

Otto-von-Guericke-Universität Magdeburg

Faculty of Computer Science



Master's Thesis

Estimation on impact of BPM using Compative Analysis of Process Mining Tools

Author:

Akhila Mummidi

Feb 7, 2021

Supervisor:

Prof. Dr. Hans-Knud Arndt

Hannes Feuersenger, M. Sc.

Department of Technical and Business Information Systems

Mummidi, Akhila:

Estimation on impact of BPM using Comparative Analysis of Process Mining Tools
Master's Thesis, Otto-von-Guericke-Universität Magdeburg, 2020.

Abstract

The main goal of this thesis was to compare and analyze process mining tools based on a well-defined framework. In Business Process Management, the stages of life-cycle are analysed traditionally. So there is a need of improvement in analysis the process with grow in data. To make Business process digitally available, Process Mining comes to effect. Process Mining is a discipline which captures the process and gives an automated output with it's techniques. In this thesis, we described how process mining improves business process management. For better understanding, Process mining tools are compared and analysed. The tools are selected towards the focus of students, so we preferred tools having academic licence. Finally the evaluation results are made on the analysis of all the tools behaviour.

Acknowledgements

I want to take this opportunity to thank my supervisor Dr. Ing. Sandro Schulze for supporting me throughout this thesis.

I am always thankful to my family for supporting & encouraging me to achieve my goals.

Thanks to my friends for helping me in many ways during my Master's course.

Contents

List of Figures	ix
List of Tables	xi
List of Code Listings	xiii
1 Introduction and Motivation	1
1.1 Motivation	1
1.2 Aim of this thesis	2
1.3 Approach	2
1.4 Outline of this thesis	2
2 Background: A an overview of Process Mining in BPM	5
2.1 Introduction to Business Process Management	5
2.2 Process Mining	7
2.2.1 Event logs format	7
2.2.2 Process Mining techniques	8
2.2.2.1 Process discovery:	8
2.2.2.2 Conformance checking:	9
2.2.2.3 Enhancement:	9
2.3 Tool Selection	10
2.4 Process Mining tools - Academic license	10
2.4.1 Disco	10
2.4.2 Celonis	12
2.4.3 myInvenio	13
2.4.4 Minit	14
2.4.5 Rapid Miner	15
3 COMPARATIVE ANALYSIS OF PROCESS MINING TOOLS	17
3.1 Use case 1: Loan Approval data	17
3.1.1 Process Discovery	18
3.1.1.1 Disco	19
3.1.1.2 Celonis	19
3.1.1.3 myInvenio	21
3.1.1.4 Minit	22
3.1.2 Conformance Checks	23
3.1.2.1 Disco	25
3.1.2.2 Celonis	25

3.1.2.3	myInvenio	26
3.1.2.4	Minit	26
3.1.3	Enhancement	26
3.1.4	Conclusion	27
3.2	Use-case 2: Purchasing data	28
3.2.1	Process Discovery	28
3.2.1.1	Disco	29
3.2.1.2	Celonis	29
3.2.1.3	myInvenio	33
3.2.1.4	Minit	36
3.2.2	Conformance Checks	37
3.2.2.1	Disco	37
3.2.2.2	Celonis	37
3.2.2.3	myInvenio	39
3.2.2.4	Minit	39
3.2.3	Enhancement	39
3.2.4	Conclusion	40
3.3	Data pre-processing	40
3.3.0.1	Disco	40
3.3.0.2	Celonis	41
3.3.0.3	myInvenio	41
3.3.0.4	Minit	41
3.3.1	Conclusion	46
4	Evaluation	47
4.1	Evaluation Criteria	47
4.1.1	Testing Setup	47
4.1.1.1	Software	47
4.1.1.2	Hardware	48
4.1.2	Evaluation Results	48
4.1.2.1	Discussion on Research question 1	48
4.1.2.2	Discussion on Research question 2	48
4.1.2.3	Discussion on Research question 3	49
4.1.2.4	Discussion on Research question 4	49
4.1.2.5	Discussion on Research question 5	49
5	Conclusion and Future Work	51
	Bibliography	53

List of Figures

1.1	Approach of process mining tool comparision	3
2.1	Business Process Management Life-cycle	6
2.2	Process Mining	8
2.3	Example of Conformance Check	9
2.4	Data import in Disco	12
2.5	Minit start page	14
2.6	Minit start page	15
3.1	Loan Approval use case sample data	18
3.2	Process map in Disco for Loan Approval process	19
3.3	Frequency in Disco for Loan Approval process	20
3.4	Automatic process discovery in Celonis for the Loan Approval process	20
3.5	Performance analysis in Celonis for Loan Approval process	21
3.6	22
3.7	Analytics in myInvenio for Loan Approval process	23
3.8	Automatic process discovery in Mint for the Loan Approval process .	23
3.9	Performance analysis in Minit for Loan Approval process	24
3.10	Loan Approval Process bpmn	24
3.11	Conformance Check description in Celonis for Loan Approval processes	25
3.12	ConformanceCheck in myInvenio for the Loan Approval process . . .	26
3.13	Conformance Check description in myInvenio for Loan Approval process	27
3.14	Conformance Check in Minit for Loan Approval process	28
3.15	Resource based analysis in Disco for pursaching data	29
3.16	Role based analysis in Disco for pursaching data	29

3.17	Resource tab indication in Celonis	30
3.18	Data selection for resource analysis in Celonis	31
3.19	Overview of data in Celonis for Purchasing data	31
3.20	Throughput time in Celonis for Purchasing data	32
3.21	Social activities application in Celonis for Purchasing data	32
3.22	Social network for specific in Celonis for Purchasing data	33
3.23	Process Dashboard in myInvenio for Purchasing data	34
3.24	Process map in myInvenio for Purchasing data	34
3.25	Rework in myInvenio for Purchasing data	35
3.26	Time analysis in myInvenio for Purchasing data	35
3.27	Resource analysis in myInvenio for Purchasing data	35
3.28	Role analysis in myInvenio for Purchasing data	36
3.29	Activity map in myInvenio for Purchasing data	37
3.30	Social activities application in Celonis for Purchasing data	37
3.31	Social network drill down in myInvenio for Purchasing data	38
3.32	pics/Role based analysis in Minit for Purchasing data	38
3.33	Conformance Check in myInvenio for Purchasing data	39
3.34	Conformance Check description in myInvenio for Purchasing data	40
3.35	Conformance Check in Minit for Purchasing data	41
3.36	Event log identification in Disco	42
3.37	Event collection in Celonis	42
3.38	Event log collection in Celonis	43
3.39	Event log collection in myInvenio	43
3.40	Minit start page	44
3.41	Minit Data Import Step 1	44
3.42	Minit Data Import Step 2	45
3.43	Minit Data Import Step 3	45
3.44	Minit Data Import Step 4	46

List of Tables

4.1	Process Mining techniques accepted by the tools in academic license.	49
4.2	Additional features in the tools for academic license.	49
4.3	Inputs provided by the vendor for academic license of the tool.	50

List of Code Listings

1. Introduction and Motivation

Business Process the executives increases the value of the association with mix of various techniques to have a solid business measures. Due to hefty rivalry and development in the organizations, the business measures became complex which needs persistent observing. In addition, improving interior business measures there is requirement for embracing the adjustments in the outer climate to have a business measures which is consistence with outside climate. Thus, to have a total standard-ised business measure and to guarantee the consumer loyalty the idea of BPM assumes key part from an association point of view. A definitive objective of BPM is to have productive and successful and efficient business measures. The business cycles can be completely programmed or automated cycle requires human mediation. In the current time of mechanization all the organizations are managing colossal measure of information which through difficulties to the administration in checking the business cycle where process mining assumes key part in mining and examination the information.

1.1 Motivation

As we know that the technology is the key and integral part of every business organization and all the businesses are driven by technology to handle voluminous transactions. The product or services which are offered by the organizations are the result of different combination of activities. It's very crucial in performance evaluation to know what is actually happening, it's easy when the business is small but when Imagine a business with lot of complexities and business rules I wondered how those organizations managed to implement all the processes in the business with lot of technology involvement. Here, I come across the combination of Business process management and Process mining. The former talks about the traditional concepts to manage the business processes and later gives us the overview of what is actually happening in the business, when we implement this combination in complex businesses we can do better performance management and achieve desired results. Finally, these two concepts motivated me to do the analysis by using real-life data with the help of various process mining tools and to project the importance of process mining in management of business.

1.2 Aim of this thesis

The main goal of thesis is to showcase the impact of process mining in BPM and differentiate the various process mining tools by its functionalities for students. This defines the approach for choosing the correct academic version process mining tools, among arising tools, that can find a accord the hidden business use case for analysis.

Research Questions

Research Question 1 : *How process mining improves Business Process Management?*

Research Question 2 : *What are the approaches of different process mining tools for pre-processing of data?*

Research Question 3 : *What is the behaviour of the tools for process mining techniques?*

Research Question 4 : *What are the additional features provided by process mining tools*

Research Question 5 : *What are the limitation of the inputs for the tool in academic license and What is the students adaptive learning environment of the process mining tools?*

1.3 Approach

To achieve aim of thesis we do perform exercise on the real life data sets, analyse and understand the various results. Here, we do plan as [Figure 1.1](#) to perform Process Discovery, Conformance, Process Redesign on the different type of data sets and projects the findings based on evaluation frame work. We selected three tools for this purpose based on a careful evaluation of prescribed criteria. This differentiation of various tools helps the practitioners and various users to understand the strengths and weakness of each and tool and also able to get which tool more compatible to perform 1) Process Discovery 2) Conformance 3) Process re-designs.

1.4 Outline of this thesis

This research is categorised into five chapters. We followed our research as follows:

In [Chapter 1](#), Chapter 1 covers the introduction part and Short description about the importance of Business process management and we discuss in short about the use of process mining in complex business environment. In this chapter, we also discuss about the motivation,aim and research questions. We describe about Business process management with growing data. How are the effects of process management to many industries. We addressed the solution as Process Mining technology. In this motivation, approach, and research questions are described.

[Chapter 2](#) We will have detailed discussion about the BPM life cycle and process mining. Here, we also discuss about the importance of event log in process mining

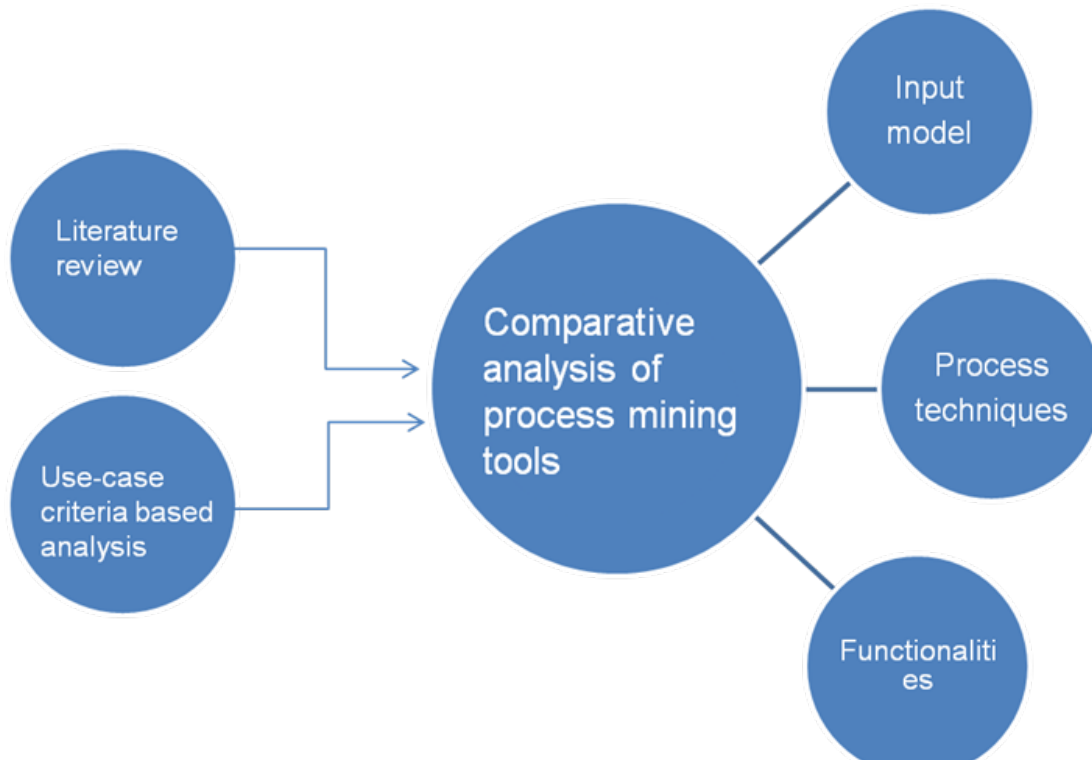


Figure 1.1: Approach of process mining tool comparison

and will have discussion on tool selection methodology. Finally, we will discuss briefly about the selected tools. Aims in describing the knowledge of Business Process Management life cycle. Recollecting tradition techniques followed in the life cycle to analyse process. Followed by describing the concept of automated technology called Process mining to make the process digitized. of various concepts that are involved in our thesis. Also include process mining techniques, process mining tools, and evaluation of selecting the tools.

Chapter 3, illustrates how our analysis is followed in different tools used for this thesis. We used use-case based criteria to analyse and the functionalities and features of the tool.

Chapter 4 aims in evaluating our analysis and compares with each other process mining tools in relating to the described research questions.

Finally, Chapter 5 depicts the conclusion of our thesis, with respect to impact of process mining in business process management. Also illustrates the about different process mining tools, its special features and limitations.

2. Background: A an overview of Process Mining in BPM

2.1 Introduction to Business Process Management

In traditional Business Process Management [1], The business process re-engineering had some important implications for how business processes are supposed to be managed. It became clear that first and foremost it is important to establish process thinking in the organization. This life cycle model is shown on this Figure 2.1 referencing from the [20].

It distinguishes six major management activities for business processes. At the top there is process identification. It is concerned with describing the overall process organization in terms of a process architecture. An individual process can be made subject to the cycle at the lower part of this figure. First process discovery comes up with an as is process model that describes the process how it currently is. This is the basis for process analysis that tries to identify weakness and the impact on the performance of the process. With these insights the process can be redesigned, the output is a to b process model. This to be process model can be implemented. Implementation here means that first we provide the information systems infrastructure for running the process. Second, we educate people working in the process. How the process is supposed to operate .when individual cases are then executed according to the process they can be monitored. Monitoring provides us insights into the conformance and performance of the process. If conformance and performance are not According to what we desire. We can continue again with another circle of process discovery analysis and design in implementation.

In the process there are various types of methods to observe the results of process life cycle. In this thesis, we mainly focus on Process analysts who are responsible for process discovery , conformance checks, process analysis and redesign of process life cycle . We illustrate all the traditional methods to analyse every activity of the process as follows.

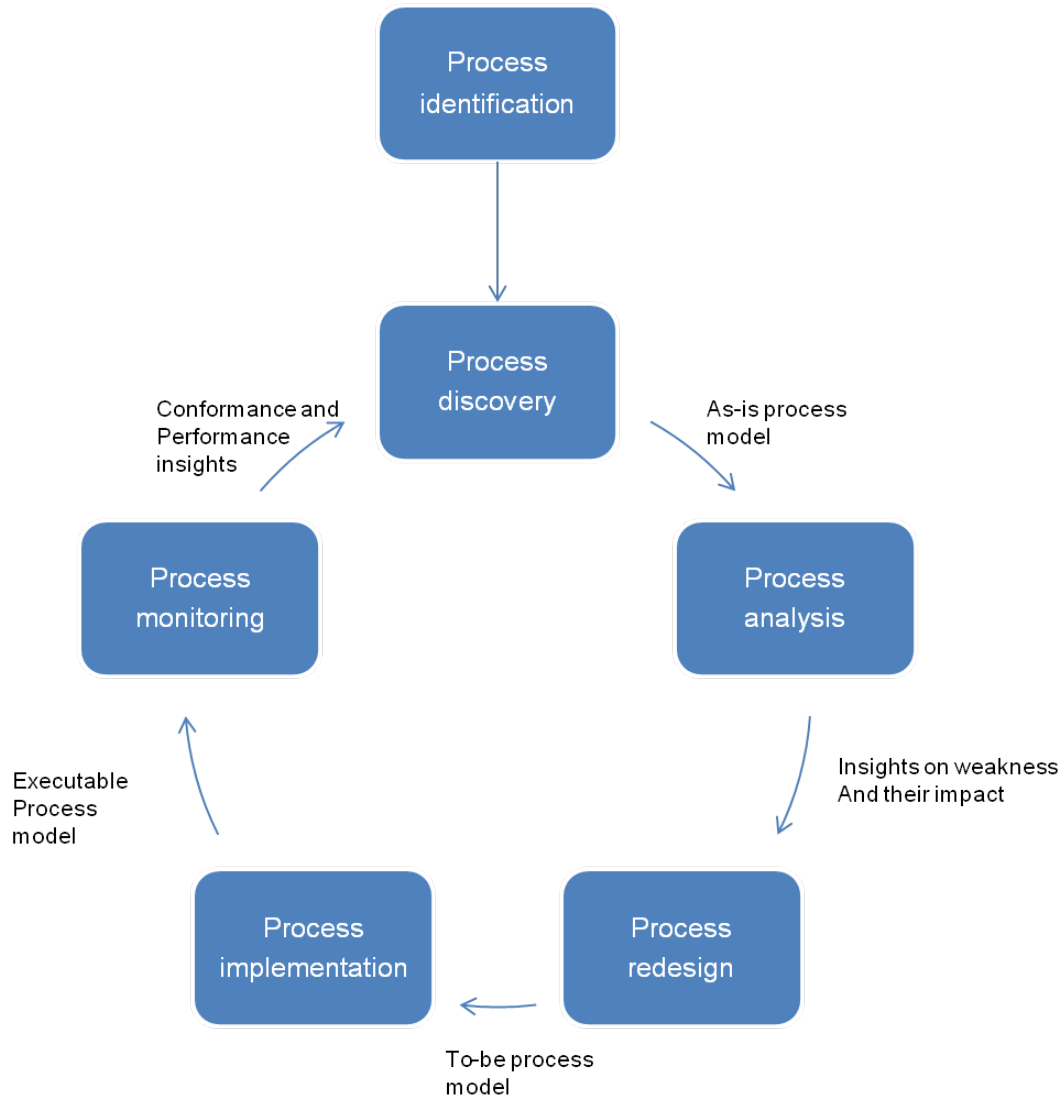


Figure 2.1: Business Process Management Life-cycle

In Process Discovery [12] [26], the one of the method is 'Document Analysis' which includes process descriptions, internal charts, organizational charts etc. In this process analyst should be ready with gather information before approaching the domain experts. It causes in refinement in the process and may not be process-oriented and trustworthy. Another method is 'Observation' which needs potential intrusive, in observing different stakeholders behave differently and as a human only few cases can be observed. Another approach to analyse process discovery is 'Interview' and 'Workshops' requires spare time of process stakeholders and requires availability of several stakeholders at the same time. In Process analysis, as an input we will have As-is process model available and as an output we will have potentially, insights on weakness and their impacts. In traditional process analysis, it consumes time to find out waiting times, delays, rework of the process.

So to adopt improvements in extensive set of cases and for objective data [19] [21], we can also discover our process by means of automated process discovery. It is technique which belongs to the family of Process Mining. The starting points are

event streams and the information stored in databases. The information systems that supports a processes store this information [17]. This will organise them in a so called event log. An event log captures, how these different cases have been processed. This information can be used to automatically discover a bpmn process model. This process model accurately represents how our process has been operating, in terms of what is visible to our information systems. Process mining offers also additional techniques of analysis. We can also have an existing process model and by means of conformance checking we can find out what are the process deviations. And we can also enhance our existing process model and get information about frequencies.

2.2 Process Mining

Process mining is a discipline which helps to monitor the performance related information with help of process models [24] [18]. The key input to these different types of process monitoring support are 'event logs'. Event logs capture data of processes that have been executed as much as they are represented in databases and in event streams. Process mining is the technology that is concerned with the techniques that can extract process information from event logs. We use event logs to automatically discover a business process model. The discovered model helps us to understand, what the data actually tells us about the different cases in terms of a model. This discovered process model is an As-is process model that captures those parts of the process that are stored in information systems. A second family of process mining techniques is 'Conformance checking'. Here we do not only take event log as input but we also take a normative process model as an input that we had created earlier. The goal of conformance checking is to understand to which degree this input process model is in line with the behaviour captured by the process mining. Also conformance checking also points us to particular cases that are not compliant with this input model. It also helps in checking the deviation of the process with the actual process. Finally we can redesign the process model by enhancement technique. It also takes an input model and an event log as input. As an output we enrich the input process model with statistical information that we find in the event log.

2.2.1 Event logs format

Often we have event logs [9] [22] available as a table. An event log contains all recorded events that relate to executed activities in a table. "A process model is an abstraction of the real world execution of a process. The event logs as a set of events that are mapped to the same case. The sequence of recorded events in a case is called trace. Process instance model describes the execution of a single process instance. A process model abstracts from the single behavior of process instances and provides a model that reflects the behavior of all instances that belong to the same process. Classifiers ensure the distinctness of cases and events by mapping unique names to each case and event. Cases and events are characterized by classifiers and attributes" [6]. There are some minimum requirements in event log data table as in and is structured as follows.

