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 and Engineering (AI & ML)

By

Sravanthi Kuruba

224G1A33A6

COMPUTER SCIENCE AND ENGINEERING

(ARTIFICIAL INTELLEGENCE & MACHINE LEARNING)

Srinivasa Ramanujan Institute of Technology
(AUTONOMOUS)

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

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

(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 and Engineering (AI & ML)

Certificate

This is to certify that the internship report entitled "Process Mining Virtual internship" is the bonafide work carried out by Sravanthi Kuruba bearing Roll Number 224G1A33A6 in partial fulfilment of the requirements for the award of the degree of Bachelor of Technology in Computer Science and Engineering (AI & ML) for 10 weeks from April – June 2024

Internship Coordinator

Mr. P. Veera Prakash M.Tech,(Ph.D).
Assistant Professor

Head of the Department

Dr.P.Chitralingappa M.Tech,,(Ph.D).
Associate professor

Date: 30/9/2024
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 Business Activities

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, 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, Computer Science and Engineering (AI&ML)**, for his kind support and for providing necessary facilities to carryout 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 project 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.

Sravanthi Kuruba
224G1A33A6

CONTENTS

List of Figures

LIST OF ABBREVIATIONS

ROI	Rate of Investment
EMS	Executive Management System
P2P	Purchase to Pay
KPI	Key Performance Indicator
PQL	Process Query Language
SQL	Structured Query Language

CHAPTER 1

Introduction

1.1 Introduction to Process Mining:

Process mining is a technique that analyzes event logs to discover, Monitor, and improve processes. It helps to understand how processes are actually executed and identify bottlenecks or inefficiencies. In process mining, event logs are used to reconstruct and visualize the actual flow of activities in a process. It can help identify process variations, bottlenecks, and areas for improvement. Various techniques and algorithms are applied to analyze the event data and provide insights into process performance and compliance. It's a powerful tool for optimizing processes and making data-driven decisions. Process mining is like having a magnifying glass for your business processes. It helps you understand how things really happen by analyzing event data. You can uncover hidden patterns, identify bottlenecks, and improve efficiency. It's like having a process detective.

What is Process Mining?

Process mining is a technique that examines a business's activities and enhances workflows for better productivity. It is an analytical system that observes, monitors and improves business functionalities. Process mining applies data science to discover, validate and improve workflows. By combining data mining and process analytics, organizations can mine log data from their information systems to understand the performance of their processes, revealing bottlenecks and other areas of improvement. Process mining, also known as performance or organisational mining, is a multidisciplinary system that detects, tracks and optimises the activities in a company for a betteryield. Process Mining combines data science and business process management using advanced algorithms and digital tools to resolve multiple systemrelated issues. Algorithms can identify and analyse a process's patterns, details and trends and form insights to reduce inefficiency.
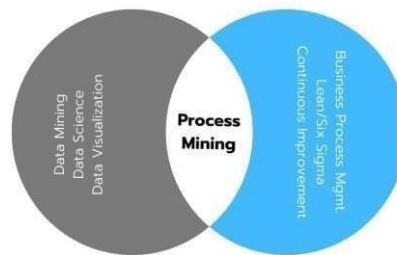
Fig 1.1: Process Mining

Process Mining achieves this union by taking the digital footprints that are created in IT systems and using them to reconstruct and visualize process flows. From here, Process Mining technology can identify patterns and deviations and ultimately eliminate bottlenecks. Now we will take a deeper look at what is required to reconstruct a process in this way.

With the help of these footprints, they can optimize, innovate and accelerate their products and services as well as the operations at the core of their business. It is therefore not surprising that companies across all kinds of industries are adopting Process Mining.

CHAPTER – 2
Foundations of Process Mining

Process mining is an analytical discipline for discovering, monitoring, and improving processes as they actually are and *not as you think they might be.* Process Mining works by extracting knowledge from event logs (also called digital footprints) readily available in today's information systems, in order to visualize business processes—and their every variation—as they run.

2.1 Review and interpret Analysis

2.1.1 Get to know Celonis Analysis:

The Celonis Execution Management System (EMS) extends process mining by executing on insights automatically and orchestrating your existing technologies.

Process, Activity, Case:

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

Fig 2.1: Process, Activity, case

Before interacting with your organization's process analysis, you'll want the right perspective on what makes business sense. With that, think about the following:

*   Think about your process' ideal set of activities, then list them in the desired order they should take place.

