

PROFESOR PATROCINANTE:

ANIBAL EDMUNDO FAUNDEZ DEL RIO

ESCUELA DE INFORMÁTICA Y TELECOMUNICACIONES

Axis Flow

Proyecto Asignatura Proceso de Portafolio de Título

Para optar al título de **Ingeniero en Informática**

ABRAHAM ISAIAS RUBILAR ALBARRAN

ADAMI BENJAMIN BERRIOS MONTECINOS

HERNAN GARRIDO

PUERTO MONTT – CHILE

2025

DEDICATORIA

Dedicamos este trabajo a nuestras familias, por su apoyo incondicional y por acompañarnos durante todo este proceso académico. Agradecemos también a nuestros docentes por su orientación, y a todas las personas que directa o indirectamente contribuyeron al desarrollo de este proyecto. Su respaldo, confianza y motivación han sido fundamentales para avanzar con compromiso y alcanzar los objetivos propuestos como equipo.

AGRADECIMIENTOS

Agradecemos a la empresa Axis por la disposición y la confianza brindada al permitirnos participar en un proyecto real, facilitando información, espacio de colaboración y retroalimentación continua durante las distintas etapas del trabajo.

Agradecemos especialmente a la Directora de Informática, Carolina Martínez, por confiarnos este proyecto, por su invaluable mediación en las gestiones entre Duoc UC y Axis, y por facilitar cada etapa del proceso.

Extendemos nuestro agradecimiento a nuestros docentes, quienes nos guiaron y apoyaron con sus orientaciones técnicas y metodológicas, permitiendo que este proyecto avanzara con claridad y sentido formativo.

Asimismo, valoramos el compromiso y la colaboración de nuestros compañeros de equipo, quienes contribuyeron desde sus fortalezas individuales al desarrollo integral de este proyecto. Su dedicación y coordinación fueron esenciales para alcanzar los resultados presentados.

Finalmente, agradecemos a nuestras familias y redes de apoyo, cuyo respaldo constante nos permitió mantener el ritmo, la motivación y el enfoque necesarios para completar esta etapa con responsabilidad y compromiso profesional.

SUMARIO

El informe final describe el desarrollo del proyecto Axis Flow, realizado con Axis Desarrollos Constructivos para mejorar la calidad, trazabilidad y utilización de datos operacionales que inicialmente estaban dispersos, dificultando su análisis y uso en la toma de decisiones.

El trabajo consistió en transformar registros crudos en un sistema de información estructurado y confiable, mediante un proceso que abarcó el diagnóstico de los datos, su limpieza y normalización, el diseño e implementación de una base de datos relacional, el desarrollo de scripts en Python para calcular tiempos y métricas, la creación de visualizaciones en Streamlit y la revisión conjunta de los entregables.

El proyecto utilizó la metodología ágil Scrum, basada en Sprints que permiten entregar incrementos funcionales revisados por el Product Owner, junto con eventos periódicos que favorecen la coordinación, la retroalimentación continua y el cumplimiento de tareas.

La metodología se alineó con los principios de CRISP-DM (Chapman et al., 2000; Shearer, 2000; Wirth & Hipp, 2000), avanzando iterativamente por etapas de entendimiento del negocio, preparación de datos, modelamiento, evaluación y despliegue parcial, aplicando competencias propias de la Ingeniería en Informática.

Los resultados incluyen un conjunto de datos depurado, una base de datos relacional (Elmasri & Navathe, 2016) que representa el proceso productivo, un modelo de Machine Learning para predicción de eventos, métricas sobre tiempos reales y desviaciones planificadas, y visualizaciones que facilitan la interpretación de indicadores clave.

El impacto esperado es una mayor capacidad para identificar cuellos de botella, evaluar procesos y apoyar la planificación, además de fortalecer la formación del equipo en un entorno real, potenciando habilidades técnicas, colaboración, comunicación y adaptación.

ÍNDICE

DEDICATORIA.....	1
AGRADECIMIENTOS.....	2
SUMARIO.....	3
ÍNDICE.....	4
ÍNDICE DE TABLAS.....	6
ÍNDICE DE FIGURAS.....	7
1. ANTECEDENTES GENERALES.....	8
1.1. Introducción.....	8
1.2. Descripción del Proyecto.....	9
1.2.1. Tema.....	9
1.2.2. Áreas de Desempeño.....	9
1.2.3. Competencias o Unidades de Competencias.....	10
1.3. Fundamentación Proyecto APT.....	12
1.3.1. Relevancia del proyecto APT.....	12
1.3.2. Descripción de Proyecto APT.....	12
1.3.3. Pertinencia del proyecto con el perfil de egreso.....	13
1.3.4. Relación con los intereses profesionales.....	14
1.3.5. Fortalezas y debilidades para desarrollar el proyecto APT.....	15
1.4. Planteamiento del problema.....	17
1.5. Objetivos.....	18
1.5.1. Objetivo General.....	18
1.5.2. Objetivos Específicos.....	18
2. DISEÑO METODOLÓGICO.....	19
2.1. ETAPA N°1 – Exploración y diagnóstico inicial.....	22
2.1.1. Objetivo de la etapa.....	22
2.1.2. Descripción de la etapa.....	22
2.1.3. Actividades realizadas.....	22
2.1.4. Procedimiento.....	22
2.2. ETAPA N°2 – Limpieza y Normalización de Datos.....	24
2.2.1. Objetivo de la etapa.....	24
2.2.2. Descripción de la etapa.....	24
2.2.3. Actividades realizadas.....	24
2.2.4. Procedimiento.....	25
2.3. ETAPA N°3 – Modelamiento y Estructuración.....	26
2.3.1. Objetivo de la etapa.....	26
2.3.2. Descripción de la etapa.....	26
2.3.3. Actividades realizadas.....	26
2.3.4. Procedimiento.....	27

2.4. ETAPA N°4 – Modelado predictivo.....	28
2.4.1. Objetivo de la etapa.....	28
2.4.2. Descripción de la etapa.....	28
2.4.3. Actividades realizadas.....	28
2.4.4. Procedimiento.....	28
2.5. ETAPA N°5 – Desarrollo técnico y procesamiento automatizado.....	30
2.5.1. Objetivo de la etapa.....	30
2.5.2. Descripción de la etapa.....	30
2.5.3. Actividades realizadas.....	30
2.5.4. Procedimiento.....	31
2.6. ETAPA N°6 – Interpretación y Visualización.....	32
2.6.1. Objetivo de la etapa.....	32
2.6.2. Descripción de la etapa.....	32
2.6.3. Actividades realizadas.....	32
2.6.4. Procedimiento.....	33
2.7. ETAPA N°7 – Revisión y Ajuste.....	34
2.7.1. Objetivo de la etapa.....	34
2.7.2. Descripción de la etapa.....	34
2.7.3. Actividades realizadas.....	34
2.7.4. Procedimiento.....	34
3. RESULTADOS.....	36
3.1 Diagnóstico Integral del Estado de los Datos.....	36
3.2 Conjunto de Datos Estandarizado y Confiable.....	37
3.3 Base de Datos Relacional para Trazabilidad Completa.....	38
3.4 Modelo Clasificador binario para predecir el Cumplimiento del Takt Time.....	39
3.5 Sistema de Métricas Automatizado para la Toma de Decisiones.....	39
3.6 Panel de Visualización Interactivo y Accesible.....	39
3.7 Solución Validada y Ajustada a la Operación Real.....	45
4. CONCLUSIONES Y RECOMENDACIONES.....	46
4.1. Conclusiones.....	46
4.2. Limitaciones y Recomendaciones.....	46
BIBLIOGRAFÍA.....	48
LINKOGRAFÍA.....	49
ANEXOS.....	50

ÍNDICE DE TABLAS

1. Diseño Metodológico

19

ÍNDICE DE FIGURAS

1. Figura 1: Distribución de Tiempos por Proceso	37
2. Figura 2: Modelo relacional de la Base de Datos (MySQL)	38
3. Figura 3: Pestaña Cronología y Distribución	40
4. Figura 4: Pestaña Análisis por Correlativo	41
5. Figura 5: Pestaña Análisis por Proceso	42
6. Figura 6: Pestaña Análisis por Operario	43
7. Figura 7: Pestaña Evolución Temporal	44
8. Figura 8: Pestaña Eficiencia por Operarios	45

1. ANTECEDENTES GENERALES

1.1. Introducción

El presente informe final expone el desarrollo del proyecto Axis Flow, una iniciativa orientada a mejorar la trazabilidad y eficiencia del proceso de producción de baños modulares en la empresa Axis Desarrollos Constructivos, ubicada en la Región de Los Lagos. Axis es una organización que ha impulsado la industrialización de la construcción mediante procesos repetibles y controlados; sin embargo, gran parte de sus registros operativos aún se realizan de manera manual o en planillas dispersas, lo que genera problemas de calidad, fragmentación de la información y dificultades para analizar el rendimiento real de los procesos productivos.

Nuestro proyecto surge como respuesta a esta necesidad, abordando el desafío de transformar datos provenientes de registros manuales en información estructurada, confiable y analíticamente útil, con el fin de apoyar la toma de decisiones y la mejora continua. Para ello, se desarrolló un trabajo integral que incluyó el análisis exploratorio de datos (EDA) (Tukey, 1977), la depuración y normalización de los registros operativos, el diseño y refactorización de una base de datos relacional, además del desarrollo de un prototipo funcional, que permita visualizar indicadores de producción en forma centralizada.

