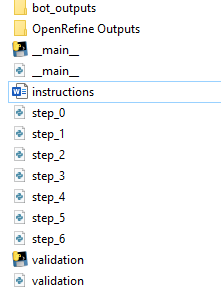
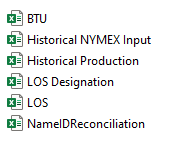
**How to use the code**

1. You should see these items in this folder excluding the instructions document; \_\_main\_\_.exe, validation.exe, bot\_outputs, OpenRefine Outputs, and the source code. If you see a pycache folder just ignore it.

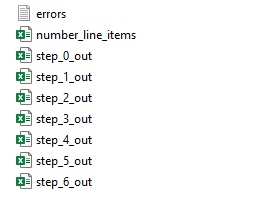


1. You should **STRICTLY** follow the naming convention of all the files that were provided in the OpenRefine Outputs. **ANY MISTAKES IN THE NAMING OF INPUT DATA FILES** will give an error when running the main file, so check the OpenRefine Outputs, and make sure that all the files that you give in the folder should follow this naming convention:



where LOS will be the initial file you will give that contains the unprocessed data so that the steps will start to perform on this file. The rest of the files should be named **EXACTLY** as you can see in the provided image, otherwise the code will **NOT** work.

1. Once you put the files in the OpenRefine Outputs folder, the rest is easy. Just run the \_\_main\_\_.exe file, and wait until the code stops.
2. After that you can go to bot\_outputs folder and there you should see something like this:



where your file of concern is the **STEP\_6\_OUT** file. The rest from step\_0\_out till step\_5\_out are just the steps that the code performs, and these files were used by the code only to process the next step. The final output file is the step\_6\_out file.

Note: Ignore the number\_line\_items file, it just stores the total number of line items of each record.

**NOTE FOR ERRORS FILE**: First run the validation exe, and then the errors log file will be generated in the bot outputs folder. See if the file is empty, then run the main.exe file, otherwise, fix any errors generated and then run the main.exe file to get your output.