*   Write a list of undesired activities that negatively impact your organization's goals so that you can spot them in the analysis and then, using Celonis Analysis tools, identify potential root causes.

*   Another angle to consider: Is it possible that a case goes through desired activities but not in the right order?

*   Do some cases go through one activity more than once? Which activity or activities are usually associated with "rework?" • Is it possible a crucial activity is missing from some cases' journey?

2.1.2 Use the variant explorer

2.1.2.1 what is variant explorer?

As the name implies, using the Variant Explorer, you can discover all the processvariants—that is all the different ways the process flows in your organization**.**

The Variant Explorer is one of the Analysis tools to help you take an "exploratory" approach to find out how your process is performing.
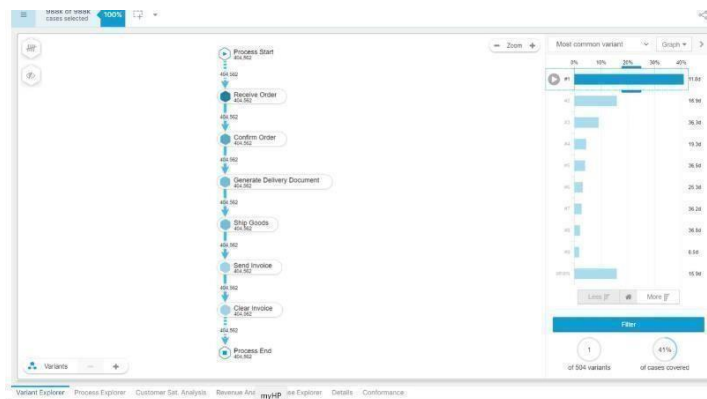
Fig 2.1.2: Variant explorer

2.1.3 Use the process Explorer

2.1.3.1 what is process Explorer?

*   The Process Explorer is another analysis tool to use when taking an exploratory approach. It's especially useful for quickly revealing activities beyond the most common ones.

- It also allows you to narrow your focus on a single activity especially useful for quickly revealing activities beyond the most common ones. It also allows you to narrow your focus on a single activity.
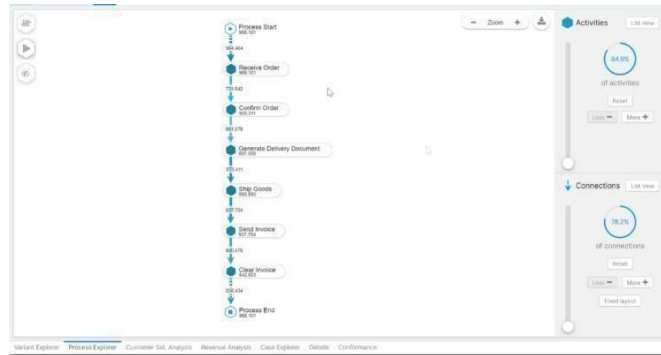

Fig 2.1.3: Process Explorer

Process Explorer is not showing the process path of any particular group of cases. It is showing the most common activities and connections, by default based on Case Frequency.

Once you learn about working with charts and tables, you'll find even more use cases for the Process Explorer. For example, after using the Process Explorer to filter on cases associated with an undesired activity (such as "Cancel Order") you can then drill into charts and tables to identify associated attributes (such as Sales Organization or Customer).

2.1.4 Use charts and Tables review KPIs

2.1.4.1 What are Analysis Charts and Tables?

- A dimension is a category of attributes; for example, the dimension "customer name" is a category for individual customer names. Other examples of dimensions, depending on the nature of the process, can include vendor name, sales organization, region, and material group.

