Process Conformance and Performance Analysis: 
TU/e Travel Expense Process 
Business Process Intelligence Challenge 2020

Chiao-Yun Li\textsuperscript{1} and Jingjing Xu\textsuperscript{2}

\textsuperscript{1} Fraunhofer FIT, Schloss Birlinghoven, 53754 Sankt Augustin, Germany 
\texttt{chiao-yun.li@fit.fraunhofer.de}

\textsuperscript{2} Computer Science Department, RWTH Aachen University, 52056 Aachen, Germany 
\texttt{jingjing.xu@fit.fraunhofer.de}

Abstract. In many organizations, sometimes employees need to travel 
for their work. The organizations pay for the expenses of such business 
trips. To ensure the necessity of the trips and that the expenses are 
reasonable, there are processes to control the traveling rates and track 
the expenses. Meanwhile, such processes require expertise to control the 
spending. The necessity for the expertise for such processes also intro-
duces additional expenses for the organizations. To identify the unnec-
essary spending, this report presents the analysis of an administrative 
process for handling business trips at a university in the Netherlands. 
We analyze to what degree the executions of the process comply with 
the policy and the efficiencies of the employees. The results show that 
the noncompliance and inefficiencies are mainly due to the design of the 
process and the interaction or communication between the stakeholders 
in the process.

Key words: BPI Challenge, Process Mining, Conformance Checking, 
Performance Mining

1 Introduction

Business trips are essential for many firms. Meanwhile, the companies need to 
pay for the expenses made by the employees conducting such trips. Although 
such expenses may be seen as an investment, on average, they account for 1% 
of the revenues of the companies \cite{1}. As such, it is necessary for companies 
to have processes for controlling the expenses of business trips. Such processes 
assess the necessity of business trips and expenses, which are still performed 
manually in most firms \cite{2}. The manual operation implies the additional costs 
for the workforce and may introduce faulty decisions that lead to expense fraud. 
Regularly analyzing the executions of these processes enables corporations to 
enhance the compliance of the policy and the efficiencies to deduct the costs or 
identify such fraud.

This paper presents the analysis of a process in Eindhoven University of 
Technology (TU/e) with the aim of identifying the opportunities to deduct the
costs in the process. These costs can be categorized into direct and indirect costs. Direct costs are the travel expenses that the university needs to pay for; indirect/hidden costs are the unnecessary spending due to the inefficiencies in the process. Based on the domain knowledge provided by the process owner, we evaluate the degree of noncompliance and inefficiencies in the process. To locate and understand the causes of the identified issues, we investigate how the process is executed from different perspectives and provide the possible reasons of the issues identified. This analysis report serves as the basis for the process owner to redesign the system or the process with the goal of reducing the costs in the process and the risks of employee fraud.

The travel expenses of the business trips are one of the most difficult operating expenditures for companies to control [3]. One of the root causes is due to the lack of the visibility of the process [4]. Process mining is a technology that enhances the visibility of processes by analyzing the records from the actual executions of business processes. Using process mining techniques, we analyze the administrative process of handling business trips in TU/e. We highlight the most significant findings as follows:

- 5739 international trips are not declared. Among these trips, €1.6 millions are paid to the travellers for the prepaid travel expenses.
- The handover of a request is performed well among the administrative staff; however, for handling different types of requests for an international trip, the staff is not informed properly with regards to the status of the process.
- The noncompliance is mostly due to the illegal submissions of the requests that are reviewed or even paid.
- The batched executions performed by the system and waiting for the submissions of the requests are the root causes of the bottlenecks.

The structure of this paper is as follows. The scope and the methodology of the analysis are introduced below in Section 1.1. Then, we analyze the attributes and explore the behavior of the process in Section 2. The compliance is evaluated in Section 4 with the decisions made in the process analyzed in Section 5. Section 6 presents the performance of the process with the diagnoses of the bottlenecks. Finally, Section 7 concludes the results.

1.1 Scope and Methodology

The scope of the analysis includes the evaluation and diagnosis of the quality and efficiency of the executions of the process. We perform the analysis based on the domain knowledge and the datasets provided. Therefore, given the information, there are limitations of the diagnosis for the noncompliance and the inefficiencies. Nevertheless, we propose the possible reasons and solutions for the issues identified. The exact root causes for certain issues require consultations with the stakeholders or more information that is not provided in the datasets. With such limitations, our methodology is based on the following phases. We describe the motivations and objectives in each phase as follows.
1. **Exploration**: We explore the process by analyzing the attributes and the predominant behavior in the process. Based on the results, we identify the business trips of which the requests are not handled properly and preprocess the datasets for further analysis.

2. **Process Modeling and Enhancement**: We model the process by incorporating the domain knowledge and the results of process discovery. The model serves as a media for process analysts to interpret the results and forms the basis for the following analysis. Therefore, we model the process at different levels of granularity of the process. The business rules provided or identified are integrated into the model.

3. **Conformance Checking**: We evaluate to what degree the executions of the process conform to the refined model and the business rules. The deviations and violations are listed with their financial impact provided.

4. **Decision Analysis**: The expenses that the university needs to pay for are determined by the decision of the necessity of the trips and the requests for reimbursements. We analyze these decisions and their impact on the process from various perspectives.

5. **Performance Analysis**: The efficiency reflects the indirect costs in the process. We evaluate the performance and identify the bottlenecks in the process. Since the process contains many manual executions, we investigate into the working patterns of the administrative staff. In addition, we analyze the workload from different aspects and identify the properties of a trip that may impact the performance of the process. For instance, we check whether the requests with more travel expenses require more time to be reviewed.

In the remainder of the report, we present the results of the analysis in each phase and conclude the most significant findings along with the suggestions to improve the compliance and efficiency of the process.

