#### 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 (Data Science)**

by

#### SHAIK GUTTURU AFRA TAHASEEN

(214G1A3202)



# DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING (DATA SCIENCE)

# 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.

2023 - 2024

# 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 (Data Science)**



## Certificate

This is to certify that the internship report entitled "Process Mining Virtual Internship" is the bonafide work carried out by SHAIK GUTTURU AFRA TAHASEEN bearing Roll Number 214G1A3202 in partial fulfilment of the requirements for the award of the degree of Bachelor of Technology in Computer Science and Engineering (Data Science) for 10 weeks from May 2023 to July 2023.

#### **Internship Coordinator**

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

#### **Head of the Department**

Dr.P. Chitralingappa, MTech., (Ph.D.)., Associate Professor

Date: EXTERNAL EXAMINER

Place: Ananthapuramu

#### **PREFACE**

#### Brief overview of the company's history:

- Celonis is a German software company founded in 2011 by Bastian Nominacher, Martin Klenk, and Alexander Rinke.
- The company specializes in process mining, which involves analyzing and visualizing business processes based on data from IT systems.
- Celonis aims to help organizations improve their operational efficiency, identify bottlenecks, and enhance decision-making by providing insights into their processes. Since its inception, Celonis has gained recognition as a leading player in the field of process mining and has expanded its operations globally.

#### **Company's Mission Statement:**

Celonis' mission statement is centered around empowering organizations to unlock the full potential of their processes through advanced data analytics and process mining technology. The company aims to help businesses improve operational efficiency, drive innovation, and make data-driven decisions to achieve transformative outcomes.

#### **Business Activities:**

- 1. Process Mining Software.
- 2. Consulting and Services
- 3. Training and Education
- 4. Research and Development
- 5. Partnerships
- 6. Global Expansion
- 7. Customer Support

#### **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 & HOD, 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 (Data Science)**, 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.

**G. Afra Tahaseen** (214G1A3202)

## **INDEX**

	Contents	Page No.
List of Figu	ires	I
List of Abb	reviations	II
Chapter 1	Introduction - Process Mining	8
Chapter 2	Technology 2.1 Process Mining Utilization 2.2 Some of those challenges 2.3 Process mining use cases	11
Chapter 3	Applications 3.1 Sales 3.2 Banking 3.3 Customer Service 3.4 Education 3.5 Healthcare	14
Chapter 4	Modules Explanation  Module 1: Process Mining Fundamentals Process, activity, and case  Module 2: Variant Explorer  Module 3: Process Explorer Predecessor and Successor activities  Module 4: Using charts and Tables, Review KPI's  Module 5: Selection View, Selection View Buttons, Selection types  Module 6: Case Study (Pizzeria's Business)	18
Chapter 5	Real time example	25
Chapter 6 Conclusion	Outcomes	27 28
Internship certificate		29
References		31

## **LIST OF FIGURES**

Fig. No.	Description	Page No.
1.1	Process Mapping vs Process Mining	8
1.2	Core Capabilities of Process Mining	9
1.3	Various stages that get processed	10
3.1	Pie Chart representation of applications	14
4.1	An entire order through series of Activities	18
4.2	Case Frequency and Activity Frequency	21
4.3	Multi Variant Representation	22
4.4	Representation of Sales	23
4.5	Selection Types	24
4.6	Three Benefits that convinced	25
4.7	Order-To-Cash Process	25
4.8	Pizzeria's Data	26
5.1	Transformation of Pepsi Co	28
6.1	Pie Chart of benefits	27

## **LIST OF ABBREVIATIONS**

KPI Key Performance Indicators

NPS Net Promoter Score

PQL Process Query Language

SQL Structured Query Language

BPM Business Process Management

ROI Return on Investment

DSO Days Sales Outstanding

# CHAPTER - 1 INTRODUCTION

Our 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.

**Process Mining** is the leading new technology when it comes to talking about **algorithmic businesses** - in other words, businesses that **use** algorithms and large amounts of real-time data to create business value.



Fig.1.1: Process Mapping vs Process Mining

Compared to the traditional process mapping approaches, Process Mining technology solves the complexity and visibility problem.