- **Key Performance Indicators (KPIs)** are used to calculate and add aggregated values; for example, case count, order value, invoice value, throughput time, and automation rate.
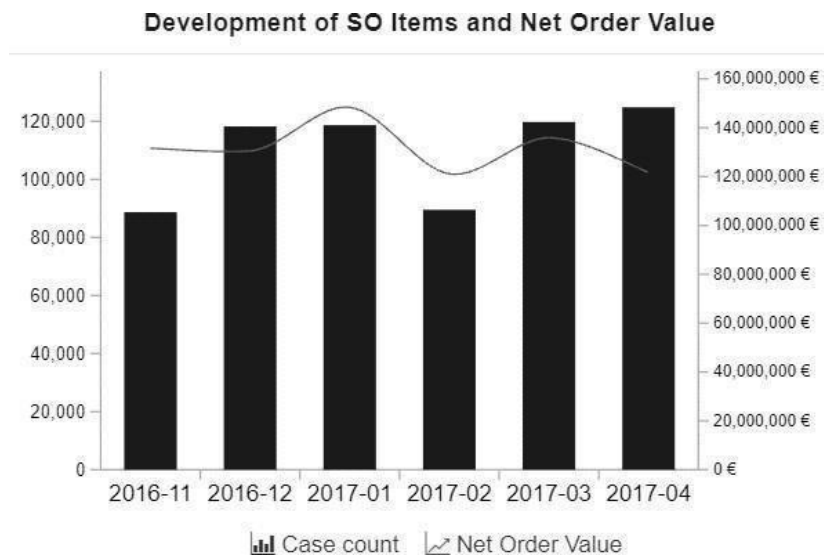
Fig 2.1.4: Development of SO items and Net Order Value

This chart shows the development of sales order items (KPI) and the corresponding net order value (KPI) over a period of time (dimension).

The x-axis displays the dimension, the creation date of sales order, grouped by months.

The two y-axes display the KPIs: The columns display the number of sales order items (case count) and the line displays the net order value.

2.1.5 Use Selection views

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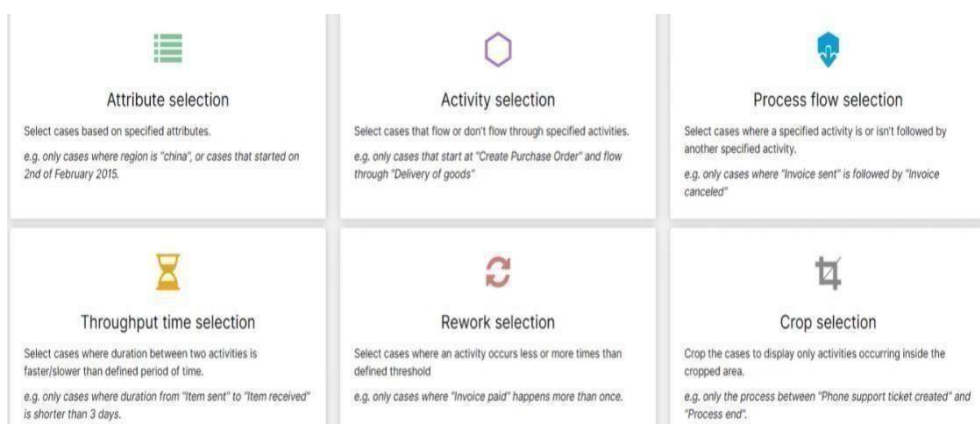

Fig 2.1.5: Selection views

2.2 Build Analyses

2.2.1 Get ready to build Analyses

As Celonis Analysis builder, the analyses you create are the interface between the data and the end-users. You help make digital business processes transparent so that users in operational and leadership roles can make data-driven decisions.

Everyone become familiar with best practices in gathering user requirements as well as in data visualization—to make data understandable, easy to work with, and visually appealing.

Beyond uncovering process execution gaps and causes of inefficiency, some of our customers choose to use Celonis services such as process automation (using Action Flows) and Apps (using Knowledge Models and Views) to address inefficiencies and maximize their organization's execution capacity. You can keep them in mind as you think about business users' needs and organizational efficiency, and what can be addressed with analyses, and what should be addressed with other solutions.
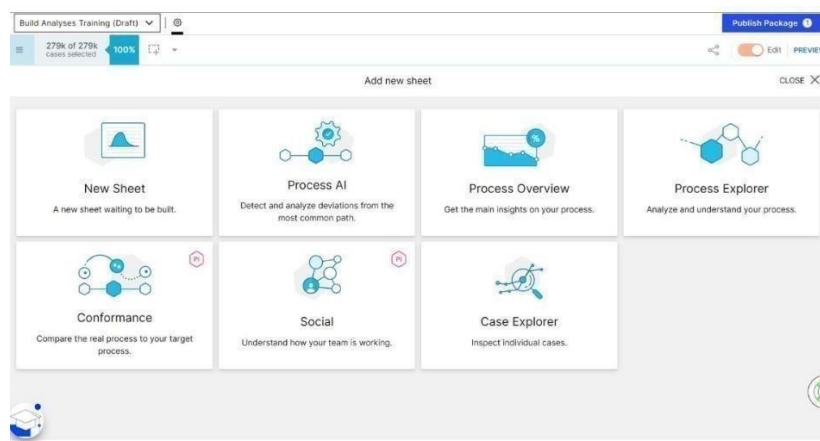