## 2 Exploration

An execution of an *activity*, i.e., a well-defined process step, is recorded as an *event* executed in the context of a process instance, i.e., *case*. The executions of the administrative process for handling business trips inTU/eare organized into five datasets. Each dataset contains events collected based on an aspect:

- *Domestic Declarations* contains the events related to the declarations of the domestic trips;
- *International Declarations* contains the events related to the applications of travel permits and the declarations of the international trips;
- *Prepaid Travel Cost* contains the events related to the applications of travel permits and the claims for the reimbursement of the prepaid travel costs;
- *Travel Permits* contains all the events related to the international trips;
- *Requests for Payment* contains events that are not related to business trips, e.g., representation costs, hardware purchased for work.
We explore the process based on two aspects, attributes and behavior. In the analysis of the attributes, we identify incomplete cases, e.g., undeclared trips, and the data quality issues in the datasets. The behavioral overview emphasizes the most predominant patterns in the process. Based on the results, we conduct data cleansing for the analysis in the following phases.

2.1 Attribute Analysis

The attributes of a process can be categorized as event attributes and case attributes. Event attributes describe the properties of events and case attributes describe the properties of cases. This section presents the analysis of event and case attributes.

Event Attributes All the events in the five datasets share the same event attributes. Table 1 shows the five attributes with the example of the values. We discover some patterns of the attributes as follows.

- The event identifier is not unique, i.e., an execution can be recorded as several events in the datasets. This is due to the way that the datasets are extracted.
- The activities Send Reminder, Payment Handled, and Request Payment are executed by the system. Other activities are executed by the staff member.
- If an activity is executed by the system, then the role executed the activity is UNDEFINED; otherwise, it is STAFF MEMBER.

Assuming that the activities executed by the system are automatic actions, 73.6% of the executions in the process are performed manually. To improve the process efficiency, the university should look into these manually executed activities and seek for the possibility for further automation. To support the objective, we analyze the patterns and the efficiencies of the executions performed by the staff member and the system in the report.

Case Attributes We classify the case attributes into three categories. By analyzing the correlation attributes, we identify different types of requests of a trip. The analysis of amount attributes quantifies the financial impact of the process to the university. Finally, the administration attributes, e.g., the project and the budget that a trip is based on, show the relationships between the cases and other aspects of the process. We present the analysis of the attributes and exclude ReqPay since it is supposed not to be related to business trips.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Example</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event ID</td>
<td>st_step 64248_0</td>
<td>Event identifier</td>
</tr>
<tr>
<td>Activity</td>
<td>Permit APPROVED by SUPERVISOR</td>
<td>A well-defined process step</td>
</tr>
<tr>
<td>Timestamp</td>
<td>2018.02.09 11:24:08</td>
<td>Timestamp of execution</td>
</tr>
<tr>
<td>Resource</td>
<td>STAFF MEMBER</td>
<td>The resource executed the activity</td>
</tr>
<tr>
<td>Role</td>
<td>SUPERVISOR</td>
<td>Role of the resource in the organization</td>
</tr>
</tbody>
</table>
Correlation Attributes: Every international trip requires a travel permit. Meanwhile, several claims for prepaid travel costs and declarations can be filed independently for one international trip. For convenience, we refer to requests for travel permits, prepaid travel costs, and declarations as “requests” in general. The declarations can be further categorized into international declarations and domestic declarations. Only declarations are necessary for the domestic trips. For other types of requests, they are submitted for the international trips. In this section, we analyze the relationships between different types of requests for the international trips by comparing the correlation attributes. The trip with travel permit 423 may be handled improperly due to the mismatch of the attributes. We exclude the records related to the trip and Fig. 1 presents the results.

As shown in Fig. 1, there are 5739 international trips that no declarations are submitted and 1457 international trips that no requests for prepaid travel expenses are submitted. Meanwhile, most international trips are handled with only one declaration or one request for prepaid travel expenses. In addition, there are 155 requests for prepaid travel costs that are not related to the international trips. They may be the claims for the prepaid travel expenses for domestic trips. However, there are no attributes to verify the hypothesis.

Amount Attributes: Whether the travel expenses are under control can be evaluated based on the expenses and the budget. In the process, the following attributes describe different types of expenses and the budget:

- Requested Budget: Amount of budget for an international trip.
- Total Reimbursed: Total amount reimbursed for an international trip.
- Amount Overspent: Amount over spent for an international trip.
- Prepaid Costs: Amount reimbursed for a request for prepaid travel costs.
- Amount Declared (International): Amount reimbursed for an international declaration.
- Amount Declared (Domestic): Amount reimbursed for a domestic declaration.

In the datasets, the amount overspent may be different from the difference of the amount of the expenses reimbursed and the budget due to the intended manipulation of the data. Assuming the currency is in euro, Fig. 2 shows the
average and total amount of the attributes. Since there are undeclared international trips as shown in Fig. 1, we further analyze these trips and compute the average by excluding zeros in Fig. 2. The results are as follows.

- There is no budget for a domestic trip.
- The expenses for domestic trips are lower than for the international trips.
- On average, the employees budget around €100 more than the actual expenses for international trips.
- Among the undeclared international trips, there are 168 trips for which around €1.6 million is paid nevertheless. The payments are due to the prepaid costs.
- The expenses for prepaid costs for the international trips are higher than the expenses declared on average. This may result from the type of the payments. For example, the flight tickets and the conference fees prepaid by the travellers are usually higher than the expenses declared, e.g., the hotel payments.
- In terms of the total expenses, the expenses declared are higher. It may be due to the fact that some international trips do not have requests for prepaid travel costs submitted.

**Administration Attributes:** Each type of requests is described by a set of attributes that reflect the structure of the organization or the policy of handling the business trips. For example, a request for travel permit should be booked on a project with a budget code recorded in the system. Then, there is a department that handles the request. We analyze the distribution of the attributes and summarize the most significant results:

- Most requests are described by a few values of the attributes. For example, 93% of the requests for prepaid costs are booked on project 503.
- There are many requests that cannot be described by some attributes, i.e., the values of the attributes are **UNKNOWN**.
- Some departments handle both travel permit and prepaid costs. For declarations, there are no departments specified.
- All domestic declarations are made with budget 86566.
- All requests for prepaid costs are based on **Cost Type 0**.

![Fig. 2: Number of international trips per number of claims for a trip.](image-url)
In the analysis of the amount attributes, we find that a domestic trip can be made without a budget. Nevertheless, the analysis of the administration attributes suggests that all the domestic trips are booked to the same budget. This implies that all the expenses of domestic trips are controlled together; however, the regulations are less strict than for the international trips.

