



Universidad Politécnica
de Madrid

**Escuela Técnica Superior de
Ingenieros Informáticos**



Grado en Ingeniería Informática

Trabajo Fin de Grado

**Herramienta de Extracción Automática
de Procesos en un Entorno
Universitario Usando Celonis**

Autor: Juan Vicente Ramos García

Tutor(a): José Antonio Mateo Cortés y Antonio Jesús Díaz Honrubia

Madrid, junio 2024

Este Trabajo Fin de Grado se ha depositado en la ETSI Informáticos de la Universidad Politécnica de Madrid para su defensa.

Trabajo Fin de Grado
Grado en Ingeniería Informática

Título: Herramienta de Extracción Automática de Procesos en un Entorno Universitario Usando Celonis
junio 2024

Autor: Juan Vicente Ramos García

Tutor: José Antonio Mateo Cortés y Antonio Jesús Díaz Honrubia
Departamento de Arquitectura y Tecnología de Sistemas Informáticos y Departamento de Lenguajes y Sistemas Informáticos e Ingeniería de Software
Escuela Técnica Superior de Ingenieros Informáticos
Universidad Politécnica de Madrid

Agradecimientos

Por ti, Papá.

Por todos aquellos momentos en los que hablábamos de la graduación y de cómo lo íbamos a pasar.

Por todas aquellas risas imaginándonos como lo íbamos a celebrar.

Ahora que ya no estás, me sigues dando fuerza para conseguirlo todo.

Que con este trabajo estemos un paso más cerca de ese momento.

Pero lo más importante, es que a pesar de que no estés, vas a celebrarlo conmigo allá donde quiera que estés.

Te quiero.

Como siempre me decías, pasase lo que pasase:

Paso corto y mala leche.

Resumen

Este Trabajo Final de Grado tiene como objetivo extraer, monitorizar y optimizar procesos de un entorno universitario, como es la Universidad Politécnica de Madrid. Para ello, se han planteado dos escenarios ficticios existentes en todas las universidades: la interacción del alumno con la plataforma Moodle y el proceso de Matriculación. En primer lugar, se analizará cómo funciona la minería de procesos y como se aplica la metodología existente para poder extraer y reflejar un proceso en una plataforma para tal efecto. En segundo lugar, se explica con detalle cómo se han transformado y normalizado los datos para conseguir una base sólida con la información necesaria para poder representar los procesos escogidos. Tras la normalización de los datos, la implementación de análisis del proceso se muestra en cuadros de mando donde se plasmarán las principales métricas, así como la segmentación de la información mediante diversos gráficos. Finalmente, tras tener la implementación terminada, se analizarán los resultados obtenidos y el impacto, concluyendo con las oportunidades de mejora y los próximos pasos a realizar con base en este trabajo.

Abstract

This Final Degree Project aims to extract, monitor, and optimize processes within a university environment, specifically at the Universidad Politécnica de Madrid. To achieve this, two fictional scenarios common to all universities have been proposed: student interaction with the Moodle platform and the enrollment process. First, the project will analyze how process mining works and how the existing methodology can be applied to extract and represent a process on a specialized platform. Second, it will explain in detail how the data has been transformed and normalized to establish a solid base with the necessary information for representing the chosen processes. After data normalization, the implementation of process analysis will be shown in dashboards that display the main metrics and segment the information using various charts. Finally, after completing the implementation, the obtained results and impact will be analyzed, concluding with the identified improvement opportunities and next steps to be taken based on this work.

Tabla de contenidos

1. Introducción	1
1.1. Motivación o necesidad del proyecto	1
1.2. Objetivos	2
1.3. Planificación	2
1.4. Estructura de la memoria	3
2. Estado del arte	5
2.1. Process Mining	5
2.2. Herramientas de Process Mining	6
2.3. Celonis	7
2.4. Requisitos de seguridad, privacidad y certificaciones	9
3. Análisis y Diseño	10
3.1. Arquitectura de la solución	10
3.2. Integración y transformación de datos en Celonis	11
3.3. Metodología de interpretación de registros de eventos	12
3.4. Modelo de datos	15
3.5. Estructura y organización de elementos de visualización	17
4. Casos de Estudio	21
4.1. Fuente de datos de Moodle	21
4.1.1. Ficheros de datos y estructura	21
4.1.2. Transformaciones	22
4.1.3. Modelo de Datos	24
4.1.4. Implementación y visualización	26
4.1.5. Oportunidades teóricas de mejora	41
4.2. Fuente de datos de Matriculación	43
4.2.1. Generación del log de eventos	43
4.2.2. Ficheros de datos y estructura	44
4.2.3. Transformaciones	46
4.2.4. Modelo de Datos	48
4.2.5. Implementación y visualización	49
4.2.6. Oportunidades teóricas de mejora	66
5. Conclusiones y trabajo futuro	68

6. Análisis de impacto	70
6.1. Impacto empresarial	70
6.2. Impacto económico	70
6.3. Impacto social y cultural	70
6.4. Impacto basándose en los Objetivos de Desarrollo Sostenible (ODS)	71
Bibliografía	72
Anexos	76
A. Moodle - Código de las transformaciones	80
A.1. Transformación Tabla de Actividades	80
A.2. Transformación Tabla de Casos	83
A.3. Transformación Tablas Auxiliares	84
A.4. Transformación Ajustes	85
B. Moodle - Código de las Vistas, métricas y KPIs	86
C. Matriculación - Código de las transformaciones	88
C.1. Transformación Tabla de Actividades	88
C.2. Transformación Tabla de Casos	89
C.3. Transformación Tablas Auxiliares	89
D. Matriculación - Código de las Vistas, métricas y KPIs	91

Capítulo 1

Introducción

En este apartado de introducción se comentan la motivación o necesidad del proyecto, los objetivos generales y específicos, así como la planificación que se ha seguido en el trabajo y la estructura de esta memoria.

1.1. Motivación o necesidad del proyecto

Hoy en día, todas o prácticamente todas las organizaciones en cualquier ámbito tienen una multitud de procesos en diversas áreas que las componen. Estos procesos abarcan desde tareas sencillas y livianos hasta procesos complejos y transversales que son el pilar de la integridad de dichas organizaciones. Esto, sumado a la era de la digitalización y al auge de iniciativas de transformación digital [1] para convertir cualquier organismo, empresa o administración en entidades orientadas a los datos o *data-driven* [2], convierte el panorama actual de todos estos entes en una organización orientada a procesos y orientada a los datos.

En un mundo donde los procesos están soportados por una variedad de sistemas distintos y heterogéneos, por procesos de todo tipo y donde las métricas cobran una importancia clave para la viabilidad y la productividad de una empresa, la optimización y mejora de dichos procesos toma un papel protagonista. Cualquier empresa puede digitalizar sus procesos y transformarse digitalmente mediante el uso de software comercial para cualquiera de sus áreas, o utilizando servicios web, portales, *chatbots*, robotización o *ETLs* (herramientas y técnicas para un flujo de *Extracción, Transformación y Carga (Load)* de datos) [3] para este fin. Pero la pregunta que hay que hacerse es si se está haciendo de manera correcta y óptima, buscando el mejor desempeño de estos procesos.

En este contexto, las universidades y, en particular, la Universidad Politécnica de Madrid, tienen multitud de procesos, de todo tipo, algunos heredados desde hace décadas. Aquí la Minería de Procesos se convierte en un factor determinante para optimizar dichos procesos. En este trabajo de fin de grado, se entrará en el mundo de los procesos universitarios, donde aún no se ha aplicado ampliamente esta técnica, buscando extraer patrones y oportunidades de mejora en procesos del entorno universitario, fusionando el gran mundo de Minería de Da-

tos (*Data Mining*) [4] con el mundo de los procesos de negocio.

1.2. Objetivos

El objetivo principal de este trabajo es la optimización de procesos y la obtención de valor en los procesos universitarios mediante el descubrimiento de oportunidades de mejora utilizando Celonis [5] como herramienta de Minería de Procesos. Esto permitirá descubrir y encontrar patrones de comportamiento ideal y no ideal en los procesos pertenecientes al entorno universitario.

Este estudio y descubrimiento permitirá identificar uno o varios casos de uso sobre un proceso de la Universidad y permitirá la búsqueda de oportunidades de mejora en dichos procesos así como la comprensión y mejora de las principales Métricas de Rendimiento o KPIs (*Key Performance Indicators*) [6] clave para la toma de decisiones de los usuarios y propietarios de estos procesos. Esto implica y requiere la actuación de varios agentes de la universidad, por lo que estas mejoras se centrarán en la productividad y en cómo aprovechar de manera más eficiente el trabajo y desempeño de dichos agentes.

1.3. Planificación

La planificación de esta tarea, se realizó pensando en un enfoque ágil de iteración en dos sprints o casos de uso [7] distribuidas a lo largo de las siguientes tareas:

1. Introducción a la minería de procesos y a Celonis: 40 horas.
2. Instalación y aprendizaje del entorno de desarrollo: 15 horas.
3. Capturar y limpiar los datos que compongan el conjunto de datos del problema (logs). Incluye reuniones con la universidad para obtener los datos: 30 horas.
4. Programar un ejemplo sencillo (prueba de concepto) de un proceso con pocos datos: 10 horas.
5. Implementación de un ejemplo completo: 70 horas.
6. Evaluación inicial del ejemplo encontrando patrones de comportamiento: 20 horas.
7. Evaluar distintas técnicas de minería de procesos: 40 horas .
8. Implementar las posibles mejoras encontradas durante las reuniones: 7 horas.
9. Reuniones de seguimiento: 10 horas.
10. Memoria del TFG: 50 horas.
11. Refinamiento: 5 horas.

1.4. Estructura de la memoria

Estas tareas se distribuyen en el tiempo mediante el cronograma mostrado en la Figura 1.1.

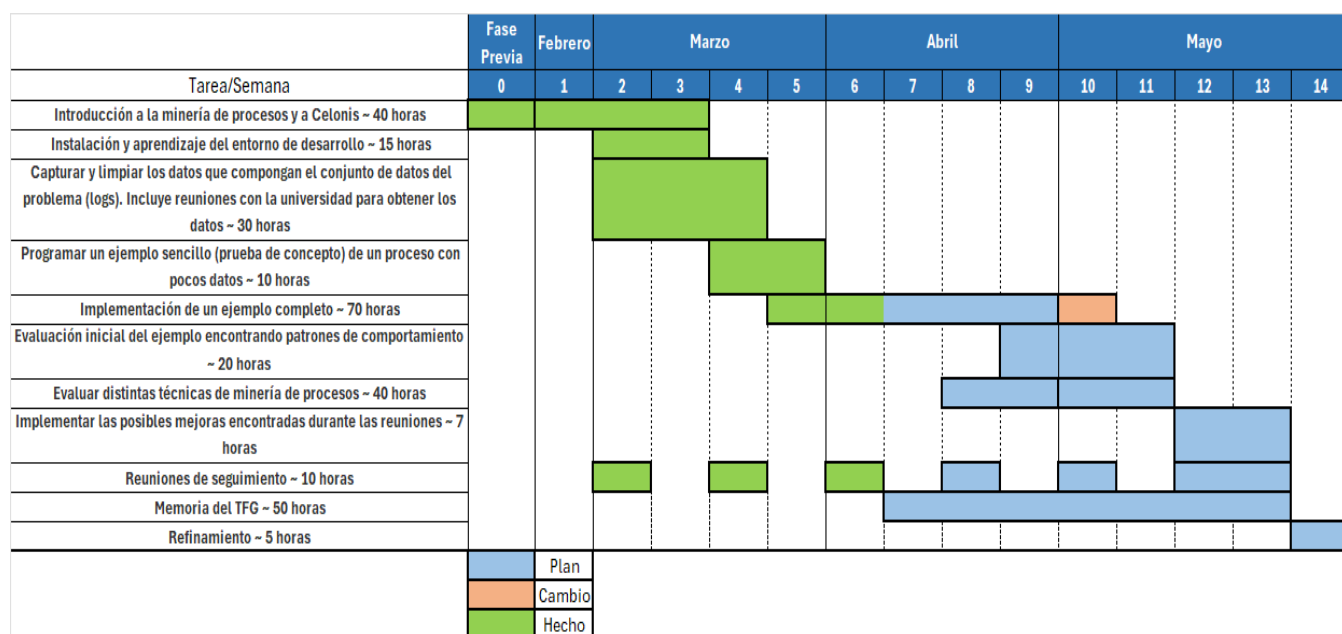


Figura 1.1: Diagrama de Gantt del proyecto

1.4. Estructura de la memoria

Esta memoria se organiza y estructura de la siguiente manera:

- El capítulo 2. Estado del arte, explica como se encuentra actualmente la situación respecto al Process Mining, como resolver la problemática actual de los procesos en las organizaciones y como Celonis se ha convertido en una solución viable a ello.
- El capítulo 3. Análisis y Diseño se muestra de manera teórica e ilustrada como se distribuye la solución y las decisiones de diseño que se han seguido.
- El capítulo 4. Casos de Estudio incluye las dos partes en las que se ha dividido el trabajo, el primero el caso de Moodle y el segundo el de matriculación. Aquí se muestra el proceso de implementación de inicio a fin, así como las oportunidades de mejora encontradas.
- El capítulo 5. Análisis de impacto del trabajo en diversos aspectos, así como el impacto en los Objetivos de Desarrollo Sostenible(ODS) [8] de la Agenda 2030.
- El capítulo 6. Conclusiones y trabajo futuro recoge los resultados y conclusiones a los que se ha llegado tras la elaboración de este proyecto y el trabajo futuro a realizar.

Capítulo 1. Introducción

- La bibliografía contendrá todos los recursos que se han consultado, tanto digitales como escritos para la elaboración de esta memoria.
- Los anexos incluyen información adicional a esta memoria, como la lista de Figuras y Tablas, y el código detallado utilizado en este trabajo.

Capítulo 2

Estado del arte

En este capítulo se describe el estado del arte de este trabajo y se mencionan trabajos relacionados, donde se ha aplicado esta técnica en el ámbito universitario.

2.1. Process Mining

Process Mining [9] se refiere a las tecnologías, técnicas y metodologías dentro de la ciencia de los datos que se centran en la intersección de Data Mining con los procesos de negocio (véase la Figura 2.1). Esta disciplina se enfoca en el análisis de los registros de eventos generados por los procesos de negocio en sus sistemas fuente o transaccionales (también conocidos como “*huella digital*”), utilizando herramientas y algoritmos propios de Process Mining con el fin de descubrir, monitorizar y mejorar los procesos empresariales.

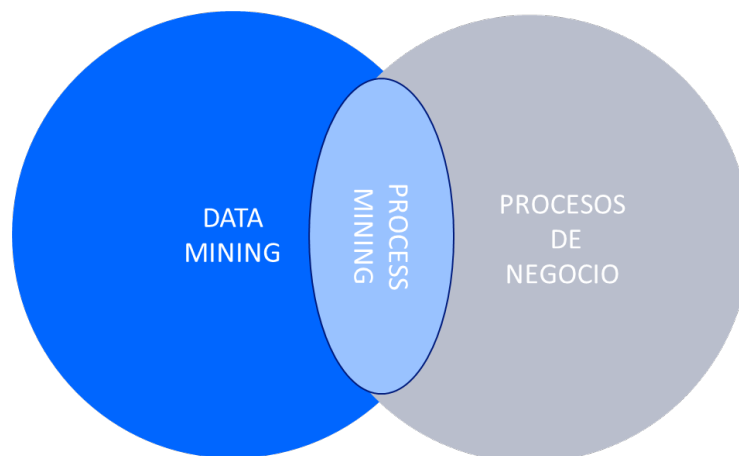


Figura 2.1: Process Mining como intersección de Procesos y Datos

La base o pilar de todo proceso existente, para poder ser apto para utilizar Process Mining en él, es el registro de eventos o *huella digital*, tal y como se muestra en la Figura 2.2. Esto consiste en una marca o traza de cada paso o etapa que

el proceso sigue a lo largo del tiempo. Esta huella digital se compone de un identificador, una etapa o actividad que describe la acción que ha realizado el proceso o el usuario en ese proceso, y un registro de tiempo, es decir, cuándo ocurrió esa acción. Con estas tres variables o entidades, se dispone del mínimo producto viable para poder representar y poder analizar cualquier proceso, independientemente de su naturaleza, complejidad y atributos.

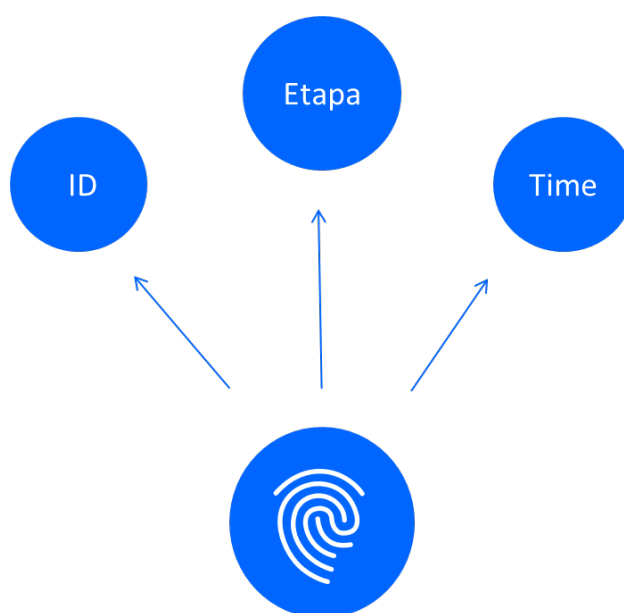


Figura 2.2: Huella digital y sus atributos

Una vez se consiga acceso a dichos datos, y como en todo sistema transaccional, aplicación cloud o ficheros planos, se requerirá de un trabajo previo de conexión o importación de los datos y una limpieza de tipos, eliminación de datos trunca-dos, anonimización y omisión de datos redundantes que permitan tener un set de datos lo más limpio y pulido posible.

Tras este trabajo de limpieza de datos, conexión a los sistemas y modelización de las bases para nuestro análisis de Process Mining, se podrá representar un "mapa" que muestra todas y cada una de las etapas, con todas las variantes del proceso de inicio a fin.

Estas variantes representan de manera objetiva y fiable todos y cada uno de los caminos que nuestro proceso sigue mientras desarrolla su actividad. Es lo que se conoce como *Árbol de Proceso* (véase la figura 2.3.).

Una vez alcanzado este hito, comenzará el trabajo de investigación, optimización y mejora, así como la investigación de las desviaciones en estos procesos y dónde intervenir. En conjunto, todo este proceso de identificación, captura y mejora entra dentro del ámbito del Process Mining.

2.2. Herramientas de Process Mining

Tras analizar las diferentes tecnologías con sus ventajas e inconvenientes, así como el potencial de cada una para nuestro objetivo, se ha concluido que Celo-

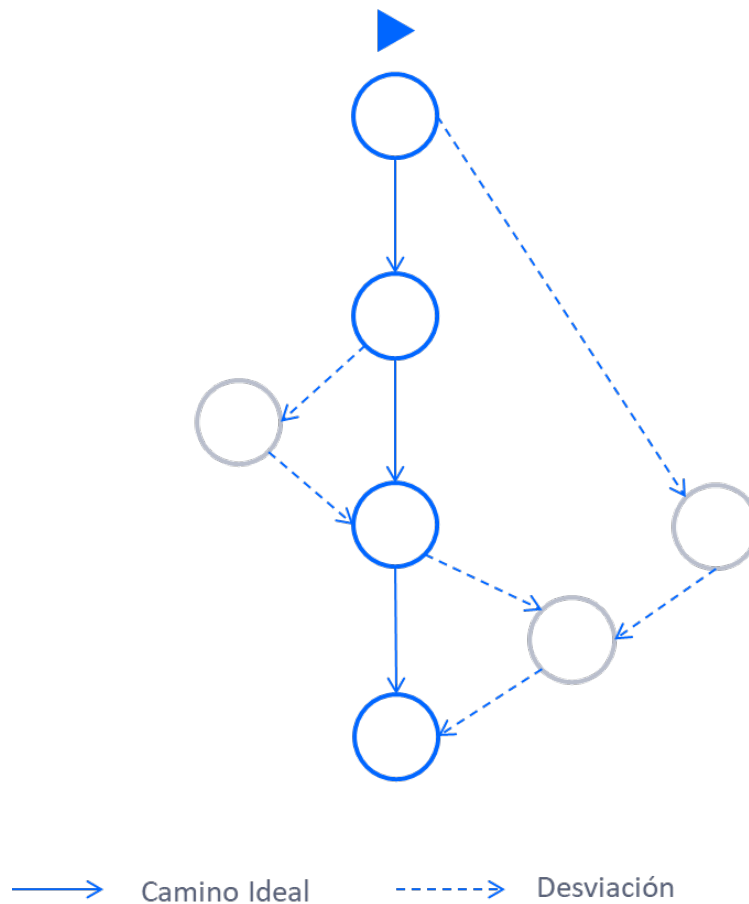


Figura 2.3: Representación del Árbol de Proceso

nis[10], como líder en esta disciplina (como se ve en la Figura 2.4), es la mejor tecnología para este proyecto.

2.3. Celonis

En esta sección se va a intentar resolver la siguiente pregunta: ¿Por qué Celonis y no otra herramienta de Análisis de datos o *Business Intelligence* [11]?

Celonis es una empresa fundada en 2011 con sede en Múnich, especializada en minería de procesos y oportunidades de mejora en los procesos empresariales. Actualmente, es uno de los líderes en esta disciplina.

Celonis nace como idea de tres estudiantes de la Universidad Técnica de Múnich, quienes, al darse cuenta del potencial de la minería de procesos, podían ayudar a abordar las dificultades que estaban ocurriendo en el entorno empresarial alemán en aquellos años. Estas dificultades causaban que las empresas no supieran a ciencia cierta cómo se ejecutaban realmente sus procesos y de qué manera poder optimizarlos. Con esta visión, fundaron Celonis, convirtiéndolo en una de las empresas más grandes actualmente y líder en este segmento.

Su actual producto no solo abarca Process Mining, sino que también ofrece

Capítulo 2. Estado del arte



Figura 2.4: Cuadrante Mágico de Gartner 2023 en Process Mining

herramientas de *Business Intelligence* similares a las que ofrecen actualmente PowerBI [12], Qlik [13] o Tableau [14].

Celonis destaca [5] respecto a otras tecnologías y herramientas en 5 puntos principales:

- **Funcionalidad Avanzada:** Celonis como tecnología pionera en Minería de Procesos, sigue ofreciendo continuamente nuevas funcionalidades y mejorando las ya existentes, incluyendo añadidos transversales de otros campos como la incorporación de *Machine Learning*[15] o máquinas y entornos de Python incorporados para enriquecer los desarrollos.
- **Experiencia intuitiva:** fácil de usar y su desarrollo, así como la comprensión de los resultados, hace de Celonis una opción más atractiva que los reportes o cuadros de mando tradicionales.
- **Integración con los sistemas más utilizados:** Celonis tiene conectores pre construidos, los sistemas más utilizados actualmente (SAP, Salesforce, Oracle...) así como con los motores de bases de datos predominantes hoy en día (Azure, SQL, PostgreSQL, etc.).

2.4. Requisitos de seguridad, privacidad y certificaciones

- **Automatización:** Celonis permite la interacción en tiempo real y ejecutar acciones correctivas gracias a los llamados *Flujos de Acción* que no requieren de programación a bajo nivel.
- **Escalabilidad:** Desde un proceso con pocos datos a escalar a niveles de miles y miles de procesos permite un crecimiento orgánico y flexible.

Por el contrario, Celonis tiene alguna que otra desventaja y que deben ser mencionadas:

- **Curva de aprendizaje:** A pesar de ser muy intuitiva, algunas funcionalidades a nivel más técnico, como la construcción de modelos de datos, requieren de tiempo y práctica para poder dominar Celonis en su totalidad.
- **Sistemas personalizados o customizados:** Muchas organizaciones y empresas tienen un alto grado de personalización en sus sistemas fuente o transaccionales, llegando en algunos casos a la posibilidad de tener sistemas contruidos desde cero y que no dependen de ningún producto de mercado. En este sentido, Celonis requiere de un mayor esfuerzo para generar y construir conectores desde cero, sin ningún tipo de conexión o integración pre construida.

En resumen, las avanzadas capacidades de Celonis, así como su posición como pioneros en Minería de Procesos, lo han convertido en una de las opciones más atractivas para las organizaciones a la hora de monitorizar y mejorar sus procesos. Más aún ahora que la era de la digitalización y la transformación digital está en auge y promete transformar de manera transversal la manera de actuar y explotar los datos en los procesos y escenarios más comunes.

2.4. Requisitos de seguridad, privacidad y certificaciones

Otro factor importante para la elección de Celonis como herramienta de Process Mining era la gran cantidad de certificaciones y regulaciones europeas con las que contaba, más aún siendo una herramienta que trata datos personales en ciertos procesos. Por ello es importante disponer de una tecnología que cumpla todos los estándares y requerimientos de seguridad y privacidad de datos europeos. Los estándares que Celonis tiene y dispone son los siguientes:

- *ISO 270001:* Privacidad de la información.[16]
- *ISO 27701:* Seguridad de la información. [17]
- *GPDR:* Reglamento General Europeo de Protección de Datos.[18]
- *CSA STAR:* Seguridad de proveedor de servicios en la nube.[19]
- *ISO 9001:* Sistemas de Gestión de la Calidad. [20]
- *SOC 1 y 2:* Informes de control de seguridad de los datos en la nube.[21].

Capítulo 3

Análisis y Diseño

En este capítulo se mostrará el análisis del trabajo a realizar, siendo agnósticos al caso de uso y haciendo foco en la arquitectura de Celonis. Este análisis y diseño se centrará en un enfoque modular, escalable y ágil, primando que cualquier persona, entidad u organización pueda replicar de manera sencilla estos casos de uso. Para ello se tratará de desglosar y modularizar los elementos lo máximo posible con este objetivo en mente.

3.1. Arquitectura de la solución

Celonis tiene una estructura de varias pilares (véase la Figura 3.1) que conforman un flujo secuencial de datos. Se divide de la siguiente manera:

1. Conexiones de datos o subida de ficheros.
2. Almacenamiento de los datos en el motor de base de datos.
3. Transformación de datos.
4. Modelización de datos.
5. El componente *Studio* [22] y explotación por los usuarios en vistas.

El flujo de los datos comenzará con la conexión de Celonis a una fuente de datos, bien a una herramienta SaaS en Cloud [23], a una base de datos (Relacional o no relacional) o bien mediante una subida de ficheros.

Estos datos, bien de un sistema fuente o bien de un fichero, se almacenan y se tratan en Celonis como si de una tabla se tratase, almacenándolos en el motor de base de datos de la Cloud de Celonis. Se manipula y trata cada registro como una fila y siguiendo todas las convenciones y consideraciones de un motor de base de datos SQL.

En este caso el lenguaje para toda la interacción de estos datos es Vertica SQL [24], podría considerarse un dialecto de SQL con particularidades respecto a otros lenguajes derivados de SQL como MySQL [25] o TransactSQL [26].

La división de la arquitectura en diferentes pilares permite una óptima abstracción de los datos a los usuarios de las vistas, permitiendo ser una arquitectura

3.2. Integración y transformación de datos en Celonis

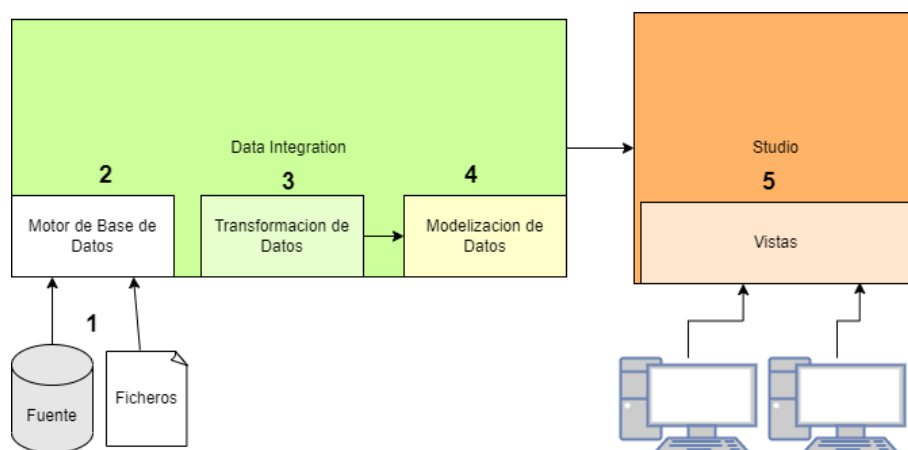


Figura 3.1: Arquitectura de la plataforma de Celonis y sus componentes

agnóstica al proceso, a los datos y a la tecnología aplicada por detrás.

3.2. Integración y transformación de datos en Celonis

La integración y almacenamiento de los datos una vez cargados permite su transformación y explotación para su posterior uso (véase la Figura 3.2). Las posibilidades que permite Vertica SQL son infinitas y desde creación de tablas, uniones entre diferentes tablas, tratamiento de fechas, condiciones, etc. permite realizar los trabajos necesarios para que nuestros datos sean legibles e interpretables por Celonis en nuestro Modelo de Datos.

