

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento
Ampicilina + Sulbactam en Sicmafarma SAS



**Modelo de analítica y visualización de datos para la predicción de la demanda del
medicamento Ampicilina + Sulbactam en Sicmafarma SAS**

Daniela Alexandra Farfán Castillo

Jhoan Fernando Pico Toro

Gabriel Jaime Valbuena Quiroga

Universidad Ean

Facultad Ingeniería

Programa Maestría en gerencia de proyectos

Programa Maestría en ciencia de datos

Bogotá, Colombia

05/12/2025

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento
Ampicilina + Sulbactam en Sicmafarma SAS

Nota de aceptación:

Firma del jurado

Firma del jurado

Firma del director del trabajo de grado

Ciudad, día/mes/año

Dedicatoria

Daniela Alexandra Farfán Castillo

Dedico este nuevo logro a mis padres, agradeciendo profundamente su apoyo incondicional durante todo este proceso. Gracias por enseñarme el valor del esfuerzo y por ser siempre mi mayor ejemplo y fuente de inspiración.

Jhoan Fernando Pico Toro

Dedico este proyecto y la culminación de este posgrado a mis padres, Luz Toro y Fernando Pico, y también a mi querida Laura Mazabel, quienes me alentaron desde el inicio a emprender este reto académico y me acompañaron con su apoyo incondicional hasta verlo realizado.

Gabriel Jaime Valbuena Quiroga

Expreso mi gratitud a mi esposa por su amor, comprensión y fortaleza; a mi papá, hoy en el cielo, seguro de su orgullo por este logro; a mi familia, por su apoyo constante; y a mi hijo, cuya sonrisa dio sentido a cada esfuerzo. A todos, gracias. Dedico a ustedes esta tesis

Agradecimientos

Agradecemos profundamente a nuestra directora de tesis, Luz Maribel Guevara Ortega, por su trabajo y su valiosa guía a lo largo de este proceso. Extendemos nuestro agradecimiento a la Universidad Ean por la formación recibida y a nuestras familias, cuyo apoyo permanente ha sido fundamental para alcanzar nuestras metas y llevar a buen término este proyecto.

Resumen

Este trabajo de grado surge de la necesidad de optimizar la comercialización del medicamento Ampicilina + Sulbactam en Sicmafarma S.A.S., dado que las decisiones comerciales carecían de soporte analítico para anticipar la demanda y gestionar adecuadamente los inventarios. Ante esta situación, se establece el propósito de diseñar un modelo de analítica y visualización de datos, el cual permita mejorar la precisión en la predicción de la demanda del medicamento.

Para lograrlo, se desarrolló una consultoría académica estructurada en tres componentes. Primero, se realizó un diagnóstico de ventas incorporando variables internas (ventas y precios propios) y externas (precios de competidores, factores epidemiológicos y climatológicos). Segundo, se comparan modelos de pronóstico, especialmente OLS, log-log y random forest, para identificar cuál se ajusta mejor al comportamiento de los datos. A partir de dicha comparación, se seleccionó el modelo log-log ya que presentó el equilibrio más adecuado entre la capacidad explicativa e interpretabilidad. Tercero, se planteó Scrum como marco metodológico para implementar el modelo, permitiendo planear iteraciones y escalar la propuesta a otros medicamentos.

Los resultados demostraron la compatibilidad entre la analítica de datos y la metodología Scrum, confirmando que su integración genera soluciones aplicables y orientadas al mejoramiento continuo dentro de la organización.

Palabras clave: Analítica, scrum, predicción de la demanda, modelo log-log, Ampicilina + Sulbactam.

Abstract

This undergraduate thesis arises from the need to optimize the commercialization of the medication Ampicillin + Sulbactam at Sicmafarma S.A.S., given that commercial decisions lacked analytical support to anticipate demand and adequately manage inventory. In response to this situation, the objective was established to design a data analytics and visualization model that would improve the accuracy of demand forecasting for the medication.

To achieve this, an academic consultancy was developed and structured into three components. First, a sales diagnosis was conducted incorporating internal variables (sales and own prices) and external variables (competitor prices, epidemiological factors, and climatological conditions). Second, forecasting models—particularly OLS, log–log, and random forest—were compared to identify which one best aligned with the behavior of the data. Based on this comparison, the log–log model was selected because it presented the most appropriate balance between explanatory power and interpretability. Third, Scrum was proposed as the methodological framework for implementing the model, enabling the planning of iterations and scaling the proposal to other medications.

The results demonstrated the compatibility between data analytics and the Scrum methodology, confirming that their integration generates applicable solutions oriented toward continuous improvement within the organization.

Keywords: Analytics, Scrum, demand forecasting, log–log model, Ampicillin + Sulbactam.

Contenido

Resumen	6
Abstract.....	7
Contenido	8
Lista de Figuras	10
Lista de Tablas.....	11
Introducción	12
Objetivos.....	15
<i>Objetivo general</i>	15
<i>Objetivos específicos</i>	15
Justificación.....	16
Marco Institucional	18
<i>Propuesta de valor</i>	18
<i>Misión</i>	18
<i>Visión</i>	18
<i>Portafolio</i>	19
<i>Distribución</i>	21
Marco Contextual y Conceptual	23

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento
Ampicilina + Sulbactam en Sicmafarma SAS

Comercializadoras farmacéuticas en Colombia	23
Inteligencia de Negocios (BI) en las empresas.....	24
Fundamentos conceptuales de inteligencia de negocios aplicados al contexto del proyecto.....	25
Estructura del modelo	26
Componentes analíticos y articulación de métodos	27
Indicadores de desempeño y gobernanza del modelo.....	28
Visualización de resultados para la acción.....	28
Modelos para la predicción de la demanda	29
Regresión lineal múltiple (OLS).....	30
Modelo Log – Log.....	31
Modelo Random Forest	31
Ventajas y limitaciones de los modelos	33
Rendimiento de los modelos de predicción de la demanda.....	35
Coeficiente de determinación (R^2)	35
Error absoluto medio (MAE).....	36
Raíz del error cuadrático medio (RMSE)	36
Scrum	36
Roles.....	38
Artefactos	39
Eventos	39
Valores	40
Beneficios estratégicos	40
Integración Business Intelligence y Scrum	40

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento
Ampicilina + Sulbactam en Sicmafarma SAS

Diseño Metodológico de la Consultoría	42
<i>Enfoque del Estudio</i>	42
<i>Tipo de Investigación</i>	43
Alcance descriptivo	43
Alcance explicativo	43
<i>Datos y cobertura</i>	44
<i>Variables y fuentes de información</i>	44
<i>Estructura de los datos</i>	45
<i>Diccionario de variables</i>	45
<i>Plan de consultoría</i>	49
Fase 1.	49
Fase 2.	50
Reproducibilidad y trazabilidad	55
Fase 3.	55
Diagnóstico del estado actual de las ventas de Ampicilina + Sulbactam 1,5 g en Sicmafarma S.A.S	57
<i>Distribución de la variable dependiente.....</i>	57
<i>Correlaciones entre variables</i>	59
<i>Evento climático ENSO y comportamiento de la demanda</i>	62
<i>Relación entre precio de la competencia y ventas</i>	64
<i>Estacionalidad de la demanda</i>	66
<i>Implicaciones para la operación</i>	68

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en la empresa Sicmafarma S.A.S.70

Modelo de Regresión Lineal Múltiple (OLS).....70

Estructura y Variables.....70

Resultados y Diagnóstico71

Forma funcional modelo ols.....71

Predicción contra datos reales.....71

Modelo de Regresión Log-Log.....72

Estructura y Variables.....72

Resultados y Diagnóstico73

Depuración de datos73

Forma funcional modelo log – log74

Predicción contra datos reales.....74

Modelo de Random Forest75

Evaluación y resultados75

Importancia de variables76

Selección del modelo78

Comparación de Métricas de Error.....79

Interpretación Estratégica del Modelo Log-Log seleccionado.....79

Síntesis de hallazgos y selección del modelo.....82

Propuesta de plan de intervención para la implementación del modelo de analítica y visualización de datos aplicando la metodología Scrum.....84

Objetivo general.....84

Objetivos específicos85

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento
Ampicilina + Sulbactam en Sicmafarma SAS

Metodología Scrum	85
Roles	85
Artefactos	87
Eventos	88
Metas	89
Indicadores de éxito (KPIs)	90
Costos estimados	93
Cronograma de implementación	94
Beneficios esperados	94
Limitaciones del modelo	96
Sostenibilidad del modelo	96
Recomendaciones Estratégicas para Sicmafarma S.A.S.	97
Análisis de los resultados	98
Conclusiones y Recomendaciones	100
Conclusiones	100
Recomendaciones	101
Anexos	102
Referencias	103

Lista de Figuras

Figura 1	20
Figura 2	38
Figura 3	41
Figura 4	57
Figura 5	59
Figura 6	63
Figura 7	65

Lista de Tablas

Tabla 1	45
Tabla 2	46
Tabla 3	78
Tabla 4	86
Tabla 5	87
Tabla 6	88
Tabla 7	93

Introducción

El sector farmacéutico en Colombia ha atravesado transformaciones importantes en las últimas décadas, impulsadas por la creciente competencia, la innovación tecnológica y la necesidad de eficiencia en la cadena de suministro. Esta dinámica ha exigido a las empresas adaptarse rápidamente para seguir siendo competitivas. Como evidencia de esta evolución, la Asociación Colombiana de la Industria Farmacéutica (ASCIF) destaca que la industria farmacéutica abastece el 80 % del mercado interno y exporta a más de 16 países, lo que refleja su capacidad de adaptación y proyección en un entorno global cada vez más exigente (ASCIF, 2025).

No obstante, la predicción precisa de la demanda de medicamentos continúa siendo un desafío estratégico para el sector. Una estimación inadecuada impacta directamente en la gestión de inventarios, generando riesgos de sobreabastecimiento, desabastecimiento o pérdidas financieras. Frente a este panorama, la analítica se presenta como una alternativa clave, al transformar datos históricos, tendencias del mercado y variables externas en pronósticos robustos que favorecen la planificación de la demanda, optimizan los niveles de inventario y respaldan la toma de decisiones en tiempo real (Turban et al., 2020).

De manera complementaria, Scrum, marco ágil orientado a generar valor mediante soluciones adaptables a problemas complejos (Schwaber & Sutherland, 2020), se ha posicionado como un enfoque estratégico en la transformación de la industria farmacéutica. La adopción de metodologías ágiles no solo impulsa la eficiencia operativa y la reducción de costos, sino que también facilita la comercialización oportuna de los medicamentos y mejora el desempeño organizacional en un entorno altamente competitivo (Rigby, Sutherland, & Noble, 2018).

En este sentido, la analítica y la metodología Scrum se consolidan como enfoques estratégicos para abordar los retos comerciales del sector farmacéutico, al favorecer la

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

eficiencia en la gestión, fortalecer la toma de decisiones y aumentar la confiabilidad en la predicción de la demanda (Rigby, Sutherland, & Noble, 2018).

En este contexto, la empresa Sicmafarma S.A.S, empresa farmacéutica especializada en la comercialización de medicamentos esenciales y productos de cobertura básica en salud (Sicmafarma, 2025), enfrenta dificultades para comercializar de manera óptima el medicamento Ampicilina + Sulbactam, debido a que las decisiones actuales se toman, en su mayoría, con base en criterios tradicionales, sin un soporte analítico sólido que permita anticiparse a la demanda o gestionar de manera ágil los inventarios.

Ante esta situación, la empresa solicita la presente consultoría, en la cual, considerando la problemática identificada, se propone elaborar un modelo de analítica y visualización de datos orientado a la predicción de la demanda del medicamento Ampicilina + Sulbactam, con el propósito de fortalecer la toma de decisiones en Sicmafarma S.A.S. Para ello, se plantea el modelo y se incorpora Scrum como enfoque metodológico que facilite su refinamiento, asegurando mayor nivel de precisión y pertinencia en los resultados.

De acuerdo con lo anterior, se delimita el alcance de la consultoría al diseño y validación conceptual de un modelo de analítica y visualización de datos para predecir la demanda mensual de Ampicilina + Sulbactam 1,5 g durante los años 2019–2023, integrando variables climáticas (ENSO), epidemiológicas (INS) y de precios (propios y de competidores). Se comparan modelos de pronóstico (OLS, log–log y Random Forest), se selecciona el modelo que mejor se ajusta al comportamiento de los datos y finalmente, se propone su implementación con Scrum para iterar entregables, y escalar el modelo a otros medicamentos.

El análisis se limita a construir y validar un modelo predictivo. Por tanto, no incluye: evaluación clínica ni farmacovigilancia del medicamento; definición de precios o asuntos regulatorios; diseño detallado de la logística ni de políticas de inventario; puesta en producción del modelo ni su integración con sistemas (ERP/CRM); pruebas de causalidad (las relaciones

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

se usan con fines predictivos); ni análisis por regiones, canales o clientes por falta de datos. Se parte de que las fuentes consultadas tienen calidad y consistencia suficientes.

En este marco, se formula la pregunta de investigación que orienta el estudio: ¿Cómo puede un modelo de analítica y visualización de datos mejorar la precisión en la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma S.A.S.?

Con el fin de dar desarrollo a la consultoría y a resolver la pregunta de investigación planteada, el documento se estructura de la siguiente manera: primero se exponen la introducción, los objetivos, la justificación, los marcos institucional, contextual y conceptual, y el diseño metodológico. Luego se presenta el diagnóstico del estado actual de las ventas, se comparan los modelos y se selecciona el más adecuado para explicar y predecir la demanda. Finalmente, se propone el plan de implementación del modelo de BI con Scrum y se formulan conclusiones y recomendaciones derivadas de la consultoría.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Objetivos

Objetivo general

Diseñar modelo de analítica y visualización de datos que permita mejorar la precisión en la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma S.A.S

Objetivos específicos

- Realizar diagnóstico del estado actual de la distribución y venta del medicamento Ampicilina + Sulbactam en la empresa Sicmafarma S.A.S.
- Elaborar modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en la empresa Sicmafarma S.A.S.
- Elaborar plan de intervención que permita la implementación del modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma S.A.S., aplicando la metodología Scrum.

Justificación

El sector farmacéutico colombiano combina control de precios, exigencias regulatorias y presiones de eficiencia que estrechan los márgenes y vuelven críticos los errores de planeación. En ese contexto, los medicamentos de bajo costo y alta rotación como la Ampicilina + Sulbactam exigen pronósticos sensibles a variaciones estacionales, epidemiológicas y logísticas, debido a que el Precio Máximo de Venta (PMV) limita el ajuste por vía de precios y amplifica el impacto de quiebres de stock o sobre inventarios sobre el acceso y la sostenibilidad financiera (CNPMDM, 2022).

En respuesta a estos desafíos, la analítica aporta el soporte para transformar datos internos y fuentes externas en información accionable, visibilidad de rotación, envejecimiento de inventarios y señales tempranas de demanda, en coherencia con los requerimientos de trazabilidad y documentación del sector. La evidencia aplicada en entornos farmacéuticos reporta mejoras en disponibilidad y reducción de desperdicios cuando se integran ventas, caducidades y patrones de consumo en modelos y tableros de seguimiento (Mendoza Erazo, 2024).

En este contexto, se define el alcance de esta consultoría, acotado deliberadamente a un solo medicamento (Ampicilina + Sulbactam) y a la construcción y validación conceptual de un modelo de estimación de demanda que integre series internas de ventas y precio con variables climatológicas, epidemiológicas y precios de la competencia, en consonancia con el esquema de PMV. La elección del caso se sustenta en su condición de producto esencial del portafolio y en su alta rotación, idónea para un piloto analítico rigurosamente acotado (Sicmafarma, 2025).

A partir de esta definición, y con el objetivo de realizar mejora continua sobre el modelo propuesto, se adopta Scrum como marco de ejecución que aporta estructura y metodología. Cada iteración permite refinar progresivamente el modelo de analítica, generando entregables

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

de alto impacto y alertas tempranas; la guía oficial describe artefactos y reglas que soportan ciclos cortos, inspección y transparencia (Schwaber & Sutherland, 2020).

En este sentido, el resultado esperado es una propuesta metodológica y una hoja de ruta para escalar gradualmente el enfoque analítico a otras referencias del portafolio mediante ciclos breves de validación y mejora continua. Asimismo, cada ajuste derivado de estos ciclos contribuye a que el modelo se refine progresivamente, incrementando su precisión y adaptabilidad, en línea con la necesidad del sector de fortalecer su eficiencia operativa y su capacidad de respuesta (ASCIF, 2025).

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Marco Institucional

Desde 2010, la empresa Sicmafarma S.A.S ha impulsado la expansión de Laboratorio Biosano, una empresa con más de 80 años de trayectoria en la industria farmacéutica chilena. La compañía se ha especializado en la comercialización de medicamentos esenciales y productos de cobertura básica en salud, enfocándose en satisfacer la demanda de productos vitales que no se encuentran disponibles o que están en desabastecimiento (Sicmafarma, 2025).

Sicmafarma S.A.S. (2025) definió su propuesta de valor, misión, visión, portafolio y distribución como se observa a continuación:

Propuesta de valor

La empresa se destaca por su compromiso con la calidad, ética e innovación, lo que la ha consolidado como líder en la transformación de la atención médica. Ofrece un portafolio integral de productos farmacéuticos, desde medicamentos genéricos hasta soluciones especializadas, cumpliendo con altos estándares internacionales y asegurando la disponibilidad de tratamientos efectivos y accesibles en el sector salud.

Misión

Ser una comercializadora farmacéutica de alcance nacional e internacional, ofreciendo productos esenciales para el sector privado y público. Su portafolio abarca medicamentos del plan básico de salud de la mayoría de los países de Latinoamérica, con un énfasis particular en la comercialización de medicamentos esenciales o vitales no disponibles, contribuyendo a cubrir una necesidad crítica dentro del sistema de salud.

Visión

Para el año 2030 posicionarse como la comercializadora farmacéutica de referencia en el sector, destacándose por su capacidad de proporcionar

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

medicamentos e insumos de alta calidad en el momento oportuno. La empresa aspira a ser líder en la comercialización de medicamentos esenciales y vitales no disponibles, cubriendo las necesidades de salud en mercados clave.

Portafolio

La línea de productos de Sicmafarma S.A.S se encuentra organizada en tres grandes categorías que reflejan su compromiso con la salud y el bienestar. En primer lugar, se destacan las áreas terapéuticas, que abarcan desde antiinfecciosos hasta soluciones cardiovasculares, anestesia, sedación y apoyo al sistema nervioso. En segundo lugar, se presentan los medicamentos, con un portafolio de 64 referencias, entre las cuales se incluye la Ampicilina + Sulbactam 1.5 g. Finalmente, se incorporan los dispositivos e insumos médicos, como el preservativo, que complementan la oferta de soluciones integrales en salud.

Sicmafarma S.A.S ofrece un portafolio terapéutico especializado, dirigido al sector hospitalario en Colombia, con cobertura en instituciones públicas y privadas a lo largo del país. Con más de 15 años de experiencia, la empresa ha establecido alianzas estratégicas con profesionales de la salud, entidades gubernamentales y distribuidores, garantizando una oferta oportuna y alineada con las necesidades críticas del sistema de salud.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

A continuación, en la Figura 1 se observa productos de Sicmafarma S.A.S:

Figura 1

Productos Sicmafarma S.A.S



Nota. Elaboración propia, adaptado de Sicmafarma S.A.S, 2025.

<https://sicmafarma.com.co/>

Dentro de su portafolio, Sicmafarma S.A.S cuenta con una línea de antiinfecciosos, que incluye antibióticos de amplio espectro, antimicóticos y tratamientos especializados para infecciones graves, respondiendo a la problemática global de la resistencia antimicrobiana. Además, la empresa lidera la comercialización de anestésicos y sedantes, esenciales para procedimientos quirúrgicos, cuidados intensivos y manejo del dolor, garantizando seguridad y eficacia.

Su línea cardiovascular está orientada al tratamiento de enfermedades crónicas no transmisibles, como hipertensión e insuficiencia cardíaca, ofreciendo medicamentos efectivos y asequibles que están presentes en protocolos hospitalarios para la atención de pacientes críticos y postoperatorios. Sicmafarma S.A.S también ofrece diluyentes y soluciones intravenosas, fundamentales en la

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

administración segura de terapias intravenosas, y una línea para el sistema nervioso, que abarca medicamentos para trastornos neurológicos y de salud mental, tales como epilepsia y depresión.

Adicional a las áreas principales, la empresa dispone de productos para el manejo del dolor agudo y crónico, así como coadyuvantes en oncología, mejorando la calidad de vida de los pacientes y optimizando la tolerancia a los tratamientos oncológicos. Con una visión de innovación constante, Sicmafarma S.A.S sigue monitoreando tendencias internacionales y demandas del sistema hospitalario colombiano, posicionándose como un actor clave en el acceso a medicamentos vitales y soluciones terapéuticas críticas.

El portafolio de medicamentos con el que cuenta la empresa está constituido por un total de 64 referencias farmacéuticas, dentro de las cuales se encuentra la Ampicilina + Sulbactam 1.5 g. Este medicamento ha sido seleccionado como objeto de análisis en el marco de la presente consultoría, dado su relevancia dentro de la categoría de antibióticos y su incidencia en la gestión comercial. Asimismo, es importante señalar que el producto cuenta con el registro sanitario INVIMA 2020M-0014504-R1, lo que garantiza su autorización para la producción, distribución y comercialización en el territorio nacional.