- Case id

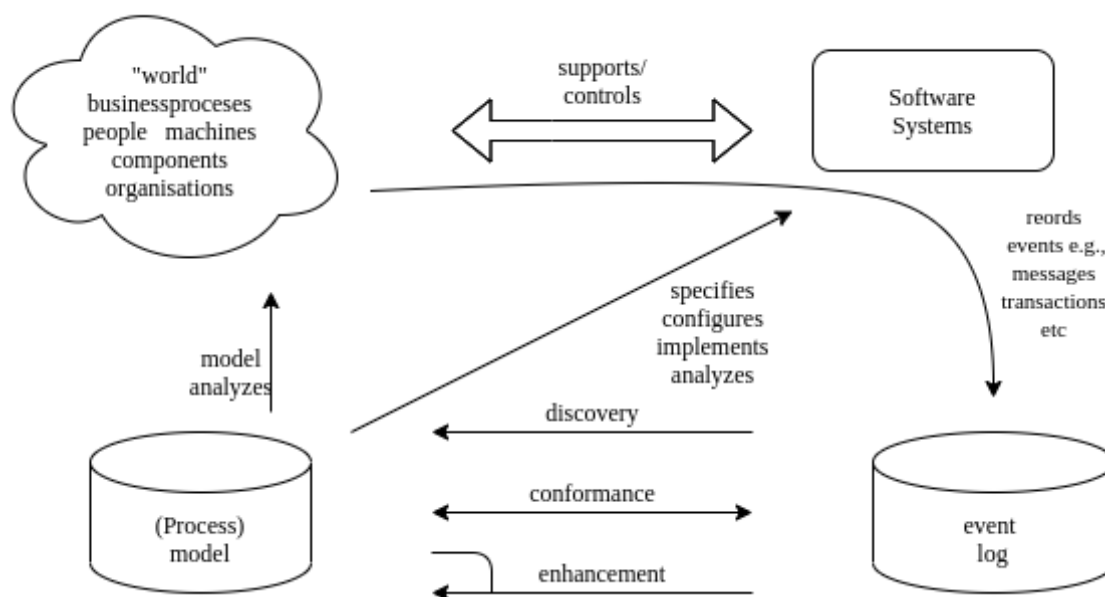


Figure 2.2: Process Mining

- Activity:
- Timestamp

Process mining [2] is a tool which takes the event log from the organization IT system and do some exercise on that logs by using various algorithms and give us the complete picture about the process actually happened and also give us various deviations throughout the process.

2.2.2 Process Mining techniques

2.2.2.1 Process discovery:

On the basis of event log this technique give us the visual picture of actual process flow happened over the period of time in a diagrammatic way. This helps the management and auditors to walk through the actual process thoroughly unlike traditional method of modelling the business process where lot of human interaction and time consumed. In BPM, understanding of the actual process is the key triggering point for designing the future plan which can be possible with this discovery technique [11] [23]. [3]

For example, a financial service provider running the business from long back where the process is drafted according to the situations prevailing at that time. At present, the organization wants to change the business process which will be consistent with present business environment for which it needs thorough analysis of the existing process and it can be possible in less time with the help of this process discovery technique.

2.2.2.2 Conformance checking:

It's another technique which mine the data recorded in event log and compares with modelled process and give us the deviations occurred in the process [16] [7]. With the help of this conformance technique the management of the organization can know whether guidelines issued for particular process are followed accurately. Because, violations to standard guidelines need careful consideration in the monitoring phase of BPM. This technique can be also used by auditors to know the degree of violation in risk assessment activity.

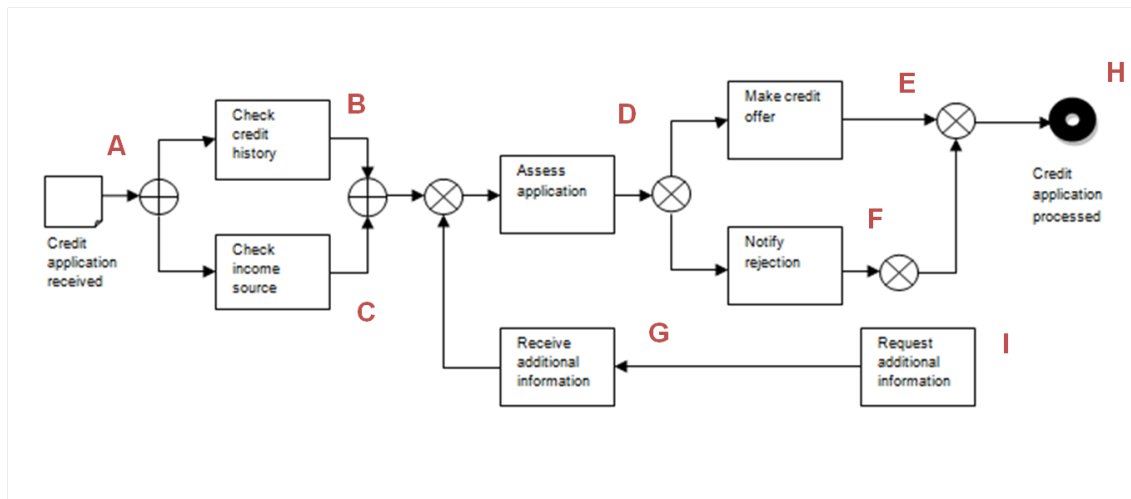


Figure 2.3: Example of Conformance Check

For example, we use credit application, The process starts with the event credit application received. The letter A indicated an abbreviation with which we can refer to events that correspond to this event as shown in Figure 2.3. Then there is an end split, activities check credit history and check income sources to be conducted letters B and C refer to these activities. After the end join there is an xor join, it is closing of a loop that means we continue the process with the activity assess application. The letter D refers to this activity. There is an xor split to either ground the application or to notify rejection. The rejection is referred to letter F. If the application is granted we make a credit offer. The letter E refer to this activity. If we make the offer the credit application is processed, refers to letter H. If we notify about the rejection there is a choice, refers to letter I or G. In return to a new assessment of the application. The corresponding event logs is as follows, we observe the following sequences: ABCDEH, ACBDEH, ABCDFH, ACBDFH, ABDEH, ABDFH. In the last two event logs C is missing. Conformance checking now provides techniques to identifies the missing activities. It also highlights that there is a part in the process model that is apparently never observed like letter G and I in our example. So conformance check helps the analysts to compare the process models that are used and defined as a norm for the process in relation to the event logs that have been observed.

2.2.2.3 Enhancement:

Based on the bottlenecks and deviations identified in conformance checking, the tool suggests the possible course of actions to be taken. As mentioned earlier, we

can execute the various phases mining with the help of process mining tool. "The extension or improvement of an existing process model using information about the actual process recorded in some event log [27]". The results of process discovery can be used at design phase to analyse the existing process thoroughly. The results of conformance checking and redesign can be used in monitoring and redesign phases of BPM. The process mining tool makes BPM professional's job easy and can take timely and structured decisions to standardise the process.

2.3 Tool Selection

As process mining is a new upcoming technology, there are many tools emerging in today's practice. To ensure reality, survey is done on tool websites, journals, literature review [13] [19] to gather all the available tools in the market. Tools selection is more focused towards students. So a first step, selection process is done as:

- selecting the vendors providing Academic alliances for students and also based on tool and based on the availability in Europe.
- We picked up the process mining tools based on four categories i.e., simple, moderate, advanced respectively.

A list of 10 tools are obtained. Further refinement is done as follows. PM4Py is an open-source Python library that which is difficult to use by business people. Similarly, we excluded Apromore. The ProM framework offers a comprehensible library of scientific techniques and algorithms but it is not illustrated in the thesis because it is already researched by many people. On top of that, there are many tools with ease. Finally 5 tools are tested in this thesis i.e., Disco, myInvenio, Celonis, Minit, RapidMiner.

2.4 Process Mining tools - Academic license

2.4.1 Disco

Disco is the simple process mining tool having very limited features when compared to other tools of process mining. In the academic license of "DISCO"¹, we can import only up to 5 million events. When we logged on to the "DISCO" for the first time we only get two options which are "Play in the sandbox" and "load Your Own Data". In "Play on the Sandbox", we can explore a demo of one example event log which is inbuilt in the tool. In the second option, we can import our own data to do our analysis and as said earlier, we can import only 5 million events and we can only import the following forms of input wide CSV, XLS, MXML, XES and FXL Figure 2.4.

As mentioned this tool is very simple and having limited features. The features of the disco are suitably grouped only under process discovery technique., three main parts of the tool are 1) Map 2) Statistics 3) Cases. The sub features of each part are discussed below:

¹<https://fluxicon.com/>

1. Map: In this part, we get complete picture of the actual process will be generated based on the given event log.
 - Frequency: Frequency is nothing but what to be projected on the path, whether to show case duration or mean time. The special feature of disco is that we can apply two frequencies at a time which is at the discretion of the process owner that how the process owner designed to see the paths.
 - Filters: Along with the two basic filters which are activity and path, disco is providing another 3 filters explained below:
 - A) Variation filter: Its known fact that the actual process will be scattered in different routes for different due to the reasons and the common reasons are employee behaviour, technical failures or lack of knowledge etc., . This filter gives us the option to filter the most frequent and exceptional route. Route is called technically as a variant.
 - B) Endpoint filter: When we are doing the analysis of the process we gain the understating of the modelled process and from which we can know what are the start and end events should be. Hence, by using this endpoint filter we can remove the cases from the analysis which do not have desired start and end events. This will give only numerical results that how many cases does not have desired start and end events, but it will not provide us the case ids which are violated the desired rule.
 - C) Performance filter: In this filter, we can filter cases by duration like we can filter long, short, medium duration cases.

Note: The above all filters only give us the statistical results like how much percentage of cases falls under particular filter and the discovery map will be updated according to the selected filters.
2. Statistics: Second part of this tool is statistics; statics part is having two features which are Overview and Activity.
 - Overview: This feature provides statistics and variety of charts of various categories which are Events over time, Active cases over time, Case variants, Events per case, Case duration. This statics only gives a various charts for observation but no drill down is possible to do in detail analysis. It also gives two types of tables which are Cases and Variants. The former gives, Case id wise detailed statics like events per case and in which variant it has grouped and duration of each case and the later gives, variant wise statistics like how many cases and how many events falls under each variant, mean time of each variant.
 - Activity: This part gives only the numerical data of each activity, number cases falls under each activity. Statics part is just gives an overview in various chart forms but it will not provide us any option to go in detail and do the analysis.
3. Cases: This part provides an option to see each case in detail like we can select case number and explore that case in detail in all possible ways.

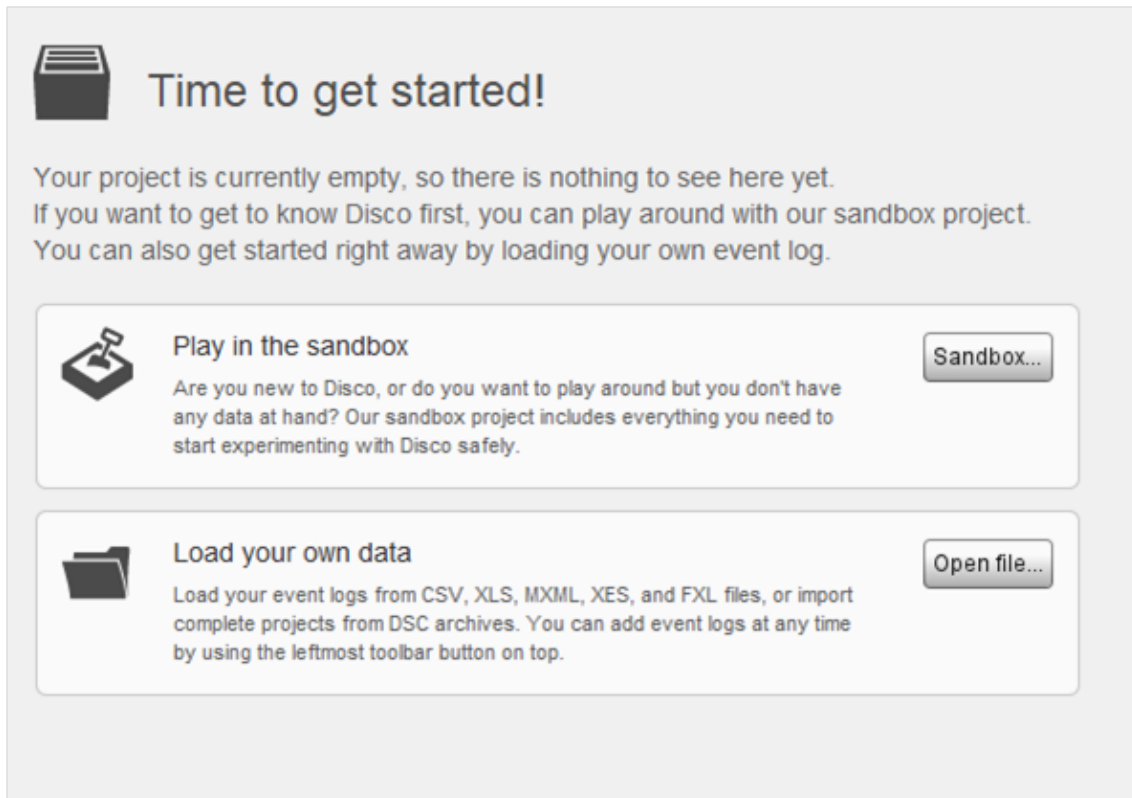


Figure 2.4: Data import in Disco

2.4.2 Celonis

Celonis is a cloud based process mining tool. Celonis provides Academic Edition license for students ². Celonis is one of the trending and pioneer software in the era of cycle mining [5].

1. Data Input: The data input process in Celonis is different from the other software. The pre-processing of data in Celonis is explained in Section 3.3. The initial step, Will have to logon to Event collection option and create data pool by uploading the data and it will take all kind of input wise .xes, .csv, .xlsx etc., and while creating data pool it will ask only case id, activity, start time, end time as the key criteria for pre-processing of data, having the option to detect and correct the errors in input data. Further step, logon to process explorer option and create a work space for our analysis by selecting the relevant data pool and another good feature of this software is we can create multiple analyses in a single work space with single data pool. Finally, the tool is ready to do the analysis.
2. The features of this web based tool are named as apps in the tool. The tool is having wide varieties of applications which are:
 - Process AI: It give us the most common path generated from the input and also gives the basic KPI like the count of activities and events in

²<https://lms.celonis.com//catalog>

such most common path and the option is provided to edit/add KPI for such common path. It also gives us the deviations in such most common path. This can be helpful when we are not having proper model to do conformance checking, we can analyse the most frequent paths and deviations from such path and take appropriate actions.

- Process overview: It is nothing but a glance view of the total data set we gave. It gives wide variety of options like cases and events per day, average throughput time, happy path nothing but most frequent path and cases count in that happy path out of total cases. Average cycle times can be configured in various forms like days, hours, and minutes. Process overview can be seen for a particular time period which can be done as per the requirement of the user. The overview can be seen in graphical format with trend analysis.
- Process Explorer: This is the key feature which will assist in discovery phase of the Process mining.
- Conformance: This is another key feature which will assist in conformance phase of the process mining. e) Social: This particular feature is to analyse the behavioural aspects of human personnel in the organization like the relationship between different personnel and how the data is flown among those persons. KPI per each person like cases per person, average cycle time per person to check the efficiency of the each person.
- Case explorer: This is another considerable feature. In case explorer feature we can explore the particular case id in detail like the number of activities, name of those activities, and total cycle time of that particular case can be seen. This will be helpful when we are dealing with huge data and wanted to explore a particular case.
- Besides above, we can add new feature/applications which are activity explorer, different varieties of charts and graphs.

2.4.3 myInvenio

About tool - This tool offers Academic Alliance Agreement ³ for students. myInvenio provides a cloud based platform for 1 year with limited features. This process mining tool allows students to upload up to 2GB of data in raw or compressed (zip, gz) CSV or XES files format. Data collection in tool - The tool is legible for newbie. Tool allows students to create and analysis only one process at a time. Creating a new process in the tool is with ease. As described in the figure Process title and process organisation A preview of the uploaded data will be displayed below. Creating a new process in the tool is with ease.

This tool can be explained in various sub features as follows:

- Dashboard: Here, we can have the bird's-eye view of the all cases, activities, resources, role, in tabular and graphical forms.

³<https://www.my-invenio.com/myinvenio-academic-alliance>

- Process Map: After the dashboard, we can peep into the model tab to view pictorial form of the process. Here, we can see the most frequent activities quickly due to the unique feature of highlighting the frequent activities in the process map.
- Rework identification: This is the special feature of MyInvenio it highlights the rework activities in any kind of processes irrespective of its size.
- Analytics: This part of the tool helps in observing the statistics and different variants process of the data.

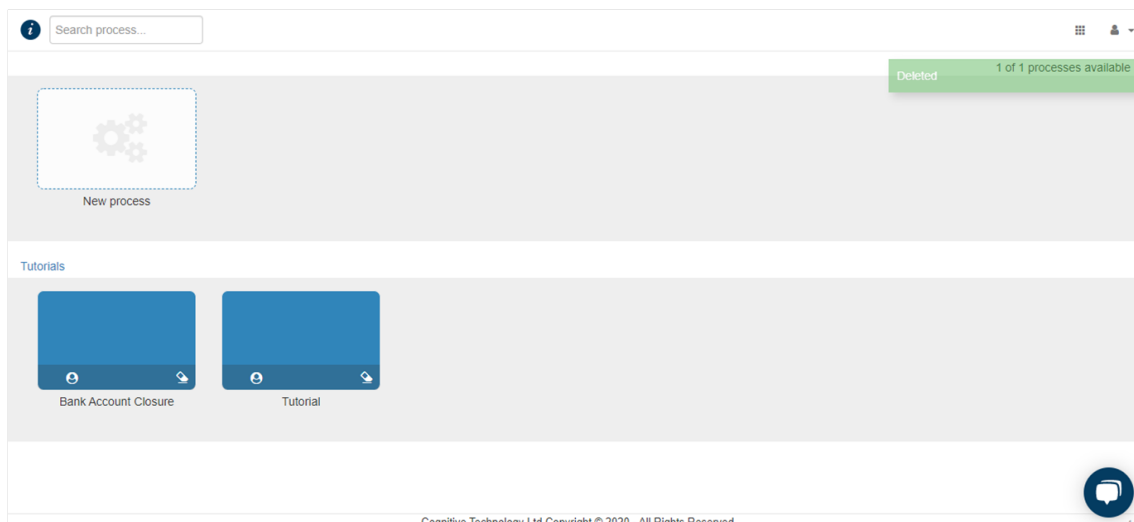


Figure 2.5: Minit start page

2.4.4 Minit

In Minit, to get a free academic license tool ⁴, we need to request a trial tool from the vendor website as an academic partner. Vendors support team contacts with the set of questions through an email. By following the steps, Minit standalone edition software download link, user guide and license number is provided by the vendor. If you run the application for the first time, a configuration wizard will guide you through the application setup. After license activation, activation number and features limitations are displayed. As a first steps, pre-imported process are stored in the tool as a demo. To get ease with the tool and its features, Minit created a learning platform, a centralized hub to find all the relevant information about the tool and insights on process mining Minit solution itself. To start the mining academy goes to lms.minit.io to register, click on the sign up button. All the created new account should be approved by the administrator. For student license it provides a set of 9 tutorials. On the home screen, you will find a couple of preselected training courses. The training content is categorized into groups search has first steps, advanced training, tricks and trips training, structure training, library contains all the webinars and other useful content. The new content is regularly updated in the course catalog section. Enroll the course to start with. Learning

⁴<https://lms.minit.io/>

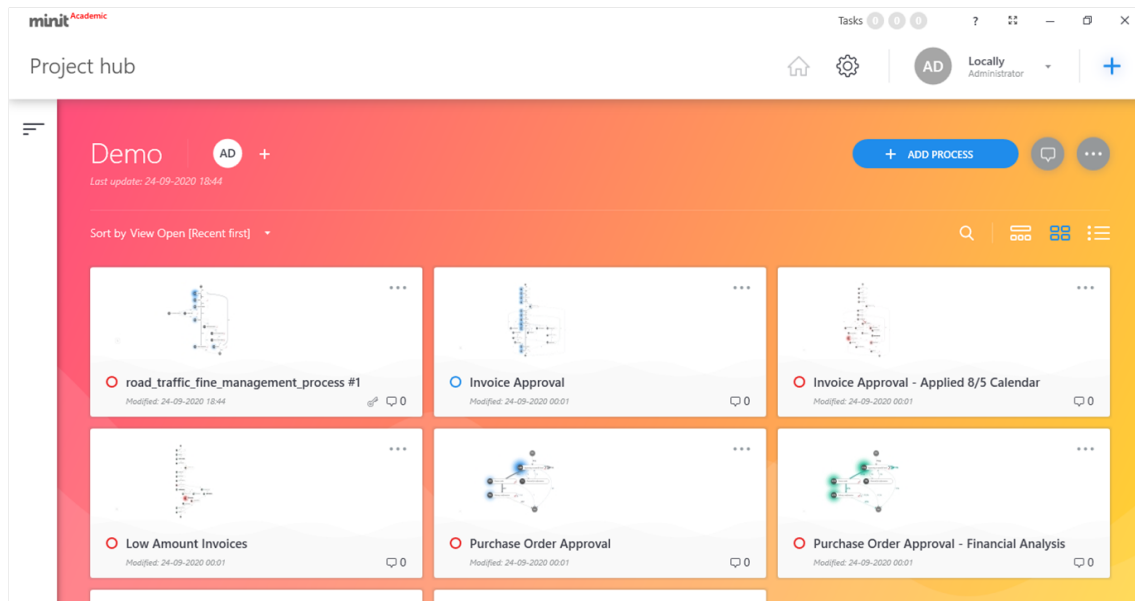


Figure 2.6: Minit start page

process, certifications etc can be checked. Badges and points are attractive for a learner to fulfill the course. There is a discussion forum for quick clarifications in the course, concept, features of the tool etc. Data Imports, data sets, types, attributes In Minit - we will start with project hub screen, which organizes your projects and processes in the project. Switch on the list on the left side. Create your newly created project. The features of the tool are as follows:

- Process Map: It gives a pictorial form of process map.
- Process Animation: Helps in understanding the flow of the process.
- Charts: Here, we can create our own KPI's.
- Statistics: This helps to understand the number of events in the cases.
- Variants: Variants show cases different paths in the process to analysis unusual process in the data.
- Filters: This helps in filtering the data for deep through analysis. All the features and functionalities are used in the analysis [Chapter 3](#).

