

# MPVIS: Statistics for Real-Life Event Log Examples

MPVIS has been used to discover and visualize process models for several publicly available real-life event logs. The resulting process models can be found at <https://bit.ly/mpvis-examples>. Table 1 summarizes the considered event logs and provides the ‘time taken for discovery’ and ‘time taken for visualization’ performance statistics, in seconds, for both MP-DFG and MD-DRT. The tests were run in Google Colaboratory without a pro license. It must be noted that these statistics were obtained by applying the default settings of both modules, but that the pruning and grouping capabilities of the library have been used for the process models shown in the above link to highlight their utility in reducing the complexity of the visualized processes.

**Table 1:** Performance of MPVIS on real-life event logs

Event Log	Case Count	Variant Count	Event Count	Activity Count	MP-DFG		MD-DRT	
					Disc. Time (sec.)	Vis. Time (sec.)	Disc. Time (sec.)	Vis. Time (sec.)
BPI Challenge 2013 - Incidents [1]	7,554	2,278	65,533	13	18.30	0.25	177.14	484.85
BPI Challenge 2015 - Municipality 1 [2]	1,119	1,170	52,217	398	23.91	-*	52.24	4,283.47
Curricular Analytics [3]	11,595	3,935	126,658	8	48.56	0.30	404.18	572.52
Device Repair Process [4]	23,906	2,727	178,078	13	60.95	0.62	1,042.67	90.96
Hospital Billing [5]	100,000	1,020	451,359	18	130.79	0.91	14,853.11	21.20
Hospital Locations [6]	328	95	2,222	36	14.72	0.38	18.12	0.94
IT Incident Log [7]	24,918	2,079	141,712	9	37.83	0.25	960.08	135.86
Italian Help Desk [8]	4,580	226	21,348	14	19.91	0.36	46.45	1.57
Production Data [9]	225	221	4,543	55	2.20	20.24	3.44	18.18
Sepsis Cases [10]	1,050	846	15,214	16	17.48	0.79	19.84	58.49
Supply Chain Management [11]	99	12	972	10	0.23	0.31	1.08	0.36

\* Unable to visualize in a reasonable time due to the number of activities and connections.

## References

- [1] W. Steeman. *BPI Challenge 2013, incidents*, Ghent University, 2013.
- [2] B. F. van Dongen, "Bpi challenge 2015," in *11th International Workshop on Business Process Intelligence (BPI 2015)*, 2015.
- [3] J. P. Salazar-Fernandez, M. Sepúlveda, J. Munoz-Gama, and M. Nussbaum, "Curricular analytics to characterize educational trajectories in high-failure rate courses that lead to late dropout," *Applied Sciences*, vol. 11, no. 4, p. 1436, 2021.
- [4] M. Stierle and M. Matzner. *Customer Service-Device Repair Process (sp2020)*, Zenodo. DOI, 2020.
- [5] F. Mannhardt, M. De Leoni, H. A. Reijers, and W. M. Van Der Aalst, "Data-driven process discovery-revealing conditional infrequent behavior from event logs," in *Advanced Information Systems Engineering: 29th International Conference, CAISE 2017, Essen, Germany, June 12-16, 2017, Proceedings 29*, 2017: Springer, pp. 545-560.
- [6] S. Namaki Araghi. *Living Lab Hospital Interpreted Location event logs*, Mendeley Data, 2023.
- [7] S. I. Shifat. *IT incident log Dataset*, Kaggle Datasets, 2020.
- [8] M. Polato. *Dataset belonging to the help desk log of an Italian Company*, University of Padova, 2017.
- [9] D. Levy, "Production analysis with process mining technology," *Dataset*, 2014.
- [10] F. Mannhardt and D. Blinde, "Analyzing the trajectories of patients with sepsis using process mining," in *RADAR+ EMISA 2017*, 2017: CEUR-ws. org, pp. 72-80.
- [11] F. A. Tridalestari, B. Warsito, A. Wibowo, and H. Prasetyo, "Analysis of E-commerce process in the downstream section of supply chain management based on process and data mining," *Ingénierie des Systèmes d’Information*, 2022.