La ejecución del proyecto se basó en la metodología CRISP-DM (Chapman et al., 2000), permitiendo avanzar de manera iterativa a través de fases de entendimiento del negocio, exploración y preparación de datos, modelado y evaluación. Asimismo, se integraron herramientas de programación, análisis de datos y visualización, empleando Python, pandas (McKinney, 2017), scikit-learn y entornos colaborativos que facilitaron el trabajo del equipo.

Este informe presenta los fundamentos del proyecto, su desarrollo técnico, los resultados obtenidos y el impacto esperado tanto para la empresa como para nuestra formación profesional. Constituye el cierre formal del trabajo realizado, integrando aspectos analíticos, tecnológicos y metodológicos propios de la Ingeniería en Informática.

1.2. Descripción del Proyecto

1.2.1. Tema

El proyecto Axis Flow se centra en la optimización del proceso de producción de baños modulares en la empresa Axis, mediante la transformación de registros operacionales manuales en un sistema organizado, trazable y orientado al análisis. El proyecto aborda la necesidad de estructurar, depurar y consolidar la información productiva para mejorar la visibilidad del rendimiento, identificar cuellos de botella y apoyar la toma de decisiones con datos confiables.

A través del análisis de datos, la creación de una base de datos relacional y el desarrollo de una aplicación de dashboards interactivos, el proyecto busca modernizar la gestión de los procesos productivos, integrando herramientas de la Ingeniería en Informática para fortalecer la eficiencia y control operativo dentro de la empresa.

1.2.2. Áreas de Desempeño

El proyecto Axis Flow se desarrolla en diversas áreas de desempeño propias de la Ingeniería en Informática. En primer lugar, abarca el análisis y gestión de datos, incluyendo la exploración inicial (EDA), la limpieza, transformación y estandarización de los registros de producción entregados por la empresa.

También se enmarca en el área de diseño, construcción y administración de bases de datos, donde se elaboró un modelo relacional que permite asegurar integridad, consistencia y trazabilidad de la información operativa, solucionando problemas derivados de registros manuales e inconsistencias en el ingreso de los datos.

Asimismo, el proyecto involucra el desarrollo de herramientas de apoyo analítico, mediante la implementación de scripts en Python y el diseño de una interfaz en Streamlit para visualizar datos, generar métricas relevantes y facilitar la interpretación de los procesos productivos.

Finalmente, el trabajo integra el área de gestión de proyectos TI, contemplando planificación, coordinación del equipo, iteraciones con el cliente, organización mediante herramientas colaborativas y adopción de principios de metodologías ágiles (Schwaber, 2004).

1.2.3. Competencias o Unidades de Competencias

El desarrollo del proyecto Axis Flow permitió aplicar y evidenciar diversas competencias disciplinares propias de la Ingeniería en Informática. Entre las principales, destacamos:

• Análisis y gestión de datos

Competencia asociada a la capacidad de recopilar, limpiar, transformar y analizar información para obtener conocimiento útil.

Esta competencia se expresó en:

- La exploración de los datos (EDA),
- La identificación de inconsistencias,
- Implementación de metodologías de análisis de datos (CRISP-DM),
- La depuración y estandarización de los registros entregados por Axis.

• Diseño y administración de bases de datos

Vinculada al modelamiento, construcción y mantenimiento de sistemas de almacenamiento de información.

Se evidenció mediante:

- El diseño del modelo relacional,
- La creación y refactorización de tablas,
- Creación de vistas de,
- La definición de claves, relaciones y restricciones,
- La preparación de la base para consultas analíticas.

• Programación y desarrollo de soluciones informáticas

Relacionada con la implementación de herramientas que procesan datos y apoyen la toma de decisiones.

Esta competencia se aplicó en:

- El desarrollo de scripts en Python,
- Cálculos de tiempos, métricas y transformaciones bajo estadística descriptiva,
- Desarrollo de aplicación de escritorio bajo el motor de un navegador web,
- Ejecución de procesos automatizados para análisis,

• Gestión de proyectos TI

Asociada a la capacidad de planificar, organizar, coordinar y evaluar proyectos tecnológicos.

Se expresó en:

- La elaboración del plan de trabajo,
- La coordinación del equipo mediante ClickUp,
- Las iteraciones incrementales con el cliente,
- Los ajustes de planificación según disponibilidad de datos.

• Evaluación y aseguramiento de la calidad

Relacionada con la capacidad de validar datos, procesos y resultados.

Fue puesta en práctica mediante:

- La detección y corrección de datos inconsistentes,
- La revisión cruzada entre integrantes,
- La validación de resultados junto al cliente.

1.3. Fundamentación Proyecto APT

1.3.1.Relevancia del proyecto APT

El proyecto Axis Flow es relevante debido a la necesidad concreta de mejorar la gestión, calidad y trazabilidad de los datos operacionales dentro de la empresa Axis, dedicada a la construcción industrializada de baños modulares. Actualmente, gran parte de los registros productivos se realizan de forma manual o mediante planillas dispersas, lo que genera inconsistencias, problemas de calidad y fragmentación de la información (Olson, 2003; English, 1999).

Esta situación afecta directamente la capacidad de la empresa para identificar cuellos de botella, evaluar la eficiencia de sus operarios y gestionar los tiempos de producción, elementos fundamentales dentro de un modelo industrializado que requiere precisión, control y retroalimentación constante. Por ello, el proyecto adquiere relevancia tanto a nivel organizacional como académico, al proponer una solución que transforma datos crudos en información estructurada, confiable y utilizable para la toma de decisiones operativas y estratégicas.

Desde la perspectiva de la Ingeniería en Informática, el proyecto permite aplicar competencias clave relacionadas con el análisis de datos, la construcción de bases de datos relacionales, la estandarización de procesos informáticos y la gestión de proyectos tecnológicos en un entorno real. Esto contribuye al desarrollo profesional del equipo al enfrentar una problemática auténtica del sector productivo, integrando metodologías, herramientas y criterios técnicos propios de la disciplina.

1.3.2. Descripción de Proyecto APT

El proyecto Axis Flow consiste en el análisis, depuración, estructuración y modelamiento de los datos operacionales asociados a la producción de baños modulares en la empresa Axis. Su propósito es transformar registros manuales y planillas desorganizadas en un sistema de información más claro, confiable y utilizable para evaluar el rendimiento de los procesos productivos utilizando la metodología LEAN.

El proyecto abarca un trabajo completo que incluye la exploración y análisis de datos (EDA) (Tukey, 1977; Pyle, 1999), la limpieza y estandarización de los registros históricos, el diseño y construcción de una base

de datos relacional que permita mejorar la trazabilidad de los procesos, y el desarrollo de herramientas en Python para generar métricas, cálculos automatizados y visualización de información relevante.

A través de estas etapas, se busca fortalecer la comprensión de los tiempos de producción, identificar patrones, inconsistencias y cuellos de botella, y entregar a la empresa un sistema estructurado que facilite la toma de decisiones basadas en datos. Todo el desarrollo fue realizado siguiendo principios de metodologías ágiles (Schwaber, 2004) y empleando tecnologías propias de la disciplina, como Python, pandas, SQL y Streamlit.

El proyecto apunta a mejorar la calidad de los datos, asegurar su integridad y disponer de información que permita analizar el desempeño real de los procesos productivos dentro de un entorno industrializado. Así mismo, fomentar la responsabilidad compartida de la estandarización y recolección de datos. Una buena cultura de datos se refleja en la colaboración de todos los niveles de la empresa, donde cada miembro aporta información consistente y oportuna que permite un análisis más efectivo y la mejora continua.

1.3.3. Pertinencia del proyecto con el perfil de egreso

El proyecto Axis Flow se encuentra plenamente alineado con el perfil de egreso de la carrera de Ingeniería en Informática, ya que permite aplicar de manera directa diversas competencias disciplinares que forman parte del quehacer profesional del área. El desarrollo del proyecto requirió integrar conocimientos y habilidades relacionadas con el análisis y gestión de datos, la construcción y administración de bases de datos, la programación de soluciones tecnológicas, y la gestión de proyectos TI, todas dimensiones esenciales del perfil de egreso.

En primer lugar, la pertinencia se evidencia en el trabajo de exploración, limpieza y estandarización de datos, actividades que demandan pensamiento analítico, rigurosidad técnica y capacidad de evaluar la calidad de la información, competencias propias del ámbito del análisis de datos. De igual manera, el diseño y refactorización de la base de datos relacional demuestran dominio en el modelamiento de información, integridad referencial y estructura lógica, aspectos íntimamente ligados al diseño de sistemas informáticos.

Asimismo, el uso de Python y herramientas analíticas para generar métricas, automatizar cálculos e interpretar los procesos productivos refleja la aplicación práctica de habilidades de programación

orientadas a resolver problemas reales. Finalmente, la organización del trabajo, el uso de metodologías ágiles, la comunicación con el cliente y la coordinación mediante herramientas colaborativas como ClickUp permiten evidenciar competencias asociadas a la gestión de proyectos de tecnologías de información, la planificación y el trabajo en equipo.

1.3.4. Relación con los intereses profesionales

El proyecto Axis Flow se vincula directamente con nuestros intereses profesionales, ya que nos permite profundizar en áreas que consideramos fundamentales para nuestro desarrollo como futuros ingenieros en informática. Dentro de estos intereses se encuentran el análisis de datos, la programación orientada al procesamiento de información, y la gestión de proyectos de tecnologías de información, todos elementos ampliamente abordados durante el desarrollo del proyecto.

