# Business Process Intelligence Challenge 2019: Process discovery and deviation analysis of purchase order handling process

Jongchan Kim, Jonghyeon Ko, and Suhwan Lee

Ulsan National Institute of Science and Technology, 50 UNIST-gil, Ulsan 44919, Republic of Korea
{jckim,whd1gus2,ghksdl6025}@unist.ac.kr

**Abstract.** Process mining is of paramount importance as the number of event logs cumulated in information systems is growing exponentially over time. Increased number of event logs implies that the amount of information that can be mined to be used to meet the stakeholders requirements through the generation of qualitative and quantitative outputs enhances. In compliance with the goal of the Business Process Intelligence Challenge to make the result of the analysis useful in the real world, we take a real-life event log of a company operating in the Netherlands in order to analyze and suggest improvements to the companys purchase order handling process for its subsidiaries. In this paper, we discover and refine business process models that encompass the four types of flows in the event log. Then, we also identify different types of deviation in the event log from the discovered processes and their extents based on alignment techniques and various statistical tests.

**Keywords:** Business Process Intelligence · Process mining · Business Process Model · Deviation analysis.

## 1 Introduction

Process mining has gained increasing traction in industry as a set of efficient techniques analyzing event logs to enhance business processes. Likewise, it is important to provide practical implications through the application of process mining techniques, which also aligns with the goal of Business Process Intelligence Challenge 2019 [1]. We were provided with the real-world problems that a company is facing with, and thereby, we strived to provide fruitful answers and explanations in detail for each question provided. Following the guidelines provided in the Process Mining Manifesto [2], our analysis concerns different phases of the BPM lifecycle [3].

**1. Process identification** : Before entering the deep-dive analysis step, the first and basic step is to acquire the knowledge drawn from the given event log, including where the event log is generated and clarifying what we want

to discover with the event log. In this case, the event log is expected to be generated from the procurement system of a painting company.

**2. Process discovery** : The second step is to extract a business process model from an event log using process mining tools (e.g. *petri net*) based on identified process. It is essential to filter out deviations from the dataset and keep regular process flows unstained based on the obtained business process model reasonably fitted to the event log.

**3. Process analysis** : For the third step, discovered process is analyzed in different levels such as activity, timestamp, and other features. We, playing the role of consultants, diagnose the current situation of the system driven from the business process model and offer the reasons of deviations explaining where and when the discrepancy between deviants and regular model appeared in detail.

Following the three phases as in above, we mainly focus on 1) discovering the process model, and 2) analyzing the deviations in duration (between activities/instances), activities and other features as to answer for three questions provided in BPIC 2019.

The rest of the report is designed as follows. Section 2 provides comprehensive analysis of the event log. In section 3, 4, and 5 we focus on answering the three questions. Specifically in Section 3, a BPMN model that is constructed to encompass various different models in the process is described. In Section 4, throughput of invoicing processes are calculated, such as the duration between different activities or instances. Moreover, deviations in durations are identified and various factors affecting these deviations are also discussed. In Section 5, various types of deviations are analyzed using an alignment technique, and multiple statistical tests. Finally in Section 6, the overall summary of the analysis along with multiple minor and major findings are presented.

## 2 Overall Analysis

### 2.1 Understanding of the background of the company

Prior to analyzing an event log, it is extremely important to understand the general situation and the process of the given data by gathering information and integrating background knowledge. The given event log contains information about how the company and its subsidiaries handle purchase orders with vendors. General flows are best characterized by activities which handle purchase order, goods receipt and invoice, and can be described as follows. Firstly, one of the subsidiaries creates purchase order document and send it to vendors. After the ordered items are delivered, the company records goods receipt to confirm whether the amount and types of items are correctly delivered according to what had been ordered. For the last step, the company registers the invoice and clears it.

### 2.2 Data selection and preprocessing

We were provided with two files, where the first file is in xes format (*log_IEEE.xes*) and the other file is in csv format (*BPI_Challenge_2019.csv*). The xes file is an original event log and the csv file is a converted version of the xes file. Between these two files, however, it was revealed that there is a difference in timestamps. Although the number of cases, events, and activities are identical in both files, mean and median values of case duration are not identical as the starting and ending time of the event log are not equal. The starting and ending time of the xes file are 27.01.1948 07:59:00, 10.04.2020 06:59:00, respectively, whereas in the csv file, the starting and ending time are 26.01.1948 23:59:00 and 09.04.2020 23:59:00, respectively. Since not only the factors affecting the difference in the timestamp between two files cannot be identified but also the difference in timestamps was not immense, we decided to use the xes file.

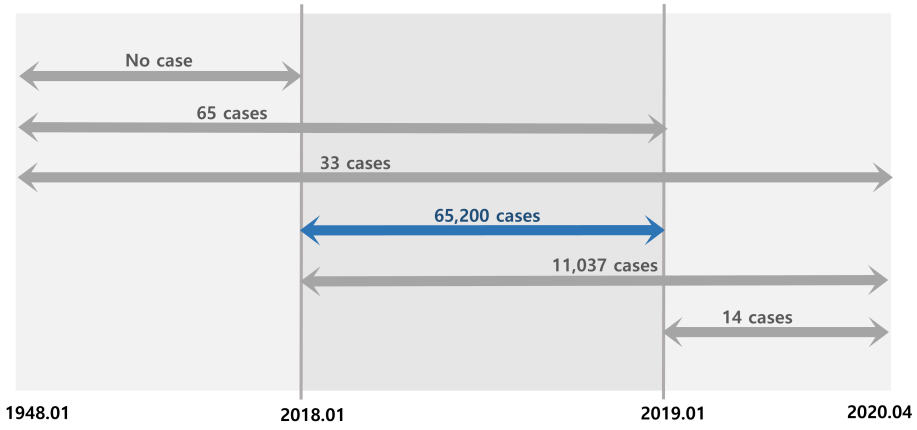



**Fig. 1.** Range of the timestamp in xes file

As mentioned earlier, the dataset has a wide range of timestamp, dating from Jan 27, 1948 to Apr 10, 2020, while the data should have been collected from events that happened in 2018 only. We figured out that the dataset can be classified into three types as in Fig. 1. The first type both starts and ends in 2018, the second type either starts or ends in 2018, and the third type neither starts nor ends in 2018. The distribution of timestamp shows that approximately 87% of all cases both started and ended in 2018, which belong to the first type. Hence, we argue that the second and the third type of the dataset are erroneous as these are out of the scope of the analysis. One of the reason the erroneous cases are introduced can be a system error as these cases are detected with an activity named, SRM: Transfer Failed (E. Sys.). Therefore, we regard these cases as out-

liers and we do not further consider these. To sum up, the filtered data contains 65,200 cases with 1,315,795 events and 41 activities.

### 2.3 Data exploration

Table 1 provides the description of variables of the dataset that are used in the analysis. As it describes, *case concept:name* corresponds to *Case ID*, *event concept:name* corresponds to *Activity*, *event time:timestamp* corresponds to *Timestamp*, and *event User* (or *event org:resource*) corresponds to *Resource*. Since the name of each variable is quite long, we changed the name of each variable by deleting the words, "case" and "event", located at the beginning of the name of each variable.

