



Best-ReMaP

Healthy Food for a Healthy Future



WP 5 – REFORMULATION AND PROCESSED FOOD MONITORING

T+1 Statistics methodology (task 5.5.1)

Part 1

Anses

20.10.2022



Training for statistics on T+1 data

1st part

20th October 2022

Introduction to R software + verifications to perform on the data collected before being able to produce indicators

Today

2nd part

24th November 2022

Creation of indicators on T+1 data and comparisons with pre-existing data + production of the report



1) Introduction of task 5.5.1 [\(page 4\)](#)

2) Installation of the necessary equipment and presentation of the Rstudio software [\(page 10\)](#)

- A. Preliminary steps [\(page 11\)](#)
- B. Installation of software [\(page 39\)](#)
- C. Introduction to R studio [\(page 48\)](#)
- D. Cleaning of the Rstudio interface [\(page 62\)](#)

3) Running of the verification programs [\(page 69\)](#)

- A. Part 1: R setup program [\(page 78\)](#)
- B. Part 2: Verification programs and template cleaning/standardization [\(page 88\)](#)
 - i. 1st verification program : '*Verifications template_step_1*' [\(page 90\)](#)
 - ii. 2nd verification program : '*Verifications template_step_2*' [\(page 119\)](#)
 - iii. 3rd verification program : '*Verifications template_step_3*' [\(page 144\)](#)
 - iv. 4th verification program : '*Verifications template_step_4*' [\(page 166\)](#)

4) Next steps [\(page 191\)](#)





1) Introduction of task 5.5.1





Reminder of the activities objectives of WP5

Activities

Task
5.4.1

- **Data collection** : pictures of products taken in shops
- **Data entry** : entry of all product information in a common template
- **Data codification** :
 - Classification of products of the template into categories and subcategories that allow homogeneous grouping of products according to recipe, ingredients, ...
 - The subcategories are defined to enable to identify best formulation and room for reformulation

Task
5.5.1

- **Production of indicators** :
 - Indicators that are defined for monitoring (analysis of the food offer, nutritional values, portion size, ...)
 - These indicators are produced at the subcategory level because analysis at the subcategory level have been identified as the “gold standard” during JANPA

Objectives

- **Data available to characterize the food offer and the nutritional quality of processed food at a given time** (number of products by subcategory, nutrient values, list of ingredients...)
- **Monitoring**: knowledge of the evolution over time (reformulation, addition of new products...)
- **Assessment of the impact of nutrition policy measures** both on the food offer and the composition of processed food (commitments with industry, thresholds, taxes)





Task 5.5.1 : Production of statistics for the 2nd snapshot (T+1)

Partners involved in batch 1
Austria
Belgium
Estonia
Germany
Hungary
Romania

+ Ireland

(involved in the first European snapshot T0)

- Partners involved in the Batch 1 of the second European Snapshot (T+1) will have **to produce statistics/indicators** by Best-ReMaP categories and subcategories of products collected during the T+1
- **Comparisons with pre-existing data** will be made in order to establish trend assessments of the nutritional quality of the processed food
- Each partners will have **to produce a report** that will be included in the WP5 deliverable (D5.3)





WORK Package 5 – Reformulation and processed food monitoring

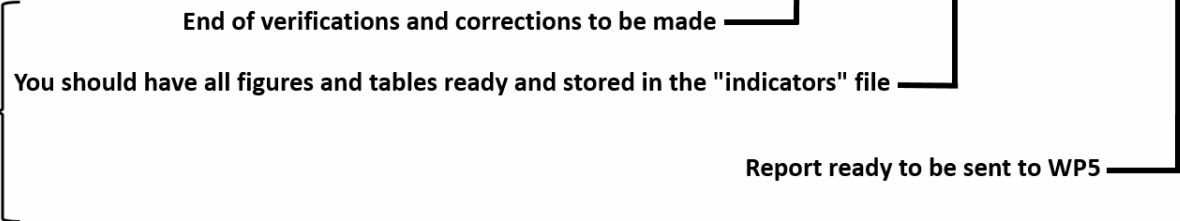
Timeline of the task 5.5.1

2022												2023					
j	f	m	a	m	j	j	a	s	o	n	d	j	f	m	a	m	j
M16	M17	M18	M19	M20	M21	M22	M23	M24	M25	M26	M27	M28	M29	M30	M31	M32	M33
Task 5.4.1 : Data collection and data treatment (data entry and encoding)																	
Mar 2022 - Feb 2023																	
		Collection of pictures Mar 2022	Data treatment (data entry and encoding) Apr 2022 - Jan 2023 (February 2023 at the latest)														
			Soft drinks	Breakfast cereals	Bread products	Fresh dairy products and desserts	Delicatessen meats and similar										
												Task 5.5.1 : Production of statistics					
												Nov 2022 - Apr 2023 (6 months)					
												Verifications steps Nov - Jan 2022	Production of indicators Feb 2023	Production of the report Mar - Apr 2023			
												<i>(3 months)</i>	<i>(1 month)</i>	<i>(2 months)</i>			

1st
Nov
2022

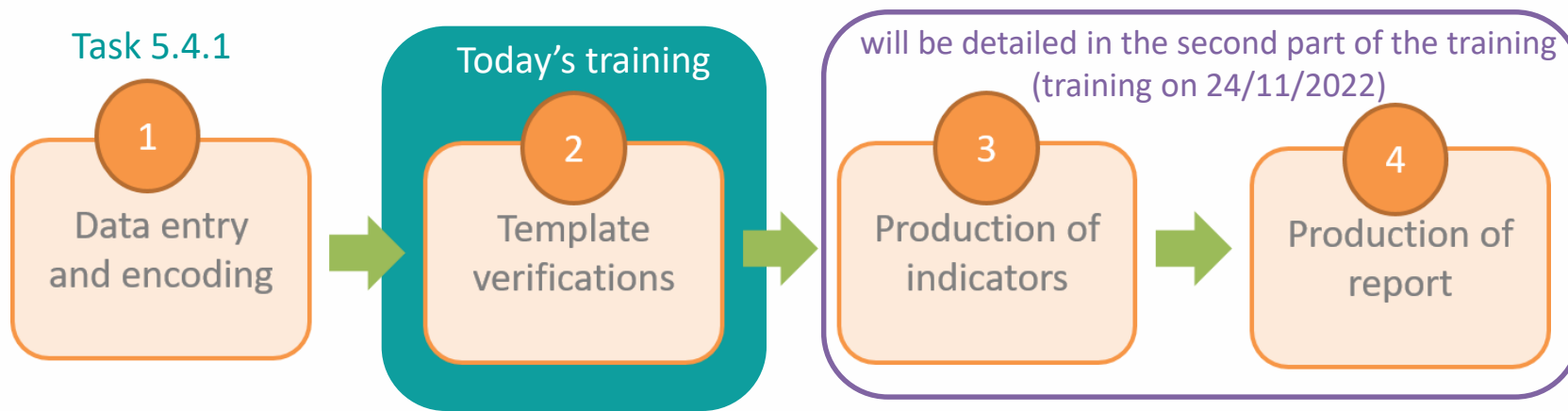
Please send us feedback to confirm that the R and Rstudio software are installed
If you have any problems before this date, please let us know so that we can help you

These deadlines are indications so that you can organize yourselves as well as possible





Summary of the steps in task 5.5.1



It is important to finish each step before moving on to the next

- **Steps 2 and 3** will be done by data processing programs on the Rstudio software which is a free software and therefore accessible to all. The programs have already been created and written in order to harmonise the work and ensure that all partners work in the same way. You will just need to run the programs on your data (you will not have to create any programs).
- **Step 3 and 4** will not be detailed in this training but in the 2nd part of the training on **24th November 2022**
- **Step 4** : you will receive all the elements so that everyone can produce a report on the analysis of his data. A sample report will be provided to ensure that the results for each country are presented in a harmonised way. You will have more information about the report during the 2nd part of the training.





Equipment needed

Tools you already have

- **T+1 collection template** : template that you filled in according to the WP5 methodology during your T+1 data collection and with data for the 5 priority food categories.
- **Pictures of the products** that you have collected for your T+1 data collection
- **Pre-existing data template** : template with your pre-existing national data that have been classified in the Best-ReMaP nomenclature

Tools you will have to download (explanation pages [40-47](#))

- **R, Rstudio** : Free statistical software that you will use to check and correct the data entered in your template and to perform indicators/statistics on your data
- **WP5 R programs** : R programs created by Anses that you will just have to run on the Rstudio software (you will be guided at each step, you will not have to create programs or develop code).
- **Support files for programs** (*wp5_verification_file*; *Best-ReMaP_nomenclature*; *Subcategory_order*; *Years of interest*) : files that you will have to download and as it will be used in the R program. These files do **not have to be filled in or modified** (except the file 'Years of interest').

The **WP5 R programs** and the **support files** are available in a **zip folder** on the project intranet by following this link:
https://portal.nijz.si/ssf/a/c/p_name/ss_forum/p_action/1/binderId/21932/entityType/folder/action/view_permalink/novl_url/1

You will need to copy this **folder** as it is on the desktop of your computer (this action is detailed in the [slide 32-33](#))



2) Installation of the necessary equipment and presentation of the Rstudio software

A. Preliminary steps ([page 11](#))

B. Installation of software ([page 39](#))

C. Introduction to R studio ([page 48](#))

D. Cleaning of the Rstudio interface ([page 62](#))





2) Installation of the necessary equipment and presentation of the Rstudio software

A. Preliminary steps

B. Installation of software

C. Introduction to R studio

D. Cleaning of the Rstudio interface





Preliminary steps

1st preliminary step : preparation of the T+1 collection template [\(page 13\)](#)

2nd preliminary step : preparation of the pre-existing data template [\(page 30\)](#)

3rd preliminary step : creation of the working folder [\(page 32\)](#)

4th preliminary step : preparation of the file "Years of interest.csv" [\(page 34\)](#)





Preliminary steps

1st preliminary step : preparation of the T+1 collection template

- You must ensure that your T+1 collection template is a **single file** with a **single tab** for all data collected during T+1 (the 5 food categories in the same tab).
- As your T+1 collection template contains many rows and drop-down menus, you will have to copy it into a **new .xlsx excel file** to keep only the filled rows and remove the drop-down menus.

see the following slides for a step-by-step explanation of this procedure

It is also possible for you to perform the verification steps category by category by working on 5 different files (1 per Best-ReMaP category).

Please note that this will result in more file handling and that you will need to gather all your T+1 data in one file at the end of your verification steps. The program to create the indicators will have to be run on the compiled file of all verified categories (see training of 24 November).





Preliminary steps

collection template T+1 France.xlsx - Excel

File Home Insert Page Layout Formulas Data Review View Tell me what you want to do... LAGUITTON Thomas Share

Clipboard Font Alignment Number Styles Cells Editing

A1 Product_code

	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV
1	Fat_as_consumed	Saturated fat_as_consumed	Carbohydrates_as_consumed	Sugar_as_consumed	Protein_as_consumed	Salt_as_consumed	Fibre_as_consumed	Comment	Category_code	Subcategory_code
2297									9	650
2298									9	651
2299									9	651
2300									9	650
2301									9	651
2302									9	651
2303									9	650
2304									9	651
2305									9	650
2306									9	650
2307									9	644
2308									9	649
2309									9	648
2310									9	656
2311									9	648
2312										
2313										
2314										

Average: 3,34146E+11 Count: 70614 Sum: 1,10743E+16 100%

1- In your T+1 collection template, **select all lines containing your data** by scrolling manually until the last line that has been filled with your data.

Make sure to select only the lines with data and all the columns (from *Product_code* to *Subcategory_code*)





WORK Package 5 – Reformulation and processed food monitoring

Preliminary steps

The screenshot shows the Microsoft Excel interface with the following details:

- File Name: collection template T+1 France.xlsx - Excel
- Current Tab: Home
- Formula Bar: Product_code
- Selected Range: A1:AV14
- Context Menu: Open over the selection, with 'Copy' circled in red.
- Spreadsheet Columns: AM, AP, AQ, AR, AS, AT, AU, AV
- Spreadsheet Headers: Fat_as_consumed, Sugar_as_consumed, Protein_as_consumed, Salt_as_consumed, Fibre_as_consumed, Comment, Category_code, Subcategory_code
- Spreadsheet Data (Rows 1-14):

	Fat_as_consumed	Sugar_as_consumed	Protein_as_consumed	Salt_as_consumed	Fibre_as_consumed	Comment	Category_code	Subcategory_code
1								
2297							9	650
2298							9	651
2299							9	651
2300							9	650
2301							9	651
2302								651
2303								650
2304								651
2305								650
2306								644
2307								649
2308								648
2309								648
2310							9	656
2311							9	648
2312								
2313								
2314								
- Text Box: 2- Right-click and copy the selection





WORK Package 5 – Reformulation and processed food monitoring

Preliminary steps

3- Open a new excel file, empty

4- Paste your selection by using the option *Paste special > Paste Values*

The screenshot shows the Excel interface with the 'Home' tab selected. The 'Paste' dropdown menu is open, showing options like 'Paste', 'Paste Values', and 'Paste Special...'. The spreadsheet grid is visible with columns C through U and rows 5 through 21. A text box is overlaid on the grid, providing instructions for steps 3 and 4.





WORK Package 5 – Reformulation and processed food monitoring

Preliminary steps

Book1 - Excel

File Home Insert Page Layout Formulas Data Review View Tell me what you want to do... LAGUITTON Thomas Share

Clipboard Font Conditional Format as Cell Insert Delete Format AutoSum Fill Sort & Find & Select

A1 Product_code

Overview of your new .xlsx file with your data (without the drop-down menus)

	A	B	C	D	E	F		T	U												
1	Product_c	Father_pr	Country	Year	Category	Subcatego	Bar_code	Assortme	Brand_na	Brand_ow	Type_of_k	Legal_nan	Legal_nan	Commerci	Commerci	FOP_label	Nutri_Sco	Ingredient	Net_weig	Net_weig	Numb
2	1650		Ireland	2021	Breakfast	Cereals w	4,09E+12	No	Kavanagh	Aldi	Hard discc	Gluten Fre	Gluten Fre	Gluten Fre	Gluten Fre	None from the list		Gluten Fre	500 g		
3	1651		Ireland	2021	Breakfast	Traditiona	4,09E+12	No	Kavanagh	Aldi	Hard discc	Wholegra	Wholegra	High Ome	High Ome	Traffic light		Jumbo Oa	500 g		
4	1653		Ireland	2021	Breakfast	Cereals w	4,09E+12	No	Kavanagh	Aldi	Hard discc	Instant Po	Instant Po	Microwav	Microwav	Traffic light		Wholegra	500 g		
5	1654		Ireland	2021	Breakfast	Cereals w	4,09E+12	No	Kavanagh	Aldi	Hard discc	Wholegra	Wholegra	Organic Ju	Organic Ju	Traffic light		Oat Flake	750 g		
6	1655		Ireland	2021	Breakfast	Cereals w	4,09E+12	No	Kavanagh	Aldi	Hard discc	Wholegra	Wholegra	Porridge C	Porridge C	Traffic light		Wholegra	1500 g		
7	1657		Ireland	2021	Breakfast	Traditiona	4,09E+12	No	Specially	Aldi	Hard discc	A Blend o	A Blend o	Very Bern	Very Bern	Traffic light		Blackcurra	500 g		
8	1658	603	Ireland	2021	Breakfast	High-fibre	4,09E+12	No	Harvest M	Aldi	Hard discc	Wholegra	Wholegra	Bran Flake	Bran Flake	Traffic light		Wholegra	750 g		
9	1659		Ireland	2021	Breakfast	Chocolate	4,09E+12	No	Harvest M	Aldi	Hard discc	Cocoa Fla	Cocoa Fla	Choco Ho	Choco Ho	Traffic light		Wholegra	375 g		
10	1661	937	Ireland	2021	Breakfast	Chocolate	4,09E+12	No	Harvest M	Aldi	Hard discc	Milk Choc	Milk Choc	Choco Ric	Choco Ric	Traffic light		Rice, Suga	375 g		
11	1663	586	Ireland	2021	Breakfast	Sweet cer	4,09E+12	No	Harvest M	Aldi	Hard discc	Corn Flake	Corn Flake	Corn Flake	Corn Flake	Traffic light		Maize, Su	500 g		
12	1665	916	Ireland	2021	Breakfast	Crunchy fr	4,09E+12	No	Harvest M	Aldi	Hard discc	Crunchy C	Crunchy C	Crisp Cere	Crisp Cere	Traffic light		Oat Flake	500 g		
13	1672		Ireland	2021	Breakfast	Crunchy n	4,09E+12	No	Harvest M	Aldi	Hard discc	Oat, Corn	Oat, Corn	Honey & N	Honey & N	Traffic light		WHOLEGR	500 g		
14	1674	597	Ireland	2021	Breakfast	Sweet cer	4,09E+12	No	Harvest M	Aldi	Hard discc	Toasted F	Toasted F	Honey Nu	Honey Nu	Traffic light		MAIZE (63	500 g		
15	1675		Ireland	2021	Breakfast	Crunchy fr	4,09E+12	No	Harvest M	Aldi	Hard discc	Low Sugar	Low Sugar	Low Sugar	Low Sugar	Traffic light		Oat flakes	500 g		
16	1676		Ireland	2021	Breakfast	Crunchy fr	4,09E+12	No	Harvest M	Aldi	Hard discc	Low Sugar	Low Sugar	Low Sugar	Low Sugar	Traffic light		Oat Flake	500 g		
17	1677	937	Ireland	2021	Breakfast	Chocolate	4,09E+12	No	Harvest M	Aldi	Hard discc	Chocolate	Chocolate	Choco Ric	Choco Ric	Traffic light		Rice (83%	375 g		
18	1680	996	Ireland	2021	Breakfast	Sweet cer	4,09E+12	No	Harvest M	Aldi	Hard discc	Rice and V	Rice and V	Original B	Original B	Traffic light		RICE (66%	500 g		
19	1681		Ireland	2021	Breakfast	Traditiona	4,09E+12	No	Harvest M	Aldi	Hard discc	Toasted H	Toasted H	Protein Gi	Protein Gi	Traffic light		Wholegra	400 g		
20	1682	850	Ireland	2021	Breakfast	Traditiona	4,09E+12	No	Harvest M	Aldi	Hard discc	Honey To	Honey To	Raisin & A	Raisin & A	Traffic light		WHOLEGR	1000 g		
21	1683		Ireland	2021	Breakfast	Traditiona	4,09E+12	No	Harvest M	Aldi	Hard discc	Oats with	Oats with	Really Nu	Really Nu	Traffic light		Wholegra	500 g		

Sheet1

Select destination and press ENTER or choose Paste

Average: 3,34146E+11 Count: 70614 Sum: 1,10743E+16





Preliminary steps

1st preliminary step : preparation of the T+1 collection template

→ Now you have an **.xlsx file** containing only the lines with your data and without the drop-down menus.

You can rename it **T+1 data collection *country*.xlsx** (with the name of your own country) for example.

→ You must **make a copy** of this file (T+1 collection template *country*.xlsx) and save it in **.csv (comma separator) format** under the name **T+1_data_collection_ *country*.csv** (with the name of your own country)

The creation of this file in .csv format is very important because it is this file that will be used in the R software for the verification and indicator creation stages.

see the following slides for a step-by-step explanation of this procedure





WORK Package 5 – Reformulation and processed food monitoring

Preliminary steps

Click on the tab **File**

T+1 collection template France.xlsx - Excel

LAGUITTON Thomas Share

Product_code	Father_product_code	Country	Year	Category_name	Subcategory_name	Bar_code	Assortment	Brand_name	Brand_owner
1650		Ireland	2021	Breakfast cereals	Cereals without added sugar	4088600257617	No	Kavanagh's	Aldi
1651		Ireland	2021	Breakfast cereals	Traditional muesli flakes	4088600052724	No	Kavanagh's	Aldi
1653		Ireland	2021	Breakfast cereals	Cereals without added sugar	4088600055985	No	Kavanagh's	Aldi
1654		Ireland	2021	Breakfast cereals	Cereals without added sugar	4088600326597	No	Kavanagh's	Aldi
1655		Ireland	2021	Breakfast cereals	Cereals without added sugar	4088600056128	No	Kavanagh's	Aldi
1657		Ireland	2021	Breakfast cereals	Traditional muesli flakes	4088600167756	No	Specially Selected	Aldi
1658	603	Ireland	2021	Breakfast cereals	High-fibre cereals	4088600118871	No	Harvest Morn	Aldi
1659		Ireland	2021	Breakfast cereals	Chocolate-flavoured cereals	4088600140261	No	Harvest Morn	Aldi
1661	937	Ireland	2021	Breakfast cereals	Chocolate-flavoured cereals	4088600118888	No	Harvest Morn	Aldi
1663	586	Ireland	2021	Breakfast cereals	Sweet cereal flakes	4088600248646	No	Harvest Morn	Aldi
1665	916	Ireland	2021	Breakfast cereals	Crunchy fruit muesli	4088600248493	No	Harvest Morn	Aldi
1672		Ireland	2021	Breakfast cereals	Crunchy muesli with nuts_seeds	4088600064529	No	Harvest Morn	Aldi
1674	597	Ireland	2021	Breakfast cereals	Sweet cereal flakes	4088600247465	No	Harvest Morn	Aldi
1675		Ireland	2021	Breakfast cereals	Crunchy fruit muesli	4088600163284	No	Harvest Morn	Aldi
1676		Ireland	2021	Breakfast cereals	Crunchy fruit muesli	4088600163260	No	Harvest Morn	Aldi
1677	937	Ireland	2021	Breakfast cereals	Chocolate-flavoured cereals	4088600166100	No	Harvest Morn	Aldi
1680	996	Ireland	2021	Breakfast cereals	Sweet cereal flakes	4088600248523	No	Harvest Morn	Aldi
1681		Ireland	2021	Breakfast cereals	Traditional muesli flakes	4088600248776	No	Harvest Morn	Aldi





Preliminary steps

T+1 collection template France.xlsx - Excel

LAGUITTON Thomas

Export

- Create PDF/XPS Document
- Change File Type

Create a PDF/XPS Document

- Preserves layout, formatting, fonts, and images
- Content can't be easily changed
- Free viewers are available on the web

Create PDF/XPS

Go to *Export*





Preliminary steps

T+1 collection template

Export

Info
New
Open
Save
Save As
Print
Share
Export
Publish
Close
Account
Options

Create PDF/XPS Document

Change File Type

Change File Type

Workbook File Types

- Workbook (*.xlsx)
Uses the Excel Spreadsheet format
- OpenDocument Spreadsheet (*.ods)
Uses the OpenDocument Spreadsheet format
- Macro-Enabled Workbook (*.xlsm)
Macro enabled spreadsheet
- Template (*.xltx)
Starting point for new spreadsheets
- Binary Workbook (*.xlsb)
Optimized for fast loading and saving

Other File Types

- Text (Tab delimited) (*.txt)
Text format separated by tabs
- Formatted Text (Space delimited) (*.prn)
Text format separated by spaces
- CSV (Comma delimited) (*.csv)
Text format separated by commas

Save As

1- Select **Change File Type**

2- Select **CSV (Comma delimited) (*.csv)**

3- Press **Save As**





WORK Package 5 – Reformulation and processed food monitoring

Preliminary steps

Save your file on your desktop under the name : *T+1_data_collection_country.csv* (with the name of your own country)

	G	H	I	J
	Bar_code	Assortment	Brand_name	Brand_owner
	4088600257617	No	Kavanagh's	Aldi
	4088600052724	No	Kavanagh's	Aldi
	4088600055985	No	Kavanagh's	Aldi
	4088600326597	No	Kavanagh's	Aldi
	4088600056128	No	Kavanagh's	Aldi
	4088600167756	No	Specially Selected	Aldi
	4088600118871	No	Harvest Morn	Aldi
	4088600140261	No	Harvest Morn	Aldi
	4088600118888	No	Harvest Morn	Aldi
	4088600248646	No	Harvest Morn	Aldi
	4088600248493	No	Harvest Morn	Aldi
	4088600064529	No	Harvest Morn	Aldi
	4088600247465	No	Harvest Morn	Aldi
	4088600163284	No	Harvest Morn	Aldi
	4088600163260	No	Harvest Morn	Aldi
	4088600166100	No	Harvest Morn	Aldi
	4088600248523	No	Harvest Morn	Aldi
	4088600248776	No	Harvest Morn	Aldi





WORK Package 5 – Reformulation and processed food monitoring

Preliminary steps

Microsoft Excel

Some features in your workbook might be lost if you save it as CSV (Comma delimited). Do you want to keep using that format?

Press Yes

Product_code	Father_product_code	Country	Year	Category_name	Subcategory_name	Bar_code	Assortment	Brand_name	Brand_owner
1650		Ireland	2021	Breakfa		57617	No	Kavanagh's	Aldi
1651		Ireland	2021	Breakfa		52724	No	Kavanagh's	Aldi
1653		Ireland	2021	Breakfa		55985	No	Kavanagh's	Aldi
1654		Ireland	2021	Breakfa		26597	No	Kavanagh's	Aldi
1655		Ireland	2021	Breakfa		56128	No	Kavanagh's	Aldi
1657		Ireland	2021	Breakfa		57756	No	Specially Selected	Aldi
1658	603	Ireland	2021	Breakfast cereals	High fibre cereals	4088600118871	No	Harvest Morn	Aldi
1659		Ireland	2021	Breakfast cereals	Chocolate-flavoured cereals	4088600140261	No	Harvest Morn	Aldi
1661	937	Ireland	2021	Breakfast cereals	Chocolate-flavoured cereals	4088600118888	No	Harvest Morn	Aldi
1663	586	Ireland	2021	Breakfast cereals				Harvest Morn	Aldi
1665	916	Ireland	2021	Breakfast cereals				Harvest Morn	Aldi
1672		Ireland	2021	Breakfast cereals				Harvest Morn	Aldi
1674	597	Ireland	2021	Breakfast cereals	Sweet cereal flakes	4088600247465	No	Harvest Morn	Aldi
1675		Ireland	2021	Breakfast cereals	Crunchy fruit muesli	4088600163284	No	Harvest Morn	Aldi
1676		Ireland	2021	Breakfast cereals	Crunchy fruit muesli	4088600163260	No	Harvest Morn	Aldi
1677	937	Ireland	2021	Breakfast cereals	Chocolate-flavoured cereals	4088600166100	No	Harvest Morn	Aldi
1680	996	Ireland	2021	Breakfast cereals	Sweet cereal flakes	4088600248523	No	Harvest Morn	Aldi
1681		Ireland	2021	Breakfast cereals	Traditional Muesli flakes	4088600248776	No	Harvest Morn	Aldi





Preliminary steps



When you save your **.xlsx** file in **.csv format**, the barcodes in the **.csv file** appear in scientific writing (e.g. 1.89E+12).

It is necessary to select the column 'bar_codes' and change the column format to '**Number**' with **0 decimal digit**.

The bar codes will appear in full and you will not lose any information. You can then save this new version of your template with the full barcodes and overwrite the old one.

Be careful, as soon as you reopen this new **.csv file**, the barcodes will be written scientifically again and you will have to repeat this procedure.

We therefore invite you to do this procedure only once when you save the **.csv file** and not to reopen the saved file.

(To be read in R, the file must not be opened so you do not need to re-open it)

see the following slides for a step-by-step explanation of this procedure





WORK Package 5 – Reformulation and processed food monitoring

Preliminary steps

	A	B	C	D	E	F	G	H	I	J
1	Product_code	Father_pr	Country	Year	Category_name	Subcategory_name	Bar_code	Assortment	Brand_name	Brand_owner
41	340		Poland	2021	Delicatessen me	Dried, smoked or c	20426156	No	SaintAlby	
42	341		Poland	2021	Breakfast cereals	Filled cereals	3,56E+12	No	carrefour	
43	342		Poland	2021	Delicatessen me	Pâté	3,56E+12	No	Reflets de France	
44	343		Poland	2021	Delicatessen me	Pâté	3,25E+12	No	Reflets de France	
45	343		Poland	2021	Delicatessen me	Dry sausage	5,41E+12	No	Marcassou	
46	345		Poland	2021	Fresh dairy prod	Classic plain fresh c	5,4E+12	No	Carrefour	
47	346		Poland	2021	Fresh dairy prod	Artificially-sweete	20916626	No	Milbona	
48	347		Poland	2021	Fresh dairy prod	Dessert creams and	3,56E+12	No	Carrefour bio	
49	348		Poland	2021	Fresh dairy prod	Dessert creams and	3,03E+12	No	Danone	Danone
50	349		Poland	2021	Fresh dairy prod	Fresh sweetened s	5,41E+12	No	Alpro	
51	350		Poland	2021	Breakfast cereals	Sweet cereal flakes	27051832	No	Goldenbridge	
52	351		Poland	2021	Delicatessen me	Sausages	20773090	No	Lidl	
53	352		Poland	2021	Delicatessen me	Sausages	27013144	No	Pand	
54	353		Poland	2021	Delicatessen me	Pâté	27097168	No	Dçlif	
55	354		Poland	2021	Fresh dairy prod	Classic sweet yoghu	5,41E+12	No	Dano	
56	355		Poland	2021	Delicatessen me	Dry sausage	20501495	No	Delu	
57	356		Poland	2021	Bread products	Plain toasted bread	8,71E+12	No	Haus	
58	357		Poland	2021	Fresh dairy prod	Classic sweetened	27053417	Yes	Milsa	
59	358		Poland	2021	Fresh dairy prod	Classic plain fresh c	20669850	Yes	Milbo	
60	359		Poland	2021	Fresh dairy prod	Classic sweet yoghu	5,41E+12	Yes	PurNatur	
61	360		Poland	2021	Bread products	Puffed cakes	3,56E+12	No	Carrefour	
62	361		Poland	2021	Fresh dairy prod	Artificially-sweete	5,41E+12	No	Danone	Danone
63	362		Poland	2021	Delicatessen me	Pâté	5,4E+12	No	Carrefour	
64	363		Poland	2021	Delicatessen me	Cured ham	5,4E+12	No	Carrefour	
65	364		Poland	2021	Bread products	Other rusks	6,12E+11	No	Nairns	
66	365		Poland	2021	Bread products	Pre-packaged break	27046197	No	Délipain	
67	366		Poland	2021	Bread products	Wholemeal_cereal	7,31E+12	No	Krisprolls	
68	367		Poland	2021	Soft drinks	Flavoured milk bev	3,43E+12	No	Lactel	
69	368		Poland	2021	Bread products	Pre-packaged break	5,41E+13	No	Biaform	
70	369		Poland	2021	Soft drinks	Sugar-sweetened f	27002766	No	River	
71	370		Poland	2021	Soft drinks	Other beverages w	8,85E+12	No	UFC	
72	371		Poland	2021	Soft drinks	Colas without adde	4,06E+12	No	Pepsi	lion's
73	372		Poland	2021	Bread products	Other rusks	3,76E+12	No	Léanature)

Bar codes appearing in scientific format in your file
T+1_data_collection_country.csv



WORK Package 5 – Reformulation and processed food monitoring

Preliminary steps

The screenshot shows the Microsoft Excel interface with the 'Home' tab selected. A spreadsheet is open with columns labeled A through J. Column G is highlighted, and a right-click context menu is open over it. The menu options include Cut, Copy, Paste Options, Paste Special..., Insert, Delete, Clear Contents, Format Cells..., Column Width..., Hide, and Unhide. The 'Format Cells...' option is highlighted. The spreadsheet data includes columns for Product_code, Father_pr, Country, Year, Category_name, Subcategory_name, Bar_code, Brand_name, and Brand_ownership.

	A	B	C	D	E	F	G	H	I	J
1	Product_code	Father_pr	Country	Year	Category_name	Subcategory_name	Bar_code		Brand_name	Brand_ownership
41	340		Poland	2021	Delicatessen me	Dried, smoked or c	2042		ntAlby	
42	341		Poland	2021	Breakfast cereals	Filled cereals	3,56		refour	
43	342		Poland	2021	Delicatessen me	Pâté	3,56		lets de France	
44	343		Poland	2021	Delicatessen me	Pâté	3,25		lets de France	
45	343		Poland	2021	Delicatessen me	Dry sausage	5,41		ccassou	
46	345		Poland	2021	Fresh dairy prod	Classic plain fresh c	5,4		refour	
47	346		Poland	2021	Fresh dairy prod	Artificially-sweete	2091		bona	
48	347		Poland	2021	Fresh dairy prod	Dessert creams and	3,56		refour bio	
49	348		Poland	2021	Fresh dairy prod	Dessert creams and	3,03		ione	Danone
50	349		Poland	2021	Fresh dairy prod	Fresh sweetened s	5,41		ro	
51	350		Poland	2021	Breakfast cereals	Sweet cereal flakes	2705		denbridge	
52	351		Poland	2021	Delicatessen me	Sausages	2077			
53	352		Poland	2021	Delicatessen me	Sausages	2701			
54	353		Poland	2021	Delicatessen me	Pâté	2709/168	No	Uclifin	
55	354		Poland	2021	Fresh dairy prod	Classic sweet yogh	5,41E+12	No	Danone	Danone
56	355		Poland	2021	Delicatessen me	Dry sausage	20501495	No	Deluxe	
57	356		Poland	2021	Bread products	Plain toasted bread	8,71E+12	No	Haust	
58	357		Poland	2021	Fresh dairy prod	Classic sweetened	27059647	Yes	Milsa	
59	358		Poland	2021	Fresh dairy prod	Classic plain fresh c	20669850	Yes	Milbo	
60	359		Poland	2021	Fresh dairy prod	Classic sweet yogh	5,41E+12	Yes	PurN	
61	360		Poland	2021	Bread products	Puffed cakes	3,56E+12	No	Carre	
62	361		Poland	2021	Fresh dairy prod	Artificially-sweete	5,41E+12	No	Danone	Danone
63	362		Poland	2021	Delicatessen me	Pâté	5,4E+12	No	Carrefour	
64	363		Poland	2021	Delicatessen me	Cured ham	5,4E+12	No	Carrefour	
65	364		Poland	2021	Bread products	Other rusks	6,12E+11	No	Nairs	
66	365		Poland	2021	Bread products	Pre-packaged brea	27046197	No	Délpain	
67	366		Poland	2021	Bread products	Wholemeal_cereal	7,31E+12	No	Krisprolls	
68	367		Poland	2021	Soft drinks	Flavoured milk bev	3,43E+12	No	Lactel	
69	368		Poland	2021	Bread products	Pre-packaged brea	5,41E+13	No	Biaform	
70	369		Poland	2021	Soft drinks	Sugar-sweetened f	27002766	No	River	
71	370		Poland	2021	Soft drinks	Other beverages w	8,85E+12	No	UFC	
72	371		Poland	2021	Soft drinks	Colas without adde	4,06E+12	No	Pepsi	
73	372		Poland	2021	Bread products	Other rusks	3,76E+12	No	Léanature	

Select the *Bar_codes* column, right click and go to *Format Cells*



WORK Package 5 – Reformulation and processed food monitoring

Preliminary steps

The screenshot shows the Microsoft Excel interface with the 'Format Cells' dialog box open. The 'Number' tab is selected, and the 'Category' is set to 'Number'. The 'Decimal places' are set to 0. The 'Negative numbers' section shows three options: '-1234', '1234', and '-1234'. The background spreadsheet has columns for Product_code, Father_pr, Country, Year, and Category.

