CSV Data Analysis Tool - Documentation

This script provides a menu-driven interface for analyzing CSV data. It uses the `whiptail` library for interactive menu and input prompts.

Requirements:

Bash shell

`whiptail` library (usually installed by default)

Usage:

1. Save the script: Save the code provided as a `.sh` file (e.g., `csv_analysis.sh`).

2. Make it executable: Run `chmod +x csv_analysis.sh` in the terminal.

3. Run the script: Execute the script from the terminal by typing `./csv_analysis.sh`.

4,or just bash csv_analysis.sh

Menu Options:

The script presents a menu with the following options:

1. Display row and column count: Shows the number of rows and columns in the CSV file.

2. List unique values in a column: Lists all unique values in a specified column.

3. Show column names (header): Displays the header row containing column names.

4. Find minimum and maximum values: Finds the minimum and maximum values in a specified column (only works for numeric data).

5. Find most frequent value: Determines the most frequent value in a specified column.

6. Calculate summary statistics: Calculates mean, median, and standard deviation for values in a specified column (only works for numeric data).

7. Filter and extract data: Filters the CSV file based on a value in a specified column and saves the filtered data to a new file.

8. Sort the CSV file: Sorts the CSV file based on the values in a specified column.

9. Save analysis results: Saves the results of all analyses to a text file.

10. Exit: Exits the script.

Input Prompts:

For most options, the script will prompt you for the following information:

CSV file path: Enter the path to your CSV file.

Column number: Specify the column number you want to analyze (starting from 1).

Filter value: Enter the value to filter data by (for option 7).

Output file name: Enter the desired name for the output file (for option 9).

Examples:

Display row and column count:

```bash
./csv_analysis.sh
```

Select option 1 from the menu and enter the path to your CSV file.

List unique values in column 3:

```bash
./csv_analysis.sh
```

Select option 2 from the menu, enter the path to your CSV file, and then enter "3" as the column number.

Find most frequent value in column 2:

```bash
./csv_analysis.sh
```

Select option 5 from the menu, enter the path to your CSV file, and then enter "2" as the column number.

Notes:

The script assumes that your CSV file uses a comma (`,`) as the delimiter.

For options 4, 6, and 9, the script attempts to identify and process numeric values only.

The script creates new files for filtered and sorted data, and for saving analysis results, if they don't already exist.

This script provides a basic framework for CSV data analysis. You can customize and expand it to include more advanced features as needed.