La necesidad de explorar, limpiar y estructurar datos operacionales provenientes de la empresa Axis ofrece una oportunidad real para fortalecer habilidades analíticas y comprender cómo la información puede transformar procesos en entornos productivos reales. Esto se alinea con nuestro interés por el data analysis y la capacidad de generar valor a partir de datos confiables.

Asimismo, el uso de Python, librerías de análisis y herramientas de modelamiento de bases de datos profundiza nuestra motivación por mejorar nuestras capacidades de programación y diseño de soluciones informáticas. Estas actividades aportan experiencia práctica en la automatización de procesos, la estructura de datos y el uso de técnicas que son altamente demandadas en el ámbito profesional actual.

Finalmente, la organización del proyecto con herramientas colaborativas, las iteraciones con el cliente y la adaptación continua de los entregables permiten reforzar nuestro interés por la gestión de proyectos TI, entendida como un proceso dinámico que exige planificación, comunicación, trabajo en equipo y capacidad de respuesta. Todo ello contribuye significativamente a nuestro crecimiento profesional y nos prepara para enfrentar desafíos reales en entornos laborales donde la tecnología y los datos cumplen un rol central.

1.3.5. Fortalezas y debilidades para desarrollar el proyecto APT

El desarrollo del proyecto Axis Flow implicó la integración de diferentes habilidades y competencias, lo que permitió identificar diversas fortalezas y también algunas debilidades que influyeron en el progreso del trabajo.

Fortalezas

Entre las principales fortalezas del equipo se encuentra la capacidad analítica, reflejada en el trabajo realizado durante el proceso de exploración, depuración y estandarización de los datos operacionales entregados por la empresa. Esta capacidad permitió detectar inconsistencias, corregir errores y estructurar la información de manera eficiente para su posterior uso en análisis y modelamiento.

Otra fortaleza relevante fue la capacidad de programación y resolución técnica, evidenciada en el desarrollo de scripts en Python, la automatización de cálculos y la creación de herramientas para la interpretación de los datos productivos. A esto se suma la competencia en diseño y administración de bases de datos, lo cual permitió construir un modelo relacional sólido y coherente con las necesidades reales del entorno.

Asimismo, el equipo demostró buen manejo de la gestión de proyectos TI, organizando las actividades mediante ClickUp, manteniendo comunicación constante y ajustándose a los requerimientos del cliente durante las iteraciones. La colaboración interna y la disposición a aprender nuevas herramientas fueron también elementos claves para sostener el avance del proyecto.

Debilidades

Entre las debilidades identificadas, una de las principales fue la dependencia de la información proporcionada por el cliente, cuya entrega no siempre se realizó en los plazos acordados, generando retrasos en actividades críticas como la validación de datos y la actualización de la base de datos. Esta situación obligó a reprogramar tareas y ajustar el plan de trabajo.

También se evidenció la necesidad de fortalecer la planeación temporal, especialmente al estimar la duración de actividades técnicas que dependían de la calidad de los datos o de iteraciones con el cliente.

En algunos casos, la carga documental superó lo previsto, lo que requirió redistribuir esfuerzos para mantener la coherencia del proyecto.

A pesar de estas debilidades, el equipo logró sortear las dificultades mediante comunicación constante, revisión conjunta de avances y flexibilidad para reorganizar responsabilidades cuando fue necesario. Esto permitió mantener la continuidad del proyecto y cumplir con los objetivos establecidos.

1.4. Planteamiento del problema

La empresa Axis registra sus procesos de producción de baños modulares mediante formatos manuales y planillas dispersas, lo que genera datos incompletos, duplicados, inconsistentes y difíciles de integrar. Esta falta de estandarización dificulta la trazabilidad real de los tiempos de producción, la medición del rendimiento de los operarios y la identificación de cuellos de botella dentro del proceso productivo.

La ausencia de un sistema unificado y confiable provoca que la información utilizada para evaluar el desempeño dependa de registros susceptibles a errores humanos, modificaciones sin control y variaciones en la forma de capturar los datos. Como consecuencia, los equipos de planificación y control de producción no cuentan con una fuente de información sólida que permita analizar con precisión los tiempos reales invertidos, comparar los procesos contra sus tiempos estándar o tomar decisiones basadas en evidencia.

Este problema afecta directamente la capacidad de la empresa para mejorar la eficiencia operativa, optimizar recursos y avanzar en la industrialización de la construcción, donde la precisión en los datos es fundamental para asegurar procesos repetibles, controlados y medibles.

Por ello, se hace necesaria la generación de una solución que permita unificar, limpiar, estructurar y analizar los datos operacionales, facilitando la obtención de métricas confiables y habilitando un sistema de información que apoye la toma de decisiones dentro de un entorno productivo cada vez más demandante.

1.5. Objetivos

1.5.1.Objetivo General

Desarrollar una solución que permita a la empresa Axis registrar, organizar y gestionar sus datos operacionales de forma confiable y sin inconsistencias, mediante la depuración y estructuración de la información productiva, con el fin de asegurar una trazabilidad completa de los procesos y generar métricas que faciliten la toma de decisiones informadas.

1.5.2.Objetivos Específicos

- Realizar un análisis exploratorio de los datos (EDA) para identificar inconsistencias, duplicidades y patrones relevantes dentro de los registros operacionales.
- Depurar, estandarizar y transformar la información productiva entregada por la empresa Axis, asegurando su calidad, coherencia y usabilidad.
- Desarrollar un modelo predictivo clasificador en base al tiempo teórico de los procesos de la empresa.
- Diseñar y construir una base de datos relacional que permita almacenar los procesos productivos de manera estructurada, garantizando integridad y trazabilidad.
- Implementar herramientas en Python que automaticen cálculos, generen métricas y faciliten la interpretación de los tiempos de producción y desempeño de los procesos.
- Desarrollar visualizaciones y reportes analíticos que permitan comprender el comportamiento de los procesos y apoyar la toma de decisiones basada en datos.
- Validar la estructura de la base de datos, las métricas generadas y los análisis obtenidos mediante iteraciones con el cliente y revisión interna del equipo.

2. DISEÑO METODOLÓGICO

Tabla del diseño metodológico.

DISEÑO METODOLÓGICO			
Objetivos específicos	Etapas		Actividades
Realizar un análisis exploratorio de datos (EDA).	N°1	Exploración y diagnóstico inicial	Revisión de los archivos entregados por Axis.- Identificación de inconsistencias, duplicados y valores faltantes.- Análisis preliminar de variables y relaciones.- Detección de problemas de formato y estructura.
Depurar, estandarizar y transformar los datos operacionales.	N°2	Limpieza y normalización de datos	Eliminación de duplicidades.- Corrección de registros erróneos.- Conversión y estandarización de formatos de fecha, texto y tiempos.- Reglas de transformación y homologación de campos.
Diseñar y construir una base de datos relacional.	N°3	Modelamiento y estructuración	Diseño del modelo entidad-relación.- Definición de llaves primarias y relaciones.- Creación y refactorización de tablas.- Migración de datos limpios al modelo final.
Construir y evaluar un modelo predictivo basado en Machine Learning	N°4	Modelado predictivo (Machine Learning)	Definición del problema de modelado.- Selección de la variable objetivo

			<p><i>(Cumple_TT).- División del dataset en entrenamiento y prueba (train/test).- Ingeniería de características en variables temporales, estructurales y operacionales.-</i></p> <p><i>Preprocesamiento mediante pipelines (One-hot Encoding, escalado y manejo de nulos).-</i></p> <p><i>Entrenamiento del modelo (Regresión Logística).-</i></p> <p><i>Obtención de métricas clave (Accuracy, Precision, Recall, F1-score, ROC-AUC).-</i></p> <p><i>Interpretación del modelo y evaluación de impacto para la toma de decisiones.</i></p>
Implementar herramientas analíticas en Python para generar métricas.	N°5	<i>Desarrollo técnico y procesamiento automatizado</i>	<p>Construcción de scripts en Python (pandas).- Cálculo automatizado de tiempos reales y comparativos.-</p> <p>Generación de métricas de desempeño.- Validación de resultados con el cliente.</p>
Desarrollar visualizaciones y reportes analíticos.	N°6	<i>Interpretación y visualización</i>	<p><i>Construcción de reportes con Streamlit.- Creación de gráficos y resúmenes descriptivos.-</i></p> <p><i>Análisis de KPIs clave.- Ajustes basados en retroalimentación del cliente.</i></p>
Validar los entregables mediante iteraciones con el cliente.	N°7	<i>Revisión y ajuste</i>	<p>Reuniones con el cliente para validar resultados.- Ajustes en BD, métricas y reportes.-</p>

			Documentación de cambios.- Preparación del informe final.
--	--	--	--

Fuente: Elaboración propia.

2.1. ETAPA N°1 – Exploración y diagnóstico inicial

2.1.1. Objetivo de la etapa

El objetivo de esta etapa fue analizar el estado inicial de los datos operacionales entregados por la empresa Axis, evaluando su estructura, calidad y completitud con el fin de identificar inconsistencias, duplicidades y problemas de formato que pudieran afectar el desarrollo del proyecto.

2.1.2. Descripción de la etapa

Durante esta etapa se realizó una revisión exhaustiva de los archivos de registro vinculados a la producción de baños modulares. Se analizaron las planillas y formatos utilizados por la empresa para documentar tiempos y actividades, con el propósito de comprender las características del proceso y los criterios de captura de información. Esta exploración permitió diagnosticar la calidad del conjunto de datos y definir los lineamientos para su limpieza y estandarización en etapas posteriores.

