Tool Windows Usage Data Analyzing

Nikita Magomedeminov

Assumptions

- Durations longer than 12 hours are considered outliers.
- Negative durations are impossible.

Analysis pipeline

The analysis pipeline consisted of:

1. Sorting events chronologically.
2. Matching pairs of open and close events.
3. Separating the dataset into “auto” and “manual” groups for comparison.
4. Delete outliers.
5. Perform statistical tests.

All points are described in details further.

Handling messy data

- Missing **next events** (end of session) were dropped.
- Duplicate consecutive events (same type repeated) were removed.

Strategy for matching open/close events

1. Group events, sorted by `timeline` per `user_id` and shifting the `event` and `timestamp` columns by -1. Save them as `next_event` and `next_timeline`.
2. Calculate the difference of and save it as `duration`

Removing outliers

1. We remove very long events at the very beginning (more than 12 hours).
2. We use rule of thumb to remove other outliers.

Described filtered arrays

Metric	Auto	Manual
Count	957	617
Mean in ms	337083	36210
Std in ms	491610	64360

Statistical tests

Here we test 2 filtered arrays with manual and auto open types.

1. Kolmogorov-Smirnov test, to check whether 2 samples are from one distributions or not. **Result**: p-value $\ll 0.05$.
2. Cliff’s delta, 95% confidence interval, calculated using bootstrap. **Result**: $[0.5978, 0.6874]$, what means that tool window open automatically lasts longer.
3. 95% confidence interval for means difference. **Result**: $[266849.267, 334896.404]$, what means that tool window open automatically lasts longer again.
4. Group data inside users by `open_type`, to see means difference again. Here the condition of minimum 10 events of each `open_type` inside the user is required. Only 14 users left, and there is

only one with `manual_duration` > `auto_duration`. If the process is considered to be random, then p-value is ~ 0.002

Conclusion

By performing different tests we see, that on average automatically opened toolwindow lasts longer. This is not something of random – high confidence levels were used in statistical tests, and even in the last test, the p-value is very small.

