PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR
FACULTAD DE HÁBITAT, INFRAESTRUCTURA Y CREATIVIDAD
CARRERA DE SISTEMAS



PROYECTO DE TITULACIÓN

Tema: DISEÑO DE UNA ARQUITECTURA EMPRESARIAL BASADA EN TOGAF INTEGRADA CON CRISP-DM PARA FORTALECER LA TOMA DE DECISIONES BASADA EN DATOS EN UNA EMPRESA DEL SECTOR TEXTIL.

AUTOR:
PABLO GABRIEL JARRIN MERA

TUTOR:
NELSON ESTEBAN SALGADO REYES

QUITO DM, JUNIO DE 2026

Dedicatoria

A mis padres porque sin ellos no estaría culminando este sueño, por su constante apoyo y motivación.

A mi hermana por permitirme ser su ejemplo a seguir, por nunca dejar de creer en mí y acompañarme en cada paso.

A todas las personas que nunca dejaron de apoyarme incluso en los momentos más difíciles.

Dedico este logro a Dios que me permitió recorrer este camino lleno de desafíos para crecer como profesional, pero, sobre todo como ser humano.

Agradecimiento

Le estoy eternamente agradecido a mis padres, porque sin ellos no habría entendido la responsabilidad que tengo conmigo mismo y con mi futuro. Me enseñaron la importancia de ser grato con las oportunidades que se me presentan.

A toda mi familia por nunca dejar de creer en mí, especialmente a mi hermana por su apoyo incondicional, por enseñarme que las caídas no definen a las personas, las vuelven más fuertes.

A ella, que vio mis errores y me eligió por ello mismo 🫧.

Agradezco a todas aquellas personas que formaron parte de este proceso, presentes o no, sumaron en mi vida más de lo que les podría retornar.

Resumen

El presente proyecto de titulación busca diseñar una estructura empresarial que combine la metodología CRISP-DM con TOGAF para mejorar los procesos en una empresa textil. Se enfoca en una empresa ecuatoriana que necesita mejorar su análisis de datos para tomar decisiones estratégicas, especialmente en cuanto a la satisfacción de los clientes después de la venta. Para desarrollar este proyecto, se utilizó la metodología CRISP-DM. Primero, se comprendió el negocio y se identificó la importancia de analizar las calificaciones que los clientes dan después de la venta. Luego, se prepararon los datos utilizando un conjunto de encuestas que incluían variables como la fecha de la encuesta, la calificación y el canal de venta. Con estos datos, se creó una serie de tiempo para predecir el comportamiento futuro de las calificaciones.

En la fase de modelado, se aplicaron técnicas de series de tiempo, como los modelos SARIMAX y Prophet, para estimar si las calificaciones futuras subirían, bajarían o se mantendrían. Los resultados mostraron que es probable que las calificaciones disminuyan, lo que indica que la empresa debe mejorar sus procesos de atención y seguimiento. Con los resultados de CRISP-DM, se desarrolló una propuesta de estructura empresarial utilizando el ciclo ADM de TOGAF. Esta propuesta permitió convertir los hallazgos analíticos en soluciones dentro de los dominios de negocio, datos, aplicaciones y tecnología.

En conclusión, este proyecto aporta una propuesta estructurada para fortalecer la gestión de la satisfacción del cliente y promover decisiones empresariales basadas en datos. Combinar CRISP-DM con TOGAF permite unir el análisis de datos con la estructura empresarial, lo que facilita que los resultados predictivos se conviertan en mejoras continuas para los procesos organizacionales.

Tabla de Contenidos

1. CAPÍTULO I: INTRODUCCIÓN 13

1.1 Tema 13

1.2 Justificación 13

1.3 Planteamiento del Problema 14

1.4 Objetivos 16

1.4.1 Objetivo General 16

1.4.2 Objetivos específicos 16

1.5 Alcance 16

2. CAPÍTULO II: MARCO TEÓRICO CONCEPTUAL 18

2.1 Antecedentes 18

2.2 Marco Conceptual 18

2.2.1 Arquitectura Empresarial 18

2.2.2 TOGAF 20

2.2.3 Método ADM de TOGAF 21

2.2.4 Dominios de TOGAF 23

2.2.5 Arquitectura Empresarial orientada por datos 24

2.2.6 Ciencia de datos 27

2.2.7 Minería de Datos 28

2.2.8 Modelos Predictivos 30

2.2.9 Evaluación de modelos predictivos 32

2.2.10 CRISP-DM 33
2.3 Integración de CRISP-DM y TOGAF para la toma de decisiones empresariales 34
2.3.1 Datos históricos 34
2.3.2 Aplicación de la metodología CRISP-DM 34
2.3.3 Generación de modelo predictivo 35
2.3.4 Integración de resultados al ciclo ADM de TOGAF 36
2.3.5 Construcción de Arquitectura TO-BE 37
2.3.6 Apoyo a la toma de decisiones empresariales 37
3. CAPÍTULO III: ARQUITECTURA EMPRESARIAL DE PONTE SELVA 38
3.1 Arquitectura de Negocio 38
3.1.1 Nivel Estratégico: 38
3.1.2 Nivel Operativo 39
3.1.3 Nivel de Apoyo 39
3.1.4 Arquitectura Actual (AS-IS) 40
3.1.5 Stakeholders y Actores 41
3.1.6 Mapa de Procesos 41
3.1.7 Proceso Comercial y Postventa 44
3.1.8 Cadena de Valor de Porter 45
3.1.9 Diagramas de Casos de uso 46
3.2 Arquitectura de Información 51
3.2.1 Catálogo de Información 51

3.2.2 Diagrama de Flujo de Información 51
3.3 Arquitectura de Aplicaciones 51
3.3.1 Catálogo de Aplicaciones 51
3.3.2 Diagrama de Integración de Aplicaciones 52
3.3.3 Matriz Aplicación – Proceso 52
3.4 Arquitectura Tecnológica 52
3.4.1 Inventario Tecnológico 52
3.4.2 Infraestructura Física 53
3.4.3 Arquitectura de red 54
3.5 Arquitectura de Seguridad 54
3.5.1 Seguridad física 54
3.5.2 Seguridad Lógica 54
3.5.3 Diagrama de Seguridad 55
3.6 Arquitectura Objetivo (TO-BE) 56
3.7 Arquitectura de Seguridad y Gobierno de Datos 57
3.7.1 57
3.8 Análisis GAP de la Arquitectura Empresarial 57
4. CAPÍTULO IV: APLICACIÓN DE LA METODOLOGÍA CRISP-DM 59
4.1 Desarrollo de aplicación web en Streamlit para CRISP-DM 59
4.1.1 Configuración de la Aplicación web 59
4.1.2 Fase 1 de CRISP-DM en aplicación web 60

4.1.3 Fase 2 de CRISP-DM en aplicación web 61
4.1.4 Fase 3 de CRISP-DM en aplicación web 62
4.1.5 Preparación de Serie de Tiempo 63
4.1.6 Fase 4 de CRISP-DM en aplicación web 64
4.1.7 Fase 5 de CRISP-DM en aplicación web 65
4.1.8 Fase 6 de CRISP-DM en aplicación web 66
4.2 Fase 1 CRISP-DM: Comprensión del negocio 67
4.3 Fase 2 CRISP-DM: Comprensión de los Datos 68
4.3.1 Descripción general de los datos 68
4.3.2 Diccionario y Calidad de los datos 69
4.3.3 Exploración inicial de variable objetivo 69
4.3.4 Exploración temporal en función del Score 70
4.4 Fase 3 CRISP-DM: Preparación de los datos 70
4.4.1 Limpieza y aplicación de reglas 70
4.4.2 Distribución de variable objetivo después de limpieza 71
4.4.3 Interpolación de la serie 72
4.5 Fase 4 CRISP-DM: Modelado 72
4.5.1 Aplicación del modelo 72
4.6 Fase 5 CRISP-DM: Evaluación 73
4.7 Fase 6 CRISP-DM: Despliegue 74
4.8 Relación de los resultados de CRISP-DM con TOGAF 74

5. CAPÍTULO V: DISEÑO DE ARQUITECTURA EMPRESARIAL TOGAF/ADM. 76

5.1 Relación entre CRISP-DM y TOGAF/ADM 76

5.2 Aplicación del ciclo ADM de TOGAF 76

5.2.1 Fase preliminar 76

5.2.2 Fase A: Visión de Arquitectura 77

5.2.3 Fase B: Arquitectura de Negocio 77

5.2.4 Fase C: Arquitectura de Datos (sistemas de información) 78

5.2.5 Fase D: Arquitectura Tecnológica 79

5.2.6 Fase E: Oportunidades y Soluciones 79

5.2.7 Fase F: Planificación de la Migración 80

5.2.8 Fase G: Gobierno de la Implementación 81

5.2.9 Fase H: Gestión de cambios de Arquitectura 82

6. CAPÍTULO VI: CONCLUSIONES Y RECOMENDACIONES 84

6.1 Conclusiones 84

6.1.1 Conclusión general 84

6.1.2 Conclusiones específicas 84

6.2 Recomendaciones 86

7. BIBLIOGRAFÍA: 87

Contenido de Tablas

Tabla 1: Fases de ADM. Fuente: Elaboración Propia 22

Tabla 2: Catálogo de Información. Fuente: Elaboración Propia 50

Tabla 3: Catálogo de Aplicaciones. Fuente: Elaboración Propia 50

Tabla 4: Matriz Aplicación - Proceso. Fuente: Elaboración Propia 51

Tabla 5: GAP entre arquitecturas. Fuente: Elaboración propia 57

Tabla 6: KPI del objetivo estratégico. Fuente: Elaboración propia 66

Tabla 7. Fuente: Elaboración Propia 72

Tabla 8. Fuente: Elaboración Propia 72

Tabla 9: Relación CRISP-DM Y TOGAF. Fuente: Elaboración Propia 75

Tabla 10: Limitaciones y procesos objetivos. Fuente: Elaboración Propia 77

Tabla 11: Componentes de la arquitectura tecnológica. Fuente: Elaboración Propia 78

Tabla 12: Planificación de la migración. Fuente: Elaboración Propia 79

Tabla 13: Architecture Roadmap. Fuente: Elaboración Propia 80

Contenido de Figuras

Figura 1. Fuente: https://albatian.com/media/filer_public_thumbnails/filer_public/2b/b8/2bb879b1-1d69-400b-9984-5defbe7a8dba/pr-1703-articulo5_ea-esp.jpg__1170x0_q85_subsampling-2_upscale.jpg 18
Figura 2. Fuente: https://www.sapbwconsulting.com/hs-fs/hubfs/ig_04a582b0c60c5c8a0168f8ab96faf4819bb2758d24295310f1.png?width=1024&height=1024&name=ig_04a582b0c60c5c8a0168f8ab96faf4819bb2758d24295310f1.png20
Figura 3. Fuente: https://chae201521700912580.wordpress.com/wp-content/uploads/2015/10/adm-v3.png 21
Figura 4. Fuente: https://cdn.sanity.io/images/dnt4jhwg/production/8af553ad67a4c7a5c1b4aabc5b14335d4ba5ed4c-720x540.webp 24
Figura 5. Fuente: https://cdn.prod.website-files.com/671603568e40dcb4a09fce45/67459c9121818a87c59e9567_673334fa817bad004287163d_Datos.jpeg 28
Figura 6. Fuente: https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcSqJIE7LVDIeSD6Y6h3fqdBb_wI0DuKJTt2DA&s 30
Figura 7. Fuente: https://i.ytimg.com/vi/aUwXWGa8jK0/maxresdefault.jpg 31
Figura 8. Fuente: https://substackcdn.com/image/fetch/$s_!Wo71!,f_auto,q_auto:good,fl_progressive:steep/https%3A%2F%2Fsubstack-post-media.s3.amazonaws.com%2Fpublic%2Fimages%2Fa76e4b10-6391-40b4-b4df-092e4e661b03_1200x600.png 35
Figura 9: AS-IS Empresa Escogida. Fuente: Elaboración Propia 39
Figura 10: Mapa de Procesos. Fuente: Elaboración Propia 42
Figura 11: Proceso Comercial y Postventa. Fuente: Elaboración Propia 43
Figura 12: Cadena de Valor de Porter. Fuente: Elaboración Propia 44
Figura 13. Fuente: Elaboración Propia 45
Figura 14. Fuente: Elaboración Propia 46
Figura 15. Fuente: Elaboración Propia 47
Figura 16. Fuente: Elaboración Propia 48
Figura 17. Fuente: Elaboración Propia 49
Figura 18. Fuente: Elaboración Propia 50
Figura 19. Fuente: Elaboración Propia 51
Figura 20. Fuente: Elaboración Propia 52
Figura 21. Fuente: Elaboración Propia 53
Figura 22. Fuente: Elaboración Propia 54
Figura 23. Fuente: Elaboración Propia 55
Figura 24. Fuente: Elaboración Propia 56
Figura 25. Fuente: Elaboración Propia 58

Figura 26. Fuente Elaboración Propia 59
Figura 27. Fuente: Elaboración Propia 60
Figura 28. Fuente: Elaboración Propia 61
Figura 29. Fuente: Elaboración Propia 62
Figura 30. Fuente: Elaboración Propia 63
Figura 31. Fuente: Elaboración Propia 64
Figura 32. Fuente: Elaboración Propia 65
Figura 33. Fuente: Elaboración propia 67
Figura 34. Fuente: Elaboración propia 68
Figura 35. Fuente: Elaboración Propia 68
Figura 36. Fuente: Elaboración Propia 69
Figura 37. Fuente: Elaboración Propia 69
Figura 38. Fuente: Elaboración Propia 70
Figura 39. Fuente: Elaboración Propia 71
Figura 40. Fuente: Elaboración Propia 71
Figura 41. Fuente: Elaboración Propia 72
Figura 42. Fuente: Elaboración Propia 79
Figura 43. Fuente: Elaboración Propia 81
Figura 44. Fuente: Elaboración Propia 81

1. CAPÍTULO I: INTRODUCCIÓN

1.1 Tema

Diseño de una Arquitectura Empresarial basada en TOGAF integrada con CRISP-DM para fortalecer la toma de decisiones basada en datos en una empresa del sector textil.

1.2 Justificación

En un entorno empresarial altamente dinámico y competitivo, las organizaciones enfrentan constantes desafíos relacionados con la optimización de sus procesos, la adaptación al mercado y la toma de decisiones basada en datos. En este contexto, la ausencia de una visión que articule la estrategia del negocio con el uso eficiente de la información limita la capacidad de las empresas para responder de manera oportuna a estos cambios, evidenciando la necesidad de nuevos enfoques trabajando de manera estructurada que permitan mejorar su desempeño organizacional.

La presente investigación adquiere relevancia al proponer la integración de la metodología CRISP-DM con un enfoque de Arquitectura Empresarial basado en TOGAF. Por un lado, CRISP-DM permite comprender el negocio, analizar los datos y generar modelos predictivos que aportan conocimiento valioso para la toma de decisiones. Por otro lado, TOGAF proporciona un marco metodológico robusto para diseñar, planificar e implementar arquitecturas empresariales alineadas con los objetivos estratégicos de la organización. La combinación de ambos enfoques sugiere una alternativa que permite no solo analizar la situación actual de la empresa, sino también orientar la transformación de sus procesos a partir de evidencia.

Por otro lado, referente a la factibilidad de este estudio, debido a la disponibilidad de marcos metodológicos ampliamente documentados como TOGAF y CRISP-DM, así como al acceso a herramientas tecnológicas que facilitan el análisis de datos y el modelado de procesos empresariales se considera viable el desarrollo de esta investigación. Además, el

enfoque propuesto puede ser aplicado de manera adaptable a diferentes organizaciones del sector, lo que incrementa su alcance y utilidad.

1.3 Planteamiento del Problema

La empresa Ponte Selva, perteneciente al sector textil ecuatoriano, cuenta con una amplia trayectoria y procesos productivos bien definidos que abarcan desde la hilatura, tejeduría y tintorería, hasta la distribución de sus productos a nivel nacional. La empresa dispone de información histórica proveniente de procesos comerciales y postventa, sin embargo, estos datos no se transforman sistemáticamente en conocimiento que permita anticipar tendencias de satisfacción del cliente ni orientar decisiones estratégicas dentro de una arquitectura empresarial integrada

El problema principal radica en la falta de integración entre el análisis de datos y la arquitectura empresarial, lo que limita la capacidad de la organización para tomar decisiones estratégicas basadas en información predictiva. Aunque la empresa dispone de controles de calidad y procesos establecidos, no se identifican mecanismos formales que permitan transformar los datos generados en conocimiento útil para prever comportamientos futuros del negocio.

De este problema central se derivan varios problemas secundarios. En primer lugar, la inexistencia de modelos predictivos que permitan anticipar resultados en procesos clave como asegurar la satisfacción del cliente. En segundo lugar, la limitada utilización de los datos generados en las distintas áreas para apoyar la toma de decisiones estratégicas. Finalmente, la ausencia de una arquitectura empresarial alineada con el análisis de datos dificulta la identificación de oportunidades de mejora y la adaptación eficiente a las condiciones del mercado.

Ponte Selva posee datos recopilados en base a evaluaciones de su desempeño y estudios de los procesos internos de la empresa, a la vez que de encuestas de satisfacción a sus clientes. De esta manera, crearon una base de datos que se podría utilizarse como herramienta de retroalimentación si se aplican los métodos de análisis pertinentes como CRISP-DM y sus resultados se los trabaja en una metodología estructurada y especializada en manejo de procesos como lo es TOGAF. Es así, que se podría orientar el soporte a la toma de decisiones y fortalecer su competitividad en el sector textil.

Por tanto, referente a esta problemática se plantea la siguiente pregunta central:

- ¿Cómo la integración de CRISP-DM con TOGAF permite transformar información histórica empresarial en capacidades analíticas alineadas con los procesos estratégicos de Ponte Selva?

Como preguntas secundarias se plantea:

- ¿Cómo pueden los datos generados en los procesos internos y en las encuestas de satisfacción de Ponte Selva ser preparados y analizados mediante la metodología CRISP-DM para obtener información relevante?
- ¿Qué modelos predictivos pueden desarrollarse para anticipar el comportamiento de la satisfacción del cliente como indicador estratégico del proceso comercial y postventa?
- ¿Qué impacto tiene la implementación de un modelo que integre CRISP-DM + TOGAF en la eficiencia operativa y competitividad de la empresa?

1.4 Objetivos

1.4.1 Objetivo General

Diseñar una arquitectura empresarial basada en TOGAF integrada con CRISP-DM que permita incorporar capacidades analíticas predictivas para apoyar la toma de decisiones estratégicas en Ponte Selva.

1.4.2 Objetivos específicos

1. Fundamentar teóricamente los conceptos de Arquitectura Empresarial, TOGAF y la metodología CRISP-DM, estableciendo las bases de conocimiento para su aplicación en el sector textil.
2. Desarrollar un modelo de análisis de datos basado en la metodología CRISP-DM aplicado a la empresa Ponte Selva, que permita obtener información relevante y generar predicciones sobre sus procesos clave.
3. Diseñar una arquitectura empresarial basada en TOGAF que integre los resultados del análisis de datos para brindar soporte en la toma de decisiones apoyando los procesos organizacionales de la empresa.
4. Evaluar la arquitectura propuesta mediante indicadores de alineación estratégica, integración tecnológica y generación de valor basada en datos.

1.5 Alcance

El alcance definido para este proyecto de titulación comprende la aplicación de las fases de la metodología CRISP-DM, desde la comprensión del negocio hasta la evaluación de modelos, utilizando datos históricos proporcionados por la empresa, tales como indicadores de desempeño, información de procesos internos y encuestas de satisfacción. A partir de este análisis, se generarán modelos predictivos orientados a procesos, en este caso, la satisfacción del cliente.

En cuanto a la arquitectura empresarial, el alcance contempla el diseño de una propuesta basada en TOGAF, considerando las capas de negocio, datos, aplicaciones y tecnología, integrando los resultados obtenidos del análisis de datos para el soporte en la toma de decisiones. La investigación se limita al diseño y propuesta de la arquitectura empresarial, sin incluir la implementación práctica de la misma.

2. CAPÍTULO II: MARCO TEÓRICO CONCEPTUAL

2.1 Antecedentes

En el campo de la analítica de datos, la metodología CRISP-DM ha sido ampliamente utilizada como un estándar para el desarrollo de proyectos de minería de datos. Diferentes trabajos han demostrado su utilidad en la generación de modelos predictivos aplicados a problemas empresariales como la estimación de la demanda, el control de calidad, el análisis del comportamiento del cliente, entre muchos otros tópicos, permitiendo una toma de decisiones basada en datos.

Recientemente, autores como Schaper (2025) han planteado la necesidad de integrar ambos enfoques, proponiendo una visión unificada en la que CRISP-DM y TOGAF no se consideran metodologías independientes, sino complementarias. En su estudio "Integrating CRISP-DM and TOGAF: A Unified Approach for Enterprise AI Architecture", se expone cómo las fases de CRISP-DM pueden alinearse con las etapas del ADM de TOGAF, logrando conectar el análisis de datos con la arquitectura empresarial y asegurando que las soluciones analíticas estén alineadas con los objetivos del negocio.

Por lo tanto, la presente investigación busca aportar a este campo mediante el diseño de una arquitectura empresarial que integre ambas metodologías, aplicada a la empresa Ponte Selva, con el fin de brindar soporte a la toma de decisiones y fortalecer su competitividad mediante el uso estratégico de los datos.

2.2 Marco Conceptual

2.2.1 Arquitectura Empresarial

La Arquitectura Empresarial o AE en varios materiales bibliográficos, es una disciplina orientada a integrar los principales componentes de una organización, tales como la misión, la visión, los objetivos estratégicos, los procesos de negocio, los datos, las aplicaciones y la infraestructura tecnológica. Su finalidad es lograr que estos elementos funcionen de manera

articulada, permitiendo que la tecnología no sea utilizada de forma aislada, sino como un soporte directo para el cumplimiento de los objetivos organizacionales. En este sentido, Conexiam (2026) habla que la Arquitectura Empresarial es un conjunto de modelos y vistas que explican la construcción de una organización y cómo modificarla de forma eficaz. La arquitectura empresarial constituye una herramienta estratégica para identificar y cerrar brechas entre los objetivos de un negocio y la realidad. Se crea la arquitectura empresarial a fin de guiar un cambio efectivo. TOGAF (The Open Group Architecture Framework) plantea que la Arquitectura Empresarial puede modelarse a través de cuatro dimensiones: negocio, datos, aplicaciones y tecnología.