Distribución

Sicmafarma S.A.S. cuenta con una red de distribución que integra una amplia diversidad de clientes, desde hospitales, clínicas y operadores logísticos hasta farmacias independientes, grandes cadenas y mayoristas, con presencia tanto en las principales ciudades del país como en zonas rurales y de difícil acceso. Esta cobertura nacional se respalda en una cadena de suministro robusta y en alianzas

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

estratégicas, que aseguran la entrega eficiente y segura de medicamentos en cualquier región de Colombia (Sicmafarma, 2025).

Marco Contextual y Conceptual

Comercializadoras farmacéuticas en Colombia

El sector farmacéutico colombiano constituye una de las industrias más estratégicas del país, tanto por su peso económico como por su relevancia social. De acuerdo con la Asociación Nacional de Empresarios (A.N.D.I., 2022), esta industria representa cerca del 12% del PIB industrial y genera más de 57.000 empleos formales, caracterizados por altos niveles de especialización y estabilidad laboral. Además, Colombia es el tercer mercado farmacéutico más grande de América Latina, lo que evidencia su importancia dentro de la región (A.S.C.I.F., 2025). En términos de producción, la mayor parte corresponde a medicamentos genéricos, que abastecen aproximadamente el 80% de las unidades consumidas en el país, aunque en términos de valor de ventas predominan los medicamentos importados e innovadores (El Tiempo, 2023).

Dentro de la cadena de valor, las comercializadoras farmacéuticas desempeñan un papel central como intermediarias entre los laboratorios fabricantes (nacionales o multinacionales) y los canales de distribución, tanto institucionales como privados. Estas compañías se encargan de la compra, almacenamiento, transporte y entrega de medicamentos, garantizando su disponibilidad oportuna y en condiciones de calidad. Algunas están orientadas al canal institucional, por ejemplo, gestionando la dispensación a pacientes del sistema asegurado a través de EPS, mientras que otras se enfocan en el canal comercial minorista (farmacias, droguerías, cadenas especializadas). Según Audifarma y Copidrogas, principales operadores del mercado, la función de estas comercializadoras es crítica para asegurar la trazabilidad, cumplimiento normativo y eficiencia logística en la distribución (El País, 2025).

El entorno regulatorio está determinado por dos ejes principales: la vigilancia sanitaria, a cargo del Instituto Nacional de Vigilancia de Medicamentos y Alimentos (INVIMA), y la regulación de precios, liderada por la Comisión Nacional de Precios de Medicamentos y

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Dispositivos Médicos (CNPMDM). El INVIMA otorga registros sanitarios, inspecciona plantas de producción y centros de distribución, regula la publicidad de medicamentos y coordina la farmacovigilancia. Por su parte, la CNPMDM fija precios máximos de venta para un grupo de medicamentos estratégicos a través de circulares anuales, utilizando metodologías de referenciación internacional (C.N.P.M.D.M., 2022). Esto ha permitido moderar precios en beneficio del sistema de salud y los pacientes, aunque también ha limitado los márgenes de rentabilidad para algunos laboratorios y comercializadoras.

Inteligencia de Negocios (BI) en las empresas

La Inteligencia de Negocios se define como el conjunto de procesos, metodologías y tecnologías que permiten transformar grandes volúmenes de datos en información útil y conocimiento accionable para la toma de decisiones. Su propósito principal es mejorar la eficiencia organizacional, facilitar el monitoreo de indicadores clave de desempeño (KPIs) y apoyar la planeación estratégica mediante la integración y análisis de datos provenientes de múltiples fuentes.

El enfoque de BI suele organizarse en tres niveles:

- a) Análisis descriptivo: resume y visualiza información histórica para comprender qué ocurrió.
- b) Análisis diagnóstico y predictivo: identifica causas de los resultados e infiere escenarios probables futuros.
- c) Análisis prescriptivo: sugiere acciones concretas basadas en modelos avanzados y simulaciones. (Turban et al., 2020).

En la práctica, BI se materializa en herramientas como tableros de control (*dashboards*), reportes automatizados y sistemas de alertas tempranas, que facilitan la visibilidad en tiempo real sobre procesos clave como ventas, inventarios, logística o desempeño financiero. Al permitir una visualización accesible de la información, BI mejora

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

la coordinación entre áreas y promueve una cultura organizacional basada en evidencia (Sanchez Obando et al., 2024).

En un mundo empresarial caracterizado por la volatilidad y la incertidumbre, la adopción de sistemas de BI otorga ventajas competitivas significativas:

- a) Agilidad en la respuesta: facilita decisiones rápidas frente a cambios en el mercado.
- b) Optimización de procesos: identifica ineficiencias y oportunidades de mejora.
- c) Soporte estratégico: conecta el análisis de datos con la planeación y ejecución de objetivos corporativos.
- d) Mejora continua: fomenta el aprendizaje organizacional al sistematizar y retroalimentar el ciclo de gestión.

De este modo, la Inteligencia de Negocios ha dejado de ser un recurso exclusivo de las áreas financieras o tecnológicas, para convertirse en un pilar transversal de la estrategia empresarial, capaz de integrarse con metodologías modernas de gestión como Agile o Scrum.

Fundamentos conceptuales de inteligencia de negocios aplicados al contexto del proyecto

Este trabajo se apoya en los siguientes pasos para la toma de decisiones: primero se reúnen e integran los datos; luego se almacenan para analizarlos; después se aplican los métodos analíticos; y, por último, se presentan y se monitorean los resultados. Para el diseño lógico del repositorio analítico se retoma la perspectiva dimensional de Kimball, con separación entre hechos (medidas) y dimensiones (por ejemplo, tiempo, producto y competidor), por su foco en interpretabilidad, cálculo de indicadores y trazabilidad (Kimball & Ross, 2011). Esta lectura es coherente con la comprensión de la inteligencia de negocios como un proceso organizacional orientado a convertir datos en información y

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

conocimiento útil para la decisión, con énfasis en la utilidad gerencial y la mejora de procesos en contextos farmacéuticos (Mendoza Erazo, 2024; Sanchez Obando et al., 2024).

Bajo ese marco, se consolidan fuentes internas (ventas y precios propios) y externas (precios de competidores, epidemiología y clima). Respecto al almacenamiento analítico se concibe con lógica de hechos y dimensiones: el hecho central son las ventas mensuales de Ampicilina + Sulbactam 1,5 g, con dimensiones de tiempo (mes y año, estacionalidad), de mercado y competencia (precios por competidor) y de contexto (indicadores epidemiológicos y climáticos). Aunque por restricciones operativas el soporte físico es un conjunto de datos integrado en Excel o CSV y no un almacén de datos formal, el diseño conceptual guía la estandarización, la construcción de variables y la lectura de indicadores clave. En cuanto a la analítica, se implementan modelos explicativos y predictivos (OLS, log log y Random Forest) y finalmente se presentan resultados mediante tablas, gráficos y resúmenes ejecutivos (Turban et al., 2020; Kimball y Ross, 2011).

Estructura del modelo

En este proyecto, el modelo de analítica se implementa como un flujo ligero y reproducible que convierte fuentes mensuales en pronósticos utilizables para la toma de decisiones. El recorrido inicia con la consolidación manual de archivos Excel o CSV que incluyen ventas internas y series públicas pertinentes. Luego se realiza depuración (verificación de tipos, búsqueda de nulos y duplicados, revisión de rangos plausibles), asegurando condiciones mínimas de calidad antes de cualquier análisis, en línea con el principio de confiabilidad de la información para la gestión.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Tras la depuración, se construyen variables temporales acordes con la frecuencia mensual (rezagos y dummies estacionales) y, cuando corresponde, se aplican transformaciones para estabilizar la varianza (por ejemplo, logaritmos). Con el conjunto de datos estructurados se entrenan los modelos candidatos respetando la secuencia temporal y se evalúan fuera de muestra. Los resultados se reportan como pronósticos mensuales y figuras de soporte. Esta estructura privilegia trazabilidad y reproducibilidad mediante herramientas accesibles; aunque no existe orquestación en tiempo real, dadas las restricciones de actualización y cultura de datos, el proceso se mantiene orientado a la decisión (Turban et al., 2020).

Delimitación conceptual breve. Este marco se circunscribe a un entorno de datos mensuales con integración de fuentes internas y señales contextuales públicas. La arquitectura adoptada responde al objetivo de traducir información depurada en indicadores y pronósticos accionables, priorizando interpretabilidad y trazabilidad sobre automatización en tiempo real.

Componentes analíticos y articulación de métodos

Los métodos se integran dentro del mismo flujo de procesamiento para garantizar comparabilidad: OLS como línea base interpretable para cuantificar relaciones; log log para leer elasticidades y manejar heterocedasticidad mediante transformaciones; y Random Forest para capturar no linealidades e interacciones entre señales externas y patrones estacionales. Ejecutarlos con el mismo conjunto de variables y la misma partición temporal permite una comparación justa y una selección sustentada en evidencia (Hyndman & Athanasopoulos, 2021; James et al., 2021).

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Esta articulación clarifica el papel de cada modelo en la decisión: OLS facilita comunicar efectos marginales; el log log expresa impactos en términos porcentuales, útiles para discutir sensibilidad de la demanda; Random Forest aporta flexibilidad y un ordenamiento de importancia de variables para priorizar señales. La coexistencia de tres modelos fortalece el proceso al habilitar comparación interna bajo condiciones homogéneas y conducir a la elección del método que mejor equilibra desempeño e interpretación.

Indicadores de desempeño y gobernanza del modelo

La evaluación utiliza métricas fuera de muestra para estimar capacidad de generalización: MAE como métrica primaria (interpretación directa en unidades) y RMSE, MAPE y R^2 como complementarias. Se combinan un corte cronológico de entrenamiento, validación y prueba y una verificación adicional de estabilidad del error a lo largo del tiempo. Desde la perspectiva de BI, estos indicadores se leen hacia valor de negocio (nivel de servicio, reducción de quiebres y cobertura), alineando la medición con la decisión operativa (Chai y Draxler, 2014; Kuhn y Johnson, 2019).

La selección del modelo se apoya en las métricas y en criterios de estabilidad y trazabilidad. Un método con error estable y lectura sencilla favorece la adopción por compras y planeación, incluso si la ventaja numérica sobre alternativas es moderada; esta conexión entre desempeño técnico y utilidad práctica es central en BI y facilita revisiones periódicas e iteraciones cuando se disponga de más datos (Turban et al., 2020).

Visualización de resultados para la acción

La visualización cierra el ciclo datos y decisión. El documento presenta comparaciones de pronóstico frente a observado por método, trayectorias mensuales y distribuciones de

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

error para revisar sincronía, tendencia y sesgos. En Random Forest se reportan importancias de variables para priorizar monitoreo, mientras que en el log log se comunican elasticidades que traducen impactos porcentuales al lenguaje de negocio. Estas salidas estandarizadas facilitan que perfiles gerenciales y de planeación interpreten y accionen el resultado del modelo seleccionado (Greene, 2018; James et al., 2021).

Con este diseño, la integración de arquitectura, métodos, indicadores y visualización muestra un flujo completo que parte de datos depurados y concluye en decisiones operativas. Los tres modelos se entrenan bajo condiciones homogéneas, se evalúan con indicadores comparables, se selecciona la especificación que ofrece el mejor balance entre desempeño e interpretación y se comunica su resultado de forma clara para su adopción. Así, los métodos no están aislados: quedan articulados dentro de un modelo de inteligencia de negocios consistente con los principios teóricos y adaptado a la escala y disponibilidad de información del proyecto (Turban et al., 2020; Kimball y Ross, 2011).

De este modo, el modelo conceptual guía la selección y uso de métodos analíticos y visualizaciones de manera consistente con los objetivos de la consultoría: mejorar la planeación de compras y reducir el riesgo de quiebres en el suministro del medicamento.

Modelos para la predicción de la demanda

Para abordar el problema de estimación y análisis de la demanda de medicamentos, se implementaron tres enfoques complementarios con fundamentos sólidos en econometría clásica y aprendizaje automático: regresión lineal múltiple (OLS), modelo log–log y Random Forest. Esta combinación permite capturar distintos niveles de complejidad en la relación entre variables, desde modelos interpretables hasta técnicas con mayor capacidad predictiva.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Regresión lineal múltiple (OLS)

La OLS se utiliza para cuantificar la relación entre una variable dependiente continua y un conjunto de predictores, permitiendo identificar el signo, la magnitud y la significancia estadística del efecto marginal de cada uno, manteniendo constantes los demás. En análisis de demanda es un método de referencia por su transparencia interpretativa y su amplia documentación en la literatura econométrica (Wooldridge, 2016).

En su forma escalar, el modelo para la observación temporal t se expresa en la Ecuación (1):

Forma funcional modelo OLS:

$$y_t = \beta_0 + \sum_{k=1}^K \beta_k x_{tk} + \varepsilon_t \quad (1)$$

En esta formulación, (y_t) denota la variable dependiente (p.\,ej., ventas mensuales en unidades); (x_{tk}) representa el valor del predictor (k) en el mes (t) (como precios propios o de competidores, indicadores epidemiológicos o marcadores climáticos/estacionales); (β_0) es el intercepto; (β_k) recoge el cambio esperado en (y_t) ante una variación unitaria de (x_{tk}) controlando por el resto; y (ε_t) resume factores no observados en (t) . Cuando (x_{tk}) es una variable indicadora (por ejemplo, una categoría de ENSO), (β_k) se interpreta como la diferencia promedio respecto a la categoría de referencia.

La notación matricial compacta de la Ecuación (1) se presenta en la Ecuación (2):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2)$$

En términos interpretativos, un $\beta_k > 0$ implica que, al aumentar x_k una unidad, el modelo espera que y aumente en β_k unidades; si $\beta_k < 0$, el efecto es inverso.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Modelo Log – Log

En este estudio se empleó un modelo log-log, que transforma tanto la variable dependiente como las independientes mediante logaritmos naturales. Este enfoque no solo estabiliza la varianza y reduce la influencia de valores atípicos, sino que además permite interpretar los coeficientes como elasticidades, lo que resulta especialmente relevante en estudios de demanda y comportamiento del consumidor (Greene, 2018). La forma funcional se presenta en la Ecuación (3).

Forma funcional modelo Log – Log

$$\ln(y_i) = \beta_0 + \sum_{k=1}^K \beta_k \ln(x_{ik}) + \sum_{m=1}^M \gamma_m D_{im} + u_i, \quad i = 1, \dots, n. \quad (3)$$

Lectura de los coeficientes:

- Continuas: β_k es una elasticidad: un aumento del 1 % en x_k se asocia con un cambio aproximado de β_k % en y .
- Dummies (D_{im}): efecto porcentual exacto sobre y : $100 \times (e^{\gamma_m} - 1)$ % si $|\gamma_m|$ es pequeño $\approx 100\gamma_m$ %).

La especificación log–log también facilita comparar el tamaño de los efectos entre variables al expresarlos en porcentajes y, cuando alguna x puede tomar valor cero, permite documentar un desplazamiento mínimo ($\ln(x + c)$ con $c > 0$) para mantener la transformación definida.

Modelo Random Forest

Por último, Se incluyó un modelo de Random Forest, el cual es un ensamble de árboles de decisión entrenados con *bootstrap* y con submuestreo aleatorio de predictores en cada división. Esta estrategia reduce varianza y capta no linealidades e interacciones

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

sin imponer una forma paramétrica; aunque su interpretabilidad es menor que en modelos lineales, suele ofrecer buena precisión y robustez, además de herramientas prácticas como la importancia por permutación y los perfiles de dependencia parcial (James et al., 2021). La idea básica del ensamble se resume en la Ecuación (4).

$$\hat{y}(x) = \frac{1}{B} \sum_{b=1}^B T_b(x) \quad (4)$$

Forma funcional e Interpretación Random Forest

Random Forest (RF) es un ensamblaje de árboles CART entrenados sobre muestras bootstrap y con submuestreo de predictores en cada división. La predicción se obtiene promediando los árboles (regresión) o por votación (clasificación). Esta estrategia reduce la varianza del estimador y capta no linealidades e interacciones sin imponer una forma paramétrica (James et al., 2021). La misma regla de agregación del ensamble en la Ecuación (4) se reitera en la Ecuación (5) para dejar explícitos los símbolos.

$$\hat{y}(x) = \frac{1}{B} \sum_{b=1}^B T_b(x) \quad (5)$$

donde $T_b(\cdot)$ es el árbol entrenado en la muestra bootstrap b y B es el número total de árboles.

Error fuera de bolsa (OOB)

Para cada observación i , se predice con los árboles que no la muestrearon (árboles “OOB”); el error OOB aproxima el desempeño fuera de muestra. Su cálculo típico se muestra en la Ecuación (6).

$$MSE_{OOB} = \frac{1}{n} \sum_{i=1}^n \text{big} \left(y_i - \widehat{y}^{(-i)}(x_i) \text{big} \right)^2 \quad (6)$$

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Importancia de variables

Se mide el aumento del error al permutar una variable x_j (romper su asociación con y), promediado sobre los árboles. La definición operativa se presenta en la Ecuación (7).

$$Imp(j) = \frac{1}{B} \sum_{b=1}^B (MSE_b^{perm(j)} - MSE_b) \quad (7)$$

- No hay “coeficientes” con interpretación directa como en OLS: la relevancia se analiza vía importancias y perfiles de dependencia parcial (PDP/ICE).
- Hiper-parámetros clave: B (número de árboles), profundidad máxima, tamaño mínimo de hoja y $max_features$ (número de predictores considerados por división). Aumentar B estabiliza; la complejidad se controla con profundidad/hojas y $max_features$.

En conjunto, estos tres modelos permiten una visión integral: el modelo OLS aporta claridad y estructura básica; el modelo log-log ofrece un enfoque más refinado y teóricamente sólido en términos de elasticidad; y el modelo Random Forest refuerza el componente predictivo, útil para decisiones operativas de corto plazo.

Ventajas y limitaciones de los modelos

A continuación, se presentan los modelos, destacando sus ventajas, limitaciones y los casos de uso recomendados.

Tabla 1

Ventajas, limitaciones y casos de uso de cada uno de los modelos

Modelo	Ventajas	Limitaciones	Casos de uso recomendados
---------------	-----------------	---------------------	----------------------------------

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

OLS	Resultado fácil de explicar; muestra cuánto cambia la demanda cuando cambia una variable; sirve como referencia para comparar otros métodos.	Resume relaciones “promedio”; puede perder precisión si las variaciones son muy irregulares o si hay factores muy relacionados entre sí.	Cuando se necesita trazabilidad y defensa del resultado ante lectores no técnicos; como línea base para evaluar mejoras.
Log–Log	Permite leer efectos en porcentaje (lenguaje natural para precio y volumen); ayuda a comparar variables en distintas escalas.	La estimación en escala log minimiza diferencias porcentuales; esto puede restar peso a errores grandes en unidades cuando el valor real es alto.	Cuando la discusión se da en términos “% arriba/abajo”; para comunicar sensibilidades porcentuales de cada variable.
Random Forest	Descubre patrones complejos y combinaciones de factores sin fijar una fórmula previa; ofrece un “ranking” de variables más influyentes.	Menos sencillo de interpretar globalmente; requiere cuidado al validar en el tiempo; su lectura es más descriptiva que explicativa.	Cuando se sospechan relaciones no lineales o interacciones; para priorizar señales y contrastar el desempeño frente a métodos más simples.

Nota. *Ventajas, limitaciones y casos de uso sintetizados para el contexto del proyecto (serie mensual 2019–2023 con covariables internas y externas). Referencias de apoyo: James et al. (2021), Greene (2018), Kuhn y Johnson (2019), Hyndman y Athanasopoulos (2021).*

La selección de OLS, Log–Log y Random Forest obedece a tres criterios alineados con el objetivo del proyecto: comunicación clara para audiencias de negocio, coherencia con la disponibilidad y la frecuencia de los datos, y contraste metodológico dentro de un mismo flujo analítico. OLS provee una línea base transparente para cuantificar asociaciones y abrir la discusión con evidencia; el modelo Log–Log permite traducir los impactos a porcentajes y discutir elasticidades, un lenguaje natural cuando se habla de

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

precios y volúmenes; y Random Forest añade una perspectiva flexible para contrastar posibles relaciones complejas y priorizar señales a partir de medidas de importancia de variables (Greene, 2018; James et al., 2021; Wooldridge, 2016).

En cuanto a alternativas, se privilegió la coherencia operativa y la solidez argumentativa del enfoque. Los modelos de series clásicas (ETS/ARIMA) son especialmente útiles cuando la dinámica interna domina y se dispone de series largas; en este trabajo el interés estuvo en integrar covariables y comunicar efectos al negocio, por lo que la tripleta OLS, Log–Log, RF resultó más pertinente (Hyndman & Athanasopoulos, 2021). Métodos de mayor complejidad (por ejemplo, boosting o redes neuronales) pueden aportar valor, pero su beneficio marginal en una muestra mensual acotada y su costo explicativo para audiencias no técnicas no ofrecían una ventaja clara en esta fase. En conjunto, la combinación elegida equilibra interpretabilidad, lenguaje económico comprensible y una opción flexible para contrastar patrones, manteniendo la trazabilidad del proceso de BI.

Rendimiento de los modelos de predicción de la demanda

Para evaluar el rendimiento de los modelos predictivos se utilizaron tres métricas estándar: Coeficiente de determinación (R^2), el error absoluto medio (MAE) y la raíz del error cuadrático medio (RMSE).

