11 TECHNICAL BOOKLETS

Recommended tests

How to prepare the samples

How to analyze the results

TASTING GUIDE :

Tools to integrate organoleptic quality criteria in breeding programs

Authors : Camille VINDRAS, Nicolas SINOIR

Reviewers : Aude COULOMBEL, Sally HOWLETT, Frédéric REY, Bruno TAUPIER-LETAGE

Layout : Nicolas SINOIR, Aude COULOMBEL

ITAB Institut Technique de l'Agriculture Biologique

Solibam

SEVENTH FRAMEWORK PROGRAMME

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE
Ministère de l'Écologie, de l'Énergie, du Développement durable et de la Mer

CONTENTS

TABLE OF CONTENTS

How to use this guide

This tasting guide provides tools to implement sensory evaluation in order to complete the organoleptic quality objectives as required by SOLIBAM.

To achieve this, a set of experiments will be carried out to meet specific objectives. Four sensory evaluation tests are proposed: the first aims to integrate gustatory criteria in the breeding process, the second is a tool used to identify and understand the consumers' expectations. The third and fourth evaluate the impact of agronomic practices on the organoleptic characters of a product.

The final objective of this guide is to provide a common methodological basis for all SOLIBAM partners involved in WP7. It contains:

- An **introductory chapter**: Introduces the methods for sensory analyses.
- Four **technical booklets** detailing recommended tests: you will be guided from the implementation of your gustatory tests through to the analysis of the data. All the tools you need to perform the various tests will be found here. These booklets contain screen captures from the open-source software, 'R'. Installation instructions for R are explained in the last sheet.
- Five **products booklets** describing how to prepare samples : these sheets focus on the product specificity from sensory descriptors through to sample preparation
- A **glossary**: Defines the specific terminology used in the tasting guide (e.g. sensory analysis terms, sensory attributes, statistical terms).

A useful tool to reach SOLIBAM objectives

SOLIBAM Objectives

- **WP7** : Deals with impacts of the interactions between crop genotypes and management innovations on crop nutritional, organoleptic and end-use quality
- **Task 7.1.** : This task has two aims:

1. **To take into consideration the organoleptic criteria in the breeding process.**

This is essential to ensure both gustatory and agronomic criteria are integrated in the breeding process.

2. **To measure the impact of agronomic and breeding practices on the gustatory quality of the product and on its acceptability.**

This requires firstly to ascertain if there is a perceived difference between SOLIBAM's produce and the control, and secondly to determine if SOLIBAM's varieties are preferred by the consumer.

In other words, the gustatory selection has to be implemented from year one (it is the first objective). Once the varieties have been selected, the methodology will be validated in two ways:

1) **by experts** in order to determine if there is a breeding effect on organoleptic quality,
2) **by consumers** to determine whether the SOLIBAM varieties are preferred by the consumer or not.

Different analyses are required for different steps of the breeding process. The figures on the following page illustrate how to select the appropriate test depending on the specific objective.

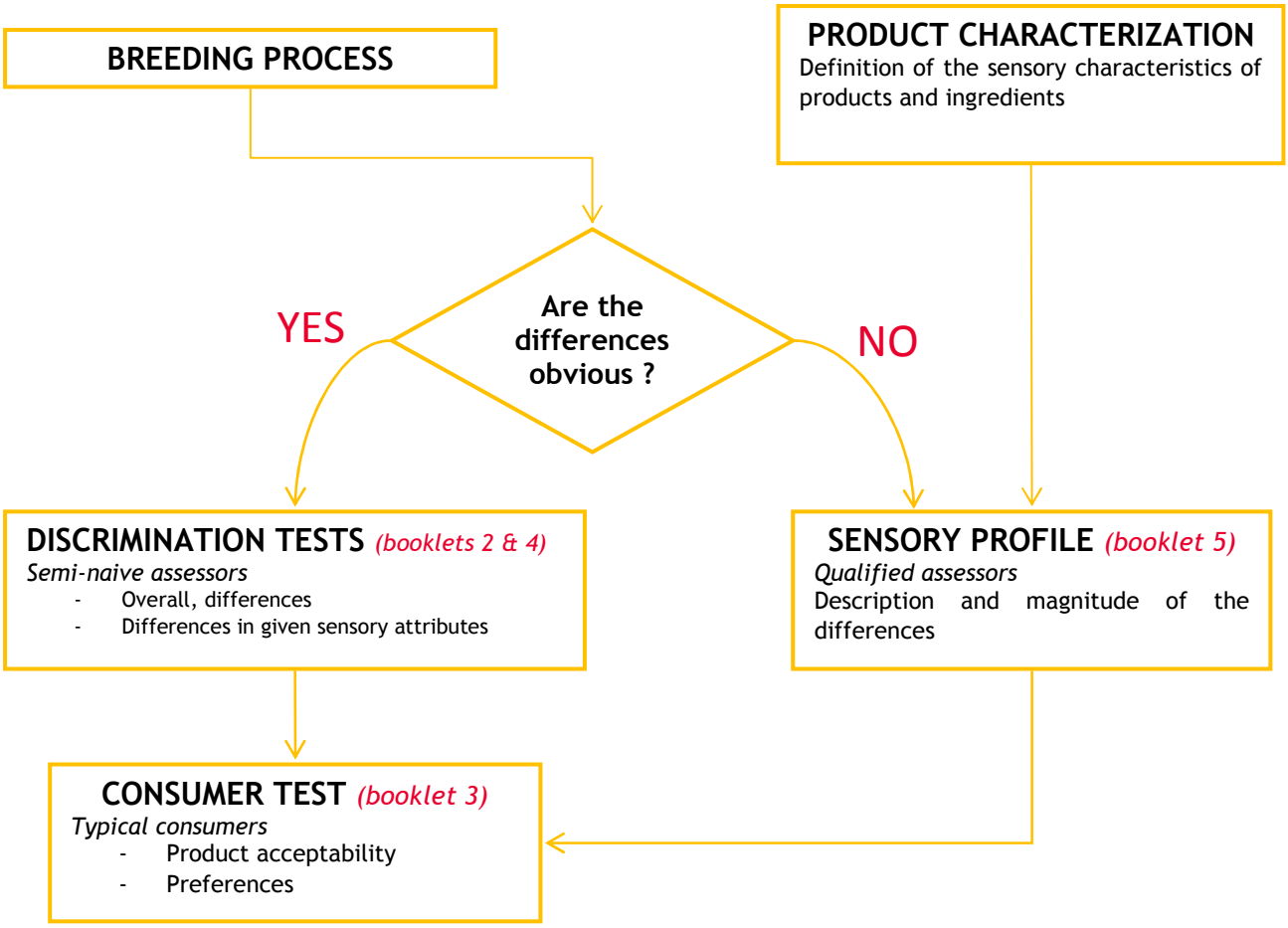


Figure 1: Selection criteria to choose suitable tests

(« De la perception à la mesure sensorielle », Fortin J. et Durand, N., Ed. La Fondation des gouverneurs, 2004)

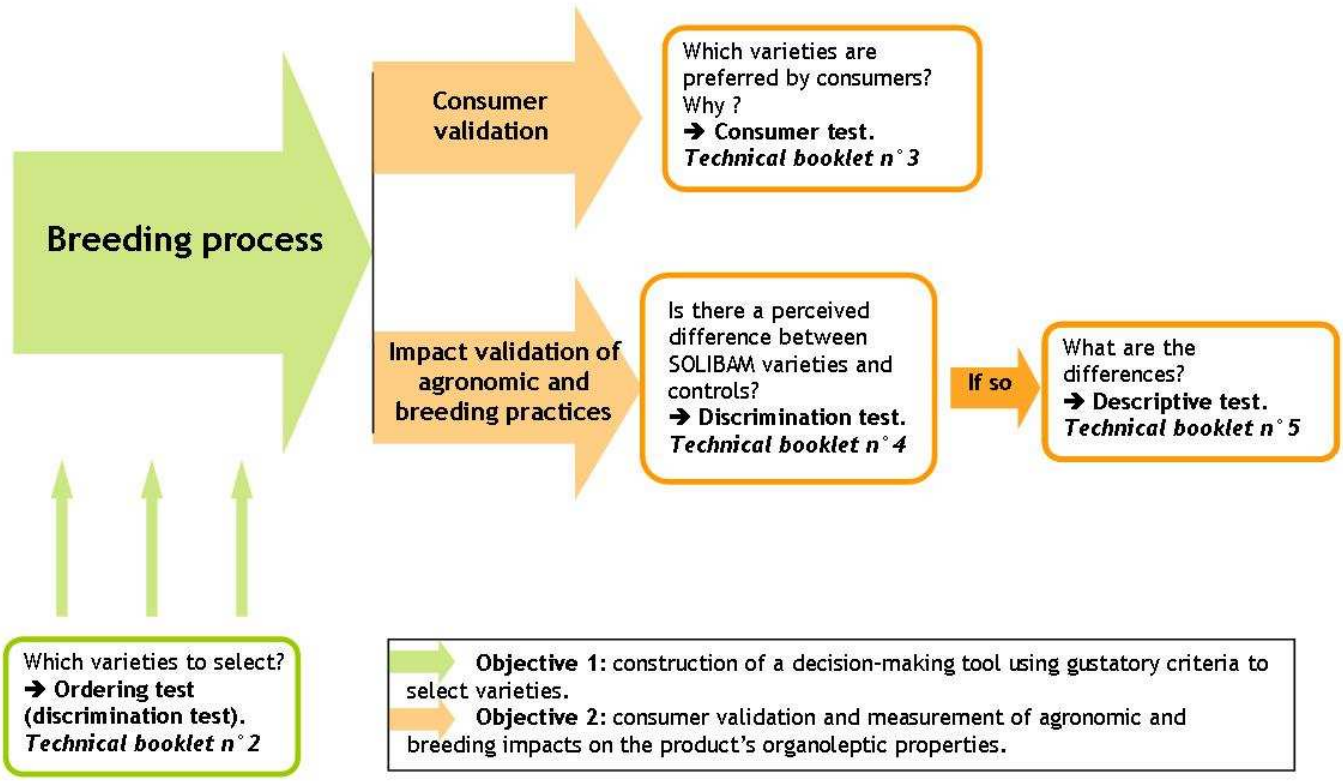


Figure 2: Global approach: construction and validation

1. Initiation to sensory analysis

From sensory perception to sensory analysis

Sensory perception results from the integration of information from multiple sensory organs:

Perception	Example
Vision	Shape, colour, appearance.
Olfaction	Odour (sweet, pungent, floral).
Gustation	Sweet/salt, sharp/bitter, flavours (savour, perfume in mouth).
Hearing	Crunchy,.
Sense of touch	Texture (smooth, rough), temperature, firmness.
Trigeminal perception	Fresh/hot sensations, astringency

Table 1: Different types of sensation in sensory perception

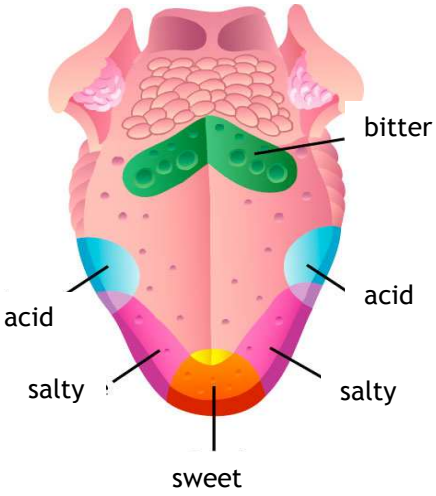


Figure 3: The gustatory chain

The gustatory chain begins with the taste buds, located on the tongue. Taste buds are grouped in specific areas corresponding to the detection of different flavours (Fig 3). There are four kinds of taste buds:

- **Caliciforms**, located at the back of the tongue
- **Fungiforms** (mushroom-shaped), located on the tip and sides of the tongue
- **Filliforms** (the most numerous type), responsible for tactile sensation (temperature and texture), but are not directly involved in taste perception. Tactile sensation completes the gustatory message.
- **Foliate papillae** (leaf-shaped), located on the tongue's edge.

Quality measurement: specificity of the sensory quality

The concept of quality can be broken down into agronomic, commercial, nutritional and gustatory components. Most of these can be measured objectively, for example, colour, firmness, juiciness, soluble dry matter, acidity analyses of nutritional compounds. The existence of automates able to make 4 or 5 of the measures at the same time highlights the instrumental technical advancement for the quality measurement. Some criteria are, however, purely subjective in nature, such as the mealiness of tomatoes or the toughness of a grapefruit skin grapefruit. These require sensory analysis methods. It is necessary to pay close attention to the selection of samples, in order to ensure the reliability and replicability of the experiment. The sample has to be representative of the batch (homogenous) and clearly described, with information on species, variety, origin, agronomic practice, harvest date, post harvest storage conditions and physical-chemical characteristics.. To compare varieties, it is of vital importance that all samples are at the same stage of maturity. They can be graded according to the appearance, the colour or the intrinsic characteristics of the product (IR, firmness, acidity).

2. Integrate the gustatory criteria in the breeding process : ranking test

Type of test	Discrimination test
Type of data	Sensory characteristics
Subject	Semi-naïve
Cost/difficulty €/*= cheap/easy, €€/ **= quite expensive/ average difficulty, €€€/***= expensive/ great difficulties.	€ / *

The test aims to integrate gustatory and agronomic characteristics in the breeding process. Easy to implement, it provides a product ranking on a given sensory attribute (for example the tomato's tenderness).

Table 2: Characteristics of the ranking test

This protocol has been developed within the framework of SOLIBAM to deal with the objectives of the organoleptic task, and is applicable to all partners and products.

Approach

Before implementing the test, a survey on the product concerned should be undertaken in order to understand factors such as market segmentation (food preferences studies…) and consumers' expectations. Institutes such as CTIFL in France can provide this kind of information.

Ranking test

A panel of assessors compares several products simultaneously and ranks them according to the perceived magnitude of a given sensory characteristic (e.g. acidity, fibrousness).
This method has the advantage of being easy to implement. The jury ideally comprises 12 semi-naive assessors (consumers initiated to sensory analyses, see below) according to the ISO 8587 standard[1], although it is possible to highlight significant differences with a smaller number of assessors.

Key Characteristics:

- Products are presented simultaneously

This requires that the whole set of samples to be tested is available at the same time. Some vegetable species show marked difference in precocity (e.g. broccoli), and therefore care should be taken to ensure that samples of the same precocity are compared.

- The assessors can taste as much as they need
- When they answer, assessors cannot put any two products at the same rank, i.e. all ranks assigned must be unique.

It is advised not to exceed 6 samples per session. For sample preparation, refer to the product sheets n°1 to n°5.

How to initiate a jury?

- The first step is to familiarise the jury with the techniques and concepts of sensory analysis: the first sheet will provide the background information necessary.

- The second step is to train the jury using one or two simple "Triangle tests" or "2 among 5" tests (see sheet n°4), on very different products.

[1] ISO 8587:2006 is a standard from International Organisation for Standardisation which describes a method for sensory evaluation with the aim of placing a series of test samples in rank order.

Results and analysis methods

Assessors/varieties	A	B	C	D
1	1	2	3	4
2	2	1	3	4
3	2	1	4	3
4	1	4	2	3
5	1	3	4	2
6	1	2	4	3
7	1	4	3	2
Mean value	1.3	2.4	3.3	3.0

Table 3: Example of ranking test results

For example, Table 3 presents the results of a ranking test: 7 assessors have classified four varieties of tomato (A-D) according to their perceived tenderness.

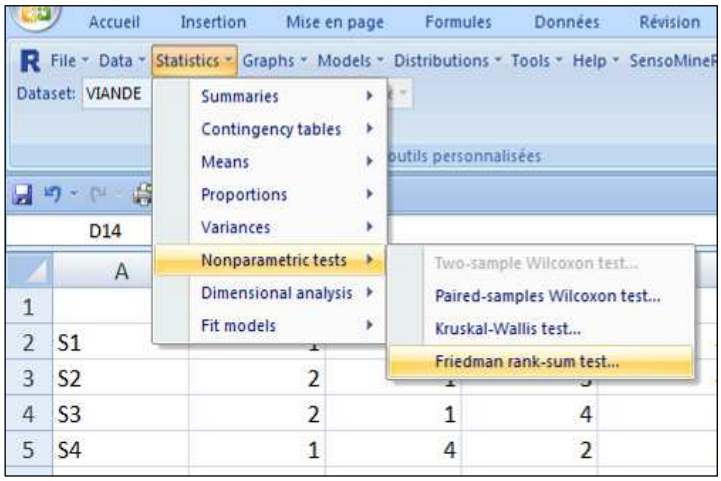
Ranks/varieties	A	B	C	D
1	5	2	0	0
2	2	2	1	2
3	0	1	3	3
4	0	2	3	2

Table 4: Example of a frequency distribution

This table shows frequency of occurrence of each rank assigned by the seven assessors, for each variety. It is derived from Table 3.

- **Step 1 : Rank the varieties according to the intensity of the given sensory characteristics**

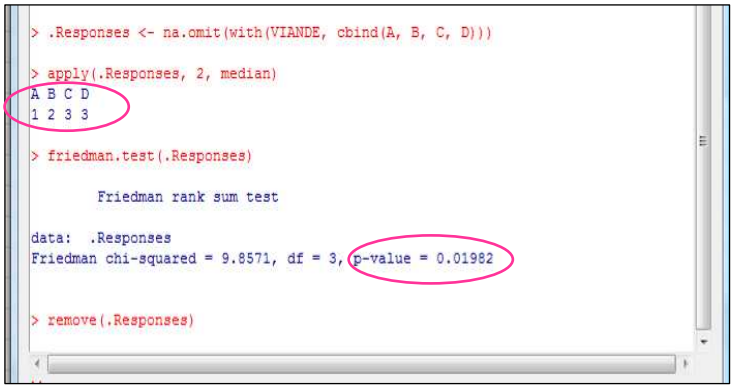
Analysis: Friedman's test



Null hypothesis (H_0): all varieties have exactly the same tenderness (rank means are equal)

Friedman's test (non parametric test on k independent samples) leads to the rejection or acceptance of this hypothesis, based on α value (<0.05).

The four variables (varieties) are selected in the dialogue box. The screenshot on the left gives an example of the results which can be obtained using R software command (see Sheet n°9).

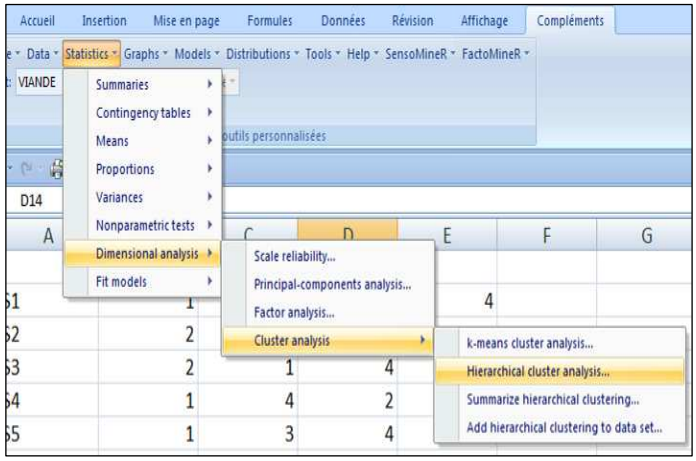


In this example, The Friedman's test results show that the varieties differ with respect to their perceived tenderness (p-value=0.019 < α =0.05): differences between variety means cannot be ascribed to random effects.

It also indicates that the assessor panel performed reasonably well as a whole, with acceptable homogeneity. Variety A is the most tender followed by variety B. It is not possible, however, to differentiate between varieties C and D. This may be due to an insufficient number of assessors (the ISO 85 87 norm recommends 12 assessors). Through this analysis, the varieties that best fit the targeted consumers' expectations can be selected.

