Instructions to research assistant

Julie Lenoir

October 2018

Terms of contract

The contract is for a term of one month, starting on October 8th, and includes the Toussaint vacation. The Research Assistant will be remunerated at 750 Eur upon completion, which corresponds to 50 hours of work.

Instructions

The RA will help building the correspondence between two different product classifications used in India. The purpose is to match the products codes from the ASIC87 classification to the ASIC08 classification.

The ASIC87 classification is organized by industry, and by type of product (input, output, intermediary items, packing items, etc). This implies that some items description can be found under different codes. Those codes are composed of 8 digits: the 4 first ones represent the industry code. The last 4 digits are product code. The first digits of those 4 (that is the 5th digit) gives an indication of the type of product:

1 Raw materials

2 Chemicals and auxiliary/ intermediary materials

3 Packing items

4 Consumable stores

5 Products and by-products

7 Machineries, components and accessories

8 Imported materials

The ASIC08 was organised by type of product (dairy, meat, fabrics, etc).

We want to find the perfect match for each ASIC87. However, most of the time, we cannot expect to find a unique match for the ASIC87. In that case, the less proposed matches, the better.

Both classifications (ASIC87 and ASIC08) are provided in a PDF format. The RA is also provided with the industry classification which corresponds to the first 4 digits of the ASIC87 codes.

Those classifications will contain typos and grammatical mistake. Please do not correct them and leave the products description provided as they are.

The RA will be provided with an Excel file which contains a macro that have been programmed to account for progression. The macro can be executed using a keyboard shortcut that is to be chosen by the RA. When the number of matches hit (or exceed) a thousand (1000), a message box will appear to warn the RA. At that point, the RA will extract the matched lines (using the descending ordering command or the sorting command), and will cut/paste those lines in a new sheet. Those lines will be then reviewed and taken away for the Excel file.

The first Excel file contains around 7300 lines and unique ASIC87 description. The next Excel file contains around 6700 lines and non unique ASIC87 descriptions.

Git

The progression will be shared using Git. Please stick to the following procedure and **do not try to solve problems by yourself**. Git is very sensitive, and a bad manipulation could lead you to involuntarily delete everything for ever.

1. Open the Terminal.

2. Navigate to your repository using `cd path`

3. Update your folder with the last changes: `git pull`

4. Work on the files and close them.

5. Back to the Terminal, check your changes: `git status`

6. If everything is right, track your changes: `git add .` for all the files, `git add filename` for one particular file

7. Save your changes: `git commit -m "commit message"` (adapt the commit message to the situation)

8. Share your changes: `git push`

File structure

The Excel file is structured in the following way:

	Number of matches					
Identifier	Progression dummy	Industry description	ASIC87 code	ASIC87 description	Hand-matching	Propositions ...
Line number. Please never modify that column.	1 indicates a match has been selected in that line.				If none of the propositions seem satisfying or you think another ASIC08 could be matching, put the ASIC08 code here.	Bold (CMD + B) the proposition(s) you think is a match.

Useful resources

Since we are dealing with industrial products, there will be some products that have specific uses or description that do not ring a bell. In that case, Wikipédia will be very useful to get more information on the production process of an item.

Here are different websites that can also be useful:

- The US government chemistry dictionary: https://pubchem.ncbi.nlm.nih.gov/

- An online dictionnary that contains indian words: https://www.merriam-webster.com/

- Glossary of indian spices : https://www.vegrecipesofindia.com/indian-spices-glossary-of-indian-spices-in-english-and-hindi/