

Biostat 510: Statistical Computing Packages

SPSS Homework 7

Due Tuesday, March 30, 2004

Be sure to include your SPSS commands for each question as part of the homework. You can either get the SPSS commands by typing them directly into the Syntax Window, or by going to the appropriate dialog box and filling it out and then pasting to the syntax window. Remember, commands that are pasted into syntax will not be run until you highlight and submit them from the syntax window.

Provide appropriate annotated output for each question (output isn't necessary for questions that require syntax only). Use Copy Objects to get the output into a word processor. (You can experiment with just Copy and Paste or Copy Objects and Paste. See which works better). Be sure to size your SPSS output in your homework, so that you don't cut off parts of the output tables. Just grab the corner with your mouse, and drag it diagonally, so it fits within the print area.

- 1) Import the Excel file from homework 1 (hrtrate_2004.xls from my web page) into SPSS. Even if you have your own version of this dataset, please download it from my web page: <http://www.umich.edu/~kwelch>, so everyone will be working with the same data.
 - a) Check the output window to be sure your data was imported correctly.
 - b) Set up missing values using syntax in the syntax window. Be sure to submit your syntax, so it will be run. Note: missing value codes for Height and Weight are 999. There are no missing values for Pulse1 and Pulse2. You will also need to set up missing value codes for Htunit and Wtunit. These are "999" and " ".
 - c) Create value labels for Group (1=Ran, 2=Did Not Run) using syntax in the syntax window.
- 2) Compute WTKG, HTM and BMI, appropriately calculated for everyone in the data set.
 - a) WTKG: The appropriate syntax for this new variable is shown below:

```
IF (wunit="LBS") wtkg = weight/2.2046 .
EXECUTE .
IF (wunit="KG") wtkg = weight .
EXECUTE .
```

 - i) If you wish to use the point-and-click method, go to the Data Window (be sure the Data view is active), select Transform...Compute. Type in WTKG as the Target Variable, and fill in **Weight/2.2046** in the Numeric expression box. Then select If...and in the dialog box that pops up, select "include if case satisfies condition", and in the box, type: wunit="LBS" . Click on "continue", then go back to the dialog box and select "Paste". Go to the syntax window and submit these commands.
 - ii) For those with weight reported in KG, go to Transform...Compute. Type in WTKG as the Target Variable, and fill in **Weight** in the Numeric expression box. Then select If...and in the dialog box that pops up select "include if case satisfies condition", and in the box type: wtunit="KG". Click on "continue",

then go back to the dialog box and select "Paste". Go to the syntax window and submit these commands.

- b) Compute height in meters (HTM), using either syntax or by filling in the appropriate Compute Menus. Remember, $HTM = \text{height}/39.37$, if height is measured in inches, and $HTM = \text{height}/100$, if height is measured in centimeters. Either fill in menus, or simply type the syntax. If you use the point and click method, be sure to paste the syntax to the syntax window.
 - c) Compute BMI as $WT/KG/HTM**2$, using either point and click or syntax. If you use the point and click method, be sure to paste your commands to the syntax window.
- 3) Use Automatic Recode to create Gender (a numeric variable) from SEX, a string or character variable.
- a) Go to the Data Window and Select Transform...Automatic Recode...Click on Sex and add it to the variable box
 - b) Type Gender into the Newname box and click on Newname, so it gets sent to the upper box. Click on "OK" or "Paste"
 - c) Note: if you have already got the variable, Gender, in your dta set, you will not be able to do the recode. You will have to remove it from your data set before you can recode it again.
- 4) List all variables for the first 10 cases in your data set. Get a dictionary of your file, that shows all of the variables, the coding and any labels you have set up.

```
list variables=all
      /cases = from 1 to 10.
display dictionary.
```

- 5) Get descriptive statistics for all numeric variables in your data. Include the output in your homework.
- a) Go to the Analyze...Descriptive Statistics menu. Select Descriptives. Highlight and select all variables in the list (all numeric variables). Note the number of observations, min, max and mean for each variable. Check your values to be sure they all make sense.
 - b) What is the listwise number of complete cases?
- 6) Get a histogram, box plot and descriptive statistics for Pulse2 for those who ran and those who did not run. To do this, go to the Analyze menu, select Explore, and fill out the appropriate boxes. You will enter Pulse2 in the "Dependent List", and enter Group in the "Factor List". When done, select paste, then go to the syntax window and submit your commands. Include the graphs and descriptive statistics in your write-up.
- 7) Calculate an independent samples t-test to compare Pulse2 for those who ran and vs. those who didn't run. To do this, go to Analyze...Compare means...Independent samples t-test and fill out the dialog box.
- a) What are the null and alternative hypotheses? What are the results of the homogeneity of variance test? Report the test statistic, degrees of freedom and p-value. Which t-test results will you use?
 - b) What is the value of the t-statistic, the degrees of freedom and the p-value?
 - c) What do you conclude?
- 8) Calculate a paired t-test to compare the means of Pulse1 and Pulse2 for all subjects.
- a) What is the correlation of these two variables?

- b) What is the value of the paired t-statistic, the degrees of freedom and the p-value?
- c) What do you conclude about the mean of Pulse1 and Pulse2 for all students?
- 9) Calculate a paired t-test to compare Pulse1 vs. Pulse2 separately for those who ran and those who did not run.
 - a) First split the file by ran. To do this, go to the Data Editor Window and select Data...Split File...click on "Organize output by groups" and select Group as the variable to use for the split file. Then paste your commands to your syntax editor and submit them.
 - b) Now, resubmit the commands for the paired t-test.
 - c) What are the means of Pulse1 and Pulse2 for those who ran? For those who did not run?
 - d) What do you conclude about the change in mean Pulse rate from time 1 to time2 for those who ran? For those who did not run? Give the value of the paired t-test t statistic, the degrees of freedom and the p-value for those who lived and for those who died.
 - e) Turn Split File Off.
- 10) Save your work
 - a) Save your syntax commands as homework7.sps. Hand them in with your writeup.
 - b) Resave your data set as heartrate.sav so that your any changes to the data set will be saved.