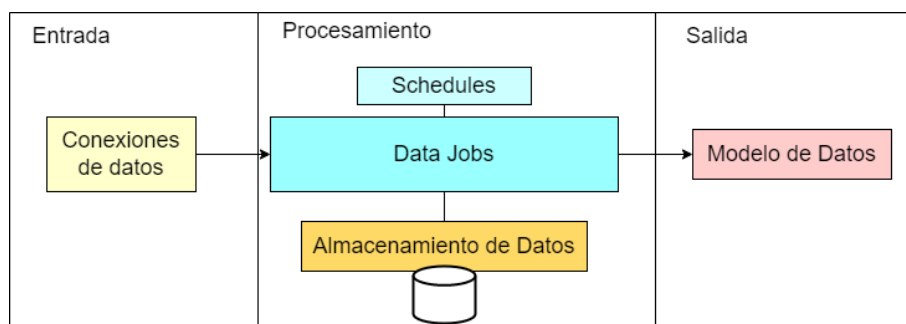


Figura 3.2: Estructura de la Integración de datos

Esto se realiza con los *Data Jobs*, que son tareas o elementos de transformación que permiten ejecutar cualquier sentencia Vertica SQL mediante la consola de ejecución donde se podrá ejecutar, crear y borrar código, así como ver la salida o resultado de la ejecución en la consola como si de un cliente de base de datos cualquiera se tratase (véase la Figura 3.3).

Una vez realizadas todas las tareas y trabajos de transformación se podrá crear un modelo de datos relacionado las diferentes tablas para que Celonis pueda interpretar cada una de ellas y poder extraer el árbol de proceso.

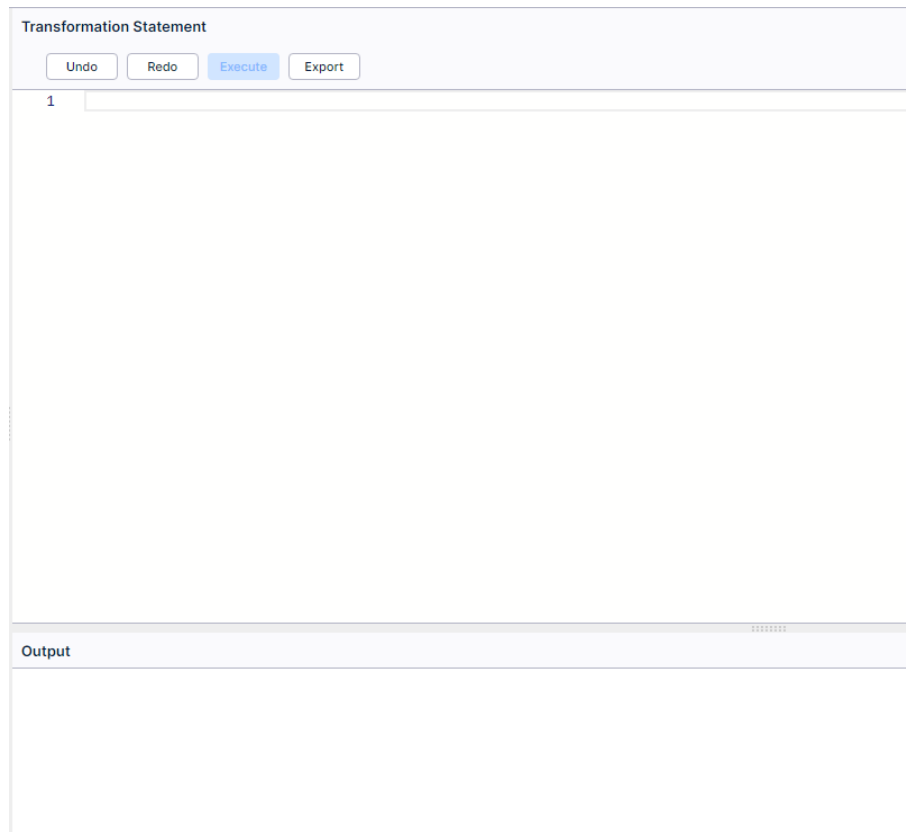


Figura 3.3: Consola de desarrollo de Vertica SQL en Celonis

Pero antes de eso, hay que saber qué jerarquía de tablas, de campos y de tipos de campo necesita Celonis para poder representar dichos datos de proceso.

3.3. Metodología de interpretación de registros de eventos

La disciplina de Minería de Procesos tiene como centro de su universo un único elemento, que es imprescindible y necesario para poder funcionar: el registro de eventos o *Event Log*[27].

Este registro de eventos es la *huella* que deja todo proceso en los sistemas, representando todas las actividades que un usuario, un automatismo u otro elemento realiza sobre un objeto o una entidad del sistema.

Imaginando un proceso sencillo, cualquiera, como puede ser una persona en su casa interactuando con su sistema de Domótica [28] en casa, utilizando luces, cerradura inteligente, etc. El registro de eventos sería todas y cada una de las actividades que va realizando el usuario A sobre los objetos del sistema en un registro de tiempo determinado.

Estas acciones o interacciones generarían un registro o log de eventos que consistiría, por ejemplo, en un fichero de texto plano o en un fichero de valores en filas con la siguiente distribución de la Tabla 3.1.

3.3. Metodología de interpretación de registros de eventos

Tabla 3.1: Ejemplo de registro de eventos

Usuario	Actividad	Huella temporal
A	Encender luces del salón	20-10-2018 20:55:43
A	Encender luces del pasillo	20-10-2018 21:15:16
A	Apagar Luces del pasillo	20-10-2018 23:30:13
A	Apagar luces del salón	20-10-2018 23:38:30
A	Bloquear cerradura inteligente	20-10-2018 23:44:55
A	Encender alarma	20-10-2018 23:45:32

Estos campos (Identificador, Nombre de Actividad y Huella temporal) serían los campos mínimos necesarios e imprescindibles para que Celonis y cualquier herramienta de Minería de Procesos pueda representar el proceso que está realizando el usuario A y poder analizar.

Obviamente, esto es una muestra muy reducida, pero en este caso sirve para poder visualizar como funciona la Minería de Procesos.

En particular, este proceso se vería o representaría según se puede apreciar en la Figura 3.4.

Partiendo de este ejemplo sencillo de un proceso ideal, lineal y secuencial, se puede representar un ejemplo a una mínima escala de lo que sería un proceso con desviaciones o variantes no deseadas y que suceden en prácticamente todos los procesos.

Prácticamente, todos los procesos, por muy optimizados que estén y por muy perfectos que sean, siempre difieren radicalmente de la teoría a la práctica.

Un ejemplo de una desviación o caso atípico para este mismo ejemplo de la domótica sería el representado en la Tabla 3.2.

Tabla 3.2: Ejemplo de registro de evento con desviaciones

Usuario	Actividad	Huella temporal
A	Encender luces del salón	20-10-2018 20:55:43
A	Encender luces del pasillo	20-10-2018 21:15:16
A	Caída del sistema	20-10-2018 21:16:00
A	Encender luces del salón	20-10-2018 21:17:33
A	Encender luces del pasillo	20-10-2018 21:18:20
A	Apagar Luces del pasillo	20-10-2018 23:30:13
A	Apagar luces del salón	20-10-2018 23:38:30
A	Bloquear cerradura inteligente	20-10-2018 23:44:55
A	Error en la conexión con la cerradura	20-10-2018 23:45:02
A	Bloquear cerradura inteligente	20-10-2018 23:47:19
A	Encender alarma	20-10-2018 23:55:32
A	Alarma desconectada	20-10-2018 23:55:40
A	Encender alarma	20-10-2018 23:58:40

Esta metodología se representa a nivel de proceso según la Figura 3.5 y es aquí en donde entra Process Mining, permitiendo el descubrimiento de estas varian-

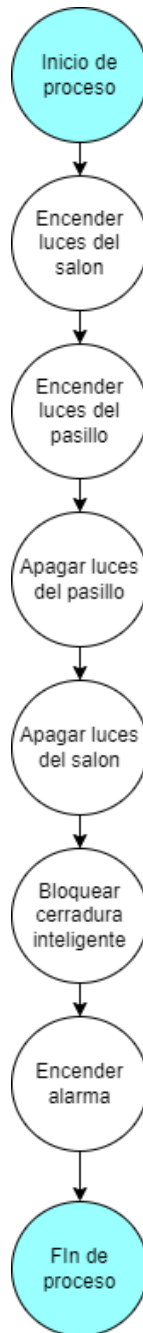


Figura 3.4: Árbol de Proceso - Camino ideal y sencillo

tes no deseadas, analizando el porqué ocurren y como solucionarlas y permitiendo cuantificar tanto el número de ocurrencias de estas desviaciones así como los beneficios esperados de la eliminación de estas.

Una vez entendido como tienen que estar formateados los datos para poder ser representados como un árbol de proceso, el siguiente paso es explicar el modelo de datos que tiene que diseñarse para poder utilizar los datos en Celonis.

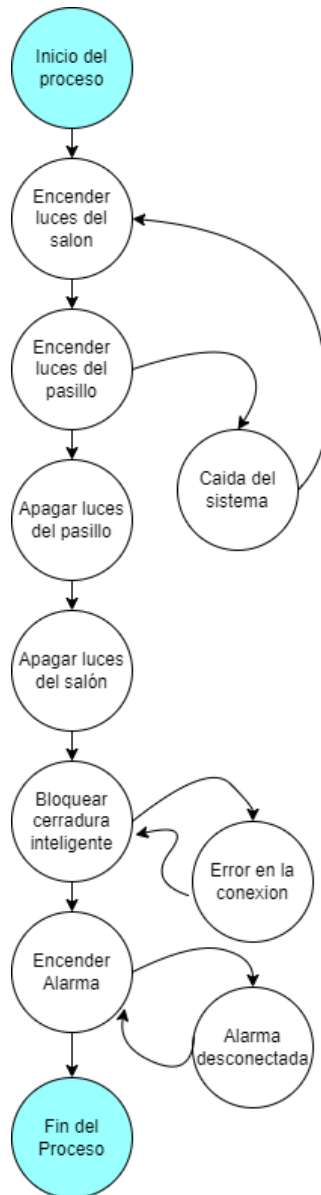


Figura 3.5: Árbol de Proceso - Camino con desviaciones

3.4. Modelo de datos

Celonis utiliza una estructura de tablas y contenido de tablas particular para los modelos de datos. Aquí entra la funcionalidad de uniones o *joins* de cualquier modelo de bases de datos relacional, siendo la funcionalidad de 1 a N la principal en este caso para relacionar diferentes tablas y datos. Se trata de la tabla de Casos y la tabla de Actividades.

Al igual que se necesitan los tres campos mínimos para la representación de un registro de eventos en Celonis, también una tabla de Actividades, centro de todo el modelo de datos. Esta tabla contendrá los datos normalizados tal cual se ha comentado en el capítulo anterior.

Capítulo 3. Análisis y Diseño

La Minería de Procesos y Celonis están basados en estas colecciones de eventos, que a su vez se basan en estos tres atributos principales que conformarán la tabla de Actividades junto a otros diversos campos que enriquecerán esta tabla y que podrán dar información adicional como tipologías, segmentación o diferentes dimensiones (véase la Tabla 3.3). La Tabla de Actividades contiene para

Tabla 3.3: Estructura de campos de registro de eventos

Usuario	Actividad	Huella temporal	...	Campo n	Campo 2
A	Actividad A	DD-MM-YYYY HH:MI:SS	...	Valor 1	Valor 2
A	Actividad B	DD-MM-YYYY HH:MI:SS	...	Valor 1	Valor 2
A	Actividad C	DD-MM-YYYY HH:MI:SS	...	Valor 1	Valor 2
B	Actividad A	DD-MM-YYYY HH:MI:SS	...	Valor 1	Valor 2
B	Actividad B	DD-MM-YYYY HH:MI:SS	...	Valor 1	Valor 2
C	Actividad A	DD-MM-YYYY HH:MI:SS	...	Valor 1	Valor 2
C	Actividad B	DD-MM-YYYY HH:MI:SS	...	Valor 1	Valor 2
C	Actividad C	DD-MM-YYYY HH:MI:SS	...	Valor 1	Valor 2
C	Actividad D	DD-MM-YYYY HH:MI:SS	...	Valor 1	Valor 2
D	Actividad A	DD-MM-YYYY HH:MI:SS	...	Valor 1	Valor 2
D	Actividad B	DD-MM-YYYY HH:MI:SS	...	Valor 1	Valor 2
E	Actividad A	DD-MM-YYYY HH:MI:SS	...	Valor 1	Valor 2
E	Actividad B	DD-MM-YYYY HH:MI:SS	...	Valor 1	Valor 2

un identificador A la huella digital que ha dejado en el proceso.

La otra tabla pilar del modelo de datos es la tabla de Casos. Esta tabla sería lo que viene siendo una tabla de Hechos, una tabla de información asociada a cada identificador que servirá para poder segmentar la información y poder filtrar y analizar basándose en las diferentes dimensiones de esta tabla de Casos.

Un ejemplo práctico tomando como referencia el usuario C de la Tabla 3.3 sería el representado a continuación en la Tabla 3.4 Estos campos pueden contener

Tabla 3.4: Ejemplo de registro de evento con desviaciones

Usuario	Tipo de Usuario	Origen	...	Edad	Fecha de Nacimiento
C	Actividad A	Editor	...	56	DD-MM-YYYY

todo tipo de información, el único requisito es que tiene que ser un registro único para cada usuario o campo identificador, para así cumplirse la relación 1:N entre la tabla de Casos y la tabla de Actividades tal y como se muestra en la Figura 3.6.

Este sería el modelo mínimo a implementar en cualquier proyecto de Celonis y mediante el cual se podrá representar prácticamente cualquier proceso en la herramienta. Hasta en el proceso, si se consigue extraer esta modelización, se podrá visualizar el proceso mediante Minería de Procesos. Se pueden añadir tantas tablas como nivel de enriquecimiento se quiera añadir a los análisis, siempre teniendo en cuenta que la funcionalidad 1:N desde la tabla de Casos o tabla principal es obligatoria. Es lo que se llama comúnmente como el esquema o diseño relacional de *Estrella* [29] (véase la Figura 3.7), uno de los modelos

3.5. Estructura y organización de elementos de visualización

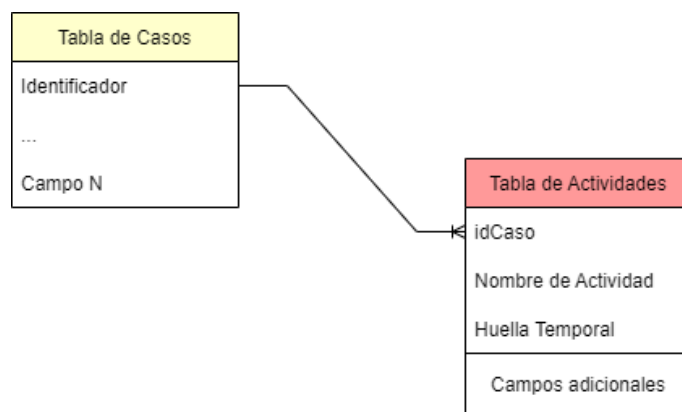


Figura 3.6: Modelo de relación entre tablas

de datos multidimensionales más sencillos de entender y analizar y aplicable a grandes cantidades de datos debido a su óptimo funcionamiento.

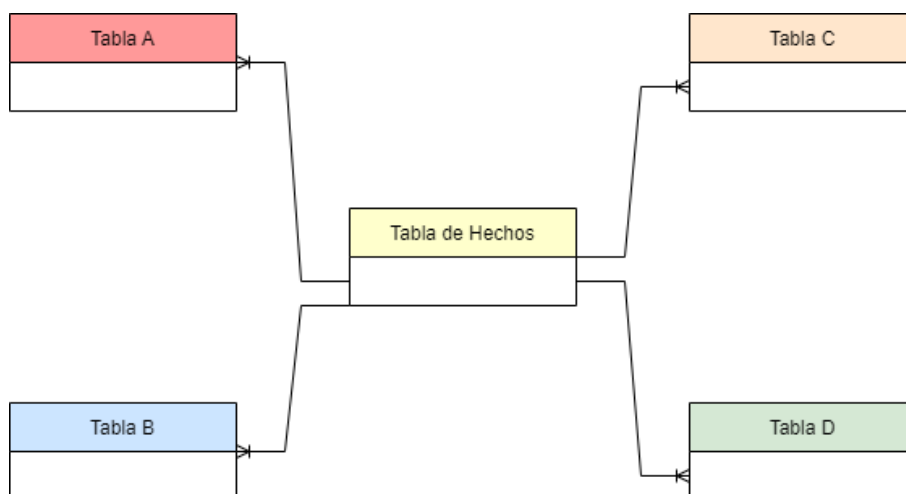


Figura 3.7: Modelo de relación entre tablas

Esta va a ser la metodología de diseño y de implementación que se va a aplicar para este trabajo, utilizando una entidad como tabla de Hechos o Casos y la tabla de Actividades unida a esta mediante una relación 1 a N para representar las actividades. Se añadirán tablas adicionales al modelo, según el caso y según el valor que aporte a los análisis posteriores.

3.5. Estructura y organización de elementos de visualización

Una vez explicada y organizada la parte inferior o parte más enfocada a un bajo nivel del trabajo, hay que aclarar también como se van a organizar y estructurar los elementos de visualización que serán la implementación a un alto nivel y que será el punto final o visible la implementación, desde donde se explotará la in-

formación. A un mayor nivel, las implementaciones, o proyectos se organizan en

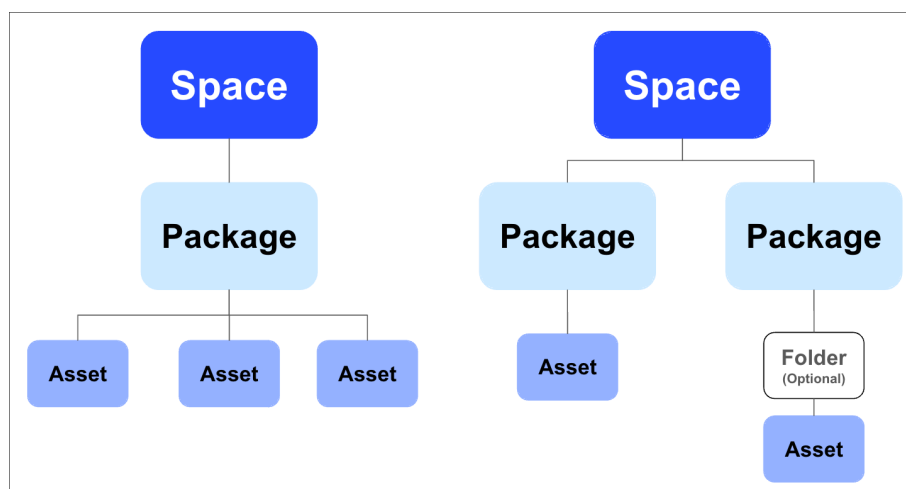


Figura 3.8: Estructuración de elementos

Espacios, elemento superior de organización de la parte de análisis de Celonis, estando referenciado este elemento a un Proceso o a un ente mayor como puede ser una empresa, un área o un territorio (por ejemplo, proyectos con un Espacio por País, un Espacio por área o incluso un Espacio por empresa a la que se le haya provisto servicios de implementación).

Por debajo, existen los paquetes, que puede ser similar a los espacios, pero con un Modelo de Datos asignado, pudiendo así tener varios modelos de datos en un mismo espacio y pudiendo diversificar las estrategias de implementación. Al igual que un espacio, se puede entender un paquete como un elemento de organización superior y los paquetes como subdivisiones de este elemento si así se considerase (véase la Figura 3.8).

El siguiente elemento, la carpeta, es opcional, pudiendo crearlas para una organización más óptima o no, ya depende del desarrollador.

Usando carpetas o no ya el siguiente nivel serían los elementos de visualización según se representa en la Figura 3.9, teniendo disponibles las siguientes funcionalidades:

- **Vista:** Vista de componentes visuales a modo de aplicación que permite visualizar los datos mediante gráficos, métricas y otros elementos.
- **Análisis:** Visualizaciones personalizadas del proceso y los datos de una manera más flexible.
- **Skill:** Funcionalidad de automatización de tareas identificadas dentro de los procesos desplegados.
- **Modelo de Conocimiento:** Almacén de conocimiento utilizable dentro del resto de elementos, como filtros, variables, métricas y código.
- **Flujo de Acción:** Elemento de automatización que dispara alarmados y automatismos que interactúan con los sistemas y los datos desplegados (p.e envío de un correo con casos que se salen de ciertos parámetros)

3.5. Estructura y organización de elementos de visualización

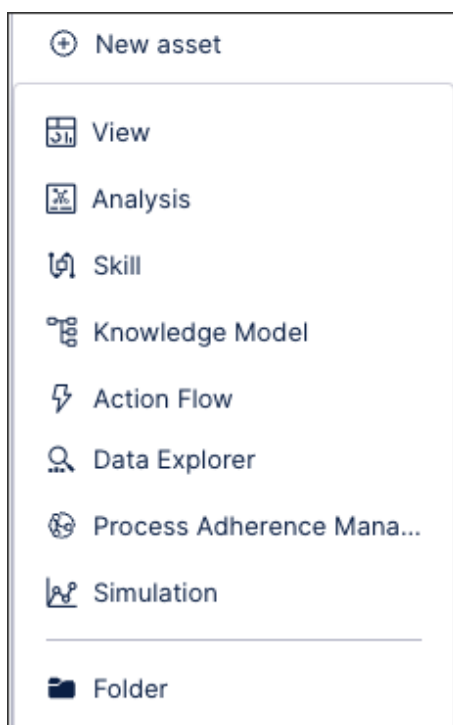


Figura 3.9: Elementos de Visualización

- Explorador de Datos: Elemento de validación de datos y de consultas de código.
- Process Adherence Manager: Evolución del árbol de proceso, permitiendo representar los eventos entre sí y su interacción en tiempo real.
- Simulación: Permite la creación de un gemelo digital del proceso, permite probar modificaciones sobre este y su impacto, ventajas y desventajas respecto al proceso original

En toda esta parte frontal o de cara al usuario, todos los componentes de las vistas se desarrollan en un lenguaje específico de Celonis llamado PQL (Process Query Language o Lenguaje de Procesos)[30]. Este lenguaje es un lenguaje de dominio específico de proceso diseñado y adaptado a un modelo de datos especial enfocado a los usuarios de negocio.

Este lenguaje permite transformar y traducir preguntas relacionadas con procesos en consultas que se ejecutan por el motor personalizado del lenguaje PQL. Esto cubre funciones específicas, desde operadores, agregaciones, operadores matemáticos, condicionales, etc. La sintaxis está inspirada en SQL, pero especializada para consultas de proceso basadas en el lenguaje natural que reduce la curva de aprendizaje y la curva de dificultad de uso de los usuarios menos técnicos que usen la herramienta.

La distribución y organización que se va a seguir en este proyecto de manera general según la Figura 3.10.

Este esquema es orientativo a nivel general de metodología de trabajo y para cada uno de los casos de uso que se implementen, se incluirá de manera detallada.

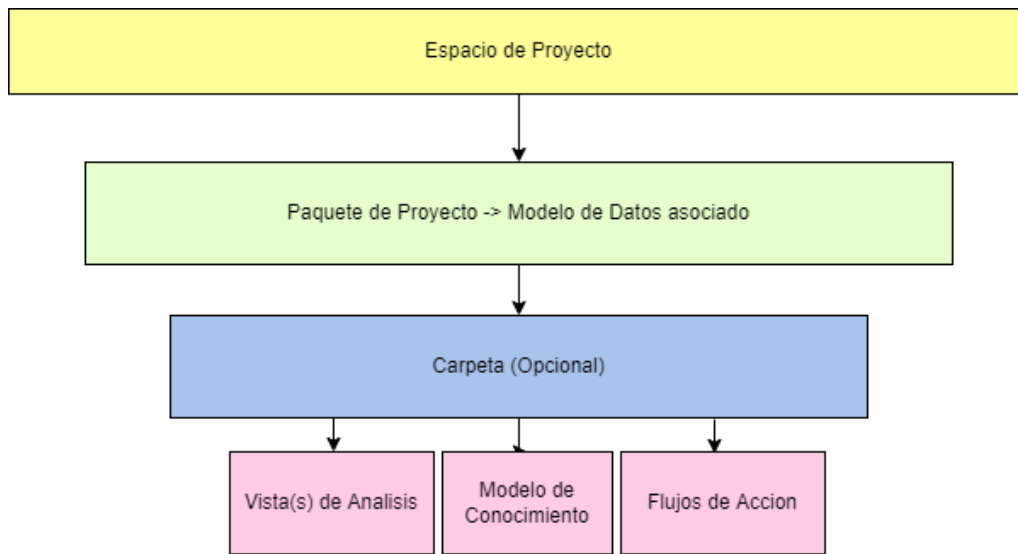


Figura 3.10: Elementos de Visualización

Capítulo 4

Casos de Estudio

En este capítulo se desarrollarán los dos casos de uso planteados. Por una parte, la aplicación de Process Mining sobre un fichero de datos de actividad de Moodle, a modo de prueba de concepto, mostrando las capacidades de Celonis y por la otra la construcción de cero de un caso existente en la universidad, el de Matriculación.

Esto servirá para identificar la idoneidad de los procesos universitarios para su aplicación de Process Mining. Se desglosarán las tareas que se han realizado desde la descripción de la fuente de datos hasta la implementación de las visualizaciones y oportunidades de mejora encontradas. Detallar aquí que tanto el caso de Moodle como el de matriculación son casos ficticios, con datos no reales, totalmente inventados, por lo que las conclusiones se ajustarán al planteamiento realizado en los objetivos generales de Process Mining de identificación de casos con desviaciones, mejora de tiempos medios y optimización de la productividad.

4.1. Fuente de datos de Moodle

El primero de los casos a implementar como prueba de concepto de Process Mining es la actividad del alumno en Moodle[31]. Consiste en el proceso que comienza cuando un alumno es dado de alta en Moodle por el administrador, el sistema o el profesor y este alumno interactúa con los elementos que tiene disponibles en la plataforma. Esto incluye actividades como ver un curso, revisar calificaciones o realizar entregas.

4.1.1. Ficheros de datos y estructura

Los datos fuente para este caso se han construido en conjunto con los tutores del trabajo, identificando las actividades que suceden en Moodle y utilizando así un set de datos lo más parecidos posible a la realidad. La fuente de datos es un conjunto de ficheros en formato Excel, uno por cada una de las asignaturas y que contienen campos desglosados según la Tabla 4.1.

A partir de aquí habrá que extender los datos para poder cumplir con la metodología de Celonis y que el caso de uso sea válido para poder visualizarlo con

Capítulo 4. Casos de Estudio

Tabla 4.1: Tabla fuente de datos por Asignatura

Campo	Tipo de Campo	Descripción	Ejemplo de valor
Fecha	Fecha	Huella temporal de la actividad	DD-MM-YYYY HH:MI:SS
Usuario	Texto	Usuario que realiza la actividad	Administrador
Usuario afectado	Texto	Usuario sobre el que se realiza la actividad	Pascualín García
Contexto del evento	Texto	Objeto sobre el que se realiza la actividad	Curso: Entrega A
Componente	Texto	Tipo de Objeto	Tarea
Nombre del evento	Texto	Acción del evento o Actividad que se realiza	Curso visto
Descripción	Texto	Descripción detallada de la actividad	Usuario '1234' realiza 'tarea' sobre elemento 'página'

Process Mining. Mediante la transformación de datos en Celonis se pueden realizar estas tareas de extensión de datos creando tablas adicionales que permitan conseguir el objetivo de un modelo de datos óptimo y estable.

4.1.2. Transformaciones

Debido a la estructura de Tabla de Casos y Tabla de Actividades, y teniendo un solo fichero con datos, hay que evaluar concienzudamente cómo se quieren enfocar las transformaciones. Lo más cómodo será utilizar la estructura mínima de tablas para esta prueba de concepto, las dos tablas anteriormente mencionadas, que habrá que extraer de este único fichero para conseguirlo.

Debido a la sencilla estructura de trabajo disponible, se va a emplear un único Trabajo de Datos (o *Data Job*) para todas las transformaciones. Y dentro de este *Data Job* se va a crear una estructura ordenada, documentada y modularizada de Transformaciones de código (véase la Figura 4.1).

La división del trabajo de datos consiste en 6 Transformaciones principales:

- 00. Test
- 01. Limpieza de Tablas
- 02. Creación de Tabla de Actividades
- 03. Creación de Tabla de Casos
- 04. Creación de tablas adicionales/auxiliares
- 05. Ajustes

La transformación 00. Test será como si fuese un banco de pruebas donde ejecutar consultas de todo tipo, pruebas, etc.