2.2 Behavioral Exploration

To obtain a “big picture” of the process, we apply Dotted Chart to explore the most dominant behavior [5, 6]. Fig. 3 shows an example of the explorations for the arrival rate of the requests for travel permits. All the events of a case in Permit are plotted horizontally based on the timeline. The cases are ordered vertically based on the timestamp of the first event, which shows the arrival rate of the first submission of the requests for travel permits. The colors of the dots represent the activities of the events. The activities executed in batch are shown predominately as a vertical line. Fig. 4 demonstrates an example of the batched executions of Send Reminder. We further explore the behavior by changing the configuration of the tool. Below are the findings.

- **Send Reminder** is executed either at around 6 am in the morning or 3:20 pm in the afternoon on the first day of each month.
- **Payment Handled** is executed mostly at around 5:30 pm in the afternoon on every Monday or Thursday.
- The throughput rate increases significantly in the end of 2017.
- The roles of ADMINISTRATION and BUDGET OWNER exist since 2018 whereas PRE_APPROVER only exists in 2017.
- An event of End trip appears in the “future” compared to the release time of the datasets.

The regular patterns of executing Send Reminder and Payment Handled may be due to the configuration of the system, which we assume that the executions are automatic. For the execution in the “future”, provided with the information by the university, it is because the trips are recorded as the estimated dates. Meanwhile, knowing that there is a process change in the end of 2017, the change of the throughput rate and the roles may be due to the change.

Fig. 3: Arrival rate of travel permits.  
Fig. 4: Send Reminder in batch.
Table 2: Overview of the datasets after data cleansing.

<table>
<thead>
<tr>
<th>Datasets</th>
<th>#Cases</th>
<th>#Activities</th>
<th>#Events</th>
<th>Time Span</th>
</tr>
</thead>
<tbody>
<tr>
<td>Permit</td>
<td>5597</td>
<td>15</td>
<td>20014</td>
<td>2018.01.06 11:13:23 - 2019.02.19 08:52:57</td>
</tr>
<tr>
<td>PreReq</td>
<td>1558</td>
<td>13</td>
<td>8799</td>
<td>2018.01.08 08:54:11 - 2019.01.28 17:31:29</td>
</tr>
<tr>
<td>IntTrip</td>
<td>5597</td>
<td>2</td>
<td>11194</td>
<td>2018.01.01 00:00:00 - 2021.09.01 00:00:00</td>
</tr>
<tr>
<td>DomDec</td>
<td>8260</td>
<td>12</td>
<td>46375</td>
<td>2018.01.06 09:41:58 - 2019.06.17 17:30:58</td>
</tr>
<tr>
<td>IntDec</td>
<td>4685</td>
<td>14</td>
<td>29133</td>
<td>2018.01.10 10:13:16 - 2019.08.20 17:31:19</td>
</tr>
<tr>
<td>ReqPay</td>
<td>5778</td>
<td>16</td>
<td>31820</td>
<td>2018.01.06 10:00:19 - 2019.08.08 14:57:18</td>
</tr>
<tr>
<td>SedRem</td>
<td>1190</td>
<td>1</td>
<td>1932</td>
<td>2018.03.01 15:20:05 - 2019.08.01 06:01:01</td>
</tr>
</tbody>
</table>

Such change may involve the change of the system or the roles in the process, e.g., the roles of the individuals are renamed or reassigned. In terms of performance, the productivity increases with the process change.

2.3 Data Preprocessing

The analysis above gives an insight of the issues in data quality. For example, the amount attributes are intended manipulated due to the privacy concern. Some issues may not be handled with data cleansing. For other issues, we process the datasets in the following steps:

1. Filter out cases starting before 2018 to focus on the latest process.
2. Create dataset IntTrip which contains the estimated dates of the trips. This differentiates the estimation and the executions that are actually performed.
3. Create dataset SenRem which contains the events of Send Reminder since it is not directly related to any types of the requests.
4. Filter out cases in PreReq if the requests are not for international trips.
5. Filter out events in the provided datasets such that the datasets contain only events for handling the type of the request.

Table 2 summarizes the statistics of the datasets after preprocessing. There are 5597 international trips since 2018 and 1190 international trips of which the reminders are sent for the declarations.

3 Process Modeling and Enhancement

We consider a business process can be best understood with a model. The general process flow after the process change is described by the process owner. Despite the fact that no normative model is provided, we model the process according to the description and enhance the model with the knowledge discovered from the datasets. We define stages for analysis and easier interpretation of the results. Section 3.1 introduces the stages defined and the process at the stage level, followed by explanation of the core subprocess, i.e., decision making of the requests, in Section 3.2.
3.1 Process at the Stage Level

The business trips at TU/e are distinguished into two types, international and domestic trips. We define four stages to describe the process at a higher granularity. Fig. 5\(^1\) shows the process for international trips. A travel permit is required before any further arrangement for an international trip (Apply stage). The reimbursement of the travel expenses is initiated by the submission of a claim or a declaration. For the prepaid costs, e.g., as the registration fees of conferences or flight tickets, the claims should be filed as soon as the payments are done and before the trip starts (Claim stage). The rest of the expenses that are paid on the spot, e.g., hotel and food, are paid to the traveller after the trip (Travel stage) if the travel declarations are approved (Declare stage). Note that, in both the Claim and the Declare stages, there can be several claims or declarations filed independently for one trip\(^2\). If no declaration is submitted, reminders may be sent (Send Reminder). For all the administrative stages, i.e., Apply, Claim, and Declare, a request may be rejected during the review and the traveller could re-submit it again. Beside the general process flow for international trips described above, below summarizes the business rules declared in the description:

1. The estimated dates for the international trips are recorded in the system when the request for travel permit submitted.
2. Declarations are only mandatory for domestics occurs.
3. The Claim stage starts before the Travel starts \(^3\).
4. International declarations should be filed within \textit{two months} after the estimated end date of a trip.

