Volvo IT Process Management: BPI 2013 Revisited

Master in Data Science, PODS Group 12
Bruna Barraquer
Walter Troiani
Nayara Costa
Haonan Jin
Hang Yao

Context: IT Management

[1]

[1] Image source: This image was created with the assistance of Generative AI

The Datasets

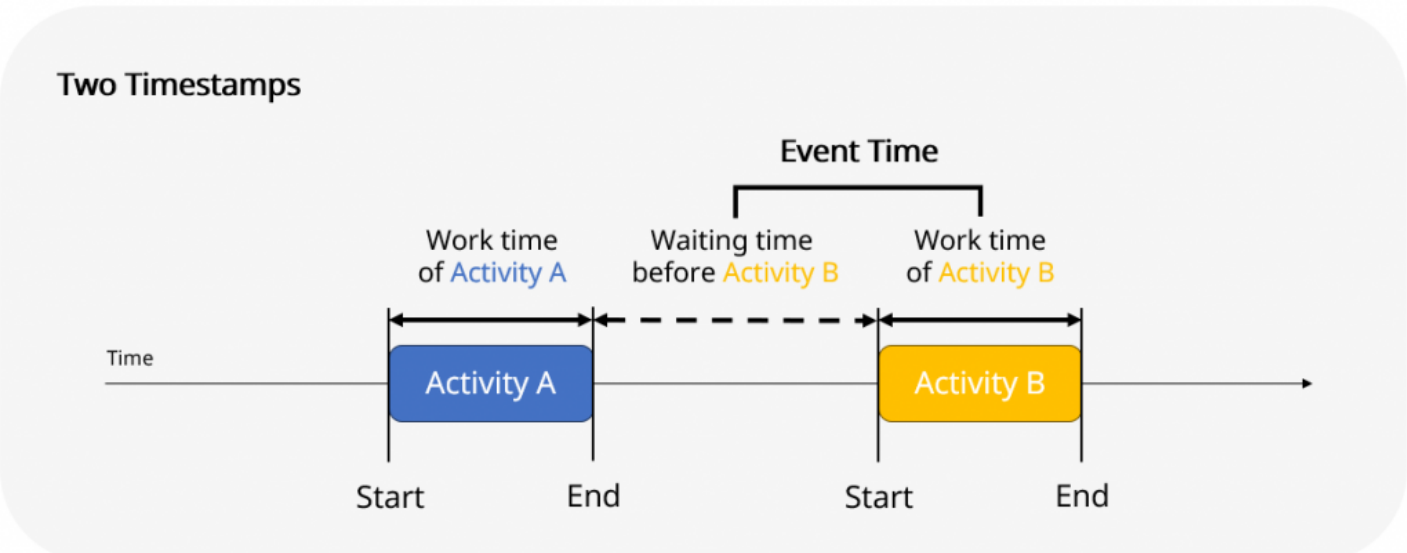
	A	B	C	D	E	F	G	H	I	J	K
1	Case_ID	Activity	Resource	concept_name	impact	org_group	org_role	organization_involved	product	resource_country	End_Time
2	1-364285768	Accepted/In Progres	Frederic	Accepted	Medium	V30	A2_4	Org line A2	PROD582	France	2010-03-31T16:59:42.000
3	1-364285768	Accepted/In Progres	Frederic	Accepted	Medium	V30	A2_4	Org line A2	PROD582	France	2010-03-31T17:00:56.000
4	1-364285768	Queued/Awaiting Ass	Frederic	Queued	Medium	V5 3rd	A2_5	Org line A2	PROD582	France	2010-03-31T17:45:48.000
5	1-364285768	Accepted/In Progres	Anne Claire	Accepted	Medium	V5 3rd	A2_5	Org line A2	PROD582	France	2010-04-06T16:44:07.000
6	1-364285768	Queued/Awaiting Ass	Anne Claire	Queued	Medium	V30	A2_4	Org line A2	PROD582	France	2010-04-06T16:44:38.000
7	1-364285768	Accepted/In Progres	Anne Claire	Accepted	Medium	V13 2nd 3rd	A2_5	Org line A2	PROD582	France	2010-04-06T16:44:47.000
8	1-364285768	Completed/Resolved	Anne Claire	Completed	Medium	V13 2nd 3rd	A2_5	Org line A2	PROD582	France	2010-04-06T16:44:51.000
9	1-364285768	Queued/Awaiting Ass	Anne Claire	Queued	Medium	V30	A2_4	Org line A2	PROD582	France	2010-04-06T16:45:07.000
10	1-364285768	Accepted/In Progres	Eric	Accepted	Medium	V30	A2_4	Org line A2	PROD582	France	2010-04-08T12:52:23.000
11	1-364285768	Queued/Awaiting Ass	Eric	Queued	Medium	V5 3rd	A2_5	Org line A2	PROD582	France	2010-04-08T12:53:35.000
12	1-364285768	Accepted/In Progres	Anne Claire	Accepted	Medium	V5 3rd	A2_5	Org line A2	PROD582	France	2010-04-20T11:07:11.000
13	1-364285768	Accepted/Assigned	Anne Claire	Accepted	Medium	V5 3rd	A2_5	Org line A2	PROD582	France	2010-04-20T11:07:19.000
14	1-364285768	Accepted/In Progres	Sarah	Accepted	Medium	V5 3rd	A2_5	Org line A2	PROD582	France	2012-04-11T17:11:17.000