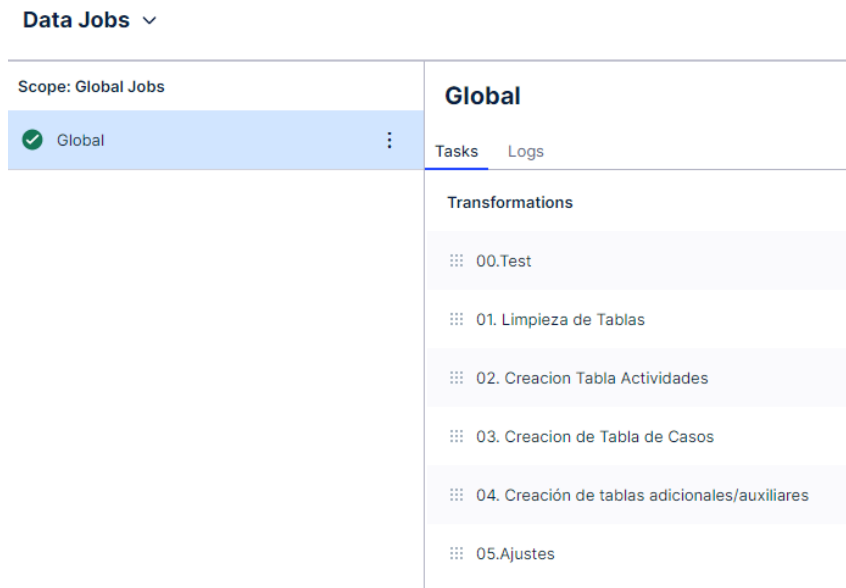


Figura 4.1: Lista de Transformaciones en la plataforma

La transformación 01. Test consiste en la limpieza de las tablas fuente en caso de requerir algún tratamiento a nivel de tipo de dato, fecha, cálculo, etc. En este caso, los cálculos no requieren de una limpieza de los datos exhaustiva, por lo que se realizara en la propia transformación de la tabla de Actividades.

La transformación 02. Creación Tabla Actividades consiste en la creación de la tabla de unión de los ficheros por asignatura, pivotando el origen del fichero en el campo *ASIGNATURA* y así poder segmentar a los alumnos por este campo. Por otra parte, se ha creado un campo *ENTREGA TIEMPO* que se rellenará si la entrega cumple con el plazo establecido o no.

En la creación de esta tabla se han anonimizado los *idUsuario* con una función MD5 [32] en caso de ser datos sensibles y de carácter personal. En este caso son datos ficticion, pero se prepara el caso para casos futuros donde no lo sean.

En la transformación 03. Creación de Tabla de Casos se cogerán los *idUsuario* de la Tabla Actividades y se generará una tabla de registros únicos, un *idUsua-*rio por cada Fila con los siguientes campos que se han creado para la obtención posterior de métricas importantes.

- *idUsuario*: Identificador
- *cantidadActividades*: Número de actividades que el alumno ha realizado en sus cursos en Moodle
- *duracionCaso*: tiempo persistente del usuario desde la primera actividad hasta la última.
- *numAsignaturas*: Número de asignaturas en las que participa el usuario.
- *ultimoEstadoCaso*: Última actividad o interacción del usuario con el sistema.

Capítulo 4. Casos de Estudio

- **ultimaConexion:** Última entrada del usuario al sistema como huella temporal.

En la transformación 04. Creación de tablas adicionales/auxiliares, se crean dos tablas adicionales para enriquecer el análisis, una es la tabla de Entregas, que contendrá un desglose de los alumnos y el n.º de entregas que han realizado. Por otra parte, una tabla ficticia de las fechas límite de entrega para poder calcular tareas y entregas fuera de tiempo, así como no entregadas.

En la transformación 05. Ajustes, se realizarán todos los cálculos de campos que no se calculen en la propia ejecución de las queries de Inserción de registros, esto incluyen actualizaciones de campos mediante sentencias UPDATE. En este caso particular se rellena el campo de entrega a tiempo, haciendo una unión de la tabla de Actividades con la de Fecha Límite de Entrega para ver con qué antelación o no se ha entregado una práctica, ejercicio o evento.

El código detallado de las transformaciones se adjunta en el Anexo para cada una de las transformaciones.

4.1.3. Modelo de Datos

Una vez aplicadas las transformaciones se dispondrán de las tablas limpias, normalizadas a falta de realizar las relaciones entre ellas. Tras este paso han quedado las tablas del modelo desglosadas de la siguiente manera:

- Tabla de Casos con la información de los alumnos (véase la Figura 4.2)
- Tabla de Actividades con la información de las actividades y eventos sucedidos (véase Figura 4.3)
- Tabla con la información de las entregas realizadas por los alumnos (véase Figura 4.4)
- Tabla con la información de las fechas límite de las entregas a realizar (véase Figura 4.5)

Tabla 4.2: Tabla de Casos

Campo	Tipo de Campo	Descripción
idUsuario	Varchar	Identificador de usuario anonimizado
cantidadActividades	Integer	N.º de eventos que hace el usuario
duracionCaso	Integer	Tiempo desde la primera actividad a la última
numAsignaturas	Integer	N.º de asignaturas en las que está matriculado el alumno
ultimoEstadoCaso	Varchar	Ultima actividad que ha realizado el usuario
ultimaConexion	Timestamp	Ultima vez que el usuario ha entrado al sistema

El modelo de datos a nivel relacional queda conformado según la figura 4.2 quedando su estructura y su volumetría confeccionada acorde a la figura 4.3.

4.1. Fuente de datos de Moodle

Tabla 4.3: Tabla de Actividades

Campo	Tipo de Campo	Descripción
USUARIO	Varchar	Identificador de usuario que realiza la actividad
USUARIO AFECTADO	Varchar	Usuario sobre el que recae la acción de la actividad, si procede
TIMESTAMP	Timestamp	Huella temporal de la actividad
CONTEXTO EVENTO	Varchar	Actividad y tipo detallado
COMPONENTE	Varchar	Tipo de actividad
ACTIVITY NAME	Varchar	Evento realizado por el usuario.
DESCRIPCIÓN	Varchar	Descripción detallada de la actividad realizada por el usuario
ORIGEN	Varchar	Tipo de acceso del usuario
ASIGNATURA	Varchar	Asignatura en la que está matriculado el alumno
ENTREGA TIEMPO	Integer	Días de diferencia de la entrega, si procede, respecto a la fecha límite de entrega.

Tabla 4.4: Tabla de entregas de alumnos

Campo	Tipo de Campo	Descripción
USUARIO	Varchar	Identificador de usuario
ASIGNATURA	Varchar	Asignatura en la que está matriculado el alumno
numEntregas	Integer	N.º de entregas que ha realizado el alumno
totalEntregas	Integer	N.º de entregas que debe de realizar el alumno

Tabla 4.5: Tabla de fechas límite de entrega

Campo	Tipo de Campo	Descripción
CONTEXTO EVENTO	Varchar	Nombre de la entrega
ASIGNATURA	Varchar	Asignatura de las entregas
FECHA ENTREGA	Timestamp	Fecha Límite de entrega del evento

Si se echa una vista atrás a la arquitectura que se desglosó con detalle de Celsonis (véase la Figura 4.4), se puede apreciar que 4 de los 5 pasos de desarrollo estarían completos al realizar esta modelización de los datos y quedarían realizar las tareas más técnicas de implementación y desarrollo y empezar a trabajar y explotar los datos buscando sacar las métricas deseadas.

Aquí entra la siguiente fase de implementación de vistas y métricas, donde mediante la utilización de todos los componentes visuales y el lenguaje PQL se puede hacer desde 0, o desde una plantilla en caso de ser un caso o un proceso estándar, todas las visualizaciones y métricas necesarias para los casos de uso.

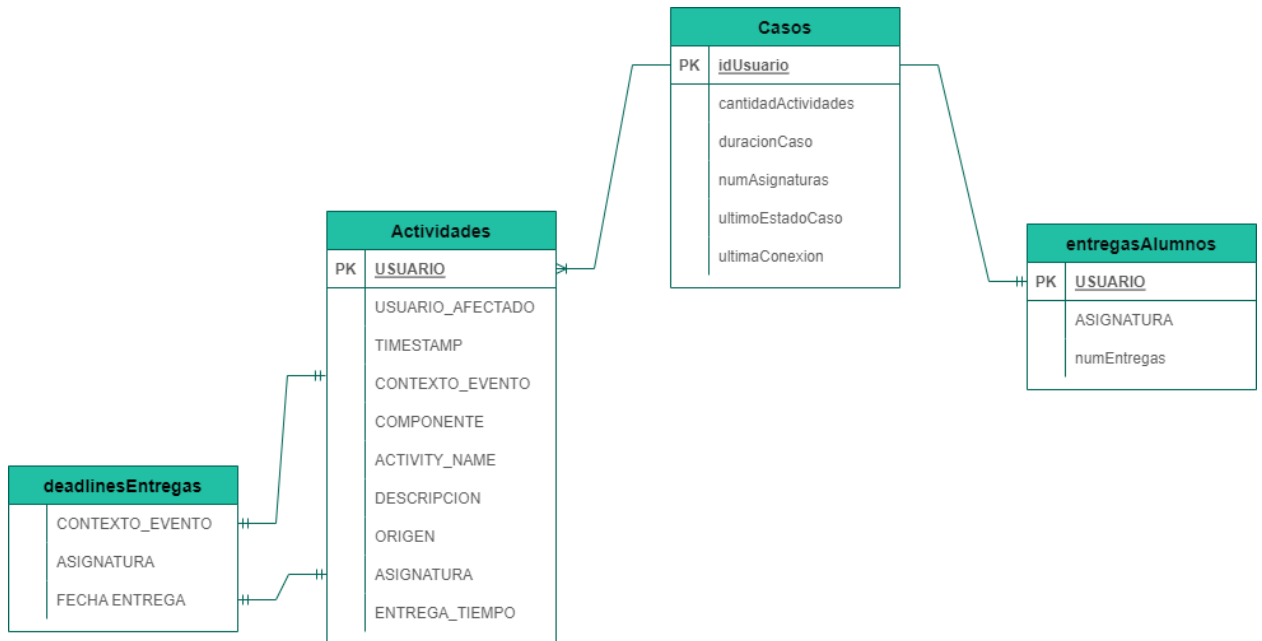


Figura 4.2: Modelo Relacional de Datos

Loaded Data Model ✕

Table sync ✔ >

Data load progress ✔ ▾

Table name ⬆️⬆️	Loaded rows ⬆️⬆️
Casos	613
Logs_Moodle_No_Hash	105 627
deadlinesEntregas	22
entregasAlumnos	484

Figura 4.3: Volumetría de las tablas cargadas en el modelo de datos

4.1.4. Implementación y visualización

Para este caso de uso, que hace una función de prueba de concepto o de valor de Celonis y Minería de procesos sobre un ejemplo sencillo y básico, pero enfocado a la universidad, se va a tratar de realizar la mínima implementación posible, para mostrar un mínimo producto viable de esta tecnología. ¿Por qué? Para demostrar que con la potencialidad y esfuerzo en Minería de Procesos en un ecosistema que tiene tantos procesos existe gran capacidad de mejora con un esfuerzo moderado. Se ha planteado una estructura estándar para este caso

4.1. Fuente de datos de Moodle

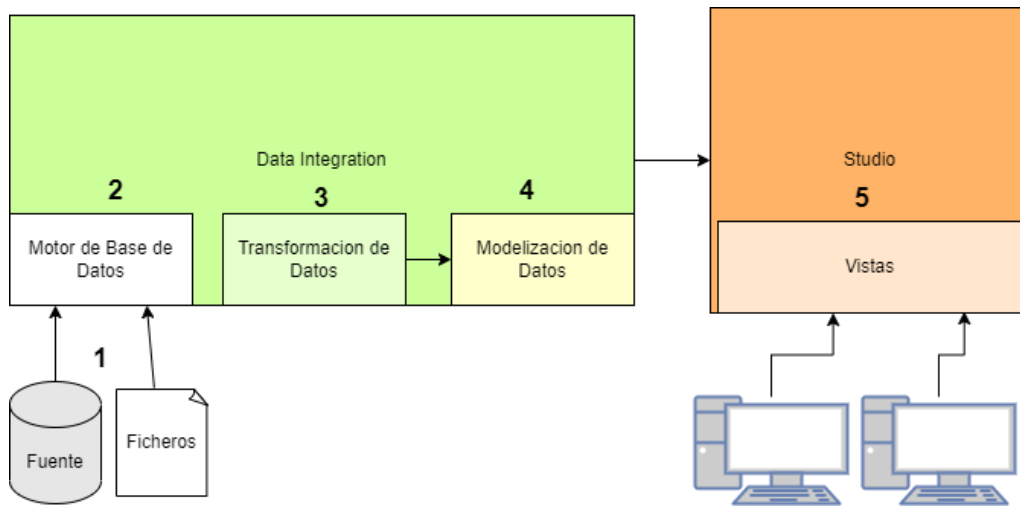


Figura 4.4: Arquitectura del proceso en Celonis

de uso que contendrá, a nivel de organización de Espacio, paquete y elemento visual, según las Figuras 4.5 y 4.6

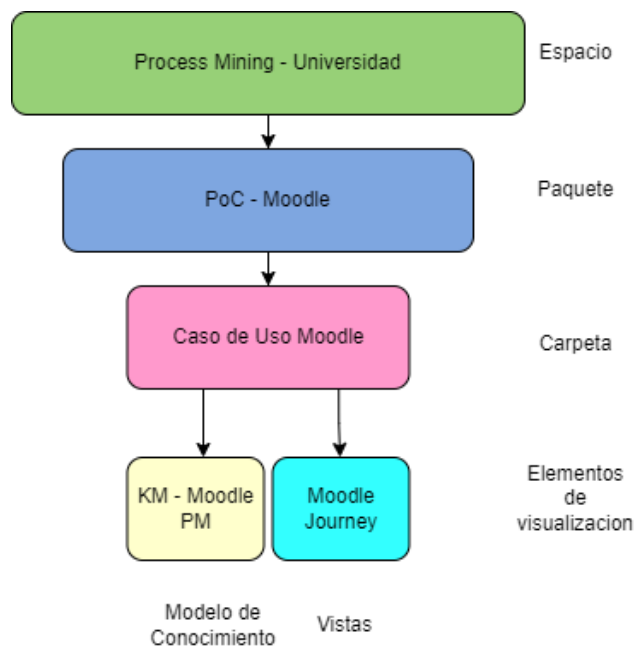


Figura 4.5: Organización de los elementos de visualización

Modelo de Conocimiento

En primer lugar, en el modelo de conocimiento se han almacenado todas las métricas que se han desarrollado para este caso de uso. Esto permite mayor flexibilidad, escalabilidad y modularización de cualquier proyecto e implementación, teniendo un contenedor de elementos y que si se reutilizase alguno de estos, en caso de cambio, solo habría que modificar una vez dicho elemento y no

Capítulo 4. Casos de Estudio

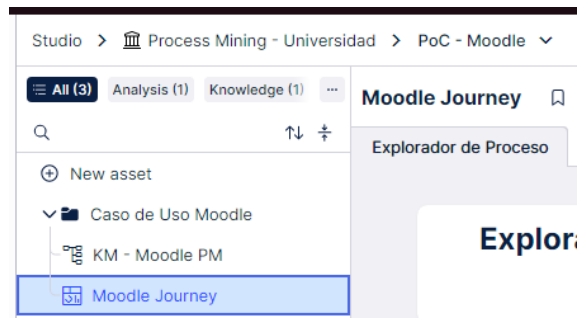


Figura 4.6: Organización de los elementos de visualización II

N veces que aparece a lo largo de todo el desarrollo (Véase la Figura 4.7).

Records	4	KPIs		
KPIs	21	Name ↕	ID ↕	PQL Formula ↕
Augmented Attributes	0	# Accesos	accesos	count(case when "Logs_Moodle_No...
Filters	0	# Alumnos	alumnos	COUNT (DISTINCT "Casos"."idUsua...
Triggers	0	# Alumnos que no han realizado ...	alumnos_que_no_han_realizado_nin...	sum(CASE WHEN PU_COUNT ("Cas...
Actions	0	# Entregas a realizar	entregas_a_realizar_	COUNT ("deadlinesEntregas"."CON...
Flags	0	# Entregas a tiempo	entregas_a_tiempo	COUNT(CASE WHEN "Logs_Moodle...
Variables	0	# Entregas fuera de tiempo	entregas_fuera_de_tiempo	COUNT(CASE WHEN "Logs_Moodle...
Event Logs	1	# Entregas realizadas	entregas_realizadas	sum (CASE WHEN "Logs_Moodle_N...
Custom Objects	0	% alumnos que no han realizado ...	alumnos_que_no_han_realizado_entr...	count(distinct CASE WHEN PU_COU...
		% sobre el total	sobre_el_total	count(distinct CASE WHEN PU_COU...
		Alumnos sin entregas	alumnos_sin_entregas	count(distinct CASE WHEN PU_COU...
		Count Table Casos	COUNT_TABLE__CASOS	COUNT_TABLE("Casos")
		Count Table DeadlinesEntregas	COUNT_TABLE__DEADLINESENTRE...	COUNT_TABLE("deadlinesEntregas")
		Count Table EntregasAlumnos	COUNT_TABLE__ENTREGASALUMN...	COUNT_TABLE("entregasAlumnos")
		Count Table Logs Moodle No Ha...	COUNT_TABLE__LOGS_MOODLE_NO...	COUNT_TABLE("Logs_Moodle_No_H...
		En tiempo	en_tiempo	sum (CASE WHEN "Logs_Moodle_N...

Figura 4.7: Modelo de conocimiento del proceso de Moodle

Estas fórmulas se van añadiendo a la vez que se va desarrollando o incluso en primer lugar se definen, calculan y prueban los KPIs antes de desarrollar las vistas. En este caso, según se ha avanzado la creación de las vistas se han ido añadiendo de manera incremental.

Process Explorer

En primer lugar, el Proceso es el núcleo del caso de estudio. Todo el trabajo previo se ha realizado con el objetivo principal de extraer y representar el proceso. Esto se logra a través de la funcionalidad del *Process Explorer* [33], que muestra,

mediante actividades y conexiones, todas las variantes del proceso, tal como se ilustra en la Figura 4.8.

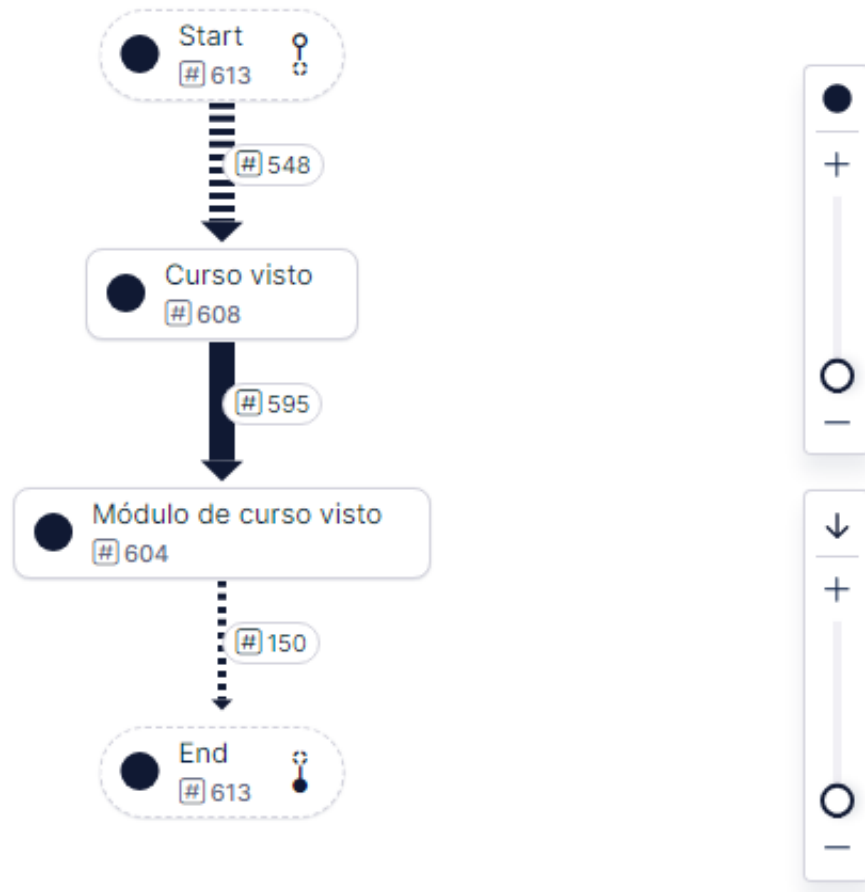


Figura 4.8: Explorador de Proceso

Este sería el camino más común, donde un caso comienza su proceso, ve un curso, accede y ve a un módulo de curso y acaba el proceso, estas actividades aparecen en al menos una parte del proceso para 608 de los 613 casos que comienzan el proceso. A partir de aquí y mediante el desplegable de Actividades y Conexiones, se pueden ampliar las diferentes variantes y posibilidades que tiene el proceso (véase figura 4.9):

Desplegando actividades se puede ver como un proceso «sencillo» pasa a ser algo más enrevesado con más actividades y conexiones. Incluso se puede desplegar al máximo actividades y conexiones y llegando a lo que se conoce comúnmente como el "Gráfico Espagueti"(Figura 4.10).

Este gráfico es totalmente inmanejable, ya que un usuario no podría establecer conclusiones precisas ni obtener ninguna información de utilidad del mismo. Lo que sí se puede ver es que de un proceso que parece sencillo, con un set

Capítulo 4. Casos de Estudio

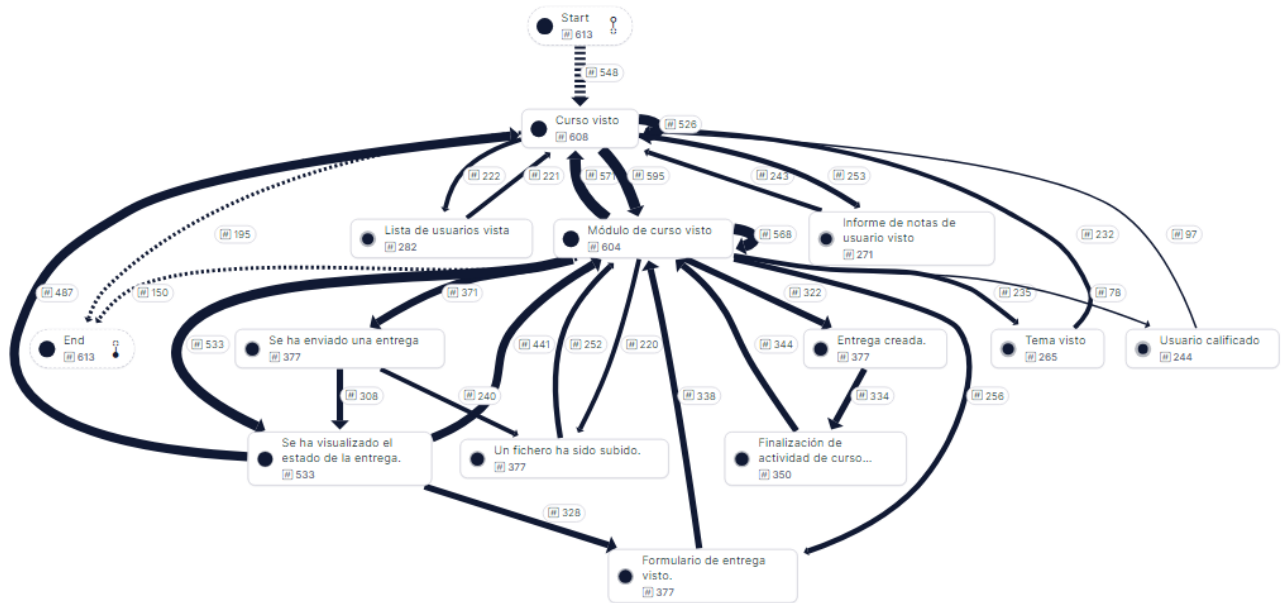


Figura 4.9: Explorador de proceso parcialmente ampliado a nivel de actividades y conexiones

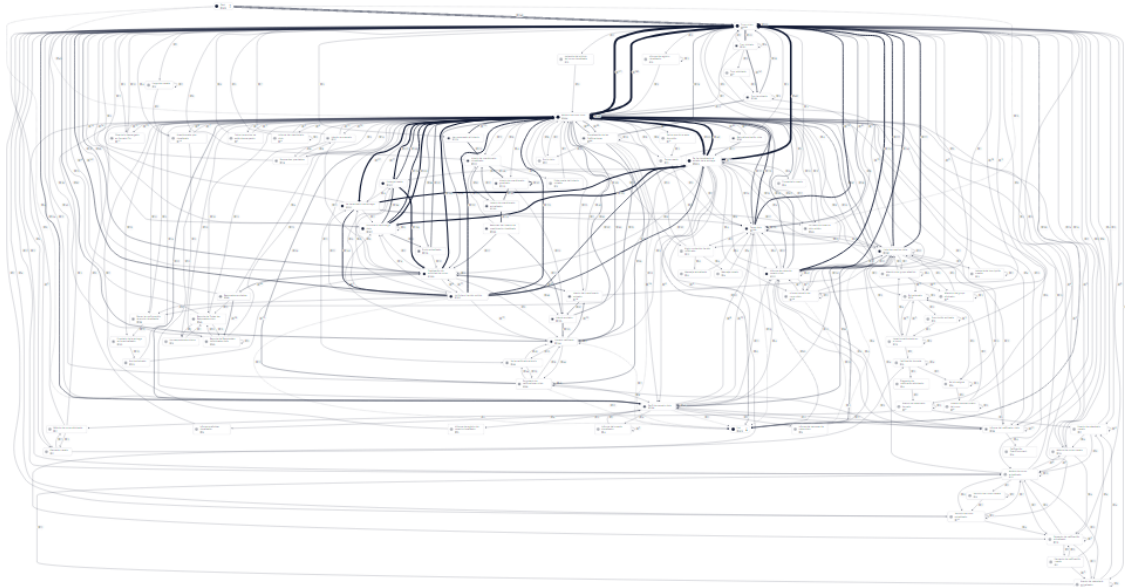


Figura 4.10: Explorador de proceso completo - Diagrama espagueti

de datos a priori bastante acotado tanto en registros como actividades, se despliega y extrae un proceso con infinitud de variantes y conexiones entre ello. La pregunta es: ¿Se puede trabajar con esto? No es recomendable, pero si se van metiendo variantes conocidas (por ejemplo un intento de cuestionario) y cambiando de métricas de Frecuencia de Casos a Tiempo Medio de Proceso se puede analizar por partes esta inmensidad para analizar casos que se salgan de norma

4.1. Fuente de datos de Moodle

o en aquellos que ya hay certeza de que desvirtúan el proceso y confirmar estas sospechas.

Esto se consigue seleccionando la actividad deseada y cambiando la métrica de visualización para adaptar el Process Explorer a la investigación en particular, como se indican en las Figuras 4.11 y 4.12.

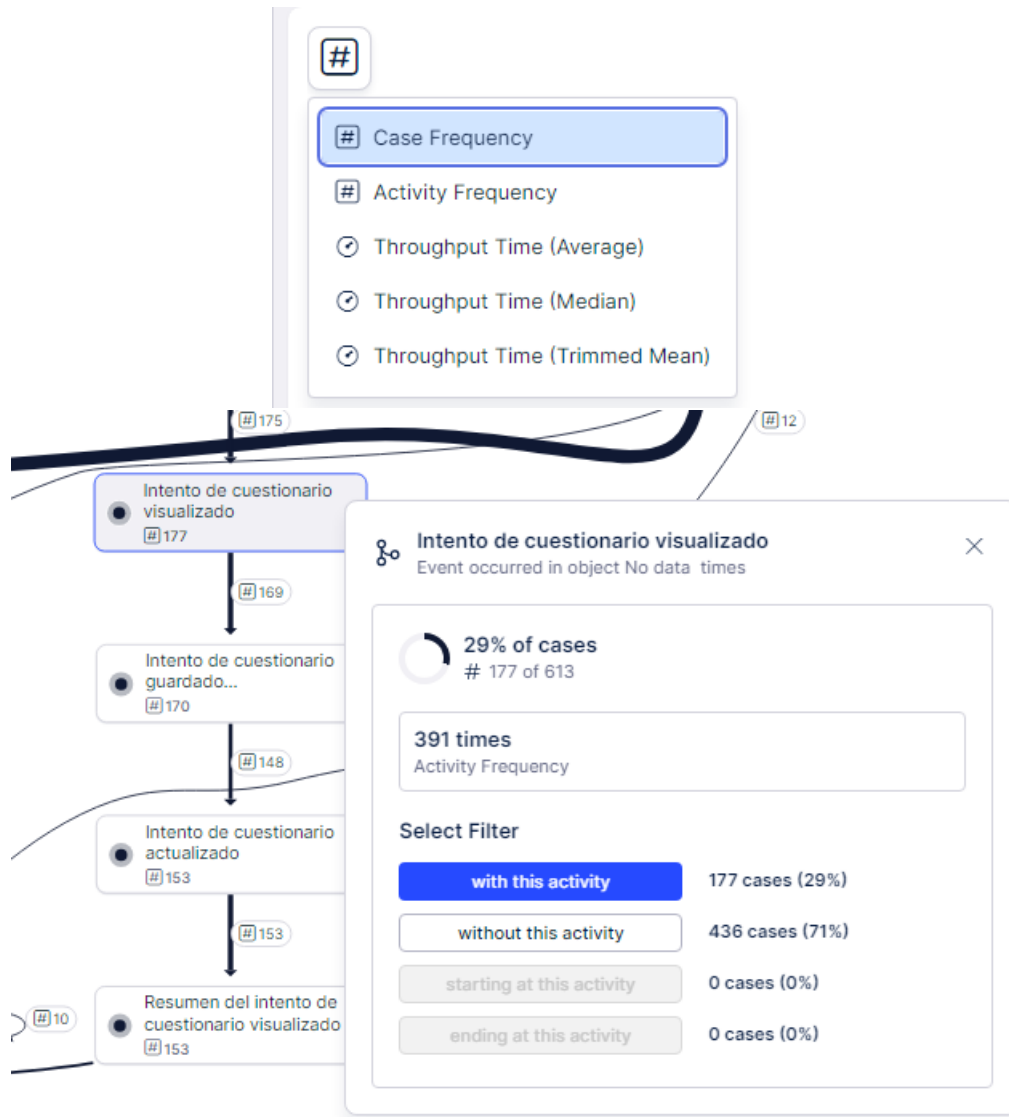


Figura 4.11: Filtros de Actividad y selección de dimensión

A partir de aquí se pueden estudiar variantes más comunes, incluir variantes que involucren a una actividad detallada, incluir conexiones específicas o excluir actividades, dinamizando así el análisis y consiguiendo que se encuentre y haga foco en las desviaciones o variantes no deseadas.

