

DATA RECONCILIATION SYSTEM

Problem statement:

To build a system to identify reconciled data among ledger data in comparison with target data. The system must be able to preprocess data: reorder target columns to match ledger columns, reorder target rows to match ledger rows and finally generate a reconciliation sheet indicating reconciled and mismatched data. The system should be flexible to all kinds of reconciliation.

Data Description:

Data used for building the system comprises two files, one, the company ledger file which contains data related to the GST bills manually entered by a team from Infocorp and two, data from GST2B filings filed by clients.

System Inputs

The system requires the following data to process the reconciliation sheet

1. Ledger data (company data)
2. GST portal data (GST2B filing)
3. Match ¹primary key column of ledger to GST portal data
4. Match ²secondary key column of ledger to GST portal data
5. Provide the list of columns to be reconciled

System Outputs

System will generate the following outputs

1. Updated company data for the reconciliation columns
2. Updated reordered and cleaned GST portal data for the reconciliation columns
3. Reconciliation sheet to indicate the entries that are reconciled and the entries that aren't reconciled

¹ primary key: Column in the company data and GST portal data for which the entries in the columns have an exact 100% match with one another

² secondary key: Any column in the company data and GST portal data which match one another

System Description

Once the system is provided with the inputs, it will first reorder the columns of the GST portal data to match the ledger data, this is done by firstly, reordering a few rows in the GST portal data, following which, similarity is checked between columns using WER(word error rate) to pick the most similar columns and match them. The next step is to reorder the rows of the GST portal data and this is done by performing a similarity check between rows of GST portal data and ledger data using WER and picking the most similar rows to match them. The reconciliation sheet is then prepared, by creating an empty sheet with the same column names and equal number of rows as the ledger data. After which, The sheet entries are filled with 'reconciled' if the entries of the ledger match the entries of the GST portal data and filled with 'Mismatch error' if the entries of the ledger do not match the entries of the GST portal data. The system design/ logic of the application is given below

System Design

