Trabajo previo a la obtención de título de Magister en

Ciencia de datos y Máquinas de Aprendizaje con Mención en Inteligencia Artificial

AUTORES:

Kevin Fernando Caiza Lema

Jaime Manuel Castro Arévalo

Patricio Daniel Galárraga Sosa

Marco Vinicio Garzón Chamorro

TUTORES:

Andrés Roberto Navas Castellanos

Gabriela Estefania Chiliquinga Jiménez

Predicción de congestión hospitalaria mediante análisis de admisiones históricas

Quito - Ecuador

Julio – 2026

REINVENTEMOS EL FUTURO

Certificación de autoría

Nosotros, **Kevin Fernando Caiza Lema, Jaime Manuel Castro Arévalo, Patricio Daniel Galárraga Sosa y Marco Vinicio Garzón Chamorro,** declaramos bajo juramento que el trabajo aquí descrito es de nuestra autoría; que no ha sido presentado anteriormente para ningún grado o calificación profesional y que se ha consultado la bibliografía detallada. Cedemos nuestros derechos de propiedad intelectual a la Universidad Internacional del Ecuador (UIDE), para que sea publicado y divulgado en internet, según lo establecido en la Ley de Propiedad Intelectual, su reglamento y demás disposiciones legales.

Validar únicamente en FirmaEC.
Firmado electrónicamente por:
KEVIN FERNANDO CAIZA LEMA

Firma

Kevin Fernando Caiza Lema

Validar únicamente en FirmaEC.
Firmado electrónicamente por:
JAIME MANUEL CASTRO AREVALO

Firma

Jaime Manuel Castro Arévalo

Validar únicamente en FirmaEC.
Firmado electrónicamente por:
PATRICIO DANIEL GALARRAGA SOSA

Firma

Patricio Daniel Galárraga Sosa

Validar únicamente en FirmaEC.
Firmado electrónicamente por:
MARCO VINICIO GARZON CHAMORRO

Firma

Marco Vinicio Garzón Chamorro

Autorización de Derechos de Propiedad Intelectual

Nosotros, **Kevin Fernando Caiza Lema, Jaime Manuel Castro Arévalo, Patricio Daniel Galárraga Sosa y Marco Vinicio Garzón Chamorro**, en calidad de autores del trabajo de investigación titulado ***Predicción de congestión hospitalaria mediante análisis de admisiones históricas***, autorizamos a la Universidad Internacional del Ecuador (UIDE) para hacer uso de todos los contenidos que nos pertenecen o de parte de los que contiene esta obra, con fines estrictamente académicos o de investigación. Los derechos que como autores nos corresponden, lo establecido en los artículos 5, 6, 8, 19 y demás pertinentes de la Ley de Propiedad Intelectual y su Reglamento en Ecuador.

D. M. Quito, (junio de 2026)

Firma

Kevin Fernando Caiza Lema

Firma

Jaime Manuel Castro Arévalo

Firma

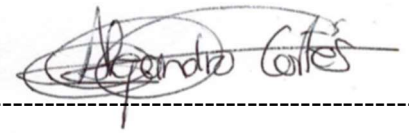
Patricio Daniel Galárraga Sosa

Firma

Marco Vinicio Garzón Chamorro

Aprobación de dirección y coordinación del programa

Nosotros, **Alejandro Cortés López Director EIG y Karla Estefanía Mora Cajas Coordinadora UIDE,** declaramos que: **Kevin Fernando Caiza Lema, Jaime Manuel Castro Arévalo, Patricio Daniel Galárraga Sosa y Marco Vinicio Garzón Chamorro** son los autores exclusivos de la presente investigación y que ésta es original, auténtica y personal de ellos.



Alejandro Cortés López

Director de la
Maestría en Ciencia de Datos y Máquinas de Aprendizaje con Mención en Inteligencia Artificial

Karla Estefanía Mora Cajas

Coordinadora de la
Maestría en Ciencia de Datos y Máquinas de Aprendizaje con Mención en Inteligencia Artificial

DEDICATORIA

Dedicamos este trabajo de titulación, en primer lugar, a Dios, por guiarnos y brindarnos la fortaleza, sabiduría y perseverancia necesarias para culminar esta importante etapa académica y profesional.

A nuestras familias, por su amor incondicional, comprensión y apoyo constante durante todo este proceso. Gracias por acompañarnos en los momentos de esfuerzo, por motivarnos a seguir adelante y por confiar en nosotros aun cuando el camino presentó dificultades. Su respaldo ha sido fundamental para alcanzar esta meta.

A nuestros seres queridos, quienes con sus palabras de aliento, paciencia y cariño contribuyeron a que este logro fuera posible. Cada muestra de apoyo representó una motivación para continuar con responsabilidad, compromiso y dedicación.

Finalmente, dedicamos esta tesis a todas las personas que, de una u otra manera, formaron parte de este proceso de aprendizaje y crecimiento. Este logro representa no solo el resultado de nuestro esfuerzo académico, sino también el reflejo de la constancia, la disciplina y el acompañamiento recibido a lo largo de este camino.

AGRADECIMIENTOS

Expresamos nuestro agradecimiento a la Universidad Internacional del Ecuador por brindarnos el espacio académico, los recursos y las condiciones necesarias para desarrollar este trabajo de titulación como parte de nuestro proceso de formación profesional.

Extendemos también nuestro reconocimiento a los docentes que formaron parte de nuestra preparación a lo largo de la maestría, cuyos conocimientos y experiencias contribuyeron al desarrollo de las competencias necesarias para abordar este proyecto con rigor, responsabilidad y sentido crítico.

De igual manera, agradecemos a nuestros compañeros de grupo por el compromiso, la colaboración y la responsabilidad demostrada durante la elaboración de esta tesis. El trabajo conjunto, la comunicación constante y la distribución de responsabilidades fueron elementos fundamentales para cumplir los objetivos planteados.

Finalmente, agradecemos a todas las instituciones y fuentes de información que hicieron posible el acceso a los datos utilizados en esta investigación, así como a quienes aportaron con criterios, revisiones o sugerencias que enriquecieron el resultado final de este trabajo.