Product_code	Father_pr	Country	Year	Category
340		Poland	2021	Delicatessen
341		Poland	2021	Breakfast cereals
342		Poland	2021	Delicatessen
343		Poland	2021	Delicatessen
345		Poland	2021	Fresh dairy products
346		Poland	2021	Fresh dairy products
347		Poland	2021	Fresh dairy products
348		Poland	2021	Fresh dairy products
349		Poland	2021	Fresh dairy products
350		Poland	2021	Breakfast cereals
351		Poland	2021	Delicatessen
352		Poland	2021	Delicatessen
353		Poland	2021	Delicatessen
354		Poland	2021	Fresh dairy products
355		Poland	2021	Delicatessen
356		Poland	2021	Bread products
357		Poland	2021	Fresh dairy products
358		Poland	2021	Fresh dairy products
359		Poland	2021	Fresh dairy products
360		Poland	2021	Bread products
361		Poland	2021	Fresh dairy products
362		Poland	2021	Delicatessen
363		Poland	2021	Delicatessen
364		Poland	2021	Bread products
365		Poland	2021	Bread products
366		Poland	2021	Bread products
367		Poland	2021	Soft drinks
368		Poland	2021	Bread products
369		Poland	2021	Soft drinks
370		Poland	2021	Soft drinks
371		Poland	2021	Soft drinks
372		Poland	2021	Bread products

In the **Number** tab, choose the 'number' category, indicate '0' for decimal places and click OK



WORK Package 5 – Reformulation and processed food monitoring

Preliminary steps

	A	B	C	D	E	F	G	H	I	J
1	Product_code	Father_pr	Country	Year	Category_name	Subcategory_name	Bar_code	Assortment	Brand_name	Brand_ownd
41	340		Poland	2021	Delicatessen me	Dried, smoked or c	20426156	No	SaintAlby	
42	341		Poland	2021	Breakfast cereals	Filled cereals	3560071013226	No	carrefour	
43	342		Poland	2021	Delicatessen me	Pâté	3560070976140	No	Reflets de France	
44	343		Poland	2021	Delicatessen me	Pâté	3245390135445	No	Reflets de France	
45	343		Poland	2021	Delicatessen me	Dry sausage	5410916000042	No	Marcassou	
46	345		Poland	2021	Fresh dairy prod	Classic plain fresh c	5400101000766	No	Carrefour	
47	346		Poland	2021	Fresh dairy prod	Artificially-sweete	20916626	No	Milbona	
48	347		Poland	2021	Fresh dairy prod	Dessert creams and	3560070917051	No	Carrefour bio	
49	348		Poland	2021	Fresh dairy prod	Dessert creams and	3033490143589	No	Danone	Danone
50	349		Poland	2021	Fresh dairy prod	Fresh sweetened s	5411188119586	No	Alpro	
51	350		Poland	2021	Breakfast cereals	Sweet cereal flakes	27051832	No	Goldenbridge	
52	351		Poland	2021	Delicatessen me	Sausages	20773090	No	Lidl	
53	352		Poland	2021	Delicatessen me	Sausages	27013144	No	Panda	
54	353		Poland	2021	Delicatessen me	Pâté	27097168	No	DÇlifin	
55	354		Poland	2021	Fresh dairy prod	Classic sweet yogh	5410146417061	No	Danone	Danone
56	355		Poland	2021	Delicatessen me	Dry sausage	20501495	No	Deluxe	
57	356		Poland	2021	Bread products	Plain toasted bread	8710445018233	No	Haust	
58	357		Poland	2021	Fresh dairy prod	Classic sweetened	27059647	Yes	Milsani	
59	358		Poland	2021	Fresh dairy prod	Classic plain fresh c	20669850	Yes	Milbona	
60	359		Poland	2021	Fresh dairy prod	Classic sweet yogh	5412971116287	Yes	PurNatur	
61	360		Poland	2021	Bread products	Puffed cakes	3560070600946	No	Carrefour	
62	361		Poland	2021	Fresh dairy prod	Artificially-sweete	5410146416026	No	Danone	Danone
63	362		Poland	2021	Delicatessen me	Pâté	5400101007635	No	Carrefour	
64	363		Poland	2021	Delicatessen me	Cured ham	5400101063365	No	Carrefour	
65	364		Poland	2021	Bread products	Other rusks	612322000738	No	Nairns	
66	365		Poland	2021	Bread products	Pre-packaged brea	27046197	No	Délipain	
67	366		Poland	2021	Bread products	Wholemeal_cereal	7311070331318	No	Krisprolls	
68	367		Poland	2021	Soft drinks	Flavoured milk bev	3428271180011	No	Lactel	
69	368		Poland	2021	Bread products	Pre-packaged brea	54107246001113	No	Biaform	
70	369		Poland	2021	Soft drinks	Sugar-sweetened f	27002766	No	River	
71	370		Poland	2021	Soft drinks	Other beverages w	8850025000026	No	UFC	
72	371		Poland	2021	Soft drinks	Colas without adde	4060800105882	No	Pepsi	
73	372		Poland	2021	Bread products	Other rusks	3760020508371	No	Léanature	

Your barcodes appear in full, you can save this table by overwriting the previous version and close it.

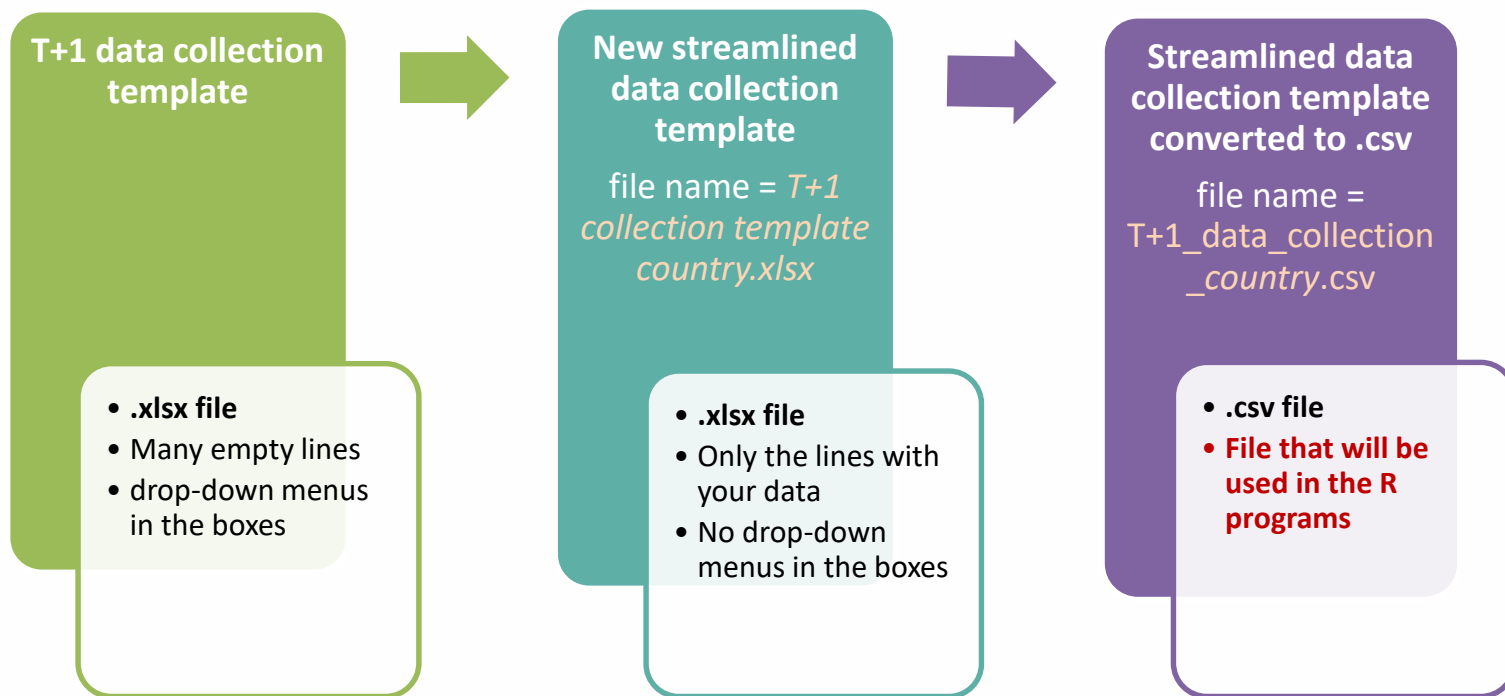
Please note! If you open this file again, you will have to do the same operation again. The numbers are automatically converted to scientific format when opening a .csv file.



Preliminary steps

1st preliminary step : preparation of the T+1 collection template

Summary of the manipulations in the 1st preliminary step





Preliminary steps

2nd preliminary step : preparation of the pre-existing data template

- You must **make a copy** of your file containing your pre-existing data that has been reclassified into the Best-ReMaP nomenclature. You must save this copy in **.csv format** under the name **pre_existing_data_country.csv** (with the name of your own country)
- You can follow the procedure on **pages [19 to 24](#)** for creating the copy in .csv format.
- You should also follow the procedure on **pages [24 to 28](#)** each time you open the file “pre_existing_data_country.csv” so that you do not lose any barcode information





Preliminary steps

2nd preliminary step : preparation of the pre-existing data template

Summary of the manipulations in the 2nd preliminary step

Your file containing your pre-existing data that has been reclassified into the Best-ReMaP nomenclature

- .xlsx file



File containing your pre-existing data converted to .csv

file name =
`pre_existing_data_country.csv`

- .csv file
- **File that will be used in the R programs**
- **This file will only be used in the 'Verifications template_step_4' program to check the matching of the father products**





Preliminary steps

3rd preliminary step : creation of the working folder

Title	Size	State	Author	Date
T+1_statistics_programs.zip	113.640 KB		Laure Bart	19 Oct 2022, 17:47

➤ You must download the zip folder called 'T+1_statistics_program.zip' from the Best-ReMaP intranet and copy it on your desktop
https://portal.nijz.si/ssf/a/c/p_name/ss_forum/p_action/1/binderId/21932/entityType/folder/action/view_permalink/novl_url/1

➤ You must unzip this folder before using it



Preliminary steps

3rd preliminary step : creation of the working folder

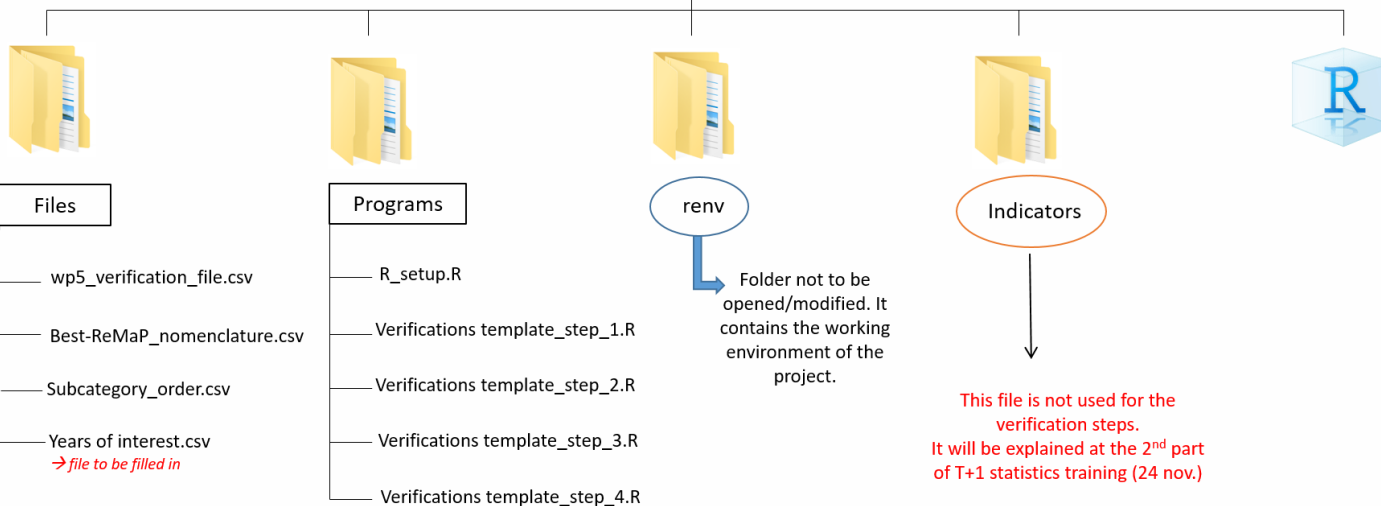
Please copy this file directly on your desktop →



T+1_statistics_programs

T+1_statistics_programs.Rproj

= This is the project you need to open for the task. All the work will be done here.



Copy here :

- your file with your collected and encoded data :
→ T+1_data_collection_country.csv
- your file with your preexisting data encoded :
→ pre_existing_data_country.csv

Files must be saved in .csv format (comma delimited)

It is very important that you don't change any names of the files that are listed here (except the file with your collected and encoded data)





Preliminary steps

4th preliminary step : preparation of the file "Years of interest.csv"

- In the working folder **T+1_statistics_programs** that you copied to your desktop in the 3rd preliminary step, there is a file called "**Years of interest.csv**" in the folder 'Files'.
- Before starting the task 5.4.1, we asked you to select which **years of each category** of your pre-existing data **will be used for pairing + comparisons/creation of indicators** with the data collected during Best-ReMaP. The selected years are the ones we called "years of interest".
- You will therefore need to edit the file "**Years of interest.csv**" manually by entering the years of interest (those that you have selected) of your pre-existing data for each Best-ReMaP category.
- Once you have modified this file, you should save the changes.
This file will be used in the 4th data verification step ([page 166](#))

see the following slide for an example





Preliminary steps

4th preliminary step : preparation of the file "Years of interest.csv"

The screenshot shows a Windows File Explorer window with the following details:

- Address bar: Ce PC > Bureau > T+1_statistics_programs > Files
- Search bar: Rechercher dans : Files
- Table of files:

Nom	Modifié le	Type	Taille
Best-ReMaP_nomenclature.csv	12/07/2022 16:48	Fichier CSV Micro...	34 Ko
pre_existing_data_France.csv	03/10/2022 15:52	Fichier CSV Micro...	334 Ko
T+1_data_collection_France.csv	10/10/2022 17:10	Fichier CSV Micro...	1 431 Ko
wp5_verification_file.csv	30/08/2022 10:29	Fichier CSV Micro...	15 Ko
Years of interest.csv	03/10/2022 17:27	Fichier CSV Micro...	1 Ko

The file 'Years of interest.csv' is highlighted in blue. A green text box with a blue arrow pointing to the file contains the following text:

The file to be modified named "Years of interest.csv" is located in the 'Files' folder in the T+1_statistics_programs folder that must have been copied to your desktop.





Preliminary steps

4th preliminary step : preparation of the file "Years of interest.csv"

	A	B	C	D
1	Category_code	Category_name	Year_of_interest_1	Year_of_interest_2
2		18 Bread products		
3		1 Breakfast cereals		
4		5 Delicatessen meats and similar		
5		3 Fresh dairy products and desserts		
6		9 Soft drinks		
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				

Overview of the file "Years of interest.csv"

- Column '**Year_of_interest_1**': you must enter for each category the selected year of interest from your "**pre_existing_data_country.csv**" file.
(other years of collection for the category will not be taking into account during the creation of indicators)
- Column '**Year_of_interest_2**': you must enter the other year of interest of the category if there is a second year of interest for the category (a collection may have occurred on two consecutive years). If there is only one year of interest for the category, it should be repeated in this column.
- If you do not have pre-existing data for one or more of the categories monitored in Best-ReMaP, you must delete the row(s) for that/those category(ies).



Preliminary steps

4th preliminary step : preparation of the file "Years of interest.csv"

	A	B	C	D	E
1	Category_code	Category_name	Year_of_interest_1	Year_of_interest_2	
2	18	Bread products	2017	2017	
3	1	Breakfast cereals	2016	2016	
4	5	Delicatessen meats and similar	2017	2017	
5	9	Soft drinks	2015	2016	
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					

Example

From your "pre_existing_data_ *country*.csv" file:

- You have selected **2017** as the year of interest for the categories : *Bread products, Delicatessen meats and similar*.
- You have selected **2016** as the year of interest for the category : *Breakfast cereals*
- You have selected **2015 and 2016** as years of interest for the category : *Soft drinks*. Indeed, products from this category were collected from December 2015 to January 2016.
- You have no pre-existing data for the sub-category *Fresh dairy products and desserts*, so the row has been deleted.



Preliminary steps

4th preliminary step : preparation of the file "Years of interest.csv"

	A	B	C	D	E
1	Category_code	Category_name	Year_of_interest_1	Year_of_interest_2	
2		18 Bread products			
3		1 Breakfast cereals			
4		5 Delicatessen meats and similar			
5		3 Fresh dairy products and desserts			
6		9 Soft drinks			
7		41 Baby food			
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					

File "Years of interest.csv"

- ❖ This file is a **key file** for the rest of the programs as it allows the identification of the categories for which indicators will be generated. If a category for which you have pre-existing data is forgotten in this file, then it will not be taken into account for the creation of indicators.
(see next training on 24 November)
- ❖ If you have other categories that are not monitored in Best-ReMaP but for which you wish to generate indicators, you must enter this category in the file. **But these indicators will not appear in the Best-ReMaP reports.**

Example:

You wish to generate indicators for another category which is *Baby food (41)*, you must enter the category_code and category_name + the years of interest of your pre-existing data for this category



2) Installation of the necessary equipment and presentation of the Rstudio software

A. Preliminary steps

B. Installation of software

C. Introduction to R studio

D. Cleaning of the Rstudio interface



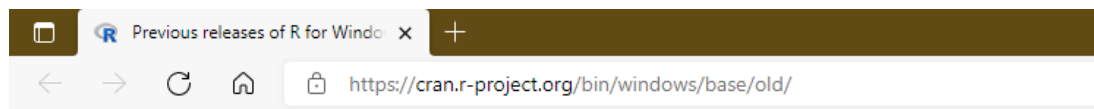


Installation of R software

- **Download the R software:**

Follow this link and select **version 4.1.2** (which is not the latest version but the version on which the programs were created):

<https://cran.r-project.org/bin/windows/base/old/>



Previous Releases of R for Windows

This directory contains previous binary releases of R for Windows.

The current release, and links to development snapshots, are available [here](#). Source code for these releases and other

In this directory:

- [R 4.2.0](#) (April, 2022)
- [R 4.1.3](#) (March, 2022)
- [R 4.1.2](#) (November, 2021)
- [R 4.1.1](#) (August, 2021)
- [R 4.1.0](#) (May, 2021)
- [R 4.0.5](#) (March, 2021)
- [R 4.0.4](#) (February, 2021)
- [R 4.0.3](#) (October, 2020)
- [R 4.0.2](#) (June, 2020)
- [R 4.0.1](#) (June, 2020)
- [R 4.0.0](#) (April, 2020)
- [R 3.6.3](#) (February, 2020)
- [R 3.6.2](#) (December, 2019)
- [R 3.6.1](#) (July, 2019)
- [R 3.6.0](#) (April, 2019)

Version of R to download (click on the link)

If you already have R on your computer, check which version of the software you have.

If it is a version **other than 4.1.2**, then you need to download version **4.1.2** as shown.

→ You will have **2 versions of R** on your computer.

see the following slide for next step



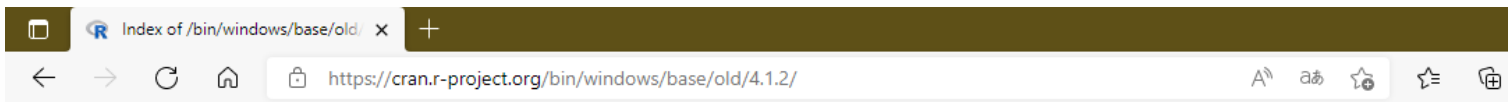


Installation of R software

- **Download the R software:**

Follow this link and select **version 4.1.2** (which is not the latest version but the version on which the programs were created):

<https://cran.r-project.org/bin/windows/base/old/>



Index of /bin/windows/base/old/4.1.2

Name	Last modified	Size	Description
Parent Directory	-		
NEWS.R-4.1.2.html	2021-11-01 19:14	105K	
R-4.1.2-win.exe	2021-11-01 20:30	86M	
README.R-4.1.2	2021-11-01 19:14	8.5K	
SVN-REVISION.R-4.1.2	2021-11-01 19:14	46	
md5sum.txt	2021-11-01 20:30	50	
release.html	2021-11-01 19:14	90	
rw-FAQ.html	2021-11-01 19:14	99K	

- Click to download this .exe file.
- Once you have downloaded this file, you can open it and click on 'Run'.
- The R software will then be installed on your computer.

Apache Server at cran.r-project.org Port 443

During the software installation, accept all the basic settings by clicking 'next' at each step





Installation of R software

Tutorial video to download and install the R software

→ *This video is available at any time on the Best ReMaP intranet in the WP5 section*
https://portal.nijz.si/ssf/a/c/p_name/ssf_forum/p_action/1/entityType/folderEntry/action/view_permalink/entryId/77697/novl_url/1





Installation of Rstudio software

- **Download the Rstudio software:**

Follow this link :

<https://www.rstudio.com/products/rstudio/download/>

Join us at **rstudio::conf(2022)** to sharpen your R skills. | July 25-28th in D.C. [LEARN MORE](#)

R Studio

DOWNLOAD SUPPORT DOCS COMMUNITY

Products Solutions Customers Resources About Pricing

Download the RStudio IDE

Choose Your Version

The RStudio IDE is a set of integrated tools designed to help you be more productive with R and Python. It includes a console, syntax-highlighting editor that supports direct code execution, and a variety of robust tools

R Studio Team

RStudio's recommended professional data science solution for every team. RStudio

When you are on this page, scroll down to find the links to all the Rstudio installers (see next slide)





Installation of Rstudio software

Download the RStudio IDE - RSti... x

https://www.rstudio.com/products/rstudio/download/

All Installers

Linux users may need to [import RStudio's public code-signing key](#) prior to installation, depending on the operating system's security policy.

RStudio requires a 64-bit operating system. If you are on a 32 bit system, you can use an [older version of RStudio](#).

OS	Download	Size	SHA-256
Windows 10/11	RStudio-2022.02.2-485.exe	177.27 MB	74187a33
macOS 10.15+	RStudio-2022.02.2-485.dmg	217.09 MB	cda82e98
Ubuntu 18+/Debian 10+	rstudio-2022.02.2-485-amd64.deb	128.58 MB	508a6e9c
Fedora 19/Red Hat 7	rstudio-2022.02.2-485-x86_64.rpm	144.66 MB	7400234c
Fedora 34/Red Hat 8	rstudio-2022.02.2-485-x86_64.rpm	144.70 MB	ad00e2c5
Debian 9	rstudio-2022.02.2-485-amd64.deb	128.91 MB	8edcf614
OpenSUSE 15	rstudio-2022.02.2-485-x86_64.rpm	129.31 MB	e54c69f6

- Click to download this .exe file.
- Once you have downloaded this file, you can open it and click on 'Run'.
- The Rstudio software will then be installed on your computer.

During the software installation, accept all the basic settings by clicking 'next' at each step





Installation of Rstudio software

Tutorial video to download and install the Rstudio software

→ *This video is available at any time on the Best ReMaP intranet in the WP5 section*
https://portal.nijz.si/ssf/a/c/p_name/ss_forum/p_action/1/entityType/folderEntry/action/view_permalink/entryId/77698/novl_url/1





Overview of R et Rstudio interfaces



'R'

- No processing will be done on this interface
- Software needed to be able to work on Rstudio



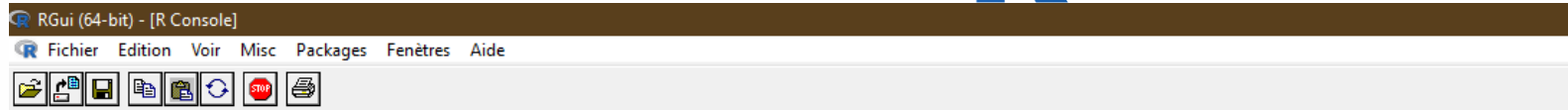
'Rstudio'

- Interface that allows the software to be used = environment that facilitates input, code execution and visualisation of results
- Programs will be running through RStudio





Overview of R



```
R version 4.1.2 (2021-11-01) -- "Bird Hippie"
Copyright (C) 2021 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R est un logiciel libre livré sans AUCUNE GARANTIE.
Vous pouvez le redistribuer sous certaines conditions.
Tapez 'license()' ou 'licence()' pour plus de détails.

R est un projet collaboratif avec de nombreux contributeurs.
Tapez 'contributors()' pour plus d'information et
'citation()' pour la façon de le citer dans les publications.

Tapez 'demo()' pour des démonstrations, 'help()' pour l'aide
en ligne ou 'help.start()' pour obtenir l'aide au format HTML.
Tapez 'q()' pour quitter R.

> |
```

Overview of the R software when
you open it

This is just for information
purposes as you will not be
working on this software but on
the **Rstudio interface**.



2) Installation of the necessary equipment and presentation of the Rstudio software

A. Preliminary steps

B. Installation of software

C. Introduction to R studio

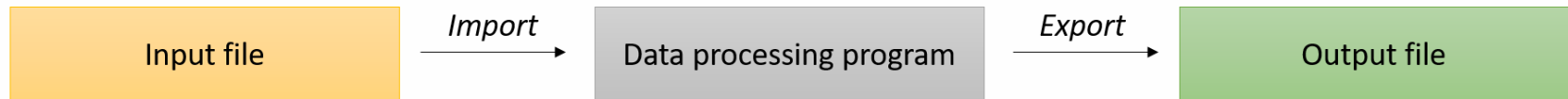
D. Cleaning of the Rstudio interface





Introduction to Rstudio software

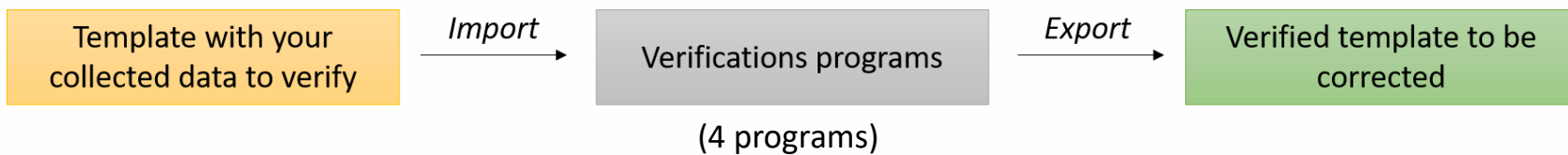
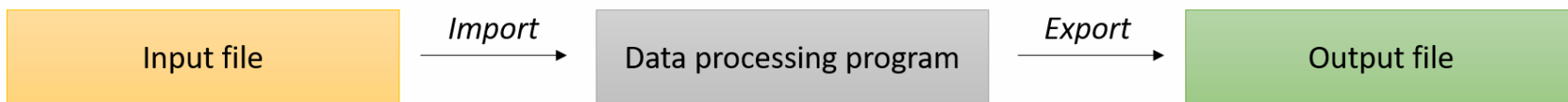
How does data processing software like Rstudio work?





Introduction to Rstudio software

How will the Rstudio data processing software work in Task 5.3.2?





WORK Package 5 – Reformulation and processed food monitoring

Introduction to Rstudio software

Overview of the Rstudio software when you open it for the first time.

The screenshot displays the RStudio interface with the following components:

- Console:** Shows the R version (4.1.2), copyright information, and project loading details. A callout box highlights this area.
- Environment:** Shows the current environment is empty. A callout box explains its function.
- Files:** Shows the file explorer with a list of files and folders in the project directory.

Name	Size	Modified
..		
.RData	6.2 KB	May 19, 2021
.Rhistory	29 KB	Oct 11, 2022
.Rprofile	25 B	May 4, 2022
Files		
Indicators		
Programs		
renv		
renv.lock	36.4 KB	May 4, 2022
T+1_statistics_programs.Rproj	218 B	Oct 11, 2022

« **Console** » that displays information about a program running, where you can see the execution of a program

« **Environment** » that displays the "working tools" (variables, tables, functions) created by the program





Introduction to Rstudio software

How to open a program in Rstudio

The screenshot shows the RStudio interface with the 'File' menu open. The 'Open File...' option is highlighted, and a blue callout box points to it with the text: "Path to the folder where the R program you want to open is located". Below the menu, a file explorer window is open, showing the path: `Users > t.jaguttin > Desktop > T+1_statistics_programs`. The file explorer lists several files and folders:

Name	Size
..	
.RData	6.2 KB
.Rhistory	30.9 KB
.Rprofile	25 B
Files	
Indicators	
Programs	
renv	
renv.lock	36.4 KB
T+1_statistics_programs.Rproj	218 B





WORK Package 5 – Reformulation and processed food monitoring

Introduction to Rstudio software

Rstudio when a program is opened

The screenshot displays the RStudio interface with the following components:

- Source Editor:** Contains an R script file named "Verifications template_step_1.R". The script includes comments and code for loading packages and setting parameters. A red box highlights the script content, with a red arrow pointing to the title bar and a red text box explaining it is the "script" of the program.
- Environment Pane:** Located on the right, it shows the "Global Environment" and states "Environment is empty". A green box highlights this pane with the text "« Environment » is still here".
- Console:** Located at the bottom, it shows the R version (4.1.2) and copyright information. A teal box highlights the console area with the text "« Console » appears here".

```
1 ##### FIRST PROGRAMME OF VERIFICATION
2
3 #Version of R used : 4.1.2
4 #Rstudio needs to be used for using the program in a simpler way : it is a tool for working on R programs that gi
5 ###IMPORTANT : R software must be installed on your computer
6
7
8 ##dplyr : 1.0.7
9 ##tidyverse : 1.3.1
10
11 #Load the necessary packages each time you open Rstudio f
12 library("dplyr")
13 library("tidyverse")
14
15
16 #SET PARAMETERS :
17 #You will need to personalize the variables below with yo
18
19 ##Input file (csv) name :
20 #Indicate ("T+1_data_collection_France.csv") - here you n
21 input_file="T+1_data_collection_France.csv"
22
23 ##Verification file (csv) name :
24 ##The csv file should be saved in the same folder as the one containing your data set in csv
25 verification_file="wp5_verification_file.csv"
26
27
```





Introduction to Rstudio software

The screenshot shows the RStudio interface with a script file named 'Verifications template_step_1.R'. The script contains the following code:

```
1 ##### FIRST PROGRAMME OF VERIFICATION
2
3 #Version of R used : 4.1.2
4 #Rstudio needs to be used for using the program in a simpler way : it is a tool for working on R programs that gi
5 ###IMPORTANT : R software must be installed on your computer in order to install, run and use Rstudio
6
7
8 ##dplyr : 1.0.7
9 ##tidyverse : 1.3.1
10
11 #Load the necessary packages each time you open Rstudio for running the program with the functions below
12 library("dplyr")
13 library("tidyverse")
14
15
16 #SET PARAMETERS :
17 #You will need to personalize the variables below with your parameters
18
19 ##Input file (csv) name :
20 #Indicate ("T+1_data_collection_France.csv") - here you need to replace "France" by
21 input_file="T+1_data_collection_France.csv"
22
23 ##Verification file (csv) name :
24 ##The csv file should be saved in the same folder as the one containing your data se
25 verification_file="wp5_verification_file.csv"
26
27
```

The callout box, titled "Description of a program", contains the following text:

- Lines preceded by a # and appearing all in green are comments in the program. These lines are there to describe the program and will not be run in the program.
- The other lines without a # are the **commands** that will run in the program.





WORK Package 5 – Reformulation and processed food monitoring

Introduction to Rstudio software

The screenshot displays the RStudio interface. The main editor window shows an R script with the following content:

```
1 ##### FIRST PROGRAMME OF VERIFICATION
2
3 #Version of R used : 4.1.2
4 #RStudio needs to be installed on R programs that give
5 ##### IMPORTANT :
6
7
8 ##dplyr : 1.0.7
9 ##tidyverse : 1.
10
11 #Load the necessary libraries
12 library("dplyr")
13 library("tidyverse")
14
15
16 #SET PARAMETERS
17 #You will need to
18
19 ##Input file (csv)
20 #Indicate ("T+1
21 input_file="T+1
22
23 ##Verification file
24 ##The csv file s
25 verification_fil
26
27
```

A red circle highlights the 'Save' icon in the top toolbar. A red arrow points from this icon to a text box containing the following instructions:

Modification of the program script

- For all programs you will have to make small changes (very few, only on the set parameters part).
- You can modify directly in the script and press the **save button** to save the modifications to the script.

The environment pane on the right shows 'Global Environment' and 'Environment is empty'. The console at the bottom displays the R version and platform information:

```
R 4.1.2 · C:\Users\t.laguitton\Desktop\T+1_statistics_programs/
R version 4.1.2 (2021-11-01) -- "Bird Hippie"
Copyright (C) 2021 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
```





Introduction to Rstudio software

The screenshot shows the RStudio interface with a script editor containing the following R code:

```
1 ##### FIRST PROGRAMME OF VERIFICATION
2
3 #Version of R used : 4.1.2
4 #Rstudio needs to be used for using the program in a simpler way : it is a tool for working on R programs that gi
5 ###IMPORTANT : R software must be installed on your computer in order to install, run and use RStudio
6
7
8 ##dplyr : 1.0.7
9 ##tidyverse : 1.3.1
10
11 #Load the necessary packages each time you open Rstudio for running the program with the functions below
12 library("dplyr")
13 library("tidyverse")
14
15
16 #SET PARAMETERS :
17 #You will need to personalize the variables below with your parameters
18
19 ##Input file (csv) name :
20 #Indicate ("T+1_data_collection_France.csv") - here you need to replace "France" by the name of
21 input_file="T+1_data_collection_France.csv"
22
23 ##Verification file (csv) name :
24 ##The csv file should be saved in the same folder as the one containing your data set in csv
25 verification_file="wp5_verification_file.csv"
26
27
```

The 'Run' button in the toolbar is circled in red. A callout box contains the following text:

How to run a command

To run a command, simply select it (as shown in the image) and click on the **'Run' button** (or Ctrl+Enter).