2.4.5 Rapid Miner

Rapid Miner is an highly demanded data science tool. : Its conventional tool of data mining in which we can perform various analysis like decision tree analysis, chart based analysis in the form various types of charts. Coming to process mining, RapidMiner does not have any inbuilt feature to perform process mining techniques. To do process mining, we should download an extension called “RapidProm” from

the “RapidMiner Marketplace” available in the “Menu Tab” of the tool. After downloading this extension, we can see various functionalities like Data discovery, Conformance, Performance analysis and this extension is also having the functionality called “Conversion” and this helps in converting the “Petri Net” to “.Bpmn”.

In our thesis, we tried to perform process mining in RapidMiner but the system was not supportive enough to perform the analysis. Based on literature review, it’s observed that unlike other process mining tools this tool having various algorithms in “Data discovery” functionality and various algorithms generate process map in different models. Unlike other tools, in “Data discovery” we can’t see the KPIs like case frequency and time on the process map and to see those KPIs we need to go to “performance analysis” functionality and follow the required step. In our observation, we noticed that the process mining feature in “RapidMiner” is having comprehensive functionalities but not more user friendly unlike other tools.

3. COMPARATIVE ANALYSIS OF PROCESS MINING TOOLS

In this current chapter, we would highly observe the performance of different process mining techniques applied to different process mining software tools [25] [10]. [17] [14] [4] [15] [8]

Going detailed in this chapter, we analyzed all the tools using different sample use-cases taken from different sectors of business. This helps to understand process mining can be used in all kinds of organizations having different variations of data. Each use-case is described in the below sub-sections to understand the application of process mining in organizations.

In different use cases, all process mining techniques are applied in the following use cases, which as a different format. This can be identified all the functionalities of the tools Disco, Celonis, myInvenio, Minit, and Rapid Miner respectively.

The observation is made on how software tool identifies bottlenecks, process improvements, process variants, time management, cost, resources, process discovery, conformance checks. The whole process of this analysis emphasised mainly on how process mining adds value to the business process management in the organization.

3.1 Use case 1: Loan Approval data

The loan approval process is one amongst the more sensitive areas of banking and therefore in the Finance sector. Because of the pursuit to make efficient processes and achieve an optimized resource consumption no matter a website, the explore for a qualitative product application is becoming vital. The business process of loan approval is an integral a part of a bank or institution. Each business process also includes a particular level of risk related to credit scoring. supported a literature review, the bank distinguishes between two groups of credit: Approval and rejection

of a loan, which involves a risk assessment process. Before the loan granting, banks verify the individual user who would like to require out a loan. Once a loan is approved by a bank, an agreement between the bank and a customer is ready, under terms which vary among banks, yet as countries.

The use cases presented in this thesis will be based on demo event logs available at [url:?](#)]. This use case Figure 3.1 will present simplified version of the process to show effective application of process mining techniques and the limits of the tools presented in this thesis.

Case ID	Activity	Start Time
Case_1	Request for a loan	01-08-2015 21:11
Case_1	Approve clients for automatic credit check	02-08-2015 00:52
Case_1	Issue a loan	02-08-2015 23:04
Case_1	Automated credit check	02-08-2015 05:21
Case_10	Request for a loan	01-08-2015 22:33
Case_10	Approve clients for automatic credit check	02-08-2015 11:48
Case_10	Automated credit check	03-08-2015 02:15
Case_10	Issue a loan	04-08-2015 11:13
Case_100	Request for a loan	01-08-2015 17:31

Figure 3.1: Loan Approval use case sample data

In analysis area of our use case example, we will be using data sets represents . We will define a problem statement to identify opportunities for process improvements by revealing process deviations and bottlenecks. We should also know the potential cost and time inefficiencies or wastes, violation of duties or frauds that we are currently experiencing.

3.1.1 Process Discovery

This part of the section depicts about how tools represent process model for loan approval use case. Observing the scope of improvement in the use case and differentiating the approach in different tools. As time is important factor in processing a loan. And in a delay in the process may occur chance to loose customers as they look for quick approaches. So based on time analysis on use-case, focuses some of the most needed areas for improvements such as average time taken for a loan application.

The tools accepts the CSV format file. With few data preparation steps, the tool show process discovery technique. For user readability, the number of paths and activities in the use case are reduced. So tool show 8 activities with total of 2099 cases and 18.2k events.

Objective:

- Process Discovery technique to Loan approval use case.

Expectations:

- The tools should show process discovery mode, with start and end activities.
- identify opportunities for process improvements by revealing process deviations and bottlenecks. We should also know the potential time inefficiencies.

3.1.1.1 Disco

In Disco, after reprocessing of data, the process map and process improvements are observed by using three major parts of the tool i.e., Map, Statistics, Cases. These are explained using loan approval use-case and it's sub features of each part are discussed below: According to our requirement, we can adjust the process map by applying the activity and path filters to the extent it's required for analysis. Since, the data set is having minimal cases, we kept activity and path filters to the maximum and the map is shown in the [Figure 3.2](#)

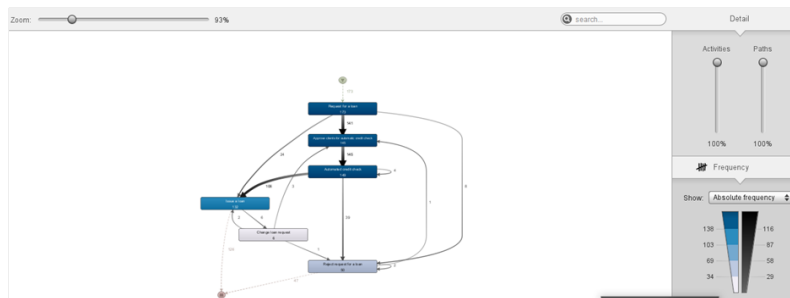


Figure 3.2: Process map in Disco for Loan Approval process

As part of my analysis, I kept two frequencies which are case frequency and mean duration. Case frequency is nothing but how many cases are flowed between the selected path and mean duration is nothing but average time taken for each case between the selected paths. These two frequencies are reasonably required as we are doing analysis of loan approval process. Because of these dual frequency settings, we can discover the unusual paths in terms of time and number of cases. In our data, the unusual path is request for loan to issue loan which takes around 4 days and 24 cases are flowed along this path which is unusual and need careful consideration. The maximum time taking path is automated credit check to issue loan which is 26.7 hours and substantial number of cases are flowing through this path which needs a consideration by top level management that whether the mean time reduced to the extent possible so that customers are attracted for speedy process as shown in [Figure 3.3](#).

3.1.1.2 Celonis

The various features of the process discovery in the Celonis are, activity filter, connection filter (Path filter). For user-friendly analysis can keep connection filter to the minimum. Another feature is grouping the activities and removal of the activities which are not relevant for particular analysis. The data set under consideration is loan approval process in an organisation which consists of 173 Cases and 655 events

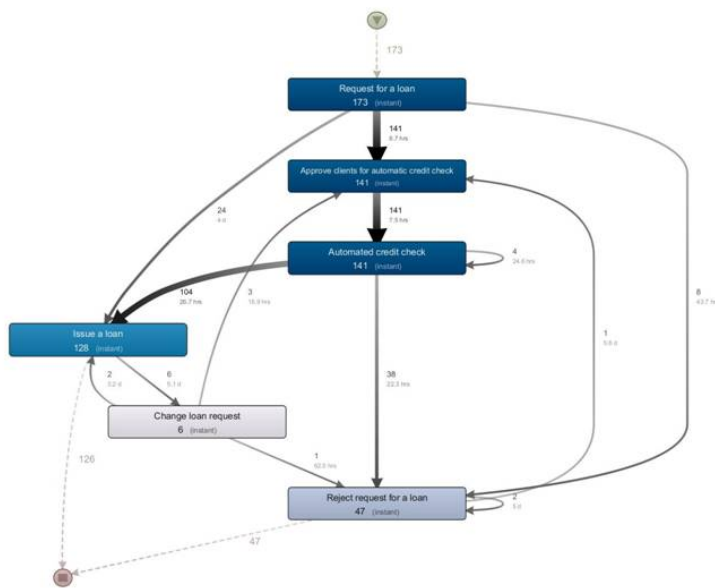


Figure 3.3: Frequency in Disco for Loan Approval process

for the particular duration. The key analysis criteria of a loan approval cycle are average cycle time and deviations in the standard path. The discovery of the process can be seen in the Figure 3.4.

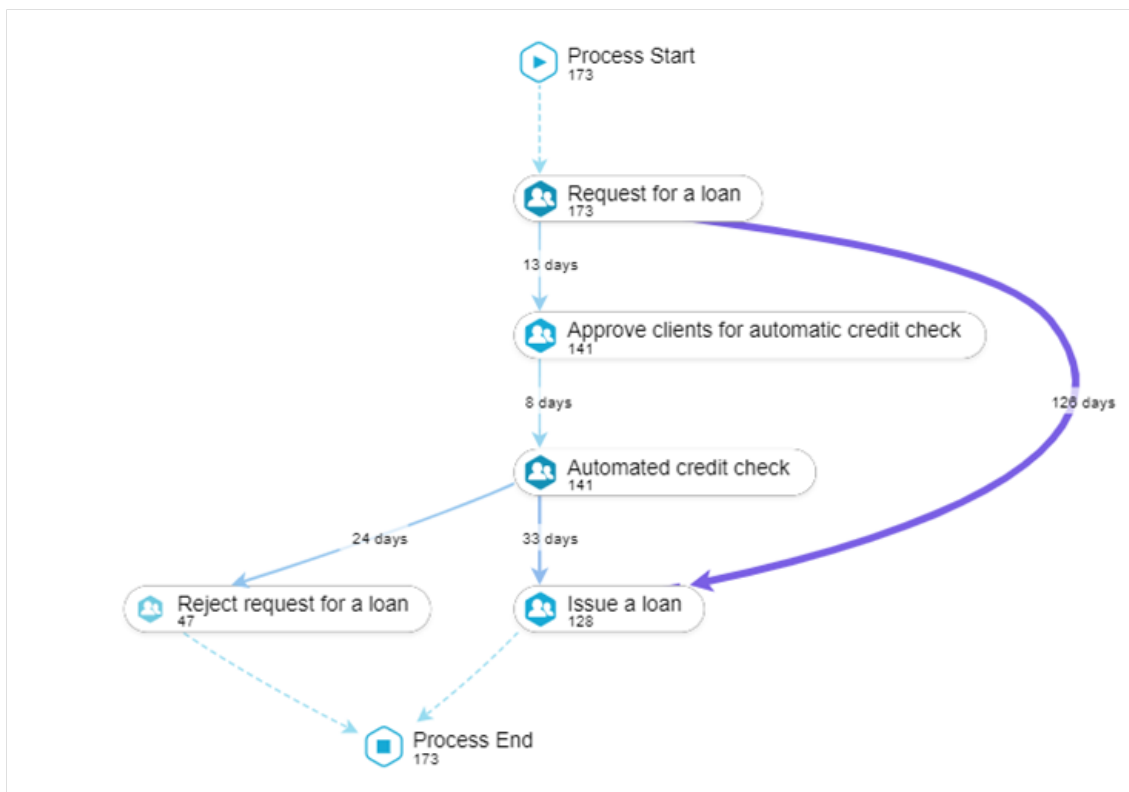


Figure 3.4: Automatic process discovery in Celonis for the Loan Approval process

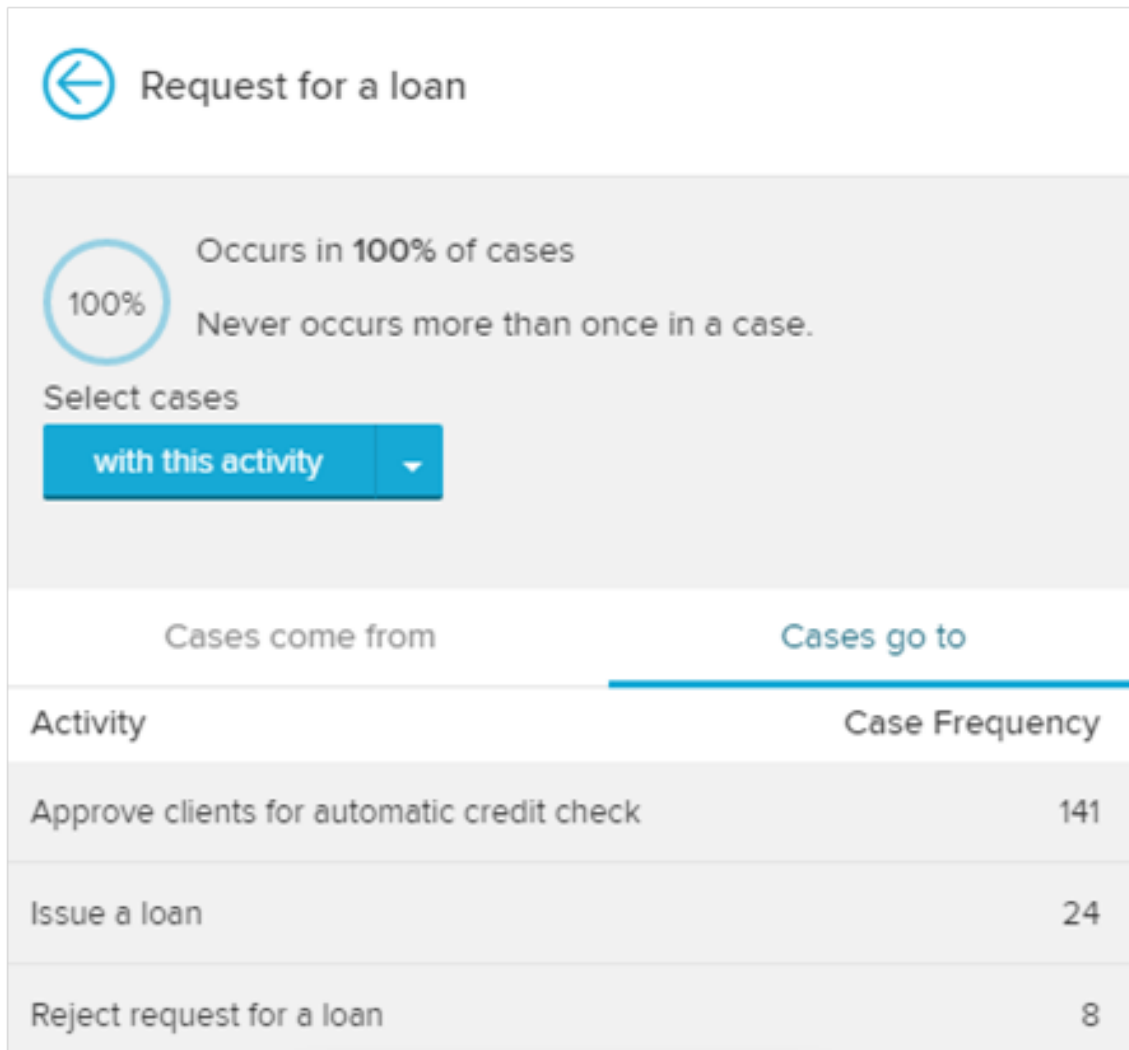


Figure 3.5: Performance analysis in Celonis for Loan Approval process

From the [Figure 3.4](#) we can see the number of events in each activity and average through put time between the activities. It can be seen that automated credit check to issue or reject loan is taking considerably much time than the other activities. The most unusual path we can see is request for a loan to issue loan directly without any approvals and credit checks, this kind of feature can be used by the business professionals to investigate and take actions by exploring the whole process in detail. Another good feature of this process discovery is to check the various incoming and outgoing routes of a particular activity which can be seen from [Figure 3.5](#).

3.1.1.3 myInvenio

In myInvenio, after the necessary data pre-processing is applied, one can discover a process graph with all the required information attached. For instance, for a given activity the notification bar display the respective frequency and performance. Process visualisation feature of myInvenio to animate processes, makes it easy to spot bottlenecks on the fly. Moreover, it provides a capability to export the animation in video format. [Figure 3.6](#)

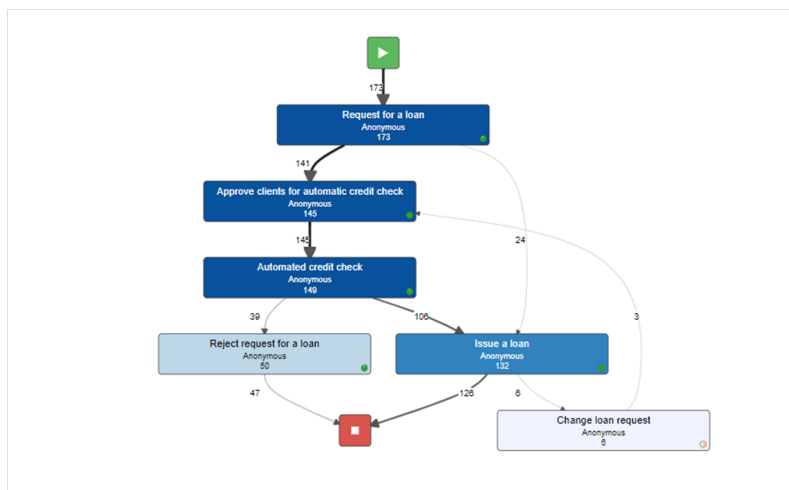


Figure 3.6:

The Analytic section in the tool helps in analysing the process. It helped in identifying the different process variants in the tool. Results from the tool showcases that 10 process variants occurred for the given use case in which five of the case have high throughput time. 43 case are following less throughput time on an average in total of 173 cases as shown in Figure 3.7. This analytic section is a unique feature of myInvenio. This is divided into different sections to show case the clear analysis. This is as follows:

- Process cases, The number of cases are shown. In figure Figure 3.7 it shows all the cases are covered.
- Case details helps us to go through every case and time taken by each case.
- Process section helps us to show the selected case or variant process in a diagrammatic format. It helps to understand each case individually.
- KPI summary indicates the minimum, average and maximum throughput time taken by the cases in the process. It shows that there are 43 cases which took very less time and most of the cases that is 125 of them are taking average time to complete the process.
- variants summary helps in identifying the process variants. So the improvements can be made less occurrence of cases. In this use-case, we have 5 process variants which are completely taking different process path with very long time.

3.1.1.4 Minit

As a part of log analysis, tool identified which activities are the starting activities and which are end activities. Tool depicts 'Request for a Loan' attribute is a start activity and 'Issue a Loan' and 'Reject request for loan' attributes are two end activities in different cases as per the given data. Based on the individual process and available data, we had to recognise which activities are considered as end activities. Figure 3.8

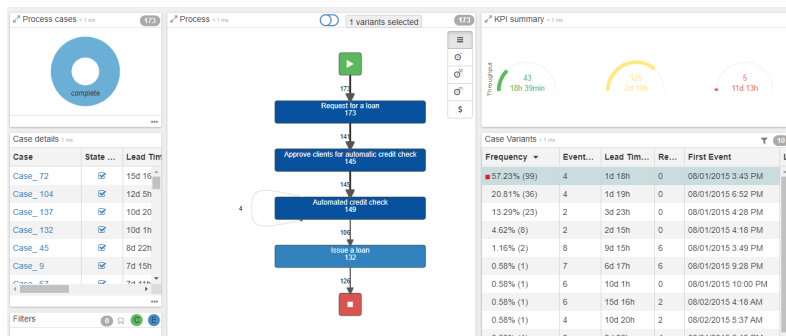


Figure 3.7: Analytics in myInvenio for Loan Approval process

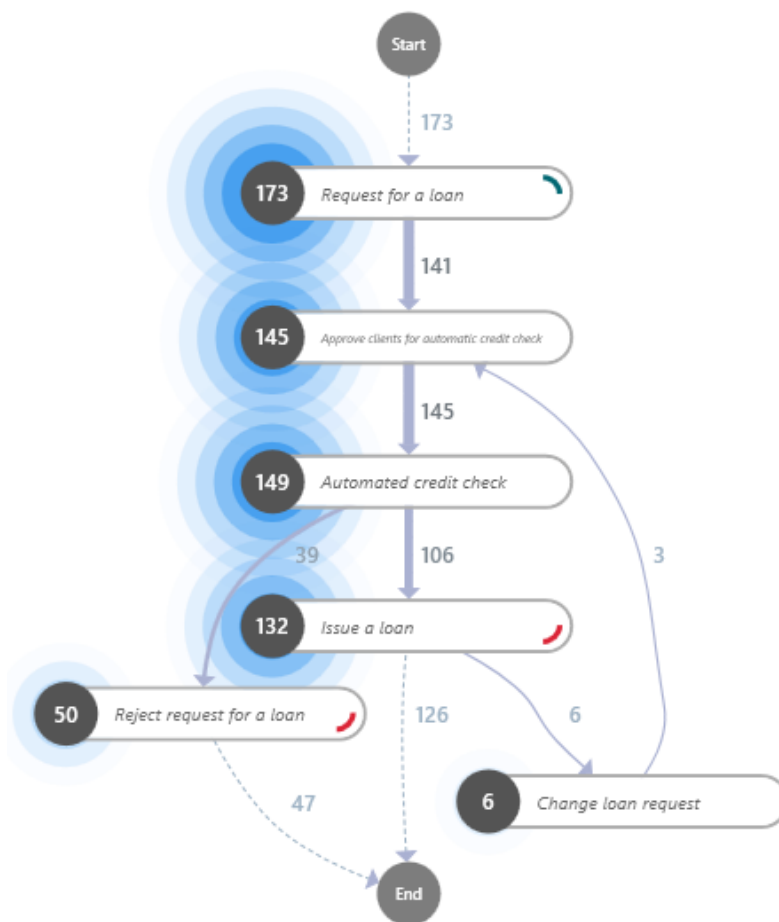


Figure 3.8: Automatic process discovery in Mint for the Loan Approval process

As per performance criteria, the tool shows that all activities that are passing through 'Issue a loan' -> 'Change loan request' take above throughput time Figure 3.9 .