Fig 2.2.1: Analysis Sheet

2.2.2 Configure tables and charts in analysis

To add dimensions and KPIs to an analysis component, you'll need to work with the data tables in the analysis. In the SAP Purchase-to-Pay (P2P) data model we work with for this training, we have four tables regardless of whether we're selecting dimensions or KPIs.
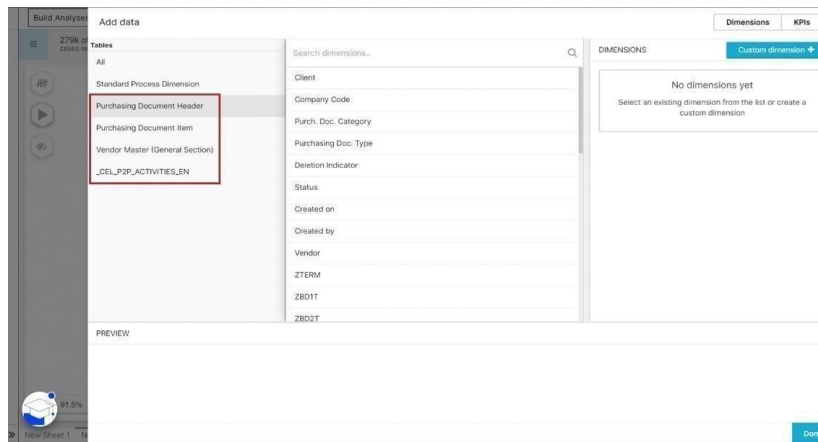
Fig 2.2.2: Data tables

• Purchasing Document Header (EKKO)

• Purchasing Document Item (EKPO)

• Vendor Master (LFA1)

• Activity Table also known as the Event Log

The principle of creating table and chart components in Celonis is always the same. Although, you'll need to select only one KPI to display in a Pie Chart. Celonis Analysis includes four types of single KPI components. The most common use cases for the single KPI component include the case count and net value. For both, you would use the Number KPI.

2.2.3 Configure single KPI, Selection and Design components

Celonis Analysis includes four types of single KPI components. The most common use cases for the single KPI component include the case count and net value. For both, you would use the Number KPI.

Fig 2.2.3: Single KPI componenets

What does Activity table contain?

At a <u>minimum,</u> the Activity Table contains these three columns.

- **Case**
- **Activity**
- **Event time**

Case: The "Case key" column specifies which object you follow through the process. As we are analyzing a Purchase-to-Pay process in this training, the objects we are following through the process are single purchase order items.

Activity: The "Activity" column contains all the activities that have been carried out for the purchase order items.

Event time: The "Event Time" column contains the point in time at which the activities have been conducted.

How is the Activity Table Created?

Unlike with the other tables in the data model, Data Engineers cannot directly extract the Activity Table from the source system. Instead, they identify the *digital footprints* in the source system and then consolidate them in the Activity Table. A digital footprint, at minimum, includes what has been done (activity), when it has been done (time), and which unique identifier is it associated with(eg sales order number).
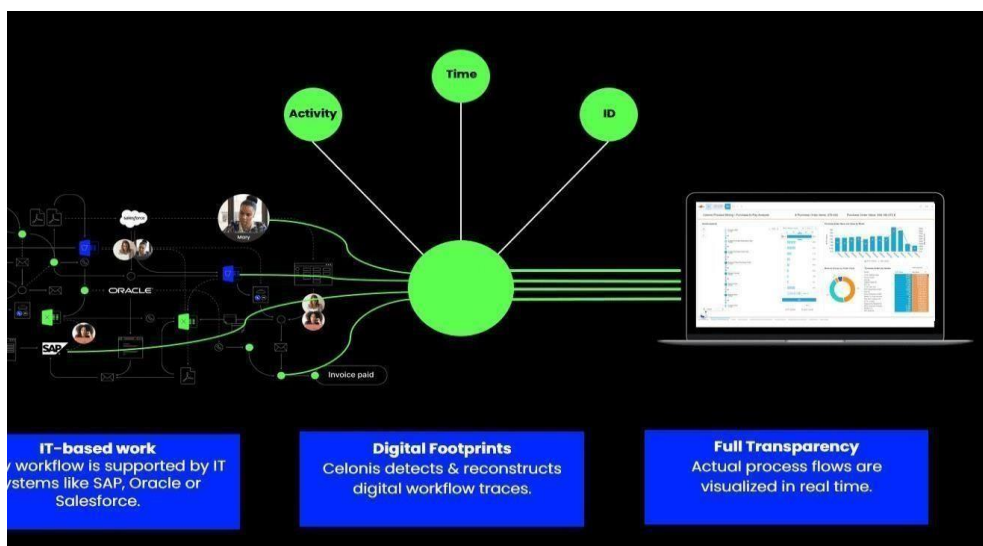

