# User Manual: RP Unit Configuration Cleaner Tool

**Link to Tool:** https://rpunitconfigcleanertool.streamlit.app/

## 1. Overview

The Unit Configuration Cleaner Tool is a web-based utility designed to automate the cleaning and formatting of property unit data. It processes Excel (.xlsx) or CSV (.csv) files to ensure unit names are standardized, free of invalid special characters, and devoid of accidental duplicates.

## 2. File Requirements

Before uploading, ensure your source file meets these criteria:

- **Format:** Must be .xlsx or .csv.

- **Column Headers:** The tool identifies data based on column names (case-insensitive). Your file **must** contain headers containing these keywords:

  - Unit (Required): The specific unit number/name.

  - Tower (Optional but recommended): The building or tower name.

  - Corporate (Optional): Corporate housing identifiers.

## 3. How to Use the Tool

### Step 1: Upload Files

1. Open the tool in your browser.

2. Click the **"Browse files"** button (or drag and drop).

3. Select one or multiple files.

4. The tool will begin processing files one by one.

### Step 2: Review Special Characters (Conditional)

The tool scans the **Unit** column for non-standard characters (e.g., symbols like @, #, $, or dates). If found, a warning will appear displaying the first 50 affected rows.

**You must select an action:**

- **Keep These Rows:** Retains the data exactly as is.

- **Delete These Rows:** Removes the rows containing special characters from the final file.

- **Cancel Processing:** Stops the process for this file so you can fix it manually offline.

Click **"Confirm Special Chars"** to proceed.

### Step 3: Review Duplicates (Conditional)

The tool checks for rows where the **Unit** and **Tower** combination is identical. If duplicates are found, a warning is displayed.

**You must select an action:**

- **Keep All (2+ Rows):** Keeps all duplicate entries (useful if duplicates are intentional).

- **Retain 1 Row:** Keeps the first instance and deletes the duplicates.

- **Cancel Processing:** Stops the process.

Click **"Confirm Duplicates"** to proceed.

### Step 4: Download Processed Files

Once validdtion is complete, the tool formats the data. It combines the **Tower** and **Unit** (and Corporate, if applicable) into a standardized format: Tower - Unit.

1. Look for the green **"✅ Processed"** message.

2. Click the **"⬇️ Download Cleaned File"** button.

3. The file will save to your computer with _cleaned appended to the filename.

### 4. Logic & Formatting Rules

- **Empty Values:** Terms like "N/A", "NA", "Blank", or empty cells are automatically removed to prevent dirty data (e.g., "Tower A - NA" becomes just "Tower A").

- **Unit Construction:** The final Unit column is created by joining the Tower and the Unit Number with a hyphen (e.g., Building A - 101).

### 5. Troubleshooting

| Issue | Cause | Solution |
|---|---|---|
| "⚠️ No 'Unit' column found" | The tool cannot find a column header with the word "Unit" in it. | Rename your column header to "Unit" or "Unit Number" and re-upload. |

| Issue | Cause | Solution |
|---|---|---|
| **"UnicodeDecodeError"** | The CSV file uses a unique encoding. | Try saving the file as a standard .xlsx (Excel) file and re-uploading. |
| **Download button not appearing** | The tool is waiting for your input. | Scroll up to see if there is a pending "Special Characters" or "Duplicates" decision box waiting for confirmation. |