Analysis of Tool Window Usage Data

Bruno Tomusiak

November 2025

1 Introduction

The purpose of this report is to present the approach and findings for the following objective:

> *Determine whether there is a significant difference in how long a tool window stays open depending on whether it was opened manually or automatically.*

To answer this, I:

- reconstructed complete "episodes" by matching open/close events,
- computed per-episode open durations,
- compared duration distributions between manual and automatic opens,
- assessed statistical significance and practical effect size.

2 Data

The raw event log contains four fields per row: `user_id`, `timestamp` (epoch milliseconds), `event` (`opened`/`closed`), and `open_type` (`manual`/`auto` for open events; missing for close events). After timestamp conversion and sorting, the dataset comprises **3503** events across users (as shown by the notebook's `DataFrame.info()` output).

3 Data cleaning

Timestamps were converted from epoch milliseconds to `datetime`. Events were globally sorted by `user_id` and `timestamp`. Missing values occur only in `open_type` on close events, which is expected given the schema. No other systematic missingness was detected.

4 Reconstructing usage episodes

Real-world logs contain irregularities (close without prior open, multiple opens in a row, open without a closing at the file end). I therefore used a simple and robust *state-machine* procedure:

> *First, I grouped events by user and sorted each user's timeline chronologically. I tracked two state variables: the timestamp of the currently open session and its open type (manual/auto). Upon encountering an* `opened` *event while no session was active, I recorded its timestamp and open type. Upon encountering the next* `closed` *event, I computed the duration as the difference between close and open times and appended it as a completed episode. Non-positive durations were discarded to guard against mis-ordered or corrupted records. Consecutive* `opened` *events without an intervening close collapse into a single ongoing session; stray* `closed` *events are ignored.*

This procedure yielded a table of episodes with columns: `user_id`, `opened_at`, `closed_at`, `open_type`, and `duration` (seconds/minutes). In total, this produced **1622** open–close episodes (as seen in the notebook summaries).

5 Descriptive statistics

Durations exhibit the typical long-right-tail shape (many short sessions with a few very long ones). For all episodes combined, the notebook shows:

- mean duration $\approx 11\,595.000\,\text{s}$ (=193.252 min),

which is inflated by extreme values, consistent with heavy skew. For this reason, I removed the outliers. As a result, I kept 99% of the records.

By open type, the groupwise summary (from the notebook's grouped `describe`) indicates:

- 75th percentile (minutes): **auto** ≈ 23.049 vs. **manual** ≈ 2.676.
- 50th percentile (minutes): **auto** ≈ 3.518 vs. **manual** ≈ 0.230.
- mean (minutes): **auto** ≈ 147.720 vs. **manual** ≈ 50.770.

These figures already suggest systematically longer auto episodes and a very long tail in both groups.

6 Statistical significance

To formally test the difference between the two duration distributions, I used the Mann–Whitney U (Wilcoxon rank-sum) test, which compares distributions without assuming normality (appropriate for skewed, heavy-tailed durations). The notebook reports:

$$p\text{-value} \;=\; 2.01 \times 10^{-58},$$

which provides *overwhelming* evidence that manual and auto duration distributions differ.

Interpretation. A p-value of 2×10^{-58} indicates that, under the null hypothesis of no difference, observing a separation at least this extreme is astronomically unlikely. Thus, we reject the null and conclude that opening mode is strongly associated with how long the tool window stays open.

7 Effect size (practical difference)

Beyond significance, I quantified the practical gap using the median difference. The notebook computes:

$$\text{median(auto)} - \text{median(manual)} \;=\; 3.287\,\text{min}.$$

Hence, a *typical* auto-opened session lasts about 3.290 min longer than a typical manually opened session in this dataset. Combined with the much larger upper quantiles for auto, this points to both a central shift and a heavier right tail for automatically opened windows.

8 Visualization

Density plots on a log-scaled x-axis clearly show the auto distribution shifted to the right relative to manual, with well-separated medians and heavier auto tails. Note that a KDE integrates to 1 in *linear* space; when displayed on a log-scaled axis the visual area is not conserved, but relative placement and separation remain informative.

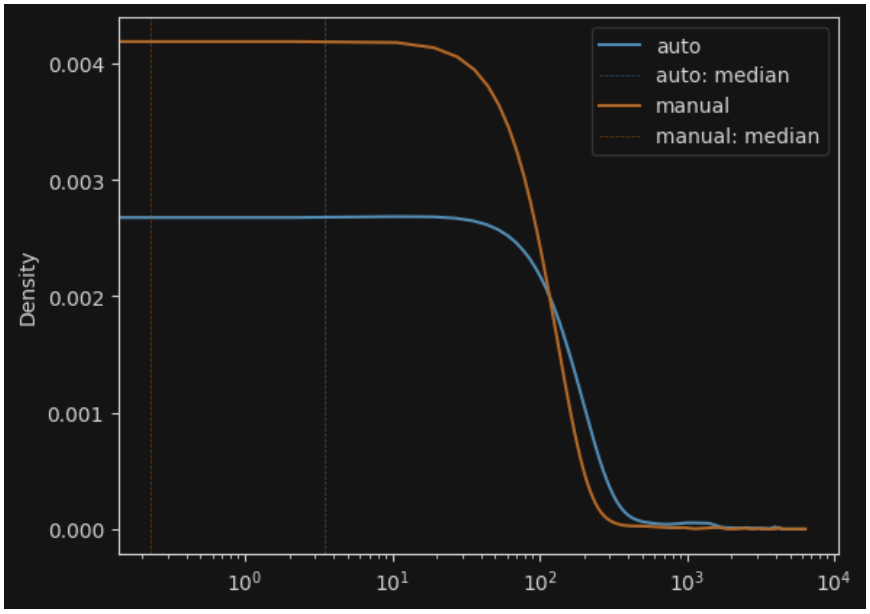

Figure 1: Kernel density estimate of tool window open durations (minutes, log scale). The orange curve represents manual openings and the blue curve automatic openings. Dashed lines mark each group's median duration.

9 Conclusions

There is a **clear and statistically decisive** difference in how long the tool window remains open depending on opening mode. In this dataset:

- distributions differ with $p \approx 2.0 \times 10^{-58}$ (Mann–Whitney U),
- the typical (median) auto-opened session is about 3.290 min longer than the typical manual session,
- upper quantiles for auto are substantially larger, indicating more frequent long-lived sessions.

Practically, automatically opened windows tend to persist longer, and the effect is both statistically and behaviorally meaningful.