Coeficiente de determinación (R^2)

Indica la proporción de la varianza de la variable dependiente explicada por el modelo; su valor oscila entre 0 y 1, donde valores cercanos a 1 implican mayor capacidad explicativa ((James et al., 2021). El R^2 ajustado $\overline{R^2}$ penaliza por número de predictores p , favoreciendo comparaciones entre especificaciones.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad \bar{R}^2 = 1 - (1 - R^2) \frac{n-1}{n-p-1} \quad (8)$$

Error absoluto medio (MAE)

Mide el promedio de las diferencias absolutas entre observado y predicho; es directo de interpretar porque conserva la escala de y .

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (9)$$

Raíz del error cuadrático medio (RMSE)

Penaliza con mayor peso los errores grandes al elevar primero al cuadrado las diferencias antes de promediarlas y luego extraer la raíz (Chai & Draxler, 2014). Si $RMSE \gg MAE$, suele haber outliers o colas pesadas que conviene revisar.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (10)$$

Scrum

Entre las metodologías ágiles, *Scrum* se ha consolidado como la más utilizada a nivel mundial. Scrum se define como un marco de trabajo empírico, ligero y adaptable que permite crear valor en entornos complejos a través de ciclos cortos de trabajo denominados *sprints* (Schwaber & Sutherland, 2020). Su fundamento es el control empírico de procesos, basado en tres pilares:

- a) Transparencia: la información esencial es visible y compartida por todos los actores.
- b) Inspección: se revisa periódicamente el progreso para identificar desviaciones.
- c) Adaptación: el equipo ajusta de manera oportuna sus procesos o entregables según lo aprendido.

La literatura reciente ha evidenciado que la adopción de Scrum contribuye a mejorar la coordinación de equipos, reducir la incertidumbre y acelerar el aprendizaje organizacional

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

en entornos de negocio y transformación digital (Patara & Raharjo, 2023). Asimismo, su énfasis en la entrega incremental de valor resulta especialmente pertinente en contextos dinámicos, donde la información evoluciona rápidamente y las empresas requieren capacidad de adaptación constante.

Scrum constituye un marco moderno y ampliamente validado de gestión de proyectos empresariales, cuyo valor radica en su capacidad para articular flexibilidad, transparencia y aprendizaje continuo en la creación de soluciones innovadoras y competitivas. En entornos dinámicos y con alta variabilidad en la demanda, como ocurre con la comercialización de medicamentos, la aplicación de principios de la gerencia de proyectos se ha consolidado como una práctica eficaz para mejorar la eficiencia operativa, la alineación estratégica y la capacidad de respuesta comercial (Berryhill, 2023).

Según Schwaber y Sutherland (2020), Scrum es un marco ágil, ligero y empírico orientado a la generación de valor en entornos complejos, a través de ciclos cortos de trabajo denominados *sprints*. Se fundamenta en el empirismo, entendido como el proceso de aprendizaje a partir de la experiencia y el ajuste continuo en consecuencia, así como en los principios de *Lean*, que promueven la concentración en lo esencial para el cliente. Su finalidad principal es optimizar la entrega de valor en contextos caracterizados por la incertidumbre y el cambio constante.

En la Figura 2 se ilustran los roles, artefactos, eventos, valores y beneficios que integran el marco de trabajo o ciclo Scrum.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Figura 2

Metodología Scrum



Nota. Adaptado de “The Scrum Framework, Illustrated” Overeem, 2020.

<https://www.scrum.org/resources/blog/scrum-framework-illustrated#:~:text=The%20poster%20shows%20the%20events%2C,build%20trust%20between%20stakeholders%20and>

A continuación, se relacionan cada uno de los roles que integran el ciclo scrum y sus principales funciones.

Roles

Scrum define tres roles principales, que conforman el equipo Scrum:

1. Product Owner (PO): representa la voz del negocio, prioriza el valor a entregar y gestiona el product backlog.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

- II. Scrum Master (SM): actúa como facilitador del marco, elimina impedimentos y promueve la cultura ágil.
- III. Developers o equipo de desarrollo: grupo multidisciplinario encargado de transformar los requisitos en un incremento usable y de calidad.

Los artefactos serán descritos a continuación, los cuales también hacen parte del ciclo Scrum.

Artefactos

Los artefactos promueven transparencia y facilitan la inspección y adaptación:

- I. Product Backlog: lista priorizada de necesidades, ideas o funcionalidades.
- II. Sprint Backlog: subconjunto de elementos del Product Backlog seleccionados para un sprint, con un plan de ejecución.
- III. Incremento: producto o resultado funcional al final de cada sprint, que cumple con la *Definition of Done*, es decir que el entregable cumplió con los criterios para considerarse como terminado.

Se listan y se da una breve explicación acerca de los eventos que componen el ciclo Scrum.

Eventos

Scrum estructura el trabajo cada sprint incluye los siguientes eventos, que generan planificación, plan de acción, verificación y mejora continua:

- I. Sprint Planning: planificación del Sprint Goal y definición de qué se hará y cómo se logrará.
- II. Daily Scrum: reunión breve de 15 minutos para sincronizar el trabajo y ajustar planes diarios.
- III. Sprint Review: presentación del incremento a los actores clave para recibir retroalimentación.
- IV. Sprint Retrospective: espacio de mejora continua para el equipo.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Valores

Compromiso, coraje, enfoque, apertura y respeto. Su práctica consistente correlaciona con mayor alineamiento, colaboración y entrega sostenida de valor (Patara & Raharjo, 2023; Schwaber & Sutherland, 2020).

Beneficios estratégicos

La literatura evidencia que Scrum genera ventajas competitivas al:

- I. Reducir la incertidumbre mediante entregas incrementales.
- II. Promover la adaptabilidad frente a cambios en el entorno.
- III. Facilitar el aprendizaje organizacional.
- IV. Alinear el trabajo de los equipos con los objetivos estratégicos (Patara & Raharjo, 2023; Rigby et al., 2016).

Integración Business Intelligence y Scrum

Responde a la necesidad de las empresas de gestionar entornos dinámicos, caracterizados por la alta variabilidad de la demanda, la incertidumbre regulatoria y la competencia creciente. En este contexto, BI aporta la capacidad de recopilar, procesar y visualizar datos en tiempo real, mientras que Scrum provee la estructura organizativa y metodológica para transformar esa información en decisiones ágiles y efectivas. Según (Schwaber & Sutherland, 2020), Scrum resulta especialmente valioso en entornos complejos, donde la transparencia, inspección y adaptación permiten a los equipos reducir incertidumbre y responder con rapidez a cambios en el mercado.

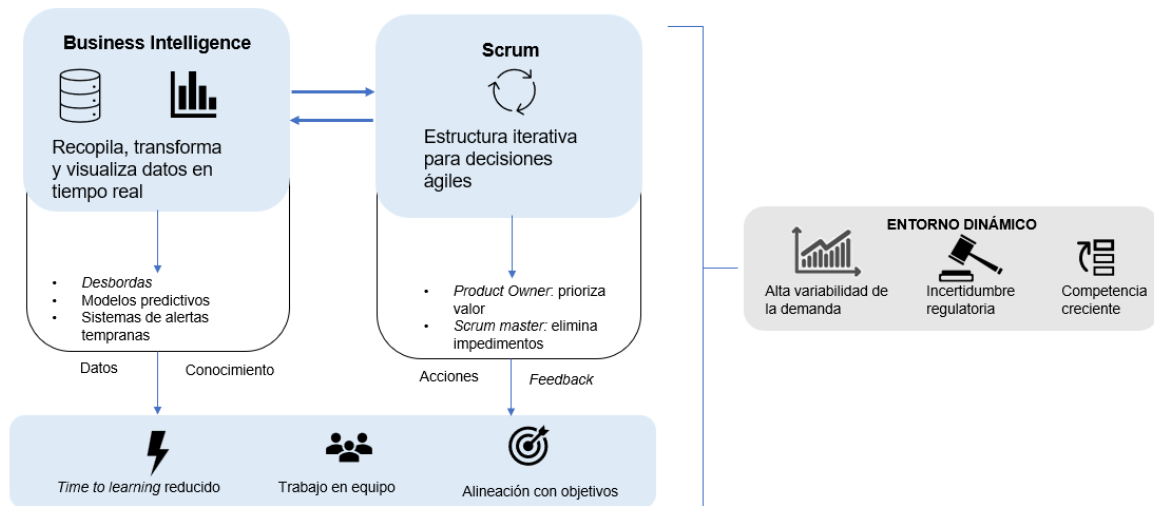
El uso conjunto de BI y Scrum favorece la entrega incremental de valor. Cada *sprint* puede orientarse al desarrollo de tableros de control, modelos predictivos o sistemas de alerta temprana, los cuales son revisados y validados con frecuencia. De esta manera, la organización no necesita esperar a la finalización de un proyecto de gran escala para obtener beneficios; en cambio, puede evaluar resultados parciales y adaptarlos en función de nuevas condiciones del mercado o del aprendizaje generado en cada iteración (Patara

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

& Raharjo, 2023). Esta dinámica contribuye a reducir el *time-to-learning*, acelerando la conversión de datos en decisiones accionables. En la Figura 3 se observa la articulación de BI y Scrum.

Figura 3

Interacción BI y Scrum



Nota. Elaboración propia adaptado de Schwaber & Sutherland (2020); Patara & Raharjo (2023); Rigby, Sutherland, & Takeuchi (2016).

La interacción entre ambos enfoques se representa mediante flujos bidireccionales, puesto que BI suministra datos procesados que se convierten en insumos estratégicos, dado que mediante la dinámica iterativa de Scrum se ajusta continuamente las necesidades analíticas, logrando que los datos generados y procesados adquieran mayor utilidad y precisión en la toma de decisiones (Patara & Raharjo, 2023).

Otro aporte relevante de Scrum a BI y a los procesos predicción de la demanda es la capacidad de alinear la analítica con los objetivos estratégicos de la organización. El rol del *Product Owner* asegura que los desarrollos analíticos (dashboards, reportes, algoritmos de predicción) respondan a necesidades reales del negocio, priorizando aquellas que generan mayor valor. Al mismo tiempo, el *Scrum Master* garantiza que el

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

marco ágil se aplique correctamente, promoviendo un entorno de trabajo colaborativo y eliminando impedimentos que puedan ralentizar la entrega de valor. Este enfoque, basado en roles claramente definidos y procesos iterativos, fortalece la conexión entre la generación de conocimiento y su aplicación práctica en decisiones comerciales (Schwaber & Sutherland, 2020).

Asimismo, Scrum facilita la adaptación continua de las tácticas comerciales frente a señales externas que afectan la demanda, como cambios en los precios de la competencia, variaciones climáticas o brotes epidemiológicos. La evidencia muestra que las organizaciones que adoptan metodologías ágiles reportan mayores niveles de coordinación y aprendizaje organizacional, lo que a su vez se traduce en una mejor capacidad de respuesta a contextos volátiles (Patara & Raharjo, 2023). En este sentido, Scrum no solo organiza el trabajo en torno a entregables analíticos, sino que también promueve un aprendizaje colectivo que permite mejorar continuamente las estrategias de mercado.

Finalmente, la escalabilidad de Scrum resulta clave en entornos empresariales modernos. Una vez consolidada la dinámica de trabajo, es posible replicar el enfoque en diferentes líneas de productos, mercados o segmentos de clientes, lo que incrementa la capacidad organizacional de absorber complejidad y gestionar múltiples frentes de manera simultánea. Así, la combinación de BI y Scrum no solo ofrece ventajas operativas inmediatas, sino que también fortalece la madurez organizacional en la gestión de proyectos de transformación digital y comercial (Rigby et al., 2016).

Diseño Metodológico de la Consultoría

Enfoque del Estudio

Esta consultoría adopta un enfoque cuantitativo. La base empírica es el histórico mensual de ventas del medicamento Ampicilina + Sulbactam 1,5 g suministrado por Sicmafarma

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

S.A.S; la empresa, en una reunión exploratoria, confirmó la relevancia operativa del producto. Sobre esa serie se integran variables externas climáticas y epidemiológicas seleccionadas por plausibilidad causal y evidencia sectorial.

Tipo de Investigación

Este estudio se enmarca en una investigación aplicada, con alcance descriptivo - explicativo, bajo un diseño no experimental, observacional y longitudinal retrospectivo (serie temporal mensual). El propósito es resolver la problemática en Sicmafarma S.A.S: mejorar la toma de decisiones comerciales sobre el medicamento Ampicilina + Sulbactam 1,5 g mediante el modelo que explique y prediga su demanda. La investigación aplicada busca generar conocimiento útil y transferible para la práctica, más que producir teoría abstracta (Saunders et al., 2019). En términos de métodos de negocio, se privilegia el uso de analítica para la decisión con resultados operativos.

Alcance descriptivo

El estudio realiza un análisis descriptivo del comportamiento de las ventas mensuales del medicamento (niveles, distribución, variabilidad y estacionalidad) y de los predictores externos de interés (clima, epidemiología y precios de competidores). Se emplean resúmenes numéricos y visualizaciones (histogramas, boxplots, promedios mensuales, mapas de calor de correlaciones) para caracterizar patrones y rasgos relevantes de cada serie, en línea con los lineamientos de investigación descriptiva (Hernández-Sampieri & Mendoza Torres, 2018).

Alcance explicativo

Se estiman asociaciones entre la variable dependiente y los predictores mediante Regresión Lineal Múltiple (OLS) y una especificación log-log para interpretar elasticidades (precio propias y cruzadas) y diferencias porcentuales asociadas a variables indicadoras. La elección privilegia interpretabilidad y validez temporal: El análisis respeta la secuencia

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

temporal y se evalúa con particiones cronológicas de entrenamiento, validación y prueba (time-series cross-validation), evitando filtraciones de información y favoreciendo una estimación honesta del desempeño predictivo (Hyndman & Athanasopoulos, 2021).

Datos y cobertura

Sicmafarma S.A.S. cuenta con un portafolio de 64 productos; el estudio se limita a Ampicilina + Sulbactam 1,5 g. Se analizan 60 observaciones mensuales entre enero de 2019 y diciembre de 2023, agregadas a nivel nacional y sin desagregación por canal ni zona, conforme a la disponibilidad de datos (Sicmafarma, 2025).

El análisis respetará la secuencia temporal y se evaluará con particiones cronológicas de entrenamiento, validación y prueba, en línea con buenas prácticas para series de tiempo (Hyndman & Athanasopoulos, 2021).

VARIABLES Y FUENTES DE INFORMACIÓN

La selección de variables obedece a pertinencia empírica, disponibilidad y respaldo sectorial, con el fin de caracterizar la demanda y estimar asociaciones temporales relevantes.

- Sicmafarma S.A.S: datos internos sobre ventas y precios del medicamento Ampicilina + Sulbactam 1,5 g y de sus principales competidores (Vitalis, Aurobindo, Pfizer, entre otros) (Sicmafarma, 2025)
- Instituto Nacional de Salud (INS): datos epidemiológicos mensuales sobre enfermedades respiratorias, infecciosas y virales, como tuberculosis, meningitis, IRAG y otras patologías relacionadas (I.N.S., 2025).
- Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM): datos sobre temperatura promedio mensual en cinco ciudades clave (Bogotá, Cartagena, Medellín, Manizales y Pereira) (I.D.E.A.M., 2025).

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

- Climate Prediction Center (CPC - National Weather Service, EE. UU.): información climática global sobre el fenómeno ENSO, clasificado mensualmente como Niño, Niña o Neutral (N.O.A.A. Climate Prediction, 2025).

Estructura de los datos

En la tabla 2 se describe el detalle de la categoría de los datos utilizados.

Tabla 2

Estructura de los datos

Categoría	Detalle
Observaciones	60 meses (n = 60), de enero de 2019 a diciembre de 2023.
VARIABLES NUMÉRICAS	18 variables (clima y salud pública).
VARIABLES DE PRECIOS	7 variables (precio propio y 6 competidores).
VARIABLE DEPENDIENTE	Ventas de Ampicilina + Sulbactam 1,5 g.
VARIABLES CATEGÓRICAS	2: mes (temporal) y ENSO (Niño/Neutral/Niña).

Nota. Elaboración propia basado en datos de Ventas y precios: (Sicmafarma S.A.S., 2025); epidemiología: (INS, 2025); clima: (IDEAM, 2025); clasificación ENSO (ONI): (NOAA CPC, 2025).

Diccionario de variables

En la tabla 3, se presenta el diccionario de todas las variables que se tuvieron en cuenta en las primeras instancias del modelado.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Tabla 3

Diccionario de variables

Variable	Descripción	Unidad/Escala	Fuente
Mes	Mes y año de observación (formato MM/AAAA).	—	—
Ampicilina + sulbactam 1,5 g	Unidades mensuales vendidas del medicamento Ampicilina + Sulbactam 1,5 g por Sicmafarma S.A.S.	Unidades	Sicmafarma S.A.S. (2025)
BOGOTA_temp_promedio	Temperatura promedio mensual en Bogotá.	°C	IDEAM (2025)
CARTAGENA DE INDIAS_temp_promedio	Temperatura promedio mensual en Cartagena de Indias.	°C	IDEAM (2025)
MANIZALES_temp_promedio	Temperatura promedio mensual en Manizales.	°C	IDEAM (2025)
MEDELLIN_temp_promedio	Temperatura promedio mensual en Medellín.	°C	IDEAM (2025)
PEREIRA_temp_promedio	Temperatura promedio mensual en Pereira.	°C	IDEAM (2025)
Dengue_Grave	Casos mensuales confirmados de dengue grave.	Casos	INS (2025)

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Infección_respiratoria_aguda_grave_(IRAG)	Casos mensuales de infección respiratoria aguda grave (requiere hospitalización).	Casos	INS (2025)
Meningitis_por_Haemophilus_Influenzae	Casos mensuales de meningitis por Haemophilus influenzae.	Casos	INS (2025)
Meningitis_Meningococica	Casos mensuales de meningitis meningocócica.	Casos	INS (2025)
Meningitis_por_Neumococo	Casos mensuales de meningitis por neumococo.	Casos	INS (2025)
Tos_ferina	Casos mensuales de tos ferina.	Casos	INS (2025)
Tuberculosis_pulmonar	Casos mensuales de tuberculosis pulmonar.	Casos	INS (2025)
Zika	Casos mensuales de zika.	Casos	INS (2025)
morbilidad_ira_tot_irag	Total, mensual de atenciones por IRA clasificadas como IRAG.	Atenciones	INS (2025)
morbilidad_ira_tot_cau_h	Total, mensual de atenciones IRA en hospitalización.	Atenciones	INS (2025)
morbilidad_ira_tot_irauci	Total, mensual de atenciones IRA en UCI.	Atenciones	INS (2025)

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

morbilidad_ira_tot_iraext	Total, mensual de atenciones IRA en consulta externa.	Atenciones	INS (2025)
morbilidad_ed	Total, mensual de atenciones por enfermedad diarreica aguda (EDA).	Atenciones	INS (2025)
precio_sicma	Precio de venta mensual del producto Ampicilina + Sulbactam 1,5 g (Sicmafarma).	\$/pesos	Sicmafarma S.A.S. (2025)
ONI	Oceanic Niño Index (Niño-3.4) mensual.	Índice (°C)	NOAA CPC (2025)
Evento_Sostenido	Clasificación mensual ENSO (Niño, Neutral o Niña).	Categorica	NOAA CPC (2025)
Precio_AUROBINDO PHARMA LIMITED	Precio del competidor AUROBINDO PHARMA LIMITED.	\$/pesos	Sicmafarma S.A.S. (2025)
Precio_BIOSELECT S.A.C.I.	Precio del competidor BIOSELECT S.A.C.I.	\$/pesos	Sicmafarma S.A.S. (2025)
Precio_FARMALOGICA S.A.	Precio del competidor FARMALOGICA S.A.	\$/pesos	Sicmafarma S.A.S. (2025)
Precio_LABORATORIOS DELTA S.A.S.	Precio del competidor LABORATORIOS DELTA S.A.S.	\$/pesos	Sicmafarma S.A.S. (2025)
Precio_PFIZER S.A.S.	Precio del competidor PFIZER S.A.S.	\$/pesos	Sicmafarma S.A.S. (2025)

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Precio_VITALIS S.A. C.I.	Precio del competidor VITALIS S.A. C.I.	\$/pesos	Sicmafarma S.A.S. (2025)
--------------------------	--	----------	--------------------------------

Nota. Elaboración propia basado en datos de Temperaturas diarias del portal de datos abiertos del IDEAM (2025) promediadas a frecuencia mensual por el autor. Clasificación ENSO basada en ONI del NOAA CPC (2025). Series epidemiológicas: INS (2025). Ventas y precios (propio y competidores): Sicmafarma S.A.S. (2025).

Plan de consultoría

El plan de consultoría se estructura en tres fases que dan respuesta a los objetivos específicos planteados en la investigación. En la primera fase se realiza el diagnóstico del estado actual de la distribución y venta del medicamento Ampicilina + Sulbactam 1,5 g, con el fin de identificar brechas y oportunidades de mejora en la gestión comercial y logística. La segunda fase se orienta a la elaboración de un modelo de analítica y visualización de datos para la predicción de la demanda de dicho medicamento, integrando variables internas y externas que influyen en su comportamiento. Finalmente, en la tercera fase se formula la propuesta de plan de intervención para la implementación del modelo de analítica mediante la metodología Scrum, asegurando un enfoque ágil, colaborativo y sostenible que facilite su adopción dentro de Sicmafarma S.A.S.

