

A Report on
INTERNSHIP

PROCESS MINING VIRTUAL INTERNSHIP

*Submitted in partial fulfillment of the requirements
for the award of the degree of*

BACHELOR OF TECHNOLOGY

in

Computer Science & Engineering(AI&ML)

by

MAHABUBBASHA S

(214G1A3349)



Department of Computer Science & Engineering(AI&ML)



Srinivasa Ramanujan Institute of Technology
(AUTONOMOUS)

Rotarypuram Village, B K Samudram Mandal, Ananthapuramu - 515 701

2023-2024



Srinivasa Ramanujan Institute of Technology (AUTONOMOUS)

Rotarypuram Village, B K Samudram Mandal, Ananthapuramu - 515 701

Department of Computer Science & Engineering(AI & ML)



Certificate

This is to certify that the internship report entitled **process Mining Virtual Internship** is the bonafide work carried out by **MAHABUBBASHA S** bearing Roll Number **214G1A3349** in partial fulfilment of the requirements for the award of the degree of **Bachelor of Technology** in **Computer Science and Engineering(Artificial Intelligence & Machine Learning)** for April 2023 to July 2023.

Internship Coordinator

Mr. P. Veera Prakash, M.Tech(ph.D),
Assistant Professor & HOD of CSE

Head of the Department

Dr. G. Hemanth kumar Yadav M.Tech(ph.D)
Assistant Professor

Date:

Place: Ananthapuramu

EXTERNAL EXAMINER

PREFACE

All India Council for Technical Education (AICTE) has initiated various activities for promoting industrial internship at the graduate level in technical institutes and Eduskills is a Non-profit organization which enables Industry 4.0 ready digital workforce in India. The vision of the organization is to fill the gap between Academic and Industry by ensuring world class curriculum access to the faculties and students. Formation of the All-India Council for Technical Education (AICTE) in 1945 by the Government of India.

Purpose: With a vision to create an industry-ready workforce who will eventually become leaders in emerging technologies, EduSkills & AICTE launches ‘Virtual Internship’ program on Process Mining. This field is one of the most in-demand, and this internship will serve as a primer.

Company’s Mission Statement: The main mission of these initiatives is enhancement of the employability skills of the students passing out from Technical Institutions.

ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany the successful completion of any task would be incomplete without the mention of people who made it possible, whose constant guidance and encouragement crowned our efforts with success. It is a pleasant aspect that I have now the opportunity to express my gratitude for all of them.

It is with immense pleasure that I would like to express my indebted gratitude to my internship coordinator **Mr. P. Veera Prakash, Assistant Professor, Department of Computer Science and Engineering**, who has supported me a lot and encouraged me in every step of the internship work. I thank him/her for the stimulating support, constant encouragement and constructive criticism which have made possible to bring out this internship work.

I am very much thankful to **Dr. G. Hemanth Kumar Yadav, Assistant Professor, Department of CSE-Artificial Intelligence and Machine Learning**, for his kind support and for providing necessary facilities to carry out the internship.

I wish to convey my special thanks to **Dr. G. Balakrishna, Principal of Srinivasa Ramanujan Institute of Technology** for giving the required information in doing my internship. Not to forget, I thank all other faculty and non-teaching staff, and my friends who had directly or indirectly helped and supported me in completing my internship in time.

I also express our sincere thanks to the Management for providing excellent facilities and support.

Finally, I wish to convey my gratitude to my family who fostered all the requirements and facilities that I need.

MAHABUBBASHA S
(214G1A3349)

CONTENTS

Page No

Chapter 1: Introduction	01 - 04
1.1 Introduction to Process Mining	
1.2 Importance of Process Mining	
1.3 Types of Process Mining	
Chapter 2: Tools used for process Mining	05 – 07
Chapter 3: Foundation of Process Mining	08 - 13
3.1 Get to know Celonis Analysis	
3.2 Navigate to an Analysis	
3.3 Use the Variant Explorer	
3.4 Use the process Explorer	
Chapter 4: Learn to get data into Celonis	14 - 19
4.1 Introduction to SQL & PQL	
4.2 SQL vs PQL	
4.3 Get Data into the EMS	
Chapter 5: Applications	20 – 22
Chapter 6: Real Time Examples	23 – 25
Chapter 6: Internship Outcomes	26
Chapter 7: Conclusion	27
Internship Certificate	28
References	29

