

Using process mining to detect undesirable phenomena in the process of submitting declaration for international and national trips

Zhilkin I.I.¹, Deyneko M.A.², Fadyushina.M.V.³, Iov R.V.⁴, Makarov V.A.⁵, Alyaev G.I.⁶, Bordachev V.Y.⁷, Karpova N.G.⁸, Nikolaev V.A.⁹, Nizamov I.V.¹⁰, Shusterzon K.A.¹¹, Stanevich A.V.¹² and Tiguntsev Y.A.¹³
Sberbank, Russian Federation, Team Baikal Seals

- ¹Auditor, Team leader, IIZhilkin@sberbank.ru
²Lead auditor, madeynenko@sberbank.ru
³Lead auditor, Fadyushina.Ma.Va@sberbank.ru
⁴Senior auditor, Iov.R.V@sberbank.ru
⁵Senior auditor, Makarov.V.Aleksand@sberbank.ru
⁶Auditor, GIAlyaev@sberbank.ru
⁷Auditor, VYBordachev@sberbank.ru
⁸Auditor, NGKarpova@sberbank.ru
⁹Auditor, VAlNikolaev@sberbank.ru
¹⁰Auditor, Nizamov.I.V@sberbank.ru
¹¹Auditor, KAShusterzon@sberbank.ru
¹²Auditor, AVStanevich@sberbank.ru
¹³Auditor, yatiguntsev@sberbank.ru

Abstract. For this year's Business Process Intelligence Challenge, we collected data on the TU/e reimbursement process. The focus of our research is the answer to all the questions put to us in the description of competition.

Keywords: Process Mining, BPIC 2020, Business Process, Fluxicon Disco, PM4PY, Process visualization

INTRODUCTION

Process mining technology includes automated construction of models of actually executed business processes based on the analysis of the event log. At the same time, the restored process models combined with data on the process execution time and organizational structure elements allow you to see all the hidden shortcomings, providing business process owners and analysts with a huge amount of materials for further improvement.

Analyzing the models of restored business processes, you can find: excessive approval cycles, cancellation of previously performed actions, "ping-pong" of performers, delays in the execution of functions, unnecessary actions in the process,

unnecessary or ineffective performers, and most importantly, exceptional situations in the processes that occur due to errors of performers and which require serious resources to fix. As part of the competition, we investigate the process of submitting various documents, including declarations (domestic and international), travel permits and request for payment.

For events of 2018, after an employee submits a request for a trip, it is sent to the business trip administration for approval.

If it's approved, then the trip request is sent to the budget owner, and only then to supervisor. Also note that if budget owner and supervisor the same person, then only one of these steps is taken. In some cases, the director also needs to approve the request.

In all cases, failure leads to one of two outcomes. Either the employee, resubmits the request for review, or the employee also rejects the request.

If the approval stream has a positive result, then the payment is requested and made.

Travel permits follow a slightly different flow, since payment is not required in this case. Instead, after all the approval stages, a ride can take place that is specified with an estimated start and end date. These dates are not the exact dates of the trip, but rather are estimated by the employee when applying for a permit. Actual travel dates are not recorded in the data, but in most cases, they should be close to the specified dates.

After the travel authorization is approved, but before the trip begins, employees may request a refund of pre-paid travel expenses. Multiple queries can be submitted independently from each other. After the end of the trip, the employee receives several reminders about the need to submit a trip declaration, and there are cases you can see several declarations, which in our opinion is normal for this process.

For event of 2017, there are some differences, since it was a pilot year, and the process changed slightly several times, so there may be some errors in solving problems.

We are also using the Disco program for the first time and decided to use it specifically in this challenge to gain experience with it. Disco contains fast process mining algorithms, and nice log management and filtering framework. More detail about our Disco experience you can read on conclusion of this work.

1 What is the throughput of a travel declaration from submission (or closing) to paying?

Median of values processing time is 125,11 Hours (5,21 Days). If we consider "returned" documents or not, it didn't change median values. It's not correct to take average time processing, because it has a lot of difference with reality. Example of process travel declaration from submission to paying without any reject you can see on Figure 1.

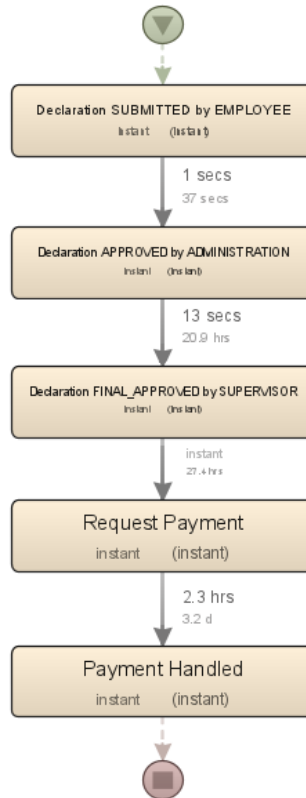


Fig. 1. Example of process travel declaration from submission to paying.

2 Is there are difference in throughput between national and international trips?

According to answer from question №1, we took time median, to use “Little’s” law

$$L = \lambda W \quad (1)$$

where:

- L – long-term average number of customers
- Gamma – long-term average effective arrival rate
- W – time that customer spends in the system.

$$\text{ABS FREQ.Domestic Declaration} = \frac{11531 \text{ documents}}{125,11 \text{ hours}} \approx 83 \text{ documents/hour}$$

$$\text{CASE FREQ.Domestic Declaration} = \frac{10365 \text{ documents}}{125,11 \text{ hours}} \approx 92 \text{ documents/hour}$$

To get things more clearly, we removed stages “START TRIP” and “END TRIP” and their connections to documents processing. Our assumption is, that if employer is on trip or he’s just given documents, there’s no point. Calculation was made only for document flow capacity.

$$\text{ABS FREQ.Domestic Declaration} = \frac{6255 \text{ documents}}{153 \text{ hours}} \approx 41 \text{ documents/hour}$$

$$\text{CASE FREQ.Domestic Declaration} = \frac{6001 \text{ documents}}{153 \text{ hours}} \approx 39 \text{ documents/hour}$$

3 Are there differences between clusters of declarations, for example between cost centers/departments/projects etc.

We removed all necessary process stages and created first process graph (Appendix 1, Figure 1).

