An Internship Report

on

PROCESS MINING VIRTUAL INTERNSHIP

Submitted 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)

by

C.Sreelatha **224G1A33B1**

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

(Artificial Intelligence & Machine Learning)

SRINIVASA RAMANUJAN INSTITUTE OF TECHNOLOGY
(AUTONOMOUS)

(Affiliated to JNTUA, accredited by NAAC with 'A' Grade, Approved by AICTE, New Delhi & Accredited by NBA (EEE, ECE & CSE))
Rotarypuram village, B K Samudram Mandal, Ananthapuramu-515701.

2024 - 2025

SRINIVASA RAMANUJAN INSTITUTE OF TECHNOLOGY (AUTONOMOUS)

(Affiliated to JNTUA, accredited by NAAC with 'A' Grade, Approved by AICTE, New Delhi & Accredited by NBA (EEE, ECE & CSE))
Rotarypuram village, B K Samudram Mandal, Ananthapuramu-515701.

Department of Computer Science & Engineering (Artificial Intelligence & Machine Learning)

Certificate

This is to certify that the internship report entitled "Process Mining Virtual Internship" is the bonafide work carried out by **C.Sreelatha** bearing Roll Number 224G1A33B1 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 10 weeks from April – June 2024.

Internship Coordinator

Mr. P. Veera Prakash, M. Tech., (Ph..D),MIEI,MCSI..

Assistant Professor

Head of the Department

Dr P. Chitralingappa, M.Tech,Ph.D.,

Associate 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 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. P. Chitralingappa, Associate Professor & HOD, Computer Science and Engineering (Artificial Intelligence & Machine Learning),** for his kind support and for providing necessary facilities to carry out the work.

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.

C.Sreelatha

(224G1A33B1)

CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

EMS	Executive Management System
KPIs	Key Performance Indicators
NPS	Net Promoter Score
PQL	Process Query Language
ROI	Rate of Investment
SQL	Structured Query Language

CHAPTER 1

INTRODUCTION

World and the organizations in it are full of processes. From purchasing to order management, organizations deal with complex, global and sometimes faulty processes on a daily basis. **Frictionless processes**, on the other hand, ensure:

- that you can find the right groceries at the grocery store,
- that planes land on time,
- that patient waiting times at hospitals are kept to a minimum.

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

The real world examples of process mining include procurement, order management, compliance, intelligent automation, digital transformation, KPI reporting, accounts payable, accounts receivable, auditing, IT development, service management, logistics, and many more.

Understanding Process Mining:

Process mining is a dynamic discipline that involves extracting meaningful insights from event logs to uncover, analyze, and enhance real-world processes within an organization. By bridging the gap between data science and process management, process mining enables us to understand the intricacies of how processes are executed and identify opportunities for improvement.

CHAPTER 2

FOUNDATION OF PROCESS MINING

- Process mining is a set of techniques used for obtaining knowledge and extracting insights from processes by the means of 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 goal.
Case: An item or object you follow through the process.
Activity: Events that take place during a process.

2.1 Get to know Celonis Analysis:

The **Review and Interpret Analyses**(opens in a new tab) 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 take a look at the **Deliver Business Value with Celonis**(opens in a new tab) 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