LIST OF FIGURES

Fig. No	Title	Page No.
1.1	Process Mining Model	01
1.2	Benefits of Process Discovery	03
3.1	Modules and Tables KPIs	09
3.2	Module of Use of Charts and Tables KPIs	09
3.3	Exercise of Process Mining	10
3.4	Example of Variant Explorer (DDA	
3.5	Six Selection Types	10
4.1	Comparison of SQL and PQL	15
4.2	PQL Engine	15
4.3	KPI's	19
5.1	Applications of Process Mining	20
6.1	Example View	23

LIST OF ABBREVIATIONS

NPS	Net Promoter Score
KPIs	Key Performance Indicators
PQL	Process Query Language
SQL	Structured Query Language

Chapter 1

Introduction

1.1 Introduction to Process Mining

Process Mining is often described as occupying the area between business process management and data mining. The speed, accuracy, and auditability these technologies deliver can result in significant cost savings and much faster time to market

- It is a technique in the field of process management that supports the analysis of business processes based on event logs and drives improved efficiency, effectiveness and compliance through its insights.
- Modern Process Mining technology quickly and reliably extracts information from event and transaction logs to visually depict real-time process models for current processes.

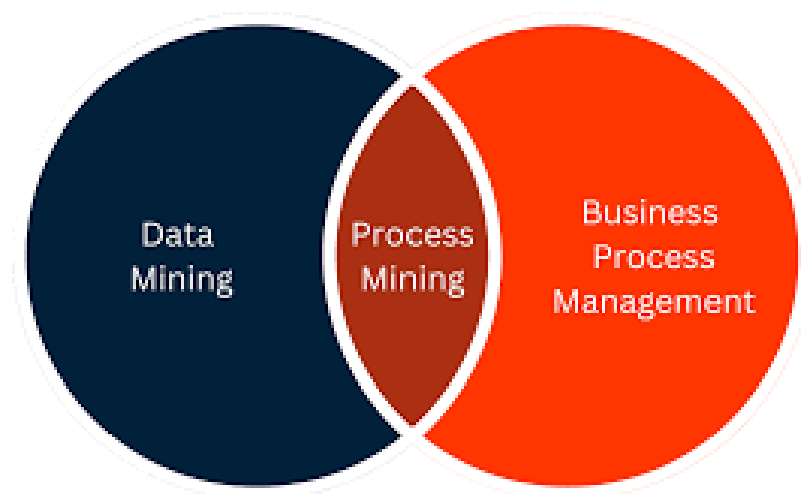


Fig 1.1: Process Mining Model

Definition

Process mining is a technique designed to **discover, monitor and improve real processes** by extracting readily available knowledge from the event logs of information systems.

1.2 Importance of Process Mining

Most of the Company leaders and executives may think they know the ins and outside of a business but lack the data to back up their claims. Process mining removes any guesswork and false assumptions by translating log data into visual models and representations. With a more accurate view of daily operations, leaders can understand what's truly working for the company and makes the understanding more efficient. Teams can use these insights to reallocate resources and better apply employees' time and energy. By making these adjustments, company leaders can improve the teams' performance and improve the customer experience, cut down on unnecessary costs and boost revenue streams.

1.3 Types of Process Mining

The three basic types of process mining

- 1.Process Discovery
- 2.Conformance Checking
- 3.Performance Analysis

1.3.1 Process Discovery

Process Discovery can be defined as a set of tools and techniques that are used to define, map and analyze the processes of an organization.

Process Discovery tools are machine learning-based tools that help organizations identify business processes, record all possible variations using machine learning algorithm and make recommendations for Automation. Not only do Process Discovery tools distinguish business processes that can be automated, but they also help design Automation workflows, making the mapping, planning and implementation of Automation quicker and more efficient



Fig 1.2: Benefits of Process Discovery

1.3.2 Conformance Checking

Conformance checking is a technique used to check process compliance by comparing event logs for a discovered process with the existing reference model (target model) of the same process. This technique is used to determine whether the target process corresponds to the actual process, highlighting deviations between the two.

Conformance checking helps ensure that all process deviations from the target process or reference model are identified, minimizing risk of audit problems or legal violations. Process deviations often require more resources than the target process or may have a negative impact on product or service quality.

1.3.3 Performance of analysis

The performance of a process or organization can be defined in different ways.

Performance analysis uses a variety of key performance indicators (KPIs), like lead time, average cost of resources used, or customer satisfaction scores. It looks at three specific dimensions: cycle time, cost, and quality. For each of these performance dimensions, different Key Performance Indicators (KPIs) can be defined.

High performing processes are based on strong process models. It's an indicator of a well-designed and well-implemented process. Process implementation cost is a key indicator of process performance.