---

\(^1\) The process is modeled in BPMN [7]
\(^2\) We assume that a case for handling an international trip archives only if all the requests filed for the trip are reimbursed.
\(^3\) The model does not incorporate the rule due to BPMN formalism. Nevertheless, we include the rule for the following analysis.
Additionally, we discover some patterns which we assume to be the business rules to comply with based on the majority of the cases:

5. **Reminders for declarations are sent every two months after the estimated end dates for an international trip.**
6. **Reminders are sent at most twice.**

Each administrative stage is responsible for handling one type of requests. In the next section, we present how the requests are reviewed.

### 3.2 Review Process

The review processes in all the administrative stages follow the control-flow as shown in Fig. 6. The review of a request is performed by three roles of the administrative staff sequentially, administration, budget owner, and supervisor. If the budget owner and supervisor are the same person, only the review by supervisor is recorded in the system. In some cases, a request may require review by a director. If a request is rejected, the employee, i.e., the traveller, terminates the round of the review and resubmits the request again. Otherwise, in the *Apply* stage, the permit is approved and the process continues as in Fig. 5; in the *Claim* and the *Declare* stages, the payment is requested and handled by the system automatically. For the *Declare* stage, there are specific rules to handle the declarations at TU/e: 1. a declaration should be automatically rerouted to the responsible supervisor if the budget owner does not handle it within 7 days; 2. a declaration should be rejected by the administration if it is the first declaration of the trip filed 2 months after the trip ends.

![Fig. 6: The review process of the stages: Apply, Claim, and Declare.](image-url)
4 Process Conformance

Section 3 gives an overview of the control-flow and the business rules of the normative process. To evaluate to what degree the executions of the process comply with the process defined, we assess the process conformance in this section. Consider the review process in all the three administrative stages, modeling the process at the activity level poses a challenge for interpreting the results. Thus, conformance checking is conducted at two levels. First, we evaluate how the requests are being handled in each administrative stage. Then, we analyze the behavior of the process at the stage level by verifying if the process is executed as described in Section 3.1.

4.1 Stage Conformance

Understanding how the requests are handled is crucial for preventing faulty executions which may lead to the expense reimbursement fraud or unnecessary spending. Therefore, for every administration stage, we evaluate the conformance of the process of handling the corresponding type of requests. Table 3 summarizes the deviations or the violations with the total expenses of the corresponding requests, which indicate the financial impact of the violations to the business.

In Table 3, the conformance of handling the permits is high. We assume that the requests saved are not submitted for the review. The most frequent violation or deviation is due to the trips that are planned but the requests for permits are never submitted. For these trips, there are no payments made. Therefore, the deviations do not cause financial loss for TU/e. However, for the requests

Table 3: Noncompliance of the administrative stages of handling the requests.

<table>
<thead>
<tr>
<th>Violation or Deviated Pattern</th>
<th>#Requests</th>
<th>Total Payments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Permit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Only • SAVED by EMPLOYEE executed</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>• FINAL_APPROVED by SUPERVISOR</td>
<td>3</td>
<td>70.98</td>
</tr>
<tr>
<td>• APPROVED by SUPERVISOR</td>
<td>2</td>
<td>1014.42</td>
</tr>
<tr>
<td>• REJECTED by EMPLOYEE not executed</td>
<td>2</td>
<td>5812.49</td>
</tr>
<tr>
<td>• SUBMITTED by EMPLOYEE → • REJECTED by EMPLOYEE</td>
<td>1</td>
<td>5136.72</td>
</tr>
<tr>
<td>• APPROVED by EMPLOYEE → • SAVED by EMPLOYEE</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>• APPROVED by BUDGET OWNER → • SAVED by SUPERVISOR</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>• SUBMITTED by EMPLOYEE → • FOR_APPROVAL by EMPLOYEE</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Request for Prepaid Travel Costs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Only • SAVED by EMPLOYEE executed</td>
<td>12</td>
<td>10977.87</td>
</tr>
<tr>
<td>• REJECTED by EMPLOYEE not executed</td>
<td>6</td>
<td>3011.28</td>
</tr>
<tr>
<td>• REJECTED by EMPLOYEE → • SAVED by EMPLOYEE</td>
<td>2</td>
<td>1401.50</td>
</tr>
<tr>
<td>• SUBMITTED by EMPLOYEE → • REJECTED by EMPLOYEE</td>
<td>1</td>
<td>89.04</td>
</tr>
<tr>
<td>Declarations (International and Domestic)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• APPROVED by SUPERVISOR → • APPROVED by BUDGET OWNER exceeds 7 days</td>
<td>320</td>
<td>176416.46</td>
</tr>
<tr>
<td>• REJECTED by EMPLOYEE not executed</td>
<td>39</td>
<td>19477</td>
</tr>
<tr>
<td>• SUBMITTED by EMPLOYEE → • REJECTED by EMPLOYEE</td>
<td>11</td>
<td>5503.72</td>
</tr>
<tr>
<td>Request Payment not executed</td>
<td>11</td>
<td>2294.72</td>
</tr>
<tr>
<td>• APPROVED by BUDGET OWNER → • REJECTED by SUPERVISOR exceeds 7 days</td>
<td>7</td>
<td>1322.09</td>
</tr>
<tr>
<td>Payment Handled → Request Payment</td>
<td>1</td>
<td>3084.26</td>
</tr>
<tr>
<td>• SUBMITTED by EMPLOYEE → • FOR_APPROVAL by SUPERVISOR</td>
<td>1</td>
<td>349.38</td>
</tr>
<tr>
<td>• REJECTED by EMPLOYEE → • SAVED by EMPLOYEE</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

* denotes the type of request.
for prepaid travel costs, similar type of violations cause nearly €11k paid for the 12 requests. Another type of violations is the cases that are rejected by the supervisors give the final approval. Nevertheless, the expenses are not significant. For the other deviations and violations, we assume they do not impact the business much. For example, since both the supervisors and the directors approve the requests (the third row for the violation of handling permits), such violations are assumed to be acceptable.