Visualizations

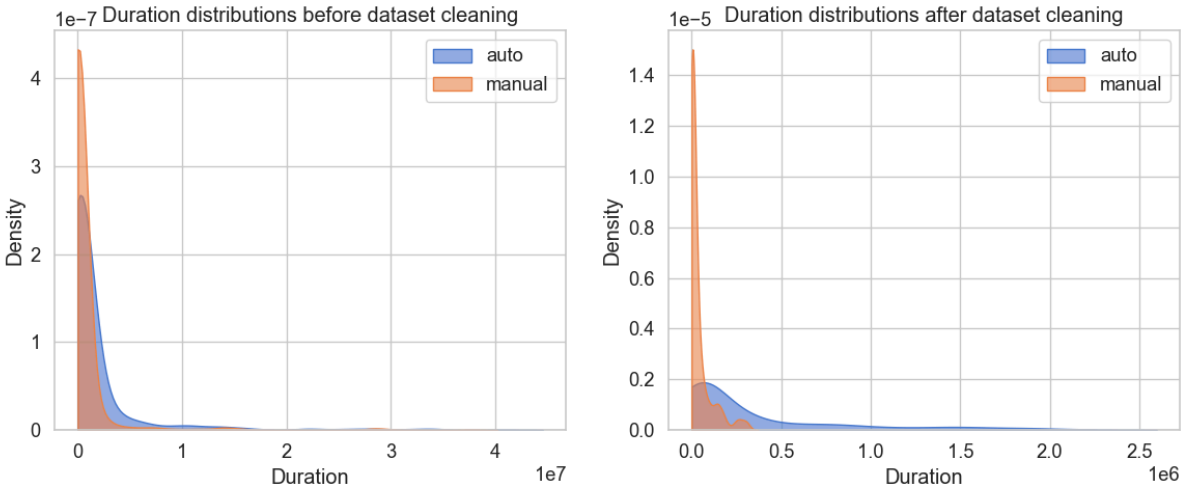
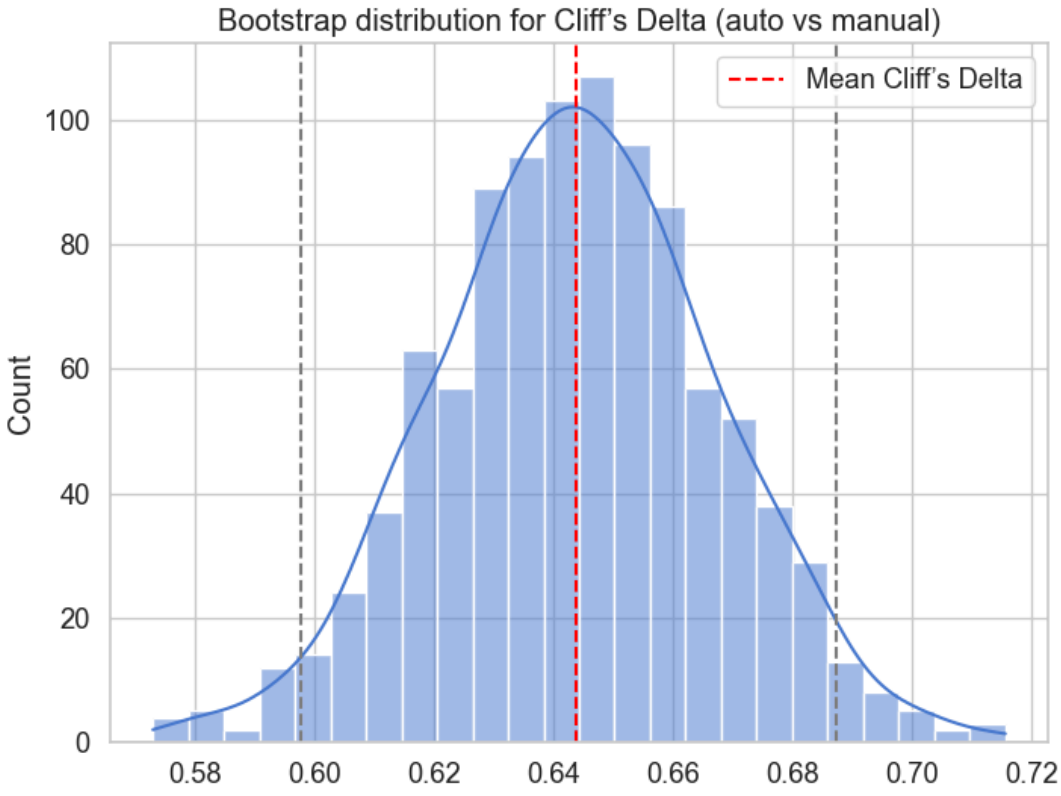
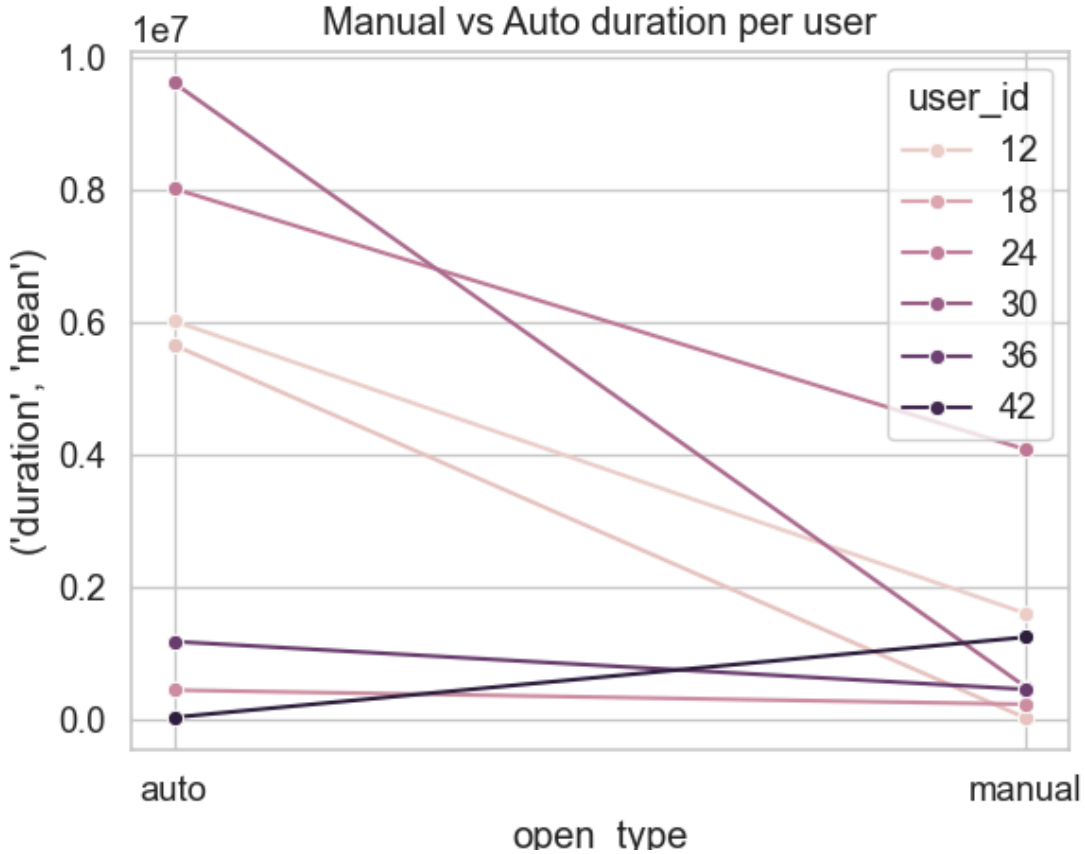


Figure 1: Initial data distributions




Manual vs Auto duration per user
1e7
1.0
0.8
0.6
0.4
0.2
0.0
('duration', 'mean')
user_id
12
18
24
30
36
42
auto
manual
open_type