Analysis of payment declarations revealed the following differences in the process clusters, that are responsible for processing declarations and organizing payments:

If the department of the employee has its own budget for travel, then the approval takes place without the approval of BUDGET_OWNER, if there is no budget, the approval of BUDGET_OWNER is required (Figure 2).

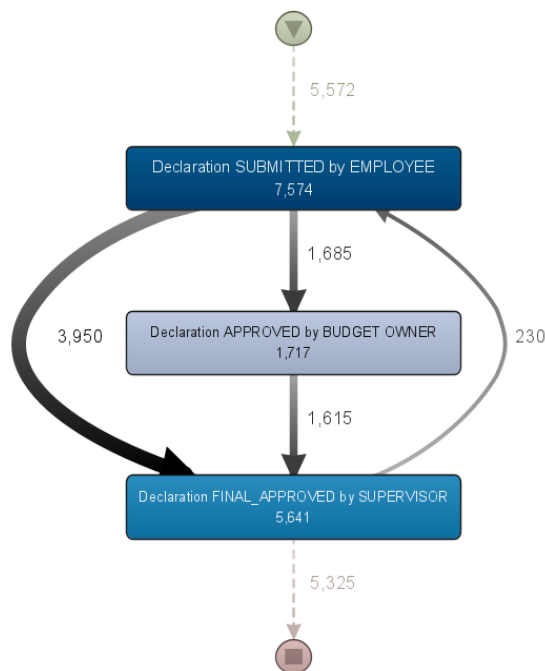


Fig. 2. Process graph from declaration submitted to declaration final approved by supervisor for answer third question.

1. Declarations agreed by PRE_APPROVER are always correct, are not returned for revision and are not rejected.
2. Some declarations require additional agreement DIRECTOR (Figure 3).

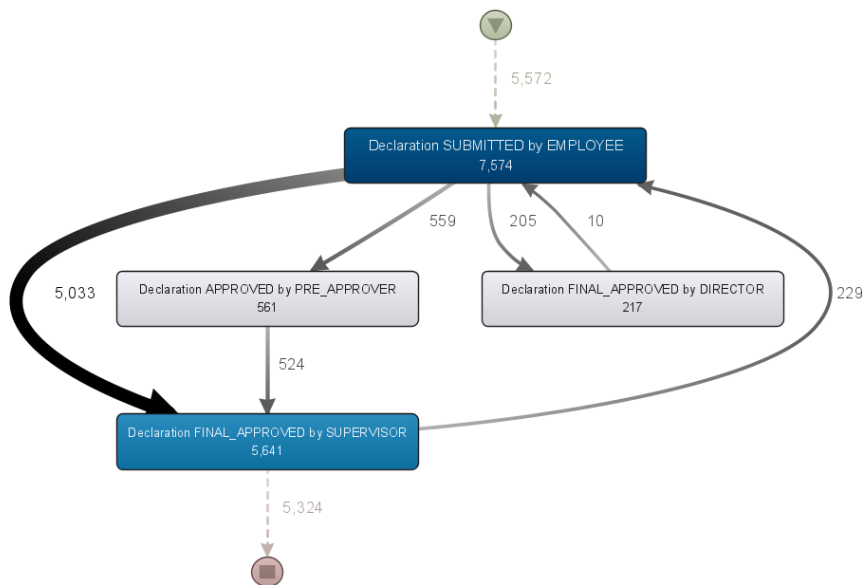


Fig. 3. Process graph from declaration submitted to declaration final approved supervisor or director for answer third.

4 What is the throughput in each of the process steps, i.e. the submission, judgement by various responsible roles and payment?

The speed of process stages was also investigated. The processing speed at the main stages is presented in the table below:

Table 1. throughput in each of the process steps.

From	To	dT median	T max	T min
Permit SUBMITTED by EMPLOYEE	Permit FINAL_APPROVED by SUPERVISOR	24.8 h	90.1 d	1 s
Permit SUBMITTED by EMPLOYEE	Permit FINAL_APPROVED by DIRECTOR	4 d	31.3 w	2 s
Permit FINAL_APPROVED by SUPERVISOR	Request For Payment SUBMITTED by EMPLOYEE	7.4 d	31.7 w	2 s

Request For Payment SUBMITTED by EMPLOYEE	Request For Payment by FINAL_APPROVED SUPERVISOR	26.6 h	35.1 d	3 s
Request For Payment SUBMITTED by EMPLOYEE	Request For Payment REJECTED by EMPLOYEE	41 h	79.9 d	11 s
Request For Payment SUBMITTED by EMPLOYEE	Request For Payment by FINAL_APPROVED DIRECTOR	3.2 d	19.6 d	15.4 min
Declaration SUBMITTED by EMPLOYEE	Declaration FINAL_APPROVED by SUPERVISOR	46.6 h	14 m	2 s
Declaration SUBMITTED by EMPLOYEE	Declaration REJECTED by EMPLOYEE	50.5 h	30.3 w	18 s
Declaration SUBMITTED by EMPLOYEE	Declaration FINAL_APPROVED by DIRECTOR	4.1 d	68.7 d	5 s
Declaration FINAL_APPROVED by DIRECTOR	Request Payment	23.9 h	19.9 d	1 s
Declaration FINAL_APPROVED by SUPERVISOR	Request Payment	28.1 h	38 d	0 s
Request For Payment FINAL_APPROVED by SUPERVISOR	Request Payment	24.6 h	43 d	0 s
Request Payment	Payment Handled	3.2 d	29.1 d	75.6 min

According to the table results, we decided to take median of time, because now it characterizes most of cases.

