Mikołaj Kamiński

Email: 280596@student.pwr.edu.pl

Report for IDE Tool window data

1. Assumptions and matching strategy

The dataset given contains unmatched open–close event pairs, so cleaning and preprocessing are required before analysis. The first step is to divide the dataset by user_id's, that way records from different users won't interfere with each other. Then for each user's event (sorted by timestamp) we execute the following strategy:

I. Initialization :

- First of all we initialize an empty list to store currently "open" events

II. Processing events:

- If we encounter a record where event = 'open', we append it to the list of open events.
- If we encounter a record where event='close', we search the list for a suitable open event, starting with the first open event we currently have in our list.
- The duration is calculated as such (timestamp_close- timestamp_open)/1000 (seconds are easier to analyse than milliseconds)
- If the calculated duration is less than 24 hours, we consider it a valid record
- If the calculated duration is over 24 hours, we discard that open record, as it won't match with any new records

III. Assumptions:

- Given the lack of information about given dataset, I assumed the 24 maximum valid time between open and close events

time_window	discarded_opens	discarded_closes	discarded_opens_percentage	discarded_closes_percentage	matched_pairs	matched_pairs_percentage
1	404	177	21.66	10.81	1461	89.19
12	287	60	15.39	3.66	1578	96.34
24	256	29	13.73	1.77	1609	98.23
48	244	17	13.08	1.04	1621	98.96
72	239	12	12.82	0.73	1626	99.27
96	237	10	12.71	0.61	1628	99.39

(generated/csv/time_window_analysis.csv)

- As the table shows, 98,23% of all events get matched. Choosing a bigger time_window, would only slightly increase matched_pairs rate, while increasing the number of outliers which don't provide much value in analysis. Sessions exceeding 24 hours most likely represent edge cases rather than typical usage patterns.

IV. Output:

- For each valid pair we create a new record with:
  - open_type: (from the original "open" event)
  - duration: (time between events)
  - user_id: (user associated with events)
- Finally, all matched pairs across users are combined into a clean dataset containing only relevant, paired events.

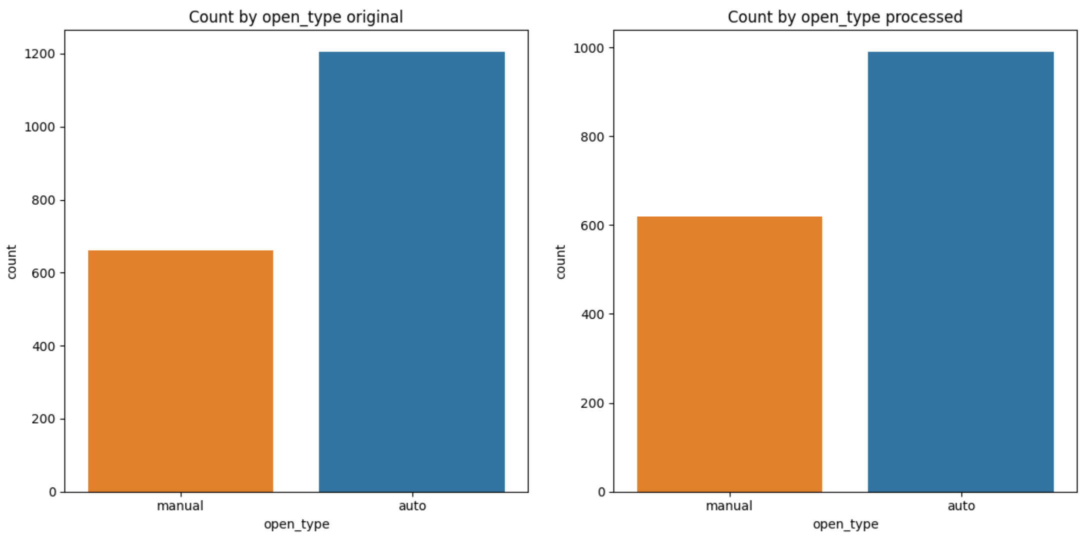
2. Analysis

2.1 Comparison with original dataset

	stage_name	total_records	manual_count	auto_count
1	initial_dataframe	3503	661	1204
2	delete_duplicates	3503	661	1204
3	final_dataframe	1609	620	989

(generated/csv/datasets_stats.csv)

The initial dataset didn't contain any duplicates. The ratio between two different open_types is roughly 2:1, but after data processing the ratio changes to roughly approximately 3:5. Of the original 661 windows opened manually, 620 successfully found a close event. With automatically opened windows 315 weren't closed or didn't find a matching closed event. Basing on that we formulate a hypothesis that automatically opened windows are more frequently left open or have longer durations that exceed tracking periods.