Fase 1. Diagnóstico del estado actual de la distribución y venta del medicamento Ampicilina + Sulbactam 1,5 g

Para dar cumplimiento a este objetivo, se llevará a cabo un análisis exploratorio de datos sobre la serie mensual agregada a nivel nacional del medicamento Ampicilina + Sulbactam 1,5 g correspondiente al año 2019–2023. El análisis caracterizará niveles, distribución y variabilidad de la demanda; identificará patrones de estacionalidad mensual; y examinará diferencias por fases ENSO (Niño, Neutral y Niña), así como asociaciones contemporáneas con epidemiología y precios de competidores. Los resultados se

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

documentarán mediante tablas y visualizaciones (histogramas, boxplots, promedios o índices mensuales, mapas de calor y diagramas de dispersión), atendiendo criterios de calidad de datos y validez temporal (Hyndman & Athanasopoulos, 2021).

Previo al análisis, se depurará la base con estandarización de nombres y códigos, tratamiento de ceros y anomalías, y verificación de consistencia temporal y metadatos. También se integrarán las fuentes externas INS, IDEAM y CPC-NOAA a frecuencia mensual, asegurando alineación de fechas y definiciones. Se preservará el orden temporal en toda la exploración, y las comparaciones por fase o por mes se presentarán con notas que delimiten su alcance descriptivo y no causal (Kuhn & Johnson, 2019).

Adicionalmente, se presentará una lectura diagnóstica de la situación actual. Se sintetizarán regularidades como picos estacionales, sensibilidad a ENSO y a indicadores epidemiológicos, y señales competitivas. También se delimitarán brechas observables en visibilidad competitiva, integración de señales externas y calidad de datos, que enmarcarán el objetivo analítico siguiente.

Fase 2. Elaborar modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam 1,5 g

Se desarrollarán y compararán tres enfoques complementarios: Regresión Lineal Múltiple como línea base interpretativa; modelo log-log para obtener elasticidades y estabilizar varianza; y Random Forest para capturar no linealidades e interacciones. También se incluirá preprocesamiento con codificación de variables indicadoras, transformaciones logarítmicas cuando corresponda y controles de multicolinealidad, además de diagnósticos de ajuste y estabilidad con soporte metodológico reciente (Kuhn & Johnson, 2019).

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

La evaluación se realizará con particiones cronológicas en entrenamiento, validación y prueba que respeten la secuencia de la serie, y con métricas fuera de muestra R^2 , MAE y RMSE. Se priorizará la interpretabilidad mediante coeficientes y elasticidades en los modelos lineales, y la robustez temporal mediante la estabilidad de errores y del desempeño en ventanas, de acuerdo con buenas prácticas para series de tiempo (Hyndman & Athanasopoulos, 2021).

Particiones y criterio para elegir el modelo final

Dado el carácter temporal de los datos, las particiones se establecen estrictamente en orden cronológico para evitar filtraciones de información futura hacia el entrenamiento. De este modo, la muestra se divide en tres bloques consecutivos: un primer tramo destinado a entrenamiento, un segundo tramo utilizado para validación y un último tramo reservado como prueba fuera de muestra. Esta configuración permite ajustar parámetros y especificaciones sobre el conjunto de entrenamiento, tomar decisiones de selección con base en el desempeño en validación y, finalmente, informar el rendimiento en condiciones no vistas usando el bloque de prueba, que actúa como aproximación a la generalización del modelo en producción. La partición cronológica también garantiza que la evaluación respete la direccionalidad temporal, coherente con las decisiones operativas que se tomarán mensualmente.

El criterio principal de comparación y selección es el MAE (error absoluto medio) calculado en validación, por su interpretación directa en unidades y su alineación con objetivos de negocio (proyecciones en el mismo orden de magnitud que las ventas). Como respaldo se reportan RMSE, MAPE y R^2 para ofrecer una lectura complementaria del tamaño de los errores, su penalización a valores extremos y la proporción de variabilidad explicada. La decisión final pondera tres elementos: el MAE de validación

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

(prioritario), la consistencia entre métricas y la interpretabilidad de la salida (coeficientes/elasticidades o importancias) para su adopción por áreas de compras y planeación. El desempeño final comunicado en el documento corresponde al bloque de prueba y no interviene en la fase de ajuste, cumpliendo así la separación entre desarrollo y evaluación.

Selección de hiperparámetros y especificaciones

En OLS y Log-Log no se realizan búsquedas de hiperparámetros; la especificación se define ex ante con base en evidencias del análisis exploratorio y criterios sustantivos: inclusión de dummies mensuales para capturar estacionalidad, rezagos de la variable objetivo cuando su aporte es justificable, y transformaciones logarítmicas en el caso Log-Log para estabilizar la varianza y habilitar la lectura de elasticidades. Esta decisión preserva la trazabilidad del efecto de cada variable y facilita la discusión con lenguaje económico (signos, magnitudes y elasticidades).

Para Random Forest, dado el tamaño muestral mensual y la necesidad de evitar sobreajuste, se aplica una búsqueda contenida y estructurada de hiperparámetros sobre el bloque de validación: número de árboles, profundidad máxima, tamaño mínimo de hoja, criterio de división, proporción de predictores por partición y uso o no de *bootstrap*. Las combinaciones exploradas son parsimoniosas y comparables entre sí; en todos los casos, el ajuste del modelo se realiza únicamente con el bloque de entrenamiento, y la comparación se basa en el MAE de validación. Se fija una semilla para asegurar replicabilidad de los resultados y, una vez elegida la configuración, se reentrena el modelo con entrenamiento+validación y se informa su desempeño en prueba. Esta estrategia permite balancear capacidad predictiva y estabilidad sin sacrificar la claridad para el usuario final (Hyndman & Athanasopoulos, 2021; James et al., 2021).

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Software y lenguajes utilizados

El desarrollo se implementa en Python 3.x por su ecosistema científico y la posibilidad de mantener un flujo reproducible mediante cuadernos. Para la preparación de datos se emplean pandas y numpy (lectura de Excel/CSV, transformación, unión por fechas y creación de variables); para la estimación de modelos lineales (OLS y Log-Log) se utiliza statsmodels, que facilita reportar coeficientes, intervalos de confianza y pruebas clásicas; para Random Forest y las métricas de evaluación se emplea scikit-learn; la visualización se realiza con matplotlib, generando figuras “Pronóstico vs Real”, distribución de errores y, cuando aplica, gráficos de importancia de variables.

Las salidas intermedias (matrices de diseño, conjuntos particionados, transformaciones aplicadas) y finales (pronósticos, métricas y gráficos) se exportan a CSV/Excel y a imágenes, de modo que el documento de tesis incorpore tablas y figuras directamente reproducibles desde los cuadernos. Esta elección tecnológica mantiene el flujo ligero, controlable y coherente con el alcance de la consultoría, a la vez que cumple el principio de BI de convertir datos en insumos accionables para la decisión (Turban et al., 2020).

Validez de datos y aseguramiento de calidad

Antes de modelar, se ejecutan controles sistemáticos de calidad y coherencia para las fuentes internas y externas. En primer lugar, se verifica el esquema y tipos de datos (fecha en formato válido, variables numéricas sin caracteres, categorías consistentes), la presencia de nulos y duplicados por período, y la alineación temporal mensual entre ventas, clima y epidemiología (todas las series agregadas a la misma frecuencia y a las mismas fechas). En segundo lugar, se revisan rangos plausibles para variables clave

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

(ventas y precios en intervalos realistas; series sanitarias y climáticas dentro de márgenes esperados), registrando cualquier ajuste correctivo con su justificación. Finalmente, se documenta un diccionario de variables con nombres estandarizados, definiciones y unidades, y se etiquetan los insumos con fecha de corte, garantizando comparabilidad entre actualizaciones. Estos pasos sostienen la validez interna del análisis y evitan sesgos originados en discrepancias de formato o cobertura (Kuhn & Johnson, 2019; Turban et al., 2020).

Manejo de valores atípicos

La identificación de atípicos se realiza durante el EDA mediante inspecciones gráficas (series temporales, boxplots, diagramas de dispersión) y resúmenes robustos. Dada la naturaleza mensual y el tamaño acotado de la muestra, la regla general es no eliminar ni modificar observaciones a menos que se trate de errores manifiestos de registro (por ejemplo, duplicación de un mes o un valor imposible por unidad de medida). En esos casos, la corrección se documenta expresamente y se conserva la versión previa del archivo para trazabilidad.

Cuando una observación extrema es verídica y sustantivamente relevante (p. ej., un pico por evento epidemiológico), se mantiene en el modelado y se informa su efecto a través de las métricas y la interpretación de resultados. Si procede, se realizan verificaciones de sensibilidad comparando métricas “con y sin” tratamiento específico (por ejemplo, transformación logarítmica) para asegurarse de que la conclusión no dependa de una decisión puntual. El objetivo es preservar la estabilidad del error y la coherencia explicativa sin introducir sesgos (Hyndman & Athanasopoulos, 2021).

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Reproducibilidad y trazabilidad

La reproducibilidad se garantiza con una organización por cuadernos: EDA (integración y depuración), OLS, Log-Log y Random Forest (preparación específica, entrenamiento, evaluación y generación de figuras). Cada cuaderno contiene celdas secuenciales que ejecutan los pasos en orden, con semillas fijas para cualquier procedimiento aleatorio y con exportación sistemática de resultados (métricas, pronósticos y gráficos) a archivos externos. Los datasets se versionan por fecha de corte, lo que permite saber exactamente qué extracción alimentó cada ejecución; además, se mantiene un registro breve de cambios donde se anota cualquier ajuste en variables, transformaciones o particiones.

Gracias a este esquema, cada tabla y figura del documento puede trazarse a una corrida específica y repetirse con los mismos insumos y decisiones de modelado. Esta trazabilidad cumple el estándar de transparencia y replicación esperado en trabajos aplicados y facilita la futura operacionalización del flujo en artefactos de BI (tableros, reportes o alertas) sin pérdida de control metodológico (Kuhn & Johnson, 2019).

Con esos criterios se seleccionará un modelo de referencia. Este modelo aportará parámetros interpretables o importancias de variables y proyecciones que alimentarán los artefactos de Business Intelligence, como tableros, reportes y alertas orientadas a la decisión. La entrega incluirá la especificación del flujo de datos con insumos, transformaciones y salidas esperadas, junto con los supuestos operativos para su uso recurrente.

Fase 3. Propuesta de plan de intervención para la implementación del modelo de analítica y visualización de datos aplicando la metodología Scrum.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

La presente fase tiene como finalidad definir el plan de intervención diseñado para implementar en Sicmafarma S.A.S el modelo de analítica y visualización de datos a través de la metodología Scrum. Más que una descripción técnica, se plantea un recorrido estructurado que permite comprender cómo se articula el plan, desde los objetivos generales y específicos hasta la definición de roles, artefactos, eventos y fases que darán soporte al proceso de implementación. De este modo, se introduce en la lógica del plan, visualizando cómo cada componente contribuye al desarrollo de un modelo ágil, predictivo y sostenible.

La estructura del capítulo busca generar una inmersión práctica, en la que se explican las metas en corto, mediano y largo plazo, los indicadores de éxito (KPIs), los costos estimados y el cronograma de actividades. Todo ello se presenta con el fin de mostrar de manera clara y progresiva cómo se materializa el modelo de BI planteado siguiendo la metodología Scrum, facilitando no solo su entendimiento, sino también proyectando sus beneficios, alcances y posibilidades de escalamiento dentro de la organización.

Diagnóstico del estado actual de las ventas de Ampicilina + Sulbactam 1,5 g en Sicmafarma S.A.S

Este capítulo ofrece una lectura ordenada del comportamiento reciente de la distribución y ventas de Ampicilina + Sulbactam 1,5 g en Sicmafarma S.A.S. A partir de la serie mensual nacional, se presentan los niveles y la variabilidad de la demanda, los patrones en el tiempo y las relaciones con el entorno (clima ENSO, morbilidad respiratoria y precios de competidores) tal como se evidencia en el EDA. El objetivo es dejar una imagen nítida de la situación vigente y de las brechas que hoy condicionan la planeación comercial y de abastecimiento.

A continuación, se presentan los principales hallazgos del EDA: cómo se distribuye la demanda, cómo cambia según la fase ENSO, cómo se asocia con los precios de la competencia y qué pauta estacional se reconoce a lo largo del año. El foco es descriptivo y comparativo, sin inferir causalidad.

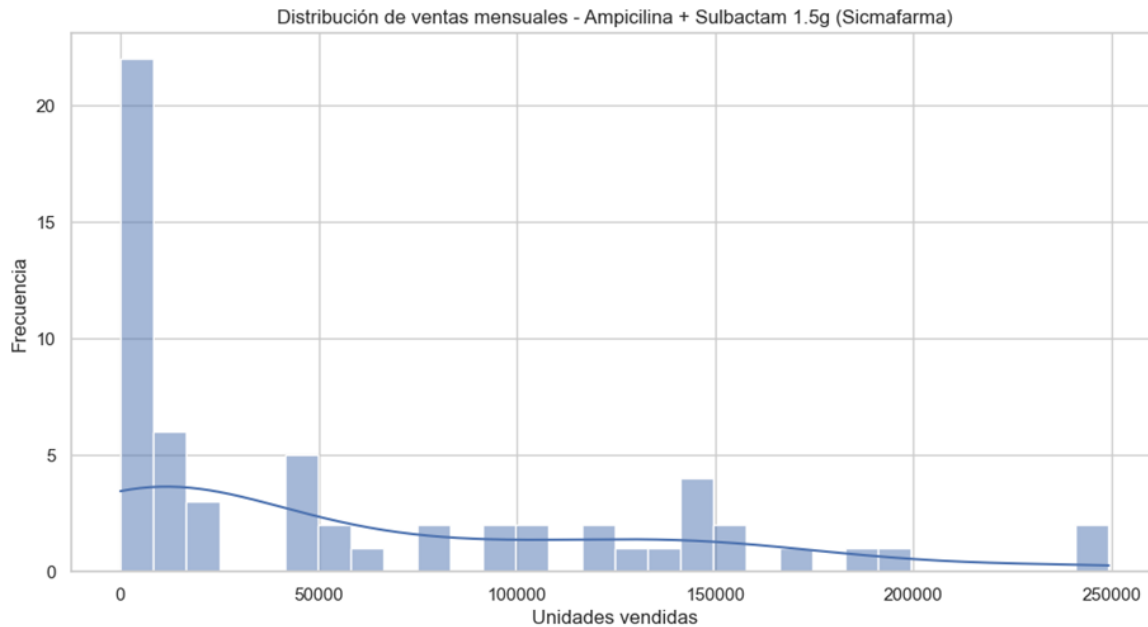
Distribución de la variable dependiente

La distribución de ventas mensuales del medicamento Ampicilina + Sulbactam 1,5 g presenta una asimetría positiva marcada (cola derecha) con fuerte concentración en valores bajos y episodios puntuales de demanda muy alta. En 22 de 60 meses (36,7%) las ventas se ubican por debajo de 12.500 unidades el intervalo más frecuente y el 50% de los meses queda por debajo de 21.350 unidades (mediana). El histograma exhibe una frecuencia decreciente a medida que aumentan las unidades vendidas, con picos aislados en los extremos superiores.

Figura 4

Distribución ventas mensuales de Ampicilina + Sulbactam 1,5 g (2019–2024)

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS



Nota. Elaboración propia con datos internos de ventas de Sicmafarma S.A.S. (2025).

En el resumen descriptivo, el valor máximo es 249.390 unidades. El percentil 75 se sitúa en 107.835 unidades, por lo que un cuarto de los meses supera ese umbral. La media es 60.524,6 unidades y la desviación estándar 69.298,4, cifras que corroboran la alta variabilidad y la distancia entre media y mediana propia de una cola derecha. El mínimo observado es 0; puede corresponder a errores de registro o a meses sin ventas por decisiones institucionales o por inventario. Dada la no normalidad y la heterogeneidad de varianza, resulta pertinente considerar transformaciones logarítmicas en el modelado para estabilizar la varianza y facilitar la lectura de elasticidades.

Desde la lectura económica, los valores extremos son coherentes con compras institucionales masivas, con picos estacionales asociados a aumentos de infecciones respiratorias o con desabastecimientos de sustitutos. La doble concentración —alta densidad de observaciones por debajo de la mediana y una fracción no trivial por encima del percentil 75— sugiere un patrón discontinuo, posiblemente ligado a calendarios presupuestales y a dinámicas logísticas del sistema de salud colombiano.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Operacionalmente, la serie se organiza en tres tramos definidos por cuantiles propios de los datos: tramo inferior, meses con ventas por debajo de la mediana, 21.350 unidades o menos; tramo intermedio, meses entre la mediana y el percentil 75, de 21.350 a 107.835; y tramo superior, meses por encima del percentil 75, más de 107.835, dentro del cual se identifican atípicos superiores según la regla de Tukey, que clasifica como atípicos los valores por encima de $Q3 + 1,5 \text{ veces el rango intercuartílico}$; de forma análoga, los inferiores se ubican por debajo de $Q1 - 1,5 \text{ veces ese rango}$. La acumulación de observaciones en el tramo inferior y la presencia de episodios en el tramo superior explican la asimetría a la derecha y los saltos puntuales del histograma. En este marco, la exposición observable se concentra en faltantes cuando la demanda se ubica en el tramo superior y en excedentes cuando se concentra en el tramo inferior. Esta lectura se desprende de la varianza elevada y de la distancia entre media y mediana.

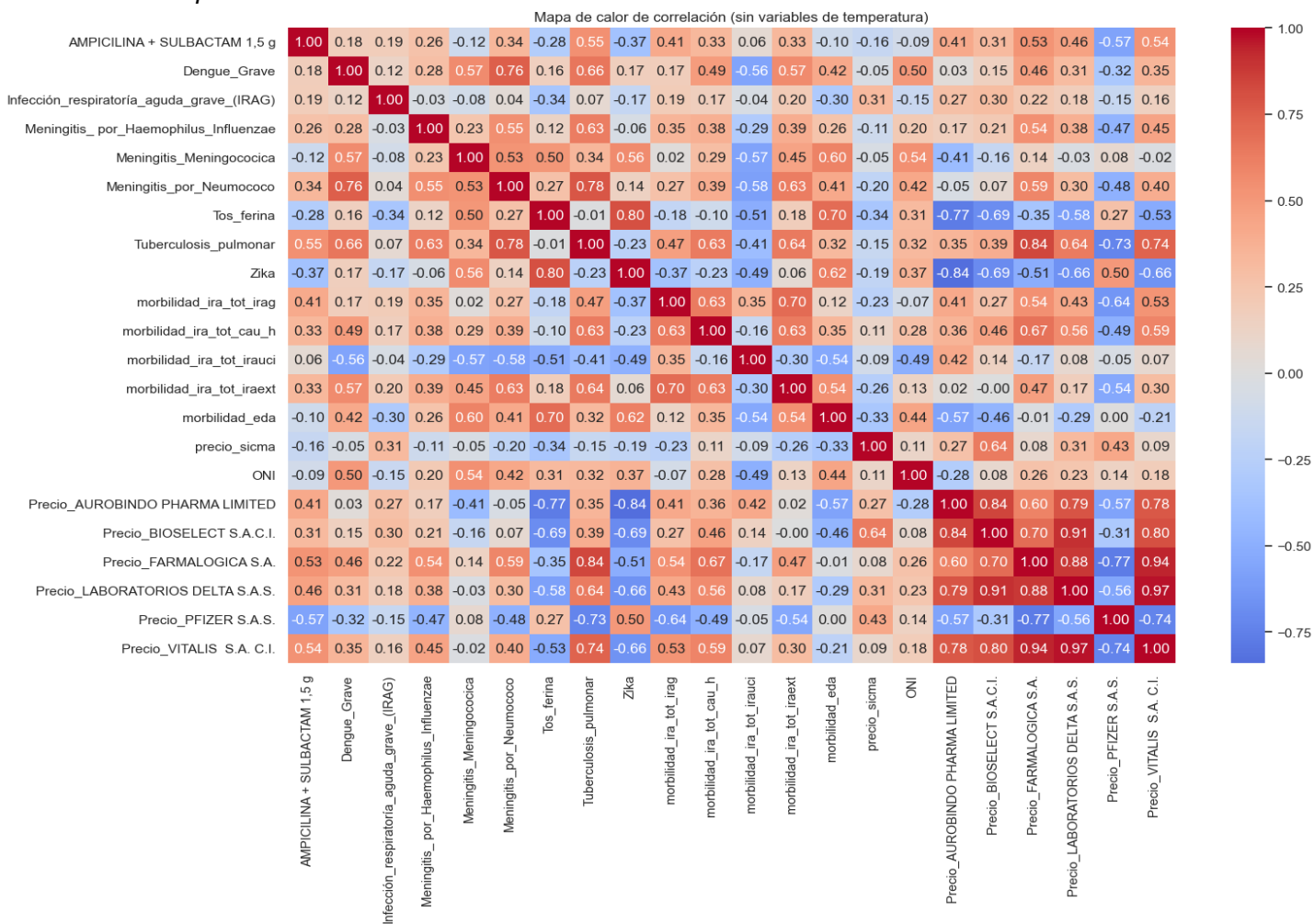
Correlaciones entre variables

El mapa de calor entre variables numéricas muestra relaciones tanto esperadas como llamativas entre los predictores y la variable dependiente. Estas correlaciones son asociaciones lineales contemporáneas en la muestra y deben leerse como señales: orientan el diagnóstico, pero no prueban causalidad ni dirección del efecto.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Figura 5

Mapa de calor de correlaciones entre variables numéricas



Nota. Elaboración propia con datos de ventas y precios de Sicmafarma S.A.S. (2025), epidemiología del INS (2025), clima (temperatura mensual) del IDEAM (2025) y clasificación mensual del fenómeno ENSO del NOAA Climate Prediction Center (2025).