• Imagine your favorite e-commerce app. Most likely, you are free to browse the products that the company is offering, and every now and then you are gently nudged to create an account in order to actually purchase something. In order to actually become a registered user, you often need to provide some information about you (the amount of information greatly varies per service). After you go through a few screens, you are ready to go and purchase. What happens in the background is the work of data scientists/analysts who analyze the user journey.

## The core capabilities of Process Mining:

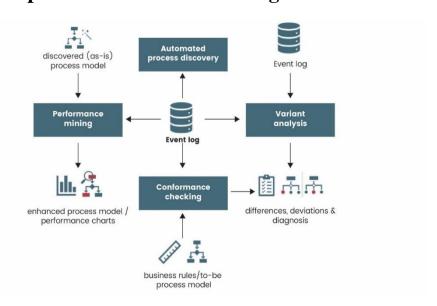


Fig.1.2: Core Capabilities of Process Mining

## 1.Process Discovery

Process Mining is the first step in process mining, which involves creating a process model from event log data without any prior knowledge of the process.

## 2. Conformance Checking

Conformance checking is the second component of process mining, which involves comparing the actual process executions with the expected process model. This component helps you identify deviations and non-compliance issues, and efficiency and standardization.

#### 3. Process Enhancement

Process enhancement is the third component of process mining, which involves improving the existing process models to optimize process performance and achieve KPIs. This component helps businesses identify root causes, predict future process behavior, and evaluate process improvement strategies.

#### 4. Variant analysis

Oftentimes, it is not enough to analyze the performance of all cases of a process at once. A lot more can be learned by comparing multiple variants of a process, like for example the variant of a process executed in North America versus Europe.



Fig.1.3: Various stages that gets processed

- 1. **Recognize the problem** Decide on a business issue that requires process mining to solve in a reasonable timeframe.
- 2. **Identify the information** Prior to getting started, it's essential that you establish where your business process data comes from and what it means.
- 3. **Start a pilot project** It is necessary to launch a new push to demonstrate the potential value of process mining. Ensure the project can be completed swiftly and produce clear, measurable results for everyone in your organization.
- 4. **Acknowledge the facts** Using process mining, you can get a clear picture based on facts, so be open and honest with your findings.

## **CHAPTER-2**

## **TECHNOLOGY**

#### 2.1 Process Mining Utilization:

- Process mining utilizes this data from IT systems to create a process model, or process graph. From here, the end-to-end process is examined, and the details of it and any variations are outlined. Specialized algorithms can also provide insight into the root causes of deviations from the norm.
- These algorithms and visualizations enable management to see if their
  processes are functioning as intended, and if they aren't, they arm them with
  the information to justify and allocate the necessary resources to optimize
  them. They can also uncover opportunities to incorporate robotic process
  automation into processes, expediting any automation initiatives for a
  company.
- Process mining sits at the intersection of Business Process
   Management (BPM) and data mining. While process mining and data mining both work with data, the scope of each dataset differs the scope of data mining is much broader, and it extends to a variety of data sets. It is used to observe and predict behaviors, having applications within customer churn, fraud detection, and market basket analysis to name a few.
- Increasing sales isn't the only way to generate revenue. Six sigma and lean methodologies also demonstrate how the reduction of operational costs can also increase your return-on-investment (ROI).
- Process mining helps businesses reduce these costs by quantifying the
  inefficiencies in their operational models, allowing leaders to make objective
  decisions about resource allocation. The discovery of these bottlenecks can not
  only reduce costs and expedite process improvement, but it can also drive
  more innovation, quality, and better customer retention. However, since
  process mining is still a relatively new discipline, it still has some hurdles to
  overcome.

#### The Technologies:

#### **Business Process Mapping (BPM):**

Provides some insight into processes, but its lengthy, subjective, and outdated as soon as its completed

#### **Robotic Process Automation:**

Helps you reduce man hours, but automating a poor process will get you poor results, faster

#### **Business Intelligence:**

Helps you understand your data, but does not analyze root causes of process problems, and so cant detect value opportunities and take action on them in real time.

#### **Data Storage and Management:**

Databases: Structured databases (SQL) or NoSQL databases store and manage event log data efficiently. Data warehouses enable centralized storage and analysis of data from various sources.