15	1-364285768	Accepted/Assigned	Sarah	Accepted	Medium	V5 3rd	A2_5	Org line A2	PROD582	France	2012-04-11T17:11:25.000
16	1-364285768	Accepted/In Progres	Loic	Accepted	Medium	V5 3rd	A2_5	Org line A2	PROD582	France	2012-05-03T11:10:10.000
17	1-364285768	Completed/Resolved	Loic	Completed	Medium	V5 3rd	A2_5	Org line A2	PROD582	France	2012-05-03T11:10:12.000
18	1-364285768	Completed/Closed	Siebel	Completed	Medium	V5 3rd	A2_5	Org line A2	PROD582	0	2012-05-11T01:26:15.000
19	1-467153946	Accepted/In Progres	Adam	Accepted	Medium	S42	V3_2	Org line C	PROD453	Sweden	2011-01-31T11:12:22.000
20	1-467153946	Accepted/In Progres	Adam	Accepted	Medium	S42	V3_2	Org line C	PROD453	Sweden	2011-01-31T11:18:44.000
21	1-467153946	Queued/Awaiting Ass	Adam	Queued	Medium	N52 2nd	V3_2	Org line C	PROD453	Sweden	2011-01-31T11:19:05.000
22	1-467153946	Accepted/In Progres	Denny	Accepted	Medium	N52 2nd	V3_2	Org line C	PROD453	Sweden	2011-01-31T12:59:46.000
23	1-467153946	Accepted/Wait - User	Denny	Accepted	Medium	N52 2nd	V3_2	Org line C	PROD453	Sweden	2011-01-31T14:37:55.000
24	1-467153946	Queued/Awaiting Ass	Denny	Queued	Medium	O3 3rd	C_6	Org line C	PROD453	Sweden	2011-02-03T08:28:58.000
25	1-467153946	Accepted/In Progres	Paul	Accepted	Medium	O3 3rd	C_6	Org line C	PROD453	Sweden	2011-02-07T12:37:33.000
26	1-467153946	Accepted/Wait - Imp	Paul	Accepted	Medium	O3 3rd	C_6	Org line C	PROD453	Sweden	2011-02-07T12:38:25.000
27	1-467153946	Accepted/In Progres	Åse	Accepted	Medium	G140 2nd	E_10	Org line C	PROD453	Sweden	2011-03-09T11:08:06.000
28	1-467153946	Accepted/Wait - Imp	Åse	Accepted	Medium	G140 2nd	E_10	Org line C	PROD453	Sweden	2011-03-09T11:27:05.000
29	1-467153946	Accepted/In Progres	Joseph	Accepted	Medium	G140 2nd	E_10	Org line C	PROD453	Sweden	2011-03-10T11:53:10.000
30	1-467153946	Accepted/Assigned	Joseph	Accepted	Medium	G140 2nd	E_10	Org line C	PROD453	Sweden	2011-03-10T11:53:22.000

- 7554 cases
- 65533 events

Ensuring data completeness

Start time?