Estructura de la Arquitectura Empresarial



Figura 1. Fuente: https://albatian.com/media/filer_public_thumbnails/filer_public/2b/b8/2bb879b1-1d69-400b-9984-5defbe7a8dba/pr-1703-articulo5_ea-esp.jpg__1170x0_q85_subsampling-2_upscale.jpg

Para desarrollar una Arquitectura Empresarial de manera ordenada, las organizaciones suelen apoyarse en frameworks o marcos de referencia. Estos proporcionan métodos, herramientas, estándares y un vocabulario común para diseñar diferentes tipos de arquitecturas. De acuerdo con The Open Group (2009, citado en Guerrero et al., 2016), un framework de arquitectura permite estructurar un sistema mediante bloques de construcción,

definiendo cómo estos se relacionan entre sí y qué estándares pueden utilizarse para su implementación.

2.2.2 TOGAF

TOGAF, por sus siglas en inglés The Open Group Architecture Framework, es un marco de referencia para el desarrollo de arquitecturas empresariales. Su propósito principal es proporcionar métodos, herramientas y lineamientos que permitan diseñar, planificar, implementar y gobernar la arquitectura empresarial dentro de una organización. De acuerdo con The Open Group, TOGAF es un framework probado de arquitectura empresarial utilizado por organizaciones líderes a nivel mundial, y se fundamenta en un modelo iterativo apoyado en buenas prácticas y activos reutilizables de arquitectura (The Open Group, 2026).

Una de las características principales de TOGAF es que modela la arquitectura empresarial a través de varios dominios o dimensiones. Estos dominios permiten analizar la organización desde distintas perspectivas: la arquitectura de negocio, que describe la estrategia, la gobernanza, la estructura y los procesos clave; la arquitectura de datos, que define la estructura física y lógica de la información; la arquitectura de aplicaciones, que identifica los sistemas que soportan los procesos del negocio; y la arquitectura tecnológica, que describe la infraestructura de hardware, software y redes necesaria para el funcionamiento organizacional (UNIR, 2021).

Dominios de TOGAF

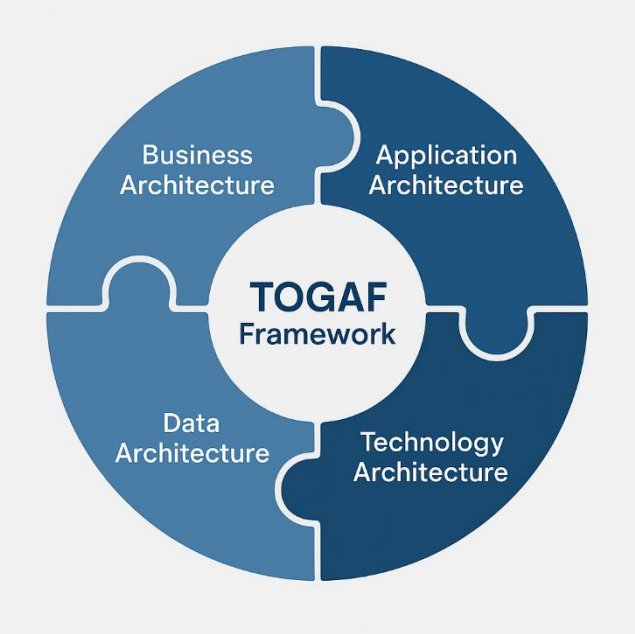


Figura 2. Fuente: https://www.sapbwconsulting.com/hs-fs/hubfs/ig_04a582b0c60c5c8a0168f8ab96faf4819bb2758d24295310f1.png?width=1024&height=1024&name=ig_04a582b0c60c5c8a0168f8ab96faf4819bb2758d24295310f1.png

2.2.3 Método ADM de TOGAF

Por sus siglas, el Architecture Development Method, es el núcleo de la metodología que plantea TOGAF, proporciona una guía estructurada e iterativa para desarrollar, gestionar y mantener una arquitectura empresarial desde su concepción hasta su evolución continua. Según The Open Group (2026), TOGAF se fundamenta en una metodología probada de arquitectura empresarial utilizada por organizaciones líderes a nivel mundial, y el ADM representa el proceso central para aplicar dicho marco en contextos organizacionales reales.

El método ADM está compuesto por una fase preliminar y ocho fases iterativas:

- La fase preliminar permite preparar a la organización para el desarrollo de la arquitectura, definiendo principios, alcance, capacidades y elementos de gobierno.
- La fase de visión de arquitectura establece una perspectiva inicial del cambio esperado y delimita el alcance del proyecto.

- La fase de arquitecturas de negocio, sistemas de información y tecnología, las cuales permiten analizar de manera integral los procesos, datos, aplicaciones e infraestructura tecnológica de la organización.
- Fase en la que se identifican oportunidades y soluciones, se planifica la migración, se gobierna la implementación y se gestiona el cambio arquitectónico de manera continua. (UNIR, 2021).

ADM

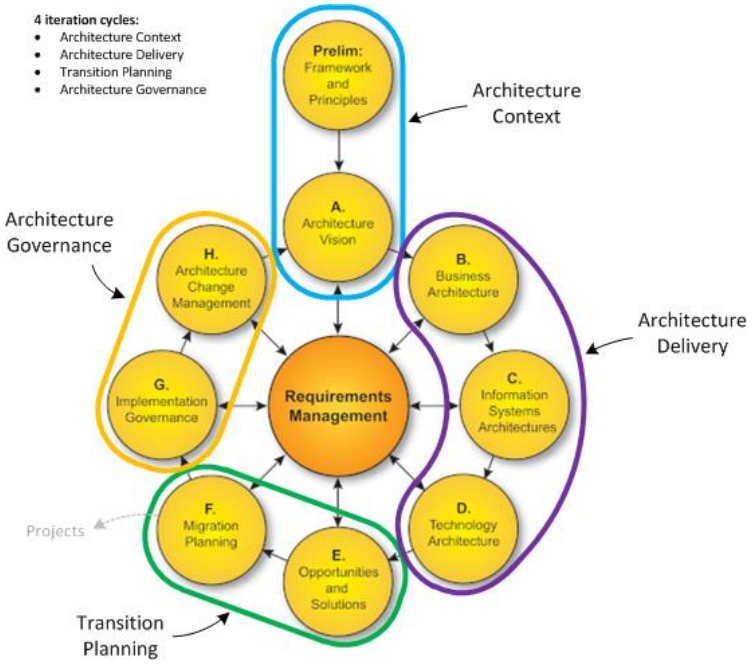


Figura 3. Fuente: https://chae201521700912580.wordpress.com/wp-content/uploads/2015/10/adm-v3.png

En el contexto de esta investigación, el método ADM resulta fundamental porque permitirá estructurar el diseño de la arquitectura empresarial para Ponte Selva. A partir del diagnóstico de los procesos y de los resultados obtenidos mediante CRISP-DM, el ADM facilitará la definición de una arquitectura objetivo que esté orientada a la mejora de la toma de decisiones y la alineación entre datos, aplicaciones, tecnología y estrategia organizacional.

Las fases específicas dentro del proceso de ADM son las siguientes tomando en cuenta cuál es su tarea enfoque:

Fases ADM

Fase de ADM	Tarea enfocada en la Fase
Fase preliminar	Prepara a la organización para adoptar TOGAF, definiendo principios, alcance y gobierno de arquitectura.
Fase A: Visión de arquitectura	Define el alcance, los interesados y la visión inicial de la arquitectura objetivo.
Fase B: Arquitectura de negocio	Analiza procesos, funciones, roles, estructura organizacional y objetivos del negocio.
Fase C: Arquitectura de sistemas de información	Desarrolla la arquitectura de datos y aplicaciones.
Fase D: Arquitectura tecnológica	Define la infraestructura tecnológica necesaria para soportar aplicaciones y procesos.
Fase E: Oportunidades y soluciones	Identifica iniciativas, soluciones y proyectos necesarios para alcanzar la arquitectura objetivo.
Fase F: Planificación de la migración	Prioriza proyectos, analiza riesgos y define la ruta de transición.
Fase G: Gobierno de la implementación	Supervisa que la implementación se mantenga alineada con la arquitectura definida.
Fase H: Gestión de cambios de arquitectura	Evalúa cambios futuros y mantiene la arquitectura actualizada.
Gestión de requerimientos	Acompaña todo el ciclo ADM para asegurar que la arquitectura responda a las necesidades del negocio.

Tabla 1: Fases de ADM. Fuente: Elaboración Propia

2.2.4 Dominios de TOGAF

TOGAF organiza la arquitectura empresarial en distintos dominios que permiten analizar la organización desde perspectivas complementarias. Estos dominios facilitan la comprensión integral de la empresa, ya que cada uno se enfoca en un conjunto específico de elementos, relaciones y dependencias. De acuerdo con Conexiam (2025), los dominios de arquitectura empresarial representan áreas parciales de una arquitectura completa y permiten estructurar el análisis de la organización en función de sus componentes principales.

- ✓ Arquitectura de negocio: describe la estrategia, estructura organizacional, procesos, funciones, capacidades, productos, servicios y actores que participan en la empresa. Este dominio permite comprender cómo opera la organización y cómo sus procesos contribuyen al cumplimiento de sus objetivos.

- ✓ Arquitectura de datos: se enfoca en la estructura, gestión, almacenamiento, calidad y flujo de la información dentro de la organización. Este dominio permite identificar qué datos posee la empresa, dónde se generan, cómo se almacenan y de qué manera pueden ser utilizados para apoyar la toma de decisiones.
- ✓ Arquitectura de Aplicaciones: describe los sistemas de información, herramientas digitales y aplicaciones que soportan los procesos del negocio. Este dominio permite analizar cómo las aplicaciones interactúan entre sí, qué funciones cumplen y qué tan alineadas están con las necesidades operativas de la empresa. Desde la perspectiva de TOGAF, este dominio contribuye a identificar si los sistemas actuales apoyan adecuadamente los procesos o si se requiere integrar nuevas soluciones tecnológicas para mejorar la eficiencia.
- ✓ Arquitectura de Tecnología: comprende la infraestructura física y lógica que soporta las aplicaciones y servicios de información. Incluye elementos como servidores, redes, bases de datos, plataformas, dispositivos, servicios tecnológicos y mecanismos de seguridad. Este dominio permite determinar si la infraestructura existente es suficiente para soportar los requerimientos actuales y futuros de la organización.

Según LinkedIn Advice (2023), los cuatro dominios de TOGAF se encuentran interrelacionados y deben alinearse con los objetivos estratégicos de la empresa para lograr una arquitectura coherente.

2.2.5 Arquitectura Empresarial orientada por datos

La transformación digital ha conducido a que las organizaciones dependen cada vez más de la disponibilidad y análisis de grandes volúmenes de información para ayudar a la toma de decisiones estratégicas. Se aprecia el nacimiento del concepto de Data-Driven Enterprise Architecture (DDEA), esta es una evolución de la Arquitectura Empresarial

tradicional que hace uso de datos actualizados, métricas y un análisis como base para la planificación y el gobierno de la organización.

Data-Driven Enterprise Architecture

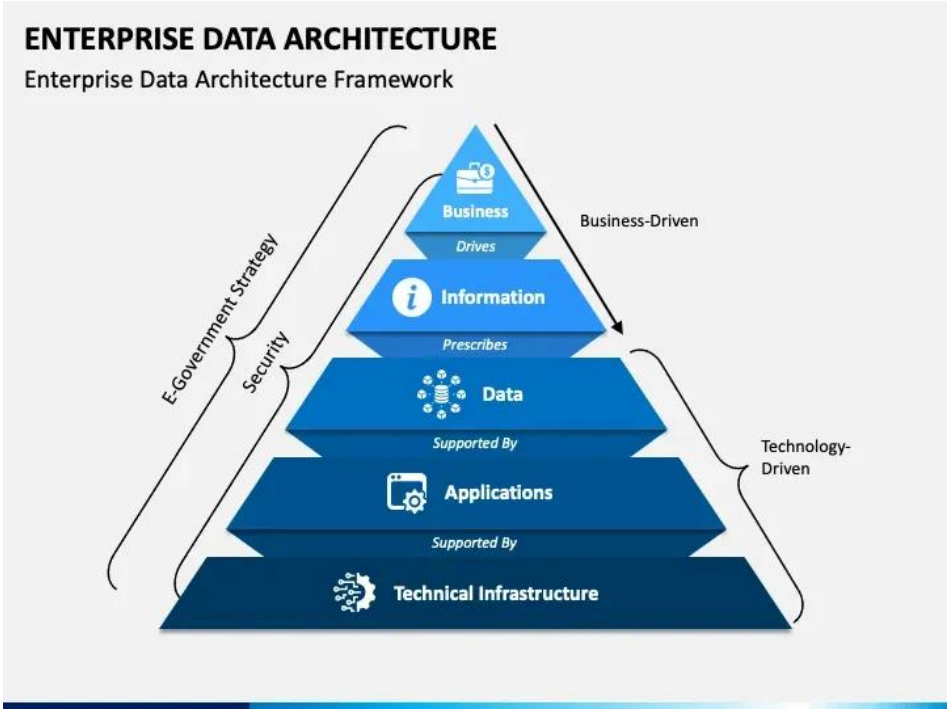


Figura 4. Fuente: https://cdn.sanity.io/images/dnt4jhwg/production/8af553ad67a4c7a5c1b4aabc5b14335d4ba5ed4c-720x540.webp

A diferencia de los enfoques tradicionales, en los que la arquitectura se documenta manualmente y tiende a quedar rápidamente desactualizada, la arquitectura basada en datos utiliza repositorios, herramientas de monitorización, analítica y fuentes de datos en tiempo real para ofrecer una representación fiel del estado actual de la organización; lo cual permite que las decisiones en el ámbito empresarial se fundamenten en evidencia objetiva, y no solo en la experiencia o la intuición de los arquitectos empresariales.

Un entorno de arquitectura empresarial orientado por datos integra información proveniente de procesos de negocio, aplicaciones, infraestructura de tecnología, capacidades de la organización y activos de la información dentro de un repositorio centralizado, y constituye una única fuente de verdad desde la cual se pueden analizar dependencias, evaluar

impactos, diseñar hojas de ruta de transformación, etc., en función de los objetivos estratégicos de la organización. BOC Group (2026)

El modelo de Data-Driven Enterprise Architecture tiene su origen en un conjunto de principios que sirven para mejorar la calidad del proceso de toma de decisiones de una organización.

1. Decisiones basadas en datos: Las decisiones arquitectónicas deben ser siempre información contrastada, métricas y análisis objetivos, así como reducir la dependencia en el 'intuitivo' o en la 'experiencia' personal que hace que se pueda errar (BOC Group, 2026).
2. Única fuente de verdad: La información por arquitectónica debe ser la guardada en un único repositorio para asegurar la continuo consistencia, integridad y disponibilidad de los datos por todas las partes interesadas de la organización (BOC Group, 2026).
3. Información actualizada: La arquitectura de la organización debe presentar siempre el estado actual, por tanto, hay que poder actualizar el proceso, la aplicación, la tecnología y los datos, evitando la documentación vieja, etc. (Tyk API Management, 2025).
4. Los procesos de trazabilidad y gestión de dependencias: Hay que mantener identificadas las relaciones entre procesos de negocio, aplicaciones, infraestructura y datos para evaluar los posibles impactos de cualquier cambio en la organización (BOC Group, 2026).
5. Alineación estratégica: Se debe asegurar que la información recogida debe utilizarse para que las iniciativas tecnológicas y las inversiones en TI tengan una relación de contribución directa en el cumplimiento de los objetivos estratégicos de la organización (Tyk API Management, 2025).

6. Transparencia y cooperación: La arquitectura debe optimizar la comunicación entre las áreas de negocio, de tecnología y dirección, aportando una visión compartida de la organización y una mayor coordinación de los diferentes actores (Ardoq, 2024).
7. Automatización de la recopilación de datos: La información arquitectónica debe ser obtenida a partir de forma automática de herramientas empresariales, repositorios, plataformas DevOps y sistemas de información, evitando el esfuerzo manual y aumentando la fiabilidad de los datos siempre que sea posible (Tyk API Management, 2025).
8. Mejoramiento continuo: La arquitectura debe evolucionar continuamente mediante la evaluación de los datos de forma continua para conseguir detectar oportunidades de mejora, de innovación y de transformación digital (BOC Group, 2026).
9. Gobierno y calidad de los datos: La pieza clave en la fiabilidad de la arquitectura es que los datos sean adecuados, consistentes y correctamente gobernados, puesto que la calidad de la información se traduce en la calidad de las decisiones (Tyk API Management, 2025).
10. Enfoque sobre el valor del negocio: La arquitectura empresarial debe proporcionar datos que permitan descubrir oportunidades de reducción de costes, innovación tecnológica y generación de nuevas capacidades que puedan aportar valor a la organización (Ardoq, 2024).

2.2.6 Ciencia de datos

Es el campo de estudio de varias disciplinas orientado a la extracción de conocimiento a partir de grandes volúmenes de información. Esta disciplina combina elementos de estadística, matemáticas, informática, programación, aprendizaje automático y conocimiento del dominio del negocio, con el propósito de analizar datos, identificar patrones y apoyar la toma de decisiones.

La ciencia de datos cobra valor en contextos empresariales donde las decisiones deben fundamentarse en evidencia y no solamente en criterios intuitivos o experiencias previas. Asimismo, Crabtree y Nehme. (2024) plantean que la ciencia de datos sirve para una extensa serie de aplicaciones, desde predecir el comportamiento de los clientes hasta mejorar los procesos empresariales. La ciencia de los datos abarca un campo muy amplio que incluye distintos tipos de análisis.

También guarda una relación directa con la Arquitectura Empresarial, ya que los resultados obtenidos del análisis de datos pueden orientar el diseño de mejoras en procesos, aplicaciones, datos y tecnología. En este sentido, los modelos analíticos no deben entenderse como soluciones aisladas, sino como insumos estratégicos que pueden integrarse dentro de una arquitectura empresarial para apoyar la transformación organizacional.

2.2.7 Minería de Datos

La minería de datos forma parte del proceso de descubrimiento de conocimiento en bases de datos, conocido como KDD (Knowledge Discovery in Databases). Este proceso no se limita únicamente a aplicar algoritmos, sino que incluye etapas previas como la recolección, selección, limpieza y transformación de los datos, así como fases posteriores de evaluación, validación, interpretación y difusión de los resultados. En este sentido, Elena (2026) explica que el KDD es un proceso que se utiliza para llevar a cabo la automatización de la extracción de conocimiento comenzando por grandes volúmenes de datos, el cual es un proceso iterativo, por lo que puede ser utilizado tantas veces como sea necesario hasta localizar la información necesaria.

Proceso de Minería de Datos

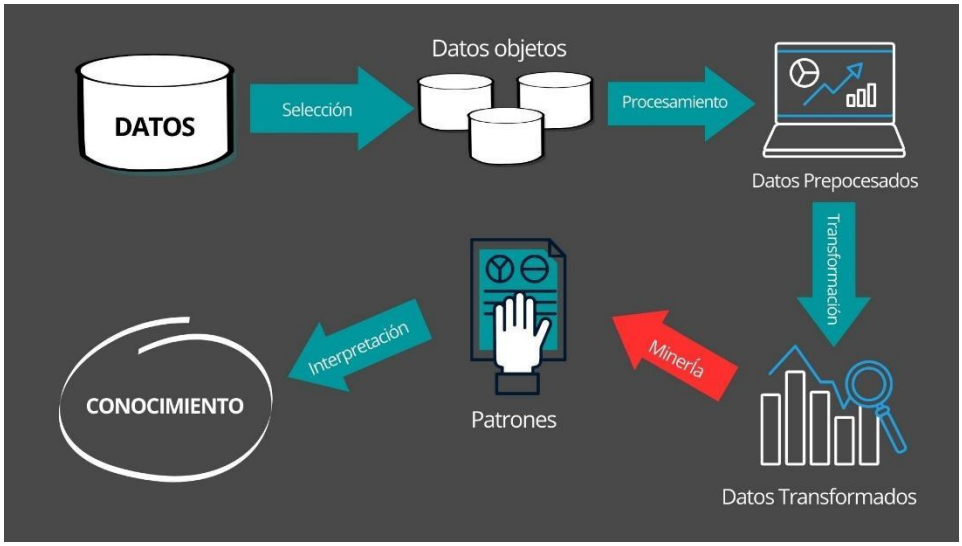


Figura 5. Fuente: https://cdn.prod.website-files.com/671603568e40dcb4a09fce45/67459c9121818a87c59e9567_673334fa817bad004287163d_Datos.jpeg

La minería de datos se apoya en técnicas provenientes de distintas áreas, como la estadística, la inteligencia artificial, el aprendizaje automático y los sistemas de bases de datos. Estas técnicas permiten analizar información histórica para identificar comportamientos repetitivos, segmentar registros, clasificar datos, estimar resultados futuros o descubrir asociaciones entre variables. Por esta razón, la minería de datos resulta especialmente útil en contextos empresariales donde se requiere transformar datos operativos en información estratégica.

Entre las principales tareas de la minería de datos se encuentran la clasificación, la regresión, el agrupamiento, la asociación, la detección de patrones, la predicción y el soporte en la toma de decisiones. Estas tareas pueden ser descriptivas, cuando buscan explicar el comportamiento de los datos, o predictivas, cuando pretenden anticipar resultados futuros a partir de información histórica. Villegas (2023) explica que emplean diversos métodos que dan lugar a modelos que ajustan los resultados esperados en función de los resultados del proceso de minería de datos de la compañía. Los modelos pueden aplicarse para justificar

las características de los datos actuales, para prever tendencias venideras o incluso para ayudar a detectar anomalías que están presentes en los datos.

2.2.8 Modelos Predictivos

Los modelos predictivos son herramientas matemáticas o computacionales que permiten estimar eventos, comportamientos o resultados futuros a partir del análisis de datos históricos y actuales. Estos modelos se basan en la identificación de patrones, correlaciones y tendencias dentro de los datos, teniendo como objetivo construir una función predictiva que permita anticipar la probabilidad de ocurrencia de un evento determinado. Según González Blanco (2024), los modelos predictivos integran técnicas estadísticas de análisis de datos y pueden apoyarse en algoritmos de aprendizaje automático e inteligencia artificial para analizar grandes conjuntos de información.

Existen diversos tipos de modelos predictivos, dependiendo del problema y de la variable que se desea estimar. Los modelos de regresión se emplean cuando el objetivo es predecir valores continuos o cuantitativos, como ventas, demanda, costos, producción o ingresos. Estos modelos buscan representar la relación entre una variable dependiente y una o más variables independientes. Entre sus variantes se encuentran la regresión lineal simple, regresión lineal múltiple, regresión polinomial y otros modelos orientados al análisis de relaciones cuantitativas (González Blanco, 2024).