#### **Data Analytics and Visualization Tools:**

Improve their processes. The choice of specific technologies may vary based on the organization's needs, available resources, and the complexity of the processes being analyzed

#### 2.2 Some of those challenges include:

- Data Quality: Finding, merging and cleaning data is usually required to
  enable process mining. Data might be distributed over various data sources. It
  can also be incomplete or contain different labels or levels of granularity.
  Accounting for these differences will be important to the information that a
  process model yields
- Concept drift: Sometimes processes change as they are being analyzed, resulting in concept drift.

#### 2.3 Process mining use cases:

Process mining techniques have been used to improve process flows across a wide variety of industries. Since process maps highlight the key performance indicators (KPIs) which impact performance, they have spurred businesses to reexamine their operational inefficiencies. Some use cases include:

- **Education:** Process mining can help identify effective course curriculums by monitoring and evaluating student performance and behaviors, such as how much time a student spends viewing class materials.
- **Finance:** Financial institutions have used process mining software to improve inter-organizational processes, audit accounts, increase income, and broaden its customer base.
- **Public works:** Process mining has been used to streamline the invoice process for public works projects, which involve various stakeholders, such as construction companies, cleaning businesses, and environmental bureaus.
- **Software Development:** Since engineering processes are typically disorganized, process mining can help to identify a clearly documented process. It can also help IT administrators monitor the process, allowing them to verify that the system is running as expected.
- **Healthcare:** Process mining provides recommendations for reducing the treatment processing time of patients.
- **E-commerce:** It can provide insight into buyer behaviors and provide accurate recommendations to increase sales.
- **Manufacturing:** Process mining can help to assign the appropriate resources depending on case—i.e. product—attributes, allowing managers to transform their business operations. They can gain insight into production times and reallocate resources, such as storage space, machines, or workers, accordingly.

#### **CHAPTER-3**

## **APPLICATIONS**

#### **3.1 Sales:**

**Increasing on-time delivery:** For customer satisfaction, on-time deliveries are essential. Companies can use process mining to uncover the reasons behind late deliveries.

**Identifying reasons that hurt monthly revenue:** Companies may lose a portion of their earnings during this process due to long-running holds or order cancellations. Process mining tools may point out the root causes of those issues and, companies can minimize those losses accordingly.

**Identifying root causes of order changes:** Customers sometimes change their orders which causes processes to take a longer time. Companies prefer to decrease these order changes to stabilize their processes.

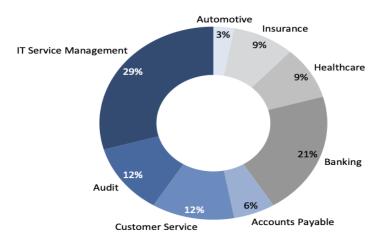


Fig.3.1: Pie Chart representation of applications

93% of business leaders aim to leverage a process mining tool. However, organizations have to be aware of its uses to apply process mining in their business.

**Benchmark the amount of returned goods:** Companies can discover the value of returned goods by using process mining. According to this insight, they can focus on improving their order-to-cash processes.

#### Sales:

#### Lead-To-Order

- Reduced sales cycle time: Lead-to-order processes can take a long time. This
  causes the payback time of marketing investments to increase. Companies can
  uncover the reasons behind this issue and, take action to reduce sales cycle
  time.
- Increased conversion rate: Converting marketing strategies into sales is critical for companies. With a process mining tool, companies can discover if they have proper strategies for increasing conversion rates.

#### 3.2 Banking:

Banks also benefit from process optimizations as most banks' processes still include legacy systems and paper-based documentation. Process mining tools can help identify bottlenecks and automation opportunities improving customer satisfaction and efficiency. Processes to optimize include:

- Mortgage: Mortgage is the most complicated B2C loan process and there are opportunities to improve it in most cases. Process mining tools can map work and communication flow of a mortgage process and detect repetitive steps and reworks leading to delays. This helps reduce customer waiting time and improve collaboration across different units.
- Other loans

#### Card operations

#### 3.3 Customer Service:

- Cross-channel analysis to identify anomalies: Process mining software can help analyze process steps across different channels to identify compliance issues and inefficiencies
- Map customer journey: Process mining tool can illustrate the customer journey in a given channel by extracting data from ticket systems. By doing so, process mining facilitates tracking customer experience, challenges that customers face, and interactions between responsible agents and customers.

