Lecture 1

L1.4. VARIABLES AND BASIC OPERATIONS

L1.4a. Compare two numbers: e^{π} and π^{e}. Print the results using "cat"
pi, exp(), "^", ">", "cat"

L1.4b. Create a vector of exponents of 2: 2^0, 2^1, 2^2, … 2^{10}
i:j, "^"

L1.4c. Output the results of 1.4b as a vector of strings with template: "2^***i*** = ***x***".
print, sprintf

L1.4d. Output the results 4b, showing only even exponents.
print, seq or "%%"

L1.5. DATA IMPORT AND EXPORT

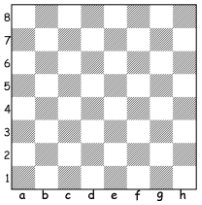
L1.5a. Dataset from http://edu.sablab.net/data/txt/shop.txt contains records about customers, collected by a women's apparel store. Check its structure. View its summary.
read.table, fix, str, summary, head

L1.5b. For the "shop" table, save into a new text file only the records for customers, who paid using Visa card.
write.table

L1.6. WORKFLOW CONTROL

L1.6a. Create a matrix 8x8. Fill it with 0. Using "for" loop change elements of the main diagonal to 1
matrix, for

L1.6b. Fill a matrix 8x8 with "1" to get a chess-board, where 0 codes a white cell and 1 – a black cell. Note, that black cells appear only if the sum of indexes is an even number (1+1=2, 1+3=4, 2+2=4, etc).



Well, in fact if you print the results you should see a rotated chess board (in chess the numeration of rows/columns starts from the bottom-left corner and in matrixes – from top-right one).
matrix, for, if

L1.6c. Calculate the sum of the first n=1000 members of the series:
s = (4/1) – (4/3) + (4/5) – (4/7) + (4/9) – (4/11) + (4/13) – (4/15) + ….

Guess, what is the limit of this series if $n \to \infty$

L1.7. DATA VISUALIZATION

L1.7a. Use dataset from http://edu.sablab.net/data/txt/mice.txt. Build distributions for male and female body weights in one plot.
plot, density

L1.7b. Draw boxplots, showing variability of bleeding time for mice of different strains.
boxplot

L1.8. DESCRIPTIVE STATISTICS

L1.1a. Use dataset from http://edu.sablab.net/data/txt/mice.txt. Calculate number of mice with bleeding time bigger than 2 minutes
read.table, sum

L1.1b. Report 5-numer summary for each column of Mice data
summary

L1.1c. For dataset Mice replace starting weight of any mice by 1000 (assume, there is a mistype). Calculate mean, median, standard deviation and median absolute deviation (MAD) of this weight. Compare the results with original measures.
mean, median, sd, mad

L1.9. DETECTION OF OUTLIERS

L1.2a. Use Mice dataset. Try to identify outlier mice on the basis of ***Weight change*** variable. Use simplest z-score, Iglewicz-Hoaglin and Grubb's methods. Build plots and report outliers.
scale, median, mad, abs, grubbs.test