Los modelos de clasificación, por su parte, se utilizan cuando el objetivo es asignar observaciones a categorías o clases previamente definidas. Este tipo de modelo es útil para problemas en los que se busca predecir si un evento ocurrirá o no, como clientes satisfechos o insatisfechos, productos conformes o defectuosos, pedidos con riesgo o sin riesgo, entre otros casos. González Blanco (2024) identifica dentro de esta categoría técnicas como regresión logística, árboles de decisión, Random Forest y máquinas de soporte vectorial, las cuales permiten clasificar registros a partir de variables de entrada.

Modelo de Clasificación

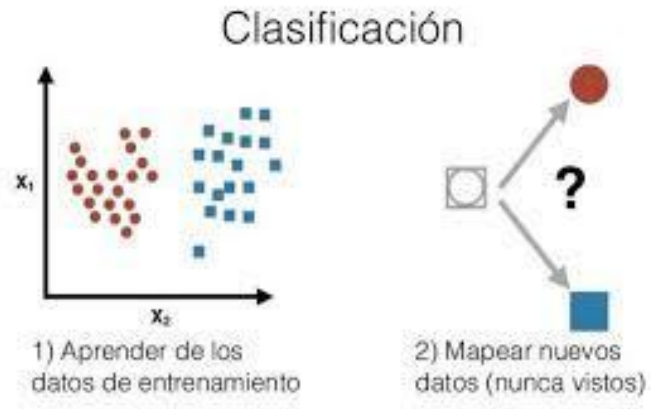


Figura 6. Fuente: https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcSqJIE7LVDIeSD6Y6h3fqdBb_wI0DuKJTt2DA&s

Además de los modelos de regresión y clasificación, existen otros enfoques predictivos como los modelos de series temporales, detección de anomalías, agrupamiento y recomendación. Los modelos de series temporales permiten estimar valores futuros basándose en observaciones registradas a lo largo del tiempo; los modelos de detección de anomalías identifican datos que se desvían del comportamiento habitual; los modelos de agrupamiento permiten descubrir estructuras o grupos en datos sin etiquetas previas; y los modelos de recomendación sugieren productos, servicios o información según preferencias o comportamientos anteriores (González Blanco, 2024).

Componentes de una Serie de Tiempo

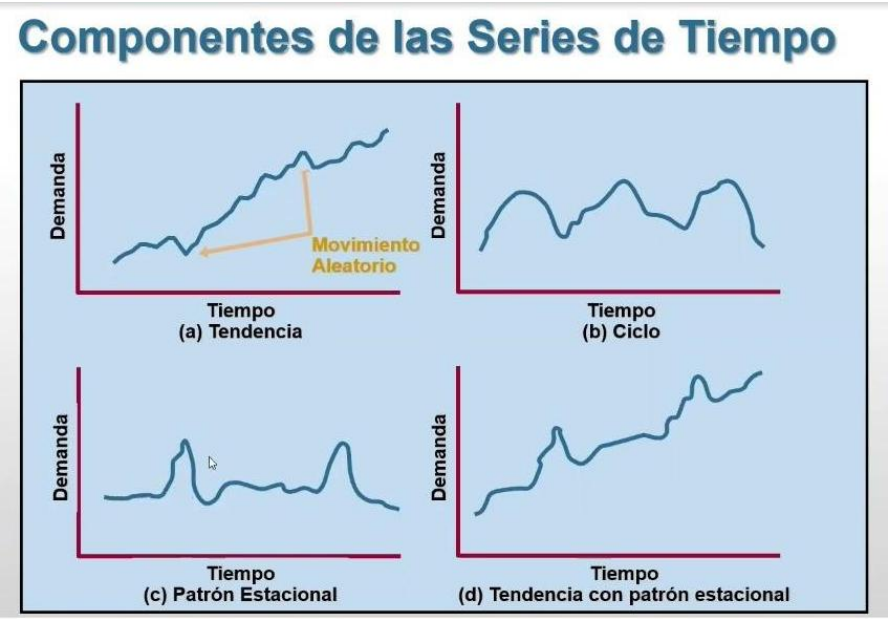


Figura 7. Fuente: https://i.ytimg.com/vi/aUwXWGa8jK0/maxresdefault.jpg

2.2.9 Evaluación de modelos predictivos

La evaluación de los modelos predictivos es una etapa fundamental, ya que permite determinar si el modelo construido es confiable y si sus resultados pueden utilizarse para apoyar la toma de decisiones. Ramalli en al. (2022) señala que, el procedimiento de evaluación establece un vínculo entre la construcción de un complejo modelo y los datos experimentales que calculan el rendimiento del modelo predictivo, comparando las predicciones del modelo con las medidas experimentales relevantes. Una práctica común consiste en dividir el conjunto de datos en un conjunto de entrenamiento, utilizado para construir el modelo, y un conjunto de prueba, utilizado para validar su desempeño.

En los problemas de regresión, donde el modelo predice valores numéricos continuos, las métricas se basan en el error existente entre el valor real y el valor predicho. Una de las métricas más utilizadas es el error absoluto medio (MAE), que calcula el promedio de las diferencias absolutas entre los valores reales y los valores estimados.

$$MAE = \frac{\sum \quad |Valor\ real - Valor\ predicho|}{n}$$

También se utiliza el error cuadrático medio (MSE), que eleva al cuadrado las diferencias entre los valores reales y predichos, penalizando con mayor fuerza los errores grandes:

$$MSE = \frac{\sum \quad (Valor\ real - Valor\ predicho)^2}{n}$$

De esta manera, los modelos predictivos permitirán analizar los datos recopilados por Ponte Selva, identificar patrones relevantes y anticipar comportamientos en la satisfacción del cliente. Los resultados obtenidos servirán como insumo para la metodología CRISP-DM y, posteriormente, serán integrados en la arquitectura empresarial basada en TOGAF, con el objetivo de fortalecer la toma de decisiones estratégicas en la empresa.

2.2.10 CRISP-DM

CRISP-DM, por sus siglas en inglés Cross Industry Standard Process for Data Mining, es una metodología orientada al desarrollo estructurado de proyectos de minería de datos. Su propósito es proporcionar una guía organizada para transformar datos en conocimiento útil, mediante un proceso compuesto por fases, tareas y resultados esperados. De acuerdo con G & G (2025), CRISP-DM abarca todas las fases de los proyectos de minería de datos, que van desde las tareas orientadas a la comprensión del problema hasta las de puesta en producción de sistemas analíticos automatizado y predictivos.

Está compuesto por seis fases principales: comprensión del negocio, comprensión de los datos, preparación de los datos, modelado, evaluación y despliegue. Estas fases no se ejecutan necesariamente de manera lineal, ya que el proyecto puede avanzar o retroceder entre ellas según los resultados obtenidos. Esta característica iterativa permite ajustar el análisis conforme se descubren nuevos problemas, necesidades o patrones en los datos (IBM, s.f.).

2.3 Integración de CRISP-DM y TOGAF para la toma de decisiones empresariales

A diferencia de una Arquitectura Empresarial como la tradicional, que plantea la estructura organizativa y la correspondiente organización tecnológica que posee la empresa, la inclusión de CRISP-DM permite enriquecer dicha arquitectura a partir de los modelos predictivos obtenidos, a su vez, usando las bases de datos o mediante los datos de la propia empresa. Por eso, la Arquitectura no sólo describe el estado actual de la empresa, sino que pasa a ser una herramienta que permite anticipar el futuro y dar soporte a la planificación empresarial.

2.3.1 Datos históricos

El proceso comienza con la recopilación de los datos históricos generados por la compañía, esto es durante su operación diaria. Para el caso de Ponte Selva, podemos sostener que la principal fuente de información es la que proporciona las encuestas de satisfacción del cliente, ya que estos datos son el principal activo del análisis predictivo. Se contienen en ellos las evidencias del comportamiento del negocio y a través de ellos es posible identificar tendencias y patrones.

2.3.2 Aplicación de la metodología CRISP-DM

Una vez recogido los datos históricos, se aplica la metodología CRISP-DM, que divide el proyecto y desarrollo en sus 6 fases:

- Comprensión del negocio.
- Comprensión de sus datos.
- Preparación de los datos.
- Modelado.
- Evaluación.
- Despliegue.

Durante estas 6 fases tiene lugar la limpieza, la transformación y la preparación de la información para la creación de los modelos predictivos que permitan responder el objetivo planteado.

2.3.3 Generación de modelo predictivo

La etapa de modelado de CRISP-DM es la fase que entrega el modelo predictivo que permite estimar la forma en la que se comportará la satisfacción del cliente en el futuro. Para la investigación se utiliza el modelo SARIMAX, que según RPubs (2021) es una estructura matemática que integra todas las propiedades de un modelo SARIMA que también pueden considerar información sobre variables que la hacen comprenderse y pronosticarse en términos de la variable de interés.

Este modelo de SARIMAX permite hacer predicciones para la evolución de los indicadores de satisfacción del cliente. De este modo, se anticipa la satisfacción que puede experimentar el cliente en el futuro y permite que la organización anticipadamente pueda conocer en qué momento puede anticiparse la caída de la calidad del servicio o cuando puede cambiar el comportamiento de los clientes.

Pronóstico Ejemplo con SARIMAX

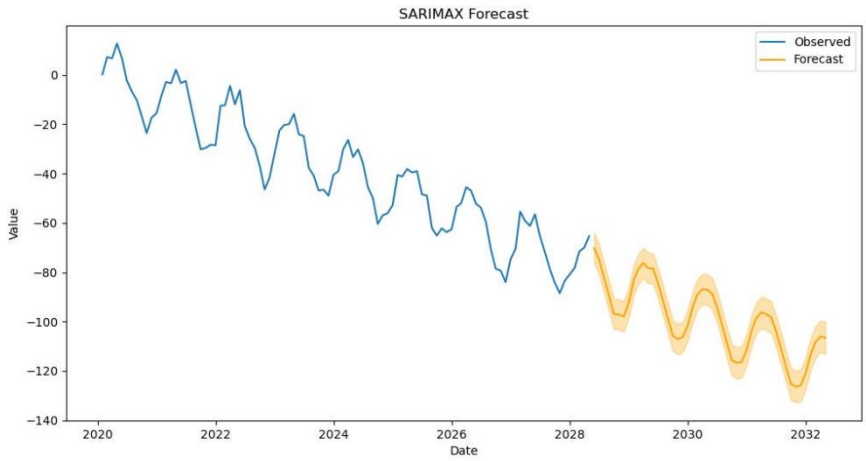


Figura 8. Fuente: https://substackcdn.com/image/fetch/$s_!Wo71!,f_auto,q_auto:good,fl_progressive:steep/https%3A%2F%2Fsubstack-post-media.s3.amazonaws.com%2Fpublic%2Fimages%2Fa76e4b10-6391-40b4-b4df-092e4e661b03_1200x600.png

De este modo, la organización puede contar con información proyectada que complementa el análisis descriptivo tradicional.

2.3.4 Integración de resultados al ciclo ADM de TOGAF

Los resultados generados por el modelo predictivo son introducidos como un recurso en el ciclo ADM de TOGAF para la definición de la arquitectura objetivo. Ya que el modelo AS-IS representa la situación actual de la empresa, la información generada desde CRISP-DM sirve para justificar las decisiones de la futura arquitectura. Los resultados predictivos dan soporte a actividades como:

- ✓ Identificación de oportunidades de mejora de los procesos de negocio
- ✓ El rediseño de las aplicaciones de negocio
- ✓ La mejora de la arquitectura de datos
- ✓ La priorización de los proyectos tecnológicos
- ✓ La planificación de las capacidades organizacionales
- ✓ La definición de iniciativas de transformación digital

Por esta razón, la fase de diseño no solo es una técnica de evaluación de la Arquitectura empresarial, sino que empieza a ser evaluada en esa evidencia obtenida a partir de la analítica de datos que se convierte en el resultado de toda esta fase.

2.3.5 Construcción de Arquitectura TO-BE

En el caso de Ponte Selva, la propuesta de arquitectura parte de una plataforma analítica basada en CRISP-DM que permite la integración de las encuestas de satisfacción y la emisión de los pronósticos, y todo ello en función de ser útil para la alta dirección en la planificación estratégica. Dicha arquitectura permite la transformación de la gestión de la empresa desde una gestión reactiva hacia un modelo predictivo y enfocado en el manejo de los datos.

2.3.6 Apoyo a la toma de decisiones empresariales

Los resultados del modelo predictivo permiten prever tendencias futuras y analizar diferentes alternativas antes de realizar inversiones, cambios de procesos o establecer nuevos sistemas. Por lo tanto, las decisiones que se relacionan con la mejora del servicio al cliente pueden nutrirse de evidencias obtenidas mediante la analítica de datos. Este trabajo en conjunto convierte la Arquitectura Empresarial en una herramienta para el gobierno de la empresa, a su vez que los datos alimentan los modelos predictivos y esos mismos modelos van orientando el proceso de transformación y apoyando las decisiones que la empresa tiene que tomar.

3. CAPÍTULO III: ARQUITECTURA EMPRESARIAL DE PONTE SELVA

La Arquitectura Empresarial permite entender cómo funciona una organización, analizando sus procesos, información, aplicaciones, tecnología y estructura organizacional. En este capítulo se desarrolla el análisis de la situación actual de Ponte Selva a través de los dominios arquitectónicos definidos por TOGAF, con el objetivo de identificar oportunidades de mejora que serán tratadas mediante una arquitectura objetivo integrada con la metodología CRISP-DM.

La organización orienta sus operaciones en la producción de prendas y artículos textiles de calidad, apoyándose en procesos de manufactura, logística y servicio al cliente que buscan garantizar la satisfacción de sus compradores.

Actualmente la empresa administra información relacionada con:

- Clientes
- Ventas
- Pedidos
- Inventario
- Producción
- Distribución
- Encuestas de satisfacción

Sin embargo, gran parte del análisis de esta información se realiza mediante herramientas tradicionales, lo cual limita la capacidad de obtener conocimiento para apoyar la toma de decisiones estratégicas.

3.1 Arquitectura de Negocio

3.1.1 Nivel Estratégico:

- Gerencia general

- Dirección estratégica
- Gestión de calidad

3.1.2 Nivel Operativo

- Comercial y ventas
- Producción
- Bodega e Inventarios
- Distribución y Logística
- Servicio al Cliente

3.1.3 Nivel de Apoyo

- Finanzas
- Talento Humano
- Compras
- Tecnologías de Información

3.1.4 Arquitectura Actual (AS-IS)

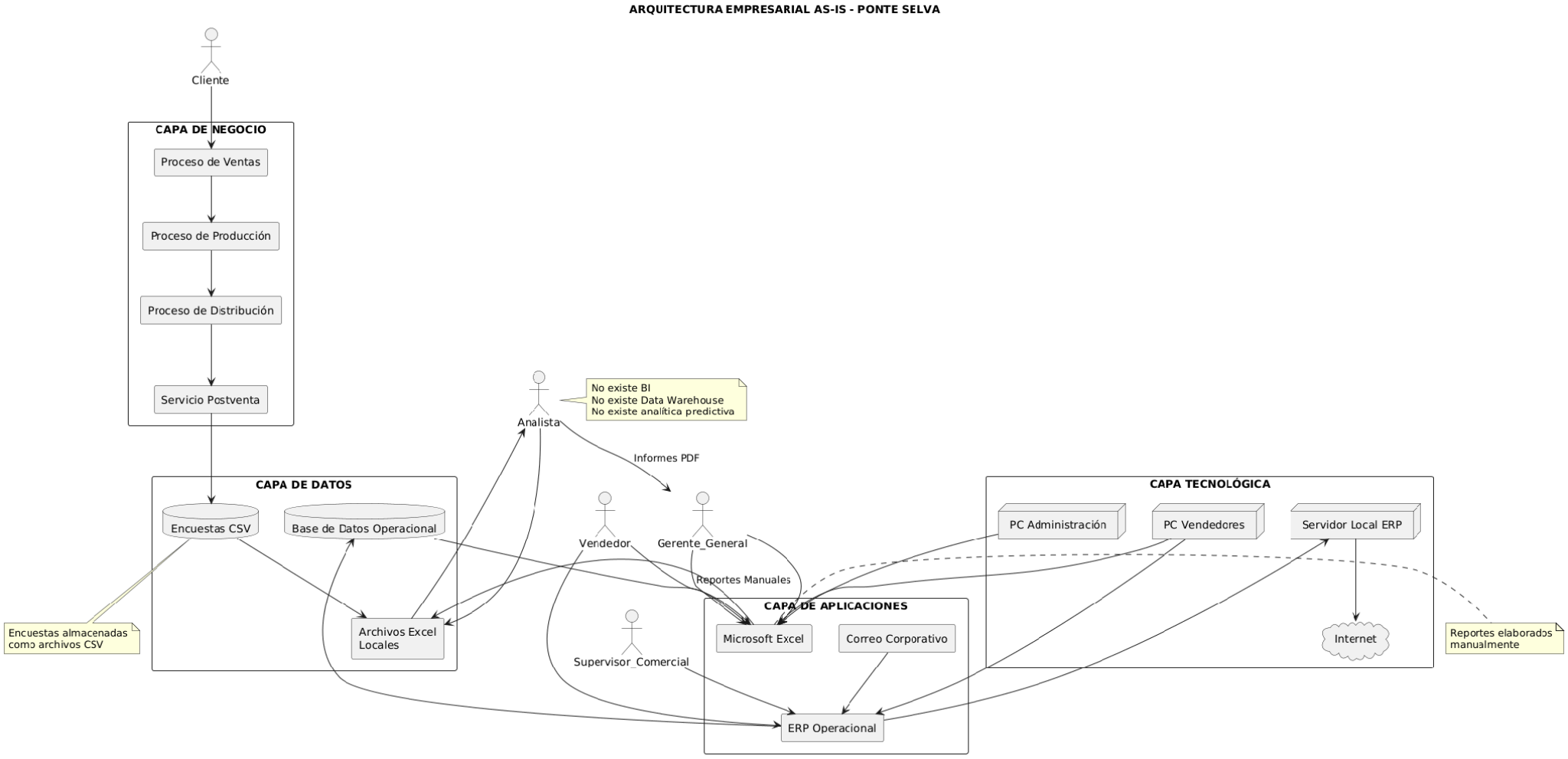


Figura 9: AS-IS Empresa Escogida. Fuente: Elaboración Propia

3.1.5 Stakeholders y Actores

- **Actores Internos**

- Gerente General
- Supervisor Comercial
- Jefe de Producción
- Bodeguero
- Operario
- Asistente Financiero
- Administrador TI
- Analista de Datos

- **Actores Externos**

- Clientes
- Proveedores
- Transportistas

3.1.6 Mapa de Procesos

- **Procesos Estratégicos**

1. Dirección Estratégica.
2. Gestión de Calidad.
3. Planificación Empresarial.
4. Gestión Comercial Estratégica.

- Procesos Misionales

1. Gestión Comercial.
2. Ventas.
3. Producción Textil.
4. Control de Calidad.

5. Bodega y Almacenamiento.
6. Distribución y Logística.
7. Servicio al Cliente.

- Procesos de Apoyo

1. Talento Humano.
2. Tecnologías de Información.
3. Finanzas y Contabilidad.
4. Compras y Gestión de Proveedores.
5. Gestión Jurídica.