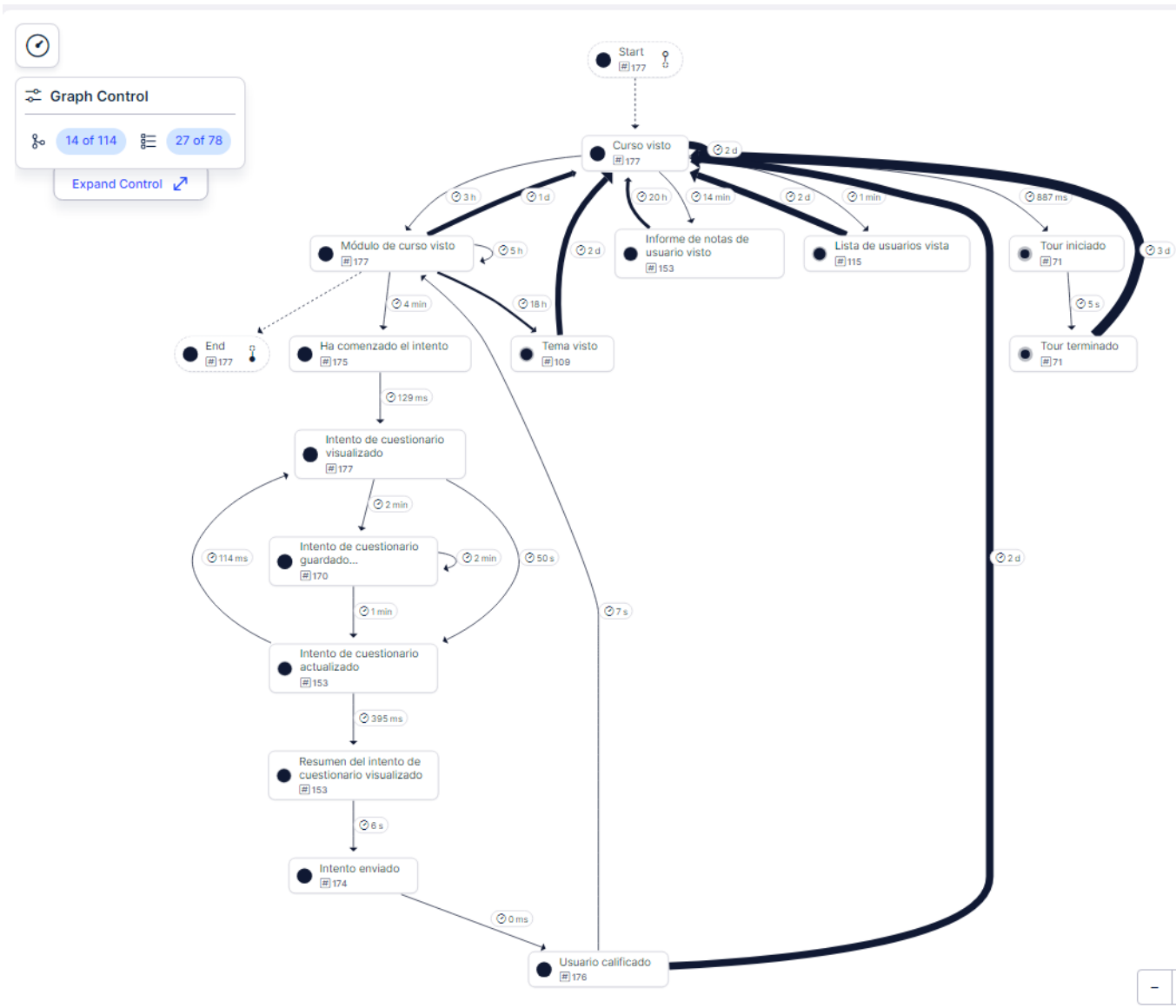


Figura 4.12: Árbol de Proceso filtrado por actividad

Explorador de Casos

En esta fase, ya disponemos de un árbol de proceso específico, pero se necesita saber más sobre cada uno de los casos que están pasando por esta variante. Al igual que se dispone de una tabla de actividades, también una tabla de casos con información adicional de cada registro que ejecuta una actividad, como puede ser el origen o cualquier dimensión que aparezca en los datos para analizar. En este caso con el *Case Explorer* o Explorador de casos se puede encontrar el detalle de estos casos y así profundizar un nivel más en el detalle del análisis (véase figura 4.13).

Con esta vista se puede identificar todo el proceso seguido por un caso actividad a actividad, por cada una de sus huellas temporales.

4.1. Fuente de datos de Moodle

Case Id	# of Activities	Throughput Time	First Activity	First Activity Timestamp
Admin	3112	248 d	Contenido del curso ...	2023-07-05 10:53:39
94524cf4ab3ab9ea5b93d03922e...	1267	224 d	Curso visto	2023-07-27 21:08:57
7da68f63f0842ea2e0e81d89a6c6...	4045	223 d	Curso visto	2023-07-28 08:33:47
8252c8f4a857984e2a45a71b63b3...	215	209 d	Curso visto	2023-08-01 08:29:42
81d326223ea7dbcb209a52a2278b...	93	195 d	Curso visto	2023-09-01 11:41:20
System	713	194 d	Rol asignado	2023-09-01 10:19:06
cca3e3648177a2464fc51b7d016f7...	48	194 d	Curso visto	2023-09-01 12:43:40
9a831b0b53f60046d05657fa9b1f6...	246	193 d	Curso visto	2023-09-02 15:01:18
ecb7e7678ad8e488b06c2807d9c6...	54	193 d	Curso visto	2023-09-03 21:53:56
e050f6f991114508ea194f18f40f80a	201	192 d	Curso visto	2023-09-01 13:59:02
718f38dcc99b61e75a6b7ae40628...	363	192 d	Curso visto	2023-09-03 00:13:42
7de66c002fcc2b8af45b3b25cf47...	197	192 d	Curso visto	2023-09-03 19:04:48
3102655d85b752ab6cc030e0ede...	398	192 d	Curso visto	2023-09-02 00:13:14

Case details: ecb7e7678ad8e488b06c2807d9c66918

Search

Activities 54 Items

- Curso visto 2023-09-03 21:53:56 8 months ago
- Curso visto 2023-10-20 11:25:22 7 months ago
- Curso visto 2024-01-16 10:57:55 4 months ago
- Curso visto 2024-01-19 15:02:33 4 months ago
- Módulo de curso visto 2024-01-19 15:02:38 4 months ago
- Módulo de curso visto 2024-01-19 15:04:15 4 months ago
- Curso visto 2024-01-22 10:00:46 4 months ago
- Curso visto 2024-01-22 14:28:17 4 months ago
- Curso visto 2024-01-23 11:02:41 4 months ago
- Curso visto 2024-02-07 17:44:14 3 months ago
- Módulo de curso visto 2024-02-07 17:44:22 3 months ago
- Curso visto 2024-02-08 13:50:16 3 months ago
- Módulo de curso visto 2024-02-08 13:50:24 3 months ago
- Curso visto 2024-02-09 12:20:02 3 months ago
- Curso visto 2024-02-09 16:24:30 3 months ago
- Módulo de curso visto 2024-02-09 16:24:32 3 months ago

Figura 4.13: Explorador de casos con el detalle por usuario

Una vez explicados los dos principales elementos de análisis utilizados, las vistas implementadas han sido las siguientes:

Capítulo 4. Casos de Estudio

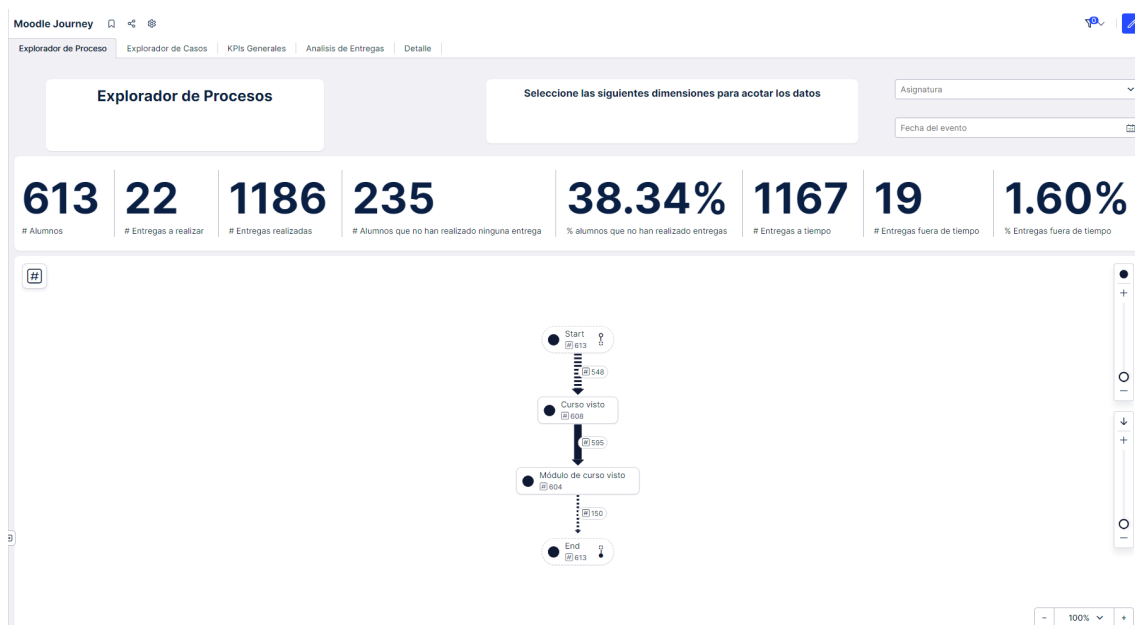


Figura 4.14: Pestaña de Explorador de Procesos

Explorador de proceso

La primera de las pestañas (véase la Figura 4.14) implementadas se centra en el Explorador de Procesos que se ha explicado anteriormente y una lista de métricas principales, que son las siguientes:

- Número de alumnos
- Número de entregas a realizar
- Número de entregas realizadas
- Número de Alumnos que no han realizado ninguna entrega
- Porcentaje de alumnos que no han realizado entregas
- Número de entregas a tiempo
- Número de entregas fuera de tiempo
- Porcentaje de entregas fuera de tiempo

Pero la primera acción que se realiza es seleccionar las dimensiones para acotar el universo de datos, a nivel de Asignatura y de Ventana temporal, acotando al gusto así a ciertas estaciones en caso de que así sea necesario para el análisis.

Filtros

La lista de métricas se desglosa como una tarjeta encima del explorador de proceso, donde se puede ver a niveles generales y donde se selecciona tanto la asignatura que se quiera filtrar como la ventana temporal deseada (véase la Figura 4.15).

4.1. Fuente de datos de Moodle

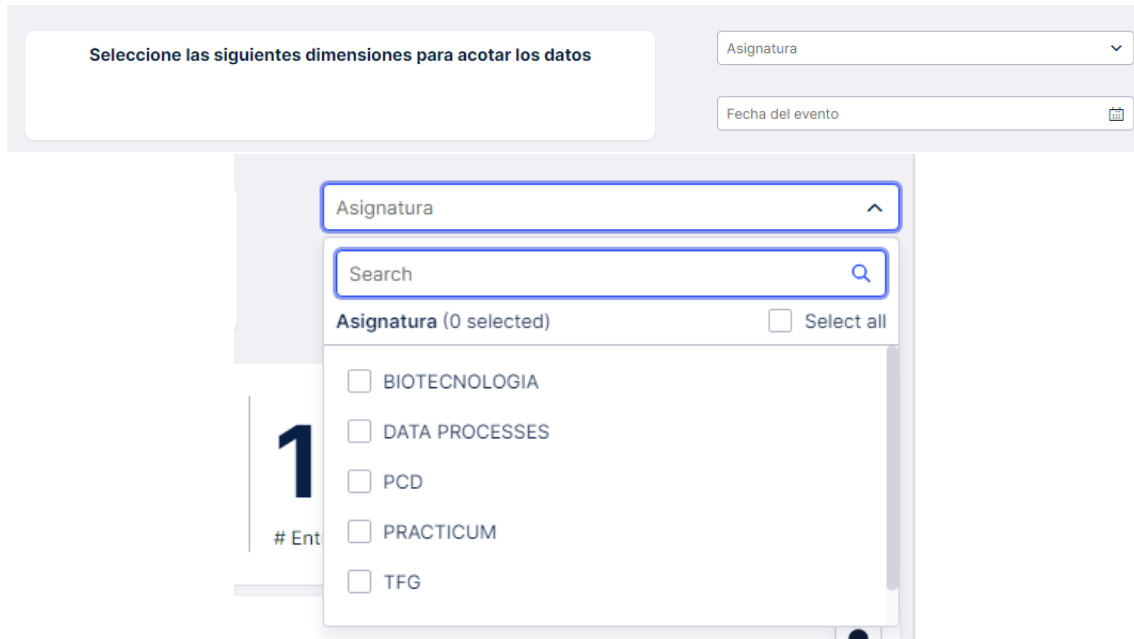


Figura 4.15: Funcionalidad de filtrado por asignatura y fecha

Lista de KPIs generales de proceso



Figura 4.16: Métricas principales del proceso

En esta tarjeta de métricas (Figura 4.16) en un simple vistazo se pueden revisar todos los números que puedan interesar respecto a los alumnos y a las entregas y apreciar como se van moviendo los porcentajes, con base en los filtros, actividades y análisis realizados. Por ejemplo, se aprecia que para los datos utilizados el porcentaje de las entregas realizadas fuera de tiempo es bastante pequeño, así como los alumnos que no han realizado ninguna entrega es alto respecto al total, pudiendo analizar las causas raíces de estos alumnos y asignaturas. Y el explorador de procesos explicado anteriormente desglosado (véase la Figura 4.17) en la parte inferior para complementar estas métricas.

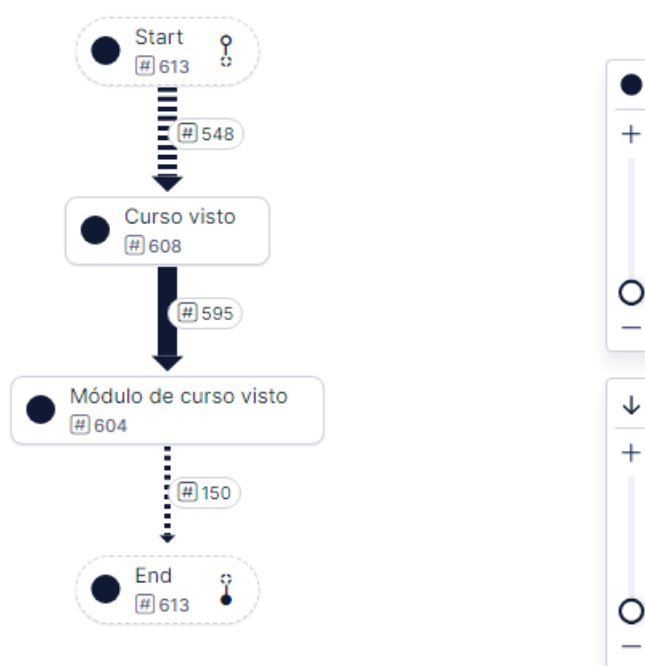


Figura 4.17: El gráfico explorador de proceso

Explorador de Casos en la visualización

La segunda pestaña contendrá el explorador de casos con el que se podrá bajar a detalle la información que se vaya filtrando (véase la Figura 4.18). Al igual que lo comentado en la sección 4.1.4 este componente ofrece la información que se ha preparado anteriormente para cada caso o identificador de alumno.

KPIs Generales

La siguiente pestaña implementada es la de KPIs generales, donde se tienen los filtros comentados anteriormente, la lista de KPIs y varios gráficos de utilidad para seguir ahondando en la explotación de los datos.

Los gráficos que se han desarrollado son:

- Desglose de alumnos por asignatura (Figura 4.20).
- Número de accesos a lo largo del tiempo (Figura 4.21).
- Distribución de las actividades y eventos (Figura 4.22).

El desglose de alumnos por asignatura, pudiendo así tener una volumetría general del número de alumnos que se tiene en clase, en este caso y como parece ser más que obvio, el Practicum y el TFG, asignaturas obligatorias y troncales acumulan un mayor número de alumnos.

Con el Número de accesos a lo largo del tiempo, se puede observar la estaciona-

4.1. Fuente de datos de Moodle

Moodle Journey

Explorador de Casos | KPIs Generales | Analisis de Entregas | Detalle

Case Id	# of Activities	Throughput Time	First Activity	First Activity Timestamp	Last Activ
Admin	3112	248 d	Contenido del curso ...	2023-07-05 10:53:39	Usuario desma
94524cfda3ab9ea5b93d03922e8...	1267	224 d	Curso visto	2023-07-27 21:08:57	Lista de usuar
7da68f63f0842ea2e0e81d89a6c67...	4045	223 d	Curso visto	2023-07-28 08:33:47	Curso visto
8252c8f4a857984e2a45a71b63b3...	215	209 d	Curso visto	2023-08-01 08:29:42	Módulo de curs
81d326223ea7dbcb209a52a2278b...	93	195 d	Curso visto	2023-09-01 11:41:20	Módulo de curs
System	713	194 d	Rol asignado	2023-09-01 10:19:06	Usuario califica
cca3e3648177a2464fc51b7016f74...	48	194 d	Curso visto	2023-09-01 12:43:40	Curso visto
9a831b0c53f60046d0e6057a9b1f6...	246	193 d	Curso visto	2023-09-02 15:01:18	Se ha visualiza
ecb7e767ad8e488b0bc2807d9c6...	54	193 d	Curso visto	2023-09-03 21:53:56	Se ha visualiza
e050f6991114508ea1941f8f04f80a	201	192 d	Curso visto	2023-09-01 13:59:02	Curso visto
718f38acc99b1e75a6b7ae40628c...	363	192 d	Curso visto	2023-09-03 00:13:42	Curso visto
7de66c002fcc2b8a45b3c25c947d...	197	192 d	Curso visto	2023-09-03 19:04:48	Módulo de curs
3102655d85b752ab6cc030e0e0ee...	398	192 d	Curso visto	2023-09-02 00:13:14	Curso visto
3ef3e6d4f952abfa8deabb9cd55a...	250	192 d	Curso visto	2023-09-03 22:28:06	Tema visto
851d359318a219dc0cdd7af59153...	280	191 d	Curso visto	2023-09-03 17:56:11	Se ha visualiza
efa2b67a78be9fc73e00b6e3e913...	205	190 d	Curso visto	2023-09-05 10:51:55	Curso visto
800ca10e671382a3f71b400849114...	121	190 d	Curso visto	2023-09-03 19:05:39	Se ha visualiza
711bc637a45beddc0b38e3ba8d83...	613	190 d	Curso visto	2023-09-03 13:43:22	Curso visto
81d1a67041f8573e25a0a7531d956...	313	190 d	Curso visto	2023-09-05 09:28:16	Módulo de curs
2ec8ef77587b331e4b31622770a6a...	930	189 d	Curso visto	2023-09-07 11:05:18	Informe de regi
a14217c1459cb52430ead9bc47977...	337	189 d	Curso visto	2023-09-04 13:49:36	Curso visto
199d7513b98dcd32d5c5927416a0...	85	189 d	Curso visto	2023-09-05 19:31:36	Curso visto
a44a4a138a885df59d127ab459a9...	140	189 d	Curso visto	2023-09-03 20:25:33	Curso visto
e935e94a562504522fbc9e202755...	174	188 d	Curso visto	2023-09-07 20:58:39	Curso visto
5ac46d5c3a30f6284036a45079fa...	176	188 d	Curso visto	2023-09-01 21:31:43	Curso visto
0ca72a91d3d3c67da50b204e89a...	276	188 d	Curso visto	2023-09-06 13:27:02	Curso visto

Case details: Admin

Search

Actividades 1000 Items

- Contenido del curso eliminado 10 months ago
- Elemento de calificación creado 10 months ago
- Elemento de calificación creado 10 months ago
- Elemento de calificación eliminado 10 months ago
- Elemento de calificación eliminado 10 months ago
- Elemento de calificación eliminado 10 months ago
- Contenido del curso eliminado 10 months ago
- Instancia de inscripción creada 10 months ago
- Sección del curso creada 10 months ago
- Sección del curso creada 10 months ago
- Sección del curso creada 10 months ago
- Sección del curso creada 10 months ago
- Sección del curso creada 10 months ago
- Sección del curso creada 10 months ago
- Sección del curso creada 10 months ago
- Elemento de calificación actualizado 10 months ago
- Elemento de calificación creado 10 months ago
- Elemento de calificación creado 10 months ago
- Elemento de calificación creado 10 months ago
- Sección del curso creada 10 months ago
- Sección del curso creada 10 months ago
- Sección del curso creada 10 months ago

Figura 4.18: Gráfico de Explorador de casos dentro de la pestaña

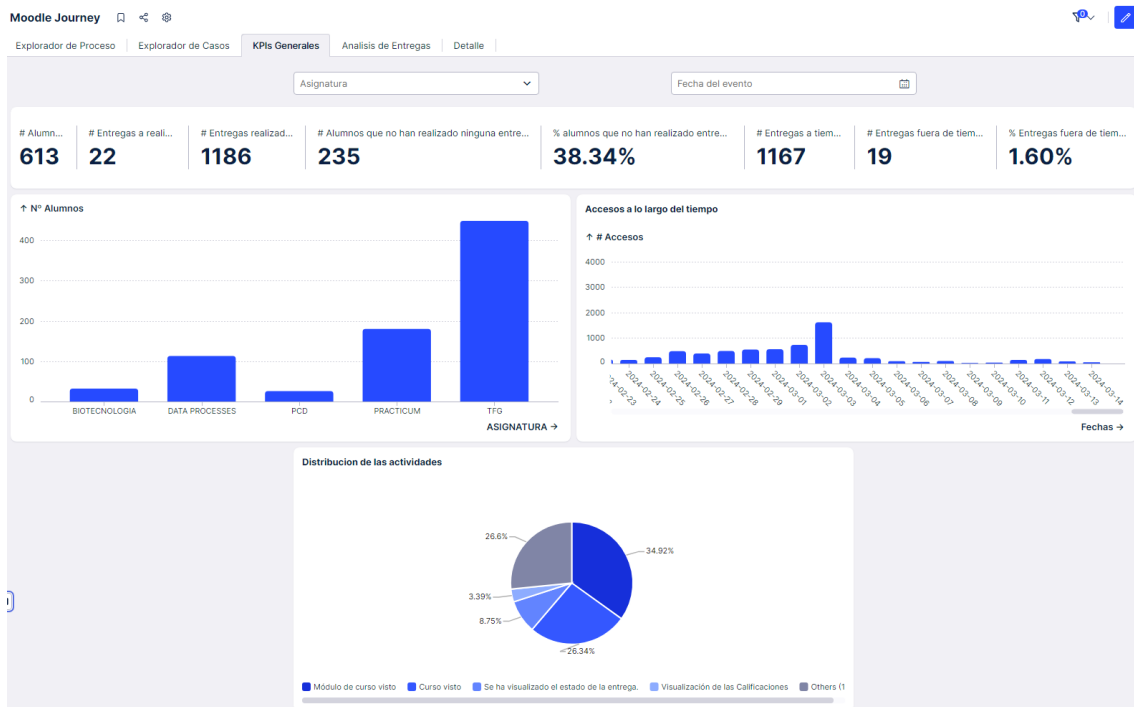


Figura 4.19: Pestaña completa de KPIs generales

alidad, y cómo están interactuando los alumnos con el Moodle. Se puede apreciar que la fecha de los primeros parciales y entregas acumula un mayor número

Capítulo 4. Casos de Estudio

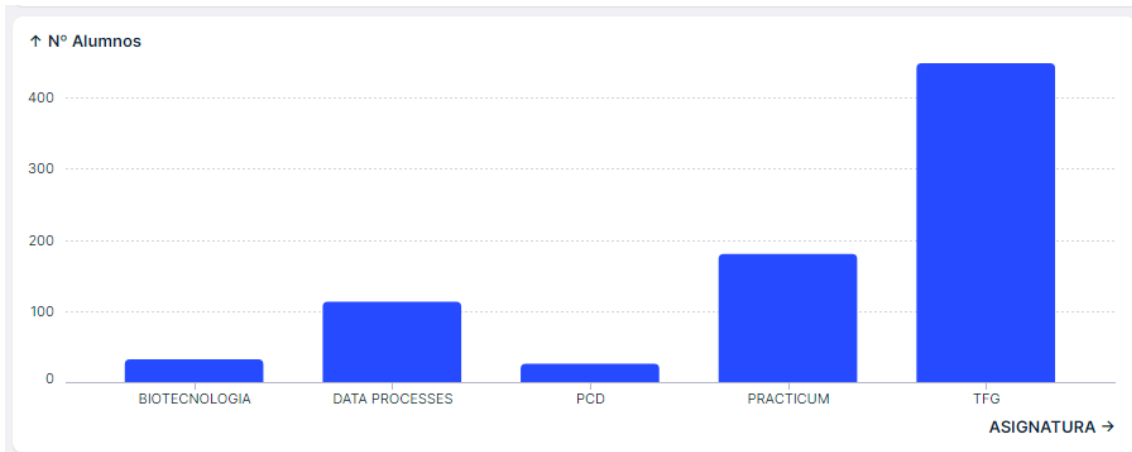


Figura 4.20: Gráfico de distribución de alumnos por asignatura

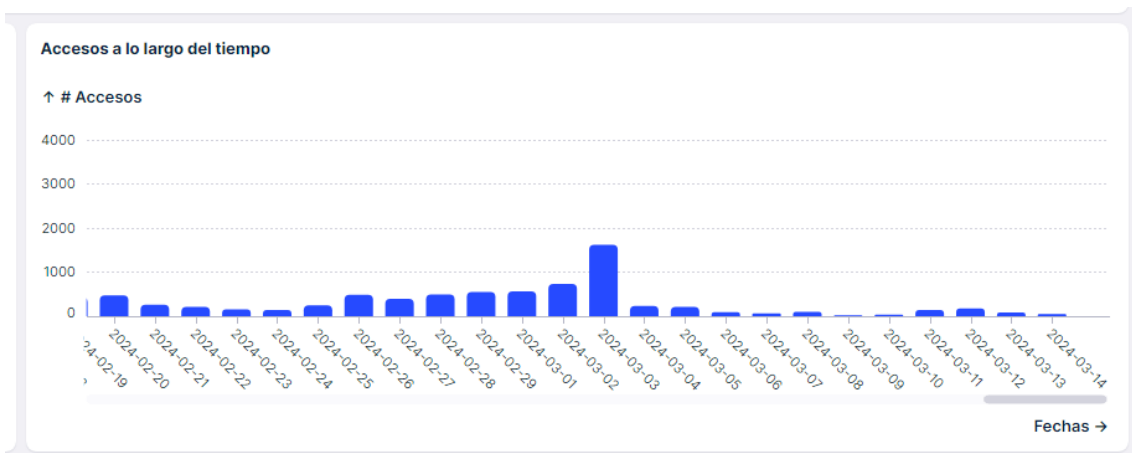


Figura 4.21: Gráfico de número de accesos a lo largo del tiempo

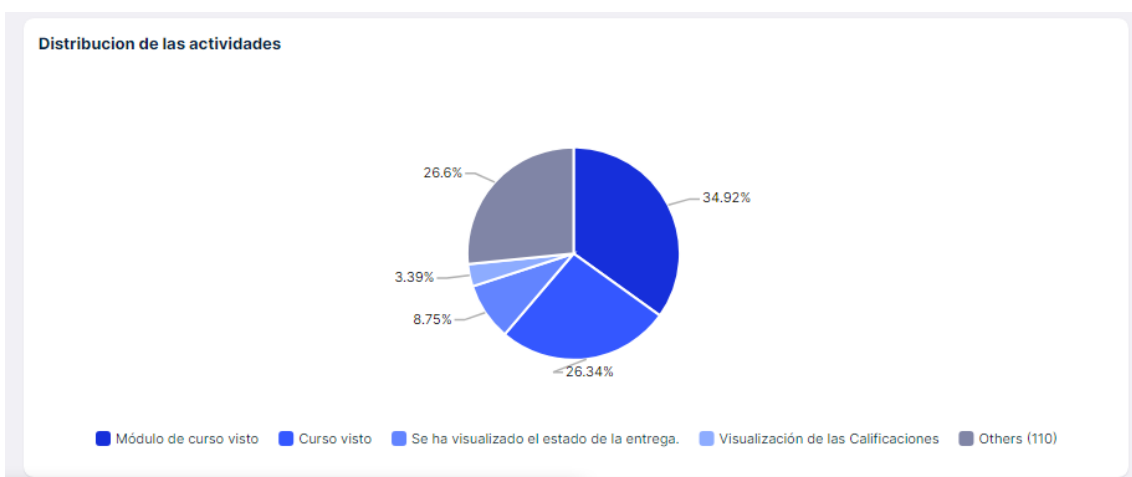


Figura 4.22: Distribución de las actividades y eventos del sistema

4.1. Fuente de datos de Moodle

de accesos correspondiente a la primera semana de marzo. Nótese que se han intentado replicar los datos de la manera más fiel posible a la realidad, donde es de esperar que la mayor parte de las entregas se realizan en los últimos días de plazo. La Distribución de las actividades y eventos que ocurren en el sistema es una métrica a analizar bastante importante, siendo en este caso las más representativas las correspondientes a los accesos de lectura y visualización de los cursos, de entrega y de calificaciones. Números que a primera vista parecen lógicos.