RESUMEN

Esta investigación desarrolló y evaluó un enfoque predictivo para estimar, con siete días de anticipación, la relación entre demanda hospitalaria y capacidad disponible en grupos de establecimientos de salud del Ecuador. Se utilizaron datos del período 2022-2024 sobre egresos hospitalarios, disponibilidad de camas y causas de egreso, partiendo de 3,434,083 registros con cobertura nacional de 633 establecimientos.

El estudio consolidó un proceso reproducible de integración, depuración, ingeniería de características, modelado, evaluación y priorización operativa. Se evaluaron escenarios con y sin outliers mediante modelos de Machine Learning, principalmente Random Forest y XGBoost, y se incorporó un experimento complementario de Deep Learning con una red neuronal densa ligera (MLP_light). En el benchmark principal, el mejor resultado puntual de Machine Learning se obtuvo con XGBoost en el escenario con outliers y conjunto completo de variables (WMAPE 0.369118; MAE 0.060626; RMSE 0.088753). Por su parte, el experimento MLP_light alcanzó el mejor WMAPE puntual del estudio (0.362742), evidenciando un desempeño competitivo bajo configuraciones específicas.

En la aplicación operativa, las predicciones fueron transformadas en niveles de prioridad para apoyar la planificación hospitalaria. El enfoque permitió identificar grupos hospitalarios con prioridad alta, media y baja, así como provincias con mayor recurrencia de alertas, entre ellas Orellana y Sucumbíos. Además, el análisis por causas y lift de alerta operativa mostró que las causas de egreso más frecuentes no siempre fueron las más asociadas a mayor prioridad, lo que aporta una lectura más precisa para identificar focos de atención. En conjunto, los resultados muestran que el modelo puede apoyar el monitoreo focalizado, la alerta temprana y la planificación preventiva de camas, personal y recursos hospitalarios.

Palabras clave:

Predicción hospitalaria; Demanda hospitalaria; capacidad hospitalaria; presión asistencial; Machine Learning (XGBoost; Random Forest); Deep Learning (MLP); lift de alerta operativa; egresos hospitalarios.

ABSTRACT

This research developed and evaluated a predictive approach to estimate, seven days in advance, the relationship between hospital demand and available capacity in groups of healthcare facilities in Ecuador. Data from the 2022-2024 period were used, including hospital discharges, bed availability, and causes of discharge, starting from 3,434,083 records with national coverage across 633 healthcare facilities.

The study consolidated a reproducible process for data integration, cleaning, feature engineering, modeling, evaluation, and operational prioritization. Scenarios with and without outliers were evaluated using Machine Learning models, mainly Random Forest and XGBoost, and a complementary Deep Learning experiment was incorporated through a lightweight dense neural network (MLP_light). In the main benchmark, the best Machine Learning result was obtained with XGBoost in the scenario with outliers and the complete set of variables (WMAPE 0.369118; MAE 0.060626; RMSE 0.088753). Meanwhile, the MLP_light experiment achieved the best WMAPE result of the study (0.362742), showing competitive performance under specific configurations.

In the operational application, the predictions were transformed into priority levels to support hospital planning. The approach made it possible to identify hospital groups with high, medium, and low priority, as well as provinces with a higher recurrence of alerts, including Orellana and Sucumbíos. In addition, the analysis by causes and operational alert lift showed that the most frequent causes of discharge were not always the ones most associated with higher priority, providing a more precise perspective for identifying areas of attention.

Overall, the results show that the model can support targeted monitoring, early warning, and preventive planning of hospital beds, staff, and resources.

Keywords:

Hospital prediction; hospital demand; hospital capacity; care pressure; Machine Learning (XGBoost; Random Forest); Deep Learning (MLP); operational alert lift; hospital discharges.

TABLA DE CONTENIDOS (Índice)

LISTA DE TABLAS (Índice de tablas)

LISTA DE FIGURAS (Índice de figuras)

CAPÍTULO 1 . INTRODUCCIÓN

1.1. Contexto del problema

La planificación hospitalaria requiere anticipar, con la mayor precisión posible, la demanda de atención para asignar recursos de forma oportuna y evitar presiones operativas sobre el sistema. Esta demanda puede expresarse mediante indicadores como consultas médicas, admisiones, atenciones de emergencia, egresos y utilización de camas. En este sentido, Pan y Yan (2026) evidencian que el análisis histórico de consultas y admisiones permite identificar tendencias en el uso de servicios médicos y proyectar necesidades futuras mediante modelos de series temporales. Entre los recursos críticos, la disponibilidad de camas constituye un componente esencial de la capacidad operativa, ya que condiciona la admisión, permanencia, traslado y egreso de pacientes. Cuando la demanda esperada no se estima adecuadamente, pueden generarse escenarios de saturación, mayores tiempos de espera, presión sobre el personal sanitario y uso ineficiente de la infraestructura hospitalaria.

En Ecuador, la información pública sobre egresos hospitalarios y camas disponibles representa una fuente valiosa para analizar la relación entre demanda observada y capacidad instalada. Sin embargo, esta información suele encontrarse distribuida en distintas estructuras, con diferencias de granularidad, cobertura y calidad, lo que dificulta su uso con fines predictivos. Estudios recientes señalan que la integración de datos históricos, variables temporales y modelos predictivos permite fortalecer la planificación de recursos hospitalarios, especialmente cuando la demanda presenta patrones no lineales, estacionales o asociados a períodos de presión asistencial (Blanco et al., 2026; Zhao, 2026). En este contexto, surge la necesidad de construir un enfoque analítico que integre dichas fuentes y genere estimaciones útiles para apoyar la planificación hospitalaria.

1.2. Situación actual de la planificación hospitalaria

En la práctica, la planificación hospitalaria se apoya con frecuencia en reportes históricos, indicadores descriptivos y criterios administrativos. Aunque estos mecanismos permiten monitorear el comportamiento pasado del sistema, presentan limitaciones para anticipar variaciones futuras de la demanda, especialmente cuando existen patrones temporales, diferencias territoriales, cambios epidemiológicos o fluctuaciones en la presión asistencial. La literatura reciente muestra que los enfoques predictivos basados en series temporales, Machine Learning y Deep Learning pueden aportar mayor capacidad de anticipación frente a métodos puramente descriptivos, al capturar relaciones complejas y comportamientos no lineales en los datos hospitalarios (Blanco et al., 2026; Zhao, 2026).