- **Step 2 : Check the homogeneity of the panel of assessors**

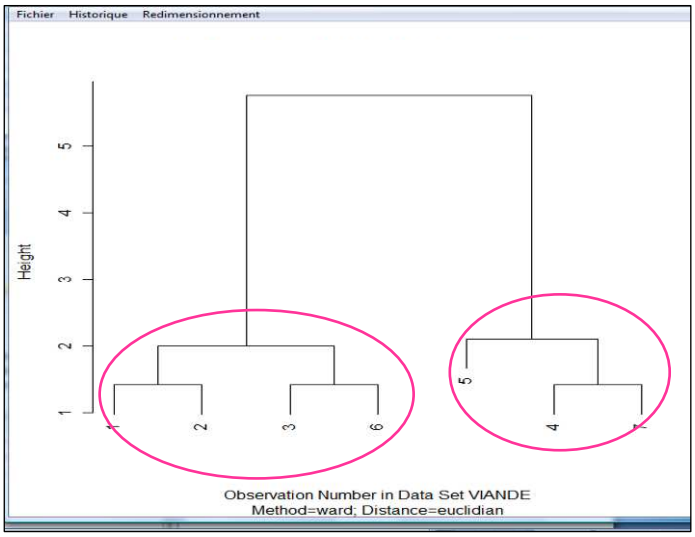
Analysis : Hierarchical Ascendant Classification (HAC)



This test can be used to evaluate the assessors: each highlighted cluster can be considered as the expression of a consensus.
In order to improve the ranking, the HAC technique indicates the presence of outliers: these individuals may not have understood the taste assessment instructions. The test can be repeated following further explanation of the instructions to ascertain if this resolves the problem.

To perform HAC, the **"Ward method"** parameters should be chosen for the classification method and **"Euclidian"** for the distance measure in the dialogue box. The four variables should then be selected.

This results in the following graph:



The HAC indicates two clusters within the assessor panel. Subjects 1, 2, 3 and 6 agree for the rank of the two first and two last samples. Subject 5, 4 and 7 differ with respect to the rank of the last three varieties.

3. Identify consumers preferences : hedonic ranking test

Type of test	Hedonic
Type of data	Preferences
Subject	Naïve (consumer)
Cost/difficulty €/*= cheap/easy, €€/ **= quite expensive/ average difficulty, €€€/***= expensive/ great difficulties.	€ / **

Simple and inexpensive, this test allows the understanding of consumers' preferences.

Table 5: Characteristics of the hedonic ranking test

This test aims to check if SOLIBAMs newly-bred varieties meet consumers' expectations better than controls. **This test is essential in the final stage of the breeding program to validate whether the aim to improve the organoleptic quality has been achieved.**
Although it requires a great number of consumers (60 to 90 per category), it can be carried out in several sessions providing that the tests conditions are exactly the same on each occasion. This test allows end-users expectations to be taken into account and can be carried out in parallel with other tests.

Hedonic evaluation test

The hedonic evaluation test involves asking consumers to rate their preference from 1 (I dislike extremely) to 9 (I like very much) for 3 to 4 sensory attributes specific to the test product.
The overall preference is ascertained at the beginning of the questionnaire in order not to influence the consumer and be closer to typical conditions of consumption. Additional information concerning sex, age and organic consumption frequency are asked at the end of the test in order to characterise the population sample.

Hedonic ranking test

A hedonic ranking test is an alternative to the hedonic evaluation test. It is based on the same principles as described in technical booklet n°2, but more closely resembles a consumer test. The ranking is based on **liking** and it requires a minimum of **60** assessors.

Results and analysis methods

	Sample number	Overall liking	taste	texture
Subject 1	Sample 1			
	Sample 2			
	Sample 3			
	Sample 4			
Subject 2	Sample 1			
	Sample 2			
	Sample 3			
	Sample 4			
Subject 3…	Sample 1			

Table 6: Dataset example, for 3 attributes and 4 samples

Subject	Variety1	Variety 2	…
S1	1	3	
S2	3	3	
…	3	2	

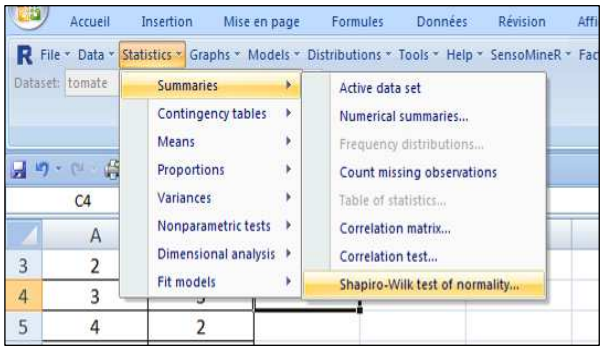
Table 7: Ranks assigned for variety 1 and 2 for the attribute 'texture'

This dataset shows results from a hedonic test carried out on tomatoes in 2006 by a French organic association (BioCIVAM 11).

One of the main objectives of hedonic test is to determine differences of appreciation for a given attribute between a set of samples (in this example differences between variety 1 and 2 for the attribute 'texture').

- **Step 1 : Check the data distribution in order to choose the most appropriate statistical tests**

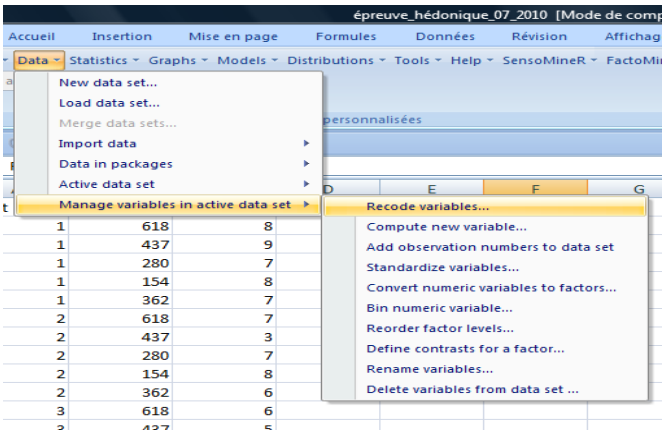
Analysis:
Test Normality with Shapiro-Wilk

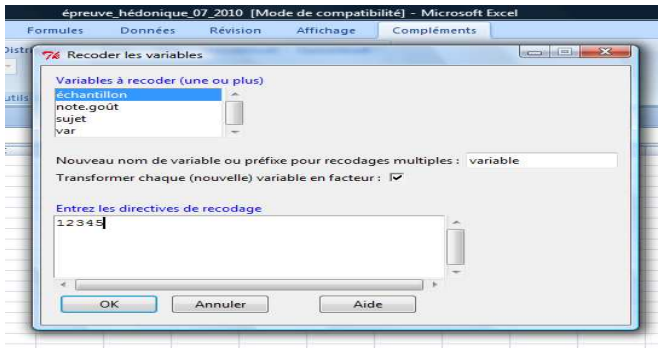


Null hypothesis (H_0): the data follow a normal distribution

The data distribution determines the type of tests that should be used to analyze the data set. If the distribution is Normal, one-way analysis of variance (ANOVA) can be performed, the source of variance being the sample, followed by multiple comparison of mean data values from each assessor.: The aim is to obtain a final ranking based on consumers' preferences.

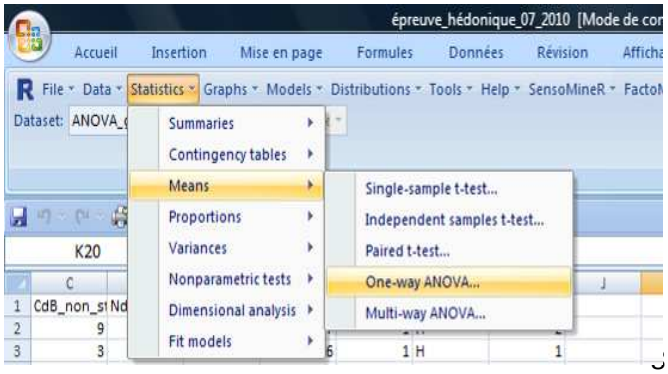
In this example, the p-value is lower than 0.05 (p=2.257e-05) which means that the data is Normally distributed. The one way ANOVA can thus be used to compare the means. Firstly, it is necessary to recode variables: samples initially called 'variables' become a factor named 'var'.





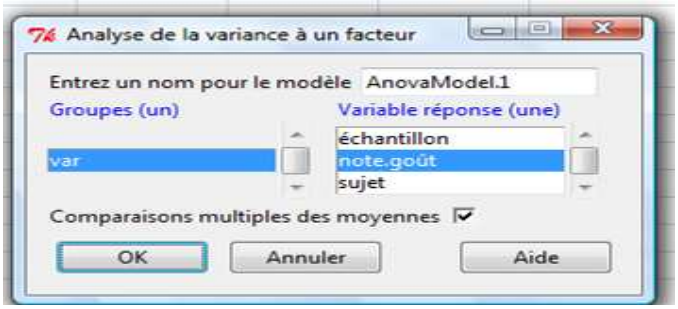
- **Step 2: Assess the consumers' preferences**

Analysis :
One-way ANOVA



Null hypothesis (H_0): Means are equal (there are no differences in preference between varieties)

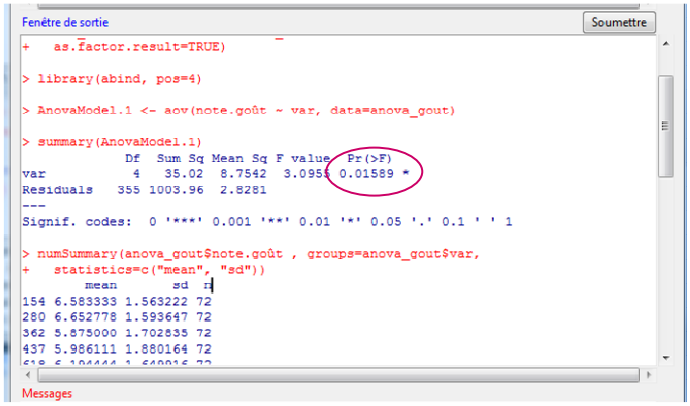
The ‘var’ factors are selected in ‘Groupes’ and the response variables are the assessor’s evaluation scores for the ‘taste’ attribute.



The output window on the right indicates that the test is significant (p-value=0.01589).

Varieties are therefore perceived differently by the different assessors. Finally, examination of the average values indicates that variety 2 is preferred to variety 1.

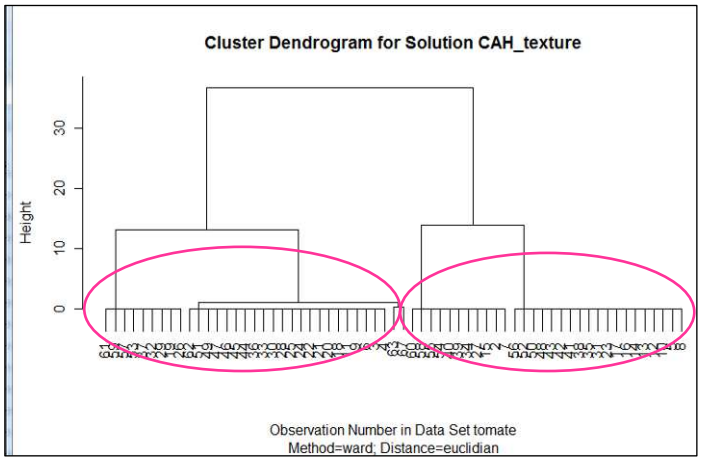
If the data set doesn’t follow a Normal distribution, a Friedman test on the rank should be used to indicate if the varieties are perceived differently by assessors.