3.1.2 Conformance Checks

Other valuable analysis is conformance check. Tools are compared by importing a .bpmn file for the Loan Approval process. In this section to check the conformance functionality of the tools, prepared a example reference model from knowledge of

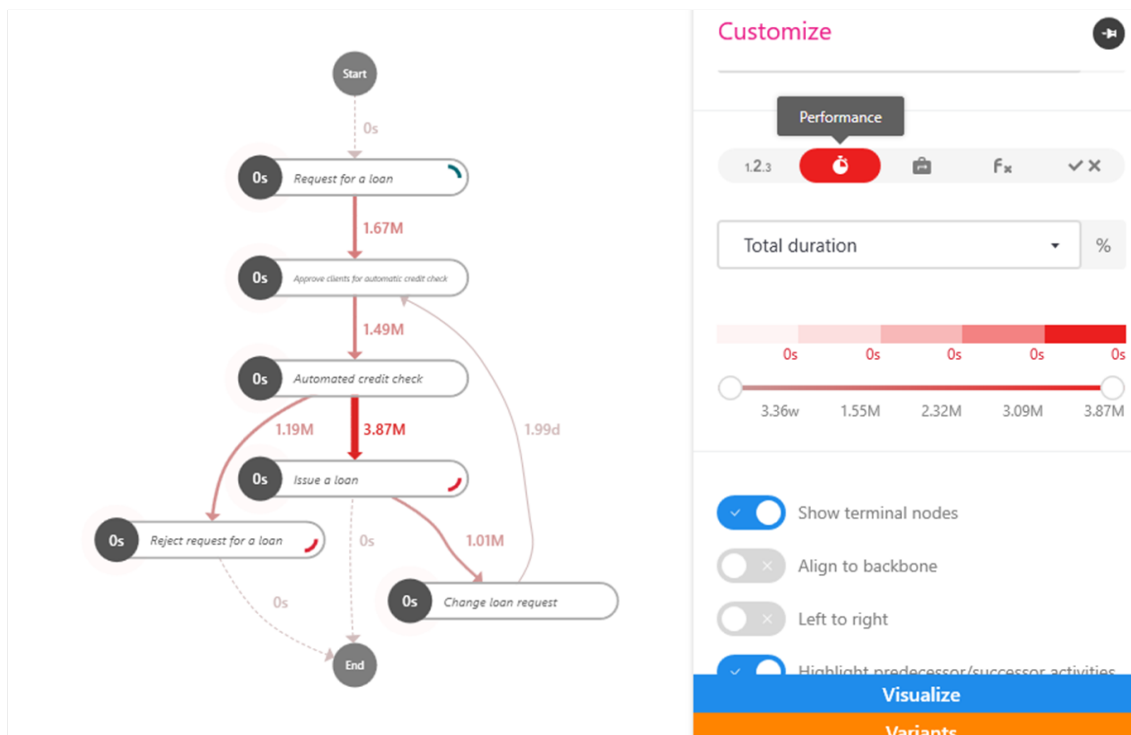


Figure 3.9: Performance analysis in Minit for Loan Approval process

literature survey. Figure 3.10 represents the bpmn model given to the tools as a real process.

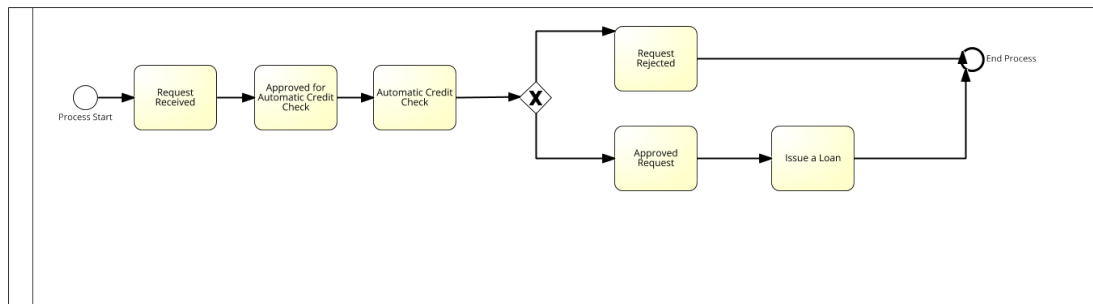


Figure 3.10: Loan Approval Process bpmn

Objective:

- The as-is process should be compared with the targeted process.

Expectations:

- The tools should show the difference in both the process. This helps keeping the process transparent.

3.1.2.1 Disco

Disco tool does not support conformance checking as it is not having the option to upload modelled process. The only possible task of process mining can be possible in this tool is Process discovery. This process discovery can be done by applying different filters and by selecting various permutations and combinations.

3.1.2.2 Celonis

The conformance activity needs the two processes which are the process actually happened which is generated based on the data given and another one is the modelled process against which conformance is needed. The modelled process can be given in 3 ways which are:

- Upload BPM file of the modelled process.
- Take the model of the process from the Celonis process repository in which you already saved the modelled process.
- The most comfortable way is to adjust the model generated from the given data to match with modelled process and do conformance checking.

In Celonis, As in [Figure 3.11](#) conformance checking is seen in descriptive format like the percentage of conformance cases, number of violations, KPIs for violated and non-violated cases and also projects effect on throughput time due to violation to modelled process and which will helpful in detecting the bottleneck activities. The drawback of the Celonis is it will not present the conformance checking results in diagrammatic form.

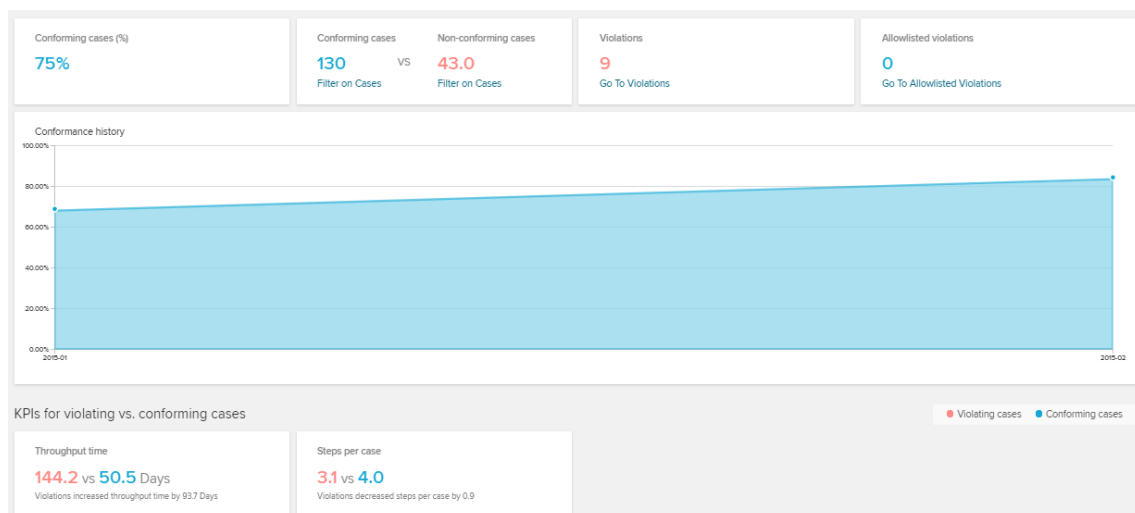


Figure 3.11: Conformance Check description in Celonis for Loan Approval process

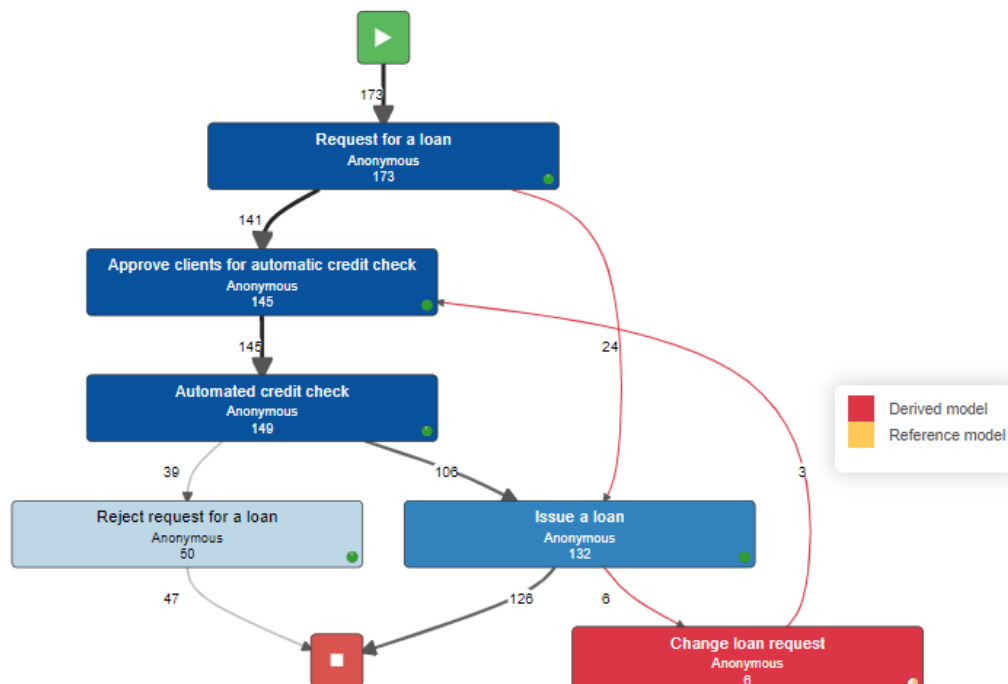


Figure 3.12: ConformanceCheck in myInvenio for the Loan Approval process

3.1.2.3 myInvenio

After loading the bpmn file for comparison, my Invenio provides conformance results in the form of comparative map by clicking the tap compare as in Figure 3.12.

Here we can identify the difference in model with actual model. The data variable 'change loan request' is deviating the process in the use-case. Tool also provides description about the conformance and non-conformance cases as in Figure 3.13. It describes that 130 cases are perfect and the other 43 are non-conformance cases. This effects in 13.9 percent of unexpected process flow. Here filters can also be applied for clear analysis of the process map.

3.1.2.4 Minit

In Minit, process can be compared with different models and also we can compare with different process variants. To do so, we need to create and need to save two views. for first view, choose variants filter, choose variant number 1 and apply. we can see a process map of most used variant 1. by clicking on saved name it and save as variant 1. For second view, go to filters and invert the same filter, by doing this we choose rest of our data set only without the first variant. click on apply and save this view as excluded variant 1. Now clicking on the process compare functionality. In this tool process compare shows a pictorial representation as in Figure 3.14. As the other tools, we can see the data variable 'change loan request' violates the process model. No other detail information is given in the tool for conformance technique.

3.1.3 Enhancement

Enhancement is nothing but redesign the process based on the results and findings from the discovery, conformance checking. This enhancement phase is handled in

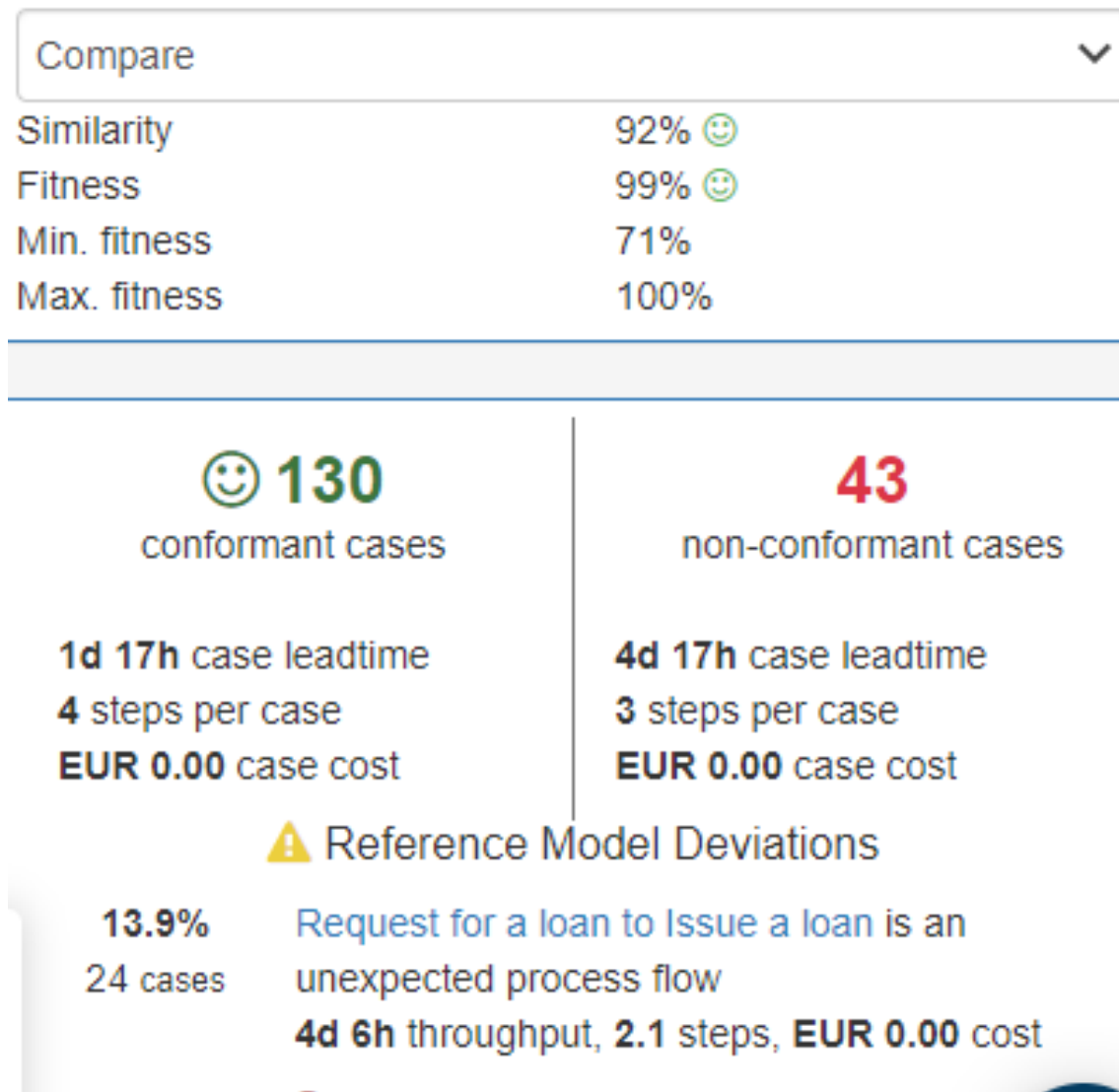


Figure 3.13: Conformance Check description in myInvenio for Loan Approval process

such a way to avoid violations, putting relevant controls and modifying the standard operating procedures. Hence, the enhancement phase is depends on the results of the above two phases and the extent of quality of enhancement is depends on the quality of output pertaining to discovery, conformance checking given by the various tools. Based on the findings discovery and conformance checks, the use-case as 5 different process variants which takes maximum throughput time. So we can improve the process time. And the data variable 'change loan request' is deviating the process. Here we can redesign the process for better improvements.

3.1.4 Conclusion

All the tools accepts the Loan Approval use-case. Also showed the process discovery and conformance techniques. Disco is easy and very easy to understand. The pictorial representation and descriptive presentation of myInvenio tool is clearly understandable compared to other tools. Celonis represents in a graphical form, KPI

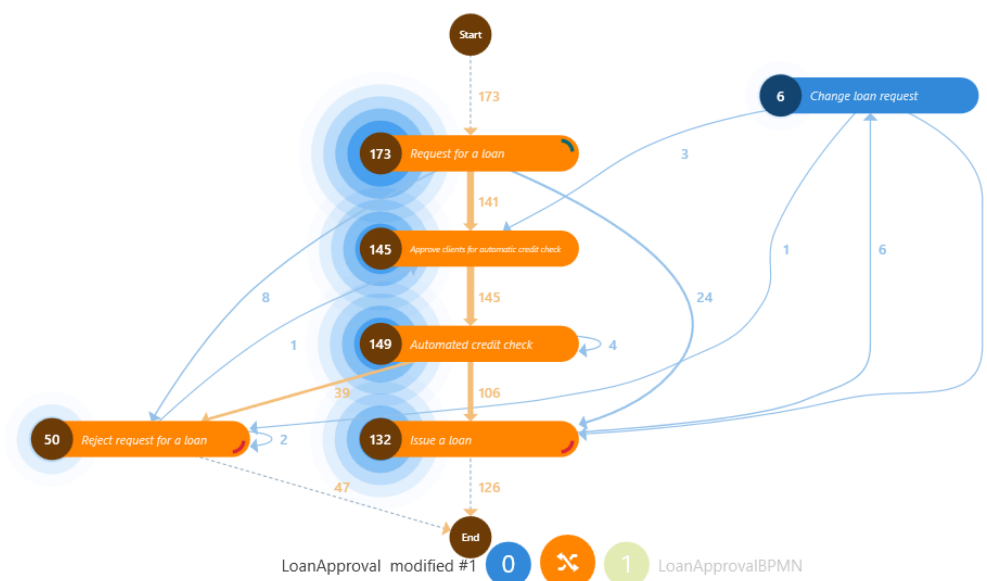


Figure 3.14: Conformance Check in Minit for Loan Approval process

and description while tool Minit as only image format presentation. And Rapid Miner extracts all the libraries from ProM which is has different libraries.

3.2 Use-case 2: Purchasing data

The second event log we are analysing is belongs to the purchasing process in an organization. This data log is taken from the website ¹ and it contains 608 cases and 9119 events. The purchasing process is very crucial process and a careful analysis is needed because this particular purchasing process involves several activities like analysing quotations and approving the vendor, release payment and settling various disputes. Now, we are going to analyse this crucial process with the help of various process mining applications.

3.2.1 Process Discovery

This part of the section depicts about how tools analysis resource and role variables in the data. Analysis the scope of improvement in the use-case and finding different ways of approach in different tools.

Objective:

- Resource and role based analysis to Purchasing use-case data.

Expectations:

- tool should identify resource and role data variables.
- tool should show all the improvements for resources and roles in the data.
- also different functionalities of the tool are explained here.

¹<https://fluxicon.com/disco/>

3.2.1.1 Disco

With respect to discovery features of disco, we already discussed the various features of disco and which are limited and not be changed based on the different use case. The pictorial form the purchasing process can be seen in the below picture and from the below picture we can see that the process ends only after Analyse Purchase Requisition and Analyse Request for Quotation which means that the purchase requisitions are not qualified as per the policy and needs careful consideration to analyse the causes for the same. Resource based analysis: While importing the data into disco, tool provided an option to select resource column. In disco, we will not get any detailed insights to do resource based analysis but we only do get the statistical information like recourse wise event count and frequency in total event log and we also have option to export this tabular data in CSV format. From this, process owners can get information that who are the crucial persons in contributing to the process and can helps in making relative strategic decisions accordingly.

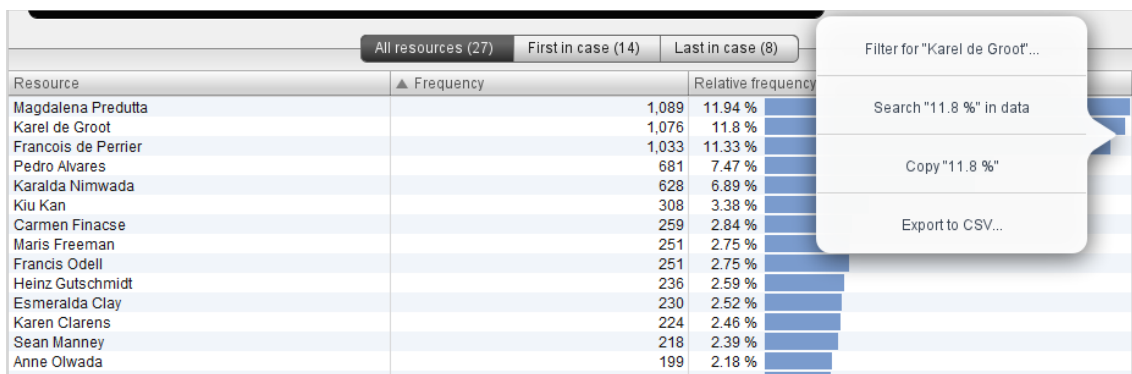


Figure 3.15: Resource based analysis in Disco for pursaching data

Role based analysis: While importing data, tool provided an option to select role column. After selecting role column, we will get the statics in tabular form as per the below picture. This table can be helpful in taking various managerial decisions by observing the most frequent and exceptional roles.