**Table 1.** Variable description

| Variable | Definition | Data type |
|---|---|---|
| eventID | ID of an event | Categorical |
| case Spend area text | Upper-level classification of an item | Categorical |
| case Company | ID of a company | Categorical |
| case Document Type | The type of a flow of an item | Categorical |
| case Sub spend area text | Lower-level classification of an item | Categorical |
| case Item Type | The type of an item | Categorical |
| case Item Category | The category of an item | Categorical |
| case GR-Based Inv. Verif. | Binary indicator signifying whether GR-Based invoicing is necessary | Binary |
| case concept:name | Case ID | Categorical |
| case Goods Receipt | Binary indicator signifying whether 3-way match is necessary | Binary |
| event org:resource | Resource ID | Categorical |
| event concept:name | Activity | Categorical |
| event Cumulative net worth (EUR) | Cumulative net worth in Euro | Numeric |
| event time:timestamp | Timestamp | Date |

**Table 2.** Number of cases by Company ID and Item Category

| Item Category | CompanyID_0000 | CompanyID_0001 | CompanyID_0002 | CompanyID_0003 |
|---|---|---|---|---|
| 3-way match, before GR | 191,135 | 0 | 2 | 0 |
| 3-way match, after GR | 13,205 | 2 | 0 | 0 |
| 2-way match | 0 | 0 | 0 | 640 |
| Consignment | 13,837 | 0 | 0 | 0 |

The event log contains information about transactions between subsidiaries and vendors, and the number of items purchased by different subsidiaries are presented in Table 2. The Table 3 shows the overall data structure and the number of items that fall into each distinctive category defined by *Item Category*, *Document Type*, and *Item Type*. In addition to these three features, *Spend area text* and *Sub spend area text* are also used to further specify the distinctive category of an item. To summarize, each item can be defined by five features, *Item Category*, *Document Type*, *Item Type*, *Spend area text*, and *Sub spend area text*, and we will use this way of defining an item based on five features throughout the remaining part of the report (see Fig. 2).

**Table 3.** Data Structure

| Item Category | Document Type | Item Type | Frequency |
|---|---|---|---|
| 3-way match, after GR (mean duration= 75 days) | Standard PO | Service | 4,547 |
| | | Standard | 7,042 |
| | | Subcontracting | 582 |
| | | Third-party | 323 |
| | EC purchase order | Service | 341 |
| | | Standard | 108 |
| | Framework order | Service | 264 |
| 3-way match, before GR (mean duration = 74.5 days) | Standard PO | Standard | 182,160 |
| | | Subcontracting | 3,664 |
| | | Third-party | 4,598 |
| | EC purchase order | Standard | 714 |
| | Framework order | Standard | 1 |
| 2-way match (mean duration = 57.5 days) | Framework order | Limit | 640 |
| Consignment (mean duration = 24.1 days) | Standard PO | Consignment | 13,837 |

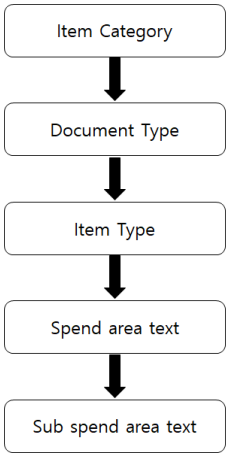



**Fig. 2.** Level of the analysis

### 2.4 Four types of flows

The event log contains four types of flows that can be defined by different combinations of three features, *case Goods Receipt*, *case GR-Based Inv. Verif*, and *Item Type*.

**(1) 3-way matching, invoice after goods receipt:** One of subsidiaries submits a purchase order for a specific item to a vendor. Then, the vendor will create an invoice and send the ordered item to the company. After receiving the item, the subsidiary will check the list of materials with the purchase order document and the invoice receipt. The company can revise the price or quantity of the item before receiving the ordered item. For the final step, the company will completely clear this order, if there is no problem on the invoicing process.
**(2) 3-way matching, invoice before goods receipt:** This flow is similar to *3-way matching, invoice after goods receipt*, except for the fact that invoices can be received before receiving goods receipts. As in *3-way matching, invoice after goods receipt*, the subsidiary will completely clear the order if there is no problem on this invoicing process.
**(3) 2-way matching:** In this process, goods receipt message is not required as all orders made in this flow are framework orders.
**(4) Consignment:** As the name of the flow implies, invoice messages are not required since it is handled in a separate process operated by the third-party company. Therefore, messages on goods receipt can only be observed in this flow.

As the name of each type of flow is long, we decided to change the name of each type of flow by deleting the word "invoice" and changing the word "matching" with "match" from each name of the type of flow for the remaining part of the report.

**Table 4.** Different *Goods Receipt & GR-based Inv. Verif* pair for four types of flows

| Item Category | case Goods Receipt | case GR-Based Inv. Verif. | Item Type |
|---|---|---|---|
| 3-way match, after GR | TRUE | TRUE | Various |
| 3-way match, before GR | TRUE | FALSE | Various |
| 2-way match | FALSE | FALSE | Various |
| Consignment | TRUE | FALSE | Consignment |

## 3 Building a business process model

**Question 1.** Is there a collection of process models which together properly describe the process in this data. Based on the four categories above, at least 4 models are needed, but any collection of models that together explain the process well is appreciated. Preferably, the decision which model explains which

purchase item best is based on properties of the item.

The BPMN model was derived from *petri net* constructed using ProM as in Fig. 3. Since the BPMN model is very complicated as it encompasses all four types of flows, subprocesses are introduced as in Fig. 4 to enhance the understandability by making the BPMN model simple and neat. Fig. 5 (a) is a subprocess called *SRM* that can be found in *3-way match, before GR*. Fig. 5 (b) is a subprocess called *SRM(Transfer followed by order)* that can be found in *3-way match, after GR*. Fig. 5 (c) is a subprocess called *Vendor Creates Invoice & Record Goods Receipt (Goods)* that can be found in more than two types of flow, and this subprocess deals with goods. Fig. 5 (d) is a subprocess called *Vendor Creates Invoice & Record Goods Receipt (Service)* that can be found in more than two types of flow, and this subprocess deals with services. Fig. 5 (e) is a subprocess called *Record Goods & Invoice receipt*that can be found in more than two types of flow. In Fig. 5 (c) to Fig. 5 (e), "&" means that any of two activities can have chance to happen earlier than the other activity.

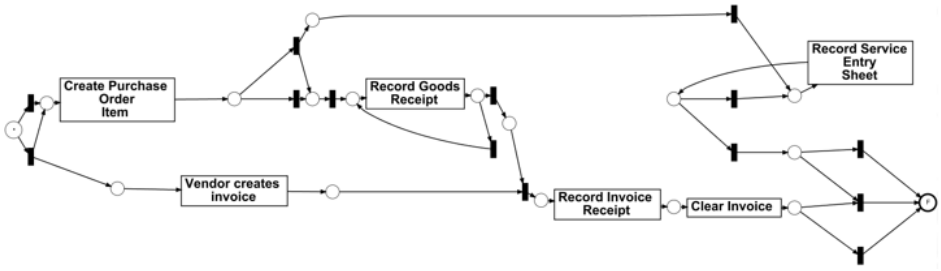



**Fig. 3.** An example of a petri net

## 4 Throughput of the invoicing process

**Question 2.** What is the throughput of the invoicing process, i.e. the time between goods receipt, invoice receipt and payment (clear invoice)? To answer this, a technique is sought to match these events within a line item, i.e. if there are multiple goods receipt messages and multiple invoices within a line item, how are they related and which belong together?

Each item contains one or more events with the information about its case duration. In a business process, case duration which means the order leading time is regarded as one of the most important factors from the perspective of a customer. Based on the business process model developed in Fig. 4, we investigate distributions of case duration for each type of 13 business process models.