There are more violations and deviations when handling the declarations. We assume that the automatic transfer of the declarations from the budget owners to the supervisors should be conducted at the 7th day after the budget owners get the declarations. Nevertheless, the most frequent violations of the rules indicates that the rule is not implemented in the system. For other violations and deviations, most of them are due to the travellers or the system. We assume that they do not have significant financial impact for TU/e.

Most deviations are due to the requests that are not submitted or occur after the review owing to the travellers or the system. In terms of the administrative staff, most handovers of the requests between different administrative roles conform to the process. The violations are mostly due to the late handovers of the declarations and the payments for the requests that are not submitted. Meanwhile, some deviations may be acceptable for the business. For example, there are 11 declarations that Request Payment is not executed; however, all the payments are made based on the approved declarations. Which of these deviations are critical requires more information for further diagnosis.

4.2 Conformance at the Stage Level

To identify the causes of additional payments and the trips that should not be made, we check the conformance of the process for international trips. For the domestic trips, only the Declare stage is mandatory and, thus, evaluated in Section 4.1. We assume that the requests that are saved by the employees are not submitted to the administrative staff. We examine whether the process conforms to the defined control-flow and other business rules described in Section 3.1. Table 4 presents the results.

The violations 1 to 4 are due to the wrong timing of the submissions of the requests. According to Section 3.1, the requests for prepaid travel costs should be filed after the travel permits are issued and before the trips start. Based on the domain knowledge, the declarations of the trips could only be filed after the trips are made, i.e., after requests for prepaid travel costs submitted. The causes of the deviations may be that the status of the trips are not well-informed between the departments of handling different types of requests. Or, according to violation 1, the requests for the prepaid costs may be mistakenly submitted as declarations. Alternatively, it could also be that the declarations are submitted as requests for prepaid costs, which may be expense fraud since the review may be less strict due to the fact that the verification of a trip is not required for the requests for prepaid costs.
The violation 5 may be owing to the configuration of the system. For the violation 6, we further investigate into the international trips of which the expenses are overspent and the same type of the requests are filed several times independently. If the budget is reasonable, i.e., the actual spending does not exceed the budget claimed, the payments made by the university may be more than the actual travel expenses. In this case, such mistake can be due to the fact that the information is not well updated across the departments such that some departments may issue payments for the expenses that have already been paid.

Finally, since the dates of the international trips are not the exact travel dates and we have no indicators for whether the trips are actually made, some of the violations may not represent the actual non-conformance in the process. For example, for violation 1., it may be that no reminders are sent since the trips did not occur. The violation 7 to 12 require the actual travel dates to confirm whether they are the actual non-compliance that causes additional spending.

4.3 Conformance Summary

Most of the non-conformance when handling a request occurs when the travellers are involved in the process and the activities conducted by the system. Depending on the view point, if the travellers are seen as external actors, then some violations may be inevitable. For instance, the university cannot control if the travellers properly reject the requests before resubmission. Nevertheless, safeguards can be implemented by rejecting the illegal submissions or reviewing them after travel permits are issued. From this perspective, the process should prohibit the administrative staff from reviewing the illegal submissions or even paying for the requests.

For the handling of different types of requests, excluding the uncertainty due to the estimated dates for the trips, the violations are mostly due to miscommunication or the information that are not updated timely between departments. Potential improvement may be automating the updates of the status of the trips.

Table 4: Noncompliance of the process for international trips at the stage level.

<table>
<thead>
<tr>
<th>Violation or Deviated Pattern</th>
<th>#Trips</th>
<th>Total Reimbursed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Declarations submitted before requests for prepaid costs submitted</td>
<td>50</td>
<td>76354.68</td>
</tr>
<tr>
<td>2. Requests for prepaid costs submitted before final decision for permits</td>
<td>175*</td>
<td>253206.98</td>
</tr>
<tr>
<td>3. Declarations submitted without permits</td>
<td>1</td>
<td>70.99</td>
</tr>
<tr>
<td>4. Payments made before final approval for permits</td>
<td>4</td>
<td>3739.43</td>
</tr>
<tr>
<td>5. Declaration submitted after due date and no reminders are sent</td>
<td>14</td>
<td>7715.41</td>
</tr>
<tr>
<td>6. More than one request for prepaid costs or declaration filed and paid</td>
<td>171</td>
<td>363724.11</td>
</tr>
<tr>
<td>and the budget is overspent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. No declarations submitted for trips approved and no reminders sent</td>
<td>335</td>
<td>68421.11</td>
</tr>
<tr>
<td>8. No declarations submitted for trips approved with reminders sent</td>
<td>833</td>
<td>1534299.22</td>
</tr>
<tr>
<td>9. Requests for prepaid costs submitted after international trip starts</td>
<td>81</td>
<td>100439.81</td>
</tr>
<tr>
<td>10. International trips start before permits are submitted or reviewed</td>
<td>1046</td>
<td>569202.05</td>
</tr>
<tr>
<td>11. Declarations submitted or saved before international trips end</td>
<td>289</td>
<td>399789.98</td>
</tr>
<tr>
<td>12. Late submissions of declarations and approved by administration</td>
<td>92</td>
<td>74759.11</td>
</tr>
</tbody>
</table>

* 1 trip that the travel permit is eventually rejected

Total: 3311722.88
5 Decision Analysis

The decisions of the administrative staff directly determine the expenses that the university should pay for. Therefore, we analyze the decisions from three aspects: per role, per submission, and per request. We only consider the requests that are related to business trips. Fig. 7 shows the frequency of decisions per role. The administration role has the highest rejection rate of handling declarations. One of the reasons may be that the administration serves as the major safeguard for the submissions that are not conforming to the rules. For example, it is specified that the administration staff should reject the first declaration of an international trip if it is submitted two months after the trip.

For each submission, we show the frequency of approval and rejection based on the final decision of the administrative staff in Fig. 8. After a request is rejected, the traveller should correct the request and resubmit it again in order to get the travel permit or the reimbursement. Assume that a resubmitted request indicates that the request is corrected. Compared to the frequency of the final decisions of the requests in Fig. 9, the rejection rates of declarations greatly decreases with the correction. There are in total 2026 declarations that have been corrected, accounting for 16% of the declarations.

