Tutorial n$^{\text{o}}$ 1
Analysis of Variance

Exercise 1. *Eat apples !*
We want to compare the vitamin C values of five apples varieties, denoted by V_1, V_2, V_3, V_4 et V_5. For every variety the vitamin C value was measured, units $mg/(100g)$, in five randomly chosen apples. We got a dataset which is summed up in the following table :

V_1	V_2	V_3	V_4	V_5
93,6	95,3	94,5	98,8	94,6
95,3	96,9	97,0	98,2	97,8
96,0	95,8	97,8	97,8	98,0
93,7	97,3	97,0	97,2	95,0
96,2	97,7	98,3	97,9	98,9

What can we deduce from the data ? To answer to this question, use the analysis of variance.

This exercise is taken from the exercise book of François Husson and Jérôme Pagès : Statistiques générales pour utilisateurs, *PUR.*

Exercise 2. *Tasting beers*
Eight marks of acidity were collected for four white beers. These marks were rounded to the integer and summed up in the following table.

	beer 1	beer 2	beer 3	beer 4
mark 1	5	0	5	0
mark 2	5	1	6	0
mark 3	5	2	6	1
mark 4	6	2	7	1
mark 5	7	3	8	2
mark 6	7	4	9	3
mark 7	8	6	10	4
mark 8	10	6	10	4

1. After having computed the summary statistics for the marks of any the four beers, use boxplots to display them.
2. We want to know whether or not these beers taste the same using the acid marks. What is an appropriate statistical tool ? Explain why.
3. Write down the corresponding statistical model.
4. State the analysis of variance table.
5. What is the percent of the variation of the acid mark explained by the beer factor ?

6. What test can you use to compare all the beers ? Write down the test procedure. Decide wether or not these beers taste the same using the acid marks with a type I error of 1 %.

Exercise 3. *Farming wheat*
An experiment is made to compare the yields (in quintals/hectares) of four varieties of wheat. These varietes are denoted by the letters A, B, C and D. The C variety is a new one yet the three other ones are known to the farmers. 4 parcels were randomly chosen for each of the four varieties and the results of the experiments were summed up in the following table :

Variety	A	B	C	D
Yields	50	52	57	53
	54	56	60	47
	53	57	59	47
	51	60	62	49

1. State the statistical models and tests it is legitimate to use in order to :
   a. assess whether the variety accounts for the cropping variations,
   b. the C variety is better than the three other ones.
2. Proceed to the computations.

Exercise 4. *Stressed rats*
Researchers try to study the effect of a new cure against stress.

First, the fifteen rats had to undergo a stressful experiment. Then they were divided in three groups, named A, B and C. An injection was made to the rats according to the group to which they belong. The rats of the group A received one dose of treatment, the rats of the group B received two doses of treatment and the rats of the group C received three dose of treatment. A measure of the stress of the rats was then carried out and the results were summed up in the following table :

group	A	B	C
variable stress	22.8	21.9	23.5
	23.1	23.4	19.6
	27.7	20.1	23.7
	27.6	27.8	20.8
	24.0	19.3	23.9