This data is incomplete, it only contains end times and lacks start times!



Ensuring data completeness

Start Time Approximation Algorithm

Apromore provides an algorithm to approximate the start time $\forall eventK$ as follows:

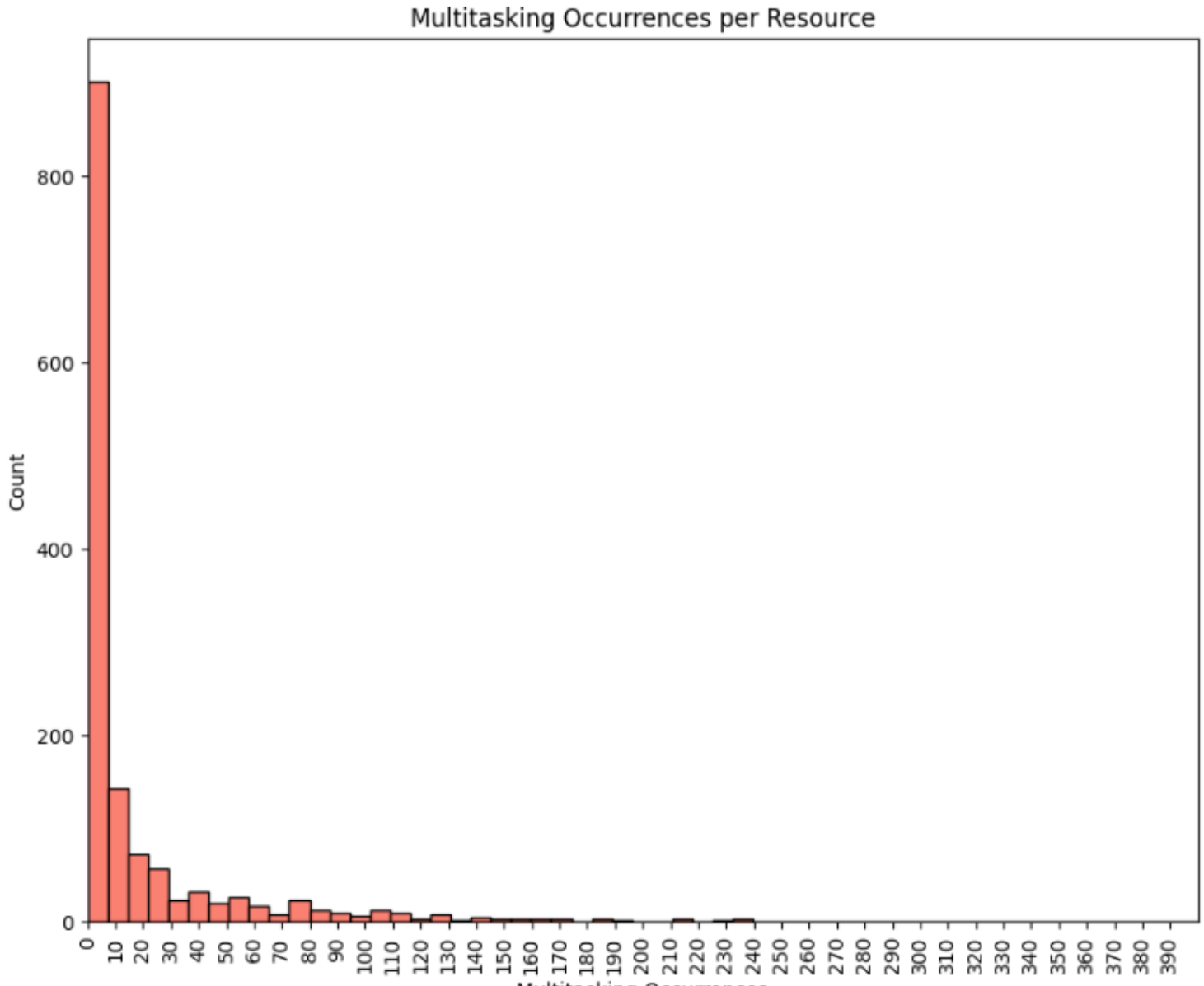
StartTime(K) $\approx \max(EndTime(K-1), ResourceAvailabilityTime(R, K))$