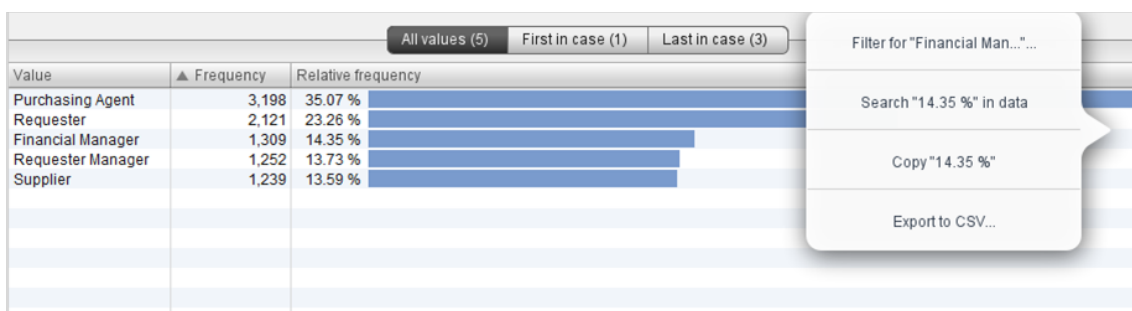


Figure 3.16: Role based analysis in Disco for pursaching data

3.2.1.2 Celonis

The discovery features in the Celonis can be seen in process explorer part and the same is discussed in the 1st use case. Now, in the present use case we are going to

analyse the event log based on resource. Resource based analysis is supported in the Celonis and this type of analysis can be seen in Social application. As we know, all the process mining tools do not give clear picture of the event log due to the fact that some algorithms do not support complex type of data but as of now Celonis is giving clear picture of our two use cases. The process flow of current use case can be seen in the below picture.

Our focus is to how this discovery phase effects the process owner to streamline the processes and from the above picture we can see that in 38 percentage of the cases amend request for quotation is followed by analyse request for quotation which results in rework and create request for quotation directly followed by create purchase requiting without analysing the same. These two findings results in rework and wastage of resources like time, money and this can be eliminated by strengthening the procedure by process owners.

As we discussed earlier, we are going to do the resource based analysis in this use case. The pre requisite to perform resource based analysis is to have resource column in the event log. Resource in terms of people in the organization who involved through the process. This resource feature can be seen in social application of Celonis. By selecting this application, it will as us to select resource column in our event log and which can be seen in [Figure 3.17](#) and [Figure 3.18](#).

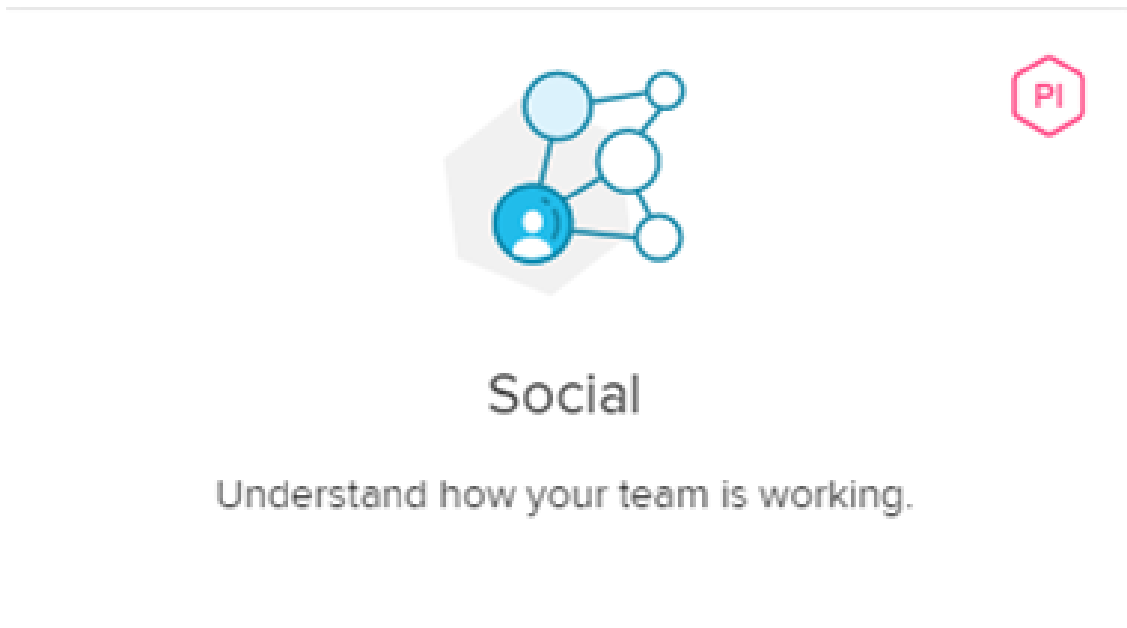


Figure 3.17: Resource tab indication in Celonis

After following these few steps, system will give us the output which helps us in doing resource based analysis and the various options of the social application are discussed below: a) Overview: It will give us key statics pertaining to performance of the personnel in the organization. The key statistics helps in taking strategic decisions for example, one of the statics is users per case which means how many users involved in performing activities for one case end to end this statics helps process owners to reduce time and cost by eliminating the non-core activities there

START TIMESTAMP	COMPLETE TIMESTAMP	ACTIVITY	RESOURCE	ROLE
Jan 1, 2011	Jan 1, 2011	Create Purchase Requisition	Kim Pessa	Requester
Jan 1, 2011	Jan 1, 2011	Create Request for Quotation R...	Kim Pessa	Requester
Jan 1, 2011	Jan 1, 2011	Analyze Request for Quotation	Karel de Groot	Purchasing
Jan 1, 2011	Jan 1, 2011	Send Request for Quotation to S...	Karel de Groot	Purchasing
Jan 1, 2011	Jan 1, 2011	Create Quotation comparison M...	M...	Purchasing
Jan 1, 2011	Jan 1, 2011	Analyze Quotation comparison ...	Immanuel Karegianni	Requester
Jan 1, 2011	Jan 1, 2011	Choose best option	Tesca Lobes	Requester
Jan 2, 2011	Jan 2, 2011	Settle conditions with supplier	Francois de Perrier	Purchasing
Jan 2, 2011	Jan 2, 2011	Create Purchase Order	Karel de Groot	Purchasing

Figure 3.18: Data selection for resource analysis in Celonis

by reducing the number of personnel involved which finally resulted in steam lining of the current process Figure 3.19.

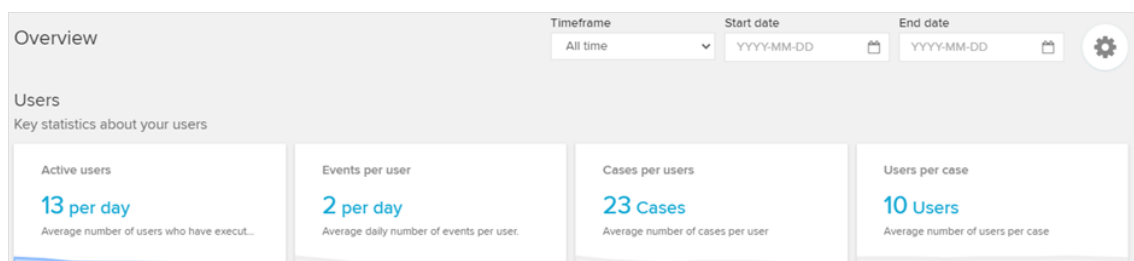


Figure 3.19: Overview of data in Celonis for Purchasing data

b) Users: This particular option of social app gives us several insights in detail, this particular feature gives the pictorial presentation of the performance of the in terms of

i) Quantum of work performed: This particular feature helps in identifying the personnel in the order of the work they performed in terms of quantum. By using this, we can conclude who is doing the more work and who is doing the less work and can take strategic decisions accordingly.

Possible decisions for the personnel doing more quantity of work: Process owner can take decisions with respect to giving more incentives and bonuses to retain such personnel and can call for the discussion personally to get to know whether they are comfortable in doing heavy quantity of work so that they can plan for sharing of burden which ultimately results in happy employee relationships.

Possible decisions for the personnel doing less quantity of work: We can't conclude straight away that the less quantity of work means inefficient employee, because less quantity of work depends on the role they have for example in our process settling dispute with supplier is happening in exceptional cases and in that exceptional cases we may have to deploy a person having specified role to handle that dispute. Hence, careful analysis is needed to conclude why they are performing less quantity of work. Some other causes for doing less quantity of work are, the personnel not having the suitable skills to perform the tasks which can be eliminated by giving proper training and still the management feels that skills can't be updated in such personnel then

they can shift them to another suitable department or can be removed from the organization, and in some cases the personnel having the required skills but not active enough in performing the tasks which results in overburden in few employees and wastage of financial resources this can be eliminated by taking corrective actions by the process owners.

ii) Throughput Time of Users: This particular features gives the pictorial form of the users based on the time of the work performed so that we can get who is performing faster and who is performing slower. Based on this, we can analyse the slow moving personnel and can give proper training to perform the work faster there by reducing the time and which ultimately results in reduction of financial resources for doing the work for more time Figure 3.20.



Figure 3.20: Throughput time in Celonis for Purchasing data

Along with the above we can also drill down the specific employee and can analyse those employee performance in detail. Another option we have in this social application is “Activities”, it will give us the activities in pictorial form and if we select specific activity it will give us the detailed information like who performed the specific activity and the importance of specific activity can be known by doing this analysis Figure 3.30.



Figure 3.21: Social activities application in Celonis for Purchasing data

For example, I have selected amend request for quotation and it gave the metrics as, key metrics are the average number of users in this activity and we can also see means what are the predecessor and successor activities. It also gives us the user’s profiles that performed the selected activity which helps in identifying whether the

users performing that particular activity are authorized and if not authorized then the process owner should take corrective actions to strengthen the controls as shown in Figure 3.22.

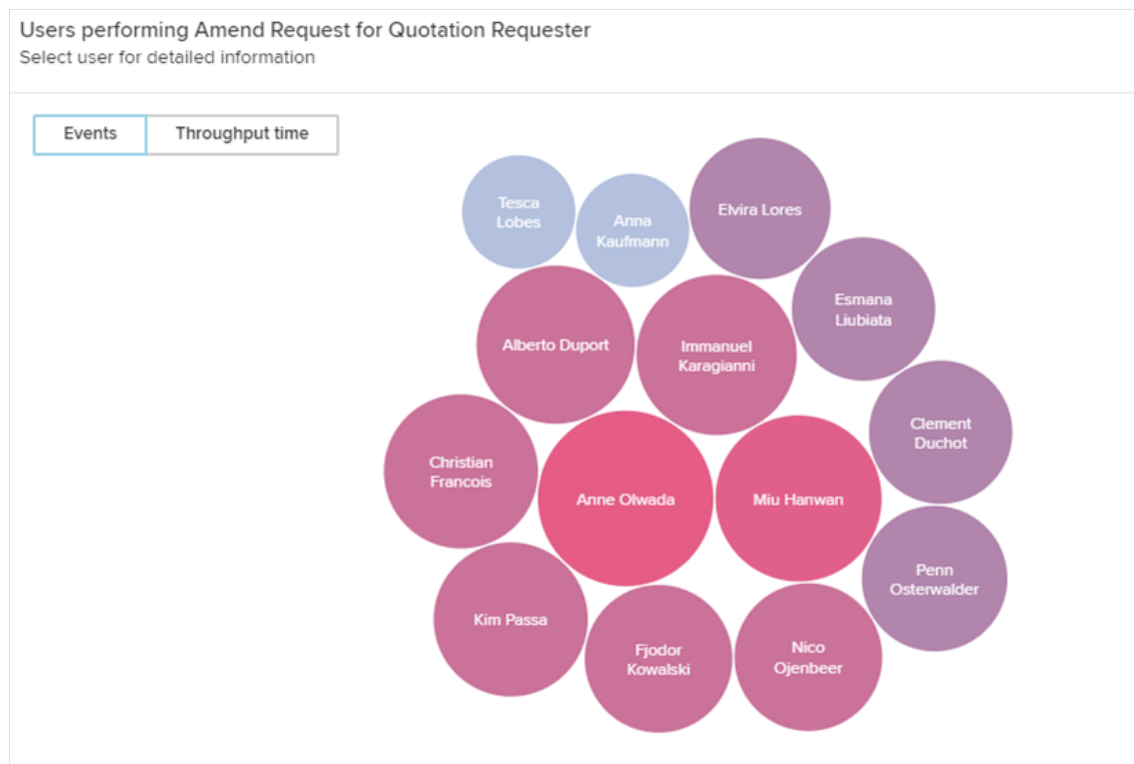


Figure 3.22: Social network for specific in Celonis for Purchasing data

Role based analysis of the event log is not supported by Celonis.

3.2.1.3 myInvenio

Data discovery is defined dashboard of the tool. Here process map is observed. Figure 3.23.

Here, we can see the most frequent activities quickly due to the unique feature of highlighting the frequent activities in the process map. In our use case “Analyse request for quotation” is having highest frequency and because of this unique feature process owners can identify easily the crucial activities in the process. The process model can be viewed in different ways by make use of the various filters provided Figure 3.24.

The easy and quick identification of rework activities helps the process owners to take corrective actions around those activities which ultimately results in minimising the time taken for the completion of the process. In our use case, the rework is highlighted as pictured below and we can see that the rework frequency is more between “Analyse requests for quotation” and “amend request for quotation” Figure 3.25.

d. Time Identification: Upon selecting the duration button, it portrays the process map with the time taken for each activity in different colours based on respective duration taken by them to perform an activity. In our use case, we have selected

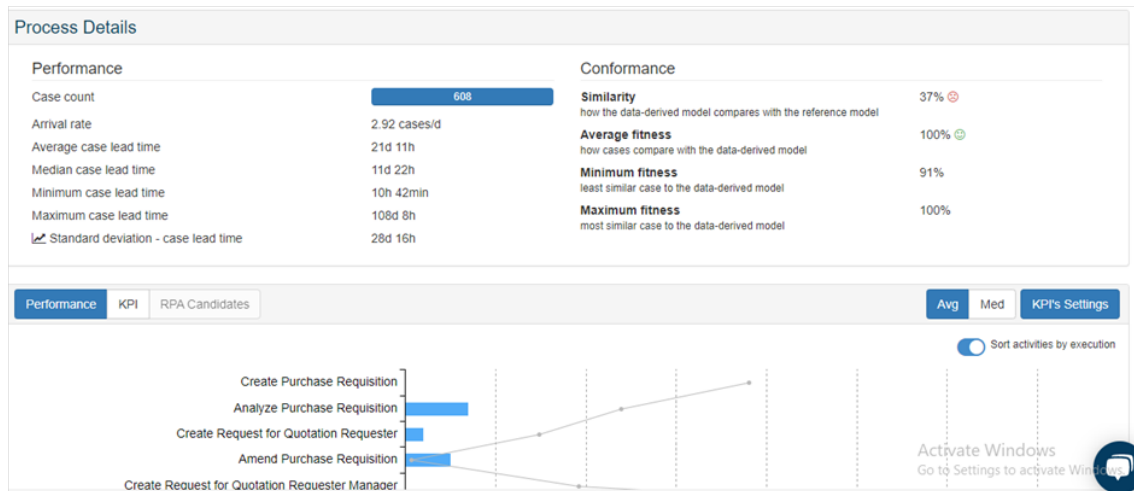


Figure 3.23: Process Dashboard in myInvenio for Purchasing data

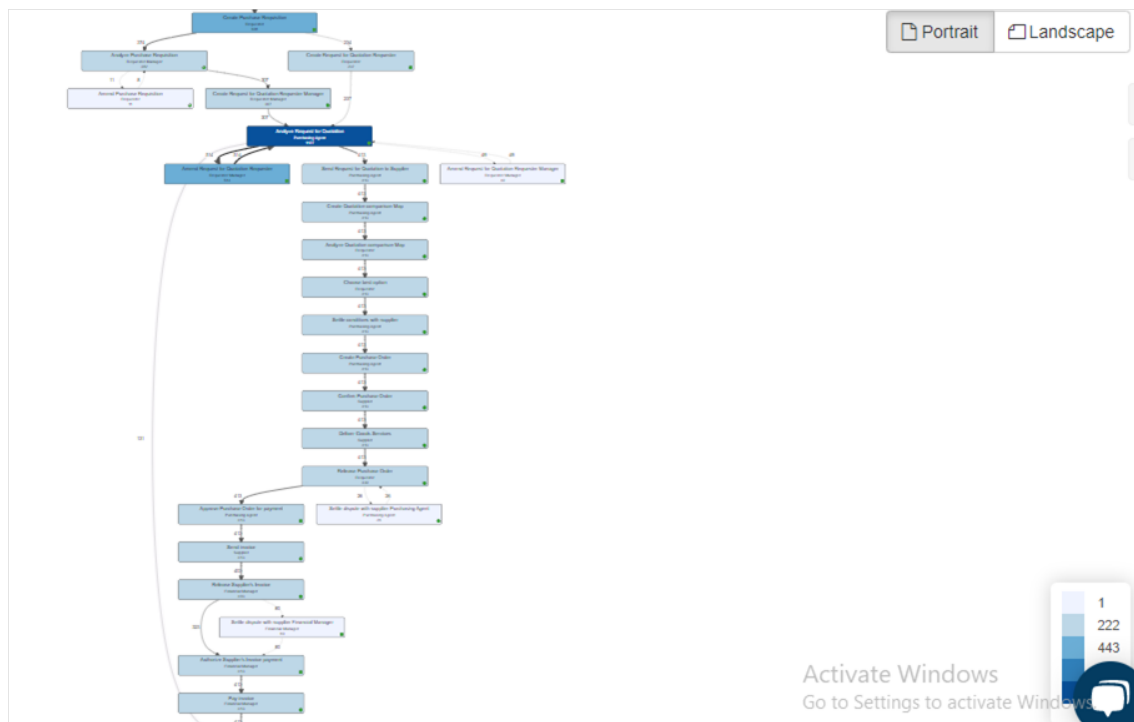


Figure 3.24: Process map in myInvenio for Purchasing data

“average time” from the drop down as it is an appropriate criteria and then we will get the output in as pictured below and from this we can quickly identify highest time taking activity which is highlighted in dark which is “Delivery Goods Service”. By using this, management can update their procedures and train the resources accordingly to perform the tasks quickly and there by getting timely delivery of goods/services Figure 3.26.

e. Resource based analysis: This tool provides an opportunity to perform resource based analysis excessively. When we go to “resource focus”, we can get an option to select required resource and then it highlights the process map with the activities

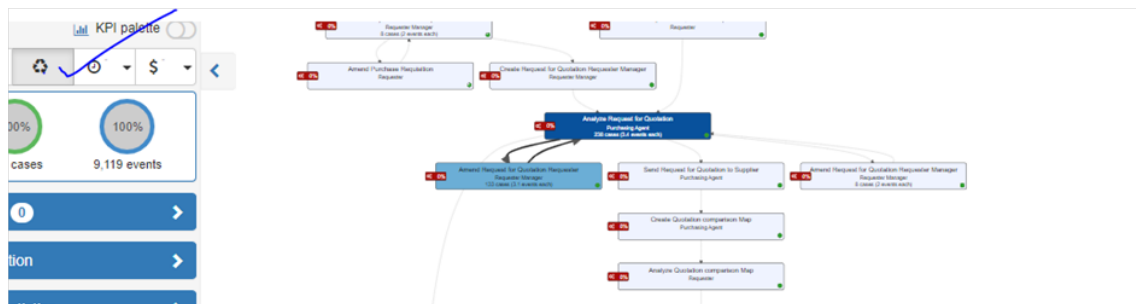


Figure 3.25: Rework in myInvenio for Purchasing data

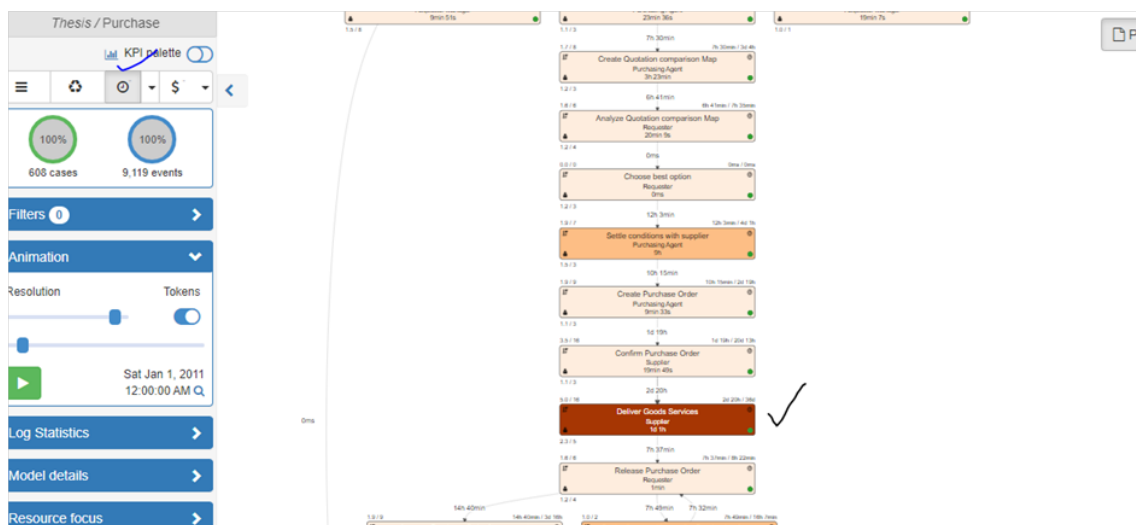


Figure 3.26: Time analysis in myInvenio for Purchasing data

performed by selected resources and we can use this feature in combination with “rework” and “time” filters so that we can see that who is responsible for delay in the process Figure 3.26.

