

## Fuzzy-Logic-Matching using Google Colab

Steps for the above :

1. installing the upgrade gspread which is a python API for google sheets.
2. Install fuzzywuzzy which is a python library that calculates the differences between sequences and patterns, also known as levenshtein distance. It works by using NLP to process text strings like words, and then clustering to group the similar text pieces together.
3. Install python-levenshtein
4. Insert the code for importing the google sheets/authentication etc., directly from code snippets "< >"
5. Import pandas library and convert the google sheets into its respective dataframes.
6. Define a function which basically converts all the uppercases into lowercases
7. Define columns for the respective tables, i.e. coClean in this case, using the function defined above.
8. Converting the series into a list and defining another function which takes a row (dictionary or pandas Series) with a specific key "coClean", performs a fuzzy match of its value against a list of choices, retrieves the confidence score of the best match, and adds this score to the row under the key "Confidence in Match", then returns the modified row.
9. 'dfMatching = dfTable.apply(fm, axis=1)' applies the fm function to each row of the DataFrame dfTable, resulting in a new DataFrame dfMatching where each row has been augmented with a confidence score indicating the quality of the fuzzy match for the value in the "coClean" column.
10. Merging the dataframes and mapping it to google spreadsheets, result is displayed on the 'Matching' sheet.