Finally, we investigate the underlying causes for the requests that are never approved. We diagnose the root causes based on their final status, which implies the reasons of the rejection. We assume that each administration role is responsible for one aspect of a request, e.g., an administration role verifies the administrative information, and the budget owner confirms whether the budget is reasonable. Therefore, if the last status of a request is rejected by, for example, a budget owner, then the root cause of the rejection results from the budget or the spending of the trip. Meanwhile, there are some requests that are never approved since they are never submitted.

Fig. 10 presents the frequency of the requests that are rejected in the end based on the possible root causes. For the reason "other", some requests are rejected by an employee without any review of the administrative staff or are
reviewed by other role such as \textsc{for.approver}. For the requests that are never approved, excluding the requests that are never submitted, these cases remain open cases based on our assumption, i.e., the administrative staff awaits the requests that are corrected an As shown in Fig. 10, \textit{most rejections are due to the errors in the administrative aspect of the requests}.

6 Process Performance

Companies lose 20\% to 30\% of their revenues every year due to process inefficiencies [8]. To identify opportunity to deduct such hidden costs in the process, we analyze the efficiencies based on the type of requests and provide an overview of the performance over stages. Since the efficiency can be related to the workload, we further analyze the amount of the tasks or requests based on various perspectives. Moreover, we identify the correlation between the performance and other aspects of the process to identify the proprieties of the process which can be utilized for better resource planning. In the following analysis, we assume that the \textit{timestamp} refers to the completion time of the executions. Therefore, only the \textit{sojourn time}, i.e., the duration from the completion of an activity to the completion of the following activity, can be used for analyzing the performance.

Fig. 10: Number of requests never approved by supervisor or director per cause.
6.1 Performance of Handling Requests

The throughput time of handling a request starts from the submission until no more executions performed for the request. The final status of a request can be a final approval, payment, or rejection (we exclude the requests that are never submitted). At the same time, once a request is submitted, there are three statuses of a request in the process: under review, waiting for resubmission, and waiting for payments after approval. The administrative staff, travellers, and the system are responsible for one of the status, respectively. We define the following performance indicators based on the possible statuses of a request.

- **Review Time**: The duration of a request being under review.
- **System Waiting Time**: The waiting time from a request being approved until the payment executed by the system. The system waiting time for handling requests for permits is zero.
- **Resubmission Waiting Time**: The waiting time for the resubmission of a rejected request.

The mean and median duration of handling each type of requests are presented in Fig. 11. The main bottleneck results from waiting for the executions by the system (yellow). Meanwhile, since reviewing requests indicates the indirect costs required for the university, we further investigate into the activities based on review time and the main bottleneck, i.e., system waiting time.

**System Waiting Time Analysis.** We assume that the activities performed by the system are instant actions, i.e., the start and completion of executing an activity is represented by an event. The sojourn time of Request Payment or Payment Handled is the waiting time to execute the activities. Fig. 12 shows the mean and median waiting time for which the types of the requests that payments are made. There is no significant difference of the bottleneck among different types of requests. For the Payment Handled, it is due to the fact that it is executed in batch on every Monday and Thursday. For the Request Payment, we do not find significant batched executions and the reasons of the delay cannot
be inferred. Assuming no specific reasons for the delay, the efficiency may be improved by configuring the system for such automatically executed activities. For example, instead of executing **Payment Handled** in batch, the payment can be made after a specific duration after the final approval of a request.

**Review Time Analysis.** The review of a request is conducted by administrative staff with a specific order of the roles. Since only sojourn time is available in the datasets, we assume the review time of an administrative employee is the sojourn time of the activities with the decision. Fig. 13 shows the statistics of the review time of the roles for each type of the requests. To fairly compare the performance and prioritize the issues, we also consider the frequency of reviews and the total review time of the roles. We summarize the observations as follows.

1. Comparing the average and median review time, the review can be performed within three days for each role. The bottleneck results from some requests that take much more time to review.
2. The administration role requires more time to review the declarations compared to requests for travel permits or prepaid travel costs on average.
3. For every role, international declarations require more time to review compared to the domestic declarations; however, since more domestic trips are made, the total review time spent on domestic declarations is more.
4. Considering the frequency of review of roles, despite the high frequency of review by the administration role, the review by the administration does not take much time in total.
5. Even though the average review time is the most for directors, such requests are infrequent and do not impact much on the total review time.
6. The role of supervisor takes the most time to review requests in total.
7. Despite the highest frequency of reviews by the administration role, the administration takes the least time to review on average.

Based on the analysis, we propose a potential solution. Assuming the control-flow of the review may not be changed, further improvement may be implemented by adjusting the distribution of resources for each role based on the type of the requests. For example, comparing the performance of the supervisor and the administration, the number of the reviews are high for both roles. It implies that more staff with the role of supervisor may be allocated for reviewing declarations.

---

**Fig. 12:** Waiting time for executing **Request Payment** and **Payment Handled**.
However, such improvement should consider the fact that some staff have the roles of both budget owner and supervisor, of which the reviews are recorded only as the supervisor in the system.

6.2 Performance over Stages

To identify other inefficiencies in other aspects of the process, we present the performance results of the process at the stage level. We present the performance of the most critical path between the stages as shown in Fig. 14. In addition, for the administrative stages, we compute the number of submissions per stage and the performance metrics, of which the ratio is shown with the colors in the figure. Note that we include the statistic for handling domestic declarations in the Declare stage to provide an overview of the process performance. Meanwhile, for international trips, the average and median throughput time of the entire process is 84.2 and 71.1 days, respectively.

The requests for prepaid travel costs should be submitted after the decisions for the travel permits. As shown in Fig. 13, it takes around 16 days on average for the requests for prepaid costs to be submitted after the travel permits issued. The severity of the delay and the noncompliance of the submissions of the requests for prepaid travel costs are up to 7 months to 1 month, respectively.

For the duration related to the Travel stage, it is based on the estimated dates for the international trips. Therefore, the statistics should be adjusted if the exact dates of the trips are provided. We assume that the estimated dates are close to the actual traveling dates. The travel permits are handled roughly 1 months before the trips and the declarations are submitted within two weeks.