Mapa de Procesos

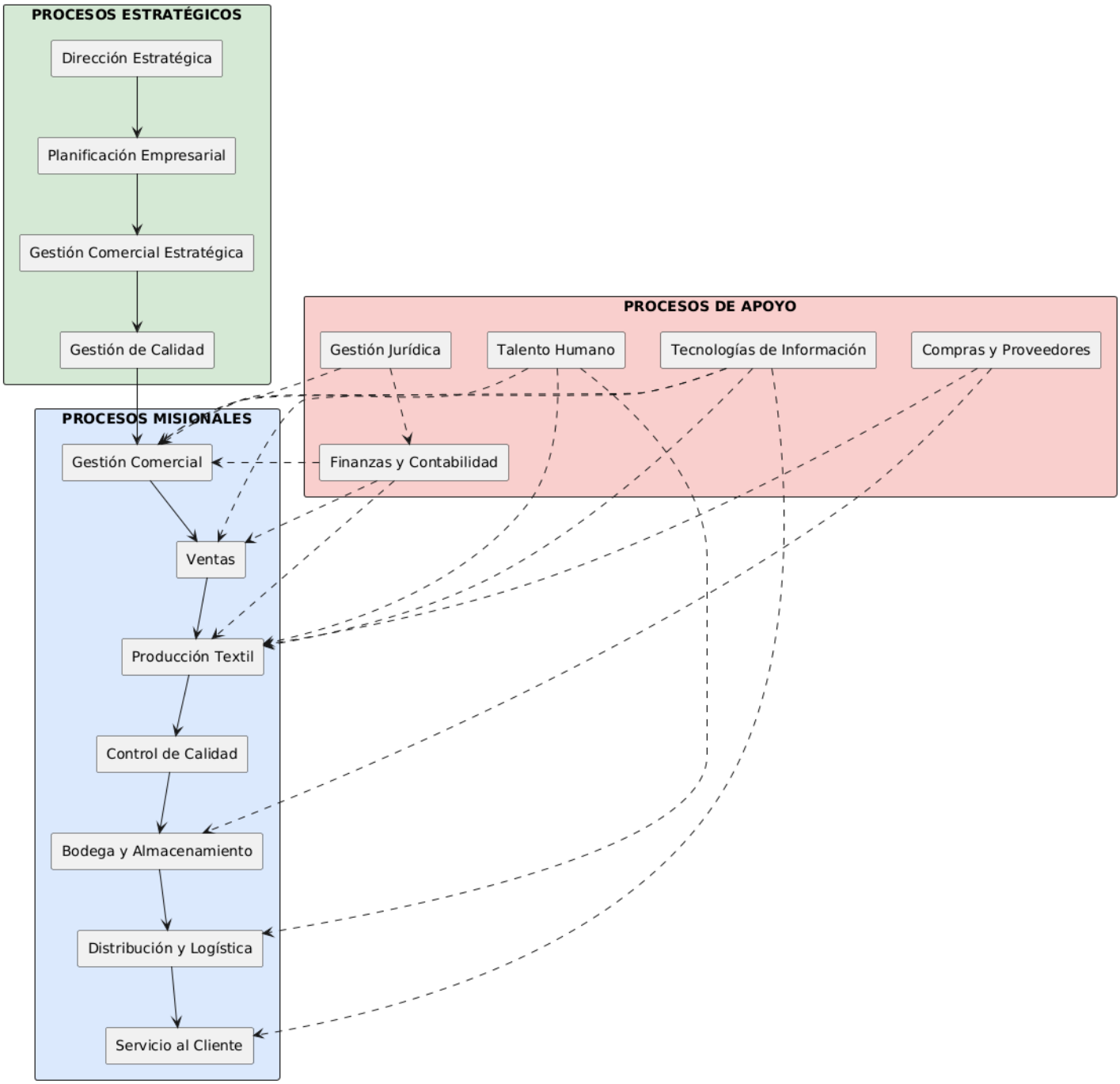


Figura 10: Mapa de Procesos. Fuente: Elaboración Propia

3.1.7 Proceso Comercial y Postventa

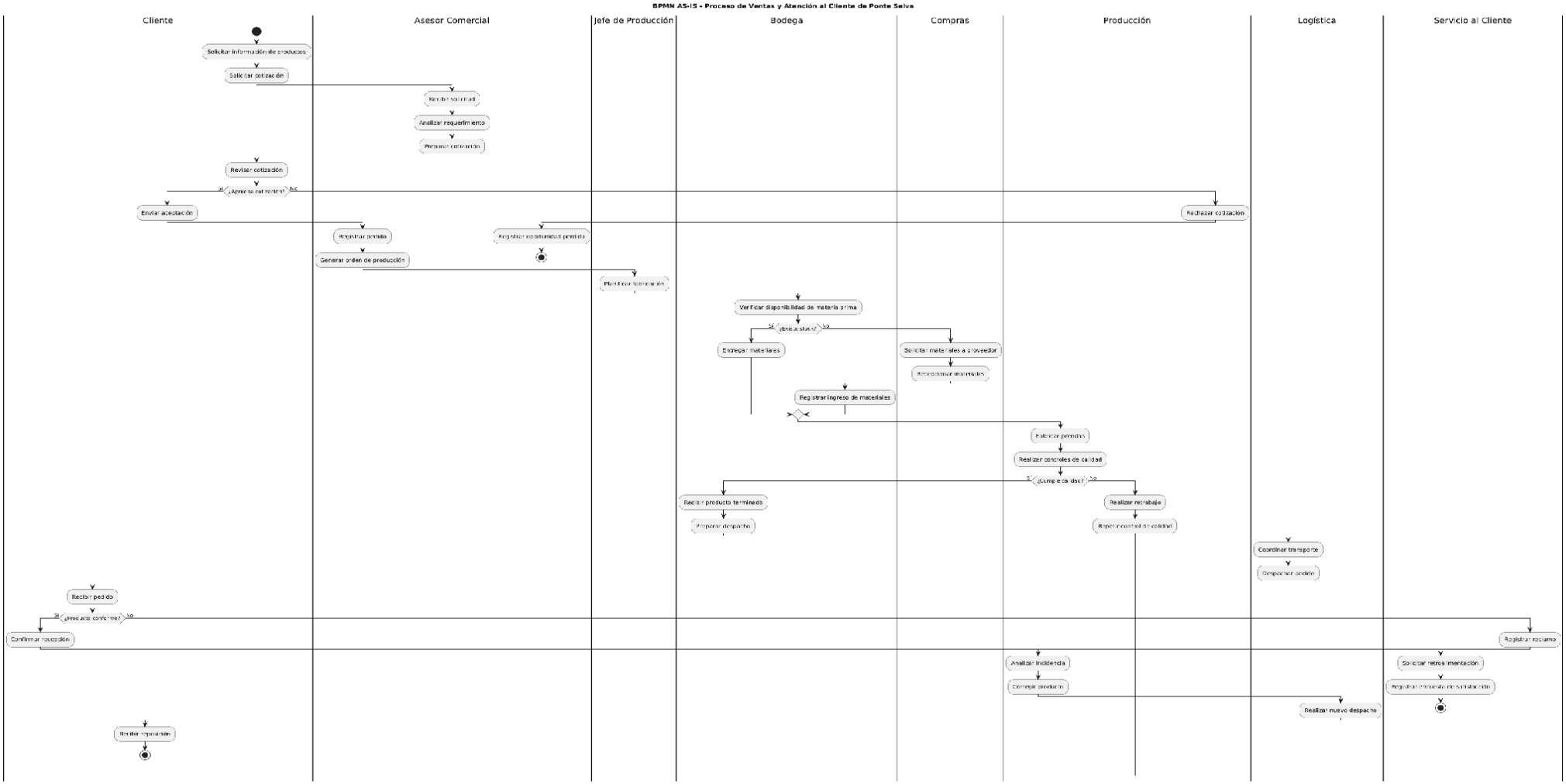


Figura 11: Proceso Comercial y Postventa. Fuente: Elaboración Propia

3.1.8 Cadena de Valor de Porter

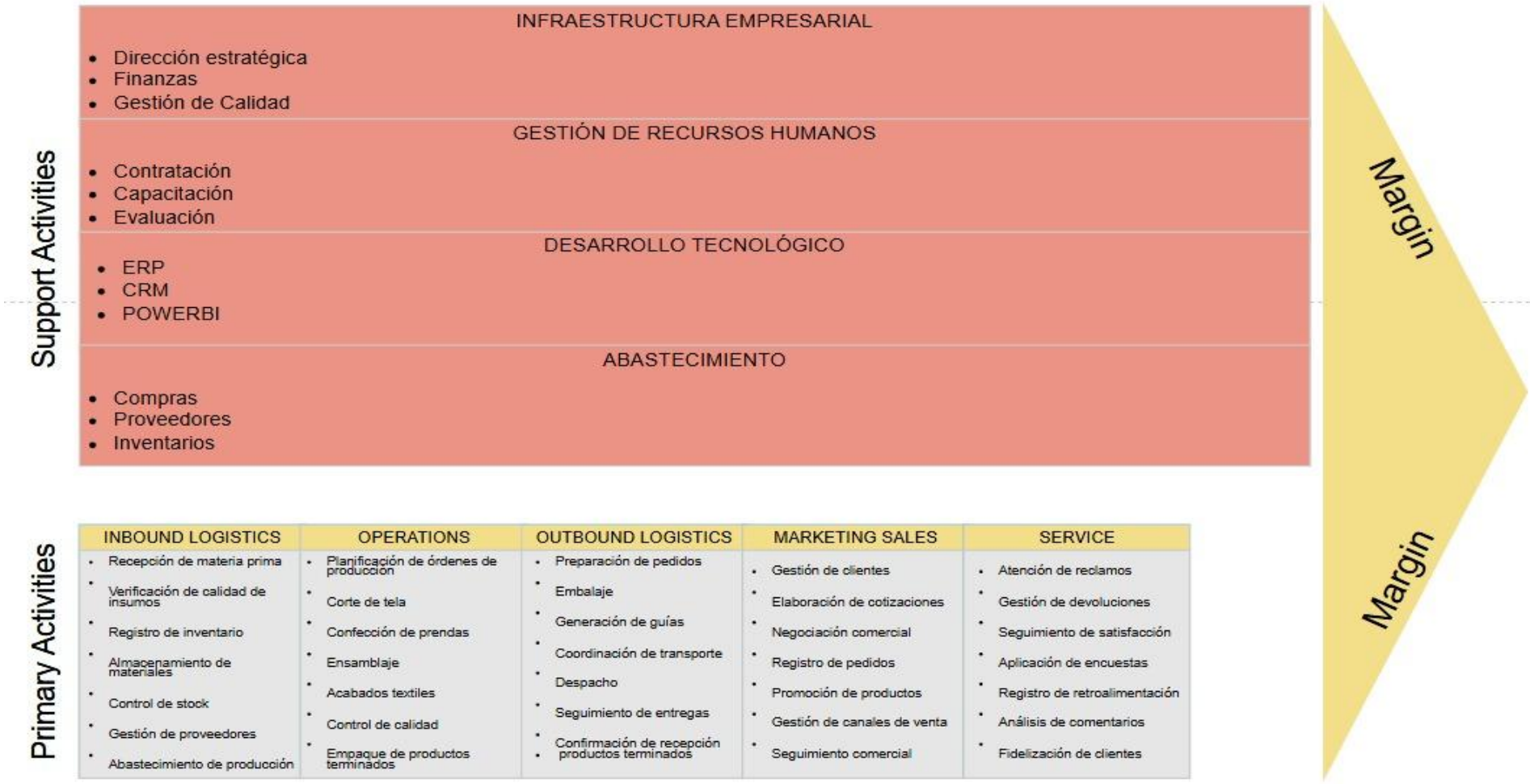


Figura 12: Cadena de Valor de Porter. Fuente: Elaboración Propia

3.1.9 Diagramas de Casos de uso

Caso de Uso Empresarial

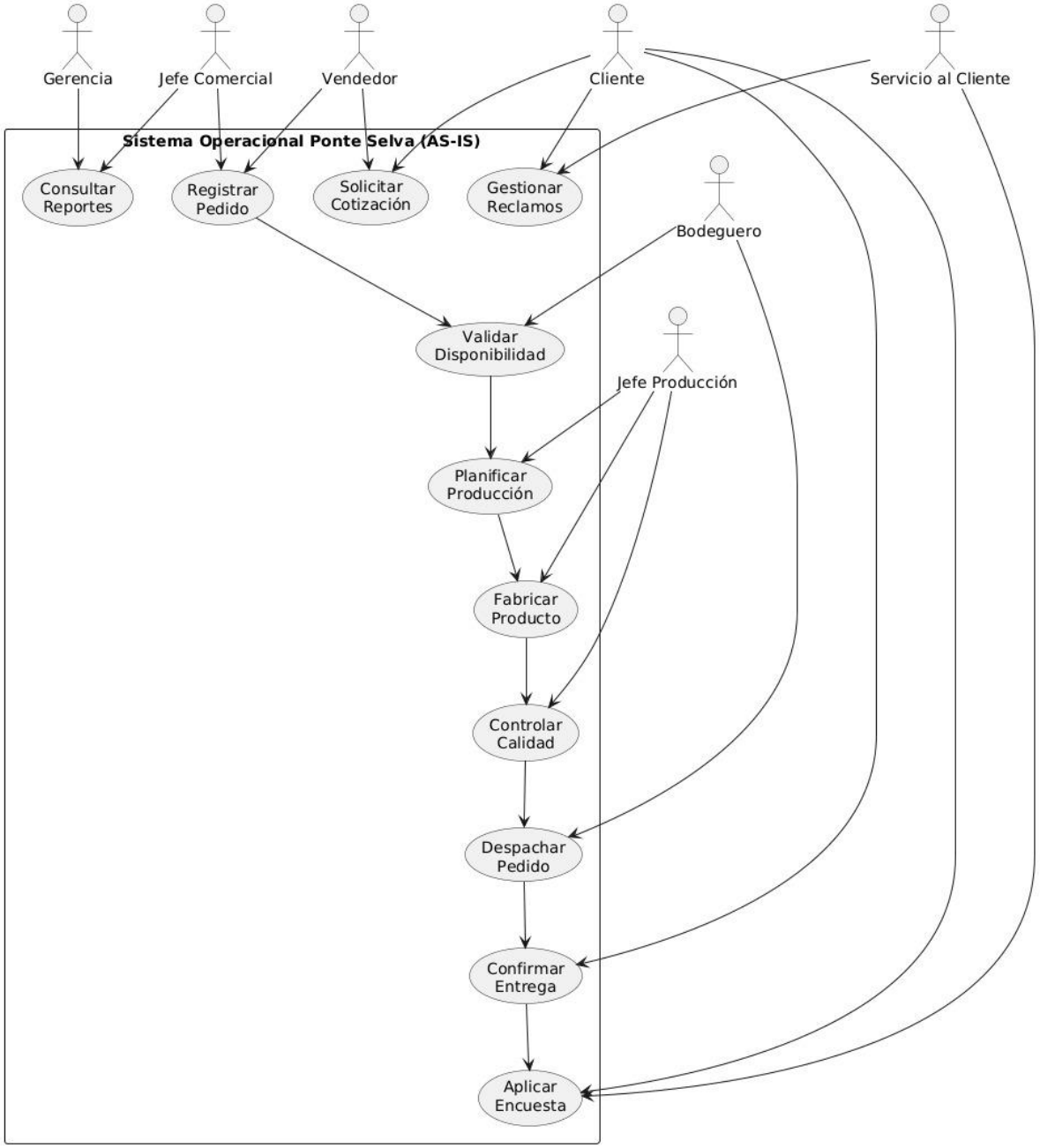


Figura 13. Fuente: Elaboración Propia

Producción Textil

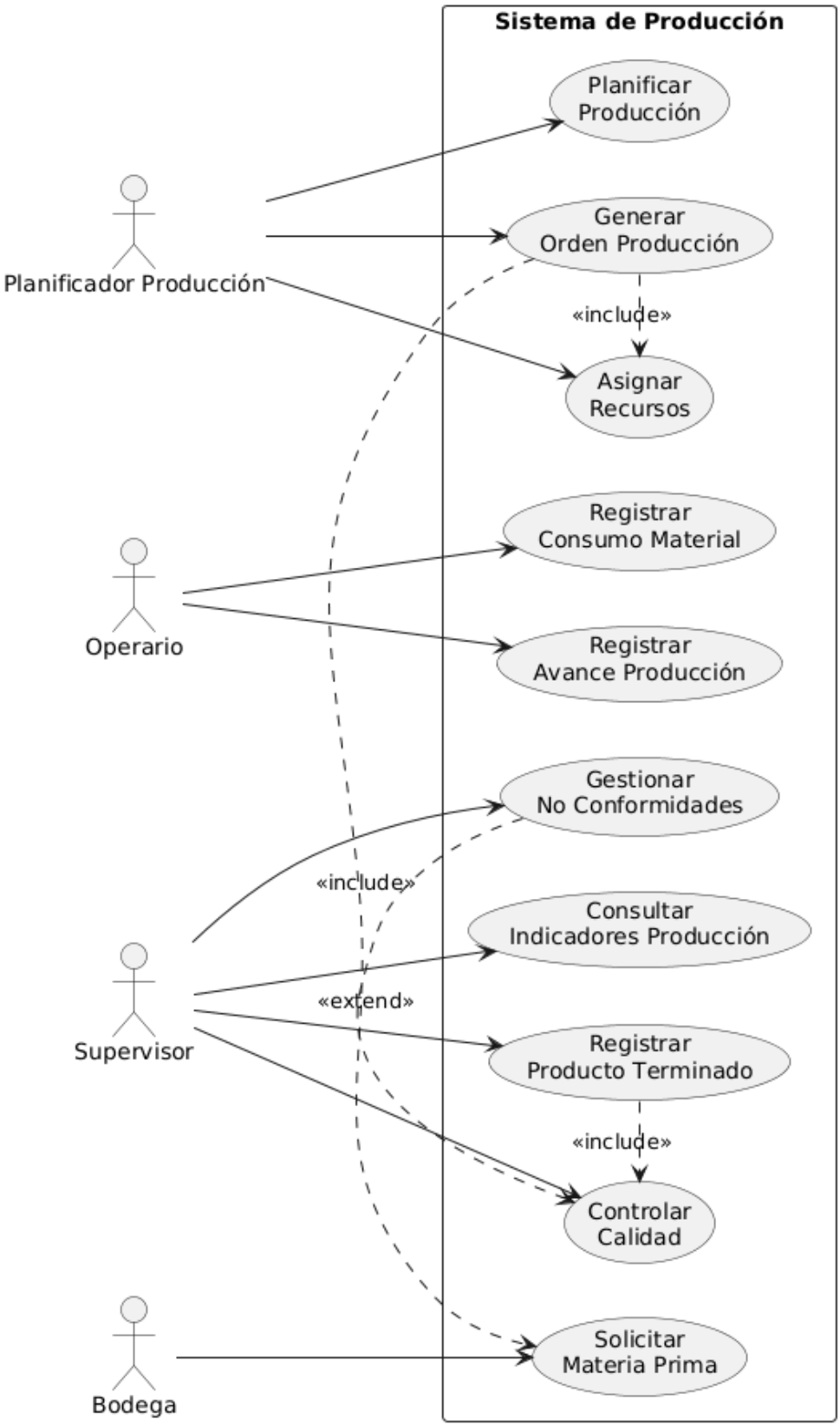


Figura 14. Fuente: Elaboración Propia

Inventario y Logística

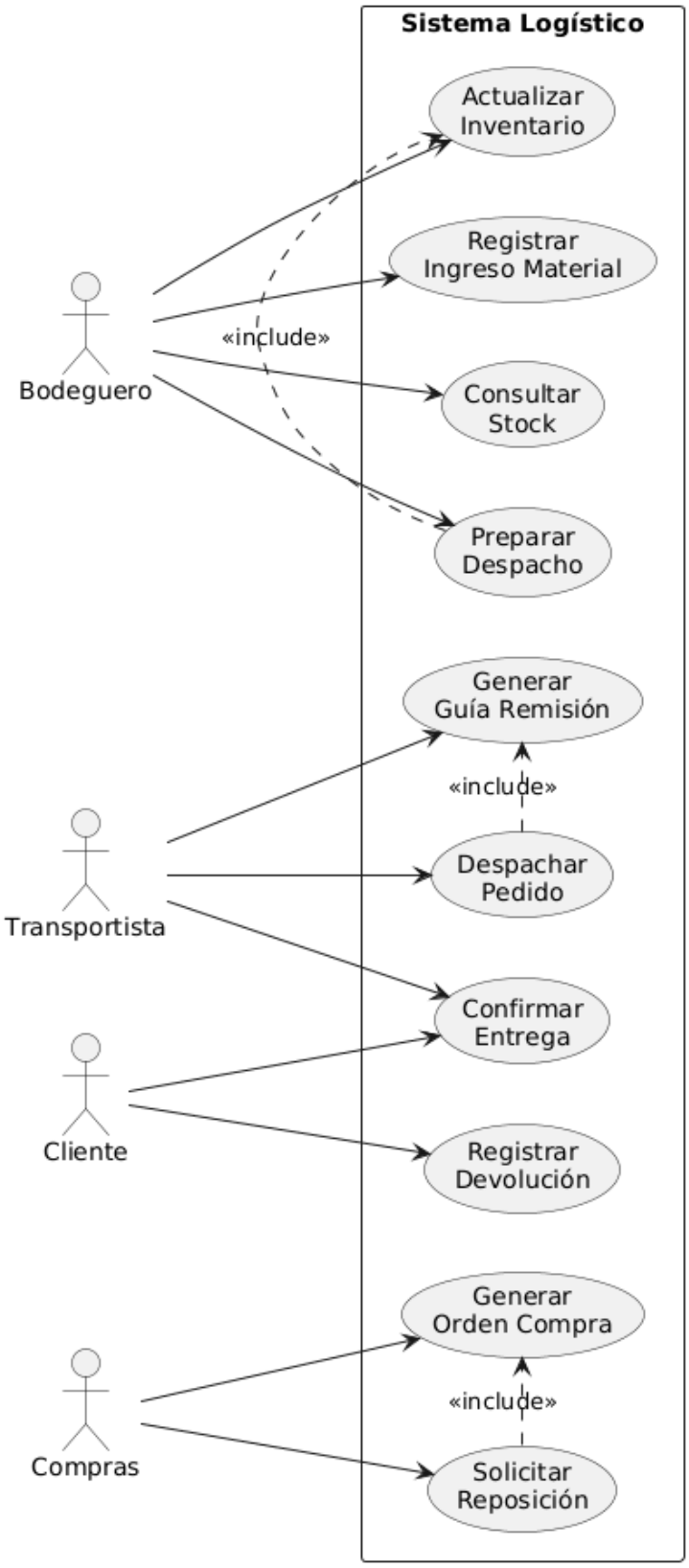


Figura 15. Fuente: Elaboración Propia

Servicio al Cliente

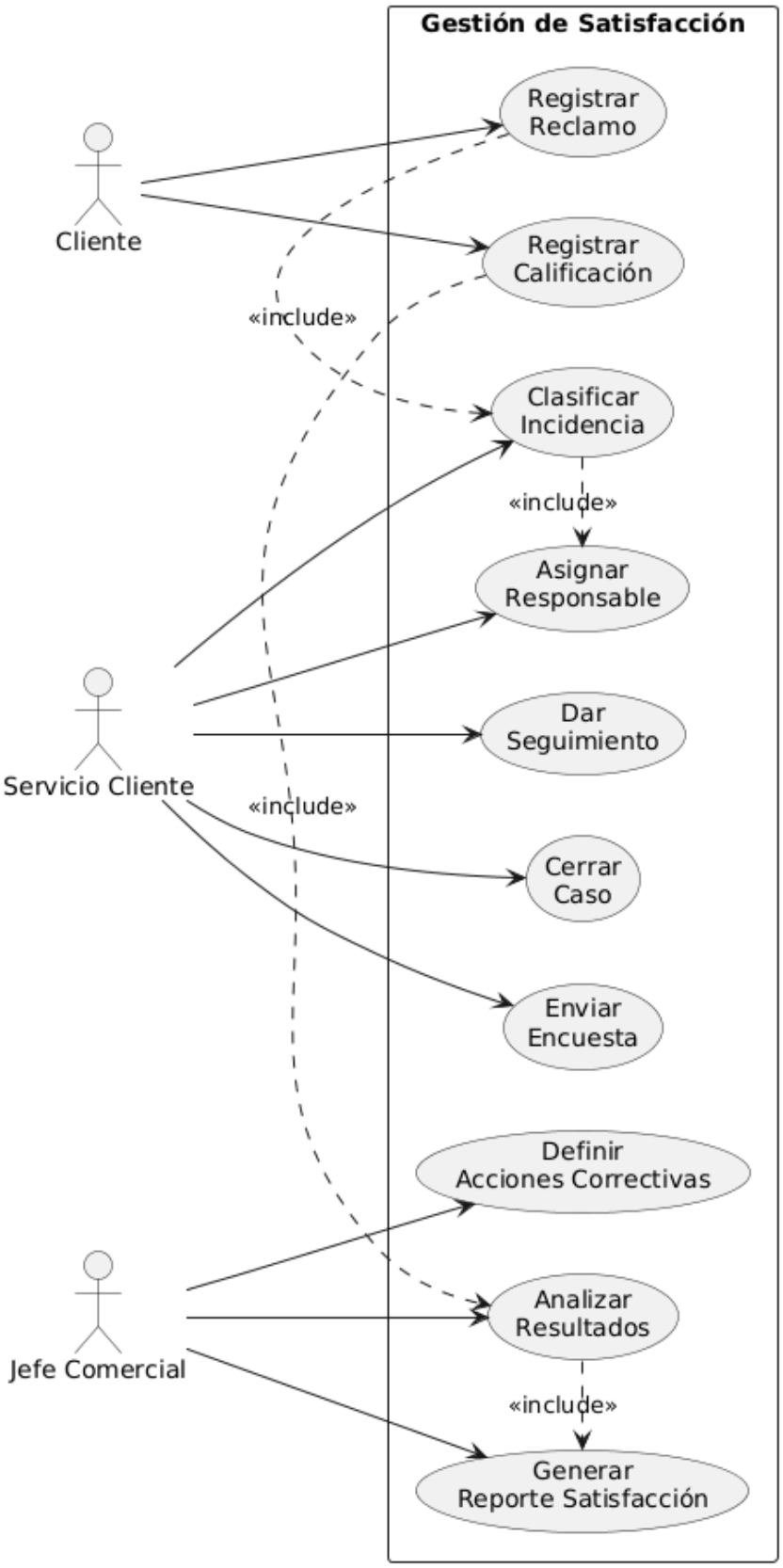


Figura 16. Fuente: Elaboración Propia

Gestión Administrativa

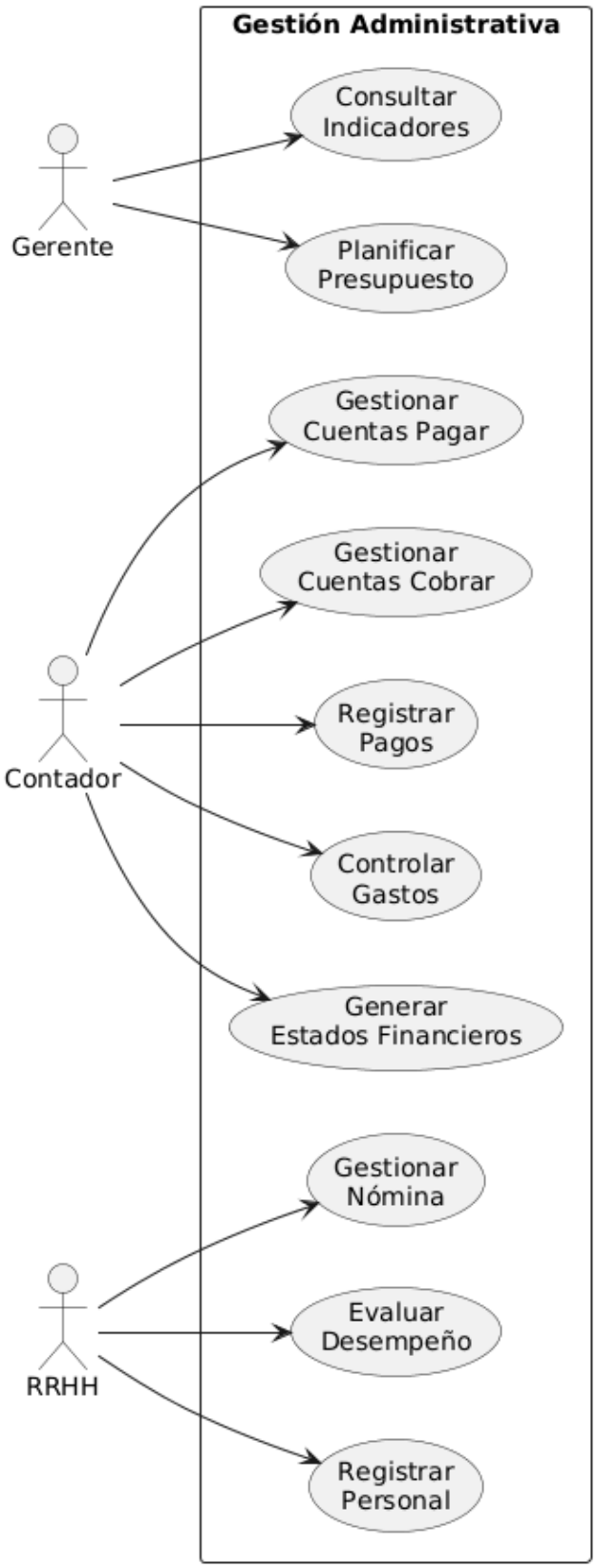


Figura 17. Fuente: Elaboración Propia

3.2 Arquitectura de Información

3.2.1 Catálogo de Información

Información	Área Responsable
Clientes	Comercial
Productos	Producción
Pedidos	Comercial
Inventario	Bodega
Proveedores	Compras
Facturación	Finanzas
Encuestas	Servicio al Cliente

Tabla 2: Catálogo de Información. Fuente: Elaboración Propia

3.2.2 Diagrama de Flujo de Información

Diagrama de Flujo de Información

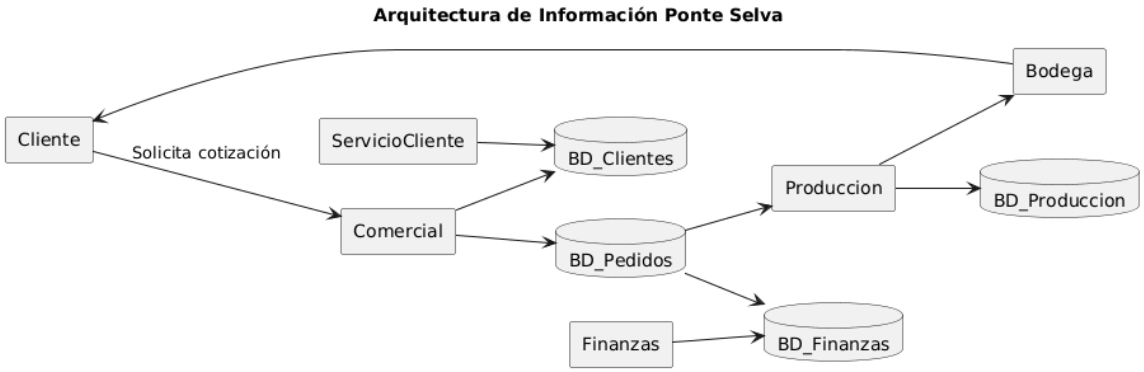


Figura 18. Fuente: Elaboración Propia

3.3 Arquitectura de Aplicaciones

3.3.1 Catálogo de Aplicaciones

Aplicación	Función
ERP	Gestión Empresarial
CRM	Clientes
Excel	Reportes
Power BI	Indicadores
Correo Corporativo	Comunicación
Sistema de Producción	Planificación

Tabla 3: Catálogo de Aplicaciones. Fuente: Elaboración Propia

3.3.2 Diagrama de Integración de Aplicaciones

Diagrama de Integración de aplicaciones