- **Step 3: Check the homogeneity of the panel of assessors**

The HAC (see Technical booklet n°2) clearly highlights two clusters of response for the ‘texture’ attribute.

In this example, the two clusters refer to consumers who prefer tender or firm tomatoes.



4. Measure the effect of agronomic practices on organoleptic characteristics : discrimination test

Type of test	Discrimination test
Type of data	Sensory properties
Subject	Semi-naïve
Cost/difficulty €/*= cheap/easy, €€/ **= quite expensive/ average difficulty, €€€/***= expensive/ great difficulties.	€ / **

In SOLIBAM, this test will be used to determine whether breeding and crop management practices influence organoleptic quality of the varieties or not.

Table 8: Characteristics of the "2 among 5" and triangle test

This suggested discrimination test should be performed as the final step in the breeding process as a means of validation. The '2 among 5' test requires few semi-naive assessors (10) and determines whether the impact of agronomic practice can be discerned in the taste of selected varieties. If so, quantitative descriptive analyses of product can be used to characterise those differences (see technical booklet n°5).

Discrimination tests : triangle test and '2 among 5' test.

The two suggested tests are based on the same assumptions. Three (triangle test) or five (2 among 5 test) product samples from two batches are presented. In the case of the triangle test, one comes from the first batch and two come from the second batch (these numbers are two and three respectively for the 2 among 5 test). They are presented simultaneously and the taster has to group the samples he perceives as identical. This is a forced choice process: the taster is compelled to answer, and the doubled (trebled) product should not be the same for each replicate.
The triangle test is more appropriate in the study of flavour attributes, but requires more precautions than the 2 among 5 test concerning batch homogeneity, as it is less statistically robust.
The Sensomine R package contains a function to create a tasting plan. It is necessary to indicate the number of assessors, tested products and replicates per subject.

Results and analysis methods

If all assessors answer at random, the distribution of the variable "number of right answers" follows a binomial distribution. As a result, it is possible to determine a threshold beyond which assessors' answers are unlikely to follow a random pattern, for a given p-value. For the '2 among 5' test, as for the 'triangle test', table 9 indicates the number of correct answers required to detect a significant perceived difference depending on the panel size, for a P-value of 0.05.
For example, in a panel of 12 assessors, 4 correct answers are sufficient to conclude that there is a significant difference, using the '2 among 5' test. However, the 'triangle test' requires 8 correct answers.

Jury size	critical workforce Kc, α=0.05	
	Triangle test	« 2 among 5 » test
5	4	3
6	5	3
7	5	3
8	6	3
9	6	4
10	7	4
11	7	4
12	8	4
13	8	4
14	9	4
15	9	5
16	9	5
17	10	5
18	10	5
19	11	5
20	11	5
21	12	6
22	12	6
23	12	6
24	13	6
25	13	6
26	14	6
27	14	6
28	15	7

Table 9:

Critical values for the discrimination tests

5. Characterize a product : descriptive analysis

Type of test	Descriptive analyses
Type of data	Sensory characteristic
Subject	Qualified
Cost/difficulty €/*= cheap/easy, €€/ **= quite expensive/ average difficulty, €€€/***= expensive/ great difficulties.	€€€ / ***

This analysis aims to delineate the difference between varieties and provides a sensory description of the product.

Table 10: Characteristics of the descriptive analysis

This test is expensive: it requires a qualified panel (ideally 10 assessors), trained for each specific product. The test has to take place in a sensory analysis laboratory with controlled conditions (temperature, humidity, red light). Moreover, it needs to comply with the European standards (ISO 11035, ISO 13299:2003).

Integrated approach

The descriptive analysis is part of the second stage of the sensory analysis and aims to validate the first step (see technical booklet n°2). It is carried out at the end of the breeding process, and only if the discrimination test has highlighted significant differences between products. This test is difficult to conduct as the training process is lengthy, including a basic training of about 5 to 10 hours followed by product specific training (generation and selection of sensory attributes, and guidance on how to use scales) of about thirty hours.

The sensory profile

In this test, the expert panel quantifies the perceived intensity of sensory descriptors on a graded scale. Each descriptor results from a consensus among the experts and the intensity measurements are then visualized on polar graph and histogram.

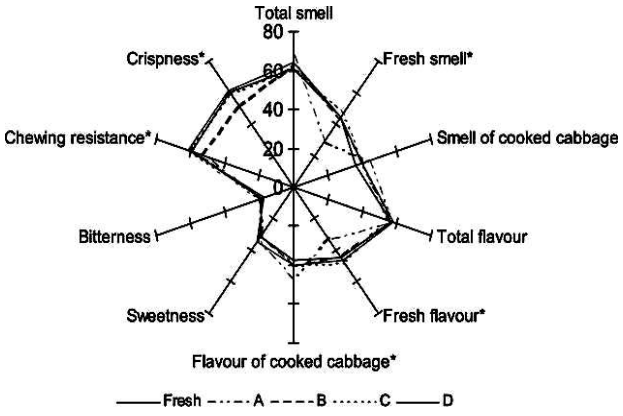


Figure 4: Cabbage organoleptic profile

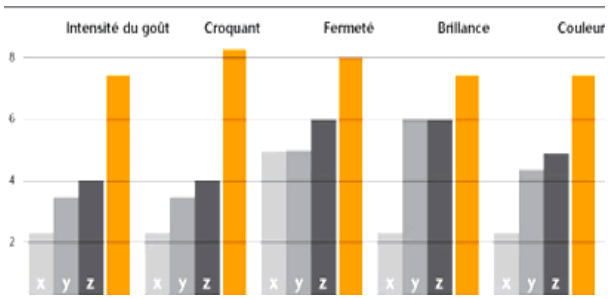


Figure 5: Intensity descriptors for 4 products and 5 sensory attributes

The final purpose of this stage is to precisely define the product's sensory characteristics, in order to develop a reliable sensory description which can be understood by all.

Analysis of Results

The data output is similar to that of the hedonic tests (see technical booklet n°3). Statistical evaluation is performed using multifactorial analysis of variance (sample, assessor, replicate), for each descriptor, to determine if the average ranks of each sample are significantly different or may have occurred at random.

o **Step 1: assess the panel performance, on three criteria**

Table 11 shows the output of the ANOVA. The 3 following points refer to different lines in the table, colour-coded for clarity. The parts in grey are not essential to interpret the results.

- **Discrimination among the sample**: This is the ability to perceive differences between samples. The main objective of a sensory profile is to determine differences in attribute intensity among samples on a specific sensory characteristic. If the product attribute has a significant effect on the variability (P-value < 0.05) it should be included in the assessment list. If the P-value is >0.05 this sensory attribute may be removed.

In the example above, the analysis of variance shows a good discrimination power for 'bitterness' (p<0.0001).

- **Panel agreement**: another important aspect of panel performance is the homogeneity among assessors in the evaluation of a sample, which can be estimated from the interaction "assessors" x "samples". When assessors differ in their scoring path (i.e. different assessors' responses for the same sample and descriptor) the probability associated with the interaction effect "assessors" x "sample" is significant (<0.05).

In table 11, we can see there is a problem of homogeneity in the assessment for the given example.

In this case, a Friedman test (table 12) will give indications as to the reasons for this lack of homogeneity: if it is significant, the panel needs further training on how to use the notation scale. If it is not significant, then the assessors are in complete disagreement and training must restart from the beginning.

In this example, the Friedman test confirms heterogeneity in the scale used.

- **Assessment replicability**:
  Precision is an important aspect of the performance of the assessors, and relates to the variability of the evaluation scores given to replicates of the same sample. The probability for the interaction "sample" x "replicates" must be greater than 0.05 to conclude that there is good replicability.

	df	SME	F (Fisher)	Probability
Product factor(P)	25	65.61	27.56	**<0.0001**
Assessors' factor (A)	12	64.58	27.12	<0.0001
Replicates' factor (R)	2	9.07	3.81	0.0227
Interaction S*A	300	4.69	197	**<0.0001**
Interaction S*R	50	1.65	0.69	**0.946**

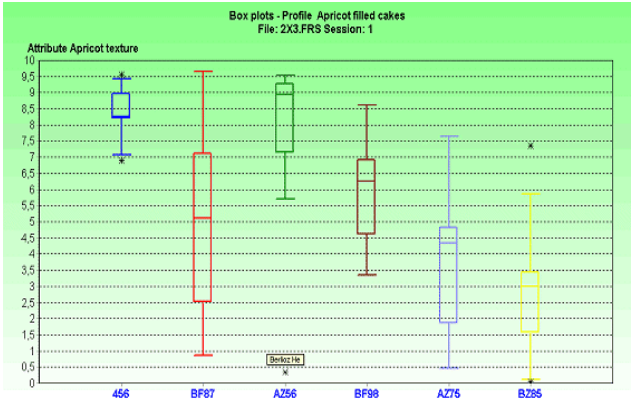
Table 11: ANOVA for the attribute 'bitterness'

$\Sigma(\Sigma(\text{ranks}))^2$	8110526,5
F (Fisher)	395,9097
p-value	**<0.0001**

Table 12: Friedman Test for the attribute 'bitterness'

o **Step 2: Summarise**

Once performance has been evaluated, average values can be compared for each descriptor scored. A multivariate analysis (HAC, or PCA) will allow the product results to be displayed relative to each other. This figure shows an example of sensory profile results (box-plot).



6. Raw products : TOMATO (example)

Tomato barometer (France)

Consumers' preferences are divided into two axes. The **first axis** concerns **flavour, aroma and juiciness** and shows consumer segmentation on **acidity**. The **second axis** deals with **texture** and shows two different preference profiles for those tasters who preferred **softness** and those who preferred **firmness and crispiness**. Although tomato consumption in France has stabilized at 12 kg per habitant, the taste satisfaction has decreased since 1998, with 1/3 of consumers unsatisfied with "tasteless tomatoes" (Baros, Journée ctifl-Inra_4/02/2010).