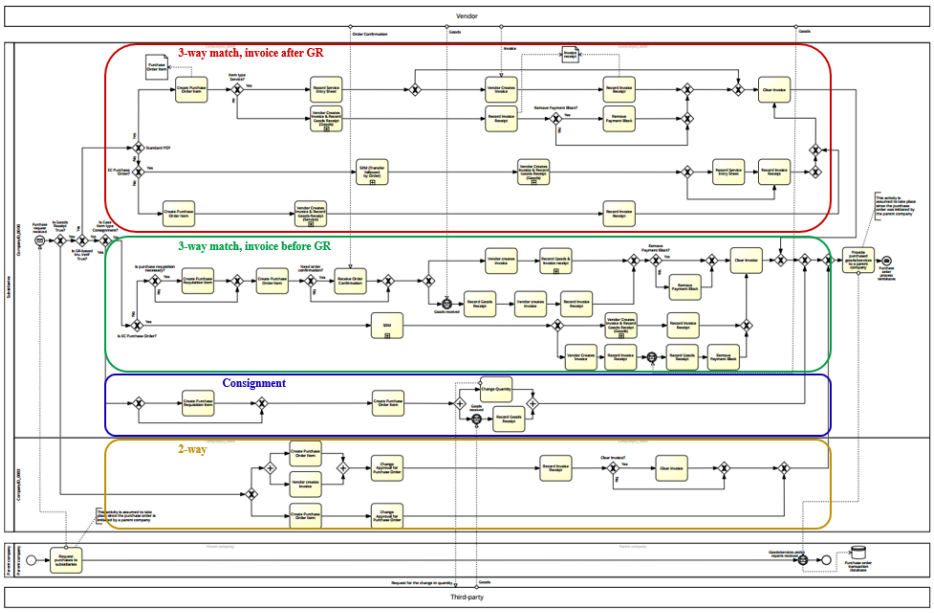



**Fig. 4.** A BPMN model

The analysis was conducted based on two approaches, *activity level approach* and *instance level approach*.

### 4.1 Activity level approach

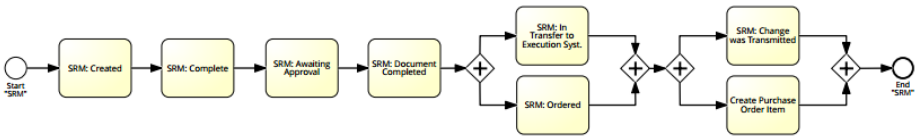
For each of 13 business process models which consists of different *Document Type* (D.T.) and *Item Type* (I.T.), we investigated durations for each activity. In order to exhibit the outcomes in a more efficient way, proportion of duration for each activity was calculated as in Fig. 6 and we focused on activities whose waiting time is longer than 30 percent of the total duration of an item.

(1) Duration in 3-way match, after GR (D.T. = EC purchase order, I.T. = service)
- from "Record Invoice Receipt" to "Clear Invoice" = 37%

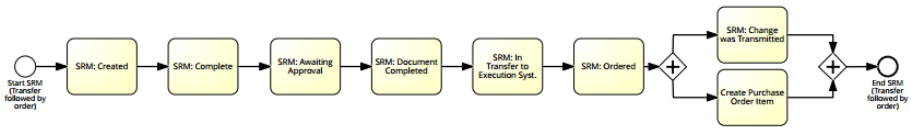
(2) Duration in 3-way match, after GR (D.T. = EC purchase order, I.T = standard)
- from "Record Goods Receipt" to "Record Invoice Receipt" = 43%
- from "Record Invoice Receipt" to "Clear Invoice" = 45%

(3) Duration in 3-way match, after GR (D.T. = Standard PO, I.T. = service)
- from "Vendor creates invoice" to "Record Invoice Receipt" = 33%
- from "Record Invoice Receipt" to "Clear Invoice" = 41%

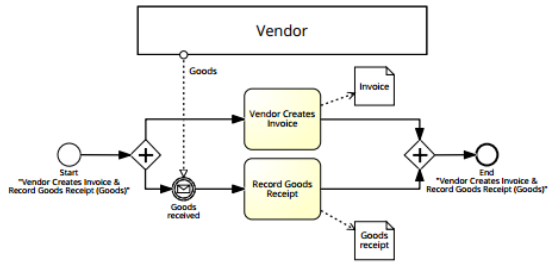
(4) Duration in 3-way match, after GR (D.T.= Standard PO, I.T. = Subcontracting)



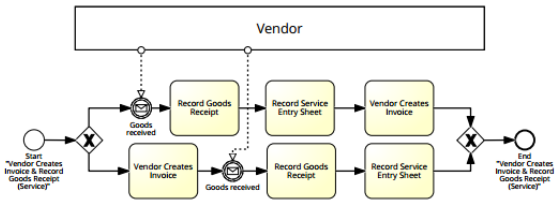

(a) SRM



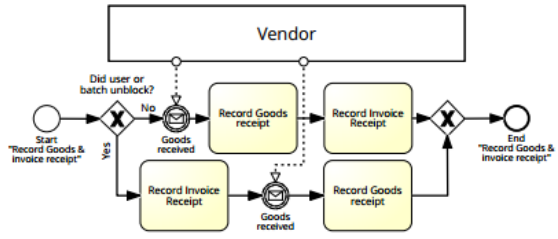

(b) SRM (Transfer followed by order)




(c) Vendor Creates Invoice & Record Goods Receipt (Goods)




(d) Vendor Creates Invoice & Record Goods Receipt (Service)




(e) Record Goods & Invoice receipt

**Fig. 5.** Subprocesses of BPMN model

- from "Record Goods Receipt" to "Record Invoice Receipt" = 100%

(5) Duration in 3-way match, after GR (D.T. = Standard PO, I.T. = Third-party)
- from "Vendor creates invoice" to "Record Invoice Receipt" = 31%
- from "Record Invoice Receipt" to "Clear Invoice" = 37%

(6) Duration in 3-way match, after GR (D.T. = Framework order, I.T. = Service)
- from "Create Purchase Order Item" to "Vendor creates invoice" = 37%

(7) Duration in 2-way match (D.T. = Framework Order, I.T. = Limit)
- from "Vendor creates invoice" to to "Create Purchase Order Item" = 65%

(8) Duration in Consignment (D.T. = Standard PO, I.T. = Consignment)
- from "Create Purchase Order Item" to "Record Goods Receipt" = 78%

### 4.2 Instance level approach

In this part, we assumed that some of distributions of case duration for each of 13 discovered process models is either similar or different according to its types such as *Document Type* (D.T.), *Item Type* (I.T.), and *Sub spend area text*, etc. In order to investigate this, some of the models were grouped.
Furthermore, we investigated some items with long order leading time, and tried to find factors in attribute level that affects the order leading time. For this, we implemented the regression analysis and extracted factors that are statistically significant. Our regression models were properly fitted under significant level=0.05 for all process models. Interestingly, we found 672 significant factor levels in 13 regression models and the summary is provided in Table 5-10. We showed maximum five factors for each attribute because of the limitation of the space of the table.

**Note:** For *Cumulative net worth*, we discovered that some of line items have multiple goods receipts. In addition, the total *Cumulative net worth* of one line item is calculated by summation of the *Cumulative net worth* in multiple good receipts.

**(1)** 3-way match ***after vs before*** goods receipt (D.T.= EC purchase order, I.T.= Standard)

**Comparison:** In terms of 3-way match with D.T.= EC purchase order and I.T.= Standard, *3-way match, after GR* has more cases with longer order leading time than *3-way match, before GR* as in Fig. 7. Additionally, *3-way match, before GR* has more transactions with vendors in different *sub spend areas* and it triggers deviation.