(generated/plots/count_by_open_type.png generated/plots/og_count_by_open_type.png)

2.2 Measure of confidence

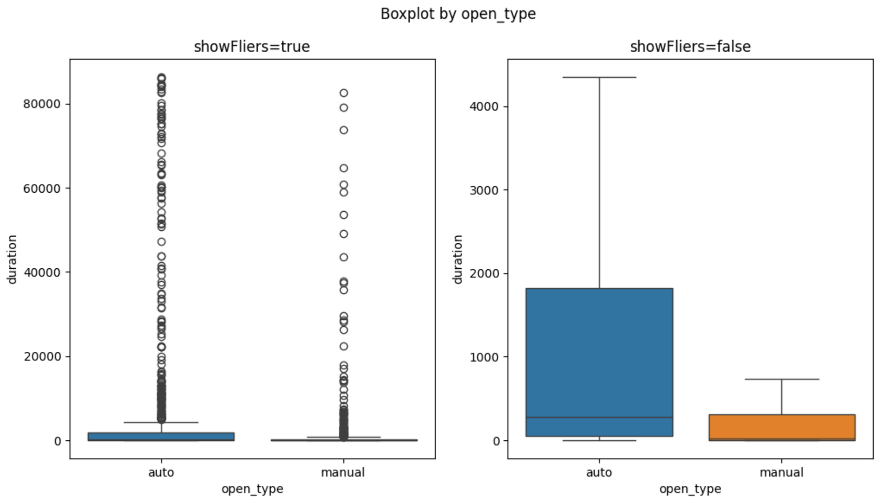
```
Summary Statistics by open_type

          count         mean           std    min     25%       50%         75%         90%          99%        max
open_type
auto      989.0  5933.515030  16597.760477  0.461  49.891  277.2780  1817.52900  12413.9806  82859.88976  86320.440
manual    620.0  1958.455889   8566.937048  0.015   2.684   17.3325   305.90275   2341.9880  52769.72187  82525.116

=== Mann-Whitney U Test ===
U statistic: 170558.00
P-value: 7.6066e-51
Rank-biserial correlation (effect size): 0.444
```

(duration_analysis.txt)

The statistics reveal a clear difference in the duration of tool window usage between automatically and manually opened instances. The average duration for automatically opened windows (mean ≈ 5933 s, median ≈ 277 s) is substantially higher than that for manually opened windows (mean ≈ 1958 s, median ≈ 17 s). Moreover, the upper quantiles for automatic openings (e.g., 90th percentile ≈ 12414 s, 99th percentile ≈ 82859 s) indicate that these windows remain active for significantly longer periods, suggesting prolonged or background usage. In contrast, manually opened windows are typically closed much sooner, reflecting shorter, more task-specific interactions.

The Mann–Whitney U test (U = 170,558, p=7,606e-51) confirms that this difference is statistically significant. The associated effect size (r = 0.444) indicates a moderate to strong practical effect, suggesting that the method of opening the window meaningfully influences how long it remains active.

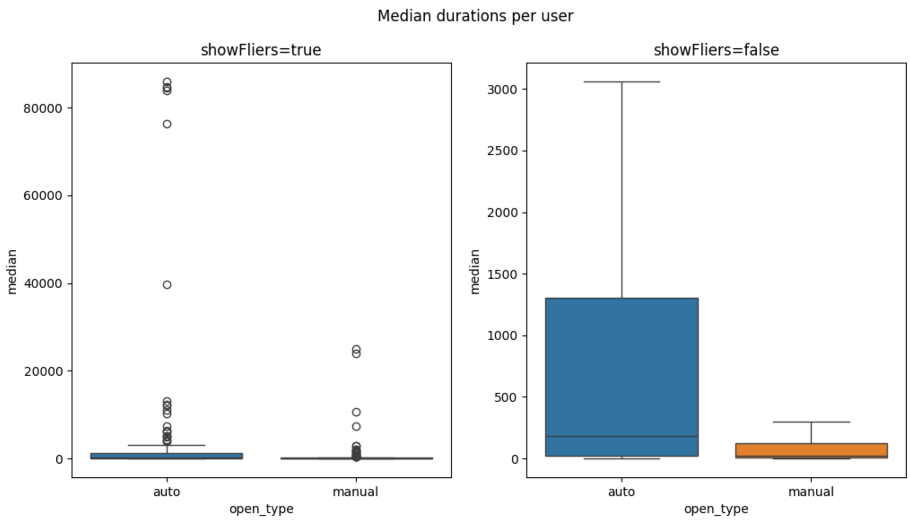
2.3 Data distribution



(generated/plots/boxplot.png)