CHAPTER – 2

TOOLS USED FOR PROCESS MINING

2.1 Tools used for Process Mining

Regardless of the industry, almost every business can benefit from process mining tools. For example, Gartner defines process mining as tools that “discover, monitor, and improve real processes (not assumed processes) by extracting knowledge from event logs readily available in today’s information systems.” Process mining is sometimes a built-in capability of a broader **Business Process Management (BPM)** solution suite, but it can also be found as standalone software.

However, choosing the right process mining capabilities for your company can be complicated. It requires in-depth research and often comes down to more than just the solution and its technical capabilities. To make your search a little easier, our editors have profiled some of the best process mining tools and software in one place. The editors have listed the companies in alphabetical order. Some of the top process mining tools are

CELONIS

Celonis claims to be the world leader in process mining.

A recent unicorn company after just 6 years, Celonis’ products have a clean and modern design.

Celonis focuses its sales on leveraging machine learning to calculate ways you can make your processes more efficient. Instead of just working out what is happening and whether there is conformance, Celonis’ Proactive Insights (PI) tries to calculate these decisions for you.

Features:

Process Discovery: Automatically creates process models by analyzing event logs, allowing you to visualize how processes are executed in reality.

Conformance Checking: Compares actual process execution with expected models, highlighting deviations and compliance issues.

Performance Analysis: Measures process efficiency, identifies bottlenecks, and provides insights for optimization.

Benefits: Celonis enables organizations to streamline operations, reduce inefficiencies, enhance compliance, and make data-driven decisions.

It brings in data from Salesforce, Oracle, Microsoft Dynamics and many more platforms.

This data is then visualized in an easy-to-understand way and presented along with recommendations for improvement.

DISCO

Explanation: Disco is known for its user-friendly interface and offers comprehensive process mining capabilities.

Features:

User-Friendly Interface: Designed for ease of use and accessibility, making it suitable for various users.

Process Visualization: Generates process maps and diagrams that visually represent actual process flows.

Root Cause Analysis: Helps identify underlying causes of process deviations and bottlenecks.

Benefits: Disco provides a simple yet powerful toolset for discovering insights from event data and driving process improvements.

QPR ProcessAnalyzer:

Explanation: QPR Process Analyzer offers process mining combined with advanced analytics capabilities.

Features:

Analytics Integration: Combines process data with advanced analytics to provide in-depth insights and predictions.

Real-Time Monitoring: Monitors processes in real time to identify deviations and opportunities for improvement.

Benefits: QPR Process Analyzer allows organizations to gain a holistic view of their processes, make data-driven decisions, and enhance operational performance. All of these are some of the popular and widely used processes mining tools offers unique features and benefits. The choice of the required tools depends on factors like your organization's specific goals, technical requirements, and the complexity of the processes you intend to analyze and optimize.

CHAPTER – 3

FOUNDATION OF PROCESS MINING

- Process mining is a set of techniques used for obtaining knowledge and extracting insights from processes by analyzing the event data generated during the execution of the process.
- This training track provides both the theoretical and applied foundations around Process Mining.
- Process mining reads this data, converts it into an event log, and then creates visualizations of the end-to-end process, along with insightful analytics.
- An event log contains each step performed during the process (the activity), the time at which the event occurred (the timestamp), and for which instance of the process (the case ID).
- Using this event log, algorithms generate a process model that shows the process as it really is - including the timing of each step and all variations

Process: A series of linked steps taken in order to achieve a particular goal.

Case: An item or object you follow through the process.

Activity: Events that take place during a process.

3.1 Get to know Celonis Analysis:

The Review and Interpret Analysis training track is designed for **data and business analysts**, **process experts**, and **process improvement specialists**. Keep in mind, this track is focused mainly on product know-how and less so on business acumen. If you'd like to complement your own experience in strategically identifying and prioritizing process inefficiencies, and planning for and implementing improvement measures, then we recommend you look at the Deliver Business Value with Celonis training track after completing this one.

they Deliver Business Value with Celonis training track after completing this one.

they Deliver Business Value with Celonis training track after completing this one.

Here's a sneak peak of what you'll experience in the Review and Interpret Analyses training track.