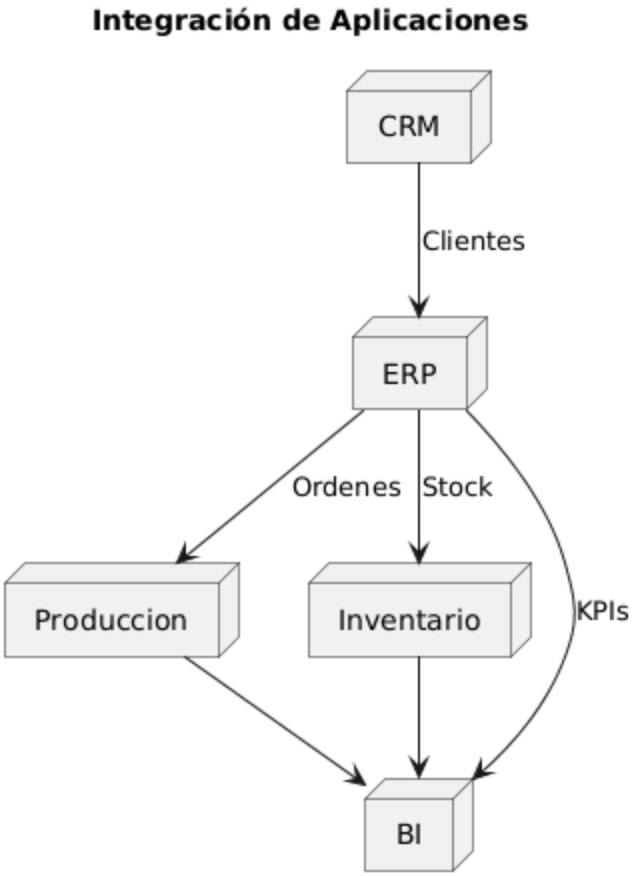


Figura 19. Fuente: Elaboración Propia

3.3.3 Matriz Aplicación – Proceso

Proceso	**Aplicación**
Ventas	ERP
Clientes	CRM
Reportes	Power BI
Producción	Sistema de Producción

Tabla 4: Matriz Aplicación - Proceso. Fuente: Elaboración Propia

3.4 Arquitectura Tecnológica

3.4.1 Inventario Tecnológico

- Hardware

- PCs Administrativas

- Laptops Gerenciales
- Servidor ERP
- Servidor Base de Datos
- Firewall
- Switches
- Impresoras

➢ Software

- Windows Server
- ERP
- CRM
- SQL Server
- Power BI

3.4.2 Infraestructura Física

Diagrama de Infraestructura Física

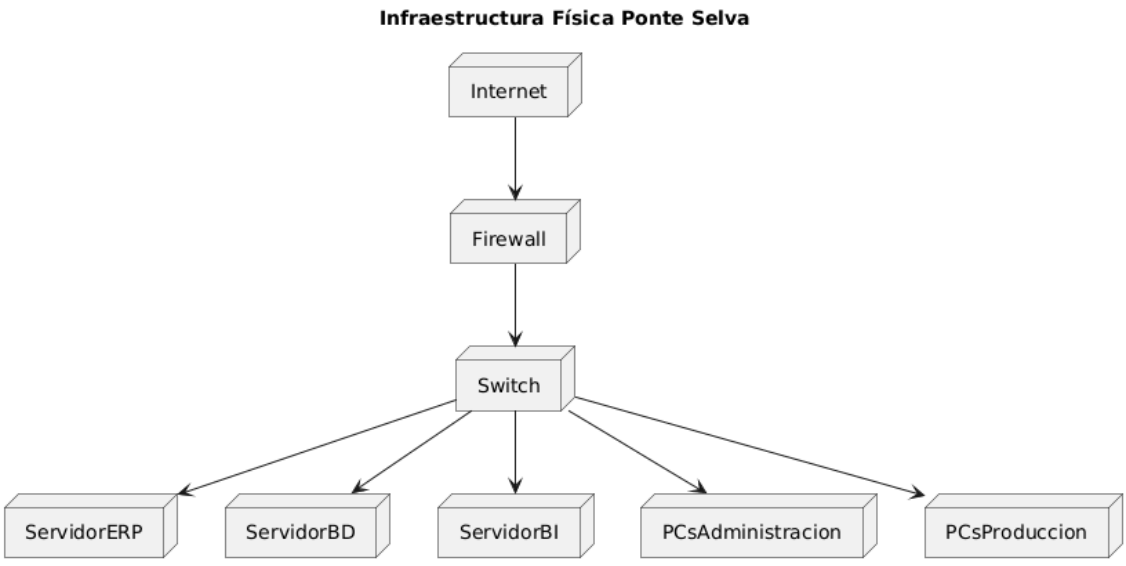


Figura 20. Fuente: Elaboración Propia

3.4.3 Arquitectura de red

Arquitectura de Red

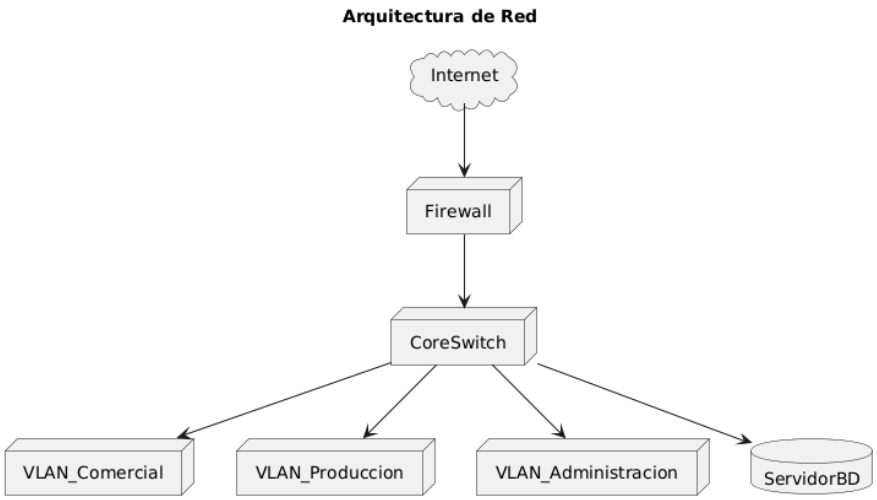


Figura 21. Fuente: Elaboración Propia

3.5 Arquitectura de Seguridad

3.5.1 Seguridad física

- Control de acceso
- CCTV
- Registro de visitantes
- Sensores de humo
- UPS
- Extintores

3.5.2 Seguridad Lógica

- Antivirus
- Firewall
- VPN
- MFA
- Gestión de accesos

- Segmentación de red

3.5.3 Diagrama de Seguridad

Diagrama de Arquitectura de Seguridad

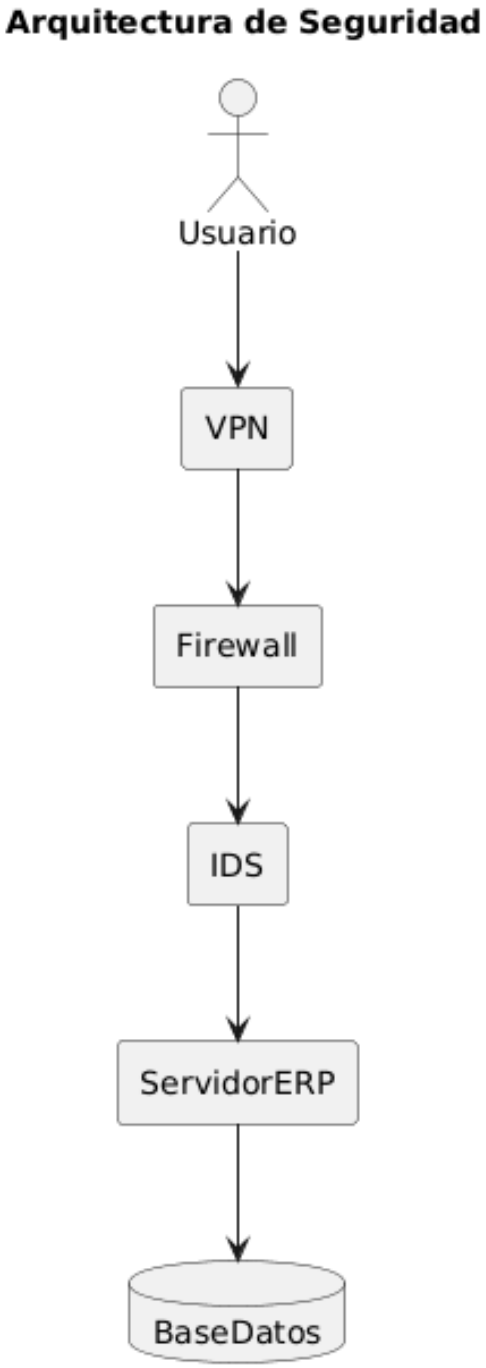


Figura 22. Fuente: Elaboración Propia

3.6 Arquitectura Objetivo (TO-BE)

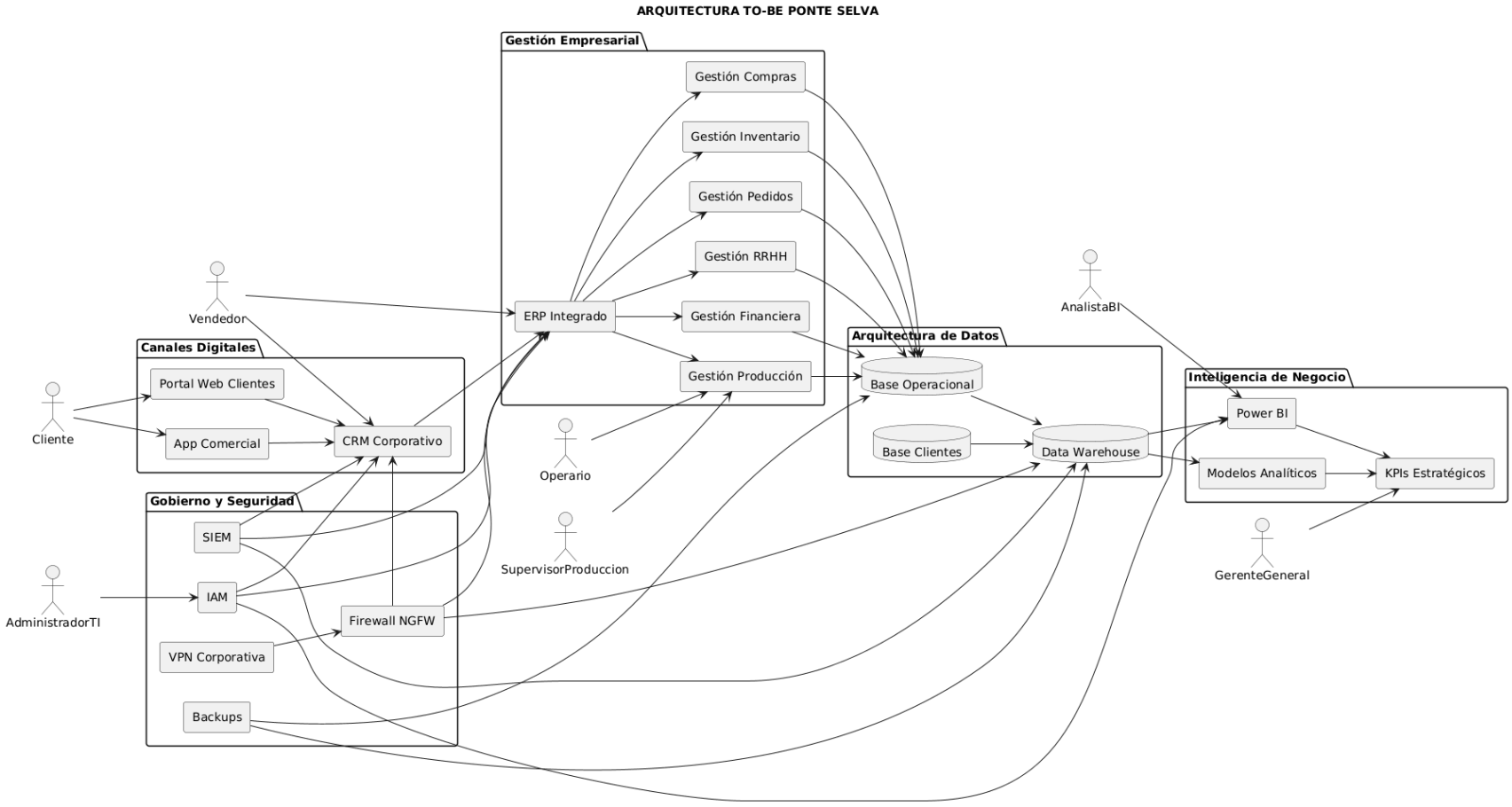


Figura 23. Fuente: Elaboración Propia

3.7 Arquitectura de Seguridad y Gobierno de Datos

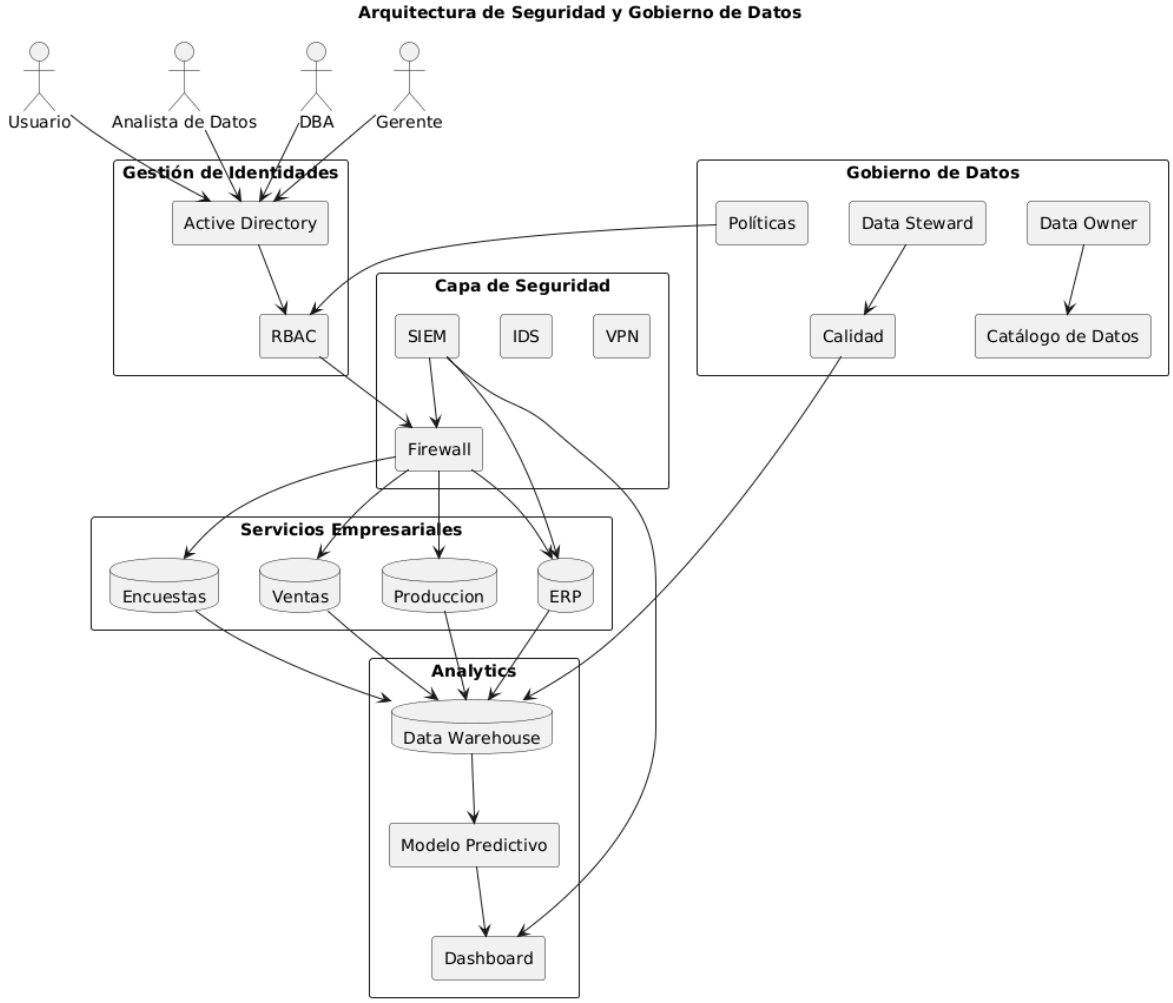


Figura 24. Fuente: Elaboración Propia

De esta forma, se identifican oportunidades de mejora relacionadas con la integración de información, la automatización de procesos y el aprovechamiento de analítica avanzada, lo cual justifica la construcción de una arquitectura TO-BE basada en TOGAF e integrada posteriormente con la metodología CRISP-DM.

