

Biostat 510
Homework 1
Due Tuesday, January 19, 2010

1. **Download the Excel files** for the six groups from the first day of class. Save them in a folder you will be able to access later. They are in the Biostat 510 section of my web page at http://www-personal.umich.edu/~kwelch/510/2010/homework_2010.html

Check the data in the Excel files, **fix any problems** and **resave** the fixed Excel files.

- a) Open each file in Excel and make sure the **first row** of the Excel file contains the variable names.
 - i. Make sure the variable names are all the same (see the list of variables below). Make sure there are no blanks in the variables names.
 - ii. Note the **Sheet name** on the tab at the bottom of the worksheet.
 - iii. Make sure that **GROUP is filled down** for all cases in each spreadsheet.
 - iv. Check the coding of variables to be sure that they conform to the coding scheme shown below.

You should have the following variables in each Excel file:

<u>Variable Name</u>	<u>Description</u>	<u>Type</u>
GROUP	Group Number	Numeric
ID	ID Number within Group	Numeric
RAN	1= Ran, 0=Did not run	Numeric
AGEYR	Age in years	Numeric
AGEMO	Age in months	Numeric
SEX	M or F	Character
HR1	Heartrate at time 1	Numeric
HR2	Heartrate at time 2	Numeric

(NB: Upper and lower case for the Variable Names do not matter, but they should all be spelled exactly the same way!!).

- b) **Resave** each of the Excel files after you have made any corrections.
 - c) **Close** all the Excel files before importing them into SAS.
2. Use the SAS Import Wizard to import the Excel files into SAS.
 - a) **See the handout on importing Excel files into SAS for specific intructions.** Save the SAS commands for importing all of the Excel files into SAS, and call your command file homework1.sas.
 - i. Be careful about naming each dataset, and be sure you have the correct sheet name for each dataset in your SAS commands.
 - ii. Include your SAS commands for importing each Excel file in your homework1.sas command file.

- iii. Be sure that you include the SAS commands, as shown below, for each group, don't just open the datasets using the Import Wizard without saving the corresponding SAS commands.

Sample SAS commands for importing Group1 data are shown below:

```
PROC IMPORT OUT= WORK.group1
            DATAFILE= "BIOS 510 Grp1 data.xlsx"
            DBMS=EXCEL REPLACE;
RANGE="Sheet1$";
GETNAMES=YES;
MIXED=NO;
SCANTEXT=YES;
USEDATE=YES;
SCANTIME=YES;
RUN;
```

3. Get descriptive statistics for the numeric variables in all six of the datasets using Proc Means. Include the SAS commands for all six groups in your SAS command file. Sample SAS commands for getting descriptive statistics for Group1 in SAS are shown below:

```
title "Descriptives for Group 1";
proc means data=group1;
run;
```

Include the SAS commands for these descriptive statistics in homework1.sas.

- ✓ Include the output for the descriptive statistics for group 1 only in your output.
4. **Combine the data sets** for all groups by **stacking** them into one temporary data set, using a **set** statement in SAS. Your SAS commands will look something like those shown below (e.g. for 3 groups):

```
data Allgroups;
    set group1 group2 group3;
run;
```

- a) Browse your Allgroups data set in SAS viewtable mode to check it out.
 - b) If there is a problem in creating your data set, go back and fix it, and recreate the data set by highlighting and resubmitting your commands.
 - c) Include the SAS commands for combining these SAS datasets in homework1.sas
- ✓ Include the portion of your SAS log showing the successful creation of the Allgroups dataset in your homework writeup.
5. Get Proc Contents for the Allgroups data set, with the variables in **creation order**, by using the **varnum** option.
 - a) How many observations are there in your dataset?

- b) How many total variables are there?
 - c) How many character variables? How many numeric variables?
- ✓ Include the output from Proc Contents in your homework write-up.
- ```

title "Contents of Allgroups Dataset";
proc contents data=allgroups varnum;
run;

```
6. Get descriptive statistics for all numeric variables in your Allgroups dataset using Proc Means.
- a) What is the average age in years in this class? What is the standard deviation of age in years? What is the minimum? What is the maximum? What is the n for age in years?
- ✓ Include this output in your homework write-up.
7. Use a **class statement** to get descriptive statistics for each GROUP. Include this output in your homework write-up.
- a) Based on these results, how many students are in each group?
  - b) What is the average age in years for each group?
- ✓ Include this output in your homework write-up.
- ```

title "Descriptives for Allgroups Dataset for each Group";
proc means data=allgroups;
class group;
run;

```
8. Get frequencies for the categorical variables GROUP, SEX, and RAN, using Proc Freq.
- a) How many students are females? What proportion are females?
 - b) How many students ran? What proportion ran?
- ✓ Include this output in your homework write-up.
9. Save your dataset as a permanent SAS dataset, using commands similar to those shown below:
- ```

libname b510 "e:\510\2010";
data b510.allgroups;
set work.allgroups;
run;

```
- ✓ Include the portion of the SAS log showing the successful creation of the permanent SAS dataset.
10. Make sure you save your SAS commands as homework1.sas. Re-run all of the commands in this command file and make sure there are no errors in your log.