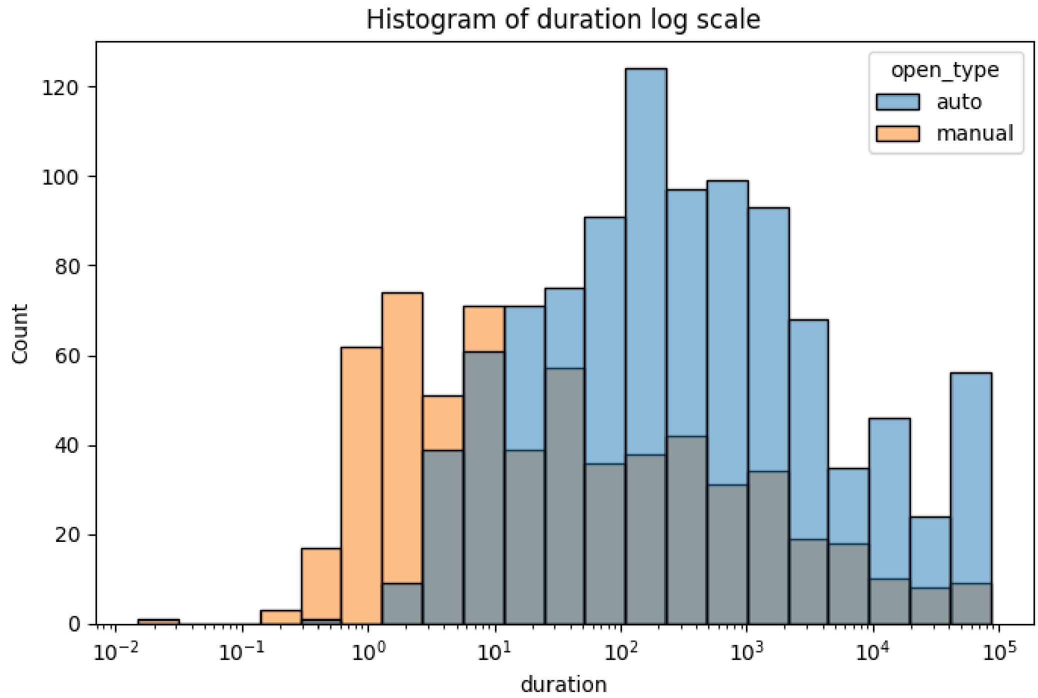
The left panel reveals that while both window types exhibit outliers, automatically opened windows display substantially more outliers with higher extreme values. When outliers are excluded, the patterns become much clearer.

The automatically opened windows demonstrate both a higher median duration and a broader interquartile range compared to the manually opened windows. This further indicates that automatically triggered openings tend to persist longer and display greater variability in their active durations. Manually open windows exhibit shorter and more uniform values .



(generated/plots/boxplot_user_median.png)

Data presented above (boxplot.png) could be skewed by a singular user with many open-close events with extremely high values. The plot (boxplot_user_median.png) shows similar properties to the previous plot: many outliers for both types, higher extreme values for automatically open windows, and lower median and IQR for manually open windows. This proves that duration differences are genuine behavioural patterns, not artifacts of individual user behaviour.



(generated/plots/histplot_log_scale.png)

Manually opened windows tend to have shorter durations, with the highest density of observations concentrated in the 0–10 second interval. This could be due to users searching for a specific tool window and accidentally clicking the wrong option, or simply because manually opened windows are typically used for a shorter period of time. In contrast, automatically opened windows exhibit a wider and more right-skewed distribution, extending to substantially higher duration values. This pattern suggests that automatically triggered openings are more likely to remain active throughout entire work sessions, whereas manually opened tool windows are usually closed shortly after being opened. The histogram supports the hypothesis that the method of opening has a significant influence on the typical duration of tool window activity, with automatic openings being both longer and more variable than manual ones.

3. Conclusion

The analysis demonstrates significant behavioural differences between automatically and manually opened tool windows. Both descriptive statistics and visualizations consistently show that automatically opened windows tend to remain active for significant periods of time and exhibit a greater variability. Conversely manually opened windows have shorter and more uniform durations. However, to draw more detailed conclusions, additional context about the dataset would be required. For instance, a less experienced user might frequently open and immediately close tool windows when searching for the correct feature, while a more experienced developer would perform fewer mistakes. Despite all that, the statistical results supported by the Mann–Whitney U test and a moderate effect size provide evidence that the opening mechanism meaningfully influences tool window activity.