Equipment needed

For the whole jury

- A chopping board and a knife
- Indelible odourless felt tip to identify the samples

For each workshop

- As many plates as varieties
- 2 tomatoes per variety, maximum 8 varieties

For each taster

- A fork
- A napkin
- A glass of water, unsalted crackers or bread to cleanse the palate between samples
- The questionnaire linked to the test

Preparation mode

Tomatoes are tasted four days after harvest; in the meantime they are stored for 48 h at 12 °C and then for 48 h at room temperature (25/30 °c).
Fruits showing irregularity and/or defects such as green colour on the upper side are excluded from the taste experiment. If the quantity is limited, however, these irregular fruits may be used providing that the defects have been removed.
After verification of the maturity and homogeneity, fruits are cut into homogeneous pieces including skin.

List of sensory attributes (non-exhaustive)

o **Odour**: Tomato aroma	o **Appearance**: Colour Grooved skin surface
o **Taste** : Salty taste Sweet taste Overall acidity	o **Texture** : Skin consistency Mealiness Softness Crispness Juiciness Firmness

Table 13: Sensory descriptors for tomato

Ranking test

o **Sample presentation**

The samples (halved tomatoes) are presented simultaneously and each is allocated a number (blind trial). Depending of the panel size, tasters are organised into groups composed of 3/4 assessors per workplace and silence is maintained at all times.

Crédit photo : Frédéric Rey

Crédit photo : Frédéric Rey

o **Questionnaire**

The assessor must taste samples and then rank them according to the perceived intensity of a given sensory characteristic (rank 1 is the most intense). It is important that assessors do not talk to each other to prevent biasing the results. It is found that the greater the number of attributes tested is, the more difficult it is to detect significant differences. In order to minimize this problem, those descriptors that can be instrumentally measured are omitted. The questionnaire proposed below (Fig 6) is an example; the attribute list in Table 13 can be used to select relevant descriptors depending on the information required.

Please taste the samples, and rank them according to the perceived intensity of the descriptors "**sweet taste**" and "**firmness**". Indicate this by entering the sample number below the appropriate rank number.

Rank number	**1**	**2**	**3**	**4**	**5**	**6**
Sweet taste						
Firmness						

Figure 6: Questionnaire for tomato (ranking test)

Hedonic ranking test

o **Sample presentation**

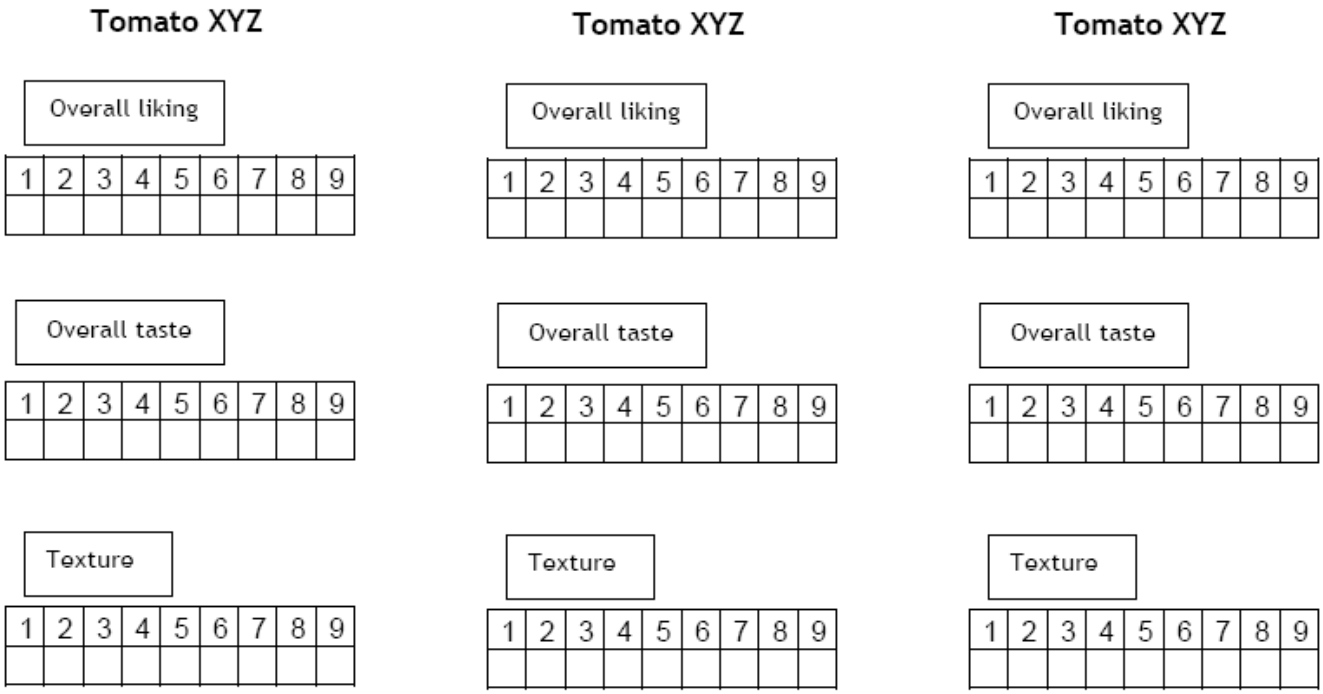
Samples are presented one by one in a precise order so that rank effect[2] is limited. SensomineR packages include special functions which take this parameter into account. Varieties are arranged on a plate (one variety per plate) identified by the sample number.

o **Questionnaire**

The questionnaire response scale ranges from 1 ("I dislike extremely") to 9("I like extremely") for the overall preference of the sensory attributes.

[2] The rank effect is the evaluation bias attributed to the product presentation order. For example, a subject can overestimate the cocoa aroma of the chocolate sample presented first.

As part of the European project SOLIBAM, we are testing tomatoes to gain a better understanding of consumer expectations. We ask that you taste … different tomatoes, and give us your opinion on their gustatory quality. Please pay close attention to the order of the samples, and fill the scale from 1 (I dislike extremely) to 9 (I like very much).



Please specify:

Gender: ☐ Male ☐ Female

Age: ☐ Less than 30 yrs ☐ 30 to 40 yrs ☐ 40 to 50 yrs ☐ Over 50 yrs

I buy organic products: ☐ At least once a week ☐ At least once a month ☐ Never

Figure 7: Questionnaire for tomato (hedonic ranking test)

Discriminative test ‘ 2 among 5’ and/or triangular test

- **Sample presentation (‘2 among 5’)**

Five product samples are presented to the assessor from two batches. Two samples come from one of the two batches and three samples come from the other. They are presented simultaneously. The trebled product should not be the same for each replicate. The SensomineR package contains a function for the creation of a tasting plan and requires the number of assessors to be entered, along with the tested products and the number of replicates per subject.

- **Questionnaire**

The taster has to group samples he perceives as identical. It is a forced choice process as the subject is compelled to answer. Two tests based on the same principle are proposed: the triangle test is more appropriate for the assessment of flavour, however, it needs to be used with more precaution due to issues concerning the batch homogeneity (it is less statistically robust than the ‘two among five’ test).

‘ 2 among 5’ test

Among these samples, two come from one batch and three from another one. Please group the samples you perceive to be identical, and indicate these by marking with a circle.

Sample 1	Sample 2	Sample 3	Sample 4	Sample 5

Figure 8: Questionnaire for tomato (‘2 among 5’ test)

Triangle test

Amongst these three samples, please identify which one is different from the other two. Write the number of the different sample in the box on the right hand side.

Sample 1	Sample 2	Sample 3	?

Figure 9: Questionnaire for tomato (triangle test)

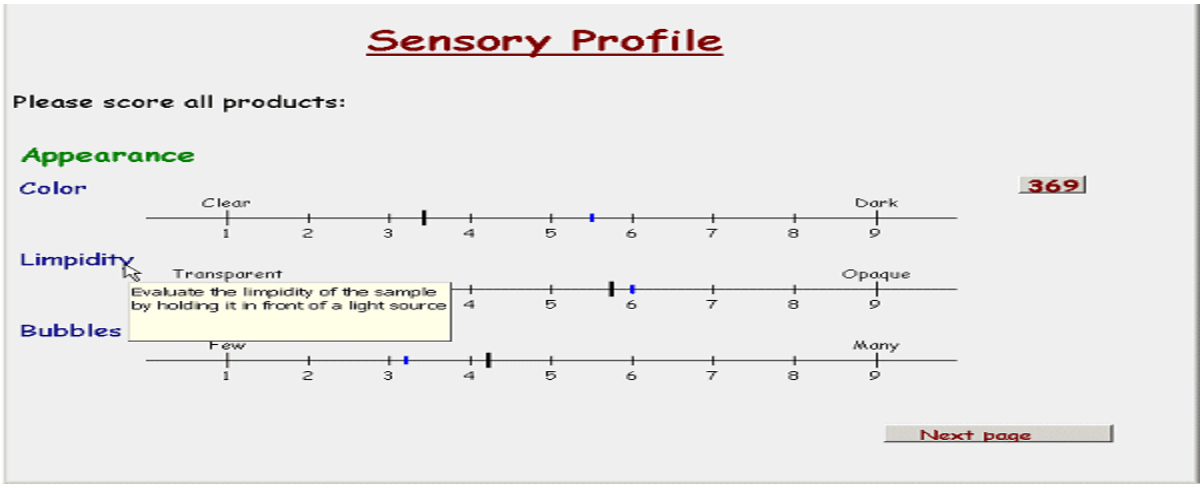
Descriptive test

o **Sample presentation**

This test is designed around a well-balanced experimental plan in order to limit rank effects. The SensomineR package contains a set of functions to create such plans and evaluate the panel performance.

o **The questionnaire**

In this test, a panel of experts has to quantify the intensity of a set of given attributes on a graded scale. The final scores awarded must be the result of a consensus between experts.



7. Cooked products : BROCCOLI (example)

Broccoli barometer (France)

After some production issues during the five years from 1990-1995), the French broccoli market continues to increase (info-Ctifl/September 2003). Nevertheless, the penetration rate remains quite low at 27%, according to the Sécodip panel for all households. Clearly it is less commonly consumed in France than many other vegetables, perhaps because less is known about its origin, provenance and history, and there is a perception of broccoli as a novelty product that is 'fast and easy to cook'. It is also considered by many to be an 'acquired taste' and this is a restraint to purchase. Other constraints include a lack of product freshness and too high a price (A&D, 04/2005-n°83).