Fig 2.2.4: Activity table

PQL allows you to further customize Analyses beyond what is possible with the Visual Editor. But remember that you can use PQL across several Celonis services, not just Analysis!

2.3 Case Study: Pizzeria mamma mia

We are looking at the digitization journey of the Pizzeria Mamma Mia from the perspective of Giovanni, the owner of the business, and Martin, his Junior Manager.

The focus will be on the Order-to-Cash process which is the core process of the Munich- based business. The journey starts with the digitization of all process steps, continues with the discovery of inefficiencies and bottlenecks, and closes with recommendations for short-term enhancements as well as for sustaining the businesses success.

The case study is based on the Celonis Execution Management System (Academic version) and a data model provided by the Celonis Academic Alliance . The course will enable you to use Process Mining in a realistic environment in order to understand the functionalities and potentials of the Celonis EMS. You will learn how to build and interpret analyses in Celonis!

The Pizzeria Mamma Mia case study is built up in a modular structure consisting of different lessons
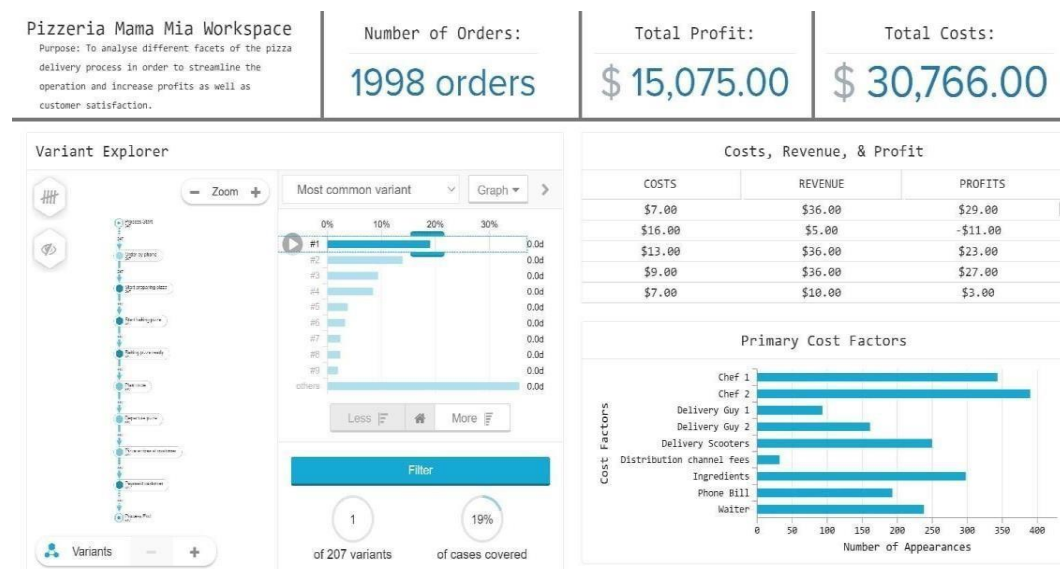
CHAPTER -3

RISING STAR   TECHNICAL

3.1 Write PQL Queries

3.1.1 Basic queries in PQL

In the course of digitization, an increasing number of log data is recorded in IT systems of companies worldwide. This data is precious, as it represents how business processes are running inside a company. Process Mining comprises data-driven methods to discover, enhance and monitor processes based on such data. The heart of Process Mining are the Event Logs.