Self-paced Reading and Video Demos

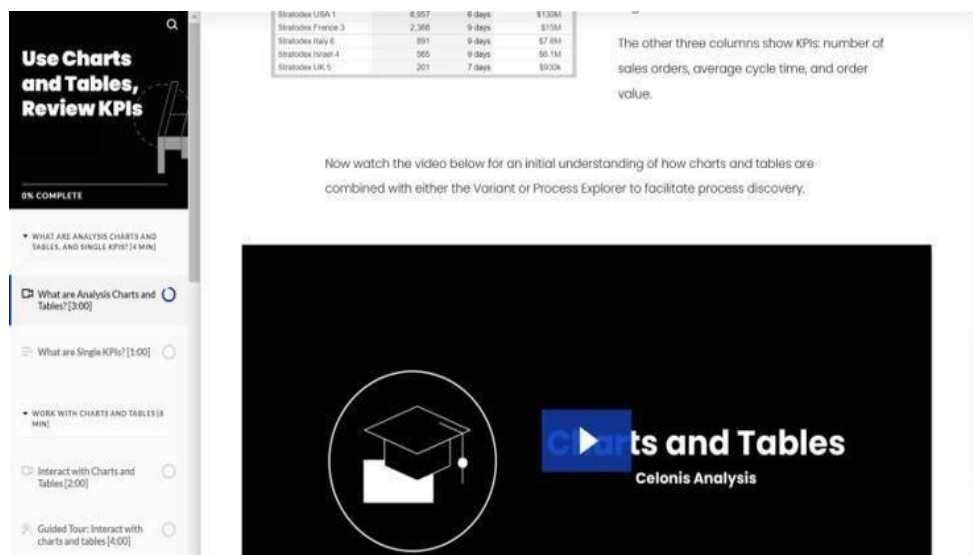


Fig 3.1: Modules and Tables KPIs

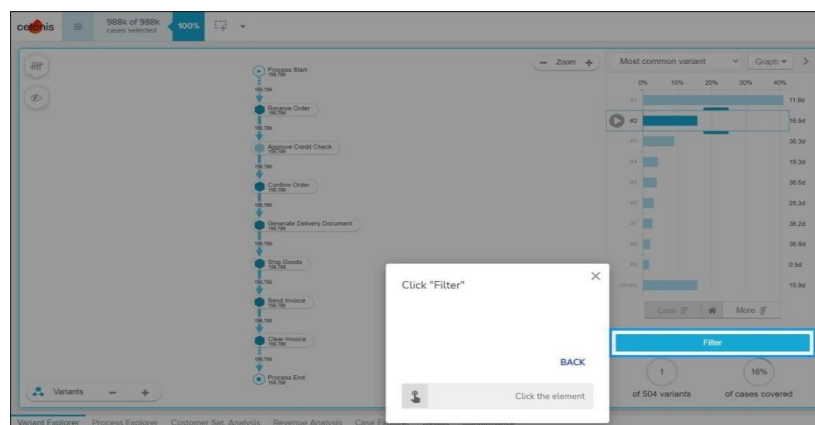


Fig 3.2: Module of Use of charts and tables KPIs

There are some of the Concepts to know in the following:

Guided Learning Tours

At certain points throughout the courses, you'll be prompted to complete public Analysis demo (no login required) that we've created just for this training. Imagine an onboarding buddy, showing you around Celonis Analysis.