Las ventas de A+S 1,5 g presentan correlación positiva moderada con el precio de VITALIS S.A. C.I. ($r = 0,54$) y con el precio de FARMALÓGICA S.A. ($r = 0,53$). Esto es compatible con un efecto sustitución: cuando esas marcas encarecen, la demanda del producto de Sicmafarma tiende a subir. En contraste, se observa correlación negativa con el precio de PFIZER S.A.S. ($r = -0,57$), lo que sugiere un patrón competitivo inverso (posible influencia de posicionamiento o disponibilidad) asociado a esa marca.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Las magnitudes (en torno a $|0,5-0,6|$) son relevantes para un mercado con múltiples determinantes simultáneos y apuntan a movimientos consistentes con dinámicas de precio-calidad y sustitución en compras institucionales. La relación negativa con PFIZER puede reflejar desplazamientos de demanda cuando esa marca gana terreno (por cobertura o abastecimiento), o comportamientos de portafolio donde los compradores rotan entre marcas líderes y alternativas.

Se identifican correlaciones positivas entre las ventas y tuberculosis pulmonar ($r = 0,55$), meningitis por *Haemophilus influenzae* ($r = 0,26$) y morbilidad total por IRAG ($r = 0,19$). Sin ser elevadas, estas asociaciones son coherentes con aumentos de demanda de antibióticos de amplio espectro en contextos de mayor carga respiratoria/bacteriana.

Las correlaciones epidemiológicas sugieren que picos sanitarios coinciden con mayores ventas del antibiótico. La señal con tuberculosis ($r = 0,55$) podría capturar movimientos estacionales y presión general sobre la red asistencial, más que un vínculo clínico directo, mientras que IRAG y *H. influenzae* aunque de menor magnitud son compatibles con mayor circulación de infecciones respiratorias. El patrón es congruente con la asimetría descrita en la distribución de las ventas: meses de base bajos y meses de evento con valores altos.

Entre predictores. Aparecen correlaciones muy altas entre precios de competidores (FARMALÓGICA-VITALIS: $r = 0,94$; BIOSELECT-DELTA: $r = 0,91$). Este movimiento “en bloque” indica multicolinealidad potencial: varias señales de precio transmiten información muy similar, lo que dificulta separar efectos en enfoques lineales estrictos y afecta la interpretabilidad de coeficientes si se buscaran lecturas causales.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

En la muestra, cuando suben VITALIS y FARMALÓGICA suele subir A+S; cuando sube PFIZER, A+S tiende a bajar; y cuando aumentan ciertas morbilidades respiratorias/bacterianas, las ventas tienden a ser mayores. Al mismo tiempo, varios precios de competidores se mueven casi juntos, por lo que sus relaciones con la demanda pueden solaparse en los análisis.

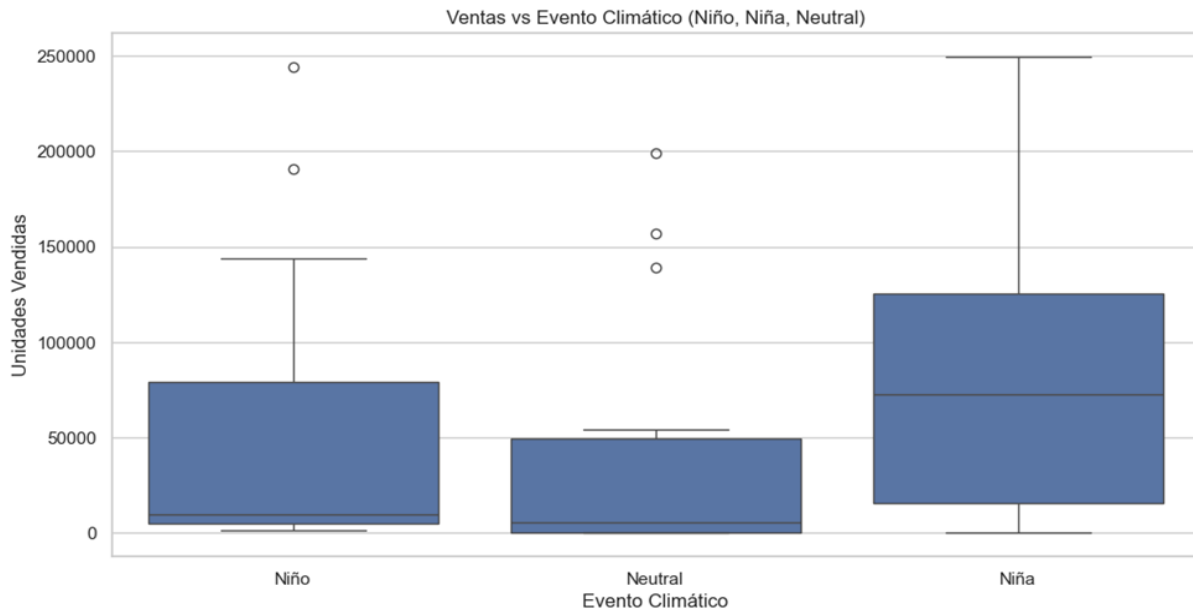
Estas señales son contemporáneas y pueden estar influidas por factores no observados (promociones, licitaciones, quiebres de stock de terceros). Su consistencia sugiere un escenario multifactorial donde el entorno sanitario y la competencia contribuyen a la intermitencia y variabilidad ya descritas.

Evento climático ENSO y comportamiento de la demanda

La variable categórica Evento Climático ENSO permite observar el comportamiento de las ventas ante La Niña, El Niño y Neutral. El boxplot muestra diferencias sustanciales en los niveles centrales por fase: durante La Niña la mediana alcanza 72.700 unidades; en Neutral la mediana es 5.625; y en El Niño se ubica en 9.785 unidades.

Figura 6

Boxplot de ventas mensuales de Ampicilina + Sulbactam 1,5 contra fenómeno ENSO



Nota. Elaboración propia con datos de ventas de Sicmafarma S.A.S. (2025) y fases ENSO (Niño/Neutral/Niña) del NOAA Climate Prediction Center (2025).

La dispersión también cambia por fase. En La Niña el rango intercuartílico va de 15.415 (Q1) a 125.239,5 (Q3) con valores extremos hasta 249.390 unidades; en El Niño el IQR va de 4.875 (Q1) a 79.117,5 (Q3); y en Neutral de 3.750 a 49.212,5 unidades. Esto indica que La Niña concentra tanto niveles más altos como mayor variabilidad.

Este patrón es coherente con la literatura epidemiológica: los eventos de La Niña suelen asociarse con mayor humedad y precipitación, condiciones que contribuyen a la propagación de infecciones respiratorias bacterianas y virales y, por tanto, una mayor demanda de antibióticos de amplio espectro como la ampicilina combinada. Es ampliamente reconocido que estos periodos traen aumentos de humedad/precipitación, lo que facilita la proliferación de agentes respiratorios; en consecuencia, el comportamiento observado es epidemiológicamente consistente y respalda considerar el ENSO como variable explicativa en análisis predictivos.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Las diferencias por fase climática también pueden reflejar reacciones del sistema de salud: EPS/IPS podrían anticipar condiciones climáticas y reforzar compras en periodos con mayor probabilidad de enfermedades respiratorias. Esto introduce una dimensión organizacional en el diagnóstico, donde el ENSO actúa como predictor indirecto de la conducta de adquisición en salud.

En la muestra, La Niña se asocia con ventas medianas más altas y mayor dispersión, mientras que El Niño y Neutral muestran niveles más bajos y menor variabilidad. En conjunto, los datos indican que el ENSO tiene un impacto meteorológico, epidemiológico y logístico sobre la demanda del producto, sensible a la estacionalidad.

Relación entre precio de la competencia y ventas

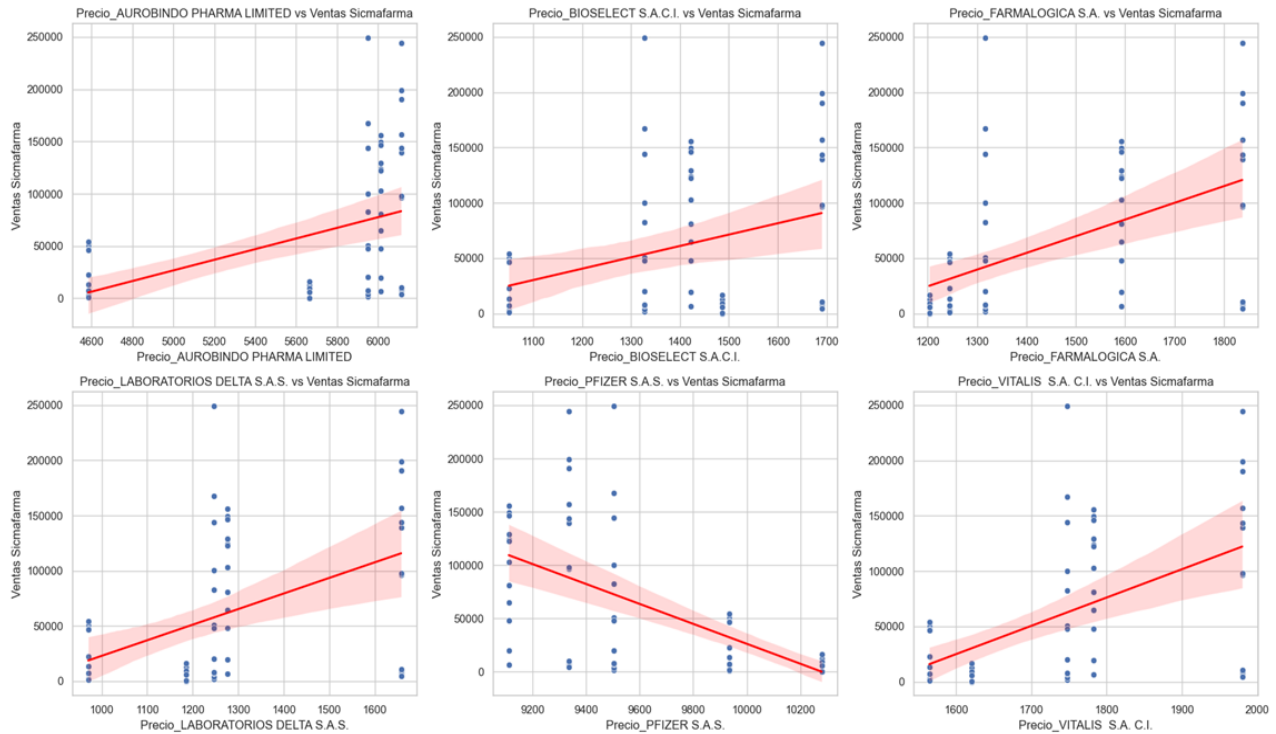
Los diagramas de dispersión permiten analizar la relación entre el precio de diferentes proveedores y las ventas del medicamento Ampicilina + Sulbactam 1,5 g por parte de Sicmafarma S.A.S.

En todos los casos, los diagramas de dispersión muestran nubes de puntos con variabilidad considerable, lo que sugiere que el precio por sí solo no explica completamente el comportamiento de la demanda, pero sí ofrece patrones relevantes cuando se analiza junto con otras variables explicativas. La finalidad de estos gráficos de dispersión no es probar una relación lineal ni causal entre precios de competidores y ventas, sino explorar el sentido (signo) de la asociación contemporánea y priorizar señales para el modelado. Por ello se incluye una línea de tendencia como resumen visual del signo (pendiente positiva o negativa) aun cuando la nube no muestre un patrón lineal marcado.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Figura 7

Dispersión de ventas mensuales de Ampicilina + Sulbactam 1,5 contra precios de las empresas competidoras



Nota. Elaboración propia con datos internos de ventas y precios (competidores) de Sicmafarma S.A.S. (2025).

En el caso del precio de PFIZER S.A.S., se observa una correlación negativa de magnitud apreciable ($r = -0,57$), lo que implica que, a mayor precio de PFIZER, las ventas de Sicmafarma S.A.S. tienden a disminuir. Esta situación podría reflejar un fenómeno de competencia directa entre marcas, donde precios altos de PFIZER coinciden con caídas en la demanda de Sicmafarma S.A.S., posiblemente por efecto de sustitución en clientes sensibles a marca. Por el contrario, se observa una correlación positiva moderada con el precio de VITALIS S.A. C.I. ($r = 0,54$), indicando que cuando VITALIS incrementa su precio, las ventas de Sicmafarma S.A.S. aumentan. Esta situación podría deberse a que VITALIS y Sicmafarma S.A.S. compiten en segmentos similares del mercado institucional, y los compradores se ven incentivados a elegir la opción más económica.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Algo similar ocurre con el precio de FARMALÓGICA S.A., cuya correlación con las ventas de Sicmafarma S.A.S. también es positiva ($r = 0,53$), lo que refuerza la idea de que los aumentos en precios de competidores pueden abrir espacio para captación de demanda en Sicmafarma S.A.S. Sin embargo, no todas las relaciones precio–venta son igualmente relevantes. Por ejemplo, el precio de AUROBINDO PHARMA LIMITED presenta una correlación positiva moderada ($r = 0,41$), pero su impacto puede estar mediado por otros factores, como disponibilidad, cobertura de EPS o reconocimiento de marca.

Este análisis gráfico valida empíricamente los coeficientes de correlación observados en el mapa de calor y aporta evidencia visual para la toma de decisiones sobre estrategias de fijación de precios y monitoreo competitivo.

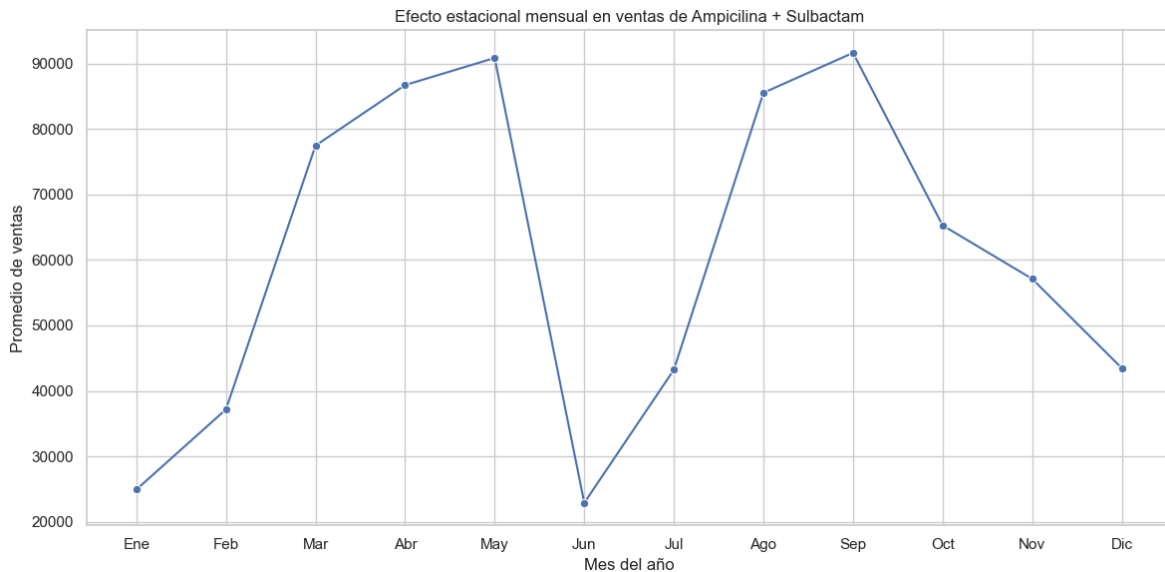
Estacionalidad de la demanda

Análisis de estacionalidad mensual revela patrones cíclicos claros en la demanda del medicamento a lo largo del año. De acuerdo con los promedios mensuales de ventas, los picos más altos se concentran en septiembre (91.626 unidades), mayo (90.850 unidades) y abril (86.707 unidades), seguidos muy de cerca por agosto (85.532 unidades) y marzo (77.468 unidades). Estos meses superan ampliamente el promedio general, evidenciando concentraciones elevadas de demanda que podrían responder a picos epidemiológicos, compras institucionales programadas o campañas sanitarias específicas.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Figura 8

Efecto estacional mensual en la demanda



Nota. Elaboración propia con datos internos de ventas de Sicmafarma S.A.S. (2025).

En contraste, los niveles más bajos de ventas se presentan en enero (24.931 unidades), junio (22.906 unidades) y febrero (37.170 unidades), indicando periodos de menor presión de consumo o desaceleración institucional. El mes de julio presenta un comportamiento intermedio (43.226 unidades), mientras que diciembre (43.487,6 unidades) y noviembre (57.112 unidades) muestran una ligera recuperación que podría estar relacionada con compras anticipadas de cierre de año. Este patrón mensual también podría reflejar la influencia de factores estacionales como las lluvias, los cambios de temperatura y la circulación de enfermedades respiratorias. La secuencia ascendente entre febrero, marzo, abril y mayo sugiere una acumulación de presión sanitaria coincidente con la temporada de transición climática en Colombia. De igual forma, la recurrencia de altos niveles en agosto y septiembre podría vincularse a brotes secundarios o estrategias institucionales de refuerzo terapéutico. Desde una perspectiva operativa, estos datos respaldan la necesidad de modelar la variable mes como un factor explicativo en los modelos predictivos. Identificar los meses de mayor y menor riesgo

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

permite optimizar la logística de producción, distribución y almacenamiento del medicamento, garantizando la disponibilidad durante los periodos críticos y evitando sobrecostos en meses de baja rotación. es un determinante estructural de la demanda y debe ser integrada tanto en el análisis cuantitativo como en la estrategia de negocio.

Implicaciones para la operación

La evidencia empírica perfila una demanda intermitente y asimétrica: muchos meses en niveles bajos y pocos episodios con valores muy altos. Ese patrón se acompaña de alta dispersión y cambios de régimen meses de base vs. meses de evento, lo que en la práctica se traduce en una exposición observable a desbalance: faltantes durante picos y excedentes en valles. Esta exposición se desprende de cifras concretas: 36,7% de los meses por debajo de 12.500 unidades, mediana en 21.350, media en 60.525 y desviación estándar en 69.298; p75 en 107.835 y máximos que alcanzan 249.390. La distancia media–mediana y la presencia de atípicos superiores, de acuerdo con la regla de Tukey, refuerzan ese riesgo operativo.

La sensibilidad a ENSO y a cargas epidemiológicas modifica tanto el nivel como la dispersión por fase: bajo La Niña la mediana asciende a 72.700 con rango intercuartílico de 15.415 a 125.240; en Neutral la mediana cae a 5.625 con IQR de 3.750 a 49.213; en El Niño la mediana es 9.785 con IQR de 4.875 a 79.118. Sumado a ello, la estacionalidad mensual acota ventanas de mayor presión: picos en abril 86.707, mayo 90.850, agosto 85.532 y septiembre 91.626; mínimos en enero 24.931 y junio 22.906, con febrero en 37.170. Diagnóticamente, una fracción relevante de las oscilaciones observadas obedece a condiciones externas que no son constantes a lo largo del año.

En el frente competitivo, coexisten relaciones opuestas PFIZER con signo negativo; VITALIS y FARMALÓGICA con signo positivo y movimientos conjuntos muy altos entre

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

precios de varias marcas. Las ventas de A+S correlacionan 0,54 con el precio de VITALIS y 0,53 con FARMALÓGICA, mientras que con PFIZER la correlación es $-0,57$. Entre competidores, FARMALÓGICA y VITALIS se mueven casi al unísono con 0,94 y BIOSELECT con DELTA con 0,91. En epidemiología, las ventas guardan asociaciones de 0,55 con tuberculosis pulmonar, 0,26 con Haemophilus influenzae y 0,19 con IRAG total. Esto configura un mercado con choques compartidos y señales solapadas; para la operación implica que cambios de ventas y de precios suelen ocurrir en bloques y que la atribución inmediata a un único competidor es frágil cuando las señales son contemporáneas y colineales.

El conjunto de evidencias configura un escenario multifactorial y sensible al entorno, con exposición observable a faltantes en picos y excedentes en valles, además de anomalías puntuales de datos (ceros y registros inconsistentes) y señales solapadas entre competidores. Con esta fotografía, el Objetivo Específico 1 diagnosticar el estado actual de la distribución y ventas del producto se considera cumplido, y se habilita el paso al siguiente objetivo, en el que estas regularidades sirven de insumo para el desarrollo analítico posterior.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en la empresa Sicmafarma S.A.S.