5 What are the bottlenecks in the process of a travel declaration?

According to diagram below, following assumptions can be made (Figure 4):

- Problems with Payment Handling.
- Problems with BUDGET OWNER. When employers asking for travel and budget is missing, they have to ask another departments to get money. This takes a lot of time.

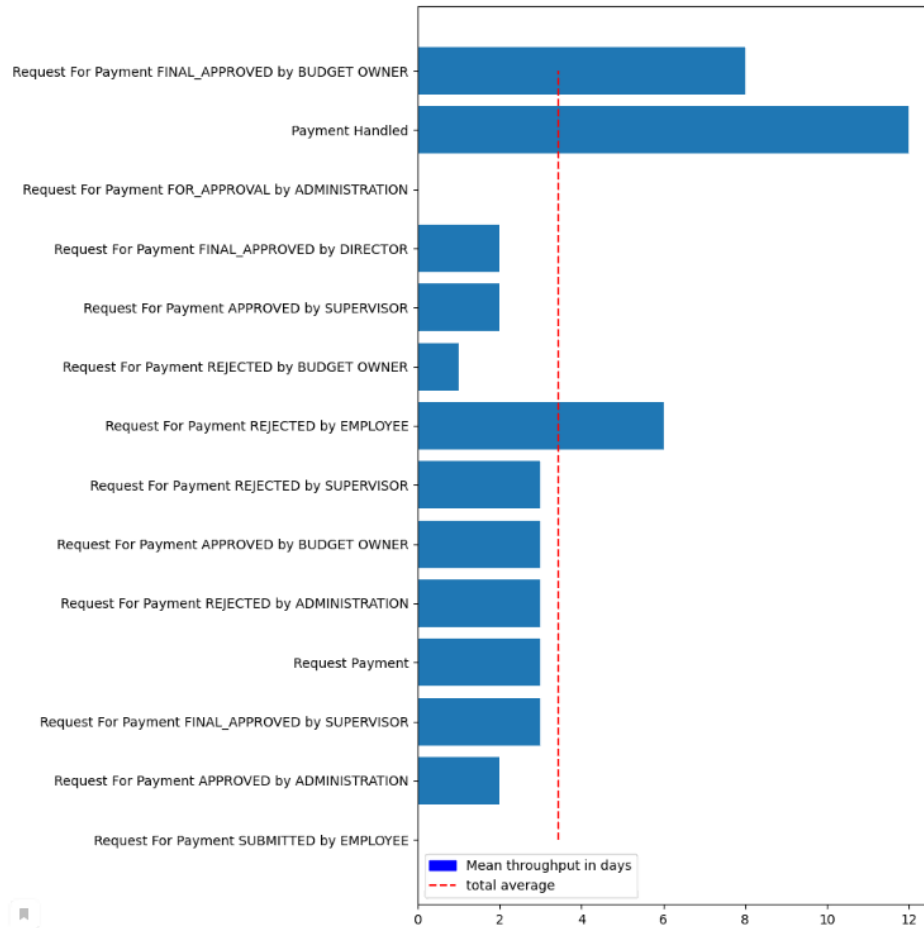


Fig. 4. Bottlenecks in the process of a travel declaration.

6 Where are the bottlenecks in the process of a travel permit (note that there can be multiple requests for payment and declarations per permit)

According to statistics – average processing time of all events – 7 days. This value will be set as normal to whole process.

After we counted all duration time of each events, the following table of bottleneck events was considered (Table 2).

Table 2. Bottlenecks in the process of a travel permit.

Stage	Average throughput
Permit REJECTED by MISSING	10-12 days
Permit REJECTED by DIRECTOR	10-12 days
Declaration SAVED by EMPLOYEE	10-15 days
Permit REJECTED by EMPLOYEE	10-13 days
Permit FINAL_APPROVED by DIRECTOR	30-35 days
PERMIT SAVED BY EMPLOYEE	50-55 days
Send Reminder	40 days
Payment Handled	20-25 days
End trip	10-12 days
Permit Final_APPROVED by SUPERVISOR	20-25 days

As following some assumptions can be made:

- There is bottleneck in DIRECTOR and SUPERVISOR approval. There is not so many cases, so that means, that company must find ways to improve speed of approval
- There is bottleneck in Payment Handled. This assumption must be checked according to banking rules and University rules.

7 How many travel declarations get rejected in the various processing steps and how many are never approved? (How many travel declarations are rejected at various stages of processing, and how many are never approved?)

When analyzing the process of issuing travel declarations, it was revealed that 4914 applications received a rejection status at various stages of verification, some of them were again sent to the previous step and received an approved status. The total number of declarations that were never approved was 2034. Most of them were rejected by the administration, or employer canceled it (declaration rejected by employee).

8 How many travel declarations are booked on projects?

To find out how many declarations are reserved for projects, you can calculate the number of project numbers that are found in the logs of international and internal declarations.

For international declarations, 824 different project numbers were found for 6255 declarations. However, there are also 2,441 declarations out of the total number of

international declarations that do not mention the project number, which is 39.02% of the total number of international declarations.

For internal declarations, there is no direct mention of the project number in the logs.

9 How many revisions were made to the declaration?

As revised declarations, we consider declarations that were rejected at a certain stage and after this stage were returned to work. We analyze internal declarations and external ones; we find the total amount of corrections:

Internal declarations (DOMESTIC DECLARATIONS):

- Rejected by Supervisor 293
- Returned to work 2
- Lost 91
- Returned to work 61
- Rejected by budget owner 59
- Returned to work 1
- Pre-approved 86 rejected
- Returned to work 1
- Rejected by Administrators 952
- Returned to work 19
- Rejected by employee 1365
- Returned to work 1081

Total: $2 + 61 + 1 + 1 + 19 + 1081 = 1165$ (10500 total declarations, 11%)

External declaration (INTERNATIONAL DECLARATIONS):

- Rejected by Supervisor 126
- Returned to work 3
- Rejected by admin 1549
- Returned to work 24
- Lost 103
- Returned to work 86
- Rejected by employee 1780
- Returned to work 1601

Total: $3 + 24 + 86 + 1601 = 1714$ (6449 total declarations, 26%)

Conclusions:

In total, the declarations were corrected by TOTAL: $1165 + 1714 = 2879$.