ResourceAvailabilityTime(R, K): time since resource R became available before event K

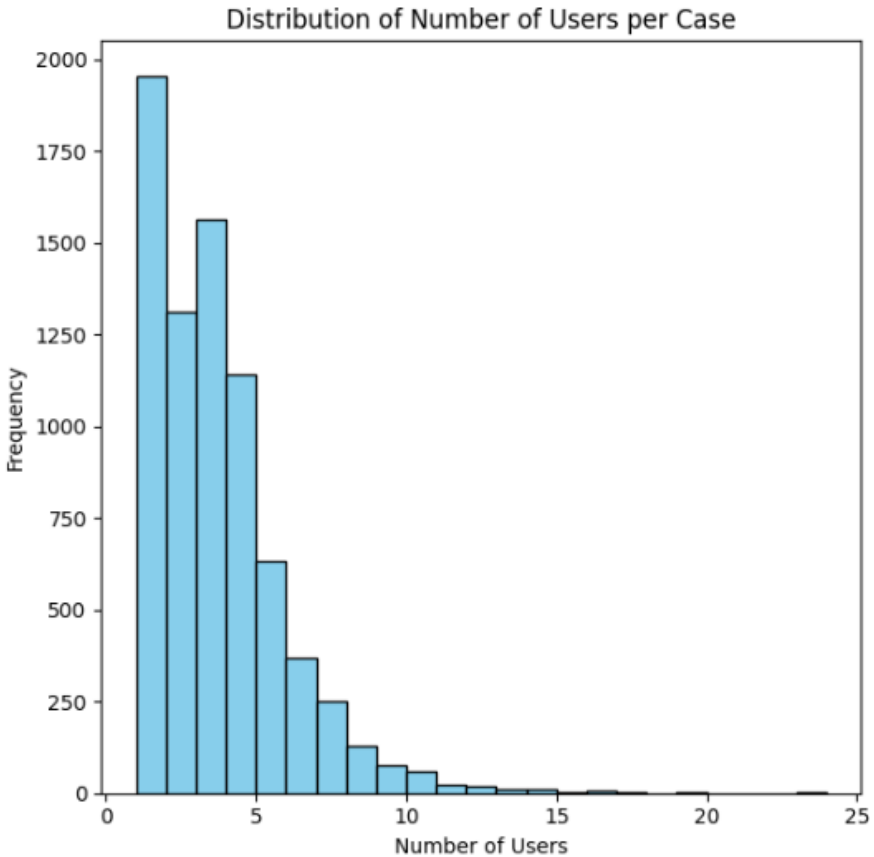
Constraints

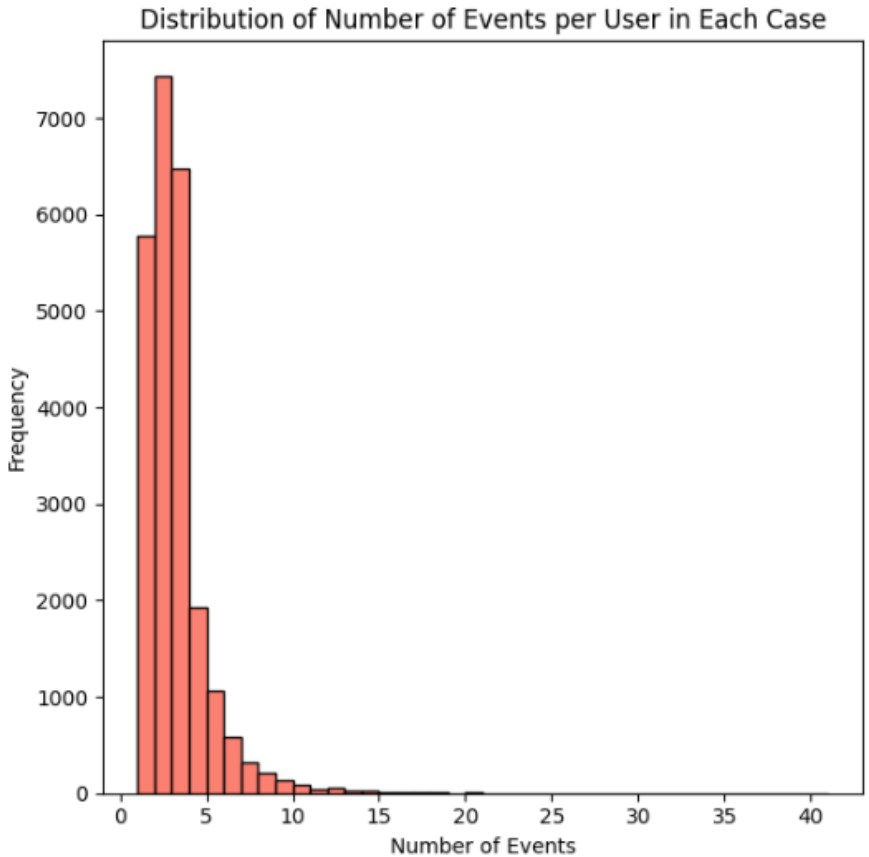
- **Sporadic Resources:** Cases include resources that are used only once or twice.
- **Frequent Multitasking:** Resources are heavily multitasked, performing multiple activities in short time lapses.

Validation


Multitasking Occurrences per Resource
Count
0
200
400
600
800
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
230
240
250
260
270
280
290
300
310
320
330
340
350
360
370
380
390

Validation


Distribution of Number of Users per Case
Frequency
Number of Users
0
250
500
750
1000
1250
1500
1750
2000
0
5
10
15
20
25


Distribution of Number of Events per User in Each Case
Frequency
Number of Events
0
1000
2000
3000
4000
5000
6000
7000
0
5
10
15
20
25
30
35
40

BPI 2013 Results and Winners

Winners analysis

Exhaustive and complete data analysis and even some prescriptions, but lacks depth and validation.

Problem

Our Question

How can the existing process be improved? by how much?

Problem

Our Question

How can the existing process be improved? by how much?

Our Proposal

Simulation of what-if scenarios to quantify how much these diverge from the original process

Tools and Methods

Tools

- Apromore
- Python + Pandas
- Excel

Methodology

Over 5 simulations of every scenario will be performed, and the one closest to the mean will be chosen to compare against the null model log.

Simulations and What-if scenarios

What-if...