Hands-on Exercises

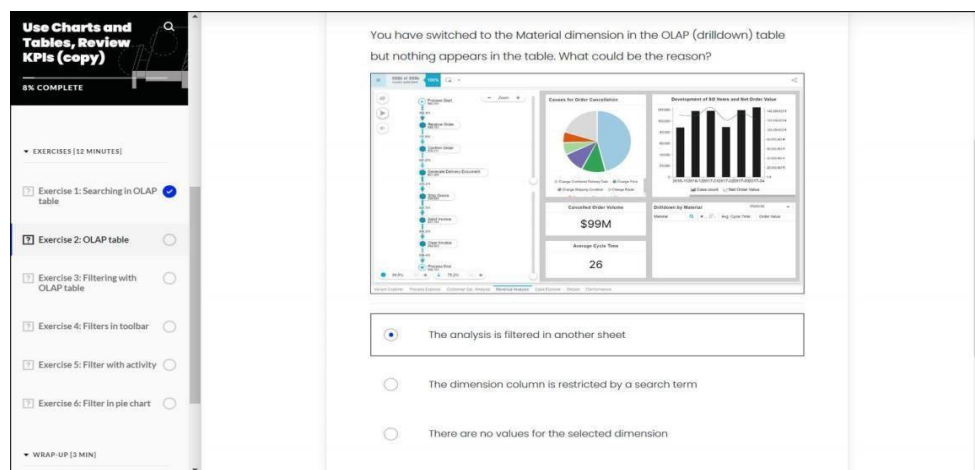


Fig 3.3: Exercise of Process Mining

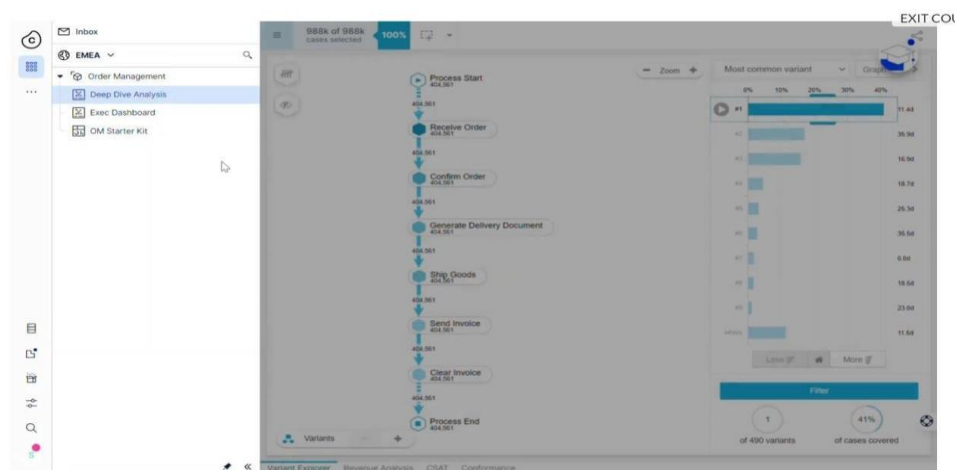


Fig 3.4 Example of variant Explorer(DDA)

Process-Specific Examples

By the end of the training track, you'll have gotten to know the most widely used Celonis Analysis components. Celonis Analysis tools **work the same way regardless of the process** but since every organization's process is unique with its own nuances, you'll need to **apply critical thinking** on your part to "connect the dots" so to say, between this training and your own process. We'll help you along the way though, prompting you to reflect and take notes on your ideas to refer to on the job

3.2 Navigate to an Analysis:

In the concept of navigate to analysis there are three words in it, that are space, package, analysis. These are arranged hierarchically.

These are used to represent the data pictorially. By seeing that and graph related to it. We can analyze the data.

These two screenshots are from two public Analysis demos. Depending on which training track you're completing, *Review and Interpret Analyses* or *Monitor KPIs in Analysis Dashboards*, you'll get to interact with one of these two demos.

3.3 USE VARIANT EXPLORER:

These two screenshots are from two public Analysis demos. Depending on which training track you're completing, *Review and Interpret Analyses* or *Monitor KPIs in Analysis Dashboards*, you'll get to interact with one of these two demos.

In short, Variant Explorer gives you a quick way to see whether most process cases follow an acceptable flow of activities or not and helps you develop your first analysis questions

3.4 USE PROCESS EXPLORER

Process Explorer is a Celonis EMS Analysis tool that helps you explore how process activities are connected. Instead of showing us specific process variants, it shows the

most common activities and connections.

Using our road trip analogy again, Process Explorer isn't showing the different routes(variants) that people took on a given trip (case). Process Explorer shows us which waypoints (activities) and roads (connections) are the most common along the journey

Process Explorer is extremely useful for finding infrequent activities, which can be difficult to spot using Variant Explore as these rare activities may not appear in common variants.

SELECTION VIEWS

Selection Views offer a more comprehensive set of options to filter on cases as compared to filtering you can do using the components in analysis sheets.

Selection Views Button

You can access the six Selection Views from anywhere in the analysis by clicking on the Selection Views button located in the analysis toolbar.

Then, you can select one of the six selection types depending on what you're looking for

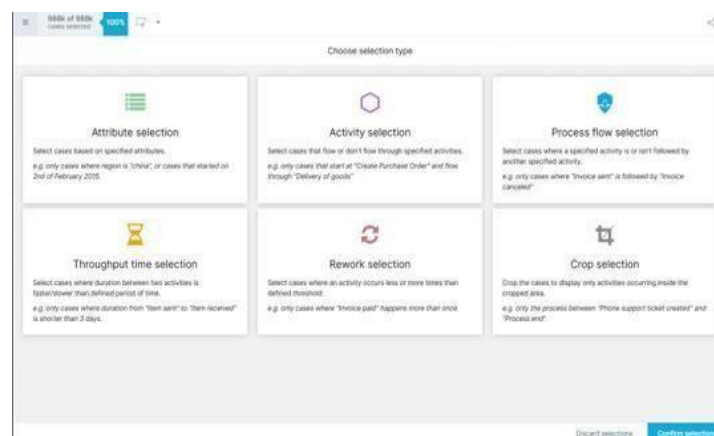


Fig 3.5: Six selection types

1. Attribute selection
2. Activity selection
3. Process flow selection
4. Throughput time
5. Rework selection
6. Crop selection

CHAPTER – 4

LEARN TO GET DATA INTO CELONIS

4.1 INTRODUCTION TO PQL AND SQL

The intention of Celonis PQL is to provide a query language for performing process mining tasks on large amounts of event data. It is based on a relational data model. The event and business data as well as all results are represented as relational data. Currently, the supported data types comprise STRING, INT, FLOAT, and DATE. Boolean values are not directly supported but can be represented as integers. Each data type can hold NULL values. In general, Celonis PQL treats NULL values as non-existing and ignores them in aggregations. Also, row-wise operations like adding the values of two columns will return NULL if one of its inputs is NULL.

