## Description of Data Cleaning Performed

The provided OpenRefine history JSON describes a series of data cleaning operations. In the uuid column, variations in case and format were standardized to "510d47db-491f-a3d9-e040-e00a18064a99\r". The full\_height column had its blank cells filled with "0", which were then transformed to "NULL". Similarly, blank cells in the full\_width column were replaced with "NULL". Additionally, blank cells in the page\_number column were filled with "N/A". These operations aimed to ensure data consistency by standardizing values and filling in blanks with appropriate placeholders.

## UUID Standardization

There were a few values that were not properly cased to the rest of the column. Standardizing this column helped to ensure data consistency. The rationale behind this was to standardize the column so when we look for a certain identifier, we are certain that every UUID can be queried with the same format.

## Full\_Height and Full\_Width Entry

There were some empty values in this column that I thought would make more sense to be filled in with NULL. The data was simply not entered into the table and as an end user using the data, it would make more sense to have these values be filled with NULL rather than nothing. This step did not support the use case, but we think this is a necessary step to answer and doubt that could have arisen initially.

## Page\_Number to N/A

There were also some values that were empty within this column. I looked at a few Menu’s where the page was linked by a foreign key and it turned out that the Menu was only 1 page long. Due to this, I filled in empty values to N/A to showcase that the page itself is the menu and not a subset of pages.

|  |  |  |  |
| --- | --- | --- | --- |
| Table | Columns | Cells | Change |
| MenuPage | uuid | 3 | Merged the same UUID's together |
| MenuPage | full\_height | 329 | Empty Cells -> NULL |
| MenuPage | full\_width | 329 | Empty Cells -> NULL |
| MenuPage | PageNumber | 1202 | Just explaining, PageNumber's that are left empty mean that they are the entire Menu as there are no pages. Moved blank cells to N/A |

## IC violations

There were no IC violations that I saw within MenuPages as every single row had its own ID. There were multiple cases of the same UUID, however this makes sense with multiple menus being updated with the same image.