- 3-way match, after GR: Transaction with 5 vendors in 4 types of *Sub spend area*

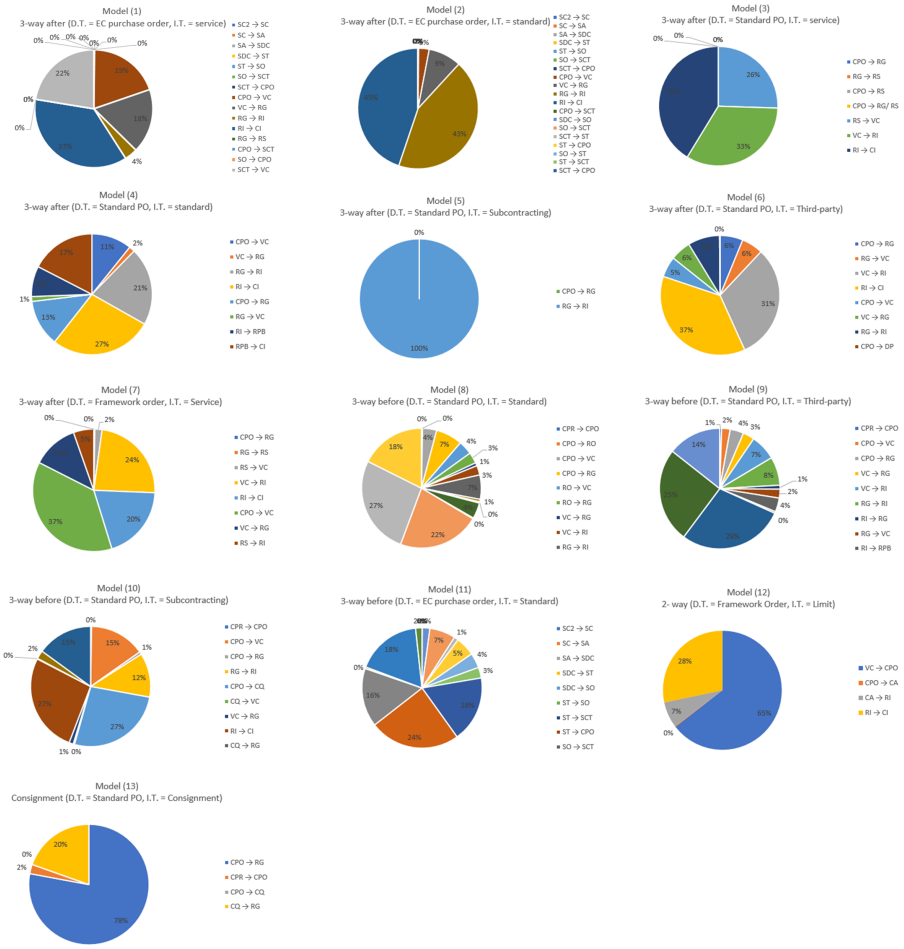



**Fig. 6.** Pie charts about proportion of duration for each activity

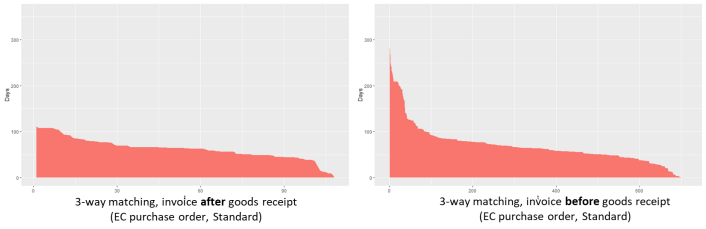



**Fig. 7.** Distribution of duration per each instance - part (1)

- 3-way match, before GR: Transaction with 70 vendors in 23 types of *Sub spend area*

**Regression for each model:** As a result of linear regression, Table 5 shows a list of significant factors that affect case duration. According to Table 5, the duration of the instance is decreased by 14.3288 days on average in case of "Laboratory Supplies  Services" factor.

- 3-way match, after GR: For the decreasing effect on case duration, one factor named "Laboratory Supplies & Services" in *Sub spend area text* and one vendor named vendorID_0029 were found to be significant.
- 3-way match, before GR: For the increasing effect on case duration, 3 factors were observed in *Sub spend area text* and 5 factors were observed in *Vendor* including vendorID_0004.

**Table 5.** Regression analysis: the list of significant factors in part (1)

| Attibutes | Factors | Estimate | P-value |
|---|---|---|---|
| ***3-way match, after GR*** | | | |
| -Sub spend area text | Laboratory Supplies & Services | -14.3288 | 0.0130 |
| -Vendor | vendorID_0029 | -17.3758 | 0.0000 |
| ***3-way match, before GR*** | | | |
| -Sub spend area text | Facility Management | 80.4243 | 0.0341 |
| | Laboratory Supplies & Services | 128.6787 | 0.0138 |
| | Other Logistics Services | 102.7780 | 0.0126 |
| -Vendor | vendorID_0004 | 89.2219 | 0.0365 |
| | vendorID_0024 | -96.2948 | 0.0173 |
| | vendorID_0027 | -38.2701 | 0.0424 |
| | vendorID_0042 | -81.3129 | 0.0110 |
| | vendorID_0045 | -91.9423 | 0.0283 |

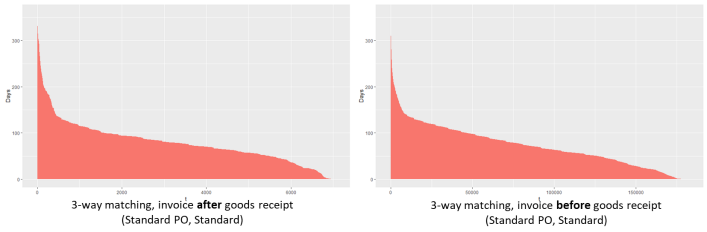



**Fig. 8.** Distribution of duration per each instance - part (2)

**Table 6.** Regression analysis: the list of significant factors in part (2)

| Attibutes | Factors | Estimate | P-value |
|---|---|---|---|
| ***3-way match, after GR*** | | | |
| Sub spend area text | Business Gifts&Promotional Items | 98.0376 | 0.0172 |
| | Chloride | -59.9164 | 0.0090 |
| | Commercial Printing | 123.2229 | 0.0026 |
| | Commodity Resins Precursors | -60.1448 | 0.0010 |
| | HR Services | 198.7098 | 0.0000 |
| Vendor | vendorID_0106 | -43.4573 | 0.0351 |
| | vendorID_0120 | -37.1949 | 0.0238 |
| | vendorID_0123 | 155.1773 | 0.0036 |
| | vendorID_0125 | -60.5384 | 0.0004 |
| | vendorID_0136 | -42.6069 | 0.0104 |
| Cumulative net worth | net worth (unit=100 EUR) | 0.0097 | 0.0000 |
| ***3-way match, before GR*** | | | |
| Sub spend area text | Additives - Other | -30.5778 | 0.0037 |
| | Adhesion Promotors | -31.7756 | 0.0059 |
| | Advertising | 93.4852 | 0.0057 |
| | Alcohol Solvents | -38.7260 | 0.0007 |
| | Aliphatic Solvents | -30.3879 | 0.0039 |
| Vendor | vendorID_0105 | -55.4461 | 0.0449 |
| | vendorID_0206 | -64.6910 | 0.0267 |
| | vendorID_0233 | 59.3381 | 0.0445 |
| | vendorID_0275 | -57.1588 | 0.0390 |
| | vendorID_0318 | -69.3700 | 0.0251 |
| Cumulative net worth | net worth(unit=100 EUR) | 0.0028 | 0.0000 |

**(2)** 3-way match ***after vs before*** goods receipt (D.T.= Standard PO, I.T.= Standard)

**Comparison:** In Fig. 8, the two plots have no clear difference in distribution of case duration even though *3-way match before GR* has much more transactions with different types of *Sub spend areas*.

