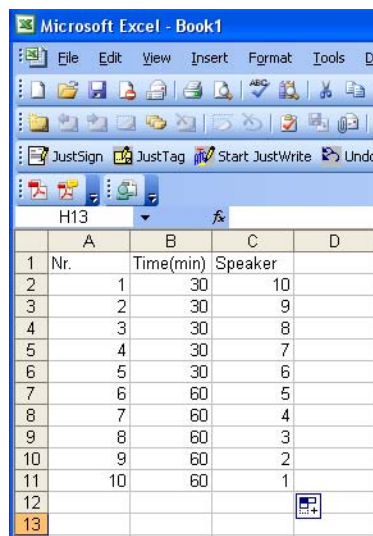# MA 222. Assignment 6

due Monday Jun. 21, 2010 at the beginning of the class.

1. For this problem you need to figure out how to input data into R. Read
   the documentation on the functions **read.table**() and **read.csv**().

Figure 1: Excel file

(a) Using Excel create a table like in the Figure 1. Save the file as a
    comma separated (.csv file). Use the functions in R to input the
    file into R. Make sure you understand the "header" option.

(b) Next we will input data downloaded from the internet. Go to the following address:
Copy and paste the table from there in Excel. Alternately, see the attached file. Save this file in csv format. Use R to read the table. Pay attention to the fact that the first two lines are not needed. Read documentation about the **read.csv**() command by typing **help**(**read.csv**). In particular find an option to *skip* the first two lines. You will need to use this data further.

2. For this problem you need to use the data you just introduced into R in problem 1b. You could use a piece of paper and a pencil. I do not advise the later option.

(a) How many variables does this data contain. Which are categorical and which are quantitative? Write a description of each variable as you understand what it represents.

(b) Make histograms of the value, revenue and income of the NBA teams. What is the shape of these distributions? Identify any outliers from the pattern. Do the same for the debt percent variable. Comment.

(c) Repeat the part above but construct boxplots instead of histograms.

Next we look at relations between variables. You want to predict team value by looking at income and at revenue. Operating income is the income produced by business after taxes and excludes income from sales of assets and investments which don't reflect the actual business. Total revenue ignores any costs incurred by the business. For a normal business revenue is less important than income for the value of the firm but NBA is not usual business.

(d) Plot team value against revenue (which one goes on the x axis?). There are some outliers. Which teams are they and in what way are they outliers? Is there a positive relationship? Is the pattern roughly linear?

(e) Now plot value against operating income. Are the same teams outliers? Which variable is a better predictor of team value?

(f) Calculate the correlation between revenue and team value as well as income and team value. Do these values reflect the plots you made earlier?

Output and include the images with your assignment. Alternately, using a calculator answer these points and create histograms yourself.