Engineering Analytics II

HW2, Due on March, 1st

The data (data HW2) were extracted from a series of consecutive patients who underwent Coronary Artery Bypass Surgery. Some of the patients were receiving a medication named "clopidogrel" prior to surgery. This medication impairs the functioning of blood platelets, cell fragments that are critically important in blood clotting. Some patients also received a daily aspirin, which also affects platelet function and may increase bleeding. The primary question to be answered is does treatment with clopidogrel increase the amount of blood lost during and after surgery.

There are many known cofactors that lead to surgical complications and data are presented for these as well: height, weight, body surface area (BSA), diabetes, preoperative mortality score. Body surface area is calculated from height and weight. The outcome measure relating to blood loss is 24EBL, the total blood loss in the first 24 hours. Also included the log transform.

1. Determine the summary descriptive statistics for each variable and plot a histogram. Compare the mean and median for each variable. Compare the Interquartile range with the standard deviation. Interpret your results.
2. For continuous variables, determine if they are normally distributed. If not, determine if the logarithmic transformation normalizes the data with an appropriate test.
3. What input variables are correlated? How can you use this information to analyze the data?
4. Calculate the covariance matrix for the input variables and for the outcome variables. Calculate the two correlation matrices and verify which variables are independent and which are correlated. Graph each pair of income variables (scatterplots) and visually establish if you agree with the correlation coefficients
5. Use the age, weight, mortality score, in PCA and describe the result.
6. Use the same variables in a Single value decomposition Analysis and interpret the results.
7. Compare blood loss between the two clopidogrel groups using any test (parametric or non-parametric test)
8. Now take a reduced subset of your predictor variables, eliminating unnecessary covariates, and determine if this changed the outcome analyses for total blood loss.
9. Finally, do you believe that clopidogrel affected the outcome (estimated blood loss 24 hours)?

Present a written report (Word Document) including the R code you use (inside the word document). Include graphs and everything you need to support your answers. Please do not submit any other file, just the Word file. Late submissions are not allowed.