**Data Cleaning Log**

**Step 1: Import CSV Data to Spreadsheet**

* Imported the CSV data into a spreadsheet.

**Step 2: Clean the Gap of One Row Between Observations**

1. Searched Google for solutions; found none for spreadsheet, most were for Excel.
2. Consulted ChatGPT for advice on removing blank rows in spreadsheets.
   * Used suggested Script App code; found it time-consuming.
3. Opted to use Excel for efficient cleaning.
   * Process: Data > From Text/CSV > browsed CSV > transform data > remove rows > removed all blank rows > close and load.
   * Saved the file as CSV and imported it into the spreadsheet.

**Step 3: Delete Extra Rows and Columns**

* Removed the last two rows and two extra columns.

**Step 4: Bold Column Headers and Check Header Consistency**

* Bolded column headers and verified consistent formatting.

**Step 5: Eyeball Data for Cleaning**

* Checked for null values in columns A, D, G, H, I, J using countif function.
  + Columns A, D, G, H, I, J had null values.
* Applied conditional formatting (A2:J10993) to highlight and identify null cells.
* Removed three rows with blank cities and non-informative description.
* Replaced three null values in the location column with the respective city names.
  + Corrected a spelling mistake in the data.

**Step 6: Fill Missing Values**

* Identified 1258 missing latitudes and longitudes, and 26 missing city\_latitude and city\_longitude values.
* Used state-wise mean to fill missing values for latitudes, longitudes, city\_latitude, and city\_longitude.

**Step 7: Check for Spelling Mistakes**

* Used the spell check tool and manually accepted suggestions.
* Deleted an incorrect record identified during spell check.

**Step 8: Identify and Delete Incorrect Records**

* Noticed two additional incorrect records and deleted them.

**Step 9: Visual Scan for Short Descriptions**

* Used conditional formatting and filtering to find rows with descriptions less than 68 characters.
* Deleted rows with incorrect or insufficient description.

**Step 10: Clean Up Suggestions on Columns C and D**

* Applied clean-up suggestions and trimmed whitespaces.

**Step 11: Reorganize Latitude and Longitude Columns**

* Reorganized columns to have latitude before longitude.

**Step 12: Check for Duplicates and Remove**

* Identified duplicate rows at positions 1445 and 1453, as well as 6819 and 6822.
* Highlighted and removed duplicates using the remove duplicate tool.