Adicionalmente, los datos de egresos y camas suelen analizarse de manera separada, lo que reduce la posibilidad de evaluar integralmente la relación entre demanda observada y capacidad disponible. Zhao (2026) destaca que la predicción de la demanda debe complementarse con el análisis de la capacidad, debido a que el incremento de pacientes puede activar umbrales críticos de uso de camas y afectar la respuesta operativa del sistema. De forma semejante, Sasikala et al. (2026) plantean que la predicción de cargas de trabajo hospitalarias puede apoyar decisiones de asignación de camas y personal, reducir tiempos de espera y mejorar el uso de recursos. Bajo este escenario, el uso de modelos predictivos sobre datos integrados puede aportar una alternativa metodológica para estimar la demanda hospitalaria y fortalecer la planificación operativa.

1.3. Formulación del problema

El problema central de esta investigación radica en la necesidad de estimar la demanda hospitalaria futura a partir de información histórica de egresos y camas, en un contexto donde las fuentes disponibles presentan diferencias de estructura y requieren procesos de integración, depuración y transformación para su uso analítico. La evidencia reciente muestra

que la demanda hospitalaria puede modelarse a partir de series históricas de admisiones, consultas o atenciones de emergencia, utilizando enfoques estadísticos clásicos y modelos de inteligencia artificial como Random Forest, XGBoost, LSTM u otras redes neuronales (Blanco et al., 2026; Pan & Yan, 2026).

En términos prácticos, la ausencia de un esquema analítico unificado limita la generación de predicciones consistentes que permitan comparar la demanda esperada con la capacidad disponible. Esto favorece una toma de decisiones más reactiva que preventiva. Por ello, se plantea desarrollar un enfoque de modelado predictivo basado principalmente en Machine Learning e incorporar, como contraste adicional, una alternativa básica de Deep Learning, utilizando datos históricos integrados del período 2022–2024 para apoyar la planificación de recursos hospitalarios en Ecuador.

1.4. Pregunta de investigación

¿En qué medida un enfoque de modelado predictivo basado en Machine Learning y contrastado de forma complementaria con Deep Learning básico, aplicado a datos históricos integrados de egresos hospitalarios y disponibilidad de camas del Ecuador durante el período 2022–2024, permite estimar la demanda hospitalaria y apoyar la planificación operativa de recursos?

1.5. Objetivo general

Desarrollar un enfoque de modelado predictivo para estimar la demanda hospitalaria a partir de datos históricos integrados de egresos y disponibilidad de camas en Ecuador, utilizando Machine Learning como estrategia principal e incorporando un modelo básico de Deep Learning como contraste complementario, con el fin de apoyar la planificación operativa de recursos en los establecimientos de salud.

1.6. Objetivos específicos

1. Integrar las fuentes de egresos hospitalarios y camas disponibles correspondientes al período 2022–2024, definiendo una unidad analítica consistente para el estudio.
2. Evaluar la calidad y cobertura de los datos para identificar criterios de depuración y transformación que aseguren su uso en el modelado predictivo.
3. Construir un dataset de modelado con variables relevantes de demanda, capacidad y temporalidad, evitando fuga de información.
4. Entrenar y comparar modelos de Machine Learning para estimar la demanda hospitalaria en el horizonte definido, e incorporar un modelo de Deep Learning como comparación complementaria.
5. Evaluar el desempeño de los modelos mediante métricas de regresión como MAE, RMSE y WMAPE, considerando que estas métricas son utilizadas frecuentemente en estudios de predicción de demanda hospitalaria y servicios de emergencia (Blanco et al., 2026; Wang et al., 2026).
6. Analizar la relación entre demanda estimada y capacidad disponible para identificar escenarios de mayor presión operativa.

1.7. Justificación

Esta investigación se justifica por su relevancia práctica, metodológica y académica. En el plano práctico, aporta una herramienta analítica para anticipar comportamientos de demanda y mejorar la planificación de camas y recursos hospitalarios. La literatura reciente evidencia que la predicción de admisiones, consultas y atenciones de emergencia puede apoyar la asignación eficiente de recursos, la preparación ante períodos de alta demanda y la reducción de ineficiencias operativas (Pan & Yan, 2026; Zhao, 2026). Además, los enfoques que combinan predicción y análisis de capacidad pueden contribuir a identificar escenarios de saturación antes de que afecten la operación hospitalaria.

En el plano metodológico, el estudio propone una ruta reproducible para transformar información histórica dispersa en datos útiles para la toma de decisiones. De esta manera, permite avanzar desde una lógica descriptiva, centrada en explicar lo ocurrido, hacia una lógica prospectiva, enfocada en estimar lo que podría ocurrir y en cómo prepararse operativamente.

En el plano académico, la investigación contribuye con evidencia aplicada sobre el uso de modelos predictivos en problemas de planificación hospitalaria. Blanco et al. (2026) señalan que los modelos de Machine Learning y Deep Learning han mostrado potencial para la predicción de demanda en departamentos de emergencia, aunque persisten desafíos relacionados con validación, generalización e interpretabilidad. De la misma manera, Wang et al. (2026) muestran que incluso redes neuronales simples pueden aportar señales predictivas útiles en contextos de escasez de datos, lo cual resulta relevante para estudios con información limitada o heterogénea. Por ello, la comparación complementaria con Deep Learning permite contrastar si una alternativa de mayor complejidad aporta mejoras efectivas frente a modelos de Machine Learning en un problema de datos hospitalarios tabulares.

1.8. Alcance y limitaciones

El alcance de la investigación se centra en el análisis y modelado predictivo de datos históricos de egresos y camas hospitalarias del período 2022–2024 en Ecuador, con fines de apoyo a la planificación operativa. El estudio no busca reemplazar decisiones clínicas ni explicar causalidad médica; su propósito es generar estimaciones cuantitativas que fortalezcan la gestión de recursos. El componente de Deep Learning se considera complementario y de contraste metodológico, sin desplazar el enfoque principal basado en Machine Learning.

Entre las limitaciones se encuentran la calidad y completitud de los registros disponibles, la heterogeneidad entre establecimientos y territorios, y la dependencia de patrones históricos

que pueden cambiar ante eventos no previstos. Esta precaución es consistente con lo señalado por Blanco et al. (2026), quienes advierten que, aunque los modelos de inteligencia artificial presentan resultados prometedores en la predicción de demanda hospitalaria, aún existen retos vinculados con la validación externa, la generalización y la interpretabilidad. En consecuencia, los resultados de esta investigación deben interpretarse como apoyo técnico para la toma de decisiones y no como predicciones determinísticas.

