The testing process was designed to ensure that all features of the claim date extraction system function correctly and reliably. The plan involved testing the program against both normal and edge-case scenarios. The following key aspects were tested:

1. **Data Loading** – Verified that the CSV file (`FACTors.csv`) loads correctly and that all rows are accessible for processing.

2. **URL Processing** – Ensured that each URL is fetched and parsed without breaking the loop if an error occurs.

3. **Claim Date Extraction** – Checked that the system correctly identifies and extracts claim dates from HTML pages, regardless of minor variations in formatting.

4. **Error Handling** – Confirmed that the program continues to run even if certain URLs fail due to timeout, invalid pages, or missing data.

5. **Output File Generation** – Validated that the processed dataset is saved back into CSV format with the new "Claim Date" column populated where available.

Testing Results

Testing was conducted using the full dataset, which contained over 100,000 URLs. The system successfully processed the majority of URLs without interruption.

During the first run, the program extracted claim dates for most records as expected. However, a small portion of URLs failed due to network errors or missing claim date elements in the HTML. These failures were gracefully handled, meaning the program skipped over problematic URLs without stopping execution.

In repeated runs, some previously failed URLs returned claim dates, showing that the issue was mostly due to temporary network timeouts rather than a flaw in the extraction logic.

The final output file contained a clean, structured dataset with claim dates filled in for all successful extractions. The missing values remained blank for URLs where the data simply wasn't available. Overall, the tool performed reliably under large-scale testing and handled problematic inputs effectively.