Fig. 13: Statistics of review time of the roles per type of requests.
on average. For the level of noncompliance, without considering the logging errors, the negative values indicate that some travellers apply for travel permits afterwards or that a trip ends earlier than the estimation. In terms of the inefficiencies, there is at most one year of delay for both the submission of an international declaration and the occurrence of a trip.

Based on the process flow and the domain knowledge, the reminders are sent for the international declarations. From the statistics between the Travel stage, the Declare stage, and the Send Reminder, it indicates that the reminders are sent based on the estimated dates for the trips. Knowing that the reminders are sent monthly for the trips of which the declarations are not submitted within two months, there are reminders that are not sent at the scheduled time. This may induce the delay of the submissions of the declarations. Meanwhile, we also identify that some reminders are sent after the payments for the declarations. It is most likely due to the fact that the executions of Send Reminder are based on the estimated dates for the trips and the status of the declarations are not updated in the system.

In conclusion, the inefficiencies are due to the waiting time between the stages. The duration of a trip mostly depends on the business instead of the executions of a process. Therefore, to minimize the overall throughput time of a case, one possible solution is to change the configuration of the system. For instance, the reminders may be sent earlier to remind the travellers for the coming due dates. Meanwhile, the missing reminders may be due to the manual logging of the status of the trip, which may be improved by automation.

6.3 Workload Analysis

The efficiency of a person performing a task is related to the workload distributed to the person [9]. To justify the efficiency of the process, we analyze the workload based on the request types, the roles, and the departments. We assume that the efficiency of the tasks performed by the system and the travellers is not affected by the workload.

To compute the workload, seeing travellers as external actors, the workload of each role is computed based on each submission. Each submission is seen as a
Table 5: Workload metrics based on different perspectives.

<table>
<thead>
<tr>
<th>Request Type</th>
<th>Arrival Rate</th>
<th>Departure Rate</th>
<th>Time under Review</th>
</tr>
</thead>
<tbody>
<tr>
<td>Travel permit</td>
<td>16.09/17.0</td>
<td>14.16/11.0</td>
<td>3.94/2.94 days</td>
</tr>
<tr>
<td>Request for prepaid travel costs</td>
<td>4.93/4.0</td>
<td>4.70/3.0</td>
<td>3.37/1.37 days</td>
</tr>
<tr>
<td>Declaration</td>
<td>42.32/42.0</td>
<td>25.79/5.0</td>
<td>4.52/1.52 days</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Administrative Role</th>
<th>Arrival Rate</th>
<th>Departure Rate</th>
<th>Time under Review</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADMINISTRATION</td>
<td>63.20/69.0</td>
<td>38.39/14.0</td>
<td>0.49/4.62 days</td>
</tr>
<tr>
<td>BUDGET OWNER</td>
<td>18.28/19.0</td>
<td>17.79/15.5</td>
<td>2.01/0.92 days</td>
</tr>
<tr>
<td>SUPERVISOR</td>
<td>33.53/9.0</td>
<td>33.64/4.0</td>
<td>2.51/1.16 days</td>
</tr>
<tr>
<td>DIRECTOR</td>
<td>1.75/1.0</td>
<td>1.72/1.0</td>
<td>3.39/1.12 days</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Department (Top 5 Frequent Organizational Entity)</th>
<th>Arrival Rate</th>
<th>Departure Rate</th>
<th>Time under Review</th>
</tr>
</thead>
<tbody>
<tr>
<td>organizational unit 65456</td>
<td>9.51/7.0</td>
<td>9.48/6.0</td>
<td>1.53/0.00 days</td>
</tr>
<tr>
<td>organizational unit 65458</td>
<td>8.78/7.0</td>
<td>8.78/7.0</td>
<td>0.82/0.82 days</td>
</tr>
<tr>
<td>organizational unit 65454</td>
<td>5.91/4.0</td>
<td>5.79/4.0</td>
<td>0.71/6.94 days</td>
</tr>
<tr>
<td>organizational unit 65459</td>
<td>5.56/3.0</td>
<td>5.49/4.0</td>
<td>1.44/0.00 days</td>
</tr>
<tr>
<td>organizational unit 65455</td>
<td>5.79/3.0</td>
<td>5.60/4.0</td>
<td>0.66/4.62 days</td>
</tr>
</tbody>
</table>

new case for the system to process, i.e., request type, role, and department. Since here is no attribute describing the departments for handling the declarations, we assume that the declarations are handled by the same departments handling the permits of the trips. The performance indicators for the workload of a system is defined as follows.

- **Arrival Rate**: The mean number of submissions arriving at the system per day.
- **Departure Rate**: The mean number of submissions leaving the system per day.
- **Time under Review**: The mean review time in the system.

Table 5 presents the metrics of each system. No matter in which perspective, the departure rate is larger than arrival rate for most cases, indicating the congestion in the system, i.e., the speed of handling the submissions cannot be kept up with the speed of the incoming submissions. Such congestion is relatively significant for the declarations and the administration role. However, despite the congestion for the administration role, the role requires relatively less time to review a submission. To locate the critical inefficiencies for process optimization, the financial cost should be included. For example, although the role of the administration has heavier workload, the costs for the staff may be less than the staff as other roles per day. Which of these causes the dominant financial loss requires further information for quantification the impact of the inefficiencies.

Additionally, we analyze the working patterns of the staff for improving the scheduling of the work. Fig. 15 present the number of executions of the roles throughput a day, week, and year. We summarize the observations as follows.

1. Most work is performed from 8 in the morning until 5 to 6 in the evening, which are the common working hours.
2. There exist breaks during lunch time and the working efficiency decreases in the afternoon for most roles compared to the morning.
3. The budget owners work from 7 in the morning to latest 10 in the evening. For other administrative roles, there exist activities throughout the nights.
4. Most work are performed on Monday during the week while some are on weekends.
5. The amount of work decreases after June in Summer.

![Fig. 15: Number of executions per hour, weekday, and month.](image)

To summarize, we analyze the process from various perspectives including trips, stage, request, and submissions of requests. The results show that waiting for the submission of a request is the main cause of the bottlenecks.