CAPÍTULO 2 . REVISIÓN DE LITERATURA.

Este capítulo presenta los fundamentos conceptuales que sustentan la investigación y explican la lógica metodológica adoptada. El marco parte de la planificación hospitalaria y de la relación entre demanda y capacidad, continúa con la analítica predictiva aplicada a la salud y culmina con los enfoques utilizados en el estudio: modelos de series temporales, Machine Learning y un modelo básico de Deep Learning como contraste complementario.

2.1. Demanda hospitalaria

La demanda hospitalaria puede entenderse como el volumen de servicios asistenciales requeridos por una población en un periodo determinado. En el contexto hospitalario, esta demanda se expresa mediante consultas, admisiones, egresos, procedimientos, días de estancia, atenciones de emergencia y utilización de camas. No se limita al número de pacientes atendidos, sino que también refleja la presión que estos generan sobre los recursos disponibles para brindar atención oportuna, segura y eficiente. Por ello, constituye un elemento central para la planificación sanitaria, la gestión de camas, la programación del personal y la asignación de recursos clínicos y administrativos (Pan & Yan, 2026; Orhan & Kurutkan, 2025).

La literatura sobre utilización de servicios de salud señala que la demanda surge de la interacción entre características individuales, condiciones sociales, factores económicos, estado de salud y accesibilidad al sistema. El Modelo Conductual de Andersen clasifica estos determinantes en factores predisponentes, facilitadores y de necesidad. Los primeros incluyen variables como edad, sexo, educación y situación laboral; los facilitadores se relacionan con recursos que permiten acceder a la atención, como ingresos, seguro médico, transporte y acceso geográfico; y los factores de necesidad representan el estado clínico real o percibido de la persona, incluyendo enfermedades crónicas, comorbilidades, discapacidad o percepción del estado de salud (Andersen, 1995; Babitsch et al., 2012; Orhan & Kurutkan, 2025).

Desde una perspectiva hospitalaria, los factores demográficos y sociales son especialmente relevantes. El envejecimiento poblacional, el crecimiento de la población, la distribución territorial, la pobreza, el desempleo, el bajo nivel educativo o la falta de apoyo familiar pueden aumentar la utilización de servicios de salud y dificultar la prevención o el control temprano de enfermedades. Esto evidencia que la demanda hospitalaria no depende únicamente de variables clínicas, sino también de condiciones sociales, económicas y territoriales (Amirsmaili et al., 2026; Rodoreda-Pallàs et al., 2026).

La demanda hospitalaria puede medirse mediante indicadores operativos como admisiones, egresos, pacientes hospitalizados, estancia promedio, tiempo de espera, ocupación de camas y rotación de camas. Estos indicadores permiten cuantificar el flujo de pacientes y valorar si el hospital cuenta con capacidad suficiente para responder a las necesidades asistenciales. La literatura sobre indicadores hospitalarios destaca que las admisiones, egresos, estancia promedio y ocupación de camas son variables críticas para evaluar la utilización de recursos y apoyar la toma de decisiones gerenciales (Fallahnezhad et al., 2024; Hadian et al., 2024).

En esta investigación, la demanda se aproxima mediante los egresos hospitalarios agregados históricamente, ya que estos representan eventos asistenciales finalizados y permiten observar la carga real atendida por cada establecimiento. Al relacionar esta demanda con la capacidad disponible, se construye un indicador de presión hospitalaria que facilita la comparación entre establecimientos y periodos temporales. Esta elección es coherente con la literatura sobre desempeño hospitalario, que utiliza indicadores de pacientes atendidos, egresos, días de estancia y camas disponibles para analizar eficiencia, utilización y presión asistencial (Fallahnezhad et al., 2024; Khalifa et al., 2015).

Tabla 2-1

Indicadores para analizar eficiencia, utilización y presión asistencial

Dimensión	Ejemplos de variables	Relación con la demanda hospitalaria
Demográfica	Edad, sexo, envejecimiento, crecimiento poblacional	Modifica la probabilidad de utilizar servicios y requerir hospitalización.
Socioeconómica	Ingresos, empleo, educación, pobreza, apoyo social	Condiciona el acceso, la prevención y la oportunidad de atención.
Clínica	Enfermedades crónicas, estado de salud percibido, comorbilidades	Genera necesidad real de atención y mayor utilización hospitalaria.
Operativa	Admisiones, egresos, estancia, ocupación de camas	Permite medir el volumen atendido y la presión sobre recursos.

2.2. Capacidad hospitalaria

La capacidad hospitalaria se refiere al conjunto de recursos físicos, humanos, tecnológicos y organizacionales disponibles para responder a la demanda de atención de una población. Aunque tradicionalmente se asocia con el número de camas hospitalarias, una visión más completa incluye personal, equipos médicos, infraestructura, servicios diagnósticos, medicamentos, sistemas de información, procesos administrativos y coordinación interna. Por tanto, la capacidad hospitalaria no representa solo una cifra estática, sino una condición operativa que determina la posibilidad real de atender pacientes con seguridad y oportunidad (Fallahnezhad et al., 2024; Khalifa et al., 2015).

Las camas hospitalarias son uno de los indicadores más utilizados para representar la capacidad instalada. Sin embargo, es importante diferenciar entre camas dotadas, funcionales y disponibles. Las camas dotadas corresponden a la infraestructura registrada; las funcionales son aquellas que cuentan con condiciones mínimas para ser utilizadas; y las disponibles son las que se encuentran efectivamente habilitadas para recibir pacientes. Esta distinción es

relevante porque un hospital puede contar con infraestructura registrada, pero no disponer de personal, equipamiento o insumos suficientes para utilizarla plenamente (Griffiths et al., 2013; Fallahnezhad et al., 2024).