2.1.3. Actividades realizadas

- Revisión de los archivos proporcionados por la empresa Axis.
- Identificación de campos relevantes y estructura general de los registros.
- Detección de valores faltantes, duplicidades y errores de digitación.
- Análisis de variaciones en formatos de fecha, texto y tiempos.
- Organización preliminar de los datos para su evaluación.
- Registro de hallazgos y criterios para la limpieza posterior.

2.1.4. Procedimiento

El procedimiento aplicado en esta etapa incluyó el uso de herramientas de análisis inicial como Microsoft Excel y Python (pandas) para inspeccionar filas, columnas, tipos de datos y patrones irregulares. Se realizaron análisis descriptivos básicos para evidenciar las inconsistencias presentes y se establecieron reglas de transformación para asegurar que los datos pudieran ser procesados correctamente en la siguiente fase.

Además, se documentaron los errores detectados y se elaboró un diagnóstico general del estado de la información, el cual sirvió como base para la etapa de limpieza y estandarización de los datos.

2.2. ETAPA N°2 – Limpieza y Normalización de Datos

2.2.1. Objetivo de la etapa

El objetivo de esta etapa fue depurar, estandarizar y normalizar los datos operacionales entregados por Axis, con el fin de construir un conjunto de información coherente, confiable y apto para análisis posteriores (Pyle, 1999). Esta limpieza permitió eliminar inconsistencias, corregir errores y asegurar que la base de datos representa con precisión los procesos productivos reales.

2.2.2. Descripción de la etapa

Durante esta etapa trabajamos sobre la data cruda proporcionada por la empresa, la cual presentaba errores de formato, valores faltantes, registros duplicados y nomenclaturas inconsistentes entre actividades, fechas y responsables.

El proceso incluyó la revisión exhaustiva de cada variable, la definición de reglas de estandarización y la aplicación de criterios validados en conjunto con el cliente.

Asimismo, tras la tercera iteración con Axis, fue necesario incorporar ajustes adicionales en la forma en que se registraban las actividades y tiempos de producción, por lo que la normalización se extendió a nivel estructural para asegurar rastreabilidad completa de los procesos.

2.2.3. Actividades realizadas

- Identificación de valores atípicos, inconsistentes o fuera de rango.
- Eliminación de duplicados detectados en los registros originales.
- Corrección de formatos de fecha, hora y campos numéricos.
- Homologación de nombres de actividades, responsables y etapas productivas.
- Estándar único para la medición de tiempos reales y tiempos planificados.
- Tratamiento de datos faltantes (eliminación o imputación según criterio).
- Revisión cruzada entre las hojas del archivo para detectar discrepancias.
- Documentación de todas las reglas de limpieza aplicadas.

2.2.4. Procedimiento

El procedimiento se desarrolló principalmente con Python y librerías asociadas, aplicando funciones específicas para la depuración y normalización.

Las acciones principales incluyeron:

- Uso de `drop_duplicates()` para eliminar registros repetidos.
- Conversión de formatos mediante `to_datetime()` y reglas manuales cuando fue necesario.
- Normalización de texto utilizando funciones de minúsculas, reemplazos y diccionarios de homologación.
- Reasignación de valores mediante condiciones controladas para corregir actividades mal registradas.
- Validación de intervalos de tiempo reales, descartando valores imposibles o invertidos.
- Integración de nuevas columnas que permitieron ordenar y clasificar correctamente cada etapa productiva.

Todos estos pasos fueron verificados en conjunto con los requerimientos entregados por Axis en la tercera reunión, garantizando que la estructura final reflejara fielmente el proceso de fabricación.

2.3. ETAPA N°3 – Modelamiento y Estructuración

2.3.1. Objetivo de la etapa

El objetivo de esta etapa fue diseñar y construir una base de datos relacional que permitiera integrar de manera estructurada y coherente los datos limpios obtenidos anteriormente, garantizando la trazabilidad de los procesos productivos y habilitando el análisis eficiente de métricas clave.

2.3.2. Descripción de la etapa

Durante esta etapa se llevó a cabo el modelamiento conceptual, lógico y físico de la base de datos aplicando principios de modelado relacional y normalización (Elmasri & Navathe, 2016). Se definieron las entidades principales del proceso productivo, sus atributos y las relaciones entre ellas.

La base se diseñó siguiendo principios de normalización para asegurar integridad, eliminar redundancias y permitir integraciones futuras. Tras la validación del modelo con el cliente, se implementaron ajustes que permitieron mejorar la trazabilidad del proceso productivo, especialmente en las actividades asociadas a tiempos de ejecución, operarios y correlativos de baño.

Finalmente, se crearon las tablas definitivas y se migraron los datos previamente limpiados, garantizando coherencia entre los registros y estructura relacional.

2.3.3. Actividades realizadas

- Identificación de entidades, atributos y relaciones presentes en el proceso productivo.
- Diseño del Modelo Entidad–Relación (MER) inicial del proyecto.
- Definición de llaves primarias, llaves foráneas y cardinalidades entre tablas.
- Aplicación de normas de normalización para evitar redundancia y dependencias parciales.
- Construcción física de las tablas en el motor de base de datos escogido.
- Refactorización del esquema tras observaciones del cliente en la tercera iteración.
- Migración de los datos previamente depurados hacia la nueva estructura.
- Creación de vistas en la base de datos.
- Validación de integridad referencial y consistencia interna de la base.

2.3.4. Procedimiento

El procedimiento contempló:

1. Modelamiento conceptual:
 - Identificación de entidades clave: baños, actividades, tiempos, operarios, correlativos.
 - Mapeo inicial del proceso completo de fabricación.
2. Modelamiento lógico:
 - Determinación de tipos de datos, llaves primarias y relaciones.
 - Elaboración del MER y su versión lógica normalizada.
3. Modelamiento físico:
 - Implementación del modelo en el motor de base de datos mySQL.
 - Creación de tablas mediante instrucciones SQL y validación de integridad.
4. Refactorización posterior:
 - Ajuste de tablas para mejorar la trazabilidad.
 - Integración de nuevas columnas solicitadas por Axis.
 - Corrección de relaciones detectadas como débiles o ambiguas.
5. Migración de datos:
 - Carga controlada de datos limpios desde archivos externos.
 - Verificación de claves, duplicados y consistencia referencial.

Todo el proceso fue desarrollado aplicando buenas prácticas de ingeniería de datos y validado en conjunto con el cliente para garantizar que el modelo representara fielmente la operación productiva real.

2.4. ETAPA N°4 – Modelado predictivo

2.4.1. Objetivo de la etapa

Definir, preparar y entrenar un modelo de clasificación binaria capaz de predecir si un proceso específico de la construcción de un baño cumplirá o no su tiempo objetivo (takt time). La meta es obtener un modelo reproducible, evaluable y basado en variables operacionales confiables.

2.4.2. Descripción de la etapa

La etapa comprende la formulación del problema de modelado, la selección de la variable objetivo, la identificación de las variables explicativas, la preparación del conjunto de datos y la implementación de un pipeline de preprocesamiento unido a un modelo de regresión logística. Incluye el manejo de valores faltantes, la codificación de variables categóricas, el escalado de variables numéricas y la partición estratificada de los datos para asegurar una evaluación representativa del desempeño.

2.4.3. Actividades realizadas

- Definición del problema como una clasificación binaria de cumplimiento del takt time.
- Selección de la variable objetivo (Cumple_TT) y análisis del leve desbalance existente entre clases.
- Elección de variables categóricas y numéricas relevantes, excluyendo identificadores específicos para evitar sobreajuste.
- Construcción de la matriz de características y del vector objetivo.
- Tratamiento de valores faltantes en variables categóricas y numéricas.
- Partición estratificada del dataset en entrenamiento (80 %) y prueba (20 %).
- Implementación de un ColumnTransformer que aplica One-Hot Encoding y escalado estándar.
- Integración del preprocesamiento y el modelo en un pipeline único.
- Entrenamiento del modelo LogisticRegression con balanceo automático de clases.

2.4.4. Procedimiento

1. Formular el problema de predicción del cumplimiento del TT como clasificación binaria.

-
2. Definir la variable objetivo y analizar su distribución para identificar desbalances.
 3. Seleccionar las variables explicativas pertinentes, evitando atributos altamente específicos.
 4. Construir X e y, y aplicar el manejo de valores faltantes según el tipo de variable.
 5. Dividir los datos mediante una partición estratificada en conjuntos de entrenamiento y prueba.
 6. Configurar el pipeline:
 - One-Hot Encoding para variables categóricas.
 - StandardScaler para variables numéricas.
 7. Añadir al pipeline el modelo de regresión logística con `class_weight="balanced"`.
 8. Entrenar el pipeline completo utilizando `X_train` e `y_train`.

2.5. ETAPA N°5 – Desarrollo técnico y procesamiento automatizado

2.5.1. Objetivo de la etapa

El objetivo de esta etapa fue implementar herramientas analíticas en Python que permitieran automatizar el cálculo de tiempos, generar métricas de desempeño y obtener información procesable a partir de los datos almacenados en la base de datos relacional, con el fin de apoyar el análisis y la toma de decisiones en la empresa Axis.

2.5.2. Descripción de la etapa

En esta etapa se desarrollaron distintos scripts en Python, utilizando principalmente la librería pandas (McKinney, 2017), para procesar los datos depurados y estructurados en la base de datos.

El trabajo se centró en:

- calcular tiempos reales de ejecución,
- compararlos con tiempos de referencia o tiempos esperados,
- generar indicadores de desempeño a nivel de proceso y producción, además de preparar la información para su posterior visualización.