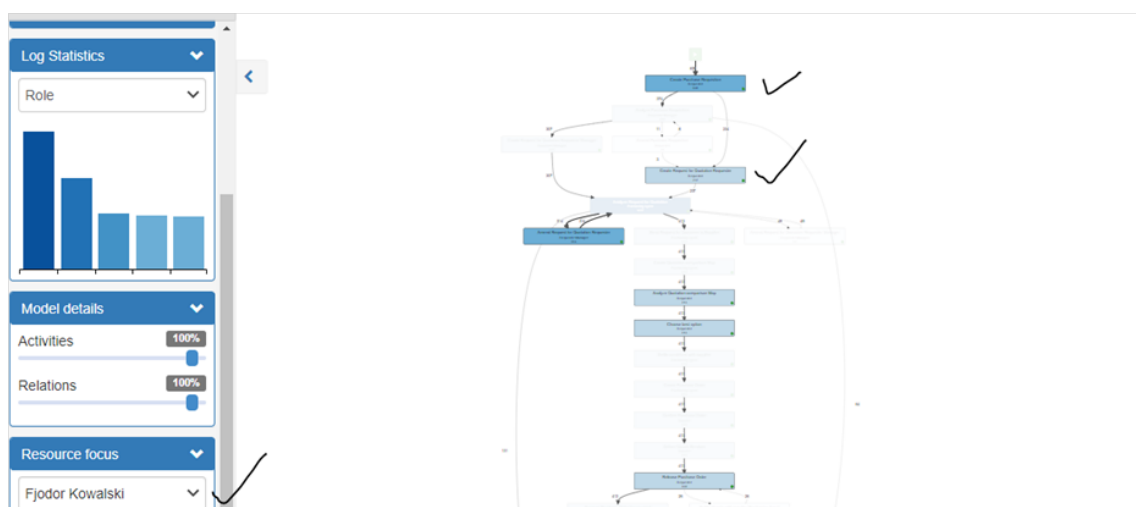


Figure 3.27: Resource analysis in myInvenio for Purchasing data

f. Role based Analysis: The special feature of this tool provides an option to view the role focused process map which helps to know the activities performed by the selected role. This is very useful in deciding segregation of duties which says that maker and checker should not be the same. In our use case, I have selected “requester” who performed various activities in the range which starts from “create purchase requisition” to “choosing the best quotation” and this can be helpful in identifying the role based violations quickly [Figure 3.28](#).

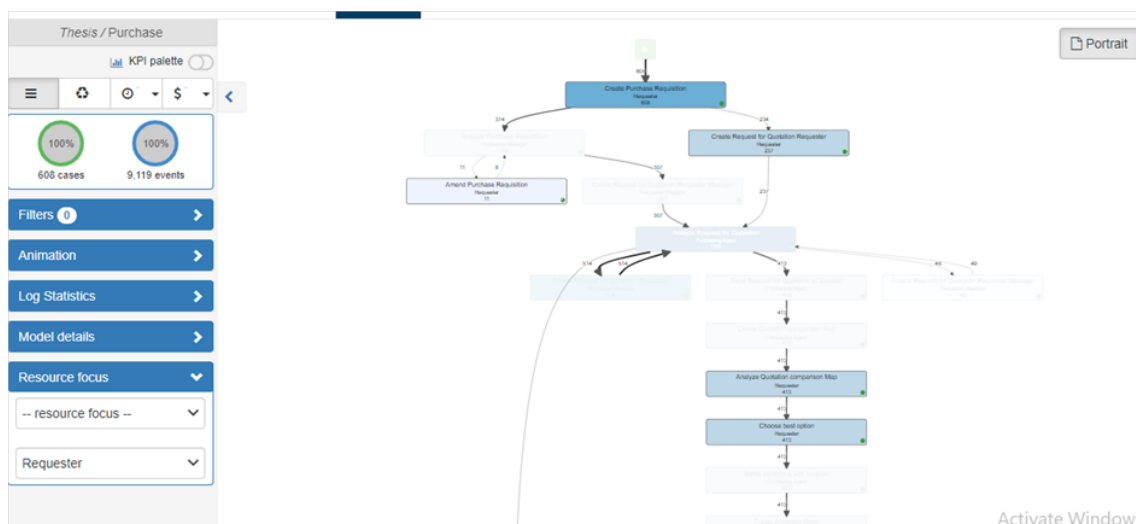


Figure 3.28: Role analysis in myInvenio for Purchasing data

g. Activity Map: “Activity map” tab of this tool provides the summarised view of all activities in the process with the respective personnel names involved in performing those activities. In every organisation, management do have expectations to perform the activities by the personnel as per the responsibilities allocated to them. By utilising this tool management can conclude whether any departure from their expectations [Figure 3.29](#).

h. Social Net: This generates the working environment of the personnel in pictorial form. From this, management can identify the personnel who are working together and doing similar tasks as pictured [Figure 3.30](#).

Another special option we have in this is we can drill down the performance history of specific person as pictured below [Figure 3.31](#).

3.2.1.4 Minit

In Minit, process discovery is defined. To analyse the roles in the use-case. We need to select Process Map icon. A social analysis, by clicking on the social chart, we can see that it will automatically changes the look of the map. It represents all the effective users and the relations between them. By clicking on the performance analysis, by choosing mean duration we can see how long average communication between the users took. by clicking on any of the paths, we find further useful information on the path panel on the right. Here Visualize we need to filter it to role based clustering. Then the role based analysis can be done. Here, The role Supplier takes more time in the process. which effects in time delay in whole process as shown

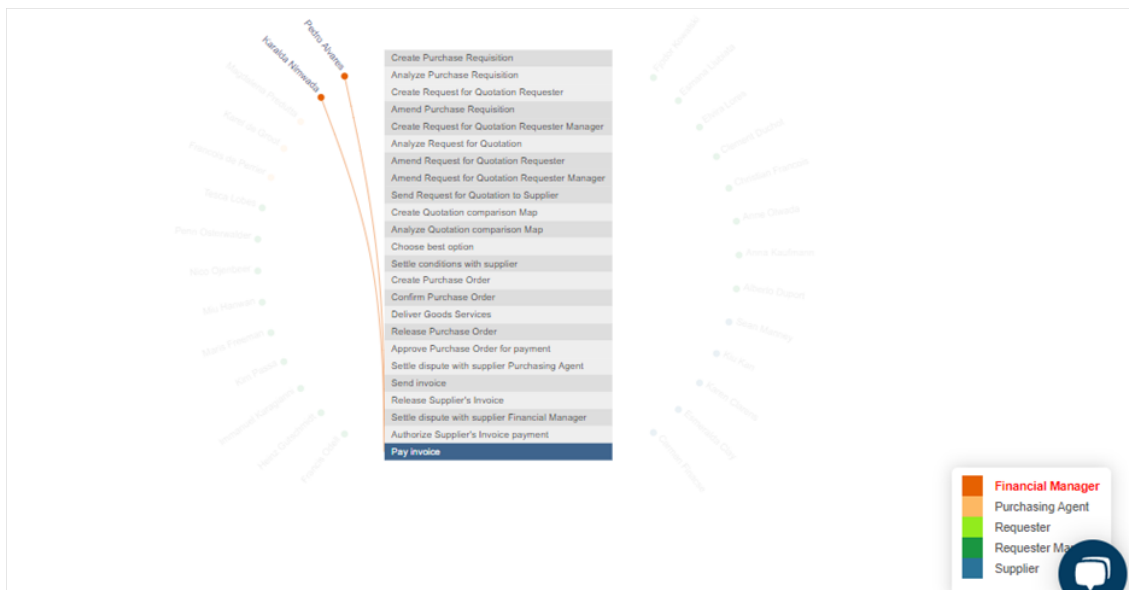


Figure 3.29: Activity map in myInvenio for Purchasing data



Figure 3.30: Social activities application in Celonis for Purchasing data

in the Figure 3.32. For resources analysis, there is a tab called socialchart. Here we can analysis all the resources. This chart is obtained in three different types, i.e., Circular form, Spring form and Multi scaling spring. Multi scaling chart helps in understanding the path of the resource in the process map clearly.

3.2.2 Conformance Checks

3.2.2.1 Disco

As said earlier, Disco does not support conformance checking. So the tool is neglected in this technique.

3.2.2.2 Celonis

In The first Use Case, we have performed conformance checking to know the deviations in the process of an organisation. In this Use Case, we are going to perform

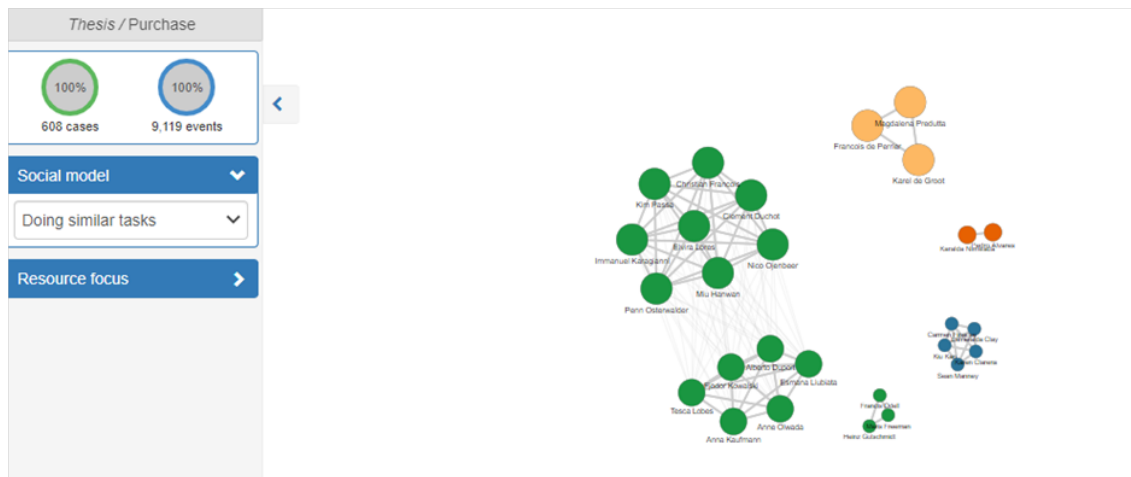


Figure 3.31: Social network drill down in myInvenio for Purchasing data

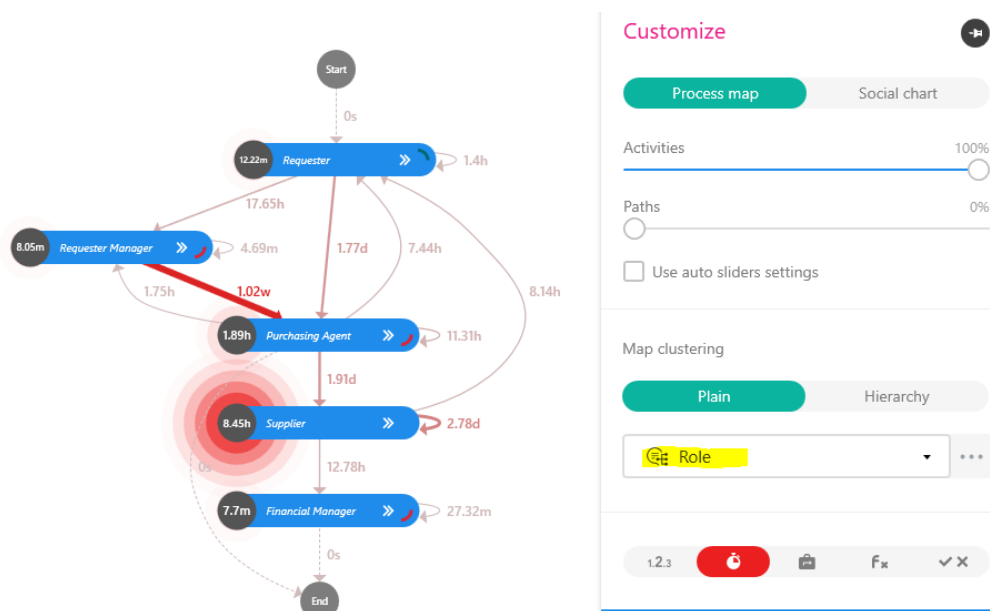


Figure 3.32: pics/Role based analysis in Minit for Purchasing data

conformance checking to know the deviations in an organisation hierarchy while performing their responsibilities. Here, we are going to demonstrate how we can use conformance checking feature to perform role based analysis because we are not having any other suitable feature in this tool to perform role based analysis.

To perform this analysis, we have to select “role column” as an activity while pre-processing of data. Thereafter, choose conformance checking feature and in that it will show various options to upload standard “.bpmn file” to perform the analysis. From those options, we go into “Mine the target process” options as we are not having any standard hierarchy structure.

After the above selection, we can view various options to create hierarchy structure to compare with the actual flow. From those options, we are going to choose option called “Most Common Variants” because as per the general understanding, the cases

which are deviated from the “Most Common Variants” are considered as exceptional or violation. In the below picture, we can see that in “Most Common Variant option” we have been provided with various variants and from that I have selected first four variants which occupies major portion out of all.

After performing the above procedure, we are going to launch the analysis and will get the output as pictured below. We can see that there are 463 non-conformance cases in total and the non-conformance might be because of various reasons and the process owners can analyse the reason for this by using the output we get from the system and thereby they can take suitable measures to overcome those violations.

3.2.2.3 myInvenio

As we discussed in the Loan processing use case, my Invenio provides conformance results in the form of comparative map [Figure 3.33](#). While we are trying to do conformance checking in the current use case, tool is giving the user-friendly and clear comparison map as this event log consist of numerous activities and events. But, we can get the conformance results in descriptive form as pictured below [Figure 3.34](#).

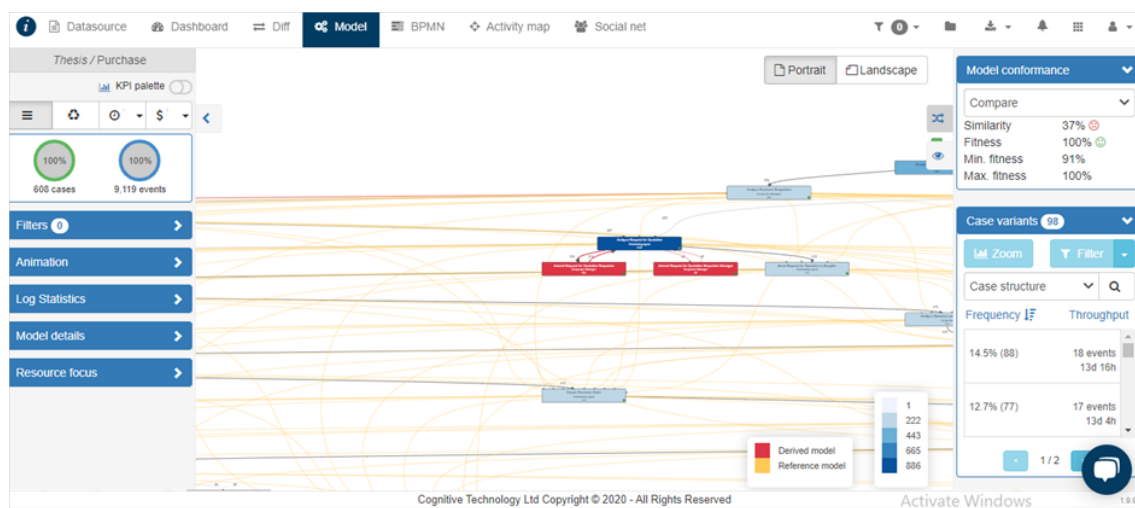


Figure 3.33: Conformance Check in myInvenio for Purchasing data

3.2.2.4 Minit

As discussed, Minit provides conformance check. In the [Figure 3.35](#) shows the model of the process map in the tool. Data variables are not visible.

3.2.3 Enhancement

As we know enhancement is a redesigning technique for the process to improve by analysis process discovery and conformance technique. In this use-case it is observed that there is a rework around the variable amend for request for quotation’ due to which results in delay in purchasing process and factors effecting this situation due to lack of trained person. We found this effect using process mining. So using this management can take care of it accordingly.

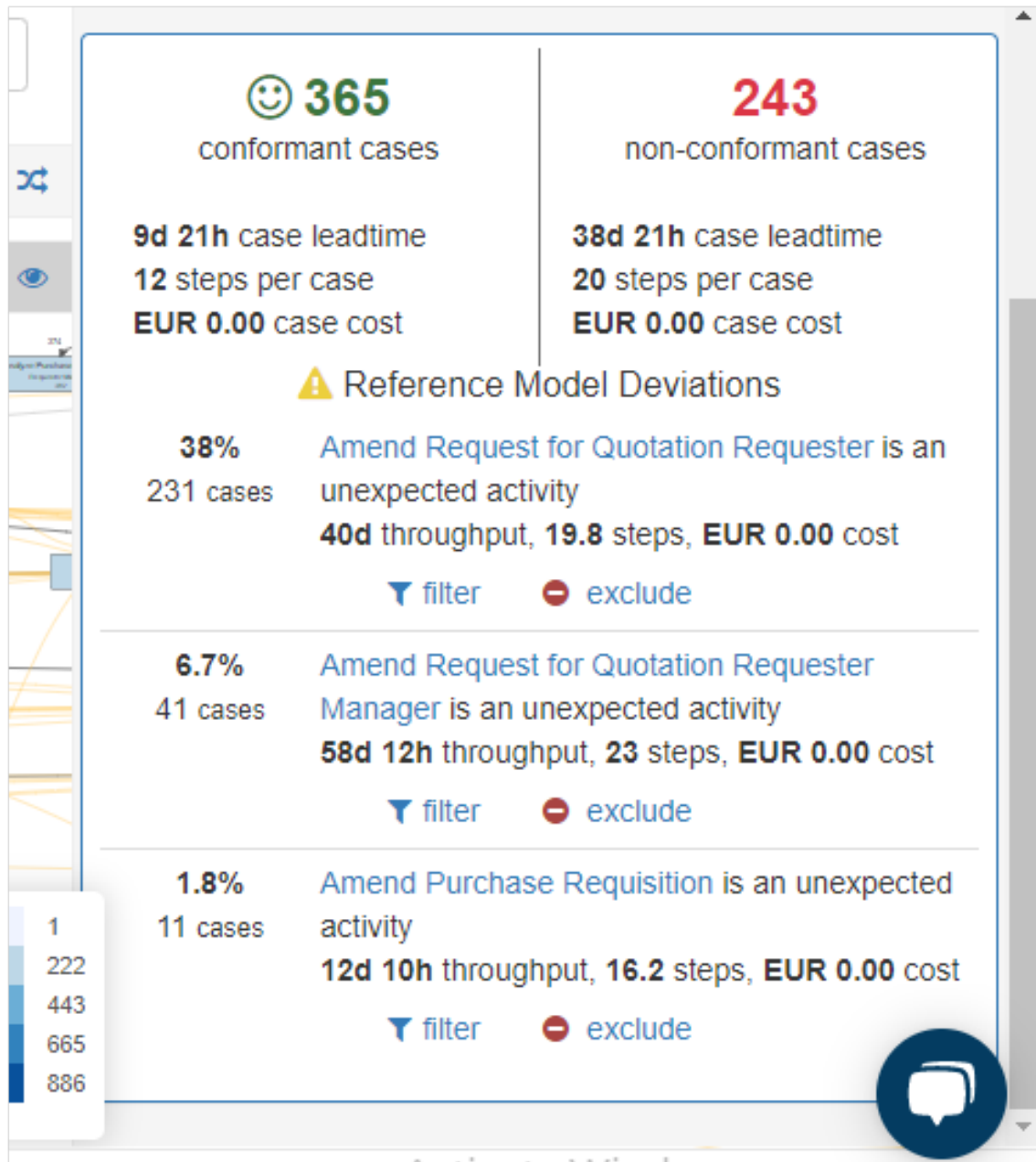


Figure 3.34: Conformance Check description in myInvenio for Purchasing data

3.2.4 Conclusion

In role and resource based analysis, the tools Disco, myInvenio and Minit shows the analysis for the given use-case. Celonis accepts only resource based analysis as a social network section. The analysis is easy and understandable in Disco and Minit.

3.3 Data pre-processing

3.3.0.1 Disco

Pre-processing of data while importing is given in the below Figure 3.36. It is very easy to understand. We can select only 5 data variables, i.e., case id, activity, timestamp, resources and one extra variable from data for analysis.