En esta sección se presenta el núcleo del análisis cuantitativo de esta investigación: el desarrollo, la comparación y la selección de modelos econométricos y de aprendizaje automático diseñados para explicar y predecir la demanda mensual del medicamento Ampicilina + Sulbactam 1,5g comercializado por Sicmafarma S.A.S. Se evalúan tres enfoques distintos: Regresión Lineal Múltiple (OLS), Regresión Log-Log y Random Forest, detallando para cada uno su arquitectura, las variables empleadas, el preprocesamiento de datos y sus resultados estadísticos. El objetivo final es identificar el modelo con el mejor balance entre capacidad explicativa, precisión predictiva e interpretabilidad para fundamentar la toma de decisiones estratégicas en la empresa.

Modelo de Regresión Lineal Múltiple (OLS)

El modelo lineal se usa como línea base para cuantificar la relación contemporánea entre las ventas mensuales en unidades y un conjunto de predictores externos. La variable dependiente es ventas_sicma. Como explicativas se incluyen precio propio, precios de competidores principales, enfermedades como tuberculosis pulmonar, morbilidad total por IRAG y las fases del ENSO codificadas con indicadores binarios para Niña y Niño, dejando Neutral como categoría de referencia. La estimación respeta el orden temporal y se evalúa fuera de muestra con particiones cronológicas.

Estructura y Variables

Se utilizaron las ventas unitarias como variable dependiente y un conjunto de variables continuas (precios de Sicmafarma S.A.S y competidores, casos de tuberculosis y morbilidad por IRAG) y categóricas (eventos El Niño/La Niña, codificadas como variables dummy).

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Resultados y Diagnóstico

El modelo obtuvo un R2 de 0.439 (Rajustado2 =0.351), indicando una capacidad explicativa modesta. El coeficiente para el precio propio (precio_sicma) fue de -51.51, lo cual es consistente con la teoría de la demanda, aunque no resultó estadísticamente significativo. Las pruebas de diagnóstico revelaron problemas estructurales severos: una estadística de Durbin-Watson de 1.275 sugirió la presencia de autocorrelación positiva en los residuos, y un Número de Condición de 1.74×10^{20} evidenció multicolinealidad extrema.

Forma funcional modelo ols

$$\begin{aligned} \text{ventas_sicma}_t = & \beta_0 + \beta_1 \text{precio_sicma}_t + \beta_2 \text{Precio_VITALIS_SA_CI}_t + \\ & \beta_3 \text{Precio_FARMALOGICA_SA}_t + \beta_4 \text{Precio_DELTA_SAS}_t + \beta_5 \text{Precio_AUROBINDO}_t + \\ & \beta_6 \text{tuberculosis_pulmonar}_t + \beta_7 \text{morbilidad_ira_tot_irag}_t + \gamma_1 \text{evento_Niña}_t + \\ & \gamma_2 \text{evento_Niño}_t + \varepsilon_t. \end{aligned} \tag{11}$$

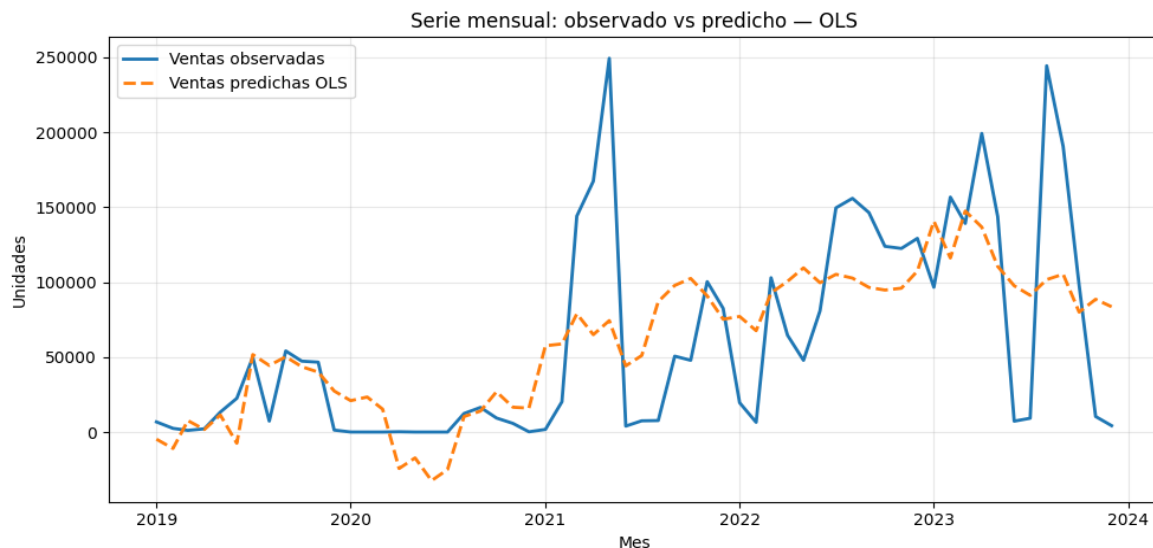
Predicción contra datos reales

La figura (9) "Serie mensual: observado vs predicho OLS" muestra dos líneas: la continua refleja las ventas observadas y la discontinua las ventas predichas por el modelo. El OLS reproduce la tendencia general y parte de las oscilaciones de media frecuencia, pero suaviza los extremos: subestima los picos de mayor magnitud y sobrestima varios valles. Este patrón es visible en los episodios de mayor demanda, donde la línea discontinua no alcanza la amplitud de la línea observada, y en meses de baja rotación, donde la predicción queda por encima del nivel real. Las métricas en escala original dan $R^2 = 0,439$, $MAE \approx 38.400$ unidades y $RMSE \approx 51.457$ unidades. El sesgo promedio es reducido, aunque los errores absolutos crecen en meses extremos, coherente con la asimetría y la alta varianza de la serie.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Figura 9

Ventas reales de Ampicilina + Sulbactam 1,5 frente a las ventas predichas por el modelo OLS



Nota. Adaptado de datos suministrados por Sicmafarma S.A.S, 2025.

Modelo de Regresión Log-Log

Este modelo transforma a logaritmo natural la variable objetivo y las explicativas continuas. Con ello se estabiliza la varianza y los coeficientes se interpretan como elasticidades. Se añaden indicadores para las fases del ENSO y para la estacionalidad mensual, dejando como base la fase Neutral y el mes de enero. La estimación respeta el orden temporal y se evalúa con particiones cronológicas.

Estructura y Variables

$\log(\text{ventas_sicma})_t$ es el logaritmo natural del Total, mensual de unidades vendidas. Las variables con logaritmo natural son continuas: precio propio, precios de competidores, e indicadores epidemiológicos. evento_Niña_t y evento_Niño_t valen 1 cuando la fase está activa. mes_m_t son dummies para febrero a diciembre con enero como referencia. En log-log, un aumento de 1% en una explicativa continua se asocia con un cambio

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

aproximado de β_k por ciento en las ventas. Para una dummy, el efecto porcentual es $100(e^\gamma - 1)$

Resultados y Diagnóstico

Este enfoque demostró ser notablemente superior. Alcanzó un R2 de 0.851 (Rajustado2 =0.702), explicando más del 85% de la variabilidad en las ventas. Las variables clave como el precio del competidor VITALIS (log_precio_VITALIS) y el evento climático La Niña (evento_Niña) resultaron altamente significativas ($p < 0.05$). El diagnóstico del modelo fue robusto, con factores de inflación de la varianza (VIF) en niveles aceptables.

Depuración de datos

Para estimar el modelo log-log se depuró la base con dos reglas claras. Primero, todas las series continuas usadas en logaritmos ventas, precios propios y de competidores, e indicadores epidemiológicos debían ser estrictamente positivas; los meses con ceros o faltantes se excluyeron. Esta práctica es estándar en trabajos con transformaciones logarítmicas porque evita valores indefinidos, estabiliza la varianza y facilita la interpretación de elasticidades, además de reducir el peso de magnitudes extremas en la estimación, según la literatura reciente. Segundo, se aplicó la regla de Tukey por variable y a la intersección: se marcaron como atípicos los valores por encima de Q3 más 1,5 IQR o por debajo de Q1 menos 1,5 IQR y se retuvieron solo los meses que quedaron dentro de esos límites en todas las variables. El rango intercuartílico IQR es la distancia entre el tercer y el primer cuartil e identifica el 50 % central de la distribución; usarlo permite delimitar outliers de forma robusta y mejorar la estabilidad de los coeficientes en presencia de colas pesadas.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

La combinación de ambas reglas redujo la muestra de 60 a 43 meses. Los excluidos se concentran al inicio del periodo por ceros y valores extremos, de ahí que las figuras del log–log arranquen más tarde que las del OLS. La ventana efectiva de estimación es, por tanto, un subconjunto de 2019–2023, decisión documentada para asegurar trazabilidad y comparabilidad. El propósito de estos filtros es estabilizar la varianza y evitar que pocos episodios anómalos dominen las elasticidades; a cambio, se reconoce que si ciertos extremos reflejan situaciones reales de mercado su eliminación puede sesgar la lectura de esos episodios. Por ello, los resultados del log–log se contrastan con un OLS sobre la muestra completa y se advierte que las ventanas de entrenamiento no coinciden.

Forma funcional modelo log – log

$$\begin{aligned} \log(\text{ventas_sicma})_t = & \beta_0 + \beta_1 \log(\text{precio_sicma})_t + \beta_2 \log(\text{Precio_VITALIS_SA_CI})_t + \\ & \beta_3 \log(\text{Precio_AUROBINDO})_t + \beta_4 \log(\text{Tuberculosis_pulmonar})_t + \\ & \beta_5 \log(\text{morbilidad_ira_tot_irag})_t + \beta_6 \log(\text{Meningitis_por_Neumococo})_t + \\ & \beta_7 \log(\text{morbilidad_ira_tot_cau_h})_t + \beta_8 \log(\text{morbilidad_ira_tot_iraext})_t + \gamma_1 \text{evento_Niña}_t + \\ & \gamma_2 \text{evento_Niño}_t + \sum_{m=2}^{12} \delta_m \text{mes}_{m,t} + \varepsilon_t. \end{aligned} \quad (12)$$

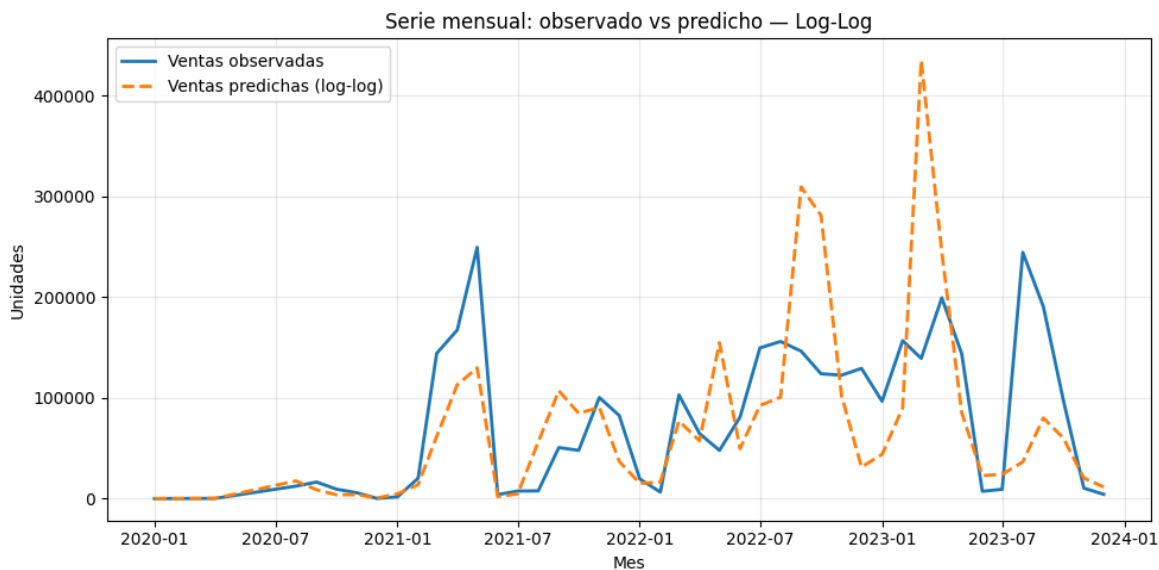
Predicción contra datos reales

La figura (10) “Comparación de ventas reales vs. predichas log–log” grafica la serie en escala original: línea continua para las ventas observadas y línea discontinua para las ventas predichas tras re transformar $\exp(\widehat{\log y})$. El modelo sigue la tendencia y mejora la sincronía de varios ascensos y descensos respecto al OLS, aunque aún suaviza algunos picos muy altos y sobrestima ciertos valles. Este comportamiento es coherente con la comprensión del logaritmo y con la colinealidad señalada. La lectura visual confirma que el log–log captura mejor la dinámica media y la estacionalidad, a la vez que reduce la sensibilidad a atípicos positivos.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Figura 10

Ventas reales de Ampicilina + Sulbactam 1,5 frente a las ventas predichas por el modelo Log-Log.



Nota. Adaptado de datos suministrados por Sicmafarma S.A.S, 2025.

Modelo de Random Forest

Se implementó un Random Forest para capturar no linealidades e interacciones entre predictores que los modelos lineales no recogen. La variable objetivo se trabajó en $\log_{1p}(\log(\text{ventas}+1))$ para suavizar la varianza y luego se retransformó a unidades para evaluar errores. Se incluyeron todas las variables numéricas y las categóricas como dummies. Para garantizar comparabilidad, se congelaron los hiperparámetros obtenidos en la búsqueda previa: $n_estimators = 477$, $max_depth = 47$, $min_samples_split = 2$, $min_samples_leaf = 1$, $max_features = \log_2$ y $bootstrap = True$ (con semilla fija para reproducibilidad).

Evaluación y resultados

Con una partición hold-out 80/20 (entrenar con el 80 % de los meses y probar con el 20 % restante), el desempeño fue $MAE \approx 25.192$ unidades, $RMSE \approx 32.104$ y $R^2 = 0,622$. Visualmente, el bosque sigue la tendencia y varias oscilaciones de magnitud media, pero

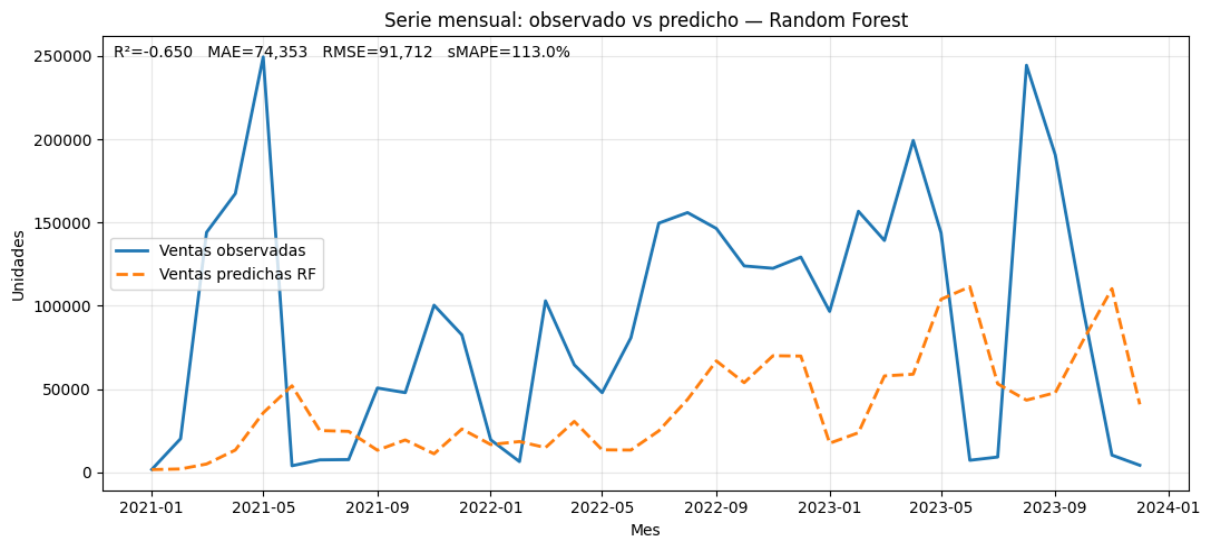
Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

suaviza los picos más altos y sobreestima algunos valles (véase Figura RF-B: Observado vs. predicho).

Para una prueba temporal más estricta se aplicó una validación *walk-forward* (entrenar hasta t y predecir $t + 1$ recorriendo toda la serie). En este esquema el rendimiento cayó a $R^2 \approx -0,65$, $MAE \approx 74.353$, $RMSE \approx 91.712$. Este deterioro es coherente con: (i) la rareza de picos muy altos, que el ensamble tiende a suavizar; (ii) el uso de menos datos en cada corte temporal; y (iii) la compresión de extremos al volver de \log_{1p} a unidades.

Figura 11

Ventas reales de Ampicilina + Sulbactam 1,5 frente a las ventas predichas por el modelo Random Forest



Nota. Adaptado de datos suministrados por Sicmafarma S.A.S, 2025.

Importancia de variables

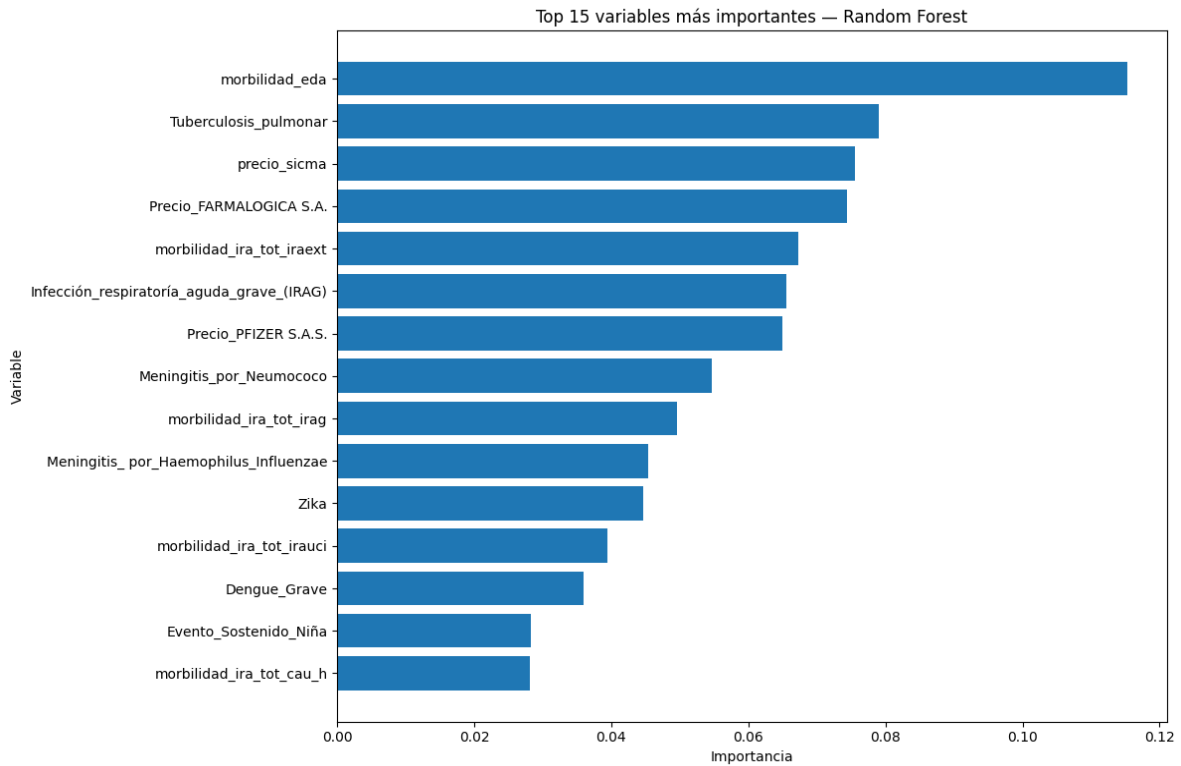
El ranking por importancia (reducción promedio de impureza) ubica como señales más influyentes a morbilidad por EDA, tuberculosis pulmonar y precio propio, seguidas por precio de FARMALÓGICA, IRAG y precio de PFIZER; la fase Niña aparece con menor peso relativo (véase Figura RF-A: Importancia de variables). En conjunto, el bosque

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

sugiere que la presión epidemiológica y el frente de precios explican buena parte de la variación que el algoritmo aprende.

Figura 12

Importancia de las variables en el modelo Random Forest



Nota. Adaptado de datos suministrados por Sicmafarma S.A.S, 2025.

El modelo Random Forest aporta buena precisión en la prueba estática y un mapa claro de factores asociados; bajo validación temporal estricta, pierde capacidad para anticipar picos poco frecuentes, por lo que su mayor valor está en capturar patrones medios y priorizar señales para monitoreo.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Selección del modelo

La selección del modelo se basó en una evaluación integral de capacidad explicativa (R2), precisión predictiva (MAE, RMSE), estabilidad estadística, interpretabilidad y aplicabilidad operativa

Tabla 4

Comparación de resultados entre los modelos

Modelo	R ²	MAE	RMSE	Interpretabilidad
Regresión Lineal (OLS)	0.439	38,400.14	51,456.97	Alta
Regresión Log-Log	0.851	<i>0.6996 (log)</i>	<i>0.8200 (log)</i>	Media-Alta
Random Forest (completo)	0.6224	25,191.52	32,103.67	Baja

Nota. Elaboración propia

Luego de realizar la validación con los 3 modelos, el modelo seleccionado es Regresión Log-Log. Esta elección se fundamenta en su superioridad explicativa (R2=0.851), que captura la dinámica del negocio de manera sobresaliente. Aunque el Random Forest ofrece una precisión predictiva ligeramente mayor en términos absolutos, el modelo log-log proporciona un marco interpretable que es crucial para la toma de decisiones. Permite no solo predecir las ventas, sino también entender *por qué* cambian, cuantificando el impacto de factores clave como precios, clima y salud pública a través de elasticidades.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Comparación de Métricas de Error

Es fundamental subrayar que las métricas de error predictivo, como el Error Absoluto Medio (MAE) y la Raíz del Error Cuadrático Medio (RMSE), no son directamente comparables entre el modelo log-log y los modelos de regresión lineal o Random Forest. Esta inconmensurabilidad surge de una diferencia fundamental en la escala de la variable objetivo. Mientras que los modelos OLS y Random Forest predicen las ventas en sus unidades originales, el modelo log-log predice el logaritmo natural de las ventas. En consecuencia, el error del primer grupo se mide en unidades físicas del producto, mientras que el error del segundo se mide en una escala logarítmica, que comprime la magnitud de las desviaciones.