3.8 Análisis GAP de la Arquitectura Empresarial

La propuesta de Arquitectura Empresarial incorpora tecnologías analíticas apoyadas en CRISP-DM que posibilitan una mejora en el sentido de ir más allá de una gestión tradicional y alcanzar un modelo organizacional orientado por datos (Data-Driven Enterprise),

permitiendo que tanto la arquitectura tecnológica como los procesos de negocio se vean fortalecidos.

Dominio de TOGAF	AS-IS	GAP entre AE	TO-BE
Negocio	Gestión basada en experiencia y reportes.	No existen mecanismos de analítica predictiva que apoyen la planificación comercial.	Gestión predictiva con soporte por parte de los modelos analíticos para anticipar tendencias de satisfacción
Datos	La información se almacena en bases de datos y hojas de cálculo.	Ausencia de gobierno enfocados en explotación analítica.	Gobierno de datos, integración de fuentes y disponibilidad de información para analítica avanzada.
Aplicaciones	ERP, sistema comercial y aplicaciones con manejo manual de información.	Baja interoperabilidad y falta de plataformas analíticas integradas.	Plataforma integrada de relacionada conexión al modelo predictivo
Tecnología	Infraestructura local con recursos limitados para procesamiento analítico.	Escasa capacidad para ejecutar modelos predictivos y escalar servicios analíticos.	Arquitectura con servicios en la nube, procesamiento analítico y despliegue automatizado del modelo predictivo.
Seguridad	Controles básicos de autenticación, respaldo y protección.	No existen mecanismos específicos para proteger modelos analíticos ni políticas de gobierno de datos.	Seguridad basada en control de accesos, cifrado de información, respaldo automatizado y monitoreo continuo.
Información	Los datos históricos únicamente sirven para consultas y reportes descriptivos.	La información no genera conocimiento para apoyar decisiones estratégicas.	Arquitectura orientada a datos donde la información provee a los modelos predictivos y dashboards ejecutivos.

Tabla 5: GAP entre arquitecturas. Fuente: Elaboración propia

4. CAPÍTULO IV: APLICACIÓN DE LA METODOLOGÍA CRISP-DM

En el presente capítulo, se desarrollarán las fases de CRISP-DM conforme a lo establecido en el Marco Teórico; esto alrededor de la empresa Ponte Selva con el propósito de analizar la retroalimentación de los clientes respecto a la calidad de los procesos de ventas. Una vez que se hayan completado las fases CRISP-DM, las conclusiones establecidas en la fase de despliegue serán las útiles para realizar TOGAF en base a la estructura de ADM.

Para organizar correctamente las fases de CRISP-DM y cada una de sus fases se utilizará el lenguaje de programación Python que es amigable con el analista de datos y con el entorno de trabajo. Streamlit será la herramienta que conectará el análisis de los datos con la proyección en una aplicación web. A continuación, se procederá con el desarrollo de la página que permitirá visualizar el proceso de Minería de Datos con CRISP-DM.

4.1 Desarrollo de aplicación web en Streamlit para CRISP-DM

4.1.1 Configuración de la Aplicación web

```python
#Importar las librerías a usar
import pandas as pd
import streamlit as st
import matplotlib.pyplot as plt
import numpy as np
#Importar librerías de modelos de Inteligencia Artificial
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score

#Establecer el nombre de la Página web
st.set_page_config(
    page_title="CRISP-DM Ponte Selva",
    layout="wide"
)

#Crear el menu lateral
st.sidebar.title("Menú principal")

opcion = st.sidebar.radio(
    "Selecciona una sección",
    [
        "Portada",
        "Carga de dataset",
        "Fase 1: Comprensión del negocio",
        "Fase 2: Comprensión de los datos",
        "Fase 3: Preparación de los datos",
        "Fase 4: Modelado",
        "Fase 5: Evaluación",
        "Fase 6: Despliegue"
    ]
)
```

Figura 25. Fuente: Elaboración Propia

4.1.2 Fase 1 de CRISP-DM en aplicación web

```python
#Crear condicional para comprobar que la opción elegida sea la Fase 1 de CRISP-DM
if opcion == "Fase 1: Comprensión del negocio":
    st.title("Fase 1 CRISP-DM: Comprensión del negocio")
#Escribir Misión
    st.header("Misión")
    st.write(
        "Ofrecer productos textiles de calidad a clientes nacionales, "
        "mediante procesos productivos y comerciales eficientes, manteniendo "
        "un servicio orientado a la satisfacción del cliente, la confiabilidad "
        "en la atención y la mejora continua de sus operaciones."
    )
#Escribir Visión
    st.header("Visión")
    st.write(
        "Ser líderes en el mercado textil reconocidos por sus productos "
        "de calidad, variedad y mejor precio, a través de recursos humanos "
        "comprometidos y satisfechos."
    )
#Escribir Objetivo Estratégico
    st.header("Objetivo estratégico")
    st.write(
        "Fabricar productos textiles de la más alta calidad mediante procesos "
        "de excelencia y mejora continua para satisfacer y exceder las "
        "expectativas de sus compradores locales e internacionales."
    )
#Escribir Objetivo de Minería de Datos
    st.header("Objetivo de minería de datos")
    st.write(
        "Analizar las calificaciones de retroalimentación de ventas de Ponte "
        "Selva para identificar patrones de satisfacción, comparar los canales "
        "de venta y predecir el comportamiento del score como apoyo a la mejora "
        "del proceso comercial."
    )
```

Figura 26. Fuente Elaboración Propia

4.1.3 Fase 2 de CRISP-DM en aplicación web

```python
#Comprobar opción del menú
if opcion == "Fase 2: Comprensión de los datos":
    st.title("Fase 2 CRISP-DM: Comprensión de los datos")

    df = st.session_state["df"]

    st.header("Descripción general del dataset")

    col1, col2, col3 = st.columns(3)
    col1.metric("Número de registros", len(df))
    col2.metric("Número de columnas", len(df.columns))
    col3.metric("Duplicados exactos", df.duplicated().sum())
#Crear vista previa del dataset
    st.subheader("Vista previa del dataset")
    st.dataframe(df.head(20))

    st.subheader("Tipos de datos")
    tipos = pd.DataFrame({
        "Columna": df.columns,
        "Tipo de dato": df.dtypes.astype(str).values
    })
    st.dataframe(tipos)
#Verificar la calidad de los datos
    st.header("Calidad de datos")

    nulos = df.isnull().sum().reset_index()
    nulos.columns = ["Columna", "Valores nulos"]
    st.dataframe(nulos)
#Detección de nulos
    canales_validos = ["WA", "PRE", "TEL", "MAIL"]

    scores_invalidos = df[
        (df["score"] < 1) | (df["score"] > 10)
    ]

    canales_invalidos = df[
        ~df["canal_venta"].isin(canales_validos)
    ]
#Detectar valores fuera de rango
    st.write("Scores fuera del rango 1 a 10:", len(scores_invalidos))
    st.write("Canales inválidos:", len(canales_invalidos))

    st.header("Análisis descriptivo del score")
    st.dataframe(df["score"].describe())
#Crear histogramas
    fig, ax = plt.subplots()
    ax.hist(df["score"].dropna(), bins=10)
    ax.set_title("Distribución de calificaciones")
    ax.set_xlabel("Score")
    ax.set_ylabel("Frecuencia")
    st.pyplot(fig)
```

Figura 27. Fuente: Elaboración Propia

4.1.4 Fase 3 de CRISP-DM en aplicación web

```python
#Verificar opción del menú
if opcion == "Fase 3: Preparación de los datos":
    st.title("Fase 3 CRISP-DM: Preparación de los datos")

    df = st.session_state["df"].copy()

# Eliminar registros con fecha o score vacio
    df_limpio = df.dropna(subset=["fecha_enc_vent", "score"])

# Filtrar scores válidos
    df_limpio = df_limpio[
        (df_limpio["score"] >= 1) &
        (df_limpio["score"] <= 10)
    ]

# Filtrar canales válidos
    canales_validos = ["WA", "PRE", "TEL", "MAIL"]
    df_limpio = df_limpio[
        df_limpio["canal_venta"].isin(canales_validos)
    ]

# Crear variables temporales
    df_limpio["anio"] = df_limpio["fecha_enc_vent"].dt.year
    df_limpio["mes"] = df_limpio["fecha_enc_vent"].dt.month
    df_limpio["trimestre"] = df_limpio["fecha_enc_vent"].dt.quarter

# Crear variable de nivel de satisfacción
    def clasificar_satisfaccion(score):
        if score >= 8:
            return "Alta"
        elif score >= 6:
            return "Media"
        else:
            return "Baja"

    df_limpio["nivel_satisfaccion"] = df_limpio["score"].apply(
        clasificar_satisfaccion
    )

    st.session_state["df_limpio"] = df_limpio

    st.success("Datos preparados correctamente")

    st.write("Registros originales:", len(df))
    st.write("Registros preparados:", len(df_limpio))

    st.dataframe(df_limpio.head())

    # Gráfico de niveles de satisfacción
    fig, ax = plt.subplots()
    df_limpio["nivel_satisfaccion"].value_counts().plot(
        kind="bar",
        ax=ax
    )
    ax.set_title("Distribución del nivel de satisfacción")
    ax.set_xlabel("Nivel de satisfacción")
    ax.set_ylabel("Cantidad")
    st.pyplot(fig)
```

Figura 28. Fuente: Elaboración Propia

4.1.5 Preparación de Serie de Tiempo

```python
# Agrupar el score por día
serie_diaria = (
    df_limpio
    .groupby("fecha_enc_vent")["score"]
    .mean()
    .reset_index()
    .sort_values("fecha_enc_vent")
)

# Crear un calendario diario completo
rango_fechas = pd.date_range(
    start=serie_diaria["fecha_enc_vent"].min(),
    end=serie_diaria["fecha_enc_vent"].max(),
    freq="D"
)

serie_diaria = (
    serie_diaria
    .set_index("fecha_enc_vent")
    .reindex(rango_fechas)
)

serie_diaria.index.name = "fecha_enc_vent"
serie_diaria = serie_diaria.reset_index()

# Identificar días con encuestas reales
serie_diaria["tiene_encuesta"] = serie_diaria["score"].notna()

# Interpolar días sin encuesta
serie_diaria["score_preparado"] = (
    serie_diaria["score"]
    .interpolate(method="linear")
    .bfill()
    .ffill()
)

# Crear índice temporal
serie_diaria["t"] = np.arange(len(serie_diaria))

# Calcular medias móviles
serie_diaria["media_movil_7"] = (
    serie_diaria["score_preparado"]
    .rolling(window=7, min_periods=1)
    .mean()
)

serie_diaria["media_movil_30"] = (
    serie_diaria["score_preparado"]
    .rolling(window=30, min_periods=1)
    .mean()
)

st.session_state["serie_diaria"] = serie_diaria
```

Figura 29. Fuente: Elaboración Propia

4.1.6 Fase 4 de CRISP-DM en aplicación web

```python
if opcion == "Fase 4: Modelado":
    st.title("Fase 4 CRISP-DM: Modelado")

    serie_diaria = st.session_state["serie_diaria"]

    # Variables del modelo
    X = serie_diaria[["t"]]
    y = serie_diaria["score_preparado"]

    # División entrenamiento/prueba
    limite = int(len(serie_diaria) * 0.8)

    X_train = X.iloc[:limite]
    y_train = y.iloc[:limite]

    X_test = X.iloc[limite:]
    y_test = y.iloc[limite:]

    # Entrenamiento del modelo
    modelo = LinearRegression()
    modelo.fit(X_train, y_train)

    # Predicción sobre datos de prueba
    y_pred = modelo.predict(X_test)

    # Entrenamiento final con todos los datos
    modelo_final = LinearRegression()
    modelo_final.fit(X, y)

    # Pronóstico futuro
    horizonte = 30

    t_futuro = np.arange(
        len(serie_diaria),
        len(serie_diaria) + horizonte
    ).reshape(-1, 1)

    fechas_futuras = pd.date_range(
        start=serie_diaria["fecha_enc_vent"].max() + pd.Timedelta(days=1),
        periods=horizonte,
        freq="D"
    )

    score_futuro = modelo_final.predict(t_futuro)
    score_futuro = np.clip(score_futuro, 1, 10)

    pronostico = pd.DataFrame({
        "fecha_enc_vent": fechas_futuras,
        "score_predicho": score_futuro
    })

    st.session_state["modelo"] = modelo
    st.session_state["y_test"] = y_test
    st.session_state["y_pred"] = y_pred
    st.session_state["pronostico"] = pronostico

    st.subheader("Pronóstico de score post venta")
    st.dataframe(pronostico)
    #Graficar histograma
    fig, ax = plt.subplots()
    ax.plot(
        serie_diaria["fecha_enc_vent"],
        serie_diaria["score_preparado"],
        label="Score histórico preparado"
    )
    ax.plot(
        pronostico["fecha_enc_vent"],
        pronostico["score_predicho"],
        label="Pronóstico"
    )
    ax.set_title("Pronóstico de calificaciones post venta")
    ax.set_xlabel("Fecha")
    ax.set_ylabel("Score")
    ax.legend()
    st.pyplot(fig)
```

Figura 30. Fuente: Elaboración Propia

4.1.7 Fase 5 de CRISP-DM en aplicación web

```python
if opcion == "Fase 5: Evaluación":
    st.title("Fase 5 CRISP-DM: Evaluación")

    y_test = st.session_state["y_test"]
    y_pred = st.session_state["y_pred"]

    # Cálculo de métricas
    mae = mean_absolute_error(y_test, y_pred)
    mse = mean_squared_error(y_test, y_pred)
    rmse = np.sqrt(mse)
    r2 = r2_score(y_test, y_pred)

    col1, col2, col3, col4 = st.columns(4)

    col1.metric("MAE", round(mae, 4))
    col2.metric("MSE", round(mse, 4))
    col3.metric("RMSE", round(rmse, 4))
    col4.metric("R²", round(r2, 4))

    st.write(
        "El MAE indica el error promedio absoluto. "
        "El RMSE permite interpretar el error en la misma escala del score. "
        "El R² muestra qué proporción de la variabilidad del score es explicada por el modelo."
    )

    # Comparación gráfica
    fig, ax = plt.subplots()
    ax.plot(y_test.values, label="Score real")
    ax.plot(y_pred, label="Score predicho")
    ax.set_title("Comparación entre score real y score predicho")
    ax.set_xlabel("Observaciones de prueba")
    ax.set_ylabel("Score")
    ax.legend()
    st.pyplot(fig)
```

Figura 31. Fuente: Elaboración Propia

4.1.8 Fase 6 de CRISP-DM en aplicación web

```python
if opcion == "Fase 6: Despliegue":
    st.title("Fase 6 CRISP-DM: Despliegue")

    serie_diaria = st.session_state["serie_diaria"]
    pronostico = st.session_state["pronostico"]

    ultimo_score = serie_diaria["score_preparado"].iloc[-1]
    score_futuro = pronostico["score_predicho"].iloc[-1]

    diferencia = score_futuro - ultimo_score
#Establecer valores de si la tendencia sube, baja o se mantiene
    if diferencia > 0.10:
        tendencia = "Subir"
    elif diferencia < -0.10:
        tendencia = "Bajar"
    else:
        tendencia = "Mantenerse"
#Entregar resultado final
    st.header("Resultado principal")
    st.write("Último score preparado:", round(ultimo_score, 2))
    st.write("Score futuro estimado:", round(score_futuro, 2))
    st.write("Tendencia estimada:", tendencia)
#Recomendaciones para la empresa
    st.header("Recomendaciones")
    st.write(
        "- Monitorear mensualmente el score promedio post venta."
    )
    st.write(
        "- Revisar los canales de venta con menor calificación promedio."
    )
    st.write(
        "- Incorporar comentarios textuales en futuras encuestas."
    )
    st.write(
        "- Crear tableros de indicadores para seguimiento comercial."
    )
    st.write(
        "- Utilizar los resultados como entrada para la arquitectura empresarial basada en TOGAF."
    )

```

Figura 32. Fuente: Elaboración Propia

Para elaborar la aplicación web, se usó Python con Streamlit, Pandas, Matplotlib y Scikit-learn. El sistema se organizó en fases como las del CRISP-DM. Esto significa que cada paso del proceso de minería de datos se puede hacer en una sección aparte. La aplicación permite cargar y analizar los datos, prepararlos, hacer una serie de tiempo, generar un pronóstico, evaluar el modelo y mostrar resultados que ayudan a tomar decisiones.

4.2 Fase 1 CRISP-DM: Comprensión del negocio

Misión: Ofrecer productos textiles de calidad a clientes nacionales, mediante procesos productivos y comerciales eficientes, manteniendo un servicio orientado a la satisfacción del cliente, la confiabilidad en la atención y la mejora continua de sus operaciones.

Visión: Ser líderes en el mercado textil reconocidos por sus productos de calidad, variedad y mejor precio, a través de recursos humanos comprometidos y satisfechos.

Valores:

- Compromiso
- Excelencia
- Trabajo en equipo
- Cumplimiento
- Orientación al cliente
- Visión de futuro

Objetivo estratégico:

Fabricar productos textiles de la más alta calidad mediante procesos de excelencia y mejora continua para satisfacer y exceder las expectativas de sus compradores locales e internacionales.

KPI:

Nombre de KPI	Métrica
Tasa de devoluciones por No Conformidad	$\frac{\#\ de\ pedidos\ devueltos\ por\ fallas\ de\ calidad}{Total\ de\ pedidos\ entregados} \times 100$
Nivel de cumplimiento de Entregas	$\frac{Pedidos\ a\ tiempo\ y\ con\ cantidad\ exacta}{Total\ de\ pedidos\ despachados} \times 100$
Índice de Retención de Clientes	$\frac{Clientes\ activos\ al\ final\ del\ periodo - clientes\ nuevos}{Clientes\ activos\ al\ inicio\ del\ periodo} \times 100$

Tabla 6: KPI del objetivo estratégico. Fuente: Elaboración propia

Objetivo de Minería de Datos:

Analizar las calificaciones de retroalimentación de ventas de Ponte Selva para identificar patrones de satisfacción, comparar los canales de venta y predecir el comportamiento del score como apoyo a la mejora del proceso comercial.

Fuente de los datos: Dataset de las encuestas a los clientes posterior a la venta.

4.3 Fase 2 CRISP-DM: Comprensión de los Datos

Se revisa la estructura inicial del dataset, se describen sus variables, se exploran los datos y se identifican problemas de calidad que deberán tratarse en la fase de preparación.

4.3.1 Descripción general de los datos

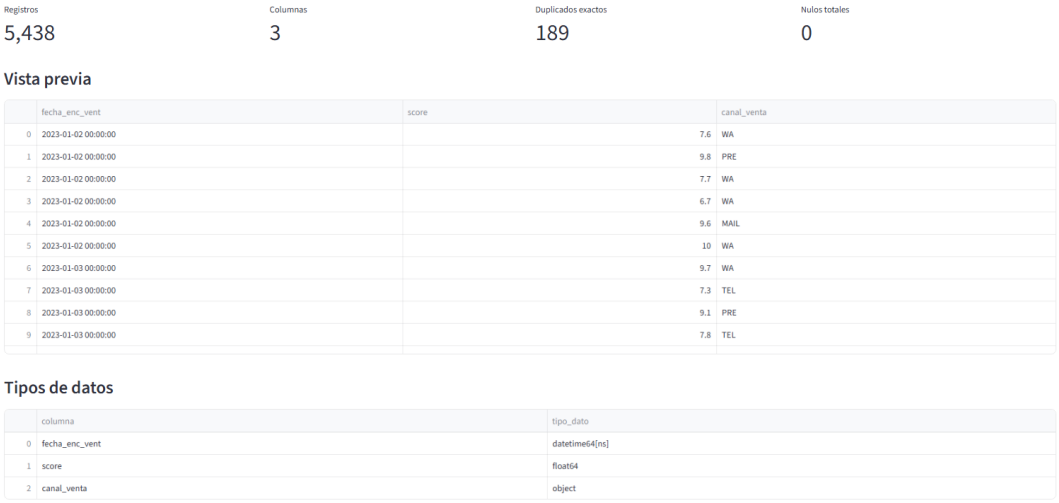


Figura 33. Fuente: Elaboración propia

Se observa que los registros se hacen de manera diaria, más de uno por día, esto se debe a que puede haber más de una venta en el día pero que se obtuvo un resultado al proceso de retroalimentación del cliente.

4.3.2 Diccionario y Calidad de los datos

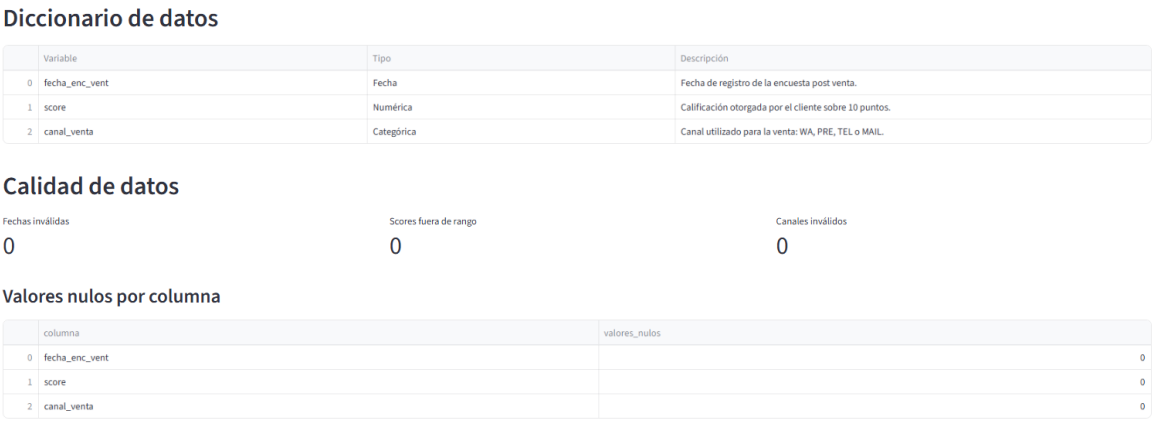


Figura 34. Fuente: Elaboración propia

Existen 3 atributos dentro de los datos: la fecha de la encuesta de valoración, la calificación del cliente y el canal por el cual se realizó la venta. No hay valores nulos dentro del dataset, posiblemente ya que los registros de descargan de manera automática por las fechas por lo que no se registran encuestas vacías.

Ya que los resultados están en un rango de 0 a 10, no es necesario realizar control de outliers, especialmente porque no hay una condición para que se manejen valores en un rango determinado en este estudio más que el de las propias calificaciones de los clientes.

