Lab 10

10.1 **(Table look-up III)** Solve the problem 2 from Lab 8 using `PROC FORMAT`.

10.2 Create a data set **a** with five numerical variables *z1,...,z5* and 50 observations. Make every element of **a** a random number from the normal distribution with the mean 100 and the variance 10. Next, basing on **a**, create a data set **stat**, with the variable *stat* and the variables *z1,...,z5* and with 55 observations. For the observations numbered $n = 1, \ldots, 50$, the variable *stat* should have (text) values Wn, and the variables *z1,...,z5* should have exactly the same values as in the set **a**. For the observations numbered $51, \ldots, 55$ the variable *stat* should take the (text) values $N, MIN, MAX, MEAN, STD$, and the variables *z1,...,z5* should take the values of those statistics for the variables *z1,...,z5*.

10.3 Modify the data set **stat** from the previous problem to contain the values of quartiles, median and interquartile range for the variables *z1,...,z5*.

10.4 Write a code that create an informat reading strings like: *January 22, 2001*, *October 3, 1956*, and so on, as genuine SAS dates.

10.5 Based on the set **grades** create a data set **averages** with the average grades of each student for each course.

10.6 The data set **data** has the variables: *group*, *x* and *y*. Find the group for which the (group) averages of *x* and *y* are closest to the global averages of *x* and *y* (global means computed for the whole set **data**).

10.7 Define a format that displays numbers of the form $m.n$ $(m, n = 0, ..., 9)$ "in words". An example: 2.8 should be formatted as *two point eight*.

10.8 Define an *outlier of the order* α as any observation which lies outside of the range

$$(med - \alpha * Range, med + \alpha * Range),$$

where *med* and *Range* are the median and interquartile range respectively. Write a code that, for a given parameter *alfa* = 1 and any given data set with one numerical variable, will find all the outliers of the order α.