- 3-way match, after GR: Transaction with 158 vendors in 48 types of *Sub spend area text*
- 3-way match, before GR: Transaction with 1263 vendors in 111 types of *Sub spend area text*

**Regression for each model:**
-3-way match, after GR: For the increasing effect on case duration, 9 factors were observed in *Sub spend area text* and 13 factors were observed in *Vendor*. On the other hand, transactions with 20 factors in *Sub spend area text* and 20 vendors have deceasing effects in case duration. About *Cumulative net worth*, the case duration is averagely increased with 9.7 days when the net worth is increased with the amount of 100,000 EUR.
-3-way match, before GR: Since the model has largest frequency among 13 business models and also has many levels of attributes, there have been observed 65 significant factors in *Sub spend area text* and 91 vendors. Among

those factors, transactions with 17 factors in *Sub spend area text* and 69 vendors were observed to have increasing effect on its case duration. For the effect of *Cumulative net worth*, the case duration was averagely increased by 2.8 days when the net worth increased by 10,000 EUR.

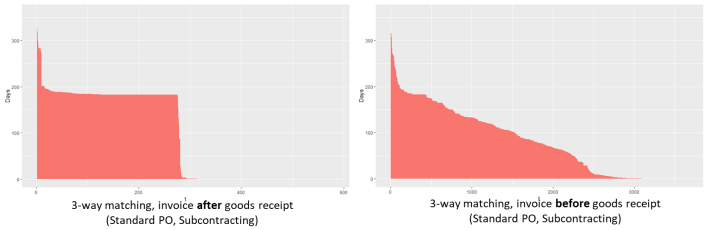



**Fig. 9.** Distribution of duration per each instance - part (3)

**(3)** 3-way match ***after vs before*** goods receipt (D.T.= Standard PO, I.T.= Subcontracting)

**Comparison:** Interestingly, in Fig. 9, distribution of case duration in *3-way match, after GR* looks quite constant while the other model is exponentially distributed. We can state that the prior process has been managed well with respect to order leading time.

- 3-way match, after GR: Transaction with 5 vendors in 3 types of *Sub spend area text*
- 3-way match, before GR: Transaction with 32 vendors in 14 types of *Sub spend area text*

**Regression for each model:**

- 3-way match, after GR: Because of the uniform distribution, there was no increasing or decreasing effect on attribute level.
- 3-way match, before GR: For the increasing effect, order leading time increased when a company had transactions with 4 vendors as in below. On the other hand, transactions with 2 factors in *Sub spend area text* and 3 vendors have increasing effects in case duration.

**(4)** 3-way match ***after vs before*** goods receipt (D.T.= Standard PO, I.T.= Third-party)

**Comparison:** Above two plots in Fig. 10 have no clear different distribution although the process model with *3-way match before GR* has much more transactions with different types of *Sub spend areas*.

**Table 7.** Regression analysis: the list of significant factors in part (3)

| Attibutes | Factors | Estimate | p-value |
|---|---|---|---|
| ***3-way match, after GR*** | | | |
| No effect | | | |
| ***3-way match, before GR*** | | | |
| Sub spend area text | Other Logistics Services | -95.4080 | 0.0257 |
| | Road Packed | -82.4135 | 0.0355 |
| Vendor | vendorID_0176 | 30.2173 | 0.0003 |
| | vendorID_0197 | 16.5945 | 0.0339 |
| | vendorID_0260 | -82.8960 | 0.0000 |
| | vendorID_0548 | -108.7705 | 0.0010 |
| | vendorID_0581 | 23.7685 | 0.0309 |

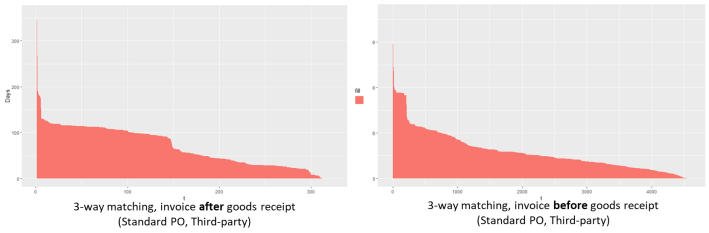



**Fig. 10.** Distribution of duration per each instance - part (4)

- 3-way match, after GR: Transaction with 8 vendors in 3 types of *Sub spend area text*
- 3-way match, before GR: Transaction with 40 vendors in 6 types of *Sub spend area text*

**Regression for each model:**

-3-way match, after GR: Among total 8 vendors, 3 vendors were observed to have significantly shorter case duration while one vendor was observed to have longer order leading time. In terms of *Cumulative net worth*, the case duration increased by 31.87 days on average, when the net worth increased by 10,000 EUR.

-3-way match, before GR: For the increasing effect in case duration, only one factor with Color Collateral showed significantly increasing effect. Regarding *Cumulative net worth*, when *Cumulative net worth* increased bof 10,000 EUR, the case duration is increased by 4.86 days on average.

**Table 8.** Regression analysis: the list of significant factors in part (4)

| Attibutes | Factors | Estimate | p-value |
|---|---|---|---|
| ***3-way match, after GR*** | | | |
| Vendor | vendorID_0299 | 53.7619 | 0.0000 |
| | vendorID_0374 | -79.3028 | 0.0020 |
| | vendorID_0381 | -68.6248 | 0.0136 |
| | vendorID_0476 | -63.5625 | 0.0130 |
| Cumulative net worth | net worth (unit=100 EUR) | 0.3187 | 0.0182 |
| ***3-way match, before GR*** | | | |
| Sub spend area text | Color Collateral | 217.1101 | 0.0002 |
| Cumulative net worth | net worth(unit=100 EUR) | 0.0486 | 0.0003 |

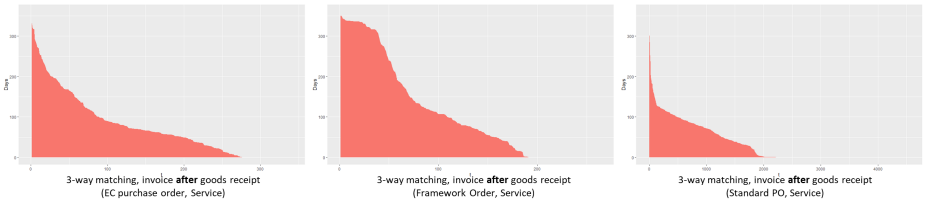



**Fig. 11.** Distribution of duration per each instance - part (5)

**(5)** 3-way match, after GR (D.T.= **EC purchase order vs Framework order vs Standard PO**, I.T.= Service)

**Comparison:** Fig. 11 shows that the ranges of case duration are similar as 3 processes have been handled with same *Item Type*. Due to different condition on *Document Type*, however, the amount of transactions is different for each model.

- EC purchase order: Transaction with 152 vendors in 32 types of *Sub spend area text*
- Framework order: Transaction with 88 vendors in 18 types of *Sub spend area text*
- Standard PO: Transaction with 64 vendors in 10 types of *Sub spend area text*

**Regression for each model:**
- EC purchase order: In *sub spend area*, 3 factors were observed to have strong increasing effect on case duration. In the meantime, 20 vendors increased the case duration while one vendor decreased it.
- Framework order: For the increasing effect, one factor about Customers in *sub spend area* and 23 vendors were observed to have significant probability. On the other hand, 6 factors in *sub spend area* and 3 vendors were observed to decrease the case duration.
- Standard PO: Two factors in *sub spend area* and 17 vendors had increasing effect on case duration while other 4 vendors had decreasing effect on it.