1. Scenario 1
2. Scenario 2
3. Scenario 3

Scenario 1: Prioritization

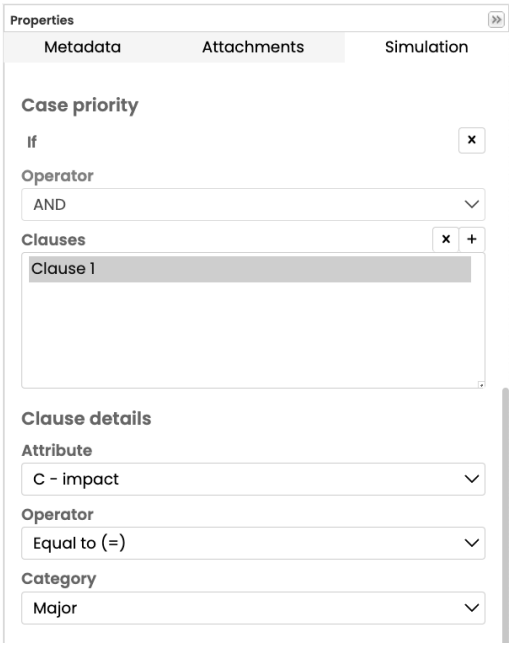


Figure: Prioritization settings:
Major > *High* > *Medium* > *Low*

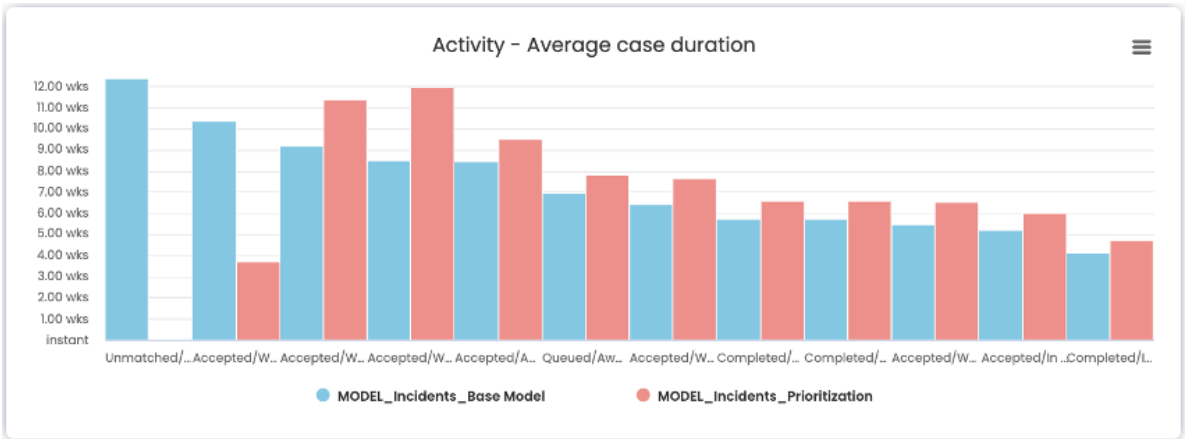
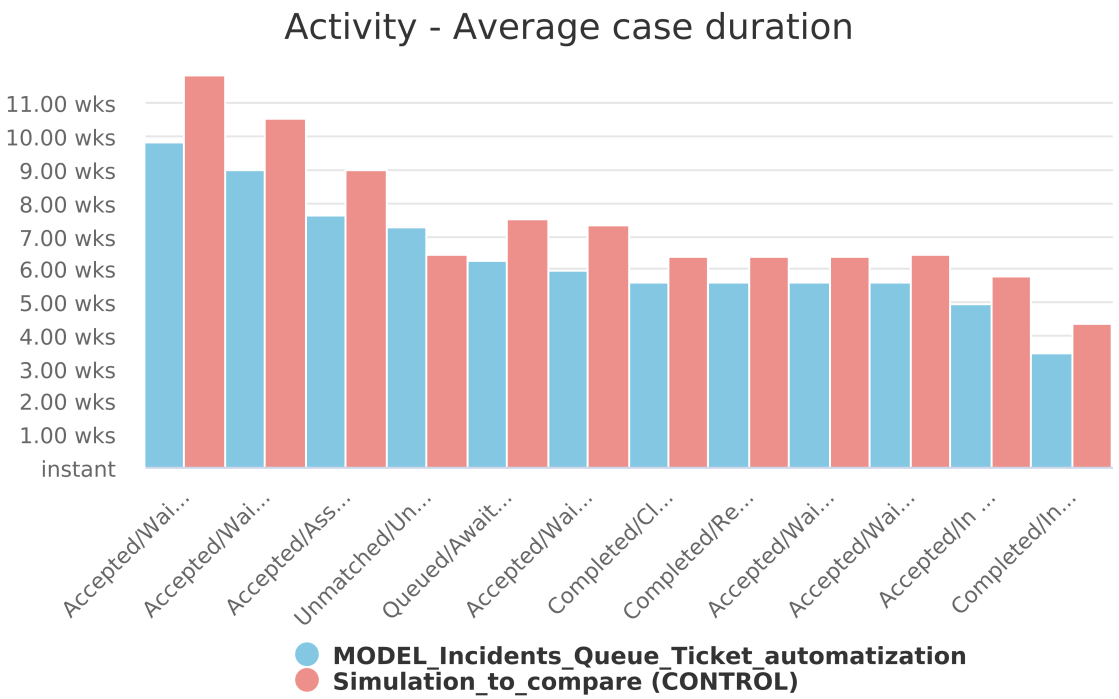
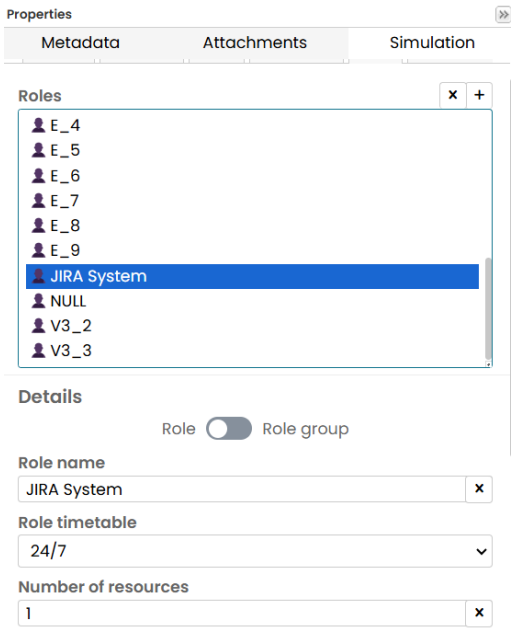


