Series 2

1. In an experimental setup oceanic bacteria were exposed to x-ray in 15 six-minutes intervalls. These are the results:

No. of bac.	355	211	197	166	142	106	104	60	56	38	36	32	21	19	15
Intervall	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

a) What does the relation between number of surviving bacteria and time of exposure look like? Is a linear regression reasonable?
Hint: `plot`.

b) Try to transform the data such that a straight line fits better.
Hint: One theory assumes that within each intervall the relative amount of bacteria killed is equal.

c) Estimate the size of the starting population of bacteria based on a regression analysis. Estimate also the relative decrease within each intervall.
Hint: `lm, summary`, see the R-Tutorial.

2. **a)** Generate a scatter plot of the following data:

x	0.34	1.38	−0.65	0.68	1.40	−0.88	−0.30	−1.18	0.50	−1.75
y	0.27	1.34	−0.53	0.35	1.28	−0.98	−0.72	−0.81	0.64	−1.59

b) Fit a straight line $y = ax + b$ using ordinary least squares (OLS) and draw it into the scatter plot.

c) Fit another straight line $x = cy + d$ using OLS and draw it also into the scatter plot.

d) Do the lines from b) and c) match? If no, why not?

3. The file `gas.dat` contains the gas consumption (in kWh) and the differences of temperature (in ℃) inside and outside of 15 houses which are heated with gas. The measures were collected over a long time span and then averaged.

a) Read in the data from the internet using
`read.table("http://stat.ethz.ch/Teaching/Datasets/gas.dat", header = TRUE)`.
Hint: Alternatively the data can be downloaded from the web using a browser and read in from the local drive using also `read.table()`. This could be necessary if you get an error reading it directly (e.g. caused by a stringent firewall).
Illustrate the data graphically. What does the relation look like?

b) Compute an ordinary linear regression `mod1` of the consumption versus the temperature difference. Compare the output when calling `mod1` and `summary(mod1)`.

c) Perform a diagnosis of the model. Does the residual analysis look satisfying?
Hints:
`plot(fitted(mod1),resid(mod1)), abline(h=0)`,
`plot(gas$temp,resid(mod1)), abline(h=0)` and
`qqnorm(resid(mod1))`, `qqline(resid(mod1))`.
Or `plot(mod1)`, which generates directly the above plots and an additional one.
If necessary, try to find an alternative model.

d) What kind of consumption do you expect when the difference in temperature is 14℃ ?
Give also the confidence intervall for the expected consumption. Hint: `predict()`.

Preliminary discussion: Monday, October 03.

Deadline: Monday, October 10.