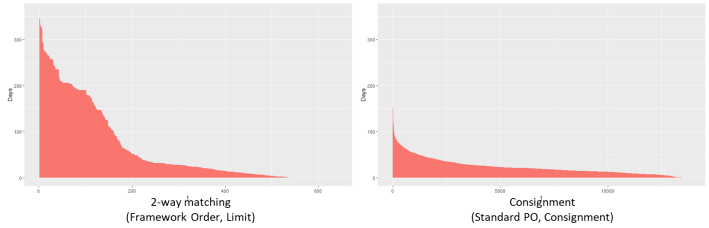



**Fig. 12.** Distribution of duration per each instance - part (6)

**(6)** The others

**Comparison:** In Fig. 12, *2-way match* process showed a huge deviation since it had transactions with large number of vendors in different *Sub spend areas*. For *Consignment* process, although some types of activities in ordering process

**Table 9.** Regression analysis: the list of significant factors in part (5)

| Attibutes | Factors | Estimate | p-value |
|---|---|---|---|
| ***EC purchase order*** | | | |
| Sub spend area text | Information Services | 208.2099 | 0.0032 |
| | MRO (components) | 223.8764 | 0.0020 |
| | Process Automation & Instrumentation | 202.1354 | 0.0042 |
| Vendor | vendorID_0019 | 163.3033 | 0.0093 |
| | vendorID_0022 | 334.6275 | 0.0000 |
| | vendorID_0025 | 205.0393 | 0.0034 |
| | vendorID_0038 | 318.3897 | 0.0000 |
| | vendorID_0043 | 123.6410 | 0.0414 |
| ***Framework order*** | | | |
| Sub spend area text | Customers | 289.7913 | 0.0238 |
| | Express | -220.9743 | 0.0372 |
| | Other Logistics Services | -232.4003 | 0.0285 |
| | Products for Resale | -272.1007 | 0.0105 |
| | Road Packed | -272.4612 | 0.0102 |
| Vendor | vendorID_0228 | 302.8015 | 0.0186 |
| | vendorID_0231 | 253.7334 | 0.0004 |
| | vendorID_0232 | 167.5334 | 0.0333 |
| | vendorID_0335 | -162.2973 | 0.0452 |
| | vendorID_0338 | 174.3054 | 0.0269 |
| ***Standard PO*** | | | |
| Sub spend area text | MRO (components) | 100.5779 | 0.0092 |
| | Other Logistics Services | 60.7551 | 0.0096 |
| Vendor | vendorID_0230 | -29.5201 | 0.0000 |
| | vendorID_0231 | 93.1153 | 0.0000 |
| | vendorID_0232 | 56.7194 | 0.0000 |
| | vendorID_0233 | 69.9751 | 0.0000 |
| | vendorID_0263 | 125.8898 | 0.0048 |

may partly be skipped, we could extract meaningful features depending on the partial information obtained from regression analysis.

- 2-way match(D.T.= Framework order, I.T.= Limit): Transaction with 250 vendors in 17 types of *Sub spend area text*
- Consignment (D.T.= Standard PO, I.T.= Consignment): Transaction with 95 vendors in 43 types of *Sub spend area text*

**Regression for each model:**
- 2-way match: For increasing effect on case duration, one factor named "QHSE Services" in *Sub spend area text* made the order leading time longer while two factors in *Sub spend area text* had decreasing effect on it. With respect to *Vendor*, transactions with 10 vendors significantly increase the order leading time.
- Consignment: In spite of partial information, several significant effects were found in *Sub spend area text* and *Vendor*. Two factors in *Sub spend area text* were observed to increase the order leading time while five factors had decreasing effect. In addition, it was observed that transactions with 30 vendors increased the order leading time while transactions with 19 vendor decreased it.

**Table 10.** Regression analysis: the list of significant factors in part (6)

| Attibutes | Factors | Estimate | p-value |
|---|---|---|---|
| ***2-way match(I.T=Limit)*** | | | |
| Sub spend area text | Government payments | -157.6528 | 0.0328 |
| | QHSE Services | 116.3163 | 0.0260 |
| | Taxation | -199.9079 | 0.0031 |
| Vendor | vendorID_1687 | 123.6710 | 0.0431 |
| | vendorID_1688 | 214.3583 | 0.0007 |
| | vendorID_1691 | 255.5862 | 0.0005 |
| | vendorID_1695 | 197.0831 | 0.0023 |
| | vendorID_1703 | 205.2750 | 0.0007 |
| ***Consignment (I.T.=Consignment)*** | | | |
| Sub spend area text | Alcohol Solvents | -23.5552 | 0.0008 |
| | Aliphatic Solvents | -24.3641 | 0.0003 |
| | Aromatic Solvents | -24.2475 | 0.0009 |
| | Chloride | 10.9585 | 0.0391 |
| | Color Collateral | -47.6652 | 0.0000 |
| Vendor | vendorID_0062 | -17.9757 | 0.0004 |
| | vendorID_0146 | 15.4614 | 0.0002 |
| | vendorID_0162 | 19.5504 | 0.0060 |
| | vendorID_0175 | 35.8772 | 0.0000 |
| | vendorID_0214 | 24.6655 | 0.0204 |

## 5 Deviation analysis based on business process model

**Question 3.** Finally, which Purchase Documents stand out from the event log? Where are deviations from the processes discovered in (1) and how severe are these deviations? Deviations may be according to the prescribed high-level process flow, but also with respect to the values of the invoices. Which customers produce a lot of rework as invoices are wrong, etc.?

### 5.1 Filtering deviant cases using Alignment technique in ProM

Our business process model is reconstructed based on *petri net* model which we assume as the baseline of model. In this problem, the *petri net* model is used for deviding the event log as 'general model' and 'deviant model' by applying *alignment* technique equipped in Prom tool.
The *alignment* is implemented for each of 13 business process models and we assumed that if a completed instance has fitness value less than 0.95, the instance is belong to 'deviant model'. The implemented result is summarized in table 11. As in Fig. 13, we observed that there are several types of deviant processes containing with missing or wrong(replicated) events in cases that fitness value is less 0.95. For the next part, we analyzed the two types of cases in 'general model' and 'deviant model' both interactively and independently.

