**Workshop Exercise 2**

**Exercise II: Recoding Variables**

**Dataset = afifi.sas7bdat**

In this exercise we will use the permanent SAS file afifi.sas7bdat, which is located in the sasdata2 folder. The codebook for the afifi.sas7bdat dataset can be found in the Appendix.

1. Save a copy of the permanent afifi dataset as work.afifi. Now use the temporary work.afifi dataset for the following exercises.
2. Obtain descriptive statistics for all variables. Does any variable seem to have missing values?
3. Obtain the n, mean, median, 25th percentile and 75th percentile for SBP1. Consider using Proc Univariate or Proc Means.

The next questions involve recodes to create new variables. To do this, you can either modify your data step to include the coding for the new variables, or create a new data set with the new variables in it.

1. The shock type variable (SHOKTYPE) has six categories. Now recode it into another variable (SHOCK) that takes only two levels. See below for the labels of the categories.

|  |  |  |
| --- | --- | --- |
| Variable name = shoktype |  | Variable name = shock |
| 2 = non-shock3 = hypovolemic4 = cardiogenic5 = bacterial6 = neurogenic7 = other | Recoded into | 0 = Non-Shock1 = Shock  (includes hypovolemic, cardiogenic, bacterial, neurogenic, or other) |

1. Create a dummy variable called LIVED (1 = lived; 0 = died) based on the variable, SURVIVE (SURVIVE is coded 1=lived; 3=died).
2. Create a dummy variable called FEMALE (1=female; 0=male) based on the variable, SEX (SEX is coded 1=Male; 2=Female).
3. Calculate a new variable called SBPDIFF, which is SBP2 – SBP1. Why do we subtract SBP1 from SBP2, and not the other way around?
4. Calculate a new variable called SBPAVG, which is the mean of SBP1 and SBP2.
5. Calculate a new variable called LOWSBP1 that takes on the value 1 if the patient’s SBP1 is less than or equal to the 25th percentile for SBP1 and 0 if above the 25th percentile.
6. Get descriptives for your new variables for all cases, and separately for those who lived and died. What is the mean change is systolic blood pressure for all cases? What is the mean change in systolic blood pressure for those patients who lived? For those who died?
7. Get a histogram of the new variable, SBPDIFF for those who lived and those who died. What is the pattern that you see?
8. Get frequency tabulations of all the new categorical variables that you created. What percent of all patients had low systolic blood pressure based on your new variable?
9. Get a crosstab of LOWSBP1 vs. LIVED. What percent of patients with low SBP1 (LOWSBP1=1) lived? What percent of patients who did not have Low SBP1 (LOWSBP1=0) lived?
10. Save your dataset, work.afifi2 as a permanent SAS dataset in your sasdata2 folder called afifi2.sas7bdat.