Análisis de Entregas

La siguiente pestaña implementada es la de Análisis de Entregas generales (véase la Figura 4.23), donde se dispondrá los filtros comentados anteriormente, la lista de KPIs y varios gráficos de utilidad para seguir ahondando en la explotación de los datos enfocados a las entregas tanto si se han realizado como si no y la distribución temporal de estas respecto a los plazos de entrega.

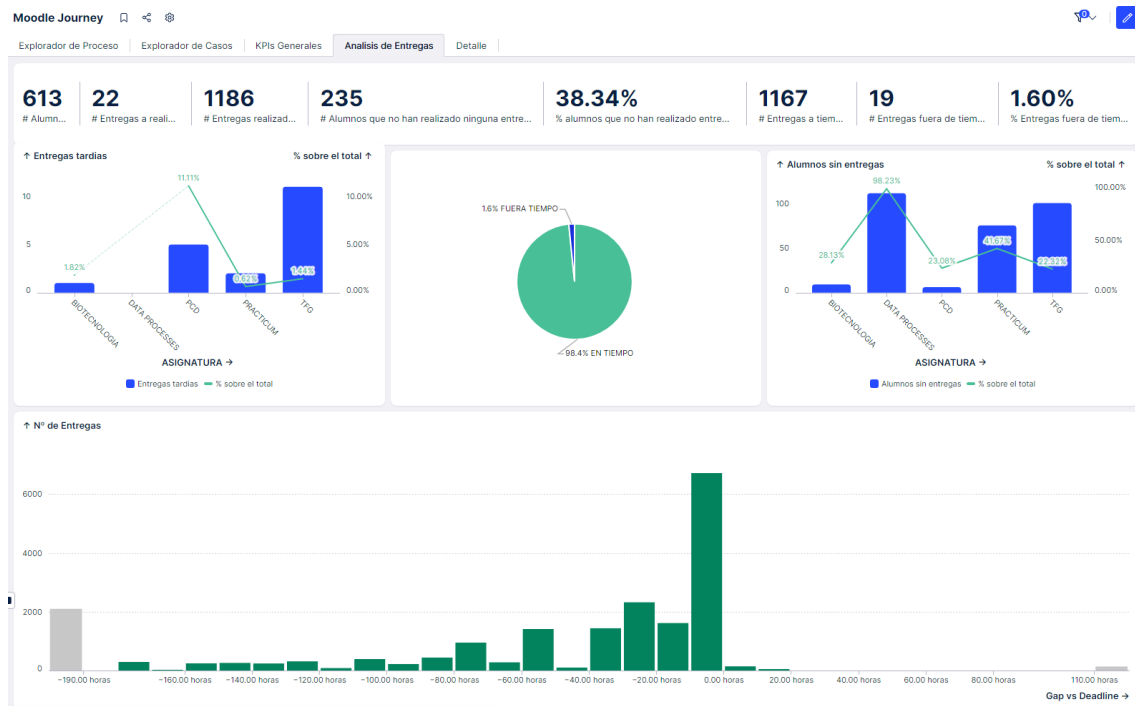


Figura 4.23: Pestaña de Análisis de entregas

Los gráficos que se han desarrollado son:

- Gráfico de entregas tardías por asignatura (Figura 4.24).
- Gráfico de entregas a tiempo o no sobre el total (Figura 4.25).
- Gráfico de alumnos sin entregas (Figura 4.26).
- Histograma en horas sobre la fecha de entrega (Figura 4.27).

Capítulo 4. Casos de Estudio

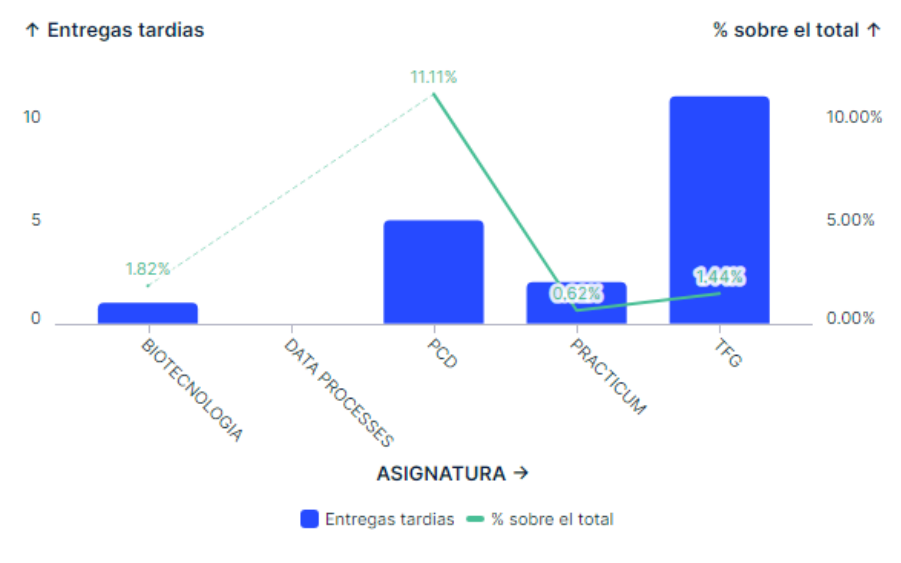


Figura 4.24: Entregas tardías por asignatura

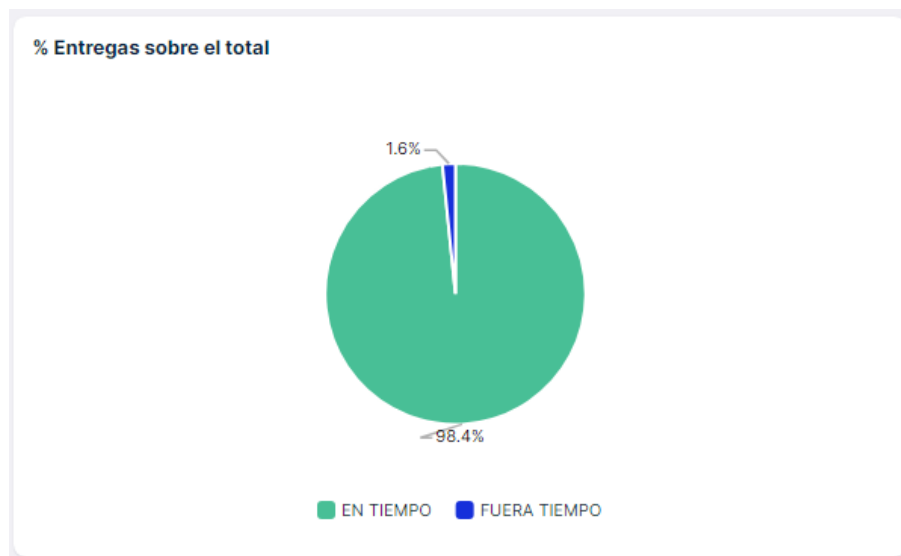


Figura 4.25: Entregas a tiempo

El Gráfico de entregas tardías por asignatura queda en dos ejes, el número de entregas que se han enviado en determinada asignatura fuera de tiempo y de este número cuanto representan sobre el total de entregas.

El Gráfico de tarta de entregas sobre el total representa el porcentaje de entregas tardías y a tiempo sobre el total.

El Gráfico de alumnos sin entregas representa la distribución del número de alumnos por asignatura, que no han realizado entregas por asignatura, así como el porcentaje de alumnos sobre los matriculados sin entregas.

Finalmente en la parte inferior se tiene una distribución en un histograma, cuya dimensión sería el número de entregas sobre un eje temporal, es decir la dife-

4.1. Fuente de datos de Moodle

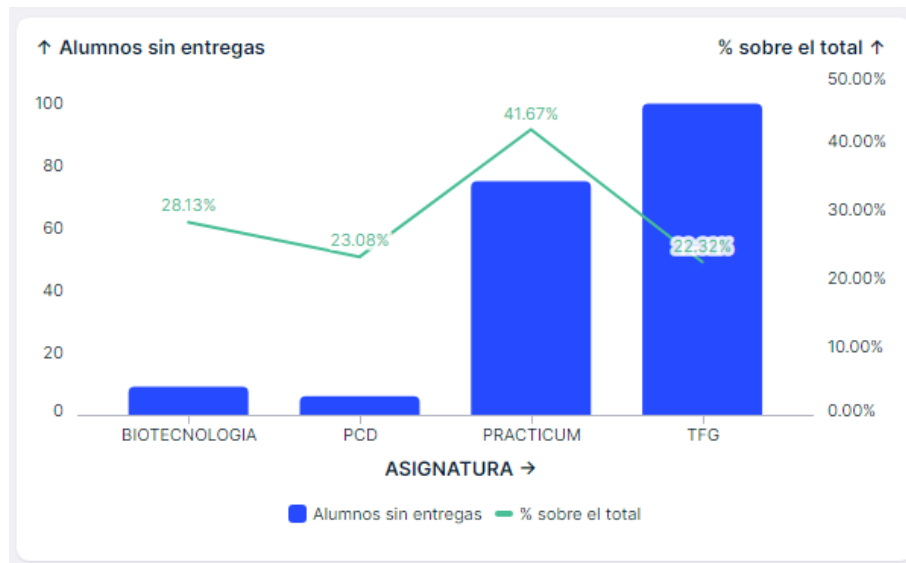


Figura 4.26: Alumnos sin entregas

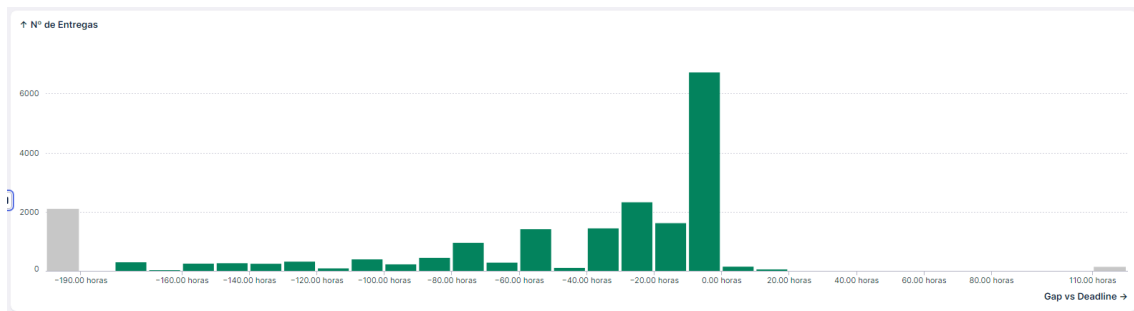


Figura 4.27: Histograma en horas sobre la fecha de entrega

rencia de tiempo de dicha entrega respecto la fecha límite, representando en un eje positivo y negativo respecto a la fecha de vencimiento, siendo aquellas con un valor negativo las entregas realizadas antes y el número de horas de entrega previas y con un valor positivo aquellas que fueron entregadas posteriormente a la fecha de entrega.

Detalle

La última pestaña implementada es la de tabla detalle, donde se puede revisar en detalle los casos y actividades con los filtros necesarios y exportar si así se quiere usar para otros propósitos (véase la Figura 4.28).

4.1.5. Oportunidades teóricas de mejora

Tomando como asunción y premisa que gran parte de este caso de uso y este proceso está realizado con datos dummy o ficticios, se van a tener en cuenta escenarios teóricos de mejora que puedan aplicarse a esta prueba de concepto, pero que sean realistas, es decir, que desde cualquier Universidad que lea este

Capítulo 4. Casos de Estudio

USUARIO	TIMESTAMP	ASIGNATURA	CONTEXTO_EVENTO	ENTREGA...	Horas respecto a la fecha de en...	FECHA ENTREGA
05db63fb81f3acf7a6915169f70dabcf	2024-03-02 17:52	TFG	Tarea: S2 - Plan de trabajo.	-1807	-30.12 horas	2024-03-03 23:59
05db63fb81f3acf7a6915169f70dabcf	2024-03-02 17:52	TFG	Tarea: S2 - Plan de trabajo.	-1807	-30.12 horas	2024-03-03 23:59
05db63fb81f3acf7a6915169f70dabcf	2024-03-02 17:52	TFG	Tarea: S2 - Plan de trabajo.	-1807	-30.12 horas	2024-03-03 23:59
05db63fb81f3acf7a6915169f70dabcf	2024-03-02 17:52	TFG	Tarea: S2 - Plan de trabajo.	-1807	-30.12 horas	2024-03-03 23:59
05db63fb81f3acf7a6915169f70dabcf	2024-03-04 15:23	TFG	Tarea: S2 - Plan de trabajo.	-1807	-30.12 horas	2024-03-03 23:59
05db63fb81f3acf7a6915169f70dabcf	2024-03-04 15:23	TFG	Tarea: S2 - Plan de trabajo.	-1807	-30.12 horas	2024-03-03 23:59
07220c061e18c2ca7d4769575c47a20c	2024-02-22 14:04	TFG	Tarea: S2 - Plan de trabajo.	-653	-10.88 horas	2024-03-03 23:59
07220c061e18c2ca7d4769575c47a20c	2024-02-23 13:29	TFG	Tarea: S2 - Plan de trabajo.	-653	-10.88 horas	2024-03-03 23:59
07220c061e18c2ca7d4769575c47a20c	2024-02-23 13:29	TFG	Tarea: S2 - Plan de trabajo.	-653	-10.88 horas	2024-03-03 23:59
07220c061e18c2ca7d4769575c47a20c	2024-02-23 13:46	TFG	Tarea: S2 - Plan de trabajo.	-653	-10.88 horas	2024-03-03 23:59
07220c061e18c2ca7d4769575c47a20c	2024-02-23 13:46	TFG	Tarea: S2 - Plan de trabajo.	-653	-10.88 horas	2024-03-03 23:59
07220c061e18c2ca7d4769575c47a20c	2024-02-23 13:46	TFG	Tarea: S2 - Plan de trabajo.	-653	-10.88 horas	2024-03-03 23:59
07220c061e18c2ca7d4769575c47a20c	2024-02-27 10:33	TFG	Tarea: S2 - Plan de trabajo.	-653	-10.88 horas	2024-03-03 23:59
07220c061e18c2ca7d4769575c47a20c	2024-02-28 09:27	TFG	Tarea: S2 - Plan de trabajo.	-653	-10.88 horas	2024-03-03 23:59
07220c061e18c2ca7d4769575c47a20c	2024-02-28 09:27	TFG	Tarea: S2 - Plan de trabajo.	-653	-10.88 horas	2024-03-03 23:59
07220c061e18c2ca7d4769575c47a20c	2024-03-03 12:46	TFG	Tarea: S2 - Plan de trabajo.	-653	-10.88 horas	2024-03-03 23:59
07220c061e18c2ca7d4769575c47a20c	2024-03-03 12:53	TFG	Tarea: S2 - Plan de trabajo.	-653	-10.88 horas	2024-03-03 23:59
07220c061e18c2ca7d4769575c47a20c	2024-03-03 12:53	TFG	Tarea: S2 - Plan de trabajo.	-653	-10.88 horas	2024-03-03 23:59
07220c061e18c2ca7d4769575c47a20c	2024-03-03 13:06	TFG	Tarea: S2 - Plan de trabajo.	-653	-10.88 horas	2024-03-03 23:59
07220c061e18c2ca7d4769575c47a20c	2024-03-03 13:06	TFG	Tarea: S2 - Plan de trabajo.	-653	-10.88 horas	2024-03-03 23:59

Figura 4.28: Tabla detalle de los alumnos

trabajo y este interesado en este caso de uso, pueda aplicar estas oportunidades de mejora en su operativa y procesos.

Se ha identificado 3 oportunidades de mejora para este caso de uso.

Oportunidad I: Gestión proactiva del alumnado

La implementación de una visualización, de un reporte, donde tener centralizado todas las dimensiones del alumno, número de accesos que han realizado, última conexión, cuándo han accedido, entregas, etc. y prácticamente cualquier tipo de dimensión permite un gran seguimiento por parte de los docentes de su curso. Tener esta implementación actualizada y dividida por asignatura y profesor, actualizada a diario o de manera semanal, permitiría a estos un seguimiento mucho más exhaustivo de lo que tienen actualmente y lo que puede ofrecer Moodle a nivel de estadísticas.

Oportunidad II: Gestión de entregas y copias.

Al igual que la oportunidad anterior, se puede hacer un foco aquí en las entregas, viendo si ciertas desviaciones en las entregas se deben a complejidad, tiempo suficiente que se ha facilitado o ver cómo corregir los plazos. Algo que no se ha comentado y que puede ser interesante es ver cómo interactúan los alumnos con Moodle en fechas concretas. Por ejemplo, durante un examen se ve que un alumno está presente en un examen que no permite la consulta de documentación de ningún tipo, pero accede al Moodle repetidas veces durante su participación en la entrega o examen. Así podrían aplicarse ciertas reglas o alarmas para identificar que dicho alumno está cometiendo un comportamiento de copia y ser sancionado o amonestado.

Oportunidad III: Estandarización de métodos de trabajo en Moodle

El seguimiento de diferentes profesores, alumnos y asignaturas permite realizar una comparativa de rendimiento o *benchmark* entre métricas de distintas asignaturas.

Aunque existen asignaturas de diferentes temas o incluso ramas, podrían extraerse ciertas buenas prácticas en las asignaturas o metodologías que tengan mejores valores en las métricas y tratar de aplicarlas en aquellas que tengan peores números. Esta estandarización y definición de buenas prácticas permitiría agilizar el trabajo de los docentes, desde la planificación de la asignatura hasta la elaboración de determinadas actividades como laboratorios o entregas.

4.2. Fuente de datos de Matriculación

Una vez explorada la potencialidad de Celonis y haber probado sus funcionalidades en un proceso acotado, se va a implementar un caso de uso puramente universitario donde el impacto buscado es más representativo para este trabajo y haciendo algo más de foco en el punto de acción de mejora sobre el proceso desde la herramienta.

El segundo caso es el caso de uso de Matriculación, proceso donde un estudiante de la universidad, solicita cita previa al sistema, se le concede y comienza la matriculación, selecciona asignaturas, horarios, grupos, avanza hasta el pago de la matrícula para finalmente formalizar la matriculación.

4.2.1. Generación del log de eventos

Para este caso de uso, se generan desde cero los ficheros de datos, teniendo en cuenta la metodología, las convenciones acordadas en todo momento, como la flexibilidad, escalabilidad y sencillez del modelo de datos. Obviamente, se necesita una tabla de datos y una tabla de actividades. Esta tabla de datos, entidad central del modelo de datos, representará los alumnos que realizan la matriculación en el sistema. La tabla de Actividades representará las actividades que los alumnos realizan en el sistema, lo que requerirá de una buena definición de estas para representar de manera correcta el proceso. Como alumnos con experiencia en matriculación, al menos en el caso de la propia universidad, se han diseñado a conciencia las actividades a representar.

En primer lugar, las actividades ideales o que aparecerían en una matriculación completada con éxito.

- Solicitud de cita previa.
- Asignación de cita previa.
- Inicio de proceso de matrícula.
- Selección de asignaturas.
- Elección de grupos y horarios.
- Confirmación de selección.

Capítulo 4. Casos de Estudio

- Desglose y pago de la matrícula.
- Confirmación de pago.
- Generación de documentos.
- Envío de documentos al alumno.
- Matriculación no completada
- Fin de Matriculación

Aquellas que se desvían de este camino ideal y que se pueden definir como actividades no deseadas.

- Pago no confirmado
- Pago cancelado
- Fallo en la generación y envío de documentos de matrícula.
- Interrupción inesperada del proceso por parte del sistema.
- Tiempo de matriculación agotado
- Contacto con Soporte por queja/duda
- Matriculación no completada
- Fin de Matriculación

Con estas actividades definidas, se tendría que generar una huella temporal aleatoria, pero con sentido, para cada una de estas actividades. Para la tabla de Casos, se han definido como universo de datos, 70 estudiantes, cantidad suficiente para una muestra de datos representativa, tanto a nivel de alumnos como de actividades. Para cada uno de estos alumnos, se necesitan algunos campos que permitan segmentar los datos y enriquecer los datos, como grado, localización, tipo de matrícula o asignaturas. De manera adicional a estas dos tablas, se añade una tabla de pagos, con la cantidad pagada y el tipo de pago asociado a cada alumno.

4.2.2. Ficheros de datos y estructura

Los datos fuente o ficheros de datos para este caso de Matriculación se han construido a mano, basándose en las especificaciones del apartado anterior y tratando de ajustar la realidad del proceso al proceso que se tiene en la propia escuela. Los datos se componen de un conjunto de ficheros en formato Excel para los alumnos, actividades, registro de eventos y pagos.

En primer lugar, se dispone de una tabla de casos, tabla que contendrá la información de todos los estudiantes y campos adicionales para enriquecer el análisis con la estructura de la Tabla 4.6.

En este sentido, se ha creado una tabla de definición de actividades que se cruzará con la tabla de actividades de manera dinámica, con estas como parámetros, en caso de querer redefinir los nombres de actividades o el contenido de

4.2. Fuente de datos de Matriculación

Tabla 4.6: Tabla de Estudiantes (Casos)

Campo	Tipo de Campo	Descripción	Ejemplo de valor
Alumno	Texto	Identificador de Alumnos	EST32
Grado	Texto	Grado que está estudiando el alumno	GMI
Localización	Texto	Localización del alumno	Torrelodones
Fecha Última Matrícula	Fecha	Última fecha de matriculación del alumno	DD/MM/YYYY HH:MI:SS
Créditos Última Matrícula	Entero	N.º. de créditos de asignaturas del alumno	10
Asignaturas	Entero	N.º. de asignaturas matriculadas	5
Tipo de Matrícula	Texto	Categorización de la matriculación basándonos en los créditos matriculados	Parcial
Estado Matrícula	Texto	Estado actual de la matrícula	Completada
Sexo	Texto	Sexo del alumno	Hombre

estas, poder cambiar únicamente esta tabla y no tener que ir registro a registro de actividades cambiando el nombre completo (véase la Tabla 4.7).

Tabla 4.7: Tabla de definición de actividades

Campo	Tipo de Campo	Descripción	Ejemplo de valor
ID Actividad	Texto	Identificador de Actividad	A
Actividad	Texto	Nombre de la actividad	Pago cancelado
Sorting	Entero	Orden de la actividad	2

Esto se consigue metiendo el valor de nombre de la actividad donde estaría la actividad de la tabla de actividades y cruzándola en las transformaciones para traernos dicho nombre (véase la Tabla 4.8).

Tabla 4.8: Tabla de Actividades

Campo	Tipo de Campo	Descripción	Ejemplo de valor
ID Actividad	Entero	Identificador de Actividad	2
Alumno	Texto	Identificador del alumno	EST22
Actividad	Texto	Identificador de Actividad	C
Timestamp	Fecha	Fecha del evento	DD/MM/YYYY HH:MI:SS
Periodo	Texto	Periodo de matriculación	Extraordinario
Dispositivo	Texto	Tipo de acceso	Móvil

También se ha generado una tabla de pagos por alumno, permitiendo centrarse en la parte financiera del proceso de matriculación (véase la Tabla 4.9).

Tabla 4.9: Tabla de Pagos

Campo	Tipo de Campo	Descripción	Ejemplo de valor
ID Pago	Entero	Identificador de Pago	65
ID Caso	Texto	Identificador del alumno	EST22
Timestamp	Fecha	Fecha del pago	DD/MM/YYYY HH:MI:SS
Cantidad	Float	Importe de pago de matrícula	349,25
Método	Texto	Método de pago	Paypal

4.2.3. Transformaciones

Las transformaciones de este caso de uso siguen la estructura metodológica de Celonis y las que se han seguido en el anterior escenario. Por otra parte, se busca la simplicidad en el modelo de datos, con una estructura mínima de tablas y extendiendo a nivel de tablas y campos cuando sea necesario. Se va a emplear un único Trabajo de Datos (o *Data Job*) para todas las transformaciones. Dentro de este trabajo de datos se va a crear una estructura ordenada y flexible de transformaciones de código (véase figura 4.29).

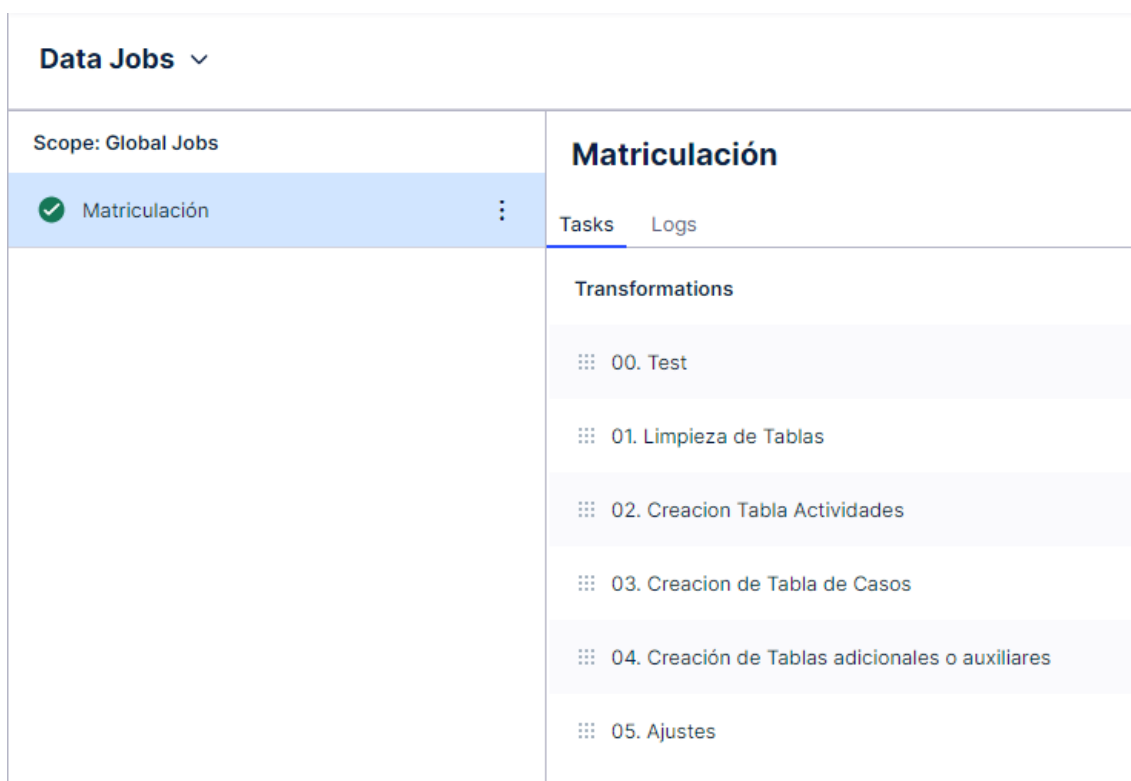


Figura 4.29: Test

Se divide en 6 Transformaciones principales:

- 00. Test
- 01. Limpieza de Tablas

4.2. Fuente de datos de Matriculación

- 02. Creación de Tabla de Actividades
- 03. Creación de Tabla de Casos
- 04. Creación de tablas adicionales/auxiliares
- 05. Ajustes

La transformación 00. Test sirve como banco de pruebas y de trabajo donde ejecutar las consultas necesarias para verificar que el contenido de las tablas subidas desde los ficheros son correctas.