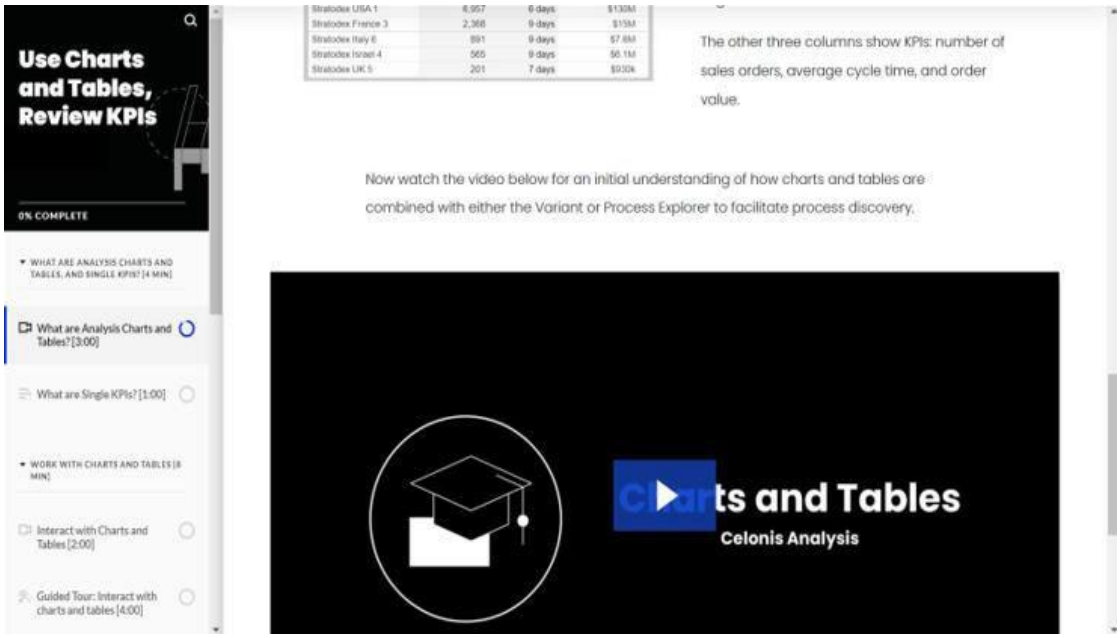

Fig 2.1 : Module of Use of Charts and Tables KPIs

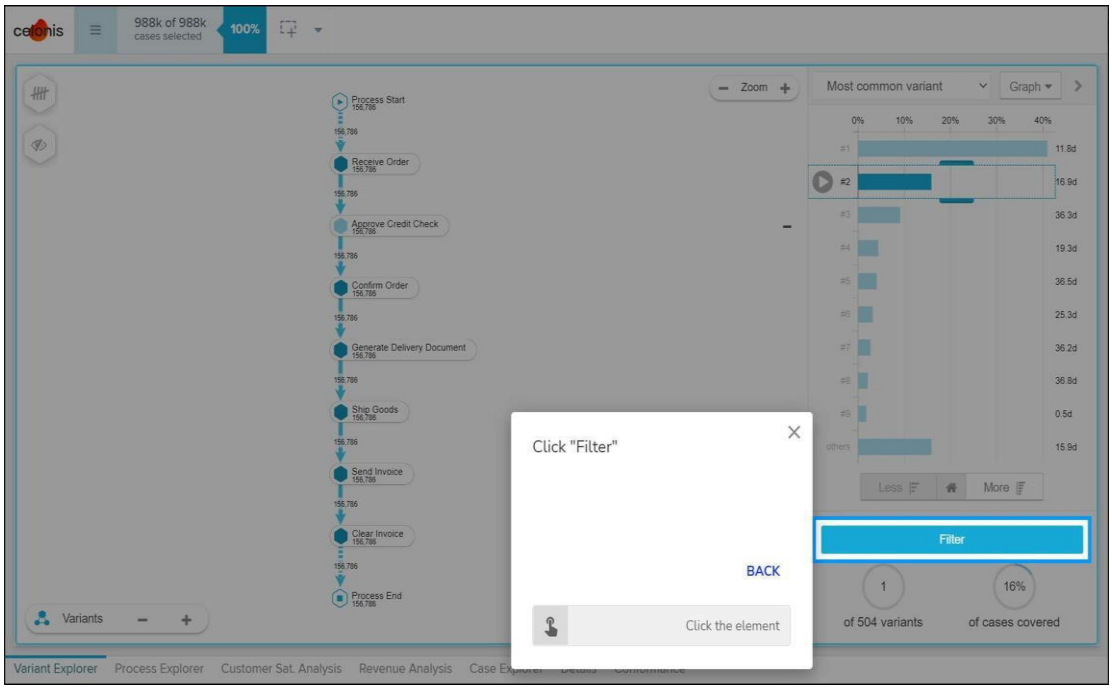

Fig 2.2: Variant Explorer

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 guided tours in a 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

At the end of most courses, you'll also be asked to answer questions in exercises that require you to reflect on the learning and even interact with the Analysis demo hand-on to answer questions.