#### 3.4 Education:

Online learning platforms: Process and task mining can reveal details on how users navigate on learning platforms to improve user experience for students.
 For example, process mining can show the potential root-causes behind student's exit rates from the given platform, such as the length of videos or organization of materials.

#### 3.5 Healthcare:

- Administrative processes: Process mining discovers event logs that contain information about healthcare processes including the personnel in-charge, steps, and cost of processes and identifies areas for improvement.
- Clinical pathways: serve for standardizing healthcare practices and detect issues that might lead to wrong treatments or delays which are crucial for many patients (e.g cancer). Process mining can be used to identify clinical pathways.

#### **Process Mining Can Improve Healthcare Outcomes**

- In healthcare, employee burnout—and the often-lagging pace of innovation—is not merely an inconvenience: it's detrimental to care delivery and saving lives. Every industry has plenty of operational inefficiencies, but optimizing processes to eliminate those inefficiencies is critical to improving outcomes for practitioners and patients.
- While it is difficult to know where to begin, given the sheer volume of data and systems providers have, process mining in healthcare is the key to identifying those discrepancies and gaps which are diminishing care.
- Process Mining enables organizations to go behind the scenes to gather valuable insights providers can use to inform decision-making, upgrade processes, and spend more time with patients.

### **CHAPTER-4**

## MODULES EXPLANTION

#### **Process Mining Fundamentals:**

When interacting with the dynamic visual representation and drilldown tools such as tables and charts, one can take an **exploratory** approach or a **confirmatory** approach.

An exploratory approach is one where you simply explore the data and see what value opportunities jump out at your Analysis tools such as the Process Explorer, the Variant Explorer, and the Conformance checker are ideal for this.

With the **confirmatory approach**, you're examining the data to see if it confirms or denies a hypothesis. Using your Analysis, you can find out whether the data confirms or denies that these perceived pain points exist and have a significant impact.

#### 4.1 Module-1

#### Process, activity, and case:

- A process is a series of linked steps taken in order to achieve a particular goal.
- An **activity** is a step that occurs in the process. Process activities are actions that initiate or terminate a process or take place during it. Each activity **consists of one or more processes** that together are a milestone in the process.

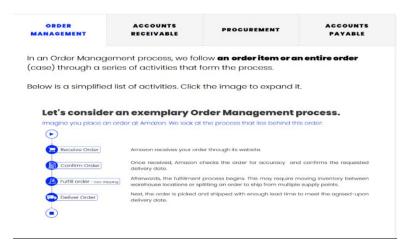


Fig.4.1: An entire order through series of Activities

A case is an "item" or "object" you follow through the process. Even for the same business process, the case differs from company to company, depending on how granular they want to get.

In order Management process, we follow **an entire order** (case) through a series of activities that form the process.

#### **4.2 Module-2:**

#### What Is the Variant Explorer?

As the name implies, using the Variant Explorer, you can discover all the process variants—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.

Represented by a number, a Key Performance Indicator (KPI) allows you to quickly assess how your process is performing.

The Case Frequency KPI reflects the number of unique cases associated with an activity or connection. In a single variant, naturally, the number is the same across the activities and connections.

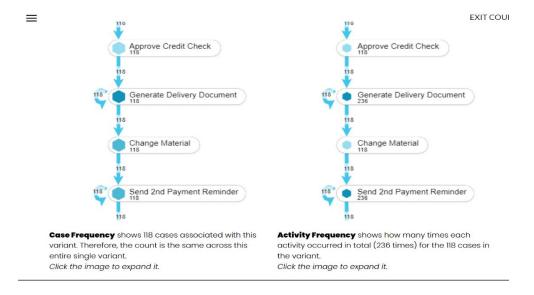


Fig.4.2: Comparison between Case Frequency and Activity Frequency

The activity frequency on "Generate Delivery Document" (236) is exactly double the case frequency (118); this reflects the fact that each case in this variant goes through "Generate Delivery Document" twice, as indicated by the loop. The same is true for the activity, "Send 2nd Payment Reminder."

#### **Multi Variant:**

Sometimes, you might want to review multiple variants at once to gain insights into how they're similar and different. In the Variant Explorer, you can superimpose adjacent and non-adjacent variants (superimposing means placing one thing on top of the other.)