Currently, Celonis PQL provides more than 150 different operators to process event data. Due to space limitations, we cannot sketch the full language. However, we can offer a brief overview of the major language features before we present selected examples to showcase the expressiveness of the language.

Language Overview

Even though Celonis PQL is inspired by SQL, there are major differences between the two query languages. Figure 6 shows these differences by comparing how to query the cases and the number of departments involved for all orders with a value of more than 1000 euros in both languages. Furthermore, it also illustrates the key concepts of Celonis PQL.

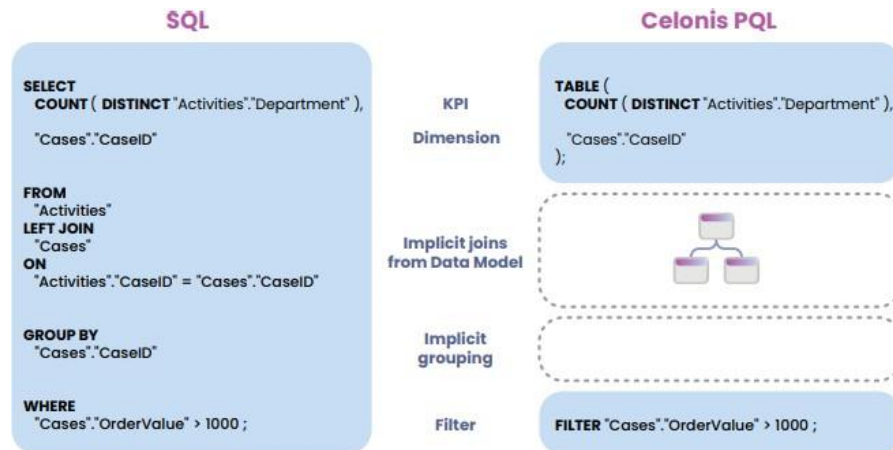


Fig 4.1 Comparison of SQL and PQL

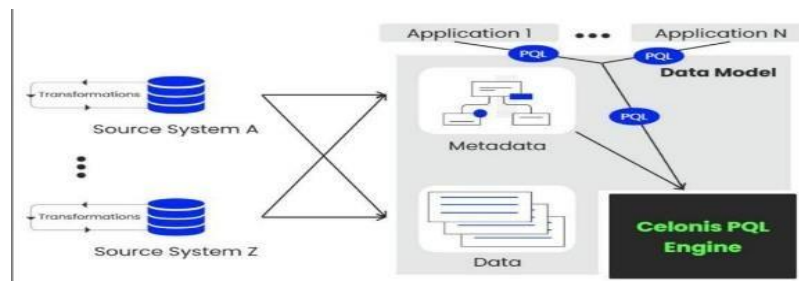


Fig 4.2 PQL Engine

SQL VS PQL

Both languages offer the possibility to filter rows. While SQL requires the user to formulate the filter condition in the WHERE clause of the query, Celonis PQL offers the FILTER statements which are separated from the TABLE statements but executed together. Splitting the data selection and the filters into different statements enables the user to define multiple filter statements in different locations inside an application, which then can be combined into the table statement to query the data.

Beyond this simple structure, Celonis PQL provides a wide range of

different operators which can be combined to answer complex business questions. The following list gives an overview of the most important classes of operators

Aggregations.

Celonis PQL offers a wide range of aggregation functions, from simple standard functions like count and average, to more advanced aggregations like standard deviation and quantiles. Most of the aggregation functions are also available as window-based functions computing the aggregation not over all values but over a user-defined sliding element window.

Data functions

These are operators like REMAP_VALUES and CASE WHEN which allow for conditional changes of values.

Date and time functions.

These functions enable the user to modify, project or round a date or time value, e.g., add a day to a date or extract the month from a timestamp. There are also functions to compute date and time differences

4.3 Get Data into the EMS

In this topic we will study about two types they

are Set up a data pipeline

Refine your Data Pipeline

In the set up a data pipeline again divide into sub parts they are

- 1.Data Integration basics
- 2.Connect to Systems
- 3.Extract Data
- 4.Transform Data
- 5.Load a Data Model

In the Refine your Data Pipeline divide into parts they are

- 1.Schedule Data Jobs
- 2.Monitor and validate your Datapipeline
- 3.Multiple Process and Systems
- 4.Boost your EMS SQL Transformations
- 5.Connect Custom processes
- 6.Quality Assuring your Data Pipeline

1.Data Integration basics

As a data engineer or analyst working in Data Integration (formerly known as Event Collection), you're responsible for bringing in clean, real-time process data into the EMS.

