Homework 2

Problem1. The file Utility.dat contains a monthly record of telephone, electricity, and fuel costs for several years. Naming the 5 columns as Month ,Year, Telephone, Electricity and Fuel.

(a) Write a SAS program to read the above data set into SAS using the INFILE statement, print out the **first ten observations** of the dataset, i.e the first ten rows of the whole data table.

(b) Create a new variable, named Total, which is the sum of Telephone, Electricity and Fuel. Print the first ten observation of Month,Year, and Total.

(c) Given the dataset from (b), first run the PROC SORT BY Year Month ; You will find out that within Year, Month is in the ridiculous initial-letter-order rather than calendar month order. By both orders I mean, for example, suppose we have three months to sort within a year: {Jan, Feb, Mar}, the calendar monthly order is (Jan, Feb, Mar), and the initial-letter-order is (**F**eb, **J**an, **M**ar). Suppose we want to sort the dataset by year and calendar month order using **PROC SORT by two variables**, how do we achieve it ? And once detest sorted by year and calendar month order please print Month, Year and Telephone ,when Year = 88 and 92. (Hint: Since SAS don't quite understand what the variable Month mean, so we should ….)

Problem2. The file China.dat contains export and import information (in dollars) by year. Naming four columns as Years Total, Exports and Imports.

(a) Read the above data set into SAS using the INFILE statement, print out the first ten observations of the dataset.

(b) Create a new variable TradeBalance = Exports – Imports. Print TradeBalance and Years in 50s(Years =1950, 1951,…,1959).

(c) Using PROC MEAN to find the mean and s.d. of TradeBalance for 50s and 80s, respectively. Report your result.

Problem3. The file Handinj.dat contains the costs (in Irish pounds) and lost work days due to hand injuries for workers in Dublin, Ireland. Naming the four columns as ID,Type, Days and Cost.

(a)Read the above data set into SAS using the INFILE statement, print out the first ten observations of the dataset.

(b)Cost is in Irish pounds. Now change it to U.S. dollars (1 Irish pound = 1.54 U.S. dollars). Print the new data set for the first ten observation.

(c)Sort the whole data set by Days in **decreasing** order. Then print ID and Days when Type = 'sport' .