Durante esta fase también se realizaron instancias de validación con el cliente, contrastando los resultados obtenidos con la experiencia operativa y los datos históricos, lo que permitió ajustar reglas de cálculo y asegurar coherencia entre el modelo analítico y la realidad del proceso productivo.

2.5.3. Actividades realizadas

- Construcción de scripts en Python para la lectura y procesamiento de los datos.
- Cálculo de tiempos reales por actividad, por baño y por correlativo.
- Generación de métricas comparativas entre tiempos reales y tiempos planificados o de referencia.
- Creación de indicadores de desempeño asociados a procesos productivos.
- Preparación de datos agregados para ser consumidos por herramientas de visualización.

-
- Validación de los resultados obtenidos en reuniones con el cliente.
 - Ajuste de fórmulas y reglas de cálculo en base a la retroalimentación recibida.

2.5.4. Procedimiento

El procedimiento seguido incluyó los siguientes pasos:

1. Conexión y carga de datos
 - Lectura de los datos desde los archivos preparados o desde la base de datos.
 - Importación de las tablas relevantes utilizando pandas.
2. Transformaciones y cálculos
 - Uso de operaciones sobre columnas para transformar tiempos y formatos.
 - Cálculo de diferencias entre tiempos teóricos y tiempos reales.
 - Agrupación por baño, proceso, fecha u operario para generar vistas resumidas.
3. Generación de métricas
 - Construcción de indicadores de desempeño por proceso.
 - Cálculo de promedios, mínimos, máximos y distribuciones de tiempos.
 - Identificación de posibles cuellos de botella a partir de los resultados.
4. Validación y ajuste
 - Presentación de los resultados al cliente para validar su interpretación.
 - Corrección de reglas de cálculo cuando se detectaron discrepancias.
 - Documentación de las fórmulas y criterios utilizados.

2.6. ETAPA N°6 – Interpretación y Visualización

2.6.1. Objetivo de la etapa

El objetivo de esta etapa fue desarrollar visualizaciones y reportes analíticos que permitieran interpretar los datos procesados, analizar los KPIs relevantes del proceso productivo y presentar los resultados de manera clara, accesible y comprensible tanto para el equipo como para el cliente.

2.6.2. Descripción de la etapa

En esta etapa se implementaron distintas visualizaciones y reportes mediante herramientas como Streamlit y bibliotecas de gráficos en Python respetando principios de diseño de dashboards (Few, 2013), con el fin de transformar las métricas calculadas en representaciones visuales útiles para la toma de decisiones.

Se generaron gráficos descriptivos, resúmenes tabulares y comparaciones entre tiempos reales y tiempos planificados, permitiendo identificar variaciones, cuellos de botella y comportamientos relevantes del proceso productivo.

Además, se incorporaron instancias de retroalimentación con el cliente, ajustando las visualizaciones iniciales para asegurar que respondieran a las necesidades operativas y facilitaran la interpretación de la información.

2.6.3. Actividades realizadas

- Construcción de reportes interactivos utilizando Streamlit.
- Creación de gráficos descriptivos (líneas, barras, histogramas y comparativos de tiempos).
- Representación visual de KPIs clave como:
 - Tiempo real vs. tiempo planificado,
 - Duración promedio por actividad,
 - Rendimiento operativo,

-
- Variaciones en tiempos entre baños o procesos.
 - Preparación de resúmenes analíticos y cuadros comparativos.
 - Ajustes de visualizaciones basados en la retroalimentación del cliente durante reuniones.
 - Validación de interpretaciones y coherencia visual con los datos procesados.

2.6.4. Procedimiento

El procedimiento aplicado fue el siguiente:

1. Selección de KPIs relevantes
 - En conjunto con el cliente se definieron los indicadores más críticos para evaluar el proceso productivo.
2. Preparación de los datos para visualización
 - Uso de pandas para generar subconjuntos de datos agregados o filtrados.
 - Organización de la información en estructuras utilizables para gráficos.
3. Construcción de visualizaciones
 - Uso de bibliotecas como matplotlib y seaborn para gráficos.
 - Implementación de interfaces simples mediante Streamlit para presentar resultados.
4. Interpretación y ajuste
 - Revisión de cada gráfico para validar su utilidad.
 - Corrección y ajuste de colores, escalas, etiquetas y títulos.
 - Inclusión de comparativos y tendencias sugeridas por el cliente.

2.7. ETAPA N°7 – Revisión y Ajuste

2.7.1. Objetivo de la etapa

El objetivo de esta etapa fue validar los entregables desarrollados —base de datos, métricas analíticas y visualizaciones preliminares— mediante iteraciones con el cliente, incorporando ajustes técnicos y metodológicos para asegurar que los resultados fueran coherentes con la operación real y respondieran a las necesidades del proceso productivo de Axis.

2.7.2. Descripción de la etapa

En esta fase se realizaron diversas instancias de revisión con el cliente para contrastar los resultados obtenidos con los conocimientos operativos del proceso de producción de baños modulares.

Durante estas reuniones, se identificaron inconsistencias en los datos originales, se levantaron nuevos requerimientos y se ajustaron tanto la estructura de la base de datos como las métricas construidas en Python.

2.7.3. Actividades realizadas

- Reuniones periódicas con el cliente para presentar avances.
- Validación de los resultados de la limpieza de datos y detección de inconsistencias.
- Ajustes en la base de datos relacional según comentarios del cliente.
- Corrección y mejora de métricas analíticas generadas en Python.
- Revisión de visualizaciones y reportes para adaptarlos al lenguaje operativo de Axis.
- Registro de observaciones, decisiones y cambios aplicados.
- Preparación de material de apoyo para el informe final.

2.7.4. Procedimiento

El procedimiento seguido incluyó:

1. Presentación de resultados preliminares

- Se mostraron al cliente las métricas, la estructura de la base de datos y las visualizaciones generadas.

2. Recepción de retroalimentación

- El cliente revisó la coherencia de los valores con su experiencia y detectó datos incorrectos o incompletos en los registros originales.

3. Aplicación de ajustes

- Modificación del modelo de datos para mejorar su trazabilidad.
- Corrección de reglas de cálculo de tiempos reales y planificados.
- Ajustes en visualizaciones para hacerlas más interpretables.

4. Validación final

- Revisión detallada con el cliente de los cambios aplicados.
- Confirmación de que los resultados representaban fielmente su proceso productivo en conjunto con jefa de producción.

5. Documentación

- Preparación de anexos técnicos para el informe final.

3. RESULTADOS

3.1 Diagnóstico Integral del Estado de los Datos

Se identificó y documentó con precisión el estado inicial de la información, detectando problemas críticos de calidad como valores incompletos, registros duplicados, inconsistencias en la nomenclatura de actividades y variaciones en los formatos de tiempo. Este diagnóstico sentó una base objetiva y técnica para todas las acciones de depuración posteriores.

3.2 Conjunto de Datos Estandarizado y Confiable

Mediante un riguroso proceso de limpieza, se obtuvo un conjunto de datos coherente y normalizado. Se eliminaron duplicidades, se corrigieron formatos de fecha y hora, y se homologaron los nombres de actividades y responsables (antes 2295 registros, después 2218 registros, es decir, pasamos de 77 a 74 baños). Como resultado, se aumentó significativamente la calidad de los registros y se eliminaron ambigüedades, permitiendo una representación fidedigna del proceso productivo.

- **Datos atípicos de los procesos:** En el gráfico se muestra una Distribución de Tiempos por Proceso. Se visualizan valores atípicos en la mayoría de los procesos (**puntos rojos**)