There are 2.5 times more corrections in international declarations than in domestic ones.

10 Are there any double payments (duplicates)?

We analyze payments:

- Prepayment Request For Payment;
- Postpaid Prepaid Travel Cost.

Processes that have a double pass of the Payment Handled event:

- are absent in Request For Payment;
- not included in Prepaid Travel Cost.

Conclusions:

There are no double payments (duplicates).

11 Are there any declarations that were not properly preceded by an approved travel authorization? Are there any declarations that are not authorized?

Based on the description of the processes, it is typical for the travel authorization process that in the log the start and end dates of the trip are not accurate, but should be close enough to the real dates. It should also be noted that a travel permit must be obtained before traveling. And this is where it gets interesting, since we will be able to see significant differences between the end date of the trip and the date of filing a travel permit, that is, when the trip first occurs, and only then the travel permit is applied. We will consider such cases for not properly submitted travel approvals. There are also cases when the final approval of a travel permit occurs immediately between the start of the trip and the end of the trip, but these cases will not be considered, since, from the above, the start and end dates of the trip presented in the log file are not accurate.

And as noted from the description of the processes, 2017 was a pilot year. It is difficult to say what specific requirements were imposed on him, but we can show the percentage when applying for a travel permit for travel in 2017, in 2018 and in two years together (as well as in the fragment of interest to us, 3 trips that ended in 2016 were identified).

To answer the question, we will analyze the PermitLog. In Figure 5, a fragment of the graph is presented in which we can see the direct relationship between the date of filing a travel permit and the date of the end of the trip. We are interested in the edge starting at the end of the trip and ending in the application of permission and will display the path of interest to us in figure 6.

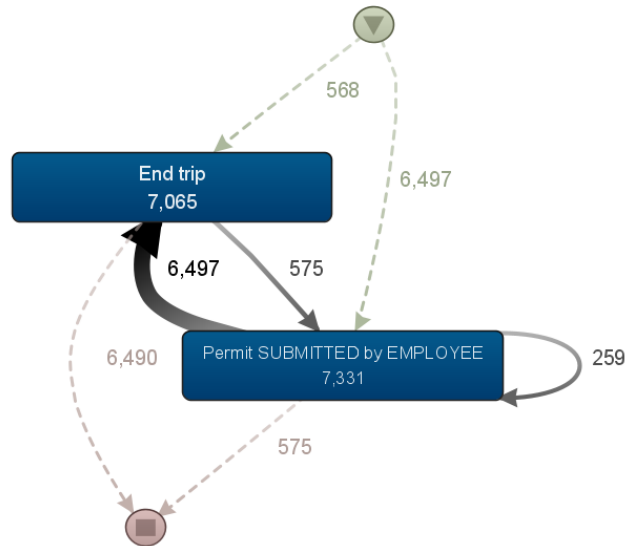


Fig. 5. Relationship between the date of filing a travel permit and the date of the end of the trip.

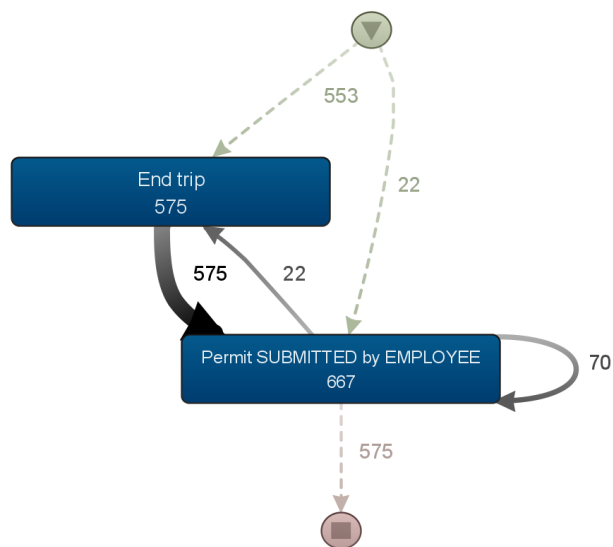


Fig. 6. Relationship between the date of filing a travel permit and the date of the end of the trip for start from End trip event.

In the following table (Table 3), we can see the main metrics for the time between END TRIP and Permit SUBMITTED by EMPLOYEE events.

Table 3. Main metrics for answer 11th question.

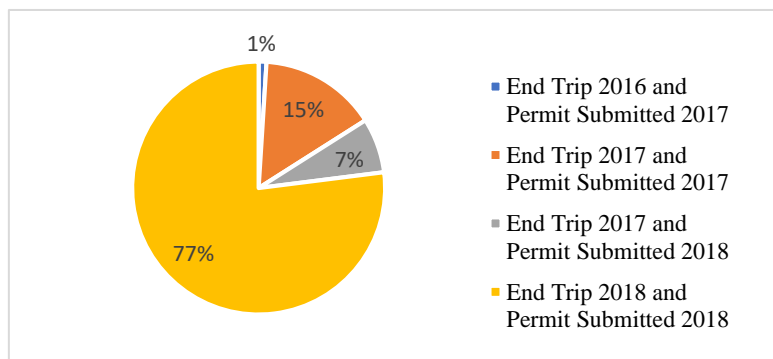
Criterion	Value	DESCRIPTION
Max	1 year, 105 days	Travel permit 25041
Min	35 mins, 55 secs	Travel permit 64727
Median	10.6 days	
> 7 days	326 CASES	
> 14 days	251 CASES	
> 28 days	145 CASES	

