**Data programming with SAS: Assignment 2**

The dataset assignment2.xlsx is an extract of data from an Oral Lipid Tolerance Test. The objective was to examine how the participants respond to a lipid challenge. Participants consumed a high fat drink. Bloods were taken at 0, 10, 20, 30, 60, 90 and 120 minutes. The blood chemistry was analysed for a number of responses. The dataset contains the following variables:

|  |  |
| --- | --- |
| ID | Participant identification code |
| time | Time (minutes) |
| SEX | Sex, 1=male; 2=female |
| AGE | Age (years) |
| BMI | Body Mass Index (kg /m2) |
| ApoA1 | Apolipoprotein A1 |
| ApoB | Apolipoprotein B |
| ApoC2 | Apolipoprotein C2 |
| ApoC3 | Apolipoprotein C3 |
| ApoE | Apolipoprotein E |
| Glucose | Glucose |
| NEFA | Non-esterified fatty acid |
| Insulin | Insulin |
| Chol | Cholesterol |
| TAG | Triglyceride |

Write SAS code to answer the following questions. Write a short report summarising your findings.

Upload both the SAS code and the report.

1. Using numerical and descriptive statistics, summarise the dataset and check for outliers.
2. Assess the correlations between the blood chemistry responses (ApoA1, ApoB, ApoC2, ApoC3, ApoE, Glucose, NEFA, Insulin, Chol, TAG).
3. What is the partial correlation between Glucose and Insulin, adjusting for BMI?
4. Selecting only baseline measurements (time 0), construct a linear model relating baseline triglyceride to age, sex and BMI. Does the mean differ for males and females?
5. Repeat the model in question 3 above, this time assuming that triglyceride follows a log-normal distribution
6. Using data from all timepoints, fit a general linear mixed model relating insulin to age, sex and bmi. Fit the repeated measures variances covariance matrix assuming an AR(1) structure. Fit the model again with an ARH(1) variance covariance matrix. Compare the fits of the two models. Is the variance homogenous across time?