Visualizing Event Type Distributions Over Time: Interactive Violin Charts

Process Mining and Visual Analytics

June 6, 2025

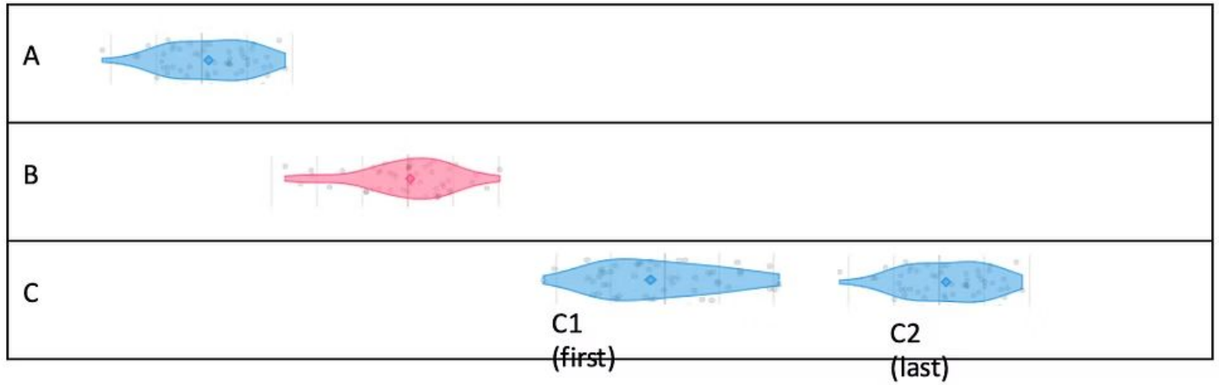
Umut Yunus Yesildal, 638201

Motivation & Roadmap

- Processing
- Sorting
- Violin Chart
- Interaction

Task 10 Event type distribution over time axis

- Processing: Calculate for each event type when they occur on a time axis relative to the start event of the sequence.
  Sort event types based on statistical parameter
  Potential visual element: violin chart
  Potential interaction: offer the user the option to select different statistical parameters for sorting: min, max, mean, median, quartils
- See Chen, W., Guo, F., & Wang, F. Y. (2015). A survey of traffic data visualization. IEEE Transactions on intelligent transportation systems, 16(6), 2970-2984.



Framework and Tools

Python

Core programming environment.

PM4Py

XES event log handling.

Framework and Tools

Plotly & Dash: Interactive Web Visualizations

Plotly creates interactive charts in Python. Dash builds web dashboards from Plotly charts—ideal for data science visualization.

Event Log Selection

Initial Attempts

- BPI 2017: Highly skewed data.
- BPI 2012: Better, but not satisfactory.

Final Selection

- Road Traffic Fines Management (2015).
- Known for good temporal distributions.

Main Insight: Data distribution significantly impacts violin plot visualization.

Design and Approach

1

Data Import

XES to DataFrame with PM4Py.

2

Preprocessing

Calculate time_since_case_start, filter outliers, optional log transformation.

3

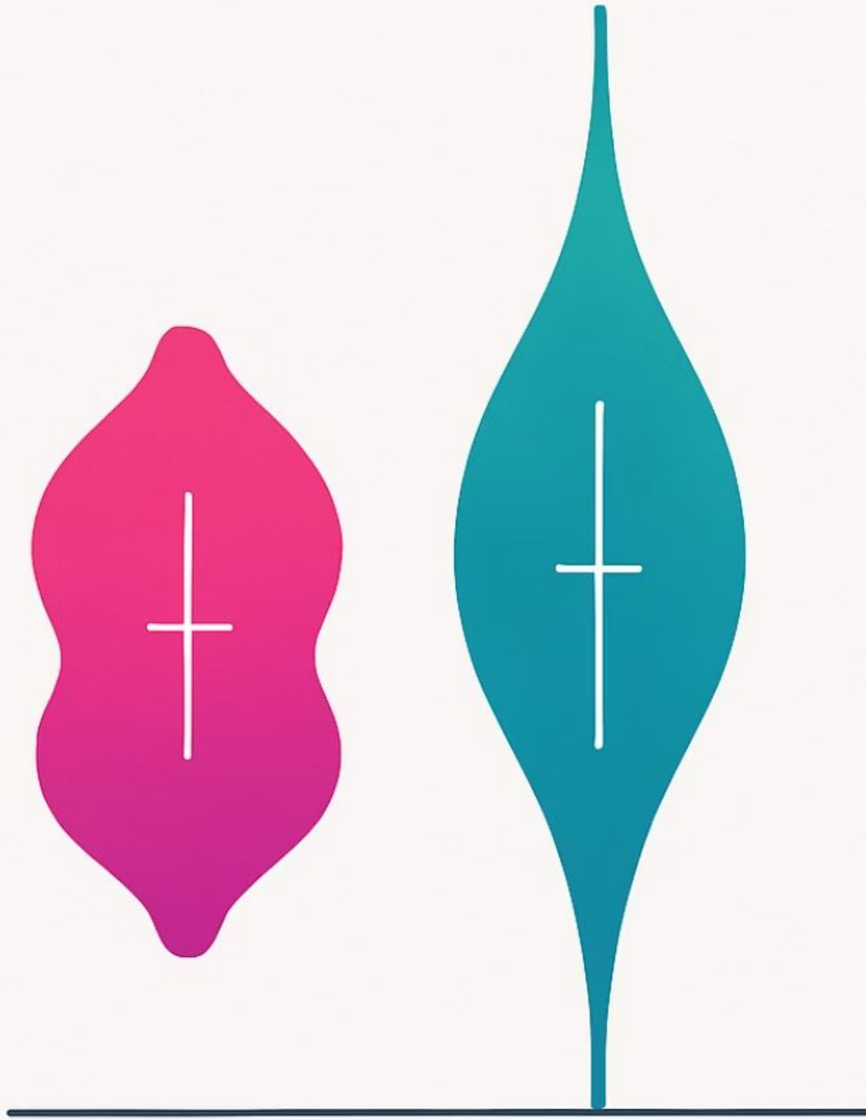
Statistical Processing

Compute min, max, mean, median, quartiles on multiple labels of data.