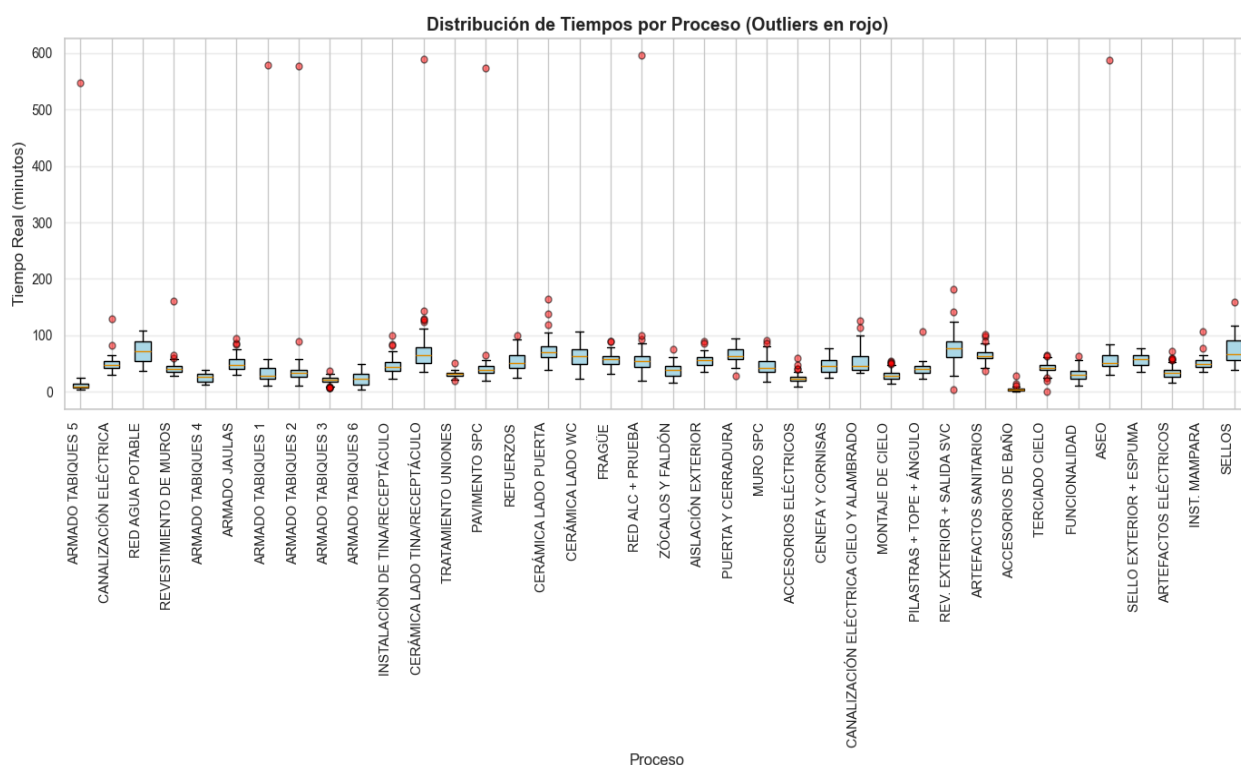


Figura 1: Distribución de Tiempos por Proceso

3.3 Base de Datos Relacional para Trazabilidad Completa

Se diseñó e implementó una base de datos relacional normalizada que integró de manera estructurada todas las entidades del proceso (baños, actividades, tiempos y operarios). Este modelo asegura la integridad de la información, elimina redundancias siguiendo buenas prácticas de diseño de bases de datos (Elmasri & Navathe, 2016), y lo más importante, proporciona una trazabilidad completa para seguir cada baño desde su inicio hasta la finalización de todas sus actividades.

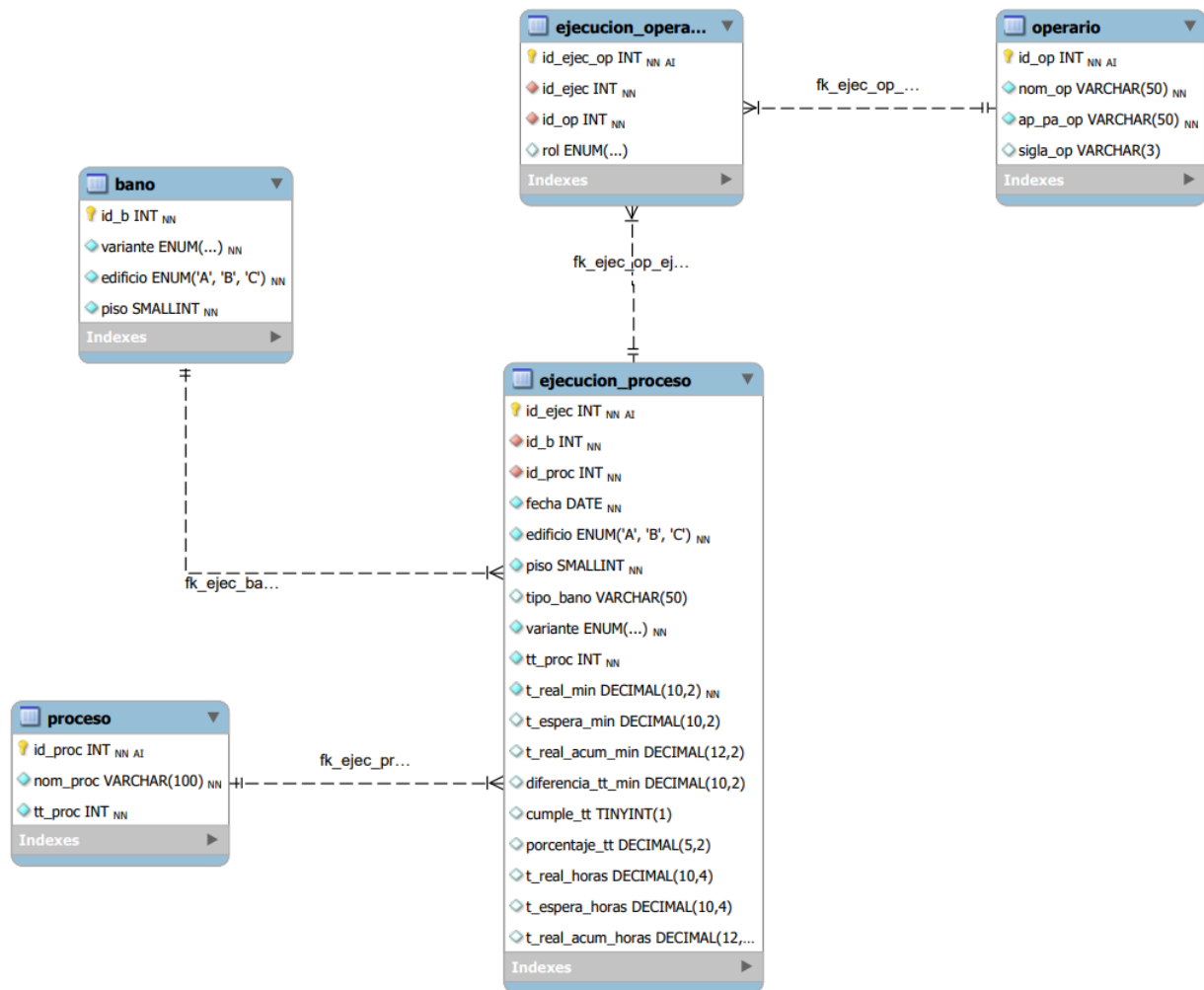


Figura 2: Modelo relacional de la Base de Datos (MySQL)

3.4 Modelo Clasificador binario para predecir el Cumplimiento del Takt Time

En el conjunto de prueba (444 registros) el modelo de regresión logística mostró un desempeño sólido: accuracy $\approx 0,79$ y ROC-AUC $\approx 0,87$, indicando buena capacidad de discriminación; la matriz de confusión muestra 128 verdaderos negativos y 35 falsos positivos sobre 163 no-cumplimientos reales, y 224 verdaderos positivos y 57 falsos negativos sobre 281 cumplimientos reales; por clase, la clase 0 (no cumple) presenta precisión $\approx 0,69$, recall $\approx 0,79$ y F1 $\approx 0,74$, mientras que la clase 1 (cumple) tiene precisión $\approx 0,86$, recall $\approx 0,80$ y F1 $\approx 0,83$; en práctica, el modelo identifica la mayoría de los incumplimientos y predice cumplimientos con alta precisión, aunque existen falsos negativos y falsos positivos cuyo impacto operacional debe evaluarse según el costo asociado a cada tipo de error.

3.5 Sistema de Métricas Automatizado para la Toma de Decisiones

Se desarrollaron scripts en Python que automatizan el cálculo de tiempos de ejecución y generan métricas clave de desempeño. Esto permite comparar los tiempos reales con los planificados, identificar cuellos de botella y analizar la eficiencia operativa con base en datos procesados y validados, facilitando un análisis continuo y objetivo empleando rutinas de análisis de datos en Python (McKinney, 2017).

3.6 Panel de Visualización Interactivo y Accesible

Con el propósito de facilitar la interpretación operativa y mejorar la toma de decisiones, se desarrolló un panel interactivo de visualización utilizando la plataforma Streamlit. Este panel permite a los usuarios explorar de forma dinámica la información relacionada con la fabricación de baños prefabricados, presentando los datos de manera clara,

La interfaz fue diseñada para que incluso personas sin conocimientos técnicos puedan comprender el desempeño productivo, revisar procesos específicos y analizar tendencias con total autonomía. Las visualizaciones integradas abarcan indicadores clave, flujos de trabajo y métricas detalladas por proceso, operativo y correlativo.

La aplicación se estructura en distintas secciones navegables mediante una barra superior, donde cada pestaña agrupa funcionalidades específicas:

- **Cronología y Distribución:** muestra un diagrama tipo Gantt que permite visualizar la secuencia temporal de los procesos de fabricación, identificando tiempos reales, tiempos de espera y progresión por correlativo.