Por lo tanto, un valor numéricamente inferior del MAE en el modelo log-log no implica una superioridad predictiva en la escala original. Para realizar una comparación equitativa, sería necesario transformar las predicciones logarítmicas de vuelta a su escala original mediante la función exponencial antes de calcular las métricas de error.

Sin embargo, la selección del modelo log-log no se fundamenta en la minimización de este error absoluto, sino en su superioridad explicativa (R^2) y su valiosa capacidad para interpretar los coeficientes como elasticidades, un atributo esencial para la toma de decisiones estratégicas que los otros modelos no ofrecen con la misma robustez (Wooldridge, 2016). La elección, por tanto, prioriza la interpretabilidad y el entendimiento del negocio sobre la precisión predictiva puntual.

Interpretación Estratégica del Modelo Log-Log seleccionado

El modelo adopta una especificación log-log para variables continuas, donde los coeficientes β se interpretan como elasticidades. Los coeficientes γ y δ , asociados a variables categóricas, indican variaciones promedio en las ventas respecto a una

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

categoría base. La forma funcional estimada que incluye únicamente las variables significativas es:

$$\log(\text{ventas_sicma}) = \beta_0 + \beta_1 \log(\text{precio_VITALIS}) + \gamma_1 \text{evento_Niña} + \gamma_2 \text{evento_Niño} + \beta_2 \log(\text{tuberculosis_pulmonar}) + \delta_3 \text{mes_marzo} + \delta_9 \text{mes_septiembre} + \varepsilon \quad (13)$$

A continuación, se presenta la interpretación de los hallazgos más relevantes desde una perspectiva de negocio:

Elasticidad Cruzada de la Demanda (Competencia)

- $\log_precio_VITALIS$ (Coeficiente: 49.11, $p=0.029$): Este es el hallazgo más significativo. Indica que por cada aumento del 1% en el precio del competidor VITALIS, las ventas de Sicmafarma S.A.S aumentan un impresionante 49.1%. Esto revela un efecto de sustitución extremadamente fuerte entre ambos productos. Los clientes son altamente sensibles al precio relativo, y Sicmafarma S.A.S puede capitalizar directamente los aumentos de precio de su principal competidor.

Factores Climáticos (ENSO)

- $evento_Niña$ (Coeficiente: 2.55, $p=0.004$): Durante los meses clasificados bajo el fenómeno de La Niña, las ventas aumentan en promedio un 2.5%. Este efecto, estadísticamente significativo, es coherente con el aumento de la humedad y las precipitaciones, que se asocian a una mayor incidencia de infecciones respiratorias bacterianas.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

- evento_Niño (Coeficiente: -1.62, $p=0.104$): De forma inversa y marginalmente significativa, los eventos de El Niño se asocian a una reducción de las ventas del 1.6%.

Factores Epidemiológicos

- log_tuberculosis_pulmonar (Coeficiente: -5.39, $p=0.099$): Este resultado, marginalmente significativo, es contraintuitivo pero revelador. Un aumento del 1% en los casos de tuberculosis se asocia a una disminución del 5.4% en las ventas de Ampicilina+Sulbactam. Se postula que esto no se debe a una relación clínica directa, sino a un efecto de desplazamiento de recursos: los protocolos para la tuberculosis requieren antibióticos específicos y de mayor control, lo que podría desviar la atención y el presupuesto de los prescriptores de antibióticos de amplio espectro como el analizado.

Estacionalidad

- Se identificaron picos de demanda estadísticamente significativos en marzo (coef: 4.56) y septiembre (coef: 4.68), en comparación con el mes de referencia, enero. Esta marcada estacionalidad valida la existencia de ciclos epidemiológicos o patrones de compra institucionales que el modelo ahora puede cuantificar y prever.

Elasticidad Precio de la Demanda (Propia)

- log_precio_sicma (Coeficiente: -4.21, $p=0.336$): Aunque el coeficiente sugiere que un aumento del 1% en el precio propio reduciría las ventas

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

en un 4.2%, este resultado no es estadísticamente significativo. La alta correlación con los precios de la competencia dificulta aislar su efecto individual.

Síntesis de hallazgos y selección del modelo

La evidencia muestra que la demanda de Ampicilina + Sulbactam 1,5 g no es aleatoria: responde de forma sistemática a un conjunto acotado de determinantes de mercado, epidemiología y clima, además de presentar una estacionalidad reconocible. Esta combinación de señales explica tanto los meses de base como los episodios de mayor demanda y habilita el uso de modelos predictivos para planear producción, compras y distribución.

Entre las alternativas evaluadas, la especificación log–log ofrece el mejor equilibrio entre capacidad explicativa e interpretabilidad. Al expresar los efectos en términos de elasticidades, permite traducir cambios porcentuales en precios o en indicadores sanitarios en variaciones porcentuales de ventas, lo que facilita escenarios “¿qué pasa sí?” y decisiones tácticas con fundamento cuantitativo. Este rasgo, sumado a su ajuste estadístico, respalda su adopción como modelo de referencia para la toma de decisiones.

En el frente competitivo, el resultado más relevante es la fuerte elasticidad cruzada con el precio de VITALIS: los movimientos de ese actor se asocian con variaciones de gran magnitud en la demanda de Sicmafarma, por encima del efecto estimado del precio propio. La implicación operativa es clara: la vigilancia sistemática de precios de competidores y la reacción oportuna de la política comercial pueden generar ganancias sustantivas en volumen.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Por último, los factores externos (fases ENSO y carga epidemiológica) dejan de ser señales cualitativas para convertirse en variables cuantificables que afectan nivel y dispersión de la demanda. Su incorporación al modelo permite anticipar ventanas de mayor riesgo y ajustar inventarios y logística con antelación.

Con estas conclusiones se cierra el segundo objetivo del estudio: desarrollar, contrastar y fundamentar la selección de un modelo predictivo con poder explicativo y utilidad práctica para Sicmafarma S.A.S., dejando como insumo operativo un marco interpretable para la decisión y un conjunto de indicadores priorizados para monitoreo continuo.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Propuesta de plan de intervención para la implementación del modelo de analítica y visualización de datos aplicando la metodología Scrum

En un entorno empresarial altamente dinámico como el farmacéutico, caracterizado por la variabilidad de la demanda, la sensibilidad a los precios y la influencia de factores epidemiológicos y climáticos, contar con herramientas de apoyo a la decisión es esencial para garantizar la competitividad y sostenibilidad. En este contexto, la integración de un modelo de analítica permite transformar grandes volúmenes de datos internos y externos en información procesada, confiable y estratégica, que facilita anticipar escenarios de mercado y optimizar la gestión de inventarios.

Sin embargo, la sola implementación de BI no es suficiente. Se requiere un marco metodológico que asegure entregas tempranas, iterativas y de alto impacto para la organización. Aquí, la metodología Scrum resulta clave, ya que proporciona un enfoque ágil y flexible que promueve la colaboración entre áreas, reduce la incertidumbre en la planificación y permite ajustar de manera continua los desarrollos analíticos a las necesidades reales del negocio. La combinación de Scrum y analítica transforma el proceso de análisis de datos en una dinámica ágil de aprendizaje y mejora constante.

A continuación, se relacionan los objetivos del plan para la implementación del modelo de analítica y visualización de datos mediante la metodología Scrum en Sicmafarma S.A.S.

Objetivo general

Implementar un modelo de analítica soportado en la metodología ágil Scrum, para predecir y gestionar la demanda del medicamento Ampicilina + Sulbactam 1,5 g, optimizando la toma de decisiones comerciales, la planificación de inventarios y la capacidad de respuesta ante factores epidemiológicos y climáticos.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Objetivos específicos

- Aplicar la metodología Scrum en el proceso de implementación del modelo de analítica, asegurando entregas iterativas, retroalimentación temprana y ajustes continuos que respondan a las necesidades comerciales y operativas de Sicmafarma S.A.S.
- Fomentar la colaboración interdisciplinaria entre las áreas comercial, de planeación, logística y de tecnología, para garantizar la alineación de objetivos y la pertinencia de los desarrollos analíticos.
- Establecer una cultura ágil de trabajo mediante la adopción de Scrum, garantizando revisiones continuas y aprendizaje organizacional.
- Definir indicadores clave (KPIs) para medir la adopción de metodología scrum en la organización.

Metodología Scrum

La implementación del modelo de analítica en Sicmafarma S.A.S. seguirá rigurosamente el marco de trabajo Scrum, cuyo propósito central es asegurar la entrega incremental y continua de valor en un entorno altamente dinámico como el farmacéutico, donde factores como estacionalidad, epidemiología y precios de la competencia influyen directamente en la demanda del producto Ampicilina + Sulbactam 1,5 g.

A continuación, se describirán los elementos (roles, artefactos, que componen la metodología Scrum y la propuesta de aplicación en Sicmafarma S.A.S.

Roles

En esta sección se detallan los roles fundamentales de Scrum aplicados al proyecto, los cuales aseguran la correcta distribución de responsabilidades, fomentan la colaboración

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

entre áreas y garantizan que los entregables generen valor tangible para Sicmafarma S.A.S.

En la tabla 4 se relacionan las funciones para cada rol.

Tabla 5

Principales funciones de cada uno de los roles.

Rol	Funciones principales
<i>Product Owner</i> (Gerencia Comercial)	Prioriza necesidades del negocio y asegura alineación con la estrategia corporativa. Es responsable de definir y establecer la prioridad de los reportes que deben generarse (ventas, ajustes de inventarios o reportes de precio). Su rol garantiza que los incrementos del modelo BI generen valor real en la gestión comercial y logística.
<i>Scrum Master</i> (Analista BI/Agile Coach)	Asegura que el marco Scrum se aplique correctamente, eliminando impedimentos como problemas de acceso a bases de datos o retrasos en la integración con el ERP. En este proyecto, actúa como puente entre el área técnica y la gerencia comercial, facilitando la adopción ágil del modelo BI.
<i>Developers:</i> Equipo Multidisciplinario (Comercial, Planeación, Logística, TI)	Ejecuta el trabajo técnico y funcional. El área de logística valida escenarios de stock, el área de TI realiza la limpieza e integración de datos y el área comercial define criterios de elasticidad precio-demanda. Su naturaleza multidisciplinaria garantiza que los entregables BI sean accionables para toda la organización.

Nota. Adaptado de "Guide to Scrum" de Schwaber & Sutherland, 2020.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Artefactos

Constituyen elementos esenciales en Scrum, pues permiten organizar, dar visibilidad y controlar el avance del proyecto. Aquí se explican los principales artefactos empleados en el modelo BI y cómo contribuyen al éxito de la implementación.

En la tabla 5 se describe los artefactos que componen la metodología Scrum.

Tabla 6

Descripción de los artefactos

Artefacto	Descripción
<i>Product Backlog</i>	Constituye una lista priorizada que contiene las iniciativas necesarias para el éxito del proyecto; por ejemplo, se determinó que la creación de reportes de ventas frente a la demanda real debía priorizarse antes que la generación de reporte de precios.
<i>Sprint Backlog</i>	Ciclo de trabajo de duración fija (generalmente entre dos y cuatro semanas) durante el cual el equipo de desarrollo lleva a cabo un conjunto de tareas priorizadas del <i>Product Backlog</i> , por ejemplo: Limpiar datos duplicados o inconsistentes de la base de datos de las ventas.
Incremento	Producto utilizable al final de cada sprint. Por ejemplo: reporte BI de demanda mensual, un algoritmo validado en ambiente de pruebas, o un reporte automatizado de stock crítico.

Nota. Adaptado de "Guide to Scrum" de Schwaber & Sutherland, 2020

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Eventos

Los eventos Scrum proporcionan un marco de reuniones estructuradas que facilitan la planificación, el seguimiento y la mejora continua. Esta sección introduce los eventos clave que guiarán el ritmo de trabajo del equipo y asegurarán la entrega incremental de valor. En la tabla 6 se observa cada evento de scrum con su respectiva descripción.

Tabla 7

Eventos de scrum

Elemento	Descripción
<i>Sprint</i>	Reunión inicial donde se decide qué historias del backlog se abordarán.
<i>planning</i>	En el caso del primer sprint se puede priorizar la conexión de datos internos (ventas e inventarios) con datos externos, factor clima y epidemiológico (ENSO, INS).
<i>Daily scrum</i>	Reuniones diarias de 15 minutos, realizadas cada mañana de forma presencial o mediante Microsoft Teams, en las que el equipo revisa el progreso alcanzado. Durante estas sesiones se responde a las preguntas: ¿qué se hizo ayer?, ¿qué se está haciendo hoy? y ¿qué se hará mañana?, además de identificar posibles impedimentos que afecten el avance del Sprint, como por ejemplo la demora en la disponibilidad de datos epidemiológicos.
<i>Sprint review</i>	Al final de cada sprint, el equipo presenta los reportes o modelos desarrollados a la gerencia comercial y logística, quienes validan si los entregables responden a sus necesidades y retroalimentan el backlog.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Espacio interno del equipo Scrum para identificar mejoras en el proceso.

Sprint Puede implicar ajustar la frecuencia de integración de datos o redefinir
retrospective criterios de aceptación para reportes más intuitivos.

Nota. Adaptado de "Guide to Scrum" de Schwaber & Sutherland, 2020

Metas

Son estructuradas en tres horizontes temporales: corto, mediano y largo plazo, con el propósito de guiar la implementación y evolución del modelo de analítica en Sicmafarma S.A.S. Este enfoque progresivo permite iniciar con resultados piloto, avanzar hacia la optimización y, finalmente, consolidar el modelo como parte integral de los procesos estratégicos de la compañía.

- Corto Plazo (0–3 meses)
 - Capacitación al 100% al equipo comercial, planeación, logística y TI en metodología Scrum.
 - Consolidación de datos que se actualice periódicamente con información de fuentes internas (ventas, inventarios y precios) y fuentes externas (factores climáticos, epidemiológicos y precios de la competencia).
- Mediano Plazo (4–8 meses)
 - Adopción de la metodología scrum por parte del equipo comercial, logística y TI, a través de la definición de roles, uso de artefactos y eventos.
 - Alcanzar una precisión predictiva $\geq 85\%$ (R^2 del modelo log–log).
 - Desplegar sistemas de alertas automáticas ENSO/IRAG para anticipar picos de demanda.
- Largo Plazo (9–12 meses)
 - Diseño de tablero BI para monitorear las ventas reales del medicamento de Ampicilina + Sulbactam vs ventas pronosticadas.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

- Escalar el modelo de analítica a dos moléculas adicionales de alto impacto en el portafolio de la organización.

Indicadores de éxito (KPIs)

Los indicadores de éxito (KPIs) son herramientas clave para medir la adopción de la metodología scrum en la implementación del modelo. A continuación, se observan los KPIs con su respectiva información de ficha técnica.

- Índice de sensibilidad de la demanda al precio (ISDP)

-Definición: Mide cómo cambia la demanda de Ampicilina + Sulbactam frente a las variaciones en el precio promedio del mercado de la competencia.

-Objetivo: Identificar el grado de sensibilidad de la demanda al precio para orientar decisiones comerciales, ajustes de precios y estrategias promocionales.

-Fórmula:

$$\text{Índice de sensibilidad de la demanda al precio} = \frac{\% \text{Variación de la demanda}}{\% \text{Variación del precio promedio del mercado}}$$

-Unidad de medida: Número índice.

- Meta: Sin valor fijo. Si **ISDP > 1**, la demanda es sensible al precio y se recomienda revisar márgenes y estrategias comerciales; si **ISDP < 1**, la demanda es inelástica.

- Frecuencia de medición: Mensual.

- Fuente de datos: Registros de ventas de Sicmafarma S.A.S. y precios promedio del mercado de la competencia.

- Responsable de medición: Scrum master (Gerencia comercial).

- Índice de impacto epidemiológico en la demanda (IIED)

-Definición: Cuantifica el efecto de eventos epidemiológicos (brotes infecciosos, temporadas de gripe u otros eventos de salud pública) sobre la demanda de Ampicilina + Sulbactam.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

-Objetivo: Detectar incrementos o disminuciones asociados a eventos epidemiológicos para anticipar necesidades de inventario y ajustar el stock de seguridad.

-Fórmula:

$$IIED = \frac{\text{Demanda durante evento epidemiológico} - \text{Demanda promedio} \times 100}{\text{Demanda promedio}}$$

-Unidad de medida: Porcentaje (%).

- Meta: Sin valor único predefinido. Si **IIED > 0**, el evento está incrementando la demanda y se requiere reforzar abastecimiento; si **IIED < 0**, indica reducción y permite evitar sobre inventarios.

- Frecuencia de medición: Según la ocurrencia del evento (semanal o quincenal durante el brote).

- Fuente de datos: Registros de ventas de Sicmafarma S.A.S. e información oficial sobre eventos epidemiológicos.

- Responsable de medición: Scrum master (Gerencia comercial).

- Índice de mejora continua (retrospectiva)

- Definición: Mide el porcentaje de acciones de mejora identificadas en las retrospectivas que fueron implementadas en los siguientes sprints, evidenciando el grado de compromiso del equipo con la mejora continua.

- Objetivo: Evaluar la efectividad de las retrospectivas y el nivel de adopción de acciones de mejora dentro de la metodología Scrum, para la implementación del modelo de BI definido para predecir y gestionar la demanda del medicamento Ampicilina + Sulbactam 1,5 g

- Fórmula:

$$\text{Índice de mejora continua} = \frac{\text{Acciones implementadas}}{\text{Acciones identificadas}} \times 100 \quad (14)$$

- Unidad de medida: Porcentaje (%).

- Meta: $\geq 70\%$.

- Frecuencia de medición: Por sprint.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

- Fuente de datos: Actas de retrospectivas y seguimiento de backlog de mejoras.
- Responsable de medición:
Scrum master (Gerencia comercial).
Developers: Equipo Multidisciplinario (Comercial, Planeación, Logística y TI)

- Índice de integración interdisciplinaria

- Definición: Nivel de participación de las áreas claves (comercial, planeación, tecnología, etc.) en la construcción del backlog y en la ejecución de entregables.
- Objetivo: Evaluar el grado de colaboración e integración entre las diferentes áreas de Sicmafarma en torno al modelo BI aplicado a través de metodología Scrum.
- Fórmula:

$$\text{Índice de integración interdisciplinaria} = \frac{\text{Áreas participantes activamente}}{\text{Áreas convocadas}} \times 100 \quad (15)$$

- Unidad de medida: Porcentaje (%).
- Meta: $\geq 75\%$.
- Frecuencia de medición: Por sprint.
- Fuente de datos: Registros de participación en backlog, actas de sprint review y reportes de colaboración interáreas.
- Responsable de medición:
Product Owner (Gerencia comercial)
Developers: Equipo Multidisciplinario (Comercial, Logística, TI)

- Ventas perdidas por quiebre

- Definición: Mide el porcentaje de demanda que no se despacha por falta de inventario, evidenciando el impacto operativo de los quiebres de stock.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

- Objetivo: Cuantificar y reducir la demanda insatisfecha para mejorar servicio, ventas y utilización de capital.
- Fórmula:
$$\text{Lost Sales} = \frac{\text{Unidades no despachadas por falta de stock}}{\text{Demanda total}} \times 100 \quad (16)$$
- Unidad de medida: Porcentaje (%).
- Meta: $\geq 2\%$.
- Frecuencia de medición: mensual.
- Fuente de datos: seguimiento mensual.
- Responsable de medición:
Product Owner (Gerencia comercial)
Developers: Equipo Multidisciplinario (Comercial, Logística, TI)

Costos estimados

La estimación de costos permite identificar y planificar los recursos financieros necesarios para llevar a cabo la implementación del modelo BI con Scrum. En esta sección se detallan las inversiones en licencias, infraestructura, consultoría y capacitación.

Tabla 8

Estimación de costos

Categoría	Costo estimado (COP)	Descripción
Licencias BI	\$18.000.000/año	Power BI Pro o Tableau (10 usuarios)
Infraestructura	\$25.000.000/año	Servidor/Cloud ETL y almacenamiento

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Consultoría inicial	\$30.000.000 (único)	Configuración modelo
Capacitación	\$8.000.000 (20 horas)	Scrum & BI para equipo (20 horas)
Costos de operación	\$40.000.000/año	Tiempo parcial equipo interno (20%)
Total Año 1	≈ \$121.000.000	—

Nota. Adaptado de valores del mercado

Cronograma de implementación

El cronograma establece las fases, actividades y entregables del proyecto en un horizonte de 12 meses. Este apartado presenta la planificación temporal que permitirá avanzar de forma estructurada y validada en cada etapa de la implementación. El plan contempla cinco fases —inicio, planeación, generación del modelo preliminar, validación y mejora del modelo, y modelo refinado—, cada una con sus respectivas actividades definidas (ver anexo 1).

