Project Description

- **The Word Statistics Calculator:**

  - Is a Java Swing application designed to analyze and display word statistics for text files within a specified directory.
  - The application provides information such as the:
    - number of words
    - longest word, shortest word
    - counts of specific words like "is," "are," and "you" for each file.
    - Additionally, global statistics for the entire directory are presented.

Project Components

1-User Interface

- **Directory Selection:**
  Users can browse and select a directory for analysis.
- **Include Subdirectories:**
  An option to include or exclude subdirectories during the analysis.
- **Start Processing Button:**
  Initiates the analysis process.
- **Tables for Statistics:**
  Two tables display per-file and directory statistics, including file names, word counts, longest and shortest words, and counts of specific words.

2- Backend Processing

- **Multithreaded Processing:**
  The application utilizes multithreading to concurrently analyze multiple files for improved performance.
- **Word Statistics Calculation:**
  It calculates word counts, longest and shortest words, and counts of specific words ('is', 'are', 'you') for each file.
- **Directory and Global Statistics:**
  Updates and displays statistics for individual directories and global statistics considering all files.
- **The FileProcessor class:**
  implemented as a separate thread, handles the concurrent processing of individual files. It updates global statistics as it analyzes each file.

Team Members Role

1-GUI

مصطفى محمد فوزي محمود

-Responsible for designing and implementing the graphical user interface using Swing components. This includes input fields, buttons, and tables for displaying file and directory statistics.

2-Main Program

حبيبة محمد سعيد

ندى هاني العمروسي

-Responsible for developing the main program logic, including the initiation of the GUI, handling user inputs, and orchestrating the overall flow of the application.

3-Threads

نورهان حسام عبد العزيز

-Responsible for implementing multi-threading functionality. This includes creating and managing threads to concurrently process multiple files within the specified directory.

4-Synchronization

دعاء مجدي محمد

مصطفى محمد فوزي محمود

-Responsible for implementing synchronization mechanisms to ensure thread safety. This includes using locks to manage access to shared data structures and prevent race conditions.

Code Documentation

WordStatisticsApp Class

- **Attributes**:
  - `directoryField`: JTextField for user input of the directory path.
  - `includeSubdirectoriesCheckbox`: JCheckBox for including subdirectories in file processing.
  - `tableModel`: DefaultTableModel for the main statistics table.
  - `statisticsTable`: JTable for displaying file statistics.
  - `start_Button`: JButton for triggering the start of file processing.
  - `table`: DefaultTableModel for the directory statistics table.
  - `statisticsTable1`: JTable for displaying directory statistics.
  - `filesToProcess`: List of files to be processed.
  - `lock`: Object for synchronization.

- **Methods:**
  - `WordStatisticsApp`: Constructor for initializing the GUI components.
  - `browseDirectory`: Method for handling directory selection using JFileChooser.
  - `startprocess`: Method for initiating the file processing.
  - `processDirectory`: Method for processing the selected directory and its files.
  - `processFilesInDirectory`: Recursive method for processing files in a directory.
  - `startProcessingThreads`: Method for starting multiple threads to process files concurrently.
  - `updateDirectoryStatistics`: Method for updating directory-specific statistics in the table.

- `updateGlobalStatistics`: Method for updating or adding a row for global statistics in the table.
- `processFile`: Method for processing individual files and calculating word statistics.
- `countWords`, `findLongestWord`, `findShortestWord`, `countOccurrences`: Helper methods for word statistics calculations.
- `main`: Entry point for the application.

FileProcessor Class

- **Methods**:
  - `run`: Overridden run method to define the behavior of file processing threads.

WordStats Class

- **Attributes**:
  - `numWords`, `longestWord`, `shortestWord`, `isCount`, `areCount`, `youCount`: Word statistics attributes.
- **Methods**:
  - `updateFromRow`: Method for updating statistics based on the provided values.