The console output at the bottom shows the R version and copyright information:

```
R 4.1.2 · C:/Users/tlaguittton/Desktop/T+1_statistics_programs/
R version 4.1.2 (2021-11-01) -- "Bird Hippie"
Copyright (C) 2021 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
```





Introduction to Rstudio software

```
1 ##### FIRST PROGRAMME OF VERIFICATION
2
3 #Version of R used : 4.1.2
4 #Rstudio needs to be used for using the program in a simpler way : it is a tool for working on R programs that gi
5 ###IMPORTANT : R software must be installed on your computer in order to install, run and use Rstudio
6
7
8 ##dplyr : 1.0.7
9 ##tidyverse : 1.3.1
10
11 #Load the necessary packages each time you open Rstudio for running the program v
12 library("dplyr")
13 library("tidyverse")
14
15
16 #SET PARAMETERS :
17 #You will need to personalize the variables below with your parameters
18
19 ##Input file (csv) name :
20 #Indicate ("T+1_data_collection_France.csv") - here you need to replace "France"
21 input_file="T+1_data_collection_France.csv"
22
23 ##Verification file (csv) name :
24 ##The csv file should be saved in the same folder as the one containing your data
25 verification_file="wp5_verification_file.csv"
26
27
```

Environment: Global Environment (153 MiB)

Environment is empty

Console: R 4.1.2 · C:/Users/t.laguillon/Desktop/T+1_statistics_programs/

R version 4.1.2 (2021-11-01) -- "Bird Hippie"
Copyright (C) 2021 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

renv.lock 36.4 KB
T+1_statistics_programs.Rproj 218 B

How to run an entire program

It is also possible to run all the commands at once. To do this, simply **select the entire program (Ctrl+A)** and click on **'Run'**. The comment lines (with # and in green) will not be run but all other commands will be run one by one.

This is what you will have to do for each program in this task! You will have to run the whole programs at once and not command by command.





Introduction to Rstudio software

The screenshot shows the RStudio environment with a script titled 'Verifications template_step_1.R'. The script contains comments and R code for loading packages and writing a CSV file. The console shows the execution of the code, including the loading of 'dplyr' and 'tidyverse' packages, and the execution of the 'write.csv2' function. The console output shows the file path and the message 'The file should appear in the folder 'Files''.

```
##### FIRST PROGRAMME OF VERIFICATION
#Version of R used : 4.1.2
#Rstudio needs to be used for using the program in a simpler way : it is a tool for working on R programs that gi
###IMPORTANT : R software must be installed on your computer in order to install, run and use Rstudio
##dplyr : 1.0.7
##tidyverse : 1.3.1
#Load the necessary packages each time you open Rstudio for running the pr
library("dplyr")
library("tidyverse")
#SET PARAMETERS :
#You will need to personalize the variables below with your parameters
##Input file (csv) name :
#Indicate ("T+1_data_collection_France.csv") - here you need to replace "F
<
2083:1 (Untitled) ↓
```

Console Terminal Jobs

```
R 4.1.2 · C:/Users/t.laguillon/Desktop/T+1_statistics_programs/Files/
+ utils::write.csv2(T0_data_f1,file= output_file,na=" ",row.names=FALSE)
+ }else if(separator==""){
+ utils::write.csv(T0_data_f1,file= output_file,na=" ",row.names=FALSE)
+ }
>
> ##utils::write.csv2 uses a comma for the decimal point and a semicolon for the
files in some western European locales.
> ##If the Excel convention is different in your country, please use the function
a_f1.2.csv",na=" ",sep="|",dec=".",row.names=FALSE)
>
> print("The file should appear in the folder 'Files'")
[1] "The file should appear in the folder 'Files'"
>
```

After running an entire program

- After running the entire program, comments are displayed in the console. In most cases, the messages in the console are displayed in **black** or **blue** and mean that the command has been completed.
- If a message appears in **red** or is preceded by a **'warning'**, it may mean that a command did not work and in this case it is probably a problem coming from an imported file. Either you can find the problem and correct it, or you can contact us for help.
- Note that the installation of packages displays a red message in the console, this is normal and is not an error.



WORK Package 5 – Reformulation and processed food monitoring

Introduction to Rstudio software

After running an entire program

- You can enlarge or shrink the console window using the quadruple arrow that appears when you move the mouse cursor over the top of the console.

```
##### FIRST PROGRAMME OF VERIFICATION
#Version of R used : 4.1.2
#Rstudio needs to be used for using the program in a simpler way : it is a
###IMPORTANT : R software must be installed on your computer in order to
##dplyr : 1.0.7
##tidyverse : 1.3.1
#Load the necessary packages each time you open Rstudio for running the pr
library("dplyr")
library("tidyverse")
#SET PARAMETERS :
#You will need to personalize the variables below with your parameters
##Input file (csv) name :
#Indicate ("T+1_data_collection_France.csv") - here you need to replace "France" by the name of your own country
2083:1 (Untitled) R Script
```

```
R 4.1.2 · C:/Users/t.laguitton/Desktop/T+1_statistics_programs/Files/
+ utils::write.csv2(T0_data_f1,file= output_file,na=" ",row.names=FALSE)
+ }else if(separator==""){
+   utils::write.csv(T0_data_f1,file= output_file,na=" ",row.names=FALSE)
+ }
>
> ##utils::write.csv2 uses a comma for the decimal point and a semicolon for the separator, the Excel convention for csv
files in some western European locales.
> ##If the Excel convention is different in your country, please use the function write.csv(T0_data_f1,file="Test_T0_dat
a_f1.2.csv",na=" ",sep="|",dec=".",row.names=FALSE)
>
> print("The file should appear in the folder 'Files'")
[1] "The file should appear in the folder 'Files'"
>
```

Name	Size
..	
.RData	6.2 KB
.Rhistory	30.9 KB
.Rprofile	25 B
Files	
Indicators	
Programs	
renv	
renv.lock	36.4 KB
T+1_statistics_programs.Rproj	218 B





Introduction to Rstudio software

The screenshot shows the RStudio interface with the following components:

- Script Editor:** Contains R code for a verification program. The code includes comments like "#Version of R used : 4.1.2" and "#Rstudio".
- Environment Pane:** Shows the current environment with variables: line_num (2255 obs. of 3 variables), Nomenclat... (618 obs. of 4 variables), T0_data (2310 obs. of 56 variables), T0_data_5 (2310 obs. of 134 variables), T0_data_f1 (2310 obs. of 50 variables), and Verificat... (119 obs. of 8 variables). Below this, a table of values is shown for variables like bar_code_..., carbohydr..., and empty_line.
- Files Pane:** Shows the file structure of the project, including folders like .RData, .Rhistory, .Rprofile, Files, Indicators, Programs, renv, and renv.lock, and files like T+1_statistics_programs.Rproj.
- Console:** Shows the output of the R script, including the message "The file should appear in the folder 'Files'" and the file path "C:/Users/.../T+1_statistics_programs/Files/".

After running an entire program

- Other elements may appear after running the entire program.
 - ❑ Tables in tabs next to the program script. These tables are working tools for the program and you don't need to look at them.
 - ❑ These program tables/tools also appear in the environment, which is no longer empty at the end of the program run.
- **The program outputs that you have to work on will automatically be saved as excel files in the 'Files' folder of the 'T+1_statistics_programs' folder.**





2) Installation of the necessary equipment and presentation of the Rstudio software

A. Preliminary steps

B. Installation of software

C. Introduction to R studio

D. Cleaning of the Rstudio interface





WORK Package 5 – Reformulation and processed food monitoring

'Cleaning' of Rstudio between each program

Table Data:

Product_code	Father_product_code	Country	Year	Category_name	Subcategory_name	Bar_code	Ass
1	1650	NA	Ireland	2021	Breakfast cereals	Cereals without added sugar	4088600257617
2	1651	NA	Ireland	2021	Breakfast cereals	Traditional muesli flakes	4088600052724
3	1653	NA	Ireland	2021	Breakfast cereals	Cereals without added sugar	
4	1654	NA	Ireland	2021	Breakfast cereals	Cereals without added sugar	
5	1655	NA	Ireland	2021	Breakfast cereals	Cereals without added sugar	
6	1657	NA	Ireland	2021	Breakfast cereals	Traditional muesli flakes	
7	1658	603	Ireland	2021	Breakfast cereals	High-fibre cereals	
8	1659	NA	Ireland	2021	Breakfast cereals	Chocolate-flavoured cereals	
9	1661	937	Ireland	2021	Breakfast cereals	Chocolate-flavoured cereals	
10	1663	586	Ireland	2021	Breakfast cereals	Sweet cereal flakes	
11	1665	916	Ireland	2021	Breakfast cereals	Crunchy fruit muesli	
12	1672	NA	Ireland	2021	Breakfast cereals	Crunchy muesli with nuts_se	
13	1674	597	Ireland	2021	Breakfast cereals	Sweet cereal flakes	
14	1675	NA	Ireland	2021	Breakfast cereals	Crunchy fruit muesli	

Console Output:

```
R 4.1.2 · C:/Users/tlaguilton/Desktop/T+1_statistics_programs/Files/ > ##utils::write.csv2 uses a comma for the decimal point and a semicolon for the separator, the Excel convention for csv files in some western European locales. > ###If the Excel convention is different in your country, please use the function write.csv(T0_data_f1,file="Test_T0_data_f1.2.csv",na=" ",sep="|",dec=",",row.names=FALSE) > > print("The file should appear in the folder 'Files'") [1] "The file should appear in the folder 'Files'" >
```

Environment Panel:

- Global Environment
- Data
 - line_num: 2255 obs. of 3 variables
- Indicators
- Programs
- renv
- renv.lock (36.4 KB)
- T+1_statistics_programs.Rproj (218 B)

Cleaning Rstudio interface (1)

Between each verification program and each run of a program, you must 'clean' the 'T+1_statistics_programs - Rstudio' interface.

First, you must delete all the tables/tools that have been created and the script of the program.

To do this, simply click on the small crosses at the top right of each tab.





WORK Package 5 – Reformulation and processed food monitoring

'Cleaning' of Rstudio between each program

The screenshot shows the RStudio interface with the following components:

- Console:** Contains R code for data cleaning, including functions for trimming spaces and writing CSV files.
- Environment:** Shows loaded objects like 'line_num', 'Nomenclat...', 'T0_data', etc. A red circle highlights the broom icon at the top right of this pane.
- Files:** A file explorer view at the bottom right showing the project directory structure.

```
R 4.1.2 · C:/Users/t.laguillon/Desktop/T+1_statistics_programs/Files/
$pb_5.b,T0_data_5$pb_6.a,T0_data_5$pb_6.b,T0_data_5$pb_7,T0_data_5$pb_8,T0_data_5$pb_9,T0_data_5$pb_10,T0_data_5$pb_11,T0_data_5$pb_12,T0_data_5$pb_13,T0_data_5$pb_14,T0_data_5$pb_15,T0_data_5$pb_16,T0_data_5$pb_17,T0_data_5$pb_18,T0_data_5$pb_19,T0_data_5$pb_20,T0_data_5$pb_21,T0_data_5$pb_22,T0_data_5$pb_23,T0_data_5$pb_24,T0_data_5$pb_25,T0_data_5$pb_26,T0_data_5$pb_27,T0_data_5$pb_28,T0_data_5$pb_29,T0_data_5$pb_30,T0_data_5$pb_31,T0_data_5$pb_32,T0_data_5$pb_33,T0_data_5$pb_34,T0_data_5$pb_35,T0_data_5$pb_36,T0_data_5$pb_37,T0_data_5$pb_38,T0_data_5$pb_39,T0_data_5$pb_40,T0_data_5$pb_41,T0_data_5$pb_42,sep=" ")
>
> #trim_string : function that will suppress all superfluous spaces
> trim_string<-function(string) base::gsub("\\s+", " ", base::gsub("^\\s+|\\s+$", "", string))
>
>
> #Apply the function trim_string on T0_data_5$Problems to remove the superfluous spaces
> T0_data_5$Problems<-trim_string(T0_data_5$Problems)
>
>
> #Add "/" between the different terms 'problem' (i.e. replace the spaces by /)
> T0_data_5$Problems<-base::gsub(" ", "/" ,T0_data_5$Problems)
>
>
> #Create the T0_data_f1 which regroups your data and the problems that needs to be cleaned
> "Problems"
> T0_data_f1<-dplyr::select(T0_data_5,"Product_code":"subcategory_code","bar_code","Problems")
> view(T0_data_f1)
>
> #Export the T0_data_f1
> if(separator==";"){
+   utils::write.csv2(T0_data_f1,file= output_file,na=" ",row.names=FALSE)
+ }else if(separator==""){
+   utils::write.csv(T0_data_f1,file= output_file,na=" ",row.names=FALSE)
+ }
>
> ###utils::write.csv2 uses a comma for the decimal point and a semicolon for the separator, the Excel convention for CSV
files in some western European locales.
> ###If the Excel convention is different in your country, please use the function write.csv(T0_data_f1,file="Test_T0_data_f1.2.csv",na=" ",sep="|",dec=".",row.names=FALSE)
>
> print("The file should appear in the folder 'Files'")
[1] "The file should appear in the folder 'Files'"
>
```

Cleaning Rstudio interface (2)

Then, you will need to clean the working environment.

To do this, you must click on the broom at the top right of the environment.





WORK Package 5 – Reformulation and processed food monitoring

'Cleaning' of Rstudio between each program

The screenshot shows the RStudio interface with a script being executed in the console. The script includes functions for trimming spaces and exporting data to CSV. A dialog box titled "Confirm Remove Objects" is displayed over the console, asking "Are you sure you want to remove all objects from the environment? This operation cannot be undone." The "Yes" button is circled in red, and the "Include hidden objects" checkbox is checked.

```
R 4.1.2 · C:/Users/tlaguitton/Desktop/T+1_statistics_programs/Files/
$pb_5.b,T0_data_5$pb_6.a,T0_data_5$pb_6.b,T0_data_5$pb_7,T0_data_5$pb_8,T0_data_5$pb_9,T0_data_5$pb_10,T0_data_5$pb_11,T0_data_5$pb_12,T0_data_5$pb_13,T0_data_5$pb_14,T0_data_5$pb_15,T0_data_5$pb_16,T0_data_5$pb_17,T0_data_5$pb_18,T0_data_5$pb_19,T0_data_5$pb_20,T0_data_5$pb_21,T0_data_5$pb_22,T0_data_5$pb_23,T0_data_5$pb_24,T0_data_5$pb_25,T0_data_5$pb_26,T0_data_5$pb_27,T0_data_5$pb_28,T0_data_5$pb_29,T0_data_5$pb_30,T0_data_5$pb_31,T0_data_5$pb_32,T0_data_5$pb_33,T0_data_5$pb_34,T0_data_5$pb_35,T0_data_5$pb_36,T0_data_5$pb_37,T0_data_5$pb_38,T0_data_5$pb_39,T0_data_5$pb_40,T0_data_5$pb_41,T0_data_5$pb_42,sep=" ")
>
> #trim_string : function that will suppress all superfluous spaces
> trim_string<-function(string) base::gsub("\\s+", " ", base::gsub("^\\s+|\\s+$", "", string))
>
> #Apply the function trim_string on T0_data_5$Problems to remove the superfluous spaces
> T0_data_5$Problems<-trim_string(T0_data_5$Problems)
>
>
> #Add "/" between the different terms 'problem' (i.e. replace spaces by slashes)
> T0_data_5$Problems<-base::gsub(" ", "/", T0_data_5$Problems)
>
> #Create the T0_data_f1 which regroups your data and the problems in one column
> T0_data_f1<-dplyr::select(T0_data_5,"Product_code":"Subcat")
> view(T0_data_f1)
>
> #Export the T0_data_f1
> if(separator==";"){
+   utils::write.csv2(T0_data_f1,file= output_file,na= " ",row.names=FALSE)
+ }else if(separator==""){
+   utils::write.csv(T0_data_f1,file= output_file,na=" ",row.names=FALSE)
+ }
>
> ##utils::write.csv2 uses a comma for the decimal point and a semicolon for the files in some western European locales.
> ##If the Excel convention is different in your country, please use the function write.csv2
> a_f1.2.csv",na=" ",sep="|",dec=".",row.names=FALSE)
>
> print("The file should appear in the folder 'Files'")
[1] "The file should appear in the folder 'Files'"
>
```

Confirm Remove Objects
Are you sure you want to remove all objects from the environment? This operation cannot be undone.
 Include hidden objects
Yes No

Cleaning Rstudio interface (3)
You need to click 'Yes' and make sure the box 'Include hidden objects' is checked.





WORK Package 5 – Reformulation and processed food monitoring

'Cleaning' of Rstudio between each program

```
R 4.1.2 · C:/Users/tlaguilton/Desktop/T+1_statistics_programs/Files/
> $pb_5.b,T0_data_5$pb_6.a,T0_data_5$pb_6.b,T0_data_5$pb_7,T0_data_5$pb_8,T0_data_5$pb_9,T0_data_5$pb_10,T0_data_5$pb_11,T0_data_5$pb_12,T0_data_5$pb_13,T0_data_5$pb_14,T0_data_5$pb_15,T0_data_5$pb_16,T0_data_5$pb_17,T0_data_5$pb_18,T0_data_5$pb_19,T0_data_5$pb_20,T0_data_5$pb_21,T0_data_5$pb_22,T0_data_5$pb_23,T0_data_5$pb_24,T0_data_5$pb_25,T0_data_5$pb_26,T0_data_5$pb_27,T0_data_5$pb_28,T0_data_5$pb_29,T0_data_5$pb_30,T0_data_5$pb_31,T0_data_5$pb_32,T0_data_5$pb_33,T0_data_5$pb_34,T0_data_5$pb_35,T0_data_5$pb_36,T0_data_5$pb_37,T0_data_5$pb_38,T0_data_5$pb_39,T0_data_5$pb_40,T0_data_5$pb_41,T0_data_5$pb_42,sep=" ")
>
> #trim_string : function that will suppress all superfluous spaces
> trim_string<-function(string) base::gsub("\\s+", " ", base::gsub("^\\s+|\\s+$", "", string))
>
>
> #Apply the function trim_string on T0_data_5$Problems to remove the superfluous spaces
> T0_data_5$Problems<-trim_string(T0_data_5$Problems)
>
>
> #Add "/" between the different terms 'problem' (i.e. replace the spaces by /)
> T0_data_5$Problems<-base::gsub(" ", "/", T0_data_5$Problems)
>
>
> #Create the T0_data_f1 which regroups your data and the problems that needs to be verified concatenated in one column
> "Problems"
> T0_data_f1<-dplyr::select(T0_data_5,"Product_code","Subcategory_code","bar_code_starts_0","Problems")
> view(T0_data_f1)
>
> #Export the T0_data_f1
> if(separator==";"){
+   utils::write.csv2(T0_data_f1,file= output_file,na=" ",row.names=FALSE)
+ }else if(separator==""){
+   utils::write.csv(T0_data_f1,file= output_file,na=" ",row.names=FALSE)
+ }
>
> ###utils::write.csv2 uses a comma for the decimal point and a semicolon for the
files in some western European locales.
> ###If the Excel convention is different in your country, please use the function
a_f1.2.csv",na=" ",sep="|",dec=".",row.names=FALSE)
>
> print("The file should appear in the folder 'Files'")
[1] "The file should appear in the folder 'Files'"
>
|
```

Cleaning Rstudio interface (4)

Then you have to clean the console by clicking on the transparent broom at the top right of the console.





WORK Package 5 – Reformulation and processed food monitoring

'Cleaning' of Rstudio between each program

The screenshot shows the RStudio interface for a project named 'T+1_statistics_programs'. The menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. The console shows the R version (4.1.2) and the current directory path. The Environment pane is empty, indicating a clean session. The Files pane shows the project structure with folders like .RData, .Rhistory, .Rprofile, Files, Indicators, Programs, renv, and renv.lock, along with the project file T+1_statistics_programs.Rproj.

Cleaning Rstudio interface (5)

Your 'T+1_statistics_programs - Rstudio' interface is cleaned up and ready for the next program.





WORK Package 5 – Reformulation and processed food monitoring

'Cleaning' of Rstudio between each program

The screenshot shows the RStudio interface with a 'Quit R Session' dialog box open. The dialog box asks: 'Save workspace image to C:/Users/t.laguillon/Desktop/T+1_statistics_programs/.RData?'. The 'Don't Save' button is circled in red. The background shows the RStudio environment with the 'Environment' pane indicating 'Environment is empty' and the 'Files' pane showing a list of files.

Name	Size
Data	6.2 KB
history	30.9 KB
profile	25 B
files	
indicators	
programs	
env	
env.lock	36.4 KB
+1_statistics_programs.Rproj	218 B

Cleaning Rstudio interface (6)

When you then close your Rstudio interface, Rstudio will offer to save the workspace image (this is different from saving the modified program script). Here you don't need to save the workspace so you can click on ***Don't Save***.





3) Running of the programs

A. Part 1 : R setup program ([page 78](#))

B. Part 2 : Verification programs and template cleaning/standardization ([page 88](#))

i. 1st verification program : 'Verifications template_step_1' ([page 90](#))

ii. 2nd verification program : 'Verifications template_step_2' ([page 119](#))

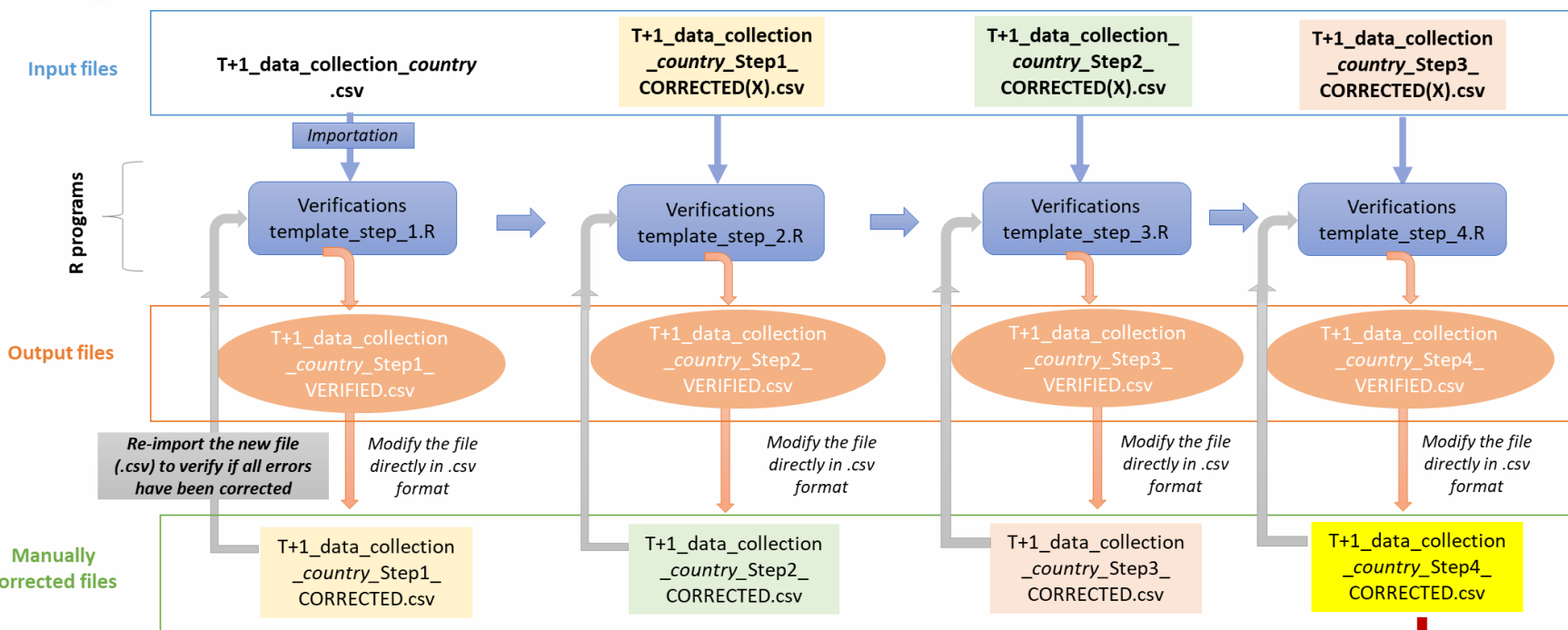
iii. 3rd verification program : 'Verifications template_step_3' ([page 144](#))

iv. 4th verification program : 'Verifications template_step_4' ([page 166](#))





Description of the different 'R' programs



→ All the R programs are already prepared and written, you will just need to make some minor parameter changes and run them.

→ The verification part with 4 programs is the longest but this part is essential to ensure the reliability of the indicators that will come out of the last program 'T+1_indicators' (next part of the training).

Final version of the template that must be :

copied in .xlsx to be sent to the JRC

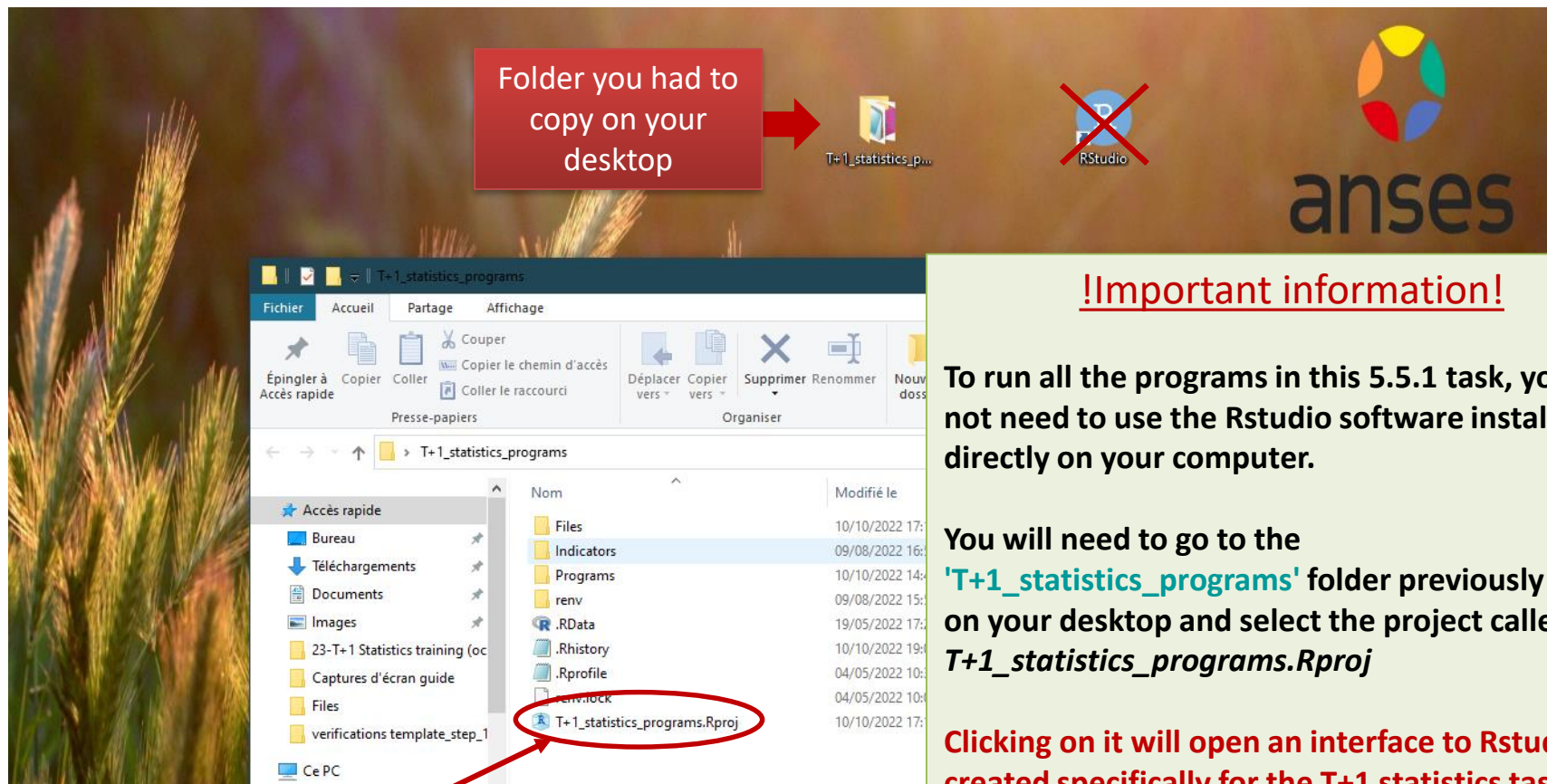
saved as .csv to be imported into the 'T+1_indicators' program



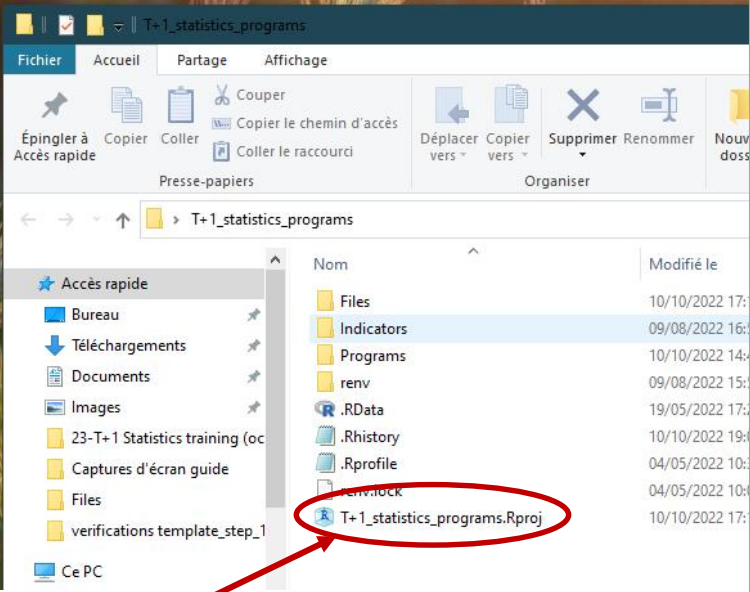


WORK Package 5 – Reformulation and processed food monitoring

Rstudio interface to use



Folder you had to copy on your desktop



Rstudio interface created specifically for the project that you always have to work with

!Important information!

To run all the programs in this 5.5.1 task, you will not need to use the Rstudio software installed directly on your computer.

You will need to go to the 'T+1_statistics_programs' folder previously saved on your desktop and select the project called *T+1_statistics_programs.Rproj*

Clicking on it will open an interface to Rstudio created specifically for the T+1 statistics task. It is very important that you run all your programs on this Rstudio interface.



WORK Package 5 – Reformulation and processed food monitoring

Rstudio interface to use

The screenshot shows the RStudio interface with the following components:

- Window Title:** T+1_statistics_programs - RStudio (highlighted with a red circle).
- Menu Bar:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Console:** Displays R version 4.1.2 (2021-11-01) -- "Bird Hippie" and project information for 'C:/Users/t.laguitton/Desktop/T+1_statistics_programs/'.
- Environment Pane:** Shows 'Global Environment' and 'Environment is empty'.
- Files Pane:** Lists files in the project directory, including 'renv.lock' and 'T+1_statistics_programs.Rproj'.

!Important information!

After opening the Rstudio interface named '*T+1_statistics_programs.Rproj*', you can check each time that you are in the correct working interface by checking that it says at the top left "**T+1_statistics_programs – Rstudio**"





WORK Package 5 – Reformulation and processed food monitoring

Rstudio interface to use

R version used by the Rstudio interface

You should also ensure that the **'T+1_statistics_programs - Rstudio'** project is working with **version 4.1.2** of R software by looking in the Rstudio console.

If the yellow highlighted message indicates a version of R other than 4.1.2, then you will need to perform the procedure **described in the following slides**.

If the yellow highlighted message indicates “R version 4.1.2”, then you can go directly to [page 78](#).





WORK Package 5 – Reformulation and processed food monitoring

Rstudio interface to use

The screenshot shows the RStudio interface with the 'Tools' menu open. The 'Global Options...' option is highlighted with a red circle. The console window displays the R version 4.1.2 (2021-11-01) and the project path 'C:/Users/t.laguitton/Desktop/T+1_statistics_programs/'.

R version used by the Rstudio interface

Go to the **Tools** tab and select **Global Options**

Files	Plots	Packages	Help	Viewer
New Folder	Delete	Rename	More	
C: > Users > t.laguitton > Desktop > T+1_statistics_programs				
Name	Size	Modified		
..				
.RData	6.2 KB	May 19, 2021		
.Rhistory	29 KB	Oct 11, 2022		
.Rprofile	25 B	May 4, 2022		
Files				
Indicators				
Programs				
renv				
renv.lock	36.4 KB	May 4, 2022		
T+1_statistics_programs.Rproj	218 B	Oct 11, 2022		





WORK Package 5 – Reformulation and processed food monitoring

Rstudio interface to use

R version used by the Rstudio interface

In the **General** section, click on **Change** next to the R version.





WORK Package 5 – Reformulation and processed food monitoring

Rstudio interface to use

R version 4.1.2 (2021-11-01) -- "Bird Hippie"
Copyright (C) 2021 The R Foundation for Statistics
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

* Project 'C:/Users/t.laguitton/Desktop/T+1_statistics_programs/'
[workspace loaded from C:/Users/t.laguitton/Desktop/T+1_statistics_programs/]

Options dialog box - R Sessions tab

Choose R Installation dialog box

RStudio requires an existing installation of R in order to work. Please select the version of R to use.