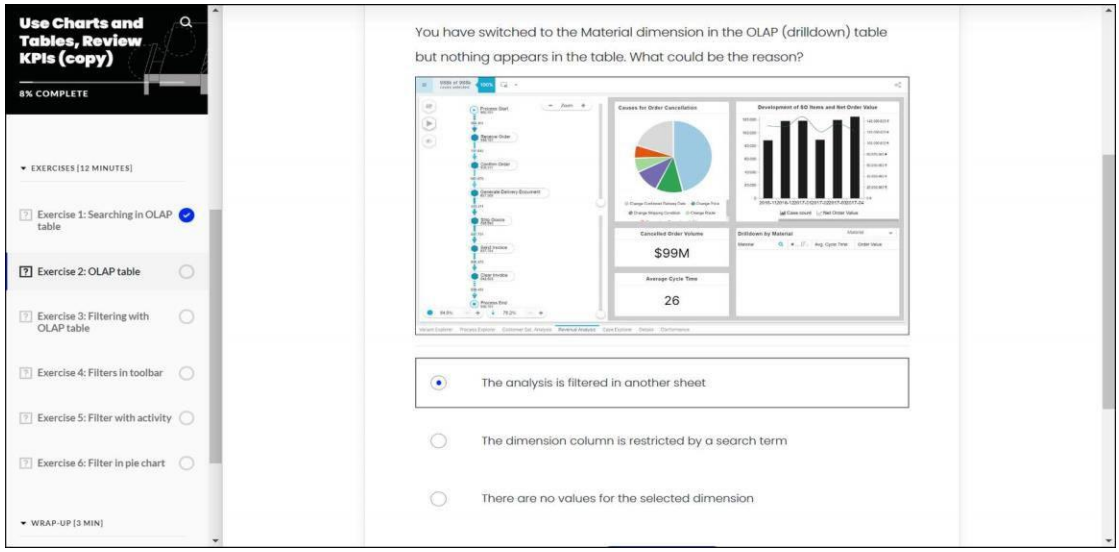

Fig 2.3: Exercise of Process Mining

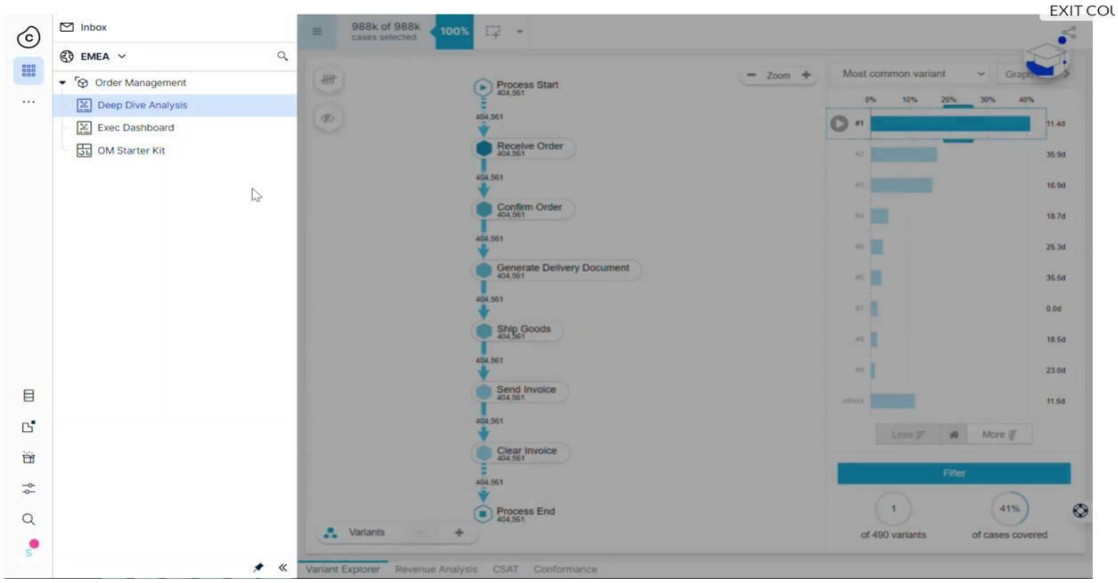

Fig 2.4: Example of Variant Explorer (DDA)

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.

2.2 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 in hierarchically.
These are used to represent the data in pictorially. By seeing that and graph related to it. We can analyse the dat.
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 actually get to interact with one of these two demos.

2.3 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 actually get to interact with one of these two demos.
Variant Explorer is a Celonis EMS Analysis tool that helps you explore how a specific process flows through your organization.

If we think about a process as a road trip, each process variant would be a potential route. Each activity within a process would be a waypoint along a route, and the connections between activities are like the roads that connect the stops. And, each trip a person makes along a particular route would be a case.

