Fundamentals of Data Analysis: Assignment #1

Deadline: 10/14/2003 (Tuesday)

Please post to the mailbox next to the IS management office (2nd floor of IS building)

1. Suppose that a researchers tries to examine the effect of "years of experience" on the annual income of Japanese female financial analyst. Answer whether each of the following quantities is "variable" or "constant".

a. sex
b. annual income
c. nationality
d. profession
e. years of experience

2. Choose the most appropriate levels of measurements for the following quantities, from "nominal level", "ordinal level", "interval level", and "ratio level". Please write the reason why you choose them, together.

a. species of flowers
b. points of TOEFL test
c. speed of five horses in a 1000m race, as measured by the horses order of finish.
d. temperature of hot water, as measured by degrees in Celsius
e. luminance of CRT monitor

3. Download the file "data01.txt" on the homepage, and analyze the "points" data (ignore the other terms, such as "department"), following the steps below.

a. Construct a frequency distribution and a relative frequency distribution when the width of categories is 50. Set the lower bound of each category multiples of 50.
b. Draw a histogram and a graph of accumulated frequency distribution, based on the result of the above question.
c. Calculate the percentile rank of data #1, using the accumulated relative frequency distribution.
d. Calculate the 50th percentile point, using the accumulated relative frequency distribution.
e. Which does this distribution have, positive skewness or negative skewness ?
f. Is this distribution unimodal ?
g. Calculate "mean", "variance" and "standard deviation", based on the original data.
h. Calculate "mean", "variance" and "standard deviation", based on the frequency distribution.

4. Please feel free to write your comments and requests on this lecture (if any).