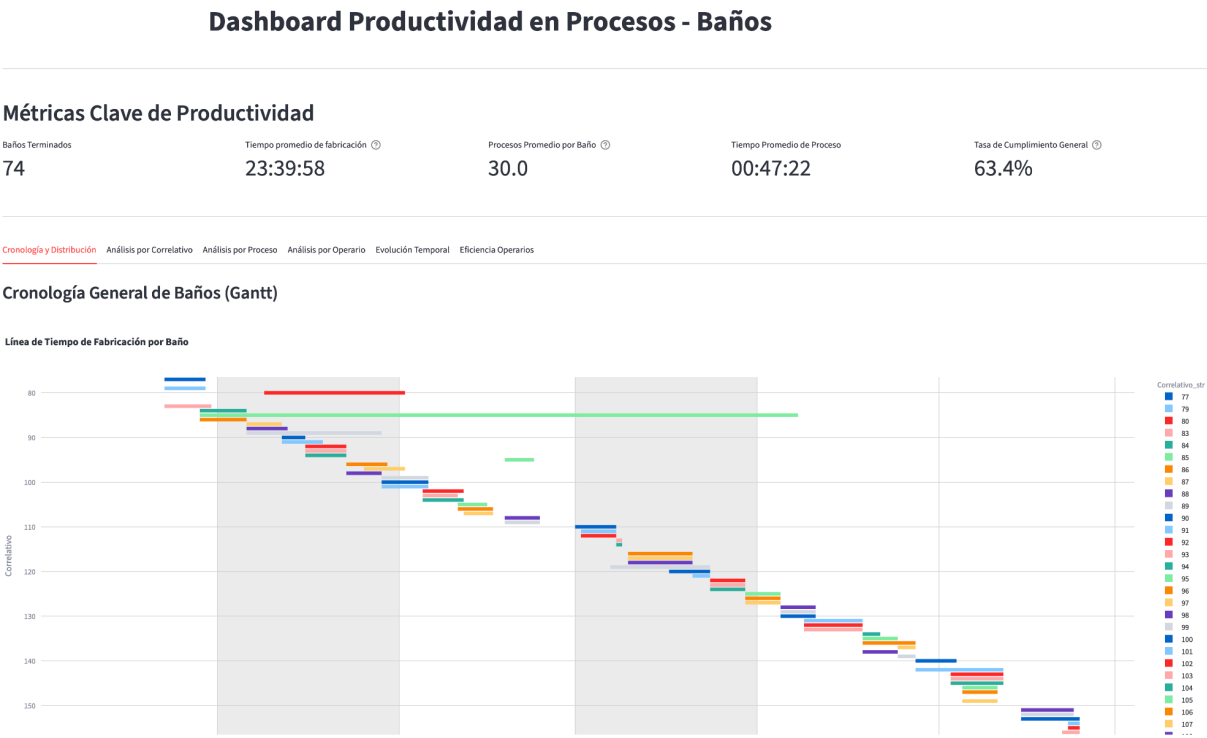


Figura 3: Pestaña Cronología y Distribución

- **Análisis por Correlativo:** permite seleccionar un baño específico y revisar sus métricas particulares, como duración total, tiempos de espera, número de procesos y operarios involucrados. La tabla de procesos se colorea automáticamente según el cumplimiento del Takt Time, facilitando una lectura inmediata de desviaciones.

Cronología y Distribución **Análisis por Correlativo** Análisis por Proceso Análisis por Operario Evolución Temporal Eficiencia Operarios

Análisis Individual por Correlativo

Seleccione un Correlativo para analizar

77

Métricas para el Correlativo: 77

Tipo de Baño

B6

Fecha de Inicio

2025-04-22

Fecha de Fin

2025-04-29

Tiempo Real Total

33:31:27

Tiempo de Espera Total

05:58:22

Número de Procesos

36

Operarios Involucrados: AF, CA, DS, JA, JC, JCO, MC, PH, RH, SN

Secuencia de Procesos

	Fecha	Proceso	T. Real	T. Espera	Takt Time	Cumple_TT	Operarios
0	2025-04-22	ARMADO TABIQUES 6	00:10:07	00:00:00	00:20:00	<input checked="" type="checkbox"/>	JCO
1	2025-04-22	ARMADO TABIQUES 5	00:08:48	00:00:00	00:20:00	<input checked="" type="checkbox"/>	CA
2	2025-04-22	ARMADO TABIQUES 4	00:24:28	00:00:00	00:20:00	<input type="checkbox"/>	CA
3	2025-04-22	ARMADO TABIQUES 3	00:26:49	00:00:00	00:20:00	<input type="checkbox"/>	JCO
4	2025-04-22	ARMADO TABIQUES 2	00:28:03	00:00:00	00:20:00	<input type="checkbox"/>	CA
5	2025-04-22	ARMADO TABIQUES 1	00:40:36	00:00:00	00:20:00	<input type="checkbox"/>	JCO
6	2025-04-22	ARMADO JALUISAS	01:24:09	00:34:36	01:00:00	<input type="checkbox"/>	CA
7	2025-04-22	REVESTIMIENTO DE MUROS	00:45:12	00:07:00	01:00:00	<input checked="" type="checkbox"/>	SN
8	2025-04-22	CANALIZACIÓN ELÉCTRICA	01:02:34	00:00:00	01:00:00	<input type="checkbox"/>	MC
9	2025-04-22	RED AGUA POTABLE	01:24:55	00:45:00	01:00:00	<input type="checkbox"/>	JC

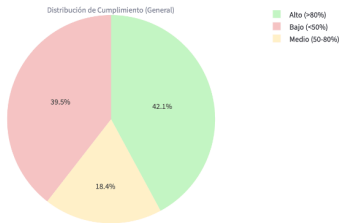
Figura 4: Pestaña Análisis por Correlativo

- **Análisis por Proceso:** presenta métricas agregadas como tasas de cumplimiento, tiempos promedio y cantidad de ejecuciones por tipo de proceso, además de gráficos circulares que resumen la distribución del desempeño general.

Cronología y Distribución Análisis por Correlativo **Análisis por Proceso** Análisis por Operario Evolución Temporal Eficiencia Operarios

Métricas Generales por Proceso

	Proceso	Tasa_Cumplimiento	Cantidad	Tiempo_Promedio	TT_Promedio
0	TERCIADO CIELO	0.0%	61	00:43:06	00:30:00
1	REV. EXTERIOR + SALIDA SVC	20.0%	65	01:17:48	01:00:00
2	SELLO EXTERIOR + ESPUMA	20.0%	29	00:56:00	00:40:00
3	ARMADO TABIQUES 2	20.0%	44	00:33:18	00:20:00
4	CERÁMICA LADO PUERTA	20.0%	71	01:13:48	01:00:00
5	ARMADO TABIQUES 1	20.0%	43	00:31:42	00:20:00
6	ARMADO TABIQUES 3	30.0%	43	00:20:30	00:20:00
7	ARMADO TABIQUES 4	30.0%	43	00:26:18	00:20:00
8	ARMADO TABIQUES 6	30.0%	28	00:25:42	00:20:00
9	ARTEFACTOS ELÉCTRICOS	30.0%	61	00:36:00	00:30:00



Análisis de Cumplimiento por Proceso - Tipo B1

	Proceso	Tasa_Cumplimiento	Cantidad	Tiempo_Promedio	TT_Promedio
0	ASLACIÓN EXTERIOR	0.0%	1	01:52:24	01:00:00
1	ARMADO TABIQUES 6	0.0%	1	00:44:48	00:20:00
2	ARTEFACTOS SANITARIOS	0.0%	1	01:02:12	01:00:00
3	REV. EXTERIOR + SALIDA SVC	0.0%	1	01:07:06	01:00:00
4	CENEFA Y CORNISAS	0.0%	1	01:10:36	01:00:00
5	CERÁMICA LADO PUERTA	0.0%	1	01:31:06	01:00:00
6	RED AGUA POTABLE	0.0%	1	01:30:00	01:00:00
7	FRAGÜE	0.0%	1	01:05:18	01:00:00
8	TERCIADO CIELO	0.0%	1	00:40:18	00:30:00
9	PUERTA Y CERRADURA	0.0%	1	01:17:12	01:00:00

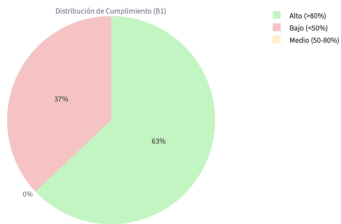


Figura 5: Pestaña Análisis por Proceso

- **Análisis por Operario:** ofrece gráficos comparativos sobre participación, tiempos reales promedio y porcentaje de cumplimiento del Takt Time para cada operario, permitiendo identificar fortalezas individuales y oportunidades de estandarización.

Cronología y Distribución Análisis por Correlativo Análisis por Proceso **Análisis por Operario** Evolución Temporal Eficiencia Operarios

Análisis de Participación por Operario

Agrupar por:

☒ Correlativo

☐ Tipo de Baño

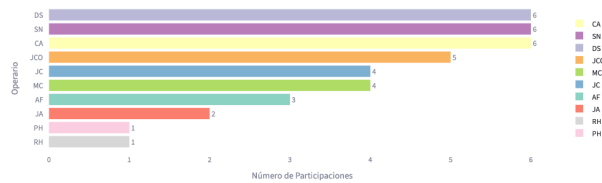
Seleccione un Correlativo

77

Participación en Correlativo 77

Total de Procesos: 36

Número de Participaciones por Operario



Porcentaje de Participación por Operario

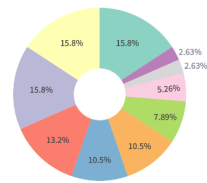


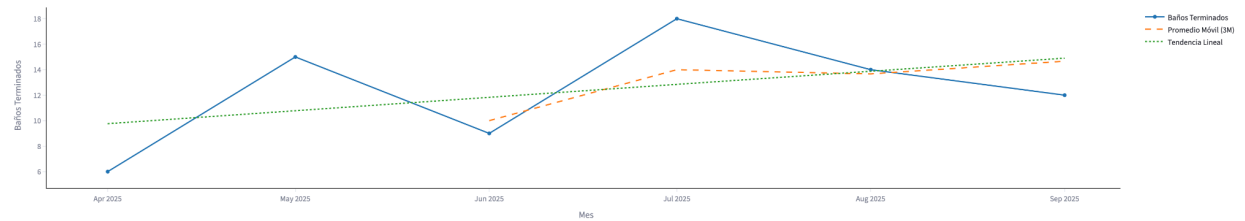
Figura 6: Pestaña Análisis por Operario

- **Evolución Temporal:** muestra tendencias mensuales de baños terminados, promedios móviles y variaciones del Lead Time diario, permitiendo evaluar la evolución del sistema productivo a lo largo del tiempo.