La gestión de la capacidad implica equilibrar la entrada, permanencia y salida de pacientes. Cuando la demanda supera la capacidad disponible, pueden generarse retrasos, listas de espera, mayor ocupación de camas, prolongación de estancias y dificultades para admitir nuevos pacientes. Por el contrario, una capacidad infrautilizada puede reflejar ineficiencia en la asignación de recursos. Por ello, la planificación hospitalaria requiere indicadores que permitan monitorear simultáneamente el volumen de pacientes, la disponibilidad de camas, la estancia promedio y la rotación de los recursos asistenciales (Khalifa et al., 2015; Hadian et al., 2024).

La capacidad también depende de la disponibilidad de recursos humanos y tecnológicos. Una cama no puede considerarse plenamente operativa si no existe personal médico, de enfermería, auxiliar y administrativo suficiente para garantizar la atención. De igual manera, la presencia de tecnologías diagnósticas, quirúrgicas y de monitoreo condiciona el nivel de complejidad que el hospital puede absorber. Por esta razón, la capacidad debe analizarse como un sistema integrado donde infraestructura, personal y procesos se articulan para sostener la atención de pacientes (Amirsmaili et al., 2026; Fallahnezhad et al., 2024).

Para esta investigación, la capacidad hospitalaria se representa principalmente mediante el total de camas disponibles. Esta variable permite normalizar el volumen de egresos y comparar la presión asistencial entre establecimientos de diferente tamaño. La relación entre egresos y camas disponibles ofrece una medida de la carga atendida por unidad de capacidad. Aunque esta aproximación no sustituye una evaluación completa de la capacidad hospitalaria, constituye un indicador útil para el modelado predictivo basado en datos históricos (Fallahnezhad et al., 2024; Khalifa et al., 2015).

Tabla 2-2

Indicadores de capacidad

Indicador de capacidad	**Descripción**	**Utilidad para la gestión**
Camas disponibles	Camas habilitadas para recibir pacientes	Permite estimar la capacidad operativa real.
Ocupación de camas	Proporción de camas utilizadas	Indica presión sobre infraestructura hospitalaria.
Estancia promedio	Tiempo medio de permanencia del paciente	Mide eficiencia del flujo y consumo de recursos.
Rotación de camas	Número de pacientes atendidos por cama en un periodo	Refleja productividad y capacidad de respuesta.

2.3. Presión hospitalaria y congestión asistencial

La presión hospitalaria expresa la relación entre la demanda de atención y la capacidad disponible para responder a esa demanda. En términos conceptuales, puede resumirse como: Demanda / Capacidad = Presión hospitalaria. Esta relación permite comprender que el problema no depende únicamente de cuántos pacientes llegan al hospital, sino de si la institución cuenta con recursos suficientes para atenderlos. Un mismo volumen de pacientes puede representar una presión baja en un hospital grande y una presión alta en un establecimiento con pocos recursos disponibles (Boyle et al., 2014; Fallahnezhad et al., 2024).

La congestión asistencial ocurre cuando la presión supera la capacidad de respuesta del sistema. En ese escenario, pueden presentarse tiempos de espera prolongados, retrasos en admisión, permanencia excesiva en emergencia, dificultades para acceder a camas, cancelación de procedimientos y afectación de la continuidad asistencial. La congestión no solo afecta la eficiencia operativa, sino también la seguridad del paciente, la satisfacción del usuario y la carga laboral del personal sanitario (Núñez et al., 2018; Wiler et al., 2015).

La presión hospitalaria está estrechamente relacionada con el flujo de pacientes. Este flujo depende de tres procesos principales: entrada, permanencia y salida. La entrada incluye admisiones, derivaciones y consultas de emergencia; la permanencia se refleja en días de estancia y ocupación de camas; y la salida se expresa en egresos, altas, transferencias o defunciones. Cuando alguno de estos procesos se retrasa, el sistema acumula pacientes y aumenta el riesgo de congestión (Núñez et al., 2018; Wiler et al., 2015).

En el presente proyecto, la presión hospitalaria se operacionaliza mediante el ratio de egresos por cama disponible. Este indicador relaciona la demanda atendida con la capacidad disponible y permite observar qué establecimientos presentan mayor carga relativa. Un valor alto del ratio sugiere que, para el número de camas disponibles, el hospital atiende un mayor volumen de egresos y puede estar sometido a mayor presión operativa. Esta variable es útil para el modelado predictivo porque resume, en un solo indicador, la interacción entre demanda y capacidad (Fallahnezhad et al., 2024; Khalifa et al., 2015).

Es importante diferenciar presión hospitalaria de congestión hospitalaria. La presión puede medirse de forma continua como una relación entre demanda y capacidad, mientras que la congestión representa una condición crítica cuando esa presión afecta el funcionamiento normal del hospital. En otras palabras, toda congestión implica presión elevada, pero no toda presión elevada necesariamente constituye congestión si el hospital logra responder mediante recursos, procesos y gestión adecuada (Blanco et al., 2026; Núñez et al., 2018).

Tabla 2-3

Conceptos de presión hospitalaria

Concepto	Definición operativa	Ejemplo de indicador
Demanda	Volumen de pacientes o servicios requeridos	Admisiones, egresos, consultas, procedimientos
Capacidad	Recursos disponibles para responder a la demanda	Camas disponibles, personal, servicios diagnósticos
Presión hospitalaria	Relación entre demanda y capacidad	Egresos por cama disponible
Congestión asistencial	Situación en la que la presión supera la capacidad de respuesta	Tiempos de espera, saturación, ocupación extrema

2.4. Analítica predictiva aplicada a la salud

La analítica predictiva aplicada a la salud utiliza datos históricos, métodos estadísticos y algoritmos de aprendizaje automático para estimar eventos futuros relacionados con pacientes, servicios, recursos y resultados clínicos. En el ámbito hospitalario, su finalidad es anticipar escenarios de demanda, ocupación, congestión o necesidad de recursos, de manera que los gestores puedan tomar decisiones preventivas y no únicamente reactivas (Orhan & Kurutkan, 2025; Blanco et al., 2026).

La predicción en salud se ha fortalecido con el crecimiento de los sistemas de información hospitalaria, los registros electrónicos, las bases administrativas y los sistemas de vigilancia. Estos repositorios permiten construir series históricas de admisiones, egresos, estancias, diagnósticos, camas y características de los pacientes. Cuando los datos se procesan adecuadamente, pueden revelar patrones temporales, estacionales, demográficos y operativos útiles para la planificación hospitalaria (Pan & Yan, 2026; Hadian et al., 2024).