Use your machine's default version of R64 (64-bit)
 Use your machine's default version of R (32-bit)
 Choose a specific version of R:

- [64-bit] C:\Program Files\R\R-4.1.2
- [32-bit] C:\Program Files\R\R-4.1.2
- [64-bit] C:\PROGRA~1\R\R-41~1.2
- [32-bit] C:\PROGRA~1\R\R-41~1.2

Rendering Engine: Auto-detect (recommended)

Plots Packages Help Viewer

Name	Size	Modified
..		
.RData	6.2 KB	May 19, 2022
.Rhistory	29 KB	Oct 11, 2022
.Rprofile	25 B	May 4, 2022
Files		
Indicators		
Programs		
renv		
renv.lock	36.4 KB	May 4, 2022
T+1_statistics_programs.Rproj	218 B	Oct 11, 2022

R version used by the Rstudio interface

Click on '**Choose a specific version of R:**' and select the **version of R 4.1.2 [64-bit]** present on your computer and then press **OK**.

(If you have any doubts about which version of R to choose at this point as there are several 4.1.2 versions on offer, please feel free to send us a screenshot of the proposals made and we will help you choose)



WORK Package 5 – Reformulation and processed food monitoring

Rstudio interface to use

The screenshot shows the RStudio interface with the 'Options' dialog box open to the 'R Sessions' tab. The 'R version' field is set to '[Default] [64-bit] C:\Program Files\R\R-4.1.2'. A 'Change R Version' dialog box is overlaid on top, displaying the message: 'You need to quit and re-open RStudio in order for this change to take effect.' The 'OK' button in this dialog is circled in red.

R version used by the Rstudio interface

Rstudio will inform you that you need to restart Rstudio for the version change to take effect.

You can click **OK**, close Rstudio and reopen it (opening the '**T+1_statistics_programs.Rproj**' project as always).

Version 4.1.2 will be now displayed in the console and you will be able to proceed to the next steps described in the following slides.



3) Running of the programs

A. Part 1 : R setup program

B. Part 2 : Verification programs and template cleaning/standardization

i. 1st verification program : 'Verifications template_step_1'

ii. 2nd verification program : 'Verifications template_step_2'

iii. 3rd verification program : 'Verifications template_step_3'

C. Part 3 : Indicators and statistics production program





'R_setup' program

Presentation of the 'R_setup' program :

Just after installing the R and R studio software and before starting the verification programs, you will need to run the 'R_setup' program just once. This will allow the installation of packages that contain functions that will be needed for data verification and the creation of indicators.

Requirements before starting the program 'R_setup' :

- Before running the 'R_setup' program, you must ensure that you have **downloaded** the **R** and **Rstudio** software.
- You must also ensure that you are working on the Rstudio interface called ***T+1_statistics_programs.Rproj*** located in the 'T+1_statistics_programs' folder on your desktop (see 2nd preliminary step [pages 32-33](#))





WORK Package 5 – Reformulation and processed food monitoring

Running the 'R_setup' program

The screenshot shows the RStudio interface with the 'File' menu open. The 'Open File...' option is highlighted. In the background, a file explorer window is open to the 'T+1_statistics_programs' directory, showing files like '.RData', '.Rhistory', '.Rprofile', 'Files', 'Indicators', 'Programs', 'renv', 'renv.lock', and 'T+1_statistics_programs.Rproj'.

Environment is empty

How to open the 'R_setup' program





WORK Package 5 – Reformulation and processed food monitoring

Running the 'R_setup' program

The screenshot shows the RStudio interface. The 'Open File' dialog box is open, displaying the contents of the 'T+1_statistics_programs' folder. The 'Programs' folder is selected. The 'Global Environment' window shows an empty environment. A callout box highlights the text 'How to open the 'R_setup' program'.

Nom	Modifié le	Type	Taille
Files	09/08/2022 17:08	Dossier de fichiers	
Indicators	09/08/2022 16:57	Dossier de fichiers	
Programs	10/10/2022 14:48	Dossier de fichiers	
renv	09/08/2022 15:52	Dossier de fichiers	
.RData	19/05/2022 17:20	R Workspace	7 Ko
.Rhistory	10/08/2022 12:12	Fichier RHISTORY	31 Ko
.Rprofile	04/05/2022 10:39	Fichier RPROFILE	1 Ko
renv.lock	04/05/2022 10:01	Fichier LOCK	37 Ko
T+1_statistics_programs.Rproj	10/10/2022 14:49	R Project	1 Ko

Name	Size	Modified
..		
.RData	6.2 KB	May 19, 2021
.Rhistory	30.2 KB	Aug 10, 2022
.Rprofile	25 B	May 4, 2022
Files		
Indicators		
Programs		
renv		
renv.lock	36.4 KB	May 4, 2022
T+1_statistics_programs.Rproj	218 B	Oct 10, 2022





WORK Package 5 – Reformulation and processed food monitoring

Running the 'R_setup' program

The screenshot shows the RStudio interface with the 'Open File' dialog box open. The dialog is displaying the contents of the 'Programs' folder in the 'T+1_statistics_programs' directory. The file 'R_setup.R' is selected. The 'Environment' pane on the right shows 'Global Environment' and 'Environment is empty'. A green callout box contains the text 'How to open the 'R_setup' program'.

Nom	Modifié le	Type	Taille
.Rhistory	19/05/2022 17:20	Fichier RHISTORY	0 Ko
R_setup.R	04/05/2022 10:55	Fichier R	1 Ko
T+1_indicators.R	18/05/2022 12:57	Fichier R	209 Ko
Verifications template_step_1.R	10/10/2022 18:30	Fichier R	90 Ko
Verifications template_step_2.R	03/10/2022 17:47	Fichier R	124 Ko
Verifications template_step_3.R	03/10/2022 17:47	Fichier R	88 Ko
Verifications template_step_4.R	04/10/2022 11:06	Fichier R	24 Ko

Name	Size	Modified
..		
.RData	6.2 KB	May 19, 2021
.Rhistory	29 KB	Oct 11, 2022
.Rprofile	25 B	May 4, 2022
Files		
Indicators		
Programs		
renv		
renv.lock	36.4 KB	May 4, 2022
T+1_statistics_programs.Rproj	218 B	Oct 11, 2022



WORK Package 5 – Reformulation and processed food monitoring

Running the 'R_setup' program

The screenshot displays the RStudio interface with the 'R_setup.R' script open. The script contains the following code:

```
1 ##### SET-UP PROGRAMME
2 #IT IS NECESSARY TO RUN THIS PROGRAM BEFORE USING RSTUDIO FOR THE TASK
3
4 renv::restore(prompt=F)
5 print("Setup successfully completed, you can close the R-setup program and proceed to the next program")
6
```

The Environment pane on the right shows 'Global Environment' and 'Environment is empty'. A file explorer window is open at the bottom right, showing the directory structure of the project:

Name	Size	Modified
..		
.RData	6.2 KB	May 19, 2022
.Rhistory	29 KB	Oct 11, 2022
.Rprofile	25 B	May 4, 2022
Files		
Indicators		
Programs		
renv		
renv.lock	36.4 KB	May 4, 2022
T+1_statistics_programs.Rproj	218 B	Oct 11, 2022

A callout box with a green border and white background contains the text: 'R_setup' program opened.





WORK Package 5 – Reformulation and processed food monitoring

Running the 'R_setup' program

The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains the R script `R_setup.R` with the following code:

```
1 ##### SET-UP PROGRAMME
2 #IT IS NECESSARY TO RUN THIS PROGRAM BEFORE USING RSTUDIO FOR THE TASK
3
4 renv::restore(prompt=F)
5 print("Setup successfully completed, you can close the R-setup program and proceed to the next program")
6
```
- Environment Panel:** Shows "Environment is empty".
- Terminal Panel:** Shows the R prompt `>` at the bottom.
- File Explorer:** Shows the directory `C:\Users\t.laguillon\Desktop> T+1_statistics_programs` with the following files:

Name	Size	Modified
..		
.RData	6.2 KB	May 19, 2021
.Rhistory	29 KB	Oct 11, 2022
.Rprofile	25 B	May 4, 2022
Files		
Indicators		
Programs		
renv		
renv.lock	36.4 KB	May 4, 2022
T+1_statistics_programs.Rproj	218 B	Oct 11, 2022

Select the entire program (Ctrl+A)
and press 'Run' (or Ctrl+Enter)
The program will run entirely.





WORK Package 5 – Reformulation and processed food monitoring

Running the 'R_setup' program

Program running

The 'R_setup' program can take **quite a long time** (a few seconds to a few minutes) to run because it installs packages in the R software.

When the program is running and has not finished its execution, a **stop sign** appears at the top right of the console. This means that the program is running and you must not close the 'T+1_statistics_programs – Rstudio' window or the program will stop running.

```
1 ##### SET-UP PRO
2 #IT IS NECESSARY TO
3
4 renv::restore(prompt
5 print("setup succes
6
```

```
xt program")
```

Environment: Global Environment

Files | Plots | Packages | Help | Viewer

Name	Size
..	
.RData	125.9
.Rhistory	31.2
.Rprofile	25 B
Files	
Indicators	
Programs	
renv	
renv.lock	36.4
T0_statistics_programs.Rproj	218 B

Console Terminal Jobs

```
R 4.1.2 · C:/Users/t.laguitton/Desktop/cccc/
OK [downloaded 1.5 Mb in 0.9 secs]
Retrieving 'https://cloud.r-project.org/bin/windows/contrib/4.1/scales_1.2.0.zip' ...
OK [downloaded 601.3 kb in 0.7 secs]
Retrieving 'https://cloud.r-project.org/bin/windows/contrib/4.1/haven_2.5.0.zip' ...
OK [downloaded 1.3 Mb in 1 secs]
Retrieving 'https://cloud.r-project.org/bin/windows/contrib/4.1/knitr_1.39.zip' ...
```





WORK Package 5 – Reformulation and processed food monitoring

Running the 'R_setup' program

```
##### SET-UP PROGRAMME
#IT IS NECESSARY TO RUN THIS PROGRAM BEFORE USING RSTUDIO FOR THE TASK
3
4 renv::restore(prompt=F)
5 print("setup successfully completed, you can close the R-setup program tab and proceed to the next pro
6
```

End of 'R_setup' execution

At the end of the 'R_setup' program execution, you should see the **sentence highlighted in yellow** in the console.

This means that the program has run correctly and the packages have been **installed successfully**.

You can now close the tab with the 'R_setup' program script, clean up the console and move on to the next program (see next pages).

```
R 4.1.2 · C:/Users/t.laguiton/Desktop/cccc/
Installing knitr [1.39] ...
  OK [installed binary]
Moving knitr [1.39] into the cache ...
  OK [moved to cache in 13 milliseconds]
Installing ps [1.7.0] ...
  OK [installed binary]
Moving ps [1.7.0] into the cache ...
  OK [moved to cache in 27 milliseconds]
Installing rmarkdown [2.14] ...
  OK [installed binary]
Moving rmarkdown [2.14] into the cache ...
  OK [moved to cache in 40 milliseconds]
Installing testthat [3.1.4] ...
  OK [installed binary]
Moving testthat [3.1.4] into the cache ...
  OK [moved to cache in 13 milliseconds]
Installing uuid [1.1-0] ...
  OK [installed binary]
Moving uuid [1.1-0] into the cache ...
  OK [moved to cache in 42 milliseconds]
> print("setup successfully completed, you can close the R-setup program tab and proceed to the next program")
[1] "Setup successfully completed, you can close the R-setup program tab and proceed to the next program"
>
```





'R_setup' program

Tutorial video to run the Rsetup program

→ *This video is available at any time on the Best ReMaP intranet in the WP5 section*
https://portal.nijz.si/ssf/a/c/p_name/ssf_forum/p_action/1/entityType/folderEntry/action/view_permalink/entryId/77696/novl_url/1





3) Running of the programs

A. Part 1 : R setup program

B. Part 2 : Verification programs and template cleaning/standardization

i. 1st verification program : 'Verifications template_step_1' ([page 90](#))

ii. 2nd verification program : 'Verifications template_step_2' ([page 119](#))

iii. 3rd verification program : 'Verifications template_step_3' ([page 144](#))

iv. 4th verification program : 'Verifications template_step_4' ([page 166](#))





Verification programs

- In this part, you will run 4 verification programs on your file : [T+1_data_collection_country.csv](#)
- This part is very important because it will check that there are no input errors in the template and that the data can be used for creating indicators.

You will need to run these programs in the following order:

- 1) [Verifications template_step_1.R](#)
 - 2) [Verifications template_step_2.R](#)
 - 3) [Verifications template_step_3.R](#)
 - 4) [Verifications template_step_4.R](#)
- The only information that verification programs cannot check is the **accuracy** of the **Best-ReMaP category and subcategory classification** of the collected products. We therefore suggest that you carefully check the classification of your products collected before starting the verification programs.





3) Running of the programs

A. Part 1 : R setup program

B. Part 2 : Verification programs and template cleaning/standardization

i. 1st verification program : 'Verifications template_step_1'

ii. 2nd verification program : 'Verifications template_step_2'

iii. 3rd verification program : 'Verifications template_step_3'

iv. 4th verification program : 'Verifications template_step_4'





1st verification program : 'Verifications template_step_1'

Presentation of the 'Verifications template step 1' program :

- This is the first out of the 4 programs of verification of the data entered in your template.
- In this program, data entry problems are checked (misspelling of information, missing mandatory fields, ...)

Requirements before starting the program 'Verifications template step 1' :

- Before running the program, you need to make sure that a **copy** of your T+1 collection template is saved in .csv format in the folder '**Files**' in the '**T+1_statistics_programs**' folder that you have saved on your desktop.
- You need to make sure that your template have been renamed :
[T+1_data_collection_country.csv](#) (with the name of your own country)
- You need to make sure that the barcodes in your file [T+1_data_collection_country.csv](#) appear in full and not in scientific format (see procedure [pages 24→28](#))

Your Rstudio interface must have been cleaned up before running the program.

All cleaning steps are described on [pages 62→68](#).





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_1' program

The screenshot shows the RStudio interface with the 'File' menu open. The 'Open File...' option is highlighted. In the background, a Windows File Explorer window is open to the project directory 'C:\Users\t.laguitton\Desktop> T+1_statistics_programs'. The file list includes:

Name	Size	Modified
..		
.RData	6.2 KB	May 19, 2022
.Rhistory	30.2 KB	Aug 10, 2022
.Rprofile	25 B	May 4, 2022
Files		
Indicators		
Programs		
renv		
renv.lock	36.4 KB	May 4, 2022
T+1_statistics_programs.Rproj	218 B	Oct 10, 2022

A text box in the center of the screenshot reads: "How to open the 'Verifications template_step_1' program".





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_1' program

The screenshot shows the RStudio interface with the 'Open File' dialog box open. The dialog is displaying the contents of the 'T+1_statistics_programs' folder. The 'Programs' folder is selected. The 'Files' panel in the background shows the contents of the 'Programs' folder.

Nom	Modifié le	Type	Taille
Files	09/08/2022 17:08	Dossier de fichiers	
Indicators	09/08/2022 16:57	Dossier de fichiers	
Programs	10/10/2022 14:48	Dossier de fichiers	
renv	09/08/2022 15:52	Dossier de fichiers	
.RData	19/05/2022 17:20	R Workspace	7 Ko
.Rhistory	10/08/2022 12:12	Fichier RHISTORY	31 Ko
.Rprofile	04/05/2022 10:39	Fichier RPROFILE	1 Ko
renv.lock	04/05/2022 10:01	Fichier LOCK	37 Ko
T+1_statistics_programs.Rproj	10/10/2022 14:49	R Project	1 Ko

Environment is empty

Files Plots Packages Help Viewer

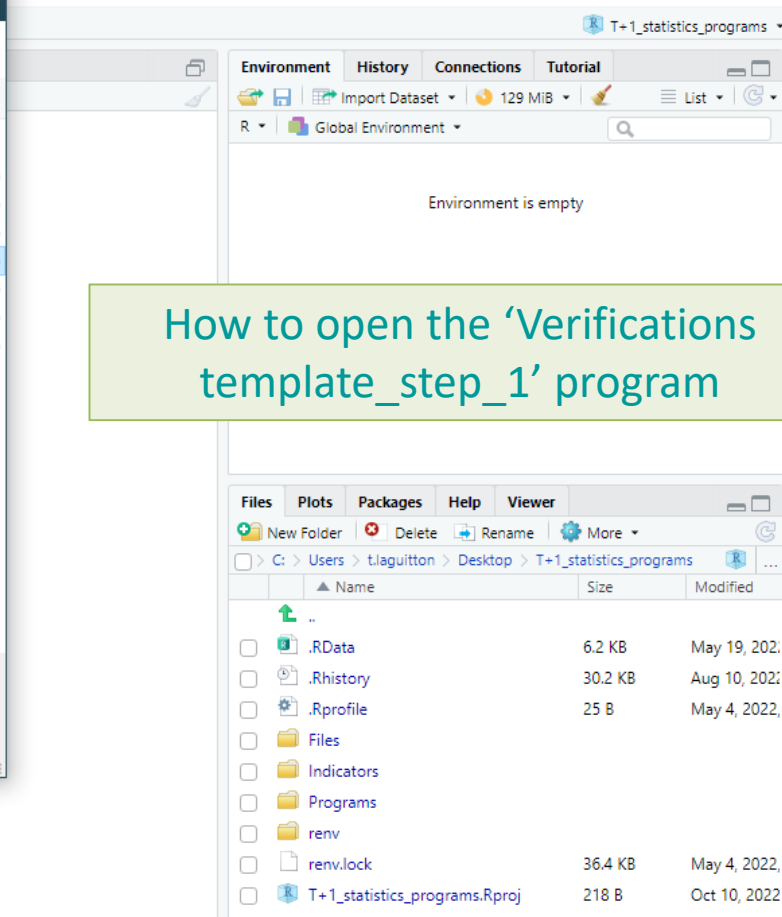
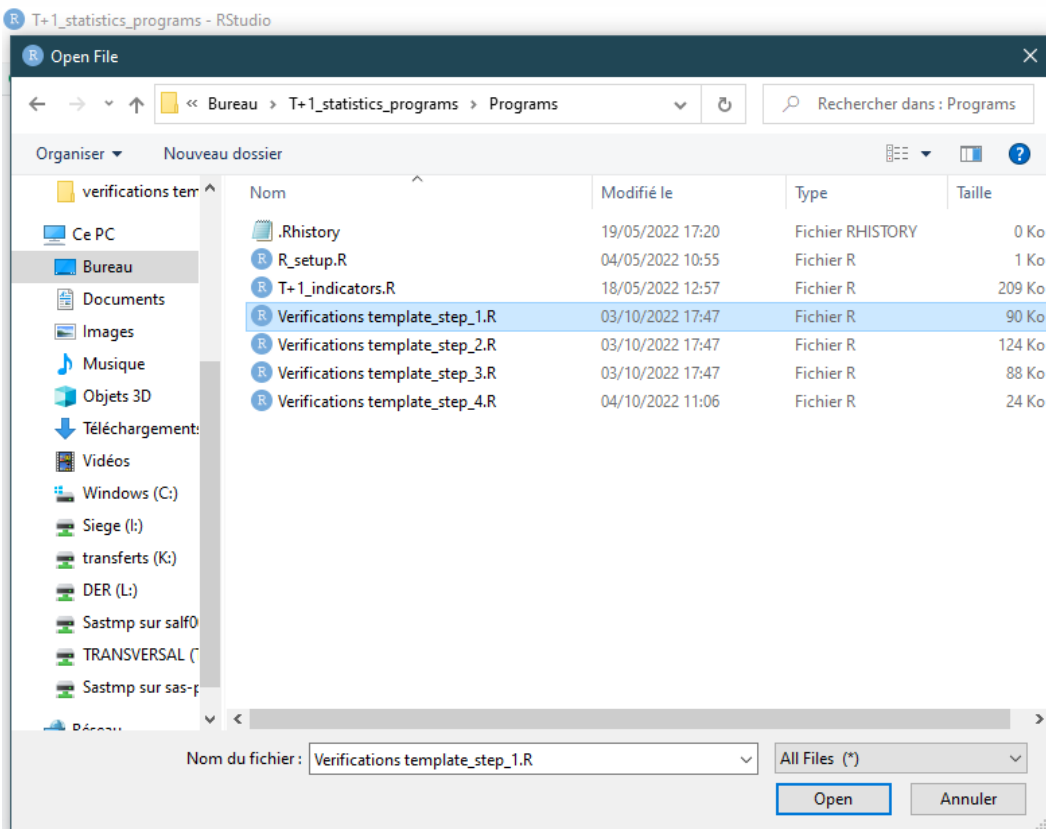
Name	Size	Modified
..		
.RData	6.2 KB	May 19, 2022
.Rhistory	30.2 KB	Aug 10, 2022
.Rprofile	25 B	May 4, 2022
Files		
Indicators		
Programs		
renv		
renv.lock	36.4 KB	May 4, 2022
T+1_statistics_programs.Rproj	218 B	Oct 10, 2022

How to open the 'Verifications template_step_1' program



WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_1' program



How to open the 'Verifications template_step_1' program



WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_1' program

The screenshot displays the RStudio interface with the 'Verifications template_step_1.R' script open. The script contains the following code:

```
1 ##### FIRST PROGRAMME OF VERIFICATION
2
3 #Version of R used : 4.1.2
4 #Rstudio needs to be used for using the program in a simpler way : it is a tool for working on R programs
5 ##IMPORTANT : R software must be installed on your computer in order to install, run and use Rstudio
6
7
8 ##dplyr : 1.0.7
9 ##tidyverse : 1.3.1
10
11 #Load the necessary packages each time you open Rstudio for running the program with the functi
12 library("dplyr")
13 library("tidyverse")
14 |
15
16 #SET PARAMETERS :
17 #You will need to personalize the variables below with your parameters
18
19 ##Input file (csv) name :
20 #Indicate ("T+1_data_collection_France.csv") - here you need to replace "France" by the name of your own
21 input_file="T+1_data_collection_France.csv"
22
23 ##Verification file (csv) name :
24 ##The csv file should be saved in the same folder as the one containing your data set in csv
25 verification_file="wp5_verification_file.csv"
26
27 ##Nomenclature file (csv) name :
28 ##The csv file should be saved in the same folder as the one containing your data set in csv
29 nomenclature_file="Best-ReMaP_nomenclature.csv"
30
31 ##Output file (csv) name :
32 output_file="T+1_data_collection_France_Step1_VERIFIED.csv"
33
34 ##Data collection year :
35 ##Enter the year at which you have collected your data for Best-ReMaP (replace "2022" with "your collecti
36 year="2021"
37
```

The Environment pane shows the Global Environment with an empty environment. A file explorer window is open, showing the directory structure of the project:

Name	Size	Modified
..		
.RData	6.2 KB	May 19, 2022
.Rhistory	30.2 KB	Aug 10, 2022
.Rprofile	25 B	May 4, 2022
Files		
Indicators		
Programs		
renv		
renv.lock	36.4 KB	May 4, 2022
T+1_statistics_programs.Rproj	218 B	Oct 10, 2022

'Verifications template_step_1' program opened



WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_1' program

```
T+1_statistics_programs - RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
Verifications template_step_1.R* x
Source on Save
13 library("tidyverse")
14
15
16 #SET PARAMETERS :
17 #You will need to personalize the variables below with your pa
18
19 ##Input file (csv) name :
20 #Indicate ("T+1_data_collection_France.csv") - here you need t
21 input_file="T+1_data_collection_France.csv"
22
23 ##Verification file (csv) name :
24 ##The csv file should be saved in the same folder as the one c
25 verification_file="wp5_verification_file.csv"
26
27 ##Nomenclature file (csv) name :
28 ##The csv file should be saved in the same folder as the one c
29 nomenclature_file="Best-ReMaP_nomenclature.csv"
30
31 ##Output file (csv) name :
32 output_file="T+1_data_collection_France_Step1_VERIFIED.csv"
33
34 ##Data collection year :
35 ##Enter the year at which you have collected your data for Bes
36 year="2021"
37 #Optional - if the data collection overlaps two consecutive ye
38 ##IF YOU HAVE ONLY ONE YEAR OF DATA COLLECTION, INDICATE THE S
39 year_2="2021"
40
41 ##Indicate the separator used for saving your Excel file in cs
42 separator=";"
43
44 ##If you have used a specific alphabet (greek, etc) in your te
45 special_alphabet="NO"
46
47 #####
48 #here=getwd() # Get the current working directory
49
31:11 (Top Level)
Console
```

Beginning of each verification program

At the beginning of each verification program, there is a set parameters part (line 16). It is only in this part that you should change information.

In the first run of the verification program 1, you need to change the name of the country with your own country name in the input file (line 21) and the output file (line 32) of the R script.

You must indicate the year in which your T+1 data was collected. If the collection was carried out in the same year, you must enter this year twice (line 36 and 39). If your collection took place in two different years, you must enter both years.

You also need to indicate the separator used in your csv file (line 42). You have the choice only between ";" and ",". In Europe, the most commonly used separator in the csv format is the ";" (pre-filled in the program).

You also need to indicate if you have used a specific alphabet in your template (line 45). This field only concerns countries with a specific alphabet and is therefore pre-filled as "NO".

Example :

- **input file** = "T+1_data_collection_Ireland.csv"
- **output file** = "T+1_data_collection_Ireland_Step1_VERIFIED.csv"
- **year** = "2021"
- **year_2** = "2022"
- **Separator** = ";"
- **Special_alphabet** = "NO"



WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_1' program

Select the entire program (Ctrl+A) and press 'Run' (or Ctrl+Enter)

The program will run entirely.

```
1 ##### FIRST PROGRAMME OF VERIFICATION
2
3 #Version of R used : 4.1.2
4 #Rstudio needs to be used for using the program in a simpler way : it is a tool for working on R programs
5 ##IMPORTANT : R software must be installed on your computer in order to install, run and use Rstudio
6
7
8 ##dplyr : 1.0.7
9 ##tidyverse : 1.3.1
10
11 #Load the necessary packages each time you open Rstudio for running the program with
12 library("dplyr")
13 library("tidyverse")
14
15
16 #SET PARAMETERS :
17 #You will need to personalize the variables below with your parameters
18
19 ##Input file (csv) name :
20 #Indicate ("T+1_data_collection_France.csv") - here you need to replace "France" by the name of your own
21 input_file="T+1_data_collection_France.csv"
22
23 ##Verification file (csv) name :
24 ##The csv file should be saved in the same folder as the one containing your data set in csv
25 verification_file="wp5_verification_file.csv"
26
27 ##Nomenclature file (csv) name :
28 ##The csv file should be saved in the same folder as the one containing your data set in csv
29 nomenclature_file="Best-ReMaP_nomenclature.csv"
30
31 ##Output file (csv) name :
32 output_file="T+1_data_collection_France_Step1_VERIFIED.csv"
33
34 ##Data collection year :
35 ##Enter the year at which you have collected your data for Best-ReMaP (replace "2022" with "your collection year")
36 year="2021"
37
```





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_1' program

The screenshot displays the RStudio interface. The top-left pane shows a data table with 13 rows and 7 columns. The top-right pane shows the Environment tab with several tables listed. The bottom-left pane shows the Console window with R code and its output.

Product_code	Father_product_code	Country	Year	Category_name	Subcategory_name	Bar_code
1	1692	NA	Ireland	2021	Fresh dairy products and desserts	Classic sweet y
2	1705	NA	Ireland	2021	Fresh dairy products and desserts	Classic sweet y
3	1706	NA	Ireland	2021	Fresh dairy products and desserts	Classic plain yo
4	1707	NA	Ireland	2021	Fresh dairy products and desserts	Gourmet plain
5	1708	NA	Ireland	2021	Fresh dairy products and desserts	Classic sweet y
6	1709	NA	Ireland	2021	Fresh dairy products and desserts	Gourmet plain
7	1727	NA	Ireland	2021	Fresh dairy products and desserts	Gourmet sweet
8	1732	550	Ireland	2021	Fresh dairy products and desserts	Gourmet sweet
9	1733	550	Ireland	2021	Fresh dairy products and desserts	Gourmet sweet
10	1736	NA	Ireland	2021	Fresh dairy products and desserts	Classic sweete
11	1746	NA	Ireland	2021	Fresh dairy products and desserts	Classic sweet y
12	1752	NA	Ireland	2021	Fresh dairy products and desserts	Fresh desserts
13	1758	NA	Ireland	2021	Fresh dairy products and desserts	Artificially-swe

```
R 4.1.2 · C:/Users/tlaguitton/Desktop/T+1_statistics_programs/Files/
+ }
>
> ##utils::write.csv2 uses a comma for the decimal point and a semicolon for the
n for CSV files in some western European locales.
> ##If the Excel convention is different in your country, please use the functi
st_T0_data_f1.2.csv",na=" ",sep="|",dec=".",row.names=FALSE)
>
> print("The file should appear in the folder 'Files'")
[1] "The file should appear in the folder 'Files'"
>
```

After running the program

Several tables appear next to the R program script. These tables are working tools for the R program but you don't need to look at them (only if there are errors or warnings in the console but in this case you have to tell us so we can help you).

At the end of the 'Verifications template_step_1' program execution, you should see the **sentence highlighted in yellow** in the console.

The program exports an excel file named T+1_data_collection_country_Step1_VERIFIED.csv to your "Files" folder in the "T+1_statistics_programs" folder.





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_1' program

The screenshot shows a Windows File Explorer window with the following details:

- Address bar: Ce PC > Bureau > T+1_statistics_programs > Files
- Files list:

Nom	Modifié le
Best-ReMaP_nomenclature.csv	12/07/2022 16
pre_existing_data_France.csv	03/10/2022 15
T+1_data_collection_France.csv	10/10/2022 15
T+1_data_collection_France_Step1_CORRECTED.csv	10/10/2022 15
T+1_data_collection_France_Step1_VERIFIED.csv	10/10/2022 15
wp5_verification_file.csv	30/08/2022 10
Years of interest.csv	03/10/2022 17

Create a copy of the file exported by R

In the 'Files' folder, you will find the file exported by the first verification program `T+1_data_collection_country_Step1_VERIFIED.csv`

→ You need to create a copy of this file and rename it “`T+1_data_collection_country_Step1_CORRECTED.csv`”

It is on this file that you will make the corrections and modifications indicated





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_1' program

Be careful to change the format of the 'Bar_code' column (see procedure [page 24-28](#)) when you are working on the file so that the barcodes appear in full and no information is lost.