Case ID	Events	Variant	Started	Finished	Duration
travel permit 25041	2	Variant 1	27.08.2017 06:00:00	10.12.2018 18:45:45	1 year, 105 days
travel permit 67652	2	Variant 1	15.02.2017 07:00:00	16.02.2018 18:20:13	1 year, 1 day
travel permit 8719	2	Variant 1	19.02.2017 07:00:00	16.02.2018 18:22:09	362 days, 11 hours
travel permit 49427	3	Variant 2	20.12.2017 07:00:00	30.11.2018 21:21:52	345 days, 14 hours
travel permit 18616	4	Variant 4	20.09.2017 06:00:00	17.08.2018 17:17:06	331 days, 11 hours
travel permit 8729	2	Variant 1	30.03.2017 06:00:00	16.02.2018 18:23:49	323 days, 12 hours
travel permit 20314	2	Variant 1	11.02.2018 07:00:00	28.12.2018 22:50:53	320 days, 15 hours
travel permit 26601	2	Variant 1	19.01.2018 07:00:00	26.11.2018 17:50:42	311 days, 10 hours
travel permit 36857	2	Variant 1	30.03.2017 06:00:00	10.01.2018 19:06:34	286 days, 13 hours
travel permit 15983	2	Variant 1	15.03.2018 07:00:00	10.12.2018 18:49:16	270 days, 11 hours
travel permit 8240	2	Variant 1	18.05.2017 06:00:00	01.02.2018 00:28:25	258 days, 18 hours
travel permit 16614	3	Variant 2	31.12.2017 07:00:00	07.09.2018 03:59:39	249 days, 20 hours
travel permit 18592	5	Variant 7	15.12.2017 07:00:00	21.08.2018 00:03:41	248 days, 17 hours
travel permit 1112	2	Variant 1	23.07.2017 06:00:00	05.03.2018 02:09:58	224 days, 20 hours
travel permit 85685	2	Variant 1	14.03.2017 07:00:00	19.10.2017 17:07:43	219 days, 10 hours
travel permit 43743	2	Variant 1	12.07.2017 06:00:00	13.02.2018 16:14:18	216 days, 10 hours
travel permit 12277	2	Variant 1	11.05.2018 06:00:00	11.12.2018 22:59:06	214 days, 16 hours
travel permit 17058	3	Variant 3	31.03.2017 17:52:18	16.10.2017 21:44:51	199 days, 3 hours
travel permit 17608	2	Variant 1	28.05.2018 06:00:00	11.12.2018 22:57:41	197 days, 16 hours
travel permit 47104	2	Variant 1	09.11.2017 07:00:00	23.05.2018 22:17:37	195 days, 15 hours
travel permit 17602	2	Variant 1	02.06.2018 06:00:00	11.12.2018 22:54:55	192 days, 16 hours
travel permit 74227	3	Variant 3	02.03.2017 00:06:01	05.09.2017 20:31:54	187 days, 20 hours
travel permit 14009	2	Variant 1	17.06.2018 06:00:00	20.12.2018 18:26:23	186 days, 12 hours
travel permit 76455	2	Variant 1	05.10.2016 06:00:00	06.04.2017 19:32:10	183 days, 13 hours
travel permit 32014	2	Variant 1	28.08.2017 06:00:00	27.02.2018 02:15:50	182 days, 20 hours

Fig. 7. Top 25 terms of duration of difference between the end of a business trip and the granting of a travel authorization made in Disco

In total, we have 145 certified trips, permits for which were submitted only 28 days after the end of the trip, 251 entries, with permits submitted 14 days after the end of the trip.

On pie below (Figure 8) showing the percentages between different years of travel endings and travel authorizations.

**Fig. 8.** Different years of travel endings and travel authorizations as a percentage.

To answer of second part of question 11, we will build a graph from the PermitLog containing the key events: Start Trip, End Trip, Permit FINAL_APPROVED by DIRECTOR, Permit FINAL_APPROVED by SUPERVISOR, Declaration FINAL_APPROVED by DIRECTOR, Declaration FINAL_APPROVED by SUPERVISOR. Using the filter, we remove all other events that do not contain at least one of the events above