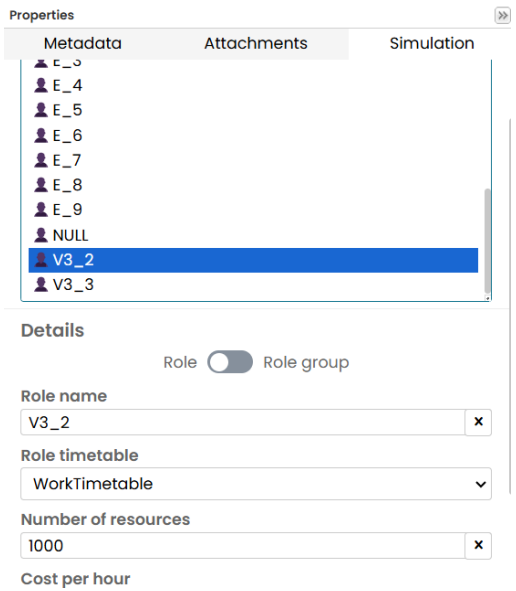
Figure: Average case duration comparison

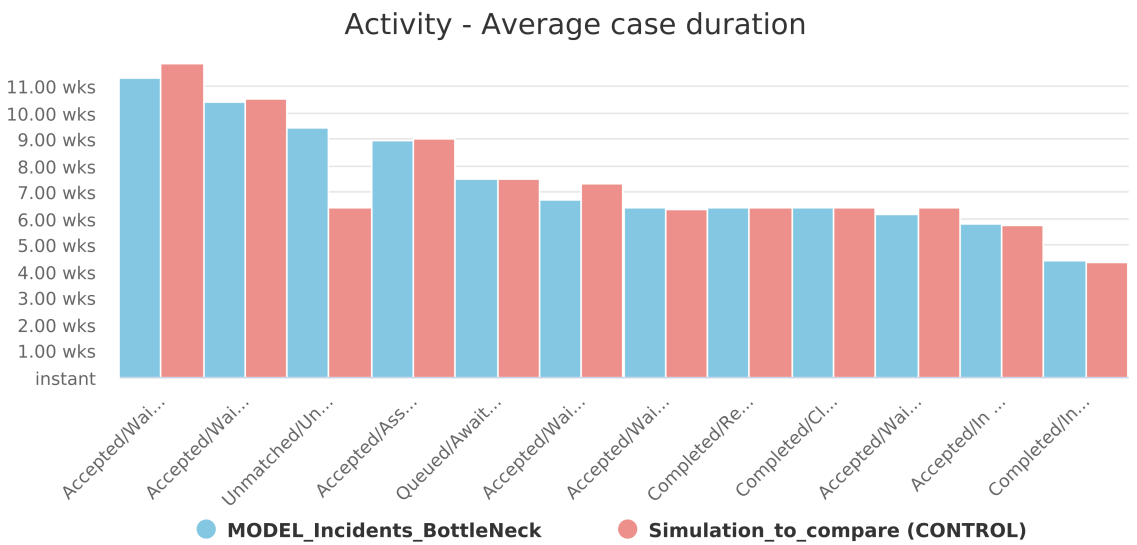
Scenario 2: Automatization Ticket





Scenario 3: Redistribution of Resources





Results

Scenario	Avg. Case Duration	Avg. Activity Duration	Avg. Resource Frequency
Original Dataset	1.33 mths	2.73 days	16.86
Prioritization by Impact	1.34 mths	2.69 days	16.93
Queue Ticket automatization	1.14 mths	2.45 days	16.89
Redistribution of resources Sim. 1	1.32 mths	2.70 days	16.78
Redistribution of Resources Sim. 2	1.33 mths	2.73 days	17.0
Case 1, 2 and 4 together	1.1375 mths	2.46 days	17.11

Reflections and conclusions

Conclusions

1. Real-world logs are often chaotic and incomplete, significantly different from the structured datasets used in class.

Reflections and conclusions

Conclusions

1. Real-world logs are often chaotic and incomplete, significantly different from the structured datasets used in class.
2. Establishing a **feedback loop with process owners** is essential for improving the process and capturing critical business insights.

Reflections and conclusions

Conclusions

1. Real-world logs are often chaotic and incomplete, significantly different from the structured datasets used in class.
2. Establishing a **feedback loop with process owners** is essential for improving the process and capturing critical business insights.
3. Apromore, while still an **immature tool** for process mining, demonstrates immense potential for future development.

Reflections and conclusions

Conclusions

1. Real-world logs are often chaotic and incomplete, significantly different from the structured datasets used in class.
2. Establishing a **feedback loop with process owners** is essential for improving the process and capturing critical business insights.
3. Apromore, while still an **immature tool** for process mining, demonstrates immense potential for future development.
4. Implementing the final prescribed scenario can yield an approximate **20% speed-up** in process performance and even conformity.

Future Work

Areas for Future Exploration

1. Conducting **additional what-if analyses**, exploring organizational granularities beyond roles: groups, departments, geographical levels...

Future Work

Areas for Future Exploration

1. Conducting **additional what-if analyses**, exploring organizational granularities beyond roles: groups, departments, geographical levels...
2. Collaborating with process owners to refine findings and initiating a more informed **process mining iteration** with enhanced business insights and improved event logs.

Future Work

Areas for Future Exploration

1. Conducting **additional what-if analyses**, exploring organizational granularities beyond roles: groups, departments, geographical levels...
2. Collaborating with process owners to refine findings and initiating a more informed **process mining iteration** with enhanced business insights and improved event logs.
3. Reporting identified problems and suggesting improvements to Apromore via the **GitHub Issues tracker**, fostering community-driven enhancements.

Future Work

Areas for Future Exploration

1. Conducting **additional what-if analyses**, exploring organizational granularities beyond roles: groups, departments, geographical levels...
2. Collaborating with process owners to refine findings and initiating a more informed **process mining iteration** with enhanced business insights and improved event logs.
3. Reporting identified problems and suggesting improvements to Apromore via the **GitHub Issues tracker**, fostering community-driven enhancements.
4. Contributing to the Apromore community by addressing and resolving the found **source code bugs** during development.