Overview of the *T+1_data_collection_country_Step1_CORRECTED.csv* file which is a new version of your T+1 data collection template in which you will have to make modifications/corrections

Product	Category	Subcategory	Bar_code	Assortment	Brand_name	Brand_owner	Type_of_packaging	Legal_name	Legal_name	Commercial_name	Commercial_name	FOP_label	Nutri_Score	Ingredient	Net_weight	Net_weight	Number	
1658	603	Ireland	2021	Breakfast	High-fibre	4,09E+12	No	Harvest M	Aldi	Hard disc	Gluten Free	Gluten Free	Gluten Free	Gluten Free	None from	Gluten Free	500 g	
1659		Ireland	2021	Breakfast	Chocolate	4,09E+12	No	Harvest M	Aldi	Hard disc	Wholegrain	Wholegrain	High Ome	High Ome	Traffic light	Jumbo Oa	500 g	
1661	937	Ireland	2021	Breakfast	Chocolate	4,09E+12	No	Harvest M	Aldi							500 g		
1663	586	Ireland	2021	Breakfast	Sweet cereals	4,09E+12	No	Harvest M	Aldi							750 g		
1665	916	Ireland	2021	Breakfast	Crunchy flakes	4,09E+12	No	Harvest M	Aldi							1500 g		
1672		Ireland	2021	Breakfast	Crunchy flakes	4,09E+12	No	Harvest M	Aldi							500 g		
1674	597	Ireland	2021	Breakfast	Sweet cereals	4,09E+12	No	Harvest M	Aldi							500 g		
1675		Ireland	2021	Breakfast	Crunchy flakes	4,09E+12	No	Harvest M	Aldi							500 g		
1676		Ireland	2021	Breakfast	Crunchy flakes	4,09E+12	No	Harvest M	Aldi							500 g		
1677	937	Ireland	2021	Breakfast	Chocolate	4,09E+12	No	Harvest M	Aldi							500 g		
1680	996	Ireland	2021	Breakfast	Sweet cereals	4,09E+12	No	Harvest M	Aldi							375 g		
1681		Ireland	2021	Breakfast	Tradition	4,09E+12	No	Harvest M	Aldi							500 g		
1682	850	Ireland	2021	Breakfast	Tradition	4,09E+12	No	Harvest M	Aldi	Hard disc	Honey To	Honey To	Raisin & A	Raisin & A	Traffic light	WHOLEGR	1000 g	
1683		Ireland	2021	Breakfast	Tradition	4,09E+12	No	Harvest M	Aldi	Hard disc	Oats with	Oats with	Really Nu	Really Nu	Traffic light	Wholegrain	500 g	





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_1' program

The screenshot shows an Excel spreadsheet with the following data table:

	AH	AI	AJ	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX
1	Salt	Fibre	Nutrient	Energy_as	Energy_as	Fat_as_co	Saturated	Carbohydi	Sugar_as	Protein_a	Salt_as_cc	Fibre_as	Comment	Category	Subcateg	bar_code_starts_0	Problems
180	0,1	0,2													3	614	
181	1,9	<0,5													5	742	
182	0,8														5	90	
183	1,6														5	90	
184	0,8														5	332	
185	1,5														5	742	
186	1,5														5	742	
187	1,9														5	742	
188	1,4														5	742	
189	2,5														5	742	
190	1,4														5	742	
191	5,2	0													5	333	duplicate_code
192	4,6	0													5	628	duplicate_code
193	3,1	0													5	520	duplicate_code
194	5	0													5	333	
195	6	0													5	333	
196	2,5														5	742	FOP_labeling_type
197	5	<0,5													5	742	
198	1,9	<0,5													5	795	
199	1,7	<0,5													5	742	

The column 'AW' (labeled 'bar_code_starts_0') is highlighted with a red border. A text box in the center of the spreadsheet reads:

Appearance of a 'bar_code_starts_0' column

This column appears at the end of the table

The use of this column is described on [page 108](#) in this document





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_1' program

The screenshot shows an Excel spreadsheet with the following data table:

	AH	AI	AJ	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX	
1	Salt	Fibre	Nutrient_Energy_as	Energy_as	Fat_as	co	Saturated	Carbohydi	Sugar_as	Protein_a	Salt_as	cc	Fibre_as	Comment	Category	Subcateg	bar_code_starts_0	Problems
180	0,1	0,2													3	614		
181	1,9	<0,5													5	742		
182	0,8														5	90		
183	1,6														5	90		
184	0,8														5	332		
185	1,5														5	742		
186	1,5														5	742		
187	1,9														5	742		
188	1,4														5	742		
189	2,5														5	742		
190	1,4														5	742		
191	5,2	0													5	333		duplicate_code
192	4,6	0													5	628		duplicate_code
193	3,1	0													5	520		duplicate_code
194	5	0													5	333		
195	6	0													5	333		
196	2,5														5	742		FOP_labeling_type
197	5	<0,5													5	742		
198	1,9	<0,5													5	795		
199	1,7	<0,5													5	742		

Appearance of a 'Problems' column which lists the fields for which the R program has found errors/inconsistencies

This column appears at the end of the table





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_1' program

Example of a problem found by the 'Verifications template_step_1' program

Product 2449 → the FOP_labeling_type is missing

The R program identifies this missing information as a problem because it is a mandatory field

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Product_c	Father_pr	Country	Year	Category	Subcategory	Barcode	Brand	Brand	Brand	Brand	Brand	Brand	Brand	Brand	FOP_labeling_type	Nutri_Sco	Ingredient
185	2430		Ireland	2021	Delicatess	Cooked p	20926939	No	Deluxe	Lidl	Hard discc	Irish carve	Irish carve	Triple Oak	Triple Oak	Smoked Ham		Irish po
186	2431		Ireland	2021	Delicatess	Cooked p	20926939	No	Deluxe	Lidl	Hard discc	Irish carve	Irish carve	Triple Oak	Triple Oak	Smoked Ham		Irish po
187	2432		Ireland	2021	Delicatess	Cooked p	20926939	No	Deluxe	Lidl	Hard discc	Irish carve	Irish carve	Triple Oak	Triple Oak	Smoked Ham		Irish po
188	2438		Ireland	2021	Delicatess	Cooked p	20926939	No	Deluxe	Lidl	Hard discc	Irish carve	Irish carve	Triple Oak	Triple Oak	Smoked Ham		Irish po
189	2439		Ireland	2021	Delicatess	Cooked p	20926939	No	Deluxe	Lidl	Hard discc	Irish carve	Irish carve	Triple Oak	Triple Oak	Smoked Ham		Irish po
190	2440		Ireland	2021	Delicatess	Cooked p	20926939	No	Deluxe	Lidl	Hard discc	Irish carve	Irish carve	Triple Oak	Triple Oak	Smoked Ham		Irish po
191	2441		Ireland	2021	Delicatess	Cured han	20622930	No	Deluxe	Lidl	Hard discc	Air dried	Air dried	Parma ham	Parma ham	P.D.O Min. 14 mo	Traffic light	Pork, sa
192	2441		Ireland	2021	Delicatess	Dried, sm	20327996	No	Deluxe	Lidl	Hard discc	Prosciutto	Prosciutto	Prosciutto	Prosciutto	Crudo Del Borgo A	None from the list	Pork, sa
193	2441		Ireland	2021	Delicatess	Dry sausag	20421556	No	Deluxe	Lidl	Hard discc	Air dried	Air dried	Parma ham	Parma ham	P.D.O Min. 14 mo	Traffic light	Pork, sa
194	2444		Ireland	2021	Delicatess	Cured han	20622930	No	Deluxe	Lidl	Hard discc	Air dried	Air dried	Parma ham	Parma ham	P.D.O Min. 14 mo	Traffic light	Pork, sa
195	2445		Ireland	2021	Delicatess	Cured han	20327996	No	Deluxe	Lidl	Hard discc	Prosciutto	Prosciutto	Prosciutto	Prosciutto	Crudo Del Borgo A	None from the list	Pork, sa
196	2449		Ireland	2021	Delicatess	Cooked p	20926939	No	Deluxe	Lidl	Hard discc	Irish carve	Irish carve	Triple Oak	Triple Oak	Smoked Ham		Irish po
197	2456		Ireland	2021	Delicatess	Cooked p	20083847	No	Dulano	Lidl	Hard discc	Hand salté	Hand salté	Black Fore	Black Forest	Ham	Traffic light	Pork, sa
198	2465		Ireland	2021	Delicatess	Sausages	20015145	No	Dulano	Lidl	Hard discc	10 Frankfu	11 Frankfu	Frankfurte	Frankfurters-	contains 10 sau	None from the list	87% por
199	2471		Ireland	2021	Delicatess	Cooked p	20013066	No	Dulano	Lidl	Hard discc	Frikadelle	Frikadelle	Pork Frika	Pork Frikadellen-	contains 5 F	Traffic light	80% mir
200	2473		Ireland	2021	Delicatess	Cured han	20126599	No	Dulano	Lidl	Hard discc	Dry cured	Dry cured	Prosciutto	Prosciutto	Ham- contains 8 sl	Traffic light	Pork, sa
201	2474		Ireland	2021	Delicatess	Cooked p	20398972	No	Dulano	Lidl	Hard discc	Slices of li	Slices of li	Roast Han	Roast Ham		Traffic light	96% por
202	2481		Ireland	2021	Delicatess	Poultry ha	20138721	No	Dulano	Lidl	Hard discc	Frikadelle	Frikadelle	Turkey & (Turkey & Chicken	Frikadellen	Traffic light	50% mir
203	2483		Ireland	2021	Delicatess	Dry sausag	20974879	No	Dulano	Lidl	Hard discc	Air dried s	Air dried s	Walnut Ai	Walnut Air Dried	Salami	Traffic light	Pork, 6%
204	2484		Ireland	2021	Delicatess	Sausages	20763657	No	Dulano	Lidl	Hard discc	Wiejska sl	Wiejska sl	Wiejska- c	Wiejska-	contains 24 slices	Traffic light	Pork, sa





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_1' program

Example of a problem found by the 'Verifications template_step_1' program

Product 2449 → the FOP_labeling_type is missing

In the 'Problems' column, it says **FOP_labeling_type** → this means that for this product, there is an inconsistency/error in the 'FOP_labeling_type' field
You will have to correct and fill in the FOP_labeling_type of this product.

	AI	AJ	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX
1	Fibre	Nutrient	Energy_as	Energy_as	Fat_as_co	Saturated	Carbohydi	Sugar_as	Protein_a	Salt_as_cc	Fibre_as	Comment	Category	Subcateg	bar_code_starts_0	<u>Problems</u>
191	0												5	333		duplicate_code
192	0												5	628		duplicate_code
193	0												5	520		duplicate_code
194	0												5	333		
195	0												5	333		
196													5	742		FOP_labeling_type
197	<0,5												5	742		





Excel file to modify after 'Verifications template_step_1' program

Terms that may appear in the 'Problems' field following the 'R verifications template_step_1' program, their meaning and what to do

- In the following slides all the terms that can appear in the problem column and how to make corrections will be presented.
- Please note that the verification program 'Verifications template_step_1' is **case sensitive** (upper and lower case letters are important). Any information entered that does not match the **spelling** and **case** of the elements in the drop-down lists of the initial data entry template will appear as an error.
- If a line is completely empty (no fields filled) but errors for all fields appear after running the program, then this line should be deleted.
Empty rows are normally deleted by the program, but some may still be imported.
- The program 'verifications template_step_1' and all other programs take into account if you have used the template containing only one 'FOP_labeling_type' field or if you have used the latest version of the template containing several 'FOP_labeling_type' fields





Excel file to modify after 'Verifications template_step_1' program

Terms that may appear in the 'Problems' field following the 'R verifications template step 1' program, their meaning and what to do

<u>Problem</u>	<u>Meaning</u>	<u>Action</u>	
Duplicate_code	Different products have the same product code	→ You must change the product code so that all products have a unique code	'Product_code' is a mandatory field
Empty_product_code	The product does not have a unique product code	→ You must create a unique product code that does not already exist for the product	
Country	Incorrect country name (i.e. not contained in the closed list of the input template) or missing country name	→ You must check the spelling of the country by comparing it with the closed list of the input template or add the country name if it is missing	'Country' is a mandatory field
Year	year different from that/those indicated in the 'set parameters' part of the program	→ You must correct the collection year that does not match your collection year(s) or add it if it is missing	'Year' is a mandatory field
Category_name	Incorrect category name (i.e. not contained in the closed list of the input template) or missing category name	→ You must check the spelling of the category name by comparing it with the closed list of the input template or add the category name if it is missing	'Category_name' is a mandatory field
Category_code	Category code that does not exist or missing category code	→ You must check the category code exists by comparing it with the classification guides of the 5 food categories or add the category code if it is missing	'Category_code' is a mandatory field



Excel file to modify after 'Verifications template_step_1' program

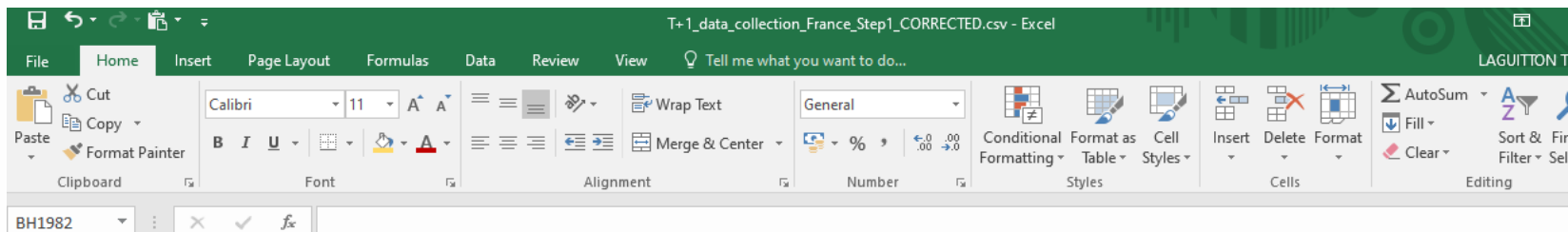
Terms that may appear in the 'Problems' field following the 'R verifications template_step_1' program, their meaning and what to do

<u>Problem</u>	<u>Meaning</u>	<u>Action</u>	
Subcategory_name	Incorrect subcategory name (i.e. not contained in the closed list of the input template) or missing subcategory name	→ You must check the spelling of the subcategory name by comparing it with the closed list of the input template or add the subcategory name if it is missing	'Subcategory_name' is a mandatory field
Subcategory_code	Category code that does not exist or missing category code	→ You must check the subcategory code exists by comparing it with the classification guides of the 5 food categories or add the subcategory code if it is missing	'Subcategory_code' is a mandatory field
Bar_code_length_or_empty	The barcode does not have 8, 12, 13, 14, or 15 digits or is missing	→ You must go back to the product pictures and correct the barcode or add the barcode if you have forgotten it. If the barcode does not appear on the product pictures or is the same as in the picture, you must indicate in the 'Comments' field: " <i>barcode checked</i> ".	
Bar_code_chr	The barcode contains characters other than numbers that are unwanted	→ You must go back to the product pictures and correct the barcode as there can be no characters other than numbers in a barcode	



WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_1' program



- If the problem **bar_code_length_or_empty** or **bar_code_chr** appears in the 'Problems' column, you need to look at the product pictures to correct the barcode entered.

If you see that the barcode to be corrected **starts with a number "0"**, you must indicate in the column 'bar_code_starts_0' → **"Yes"**

You don't need to enter the 0 at the beginning of the barcode in the 'Bar_code' field as this will be removed by Excel (the csv format doesn't take 0's at the beginning of a number into account) but this important information will be kept thanks to the 'bar_code_starts_0' column.

AT	AU	AV	AW	AX	
	Comment	Category	Subcategory	bar_code_starts_0	Problems
		9	653		
		9	653		
		9	653		
		9	648		
		9	648		
		9	648		bar_code_length_or_empty
		9	648		bar_code_length_or_empty
		9	078	Yes	bar_code_length_or_empty
		9	646		
		9	669		
		9	669		
		9	668		
		9	646		
		9	646		
		9	646	Yes	bar_code_chr
		9	646		
		9	646		
		9	645		
		9	645		
		9	645		





Excel file to modify after 'Verifications template_step_1' program

Terms that may appear in the 'Problems' field following the 'R verifications template_step_1' program, their meaning and what to do

<u>Problem</u>	<u>Meaning</u>	<u>Action</u>	
Brand_name	Brand name is missing	→ You must go back to the product pictures and add the brand name of the product if you have forgotten it. If the brand name does not appear on the product pictures or is unreadable, you must indicate in the 'Comments' field: " <i>brand name checked and unreadable/not existing</i> ".	
Type_of_brand	Incorrect type of brand (i.e. not contained in the closed list of the input template) or missing type of brand	→ You must check the spelling of the type of brand by comparing it with the closed list of the input template or add the type of brand if it is missing by looking at the product's brand name	'Type_of_brand' is a mandatory field
<ul style="list-style-type: none"> • Legal_name • Legal_name_english 	Legal name is missing Legal name in english is missing	→ You must go back to the product pictures and add the legal name of the product if you have forgotten it. If the legal name does not appear on the product pictures or is unreadable, you must indicate in the 'Comments' field: " <i>legal name checked and unreadable/not existing</i> ". → If the legal name in English is missing, you must translate the legal name and add it to the template	



Excel file to modify after 'Verifications template_step_1' program

Terms that may appear in the 'Problems' field following the 'R verifications template_step_1' program, their meaning and what to do

<u>Problem</u>	<u>Meaning</u>	<u>Action</u>	
<ul style="list-style-type: none"> • Commercial_name • Commercial_name_english 	<p>Commercial name is missing</p> <p>Commercial name in english is missing</p>	<p>→ You must go back to the product pictures and add the commercial name of the product if you have forgotten it. If the commercial name does not appear on the product pictures or is unreadable, you must indicate in the 'Comments' field: <i>"commercial name checked and unreadable/not existing"</i>.</p> <p>→ If the commercial name in English is missing, you must translate the commercial name and add it to the template</p>	
<ul style="list-style-type: none"> • FOP_labeling_type <div style="border: 1px solid red; padding: 5px; margin: 5px 0;"> <ul style="list-style-type: none"> • FOP_labeling_type_2 • FOP_labeling_type_3 • FOP_labeling_type_4 </div> <p style="color: red; text-align: center;">↓ can appear if you use the latest version of the template</p>	<p>Incorrect FOP labeling type (i.e. not contained in the closed list of the input template) or missing FOP labeling type</p>	<p>→ You must check the spelling of the FOP labeling type by comparing it with the closed list of the input template or add the FOP labeling type if it is missing by looking at the product's pictures. If there is no FOP labeling type of interest on the pictures of the product, you must enter 'None from the list'.</p> <p>→ When this problem occurs, it can affect the 'FOP_labeling_type' field and also the 'FOP_labeling_type_2/_3/_4' fields for those using the latest version of the template. In this case, it is necessary to check the spelling of the field concerned.</p>	<p>'FOP_labeling_type' is a mandatory field</p>
<p>Nutri_score</p>	<p>Incorrect nutri-score (not a letter between A and E)</p>	<p>→ You must go back to the product pictures and find the correct nutri-score of the product and enter it in the template</p>	



Excel file to modify after 'Verifications template_step_1' program

Terms that may appear in the 'Problems' field following the 'R verifications template_step_1' program, their meaning and what to do

<u>Problem</u>	<u>Meaning</u>	<u>Action</u>
Ingredient_list	Ingredient list is missing	→ You must go back to the product pictures and add the ingredient list of the product if you have forgotten it. If the ingredient list does not appear on the product pictures or is unreadable, you must indicate in the 'Comments' field: <i>"ingredient list checked and unreadable/not existing"</i> .
Net_weight	The net weight contains characters other than numbers that are unwanted	→ You must go back to the product pictures and correct the net weight as there can be no characters other than numbers in the 'net_weight' field.
Net_weight_unit	The net weight unit is different from « g » or « mL » (i.e. not contained in the closed list of the input template)	→ You must check the spelling of the net weight unit by paying attention to upper and lower case. It should be entered as "g" or "mL".
Number_of_units	The number of units contains characters other than numbers that are unwanted	→ You must correct the number of units as there can be no characters other than numbers in the 'number_of_units' field
Portion_size	The portion size contains characters other than numbers that are unwanted	→ You must go back to the product pictures and correct the portion size as there can be no characters other than numbers in the 'portion_size' field.



Excel file to modify after 'Verifications template_step_1' program

Terms that may appear in the 'Problems' field following the 'Verifications template_step_1' program, their meaning and what to do

<u>Problem</u>	<u>Meaning</u>	<u>Action</u>
Portion_size_unit	The portion size unit is different from « g » or « mL »	→ You must check the spelling of the net weight unit by paying attention to upper and lower case. It must be entered "g" or "mL" and not something else.
Nutrient_content_expression_unit	The nutrient content expression unit is different from « 100g » or « 100mL »	→ You must check the spelling of the nutrient content expression unit. It must be entered « 100g » or « 100mL » and not something else.
<ul style="list-style-type: none"> • Energy_kCal • Energy_kJ • Fat • Saturated_fat • Carbohydrates • Sugar • Protein • Salt • Fibre 	The fields contain characters other than numbers (except "<" and "traces") that are unwanted.	<p>→ You must correct so that only numbers remain and no other characters</p> <p>→ If you have any doubt about the values when correcting, go back to the product photos</p>
Nutrient_content_expression_unit_as_consumed	The nutrient content expression unit for products to be reconstituted is different from « 100g » or « 100mL »	→ You must check the spelling of the nutrient content expression unit as consumed. It must be entered « 100g » or « 100mL » and not something else



WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_1' program

Terms that may appear in the 'Problems' field following the 'R verifications template_step_1' program, their meaning and what to do

<u>Problem</u>	<u>Meaning</u>	<u>Action</u>
<ul style="list-style-type: none">• Energy_as_consumed_kCal• Energy_as_consumed_kJ• Fat_as_consumed• Saturated_fat_as_consumed• Carbohydrates_as_consumed• Sugar_as_consumed• Protein_as_consumed• Salt_as_consumed• Fibre_as_consumed	The fields contain characters other than numbers (except "<" and "traces") that are unwanted	<p>→ You must correct so that only numbers remain and no other characters</p> <p>→ If you have any doubt about the values when correcting, go back to the product photos</p>





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_1' program

- Fields that are indicated as mandatory in the previous table and that show a problem when information is missing
→ **You must correct or enter information for these fields.**
- For nutritional values containing the word “traces”, it is also necessary to **check by hand** and **standardize** the spelling of "**traces**" (so you don't have "Traces", "trace", ...)
- For products where a problem appears but the information entered is correct or missing, you must indicate in the '*Comments*' field that the problem has been checked (see the '*Action*' column in the previous tables).
→ **This way, when you will run the verification program again, you will be able to identify problems that appear without being a problem (and have already been verified).**





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_1' program

- **Be careful! When saving the file, you must be sure that the barcodes appear in full and not in scientific format. If this is the case, you must repeat the procedure on [pages 24→28](#).**
- Once the corrections have been made and the barcodes are in the correct format, you can save the file *T+1_data_collection_country_Step1_CORRECTED.csv* and close it.
- You will have to run this corrected file again in the 'Verifications template_step_1' program to make sure you haven't missed a check.

Your Rstudio interface must be cleaned up before running the 'Verifications template_step_1' program again.

All cleaning steps are described on [pages 62→68](#).





WORK Package 5 – Reformulation and processed food monitoring

2nd running of 'Verifications template_step_1' program

```
13 library("tidyverse")
14
15
16 #SET PARAMETERS :
17 #You will need to personalize the variables below with your parameters
18
19 ##Input file (csv) name :
20 #Indicate ("T+1_data_collection_France.csv") - here you need to replace "France"
21 input_file="T+1_data_collection_France.csv"
22
23 ##Verification file (csv) name :
24 ##The csv file should be saved in the same folder as the one containing your data
25 verification_file="wp5_verification_file.csv"
26
27 ##Nomenclature file (csv) name :
28 ##The csv file should be saved in the same folder as the one containing your data
29 nomenclature_file="Best-ReMaP_nomenclature.csv"
30
31 ##output file (csv) name :
32 output_file="T+1_data_collection_France_Step1_VERIFIED.csv"
33
34 ##Data collection year :
35 ##Enter the year at which you have collected your data for Best-ReMaP (replace
36 year="2021"
37 #Optional - if the data collection overlaps two consecutive years (example 2021
38 ##IF YOU HAVE ONLY ONE YEAR OF DATA COLLECTION, INDICATE THE SAME YEAR IN year_
39 year_2="2021"
40
41 ##Indicate the separator used for saving your Excel file in csv ("," or ";")
42 separator=";"
43
44 ##If you have used a specific alphabet (greek, etc) in your template, please in
45 special_alphabet="NO"
46
47 #####
48 #here-getwd() # Get the current working directory
49
```

Setting parameters of the second running of 'Verifications template_step_1' program

For this second running of the 'Verifications template_step 1', the only fields you need to change are the names of the input file and the output file.

Input_file =
"T+1_data_collection_country_Step1_CORRECTED.csv"

Output_file =
"T+1_data_collection_country_Step1_VERIFIED2.csv"

You must also indicate the year/years of your collection, the separator and if you have a specific alphabet as in the first run of the program.





WORK Package 5 – Reformulation and processed food monitoring

2nd running of 'Verifications template_step_1' program

- At the end of this second run, you get in the "**Files**" folder a file called:
"T+1_data_collection_country_**Step1**_VERIFIED2.csv"
 - You must create a copy of this file and call it :
"T+1_data_collection_country_**Step1**_CORRECTED2.csv"
- It is on this file that you will make the modifications following the checks





WORK Package 5 – Reformulation and processed food monitoring

2nd running of 'Verifications template_step_1' program

On the file : *T+1_data_collection_country_Step1_CORRECTED2.csv*

If no problems appear

Or

If problems that appear are indicated as "checked"/"verified" in the 'Comments' field

If problems that has not been checked appear

Your file has been corrected successfully!

You must correct the new problems

You can save the file
(be careful of bar codes format!)
then proceed to the next data verification program 'Verifications template_step_2'
(see next slides)

You must run the program a third time :

Input file = *T+1_data_collection_country_Step1_CORRECTED2.csv*
Output file = *T+1_data_collection_country_Step1_VERIFIED3.csv*

New file : *T+1_data_collection_country_Step1_CORRECTED3.csv*

...

You will need to run the 'Verifications template_step_1' program until no more problems appear or only problems that have been indicated as already verified/checked in the 'Comments' field.





3) Running of the programs

A. Part 1 : R setup program

B. Part 2 : Verification programs and template cleaning/standardization

i. 1st verification program : 'Verifications template_step_1'

ii. 2nd verification program : 'Verifications template_step_2'

iii. 3rd verification program : 'Verifications template_step_3'

iv. 4th verification program : 'Verifications template_step_4'





2nd verification program : 'Verifications template_step_2'

Presentation of the 'Verifications template step 2' program :

- In this second verification program, consistency problems will be highlighted: consistency of categories and their codes, sub-categories and their codes, values and their units, etc.

Requirements before starting the program 'Verifications template step 2' :

- The program 'Verifications template_step_1' should have been run on your data
- You should no longer have any problems appearing or only problems that have been notified as verified after running the program 'Verifications template_step_1'
- You must have your template in your possession and it must now be called:
T+1_data_collection_ **country** _Step1_CORRECTED(X).csv (with the name of your own country)
(X) is the number of the last file exported and corrected after the last run of the first verification program)
- You need to make sure that the barcodes in your file
T+1_data_collection_ **country** _Step1_CORRECTED(X).csv appear in full and not in scientific format (see procedure [pages 24→28](#))

Your Rstudio interface must have been cleaned up before running the program.

All cleaning steps are described on [pages 62→68](#).





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_2' program

The screenshot shows the RStudio interface with the 'File' menu open. The 'Open File...' option is highlighted. In the background, a file explorer window is open to the directory 'C:\Users\t.laguillon\Desktop\T+1_statistics_programs'. The file explorer shows a list of files and folders:

Name	Size	Modified
..		
.RData	6.2 KB	May 19, 2022
.Rhistory	30.2 KB	Aug 10, 2022
.Rprofile	25 B	May 4, 2022
Files		
Indicators		
Programs		
renv		
renv.lock	36.4 KB	May 4, 2022
T+1_statistics_programs.Rproj	218 B	Oct 10, 2022

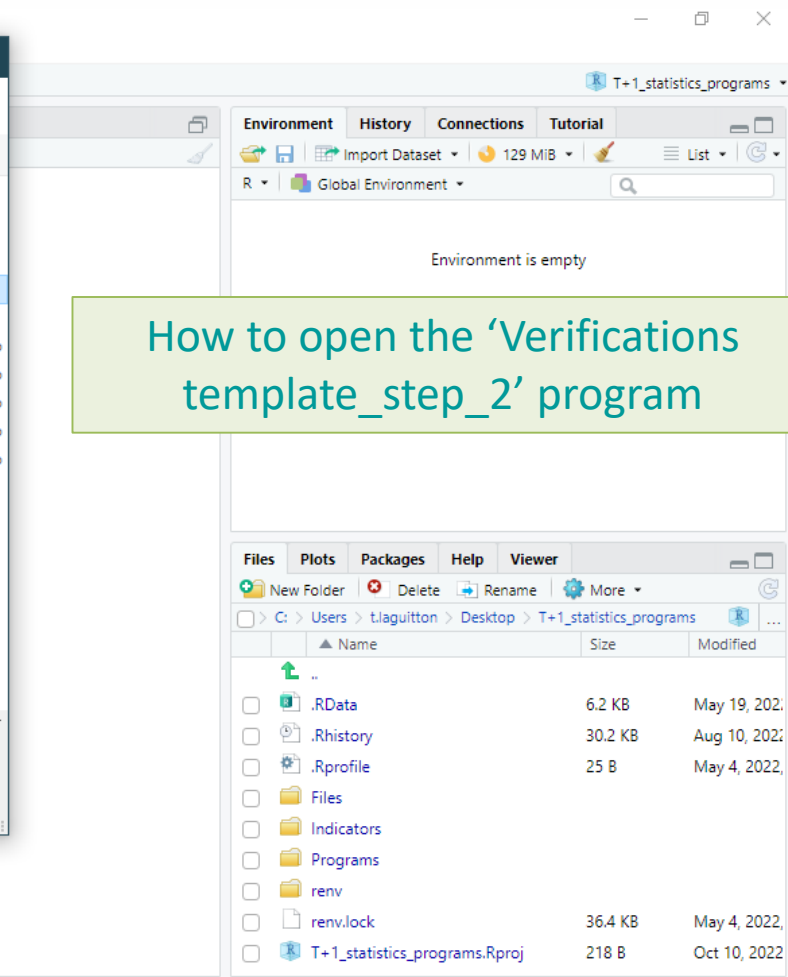
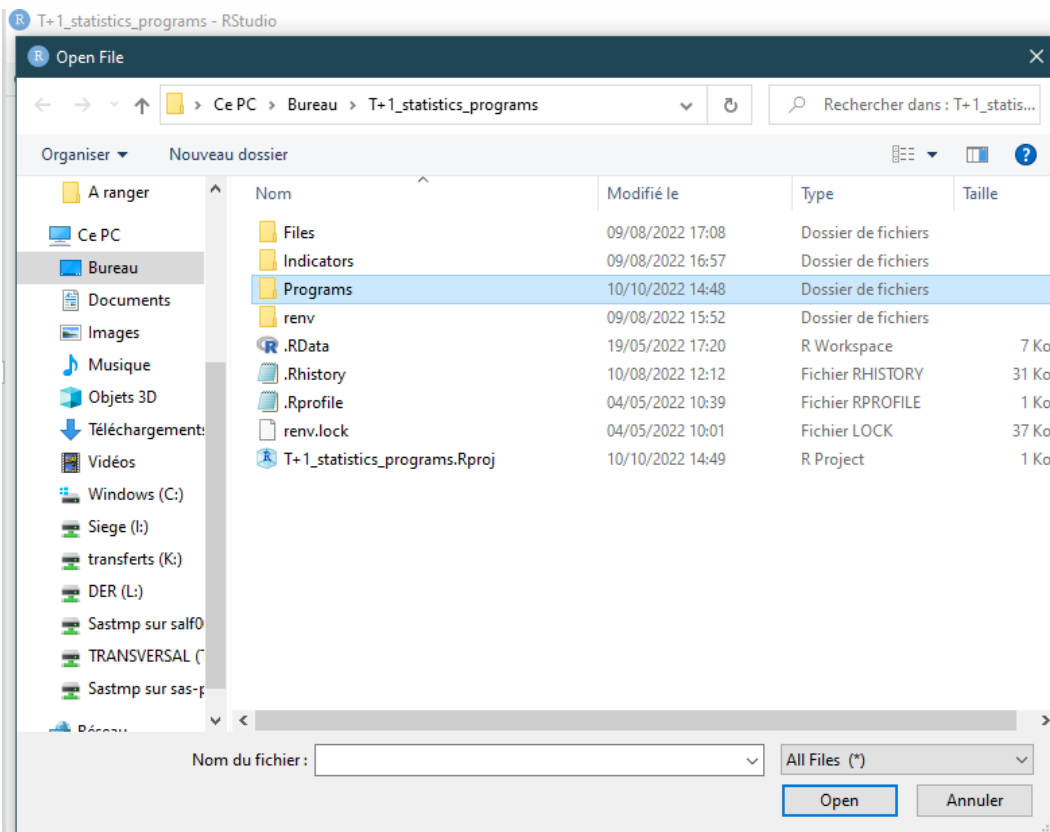
A text box in the center of the screenshot reads: 'How to open the 'Verifications template_step_2' program'.





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_2' program



How to open the 'Verifications template_step_2' program





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_2' program

The screenshot shows the RStudio interface with the 'Open File' dialog box open. The dialog is displaying the contents of the 'Programs' folder in the 'T+1_statistics_programs' directory. The file 'Verifications template_step_2.R' is selected. The 'Files' panel on the right shows the contents of the 'T+1_statistics_programs' folder, including the selected file.

Open File Dialog Contents:

Nom	Modifié le	Type	Taille
.Rhistory	19/05/2022 17:20	Fichier RHISTORY	0 Ko
R_setup.R	11/10/2022 11:42	Fichier R	1 Ko
T+1_indicators.R	18/05/2022 12:57	Fichier R	209 Ko
Verifications template_step_1.R	10/10/2022 18:30	Fichier R	90 Ko
Verifications template_step_2.R	03/10/2022 17:47	Fichier R	124 Ko
Verifications template_step_3.R	03/10/2022 17:47	Fichier R	88 Ko
Verifications template_step_4.R	04/10/2022 11:06	Fichier R	24 Ko

Files Panel Contents:

Name	Size	Modified
..		
.RData	6.2 KB	May 19, 2021
.Rhistory	29 KB	Oct 11, 2022
.Rprofile	25 B	May 4, 2022
Files		
Indicators		
Programs		
renv		
renv.lock	36.4 KB	May 4, 2022
T+1_statistics_programs.Rproj	218 B	Oct 12, 2022

How to open the 'Verifications template_step_2' program





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_2' program

The screenshot displays the RStudio interface. The main editor window shows the R script 'Verifications template_step_2.R' with the following content:

```
1 ##### SECOND PROGRAMME OF VERIFICATION
2
3 #Version of R used : 4.1.2
4
5 #Rstudio needs to be used for using the program in a simpler way : it is a tool for working on R programs
6 ###IMPORTANT : R software must be installed on your computer in order to install, run and use Rstudio
7
8
9 ##dplyr : 1.0.7
10 ##tidyverse : 1.3.1
11
12 #Load the necessary packages each time you open Rstudio for running the program with the function
13 library("dplyr")
14 library("tidyverse")
15
16
17 #SET PARAMETERS :
18 #You will need to personalize the variables below with your parameters
19
20
21 ##Input file (csv) name :
22 #Indicate ("T0_data_collection_France_step1_CORRECTED.csv") - here you need to replace "France" by the name
23 input_file="T+1_data_collection_plf_Step1_VERIFIED.csv"
24 #The template you will import is the one you have created via verifications template_step_1 ; it should a
25
26 ##Verification file (csv) name :
27 ##The csv file should be saved in the same folder as the one containing your data set in csv - here you n
28 verification_file="wp5_verification_file.csv"
29
30 ##Nomenclature file (csv) name :
31 ##The csv file should be saved in the same folder as the one containing your data set in csv
32 nomenclature_file="Best-ReMaP_nomenclature.csv"
33
34 #Enter the name of your country exactly as entered in the file (respect the capital letter)
35 #Replace Ireland with your own country (keep the quoting marks)
36 own_country="Ireland"
37
```

The Environment pane on the right shows 'Global Environment' and 'Environment is empty'. A file explorer window is open at the bottom right, showing the directory 'C:\Users> t.laguillon > Desktop > T+1_statistics_programs'. The file explorer contains the following files and folders:

Name	Size	Modified
..		
.RData	6.2 KB	May 19, 2021
.Rhistory	29 KB	Oct 11, 2022
.Rprofile	25 B	May 4, 2022
Files		
Indicators		
Programs		
renv		
renv.lock	36.4 KB	May 4, 2022
T+1_statistics_programs.Rproj	218 B	Oct 12, 2022

A green callout box with the text 'Verifications template_step_2' program opened' is overlaid on the RStudio interface.





Running of 'Verifications template_step_2' program

```
T+1_statistics_programs - RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
Verifications template_step_2.R x
Source on Save
16 #SET PARAMETERS :
17 #You will need to personalize the variables below with your paramet
18
19
20
21 ##Input file (csv) name :
22 #Indicate ("T0_data_collection_France_Step1_CORRECTED.csv") - here
23 input_file="T+1_data_collection_France_Step1_CORRECTED(X).csv"
24 #The template you will import is the one you have created via verif
25
26 ##Verification file (csv) name :
27 ##The csv file should be saved in the same folder as the one contain
28 verification_file="wp5_verification_file.csv"
29
30 ##Nomenclature file (csv) name :
31 ##The csv file should be saved in the same folder as the one contain
32 nomenclature_file="Best-ReMaP_nomenclature.csv"
33
34 #Enter the name of your country exactly as entered in the file (res
35 #Replace Ireland with your own country (keep the quoting marks)
36 own_country="France"
37
38 #Define the output file name
39 output_file="T+1_data_collection_France_Step2_VERIFIED.csv"
40
41 ##Indicate the separator used for saving your Excel file in csv (",
42 separator=";"
43
44 ##If you have used a specific alphabet (greek, etc) in your templat
45 special_alphabet="NO"
46
47 #####
48
32:31 (Top Level)
Console Terminal x Jobs x
R 4.1.2 · C:/Users/t.laguitton/Desktop/T+1_statistics_programs/
> |
```

Setting parameters of the 2nd verification program

In the first run of the 2nd verification program, you need to change the name of the country with your own country name in the input file (line 23) and the output file (line 39) of the R script + **line 36 ('own_country=')**

You must also indicate the separator and if you have a specific alphabet as in the program 'Verifications template_step_1'.