La transformación 01. Limpieza de Tablas, consiste en los trabajos de normalización de los ficheros fuente en caso de ser necesario, al contrario que en caso de Moodle, los datos vienen bastante bien estructurados y formateados, por lo que aquí entrarán algunos ajustes previos si fuesen necesarios.

La transformación 02. Creación de Tabla de Actividades, se realizará la composición de la tabla de eventos de Matriculación. A diferencia del caso anterior, aquí como Actividad se tiene un valor de texto de la A a la R que se corresponderá con un valor del fichero de descripción de actividades. Este fichero contiene el nombre y el orden de las actividades, permitiendo así dinamizar y flexibilizar los datos de evento y actividades de los ficheros fuente, en caso de tener que realizar alguna modificación. Para ello se hará una unión de la tabla de Actividades con la de descripción de actividades por este valor, trayendo el nombre y el orden.

La transformación 03. Creación de Tabla de Casos se cargará directamente desde la tabla de Casos de Alumnos que se matriculan, conteniendo esta tabla cada Alumno de manera única con los campos descriptivos para segmentar la información de Casos y de actividades.

En la transformación 04. Creación de tablas adicionales y auxiliares, se crean la tabla de Pagos, información proveniente de los ficheros fuente que contiene los pagos de cada alumno para la matrícula, así como su cantidad y método de pago. La transformación 05. Ajustes contendrá los ajustes necesarios y actualizaciones de campos en caso de necesitarlos.

El código detallado de las transformaciones se adjunta en el Anexo para cada una de ellas.

Capítulo 4. Casos de Estudio

4.2.4. Modelo de Datos

Una vez aplicadas las transformaciones, estarían las tablas preparadas para relacionarlas y conformar el modelo de datos previo a las visualizaciones. Los datos en este punto quedarían según las Tablas 4.10, 4.11 y 4.12.

Tabla 4.10: Tabla de Casos

Campo	Tipo de Campo	Descripción
CASE ID	Varchar	Identificador de Alumno
GRADO	Varchar	Grado que cursa el alumno
LOCALIZACION	Varchar	Localización donde reside el alumno
FECHA ULT MATRICULA	Fecha	Fecha de la última matriculación del alumno
CREDITOS ULT MATRICULA	Integer	Número de créditos matriculados del alumno
ASIGNATURAS	Integer	Número de asignaturas matriculadas del alumno
TIPO MATRICULA	Varchar	Matrícula Parcial o Completa del alumno
ESTADO MATRICULA	Varchar	Si se ha completado con éxito la matrícula del año actual

Tabla 4.11: Tabla de Actividades

Campo	Tipo de Campo	Descripción
ACTIVITY ID	Integer	Identificador de la actividad
CASE ID	Varchar	Identificador de Alumno
ACTIVITY NAME	Varchar	Nombre de la actividad que realiza el alumno
TIMESTAMP	Timestamp	Huella temporal de la actividad
SORTING	Integer	Orden entre actividades
PERIODO	Varchar	Periodo de matriculación
DISPOSITIVO	Varchar	Desde donde realiza la actividad el alumno

Tabla 4.12: Tabla de Pagos

Campo	Tipo de Campo	Descripción
PAGO ID	Integer	Identificador de Pago
CASEID	Varchar	Identificador de Alumno
FECHA PAGO	Timestamp	Huella temporal del pago
CANTIDAD	Float	Importe de pago de matrícula
METODO	Varchar	Metodo de pago

El Modelo de Datos a nivel relacional se desglosa según la Figura 4.30

Una vez cargado el modelo se dispondrían de las tablas relacionadas y normalizadas con la volumetría de la Figura 4.31

4.2. Fuente de datos de Matriculación

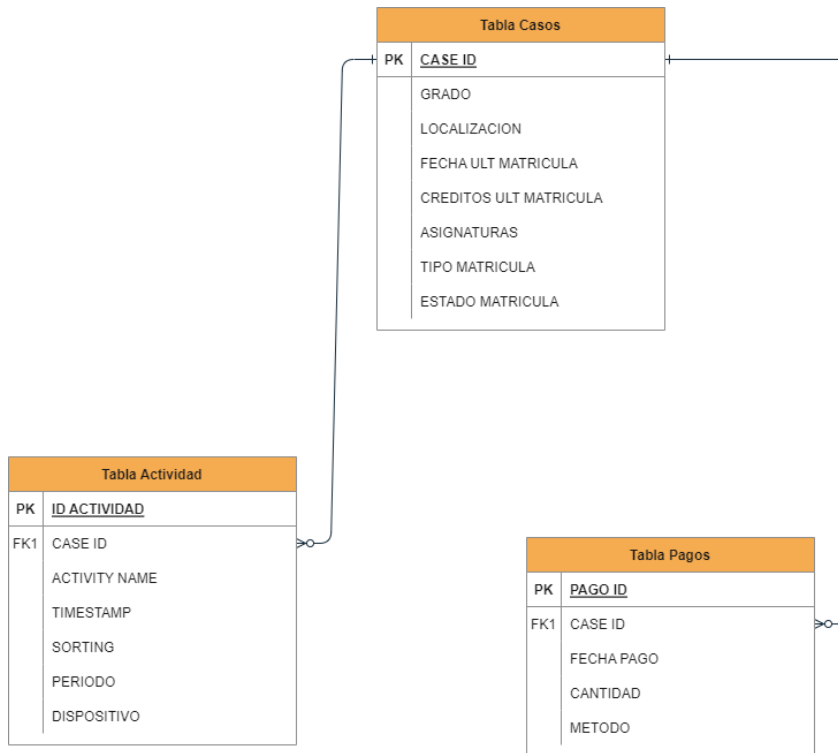


Figura 4.30: Modelo de Datos relacional del caso de matriculación

Loaded Data Model

Table name	Load type	Progress
MATRICULACION_ACTIVIDADES	FULL	✓
MATRICULACION_CASOS	FULL	✓
MATRICULACION_PAGOS	FULL	✓

Table name	Loaded rows
MATRICULACION_ACTIVIDADES	918
MATRICULACION_CASOS	70
MATRICULACION_PAGOS	70

Figura 4.31: Volumetría de los datos del modelo

4.2.5. Implementación y visualización

Para este caso de uso se ha planteado una estructura de elementos y carpetas similar al anterior, aunque añadiendo elementos adicionales como los flujos de acción (véase Figura 4.32 y 4.33).

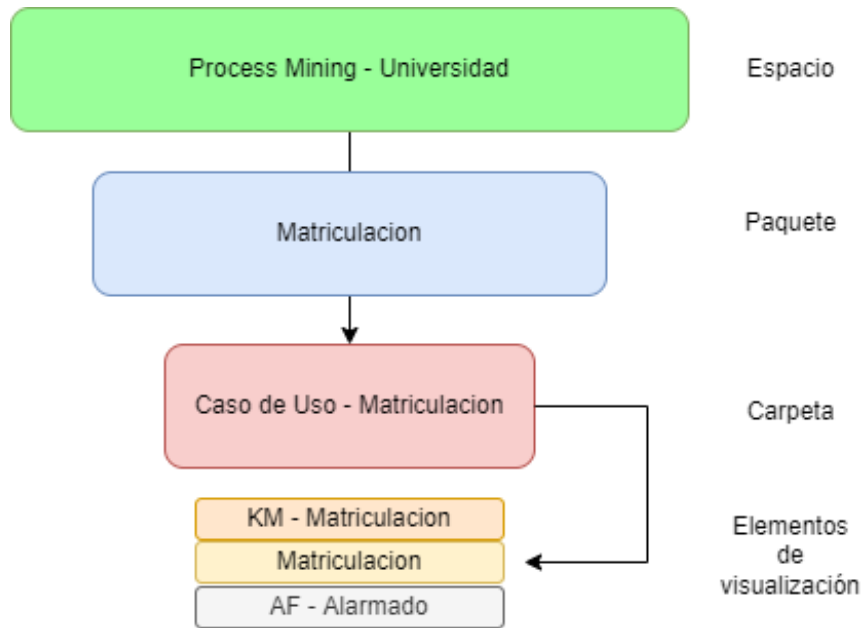


Figura 4.32: Estructura de elementos

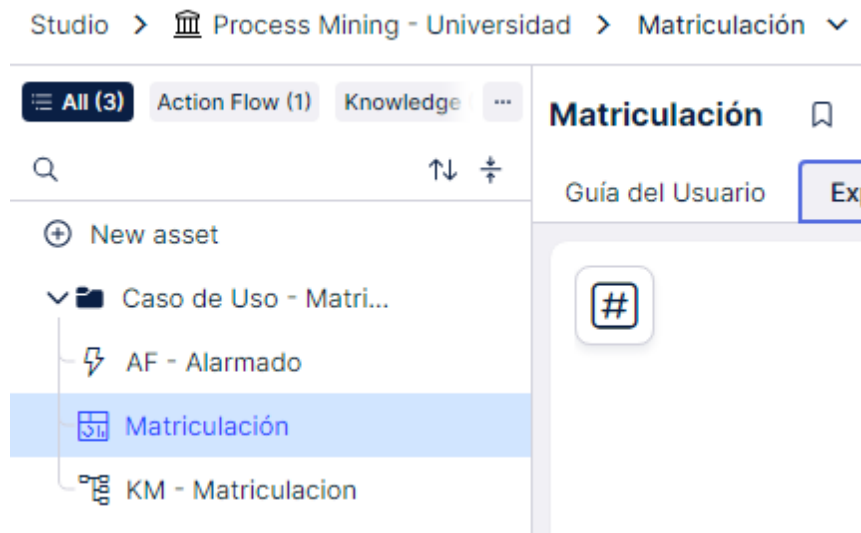


Figura 4.33: Visualización de la organización de elementos en Celonis

Modelo de Conocimiento

En primer lugar, y como buena práctica en términos de flexibilidad y escalabilidad de la implementación, se ha creado el modelo de conocimiento, almacén de todos los elementos de código utilizados en las visualizaciones. Aquí se incluyen Registros, KPIs, Filtros, Atributos, Variables, entre otros. Esto permite una mejor gestión del cambio, haciendo que las posibles modificaciones futuras puedan agilizarse para el ejecutor del cambio.¹

¹Desglose de las métricas y sus fórmulas en el apartado correspondiente del Anexo 1

4.2. Fuente de datos de Matriculación

Records		KPIs		
Name	ID	PQL Formula		
% Alumnos	alumnos	count("MATRICULACION_PAGOS"...		
Avg Events Per Case Matriculaci...	AVG_EVENTS_PER_CASE_MATRICU...	AVG(PU_COUNT("MATRICULACION_...		
Count Table Matriculacion Activi...	COUNT_TABLE__MATRICULACION_A...	COUNT_TABLE("MATRICULACION_A...		
Count Table Matriculacion Casos	COUNT_TABLE__MATRICULACION_...	COUNT_TABLE("MATRICULACION_C...		
Count Table Matriculacion Pagos	COUNT_TABLE__MATRICULACION_P...	COUNT_TABLE("MATRICULACION_P...		
Filtered Count	FILTERED_COUNT	COUNT(CASE WHEN {p1} THEN 0 E...		
Importe medio por alumno	importe_medio_por_alumno	AVG("MATRICULACION_PAGOS""CA...		
N.º de Alumnos con Incidencia/...	n_de_alumnos_con_incidencia_duda	COUNT(CASE WHEN "MATRICULAC...		
N.º Alumnos con matricula parcial	n_alumnos_con_matricula_parcial	COUNT (CASE WHEN "MATRICULA...		
N.º de Alumnos que comienzan l...	n_de_alumnos_que_comienzan_la_m...	COUNT("MATRICULACION_CASOS"...		
N.º de veces que no se complet...	n_de_veces_que_no_se_completa_la...	COUNT(distinct CASE WHEN "MATR...		
Nº de Procesos Iniciados	n_de_procesos_iniciados	COUNT(CASE WHEN "MATRICULAC...		
Number Of Process Variants Mat...	NUMBER_OF_PROCESS_VARIANTS_...	COUNT(DISTINCT SHORTENED(VAR...		
Process Variants Matriculacion A...	PROCESS_VARIANTS__MATRICULAC...	SHORTENED(VARIANT("MATRICULA...		
Ratio	RATIO	AVG(CASE WHEN {p1} THEN 1 ELSE ...		

Figura 4.34: Modelo de conocimiento del caso de uso de Matriculación

Al igual que en el caso anterior, se han ido añadiendo de manera incremental, añadiéndose según se iban desarrollando las vistas.

Panel de filtrado

Se ha desarrollado un panel de filtros (véase la Figura 4.35) que aplica al análisis completo y que permite llevar estos filtros de unas hojas a otras y encauzar la información a aquellos objetivos buscados. Este panel se desglosa en el lateral de la página y se aplica sobre todo el análisis.

Guía del Usuario

En esta primera pestaña, se facilita al usuario la navegación de manera rápida por todas las pestañas, añadiendo una breve descripción de cada una y un botón para llegar hasta a ella (véase la Figura 4.36).

Explorador de Proceso

Con los datos formateados perfectamente para su visualización, se puede visualizar el Proceso y sus diferentes variantes, siendo el de la Figura 4.37 el camino más común o el camino ideal.

La pestaña de explorador de proceso contiene la representación del proceso tal cual se ha indicado en la Figura 4.37, y aquí se pueden añadir conexiones y

Capítulo 4. Casos de Estudio

Panel de Filtros

Grado

Estado de Matricula

Tipo de Matricula

Periodo de Matricula

Método de Pago

Filtros temporales

Ventana Temporal

Detailed description: This is a vertical filter panel. It starts with a header 'Panel de Filtros'. Below it are six dropdown menus: 'Grado', 'Estado de Matricula', 'Tipo de Matricula', 'Periodo de Matricula', and 'Método de Pago'. Below these is a section titled 'Filtros temporales' which contains a 'Ventana Temporal' button with a calendar icon.

Figura 4.35: Panel de filtros del análisis

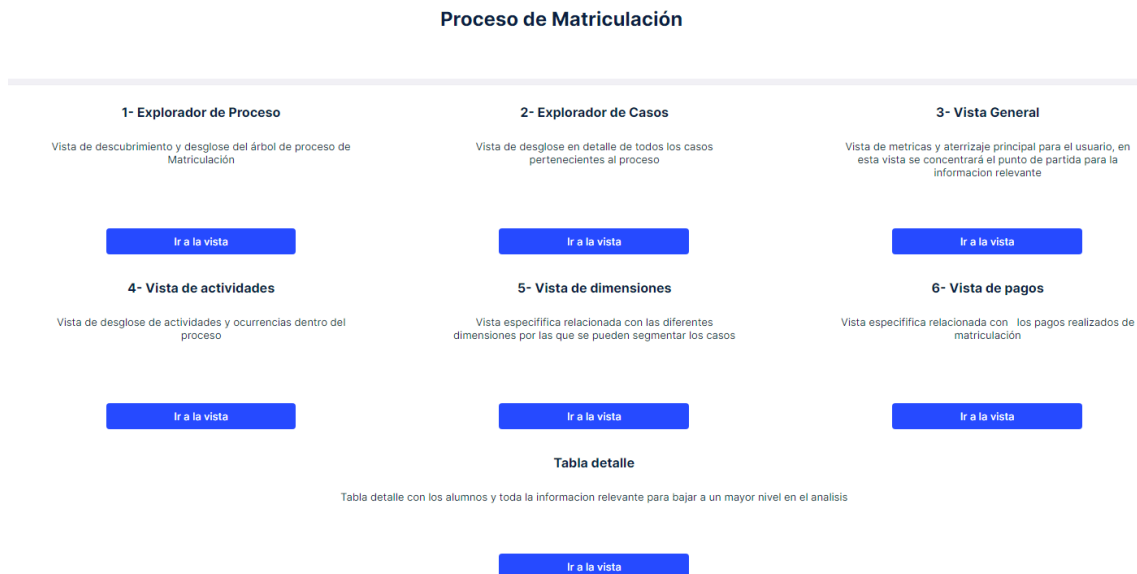


Figura 4.36: Guía de Usuario para el proceso de matriculación

actividades para ir acotando o extendiendo el proceso según las necesidades. En la figura 4.38 se puede apreciar como al añadir actividades aparecen caminos y desviaciones a priori no deseados.

Desplegando todas y cada una de las actividades y conexiones entre actividades del proceso se tiene el gráfico espagueti para el caso de matriculación.

4.2. Fuente de datos de Matriculación

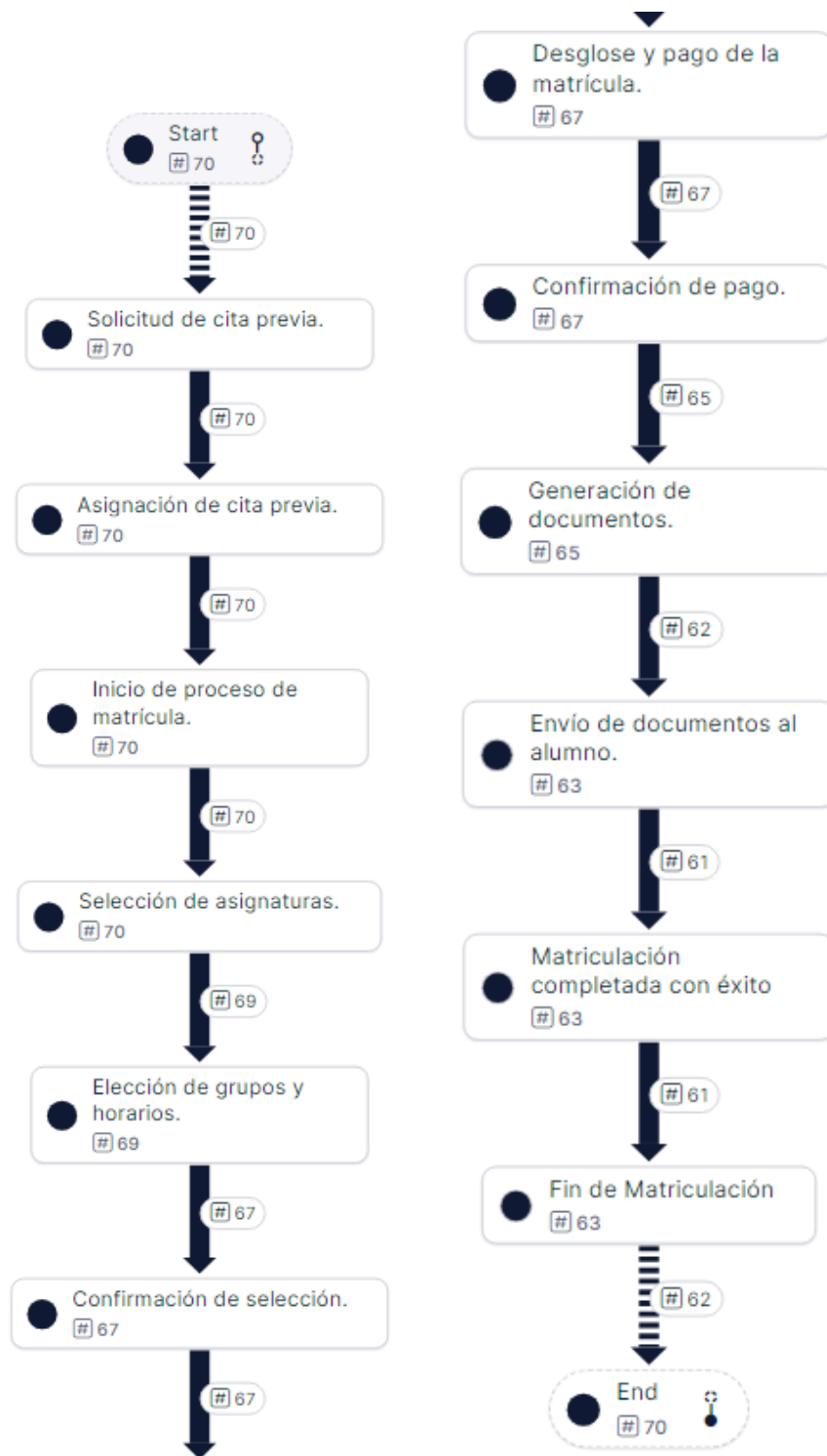


Figura 4.37: Explorador de proceso de Matriculación

Explorador de Casos

Esta pestaña incluirá el explorador de casos, aquí se podrá bajar a detalle la información con los filtros, actividades y caminos por los que vamos moviendo

Capítulo 4. Casos de Estudio

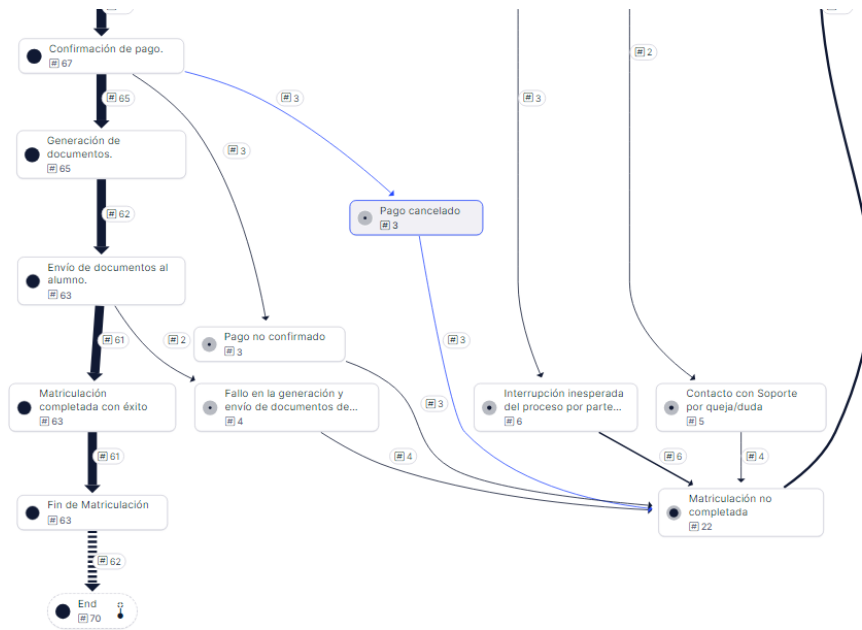


Figura 4.38: Explorador de proceso entendido

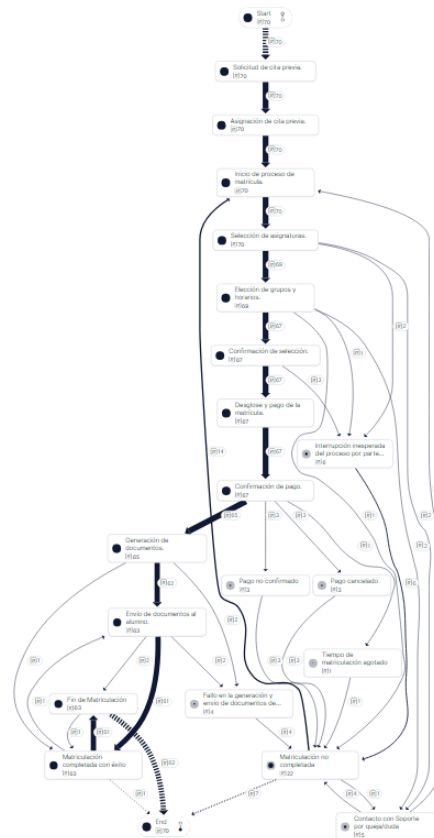


Figura 4.39: Diagrama espagueti para el proceso de matriculación

4.2. Fuente de datos de Matriculación

el análisis (véase la Figura 4.40).

Case Id	# of Activities	Throughput Time	First Activity	First Activity Timestamp	Last Activity	Last Activity Timestamp
EST47	12	14 d	Solicitud de cita pre...	2024-01-15 16:03:45	Fin de Matriculación	2024-01-29 10:30:35
EST53	18	13 d	Solicitud de cita pre...	2024-01-03 17:51:25	Fin de Matriculación	2024-01-16 18:23:56
EST59	22	13 d	Solicitud de cita pre...	2024-01-04 18:52:34	Fin de Matriculación	2024-01-17 12:44:31
EST32	12	12 d	Solicitud de cita pre...	2024-01-03 09:33:14	Fin de Matriculación	2024-01-15 19:03:03
EST43	12	12 d	Solicitud de cita pre...	2024-01-05 10:53:45	Fin de Matriculación	2024-01-17 15:43:33
EST40	12	12 d	Solicitud de cita pre...	2024-01-03 11:32:56	Fin de Matriculación	2024-01-15 12:56:04
EST63	16	12 d	Solicitud de cita pre...	2024-01-06 13:23:12	Fin de Matriculación	2024-01-18 13:11:45
EST62	17	12 d	Solicitud de cita pre...	2024-01-03 15:14:36	Fin de Matriculación	2024-01-15 14:11:21
EST68	6	12 d	Solicitud de cita pre...	2024-01-04 19:24:36	Matriculación no co...	2024-01-16 17:58:31
EST10	12	12 d	Solicitud de cita pre...	2024-01-04 22:43:22	Fin de Matriculación	2024-01-16 18:43:32
EST1	12	12 d	Solicitud de cita pre...	2024-01-03 20:44:27	Fin de Matriculación	2024-01-15 14:43:56
EST19	12	12 d	Solicitud de cita pre...	2024-01-03 20:11:04	Fin de Matriculación	2024-01-15 09:35:33
EST56	20	11 d	Solicitud de cita pre...	2024-01-05 09:43:21	Fin de Matriculación	2024-01-16 18:55:02
EST16	12	11 d	Solicitud de cita pre...	2024-01-04 13:12:16	Fin de Matriculación	2024-01-15 19:54:32
EST50	16	11 d	Solicitud de cita pre...	2024-01-04 14:31:50	Fin de Matriculación	2024-01-15 18:17:30
EST5	12	11 d	Solicitud de cita pre...	2024-01-04 17:22:09	Matriculación compl...	2024-01-15 20:32:03
EST41	12	11 d	Solicitud de cita pre...	2024-01-06 09:21:25	Fin de Matriculación	2024-01-17 12:31:07
EST65	11	11 d	Solicitud de cita pre...	2024-01-05 11:34:22	Matriculación no co...	2024-01-16 12:13:56
EST21	12	11 d	Solicitud de cita pre...	2024-01-16 15:18:49	Fin de Matriculación	2024-01-21 14:23:00
EST37	12	11 d	Solicitud de cita pre...	2024-01-04 17:24:14	Fin de Matriculación	2024-01-15 14:53:04
EST2	12	11 d	Solicitud de cita pre...	2024-01-04 20:59:55	Fin de Matriculación	2024-01-15 17:31:02
EST33	12	11 d	Solicitud de cita pre...	2024-01-06 18:22:47	Fin de Matriculación	2024-01-17 12:17:31
EST29	12	11 d	Solicitud de cita pre...	2024-01-04 20:22:33	Fin de Matriculación	2024-01-15 10:40:56

Case details: EST63

Search

Activities 16 Items

- Solicitud de cita previa. 2024-01-06 13:23:12 5 months ago
- Asignación de cita previa. 2024-01-06 13:51:42 5 months ago
- Inicio de proceso de matrícula. 2024-01-17 18:00:45 4 months ago
- Selección de asignaturas. 2024-01-17 18:02:02 4 months ago
- Contacto con Soporte por queja/duda 2024-01-17 18:20:52 4 months ago
- Matriculación no completada 2024-01-17 18:23:04 4 months ago
- Inicio de proceso de matrícula. 2024-01-18 12:01:43 4 months ago
- Selección de asignaturas. 2024-01-18 12:03:52 4 months ago
- Elección de grupos y horarios. 2024-01-18 12:10:19 4 months ago
- Confirmación de selección. 2024-01-18 12:17:04 4 months ago
- Desglose y pago de la matrícula. 2024-01-18 12:22:28 4 months ago
- Confirmación de pago. 2024-01-18 12:30:15 4 months ago
- Generación de documentos. 2024-01-18 12:37:02 4 months ago
- Envío de documentos al alumno. 2024-01-18 12:44:19 4 months ago
- Matriculación completada con éxito 2024-01-18 13:10:45 4 months ago
- Fin de Matriculación 2024-01-18 13:11:45 4 months ago

Figura 4.40: Explorador de casos para el proceso de matriculación

Capítulo 4. Casos de Estudio

Vista general

La pestaña de Vista general, se incluyen las métricas principales del proceso y los gráficos asociados a las principales dimensiones para la segmentación de la información (véase la Figura 4.41).