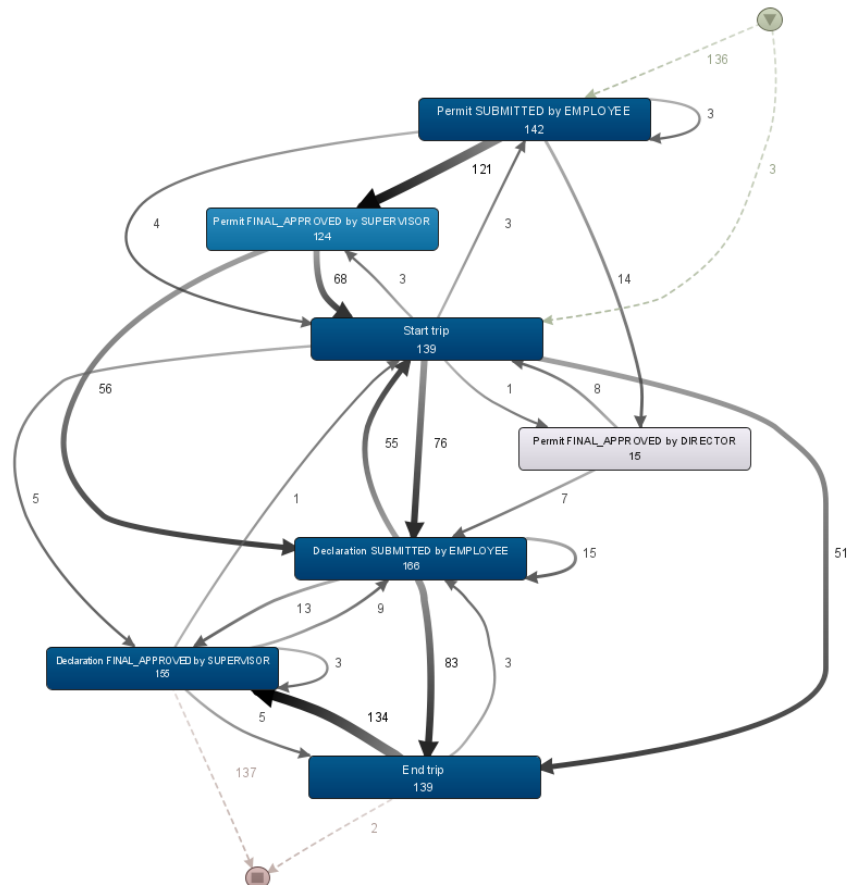


Fig. 9. Graph from the PermitLog for answer of 11th question.

From the graph (Figure 9), we can see that out of 139 trips, there are 15 travel permits with final DIRECTOR approval and 124 travel permits with final SUPERVISOR approval, which tells us that there are no completed declarations without a travel permit, since the sum of final approvals is equal to the number of trips.

12 How many travel declarations are submitted by the traveler, and how many - by the authorized person?

To answer this question, we assume that the traveler definitely cannot submit a declaration during his trip - an authorized person does it for him. Then we need to find all the events of filing a declaration by an employee that occur between the beginning of the trip and the end of the trip. Leaving only the events of interest to us, we get the graphs of the processes (Figure 10).

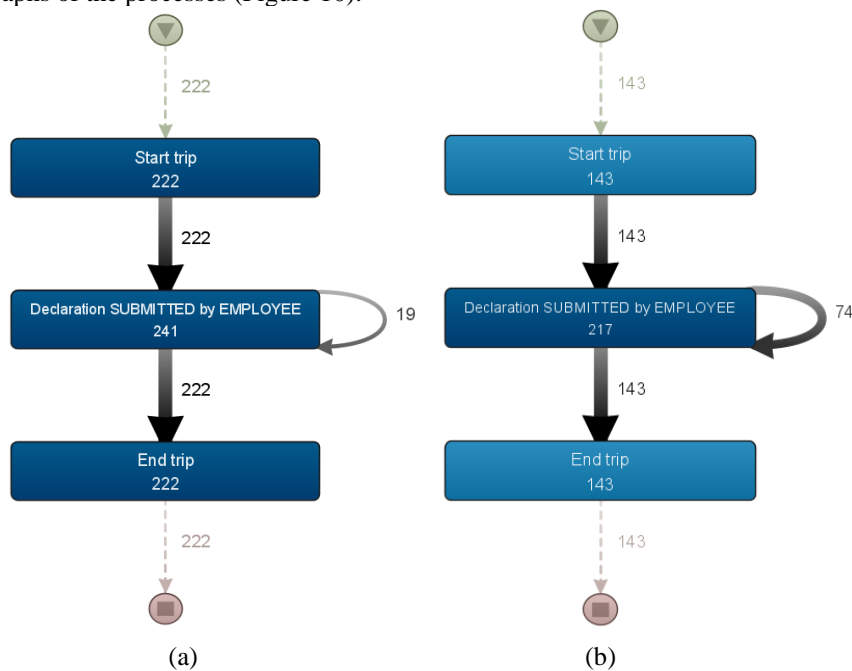


Fig. 10. Graphs from the International Declarations (a) and PermitLog (b) for answer of 12th question.

The Domestic Declarations log does not contain the trip start and trip end events, so we excluded the trip within the home country from the analysis, but instead took the Permit Log to test the validity of our hypotheses. As a result, we got 222 numbers of declarations unique to the International Declarations log and 198 numbers of declarations unique to the Permit Log. By combining unique numbers and removing duplicate values, we got 222 unique values, which we observed in the International Declarations log. This number tells us about the number of declarations filed by the authorized person, not the traveler.

13 How many travel declarations are first rejected because they are submitted more than 2 months after the end of a trip and are then re-submitted?

In the process of declaring international travel after the condition from the assignment came, the employee rejected and resubmitted 16 travel declarations (Figure 11). In the presented Permit log, the number of travel declarations rejected and resubmitted by the employee is 14.

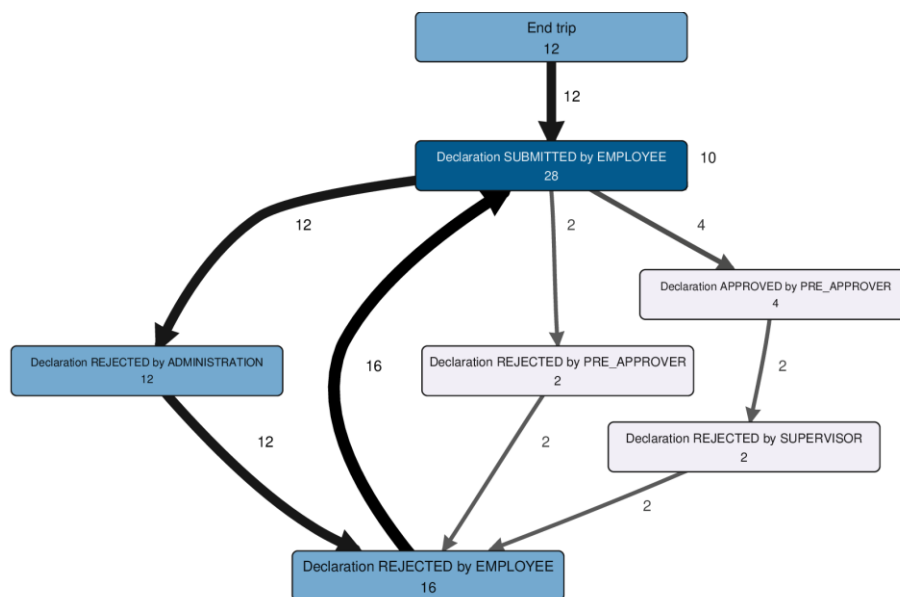


Fig. 11. Shows fragments of the international travel declaration graph, filtered by the required parameters

14 Is this different between departments?

The presented initial data are insufficient for a correct assessment and substantiation of differences in the “paths” of various departments. From the data provided, it can be said for sure that there are no differences in the process of declaring internal business trips between different structural units of the "Organizational entity". In the other three processes, there is a clear difference in the debugging of work with department declarations, which affects the number of stages and cycles passed at these stages, and subsequently affects the time spent on the process. An excellent example of a correct declaration process is “Unit 65488”, which had only one business trip, but was consistent with the ideal process.