Example :

- **Input_file** =
"T+1_data_collection_Ireland_Step1_CORRECTED(X).csv"
(X) is the number of the last file exported and corrected after the last run of the first verification program)
- **Own_country** = "Ireland" (same spelling as in the country name in the template)
- **Output_file** = "T+1_data_collection_Ireland_Step2_VERIFIED.csv"
- **Separator** = " ; "
- **Special_alphabet** = "NO"



WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_2' program

Select the entire program (Ctrl+A) and press 'Run' (or Ctrl+Enter)

The program will run entirely.

```
1 ##### SECOND PROGRAMME OF VERIFICATION
2
3 #Version of R used : 4.1.2
4
5 #Rstudio needs to be used for using the program in a simpler way : it is a tool for
6 ###IMPORTANT : R software must be installed on your computer in order to install,
7
8
9 ##dplyr : 1.0.7
10 ##tidyverse : 1.3.1
11
12 #Load the necessary packages each time you open Rstudio for running the program with
13 library("dplyr")
14 library("tidyverse")
15
16
17 #SET PARAMETERS :
18 #You will need to personalize the variables below with your parameters
19
20
21 ##Input file (csv) name :
22 #Indicate ("T0_data_collection_France_step1_CORRECTED.csv") - here you need to replace "France" by the name of
23 input_file="T+1_data_collection_France_step1_CORRECTED(X).csv"
24 #The template you will import is the one you have created via verifications template_step_1 ; it should be
25
26 ##Verification file (csv) name :
27 ##The csv file should be saved in the same folder as the one containing your data set in csv - here you need to
28 verification_file="wp5_verification_file.csv"
29
30 ##Nomenclature file (csv) name :
31 ##The csv file should be saved in the same folder as the one containing your data set in csv
32 nomenclature_file="Best-ReMaP_nomenclature.csv"
33
```

Files | Plots | Packages | Help | Viewer

Name	Size	Modified
..		
.RData	6.2 KB	May 19, 2022
.Rhistory	29 KB	Oct 11, 2022
.Rprofile	25 B	May 4, 2022
Files		
Indicators		
Programs		
renv		
renv.lock	36.4 KB	May 4, 2022
T+1_statistics_programs.Rproj	218 B	Oct 12, 2022





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_2' program

The screenshot shows the RStudio interface. The top pane displays a data table with columns: Product_code, Father_product_code, Country, Year, Category_name, and Subcategory_name. The bottom pane shows the R console with the following code and output:

```
R 4.1.2 · C:/Users/tlaguitton/Desktop/T+1_statistics_programs/Files/  
+ }else if(separator==""){  
+   utils::write.csv(Step2_T0_data_f,file=output_file,na=" ",row.names=FALSE)  
+ }  
>  
>  
> print("The file should appear in the folder 'Files'")  
[1] "The file should appear in the folder 'Files'"  
>
```

After running the program

Several tables appear next to the R program script. These tables are working tools for the R program but you don't need to look at them (only if there are errors or warnings in the console but in this case you have to tell us so we can help you).

At the end of the 'Verifications template_step_2' program execution, you should see the **sentence highlighted in yellow** in the console.

The program exports an excel file named **T+1_data_collection_country_Step2_VERIFIED.csv** to your "Files" folder in the "T+1_statistics_programs" folder.





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_2' program

Create a copy of the file exported by R

In the 'Files' folder, you will find the file exported by the 2nd verification program `T+1_data_collection_country_Step2_VERIFIED.csv`

→ You need to create a copy of this file and rename it “`T+1_data_collection_country_Step2_CORRECTED.csv`”

It is on this file that you will make the corrections and modifications indicated

Nom	Modifié le	Type
Best-ReMaP_nomenclature.csv	12/07/2022 16:48	Fichier CSV Micro...
pre_existing_data_France.csv	03/10/2022 15:52	Fichier CSV Micro...
T+1_data_collection_France.csv	10/10/2022 17:10	Fichier CSV Micro...
T+1_data_collection_France_Step1_CORRECTED(X).csv	10/10/2022 17:04	Fichier CSV Micro...
T+1_data_collection_France_Step1_VERIFIED.csv	10/10/2022 17:13	Fichier CSV Micro...
T+1_data_collection_France_Step2_CORRECTED.csv	12/10/2022 16:46	Fichier CSV Micro...
T+1_data_collection_France_Step2_VERIFIED.csv	12/10/2022 16:46	Fichier CSV Micro...
wp5_verification_file.csv	30/08/2022 10:29	Fichier CSV Micro...
Years of interest.csv	03/10/2022 17:27	Fichier CSV Micro...





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_2' program

T+1_data_collection_France_Step2_CORRECTED.csv - Excel

File Home Insert Page Layout Formulas Data Review View Tell me what you want to do... LAGUITTON Thomas Share

Clipboard Font Alignment Number Styles Cells Editing

Product_code

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
Product_code	Father_pr	Country	Year	Category	Subcategory	Bar_code	Assortme	Brand_na	Brand_ow	Type_of	Legal_nan	Legal_nan	Commerci	Commerci	FOP_label	Nutri_Sco	Ingredient	Net_weig	Net_weig	Numb
1650		Ireland	2021	Breakfast	Cereals w	4,09E+12	No	Kavanagh'	Aldi	Hard discc	Gluten Fr	Gluten Fr	Gluten Fr	Gluten Fr	None fron		Gluten Fr	500 g		
1651		Ireland	2021	Breakfast	Traditiona	4,09E+12	No	Kavanagh'	Aldi									500 g		
1653		Ireland	2021	Breakfast	Cereals w	4,09E+12	No	Kavanagh'	Aldi									500 g		
1654		Ireland	2021	Breakfast	Cereals w	4,09E+12	No	Kavanagh'	Aldi									750 g		
1655		Ireland	2021	Breakfast	Cereals w	4,09E+12	No	Kavanagh'	Aldi									1500 g		
1657		Ireland	2021	Breakfast	Traditiona	4,09E+12	No	Specially	Aldi									500 g		
1658	603	Ireland	2021	Breakfast	High-fibre	4,09E+12	No	Harvest M	Aldi									750 g		
1659		Ireland	2021	Breakfast	Chocolate	4,09E+12	No	Harvest M	Aldi									375 g		
1661	937	Ireland	2021	Breakfast	Chocolate	4,09E+12	No	Harvest M	Aldi									375 g		
1663	586	Ireland	2021	Breakfast	Sweet cer	4,09E+12	No	Harvest M	Aldi									500 g		
1665	916	Ireland	2021	Breakfast	Crunchy fr	4,09E+12	No	Harvest M	Aldi									500 g		
1672		Ireland	2021	Breakfast	Crunchy n	4,09E+12	No	Harvest M	Aldi									500 g		
1674	597	Ireland	2021	Breakfast	Sweet cer	4,09E+12	No	Harvest M	Aldi									500 g		
1675		Ireland	2021	Breakfast	Crunchy fr	4,09E+12	No	Harvest M	Aldi									500 g		
1676		Ireland	2021	Breakfast	Crunchy fr	4,09E+12	No	Harvest M	Aldi									500 g		
1677	937	Ireland	2021	Breakfast	Chocolate	4,09E+12	No	Harvest M	Aldi									500 g		
1680	996	Ireland	2021	Breakfast	Sweet cer	4,09E+12	No	Harvest M	Aldi									500 g		
1681		Ireland	2021	Breakfast	Traditiona	4,09E+12	No	Harvest M	Aldi	Hard discc	Gluten Fr	Gluten Fr	Gluten Fr	Gluten Fr	None fron		Wholegra	400 g		
1682	850	Ireland	2021	Breakfast	Traditiona	4,09E+12	No	Harvest M	Aldi	Hard discc	Honey To	Honey To	Raisin & A	Raisin & A	Traffic lig		WHOLEGR	1000 g		
1683		Ireland	2021	Breakfast	Traditiona	4,09E+12	No	Harvest M	Aldi	Hard discc	Oats with	Oats with	Really Nu	Really Nu	Traffic lig		Wholegra	500 g		

Overview of the **T+1_data_collection_country_Step2_CORRECTED.csv** file which is a new version of your T+1 data collection template in which you will have to make modifications/corrections





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_2' program

The screenshot shows an Excel spreadsheet with the following data table:

	AI	AJ	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX
1	Fibre	Nutrient_	Energy_as	Energy_as	Fat_as_co	Saturated	Carbohydi	Sugar_as_	Protein_a	Salt_as_cc	Fibre_as_	Comment	Category_	Subcategc	bar_code_starts_0	Problems
50	<0,5												3	611		
51	<0,5												3	612		
52	0,6												3	612		
53	<0,5												3	611		duplicate_bar_code
54	<0,5												3	613		
55	<0,5												3	612		
56	<0,5												3	613		
57	<0,5												3	614		verify_units_mL_100mL
58	<0,5												3	614		verify_units_mL_100mL
59													3	711		
60	2,9												3	615		
61	2,8												18	401		
62	4,9												18	401		
63	7,5												18	396		portion_size_&_units
64	5,7												18	396		portion_size_&_units
65	17												18	396		portion_size_&_units
66	16												18	396		portion_size_&_units
67	3,3												18	405		
68	4,3												18	399	Yes	
69	3,1												18	405		

Appearance of a 'Problems' column which lists the fields for which the R program has found errors/inconsistencies

This column appears at the end of the table





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_2' program

The screenshot shows an Excel spreadsheet with the following data table:

	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM
1	Net_weig	Net_weig	Number	Portion_s	Portion_s	Portion_s	Preservat	Nutrient	Energy_kJ	Energy_kC	Fat	Saturated	Carbohydr	Sugar	Protein	Salt	Fibre	Nutrient	Energy_as	Energy_as	Fat_as
98	750 g		1	30 g			Ambiant	100g	1565	371		4,7	2,4	69,4	21,6	8,5	0,7	8,4			
99	400 g		1	45 g			Ambiant	100g	1851	44											
100	400 g		1	45 g			Ambiant	100g	1882	44											
101	375 g		1	30 g			Ambiant	100g	1627	38											
102	500 g		1	45 g			Ambiant	100g	1897	45											
103	500		1	30 g			Ambiant	100g	1662	39											
104	500 g		1	45 g			Ambiant	100g	2086	49											
105	625 g		1	30 g			Ambiant	100g	1533	36											
106	750 g		1	45 g			Ambiant	100g	1559	36											
107	750 g		1	45 g			Ambiant	100g	1529	36											
108	720 g		1	40 g			Ambiant	100g	1530	36											
109	500 g		1	45 g			Ambiant	100g	1612	38											
110	500 g		1	45 g			Ambiant	100g	1876	44											
111	500 g		1	50 g			Ambiant	100g	1570	37											
112	57 g		1	57 g			Ambiant	100g	1555	36											
113	50 g		1	50 g			Ambiant	100g	1494	35											
114	1500 g		1	40 g			Ambiant	100g	1560	37											
115	500 g		1	30 g			Ambiant	100g	1560	37											
116	280 g		4	70 g			Ambiant	100g	1406	335	11,1	5,9	45,6	7,2	11,4	1,3	3,3				
117	480 g		12	40 g			Ambiant	100g	1145	270	1,9	0,4	51,1	2,8	10,5	0,6	3,4				

The callout box contains the following text:

Example of a problem found by the 'Verifications template_step_2' program

Line 103 → the net_weight_unit is missing while there is a value for the net_weight

The R program identifies this missing information as a consistency problem





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_2' program

	AH	AI	AJ	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX
1	Salt	Fibre	Nutrient_Energy_as	Energy_as	Fat_as_co	Saturated	Carbohydr	Sugar_as	Protein_a	Salt_as_cc	Fibre_as	Comment	Category	Subcategory	bar_code_starts_0	<u>Problems</u>	
98		0,7	8,4											1	676		
99		0,9	6											1	679		
100		0,6	5,6											1	680		
101		0,2	6,3											1	142		
102		0,7	3,7											1	678		
103		0,7	3,6											1	680		net_weight_&_units
104		0	6,5											1	680		
105		0,4	13,1											1	143		
106		0,1	7,1											1	386		
107		0	7,4											1	386		
108		0,2	9,5											1	143		
109																	type_of_brand

Example of a problem found by the 'Verifications template_step_2' program

Line 103 → the net_weight_unit is missing

In the 'Problems' column, it says **net_weight_&_units** → this means that for this product, there is an inconsistency/error in the 'net_weight_unit' or 'net_weight' fields.

You will have to fill in the net_weight_unit of this product





Excel file to modify after 'Verifications template_step_2' program

Terms that may appear in the 'Problems' field following the 'Verifications template_step_2' program, their meaning and what to do

<u>Problem</u>	<u>Meaning</u>	<u>Action</u>
nomenclature	Wrong association between 'Category_name', 'Category_code', 'Subcategory_name' and 'Subcategory_code'	→ Check the 4 fields and correct those (or the one) that are not correctly associated
Net_weight_&_units	<ul style="list-style-type: none"> The net weight is filled but there is no associated net weight unit <p><u>or</u></p> <ul style="list-style-type: none"> The net weight unit is filled but there is no associated net weight 	→ If there is a value in the 'net_weight' field, you must add the unit « g » or « mL » in the 'net_weight_unit' field → If there is a unit in the field 'net_weight_unit', you must go back to the pictures of the product and add the value of the portion size. If there is no net weight for this product, it must be an input error and you can delete the unit. → When you have any doubts, you must go back to the pictures of the product to be sure of what to enter
Portion_size_&_units	<ul style="list-style-type: none"> The portion size is filled but there is no associated portion size unit <p><u>or</u></p> <ul style="list-style-type: none"> The portion size unit is filled but there is no associated portion size 	→ If there is a value in the 'portion_size' field, you must add the unit « g » or « mL » in the 'portion_size_unit' field → If there is a unit in the field 'portion_size_unit', you must go back to the pictures of the product and add the value of the portion size. If there is no portion size for this product, it must be an input error and you can delete the unit. → When you have any doubts, you must go back to the pictures of the product to be sure of what to enter





Excel file to modify after 'Verifications template_step_2' program

Terms that may appear in the 'Problems' field following the 'Verifications template_step_2' program, their meaning and what to do

<u>Problem</u>	<u>Meaning</u>	<u>Action</u>
Nutritional_values _&_units	<ul style="list-style-type: none"> The nutrient content expression unit is filled but there are no associated nutritional values for the nutrients <p><u>or</u></p> <ul style="list-style-type: none"> There are nutritional values for the nutrients but there is no associated nutrient content expression unit 	<p>→ If there are nutritional values for the nutrients, you must add the unit « 100g » or « 100mL » in the 'nutrient_content_expression_unit' field</p> <p>→ If there is a unit in the field 'nutrient_content_expression_unit', you must go back to the pictures of the product and add the nutritional values of each nutrient. If there is nutritional values for this product, it must be an input error and you can delete the unit.</p> <p>→ When you have any doubts, you must go back to the pictures of the product to be sure of what to enter</p>
Nutritional_values _as_consumed_&_units	<ul style="list-style-type: none"> The nutrient content expression unit for products to be reconstituted is filled but there is no associated nutritional values for the nutrients as consumed <p><u>or</u></p> <ul style="list-style-type: none"> There are nutritional values for the nutrients as consumed but there is no associated nutrient content expression unit for products to be reconstituted 	<p>→ If there are nutritional values as consumed for the nutrients, you must add the unit « 100g » or « 100mL » in the 'nutrient_content_expression_unit_as_consumed' field</p> <p>→ If there is a unit in the field 'nutrient_content_expression_unit_as_consumed', you must go back to the pictures of the product and add the nutritional values as consumed of each nutrient. If there is no nutritional values as consumed for this product, it must be an input error and you can delete the unit.</p> <p>→ When you have any doubts, you must go back to the pictures of the product to be sure of what to enter</p>



Excel file to modify after 'Verifications template_step_2' program

Terms that may appear in the 'Problems' field following the 'Verifications template_step_2' program, their meaning and what to do

<u>Problem</u>	<u>Meaning</u>	<u>Action</u>
Wrong_country	This is not the name of your country	→ You must enter the name of your own country, making sure that this is the same spelling as the closed list of the input template
Duplicate_bar_code	Same bar code has been found for 2 or more products	<p>→ If the products have the same bar code and are similar (= duplicates = same bar code + same information for all the fields), you must delete one of the products to keep only one.</p> <p>→ If the products have the same bar code but are different, you must check if it is an input error by going back to the pictures of the products.</p> <ul style="list-style-type: none"> • If it is an input error, you must enter the correct bar code. • If the products really have the same bar code, you must keep them in the template and indicate in the 'Comments' field : « bar code checked and same for several different products »
Type_of_brand	The same brand has been associated with several types of brand. (This problem appears for all products of a same brand if they have been associated with different types of brand)	→ You must filter in the Excel file on a brand name that shows the problem « <i>Type_of_brand</i> », then you must select the correct type of brand and apply it to all the products of the same brand name.



Excel file to modify after 'Verifications template_step_2' program

Terms that may appear in the 'Problems' field following the 'Verifications template_step_2' program, their meaning and what to do

<u>Problem</u>	<u>Meaning</u>	<u>Action</u>
Carbohydrates_or_sugar_content	The sugar content is greater than the carbohydrates content	<ul style="list-style-type: none"> → You must go back to the pictures of the product and look at the sugar and carbohydrates content to correct this error. → If the error is on the product label, do not change anything and leave what is written on the product
Carbohydrates_or_sugar_as_consumed_content	The sugar as consumed content is greater than the carbohydrates as consumed content for products to be reconstituted	<ul style="list-style-type: none"> → You must go back to the pictures of the product and look at the sugar and carbohydrates content as consumed to correct this error. → If the error is on the product label, do not change anything and leave what is written on the product
Fat_or_saturated_fat_content	The saturated fat content is greater than the fat content	<ul style="list-style-type: none"> → You must go back to the pictures of the product and look at the fat and saturated fat content to correct this error. → If the error is on the product label, do not change anything and leave what is written on the product
Fat_or_saturated_fat_as_consumed_content	The saturated fat as consumed content is greater than the fat as consumed content for products to be reconstituted	<ul style="list-style-type: none"> → You must go back to the pictures of the product and look at the fat and saturated fat content as consumed to correct this error. → If the error is on the product label, do not change anything and leave what is written on the product



Excel file to modify after 'Verifications template_step_2' program

Terms that may appear in the 'Problems' field following the 'Verifications template_step_2' program, their meaning and what to do

<u>Problem</u>	<u>Meaning</u>	<u>Action</u>
Verify_units_g_100g	a unit in “g” appears for a product belonging to the 'Soft drinks' category	<p>→ You must look at all the fields of the product that have units and find the unit “g”. (Net_weight_unit, Portion_size_unit, Nutrient_content_expression_unit, Nutrient_content_expression_unit_as_consumed)</p> <p>→ You must compare with the pictures of the product to check if this is an input error and correct it if necessary.</p> <p>→ It is not necessarily an input error as some milk beverages can have units in g.</p>
Verify_units_mL_100mL	a unit in “mL” appears for a product belonging to a category other than the 'Soft drinks' category	<p>→ You must look at all the fields of the product that have units and find the unit “mL”. (Net_weight_unit, Portion_size_unit, Nutrient_content_expression_unit, Nutrient_content_expression_unit_as_consumed)</p> <p>→ You must compare with the pictures of the product to check if this is an input error and correct it if necessary.</p> <p>→ It is not necessarily an input error as some yoghourts can have units in mL.</p>



WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_2' program

Terms that may appear in the 'Problems' field following the 'Verifications template_step_2' program, their meaning and what to do

<u>Problem</u>	<u>Meaning</u>	<u>Action</u>
Enter_Nutri_score	The 'FOP labeling type' field indicates <i>Nutriscore</i> but there is no associated nutri-score in the 'Nutriscore' field	<ul style="list-style-type: none"> → You must go back to the pictures of the product and enter the letter of the nutri-score that appears on the package. → If there is no nutri-score on the picture, you must correct the entry in the field 'FOP_labeling_type' by choosing another FOP labeling type or <i>none of the list</i> (mandatory field)
Remove_Nutri_score	<ul style="list-style-type: none"> • A nutri score is filled in the 'Nutriscore' field but the 'FOP labeling type' does not indicate <i>Nutriscore</i> 	<ul style="list-style-type: none"> → You must go back to the pictures of the product and check if there is a nutri-score on the package <ul style="list-style-type: none"> • If there is a nutri-score on the package, you must indicate <i>Nutriscore</i> in the 'FOP_labeling_type' field and check that the letter of the nutri-score entered is the right one • If there is no nutri-score on the picture, you must delete the letter in the 'Nutriscore' field and choose a FOP labeling type or <i>none of the list</i> in the field 'FOP_labeling_type' (mandatory field)





Excel file to modify after 'Verifications template_step_2' program

Terms that may appear in the 'Problems' field following the 'Verifications template_step_2' program, their meaning and what to do

This problem can only occur if you are working on the latest version of the template which includes several fields for FOP labeling types (FOP_labeling_type2, FOP_labeling_type3, FOP_labeling_type4)

<u>Problem</u>	<u>Meaning</u>	<u>Action</u>
Incorrect_FOPs	<p>The first field 'FOP_labeling_type' indicates <i>None from the list</i> but not the other fields 'FOP_labeling_type2/3/4'.</p> <p>When the first field 'FOP_labeling_type' indicates <i>None from the list</i>, the other fields 'FOP_labeling_type2/3/4' must also indicate <i>None from the list</i>.</p>	<p>→ If one or more of the fields 'FOP_labeling_2/3/4' is empty, then you must enter <i>None from the list</i> in those fields.</p> <p>→ If another FOP label (nutriscore, reference intake, ...) is indicated in the field 'FOP_labeling_type2/3 or 4', you have to check on the product pictures that this is not an error and if there is indeed an FOP label on the product, it must be indicated in the first field 'FOP_labeling_type' and the other 2/3/4 must indicate 'none from the list'.</p>



WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_2' program

- **Be careful! When saving the file, you must be sure that the barcodes appear in full and not in scientific format. If this is the case, you must repeat the procedure on [pages 24→28](#).**
- Once the corrections have been made and the barcodes are in the correct format, you can save the file *T+1_data_collection_country_Step2_CORRECTED.csv* and close it.
- You will have to run this corrected file again in the 'Verifications template_step_2' program to make sure you haven't missed a check.

Your Rstudio interface must be cleaned up before running the 'Verifications template_step_2' program again.

All cleaning steps are described on [pages 62→68](#).





WORK Package 5 – Reformulation and processed food monitoring

2nd running of 'Verifications template_step_2' program

```
13 library("dplyr")
14 library("tidyverse")
15
16
17 #SET PARAMETERS :
18 #You will need to personalize the variables below with your parameters
19
20
21 ##Input file (csv) name :
22 #Indicate ("T0_data_collection_France_Step1_CORRECTED.csv") - here you need to
23 input_file="T+1_data_collection_France_Step2_CORRECTED.csv"
24 #The template you will import is the one you have created via verifications te
25
26 ##Verification file (csv) name :
27 ##The csv file should be saved in the same folder as the one containing your d
28 verification_file="wp5_verification_file.csv"
29
30 ##Nomenclature file (csv) name :
31 ##The csv file should be saved in the same folder as the one containing your d
32 nomenclature_file="Best-ReMaP_nomenclature.csv"
33
34 #Enter the name of your country exactly as entered in the file (respect the ca
35 #Replace Ireland with your own country (keep the quoting marks)
36 own_country="Ireland"
37
38 #Define the output file name
39 output_file="T+1_data_collection_France_Step2_VERIFIED2.csv"
40
41 ##Indicate the separator used for saving your Excel file in csv ("," or ";")
42 separator=";"
43
44 ##If you have used a specific alphabet (greek, etc) in your template, please i
45 special_alphabet="NO"
46
47 #####
48 here-getwd() # Get the current working directory
49
```

Setting parameters of the second running of 'Verifications template_step_2' program

For this second running of the 'Verifications template_step_2', the only fields you need to change are the names of the input file and the output file.

Input_file =
"T+1_data_collection_country_Step2_CORRECTED.csv"

Output_file =
"T+1_data_collection_country_Step2_VERIFIED2.csv"

Make sure the country name on line 36 is yours and if not, enter it.

You must also indicate the separator and if you have a specific alphabet as in the first run of the program.





WORK Package 5 – Reformulation and processed food monitoring

2nd running of 'Verifications template_step_2' program

- At the end of this second run, you get in your "files" folder a file called:

"T+1_data_collection_country_Step2_VERIFIED2.csv"

- You must create a copy of this file and call it :

"T+1_data_collection_country_Step2_CORRECTED2.csv"

→ It is on this file that you will make the modifications following the checks





WORK Package 5 – Reformulation and processed food monitoring

2nd running of 'Verifications template_step_2' program

On the file : *T+1_data_collection_country_Step2_CORRECTED2.csv*

If no problems appear
Or
If problems that appear are indicated as
"checked"/"verified" in the 'Comments' field

Your file has been corrected successfully!

You can save the file
(be careful of bar codes format!)
then proceed to the next data verification
program 'Verifications template_step_3'
(see next slides)

If problems that has not
been checked appear

You must correct the new problems

You must run the program a third time :

Input file = *T+1_data_collection_country_Step2_CORRECTED2.csv*
Output file = *T+1_data_collection_country_Step2_VERIFIED3.csv*

New file : *T+1_data_collection_country_Step2_CORRECTED3.csv*

...

You will need to run the 'Verifications template_step_2' program until no more problems appear or only problems that have been indicated as already verified/checked in the 'Comments' field.





3) Running of the programs

A. Part 1 : R setup program

B. Part 2 : Verification programs and template cleaning/standardization

i. 1st verification program : 'Verifications template_step_1'

ii. 2nd verification program : 'Verifications template_step_2'

iii. 3rd verification program : 'Verifications template_step_3'

iv. 4th verification program : 'Verifications template_step_4'





3rd verification program : 'Verifications template_step_3'

Presentation of the 'Verifications template step 3' program :

- Third verification program : verification of outliers in your nutritional values
 - For each nutrient in each subcategory, the following position indicators will be calculated :
 - 1st quartile (Q1)
 - 3rd quartile (Q3)
 - Interquartile range (IQR=Q3-Q1)
- Nutrient values will be considered outliers if they are below $Q1 - (IQR * 1.5)$ and above $Q3 + (IQR * 1.5)$.





3rd verification program : 'Verifications template_step_3'

Requirements before starting the program 'Verifications template_step_3' :

- The programs 'Verifications template_step_1' and 'Verifications template_step_2' should have been run on your data
- You should no longer have any problems appearing or only problems that have been notified as verified after running the program 'Verifications template_step_2'
- You must have your template in your possession and it must now be called:
T+1_data_collection_ **country** _Step2_CORRECTED(X).csv (with the name of your own country)
((X) is the number of the last file exported and corrected after the last run of the first verification program)
- You need to make sure that the barcodes in your file
T+1_data_collection_ **country** _Step2_CORRECTED(X).csv appear in full and not in scientific format (see procedure [pages 24→28](#))

Your Rstudio interface must have been cleaned up before running the program.
All cleaning steps are described on [pages 62→68](#).





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_3' program

The screenshot shows the RStudio interface with the 'File' menu open. The 'Open File...' option is highlighted. The Environment pane on the right shows 'Environment is empty'. A File Explorer window is open at the bottom right, showing the project directory 'T+1_statistics_programs' with files like '.RData', '.Rhistory', '.Rprofile', 'Files', 'Indicators', 'Programs', 'renv', 'renv.lock', and 'T+1_statistics_programs.Rproj'.

How to open the 'Verifications template_step_3' program





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_3' program

The screenshot shows the RStudio interface with the 'Open File' dialog box open. The dialog box is displaying the contents of the 'T+1_statistics_programs' folder. The 'Programs' folder is selected. The R Environment pane shows 'Global Environment' and 'Environment is empty'. A green callout box contains the text: 'How to open the 'Verifications template_step_3' program'.

Nom	Modifié le	Type	Taille
Files	09/08/2022 17:08	Dossier de fichiers	
Indicators	09/08/2022 16:57	Dossier de fichiers	
Programs	10/10/2022 14:48	Dossier de fichiers	
renv	09/08/2022 15:52	Dossier de fichiers	
.RData	19/05/2022 17:20	R Workspace	7 Ko
.Rhistory	10/08/2022 12:12	Fichier R HISTORY	31 Ko
.Rprofile	04/05/2022 10:39	Fichier R PROFILE	1 Ko
renv.lock	04/05/2022 10:01	Fichier LOCK	37 Ko
T+1_statistics_programs.Rproj	10/10/2022 14:49	R Project	1 Ko

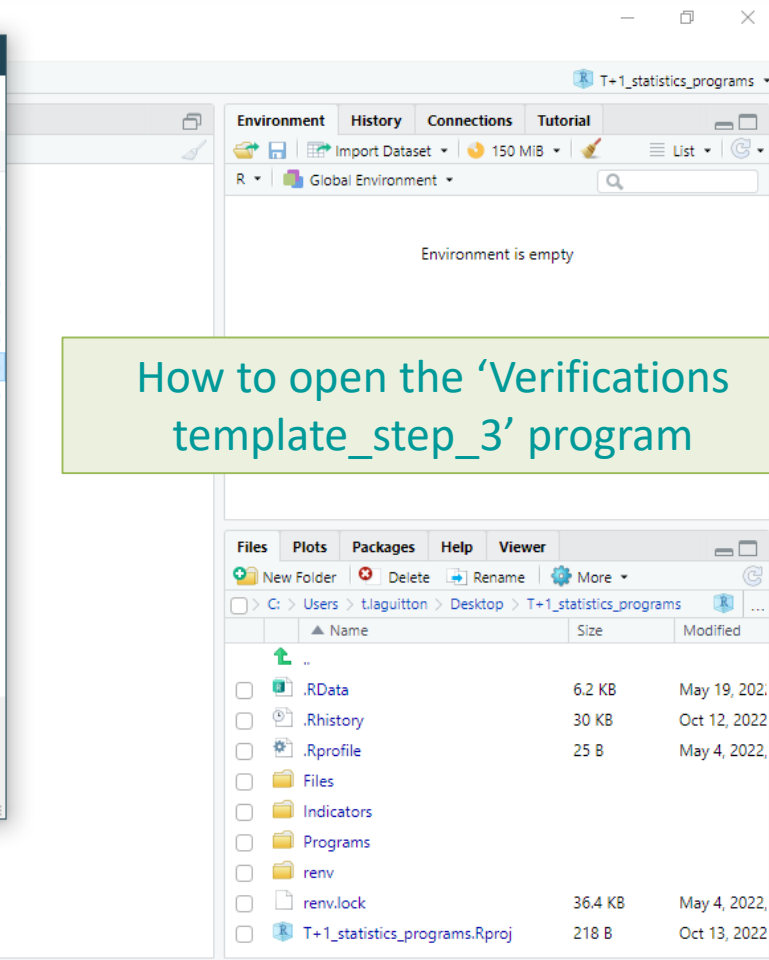
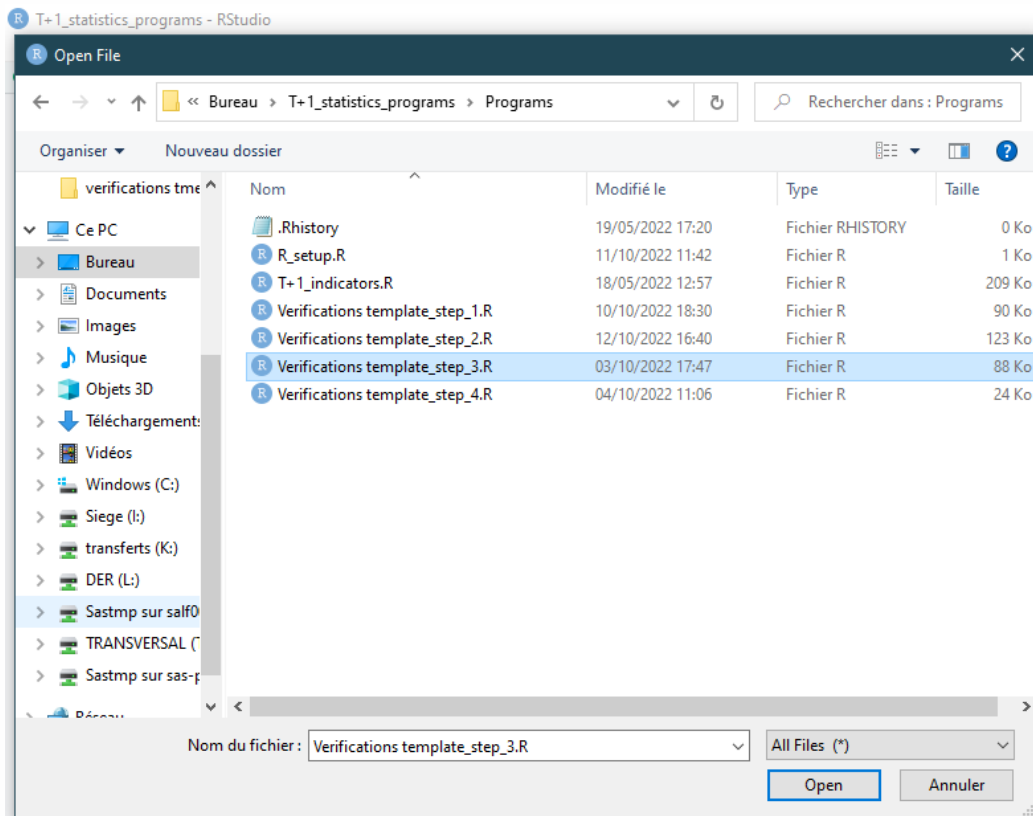
Name	Size	Modified
..		
.RData	6.2 KB	May 19, 2022
.Rhistory	30.2 KB	Aug 10, 2022
.Rprofile	25 B	May 4, 2022
Files		
Indicators		
Programs		
renv		
renv.lock	36.4 KB	May 4, 2022
T+1_statistics_programs.Rproj	218 B	Oct 10, 2022





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_3' program



How to open the 'Verifications template_step_3' program





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_3' program

The screenshot displays the RStudio interface with the following components:

- Source Editor:** Shows the R script 'Verifications template_step_3.R' with the following code:

```
1 ##### THIRD PROGRAMME OF VERIFICATION (DISTRIBUTION)
2
3 #Version of R used : 4.1.2
4 #Rstudio needs to be used for using the program in a simpler way : it is a tool for working on R programs
5 ###IMPORTANT : R software must be installed on your computer in order to install, run and use Rstudio
6
7 #Load the necessary packages each time you open Rstudio for running the program with the functions below
8 library("dplyr")
9 library("tidyverse")
10
11
12 #SET PARAMETERS :
13 #You will need to personalize the variables below with your parameters
14
15 ##Input file (csv) name :
16 #Indicate ("T0_data_collection_France_Step2_CORRECTED.csv") - here you need to replace "France" by the na
17 input_file="T+1_data_collection_France_Step2_CORRECTED(X).csv"
18 #The template you will import is the one you have created via verifications template_step_2 ; it should a
19
20
21 ##Output file (csv) name :
22 output_file="T+1_data_collection_France_Step3_VERIFIED.csv"
23
24 ##Indicate the separator used for saving your Excel file in csv ("," or ";")
25 separator=";"
26
27 ##If you have used a specific alphabet (greek, etc) in your template, please indicate "YES"
28 special_alphabet="NO"
29
30
31 #####
32 here=getwd() # Get the current working directory
33
34
```
- Environment Panel:** Shows 'Global Environment' and 'Environment is empty'.
- Files Panel:** Shows a file explorer view of the project directory 'C:\Users\t.laguillon\Desktop\T+1_statistics_programs'. The file list is as follows:

Name	Size	Modified
..		
.RData	6.2 KB	May 19, 2022
.Rhistory	30 KB	Oct 12, 2022
.Rprofile	25 B	May 4, 2022
Files		
Indicators		
Programs		
renv		
renv.lock	36.4 KB	May 4, 2022
T+1_statistics_programs.Rproj	218 B	Oct 13, 2022
- Console:** Shows the R prompt 'R 4.1.2 · C:\Users\t.laguillon\Desktop\T+1_statistics_programs/'.

