstract each baby's plurality from his/her ID variable. Remember, the ID has a 6-digits; the last digit is 1 if the baby is a singleton, 2 if a twin, and 3 if a triplet.

Use PROC MEANS (ask for options mean, min, max, n, t and probt) with your 3-level gestational age category variable as the 'by' variable to see if the three weight-adjusted fluid differences and the three percent weight change differences are different from 0 in babies in the gestational age categories (these will be paired t-tests within levels of gestational age category).

Use PROC MEANS (without t and probt) to see the mean birth weight, t4 value, apgar ratio, lowest pCO_2 and highest pCO_2 (these last 3 variables were created in homework 2) for singletons, twins, and triplets.

Create a ***temporary*** SAS dataset where you ***change the shape of the data set*** so the ***unit of observation*** is a ***'baby-day'*** (at the moment, it's a ***baby***). The dataset should contain only ***four*** variables named WT, MAP, PCO2, and DAY and 2,264 observations. Use PROC MEANS with DAY as the 'by' variable to see summary statistics fore each day for WT, MAP and PCO2.

Create another ***temporary*** SAS dataset of babies with 1 minute apgar scores < 2 who had an IVH. Then use PROC SORT, PROC PRINT, PROC FORMAT, LABEL, TITLE, and FOOTNOTE to make a report ***as similar as possible*** to the one on the next page (consider observations, order of variables and observations, labeling and titles).

```
Babies with 1 minute Apgar < 2 and IVH

Caesarian delivery=No

                 Gestational    Birth     Antenatal      Duration     Length
      id   sex      age        weight   corticosteroid   of labor    of stay

    137062   F       24          680        Complete        47.7        111
    220391   M       24          784        None             1.0         87
    112341   M       24          810        Complete         9.0         74
    136711   M       25          730        None             5.7        110
    210811   M       25          861        Partial          .           98
    112291   M       26          810        Partial          9.0         58
    137252   M       27         1093        None             7.9         80
    210121   F       30          830        None             2.0        117
    111641   M       30         1470        Partial         21.1        112
    ------                                                             -------
      cs                                                                 847


Caesarian delivery=Yes

                 Gestational    Birth     Antenatal      Duration     Length
      id   sex      age        weight   corticosteroid   of labor    of stay

    220561   F       24          930        None             0.0         44
    135431   F       25          650        Complete         1.2         54
    124861   M       25          950        None             0.0        206
    124021   F       26          630        Complete        13.0         33
    210891   F       26          683        None             .          122
    124471   M       26          880        Complete         0.0          0
    210061   F       26          943        None             .           72
    135121   M       26          953        None             0.8         91
    137201   M       27          940        None             0.0         73
    112351   M       27         1200        None             0.0         37
    110971   F       28         1180        Complete        53.0         43
    136971   M       28         1330        None             4.0         87
    112431   M       30         1100        None            14.0         67
    ------                                                             -------
      cs                                                                 929
                                                                   =======
                                                                     1776


DEN Study Data
```

Return to the *permanent* large dataset (N=566). Use PROC FORMAT to create formats for the variables pih, hosp and acs. Then use PROC FREQ to make a *3-way* table of pih, hosp and acs (make pih the 'paging' variable and hosp the 'row' variable). Specify 'cmh' as an option on the TABLES statement to get Cochran-Mantel-Haenszel statistics. Use the formats created with PROC FORMAT to label the values of the variables in the tables.

You may use PROC PLOT or PROC GPLOT for the plots.

Plot the fluid on day 2 ('y' variable) against the change in weight on day 2 (pctwt2 created in homework 2). Code **href=0** (plot y-var*x-var / href=0;) to draw a vertical line at x=0.

Finally, plot the length of stay ('y' variable) vs birth weight and request that ivh be the plotting symbol (create a format for ivh and use that if you don't like the 0's and 1's).