Power Excel Reader

Path: com.opus.backend.features.PowerExcelReader.class

This class provides functionality to read, merge, and cleanse Excel files. The class uses the Apache POI library for reading and writing Excel files and the Spire.XLS library for merging files.

1. Class fields: The class has several fields to store the input/output file names, paths, primary key, Excel workbook, columns, and regular expressions for cleansing the data.
2. Constructor: The constructor (**PowerExcelReader()**) takes five parameters: primary\_key, outputFileName, filespath, outpath, and dictionary\_file. It initializes the class fields, merges the input Excel files, and sets up the workbook object and the regular expressions HashMap based on the input dictionary file.
3. mergeFiles(): This method uses the Spire.XLS library to merge multiple Excel files into a single output file.
4. getDictionary(): This method reads a dictionary file (a CSV file containing column names and regular expressions) and returns an ArrayList of String arrays. The first element of each String array is the column name, and the second element is the corresponding regular expression.
5. getTableRowCol(): This method takes a sheet number and returns an int array containing the row and column indices where the table starts in the sheet. It does this by matching cell values with the regular expressions in the **regularExp** HashMap.
6. excelCleansingTable(): This method cleanses the data in the Excel file by exporting only the expected columns for calculations. It iterates through each sheet, finds the start of the table, and exports the required columns to a new sheet in a new workbook called **newWorkbook**. It also handles various cell types, such as dates, times, and cells containing formulas.
7. isRowEmpty(): This helper method checks if a given row is empty or not.
8. getNumberOfSheetAvaliable(): This method returns the number of sheets in the workbook.
9. close(): This method writes the **newWorkbook** to the output file path, closes the workbook and the OPCPackage objects, and deletes the temporary files.

The **PowerExcelReader** class provides a high-level API to read, merge, and cleanse Excel files by leveraging Apache POI and Spire.XLS libraries. This class can be used to process Excel files with complex structures, extract specific columns based on dictionary file patterns, and create a cleansed output Excel file.