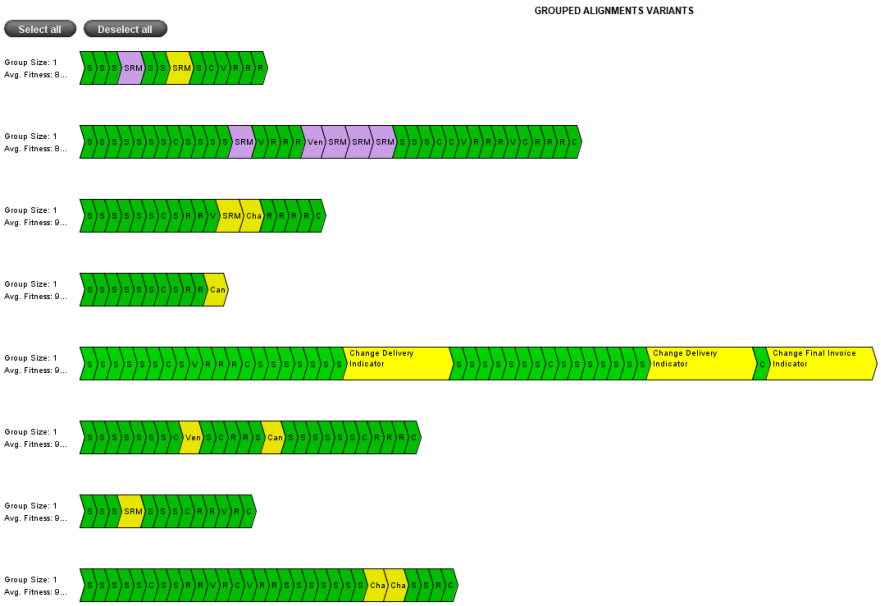



**Fig. 13.** An example of variants in alignment

**Table 11.** Alignment Regular and Deviation by Item Category

| Item Category | Document Type | Item Type | # of Traces | Threshold of inductive miner | Avg. fitness | Min. fitness | General Cases (# of variants) | Deviant Cases (# of variants) |
|---|---|---|---|---|---|---|---|---|
| 3-way match, after GR | EC Purchase order | Service | 341 | 10% | 99.6% | 88.2% | 309 | 32 |
| | | Standard | 108 | 10% | 99.3% | 78.8% | 98 | 10 |
| | Framework order | Service | 264 | 10% | 98.3% | 72.3% | 221 | 43 |
| | Standard PO | Service | 4547 | 10% | 97.3% | 67.2% | 3548 | 999 |
| | | Standard | 7042 | 10% | 93.4% | 17.9% | 5468 | 1574 |
| | | Subcontracting | 582 | 10% | 99% | 71.4% | 552 | 30 |
| | | Third-Party | 323 | 10% | 95.3% | 57.1% | 274 | 49 |
| 3-way match, before GR | EC Purchase order | Standard | 714 | 10% | 98.7% | 55.1% | 618 | 93 |
| | Standard PO | Standard | 182160 | 10% | 91.7% | 12.2% | 125479 | 65293 |
| | | Subcontracting | 3664 | 10% | 99.7% | 78.2% | 2881 | 783 |
| | | Third-Party | 4598 | 10% | 97.9% | 50% | 3733 | 865 |
| 2-way match | Framework order | Limit | 640 | 10% | 99.6% | 77.4% | 614 | 26 |
| Consignment | Standard PO | Consignment | 13837 | 5% | 96.3% | 43.7% | 13109 | 728 |

### 5.2 Comparison of distribution of features between regular and deviant cases

It is commonly expected that distributions of features between regular and deviant cases is not similar to each other. In order to test whether distributions of features of regular and deviant cases differ, various statistical tests were conducted.

Firstly, categorical features of regular and deviant cases are compared using *Chi-square test* and *Fisher's exact test*. It is widely known that *Chi-square test* can be used when the size of a contingency matrix is bigger than 22 and each cell of a matrix is bigger than or equal to 5, while Fishers exact test is acknowledged to be used when the size of a contingency matrix is 2x2 and at least one of a cell of a matrix is smaller than 5. In case where both the size of a contingency matrix is bigger than 2x2 and at least one of a cell of a matrix is smaller than 5, Monte-Carlo simulation is used to calculate the p-value. The Wikipedia says, "For hand calculations, the test is only feasible in the case of a 2  2 contingency table. However the principle of the test can be extended to the general case of an m  n table, and some statistical packages provide a calculation (sometimes using a Monte Carlo method to obtain an approximation) for the more general case." There are few assumptions to be satisfied for Chi-square test and Fisher's exact test. These assumptions are that the value in a cell of a matrix should be the count of occurrences and the sum of all counts in a matrix should be equal to the size of a feature vector, which are all satisfied.

Secondly, a numeric feature of regular and deviant cases, *Cumulative net worth*, can be compared using *Kolmogorov-Smirnov test* and *Mann-Whitney-Wilcoxon test*. In order to conduct Kolmogorov-Smirnov test, normality assumption should be satisfied. To check normality, Shapiro-Wilk normality test was conducted as shown in the table below.

In our case, however, the normality assumption was not satisfied. Therefore, *Mann-Whitney-Wilcoxon test* was used to compare the distribution of a numeric feature between regular and deviation cases. The result verifies that the distribution of all four categorical features are different from regular cases to deviant cases.

**Table 12.** Shapiro-Wilk normality test

| Item Category | Document Type | Item Type | Regular | | Deviant | |
|---|---|---|---|---|---|---|
| | | | W | p-value | W | $p$-value |
| 3-way match, before GR | Standard Purchase Order | Standard | 0.30 | <0.001 | 0.33 | <0.001 |
| | | Third party | 0.39 | <0.001 | 0.51 | <0.001 |
| | | Subcontracting | 0.69 | <0.001 | 0.71 | <0.001 |
| | EC Purchase Order | Standard | 0.43 | <0.001 | 0.19 | <0.001 |
| 3-way match, after GR | Standard Purchase Order | Standard | 0.38 | <0.001 | 0.26 | <0.001 |
| | | Third party | 0.54 | <0.001 | 0.64 | <0.001 |
| | | Subcontracting | 0.62 | <0.001 | 0.81 | <0.001 |
| | | Service | 0.06 | <0.001 | 0.08 | <0.001 |
| | EC Purchase Order | Standard | 0.69 | <0.001 | 0.68 | <0.001 |
| | | Service | 0.66 | <0.001 | 0.55 | <0.001 |
| | Framework Order | Service | 0.30 | <0.001 | 0.30 | <0.001 |
| 2-way match | Framework Order | Limit | 0.59 | <0.001 | 0.48 | <0.001 |
| Consignment | | | NA | NA | NA | NA |

**Table 13.** Mann-Whitney-Wilcoxon test

| Item Category | Document Type | Item Type | W | $p$-value |
|---|---|---|---|---|
| 3-way match, before GR | Standard Purchase Order | Standard | 1.02e+11 | <0.001 |
| | | Third party | 4.02e+7 | <0.001 |
| | | Subcontracting | 4.40e+7 | <0.001 |
| | EC Purchase Order | Standard | 6.60e+6 | <0.001 |
| 3-way match, after GR | Standard Purchase Order | Standard | 2.14e+8 | <0.001 |
| | | Third party | 2.15e+5 | 0.001 |
| | | Subcontracting | 1.36e+5 | <0.001 |
| | | Service | 1.45e+9 | <0.001 |
| | EC Purchase Order | Standard | 8.61e+4 | 0.001 |
| | | Service | 1.04e+6 | <0.001 |
| | Framework Order | Service | 1.28e+6 | <0.001 |
| 2-way match | Framework Order | Limit | 5.67e+5 | 0.3358 |
| Consignment | | | NA | NA |

**Table 14.** Chi-square test / Fisher's exact test

| Item Category | Document Type | Item Type | $p$-value | | | |
|---|---|---|---|---|---|---|
| | | | Resource | Spend area | Sub spend area | Vendor |
| 3-way match, before GR | Standard Purchase Order | Standard | 0.001 | 0.001 | 0.001 | 0.001 |
| | | Third party | 0.001 | 0.001 | 0.001 | 0.001 |
| | | Subcontracting | 0.001 | 0.001 | 0.001 | 0.001 |
| | EC Purchase Order | Standard | 0.001 | 0.001 | 0.001 | 0.001 |
| 3-way match, after GR | Standard Purchase Order | Standard | 0.001 | 0.001 | 0.001 | 0.001 |
| | | Third party | 0.001 | 0.001 | 0.001 | 0.001 |
| | | Subcontracting | 0.001 | 0.001 | 0.001 | 0.001 |
| | | Service | 0.001 | 0.001 | 0.001 | 0.001 |
| | EC Purchase Order | Standard | 0.001 | 0.001 | 0.001 | 0.001 |
| | | Service | 0.001 | 0.001 | 0.001 | 0.001 |
| | Framework Order | Service | 0.001 | 0.001 | 0.001 | 0.001 |
| 2-way match | Framework Order | Limit | 0.001 | 0.001 | 0.001 | 0.001 |
| Consignment | | | 0.007 | 0.001 | 0.001 | 0.001 |

### 5.3 Rework Attributes

With cases whose fitness was found to be low according to the *alignment* technique, we classified those cases into two groups based on whether lots of rework processes had occurred or not. The rework process can be discriminated based on two conditions. First condition is whether the case has at least one activity whose name has 'Change' in it (e.g. change quantity). The second condition is whether each activity occurred more than once. Using cases filtered with two conditions, we figured out the difference between rework and non-rework with respect to *Spend area text* and *Sub spend area text*. Table 15 is the list of values in (*Spend area text* and *Sub spend area text*) that only rework group has. Among 13 case groups, there were 3 groups that all the attributes in the rework also appeared at the non-rework group.