Beneficios esperados

La aplicación del modelo propuesto de analítica a través de la metodología de Scrum generará múltiples beneficios estratégicos, operativos y culturales para Sicmafarma S.A.S., especialmente en la gestión de productos farmacéuticos de objeto de análisis. A continuación, se detallan los beneficios más relevantes:

- **Reducción de incertidumbre en la toma de decisiones comerciales.**

La capacidad de procesar y analizar datos históricos y variables externas permite anticiparse a la evolución del mercado, reduciendo la dependencia de decisiones basadas en intuición o experiencia individual. Con modelos predictivos alimentados por datos

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

epidemiológicos, climáticos y de precios, la organización podrá formular estrategias comerciales con mayor sustento técnico y menor margen de error.

- **Mayor alineación entre ventas, inventarios y comportamiento de la demanda.**

El modelo BI permite visualizar el comportamiento de la demanda, lo que posibilita una mejor sincronización entre los niveles de inventario y las proyecciones de consumo. Esto evita tanto el desabastecimiento como la sobreacumulación de productos, mejorando la eficiencia operativa, reduciendo costos logísticos y fortaleciendo el servicio al cliente.

- **Mejora de la capacidad de respuesta ante eventos epidemiológicos y climático**

Al incorporar variables externas como el fenómeno ENSO (El Niño/La Niña) y la incidencia de enfermedades respiratorias, el modelo facilita la adopción de medidas proactivas frente a escenarios de riesgo. Esto permite, por ejemplo, ajustar campañas de distribución, abastecer zonas vulnerables con antelación o rediseñar las estrategias de comercialización frente a alertas sanitarias.

- **Fortalecimiento de la cultura de mejora continua y gestión basada en datos.**

El uso de Scrum como marco ágil impulsa ciclos iterativos de planificación, acción y retroalimentación que promueven una cultura de aprendizaje constante y adaptación rápida. Cada sprint se convierte en una oportunidad para evaluar resultados y ajustar la estrategia en función de los datos analizados.

- **Optimización de la gestión del conocimiento organizacional**

El modelo BI sistematiza la información relevante y permite que esta sea accesible de manera estructurada por los diferentes niveles de la organización. Esto mejora la transferencia de conocimiento, facilita el entrenamiento de nuevos colaboradores y promueve una visión compartida de los objetivos comerciales basados en evidencia.

- **Mejora de la eficiencia en campañas y decisiones de precio**

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Con el soporte del análisis de elasticidad precio-demanda, la empresa podrá implementar estrategias de precio más precisas y diferenciadas por segmento, momento del año o región geográfica. Esto incrementa la rentabilidad del portafolio y permite tomar decisiones más informadas sobre promociones, descuentos o ajustes de precio frente a la competencia.

- **Mayor capacidad para escalar el modelo a otras líneas de producto**

El modelo propuesto, aunque centrado inicialmente en Ampicilina + Sulbactam, está diseñado para ser replicable a otras moléculas o productos del portafolio de Sicmafarma S.A.S. Esto representa una oportunidad para estandarizar buenas prácticas de inteligencia de negocios y gestión ágil en otras unidades de negocio, fortaleciendo la sostenibilidad de la empresa a mediano y largo plazo.

Limitaciones del modelo

La implementación del modelo de analítica enfrenta algunos desafíos importantes, entre ellos la dependencia del acceso continuo a datos confiables y actualizados, condición indispensable para garantizar la precisión de los análisis. Asimismo, se debe considerar la curva de aprendizaje en el uso de herramientas BI por parte del equipo, la cual requiere procesos de capacitación y acompañamiento adecuados. Finalmente, resulta fundamental contar con un compromiso organizacional sólido para adoptar y mantener prácticas ágiles de manera sostenible, asegurando que la metodología Scrum se convierta en parte de la cultura de trabajo y no en un esfuerzo aislado.

Sostenibilidad del modelo

El modelo puede mantenerse en el tiempo siempre que se formalice como parte integral del proceso comercial, lo que implica establecer prácticas que aseguren su sostenibilidad. Para ello, resulta indispensable la capacitación continua del personal en el uso de herramientas BI, así como la designación de responsables que lideren el ciclo Scrum y el mantenimiento del modelo analítico. De igual manera, es necesario garantizar

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

la integración con los sistemas ERP o CRM ya existentes en la empresa, de modo que la información fluya de manera eficiente y coherente dentro de la organización.

Recomendaciones Estratégicas para Sicmafarma S.A.S.

Basado en las conclusiones anteriores, se emiten las siguientes recomendaciones:

- **Implementar un Sistema de Inteligencia Competitiva:** Dada la extrema sensibilidad al precio de VITALIS, es imperativo que Sicmafarma S.A.S desarrolle un monitoreo activo y en tiempo real de los precios de sus competidores clave. Esta información debe ser el insumo principal para ajustar su propia estrategia de precios de forma dinámica.
- **Integrar el Modelo Log-Log en la Planificación de la Demanda (S&OP):** Se recomienda adoptar operativamente el modelo log-log para generar pronósticos de ventas mensuales. Esto permitirá optimizar los niveles de inventario, evitando quiebres de stock durante los picos estacionales (marzo, septiembre) o en periodos proyectados de La Niña, y reduciendo costos de almacenamiento en meses de baja demanda.
- **Utilizar Pronósticos Climáticos y Epidemiológicos:** Sicmafarma S.A.S debe incorporar proactivamente los boletines del IDEAM sobre el fenómeno ENSO y los informes del Instituto Nacional de Salud en las variables de entrada del modelo. Anticipar un evento de La Niña, por ejemplo, debería desencadenar un aumento planificado de la producción y el stock.
- **Realizar una Recalibración Periódica del Modelo:** El mercado farmacéutico es dinámico. Se recomienda reentrenar y validar el modelo de forma anual, incorporando nuevos datos para asegurar que sus coeficientes y su precisión predictiva se mantengan vigentes a lo largo del tiempo.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Análisis de los resultados

El análisis muestra que la demanda mensual de Ampicilina + Sulbactam 1,5 g es altamente variable y asimétrica: la mediana ronda las 21.350 unidades, mientras la media se acerca a 60.525 con desviación estándar elevada y máximos cercanos a 249.390. Se identifican meses críticos de alta demanda (marzo–mayo y agosto–septiembre) y periodos bajos (enero y junio–julio), por lo que el mes debe tratarse como factor explicativo. El clima incide de manera diferenciada: durante La Niña los niveles centrales son mucho más altos que en fases Neutral o El Niño, en línea con mayores cuadros respiratorios. También se observan patrones competitivos y epidemiológicos consistentes: las ventas de Sicmafarma aumentan cuando suben los precios de VITALIS y FARMALÓGICA y disminuyen cuando lo hace PFIZER, y guardan asociación con indicadores como tuberculosis pulmonar, H. influenzae e IRAG. En suma, la demanda no es aleatoria; responde de forma sistemática a estacionalidad, clima, salud pública y precios.

Entre las fortalezas del estudio destaca la integración de fuentes internas y externas (ventas, precios, ENSO y epidemiología) y la identificación de patrones que orientan la gestión comercial y de abastecimiento. La comparación de enfoques econométricos y de aprendizaje automático ubica a la regresión log–log como el mejor equilibrio entre ajuste e interpretabilidad ($R^2 \approx 0,85$), permitiendo estimar elasticidades y simular escenarios “¿qué pasa si...?” útiles para la decisión táctica. Este modelo reproduce con mayor fidelidad tendencias y ciclos medios que la línea base OLS; aunque Random Forest reduce algunos errores puntuales, su menor interpretabilidad lo hace menos idóneo para la toma de decisiones. A nivel organizacional, la empresa cuenta con áreas capaces de operar tableros, alertas y rutinas de mejora continua, lo que facilita llevar estos hallazgos a prácticas de Business Intelligence (analítica) y gestionarlos con Scrum.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Las oportunidades de mejora se concentran en tres frentes. En datos, conviene fortalecer la gobernanza y calidad (tratamiento de ceros y atípicos, metadatos y trazabilidad) y automatizar la ingesta desde INS, IDEAM y CPC-NOAA mediante un pipeline ETL/ELT reproducible que alimente modelo y *dashboards*. En variables, se recomienda mayor granularidad por canal, región, licitaciones y quiebres de stock de terceros, así como calendarios de compra de EPS, para mejorar la atribución y la anticipación de picos. En analítica y gestión, se sugiere reentrenar el modelo con validación temporal (ventanas móviles o crecientes), complementar el log-log con métodos no lineales para capturar efectos locales, desplegar tableros operativos con pronóstico vs. real, riesgo estacional y señales ENSO/IRAG, y activar un monitoreo casi en tiempo real de precios de competidores, dado el peso de la elasticidad cruzada observada con VITALIS.

Como siguiente paso, Scrum es el marco idóneo para llevar el modelo a producción y replicarlo en otros productos del portafolio. Con un Product Backlog (tableros, reglas de inventario, alertas y pricing), sprints de 2–4 semanas y reviews periódicas, la organización puede iterar, medir valor y estandarizar lo que funcione. Los artefactos (Product/Sprint Backlog e Incremento) y eventos (Planning, Daily, Review, Retrospective), junto con los roles (Product Owner, Scrum Master y equipo multidisciplinario), brindan trazabilidad y foco de negocio. Se recomienda comenzar con Ampicilina + Sulbactam y, una vez estable, reutilizar la misma ruta en otras moléculas de alta rotación, creando una “línea de montaje” ágil de proyectos BI con métricas comparables de adopción y resultado.

Conclusiones y Recomendaciones

Conclusiones

A continuación, se presentan las conclusiones obtenidas a partir del desarrollo de cada uno de los objetivos formulados.

Conclusión 1: La demanda de Ampicilina + Sulbactam es altamente variable y asimétrica, con meses de “valles” frecuentes y picos concentrados que elevan el riesgo de quiebres/sobre-inventario. Esta dinámica se explica en parte por factores externos: estacionalidad (meses pico), ENSO (La Niña), señales epidemiológicas (IRAG y otros) y precios de competidores (sustitución marcada, especialmente frente a VITALIS). El diagnóstico deja una “foto operativa” clara de los determinantes y brechas (calidad/integración de datos), útil para orientar el modelado y la gestión comercial.

Conclusión 2: El modelo log–log seleccionado es el referente por su alto poder explicativo y capacidad de interpretación mediante elasticidades. Permite cuantificar impactos de precio propio y competitivo, de ENSO y de indicadores sanitarios, y capta la estacionalidad relevante, habilitando escenarios “¿qué pasa si?”. Frente a alternativas (OLS, Random Forest), ofrece el mejor equilibrio entre precisión e interpretabilidad, lo que mejora la calidad y oportunidad de las decisiones comerciales.

Conclusión 3: La propuesta Scrum (roles, artefactos, eventos, DoD y métricas) convierte el análisis en valor continuo mediante sprints cortos: tablero único con KPIs, alertas tempranas, revisión y retrospectiva para ajustar tácticas de precio, inventario y cobertura. La guía es replicable a otras referencias del portafolio y reduce el time-to-learning, institucionalizando la toma de decisiones basada en datos.

Conclusión 4: La integración de un modelo predictivo interpretable con una ejecución ágil ofrece a Sicmafarma S.AS un mecanismo probado para elevar la precisión de la planeación comercial y disminuir la incertidumbre operativa. Con gobierno de datos, reentrenamiento periódico y uso disciplinado del marco Scrum, la organización queda con

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

una ruta sostenible de escalamiento a otros productos y con capacidades internas para decidir mejor, más rápido y con evidencia.

Recomendaciones

Se presentan las recomendaciones derivadas del análisis realizado

Recomendación 1: se recomienda que la empresa ponga a funcionar el modelo en un plazo corto (p. ej., 90 días). Para ello, debería consolidar un tablero único con ventas reales vs. pronóstico y un semáforo de error; activar alertas por clima (ENSO), epidemiología (IRAG) y cambios de precios de competidores; y fijar un calendario de actualización del modelo (trimestral). También conviene dejar claros los roles: Gerencia Comercial define prioridades, el equipo de BI mantiene el modelo y Planeación ajusta inventarios. El desempeño puede seguirse con tres métricas simples: error promedio del pronóstico, quiebres de stock y tiempo de reacción ante alertas.

Recomendación 2: se sugiere ordenar la base de datos para que el modelo opere con calidad. La empresa debería crear un diccionario de datos (definición y fuente de cada campo), aplicar reglas básicas de calidad (evitar ceros imposibles, revisar fechas y outliers) y automatizar la carga mensual desde INS/IDEAM/competencia y sistemas internos. Es útil asignar un responsable por fuente y monitorear tres señales: porcentaje de registros válidos, fecha de última actualización y número de incidencias corregidas por mes.

Recomendación 3: se propone usar Scrum para mejorar y replicar el enfoque en otros productos. La organización puede trabajar en sprints de 2–4 semanas con una Definition of Done sencilla (modelo actualizado, tablero publicado y acuerdos de acción), iniciar con un backlog corto (pricing por competidor, políticas de stock para meses pico y segmentación de clientes) y pilotear en 1–2 moléculas de alta rotación. Al cierre de cada sprint, se recomienda realizar una demo al equipo comercial y una retrospectiva con una mejora concreta para el siguiente ciclo.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

Anexos

Anexo 1

Cronograma de implementación

Fase / Actividades	Entregables	Duración del proyecto (M= Mes)												
		M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	
Inicio														
Realizar capacitación metodología Scrum	Equipo de trabajo capacitado													
Planeación														
Asignar roles	Roles asignados													
Definir backlog inicial	Backlog definido Sprint planning													
Realizar integración de fuentes internas y externas	Unificación de las fuentes internas (ventas, inventarios y precios) y fuentes externas (clima, epidemiología y precios de la competencia)													
Realizar limpieza de datos	Depuración de datos													
Generación del modelo preliminar														
Aplicación del modelo log-log versión preliminar	Tablero BI (Versión preliminar)													
Validación y mejora del modelo														
Realizar sesiones de trabajo (Dailys, sprint review y sprint retrospective) Demostración del tablero	Avances de backlog Retroalimentación documentada Tablero BI ajustado Finalización del backlog													
Modelo refinado														
Aplicación del modelo log-log	Tablero BI refinado													
Replicación del modelo BI (log-log) a dos moléculas	Modelo aplicado a dos moléculas adicionales													

Referencias

- A.N.D.I. (2022, junio 14). *Radiografía del mercado farmacéutico colombiano*.
<https://www.andi.com.co/Home/Noticia/17274-radiografia-del-mercado-farmaceutico-co>
- A.S.C.I.F. (2023). *Informe sectorial de la industria farmacéutica en Colombia*. ASCIF.
- A.S.C.I.F. (2025). *Sitio oficial de la Asociación Colombiana de la Industria Farmacéutica*.
<https://ascif.org.co>
- A.S.C.I.F. (2025, mayo 15). *Colombia tiene la capacidad para convertirse en potencia exportadora de medicamentos de alta calidad*. <https://ascif.co/colombia-tiene-la-capacidad-para-convertirse-en-potencia-exportadora-de-medicamentos-de-alta-calidad/>
- Berryhill, A. (2023). *Agile in Pharmaceutical Industry: Accelerating R&D. Influential Agile Leader*. <https://www.influentialagileleader.com/accelerating-rd-with-agile-in-pharmaceutical-industry>
- Binariks. (2025, febrero 9). *5 Business Intelligence Use Cases in the Pharmaceutical Industry*.
<https://binariks.com/blog/business-intelligence-pharma-use-cases>
- Cano, J. L. (2007). *Business Intelligence: Competir con información*. Banesto, Fundación Cultural.
- Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? *Geoscientific Model Development*, 7, 1247-1250.
- C.N.P.M.D.M. (2022). *Circular 013 de 2022: Metodología para la regulación de precios de medicamentos (Precio Máximo de Venta, PMV)*. Ministerio de Salud y Protección Social.
<https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/DE/PS/circular-013-2022-cnpm.pdf>
- Grant, R. (2019). *Contemporary Strategy Analysis* (10.a). Wiley.
- Greene, W. H. (2018). *Econometric Analysis*. Pearson.
- Hernández-Sampieri, R., & Mendoza Torres, C. P. (2018). *Metodología de la investigación: Las*

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento
Ampicilina + Sulbactam en Sicmafarma SAS

rutas cuantitativa, cualitativa y mixta. McGraw-Hill Education.

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3rd ed.).
OTexts). <https://otexts.com/fpp3/>

I.D.E.A.M. (2025). *Temperatura diaria por ciudad en Colombia, 2019–2023. Datos abiertos*
IDEAM. <https://www.ideam.gov.co/transparencia/datos-abiertos>

I.N.S. (2025). *Portal institucional del INS (series epidemiológicas y SIVIGILA*.
<https://www.ins.gov.co/>

I.N.V.I.M.A. (2016). *Guía BPM Medicamentos (Resolución 1160 de 2016*.
<https://www.invima.gov.co/sites/default/files/medicamentos-y-productos-biologicos/tecnico-medicamentos/listas-y-guias/BPX/Gu%C3%ADa%20BPM%20Medicamentos.pdf>

I.N.V.I.M.A. (2023). *Informe de gestión 2022–2023*. INVIMA.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning: With Applications in R* (2nd ed.). Springer.

Kimball, R., & Ross, M. (2011). *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling* (3rd ed.). Wiley.

Kuhn, M., & Johnson, K. (2019). *Feature Engineering and Selection: A Practical Approach for Predictive Models*. CRC Press.

Mendoza Erazo, M. B. (2024). Modelo estructural de inteligencia de negocios para inventarios en farmacéutica. *Multidisciplinar y Latin American Journal*, 1-15. <https://mlaj-revista.org/index.php/journal/article/view/9>

Microsoft. (2025). *Power BI: Plan de precios*. <https://www.microsoft.com/es-es/power-platform/products/power-bi/pricing>

N.O.A.A. Climate Prediction. (2025). *Portal CPC*. <https://www.cpc.ncep.noaa.gov/>

Olarte, A., & Moure, S. (2021). Régimen de control de precios de medicamentos en Colombia. *Boletín Jurídico Farmacéutico*, 5-12.

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento Ampicilina + Sulbactam en Sicmafarma SAS

- País, E. (2025, febrero 10). *Laboratorios, distribuidores, Gobierno y EPS: así falla la cadena de medicamentos en Colombia*. <https://elpais.com/america-colombia/2025-03-29/laboratorios-distribuidores-gobierno-y-eps-asi-falla-la-cadena-de-medicamentos-en-colombia.html>
- Patara, G. S., & Raharjo, T. (2023). Comparing Scrum maturity of digital and business process reengineering groups: A case study at an Indonesia's state-owned bank. *International Journal of Advanced Computer Science and Applications*, 14(8), 126-134.
- Petersen, K., Wohlin, C., & Baca, D. (2021). The use of agile in regulated environments: A multiple-case study. *Empirical Software Engineering*, 32. <https://doi.org/10.1007/s10664-020-09896-2>
- Rigby, D., Sutherland, J., & Noble, A. (2016). Embracing Agile. *Harvard Business Review*, 94(5), 40-50.
- Rigby, D., Sutherland, J., & Noble, A. (2018). Agile at Scale. *Harvard Business Review*. <https://hbr.org/2018/05/agile-at-scale>
- Sagala, P., Wasesa, M., & Sunitiyoso, Y. (2024). The data divide in pharma: A comparative case study of business analytics capabilities impact on performance. *IEEE Access*, 1-23.
- Sanchez Obando, J. W., Duque Méndez, N. D., & Bedoya Herrera, O. M. (2024). Construction of the Invoicing Process through Process Mining and Business Intelligence in the Colombian Pharmaceutical Sector. *Computers*, 13, 245.
- Saunders, M., Lewis, P., & Thornhill, A. (2019). *Research Methods for Business Students* (8th ed.). Pearson Education Limited.
- Scalable, A. I. (2024, septiembre 2). *Leveraging Mobile BI for agile decision-making in pharmaceutical sales*. <https://insights.scalableai.com/index.php/leveraging-mobile-bi-for-agile-decision-making-in-pharmaceutical-sales/>
- Schwaber, K., & Sutherland, J. (2020). *The Scrum Guide: The Definitive Guide to Scrum—The Rules of the Game*. <https://scrumguides.org/>

Modelo de analítica y visualización de datos para la predicción de la demanda del medicamento
Ampicilina + Sulbactam en Sicmafarma SAS

Sicmafarma. (2025). *Compañía*. <https://sicmafarma.com.co/compania/>

Tiempo, E. (2023, marzo 12). *¿Colombia debería apostar por la producción de sus propios medicamentos?* <https://www.eltiempo.com/salud/colombia-deberia-apostar-por-la-produccion-de-sus-propios-medicamentos-sector-farmaceutico-explica-si-realmente-es-viable-3360974>

Turban, E., Sharda, R., & Delen, D. (2020). *Analytics, Data Science, & Artificial Intelligence: Systems for Decision Support* (11th ed.). Pearson.

Wooldridge, J. M. (2016). *Introductory Econometrics: A Modern Approach* (6th ed.). Cengage Learning.