La analítica predictiva puede emplearse en distintos niveles. En el nivel clínico, se utiliza para estimar riesgo de mortalidad, reingreso, complicaciones o evolución de enfermedades. En el nivel operativo, permite anticipar demanda de emergencia, ocupación de camas, uso de insumos y congestión asistencial. En el nivel estratégico, apoya la planificación de recursos, expansión de servicios y asignación presupuestaria. En esta investigación, el interés principal

se ubica en el nivel operativo, ya que se busca estimar la demanda hospitalaria y relacionarla con la capacidad disponible (Blanco et al., 2026; Orhan & Kurutkan, 2025).

Los modelos predictivos pueden dividirse en enfoques estadísticos y enfoques basados en aprendizaje automático. Los modelos estadísticos, como ARIMA y SARIMA, son útiles para series temporales porque describen autocorrelación, tendencia y estacionalidad. Los modelos de aprendizaje automático, como Random Forest, XGBoost y redes neuronales, pueden capturar relaciones no lineales e interacciones complejas entre variables. La elección del modelo depende de la naturaleza de los datos, el horizonte de predicción, la interpretabilidad requerida y el desempeño esperado (Hyndman & Athanasopoulos, 2021; Chen & Guestrin, 2016; Breiman, 2001).

En gestión hospitalaria, la analítica predictiva permite transformar indicadores históricos en herramientas de anticipación. Por ejemplo, si se conocen los patrones de egresos, estancias, ocupación y camas disponibles, es posible estimar periodos de mayor presión y planificar acciones como redistribución de camas, ajuste de personal o activación de protocolos de derivación. De esta forma, la predicción no reemplaza la gestión hospitalaria, sino que la complementa con evidencia cuantitativa (Fallahnezhad et al., 2024; Blanco et al., 2026).

2.5. Modelos de series temporales para predicción hospitalaria

Los modelos de series temporales son herramientas útiles cuando la variable de interés se observa secuencialmente a lo largo del tiempo. En el contexto hospitalario, muchas variables poseen estructura temporal, como admisiones diarias, egresos mensuales, ocupación de camas, días de estancia, consultas de emergencia y presión asistencial. Analizar estas series permite identificar tendencias, ciclos, estacionalidad y dependencia temporal, elementos relevantes para anticipar escenarios futuros (Box et al., 2015; Hyndman & Athanasopoulos, 2021).

2.5.1. Series temporales

Una serie temporal es una secuencia de observaciones ordenadas cronológicamente. A diferencia de un conjunto de datos transversal, en una serie temporal el orden importa porque los valores actuales pueden estar relacionados con valores pasados. En hospitales, esta dependencia se observa cuando la demanda de un periodo se parece a la de periodos anteriores o cuando existen patrones semanales, mensuales o estacionales (Hyndman & Athanasopoulos, 2021).

Las series temporales pueden contener tendencia, estacionalidad, ciclos y ruido. La tendencia representa un cambio sostenido en el tiempo; la estacionalidad corresponde a patrones repetitivos en intervalos regulares; los ciclos reflejan fluctuaciones de más largo plazo; y el ruido representa variaciones aleatorias. En el caso hospitalario, estos componentes pueden asociarse con cambios epidemiológicos, feriados, brotes infecciosos, variaciones de fin de semana o ciclos administrativos de atención (Box et al., 2015; Pan & Yan, 2026).

El análisis de series temporales requiere evaluar propiedades como estacionariedad y autocorrelación. Una serie estacionaria mantiene media y varianza relativamente constantes en el tiempo, mientras que la autocorrelación permite identificar hasta qué punto los valores pasados explican los valores presentes. Estos procedimientos ayudan a seleccionar modelos adecuados y a evitar predicciones basadas en relaciones poco consistentes (Hyndman & Athanasopoulos, 2021).

2.5.2. ARIMA

ARIMA, sigla de AutoRegressive Integrated Moving Average, es uno de los modelos estadísticos más utilizados para el pronóstico de series temporales. Su estructura combina tres componentes: el autorregresivo, que utiliza valores pasados de la propia serie; el integrado, que aplica diferenciación para lograr estacionariedad; y el de media móvil, que incorpora errores pasados para mejorar la predicción (Box et al., 2015).

El modelo se representa como ARIMA(p,d,q), donde p indica el número de rezagos autorregresivos, d el número de diferenciaciones necesarias para estabilizar la serie y q el número de rezagos de errores incluidos en el componente de media móvil. Esta estructura permite modelar series con dependencia temporal sin requerir necesariamente variables explicativas externas, por lo que suele emplearse como línea base en estudios de predicción de demanda hospitalaria (Hyndman & Athanasopoulos, 2021; Pan & Yan, 2026).

En el ámbito hospitalario, ARIMA resulta útil cuando la variable objetivo presenta patrones temporales relativamente estables. Puede aplicarse para predecir admisiones, egresos, consultas, ocupación de camas o presión asistencial agregada. Sin embargo, su principal limitación es que no captura de forma explícita relaciones no lineales complejas ni interacciones entre múltiples variables, por lo que puede ser superado por modelos de aprendizaje automático cuando existen varios predictores o estructuras no lineales (Pan & Yan, 2026; Hyndman & Athanasopoulos, 2021).

2.5.3. SARIMA

SARIMA, o Seasonal ARIMA, extiende el modelo ARIMA al incorporar componentes estacionales. Se expresa como SARIMA(p,d,q)(P,D,Q)s, donde los primeros parámetros representan la parte no estacional, los segundos la parte estacional y s define la periodicidad de la estacionalidad, por ejemplo 7 para patrones semanales en datos diarios o 12 para patrones mensuales en datos anuales (Box et al., 2015; Hyndman & Athanasopoulos, 2021).

La incorporación de estacionalidad es importante en salud, porque la demanda hospitalaria puede variar según días de la semana, meses del año, temporadas respiratorias, periodos vacacionales o brotes epidemiológicos. En estos casos, SARIMA permite modelar patrones que se repiten regularmente y mejorar la capacidad predictiva frente a un ARIMA simple (Pan & Yan, 2026).