Those Event Logs are a collection of process events that can be described by the following attributes:

- Case
- Activity
- TimeStamp

Case:

The case attribute indicates which process instance the event belongs to. A process instance is called a **case**, usually consisting of multiple events.

Activity:

The activity attribute describes the action that is captured by the event**.**

Time Stamp:

Each activity leaves a digital footprint with a timestamp, indicating precisely when each event took place. With the help of timestamps, we know precisely in which chronological order the different activities have runoff.

Celonis Software Architecture

The 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.Hover over the different components in the architecture overview to find out more.

Click on "example" to explore a data model of a procurement process.

- o Source System
- o Data Model
- o Data Model
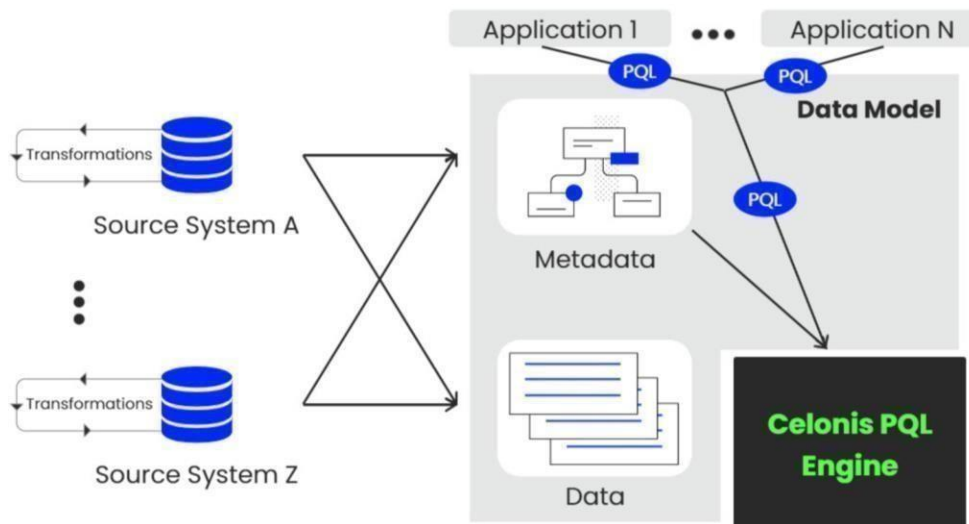- o Celonis PQL Engine
- o Applications

Fig 3.1: PQL Engine

SQL vs. PQL:

The main difference between PQLs and SQLs is that PQLs are qualified based on their behavior, while SQLs are qualified based on their intent. PQLs are more likely to convert into customers because they have already shown that they are interested in your product or service.

Once you understand the difference between PQLs and SQLs, you can use them to improve your sales process. You can use PQLs to identify leads that are most likely to convert into customers and then focus your sales efforts on those leads. You can also use SQLs to identify leads that are ready to buy and then close the deal as quickly as possible.By understanding the difference between PQLs and SQLs, you can create a more personalized sales experience for your leads. For example, you can send PQLs content that is relevant to their interests and needs, and you can send SQLs content that is focused on closing the deal.
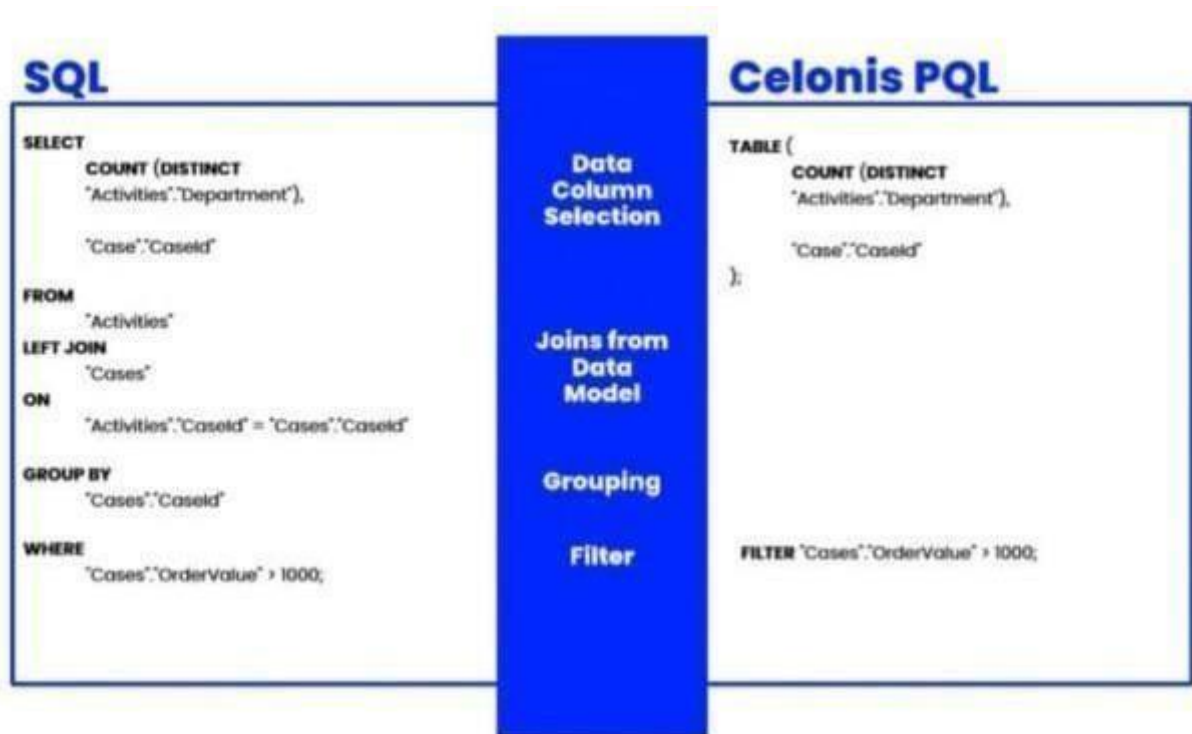
Fig 3.2 SQL vs. PQL