Equipment needed

For the sample preparation

- A steamer
- A chopping board and a knife
- A balance
- Salt [3]
- Water (for cooking)
- Indelible odourless felt tip to identify the samples

Per workshop

- A small pot for each sample
- 35g of broccoli per variety, no more than 6 varieties

Per subject

- A fork
- A napkin
- A glass of water, unsalted crackers or bread to cleanse the palate between tasting
- The associated questionnaire

List of sensory attributes (non-exhaustive)

o **Odour :** Weed-like Cooked cabbage Iodized Nutty Earthy	o **Texture :** Spongy Tender Firm Crisp
o **Taste :** Cooked cabbage Sweet Bitterness	o **Appearance :** Colour (green, brown) Compactness

Table 14 : Sensory descriptors for broccoli

Preparation method

Cooked products must be prepared in a standardised, replicable manner. The cooking method is important to ensure the samples are comparable between tasting sessions and in some cases cooking tests are necessary to determine the most appropriate method for the tasted product. In addition to broccoli, cabbage and bean cooking methods are detailed in this sheet:

o **Broccoli**

[3] It is important to use the same salt for all the samples in order to standardise the tasting.

Broccoli heads are detached and any damaged parts are removed. The heads are then cut into uniform pieces. Depending on the experimental aims, the broccoli may be tasted raw or steamed. If it is the latter, the heads are steamed for about 15 minutes, taking care to check during cooking.

o **Cabbage**

Undamaged leaves are taken evenly from the centre and the periphery of the plant and cut into pieces 2 cm wide and 15 cm long. Depending of the aim of the experiment, cabbage can be tasted raw or cooked. If cooked, the leaves are steamed for about 25 min.

o **Dried bean**

For a jury of 12 assessors, 250 g of dried beans are dipped into fresh water (3 times their volume), and boiled. They are then rinsed and impurities or damaged beans are removed. The beans are then cooked in boiled water in a pressure cooker for a period of 50 min after the valve starts rotating. The cooking time varies with variety, and it is necessary to test the cooking time for each sample. After that, the beans are rinsed in a colander and salted with a standardised amount of salt to approximate the typical conditions of consumption (5g salt for 500g beans).

For cooked products, samples must still be hot when served to the tasters (40-70°C). Salt addition (1g per 100g of product) is recommended as the experience is then closer to typical conditions of consumption, making the tasting more pleasant.
For the boiled products, it is advised to avoid tap water, which can vary in taste. Spring water should be used for cooking if possible.

Ranking test

o **Sample presentation**

Samples weighing 35 g each are placed in pots marked with the sample number, and are presented to the tasters simultaneously. Depending of the panel size, tasters are organised into groups composed of 3/4 assessors per workplace and silence is maintained at all times.


Credit : Frédéric Rey


Credit : Frédéric Rey

o **Questionnaire**

The assessor must taste samples and then rank them according to the perceived intensity of a given sensory characteristic (rank 1 is the most intense). It is important that assessors do not talk to each other to prevent biasing the results. It is found that the greater the number of attributes tested is, the more difficult it is to detect significant differences. In order to minimize this problem, those descriptors that can be instrumentally measured are omitted. The questionnaire proposed below (Fig 10) is an example; the attribute list (Table 14) can be used to select relevant descriptors depending on the information required.

Please taste the samples and rank them according to the perceived intensity of the descriptors **"bitterness"** and **"tender texture"**. Indicate this by entering the sample number below the appropriate rank number.

Rank n°	**1**	**2**	**3**	**4**	**5**	**6**
Bitterness						
Tender texture						

Figure 10: Questionnaire for broccoli (ranking test)

Hedonic ranking test

o **Sample presentation**

Samples are presented one by one in a precise order so that rank effect[4] is limited. SensomineR packages include special functions which take this parameter into account. Varieties are presented in small pots (one variety per pot) identified by the sample number.

o **Questionnaire**

The questionnaire response scale ranges from 1 ("I dislike extremely") to 9("I like extremely") for the overall preference of the sensory attributes.

As part of the European project SOLIBAM, we are testing broccolis to gain a better understanding of consumer expectations. We ask that you taste … different broccolis, and give us your opinion on their gustatory quality. Please pay close attention to the order of the samples, and **fill the scale from 1 (I dislike extremely) to 9 (I like very much).**

Broccoli XYZ	Broccoli XYZ	Broccoli XYZ	Broccoli XYZ
Overall appreciation 1 2 3 4 5 6 7 8 9	Overall appreciation 1 2 3 4 5 6 7 8 9	Overall appreciation 1 2 3 4 5 6 7 8 9	Overall appreciation 1 2 3 4 5 6 7 8 9
Overall taste 1 2 3 4 5 6 7 8 9	Overall taste 1 2 3 4 5 6 7 8 9	Overall taste 1 2 3 4 5 6 7 8 9	Overall taste 1 2 3 4 5 6 7 8 9
Sweet taste 1 2 3 4 5 6 7 8 9	Sweet taste 1 2 3 4 5 6 7 8 9	Sweet taste 1 2 3 4 5 6 7 8 9	Sweet taste 1 2 3 4 5 6 7 8 9
bitterness 1 2 3 4 5 6 7 8 9	bitterness 1 2 3 4 5 6 7 8 9	bitterness 1 2 3 4 5 6 7 8 9	bitterness 1 2 3 4 5 6 7 8 9
tenderness 1 2 3 4 5 6 7 8 9	tenderness 1 2 3 4 5 6 7 8 9	tenderness 1 2 3 4 5 6 7 8 9	tenderness 1 2 3 4 5 6 7 8 9

Please specify:

Gender: ☐ Male ☐ Female

Age: ☐ Less than 30 yrs ☐ 30 to 40 yrs ☐ 40 to 50 yrs ☐ Over 50 yrs

I buy organic products: ☐ At least once a week ☐ At least once a month ☐ Never

Figure 11: Questionnaire for broccoli (hedonic ranking test)

[4] The rank effect is the evaluation bias attributed to the product presentation order. For example, a subject can overestimate the cocoa aroma of the chocolate sample presented first.

Discriminative test ‘ 2 among 5’ and/or triangular test

o **Sample presentation (« 2 among 5 »)**

Five product samples are presented to the assessor from two batches. Two samples come from one of the two batches and three samples come from the other. They are presented simultaneously. The trebled product should not be the same for each replicate. The SensomineR package contains a function for the creation of a tasting plan and requires the number of subjects to be entered, along with the tested products and the number of replicates per subject.

Credit : Frédéric Rey

o **Questionnaire**

The taster has to group samples he perceives as identical. It is a forced choice process as the subject is compelled to answer. Two tests based on the same principle are proposed: the triangle test is more appropriate for the assessment of flavour, however, it needs to be used with more precaution due to issues concerning the batch homogeneity (it is less statistically robust than the ‘two among five’ test).

« 2 among 5 » test

Among these samples, two come from one batch and three from another one. Please group the samples you perceive to be identical, and indicate these by marking with a circle.

Sample 1	Sample 2	Sample 3	Sample 4	Sample 5

Figure 12: Questionnaire for broccoli (“2 among 5” test)

Triangle test

Amongst these three samples, please identify which one is different from the other two. Write the number of the different sample in the box on the right hand side.

Sample 1	Sample 2	Sample 3	?

Figure13: Questionnaire for broccoli (triangle test)

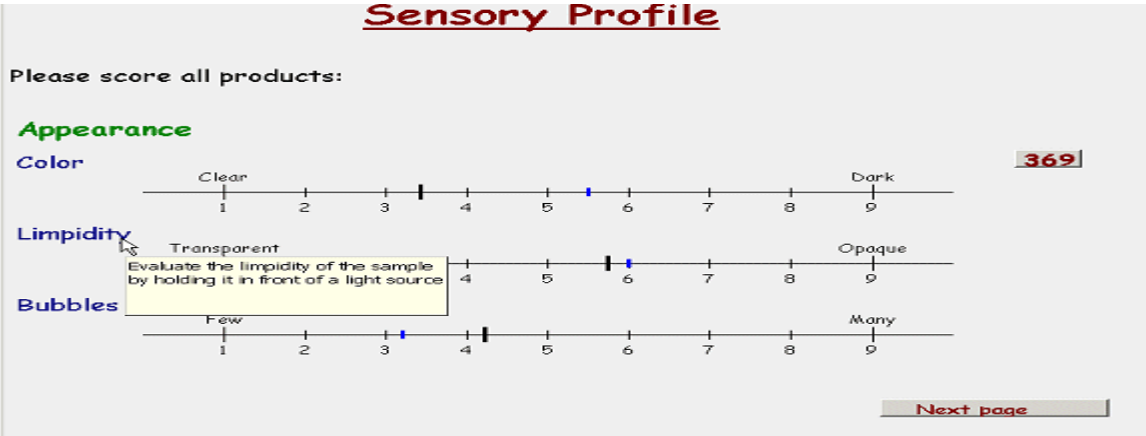
Descriptive test

o **Sample presentation**

This test goes on a well balanced experimentation plan in order to limit the rank effect. The SensomineR package contains a set of functions to create experimentation plans and evaluate the panel performance.

o **The questionnaire**

In this test, a panel of experts has to quantify the intensity of a set of given attributes on a graduate scale. Those descriptors are the result of a consensus between experts. They have to be defined during the specific panel training.



Credit : Frédéric Rey

Credit : Frédéric Rey

8. Processed products : BREAD (example)

Organic bread barometer (France)

Bread is seen as an emblematic food of French culture. When it is organically produced, it is thought to be a *natural product*, *nourishing and healthy (*Allessandrin et al., 2007[5]*)* and is produced traditionally, as opposed to industrial processing techniques embodied by the white baguette sold at low prices in supermarkets. On the other hand, criticisms of organically produced bread include concerns regarding the taste (too sour), the shelf-life and price (too high).

Equipment needed

Product preparation

- A chopping board and a knife
- Indelible felt tip to identify the samples
- A whole loaf of bread per sample.

Per sample

- A plate per sample
- Two pieces per sample

Per subject

- A napkin
- A glass of water
- The questionnaire

Preparation mode

Ideally, all tested breads are cooked in the same way at the tasting site. It is important to reduce sources of variation and centralising production such as this is a practical way to standardise preparation. List of sensory attributes

o **Appearance :** *Crust* Browning Shininess Thickness *Crumb* Colour Alveolus regularity Alveolus average size Alveolus density/crumb airing	o **Taste :** Salty Sweet Bitterness Acidity
o **Aroma:** Bread Roasting	o **Texture :** Crispiness Tenderness Elasticity

Table 15: Sensory descriptors for bread

Ranking test

o **Sample presentation**

[5] Allessandrin, A., Desmont, M.H. (2007) Qualités des blés biologiques et qualités nutritionnelles et organoleptique des pains biologiques, Final report of the project « Organic Bread » (ACTA, ITAB, INRA, Intercéréales).