En esta investigación, SARIMA es relevante como línea base estadística para la predicción del índice de presión hospitalaria. Su ventaja principal es la interpretabilidad y su capacidad para modelar dependencia temporal y estacionalidad. Sin embargo, su desempeño puede disminuir cuando la serie está influida por múltiples factores externos, cambios estructurales o relaciones no lineales. Por ello, resulta conveniente compararlo con modelos de aprendizaje automático como Random Forest, XGBoost o MLP (Hyndman & Athanasopoulos, 2021; Pan & Yan, 2026).

2.5.4. Modelo Naïve

El modelo Naïve es uno de los métodos de pronóstico más simples y utilizados como línea base o benchmark. Su principio consiste en asumir que el mejor pronóstico para el siguiente periodo corresponde al último valor observado de la serie temporal. Matemáticamente, puede expresarse como: $\hat{y}_{t+1} = y_t$, donde $\hat{y}_{t+1}$ representa el valor pronosticado e y_t corresponde a la observación más reciente disponible (McRae, 2021).

Aunque su formulación es sencilla, el modelo Naïve tiene importancia metodológica porque establece una referencia mínima que cualquier modelo más avanzado debería superar para demostrar valor predictivo adicional. En estudios de pronóstico de demanda hospitalaria, este tipo de modelo se utiliza como comparación frente a métodos más complejos, como modelos de regresión, redes neuronales o algoritmos de aprendizaje automático (McRae, 2021; Shahidian et al., 2025).

En esta investigación, el modelo Naïve se utiliza como modelo base para comparar el desempeño de los modelos predictivos posteriores. Su incorporación permite contar con una referencia inicial mediante métricas como MAE, RMSE y WMAPE, facilitando una evaluación más objetiva del aporte de los modelos de aprendizaje automático y aprendizaje profundo en la predicción de la presión hospitalaria (McRae, 2021).

2.5.5. Modelo Trend Naïve (T-Naïve)

El modelo Trend Naïve, también denominado Naïve con tendencia, extiende el modelo Naïve tradicional al incorporar la variación observada entre las dos últimas observaciones de la serie temporal. A diferencia del método Naïve, que supone que el siguiente valor será igual al último registrado, el modelo Trend Naïve proyecta hacia el futuro la dirección reciente de crecimiento o disminución de la serie (McRae, 2021).

Su formulación puede expresarse como:

$$\hat{y}_{t+1} = y_t + (y_t - y_{t-1})$$

donde la diferencia entre los dos últimos valores observados representa la variación reciente de la serie, la cual se proyecta hacia el siguiente período asumiendo la continuidad de la tendencia (McRae, 2021).

Este modelo resulta útil cuando la demanda hospitalaria presenta una tendencia creciente o decreciente relativamente estable en el corto plazo. Su principal ventaja es que mantiene una estructura sencilla, interpretable y computacionalmente eficiente, lo que permite evaluar si modelos más sofisticados aportan mejoras reales en la capacidad predictiva (McRae, 2021; Shahidian et al., 2025).

En el presente estudio, el modelo Trend Naïve se utiliza como un segundo modelo benchmark para comparar el desempeño de los modelos implementados frente a una referencia clásica y sencilla. Este método considera tanto el valor más reciente como la tendencia inmediata de la demanda hospitalaria, lo que permite evaluar si modelos más avanzados, como Random Forest, XGBoost y MLP_light, aportan una mejora real en la capacidad predictiva (McRae, 2021; Shahidian et al., 2025).

2.6. Modelos predictivos basados en aprendizaje automático

El aprendizaje automático permite construir modelos capaces de aprender patrones a partir de datos y generar predicciones sobre nuevas observaciones. A diferencia de algunos modelos estadísticos clásicos, los algoritmos de Machine Learning pueden manejar múltiples variables, relaciones no lineales, interacciones complejas y grandes volúmenes de información. En salud, estos modelos se han aplicado para predecir demanda, ocupación, riesgo clínico, reingresos, mortalidad, consumo de recursos y congestión hospitalaria (Orhan & Kurutkan, 2025; Blanco et al., 2026).

2.6.1. Random Forest

Random Forest es un algoritmo de aprendizaje supervisado basado en el ensamblaje de múltiples árboles de decisión. Cada árbol se entrena con una muestra aleatoria del conjunto de datos y con una selección aleatoria de variables, lo que reduce la varianza y mejora la capacidad de generalización. La predicción final se obtiene promediando los resultados de todos los árboles en problemas de regresión o mediante votación en problemas de clasificación (Breiman, 2001).

Una de sus principales ventajas es la capacidad para capturar relaciones no lineales sin requerir supuestos estrictos sobre la distribución de los datos. Además, proporciona medidas de importancia de variables, lo que permite identificar qué predictores tienen mayor influencia sobre la variable objetivo. En el contexto hospitalario, esto permite analizar si la presión asistencial se explica principalmente por egresos, camas disponibles, estancia, diagnósticos, temporalidad u otras variables operativas (Breiman, 2001; Orhan & Kurutkan, 2025).

Random Forest suele ser robusto frente a valores extremos y relaciones complejas; sin embargo, puede perder interpretabilidad frente a modelos lineales simples y requiere validación cuidadosa para evitar sobreajuste. En esta investigación, se considera apropiado

porque permite modelar la presión hospitalaria a partir de múltiples variables históricas y evaluar la importancia relativa de los factores asociados con demanda y capacidad (Breiman, 2001; Pedregosa et al., 2011).

2.6.2. XGBoost

XGBoost, o Extreme Gradient Boosting, es un algoritmo de ensamble basado en boosting. A diferencia de Random Forest, que construye árboles de manera relativamente independiente, XGBoost construye árboles de forma secuencial, donde cada nuevo árbol intenta corregir los errores de los anteriores. Esta estrategia permite obtener modelos de alta precisión, especialmente en problemas tabulares con variables numéricas y categóricas transformadas (Chen & Guestrin, 2016).

El algoritmo incorpora regularización, manejo eficiente de datos, optimización del gradiente y control de complejidad del modelo. Estas características explican su amplio uso en ciencia de datos aplicada. En salud, XGBoost se ha utilizado para predicción de demanda, riesgo clínico, reingresos, ocupación y otros eventos hospitalarios debido a su capacidad para capturar relaciones no lineales y jerarquizar variables importantes (Chen & Guestrin, 2016; Orhan & Kurutkan, 2025).