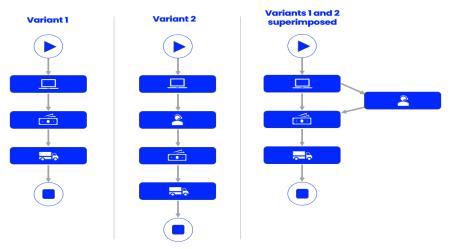


Fig.4.3: Multi Variant Representation

In the Variant Explorer, you can superimpose adjacent and non-adjacent variants

#### **4.3 Module-3:**

#### What is the 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, for example an undesired activity, to see which activities cases typically come from and which activities they're going to.

Another useful feature in the Process Explorer is the **List view** of activities and connections. This way, you can **quickly scroll through a list** of the most common activities and connections.

#### Review predecessor and successor activities:

Being able to explore predecessor and successor activities helps you to better understand how your process is unfolding in your organization.

#### **4.4 Module-4:**

#### Using charts and Tables, Review KPI's:

**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.

#### Examples:

- Number of sales order items
- Net order value

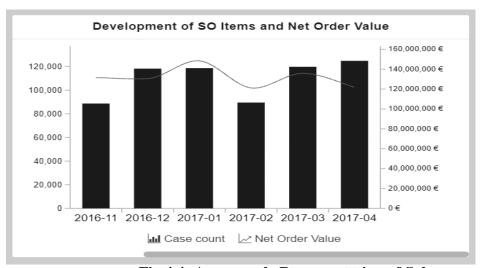


Fig.4.4: An example Representation of Sales

- 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.
- In charts, you can select one or more of the displayed timeframes to narrow the scope of the analysis. In an OLAP table, you can adjust column width just like in tools like Microsoft Excel or Google Sheets.

#### **Module-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.

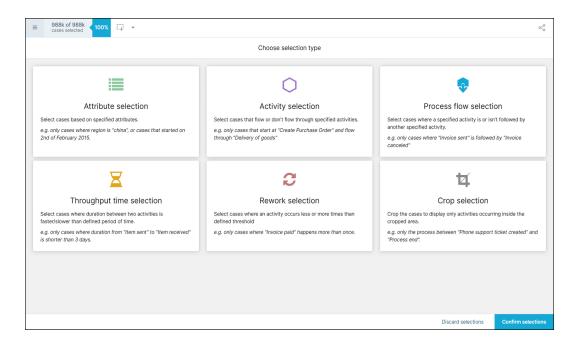


Fig.4.5: Six Selection Types

- Attribute selection
- Activity selection
- Process flow selection
- Throughput time selection
- Rework selection
- Crop selection

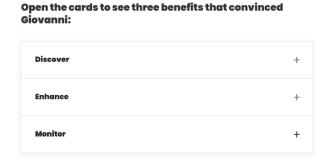
#### **Module-6:**

#### **CASE STUDY**

#### The Pizzeria's Business

#### The situation is...

The Pizzeria Mamma Mia was opened about a year ago and the business is generally going well with customers praising the good quality and taste of the pizza. However, their **customer ratings have been very low**, and they are making negative profits for some of their deliveries.



Three that convinced

Fig.4.6: Benefits

#### The Pizzeria's Order-to-Cash Process



Fig.4.7: Order-To-Cash Process

What you see above is one variant of what the **pizza order process** could look like. The starting point of the process is the customer's order and the end point the final payment. Between process start and process end, various activities take place until the customer receives his pizza. Unfortunately, the process does not always run smoothly. Activities could be missing, swapped or new activities could be added.