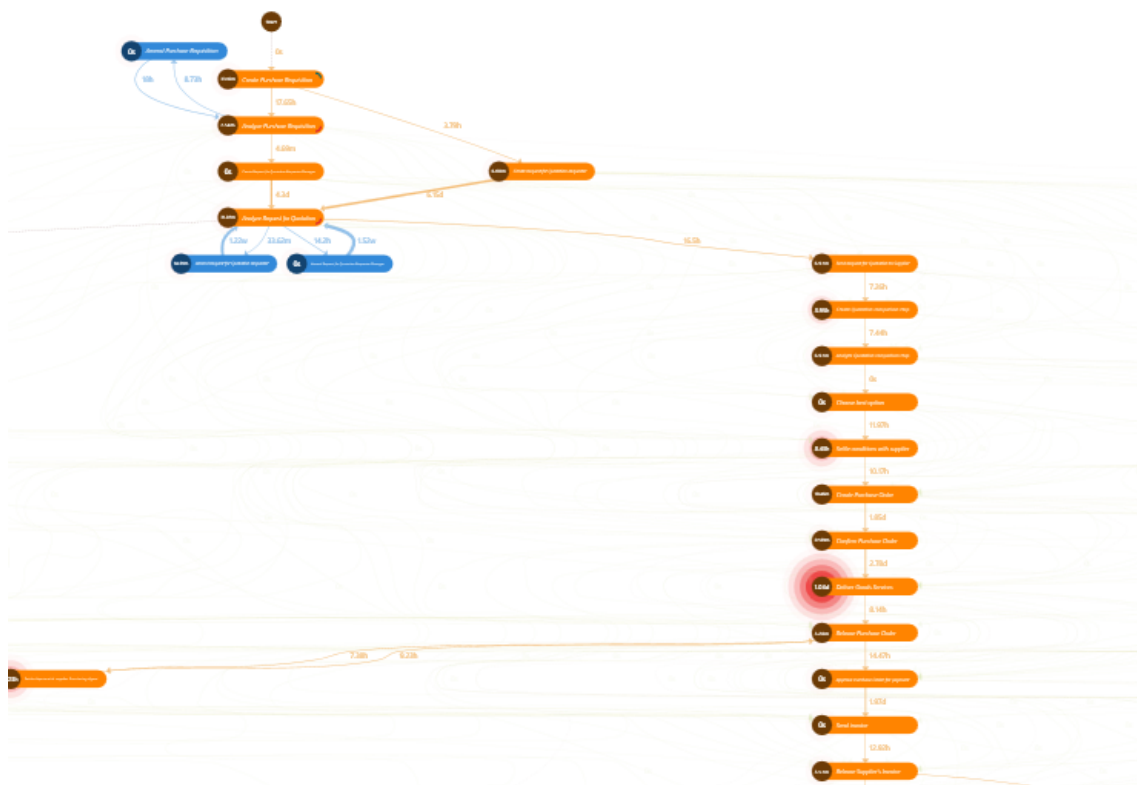


Figure 3.35: Conformance Check in Minit for Purchasing data

3.3.0.2 Celonis

In Celonis, the data pre-processing is available in 'Event collection' Figure 3.37. Here 'data pools' is a section where our data samples are stored. we can upload the data file upload Process data Models and it is convert to event log data. The events are identified automatically. It also as a 4 step process. We can select only type of the four data variables, those are case id, activity, timestamp and one of the sorting variable from the data Figure 3.38. Then to analysis the stored data, we need to move to Process analytic section. There is no limit to store the data.

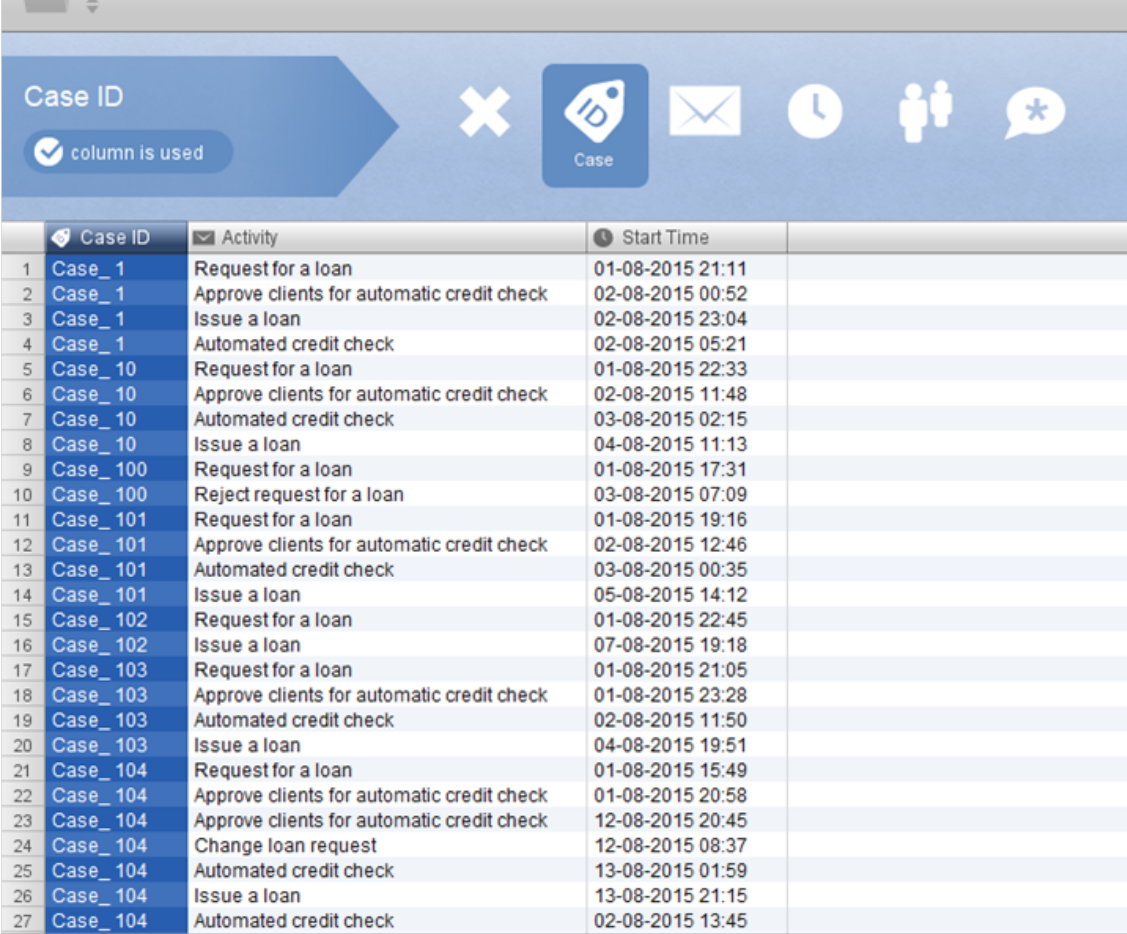
3.3.0.3 myInvenio

To import data, click on the 'Select data source' and upload it as shown in Figure 3.39. The steps are clearly explained in the tool. After the we map the relevant columns for our analysis. myInvenio provides other twelve extra data variables to select for analysis apart from case id, activity, timestamp, resource and role.

3.3.0.4 Minit

To import data, click on the 'Add process button'. This will start a process import wizard, which will guide you through all the necessary steps in order to import data.

Minit supports different types of data sources. We import data sources from files like csv, txt, xls, xlsx, xes, xml. If data is in zip format, they are compressed to lower the volume of the data, you do not need to uncompress it, can be directly loaded. Supporting native sql server imports. So if the data is in microsoft sql database then



	Case ID	Activity	Start Time
1	Case_1	Request for a loan	01-08-2015 21:11
2	Case_1	Approve clients for automatic credit check	02-08-2015 00:52
3	Case_1	Issue a loan	02-08-2015 23:04
4	Case_1	Automated credit check	02-08-2015 05:21
5	Case_10	Request for a loan	01-08-2015 22:33
6	Case_10	Approve clients for automatic credit check	02-08-2015 11:48
7	Case_10	Automated credit check	03-08-2015 02:15
8	Case_10	Issue a loan	04-08-2015 11:13
9	Case_100	Request for a loan	01-08-2015 17:31
10	Case_100	Reject request for a loan	03-08-2015 07:09
11	Case_101	Request for a loan	01-08-2015 19:16
12	Case_101	Approve clients for automatic credit check	02-08-2015 12:46
13	Case_101	Automated credit check	03-08-2015 00:35
14	Case_101	Issue a loan	05-08-2015 14:12
15	Case_102	Request for a loan	01-08-2015 22:45
16	Case_102	Issue a loan	07-08-2015 19:18
17	Case_103	Request for a loan	01-08-2015 21:05
18	Case_103	Approve clients for automatic credit check	01-08-2015 23:28
19	Case_103	Automated credit check	02-08-2015 11:50
20	Case_103	Issue a loan	04-08-2015 19:51
21	Case_104	Request for a loan	01-08-2015 15:49
22	Case_104	Approve clients for automatic credit check	01-08-2015 20:58
23	Case_104	Approve clients for automatic credit check	12-08-2015 20:45
24	Case_104	Change loan request	12-08-2015 08:37
25	Case_104	Automated credit check	13-08-2015 01:59
26	Case_104	Issue a loan	13-08-2015 21:15
27	Case_104	Automated credit check	02-08-2015 13:45

Figure 3.36: Event log identification in Disco

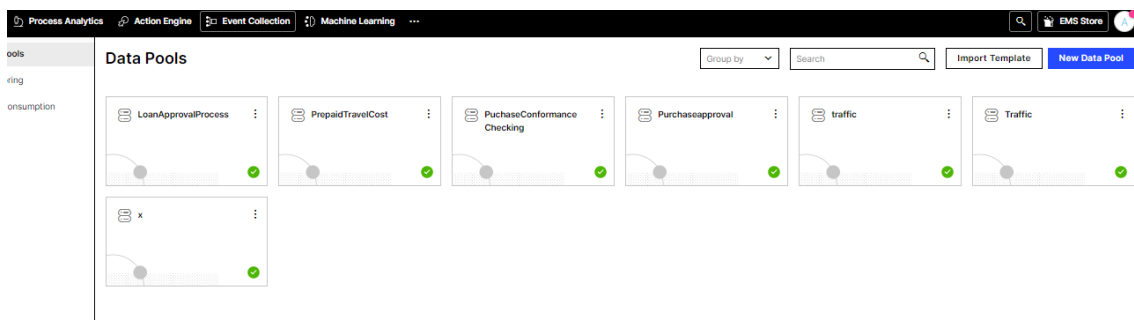


Figure 3.37: Event collection in Celonis

you can directly connect the database, choose a table or view or a stored procedure. Also mining supports ODBC driver, If we have an existing ODBC driver. Here, in this thesis, the data is imported from the file as csv for the analysis of process mining techniques in the tool.

After importing, mining tries to analyze the data. It tries to set the encoding, delimited and different parameters, which can be changed. It also show a preview of our imported data. In the next step, it will show you a preview of event log. Here we need to set all the process mining attributes that are necessary for process

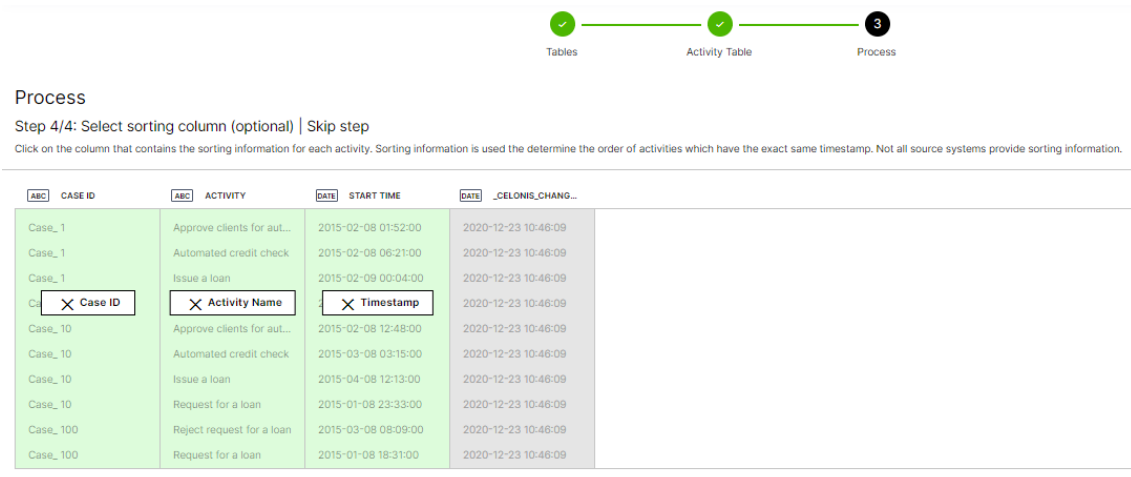


Figure 3.38: Event log collection in Celonis

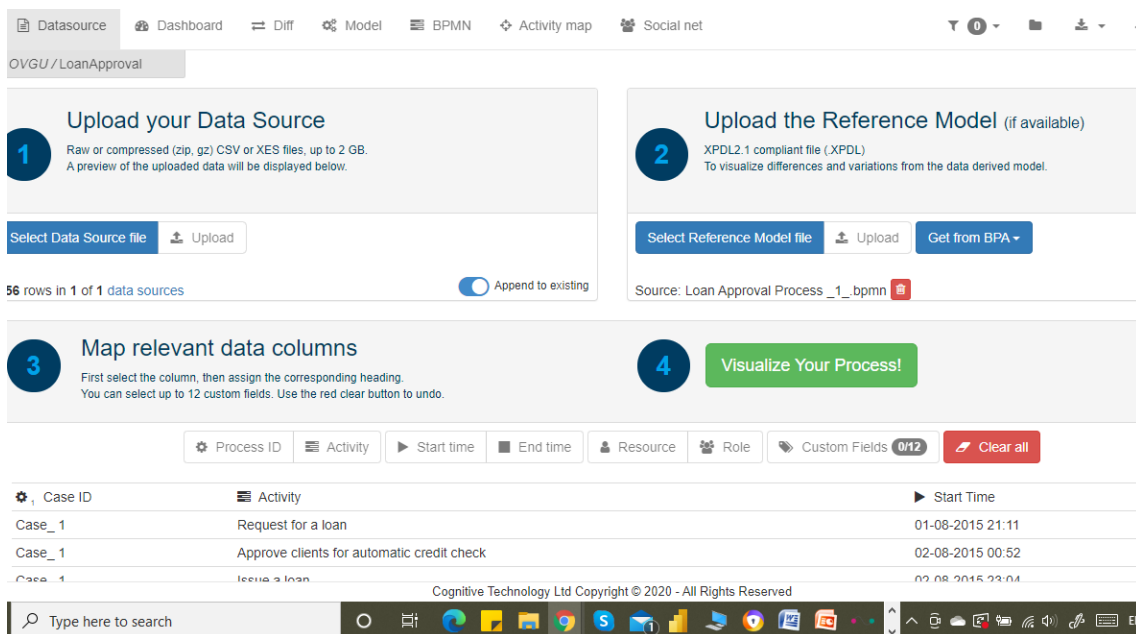


Figure 3.39: Event log collection in myInvenio

mining algorithms to work and for analysis. Minit supports all the standard and necessary attributes like case ID, activity, start and end timestamp, resources. Also we can enhance our analysis with more number of attributes. These attributes can be defined as event level attributes or case level attributes. The sense of event level attributes is that they change per event or through the case. The case level attributes on the other hand have the same value in the whole case. It also supports financial attributes. These attributes can be defined in 3 different levels, can be per case (e.g. amount on invoice which as the same value through whole case) or per event (e.g. fixed cost for executing certain event or activity) or per resources (e.g. human work rate per hour). The minit will automatically calculate the cost for the resource. Minit will automatically detects the attributes, if not we can also assign the attributes. Unwanted attributes can be excluded.

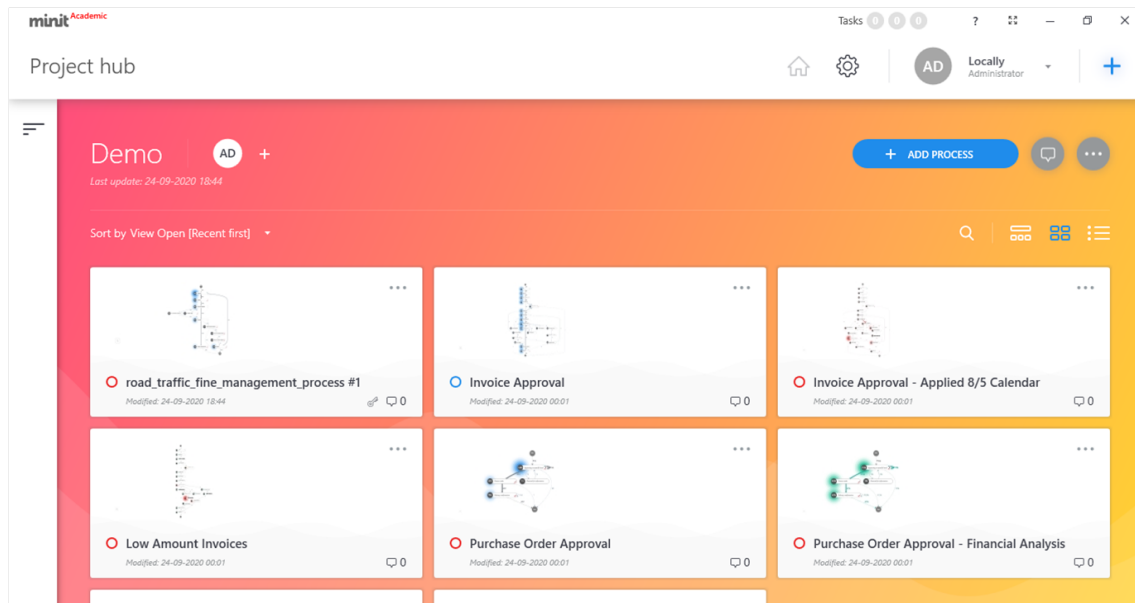


Figure 3.40: Minit start page

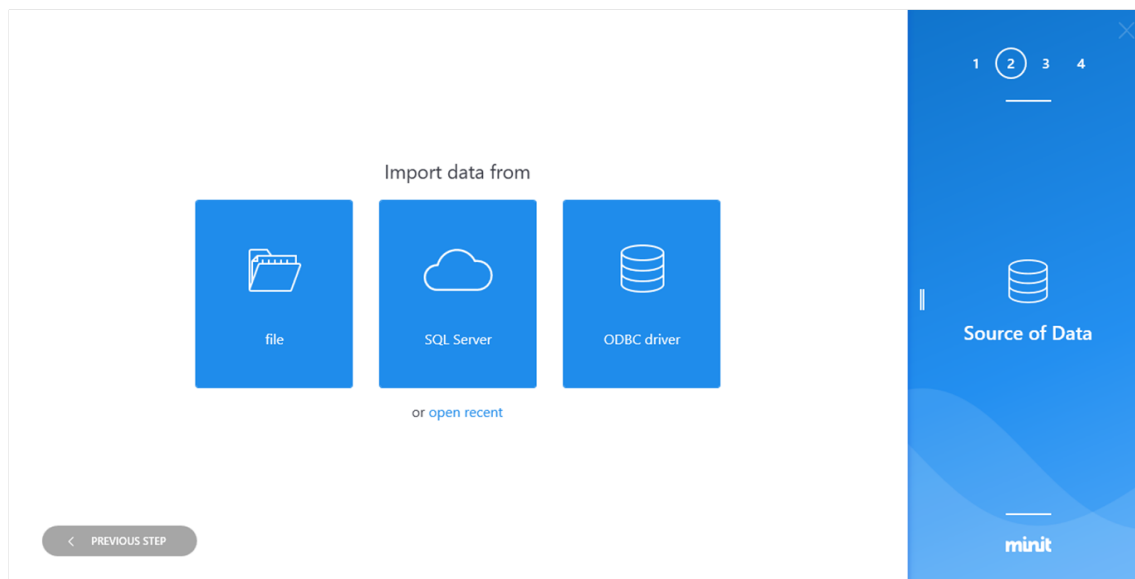


Figure 3.41: Minit Data Import Step 1

Minit will take 1000 rows from the given data, from this data it performs basic validation for the consistency, formatting, conflicts etc. This validation report will be shown in different column names. On the next page we have a review of attributes, here is the possibility to change attributes between case or event level. While setting attributes, we can set the data types between numeric and string. Also we can adjust timestamp pattern. Finally giving a name to the process. Here we have option with import full validation. Without validation is checking only the basic attributes.