Cronología y Distribución Análisis por Correlativo Análisis por Proceso Análisis por Operario **Evolución Temporal** Eficiencia Operarios

Evolución de Productividad y Ciclo

Unidades Producidas (Baños Terminados) por Mes



Tiempo de Ciclo (Lead Time) Promedio Diario

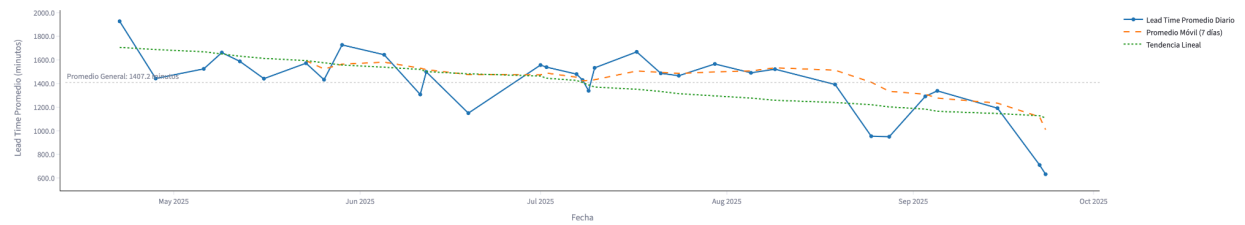
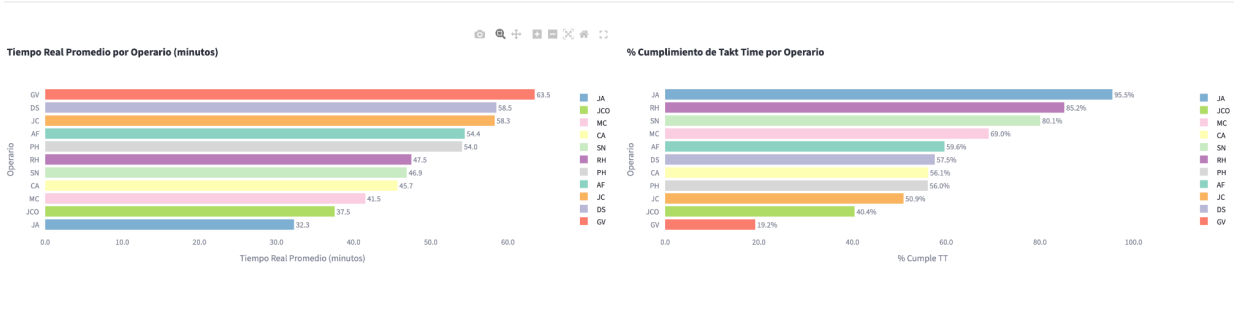


Figura 7: Pestaña Evolución Temporal

- **Eficiencia por Operarios:** detalla métricas de desempeño individuales y por tipo de baño, integrando gráficos de barras que permiten comparar tiempos reales promedio y porcentajes de cumplimiento.

Análisis de Eficiencia por Operario

Métricas Generales por Operario



Métricas por Tipo de Baño

Tipo de Baño: B1



Figura 8: Pestaña Eficiencia por Operarios

3.7 Solución Validada y Ajustada a la Operación Real

Gracias a un proceso iterativo de revisión con el cliente, se refinaron todos los componentes del proyecto (desde la estructura de la base de datos hasta las métricas y visualizaciones) para garantizar su total coherencia con la operación real de Axis. El resultado final es una solución robusta, técnicamente sólida y alineada con las necesidades específicas del proceso productivo del cliente.

4. CONCLUSIONES Y RECOMENDACIONES

4.1. Conclusiones

El proyecto Axis Flow permitió ordenar y dar sentido a un conjunto de registros que, al inicio, estaban dispersos, con formatos distintos y con varios problemas de calidad. A partir de la metodología CRISP-DM (Shearer, 2000), se avanzó de manera sistemática desde la comprensión del negocio y de los datos hasta la preparación, el modelamiento, la evaluación y despliegue de una solución analítica, logrando como resultado una base de datos relacional que refleja con claridad el proceso real de producción de baños modulares en Axis Desarrollos Constructivos. Esta base de datos se entrega a la empresa con su estructura definida y documentada, incluyendo las tablas principales, sus relaciones y un conjunto de vistas orientadas a las necesidades del área de producción, de modo que el acceso a la información relevante sea directo y no dependa de conocimientos técnicos avanzados. La automatización de cálculos en Python facilitó la obtención de tiempos reales, métricas comparativas e indicadores de desempeño, reduciendo el trabajo manual y haciendo más consistente el análisis. Las visualizaciones y reportes desarrollados en Streamlit entregan a la empresa una mirada concreta sobre los tiempos de producción, las diferencias frente a lo planificado y los puntos donde se concentran las principales demoras, aportando insumos directos para la toma de decisiones. Desde la formación profesional, el proyecto permitió aplicar de forma integrada conocimientos de análisis y gestión de datos, modelamiento de bases de datos, machine learning, programación, aseguramiento de la calidad y trabajo con un cliente real, fortaleciendo la capacidad del equipo para abordar proyectos similares en contextos productivos.

4.2. Limitaciones y Recomendaciones.

Los resultados obtenidos deben leerse considerando ciertas limitaciones propias del contexto y de la información disponible. La calidad de los análisis está ligada a la calidad de los datos de origen, tal como se advierte en la literatura sobre calidad de datos (Olson, 2003; English, 1999), que en varios casos presentaban vacíos, inconsistencias y registros incompletos que, aun después del proceso de limpieza, acotan la profundidad de algunas conclusiones y obligan a interpretar los indicadores con cautela. El alcance del trabajo se concentró en un grupo de procesos y periodos definidos en conjunto con la empresa, por lo que la incorporación de otros procesos o etapas históricas podría requerir ajustes

adicionales tanto en las reglas de depuración como en el diseño de la base de datos. Además, la solución desarrollada corresponde a un prototipo analítico funcional, pensado para explorar y visualizar información más que para operar como sistema productivo integrado; su uso a gran escala dentro de la infraestructura tecnológica de Axis exige una fase posterior de ingeniería de software, pruebas, robustecimiento y despliegue gradual.

A partir de la experiencia del proyecto y de las necesidades detectadas, se recomienda a Axis Desarrollos Constructivos avanzar en una mejora progresiva de la forma en que se capturan y registran los datos en terreno, utilizando formatos más estructurados o herramientas digitales que disminuyan errores de digitación y aseguren que la información clave quede siempre registrada. También resulta conveniente evaluar la integración de la base de datos y de los procesos de cálculo desarrollados con sistemas internos de gestión o planificación, de modo que la actualización de los datos sea más automática y la información sea coherente en toda la organización. Una vez que la calidad de los registros se mantenga de forma estable, tiene sentido ampliar el conjunto de indicadores e incorporar análisis más específicos por tipo de baño, turno, equipo o cualquier otra dimensión relevante, e incluso explorar modelos predictivos que permitan estimar tiempos de producción y anticipar retrasos. Por último, es recomendable que los dashboards y reportes generados se incorporen a las rutinas habituales del área de producción, con responsables claramente definidos, espacios formales para revisar los indicadores y un esquema de mantenimiento de la base de datos y de los scripts, de manera que el trabajo realizado se mantenga vigente y pueda seguir evolucionando.

BIBLIOGRAFÍA

- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0: Step-by-step data mining guide*. SPSS Inc.
- Shearer, C. (2000). The CRISP-DM model: The new blueprint for data mining. *Journal of Data Warehousing*, 5, 13-22.
- Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. En *Proceedings of the 4th International Conference on the Practical Application of Knowledge Discovery and Data Mining* (pp. 29-39).
- Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley.
- Pyle, D. (1999). *Data preparation for data mining*. Morgan Kaufmann.
- Olson, J. E. (2003). *Data quality: The accuracy dimension*. Morgan Kaufmann.
- Elmasri, R., & Navathe, S. B. (2016). *Fundamentals of database systems* (7.^a ed.). Pearson.
- English, L. P. (1999). *Improving data warehouse and business information quality: Methods for reducing costs and increasing profits*. John Wiley & Sons.
- McKinney, W. (2017). *Python for data analysis: Data wrangling with pandas, NumPy, and IPython* (2.^a ed.). O'Reilly Media.
- Few, S. (2013). *Information dashboard design: Displaying data for at-a-glance monitoring* (2.^a ed.). Analytics Press.
- Schwaber, K. (2004). *Agile project management with Scrum*. Microsoft Press.

LINKOGRAFÍA

Python Software Foundation. Python documentation. Disponible en:

<https://docs.python.org/3/>

The pandas development team. *pandas documentation*. Disponible en:

<https://pandas.pydata.org/docs/>

scikit-learn developers. *scikit-learn: Machine learning in Python*. Disponible en:

<https://scikit-learn.org/>

Streamlit Inc. *Streamlit documentation*. Disponible en:

<https://docs.streamlit.io/>

Matplotlib development team. *Matplotlib documentation*. Disponible en:

<https://matplotlib.org/>

Seaborn contributors. *seaborn: statistical data visualization*. Disponible en:

<https://seaborn.pydata.org/>

Oracle Corporation. *MySQL Documentation*. Disponible en:

<https://dev.mysql.com/doc/>

Microsoft Corporation. *Microsoft Excel*. Disponible en:

<https://www.microsoft.com/en-us/microsoft-365/excel>

ClickUp. *ClickUp – The everything app for work*. Disponible en:

<https://clickup.com/>

Axis Desarrollos Constructivos S.A. *Sitio oficial de Axis Desarrollos Constructivos*. Disponible en:

<http://www.axisdc.cl>

ANEXOS

ANEXO 1: analisis_axisflow.ipynb

ANEXO 2: Base de Datos - AXIS FLOW.docx

ANEXO 3: app_streamlit.py

ANEXO 4: Script INSERT INTO.ipynb

ANEXO 5: axis_flow_tables.sql