## The Pizzeria's Data:

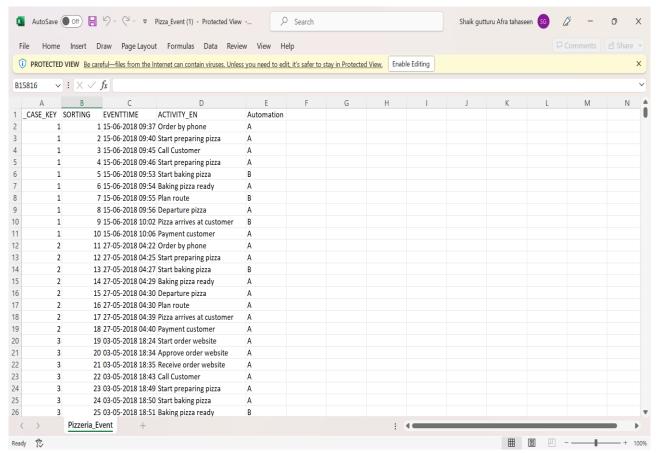


Fig.4.8: Pizzeria's Data

Digital information already exists in the checkout system as well as the order system through the website of the Pizzeria Mamma Mia.

This data can then be transformed into an event log. An event log can help us to reconstruct the process as how it really occurred.

#### **CHAPTER 5**

## **REAL TIME EXAMPLES**

PepsiCo's digital transformation strategy aims to create 'a totally different' company. They use Celonis as a key enabler to get there – and are unlocking millions of dollars in the process.

Faster, stronger, better. That's the mission that unites everyone at PepsiCo, one of the world's largest food and beverage companies. As part of their transformation journey, the company is using data and technologies such as predictive analytics, AI, robotics, process mining, and automation to transform the very fabric of their business and their processes.

#### Transforming PepsiCo process by process:

"Everyone is trying to be smarter, better, faster. My job is to enable the business to be just that: by taking the work out of the work, driving productivity, and ultimately generating cost savings and new revenue."

That's when he discovered process mining and Celonis. "It was originally a news headline that led into it. Process mining used to be a theoretical offline discussion, until Celonis came around and turned it into actual technology. And by the time picked it up at PepsiCo, Celonis had established itself as the leader in the space. Today, the company is using Celonis across nine processes — such as Accounts Receivable and Accounts Payable, Procure-to-Pay, Order-to-Cash, and Make-to-Deploy.

In Accounts receivable, PepsiCo struggled with the same challenges so many big corporations have: organizational silos, a **complex billing process, and difficulties prioritizing**. But bringing in Celonis changed the rules of the game almost overnight.

Using Celonis out-of-the-box capabilities, the Collections team learned that improving their Days Sales Outstanding (DSO) had the highest value potential — and identified which actions would quickly move the needle in terms of collecting customer payments faster.



Fig.5.1: Transformation of Pepsi Co

They've set out to speed up inventory movement, reduce shipping costs, and optimize working capital – all while accounting for regulations and potential shipping embargos.

But bringing **cash through the door** quicker is just one of the benefits. PepsiCo's Collections teams now work smarter, not harder – driving better results with less effort.

Using Celonis intelligent automation capabilities, dubbed action flows, PepsiCo automated the creation and sending of dunning notes for customers to collect on overdue receivables, helping the teams save valuable time.

## CHAPTER -6

## **OUTCOMES**

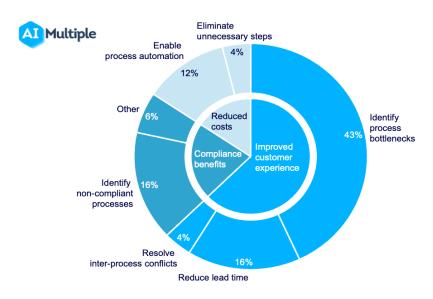


Fig.6.1: Pie Chart of benefits

#### 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.

## **Conclusion**

- Process mining aims to extract information from event logs, which are recorded from running business processes. Process mining projects may go through multiple phases in which different process mining techniques are used: process discovery, conformance checking and model enhancement, to all of which we contributed concepts and techniques. In this chapter, we summarize the previous chapters and reflect on how well these techniques address the challenges of process mining identified in 3 and point out open problems and future work.
- Process mining focuses on different perspectives, such as control-flow, organizational, case, and time. While much of the work around process mining focuses on the sequence of activities—i.e., control-flow—the other perspectives also provide valuable information for management teams.
   Organizational perspectives can surface the various resources within a process, such as individual job roles or departments, and the time perspective can demonstrate bottlenecks by measuring the processing time of different events within a process.

## INTERNSHIP CERTIFICATE





## **REFERENCES**

- [1] https://www.processmining.org
- [2] https://www.celonis.com/process-mining/what-is-process-mining