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.

2.4 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 actually 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.

2.5 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.

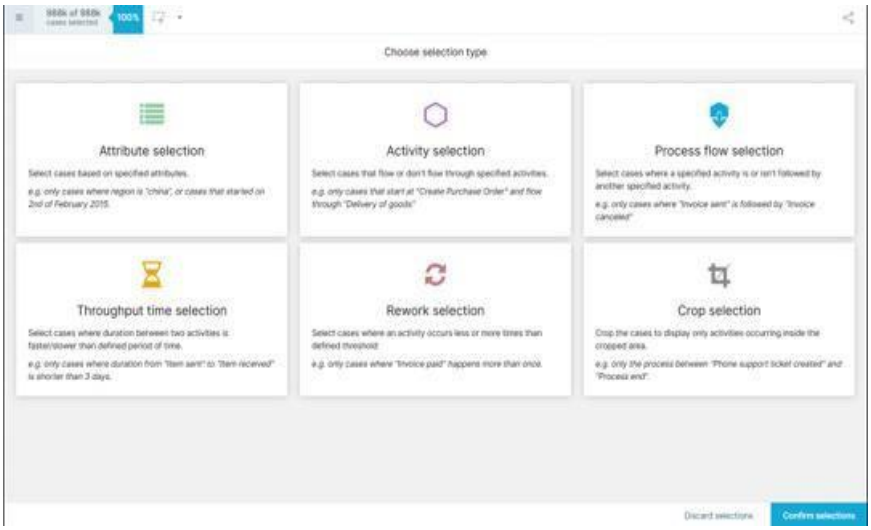

Fig. 2.5 Six selection types

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

THERE ARE FOUR IMPORTANT STEPS IN PROCESS MINING

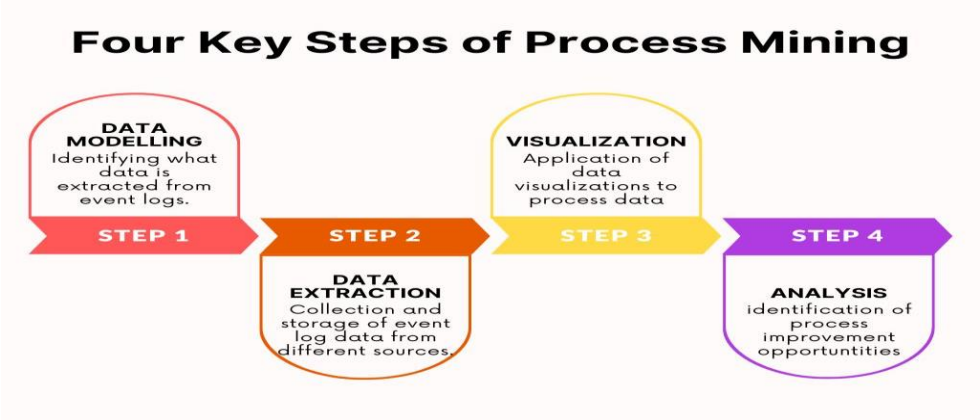

Fig. 2.6 Steps of process Mining

1.Data Modeling:

Identifying the data from which you extracted it from eventlog(where each event refers to a case,an activity,and a point in time)

2.Data Extraction:

Collection and storge of eventlog data from different sources.

3.Visualization:

To visualize the data it would be helpful.

4.Analysis:

Identification of process improvement opportunities

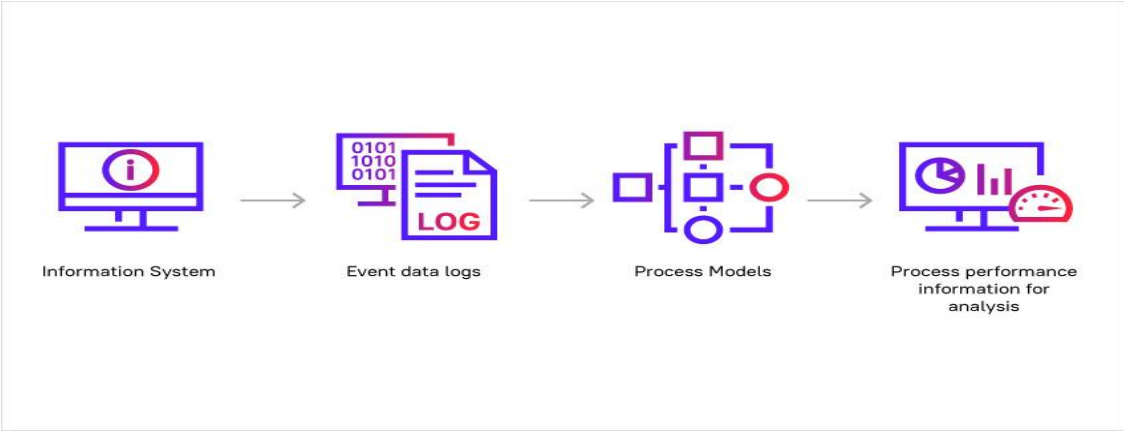

Fig. No.2.7 Process Flow

CHAPTER 3
Enhance your analysis building skills and learn to get into Celonis

3.1Write PQL Queries:

Design goals and the history of PQL:

PQL: Process Query Language

Celonis PQL is created along five design goals:

1.Simplicity
2.Flexibility
3.Eventlog Centered
4.Business focus
5.Frontend interaction

1.Simplicity:

i. Easy to use from business user to data engineer.
ii. Translates complex process questions into data queries.
iii. Makes process Mining accessible to every celonis user.

2.Flexibility:

i. Set of generic functions and operators.
ii. Combines wide range of combinations.
iii. Formulates any question, regardless of the process.

3.Event log centered:

i. Supports process mining functionalities**.**
ii. Associate cases to event.
iii. Dedicated process functions operate centered on the given event log.

3.Business Focus:

i. Combines process Mining and Business intelligence capabilities.
ii. augments event data with additional business information.
iii. variety of SQL function.

4.Frontend interaction:

i. Support of a graphical user interface(GUI)
ii. Leads to high acceptance, usage and user adaption.

Celonis PQL Engine:

As you can observe in the graphic below, Celonis PQL is an integral component of the Celonis Software Architecture. All Celonis applications use this language to query data from a data model.

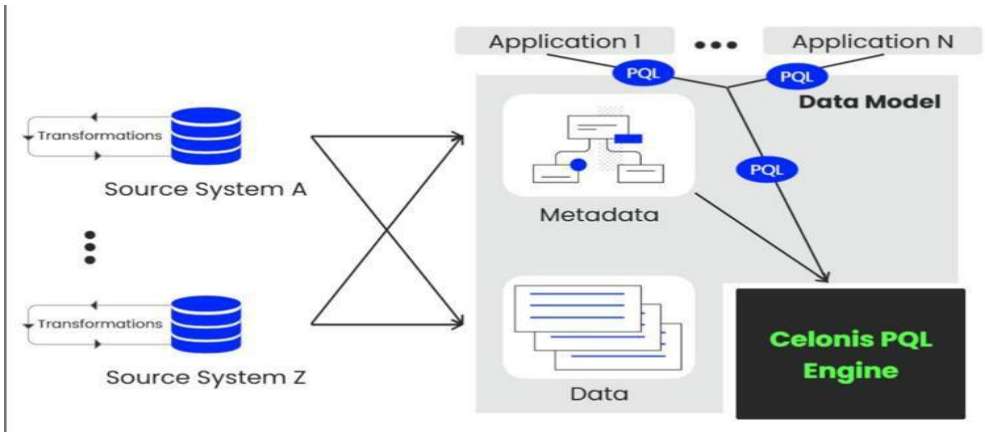

Fig 3.1: Software PQL Architecture

SQL VS PQL:

Both are different ,PQL cannot perform data manipulation and data definition language. Because it designed based on requirements.
PQL satisifies only the Aggregate functions.

3.1 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 Data pipeine
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

What is Data Integration for?
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. In other words, you build the data pipeline.

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!

1. 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 time.

In short, every Purchase Order Item has a unique case key that goes through various activities at different points in time. Together these three columns build the core of your process flow.

2. 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.

Data") that one case refers to one line in the “EKPO” table. That's why we have to select this table as the Case table.

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 a screenshot from the Studio showing what is behind the KPI.

Fig 3.2 : KPIs

CHAPTER 4
APPLICATIONS

Celonis:

Explanation: Celonis is a leading process mining platform that helps organizationsvisualize and analyze their processes using data from various systems. It offers process discovery, conformance checking, and performance analysis.