**Table 15.** Rework Attribute

| Item Category | Document Type | Item Type | Spend area text | Sub spend area text |
|---|---|---|---|---|
| 3-way match, before GR | EC Purchase Order | Standard | Order | MRO (components), Facility Management, Advertising, Packaging, Other Logistic Services |
| | Standard PO | Standarxd | Solvents | Styrene Acrylics, Paperboard |
| | | Subcontracting | Titanium Dioxides, Pigments & Colorants, Marketing | Sulphate, MRO (components), Colorants, Point of Sales, Warehousing |
| | | Thirdparty | - | - |
| 3-way match, after GR | EC Purchase Order | Service | Enterprise Services | Consulting, HR Services, QHSE Services, Third Party Labor, Marketing Support Services |
| | | Standard | - | - |
| | Framework Order | Service | CAPEX & SOCS, Workforce Services | Sea, Packaging, Transport & Hoisting Equipment, HR Services |
| | Standard PO | Service<br>Standard<br>Subcontracting<br>Thirdparty | CAPEX & SOCS<br>Specialty Resins, Commodity Resins<br>Logistics<br>- | MRO (components)<br>Pure Acrylics, Vinyl Acrylics<br>Road Packed<br>- |
| 2-way match | Framework Order | Limit | Real Estate, Energy | Real estate services, Real estate brokers or agents, Business park, Government payments, Electricity |
| Consignment | Standard PO | Consignment | Latex & Monomers | Opaque Polymers, Styrene Acrylics, Polyurethane Resins, Aliphatic Solvents, Alkyd Resins |

## 6 Discussion

In the paper, we focused on discovering the business process model and analyzing the deviations in cases. Even though all four types of flows have disparate processes, we tried to find a comprehensive business process model that best describes the majority of the cases that belong to each type of flows. In deviation analysis, we tried to figure out whether there are deviations both in terms of the duration of time it takes from activity/instance level, and in terms of the order of activities based on business process model. In analyzing deviations in the duration of time, we investigated the deviations based on the distribution of duration for each type of processes using regression method. In analyzing deviations in activities, we applied Alignment plug-in made available in ProM to filter deviant cases, where the deviation is defined based on business process model discovered in Chapter 3. Furthermore, the extent of deviation is calculated using statistical tests, such as *Kolmogorov-Smirnov test*

and *Mann-Whitney-Wilcoxon test.*
In addition to the analysis we have conducted throughout this paper, we also came up with several findings as below.

### 6.1 Findings from process discovery

-According to the explanation of *3-way match, before GR* provided in BPIC 2019 guidelines, a specific user or a batch process enables invoices to be entered and receipt before the goods are receipt, while it is required to have invoices to be receipt after goods are receipt as in *3-way match, after GR*. Therefore, we could find some cases where invoices are received before goods receipt in *3-way match, before GR*. *In 3-way match, after GR*, however, we also could find few cases where invoices are receipt before goods receipt, which is not allowed in *3-way match, after GR*. Therefore, we assumed these cases to be outliers.
-In *2-way match*, "Create Purchase Order Item" does not always come first in *2-way match* (which should come first in usual cases). This is because most of the *spend areas* of items in *2-way match* are about real estate and the maintenance of a company, such as government payments, insurances and electricity. In case of the purchase of items related to the maintenance of a company, a company may have to pay without creating purchase order first, since the payment of these items are something compulsory, but not optional. However, the reason why "Create Purchase Order Item" comes first in case the *spend area* is real estate is not identified.
-"Change approval for purchase order" should come prior to "Delete purchase order item". This does not mean that "Delete purchase order item" should follow "Change approval for purchase order".
-In *3-way match, after GR*, there are a group of cases which has multiple "Record Goods Receipt" and "Record Service Entry Sheet" activities. If those cases are completed with the proper invoicing process activities, for instance, "Record invoice Receipt" and "Clear Invoice", there is a relationship between number of executed activities and *Cumulative net worth* (CNW). CNW in the "Record Invoice Receipt" is CNW of "Record Service Entry Sheet" multiplies the how many times the activity is executed. In case of the other invoicing activity, CNW of "Vendor creates invoice" and "Clear Invoice" are CNW of "Record Service Entry Sheet"
-When we take a look at the feature '(case) Company', "companyID_003" is in charge of "Real estate & maintenance system", which can only be found in *2-way match*. From the point of view of real estate related purchasing document, the values at the *Sub spend area text* are "Real estate brokers or agents"(27.4%) and "Real estate services"(25.58%). The values related with maintenance system are "Government payments"(23.87%), "Insurance"(3.43%), and "Escrow and title services"(3.06%), which can be considered as framework orders.

### 6.2 Findings from deviation analysis

-In answering the second question, we analyzed deviations of waiting time and found causes from the perspective of activity and instance level. Regarding the duration of each activity, waiting time of "Record Invoice Receipt" and "Clear Invoice" are observed to be longer than other activities for most of 13 process models. Unlike others, in case of *3-way match, after GR* with D.T.= framework order and I.T.= service, duration from "Create Purchase Oder Items" to "Vendor creates invoice" was the longest. Using these results, we recommend subsidiaries to keep track of these activities to decrease order leading time.
-In terms of case duration, *Sub spend area text*, *vendor*, and total *Cumulative net worth* factors were found to significantly affect the deviation in duration for each process model. As in Table 8 to Table 13, the list of factors shows which factors increase or decrease the case duration and how strongly or weakly they affect it. From the result, we were successful in finding the causes in attribute level and could make companies be aware of which factors in *Sub spend area text* or vendors make the ordering process longer. This finding will give an adequate solution for the problem of long order leading time in attribute levels.
-There exists a huge difference in distribution of features between regular and deviant cases. Almost all features are found to have heterogeneous distribution between regular and deviant cases.

## 7 Conclusion

The recommendation would be that the company must be able to clearly filter out deviant cases from event logs as not only deviant cases are found to have significantly different attributes compared to regular cases, but the difference in attributes may cause tremendous negative outcome to the company. Since the event log has handled different transactions between subsidiaries and 431 different vendors, it was challenging to make generic process model and give helpful results from the analysis for all subsidiaries. However, in this paper, we managed to provide generic results covering all participating subsidiaries and give solutions to process managers.

## References

1. van Dongen, B.F. *Dataset BPI Challenge 2019.* Centre for Research Data. https://doi.org/10.4121/uuid:d06aff4b-79f0-45e6-8ec8-e19730c248f1.
2. VAN DER AALST, Wil, et al. *Process mining manifesto.* In: International Conference on Business Process Management. Springer, Berlin, Heidelberg, 2011. p. 169-194.
3. DUMAS, Marlon, et al. *Fundamentals of business process management.* Heidelberg: Springer, 2013.