Depending on the sensory attributes that have been selected, bread will be presented either as a whole loaf (global appearance) or sliced (texture, alveolus description, taste criteria).

o **Questionnaire**

The assessor must taste samples and then rank them according to the perceived intensity of a given sensory characteristic (rank 1 is the most intense). It is important that assessors do not talk to each other to prevent biasing the results. It is found that the greater the number of attributes tested is, the more difficult it is to detect significant differences. In order to minimize this problem, those descriptors that can be instrumentally measured are omitted. The questionnaire proposed below (Fig 14) is an example; the attribute list in Table 15 can be used to select relevant descriptors depending on the information required.

Please taste the sample, and rank them according to the perceived intensity of the descriptors "**Cumb airing**" and "**Acidity**". To do so, fill the number associated to the sample in the right chosen rank.

Rank n°	1	2	3	4	5	6
Cumb airing						
acidity						

Figure 14: Questionnaire for bread (ranking test)

Hedonic ranking test

o **Sample presentation**

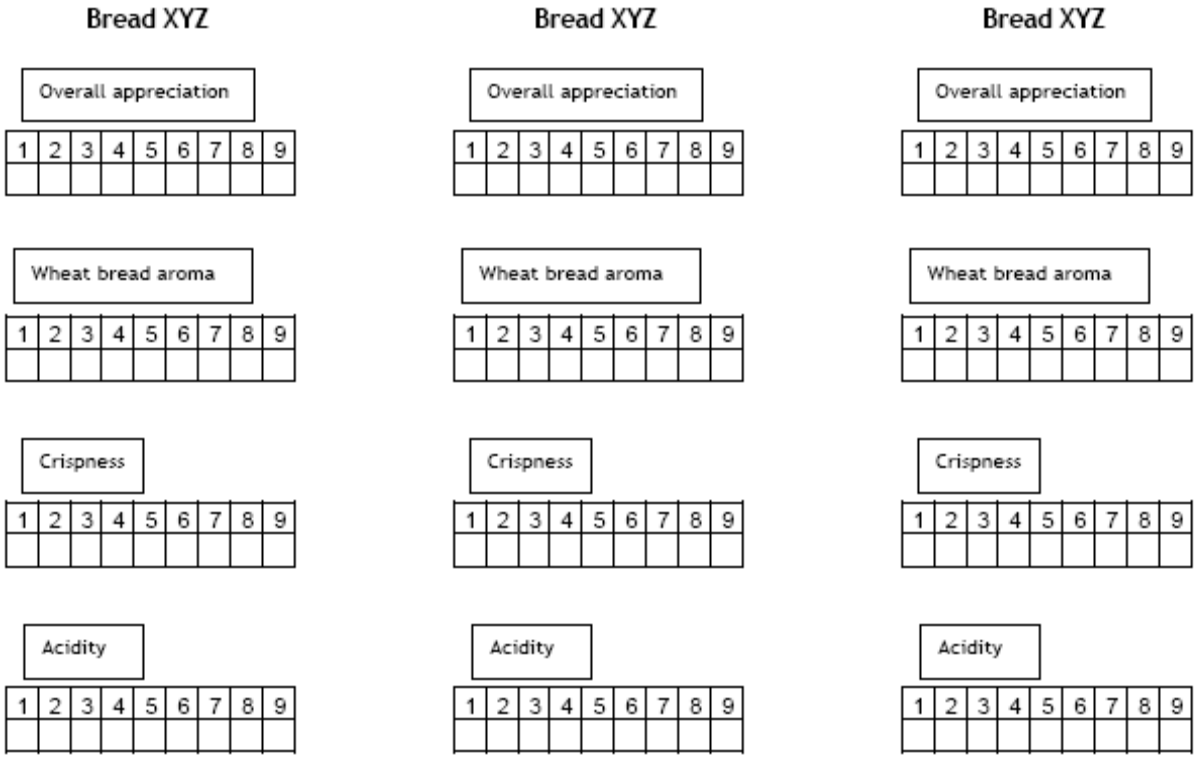
Samples are presented one by one in a precise tasting rank so that rank effect[6] is limited (well balanced experimentation plan). SensomineR packages show special functions which take this parameter into account.

o **Questionnaire**

The questionnaire response scale ranges from 1 ("I dislike extremely") to 9("I like extremely") for the overall preference of the sensory attributes.

[6] The rank effect is the evaluation bias attributed to the product presentation rank. For example, subject can overestimate the cocoa aroma for the chocolate at the first position.

As part of the European project SOLIBAM, we are testing breads to gain a better understanding of consumer expectations. We ask that you taste … different breads, and give us your opinion on their gustatory quality. Please pay close attention to the order of the samples, and **fill the scale from 1 (I dislike extremely) to 9 (I like very much).**



Please specify:

Gender: ☐ Male ☐ Female

Age: ☐ Less than 30 yrs ☐ 30 to 40 yrs ☐ 40 to 50 yrs ☐ Over 50 yrs

I buy organic products: ☐ At least once a week ☐ At least once a month ☐ Never

Figure15: Questionnaire for bread (hedonic ranking test)

Discriminative test ‘ 2 among 5’ and/or triangular test

o **Sample presentation**

All the samples must have the same shape (round, baguette,...). Five product samples (a sample = a slice) are presented to the assessor from two batches. Two samples come from one of the two batches and three samples come from the other. They are presented simultaneously. The trebled product should not be the same for each replicate. The SensomineR package contains a function for the creation of a tasting plan and requires the number of subjects to be entered, along with the tested products and the number of replicates per subject.

o **Questionnaire**

The taster has to group samples he perceives as identical. It is a forced choice process as the subject is compelled to answer. Two tests based on the same principle are proposed: the triangle test is more appropriate for the assessment of flavour, however, it needs to be used with more precaution due to issues concerning the batch homogeneity (it is less statistically robust than the ‘two among five’ test).

« 2 among 5 » test

Among these samples, two come from one batch and three from another one. Please group the samples you perceive to be identical, and indicate these by marking with a circle.

Sample 1	Sample 2	Sample 3	Sample 4	Sample 5

Figure 16: Questionnaire for bread ("2 among 5" test)

Triangle test

Amongst these three samples, please identify which one is different from the other two. Write the number of the different sample in the box on the right hand side.

Sample 1	Sample 2	Sample 3	?

Figure 17: Questionnaire for bread (triangle test)

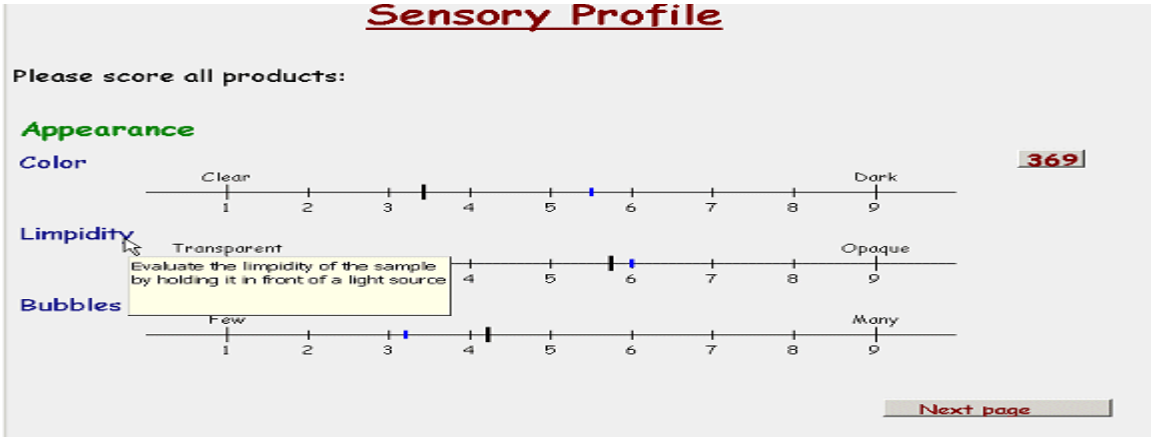
Descriptive test

o **Sample presentation**

This test is designed around a well-balanced experimental plan in order to limit rank effects. The SensomineR package contains a set of functions to create such plans and evaluate the panel performance.

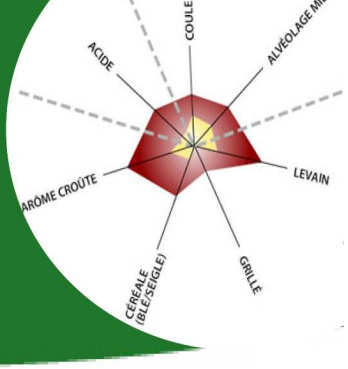
o **The questionnaire**

In this test, a panel of experts has to quantify the intensity of a set of given attributes on a graded scale. The final scores awarded must be the result of a consensus between experts. They have to be defined during the specific panel training.



Credit : European Commission

9. Installation and use of R software : SensomineR and FactomineR packages



SensomineR / FactomineR : presentation

The software suggested for analysis is a programme named “R”. This open source software is downloadable for free and can be modified. Until now, specific software for to sensory analyses has been expensive and only available for industry. The additional packages within R adapted to our study are **SensomineR**, dedicated to sensory analyses, and **FactomineR** which focuses on multivariate analyses. It provides simple, clear results represented as graphs and statistical summaries.

How to download / install the software

- Go to the web page http://sensominer.free.fr/Excel.html, download and install the **RandFriends** software on the suggested link. Follow the instructions to install R 2.11.1.

Once everything has been installed, a new icon will appear on your desktop. This icon integrates RExcel with the Rcmdr menu in Excel.