Figura 4.41: Vista completa pestaña general

En primer lugar, una tarjeta con los principales números (véase la Figura 4.42), y donde se han incluido las siguientes métricas:

- N.º de Alumnos que comienzan la matriculación.
- N.º de veces que no se completa la matriculación.
- N.º de alumnos con algún tipo de incidencia o duda.
- N.º de alumnos con matrícula parcial

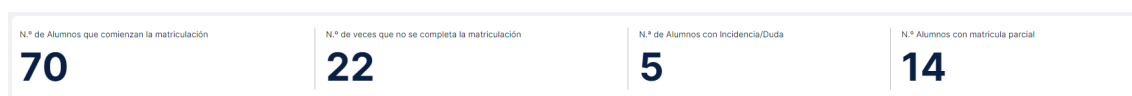


Figura 4.42: Métricas principales para la vista general del proceso

Los gráficos que se han desarrollado en esta pestaña son los siguientes:

- Distribución de los casos por fecha de matrícula (Figura 4.43).
- Distribución de los casos por tipo de matrícula (Figura 4.44).
- Distribución de los alumnos por periodos de matrícula (Figura 4.45).
- Distribución de los alumnos por estado de matrícula (Figura 4.46).
- Distribución de los alumnos por Sexo (Figura 4.47).

4.2. Fuente de datos de Matriculación

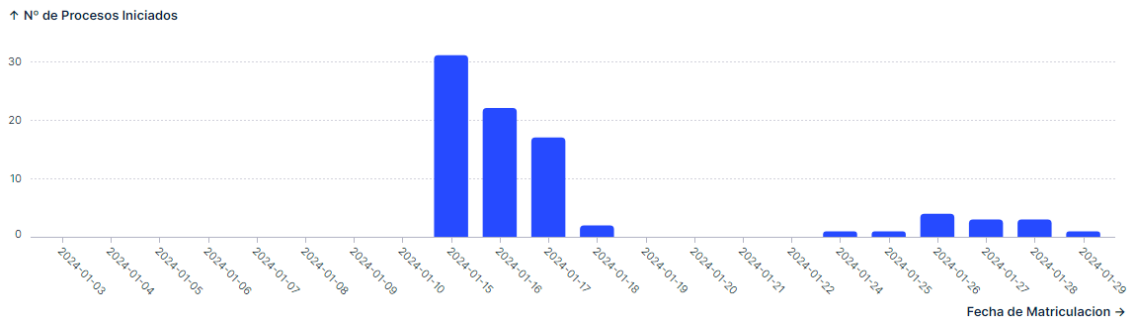


Figura 4.43: Distribución por tipo de matrícula.

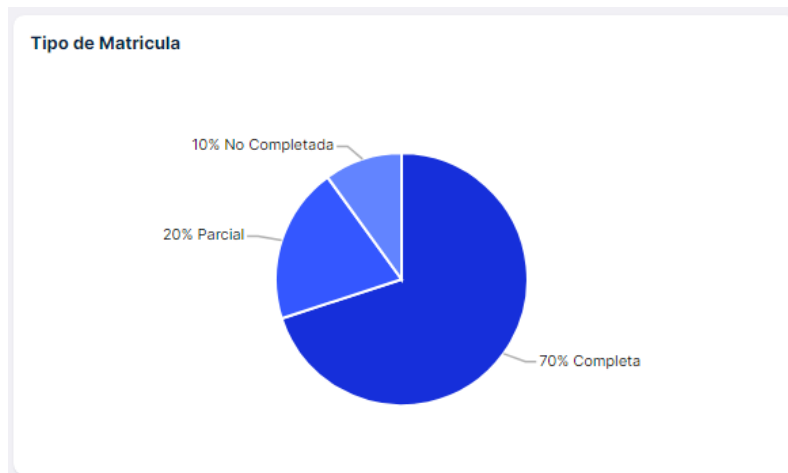


Figura 4.44: Distribución por tipo de matricula.

Periodos de matricula

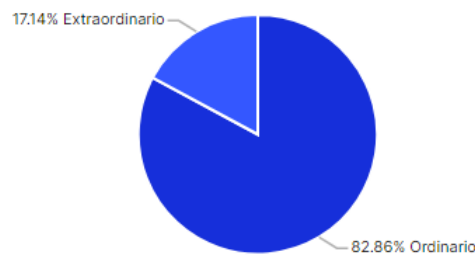


Figura 4.45: Distribución por periodo de matricula.

Vista de actividades

La pestaña de Vista de actividades incluirá un análisis específico general, donde se incluyen las métricas principales de las actividades del proceso y los gráficos asociados a las principales actividades y dimensiones (véase la Figura 4.48).

Capítulo 4. Casos de Estudio

Estado de Matricula



Figura 4.46: Distribución por estado de matricula.

Sexo de los matriculados

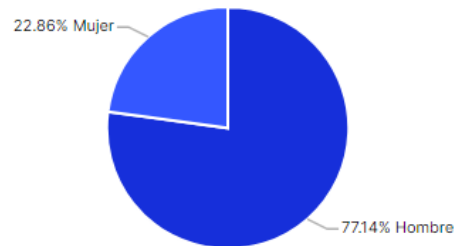


Figura 4.47: Distribución por sexo del alumno.

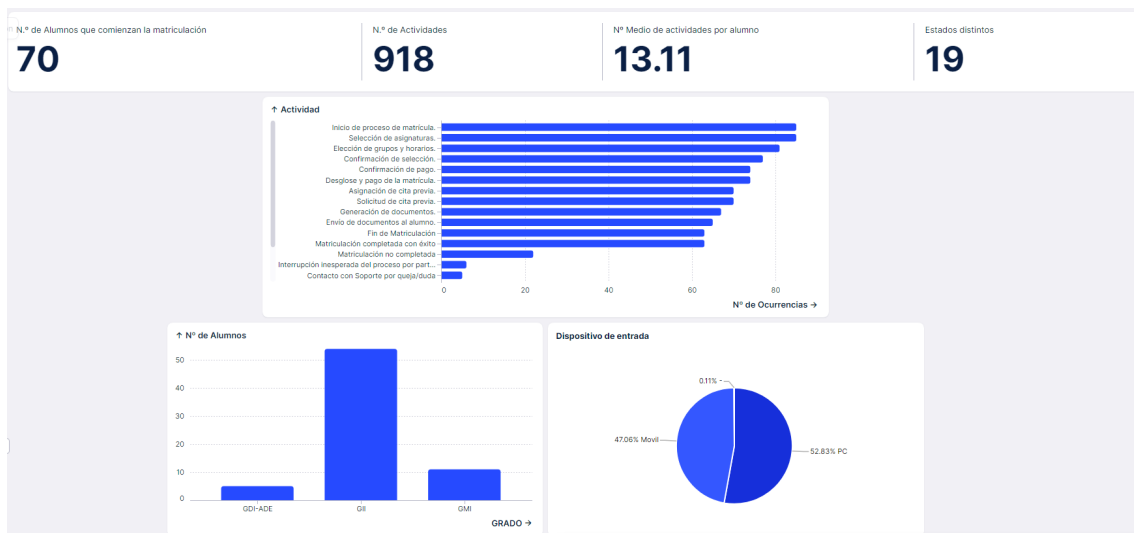


Figura 4.48: Vista completa pestaña de actividades

4.2. Fuente de datos de Matriculación

En primer lugar, una tarjeta con los principales números (véase la Figura 4.49), y donde se han incluido las siguientes métricas:

- N.º de Alumnos que comienzan la matriculación.
- N.º de actividades que aparecen en el proceso.
- N.º medio de actividades por alumno.
- Actividades distintas.



Figura 4.49: Explorador de casos para el proceso de matriculación

Los gráficos que se han desarrollado en esta pestaña son los siguientes:

- Distribución de ocurrencias de actividades (Figura 4.50).
- Distribución de los alumnos por grado (Figura 4.51).
- Distribución de las actividades por dispositivo de entrada (Figura 4.52).

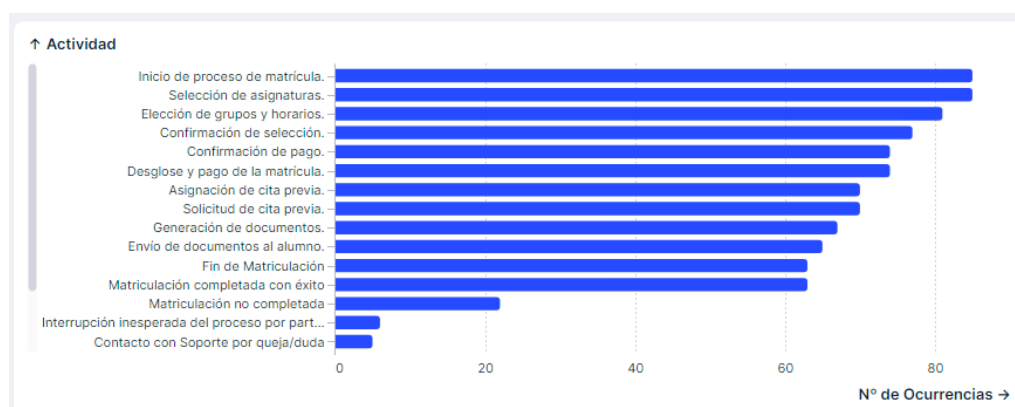


Figura 4.50: Distribución y ocurrencias de las actividades

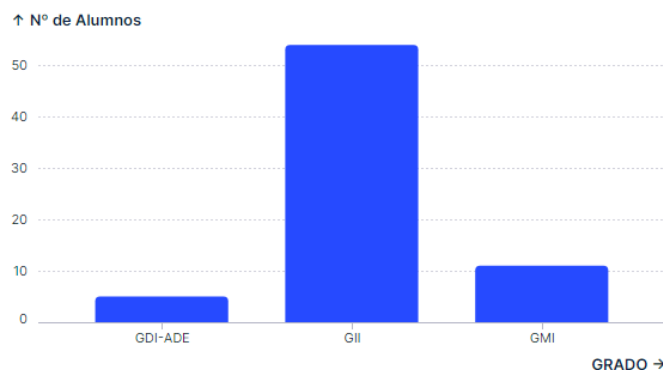


Figura 4.51: Distribución de los alumnos por grado

Capítulo 4. Casos de Estudio

Dispositivo de entrada

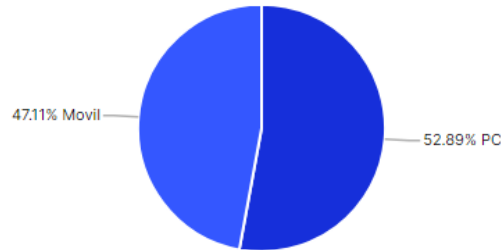


Figura 4.52: Distribución de las actividades por tipo de dispositivo

Vista de tiempos y localización

La pestaña de Vista de tiempos y localización incluye un análisis específico de los tiempos, viendo a grandes rasgos cuanto duran las actividades, donde se incluyen las métricas principales de los tiempos y localización del proceso y los gráficos asociados a las principales dimensiones de tiempo y localización (véase la Figura 4.53).



Figura 4.53: Vista general pestaña de pagos

En primer lugar, una tarjeta con los principales números (véase figura 4.53), y donde se han incluido las siguientes métricas:

- N.º de Alumnos que comienzan la matriculación.
- Tiempo medio del proceso desde cita previa.
- Tiempo medio del proceso de matriculación.
- Duración media de la actividad.

Capítulo 4. Casos de Estudio

Vista de pagos

La pestaña de Vista de pagos incluye un análisis específico de la parte de la matriculación enfocada a las finanzas, a nivel de números, viendo a grandes rasgos cuanto se ha recaudado con el proceso de matriculación, importe medio por alumno y métodos de pago (véase figura 4.58).



Figura 4.58: Explorador de casos para el proceso de matriculación

En primer lugar, una tarjeta con los principales números (véase figura 4.59), y donde se han incluido las siguientes métricas:

- N.º de Alumnos que comienzan la matriculación.
- Importe total pagado.
- Importe medio por alumno.
- N.º de pagos cancelados.
- N.º de pagos no confirmado



Figura 4.59: Explorador de casos para el proceso de matriculación

Los gráficos que se han desarrollado en esta pestaña son los siguientes:

- Tabla detalle por alumno con la cantidad pagada, cuando se ha pagado y como (Figura 4.60).
- Distribución de número de alumnos basándose en los créditos matriculados (Figura 4.61).
- Distribución de número de pagos por fecha (Figura 4.62).

4.2. Fuente de datos de Matriculación

- Distribución (Porcentaje) de alumnos por método de pago (Figura 4.63).

Alumno	CANTIDAD	FECHA_PAGO	METODO
EST1	349.27	2024-01-15 14:25	Bizum
EST11	891.14	2024-01-16 10:19	Paypal
EST13	1001.74	2024-01-17 13:52	Tarjeta
EST23	1032.1	2024-01-28 10:52	Paypal
EST26	1087.75	2024-01-16 10:12	Bizum
EST28	1106.94	2024-01-17 19:48	Bizum
EST42	947.63	2024-01-15 17:15	Paypal
EST52	1000.97	2024-01-15 14:19	Bizum
EST57	0	2024-01-15 18:02	Paypal

Figura 4.60: Distribución de alumnos por créditos matriculador

Relacion de Alumnos/Creditos

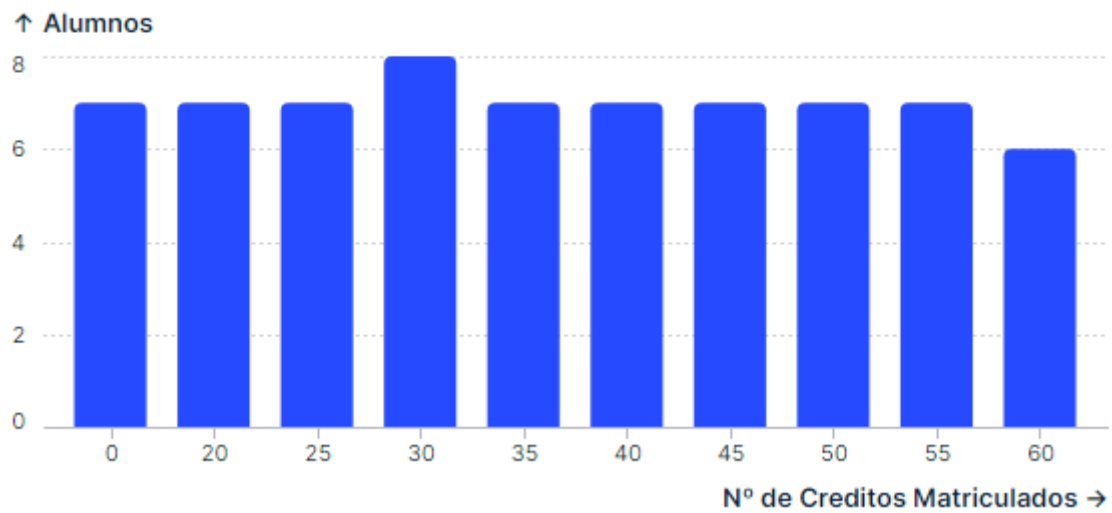


Figura 4.61: Distribución de alumnos por créditos matriculador

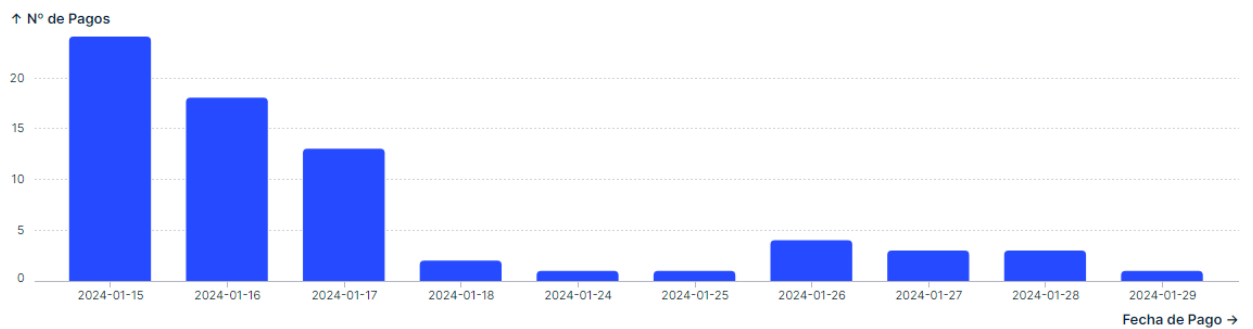


Figura 4.62: Distribución de pagos por fechas de pago

Capítulo 4. Casos de Estudio

Metodos de Pago

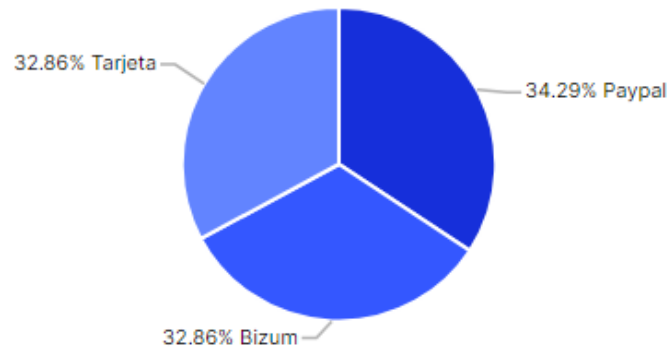


Figura 4.63: Distribución de matriculaciones por método de pago

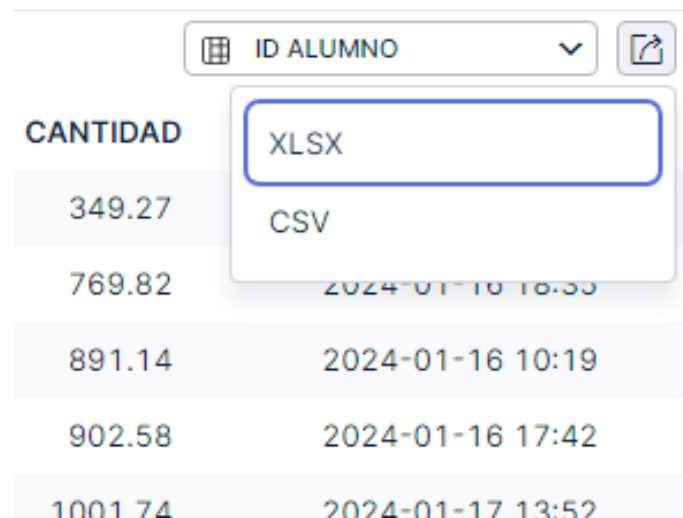
Tabla detalle

En la última pestaña se ha implementado una tabla de detalle de los alumnos exportable como fichero, en caso de necesitar la información filtrada o no para otras tareas, uso en otras herramientas o como prueba de la búsqueda en los datos (véanse las figuras 4.64 y 4.65).

Tabla de Detalle							
ID ALUMNO	GRADO	TIPO_MATRICULA	ESTADO_MATRICULA	LOCALIZACION	SEXO	CANTIDAD	FECHA_PAGO
EST1	GII	Completa	Completada	Madrid	Hombre	349.27	2024-01-15 14:25
EST10	GII	Completa	Completada	Madrid	Hombre	769.82	2024-01-16 18:35
EST11	GII	Parcial	Completada	Madrid	Hombre	891.14	2024-01-16 10:19
EST12	GII	Completa	Completada	Madrid	Hombre	902.58	2024-01-16 17:42
EST13	GII	Completa	Completada	Madrid	Mujer	1001.74	2024-01-17 13:52
EST14	GII	Completa	Completada	Madrid	Hombre	536.09	2024-01-15 10:54
EST15	GMI	Completa	Completada	Madrid	Hombre	790.37	2024-01-17 12:57
EST16	GII	Completa	Completada	Madrid	Hombre	218.53	2024-01-15 19:47
EST17	GII	Parcial	Completada	Madrid	Mujer	645.21	2024-01-15 14:47
EST18	GII	Completa	Completada	Madrid	Hombre	974.45	2024-01-17 11:39
EST19	GMI	Parcial	Completada	Madrid	Hombre	836.97	2024-01-15 09:25
EST2	GII	Parcial	Completada	Madrid	Hombre	874.65	2024-01-15 17:10
EST20	GII	Completa	Completada	Madrid	Hombre	1019.36	2024-01-25 09:47
EST21	GII	Completa	Completada	Madrid	Hombre	713.71	2024-01-27 14:13
EST22	GDI-ADE	Completa	Completada	Madrid	Hombre	830.2	2024-01-24 17:48
EST23	GII	Completa	Completada	Madrid	Hombre	1032.1	2024-01-28 10:52
EST24	GII	Completa	Completada	Madrid	Hombre	920.62	2024-01-27 11:27
EST25	GII	Completa	Completada	Madrid	Mujer	927.84	2024-01-16 16:09
EST26	GMI	Parcial	Completada	Madrid	Hombre	1087.75	2024-01-16 10:12
EST27	GII	Completa	Completada	Madrid	Hombre	1165.18	2024-01-15 12:47

Figura 4.64: Explorador de casos para el proceso de matriculación

4.2. Fuente de datos de Matriculación



CANTIDAD	
349.27	
769.82	2024-01-16 16:33
891.14	2024-01-16 10:19
902.58	2024-01-16 17:42
1001.74	2024-01-17 13:52

Figura 4.65: Exportación de tabla para análisis

Flujo de acción: alarmado por correo

Otra de las funcionalidades que ofrece Celonis es la utilización de una plataforma *Low-Code* para la interacción con los usuarios mediante automatizaciones. Se pueden realizar multitud de tareas como puede ser generar un Excel, enviar un correo, modificar una base de datos sin la manualidad del usuario, etc.

En este caso se va a generar un alarmado por correo que avise a los responsables de matriculación en caso de que algún alumno tenga alguna duda o soporte. Esta interacción será en tiempo real, permitiendo así que el alumno, que tenía la cita, no tenga la oportunidad de perder el hueco en alguna asignatura que desee matricular por la demora de la respuesta.

La implantación del alarmado será la representada en la Figura 4.66



Figura 4.66: Alarmado a responsables en caso de duda

Esta automatización de alarmado tiene tres componentes:

Capítulo 4. Casos de Estudio

- El elemento de recopilación de datos recogería los datos de la solicitud de duda del alumno en tiempo real.
- Construye mediante las herramientas de tratamiento de tablas o texto un registro de esa petición.
- Envía un correo al responsable para que sea atendido con la mayor urgencia posible.

4.2.6. Oportunidades teóricas de mejora

Tomando como asunción y premisa que el caso de uso y este proceso está realizado de manera ficticia, asumiendo las actividades del proceso así como las actividades no deseadas, se van a tener en cuenta escenarios teóricos de mejora que puedan aplicarse al proceso de matriculación, pero que sean realistas, es decir, que desde cualquier Universidad que lea este trabajo y este interesado en este caso de uso, pueda aplicar estas oportunidades de mejora en su operativa y procesos.

Para este caso de uso se han identificado 3 oportunidades de mejora.

Oportunidad I: Gestión proactiva del proceso de matriculación

La implementación de una visualización o reporte, donde tener centralizado todas las matriculaciones, tanto a nivel de volumetrías, de actividades, de tiempos y de finanzas, permite un gran seguimiento por parte de los responsables de matriculación de la universidad. Tener esta implementación actualizada y dividida por grado y otras dimensiones permite un seguimiento mucho más exhaustivo de lo que se puede tener actualmente y lo que puede ofrecer este proceso con reportes en hojas de cálculo o preguntando a los responsables del proceso.

Oportunidad II: Herramienta de alarmado en tiempo real.

Existen casos en los que los alumnos reservan la cita de matriculación, comienzan el proceso y a la hora de elegir grupos o asignaturas resulta que tienen algún tipo de duda, consulta o pregunta sobre cualquier aspecto de las clases o de las asignaturas. Si el alumno decide no continuar con la matriculación hasta resolver esa duda, es posible que en aquellas asignaturas donde el cupo es más limitado y tienen mayor demanda, pueda quedarse sin hueco. Por ello, una herramienta de agilización de la interacción para resolver la duda puede acortar los tiempos de matriculación y mejorar la experiencia del alumno.

Oportunidad III: Enfoque de campañas de matriculación por dimensiones de alumno

Las entidades universitarias destinan mucho presupuesto a campañas para fomentar la matriculación a ciertos sectores y en ciertas localizaciones. Tener monitorizado cuáles son los tipos de estudiantes que se matriculan, el sexo de los alumnos, incluso si prefieren una matrícula parcial para compaginar con su

4.2. Fuente de datos de Matriculación

trabajo, permite conocer mejor el tipo de alumno y como ofrecer campañas más efectivas.

Capítulo 5

Conclusiones y trabajo futuro

El objetivo principal de este proyecto era poder extraer y representar los procesos del entorno universitario mediante Minería de Procesos.

El resultado tras este trabajo ha sido satisfactorio, descubriendo bastante valor e impacto de la minería de procesos tras dos casos de uso sencillos en dos apartados muy dispares del ecosistema universitario (la plataforma Moodle y el proceso de matriculación).

Lo más representativo de esta optimización no es solo los hallazgos encontrados tras la implementación sino el diverso impacto que puede tener, ya que no solo se trata de impacto económico o impacto en recursos, sino que la introducción de minería de procesos en las universidades puede generar bastante impacto de manera transversal en los procesos.

Los objetivos específicos que se habían propuesto eran los siguientes:

- Recuperación de información de diversas fuentes de datos.
- Implementar diversas técnicas de minería de procesos en el ámbito universitario.
- Extraer conocimiento al realizar el descubrimiento basado en los datos objetivos de dichos procesos.
- Elaboración de mejoras de los procesos evaluados mediante el planteamiento de hipótesis de los datos obtenidos.

Tras este trabajo se puede afirmar que con base en estos objetivos planteados se puede ofrecer respuesta a todos ellos, ya que se ha conseguido de manera exitosa recuperar datos de diversos procesos y mediante la implementación de técnicas de Minería de Procesos extraer conocimiento de los datos y procesos y elaborar oportunidades de mejora a todos ellos.

Respecto al trabajo futuro, habiendo demostrado que con las técnicas explicadas en este trabajo de cómo tratar los datos para ser viables y analizados mediante cualquier técnica o herramienta de process mining, se puede asumir que prácticamente cualquier proceso con huella digital puede ser objeto de extensión del trabajo realizado.

Uno de los procesos más importantes que existe en cualquier universidad y que

mayor impacto puede generar al ser estándar y parecido entre las diferentes universidades de toda Europa sería el proceso de Erasmus. Un proceso muy interesante, extenso y del que se podría extraer muchísimo impacto a las técnicas y metodologías aplicadas en este trabajo.

Capítulo 6

Análisis de impacto

En este capítulo se analiza el impacto potencial de los resultados obtenidos durante el trabajo, destacando los puntos fuertes y efectos adversos si así fuese la situación. Estos resultados de impacto están alineados con los Objetivos de Desarrollo Sostenible (ODS) de la Agenda 2030 [8].

6.1. Impacto empresarial

Los beneficios esperados en un trabajo de estas características es bastante elevado, ya que implica una optimización de procesos, en particular en las instituciones educativas. Utilizando una metodología de trabajo y toma de decisiones orientadas al dato o basada en datos, aplicando así una asunción de la veracidad de los datos que hace de toda decisión una decisión objetiva.

Por otra parte, al ser una nueva herramienta, un posible efecto adverso sería la resistencia al uso y al cambio por parte del personal. La palabra optimización y cambio puede generar crispación y rechazo por los empleados. Esto podría solventarse con una planeada formación y capacitación continua del personal para convertir esta resistencia al cambio en una adopción progresiva y creación de recursos adeptos y expertos en minería de procesos.

6.2. Impacto económico

Aunque en estos casos de uso se ha trabajado solo con los pagos de matriculación y en cantidades ficticias bastante pequeñas, el uso de minería de procesos en casos de uso de finanzas puede generar un retorno de la inversión y un descubrimiento de capital escondido o capital que puede ser optimizado en gran medida. Esto puede generar grandes beneficios económicos.

6.3. Impacto social y cultural

La educación en todo su espectro, partiendo de las escuelas infantiles hasta la enseñanza universitaria como de postgrado, son uno de los pilares de la socie-

6.4. Impacto basándose en los Objetivos de Desarrollo Sostenible (ODS)

dad, generando expertos en toda materia, desde la más básica hasta la más complicada. Aplicando Minería de Procesos sobre los procesos universitarios conseguimos una mejora en la calidad del servicio educativo y un aumento de la transparencia de estos procesos. Por otra parte, se ha visto que históricamente la participación de la mujer en las carreras de ingeniería es inferior a la de los hombres. Con herramientas de Process Mining puede identificarse los grados dónde la presencia de la mujer es menor y enfatizar tanto campañas como iniciativas que poco a poco puedan hacer de estos grados y de la participación de la mujer en ellos un número similar al de los hombres.

6.4. Impacto basándose en los Objetivos de Desarrollo Sostenible (ODS)

El 25 de septiembre de 2015, los líderes mundiales adoptaron un conjunto de objetivos globales para erradicar la pobreza, proteger el planeta y asegurar la prosperidad para todos como parte de una nueva agenda de desarrollo sostenible, con metas específicas que deben alcanzarse en los próximos 15 años. Para lograr estos objetivos, es necesario que todos participen: gobiernos, sector privado, sociedad civil y personas individuales [8].

Este es el eslogan principal de los 17 Objetivos de Desarrollo Sostenible, tal y como se muestran en la Figura 6.1.