4.3.3 Exploración inicial de variable objetivo

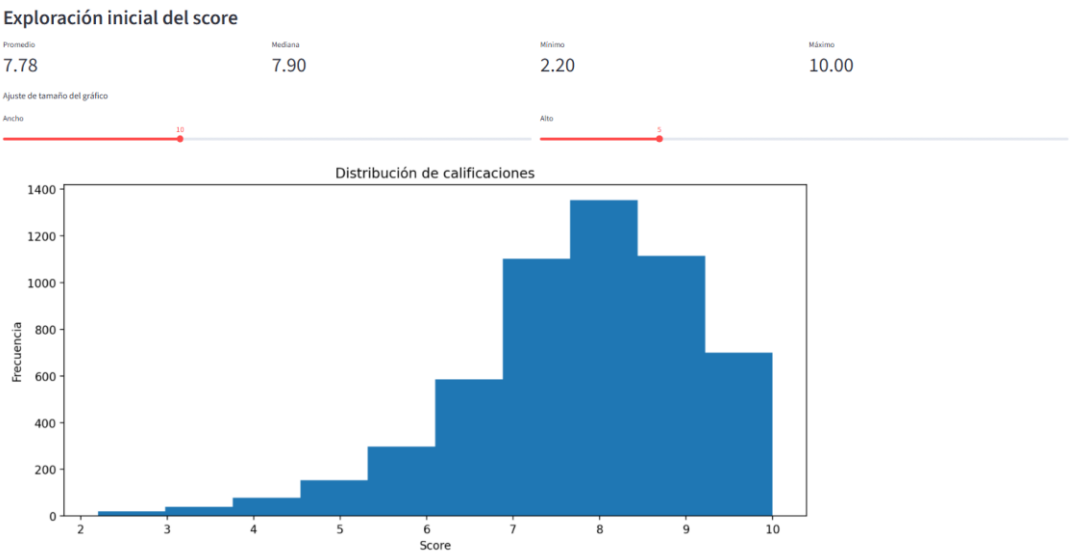


Figura 35. Fuente: Elaboración Propia

Existe en promedio una calificación de 7.78 sobre 10 al servicio de venta de la empresa. El valor más bajo ha sido de 2.20 sobre 10 y el máximo ha llegado a ser de 10/10; además, el score más común de 7.78 tiene una frecuencia de 1375 registros aproximadamente.

4.3.4 Exploración temporal en función del Score

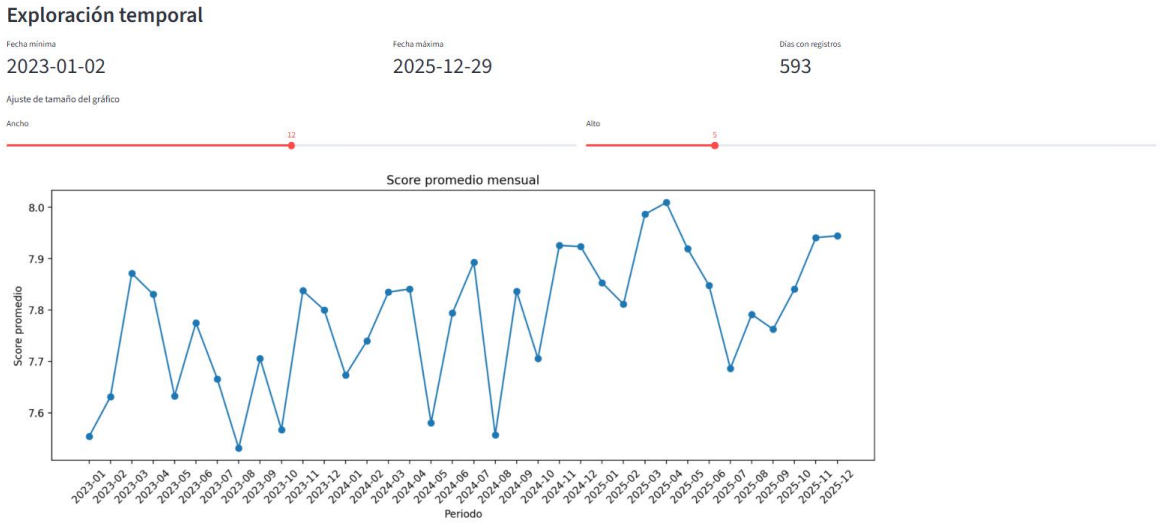
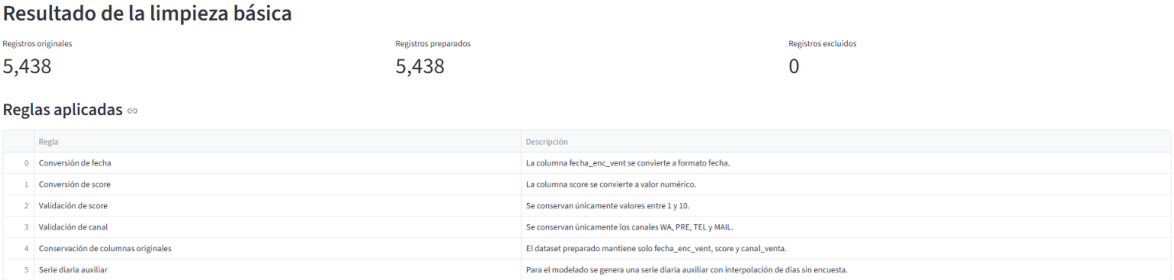


Figura 36. Fuente: Elaboración Propia

Se presenta el gráfico de serie temporal donde se muestra el periodo mensual contra su Score Promedio dado por el cliente post venta. Basado en este diagrama, se ve una tendencia general al alza del score, a pesar de tener un punto más bajo en enero de 2023. Sin embargo, su punto más alto se lo presenta en abril del 2025.

4.4 Fase 3 CRISP-DM: Preparación de los datos

4.4.1 Limpieza y aplicación de reglas



	Regla	Descripción
0	Conversión de fecha	La columna fecha_enc_vent se convierte a formato fecha.
1	Conversión de score	La columna score se convierte a valor numérico.
2	Validación de score	Se conservan únicamente valores entre 1 y 10.
3	Validación de canal	Se conservan únicamente los canales WA, PRE, TEL y MAIL.
4	Conservación de columnas originales	El dataset preparado mantiene solo fecha_enc_vent, score y canal_venta.
5	Serie diaria auxiliar	Para el modelado se genera una serie diaria auxiliar con interpolación de días sin encuesta.

Figura 37. Fuente: Elaboración Propia

Ya que los datos venían registrados de manera diaria y con la variable score en un rango predeterminado, no fue necesario realizar una limpieza de los datos, no existían nulos ni registros con un formato incompatible. Por otro lado, es hizo la conversión de los tipos de los datos que podrían causar problema en el modelado, ya que no podrían llegar a ser no leíbles por el modelo.

El atributo de la fecha de convirtió a formato “Date”. El Score o calificación del cliente se convirtió en numérico decimal. Se comprobó que no existan valores fuera del rango de 1-10, además de comprobar que los canales de venta sean únicamente los indicados como correctos.

4.4.2 Distribución de variable objetivo después de limpieza

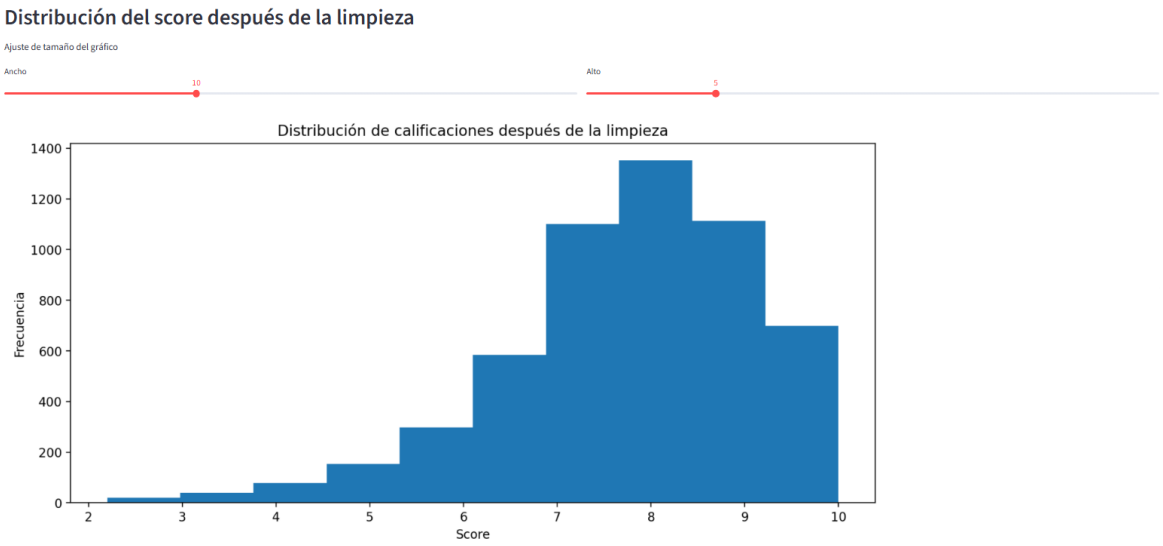


Figura 38. Fuente: Elaboración Propia

La frecuencia es mayor en aquellos valores de Score cercanos a 8, siguiendo el patrón de su media previamente calculado. Los valores con menor frecuencia son entre 2 y 4, manteniendo el valor mínimo que se obtuvo en la Fase 2 de CRISP-DM.

4.4.3 Interpolación de la serie

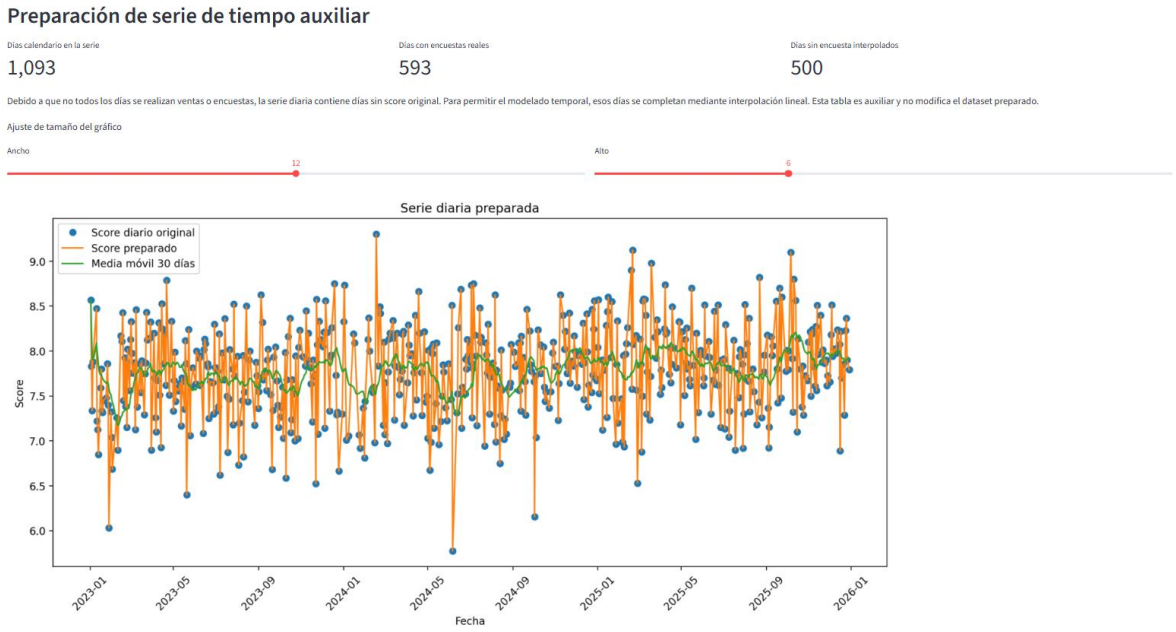


Figura 39. Fuente: Elaboración Propia

Ya que los datos fueron tomados diariamente, existen días en los que no se realizaron ventas, por tanto, no hay registro de retroalimentación del cliente. Es así que, es necesario crear los registros de los datos, a pesar de que no contengan valores reales para que se mantenga la lógica de registros hechos de manera consecutiva. Esto asegura una mayor calidad en las predicciones y mejores valores en las métricas de evaluación.

4.5 Fase 4 CRISP-DM: Modelado

4.5.1 Aplicación del modelo



Figura 40. Fuente: Elaboración Propia

La tendencia del modelo una vez hechas las predicciones es de bajar las calificaciones de sus clientes. Actualmente se tiene un score promedio de 7.94 con los datos entregados, sin embargo, se espera un score futuro estimado de 7.80.

- Pronóstico a 6 meses

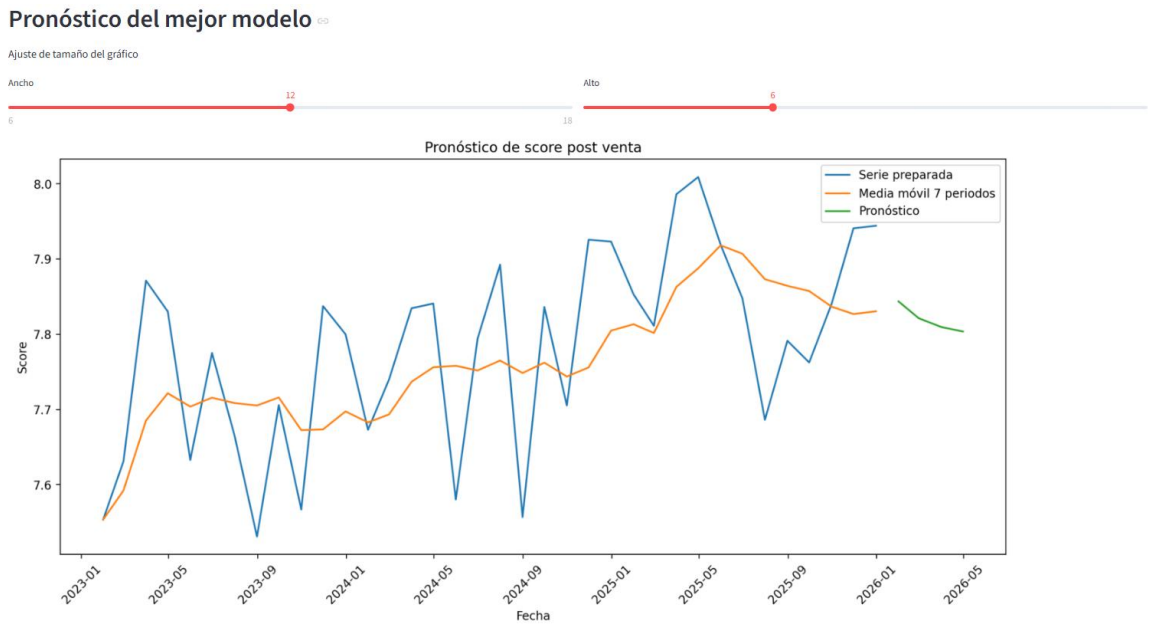


Figura 41. Fuente: Elaboración Propia

- Registro de pronósticos

fecha_enc_vent	score
2026-01-31 00:00:00	7.843721776609348
2026-02-28 00:00:00	7.821275613220569
2026-03-31 00:00:00	7.809642279969676
2026-04-30 00:00:00	7.803612989342932

Tabla 7. Fuente: Elaboración Propia

Interpretación del modelo: El modelo estima que las calificaciones post venta tenderán a disminuir en el horizonte seleccionado.

4.6 Fase 5 CRISP-DM: Evaluación

Comparación de resultados SARIMAX y Prophet

Comparación de métricas SARIMAX y Prophet

	Modelo	Tipo	MAE	MSE	RMSE	R2
0	SARIMAX(1, 0, 1)x(0, 0, 0, 0) - sin variables externas	SARIMAX	0.1487	0.0377	0.1941	-0.1393
1	SARIMAX(1, 0, 1)x(1, 0, 0, 4) - sin variables externas	SARIMAX	0.1514	0.0379	0.1948	-0.1475
2	SARIMAX(1, 0, 1)x(1, 0, 0, 4) - volumen y canales	SARIMAX	0.1524	0.0403	0.2009	-0.2206
3	SARIMAX(1, 0, 1)x(0, 0, 0, 0) - volumen y canales	SARIMAX	0.1556	0.0421	0.2051	-0.2731
4	SARIMAX(1, 1, 1)x(0, 0, 0, 0) - sin variables externas	SARIMAX	6.8348	46.747	6.8372	-1413.1698
5	SARIMAX(1, 1, 1)x(0, 0, 0, 0) - volumen y canales	SARIMAX	6.8348	46.747	6.8372	-1413.1698

Tabla 8. Fuente: Elaboración Propia

MAE: 0.0089

RSME: 0.0943

R^2: -0.2004

Las predicciones realizadas por el modelo se alejan menos de 0.10 puntos respecto a las predicciones reales en su escala de 1 a 10. Por tanto, el modelo resulta útil para aproximar el posible comportamiento futuro de las calificaciones de sus clientes postventa. Sin embargo, el valor de R^2 se mantiene en negativos, quiere decir que no se puede interpretar de manera definitiva los resultados de comportamiento del cliente ya que no existen suficientes atributos en el dataset que muestren el porqué de los resultados de score.

4.7 Fase 6 CRISP-DM: Despliegue

La tendencia futura muestra que las calificaciones de los clientes postventa se reducirán en meses futuros. Quiere decir que es necesario implementar cambios en los procesos de la empresa para mejorar la relación con los clientes. Entre estas decisiones están priorizar las ventas, la calidad del servicio, monitorear constantemente los resultados de las encuestas, entre otros aspectos que se deben tratar directamente con el área de servicio al cliente de Ponte Selva.

Enlace a la aplicación web desplegada: https://crispdm.streamlit.app/

4.8 Relación de los resultados de CRISP-DM con TOGAF

Como resultado de la fase de implementación de CRISP-DM, se identificó que la tendencia futura de las calificaciones postventa podría mostrar una reducción en los próximos períodos. Este comportamiento indica que, si no se toman acciones correctivas, la percepción de los clientes sobre la calidad del proceso de ventas podría verse afectada. Además, los resultados que se obtuvieron no son aislados, sus pronósticos son valiosos para la toma de decisiones estratégicas en la empresa.

Es en este punto que se crea la relación entre CRISP-DM y TOGAF con ADM (Architecture Development Method). Por un lado, la metodología de Minería de Datos permitió el análisis, identificación de patrones y finalmente anticipar la posible disminución del score de los clientes, TOGAF permitirá transformar esos resultados en una propuesta estructurada para la arquitectura empresarial. Este pronóstico puede ser incorporado como una necesidad de cambio dentro del ciclo de ADM. Es así como se este análisis se convierte en un elemento clave para orientar la transformación empresarial.

5. CAPÍTULO V: DISEÑO DE ARQUITECTURA EMPRESARIAL TOGAF/ADM

5.1 Relación entre CRISP-DM y TOGAF/ADM

Resultado de CRISP-DM	Necesidad a tratar	Relación establecida
Tendencia a decrecer en el score postventa	Mejorar la experiencia del cliente	Evidencia la necesidad de proponer una arquitectura orientada a fortalecer la satisfacción del cliente y mejorar el proceso postventa.
Necesidad de monitoreo de score continuo	Controlar los indicadores de desempeño	Se relaciona con la Arquitectura de Datos y la Arquitectura de Aplicaciones para mejorar el análisis y visualización periódica.
Mantener alta calidad de servicio	Fortalecer los procesos de atención y seguimiento al cliente	Se relaciona con la Arquitectura de Negocio, porque implica revisar y mejorar los procesos comerciales y servicio al cliente, así como los mecanismos de respuesta ante calificaciones bajas.

Tabla 9: Relación CRISP-DM Y TOGAF. Fuente: Elaboración Propia

5.2 Aplicación del ciclo ADM de TOGAF

5.2.1 Fase preliminar

En esta etapa se define el alcance, los actores involucrados, los principios de arquitectura y el propósito general de la propuesta. El objetivo no es rediseñar toda la organización, sino establecer una arquitectura que permita monitorear de manera continua la percepción del cliente y responder oportunamente ante una posible reducción en la calidad.

Los actores involucrados en este proceso son:

- Gerencia general
- Área comercial
- Servicio al cliente
- Área de tecnología de la información

- Analista de datos
- Clientes

Cada uno de ellos cumple un papel importante dentro de la arquitectura. La gerencia general se encarga de aprobar las decisiones estratégicas. Mientras que el área comercial y el servicio al cliente trabajan juntos para ejecutar acciones que mejoren el servicio. El área de TI brinda soporte y apoyo tecnológico a las soluciones. El analista de datos se enfoca en generar información valiosa. Por último, los clientes juegan un papel crucial al entregar retroalimentación a través de encuestas. Esta información es vital para la mejora continua.

5.2.2 Fase A: Visión de Arquitectura

Diseñar una arquitectura empresarial que integre el monitoreo del score postventa con los procesos de atención y mejora continua de Ponte Selva, lo que permitirá anticipar disminuciones en la satisfacción del cliente y orientar acciones correctivas basadas en datos.

- Problema a resolver: Posible disminución del score post venta.
- Necesidad principal: Mejorar la experiencia del cliente
- Objetivo de la Arquitectura: Integrar procesos, aplicaciones y tecnología para controlar la satisfacción
- Resultado esperado: Arquitectura capaz de monitorear y mejorar el servicio post venta
- Beneficio esperado: Aumento de la satisfacción del cliente

5.2.3 Fase B: Arquitectura de Negocio

La arquitectura de negocio define los procesos, roles y actividades que deben trabajarse para alcanzar la visión propuesta. En la empresa, el proceso central es el seguimiento postventa y la gestión de la satisfacción del cliente. Ante la tendencia decreciente del score, se plantea formalizar un proceso de monitoreo continuo, que permita

revisar periódicamente las calificaciones, identificar variaciones negativas y establecer acciones correctivas desde las áreas responsables.

Proceso actual	Limitación identificada	Proceso objetivo
Registro de encuestas postventa	La información no se revisa de manera periódica	Revisión continua del score postventa
Atención posterior a la venta	Respuesta únicamente al detectar un bajo valor	Seguimiento activo ante disminuciones del score
Evaluación de satisfacción	Análisis limitado de los resultados	Control constante mediante indicadores
Gestión de mejora	Acciones independientes a los datos de las encuestas	Acciones correctivas basadas en el comportamiento del score

Tabla 10: Limitaciones y procesos objetivos. Fuente: Elaboración Propia

Como resultado de la fase B de ADM, se plantea que el área comercial y servicio al cliente incorpore una rutina formal de revisión del score postventa. Esta rutina puede ser ejecutada mensualmente, generando reportes de satisfacción, identificando caídas relevantes y definiendo responsables para aplicar acciones de mejora.