6.4 Deep Dive Analysis

To identify the other reasons causing the unnecessary costs, we analyze the relationships between different aspects: the amount attributes representing the direct costs; the review time implying the hidden costs, and the administration attributes that may determine the review time and the payments for a trip.

For an international trip, the process is initiated with the submission for a request for a travel permit. We assume that the attributes for handling a travel permit describe the property of an international trip. Fig. 16 shows the degree of correlation of the administration attributes, the amount related to a international trip, and the total review time of each type of request and the roles of an international trip. The darker the color is, the more related the two attributes are. Based on Fig. 16, the review time of a request and the amounts of the budget and the travel expenses depend on the project and the budget that an international trip is based on. This means that the requests related to some specific projects or budgets may cause more spending and longer review time. Such requests can be prioritized when the travel permits are submitted. Meanwhile, the requested budget is, surprisingly, less relevant to the review time for a request for travel permit; however, more related to the review time of the Claim stage.
Meanwhile, we further analyze the correlation between the attributes for the process handling requests for prepaid travel costs and declarations. If the administration attributes do not exist in the datasets to describe the process of handling the request, we adopt the attributes from handling the travel permit of the trips that the requests are made for. For instance, without a project number for an international declaration, we assume that the declaration is based on the project that the trip is booked on (based on travel permit). Fig. 17 shows the correlation between the administrative attribute, the amount paid for the request, and the review time of the request. Note that since there is only one amount attribute for each type of the requests, we combine the amount attributes into the plots.

Based on Fig. 17, the review time of domestic trips does not depend on the travel expenses. For international declarations, the project that a trip is booked on is correlated to the amount declared and the review time, especially the review time of the budget owner. For the requests for prepaid travel costs, the review time spent on the requests depend on the tasks and independent of the amount claimed. As shown in Fig. 2, the average amount of payment for a request for prepaid travel costs is higher than for a declaration. However, when reviewing a declaration, the amount declared is taken in the consideration consideration while it is not the case when reviewing a request for prepaid travel costs. Along with the violations for the payments for the prepaid travel costs observed in Table 3, the requests for prepaid travel costs are not properly handled and cause much financial loss.

7 Conclusion

We model the process based on the description in the challenge and identify the relationships among multiple datasets given. The analysis of the attributes shows the data quality issues and the challenges remained in the following analysis despite data cleansing.

For the quality of the process, the order of reviewing a request is highly conformed to the model. The main non-conformance results from the timing that
Fig. 17: The level of correlation of the attributes for the requests for prepaid travel costs and the declarations. (Review Time as RT)

a request is being handled. For instance, there should not be requests for travel expenses submitted nor reviewed before travel permits are issued. Such non-conformance or non-compliance leaves the room for improvement in the quality of the process. In addition, the decisions made in the process are evaluated based on the type of the requests and the roles involved. Most of the problems occur at the administration perspective of the requests. However, most of them are corrected and resubmitted, resulting in the decrease of rejection rate after several rounds of review.

For the process performance, we analyze the process based on the duration of executions at different levels of the process. It shows that the time for waiting for submission and the executions by system account for the most time in the whole process. Possible performance may be changing the configuration of the system or sending more reminders to the employees (seeing employees as external actor in the process). Since most activities are executed manually in the process, we further analyze the workload and discover the working patterns of the administrative staff. The results show the performance is the best in the morning, on Monday, and in June.

The analysis is limited to the data quality. For example, the efficiency should be analyzed together with the number of staff. However, due to the data privacy, such information is unavailable in the datasets. Similarly, some of the diagnosis of the deviations and inefficiencies is based on the assumption of the circumstances. An example would be that we assume that the wrong order of the execution of Request Payment and Payment Handled is acceptable for the business since they are made based on approved requests. Therefore, the validity of the results and the analysis may be enhanced with more information.

References

Appendix

Frequency of *administration* attributes in the format of *dataset:attribute*

<table>
<thead>
<tr>
<th>Mismatch Pattern</th>
<th>Possible Cause</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. 1457 cases in Permit have no attributes for declarations</td>
<td>No declarations submitted</td>
</tr>
<tr>
<td>2. 5740 cases in Permit have no attributes for claims for prepaid travel costs</td>
<td>No claims for prepaid costs submitted</td>
</tr>
<tr>
<td>3. 155 cases in PreReq have Permit id as UNKNWON</td>
<td>Not requests for international trips</td>
</tr>
<tr>
<td>4. 85 cases in PreReq for travel permit 423 have no attributes in Permit</td>
<td>Process change</td>
</tr>
<tr>
<td>5. 48 cases in IntDec for travel permit 423 have no attributes in Permit</td>
<td>Process change</td>
</tr>
<tr>
<td>6. 168 cases in Permit have more events for Send Reminder, Start trip, and End trip recorded than the corresponding cases in IntDec</td>
<td>Incomplete records</td>
</tr>
</tbody>
</table>

Possible causes of the data quality issues.

<table>
<thead>
<tr>
<th>Possible Cause</th>
<th>Resulting Issues in Data Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process Change</td>
<td><strong>Change of behavior:</strong> The throughput rate increases after the process change.</td>
</tr>
<tr>
<td></td>
<td><strong>Change of roles:</strong> There is difference in the roles before and after 2018.</td>
</tr>
<tr>
<td></td>
<td><strong>Incomplete records:</strong> The records for travel permit 423 may be logged incorrectly.</td>
</tr>
<tr>
<td>Privacy</td>
<td><strong>Manipulated data:</strong> The amount attributes are intentionally manipulated.</td>
</tr>
<tr>
<td></td>
<td><strong>Unknown actors:</strong> The identity of the actor is unknown in the datasets.</td>
</tr>
<tr>
<td>Business Policy or Configuration</td>
<td><strong>Lack of records for the start of the execution:</strong> The timestamps of the start of the reviews are not recorded.</td>
</tr>
<tr>
<td></td>
<td><strong>Unknown records:</strong> Some requests may not be described by the attributes.</td>
</tr>
<tr>
<td></td>
<td><strong>Incomplete records:</strong> Executions of SYSTEM are sometimes missing in some cases.</td>
</tr>
</tbody>
</table>