Figura 6.1: Objetivos de Desarrollo Sostenible

En este trabajo, el potencial impacto se consigue en los siguientes objetivos:

- 4. Educación de Calidad: consiguiendo una mejora de los procesos educativa para una educación digna y de calidad.
- 5. Igualdad de género: consiguiendo mediante iniciativas en los procesos y campañas una mayor inclusión de la mujer en estos grados universitarios.

Bibliografía

- [1] J. L. del Val Román, «Industria 4.0: la transformación digital de la industria», en *Valencia: Conferencia de Directores y Decanos de Ingeniería Informática, Informes CODDII*, 2016.
- [2] D. Patil y H. Mason, *Data Driven*. O'Reilly Media, Inc., 2015.
- [3] P. Vassiliadis, «A survey of extract-transform-load technology», *International Journal of Data Warehousing and Mining (IJDWM)*, vol. 5, n.º 3, págs. 1-27, 2009.
- [4] C. Pérez López y D. Santin González, *Minería de datos. Técnicas y herramientas: técnicas y herramientas*. Ediciones Paraninfo, SA, 2007.
- [5] C. Ullrich, T. Lata y J. Geyer-Klingeberg, «Celonis Studio-A Low-Code Development Platform for Citizen Developers.», en *BPM (PhD/Demos)*, 2021, págs. 102-105.
- [6] B. Marr, *Key Performance Indicators (KPI): The 75 measures every manager needs to know*. Pearson UK, 2012.
- [7] A. Srivastava, S. Bhardwaj y S. Saraswat, «SCRUM model for agile methodology», en *2017 International Conference on Computing, Communication and Automation (ICCCA)*, IEEE, 2017, págs. 864-869.
- [8] C. G. Gil, «Objetivos de Desarrollo Sostenible (ODS): una revisión crítica», *Papeles de relaciones ecosociales y cambio global*, n.º 140, págs. 107-118, 2018.
- [9] W. Van Der Aalst, «Process mining: Overview and opportunities», *ACM Transactions on Management Information Systems (TMIS)*, vol. 3, n.º 2, págs. 1-17, 2012.
- [10] Gartner. «Cuadrante Mágico de Gartner». (2023), dirección: <https://www.gartner.es/es/metodologias/magic-quadrant> (visitado 24-05-2024).
- [11] J. L. Cano, «Business Intelligence: competir con información», 2007.
- [12] Microsoft. «PowerBI». (2024), dirección: <https://www.microsoft.com/es-es/power-platform/products/power-bi> (visitado 24-05-2024).
- [13] Qlik. «QlikSense, Herramienta de visualización de datos». (2024), dirección: <https://www.qlik.com/es-es/products/qlik-sense> (visitado 24-05-2024).
- [14] Tableau. «Tableau, herramienta de visualización de datos». (2024), dirección: <https://www.tableau.com/es-es> (visitado 24-05-2024).

- [15] Z.-H. Zhou, *Machine learning*. Springer nature, 2021.
- [16] A. Calder y S. Watkins, *IT Governance: an international guide to data security and ISO27001/ISO27002*. Kogan Page Publishers, 2012.
- [17] O. Fal', «Documentation in the ISO/IEC 27701 Standard», *Cybernetics and Systems Analysis*, vol. 57, n.º 5, págs. 796-802, 2021.
- [18] P. Cababie, «Análisis de la Política de Protección General de Datos en Europa (GPDR General Data Protection Regulation)», Tesis doct., Universidad de Buenos Aires. Facultad de Ciencias Económicas., 2017.
- [19] A. Lerma y M. J. Benito, «STAR, certificación de la seguridad de los proveedores de servicio desde la nube», *Revista SIC: ciberseguridad, seguridad de la información y privacidad*, n.º 108, págs. 100-101, 2014.
- [20] E. Vértice et al., *Gestión de la calidad (ISO 9001/2008)*. Editorial Vértice, 2010.
- [21] AICPA. «SOC1, SOC2, SOC3». (2024), dirección: <https://us.aicpa.org/content/dam/aicpa/interestareas/frc/assuranceadvisoryservices/downloadabledocuments/soc-reports-flyer-final.pdf> (visitado 24-05-2024).
- [22] Celonis. «Celonis Product Documentation - Studio». (2024), dirección: <https://docs.celonis.com/en/studio.html> (visitado 27-05-2024).
- [23] D. Ma, «The business model of "software-as-a-service"», en *Ieee international conference on services computing (scc 2007)*, IEEE, 2007, págs. 701-702.
- [24] VERTICA. «Vertica SQL Reference». (2024), dirección: <https://docs.vertica.com/24.2.x/en/sql-reference/> (visitado 24-05-2024).
- [25] P. DuBois, *MySQL*. New riders publishing, 1999.
- [26] L. Scott, *Relational Database and Transact-SQL*, 2017.
- [27] H. M. Marin-Castro y E. Tello-Leal, «Event log preprocessing for process mining: a review», *Applied Sciences*, vol. 11, n.º 22, pág. 10556, 2021.
- [28] S. Junestrand, X. Passaret y D. VAZQUEZ ALVAREZ, *Domótica y hogar digital*. Ediciones Paraninfo, SA, 2004.
- [29] L. Corr y J. Stagnitto, *Agile data warehouse design: Collaborative dimensional modeling, from whiteboard to star schema*. DecisionOne Consulting, 2011.
- [30] T. Vogelgesang, J. Ambrosy, D. Becher, R. Seilbeck, J. Geyer-Klingeberg y M. Klenk, «Celonis PQL: A query language for process mining», en *Process Querying Methods*, Springer, 2021, págs. 377-408.
- [31] I. R. M. de Lahidalga, «Moodle, la plataforma para la enseñanza y organización escolar», *Ikastorratza, e-Revista de didáctica*, n.º 2, págs. 3-12, 2008.
- [32] H. Dobbertin, «Cryptanalysis of MD5 compress», *rump session of Eurocrypt*, vol. 96, págs. 71-82, 1996.
- [33] F. Mannhardt, M. De Leoni y H. A. Reijers, «The multi-perspective process explorer», en *13th International Workshops on Business Process Management Workshops (BPM 2015)*, 2015, págs. 130-134.

Anexos

Índice de figuras

1.1. Diagrama de Gantt del proyecto	3
2.1. Process Mining como intersección de Procesos y Datos	5
2.2. Huella digital y sus atributos	6
2.3. Representación del Árbol de Proceso	7
2.4. Cuadrante Mágico de Gartner 2023 en Process Mining	8
3.1. Arquitectura de la plataforma de Celonis y sus componentes	11
3.2. Estructura de la Integración de datos	11
3.3. Consola de desarrollo de Vertica SQL en Celonis	12
3.4. Árbol de Proceso - Camino ideal y sencillo	14
3.5. Árbol de Proceso - Camino con desviaciones	15
3.6. Modelo de relación entre tablas	17
3.7. Modelo de relación entre tablas	17
3.8. Estructuración de elementos	18
3.9. Elementos de Visualización	19
3.10 Elementos de Visualización	20
4.1. Lista de Transformaciones en la plataforma	23
4.2. Modelo Relacional de Datos	26
4.3. Volumetría de las tablas cargadas en el modelo de datos	26
4.4. Arquitectura del proceso en Celonis	27
4.5. Organización de los elementos de visualización	27
4.6. Organización de los elementos de visualización II	28
4.7. Modelo de conocimiento del proceso de Moodle	28
4.8. Explorador de Proceso	29
4.9. Explorador de proceso parcialmente ampliado a nivel de actividades y conexiones	30
4.10 Explorador de proceso completo - Diagrama espagueti	30
4.11 Filtros de Actividad y selección de dimensión	31
4.12 Árbol de Proceso filtrado por actividad	32
4.13 Explorador de casos con el detalle por usuario	33
4.14 Pestaña de Explorador de Procesos	34
4.15 Funcionalidad de filtrado por asignatura y fecha	35
4.16 Métricas principales del proceso	35
4.17 El gráfico explorador de proceso	36
4.18 Gráfico de Explorador de casos dentro de la pestaña	37

4.19	Pestaña completa de KPIs generales	37
4.20	Gráfico de distribución de alumnos por asignatura	38
4.21	Gráfico de número de accesos a lo largo del tiempo	38
4.22	Distribución de las actividades y eventos del sistema	38
4.23	Pestaña de Análisis de entregas	39
4.24	Entregas tardías por asignatura	40
4.25	Entregas a tiempo	40
4.26	Alumnos sin entregas	41
4.27	Histograma en horas sobre la fecha de entrega	41
4.28	Tabla detalle de los alumnos	42
4.29	Test	46
4.30	Modelo de Datos relacional del caso de matriculación	49
4.31	Volumetría de los datos del modelo	49
4.32	Estructura de elementos	50
4.33	Visualización de la organización de elementos en Celonis	50
4.34	Modelo de conocimiento del caso de uso de Matriculación	51
4.35	Panel de filtros del análisis	52
4.36	Guía de Usuario para el proceso de matriculación	52
4.37	Explorador de proceso de Matriculación	53
4.38	Explorador de proceso entendido	54
4.39	Diagrama espagueti para el proceso de matriculación	54
4.40	Explorador de casos para el proceso de matriculación	55
4.41	Vista completa pestaña general	56
4.42	Métricas principales para la vista general del proceso	56
4.43	Distribución por tipo de matrícula.	57
4.44	Distribución por tipo de matrícula.	57
4.45	Distribución por periodo de matrícula.	57
4.46	Distribución por estado de matrícula.	58
4.47	Distribución por sexo del alumno.	58
4.48	Vista completa pestaña de actividades	58
4.49	Explorador de casos para el proceso de matriculación	59
4.50	Distribución y ocurrencias de las actividades	59
4.51	Distribución de los alumnos por grado	59
4.52	Distribución de las actividades por tipo de dispositivo	60
4.53	Vista general pestaña de pagos	60
4.54	Tarjeta de métricas de la vista de pagos	61
4.55	Distribución de matriculaciones por localización	61
4.56	Distribución de actividades y tiempo de actividad	61
4.57	Distribución de alumnos/matriculaciones por periodo de matriculación	61
4.58	Explorador de casos para el proceso de matriculación	62
4.59	Explorador de casos para el proceso de matriculación	62
4.60	Distribución de alumnos por créditos matriculador	63
4.61	Distribución de alumnos por créditos matriculador	63
4.62	Distribución de pagos por fechas de pago	63
4.63	Distribución de matriculaciones por método de pago	64
4.64	Explorador de casos para el proceso de matriculación	64

ÍNDICE DE FIGURAS

4.65Exportación de tabla para análisis	65
4.66Alarmado a responsables en caso de duda	65
6.1. Objetivos de Desarrollo Sostenible	71

Índice de tablas

3.1. Ejemplo de registro de eventos	13
3.2. Ejemplo de registro de evento con desviaciones	13
3.3. Estructura de campos de registro de eventos	16
3.4. Ejemplo de registro de evento con desviaciones	16
4.1. Tabla fuente de datos por Asignatura	22
4.2. Tabla de Casos	24
4.3. Tabla de Actividades	25
4.4. Tabla de entregas de alumnos	25
4.5. Tabla de fechas límite de entrega	25
4.6. Tabla de Estudiantes (Casos)	45
4.7. Tabla de definición de actividades	45
4.8. Tabla de Actividades	45
4.9. Tabla de Pagos	46
4.10.Tabla de Casos	48
4.11.Tabla de Actividades	48
4.12.Tabla de Pagos	48

Apéndice A

Moodle - Código de las transformaciones

A.1. Transformación Tabla de Actividades

```
-- Compango de cursos en tabla unica para su limpieza  
DROP TABLE IF EXISTS Logs_Moodle_No_Hash;
```

```
CREATE TABLE Logs_Moodle_No_Hash (  
  USUARIO VARCHAR(150),  
  USUARIO_AFECTADO VARCHAR(150),  
  TIMESTAMP TIMESTAMP,  
  CONTEXTO_EVENTO VARCHAR(150),  
  COMPONENTE VARCHAR(30),  
  ACTIVITY_NAME VARCHAR(100),  
  DESCRIPCION VARCHAR(800),  
  ORIGEN VARCHAR(10),  
  ASIGNATURA VARCHAR(30),  
  ENTREGA_TIEMPO INTEGER  
);
```

```
INSERT INTO Logs_Moodle_No_Hash  
SELECT  
CASE WHEN  
  "NOMBRE COMPLETO DEL USUARIO" = 'User Admin'  
then 'Admin'  
when "NOMBRE COMPLETO DEL USUARIO" = '-' then 'System'  
else MD5("NOMBRE COMPLETO DEL USUARIO")  
END,  
CASE WHEN  
  "USUARIO AFECTADO" = 'User Admin'  
then 'Admin'  
when "USUARIO AFECTADO" = '-' then 'System'
```

A.1. Transformación Tabla de Actividades

```
else MD5("USUARIO AFECTADO")
END,
HORA AS TIMESTAMP,
"CONTEXTO DEL EVENTO" AS CONTEXTO_EVENTO,
COMPONENTE,
"NOMBRE EVENTO" AS ACTIVITY_NAME,
"DESCRIPCIÓN",
ORIGEN,
'BIOTECNOLOGIA' AS "ASIGNATURA"
FROM "logs_BD-BIOTECNOLOGIA"

UNION

SELECT
CASE WHEN
"NOMBRE COMPLETO DEL USUARIO" = 'User Admin'
then 'Admin'
when "NOMBRE COMPLETO DEL USUARIO" = '-' then 'System'
else MD5("NOMBRE COMPLETO DEL USUARIO")
END,
CASE WHEN
"USUARIO AFECTADO" = 'User Admin'
then 'Admin'
when "USUARIO AFECTADO" = '-' then 'System'
else MD5("USUARIO AFECTADO")
END,
HORA AS TIMESTAMP,
"CONTEXTO DEL EVENTO" AS CONTEXTO_EVENTO,
COMPONENTE,
"NOMBRE EVENTO" AS ACTIVITY_NAME,
"DESCRIPCIÓN",
ORIGEN,
'DATA PROCESSES' AS "ASIGNATURA" FROM "logs_Data_Processes"

UNION

SELECT
CASE WHEN
"NOMBRE COMPLETO DEL USUARIO" = 'User Admin'
then 'Admin'
when "NOMBRE COMPLETO DEL USUARIO" = '-' then 'System'
else MD5("NOMBRE COMPLETO DEL USUARIO")
END,
CASE WHEN
"USUARIO AFECTADO" = 'User Admin'
then 'Admin'
```

Capítulo A. Moodle - Código de las transformaciones

```
when "USUARIO AFECTADO" = '-' then 'System'
else MD5("USUARIO AFECTADO")
END,
HORA AS TIMESTAMP,
"CONTEXTO DEL EVENTO" AS CONTEXTO_EVENTO,
COMPONENTE,
"NOMBRE EVENTO" AS ACTIVITY_NAME,
"DESCRIPCIÓN",
ORIGEN,
'PRACTICUM' AS "ASIGNATURA" FROM "logs_PRACTICUM"

UNION

SELECT
CASE WHEN
"NOMBRE COMPLETO DEL USUARIO" = 'User Admin'
then 'Admin'
when "NOMBRE COMPLETO DEL USUARIO" = '-' then 'System'
else MD5("NOMBRE COMPLETO DEL USUARIO")
END,
CASE WHEN
"USUARIO AFECTADO" = 'User Admin'
then 'Admin'
when "USUARIO AFECTADO" = '-' then 'System'
else MD5("USUARIO AFECTADO")
END,
HORA AS TIMESTAMP,
"CONTEXTO DEL EVENTO" AS CONTEXTO_EVENTO,
COMPONENTE,
"NOMBRE EVENTO" AS ACTIVITY_NAME,
"DESCRIPCIÓN",
ORIGEN,
'PCD' AS "ASIGNATURA" FROM "logs_P_C_D"

UNION

SELECT
CASE WHEN
"NOMBRE COMPLETO DEL USUARIO" = 'User Admin'
then 'Admin'
when "NOMBRE COMPLETO DEL USUARIO" = '-' then 'System'
else MD5("NOMBRE COMPLETO DEL USUARIO")
END,
CASE WHEN
"USUARIO AFECTADO" = 'User Admin'
then 'Admin'
```

A.2. Transformación Tabla de Casos

```
when "USUARIO AFECTADO" = '-' then 'System'
else MD5("USUARIO AFECTADO")
END,
FECHA AS TIMESTAMP,
"CONTEXTO DEL EVENTO" AS CONTEXTO_EVENTO,
COMPONENTE,
"NOMBRE EVENTO" AS ACTIVITY_NAME,
"DESCRIPCIÓN",
ORIGEN,
'TFG' AS "ASIGNATURA"FROM "logs_TFG";
```

A.2. Transformación Tabla de Casos

```
DROP TABLE IF EXISTS Casos;
```

```
CREATE TABLE Casos (
    idUsuario VARCHAR(100) PRIMARY KEY,
    cantidadActividades INTEGER,
    duracionCaso INTEGER,
    numAsignaturas INTEGER,
    ultimoEstadoCaso VARCHAR(50),
    ultimaConexion TIMESTAMP
);
```

```
insert into Casos
SELECT DISTINCT USUARIO, count(ACTIVITY_NAME) as "cantidadActividades"
FROM "Logs_Moodle_No_Hash"
group by USUARIO;
```

```
UPDATE Casos
set duracionCaso= L.duracionCaso,
numAsignaturas= L.numAsignaturas
from
(SELECT DISTINCT USUARIO AS USUARIO,
DATEDIFF(dd,MIN("TIMESTAMP"),MAX("TIMESTAMP")) as duracionCaso ,
COUNT(DISTINCT ASIGNATURA) AS numAsignaturas
FROM "Logs_Moodle_No_Hash"
group by USUARIO) L
where idUsuario=L.USUARIO;
```

```
UPDATE Casos
set ultimoEstadoCaso= L.ultestado,
ultimaConexion= L.ultconex
from
(SELECT USUARIO AS USUARIO, ACTIVITY_NAME AS ultestado,
"TIMESTAMP" as ultconex
```

```
FROM (
    SELECT
        USUARIO,
        ACTIVITY_NAME,
        "TIMESTAMP",
        ROW_NUMBER() OVER (PARTITION BY USUARIO ORDER BY "TIMESTAMP" DESC)
        AS rn
    FROM
        "Logs_Moodle_No_Hash"
) AS ranked
WHERE
    rn = 1) L
where idUsuario=L.USUARIO;
```

A.3. Transformación Tablas Auxiliares

```
--ENTREGAS
DROP TABLE IF EXISTS entregasAlumnos;

CREATE TABLE entregasAlumnos AS
SELECT DISTINCT USUARIO, ASIGNATURA, count (ACTIVITY_NAME) as numEntregas
FROM Logs_Moodle_No_Hash
WHERE ACTIVITY_NAME IN ('Se ha enviado una entrega',
'El estado de la entrega se ha actualizado.',
'Se ha entregado una extensión.')
GROUP BY USUARIO, ASIGNATURA;

--DEADLINES
DROP TABLE IF EXISTS deadlinesEntregas;

CREATE TABLE deadlinesEntregas AS
SELECT *
FROM "DLS";

-- Numero de Entregas que deben hacerse
DROP TABLE IF EXISTS entregasAsignatura;

CREATE TABLE entregasAsignatura (
    ASIGNATURA VARCHAR (100),
    numEntregasTotales INTEGER,
    numEntregasObligatorias INTEGER,
    numEntregasOpc INTEGER
);

INSERT INTO entregasAsignatura (ASIGNATURA,numEntregasTotales,numEntregasOblig
VALUES
('PRACTICUM', 4, 3, 1),
```

```
('BIOTECNOLOGIA',2,2,0),
('PCD',4,3,1),
('TFG',4,4,0);
```

A.4. Transformación Ajustes

```
--Entrega a tiempo o no
UPDATE "Logs_Moodle_No_Hash" C
SET "ENTREGA_TIEMPO" = b.tiempo
FROM (SELECT a.USUARIO,a.ASIGNATURA, a.CONTEXTO_EVENTO,
datediff(mi,a."FECHA ENTREGA",a."TIMESTAMP") AS tiempo
FROM
(SELECT DISTINCT USUARIO, m.ASIGNATURA, M.CONTEXTO_EVENTO,
"TIMESTAMP", DL."FECHA ENTREGA"
FROM "Logs_Moodle_No_Hash" M
LEFT JOIN "deadlinesEntregas" DL
ON M.ASIGNATURA=DL.ASIGNATURA AND M.CONTEXTO_EVENTO=DL.CONTEXTO_EVENTO
WHERE
M."ACTIVITY_NAME" IN ('Se ha enviado una entrega',
'El estado de la entrega se ha actualizado.',
'Se ha entregado una extensión.')) a)b
where C.USUARIO = b.USUARIO and C.ASIGNATURA=b.ASIGNATURA
AND C.CONTEXTO_EVENTO=b.CONTEXTO_EVENTO;

-- Numero de Entregas por Asignatura
UPDATE "entregasAlumnos" E
SET totalEntregas = A.numEntregasObligatorias
FROM entregasAsignatura A
WHERE E.ASIGNATURA=A.ASIGNATURA;
```

Apéndice B

Moodle - Código de las Vistas, métricas y KPIs

Número de accesos

```
count(case when "Logs_Moodle_No_Hash"."ACTIVITY_NAME" like '%visto%'
then "Logs_Moodle_No_Hash"."ACTIVITY_NAME"
end)
```

Número de alumnos

```
COUNT ( DISTINCT "Casos"."idUsuario")
```

Número de alumnos que no han realizado ninguna entrega

```
sum(CASE WHEN
PU_COUNT ( "Casos", "Logs_Moodle_No_Hash"."ACTIVITY_NAME",
"Logs_Moodle_No_Hash"."ACTIVITY_NAME" IN ('Se ha enviado una entrega',
'El estado de la entrega se ha actualizado.',
'Se ha entregado una extensión.')
)=0
then 1
else 0
end)
```

Porcentaje de alumnos que no han realizado ninguna entrega sobre el total

```
count(distinct CASE WHEN
PU_COUNT ( "Casos", "Logs_Moodle_No_Hash"."ACTIVITY_NAME",
"Logs_Moodle_No_Hash"."ACTIVITY_NAME" IN ('Se ha enviado una entrega',
'El estado de la entrega se ha actualizado.',
'Se ha entregado una extensión.')
)=0
then "Casos"."idUsuario"
end)
/
count(distinct "Casos"."idUsuario")
```

N.º de Entregas a Realizar

```
COUNT (
    "deadlinesEntregas"."CONTEXTO_EVENTO"
)
```

N.º de Entregas realizadas

```
sum (CASE WHEN
    "Logs_Moodle_No_Hash"."ACTIVITY_NAME" IN ('Se ha enviado una entrega',
    'El estado de la entrega se ha actualizado.',
    'Se ha entregado una extensión.')
then 1
else 0
end)
```

N.º de Entregas a tiempo

```
COUNT(CASE WHEN
    "Logs_Moodle_No_Hash"."ACTIVITY_NAME" IN ('Se ha enviado una entrega',
    'El estado de la entrega se ha actualizado.',
    'Se ha entregado una extensión.')
AND ("Logs_Moodle_No_Hash"."ENTREGA_TIEMPO" <= 0
or "Logs_Moodle_No_Hash"."ENTREGA_TIEMPO" is null)
THEN "Logs_Moodle_No_Hash"."CONTEXTO_EVENTO"
END)
```

N.º de Entregas fuera de tiempo

```
COUNT(CASE WHEN
    "Logs_Moodle_No_Hash"."ACTIVITY_NAME" IN ('Se ha enviado una entrega',
    'El estado de la entrega se ha actualizado.',
    'Se ha entregado una extensión.')
AND "Logs_Moodle_No_Hash"."ENTREGA_TIEMPO" > 0
THEN "Logs_Moodle_No_Hash"."CONTEXTO_EVENTO"
END)
```

Apéndice C

Matriculación - Código de las transformaciones

C.1. Transformación Tabla de Actividades

```
-- Creacion de tabla de actividades
DROP TABLE IF EXISTS MATRICULACION_ACTIVIDADES;

CREATE TABLE MATRICULACION_ACTIVIDADES (
  _ACTIVITY_ID INTEGER,
  _CASE_ID VARCHAR(10),
  _ACTIVITY_NAME VARCHAR(100),
  _TIMESTAMP TIMESTAMP,
  SORTING INTEGER,
  PERIODO VARCHAR(100),
  DISPOSITIVO VARCHAR(100)
);

INSERT INTO MATRICULACION_ACTIVIDADES
SELECT
  L."IDACTIVIDAD",
  "ALUMNO",
  A.ACTIVIDAD,
  "TIMESTAMP",
  A.SORTING,
  "PERIODO",
  "DISPOSITIVO"
FROM "LOG_MATRICULACION" L
LEFT JOIN
"ACTIVIDADES_MATRICULACION" A
ON L.ACTIVIDAD=A.IDACTIVIDAD;
```

C.2. Transformación Tabla de Casos

```
-- Creacion de tabla de actividades
DROP TABLE IF EXISTS MATRICULACION_CASOS;

CREATE TABLE MATRICULACION_CASOS (
_CASE_ID VARCHAR(10),
GRADO VARCHAR(100),
LOCALIZACION VARCHAR(100),
FECHA_ULT_MATRICULA TIMESTAMP,
CREDITOS_ULT_MATRICULA INTEGER,
ASIGNATURAS INTEGER,
TIPO_MATRICULA VARCHAR(100),
ESTADO_MATRICULA VARCHAR(100)
);

INSERT INTO MATRICULACION_CASOS
SELECT
"ALUMNO",
"GRADO",
"LOCALIZACION",
"FECHA ULTIMA MATRICULA",
"CREDITOS ULTIMA MATRICULA",
"ASIGNATURAS ",
"TIPO DE MATRICULA",
"ESTADO MATRICULA"
FROM "CASOS_MATRICULACION";
```

C.3. Transformación Tablas Auxiliares

```
-- Creacion de tabla auxiliar de pagos de alumnos
DROP TABLE IF EXISTS MATRICULACION_PAGOS;

CREATE TABLE MATRICULACION_PAGOS (
_PAGO_ID INTEGER,
_CASE_ID VARCHAR(10),
FECHA_PAGO TIMESTAMP,
CANTIDAD FLOAT,
METODO VARCHAR(100)
);

INSERT INTO MATRICULACION_PAGOS
SELECT
"ID_PAGO",
"ID_CASO",
"FECHA_PAGO",
"CANTIDAD",
```

Capítulo C. Matriculación - Código de las transformaciones

```
"METODO"  
FROM "PAGOS_MATRICULACION";
```

Apéndice D

Matriculación - Código de las Vistas, métricas y KPIs

Importe medio por alumno

```
AVG ("MATRICULACION_PAGOS"."CANTIDAD")
```

Numero de Alumnos con Incidencia/Duda

```
COUNT (CASE WHEN  
"MATRICULACION_ACTIVIDADES"."_ACTIVITY_NAME"  
= 'Contacto con Soporte por queja/duda'  
THEN  
"MATRICULACION_ACTIVIDADES"."_CASE_ID"  
END)
```

Numero de Alumnos con matricula parcial

```
COUNT ( CASE WHEN  
"MATRICULACION_CASOS"."TIPO_MATRICULA" = 'Parcial'  
then "MATRICULACION_CASOS"."_CASE_ID" end )
```

Numero de Alumnos que comienzan la matriculación

```
COUNT ("MATRICULACION_CASOS"."_CASE_ID")
```

Numero de veces que no se completa la matriculación

```
COUNT (distinct  
CASE WHEN  
"MATRICULACION_ACTIVIDADES"."_ACTIVITY_NAME" LIKE '%no comp%'  
then  
"MATRICULACION_CASOS"."_CASE_ID"  
end)
```

Numero de Procesos Iniciados


Capítulo D. Matriculación - Código de las Vistas, métricas y KPIs

```
COUNT (CASE WHEN
"MATRICULACION_ACTIVIDADES"."_ACTIVITY_NAME" =
'Inicio de proceso de matrícula.'
THEN "MATRICULACION_ACTIVIDADES"."_CASE_ID"
END)
```

Tiempo Medio

```
AVG (
  case WHEN
    "MATRICULACION_ACTIVIDADES"."_ACTIVITY_NAME"
    not like '%cita previa%'
  then
    DATEDIFF ( Mi , "MATRICULACION_ACTIVIDADES"."_TIMESTAMP" ,
    ACTIVITY_LEAD ("MATRICULACION_ACTIVIDADES"."_TIMESTAMP") ) end )
```

Este documento esta firmado por



Firmante	CN=tfgm.fi.upm.es, OU=CCFI, O=ETS Ingenieros Informaticos - UPM, C=ES
Fecha/Hora	Mon Jun 03 23:32:31 CEST 2024
Emisor del Certificado	EMAILADDRESS=camanager@etsiinf.upm.es, CN=CA ETS Ingenieros Informaticos, O=ETS Ingenieros Informaticos - UPM, C=ES
Numero de Serie	561
Metodo	urn:adobe.com:Adobe.PPKLite:adbe.pkcs7.sha1 (Adobe Signature)