2.Connect to Systems

Connecting to source systems is your very first step to pull process data into the Celonis EMS. The EMS utilizes a broad set of technologies like message queues, Restful APIs, Soap APIs, direct database access, or system-specific solutions to connect.

Data Integration in the EMS

Data Integration is where you set up connections and your data pipeline. The main ways you can bring data into the EMS are:

Process Connectors Extractors Data Connections Extractor Builder File Uploads Data Push API Celoxtractor.

3.Extract Data:

No matter which system you're working with when extracting data, it's a good idea to first understand the business process to know exactly which tables you need.

Why don't we extract entire databases and make our lives simple?

For simple reasons—entire database extractions would:

take too long, be taxing on source systems, take up unnecessary cloud storage, and be

expensive

4.Transform Data:

The Activity table represents your process and always contains at least these three columns that map your process:

The object ID or case key,

the process steps or activities that took place for the different case keys and the timestamps or event time of each activity

In the Purchase-to-Pay process, the Purchase Order Item Number is the central case key we follow. Every Purchase Order Item goes through different activities such as creating the request, creating the item, receiving goods, and paying the invoice. And every activity has a corresponding event.

5.Load a Data Model

Just the Activity table on its own in a Data Model is not enough. To be able to drill down into case information, we need the Case table and other master data tables.

As you know, in Celonis, the Case table is a table containing one row for each case. In other words, this table contains a row for each "process path" (a path following a case) being analyzed in the application.

By specifying a Case table, you're able to use predefined KPIs in the Celonis analysis, such as a case count. The case count now specifically refers to table "EKPO" and will always count the number of entries in this table with respect to the applied filters. This is a screenshot from the Studio showing what is behind the KPI.

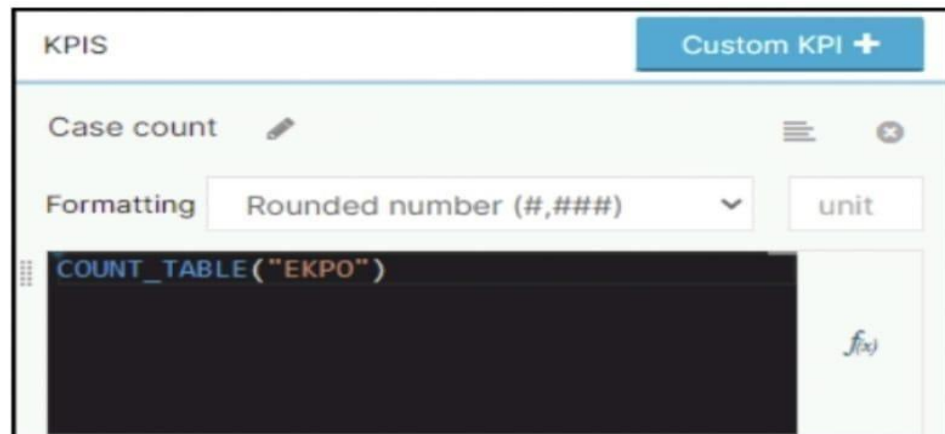


Fig 4.3 : KPIs

CHAPTER- 5

APPLICATIONS

Financial services, telecommunications, healthcare, and retail are just a few examples of industries where process mining can be used for business process management and process improvement. These sectors have a wealth of data that can be used as a starting point, and process deviations from their intended behaviour can have expensive repercussions



Fig 5.1 : Applications of Process Mining

1 . Financial Services:

Because of the rise in transaction volume and the digitization of more industries, aberrant activity is harder to detect using manual methods. Companies in the financial services sector have the chance to continually and thoroughly identify issues within high-volume processes thanks to process mining, which is a solution

to the increased regulatory and audit requirements.

2. Telecommunications:

As subscriber quantities increase and activations become more and more automated, there is a greater danger of unsuccessful activations. When telecom companies get more orders, process mining gives them the chance to identify pricey issues and client blowback in their Order-to-Activation processes.

3. Healthcare:

The risks associated with preserving population health and achieving individual patient journey objectives rise as data about patient experiences and results keep growing. Process mining supports the delivery of effective and high-quality end-to-end patient journeys for healthcare organizations dealing with the exponential growth of data, from before a first doctor appointment through treatment regimens to closed treatment cases.

