**NAM Python Script**

*PO\_Delivery/Material Master Data Column Values*

(Fuzzy Wuzzy/ Levehnstein distance String Matching and extraction algorithm)

1. **Technical description:**

Package installation guidelines:

1. Unzip NAM\_PythonProject.7z to a desired destination folder;

**NB**: It is mandatory that all script dependencies (venv) as well as .xlsx source spreadsheet files (PODelivery\_Data.xlsx, Materials\_Master.xlsx) are positioned in the same folder as the.exe (NAM\_PythonScript.exe) file;

1. Execute NAM\_PythonScript.exe (cmd window will automatically appear), wait for the .exe file to prompt for an input on the user’s behalf. This might take up to 90 seconds, depending on the volume of data inside source files.
2. The text string is in the cmd window is the following: “Please, input a percentage threshold value: ” - user must insert a number ranging from 0 – 100 (number indicates the first percentage match between the Material Group from Materials\_Master.xlsx and Short text from PODelivery\_Data.xlsx;
3. Press Enter key, wait for the NAM\_PythonScript.exe script to finish execution (cmd window will disappear once finished);
4. Result - two excel spreadsheets (BestMatch\_Mat\_Number\_MatchPercentages.xlsx, Filtered\_BestMatch\_Mat\_Number\_MatchPercentages.xlsx) will be produced as final result;
5. Total output columns in both files are:

* **Material number** (string) - material number values inserted from the Material\_Master.xlsx spreadsheet into the result document;
* **Material description** (string) – column comprising Manufacturer Part number values from Material\_Master.xlsx spreadsheet;
* **Material first match** (string) – cross-referencing a string utilizing the Fuzzy Wuzzy algorithm, finding closest match in terms of percentage;
* **First best match %** (integer) – first best match expressed in % numbers;
* **Material second match** (string) - cross-referencing a string utilizing the Fuzzy Wuzzy algorithm, finding second closest match in terms of percentage;
* **Second best match %** (integer) - second best match expressed in % numbers;
* **Material third match** (string) - cross-referencing a string utilizing the Fuzzy Wuzzy algorithm, finding third closest match in terms of percentage;
* **Third best match %** (integer) - third best match expressed in % numbers.

1. Prerequisite for the script to run and produce result - all files have to be unzipped in a single folder by the user;
2. **User data output:**

Description of result data

* Files produced are in .xlsx format;
* Names of files - BestMatch\_Mat\_Number\_MatchPercentages.xlsx, Filtered\_BestMatch\_Mat\_Number\_MatchPercentages.xlsx
* First file – comprises all string matches cross-referenced between column values from PODelivery\_Data.xlsx (Short text) and Material\_Master.xlsx (Manufacturer Part No) spreadsheet;
* Second file – contains filtered data on the basis of the integer value entered by the user in the cmd window during the process of execution. Best matches are organized in descending order, starting from the highest match and closest to 100 % and moving towards the lesser value.