3.2. Get Data into EMS

In this topic we will study about two types they are :

- Set up a data pipeline
- Refine your data

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

- Data Integration basics

- Connect to Systems

- Extract Data

- Transform Data

- Load a Data Model

In the Refine your Data Pipeline divide into parts they are
- Schedule Data Jobs

- Monitor and validate your Data pipeine

- Multiple Process and Systems

- Boost your EMS SQL Transformations

- Connect Custom processes

▪ Quality Assuring your Data Pipeline

3.2.1 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.You're probably thinking, "why is that important"? Well, it's because the process data you bring in serves as the very foundation for others to analyze and act on.Without this data, no other activities can take place within the EMS and you won't be able to mine, improve, act on, or automate your processes.
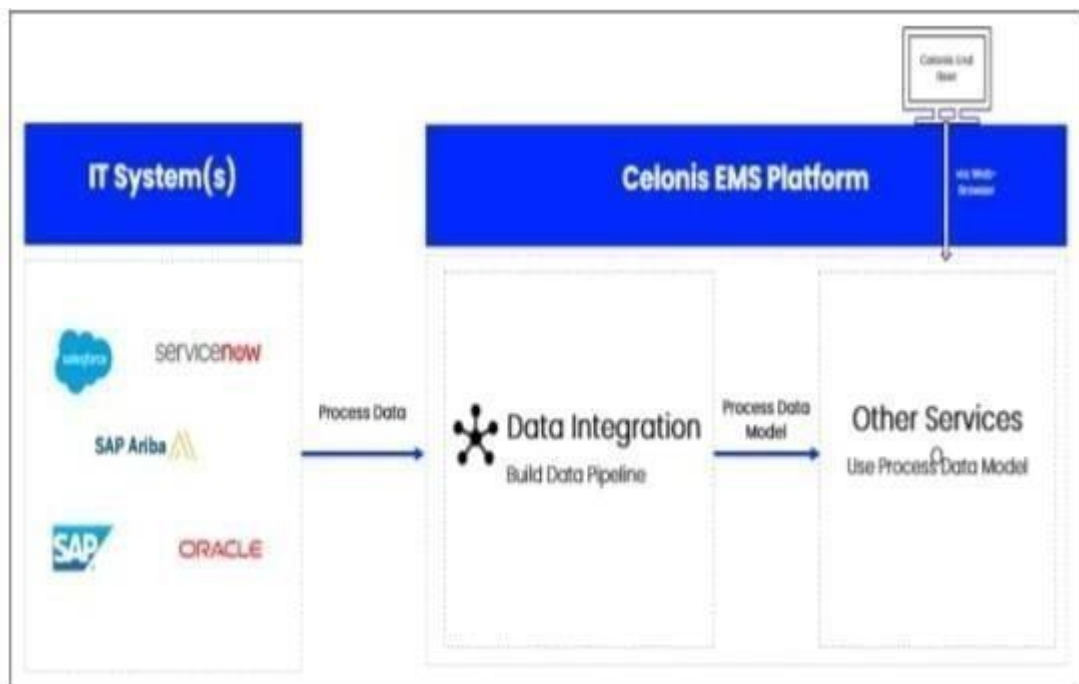
Fig 3.2.1 Data Integration