4. Retail:

Due to technology or process problems, retail businesses have seen expensive consumer fallout from complicated e-commerce operations. Process mining assists merchants in ensuring that consumers can complete transactions efficiently and without issues despite increasing transaction volumes.

5 .Digital Transformation:

Process mining is frequently used in larger-scale digital transformation initiatives because it can give you the precise insights needed for process improvement, allowing systems to run more quickly, smoothly, and efficiently, as well as objective data-driven insights into the causes of delays and inefficiencies within business processes.

As a result, process mining may assist in identifying the digital transformation

opportunities with the greatest potential for value addition and determining whether or not transformation activities have really produced the desired results. To optimize returns on investments in projects for digital transformation, process mining becomes a crucial instrument.

CHAPTER – 6

REAL TIME EXAMPLES

1.HR Process Mining: Unlock the Power of Data-Driven HR Operations

Process mining is a game-changer for HR departments, enabling data-driven insights into employee performance, recruitment, and retention.

By analysing HR processes, organizations can identify bottlenecks, streamline workflows, and improve employee engagement.

For example, process mining can help HR departments:

- Identify the most effective recruitment channels
- Optimize onboarding processes
- Reduce the time to fill open positions



Fig 6.1 : Example View

2. Finance Process Mining: Transform Your Financial Operations with Data Analysis

Process mining is transforming financial operations, enabling organizations to gain insights into their financial processes and identify areas for improvement.

By analyzing financial data, organizations can optimize cash flow, reduce costs, and improve compliance.

For example, process mining can help finance departments

- Identify billing errors
- Streamline invoice processing
- Improve financial forecasting accuracy

3. Procurement Process Mining: Optimize Your Procurement Processes with Data Analytics

By analyzing procurement data, you can optimize supplier relationships, reduce costs, improve supplier performance, and improve compliance.

For example, process mining can help procurement departments

- Identify the most cost-effective suppliers
- Streamline purchase order processing
- Improve contract management

4. Sales Process Mining: Drive Sales Growth with Data-Driven Insights

Process mining is a powerful tool for driving sales growth, enabling organizations to gain insights into their sales processes and identify areas for improvement.

By analyzing sales data, organizations can optimize sales performance, improve

customersatisfaction, and increase revenue.

For example, process mining can help sales teams

- Identify the most effective sales channels
- Optimize lead generation processes
- Improve sales forecasting accuracy.

CHAPTER – 7

OUTCOMES

After completing this Training Track, you will be able to:

- Interpret process visualizations and leverage analyses to identify process inefficiencies.
- Conceptualize your process in terms of activities and cases.
 - Save an analysis selection for future reference and share it with your team; export visualizations and process data.
- Perform the basic tasks necessary to build Celonis analyses.
- Become familiar with Analysis Settings and Permissions.
- Publish analyses using best practices in version control.
 - Put your knowledge about the theoretical foundations of Process Mining into practice.

CHAPTER-8

CONCLUSION

In conclusion, process mining is a powerful and versatile technology that offers valuable insights into the inner workings of organizational processes. By analyzing event data generated during the execution of processes, process mining uncovers hidden patterns, identifies inefficiencies, and provides actionable recommendations for process optimization. This technology has the potential to drive improvements across a wide range of industries, including manufacturing, healthcare, finance, logistics, customer service, and more.

Process mining's ability to visualize process flows, detect bottlenecks, and pinpoint deviations from the ideal path enables organizations to make informed decisions aimed at enhancing efficiency, reducing costs, and improving overall performance. The real-time applications of process mining are particularly noteworthy, as they empower businesses to respond promptly to changing circumstances, address issues as they arise, and ensure that processes operate at their optimal levels.

As technology continues to advance, process mining techniques are likely to become even more sophisticated and integrated with other data-driven approaches, further enhancing their ability to drive process excellence. However, successful implementation of process mining requires a comprehensive understanding of both the technology and the underlying business processes. Organizations that embrace process mining stand to gain a competitive edge by harnessing the power of data-driven insights to continuously refine their operations and achieve higher levels of efficiency and effectiveness.

INTERNSHIP CERTIFICATE



References:

- [1] [Ref website:https://academy.celonis.com/learning-paths/execution-management](https://academy.celonis.com/learning-paths/execution-management).
- [2] www.analyticssteps.com/blogs/types-and-applications-process-mining
- [3] Login page : <https://academy-login.celonis.com/s/login/>.
- [4] [9 Hands-On Business Use Cases For Process Mining \(datarundown.com\)](https://datarundown.com/9-hands-on-business-use-cases-for-process-mining)