5.2.4 Fase C: Arquitectura de Datos (sistemas de información)

Define cómo debe organizarse, almacenarse y utilizarse la información necesaria para formar la arquitectura. En este caso, los datos son las encuestas postventa, especialmente la fecha de registro y el score asignado por los clientes. Estos datos tienen valor estratégico porque permiten anticipar tendencias futuras de satisfacción, por ello, la empresa debe asegurar que la información postventa sea registrada de forma ordenada, completa y confiable.

Datos de alto valor:

- Fecha de encuesta postventa: Permite un análisis de tiempo y seguimiento histórico de los resultados.
- Score post venta: Mide la percepción del cliente acerca del servicio recibido.

- Historial de los scores: Identifica las tendencias y variaciones en función del tiempo.

Con estos datos, para establecer la arquitectura de datos, es necesario consolidar un repositorio de registros post venta que permita almacenar las encuestas, conservar el historial de calificaciones y generar indicadores periódicos.

5.2.5 Fase D: Arquitectura Tecnológica

Esta fase define la infraestructura necesaria para soportar las aplicaciones, datos y procesos propuestos. En este caso, se requiere una infraestructura que permita ejecutar la aplicación de CRISP-DM Y almacenar los datos postventa.

Componentes en la tecnología

Aplicación web	Permitir el acceso al análisis CRISP-DM
Repositorio de datos	Almacenar el historial de encuestas postventa
Analizador de registros en Python	Ejecutar preparación de datos, modelado y evaluación
Hosting para aplicación web	Publicar la aplicación para acceso de usuarios (analista de datos)
Generador de reportes	Exportar resultados y apoyar la toma de decisiones

Tabla 11: Componentes de la arquitectura tecnológica. Fuente: Elaboración Propia

Esta arquitectura tecnológica permitirá que la empresa realice un monitoreo periódico del score postventa y pueda actualizar el análisis conforme se incorporen nuevos registros.

5.2.6 Fase E: Oportunidades y Soluciones

En la Fase E se identifican las soluciones que permitirá a la empresa avanzar desde la situación actual hacia la arquitectura objetivo. Las oportunidades se relacionan con volver

formal y técnico el seguimiento postventa, la creación de un repositorio de datos, el uso de una aplicación para análisis y la generación de reportes.

Oportunidades y Soluciones

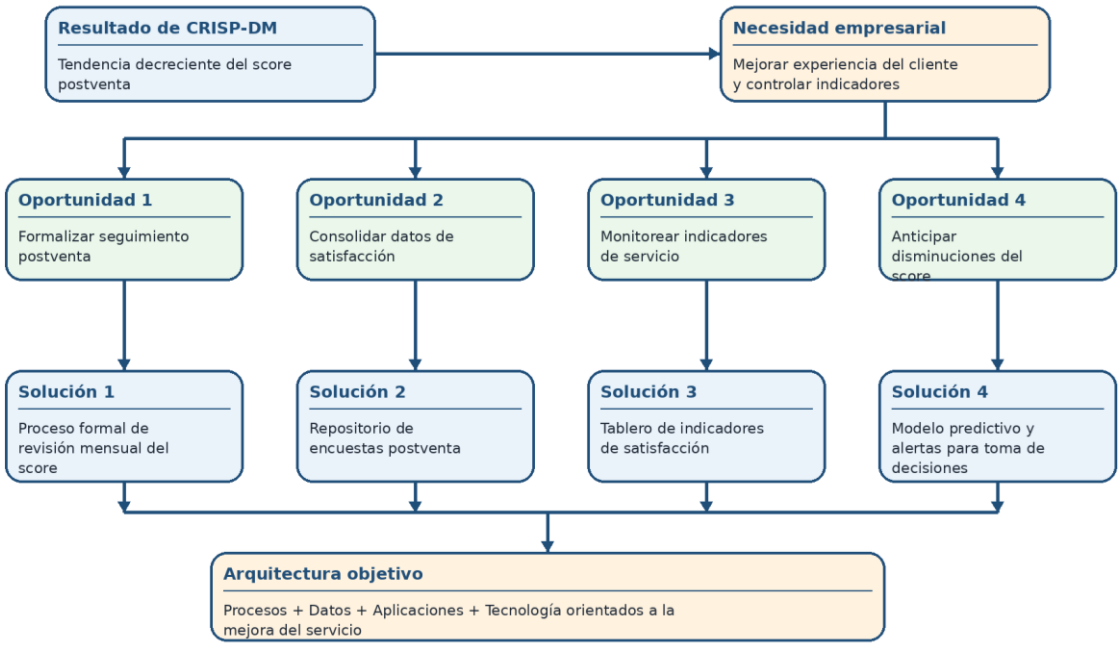


Figura 42. Fuente: Elaboración Propia

5.2.7 Fase F: Planificación de la Migración

Establece cómo implementar progresivamente las soluciones propuestas. Se realiza una implementación por etapas, iniciando con el ordenamiento de los datos postventa y finaliza con la operación continua del sistema de monitoreo.

No. Etapa	Actividad a Realizar	Resultado Esperado
1.	Reunir los datos de las encuestas post ventas	Bases sólidas para análisis
2.	Validar la calidad de los datos	Datos confiables para el sistema
3.	Implementación del aplicativo para análisis	Herramienta para monitoreo y modelado
4.	Generar los reportes cada determinado plazo	Seguimiento continuo del score post venta
5.	Definir los responsables de la revisión de reportes	Manejo de las acciones de mejora
6.	Evaluar los resultados de mejoras	Retroalimentación posterior al proceso. Útil para posibles futuros ciclos ADM

Tabla 12: Planificación de la migración. Fuente: Elaboración Propia

Architecture Roadmap

Proyecto	Prioridad	Tiempo
Gobierno de datos	Muy alta	3 meses
Limpieza de datos históricos	Muy alta	3 meses
Integración ERP con BD	Alta	4 meses
Data Warehouse	Alta	5 meses
Dashboards para BI	Alta	6 meses
Implementación CRISP-DM	Muy alta	7 meses
Desarrollo modelo SARIMAX	Alta	8 meses
Integración del modelo predictivo	Alta	10 meses
Plataforma de monitoreo	Media	11 meses

Tabla 13. Fuente: Elaboración Propia

5.2.8 Fase G: Gobierno de la Implementación

En esta fase se asegura que las soluciones ejecutadas se mantengan alineadas con la arquitectura definida. Es así como, se debe establecer seguimiento para verificar que el monitoreo del score postventa se realice de manera continua y que los resultados sean utilizados por las áreas pertinentes, normalmente las de análisis de datos.

Partes del Gobierno de Implementación



Figura 43. Fuente: Elaboración Propia

5.2.9 Fase H: Gestión de cambios de Arquitectura

Si el score continúa disminuyendo, la arquitectura deberá activar un nuevo ciclo de revisión para ajustar procesos, mejorar la captura de datos, incorporar nuevas variables o fortalecer las herramientas de análisis. De esta forma, la arquitectura empresarial se mantiene como un proceso de mejora continua. Esta fase es importante porque la satisfacción del cliente puede cambiar con el tiempo, por lo que el modelo predictivo, los indicadores y los procesos de seguimiento deben revisarse periódicamente.

Gestión de cambio de Arquitectura

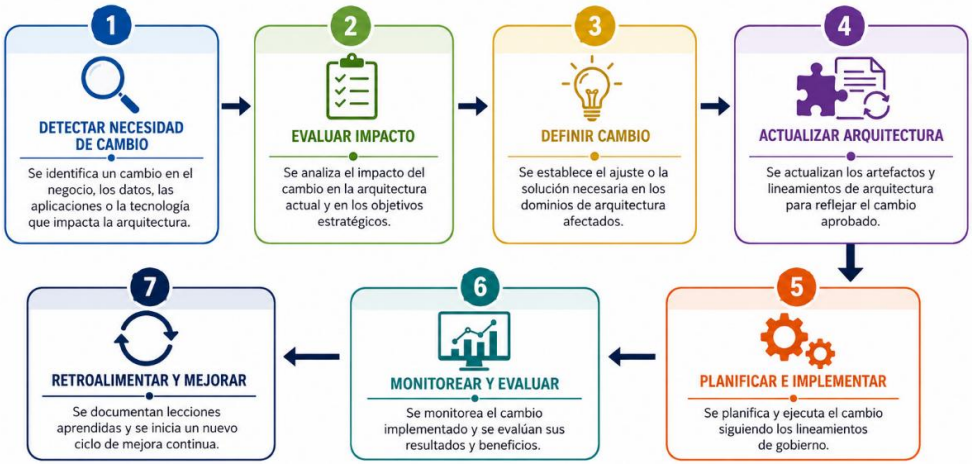


Figura 44. Fuente: Elaboración Propia

Con el planteamiento de la Fase H se concluye con el proceso de ADM en el marco de trabajo TOGAF. De esta manera, se incorporan los resultados de las fases de CRISP-DM en la elaboración de una arquitectura empresarial enfocada en los 3 aspectos primordiales de la misma: Aplicaciones, Tecnología y Datos. De esta manera, la empresa es capaz de utilizar su información no solo para toma de decisiones, sino también para fortalecer la estructura del negocio y afrontar posibles obstáculos al buscar cumplir sus objetivos estratégicos.

6. CAPÍTULO VI: CONCLUSIONES Y RECOMENDACIONES

En este capítulo se presentan las conclusiones y recomendaciones derivadas del desarrollo del proyecto de titulación. Para ello, se consideran los resultados obtenidos en la aplicación de la metodología CRISP-DM y su posterior relación con el diseño de arquitectura empresarial basado en TOGAF y ADM. Por otro lado, reúne el análisis de datos, la utilidad del modelo predictivo aplicado a los resultados postventa y la forma en que dichos datos permitieron orientar una propuesta de arquitectura empresarial enfocada en la mejora de procesos, el monitoreo y la toma de decisiones basada en datos.

6.1 Conclusiones

6.1.1 Conclusión general

El diseño de una arquitectura empresarial basada en TOGAF, junto con la metodología CRISP-DM, permitió crear una propuesta completa para mejorar la gestión de la satisfacción del cliente en Ponte Selva. Al combinar estas dos metodologías, se pudieron convertir los resultados del análisis predictivo en información útil para diseñar la arquitectura empresarial. Esto llevó a identificar mejoras en varios ámbitos, como el negocio, los datos, las aplicaciones y la tecnología. La investigación mostró que el conocimiento obtenido a través de la minería de datos es fundamental para tomar decisiones informadas y guiar la transformación de la empresa. La arquitectura empresarial basada en TOGAF y la metodología CRISP-DM demostraron ser herramientas valiosas para lograr este objetivo en Ponte Selva.

6.1.2 Conclusiones específicas

El desarrollo del Capítulo II permitió establecer las bases conceptuales que sustentan la investigación, evidenciando el trabajo en conjunto entre TOGAF y CRISP-DM. Mientras TOGAF proporciona una metodología para planificar y gestionar la transformación empresarial, CRISP-DM aporta un proceso estructurado para convertir

datos en conocimiento útil para la toma de decisiones. Por tanto, se cumplió el primer objetivo específico al construir el fundamento teórico que soporta la propuesta metodológica desarrollada.

La aplicación de CRISP-DM llevó a cabo el desarrollo ordenado de las fases de comprensión del negocio, comprensión de los datos, preparación, modelado, evaluación y despliegue, para conseguir un modelo predictivo para la estimación de la satisfacción del cliente, aplicado a través de la metodología del modelado de series de tiempo. Con ello, se alcanzó el segundo objetivo específico.

Los hallazgos obtenidos durante la fase analítica fueron utilizados como recurso para desarrollar el ciclo ADM de TOGAF, permitiendo identificar brechas en la arquitectura actual y definir una arquitectura objetivo orientada a mejorar la gestión de la satisfacción del cliente. La propuesta agrega mejoras en los dominios de negocio, datos, aplicaciones y tecnología, integrando herramientas para el monitoreo continuo, la generación de información estratégica y una mejor toma de decisiones. Por tanto, se cumplió el tercer objetivo específico mediante el diseño de una arquitectura empresarial alineada con las necesidades identificadas.

La validación de la propuesta evidenció que la arquitectura diseñada favorece una gestión más estructurada de la información respecto a la satisfacción del cliente incorporando indicadores y mecanismos de seguimiento para apoyar la toma de decisiones estratégicas. Asimismo, el análisis de los resultados permitió comprobar la efectividad entre el modelo predictivo desarrollado y las mejoras que se proponen dentro de la arquitectura empresarial.

6.2 Recomendaciones

- Ampliar el dataset utilizado en futuras etapas, incorporando variables adicionales como tipo de producto, tiempo de entrega, motivo de calificación, reclamos, observaciones del cliente, entre otros.
- Realizar una revisión mensual del score postventa, con reportes dirigidos a las áreas responsables. Esta práctica haría posible detectar disminuciones en la satisfacción del cliente y aplicar acciones correctivas de forma oportuna.
- Reentrenar los modelos de series temporales a medida que se vayan incorporando nuevos registros de las encuestas. Esto permitirá que las predicciones reflejen de mejor manera el comportamiento más reciente de la satisfacción postventa.
- Considerar la arquitectura empresarial como un proceso de mejora continua. Si se detectan nuevas necesidades, cambios en los indicadores o limitaciones en los modelos predictivos, se debe iniciar un nuevo ciclo ADM para ajustar la arquitectura propuesta.

7. BIBLIOGRAFÍA:

- Arriagada-Benítez, M. (2020). Ciencia de datos: hacia la automatización de las decisiones. *Ingeniare. Revista Chilena de Ingeniería, 28*(4), 556-557. https://doi.org/10.4067/S0718-33052020000400556
- Arroyo, E. (2015). Definición de principios de arquitectura para arquitectura empresarial de la organización. En M. Ramos-Escamilla y P. F. Solares-Soto (Eds.), *Ciencias de la Tecnología de Información* (pp. 35-44). México D.F., México: ECORFAN.
- Beltrán Martínez, B. (s.f.). *Minería de datos*. Benemérita Universidad Autónoma de Puebla. Obtenido de: https://www.cs.buap.mx/~bbeltran/NotasMD.pdf
- Carracedo Garnateo, P., y Terrádez, M. (2016). *Minería de datos. Introducción y guía de estudio*. Universitat Oberta de Catalunya. Obtenido de: https://openaccess.uoc.edu/server/api/core/bitstreams/399e7ecc-dbfc-47e4-826d-cb3f39310bc3/content
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., y Wirth, R. (2000). *CRISP-DM 1.0: Guía paso a paso de minería de datos*. Consorcio CRISP-DM.
- Conexiam. (s.f.). *Dominios de arquitectura empresarial*. Obtenido de: https://conexiam.com/es/dominios-de-arquitectura-empresarial/
- Corona Navarro, J. P., Almuna Herrera, E. M., Retamal González, B. D., y Gajardo Francino, H. A. (2018). *Modelos predictivos para fuga de clientes: Data Science*. Universidad de Chile. Obtenido de: https://repositorio.uchile.cl/bitstream/handle/2250/195198/Modelos-predictivos.pdf

- Crabtree, M., & Nehme, A. (2024, April 15). *¿Qué es la ciencia de datos? Definición, ejemplos, herramientas y más*. https://www.datacamp.com/es/blog/what-is-data-science-the-definitive-guide
- Espino Timón, C. (2017). *Análisis predictivo: técnicas y modelos utilizados y aplicaciones del mismo - herramientas Open Source que permiten su uso* (Trabajo de Fin de Grado). Universitat Oberta de Catalunya. Obtenido de: https://openaccess.uoc.edu/server/api/core/bitstreams/91aa36a3-f15c-40f3-bab9-a156722e81de/content
- Villegas, F. (2023, September 7). *Data Mining Techniques: What It Is & Importance*. QuestionPro. https://www.questionpro.com/blog/data-mining-techniques/#Defining_data_mining_techniques
- García, J., Molina, J. M., Berlanga, A., Patricio, M. A., Bustamante, Á. L., y Padilla, W. R. (2018). *Ciencia de datos: técnicas analíticas y aprendizaje estadístico. Un enfoque práctico*. Bogotá, Colombia: Alfaomega.
- González Blanco, N. (2024). *Bloque 0: Introducción a los modelos predictivos*. Universidad Loyola. Obtenido de: https://repositorio.uloyola.es/bitstream/handle/20.500.12412/6724/Bloque_0.pdf
- Grus, J. (2020). *Ciencia de datos desde cero: principios básicos con Python* (2.ª ed.). Madrid, España: Anaya Multimedia.
- Guerrero Silva, W. G., Rojas Contreras, W. M., Sánchez Delgado, M. P., y Villamizar Estrada, A. (2016). Arquitectura empresarial – dominios y beneficios. *FACE: Revista de la Facultad de Ciencias Económicas y Empresariales, 16*(1), 87-92. https://doi.org/10.24054/face.v16i1.1247
- IBM. (s.f.). *Guía de CRISP-DM de IBM SPSS Modeler*. IBM.

- LinkedIn Advice. (2023). *TOGAF: cuatro dominios de arquitectura e integración*. Obtenido de: https://www.linkedin.com/advice/3/how-do-you-integrate-four-types-architecture-togaf-other?lang=es
- Ramos-Escamilla, M., y Solares-Soto, P. F. (Eds.). (2015). *Ciencias de la Tecnología de Información*. México D.F., México: ECORFAN.
- G, I. R., & G, I. R. (2025, February 21). *CRISP-DM, Metodología de Datos*. iPMOGuide. https://ipmoguide.com/crisp-dm-metodologia-de-datos/
- Riquelme, J. C., Ruiz, R., y Gilbert, K. (2006). Minería de datos: conceptos y tendencias. *Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artificial, 10*(29), 11-18. Obtenido de: https://www.redalyc.org/pdf/925/92502902.pdf
- Sánchez Trujillo, M. G., y Pérez Hernández, J. Á. (2021). *Metodología CRISP-DM en la gestión de proyecto de Data Mining. Caso enfermedades dermatológicas*. International Conference on Project Management 2021.
- SAP BW Consulting. (s.f.). *TOGAF Framework* [Imagen]. Obtenido de: https://www.sapbwconsulting.com/hs-fs/hubfs/ig_04a582b0c60c5c8a0168f8ab96faf4819bb2758d24295310f1.png
- Schaper, Y. (2025). *Personal insights: Integrating CRISP-DM and TOGAF: A unified approach for enterprise AI architecture*. SAP Community. Obtenido de: https://community.sap.com/t5/technology-blog-posts-by-sap/personal-insights-integrating-crisp-dm-and-togaf-a-unified-approach-for/ba-p/14105057
- Sofyana, L., y Putera, A. R. (2019). Business architecture planning with TOGAF framework. *Journal of Physics: Conference Series, 1375*(1), 012056. https://doi.org/10.1088/1742-6596/1375/1/012056

- The Knowledge Academy. (2026). *Top 8 TOGAF® architecture principles you need to know*. Obtenido de: https://www.theknowledgeacademy.com/blog/togaf-architecture-principles/
- The Open Group. (s.f.). *The TOGAF Standard*. Obtenido de: https://www.opengroup.org/togaf
- UNIR. (2021). *Método TOGAF: en qué consiste y cuál es su objetivo*. Obtenido de: https://www.unir.net/revista/empresa/togaf-que-es/
- Universidad Europea. (2025). *¿Qué es el marco TOGAF?* Obtenido de: https://universidadeuropea.com/blog/togaf/
- *INDUSTRIA PIOLERA PONTE SELVA S.A. | Analisis Financiero Ecuador - TFC Smart*. (n.d.). TFC Smart. https://www.tfcsmart.com/demo/empresas/industria-piolera-ponte-selva-sa
- *Data-Driven Enterprise Architecture: A Practical Guide*. (2026, April 16). www.boc-group.com. https://www.boc-group.com/en/blog/ea/data-driven-enterprise-architecture-guide/
- *What is data-driven enterprise architecture?* (2025, December 1). Tyk API Management. https://tyk.io/learning-center/what-is-data-driven-enterprise-architecture/
- Prezi, J. Q. O. (2015, May 29). *industria piolera ponte selva*. prezi.com. https://prezi.com/evsav7wbw18a/industria-piolera-ponte-selva/
- Changoluisa González, E. S. (2023). *Propuesta de un modelo de planificación de inventarios aplicado a la Industria Ponte Selva S.A.* [Trabajo de titulación, Universidad Central del Ecuador]. Universidad Central del Ecuador.
- IBM. (s. f.). *Guía de CRISP-DM de IBM SPSS Modeler*. IBM.

- Véliz Córdova, S. M. (2015). *Plan de negocios para una tienda virtual de ropa y accesorios femeninos con asesoramiento de imagen para el año 2015* [Trabajo de titulación, Universidad Católica de Santiago de Guayaquil]. Universidad Católica de Santiago de Guayaquil.
- Conexiam. (2026). *Arquitectura empresarial - Conexiam*. https://conexiam.com/es/arquitectura-empresarial/
- International Organization for Standardization. (2022). *ISO/IEC 27001:2022 Information security, cybersecurity and privacy protection — Information security management systems — Requirements*. https://www.iso.org/standard/27001.html
- DAMA International. (2017). *The DAMA Guide to the Data Management Body of Knowledge (DAMA-DMBOK2)* (2nd ed.). Technics Publications.

Anexos

- Configuración de aplicación web

```
st.set_page_config(
   page_title="CRISP-DM Ponte Selva",
   page_icon=None,
   layout="wide"
)
```

- Definición de estilos

```
st.markdown(
   """
   <style>
   .main-title {
      text-align: center;
      font-size: 34px;
      font-weight: 800;
      color: #0B2E5F;
      margin-bottom: 0px;
   }
   .subtitle {
      text-align: center;
      font-size: 20px;
      font-weight: 600;
      color: #0B2E5F;
      margin-top: 10px;
      line-height: 1.45;
   }
   .author {
      text-align: center;
      font-size: 18px;
      font-weight: 700;
      color: #1F2937;
      margin-top: 20px;
   }
   .section-box {
      border: 1px solid #D0D7E2;
      border-radius: 10px;
      padding: 18px;
      background-color: #F8FAFC;
      margin-bottom: 18px;
   }
   </style>
   """,
   unsafe_allow_html=True
)
```

- Constantes

```
REQUIRED_COLUMNS = ["fecha_enc_vent", "score", "canal_venta"]
VALID_CHANNELS = ["WA", "PRE", "TEL", "MAIL"]
DATASET_FILE = Path("surveyPonteSelva.csv")
FREQUENCY_OPTIONS = ["Diaria", "Semanal", "Mensual"]
FREQ_MAP = {"Diaria": "D", "Semanal": "W", "Mensual": "M"}
```