

CSV File Splitter

A scalable Node.js tool to split large CSV files into smaller chunks, with advanced email and telephone processing. Supports both API and CLI usage.

Features

- Splits huge CSVs into multiple files with a configurable number of records per file
- Handles semicolon-separated emails and telephones, creating separate rows/fields as needed
- Skips compliance notice lines and empty email records
- Outputs a detailed summary in `summary.txt`

CLI Usage

1. Install dependencies:

```
npm install
```

2. Run the splitter on a CSV file:

- Add file at this path `src/uploads/`

```
node cli-split.js <filename>
```

- The output files and `summary.txt` will be created in a timestamped folder under `uploads/splits/` by default.
- You can override chunk size by setting environment variables:
 - `CSV_SPLIT_SIZE` (records per file)

API Usage

1. Start the server:

```
npm run dev
```

2. Call the split endpoint:

- POST to `/api/split-company-contacts`
- The input file is taken from `src/uploads/input.csv` by default, or set `CSV_INPUT_PATH` in your `.env`.
- Example with `curl`:

```
curl -X POST http://localhost:4000/api/split-company-contacts
```

- The response will include summary details and the output directory path.

Example Summary Output

Below is an example of the summary table generated after splitting a CSV file as `summary.txt` :

```
=====
CSV Split Summary
=====
Input:          D:\CSV_file_splitter\src\uploads\input-test.csv
Output Directory: D:\CSV_file_splitter\src\uploads\splits\input-test_20250829095306
Files Created:    11
Total Records:   1086548
Records Without Email: 1058
Records Processed: 1085490
Chunk Size:      100000
=====
```

Environment Variables

-
- `CSV_INPUT_PATH` - Path to the input CSV file (default: `src/uploads/main_file.csv`)
 - `CSV_SPLIT_SIZE` - Number of records per output file (default: 100000)
-

Output

-
- Split CSV files with headers in a `uploads/split/`
 - `summary.txt` in the output directory with a formatted summary table
-

Example

CLI:

```
node csv-split.js input.csv
```

API:

```
curl -X POST http://localhost:4000/api/split-company-contacts
```