A green callout box with the text 'Verifications template_step_3' program opened' is overlaid on the right side of the RStudio window.





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_3' program

```
T+1_statistics_programs - RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
Verifications template_step_3.R x
Source on Save
1 ##### THIRD PROGRAMME OF VERIFICATION (DISTRIBUTION)
2
3 #Version of R used : 4.1.2
4 #Rstudio needs to be used for using the program in a simpler way : it is a
5 ###IMPORTANT : R software must be installed on your computer in order to in
6
7 #Load the necessary packages each time you open Rstudio for running the pro
8 library("dplyr")
9 library("tidyverse")
10
11
12 #SET PARAMETERS :
13 #You will need to personalize the variables below with your parameters
14
15 ##Input file (csv) name :
16 #Indicate ("T0_data_collection_France_Step2_CORRECTED.csv") - here you need
17 input_file="T+1_data_collection_France_Step2_CORRECTED(X).csv"
18 #The template you will import is the one you have created via Verifications
19
20
21 ##Output file (csv) name :
22 output_file="T+1_data_collection_France_Step3_VERIFIED.csv"
23
24 ##Indicate the separator used for saving your Excel file in csv ("," or ";"
25 separator=";"
26
27 ##If you have used a specific alphabet (greek, etc) in your template, pleas
28 special_alphabet="NO"
29
30 |
31 #####
32 here=getwd() # Get the current working directory
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
2566
2567
2568
2569
2570
2571
2572
2573
2574
2575
2576
2577
2578
2579
2580
2581
2582
2583
2584
2585
2586
2587
2588
2589
2590
2591
2592
2593
2594
2595
2596
2597
2598
2599
2600
2601
2602
2603
2604
260
```



WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_3' program

Select the entire program (Ctrl+A) and press 'Run' (or Ctrl+Enter)

The program will run entirely.

```
1 ##### THIRD PROGRAMME OF VERIFICATION (DISTRIBUTION)
2
3 #Version of R used : 4.1.2
4 #Rstudio needs to be used for using the program in a simpler way : it is a tool for working on R programs
5 ##IMPORTANT : R software must be installed on your computer in order to install, run and use Rstudio
6
7 #Load the necessary packages each time you open Rstudio for running the program with
8 library("dplyr")
9 library("tidyverse")
10
11
12 #SET PARAMETERS :
13 #You will need to personalize the variables below with your parameters
14
15 ##Input file (csv) name :
16 #Indicate ("T0_data_collection_France_Step2_CORRECTED.csv") - here you need to replace with your file name
17 input_file="T+1_data_collection_France_Step2_CORRECTED(X).csv"
18 #The template you will import is the one you have created via Verifications template_step_2 ; it should be
19
20
21 ##Output file (csv) name :
22 output_file="T+1_data_collection_France_Step3_VERIFIED.csv"
23
24 ##Indicate the separator used for saving your Excel file in csv ("," or ";")
25 separator=";"
26
27 ##If you have used a specific alphabet (greek, etc) in your template, please indicate "YES"
28 special_alphabet="NO"
29
30
31 #####
32 here=getwd() # Get the current working directory
33
34
```

Environment: Global Environment, 152 MIB, Environment is empty.

Files: C:\Users\tlaguitton\Desktop> T+1_statistics_programs

Name	Size	Modified
..		
.RData	6.2 KB	May 19, 2022
.Rhistory	30 KB	Oct 12, 2022
.Rprofile	25 B	May 4, 2022
Files		
Indicators		
Programs		
renv		
renv.lock	36.4 KB	May 4, 2022
T+1_statistics_programs.Rproj	218 B	Oct 13, 2022





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_3' program

Program running

The 'Verifications template_step_3' program can take quite a long time (a few seconds to a few minutes) to run

When the program is running and has not finished its execution, a **stop sign** appears at the top right of the console. This means that the program is running and you must not close the 'T+1_statistics_programs – Rstudio' window or the program will stop running.

	lim_sup
0000000	1617.750000
0000000	385.250000
2894573	8.800000
9777955	1.550000
0000000	75.287500
1221245	1.650000
4210855	13.750000
0222045	0.250000
3552714	13.275000
0000000	224.000000
0000000	53.000000
0000000	1.000000
9888978	0.300000
7105427	8.800000
0111022	0.200000

```
R 4.1.2 · C:/Users/t.laguillon/Desktop/T+1_statistics_programs/
+ }
+ pb_outliers<-base::paste(pb_outliers,sep=" ", collapse=" ") #concatenate all the information from the pb_o
utliers list
+ step3_T0_data[i,"Problems_outlier"]<-pb_outliers #add this concatenated information in the right column/li
ne from your Step3_T0_data table
+ }
+ }
```





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_3' program

	Subcategory_name_VN	Q1	Q3	IQR	IQRx1.5	lim_inf
1	Cereals without added sugar/Energy_kJ_num	1531.50000	1566.000	34.50000	51.750000	1479.750000000
2	Cereals without added sugar/Energy_kCal_num	361.50000	371.000	9.50000	14.250000	347.250000000
3	Cereals without added sugar/Fat_num	4.67500	6.325	1.65000	2.475000	2.199999999
4	Cereals without added sugar/Saturated_fat_num	0.80000	1.100	0.30000	0.450000	0.349999999
5	Cereals without added sugar/Carbohydrates_num	60.22500	66.250	6.02500	9.037500	51.187500000
6	Cereals without added sugar/Sugar_num	0.90000	1.200	0.30000	0.450000	0.450000000
7	Cereals without added sugar/Protein_num	11.00000	12.100	1.10000	1.650000	9.350000000
8	Cereals without added sugar/Salt_num	0.00000	0.100	0.10000	0.150000	-0.150000000
9	Cereals without added sugar/Fibre_num	8.27500	10.275	2.00000	3.000000	5.275000000
10	Cereals without added sugar/Energy_as_consumed_kJ_num	224.00000	224.000	0.00000	0.000000	224.000000000
11	Cereals without added sugar/Energy_as_consumed_kCal_num	53.00000	53.000	0.00000	0.000000	53.000000000
12	Cereals without added sugar/Fat_as_consumed_num	1.00000	1.000	0.00000	0.000000	1.000000000
13	Cereals without added sugar/Saturated.fat_as_consumed_nu...	0.30000	0.300	0.00000	0.000000	0.299999999
14	Cereals without added sugar/Carbohydrates_as_consumed_...	8.80000	8.800	0.00000	0.000000	8.800000000
15	Cereals without added sugar/Sugar_as_consumed_num	0.20000	0.200	0.00000	0.000000	0.200000000

```
R 4.1.2 · C:/Users/t.laguilton/Desktop/T+1_statistics_programs/Files/
+ }else if(separator==""){
+   utils::write.csv(Step3_T0_data_f,file=output_file,na=" ",row.names=FALSE)
+ }
>
>
> print("The file should appear in the folder 'Files'")
[1] "The file should appear in the folder 'Files'"
> |
```

After running the program

Several tables appear next to the R program script. These tables are working tools for the R program but you don't need to look at them (only if there are errors or warnings in the console but in this case you have to tell us so we can help you).

At the end of the 'Verifications template_step_3' program execution, you should see the **sentence highlighted in yellow** in the console.

The program exports an excel file named `T+1_data_collection_country_Step3_VERIFIED.csv` to your "Files" folder in the "T+1_statistics_programs" folder.





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_3' program

Files

Fichier Accueil Partage Affichage

Rechercher dans : Files

Ce PC > Bureau > T+1_statistics_programs > Files

Nom	Modifié le	Type
Best-ReMaP_nomenclature.csv	12/07/2022 16:48	Fichier CSV Micro...
pre_existing_data_France.csv	03/10/2022 15:52	Fichier CSV Micro...
T+1_data_collection_France.csv	10/10/2022 17:10	Fichier CSV Micro...
T+1_data_collection_France_Step1_CORRECTED(X).csv	10/10/2022 17:04	Fichier CSV Micro...
T+1_data_collection_France_Step1_VERIFIED.csv	10/10/2022 17:13	Fichier CSV Micro...
T+1_data_collection_France_Step2_CORRECTED(X).csv	03/10/2022 15:11	Fichier CSV Micro...
T+1_data_collection_France_Step2_VERIFIED.csv	12/10/2022 16:46	Fichier CSV Micro...
T+1_data_collection_France_Step3_CORRECTED.csv	13/10/2022 10:27	Fichier CSV Micro...
T+1_data_collection_France_Step3_VERIFIED.csv	13/10/2022 10:27	Fichier CSV Micro...
wp5_verification_file.csv	30/08/2022 10:29	Fichier CSV Micro...
Years of interest.csv	03/10/2022 17:27	Fichier CSV Micro...

11 élément(s) | 1 élément sélectionné 1,56 Mo

Create a copy of the file exported by R

In the 'Files' folder, you will find the file exported by the 3rd verification program **T+1_data_collection_country_Step3_VERIFIED.csv**

→ You need to create a copy of this file and rename it “**T+1_data_collection_country_Step3_CORRECTED.csv**”

It is on this file that you will make the corrections and modifications indicated





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_3' program

T+1_data_collection_France_Step3_CORRECTED.csv - Excel

File Home Insert Page Layout Formulas Data Review View Tell me what you want to do... LAGUITTON Thomas Share

Clipboard Font Alignment Number Styles Cells Editing

Product_code

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U		
1	Product_code	Father_pr	Country	Year	Category	Subcategory	Bar_code	Assortme	Brand_nar	Brand_ow	Type_of	Legal_nan	Legal_nan	Commerci	Commerci	FOP_label	Nutri_Sco	Ingredi	Net_weig	Net_weig	Num	
2	1650		Ireland	2021	Breakfast	Cereals w	4,09E+12	No	Kavanagh	Aldi	Hard										0 g	
3	1651		Ireland	2021	Breakfast	Traditiona	4,09E+12	No	Kavanagh	Aldi	Hard											0 g
4	1653		Ireland	2021	Breakfast	Cereals w	4,09E+12	No	Kavanagh	Aldi	Hard											0 g
5	1654		Ireland	2021	Breakfast	Cereals w	4,09E+12	No	Kavanagh	Aldi	Hard											0 g
6	1655		Ireland	2021	Breakfast	Cereals w	4,09E+12	No	Kavanagh	Aldi	Hard											0 g
7	1657		Ireland	2021	Breakfast	Traditiona	4,09E+12	No	Specially	Aldi	Hard											0 g
8	1658	603	Ireland	2021	Breakfast	High-fibre	4,09E+12	No	Harvest M	Aldi	Hard											0 g
9	1659		Ireland	2021	Breakfast	Chocolate	4,09E+12	No	Harvest M	Aldi	Hard											5 g
10	1661	937	Ireland	2021	Breakfast	Chocolate	4,09E+12	No	Harvest M	Aldi	Hard											5 g
11	1663	586	Ireland	2021	Breakfast	Sweet cer	4,09E+12	No	Harvest M	Aldi	Hard											0 g
12	1665	916	Ireland	2021	Breakfast	Crunchy fr	4,09E+12	No	Harvest M	Aldi	Hard											0 g
13	1672		Ireland	2021	Breakfast	Crunchy m	4,09E+12	No	Harvest M	Aldi	Hard											0 g
14	1674	597	Ireland	2021	Breakfast	Sweet cer	4,09E+12	No	Harvest M	Aldi	Hard											0 g
15	1675		Ireland	2021	Breakfast	Crunchy fr	4,09E+12	No	Harvest M	Aldi	Hard											0 g
16	1676		Ireland	2021	Breakfast	Crunchy fr	4,09E+12	No	Harvest M	Aldi	Hard											0 g
17	1677	937	Ireland	2021	Breakfast	Chocolate	4,09E+12	No	Harvest M	Aldi	Hard											5 g
18	1680	996	Ireland	2021	Breakfast	Sweet cer	4,09E+12	No	Harvest M	Aldi	Hard											0 g
19	1681		Ireland	2021	Breakfast	Traditiona	4,09E+12	No	Harvest M	Aldi	Hard disc	Toasted H	Toasted H	Protein G	Protein G	Traffic lig			Wholegra			400 g
20	1682	850	Ireland	2021	Breakfast	Traditiona	4,09E+12	No	Harvest M	Aldi	Hard disc	Honey To	Honey To	Raisin & A	Raisin & A	Traffic lig			WHOLEGR			1000 g
21	1683		Ireland	2021	Breakfast	Traditiona	4,09E+12	No	Harvest M	Aldi	Hard disc	Oats with	Oats with	Really Nui	Really Nui	Traffic lig			Wholegra			500 g

T+1_data_collection_France_Step

Ready

Overview of the *T+1_data_collection_country_Step3_CORRECTED.csv* file which is a new version of your T+1 data collection template in which you will have to make modifications/corrections





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_3' program

The screenshot shows an Excel spreadsheet with the following data table:

	AI	AJ	AK	AL	AU	AV	AW	AX	AY
1	Fibre	Nutrient_Energy_as	Energy_as	F	category_	Subcategory	bar_code_starts_0	Problems_outlier	
2	13				1	739			
3	10				1	386			
4	8,8				1	739			
5	10,5				1	739			
6	11,5				1	739			
7	7,7				1	386			
8	11,9				1	143		Carbohydrates	
9	9,1				1	135		Energy_kj / Energy_kCal	
10	5				1	135			
11	1,8				1	745			
12	7				1	679			
13	3,5				1	680		Salt	
14	1,9				1	745		Fat / Saturated_fat	
15	8,3				1	679			
16	7,9				1	679			
17	3,1				1	135			
18	4				1	745		Carbohydrates / Protein	
19	5,6				1	386		Carbohydrates / Protein / Salt	
20	6,9				1	386			
21	7,1				1	386		Saturated_fat	

Appearance of a 'Problems_outlier' column which lists the fields for which the R program has found outliers

This column appears at the end of the table





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_3' program

	AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX	
1	Protein	Salt	Fibre	Nutrient_Energy_as	Energy_as	Fat_as	co	Saturated	Carbohydr	Sugar_as	Protein_a	Salt_as	cc	Fibre_as	Comment	Category	Subcategory	bar_code_starts_0	Problems_outlier
26	3	0,1	<0,5												3	614			
27	8,1	0,1	<0,5												3	614		Protein	
28	10	0,1	<0,5											no ingredi	3	612		Protein	
29	6,4	0,2	<0,5											no ingredi	3	613			
30	5,9	0,2	<0,5											no ingredi	2	614		Salt	
31	3,1	0,1	<0,5													613		Protein	
32	2,7	0,1	<0,5													615			
33	3,3	0,1	<0,5													615			
34	2,7	0,1		1,3												615			
35	8,2	0,1	<0,5													719			
36	2,8	0,1	<0,5													614			
37	3,6	0,1		0,8												215			
38	5,3	0,1	<0,5													611			
39	5	0,2	<0,5													611			
40	4,7	0,2		0,8												611			
41	26	5,2		0												333			
42	5,2	0,1	<0,5													611			
43	5,3	0,1	<0,5													611			
44	4	0,1	<0,5													615			
45	4,6	0,1	<0,5													613			

Example of problem found by the 'Verifications template_step_3' program

Line 28 → the column 'Problems_outlier' indicates a problem with an outlier for the protein content.

The R program identifies the protein content value for this product as different/aberrant from the values of other products in the same subcategory.





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_3' program

Example of problem found by the 'Verifications template_step_3' program

line 28 → the product belongs to the category 'Fresh dairy products desserts' and the subcategory 'Classic plain yoghurts and fermented milks with no added sugar'

	A	B	C	D	E	F	G	H	I	J	K	L
	Product_code	Father_product	Country	Year	Category_name	Subcategory_name	Bar_code	Assortment	Brand_name	Brand_owner	Type_of_packaging	Legal_status
26	1692		Ireland	2021	Fresh dairy products and desserts	Classic sweet yoghurts and fermented milks	4,09E+12	No	Brooklea	Aldi	Hard disc	Low Fat
27	1705		Ireland	2021	Fresh dairy products and desserts	Classic sweet yoghurts and fermented milks	4,09E+12	No	Brooklea	Aldi	Hard disc	Fat Free
28	1706		Ireland	2021	Fresh dairy products and desserts	Classic plain yoghurts and fermented milks with no added sugar	4,09E+11	No	Brooklea	Aldi	Hard disc	Natural
29	1707		Ireland	2021	Fresh dairy products and desserts	Gourmet plain yoghurts and fermented milks with no added sugar	4,09E+11	No	Brooklea	Aldi	Hard disc	Natural
30	1708		Ireland	2021	Fresh dairy products and desserts	Classic sweet yoghurts and fermented milks	4,09E+12	No	Brooklea	Aldi	Hard disc	Fat Free
31	1709		Ireland	2021	Fresh dairy products and desserts	Gourmet plain yoghurts and fermented milks with no added sugar	4,09E+12	No	Brooklea	Aldi	Hard disc	Greek S
32	1727		Ireland	2021	Fresh dairy products and desserts	Gourmet sweet yoghurts and fermented milks	4,09E+12	No	Brooklea	Aldi	Hard disc	Peach a
33	1732	550	Ireland	2021	Fresh dairy products and desserts	Gourmet sweet yo						cc Raspber
34	1733	550	Ireland	2021	Fresh dairy products and desserts	Gourmet sweet yo						cc Raspber
35	1736		Ireland	2021	Fresh dairy products and desserts	Classic sweetened						cc Fat Free
36	1746		Ireland	2021	Fresh dairy products and desserts	Classic sweet yogh						cc Strawbe
37	1752		Ireland	2021	Fresh dairy products and desserts	Fresh desserts wit						cc Clotted
38	1758		Ireland	2021	Fresh dairy products and desserts	Artificially-sweete						cc 2x fat fr
39	1762	405	Ireland	2021	Fresh dairy products and desserts	Artificially-sweete						cc Fat Free
40	1764		Ireland	2021	Fresh dairy products and desserts	Artificially-sweete						cc Fat free
41	24411		Ireland	2021	Delicatessen meats and similar	Cured ham						cc A select
42	1766		Ireland	2021	Fresh dairy products and desserts	Artificially-sweete						cc Fat free
43	1767	405	Ireland	2021	Fresh dairy products and desserts	Artificially-sweete						cc Fat free
44	1770		Ireland	2021	Fresh dairy products and desserts	Gourmet sweet yo						cc Greek s
45	1771		Ireland	2021	Fresh dairy products and desserts	Gourmet plain yog						cc Greek S



WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_3' program

	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV	AW
1	Saturated	Carbohydr	Sugar	Protein	Salt	Fibre	Nutrient	Energy_as	Energy_as	Fat_as_co	Saturated	Carbohydr	Sugar_as	Protein_a	Salt_as_cc	Fibre_as	Comment	Category	Subcategory	bar_code_starts
26	2,1	13	12	3	0,1	<0,5												3	614	
27	<0,1	13	11	8,1	0,1	<0,5												3	614	
28	<0,1	4,9	3,5	10	0,1	<0,5											no ingredi	3	612	
29	7,6	3,7	3,2	6,4	0,2	<0,5														
30	0,1	9,1	7,7	5,9	0,2	<0,5														
31	6,4	4,3	3,4	3,1	0,1	<0,5														
32	4,5	16	14	2,7	0,1	<0,5														
33	4,6	16	15	3,3	0,1	<0,5														
34	5,1	13	12	2,7	0,1		1,3													
35	0,2	11	10	8,2	0,1	<0,5														
36	0,7	11	10	2,8	0,1	<0,5														
37	7,9	17	6,6	3,6	0,1		0,8													
38	0,1	8,3	6,8	5,3	0,1	<0,5														
39	0,1	6,6	5,4	5	0,2	<0,5														
40	0,1	6,3	5,2	4,7	0,2		0,8													
41	6,5	0	0	26	5,2		0													
42	0,2	8,5	6,8	5,2	0,1	<0,5														
43	0,1	8,3	6,8	5,3	0,1	<0,5														
44	3,7	12	11	4	0,1	<0,5														
45	4,4	4,4	3,5	4,6	0,1	<0,5														

Example of a problem found by the 'Verifications template_step_3' program

Line 28 → the column 'Problems_outlier' indicates a problem with an outlier for the protein content.

You must check that the value entered for the protein content is correct or/and that the classification (category and/or subcategory) of the product is correct.

If correction is required, you will then need to correct the protein content value or the category/subcategory for this product.





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_3' program

Terms that may appear in the 'Problems outlier' field following the 'Verifications template_step_3' program, their meaning and what to do

<u>Problem</u>	<u>Meaning</u>	<u>Solution</u>
<ul style="list-style-type: none"> • Energy_kCal • Energy_kJ • Fat • Saturated_fat • Carbohydrates • Sugar • Protein • Salt • Fibre • Energy_as_consumed_kCal • Energy_as_consumed_kJ • Fat_as_consumed • Saturated_fat_as_consumed • Carbohydrates_as_consumed • Sugar_as_consumed • Protein_as_consumed • Salt_as_consumed • Fibre_as_consumed 	<ul style="list-style-type: none"> • The nutritional value of the product for this nutrient appears to be an outlier compared to the nutritional value for this nutrient of other products in the same subcategory. 	<p>→ Check the pictures of the product, to be sure that the value entered in the template is the correct one. If not, you must enter the correct value directly in the Excel file in .csv format.</p> <p><u>And/or</u></p> <p>→ Check that the category and subcategory associated to the product are the correct ones (you can help you with the WP5 classification guidelines that have been created for each food category). If not, you must enter the correct information (category name + code and/or subcategory name + code) directly in the Excel file in .csv format.</p> <p><u>Else</u></p> <p>→ Nutritional value and subcategory entered for this product are the correct ones, no correction is needed. You must indicate in the 'Comments' field: <i>"outliers checked"</i>.</p> <p>Be careful! A product can have wrong values + wrong classification, it is important to check both for the product.</p>



WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_3' program

- **Be careful! When saving the file, you must be sure that the barcodes appear in full and not in scientific format. If this is the case, you must repeat the procedure on [pages 24→28](#).**
- Once the corrections have been made and the barcodes are in the correct format, you can save the file *T+1_data_collection_country_Step3_CORRECTED.csv* and close it.
- You will have to run this corrected file again in the 'Verifications template_step_3' program to make sure you haven't missed a check.

Your Rstudio interface must be cleaned up before running the 'Verifications template_step_3' program again.

All cleaning steps are described on [pages 62→68](#).





WORK Package 5 – Reformulation and processed food monitoring

2nd running of 'Verifications template_step_3' program

```
1 ##### THIRD PROGRAMME OF VERIFICATION (DISTRIBUTION)
2
3 #Version of R used : 4.1.2
4 #Rstudio needs to be used for using the program in a simpler way : it is a tool for working on R programs
5 ###IMPORTANT : R software must be installed on your computer in order to instal
6
7 #Load the necessary packages each time you open Rstudio for running the program
8 library("dplyr")
9 library("tidyverse")
10
11
12 #SET PARAMETERS :
13 #You will need to personalize the variables below with your parameters
14
15 ##Input file (csv) name :
16 #Indicate ("T0_data_collection_France_Step2_CORRECTED.csv") - here you need to
17 input_file="T+1_data_collection_France_Step3_CORRECTED.csv"
18 #The template you will import is the one you have created via verifications tem
19
20
21 ##Output file (csv) name :
22 output_file="T+1_data_collection_France_Step3_VERIFIED2.csv"
23
24 ##Indicate the separator used for saving your Excel file in csv ("," or ";")
25 separator=";"
26
27 ##If you have used a specific alphabet (greek, etc) in your template, please in
28 special_alphabet="NO"
29
30
31 #####
32
```

Setting parameters of the second running of 'Verifications template_step_3' program

For this second running of the 'Verifications template_step 3', the only fields you need to change are the names of the input file and the output file

Input_file =
"T+1_data_collection_country_Step3_CORRECTED.csv"

Output_file =
"T+1_data_collection_country_Step3_VERIFIED2.csv"

You must also indicate the separator and if you have a specific alphabet as in the first run of the program.





WORK Package 5 – Reformulation and processed food monitoring

2nd running of 'Verifications template_step_3' program

- At the end of this second run, you get in your "files" folder a file called:
"T+1_data_collection_country_**Step3**_VERIFIED2.csv"
- You must create a copy of this file and call it :
"T+1_data_collection_country_**Step3**_CORRECTED2.csv"

→ It is on this file that you will make the modifications following the checks





WORK Package 5 – Reformulation and processed food monitoring

2nd running of 'Verifications template_step_3' program

On the file : *T+1_data_collection_country_Step3_CORRECTED2.csv*

If no problems appear
Or
If problems that appear are indicated as
"checked"/"verified" in the 'Comments' field

Your file has been corrected successfully!

You can save the file
(be careful of bar codes format!)
then proceed to the next data verification
program 'Verifications template_step_4'
(see next slides)

If problems that has not
been checked appear

You must correct the new problems

You must run the program a third time :

Input file = *T+1_data_collection_country_Step3_CORRECTED2.csv*
Output file = *T+1_data_collection_country_Step3_VERIFIED3.csv*

New file : *T+1_data_collection_country_Step3_CORRECTED3.csv*

...

You will need to run the 'Verifications template_step_3' program until no more problems appear or only problems that have been indicated as already verified/checked in the 'Comments' field.





3) Running of the programs

A. Part 1 : R setup program

B. Part 2 : Verification programs and template cleaning/standardization

i. 1st verification program : 'Verifications template_step_1'

ii. 2nd verification program : 'Verifications template_step_2'

iii. 3rd verification program : 'Verifications template_step_3'

iv. 4th verification program : 'Verifications template_step_4'





4th verification program : 'Verifications template_step_4'

Presentation of the 'Verifications template step 4' program :

- The 4th verification program allows you to highlight problems in the pairing of your T+1 data with your pre-existing data.
- This program will allow, for example, to :
 - Check that the information entered matches between 2 paired products (Brand, category name and code, subcategory name and code)
 - Check that the father_product_codes indicated in the T+1 data do exist in the pre-existing data file.
 - Check that unique products_code given in your T+1 data doesn't 'exist in your pre-existing data
 - ...





WORK Package 5 – Reformulation and processed food monitoring

4th verification program : 'Verifications template_step_4'

Requirements before starting the program 'Verifications template_step_4' :

- The programs 'Verifications template_step_1', 'Verifications template_step_2' and 'Verifications template_step_3' should have been run on your data
- You should no longer have any problems appearing or only problems that have been notified as verified after running the program 'Verifications template_step_3'
- You must have your template in your possession and it must now be called:
T+1_data_collection_ **country** _Step3_CORRECTED(X).csv (with the name of your own country)
(X) is the number of the last file exported and corrected after the last run of the first verification program)

You need to make sure that the barcodes in this file appear in full and not in scientific format (see procedure [pages 24→28](#))

- You must also have in your possession your pre-existing data file in **.csv format** called pre_existing_data_ **country** .csv and your file Years_of_interest.csv that must have been filled in indicating the years chosen in your pre-existing data for pairing with the T+1 data.

Your Rstudio interface must have been cleaned up before running the program.
All cleaning steps are described on [pages 62→68](#).



WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_4' program

The screenshot shows the RStudio interface with the 'File' menu open. The 'Open File...' option is highlighted. In the background, a Windows File Explorer window is open to the directory 'C:\Users\t.laguitton\Desktop\T+1_statistics_programs'. The Explorer shows a list of files and folders:

Name	Size	Modified
..		
.RData	6.2 KB	May 19, 2022
.Rhistory	30.2 KB	Aug 10, 2022
.Rprofile	25 B	May 4, 2022
Files		
Indicators		
Programs		
renv		
renv.lock	36.4 KB	May 4, 2022
T+1_statistics_programs.Rproj	218 B	Oct 10, 2022

A text box in the center of the RStudio window contains the text: 'How to open the 'Verifications template_step_4' program'.





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_4' program

The screenshot shows the RStudio interface with the 'Open File' dialog box open. The dialog box is displaying the contents of the 'T+1_statistics_programs' folder. The 'Programs' subfolder is selected. The R Environment pane shows 'Environment is empty'. A green callout box contains the text 'How to open the 'Verifications template_step_4' program'.

Nom	Modifié le	Type	Taille
Files	09/08/2022 17:08	Dossier de fichiers	
Indicators	09/08/2022 16:57	Dossier de fichiers	
Programs	10/10/2022 14:48	Dossier de fichiers	
renv	09/08/2022 15:52	Dossier de fichiers	
.RData	19/05/2022 17:20	R Workspace	7 Ko
.Rhistory	10/08/2022 12:12	Fichier R HISTORY	31 Ko
.Rprofile	04/05/2022 10:39	Fichier R PROFILE	1 Ko
renv.lock	04/05/2022 10:01	Fichier LOCK	37 Ko
T+1_statistics_programs.Rproj	10/10/2022 14:49	R Project	1 Ko

Name	Size	Modified
..		
.RData	6.2 KB	May 19, 2022
.Rhistory	30.2 KB	Aug 10, 2022
.Rprofile	25 B	May 4, 2022
Files		
Indicators		
Programs		
renv		
renv.lock	36.4 KB	May 4, 2022
T+1_statistics_programs.Rproj	218 B	Oct 10, 2022





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_4' program

How to open the 'Verifications template_step_4' program

Nom	Modifié le	Type	Taille
.Rhistory	19/05/2022 17:20	Fichier RHISTORY	0 Ko
R_setup.R	11/10/2022 11:42	Fichier R	1 Ko
T+1_indicators.R	18/05/2022 12:57	Fichier R	209 Ko
Verifications template_step_1.R	10/10/2022 18:30	Fichier R	90 Ko
Verifications template_step_2.R	12/10/2022 16:40	Fichier R	123 Ko
Verifications template_step_3.R	13/10/2022 10:51	Fichier R	88 Ko
Verifications template_step_4.R	04/10/2022 11:06	Fichier R	24 Ko

Name	Size	Modified
..		
.RData	6.2 KB	May 19, 2022
.Rhistory	30 KB	Oct 12, 2022
.Rprofile	25 B	May 4, 2022
Files		
Indicators		
Programs		
renv		
renv.lock	36.4 KB	May 4, 2022
T+1_statistics_programs.Rproj	218 B	Oct 13, 2022





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_4' program

The screenshot displays the RStudio interface with the 'Verifications template_step_4.R' script open. The script contains the following code:

```
1 ##### FOURTH PROGRAMME OF VERIFICATION (VERIFICATION FOR PAIRED PRODUCTS)
2 #####THIS PROGRAM SHOULD BE RUN AFTER THE THREE STEPS OF VERIFICATION HAVE BEEN COMPLETED
3
4 #Version of R used : 4.1.2
5 #Rstudio needs to be used for using the program in a simpler way : it is a tool for working on R programs
6 ###IMPORTANT : R software must be installed on your computer in order to install, run and use Rstudio
7
8
9 ##dplyr : 1.0.7
10 ##tidyverse : 1.3.1
11
12 #Load the necessary packages each time you open Rstudio for running the program with the functi
13 library("dplyr")
14 library("tidyverse")
15
16
17 #SET PARAMETERS :
18 #You will need to personalize the variables below with your parameters
19
20 ##Input file (csv) name :
21 #Indicate ("T+1_data_collection_France_Step3_CORRECTED.csv") - here you need to replace "France" by the na
22 input_file="T+1_data_collection_France_Step3_CORRECTED(X).csv"
23
24 ##Pre-existing data file (csv) name :
25 #Indicate ("Name_of_pre-existing_data_file.csv")
26 preexisting_data_file="pre_existing_data_France.csv"
27
28 ##Nomenclature file (csv) name :
29 ##The csv file should be saved in the same folder as the one containing your data set in csv
30 nomenclature_file="Best-ReMaP_nomenclature.csv"
31
32
```

The R console at the bottom shows the R version and the current working directory:

```
R 4.1.2 · C:/Users/tlaguilton/Desktop/T+1_statistics_programs/Files/
```

The file explorer window on the right shows the directory structure:

Name	Size	Modified
..		
.RData	6.2 KB	May 19, 2021
.Rhistory	30 KB	Oct 12, 2022
.Rprofile	25 B	May 4, 2022
Files		
Indicators		
Programs		
renv		
renv.lock	36.4 KB	May 4, 2022
T+1_statistics_programs.Rproj	218 B	Oct 13, 2022

A green callout box with the text 'Verifications template_step_4' program opened' is overlaid on the right side of the RStudio window.