4

Visualization Design

Interactive violin plots, user interaction, clear labels.



Mockups

Interactive UI Sketches

Dropdown for sorting parameters.

Horizontal Violin Plots

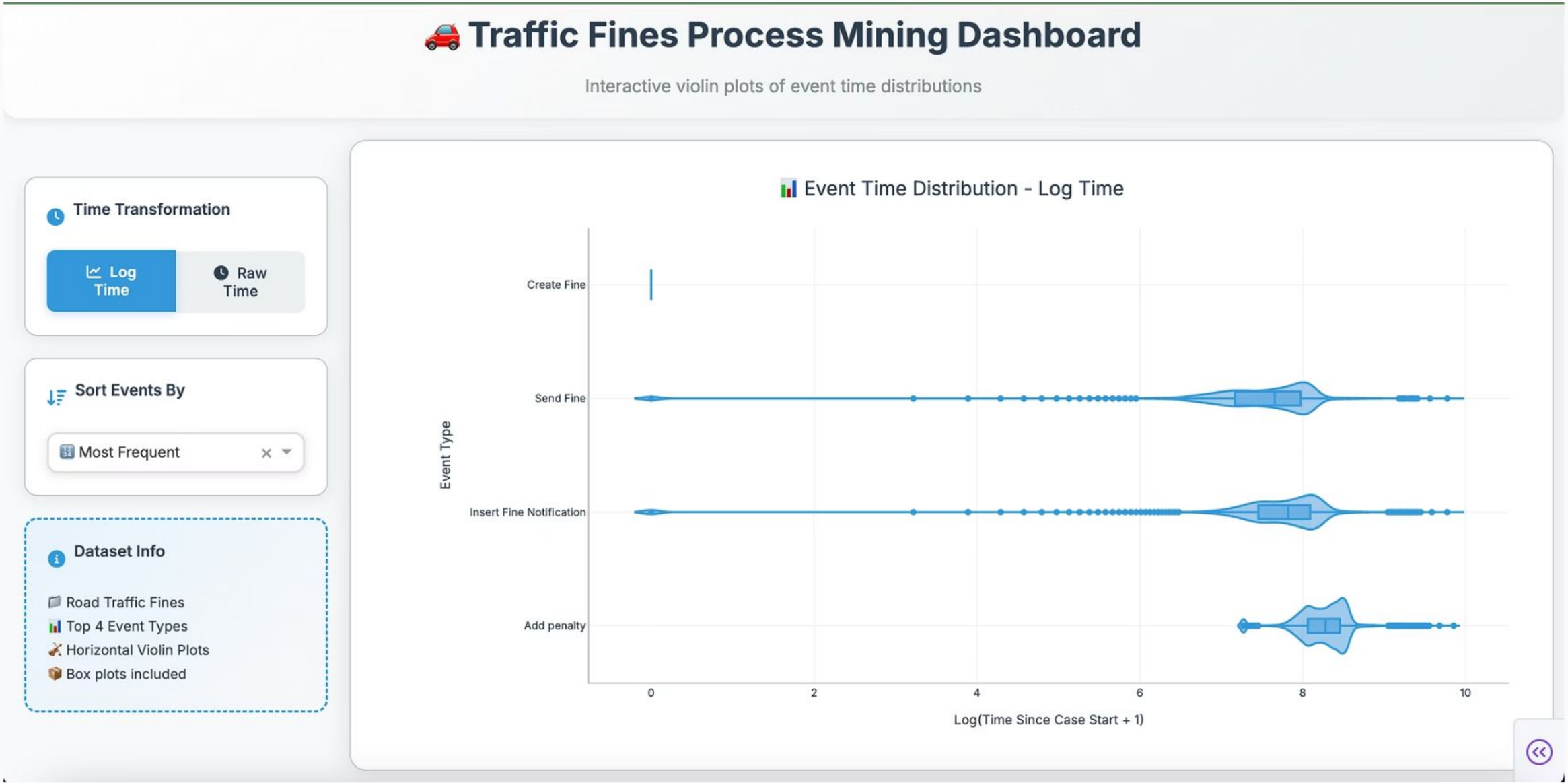
Clearly aligned per event type.

Clear Labels

Axes, title, and labels defined.

Initial Results

Initial Results



Next Steps

Processing Updates

Additional data preprocessing structure and processing approach.

Robust Stats for various datasets

Implementing processing for multiple datasets to visualize violin charts easily.

Interactive Improvements

Updating interactiveness such as tooltips, dropdown, buttons.

References

- Chen, W., Guo, F., & Wang, F.-Y. (2015). A Survey of Traffic Data Visualization. IEEE Transactions on Intelligent Transportation Systems, 16(6), 2970-2984.
- van der Aalst, W. (2016). Process Mining: Data Science in Action.
- Mendling et al. (2018). Fundamentals of Business Process Management.

- Course materials and resources.