Paper 475.330 **Assignment 1** **Due: Friday 11 August**

For the "Introduction to the Advanced Lab" tutorial you explored data obtained from the *U.N.E.S.C.O. 1990 Demographic Year Book* and *The Annual Registrar 1992* giving birth rates, death rates, life expectancies, and Gross National Products for 91 countries. In this assignment you are asked to explore some aspects of this dataset and to write a report that clearly communicates your findings.

The data consists of the following variables:

birth	live birth rate per 1,000 of population
death	death rate per 1,000 of population
i.death	infant deaths per 1,000 of population (< 1 year old)
life.m	life expectancy at birth for males
life.f	life expectancy at birth for females
gnp	gross national product
group	1=Eastern Europe, 2=South America and Mexico, 3=Western Europe, North America, Japan, Aus., 4=Middle East, 5=Asia, 6=Africa

This assignment consists of two parts:

1. Use graphical techniques to explore the relationships between birth rates, death rates, and infant death rates. Also explore the way these variables and the relationships between these variables are affected by GNP and region?

   You do not need to fit any regression models. Discover as much as you can by using graphical techniques. Identify what you think are the key (most interesting) aspects of the data and describe these in a report. You should create a few informative plots that help you communicate your ideas. Do not clutter your report with a description of everything you did. Instead, include a concise statistical appendix to indicate what techniques you used.

2. Create a simple regression model that relates the life expectancy of women to men (ignoring all other variables). What would you conclude about the relative life expectancies of women and men based on this fitted model? Make sure that the method you use to compare the life expectancies of women and men that takes into account the uncertainty in the fitted model.

   Perform standard diagnostic procedures and comment on the suitability of this model. Note you are not asked to "fix up the model" - just identify any problems. Finally use trellis plots to evaluate whether the relationship between life expectancies is affected by group or by Gross National Product.

This assignment should be handed in to the appropriate box in the basement of the Maths/Physics building by the SMIS Resource Centre, by 4pm on Friday, 11 April.