15 How many travel declarations are not approved by budget holders on time (7 days) and then automatically redirected to regulatory authorities?

We are interested in all travel declarations approved by the administrator, but in which there were no events related to the holder of the budget. For analysis, we will build a graph of this process, excluding all events with the actions of the budget holder for international declarations (Figure 12) and internal declarations (Figure 13) in advance. Next, let's remove all supervisor's actions that are not final for international declarations (Figure 14) and internal declarations (Figure 15).

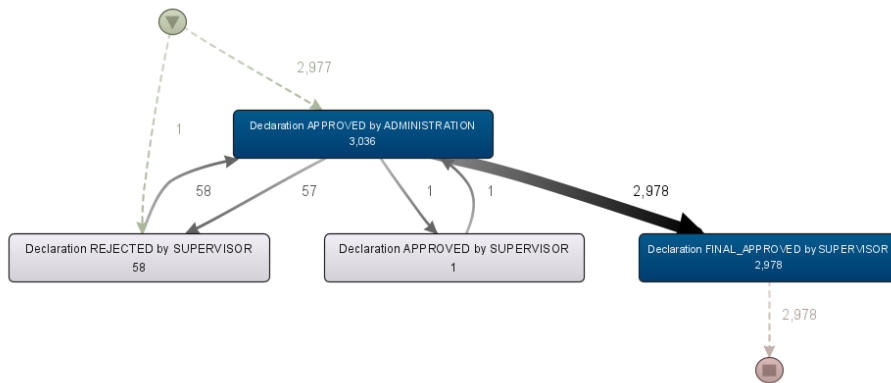


Fig. 12. Constructed graphs of the approval process for international declarations without the participation of the budget holder

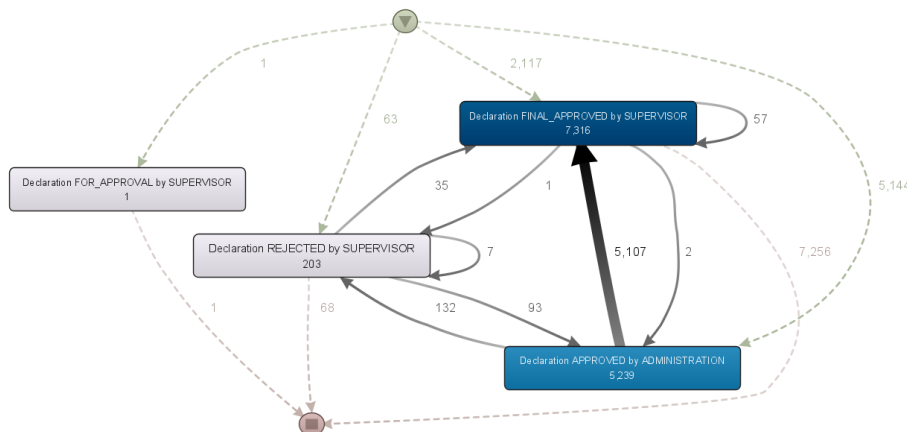


Fig. 13. Constructed graphs of the approval process for domestic declarations without the participation of the budget holder

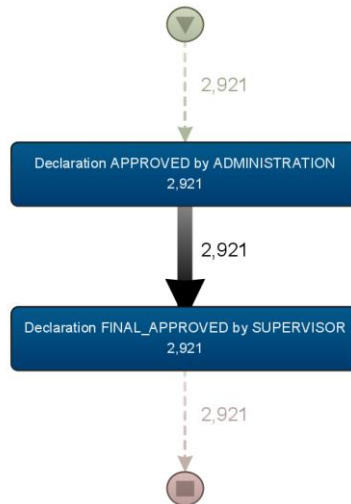


Fig. 14. Processed columns of the approval process for international declarations without participation of the budget holder

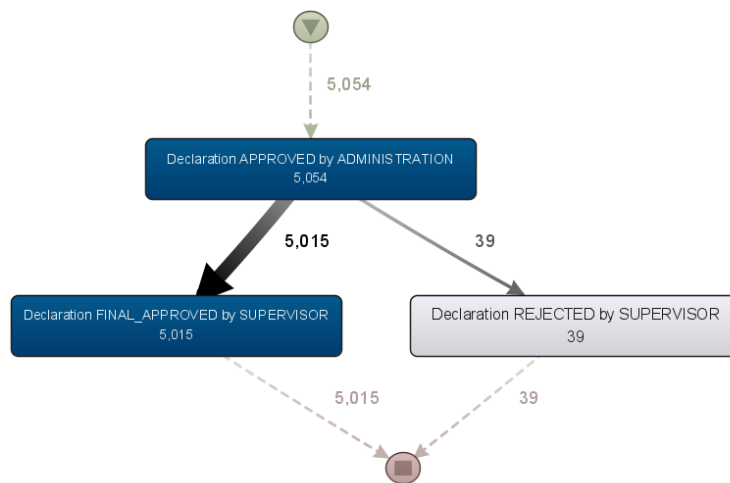


Fig. 15. Processed columns of the approval process for domestic declarations without participation of the budget holder

We found 2921 events for international travel declarations and among them only 321 events have a 7-day difference between administrator approval and supervisor approval without direct involvement of the budget holder.