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, andprovides insights for optimization.

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

Disco (Fluxicon):

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 itsuitable for various users.

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

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

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

PROM

Explanation: ProM is an open-source process mining toolset used in both research and industry.

Features

Plugin-Based Architecture: Offers a wide range of plugins for different process mining tasks, allowing users to choose specific techniques.

Flexibility: Allows customization and experimentation with various process mining algorithms.

ProcessGold:

Explanation: ProcessGold offers process mining solutions with a focus on understanding and optimizing processes.

Features:

Advanced Analytics: Combines process mining with advanced analytics for deeper insights and predictive capabilities.

Process Discovery: Automatically creates visual process models from event data, helping organizations understand their operations.

Benefits: ProcessGold empowers organizations to enhance their processes, identify improvement opportunities, and predict future outcomes.

Minit:

Explanation: Minit specializes in simplifying process discovery and analysis for organizations.

Features:

User-Friendly Interface: Designed for non-technical users, making it accessible for process analysts and business users.

Performance Metrics: Provides performance indicators and metrics to measure process efficiency.

Benefits: Minit aims to make process mining accessible to a broader audience, enabling process optimization and efficiency gains.

QPR ProcessAnalyzer:

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

Features:

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

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

Benefits: QPR ProcessAnalyzer allows organizations to gain a holistic view of their processes, make data-driven decisions, and enhance operational performance.

Each of these process mining applications offers unique features and benefits. The choice of the right application depends on factors like your organization's specific goals, technical requirements, and the complexity of the processes you intend to analyze and optimize.

Some more real time applications are

1. **Supply Chain Optimization**: Process mining can be used to analyze and optimize supply chain processes in real-time. It helps identify bottlenecks, inefficiencies, and deviations from the ideal process flow, enabling organizations to make informed decisions and adjustments on the fly.

2**. Healthcare Process Improvement**: In healthcare, process mining can be applied to analyze patient treatment pathways, identify variations, and optimize resource allocation in real-time. This can lead to improved patient care and reduced wait times.

3**. Manufacturing Process Analysis**: Process mining can monitor and analyze manufacturing processes in real-time to ensure that production is running smoothly, detect anomalies or deviations from the standard process, and make immediate adjustments to prevent defects or disruptions.

4. **Financial Transaction Monitoring**: Process mining can be used in the financial sector to monitor and detect fraudulent activities in real-time. By analyzing transaction logs, it can identify patterns of suspicious behavior and trigger alerts for further investigation.

5. **IT Service Management**: Process mining can help IT departments optimize their service management processes by monitoring the flow of IT service requests, identifying bottlenecks, and streamlining incident resolution and request fulfillment.

6**. Logistics and Transportation**: Real-time process mining can be applied to logistics and transportation operations to monitor the movement of goods, track delivery routes, and optimize the overall supply chain to ensure timely and efficient deliveries.

7. **Customer Journey Analysis**: Process mining can provide insights into customer interactions and behaviors across various touchpoints in real-time.
This information can be used to improve customer experiences and optimize marketing and sales strategies.

8. **Energy Management**: Process mining can monitor energy consumption patterns in real-time to identify energy wastage, optimize energy usage, and reduce operational costs for industrial facilities and buildings.

9**. Emergency Response and Disaster Management**: During emergency situations, process mining can help emergency responders analyze data in real-time to allocate resources effectively, optimize response times, and make informed decisions to manage crises.

10. **Compliance Monitoring**: Process mining can assist organizations in real-time compliance monitoring by identifying deviations from regulatory processes and providing alerts to prevent potential compliance violations.

11**. Human Resources Process Optimization**: Real-time process mining can be used to analyze HR processes such as employee onboarding, performance evaluations, and leave

management. It helps ensure efficient and consistent HR practices.

12. **Retail Operations Optimization**: Process mining can analyze real-time data from retail operations to optimize inventory management, supply chain operations, and store layouts for better customer experiences and increased profitability.

These are just a few examples of how process mining can be applied in real-time across various industries. The technology continues to evolve, enabling organizations to gain insights, improve efficiency, and make informed decisions based on up-to-date process data.

CHAPTER -5
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 6
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:

- Ref website:https://academy.celonis.com/learning-paths/execution-management.
- [2] Execution management link : https://tinyurl.com/2pvuuzmd.
- [3] Login page : https://academy-login.celonis.com/s/login/.