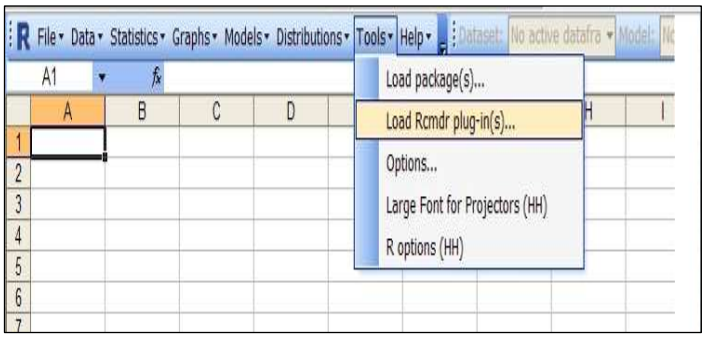
- Open RExcel. R is opened automatically at the same time. A new tab "RExcel" (or "Complements" depending on your Microsoft Office version) has been added. It provides the Rcmdr menu for easy use.

For other versions of Office, the packages are automatically installed. If you need more information about the installation, please see http://sensominer.free.fr.

How to install the SensomineR / FactomineR packages

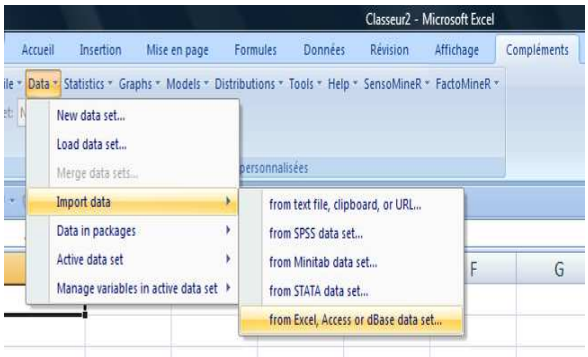
If SensomineR and FactomineR are not available on your new tab “RExcel”, it means you have to install the 2 packages. In this tab, choose tools, and then "Load Rcmdr plug-in(s)...". Select the 2 packages in the list. If they do not appear on the menu, restart RExcel.

SensomineR and FactomineR menus are now available.

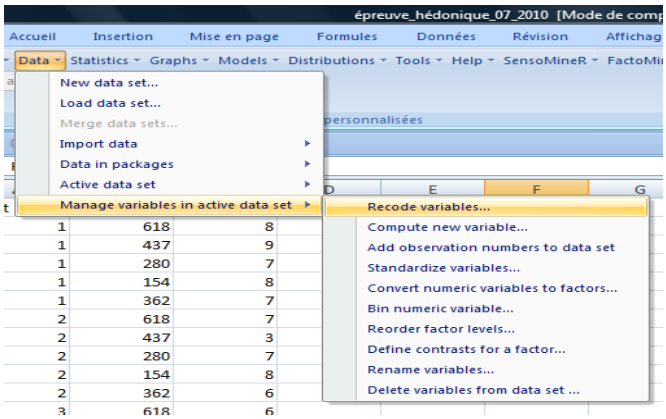


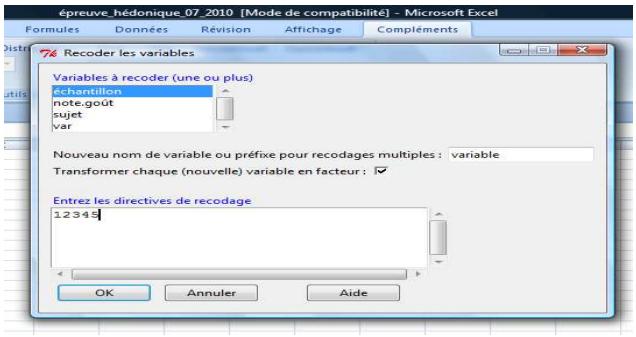
Data import

Once the software is installed, you have to import data. To do so, choose **“data”** in the tab **Rexcel** (or **Complements**). Then choose **“Import data, from Excel, Access or DBase data set… ”**, and browse to locate the data set you want to analyse.



You then need to recode variables in order to indicate which are factors. As the factor is the variety, each sample is recoded as a factor.





The software is now ready for analysis of results

Schedule : Glossary

Sensory Analysis

o **Vocabulary specific to the analysis**

Aroma: the sensation perceived by retro-olfaction when eating.

Customer acceptance: the degree of compliance with customer quality's expectations of a product.

Expert (subject): a qualified subject who has excellent sensory acuity and who has been trained to use sensory evaluation tests. Their reliability has been assured.

Flavour: Flavour is the sensory impression of a food or other substance, and is determined mainly by the chemical senses of taste and smell. The "trigeminal senses", which detect chemical irritants in the mouth and throat as well as temperature and texture, are also very important to the overall flavour perception.

Hedonic: of, related to, or marked by pleasure.

Naive (subject): Untrained subject **(**synonymous with the consumer).

Organoleptic characteristic: this term includes all of the product properties that can be perceived by sensory organs.

Retro-olfaction: the aroma perceived by the retro-nasal function (retro-olfaction - passing internally from the mouth cavity to the nasal passages).We differentiate between this and the odour perceived by direct nasal function (i.e. by breathing in through the nostrils).

Semi naive (subject): person who has already taken part in a discriminatory sensory test. One previous testing experience is the minimum training required for a discrimination test.

Sensory attribute: Sensory properties specific to a product**.**

Trigeminal perception: this term includes sensations such as irritation, tickling, burning, cooling etc

o **Description attributes sensory**

Tomato

- Firmness : resistance to mastication
- Juiciness: juice perception when chewing the product
- Skin persistence: this characteristic is linked to the skin thickness and indicates the persistence of the skin in the mouth after ingestion
- Crispiness: sound perceived when eating the product

Broccoli

Compactness: compact nature of the broccoli heads.

Cabbage

- Fibrousness: fibrous properties of the leaves
- Fruitiness: sweet taste with a fruity note
- Long finish: persistence of sensation in the mouth after eating
- Tickling sensation: this term includes the flash effect after 'hot' or 'spicy' food (e.g. wasabi) and the after taste effect.

Bean

- Taste: defines the taste of which the bean is reminiscent (e.g. sweet chestnuts, nutty, milk concentrate)
- Viscosity: relating to the thickness and stickiness of a substance

Bread: Crust

- Shininess: reflectiveness of surface
- Tanning: intensity of the crust colour
- Thickness: Depth of the peripheral part which has a more tanned colour compared to the crumb.

Bread: Crumb

- Alveolus (alveoli pl): air sac(s) within the bread
- Alveolus regularity: uniformity of the alveolus size
- Alveolus average size: average diameter of alveoli
- Alveolus density: number of visible alveoli per unit area
- Colour: sensory perception linked to the eyes reception of a light beam reflected by a body
- Elasticity: extent to which bread recovers its initial shape after light and even pressure
- Tenderness: softness upon touch.

Statistical analysis

Parametric test: parametric tests make assumptions about the spread of data and are used when it follows a Normal distribution, which can be described with parameters such as average and standard deviation.

Non parametric test: non-parametric tests make few or no assumptions about the distribution of data and are used when comparing discrete variables or when the sample size (n) of a data set is low (e.g. less than 60). In this case, the average and standard deviation parameters cannot be used to describe the data and its distribution is not Normal.

Hypothesis test: A hypothesis test is carried out to determine whether an asserted hypothesis can be accepted or rejected based on statistical probability. The **null hypothesis** states that there is no difference between the two groups under consideration (i.e. parameters such as average and standard deviation are equal).

Variance analyses: Analysis of variance (ANOVA) is a statistical test which compares means from two or more sets of data. It indicates how much of the observed variation is due to 'true' population differences and how much is due to random effects. ANOVA tests the null hypothesis that all samples come from the same population. There are two assumptions that must be fulfilled: observations must be independent both within and between samples and the data must be Normally distributed.

Principal Component Analyses (PCA): PCA is also known as factor analysis. It is a multivariate statistical test which weights variables in order to maximise the differences between individuals. The assumption behind this test is that individuals must have two or more observations assigned to them and should be measured on a continuous scale.

Hierarchical Ascendant Classification (HAC): this is a multivariate technique which consists of progressively grouping observation or assessors according to their similarity, as measured by a similarity index. Aggregation criteria must be defined for use as a similarity measurement.

Schedule : Barometers and descriptors

For each product concerned in this aspect of SOLIBAM (tomato, cabbage, broccoli, beans and bread), a bibliographic study has been carried out concerning consumer preferences. This will help to choose the sensory attributes upon which to focus in taste tests. The following section includes all the information found for all species concerned and expands upon the examples contained within the earlier sections of this guide.

Cabbage

o **Barometer**

The main criticism about cabbage from consumers is the smell released when cooking. In a CTIFL study carried out in 1997, 34% of consumers questioned claimed that they would eat more cabbage if it produced fewer odours during cooking. Therefore, consumer preferences should be explained by the descriptor such as "cabbage odour", "potato flavoured", "rancid aroma", "pungent flavour". Most consumers preferred cabbage with high notes of cabbage and potato flavours and with few pungent and rancid notes (A&D, 01/2004-n°78).

o **List of sensory attributes**

Raw cabbage

o **Odour:** Onion Sewage/gas Cresson/nasturtium/mustard Earthy/wasabi Green apples/blackberry Fresh nuts/pumpkin Marine odour Cucumber Fruity/citrus Weed-like	o **Taste :** Sweet Bitter Pungent o **Texture :** Juiciness Crispness Fibrousness Long finish Firmness Elasticity

Cooked cabbage

o **Odour:** Cabbage Potatoes Rancid	o **Taste :** Sweet Bitter Pungent o **Texture :** Fibrousness Tender

Table 16: Sensory descriptors for cabbage (raw / cooked)

When evaluating the protocol, three composite descriptors were identified as essential to differentiate and characterise the raw and cooked samples. These are 'fruity **taste, crispness, pungent sensation'** and **'cabbage taste, fruity taste, crispness'.**

Bean

o **Barometer**

The French production of beans is about 110 000 tonnes and this is located mainly in Bretagne, Nord Pas de Calais, Centre and Picardie. Consumption of pulses clearly fell between 1920 (7.3 kg/person/years) and

1985 (1.4 kg/person/years). It has, however, stabilized and increased slightly since then due to industrial preparation and the development of canned foods (1996, 1.6 kg/person/years).

- **List of sensory attributes**

Aroma/taste : Astringent Sweet chestnuts Artichoke Vanilla Nutty Green bean Milk concentrate Orange blossom	**Texture :** Skin persistence Toughness of skin Fudge Graininess Mealiness Stickiness
	Appearance: Skin surface

Table 17: Sensory descriptors for bean (raw / cooked)

Schedule : Tables and Figures

List of figures

For obvious lay-out reasons, screen captures from the R software are not identified. However, the text associated with these pictures is clear enough to identify them.

List of tables