For internal declarations, we found 5054 events, among which only 191 events have a difference of 7 days between the approval by the administrator and the approval by the supervisor without the direct participation of the budget holder

16 Next to travel declarations, there are also requests for payment. They are only intended for non-TU / e employees. Are there any TU / e employees who have applied for payment instead of a travel declaration?

To search for TU / e employees who, instead of filing a travel declaration, used a request for payment, we found all unique Project Ids from the RequestForPaying log and compared them with the unique Project Ids from the International Declarations and Permit Logs, assuming in advance that if the project is a result work of TU / e, then it could be found at various stages in the logs we are considering. As a result of the analysis, no duplicate Id Projects were found that would be in the RequestForPaying log or in one of the International Declarations / Permit Logs, from which we can conclude that the TU / e employees who applied for payment instead of the travel declaration, not found.

CONCLUSION

As part of the BPI 2020 challenge project, our team conducted a study of log files for filing various TU / e documents related to employee travel, including travel declarations (domestic and international), travel permits, payment requests, and prepaid travel expenses. Due to the fact that our team took log files for absolute reliability (excluding data anonymization), some questions posed by the organizers of the BPI 2020 challenge were not given accurate answers. It should also be noted that the description does not contain internal regulatory documents of Tu / e that regulate the process of forming declarations, so you can only accept this process as ideal based on the principle of majority and a brief description of the process.

Our team in this challenge really enjoyed using the tool for process mining Disco. It was the first time we worked with it and we were impressed with the user-friendly interface, which provided a low complexity of learning in the program. Disco has a wide range of data visualization and filtering, which provides a quick understanding of the ongoing processes.

As mentioned above, our team answered all the questions and showed the bottlenecks and shortcomings in the process of Tu/e employees when working with business trip declarations and other documents.

References

1. BRI Challenge 2020, <https://icpmconference.org/2020/bpi-challenge>, last accessed 2020/08/20
2. Wil van der Aalst: Process Mining: Discovery, Conformance and Enhancement of Business Processes. Springer (2011)
3. Process Mining and Automated Process Discovery Software for Professionals - Fluxicon Disco. <https://fluxicon.com/disco/>, last accessed 2020/08/20

APPENDIX 1

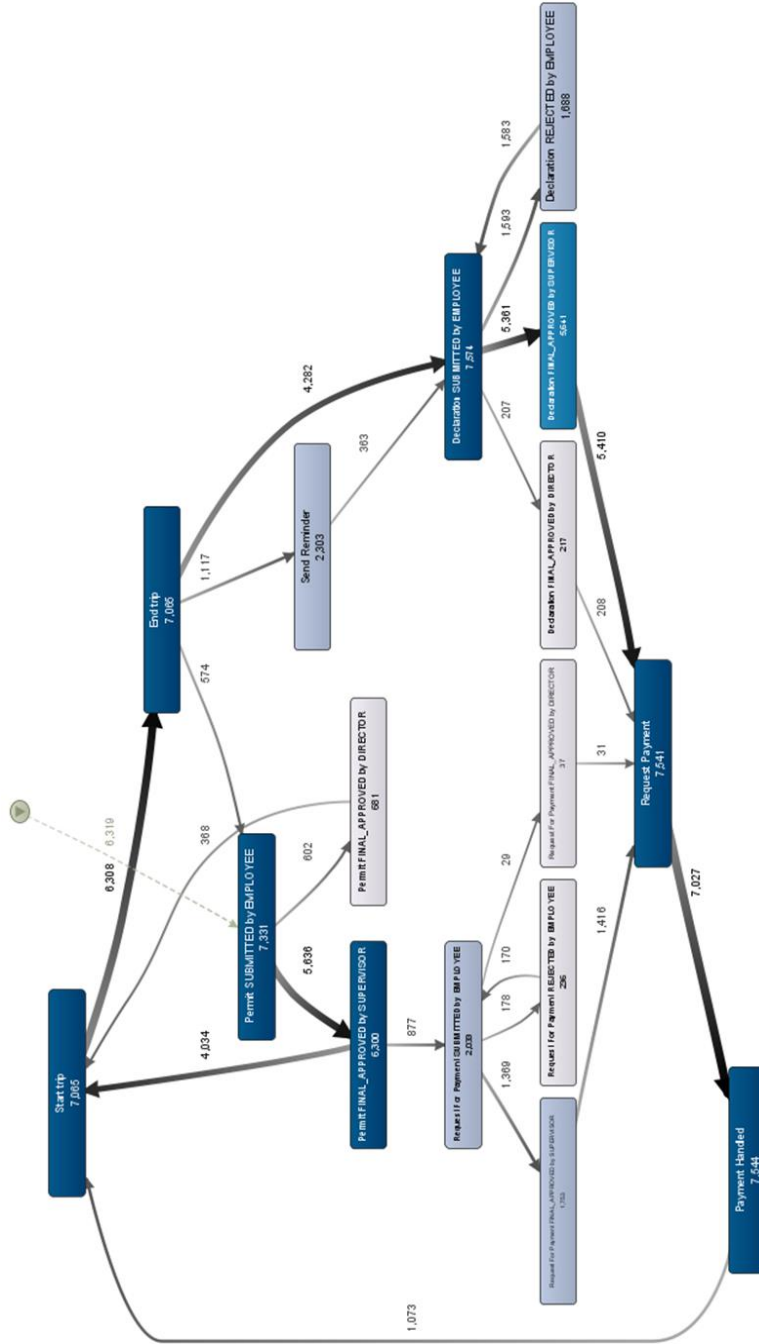


Fig. 1. Cleared process graph of trip permit and declaration