

ST 307: Topic 9 Activity

For this activity you will create a SAS program and upload that program to Moodle. Everyone must submit their own code. Be sure that your SAS file adheres to the SAS file submission guidelines (available on Moodle).

The data set for this activity is available on Moodle in the Activity 9 link. You should download this and put it in a folder on the VCL that you will use as a SAS library.

In this in-class activity we will be working with blood pressure data set contains data from random sample of individuals with high blood pressure. The bloodpressure.csv file contains the following variables:

- Subject: subject code
- Age: subject's age
- BaselineBP: subject's baseline diastolic blood pressure
- NewBP: subject's diastolic blood pressure one month after starting to take medication

1. Create a new library called "Week9".
2. Read in the bloodpressure.csv with DATA step and the following specifications:
 - a. Save the data set as "bloodpressure" in the Week9 library.
 - b. Create a new variable called "agegroup" which separates subjects into 2 age groups. If the age is less than the median value 61 then agegroup = "0" otherwise agegroup="1".
 - c. Create labels for all variables (including the new variable you created) based on variable information.
3. Use PROC TTEST procedure to answer the following questions:
 - a. Researchers expect that the medication was effective in reducing blood pressure. State your null and alternative hypotheses in your comment.
 - b. Test whether there is difference between the baseline and new blood pressure at the 0.1 significance level.
 - c. Report p-value and test statistic in your comment. State your conclusion.
 - d. Comments on the normality assumption.
4. Use PROC TTEST procedure to answer the following questions:
 - a. We want to see if baseline blood pressure has difference between age groups. Test the equality of average baseline blood pressure between two age groups at the 0.1 significance level.
 - b. Comments on whether we will be using a pooled or un-pooled variance estimate.
 - c. Report p-value and test statistic and state your conclusion.

5. Create a new data set called "groupbp" to library "Week9" and answer the following questions:
- Drop variable agegroup.
 - Create a new variable called "idgroup" that has only the first letter of subject code. Subject code starts from alphabet "N" or "A". If the first letter is "N", idgroup="N" otherwise, idgroup="A". Hint: Use SUBSTR() function.
 - Sort groupbp data set by subject group.
 - Test again whether there is difference between the baseline and new blood pressure at the 0.05 significance level by two subject groups separately.
 - Report p-value and test statistic for each subject group and state your conclusion.