

This project aimed to find matches between two databases to complete a larger dataset. df1 contained the values to be searched, and df2 contained all the information. The code was developed because it was not possible to find exact matches because the first column of df2 was not completely clean and contained unnecessary characters. The code searches for text matches in df2 using one or more columns of df1. If a match is found, it returns the value of “Result ID” in df1.

**Here are the questions I was interested in answering**

How can I perform a bulk partial match analysis since Excel is unsuitable for this task?

What variables are interesting to perform the analysis, especially if there are more than two matches?

What should be the structure of my final database?

**I took the following steps to create my analysis**

Define the variables for which you want to complete the information.

Determine the type of match required as a result.

Establish the structure of the final database.

**Here are my takeaways**

The matching code allows for a significant reduction in manual activity.

The database is completed in an automated way.

The code is adjustable to up to 3 variables.