Finally, the process is ready to analysis in the process explorer. This tool can be explained in various features as follows:

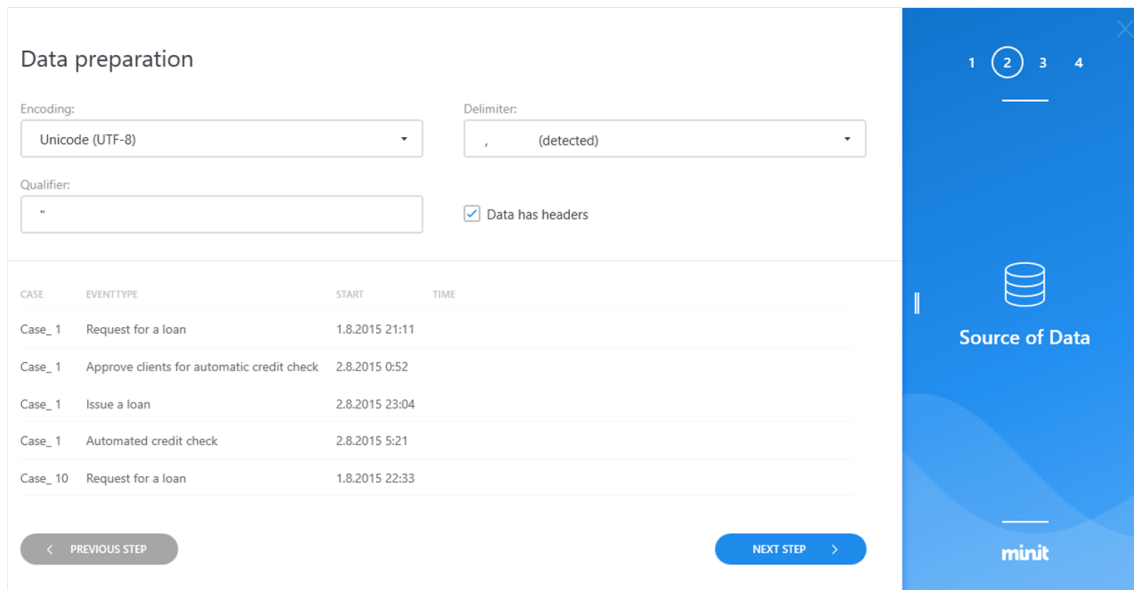


Figure 3.42: Minit Data Import Step 2

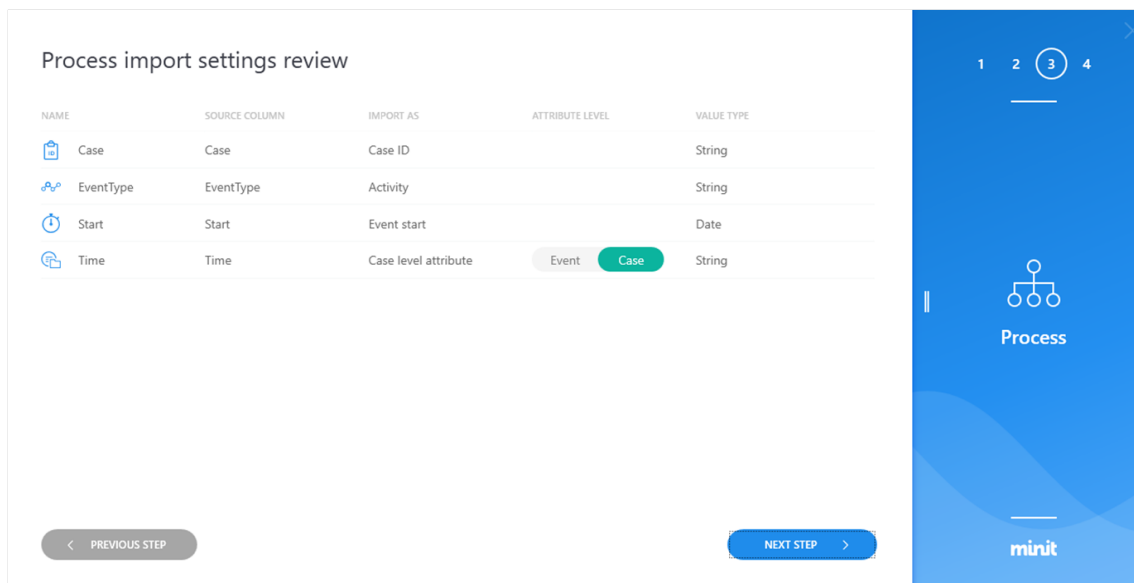


Figure 3.43: Minit Data Import Step 3

- **Dashboard:** Here, we can have the bird's-eye view of the all cases, activities, resources, role, in tabular and graphical forms.
- **Process Map:** The process map is showing frequency analysis. There are blue helix effects around the activities having different thickness and colouring effect represents the most frequent activities and paths.
- **Social Analysis:** by clicking on the social chart, we can see that it will automatically changes the look of the map. It represents all the effective users and the relations between them.

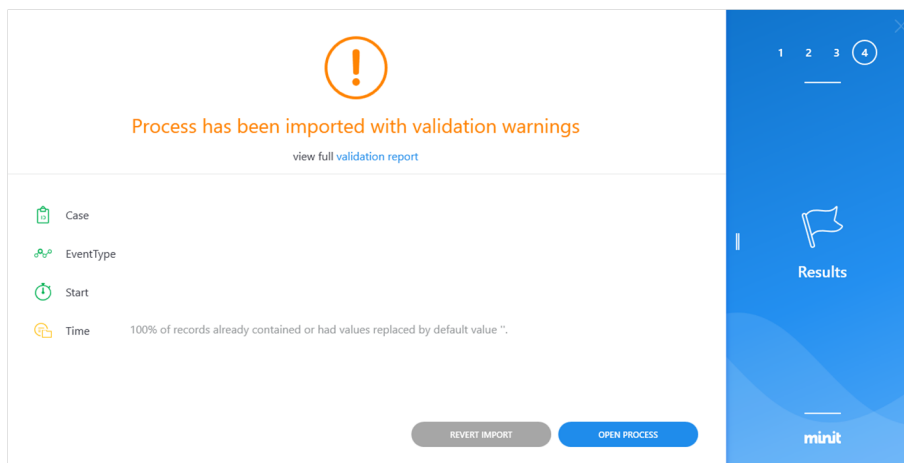


Figure 3.44: Minit Data Import Step 4

- Conformance check: we can either compare the process with reference model or we can compare with used process variants.

The process map is showing frequency analysis. There are blue helix effects around the activities having different thickness and colouring effect represents the most frequent activities and paths. To reduce the complexity of the process map and to focus on specific activities, a customize panel is present. To simplify the current map we can use the activities and path sliders that are in the customize panel. Customize panel is open by default when creating a process map. My reducing the path slider to zero percent, the process complexity is reduced which hides less frequent paths. The same is done with activity slider, hiding less frequent activities by reducing the activity slider to zero. So by reducing both the paths to zero percent, we can get the backbone of the process. The whole process map is interactive. We can drill down different activities or different paths by clicking on it to find in-depth details. For example, if we click more activities we can see predecessor and successor of the activities.

3.3.1 Conclusion

The data import and pre-processing of data is every simple in the tools. Most of the tools accepts CSV and XES file format. Celonis and Minit has integrated with high level of database connections.

4. Evaluation

4.1 Evaluation Criteria

As process mining is a new upcoming technology, there are many tools emerging in today's practice. So the selection of the tools are made on academic criteria and also based on tool availability in Europe.

At the first phase, a literature review was done to identify potential criteria from previous studies from the academic search engine sites SpringerLink, Google Scholar, Gartner monthly updates . The following search terms while doing literature review are as follows: Business process management and process mining, effects of process mining in Business process management, process mining, process mining tools, process mining tool comparison.

Second, the tools were tested with different use-cases in order to explore all the functionalities and to better understand what features and capabilities the tools offer. For the access of the tools, vendors were asked to grant access to all features in the academic license. The analysis is also based on all the available product documentation provided by the vendor.

In phase three, all the analysed results were compared and evaluated.

4.1.1 Testing Setup

In this section, we provide the necessary configuration that is essential to carry out our experiments. It includes details of both software packages and hardware components.

4.1.1.1 Software

The software tools that are selected in this thesis are cloud based and standalone softwares. The software tool testing was conducted using event logs of different use cases focusing on different industries processes with their respective happy path

reference models in BPMN format. In the current work, we focus on software functionalities and performance process of five different software systems i.e., Disco, Celonis, myInvenio, Minit, RapidMiner. We visualize the various software evolution of these process mining tools.

4.1.1.2 Hardware

All the experiments of our work were successfully analysed by downloading tools on a windows based systems and some tools are cloud based. We show the minimum hardware specifications of the tools provided by the vendors for academic licenses that are being used for our experiment [Table 4.3](#) .

4.1.2 Evaluation Results

In this section, we will evaluate our research questions based on the analysis of the process mining tools and from literature review. In particular, use-case 1 and use-case 1 from [Chapter 3](#) helps us to answer our first, second and third research questions, while literature survey from [Chapter 2](#) helps us to answer the next two research questions.

Here, we recollect our research questions of the thesis and they are as follows:

Research Question 1 : *How process mining improves Business Process Management?*

Research Question 2 : *What are the approaches of different process mining tools for pre-processing of data?*

Research Question 3 : *What is the behaviour of the tools for process mining techniques?*

Research Question 4 : *What are the additional features provided by process mining tools*

Research Question 5 : *What are the limitation of the inputs for the tool in academic license and What is the students adaptive learning environment of the process mining tools?*

4.1.2.1 Discussion on Research question 1

The traditional methods in business process management consumes a plenty of time and also depends on human biased. Process mining helps the business organisations to transform digitally. The evaluation is made from the analysis of literature survey. By using process mining tools and techniques in business organisations, we improves the operational data from the agent's machine so called tools i.e., process discovery, conformance check and enhancement which coverts automatically into process map for better analysis and targeting useful insights.

4.1.2.2 Discussion on Research question 2

The following part summarises the pre-processing of data in the process mining tools from analysis part [Section 3.3](#). The description of this given in three sub parts, i.e., extraction of data, transforming of data into event logs, and loading of data to the tool in .

Process Mining Tools	Process Discovery	Conformance Check	Enhancement
Celonis	Yes	Yes	-
Minit	Yes	Yes	-
myInvenio	Yes	Yes	-
Disco	Yes	-	-

Table 4.1: Process Mining techniques accepted by the tools in academic license.

Process Mining Tools	Flitering of data	Parallel Process Identification	Process Animation	Process Analytics & Statistics	Social Network	Dashboard Creation
Celonis	Yes	Yes	Yes	Yes	Yes	-
Minit	Yes	-	Yes	Yes	Yes	Yes
myInvenio	Yes	Yes	Yes	Yes	Yes	-
Disco	Yes	-	Yes	Yes		-

Table 4.2: Additional features in the tools for academic license.

4.1.2.3 Discussion on Research question 3

The analysis of process mining techniques from academic license tools are done in [Section 3.1](#). In summary, RapidMiner fetches the proM library. Here different algorithms are used in every technique. Remaining the tools and showed the automatic process discovery technique with different options. All the tools are ease to access the discovery technique. For conformance check, the tool Disco does not have the feature. All the other tools as conformance check technique functionality. Minit as a special feature as it can compare the same process map differentiated by different filters. The following [Table 4.1](#) summaries the research question.

4.1.2.4 Discussion on Research question 4

The analysis of different features of the tools are done in [Section 3.2](#) and [Section 3.2](#). The resources and role based analysis in use-case is performed. This helped to deep dive into every feature of the tool. All the tools provide filtering of data. The specific feature of myInvenio and Celonis is identifying parallel processing of data which helps in analysis two tasks happening at a time. All the tools provide process animation to identify step by step process with time. Also process analytics and social network are offered by all the tools. Minit has a extra feature, which helps the users to create own dashboard with preffer data. The [Table 4.2](#) describes the summary of all the features available in the process mining tools.

4.1.2.5 Discussion on Research question 5

The following section presents a summary of the findings as depicted in tables. In [Table 4.3](#), evaluation is described on the inputs provided by the literature survey, vendor of the tool for students in academic license. Furthermore, the support of a tool for a given feature is labeled by the 'Yes' and features that are not supported are labeled by the mark '-'. And learning environment for students on each tool are also described in the table.

Process Mining Tools	Data Input format	Data storage	Type of tool	License available period	Tool knowledge
Celonis	CSV, XLSX JSON, XES	Unlimited	Cloud-based	Unlimited	Academic Tutorials provided by vendor
Minit	CSV, XLSX XES	given based on student requirement.	Standalone based tool	Valid till date of enrollment by University	Academic Tutorials provided by vendor
myInvenio	CSV, XES	2 GB	Cloud-based	1 year	Through vendor website and demo sessions
Disco	CSV, XES	Unlimited	Standalone based tool	Unlimited	Through vendor website and demo sessions

Table 4.3: Inputs provided by the vendor for academic license of the tool.

5. Conclusion and Future Work

The main goal of this thesis was to compare and analyze process mining tools based on a well-defined framework. In order to approach the problem, a list of process mining problems with possible operations were acquired in the first round. Later, the mapping of problem to operation was categorized based on a case studies in the field of process mining. The criteria for the selection of the cases studies was based on its relevance to the categorization in to a scenario having a specific problem and implemented process mining operation. In another words, the studies mention a certain problem and utilize process mining tool in order to solve the problem. With the summary of the problem, possible operations and case studies, a framework has been developed. The developed framework select and focus mainly on the operations that use process mining tool. Since, our focus was to approach the comparative analysis on the basis of the supported/unsupported operations. Therefore, we were able to compare ProM, Disco and Celonis using a telephone repair event log sample to show whether the operation is supported or not in the respective tools. Despite the lack of in depth analysis on a real life log, the framework can be used to understand process mining tools from a different perspective unlike similar works. The approach of analyzing process mining tools based on process mining techniques/operations is much more precise for a practitioner expecting to understand which specific operations the tools can perform. Although the research has reached its aim, there were some unavoidable limitations. First, we evaluated the tool using one specific scenario which is based on core techniques of process mining tools as implemented by case studies. Therefore, it was not possible to generalize the results. Second, the fact that we analyzed the tools using the documentation that is insufficient in availability so that the interpretation of the findings might be affected. Finally, the conclusions of the comparison might be different if the evaluation has been performed by expert users of the respective tools or by the tool vendors themselves.

Disco provides a filtering mechanism that is more transparent and clearly represented compared to the other tools. ProM has a variety of filtering options compared to the other tools. Process discovery is one of the common operations that can be performed by any of the tools. It is possible to present the discovered model in various formats based on the end user preference. In case of ProM, the discovered model can

be converted to BPMN diagram, Petri net, heuristics model, transition system, inductive mined model and fuzzy model format. but for Disco and Celonis the output uses fuzzy mining model to display the discovered model. On the other hand, delta analysis can be performed regardless of the tools used since it is dependent on either the workflow model or the model originally used to configure the system initially. The usage of the process mining tools at this stage is for discovering the model that aligns together with the former model which is required to make delta analysis. In this study, we have addressed the research question by discovering possible problems that can be answered using process mining, followed by discovery of a list of process mining operations that solve the problem itself. In order to select core problems and operations, mapping aligned to a set of case studies was used. The same mapping is used to compare the tools. Based on the results we have acquired, ProM can support all core operations. Disco and Celonis cover partial operations. Disco is more suitable for both a beginner or expert users due to its simplicity to use and fast processing of event logs. Nevertheless, in the Disco point of view, Celonis can do similar tasks except Celonis provides live analysis of process and a visualization. On the other hand, Celonis convert event logs to database elements before analysing the data. Moreover, the timestamp must include a full set of activity timespan rather than starting and ending time in case of a given activity. The data backed up in an online or on premise database server. Based on overall observation, if there is a data that must be monitored on regular basis, Celonis is preferred. Nevertheless for a usage in projects that are either small or big with a requirement of fast, efficient and easy end user experience, Disco is preferable. Moreover, any of the tools can be used together while making the process discovery, bottleneck analysis and later can be exported as xes format in to ProM for further analysis. Regarding ProM, it can support all the operations specified and can be used with a variety of available mining algorithms. The results can be used by the practitioners and other users that are interested in using one of these 3 process mining tools.

Bibliography

- [1] W. Aalst. *Business Process Management*. 01 2017. doi: 10.1007/978-1-4899-7993-3_1179-2. (cited on Page 5)
- [2] R. Ahmed, M. Faizan, and A. Burney. Process mining in data science: A literature review. pages 1–9, 12 2019. doi: 10.1109/MACS48846.2019.9024806. (cited on Page 8)
- [3] C. Alvarez, E. Rojas, M. Arias, J. Munoz-Gama, M. Sepúlveda, V. Herskovic, and D. Capurro. Discovering role interaction models in the emergency room using process mining. *Journal of Biomedical Informatics*, 78:60–77, 2018. ISSN 1532-0464. doi: <https://doi.org/10.1016/j.jbi.2017.12.015>. URL <https://www.sciencedirect.com/science/article/pii/S153204641730285X>. (cited on Page 8)
- [4] C. W. G. A. J. M. M. W. W. M. P. A. Anne Rozinat, Ana Karla Alves Medeiros. *The Need for a Process Mining Evaluation Framework in Research and Practice*. Springer. (cited on Page 17)
- [5] G.-K. J. Badakhshan Peyman. Celonis process repository: A bridge between business process management and process mining. 9 2020. (cited on Page 12)
- [6] S. B. Sonawane and R. Patki. Process mining by using event logs. *International Journal of Computer Applications*, 116:31–35, 04 2015. doi: 10.5120/20447-2798. (cited on Page 7)
- [7] J. Caldeira and F. Brito e Abreu. Software development process mining: Discovery, conformance checking and enhancement. pages 254–259, 09 2016. doi: 10.1109/QUATIC.2016.061. (cited on Page 9)
- [8] W. Chomyat and W. Premchaiswadi. Process mining on medical treatment history using conformance checking. pages 77–83, 11 2016. doi: 10.1109/ICTKE.2016.7804102. (cited on Page 17)
- [9] K. Diba, K. Batoulis, M. Weidlich, and M. Weske. Extraction, correlation, and abstraction of event data for process mining. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, 10(3), 2020. doi: 10.1002/widm.1346. URL <https://doi.org/10.1002/widm.1346>. (cited on Page 7)
- [10] A. Esmita and E. Gupta. Process mining a comparative study. 12 2014. doi: 10.17148/ijarce. (cited on Page 17)

-
- [11] M. Faizan. Challenges and use cases of process discovery in process mining. *International Journal of Advanced Trends in Computer Science and Engineering*, 9:5164–5171, 08 2020. doi: 10.30534/ijatcse/2020/141942020. (cited on Page 8)
- [12] M. Glykas. *Business process management: theory and applications*. Springer, 2013. (cited on Page 6)
- [13] L. Ivančić. Business process modelling vs. process mining: A comparative study of methods and software tools. 05 2019. (cited on Page 10)
- [14] E. Kouzari and I. Stamelos. Process mining applied on library information systems: A case study. *Library Information Science Research*, 40(3):245–254, 2018. ISSN 0740-8188. doi: <https://doi.org/10.1016/j.lisr.2018.09.006>. URL <https://www.sciencedirect.com/science/article/pii/S0740818817303766>. (cited on Page 17)
- [15] A. Kurniati, O. Johnson, D. Hogg, and G. Hall. Process mining in oncology: a literature review. 12 2016. doi: 10.1109/INFOCOMAN.2016.7784260. (cited on Page 17)
- [16] J. Munoz-Gama. *Conformance Checking and Diagnosis in Process Mining*, volume 270. 01 2016. ISBN 978-3-319-49450-0. doi: 10.1007/978-3-319-49451-7. (cited on Page 9)
- [17] M. Pena and S. Bayona-Ore. Process mining and automatic process discovery. pages 41–46, 10 2018. doi: 10.1109/CIMPS.2018.8625621. (cited on Page 7 and 17)
- [18] A. Pourmasoumi and E. Bagheri. *Business process mining*, pages 117–124. 10 2017. ISBN 978-981-322-791-0. doi: 10.1142/9789813227927_0014. (cited on Page 7)
- [19] H. R’bigui and C. Cho. The state-of-the-art of business process mining challenges. *International Journal of Business Process Integration and Management*, 8:285, 01 2017. doi: 10.1504/IJBPIIM.2017.10009731. (cited on Page 6 and 10)
- [20] M. D. M. L. R. J. M. H. A. Reijers. *Fundamentals of Business Process Management*. 2018. ISBN 978-3-662-56508-7. doi: 10.1007/978-3-662-56509-4. (cited on Page 5)
- [21] R. Saylam and O. Sahingoz. Process mining in business process management: Concepts and challenges. pages 131–134, 11 2013. ISBN 978-1-4799-3343-3. doi: 10.1109/ICECCO.2013.6718246. (cited on Page 6)
- [22] S. Shershakov. Dpmine: Modeling and process mining tool. 01 2013. doi: 10.15514/SYRCOSE-2013-7-3. (cited on Page 7)
- [23] M. Stierle, S. Zilker, S. Dunzer, J. Tenschert, and G. Karagegova. Design principles for comprehensible process discovery in process mining. 06 2020. (cited on Page 8)

-
- [24] Y. S. K. Sungbum Park. A study of process mining-based business process innovation. *Procedia Computer Science*, pages 734–743, 2016. ISSN 1877-0509. doi: j.procs.2016.07.066. (cited on Page 7)
- [25] D. Viner, M. Stierle, and M. Matzner. A process mining software comparison. 10 2020. (cited on Page 17)
- [26] M. Weske. *Business Process Management - Concepts, Languages, Architectures, Third Edition*. Springer, 2019. ISBN 978-3-662-59431-5. doi: 10.1007/978-3-662-59432-2. URL <https://doi.org/10.1007/978-3-662-59432-2>. (cited on Page 6)
- [27] F. Yasmin, F. Bukhsh, and P. Silva. Process enhancement in process mining: A literature review. 12 2018. (cited on Page 10)