The Activity Table

Every activity or step happening to an ordered item is recorded in IT systems and leaves behind a Digital Footprint. Sometimes the activity data comes from one system but in most cases, it's in multiple systems. Either way, you bring the activities together in an "Activity Table". This Activity Table also called an event log, is the centerpiece of your process data and is what you build in Data Integration.

An Activity Table (Event Log)

Case Key	Activity	Event Time
123	Create Order	12-02-21
123	...	...
123	Pack Order	13-02-21
123	...	...
123	Ship Order	14-02-21
123	...	...

The Final Result is a Data Model

On the data side of things, that means you create one or more activity tables, case tables, and master data tables. You then connect them to one another in a Data Model.

Data Model

Activity Table **Order Management**

Case Key	Activity	Event Time
123	Create Order	12-02-21
123	--	--
123	Pack Order	13-02-21
123	--	--
123	Ship Order	14-02-21
123	--	--

Activity Table **Accounts Receivable**

Case Key	Activity	Event Time
456	Record Goods Issue	14-02-21
456	Create Invoice	14-02-21
456	Send Invoice	15-02-21
456	Receive Payment	16-02-21

Case Table 1

Case Table 2

Master Data Tables

Master Data Tables

CHAPTER-4
Real time examples of process mining

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 behavior can have expensive repercussions.

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.

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.

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.

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 rising transaction volumes.

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-5

Learning outcomes of internship

- Understand what process mining is and the basics of how it works.
- Understanding how process mining helps in Business world.
- Summarize what an event log is and why we need it for processing.
- Identify business use cases for process mining.
- Learn how to find training courses to get started.
- Understanding how to discover, analyses, and improve business process using data driven techniques.
- You will learn to extract insights from event logs, identify bottlenecks, inefficiencies, and opportunities for optimization

CHAPTER-6

Conclusion

In conclusion, process mining has the potential to bring significant benefits to organizations by improving efficiency, increasing visibility, and enabling data-driven decision making.

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.

By analyzing data from information systems, process mining can identify areas of inefficiency and help streamline processes. Process mining provides a detailed view of business processes, making it easier to identify bottlenecks and areas for improvement.Process mining can help organizations ensure they are following established procedures and regulations.

INTERNSHIP CERTIFICATE

N·E·A·T
National Educational Alliance for Technology

AICTE

अखिल भारतीय तकनीकी शिक्षा परिषद्
All India Council for Technical Education

EduSkills®
Nation Building Through Skills

AICTE - EduSkills
2024
VIRTUAL INTERNSHIP

Certificate of Virtual Internship

This is to certify that

Shahid Basha

Srinivasa Ramanujan Institute of Technology

has successfully completed 10 weeks
Process Mining Virtual Internship
During April - June 2024

Supported By **celonis**

A. Gebert
Angela-Sophia Gebert
Global Head of Academic Alliance
Celonis

Shri Buddha Chandrasekhar
Chief Coordinating Officer (CCO)
NEAT Cell, AICTE

Dr. Satya Ranjan Biswal
Chief Technology Officer (CTO)
EduSkills

Certificate ID :d9d69cd355c9541509dfcda74d14a92c
Student ID :STU65892f8299cfa1703489410

INTERNSHIP GRADE
B

GRADE- O (Outstanding):90-100 | E (Excellent):80-89 | A (Very Good):70-79 | B (Good): 60-69 | C (Fair): 50-59 | D (Average): 40-49 | P (Pass): 30-39 | F (Fail): Below 30

<u>REFERENCES</u>

- **Login page:** https://academy-login.celonis.com/s/login/
- https://academy.celonis.com/learn/course/introduction-to-process-mining/introduction-to-process-mining/course-outline?client=academic-alliance-celonis&page=2
- https://www.ibm.com/topics/píocess-mining