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_4' program

```
T+1_statistics_programs - RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
Verifications template_step_4.R x
Source on Save
16
17 #SET PARAMETERS :
18 #You will need to personalize the variables below with your parameters
19
20 ##Input file (csv) name :
21 #Indicate ("T+1_data_collection_France_Step3_CORRECTED.csv") - here you
22 input_file="T+1_data_collection_France_Step3_CORRECTED(X).csv"
23
24 ##Pre-existing data file (csv) name :
25 #Indicate ("Name_of_pre-existing_data_file.csv")
26 preexisting_data_file="pre_existing_data_France.csv"
27
28 ##Nomenclature file (csv) name :
29 ##The csv file should be saved in the same folder as the one containing
30 nomenclature_file="Best-ReMaP_nomenclature.csv"
31
32 ##Output file (csv) name :
33 output_file="T+1_data_collection_France_Step4_VERIFIED.csv"
34
35 ##Years of interest file :
36 years_of_interest_file="Years of interest.csv"
37
38 ##Indicate the separator used for saving your Excel file in csv ("," or ";")
39 separator=";"
40
41 ##If you have used a specific alphabet (greek, etc) in your template,
42 special_alphabet="NO"
43
44 #####
45 here=getwd() # Get the current working directory
46
47
34:1 (Top Level)
Console Terminal Jobs x
R 4.1.2 · C:/Users/laguillon/Desktop/T+1_statistics_programs/Files/
> |
```

Setting parameters of the 4th verification program

In the first run of the 4th verification program, you need to change the name of the country with your own country name in the input file (line 17) and the output file (line 33)

In this program, you will also need to import your pre-existing data file. To do this, you need to enter the name of your own country in the file name line 26.

You must also indicate the separator and if you have a specific alphabet as in the program 'Verifications template_step_1', 'Verifications template_step_2' and 'Verifications template_step_3'.

Example :

- **Input_file** = "T+1_data_collection_Ireland_Step2_CORRECTED(X).csv"
(X) is the number of the last file exported and corrected after the last run of the second verification program)
- **Preexisting_data_file** = "pre_existing_data_Ireland.csv"
- **Output_file** = "T+1_data_collection_Ireland_Step3_VERIFIED.csv"
- **Separator** = ";"
- **Special_alphabet** = "NO"





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_4' program

Select the entire program (Ctrl+A)
and press 'Run' (or Ctrl+Enter)
The program will run entirely.

```
1 ##### FOURTH PROGRAMME OF VERIFICATION (VERIFICATION FOR PAIRED PRODUCTS)
2 ###THIS PROGRAM SHOULD BE RUN AFTER THE THREE STEPS OF VERIFICATION HAVE BEEN COMPLETED
3
4 #Version of R used : 4.1.2
5 #Rstudio needs to be used for using the program in a simpler way : it is a tool for working on R programs
6 ###IMPORTANT : R software must be installed on your computer in order to install, r
7
8
9 ##dplyr : 1.0.7
10 ##tidyverse : 1.3.1
11
12 #Load the necessary packages each time you open Rstudio for running the program with
13 library("dplyr")
14 library("tidyverse")
15
16
17 #SET PARAMETERS :
18 #You will need to personalize the variables below with your parameters
19
20 ##Input file (csv) name :
21 #Indicate ("T+1_data_collection_France_Step3_CORRECTED.csv") - here you need to replace "France" by the name
22 input_file="T+1_data_collection_France_Step3_CORRECTED(x).csv"
23
24 ##Pre-existing data file (csv) name :
25 #Indicate ("Name_of_pre-existing_data_file.csv")
26 preexisting_data_file="pre_existing_data_France.csv"
27
28 ##Nomenclature file (csv) name :
29 ##The csv file should be saved in the same folder as the one containing your data set in csv
30 nomenclature_file="Best-ReMaP_nomenclature.csv"
31
32
```





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_3' program

The screenshot shows the RStudio interface. The top-left pane displays a table with 6 rows and 5 columns. The top-right pane shows the Environment, History, Connections, and Tutorial tabs. The bottom-left pane shows the Console with R code and its output.

Category_code	Category_name	Year_of_interest_1	Year_of_interest_2
1	Bread products	2016-2017	2016-2017
2	Breakfast cereals	2016-2017	2016-2017
3	Delicatessen meats and similar	2016-2017	2016-2017
4	Fresh dairy products and desserts	2016-2017	2016-2017
5	Soft drinks	2016-2017	2016-2017
6	Baby food	2017	2017

```
R 4.1.2 · C:/Users/Elaguitton/Desktop/T+1_statistics_programs/Files/
> #Create the final table T1_data_f to export
> T1_data_f<-dplyr::select(T1_data,"Product_code","subcategory_code","bar_code_starts_0
>
> #Export the T1_data_f table
> if(separator==";"){
+   utils::write.csv2(T1_data_f,file= output_file,na=" ",row.names=FALSE)
+ }else if(separator==""){
+   utils::write.csv(T1_data_f,file= output_file,na=" ",row.names=FALSE)
+ }
>
> print("The file should appear in the folder 'Files'")
[1] "The file should appear in the folder 'Files'"
>
```

After running the program

Several tables appear next to the R program script. These tables are working tools for the R program but you don't need to look at them (only if there are errors or warnings in the console but in this case you have to tell us so we can help you).

At the end of the 'Verifications template_step_4' program execution, you should see the **sentence highlighted in yellow** in the console.

The program exports 2 excel files named `T+1_data_collection_country_Step4_VERIFIED.csv` and `Brand_names_extraction.csv` to your "Files" folder in the "T+1_statistics_programs" folder.





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_4' program

Nom	Modifié le	Type	Taille
Best-ReMaP_nomenclature.csv	12/07/2022 16:48	Fichier CSV Micro...	34 Ko
Brand_names_extraction.csv	14/10/2022 11:36	Fichier CSV Micro...	5 Ko
pre_existing_data_France.csv	03/10/2022 15:52	Fichier CSV Micro...	334 Ko
T+1_data_collection_France.csv	10/10/2022 17:10	Fichier CSV Micro...	1 431 Ko
T+1_data_collection_France_Step1_CORRECTED(X).csv	10/10/2022 17:04	Fichier CSV Micro...	1 459 Ko
T+1_data_collection_France_Step1_VERIFIED.csv	10/10/2022 17:13	Fichier CSV Micro...	1 614 Ko
T+1_data_collection_France_Step2_CORRECTED(X).csv	03/10/2022 15:11	Fichier CSV Micro...	1 466 Ko
T+1_data_collection_France_Step2_VERIFIED.csv	12/10/2022 16:46	Fichier CSV Micro...	1 621 Ko
T+1_data_collection_France_Step3_CORRECTED(X).csv	14/10/2022 11:35	Fichier CSV Micro...	
T+1_data_collection_France_Step3_VERIFIED.csv	13/10/2022 10:27	Fichier CSV Micro...	
T+1_data_collection_France_Step4_VERIFIED.csv	14/10/2022 11:36	Fichier CSV Micro...	
wp5_verification_file.csv	30/08/2022 10:29	Fichier CSV Micro...	
Years of interest.csv	03/10/2022 17:27	Fichier CSV Micro...	

New files appearing

In the 'Files' folder, you will find 2 files exported by the 4th verification program :

- Brand_names_extraction.csv
- and
- T+1_data_collection_country_Step4_VERIFIED.csv

13 élément(s) | 2 éléments sélectionnés 1,56 Mo





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_4' program

	A	B	C	D
1	Product_code	Father_product_code	Brand_name_T+1	Brand_name_preexisting
2	3966		3 Muller Corner	Muller
3	39701		4 Muller Corner	Muller
4	39702		4 Muller Corner	Muller
5	39641	26	Muller Corner	Muller
6	39642	26	Muller Corner	Muller
7	3969	29	Muller Corner	Muller
8	2613	51	Muller Corner	Muller
9	3968	54	Muller Corner	Muller
10	39551	57	Muller Corner	Muller
11	39552	57	Muller Corner	Muller
12	3962	59	Muller Corner	Muller
13	3797	68	Activia	Danone
14	2554	70	Activia	Danone
15	3799	70	Activia	Danone
16	3800	71	Activia	Danone
17	3796	72	Activia	Danone
18	2345	86	Milbona	Lidl Milbona
19	3868	119	Fage Total	Fage
20	4061	120	Tesco Finest	Tesco
21	1776	149	Duneeen Dairy	Duneeen

“Brand_names_extraction.csv” file

This file shows the products paired between your pre-existing data and your T+1 data that have ‘brand_names’ which are not identical.

→ This file is only used to check if there are **no major pairing errors**.

Example :

brand_name_preexisting_data = *Kellogg’s*

brand_name_T+1 = *Carrefour*

In this case, you need to check the matching of the two products: maybe the product has been matched to the wrong product in the pre-existing data and you need to find the right product in the pre-existing data, maybe there is a typing error in the father_product_code, etc.

→ If two brand names are different but it is a **slight difference** (“Lidl Milbona” vs. “Milbona”, “Muller” vs. “Muller Corner”, “Kellogg’s” vs. “Kelloggs”, etc.), then you do **not need to modify** the information.

We advise you not to change the 'brand_name' in your pre-existing data as you may no longer have the product pictures in your possession to be sure of what you are changing.





WORK Package 5 – Reformulation and processed food monitoring

Running of 'Verifications template_step_4' program

Windows File Explorer window showing the directory structure and file list. The selected file is `T+1_data_collection_France_Step4_CORRECTED.csv`.

Nom	Modifié le	Type	Taille
Best-ReMaP_nomenclature.csv	12/07/2022 16:48	Fichier CSV Micro...	34 Ko
Brand_names_extraction.csv	14/10/2022 11:36	Fichier CSV Micro...	5 Ko
pre_existing_data_France.csv	03/10/2022 15:52	Fichier CS	
T+1_data_collection_France.csv	10/10/2022 17:10	Fichier CS	
T+1_data_collection_France_Step1_CORRECTED(X).csv	10/10/2022 17:04	Fichier CS	
T+1_data_collection_France_Step1_VERIFIED.csv	10/10/2022 17:13	Fichier CS	
T+1_data_collection_France_Step2_CORRECTED(X).csv	03/10/2022 15:11	Fichier CS	
T+1_data_collection_France_Step2_VERIFIED.csv	12/10/2022 16:46	Fichier CS	
T+1_data_collection_France_Step3_CORRECTED(X).csv	14/10/2022 11:35	Fichier CS	
T+1_data_collection_France_Step3_VERIFIED.csv	13/10/2022 10:27	Fichier CS	
T+1_data_collection_France_Step4_CORRECTED.csv	14/10/2022 11:36	Fichier CS	
T+1_data_collection_France_Step4_VERIFIED.csv	14/10/2022 11:36	Fichier CS	
wp5_verification_file.csv	30/08/2022 10:29	Fichier CS	
Years of interest.csv	03/10/2022 17:27	Fichier CS	

Create a copy of the file exported by R

→ You need to create a copy of the file `T+1_data_collection_country_Step4_VERIFIED.csv` and rename it `"T+1_data_collection_country_Step4_CORRECTED.csv"`

It is on this file that you will make the corrections and modifications indicated





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_4' program

T+1_data_collection_France_Step4_CORRECTED.csv - Excel

File Home Insert Page Layout Formulas Data Review View Tell me what you want to do... LAGUITTON Thomas Share

Clipboard Font Alignment Number Styles Cells Editing

Product_code

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	
1	Product_c	Father_pr	Country	Year	Category	Subcateg	Bar_code	Assortme	Brand_na	Brand_ow	Type_of	Legal_nan	Legal_nan	Commerci	Commerci	EOP_label	Nutri_Sco	Ingredi	Net	weig	Net_weig	Numb
2	1650		Ireland	2021	Breakfast	Cereals w	4,09E+12	No	Kavanagh'	Aldi	Ha										500 g	
3	1651		Ireland	2021	Breakfast	Traditiona	4,09E+12	No	Kavanagh'	Aldi	Ha										500 g	
4	1653		Ireland	2021	Breakfast	Cereals w	4,09E+12	No	Kavanagh'	Aldi	Ha										500 g	
5	1654		Ireland	2021	Breakfast	Cereals w	4,09E+12	No	Kavanagh'	Aldi	Ha										750 g	
6	1655		Ireland	2021	Breakfast	Cereals w	4,09E+12	No	Kavanagh'	Aldi	Ha										1500 g	
7	1657		Ireland	2021	Breakfast	Traditiona	4,09E+12	No	Specially	Aldi	Ha										500 g	
8	1658	6003	Ireland	2021	Breakfast	High-fibre	4,09E+12	No	Harvest M	Aldi	Ha										750 g	
9	1659		Ireland	2021	Breakfast	Chocolate	4,09E+12	No	Harvest M	Aldi	Ha										375 g	
10	1661	937	Ireland	2021	Breakfast	Chocolate	4,09E+12	No	Harvest M	Aldi	Ha										375 g	
11	1663	586	Ireland	2021	Breakfast	Cereals w	4,09E+12	No	Harvest M	Aldi	Ha										500 g	
12	1665	916	Ireland	2021	Breakfast	Crunchy fr	4,09E+12	No	Harvest M	Aldi	Ha										500 g	
13	331		Ireland	2021	Breakfast	Crunchy m	4,09E+12	No	Harvest M	Aldi	Ha										500 g	
14	1674	597	Ireland	2021	Breakfast	Sweet cer	4,09E+12	No	Harvest M	Aldi	Ha										500 g	
15	1675		Ireland	2021	Breakfast	Crunchy fr	4,09E+12	No	Harvest M	Aldi	Ha										500 g	
16	1676		Ireland	2021	Breakfast	Crunchy fr	4,09E+12	No	Harvest M	Aldi	Ha										500 g	
17	1677	937	Ireland	2021	Breakfast	Chocolate	4,09E+12	No	Harvest M	Aldi	Ha										375 g	
18	1680	996	Ireland	2021	Breakfast	Sweet cer	4,09E+12	No	Harvest M	Aldi	Ha	Hard discc	Rice and V	Rice and V	Original B	Original B	Traffic ligh		RICE (66%		500 g	
19	1681		Ireland	2021	Breakfast	Traditiona	4,09E+12	No	Harvest M	Aldi	Ha	Hard discc	Toasted H	Toasted H	Protein Gi	Protein Gi	Traffic ligh		Wholegra		400 g	
20	1682	850	Ireland	2021	Breakfast	Traditiona	4,09E+12	No	Harvest M	Aldi	Ha	Hard discc	Honey To	Honey To	Raisin & A	Raisin & A	Traffic ligh		WHOLEGR		1000 g	
21	1683		Ireland	2021	Breakfast	Traditiona	4,09E+12	No	Harvest M	Aldi	Ha	Hard discc	Oats with	Oats with	Really Nu	Really Nu	Traffic ligh		Wholegra		500 g	

T+1_data_collection_France_Step

Ready

Overview of the *T+1_data_collection_country_Step4_CORRECTED.csv* file which is a new version of your T+1 data collection template in which you will have to make modifications/corrections





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_4' program

The screenshot shows an Excel spreadsheet with the following data table:

	AI	AJ	AK	AL	AM	AN	AO	AU	AV	AW	AX
1	Fibre	Nutrient	Energy_as	Energy_as Fat				category	Subcategory	bar_code_starts_0	Problems
2	13							1	739		
3	10							1	386		
4	8,8							1	739		
5	10,5							1	739		
6	11,5							1	739		
7	7,7							1	386		
8	11,9							1	143		nonexistent_father_product_code
9	9,1							1	135		
10	5							1	135		
11	1,8							1	739		verify_subcategory_name / verify_subcategory
12	7							1	679		
13	3,5							1	680		duplicated_code
14	1,9							1	745		
15	8,3							1	679		
16	7,9							1	679		
17	3,1							1	135		
18	4							1	745		
19	5,6							1	386		
20	6,9							1	386		
21	7,1							1	386		

A green text box in the center of the spreadsheet contains the following text:

Appearance of a 'Problems' column which lists the errors/inconsistencies the R program has found

This column appears at the end of the table





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_4' program

The screenshot shows an Excel spreadsheet with the following data in the 'Problems' column:

Row	Problems
1	
2	
3	
4	
5	
6	
7	
8	nonexistent_father_product_code
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	

Example of problem found by the 'Verifications template_step_4' program

Line 8 → the column '*Problems*' indicates **nonexistent_father_product_code**

The R program does not recognise the father_product_code for this product. This means that the code entered in the field 'fater_product_code' does not exist in your pre-existing data file.





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_4' program

Example of a problem found by the 'Verifications template_step_4' program

Line 8 → the column '*Problems*' indicates **nonexistent_father_product_code**

The code "6003" doesn't exist in your pre-existing data file.

You must correct this information by checking if the product in your T+1 collection (1658) has a father product in your pre-existing data and finding the correct father_product_code.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Product_code	Father_product_code	Country	Year	Category	Subcategory	Bar_code	Assortment	Brand_name	Brand_ow	Type_of	Legal_nan	Legal_nan	Commercial	Commercial	FOP_label	Nutri_Sco	Ingredient	Net_wei
2	1650		Ireland	2021	Breakfast	Cereals w	4,09E+12	No	Kavanagh	Aldi	Hard disc	Gluten Fre	Gluten Fre	Gluten Fre	Gluten Fre	None fron		Gluten Fre	50
3	1651		Ireland	2021	Breakfast	Traditiona	4,09E+12	No	Kavanagh	Aldi	Hard disc	Wholegra	Wholegra	High Ome	High Ome	Traffic ligh		Jumbo Oa	50
4	1653		Ireland	2021	Breakfast	Cereals w	4,09E+12	No	Kavanagh	Aldi	Hard disc	Instant Po	Instant Po	Microwav	Microwav	Traffic ligh		Wholegra	50
5	1654		Ireland	2021	Breakfast	Cereals w	4,09E+12	No	Kavanagh	Aldi	Hard disc	Wholegra	Wholegra	Organic Ju	Organic Ju	Traffic ligh		Oat Flakes	75
6	1655		Ireland	2021	Breakfast														
7	1657		Ireland	2021	Breakfast														
8	1658	6003	Ireland	2021	Breakfast														
9	1659		Ireland	2021	Breakfast														
10	1661	937	Ireland	2021	Breakfast														
11	1663	586	Ireland	2021	Breakfast														
12	1665	916	Ireland	2021	Breakfast														
13	331		Ireland	2021	Breakfast														
14	1674	597	Ireland	2021	Breakfast														
15	1675		Ireland	2021	Breakfast														
16	1676		Ireland	2021	Breakfast														
17	1677	937	Ireland	2021	Breakfast														
18	1680	996	Ireland	2021	Breakfast														
19	1681		Ireland	2021	Breakfast														
20	1682	850	Ireland	2021	Breakfast														
21	1683		Ireland	2021	Breakfast														





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_4' program

Terms that may appear in the 'Problems' field following the 'Verifications template step 4' program, their meaning and what to do

<u>Problem</u>	<u>Meaning</u>	<u>Action</u>
Nonexistent_father_product_code	The father_product_code associated with this product does not exist in your pre-existing data.	<p>→ Check if the product has a father product in the pre-existing data file using information such as the product's subcategory, brand name, commercial_name elements, etc.</p> <p>→ If the product has a father product, enter the correct product_code from the pre-existing data in the 'Father_product_code' field.</p> <p>→ If the product does not have a father product, delete the wrong <i>father_product_code</i> that has been entered.</p>
Duplicated_code	The product_code of this product is already assigned to a product in your pre-existing data.	<p>→ You must assign a new product code to this product, making sure that it does not exist in your 'T+1_data_collection_country_step_4_CORRECTED.csv' file or in your pre-existing data file.</p>
Not_paired_with_year_of_interest	<ul style="list-style-type: none"> The father product does not belong to the pre-existing data collection year of interest chosen to make comparisons/indicators with T+1 data. This means that this pairing will not be taken into account for making the indicators. 	<p>→ You must check if the T+1 product cannot be paired with another product in your preexisting data file which has been collected on the year of interest for the category.</p> <p><i>Please note that the products not paired with the year of interest will not be considered as paired products for the computing of Indicators (the pairing will not be taken into account for this work)</i></p>



Excel file to modify after 'Verifications template_step_4' program

Terms that may appear in the 'Problems' field following the 'Verifications template_step_4' program, their meaning and what to do

<u>Problem</u>	<u>Meaning</u>	<u>Action</u>
<ul style="list-style-type: none">• Verify_category_name• Verify_category_code	The product and its father product have different category names and codes	<ul style="list-style-type: none">→ You must check that you have classified the product of your T+1 Best-ReMaP collection in the right category.→ If the problem is with a product in your pre-existing data that is classified in the wrong category, you must correct the error in your pre-existing data and return the file to us.
<ul style="list-style-type: none">• Verify_subcategory_name• Verify_subcategory_code	The product and its father product have different subcategory names and codes	<ul style="list-style-type: none">→ You must check that you have classified the product of your T+1 Best-ReMaP collection in the right subcategory.→ If the problem is with a product in your pre-existing data that is classified in the wrong category, you must correct the error in your pre-existing data and return the file to us.



WARNING! This is not necessarily an error, the product of the T+1 Best-ReMaP collection may have changed subcategory compared to the pre-existing data.

Example: the product has been reformulated and the sugar has been removed, etc

You must still check that you have not made a classification mistake.





WORK Package 5 – Reformulation and processed food monitoring

Excel file to modify after 'Verifications template_step_4' program

- **Be careful! When saving the file, you must be sure that the barcodes appear in full and not in scientific format. If this is the case, you must repeat the procedure on [pages 24→28](#).**
- Once the corrections have been made and the barcodes are in the correct format, you can save the file *T+1_data_collection_country_Step3_CORRECTED.csv* and close it.
- You will have to run this corrected file again in the 'Verifications template_step_4' program to make sure you haven't missed a check.

Your Rstudio interface must be cleaned up before running the 'Verifications template_step_4' program again.

All cleaning steps are described on [pages 62→68](#).





WORK Package 5 – Reformulation and processed food monitoring

2nd running of 'Verifications template_step_4' program

```
11
12 #Load the necessary packages each time you open Rstudio for running the program with the functions below
13 library("dplyr")
14 library("tidyverse")
15
16
17 #SET PARAMETERS :
18 #You will need to personalize the variables below with your parameters
19
20 ##Input file (csv) name :
21 #Indicate ("T+1_data_collection_France_Step3_CORRECTED.csv") - here you need to
22 input_file="T+1_data_collection_France_Step4_CORRECTED.csv"
23
24 ##Pre-existing data file (csv) name :
25 #Indicate ("Name_of_pre-existing_data_file.csv")
26 preexisting_data_file="pre_existing_data_France.csv"
27
28 ##Nomenclature file (csv) name :
29 ##The csv file should be saved in the same folder as the one containing your data
30 nomenclature_file="Best-REMAP_nomenclature.csv"
31
32 ##Output file (csv) name :
33 output_file="T+1_data_collection_France_Step4_VERIFIED2.csv"
34
35 ##Years of interest file :
36 years_of_interest_file="Years of interest.csv"
37
38 ##Indicate the separator used for saving your Excel file in csv ("," or ";")
39 separator=";"
40
41 ##If you have used a specific alphabet (greek, etc) in your template, please indicate
42 special_alphabet="NO"
43
44 #####
45 here=getwd() # Get the current working directory]
46
47
```

Setting parameters of the second running of 'Verifications template_step_4' program

For this second running of the 'Verifications template_step_4', the only fields you need to change are the names of the input file and the output file

Input_file =
"T+1_data_collection_country_Step4_CORRECTED.csv"

Output_file =
"T+1_data_collection_country_Step4_VERIFIED2.csv"

You must also indicate the separator and if you have a specific alphabet as in the first run of the program.





WORK Package 5 – Reformulation and processed food monitoring

2nd running of 'Verifications template_step_4' program

- At the end of this second run, you get in your "files" folder a file called:
"T+1_data_collection_country_**Step4**_VERIFIED2.csv"

- You must create a copy of this file and call it :
"T+1_data_collection_country_**Step4**_CORRECTED2.csv"

→ It is on this file that you will make the modifications following the checks





WORK Package 5 – Reformulation and processed food monitoring

2nd running of 'Verifications template_step_4' program

On the file : *T+1_data_collection_country_Step4_CORRECTED2.csv*

If no problems appear

Or

If problems that appear are indicated as "checked"/"verified" in the 'Comments' field

If problems that has not been checked appear

Your file has been corrected successfully!

You must correct the new problems

You can save the file
(be careful of bar codes format!)

You must run the program a third time :

Input file = *T+1_data_collection_country_Step4_CORRECTED2.csv*

Output file = *T+1_data_collection_country_Step4_VERIFIED3.csv*

New file : *T+1_data_collection_country_Step4_CORRECTED3.csv*

...

You will need to run the 'Verifications template_step_4' program until no more problems appear or only problems that have been indicated as already verified/checked in the 'Comments' field.





End of the 4 verification programs

→ You should now have a file called : *T+1_data_collection_country_Step4_CORRECTED(X).csv*
(X) is the number of the last file exported and corrected after the last run of the second verification program)

This file is the final version of your data collection template after you have done all the checks and corrected all the errors.

→ You must create a copy of this file and save it in **.xlsx format**
(You can call this file : *T+1_data_collection_country_final.xlsx* for example)

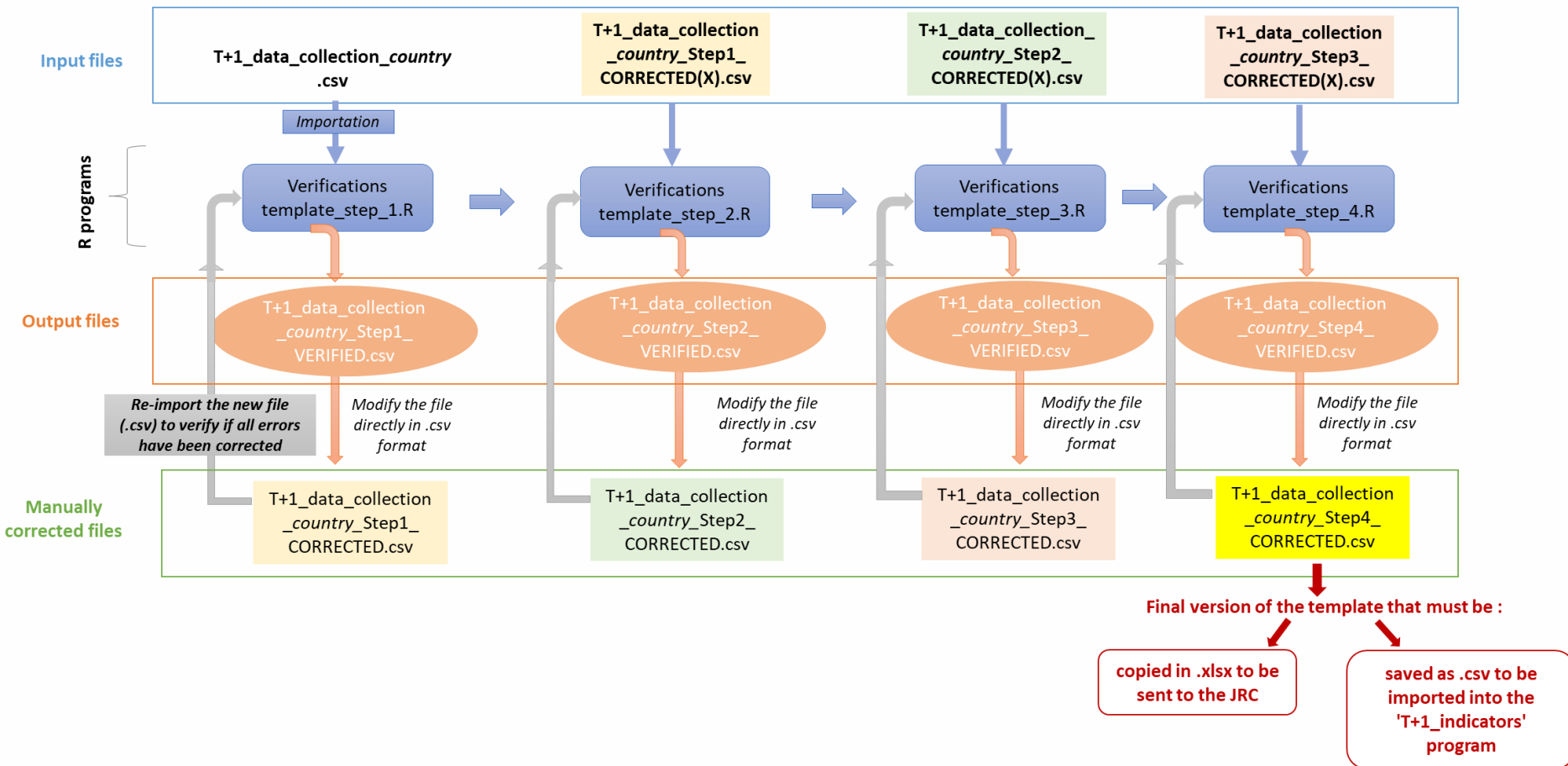
This copy in **.xlsx format** will be the final version of your data collection template that will be transmitted to the **JRC**.





WORK Package 5 – Reformulation and processed food monitoring

Summary of the 4 verification programs





4) Next steps





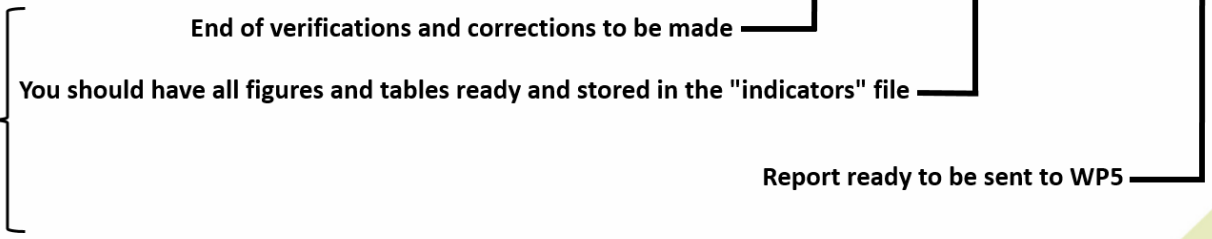
Timeline of the task 5.5.1

2022												2023								
j	f	m	a	m	j	j	a	s	o	n	d	j	f	m	a	m	j			
M16	M17	M18	M19	M20	M21	M22	M23	M24	M25	M26	M27	M28	M29	M30	M31	M32	M33			
Task 5.4.1 : Data collection and data treatment (data entry and encoding) Mar 2022 - Feb 2023																				
Collection of pictures Mar 2022		Data treatment (data entry and encoding) Apr 2022 - Jan 2023 (February 2023 at the latest)																		
		Soft drinks			Breakfast cereals			Bread products			Fresh dairy products and desserts			Delicatessen meats and similar						
												Task 5.5.1 : Production of statistics Nov 2022 - Apr 2023 (6 months)								
												Verifications steps Nov - Jan 2022 <i>(3 months)</i>			Production of indicators Feb 2023 <i>(1 month)</i>			Production of the report Mar - Apr 2023 <i>(2 months)</i>		

1st Nov 2022

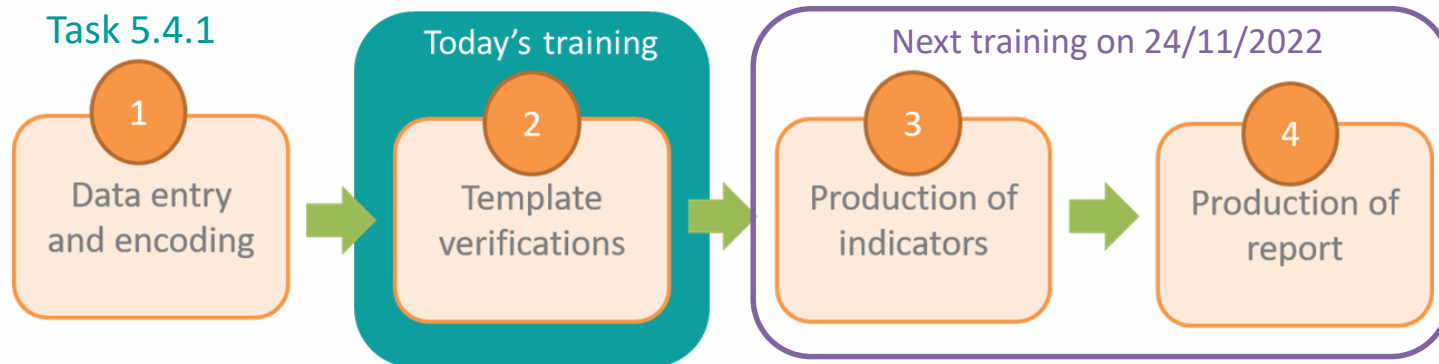
Please send us feedback to confirm that the R and Rstudio software are installed
 If you have any problems before this date, please let us know so that we can help you

These deadlines are indications so that you can organize yourselves as well as possible





Next steps



It is important to finish each step before moving on to the next

- Downloading of the R and Rstudio software and the documents needed for task 5.5.1 (working folder on the intranet) **before 1st November 2022**.
Please give us feedback so that we can help you as much as possible if you have difficulties downloading.
- A next training is planned for **24th November 2022** to present steps 3 and 4 (production of indicators and report)
- If you have any questions or need help at any time during the completion of the steps in Task 5.3.2, you can contact us at: wp5_bestremap@anses.fr
We will also be able to hold individual meetings to best assist you during this task.





Best-ReMaP
Healthy Food for a Healthy Future



**Thank you
for your attention!**

ANSES

wp5_bestremap@anses.fr

The Joint Action focusing on the implementation of validated best practices in
nutrition – Best-ReMaP

This presentation arises from the Joint Action Best-ReMaP. This JA is addressing the adaption, replication and implementation of effective health interventions, based on practices that have been proven to work in the areas of food reformulation, framing of food marketing and public procurement of healthy food in public settings. This presentation was funded by the European Union's Health Programme (2014-2020). The content of this presentation represents the views of the author only and is his/her sole responsibility; it cannot be considered to reflect the views of the European Commission and/or the Consumers, Health, Agriculture and Food Executive Agency (CHAFEA) or any other body of the European Union. The European Commission and the Agency do not accept any responsibility for use that may be made of the information it contains.