En el contexto de la predicción de presión hospitalaria, XGBoost puede ser útil cuando la variable objetivo depende de interacciones complejas entre demanda, capacidad, temporalidad y características clínicas agregadas. No obstante, requiere una adecuada optimización de hiperparámetros, ya que parámetros como profundidad del árbol, tasa de aprendizaje, número de estimadores y regularización pueden modificar sustancialmente el desempeño del modelo (Chen & Guestrin, 2016; Bergstra & Bengio, 2012).

2.6.3. Red neuronal artificial tipo MLP

Una red neuronal artificial tipo MLP, o perceptrón multicapa, es un modelo compuesto por capas de neuronas artificiales conectadas entre sí. Generalmente incluye una capa de entrada, una o más capas ocultas y una capa de salida. Cada neurona realiza una combinación ponderada de las entradas, aplica una función de activación y transmite el resultado a la siguiente capa. Durante el entrenamiento, la red ajusta sus pesos para minimizar el error de predicción (Rumelhart et al., 1986; Goodfellow et al., 2016).

El MLP puede aproximar funciones no lineales complejas, por lo que resulta útil cuando la relación entre las variables predictoras y la variable objetivo no puede representarse de manera simple. En problemas hospitalarios, puede capturar patrones entre demanda, capacidad, temporalidad y condiciones clínicas agregadas. Sin embargo, a diferencia de Random Forest o XGBoost, suele requerir normalización de variables, selección cuidadosa de arquitectura, control de sobreajuste y mayor volumen de datos para obtener resultados estables (Goodfellow et al., 2016).

En esta investigación, el MLP se considera como modelo comparativo de aprendizaje profundo simple, sin abordar arquitecturas más especializadas como redes recurrentes, convolucionales o modelos transformer. Su inclusión permite evaluar si una arquitectura neuronal básica puede mejorar la predicción de la presión hospitalaria frente a modelos estadísticos y modelos de árboles de decisión. Esta comparación es relevante porque los modelos más complejos no siempre superan a modelos más simples cuando se trabaja con series agregadas o conjuntos de datos tabulares (Goodfellow et al., 2016; Orhan & Kurutkan, 2025).

2.7. Optimización de hiperparámetros

La optimización de hiperparámetros es una etapa clave en el desarrollo de modelos predictivos. Mientras los parámetros se aprenden durante el entrenamiento, los hiperparámetros se definen antes o durante el proceso de ajuste y controlan aspectos como la complejidad, velocidad de aprendizaje, profundidad, regularización o tamaño de la arquitectura. Una configuración inadecuada puede producir sobreajuste, subajuste o bajo desempeño predictivo (Bergstra & Bengio, 2012; Pedregosa et al., 2011).

2.7.1. Hiperparámetros

Los hiperparámetros son valores externos al aprendizaje directo del modelo que regulan su comportamiento. En Random Forest, algunos hiperparámetros importantes son el número de árboles, la profundidad máxima, el número mínimo de muestras por hoja y la cantidad de variables consideradas en cada división. En XGBoost, destacan la tasa de aprendizaje, el número de estimadores, la profundidad máxima, la regularización y la proporción de columnas o filas utilizadas. En un MLP, incluyen número de capas, número de neuronas, función de activación, tasa de aprendizaje, tamaño de lote y número de épocas (Chen & Guestrin, 2016; Goodfellow et al., 2016).

La optimización busca encontrar combinaciones que mejoren el desempeño del modelo en datos no vistos. Para ello, se utilizan procedimientos de validación que separan datos de entrenamiento y prueba, evitando evaluar el modelo únicamente sobre los datos con los que fue entrenado. En series temporales, esta separación debe respetar el orden cronológico, porque utilizar datos futuros para entrenar el modelo produciría fuga de información y resultados artificialmente optimistas (Hyndman & Athanasopoulos, 2021; Pedregosa et al., 2011).

2.7.2. RandomizedSearchCV

RandomizedSearchCV es una técnica de búsqueda aleatoria de hiperparámetros implementada ampliamente en bibliotecas de aprendizaje automático. A diferencia de GridSearchCV, que evalúa todas las combinaciones posibles dentro de una cuadrícula predefinida, RandomizedSearchCV selecciona aleatoriamente un número limitado de combinaciones desde distribuciones o listas de valores. Esta estrategia reduce el costo computacional y permite explorar espacios de búsqueda amplios de manera eficiente (Bergstra & Bengio, 2012; Pedregosa et al., 2011).

La búsqueda aleatoria ha demostrado ser eficiente porque, en muchos modelos, no todos los hiperparámetros tienen la misma influencia sobre el desempeño. Por ello, evaluar combinaciones aleatorias puede encontrar configuraciones competitivas con menor tiempo de cómputo que una búsqueda exhaustiva. En esta investigación, RandomizedSearchCV se utiliza para optimizar modelos como Random Forest y XGBoost, respetando la división temporal de entrenamiento y prueba. El objetivo no es únicamente obtener el menor error posible, sino construir modelos generalizables que mantengan buen desempeño ante periodos futuros de presión hospitalaria (Bergstra & Bengio, 2012; Pedregosa et al., 2011; Hyndman & Athanasopoulos, 2021).

2.8. Métricas de evaluación de modelos predictivos

La evaluación de modelos predictivos requiere métricas que permitan cuantificar la diferencia entre los valores reales y los valores estimados. En problemas de regresión y series temporales, estas métricas permiten comparar modelos, seleccionar el mejor enfoque y valorar si el desempeño es suficiente para apoyar decisiones operativas. En predicción hospitalaria, su interpretación debe relacionarse con el impacto práctico de los errores, especialmente durante periodos de mayor presión asistencial (Hyndman & Koehler, 2006; Willmott & Matsuura, 2005).

2.8.1. MAE

El Error Absoluto Medio, conocido como MAE por sus siglas en inglés, calcula el promedio de las diferencias absolutas entre los valores reales y los valores predichos. Su fórmula general es: $MAE = (1/n) \Sigma |y_i - \hat{y}_i|$. Esta métrica se expresa en las mismas unidades de la variable objetivo, lo que facilita su interpretación práctica (Hyndman & Koehler, 2006).

Una ventaja del MAE es que trata todos los errores de forma lineal, sin penalizar excesivamente los errores grandes. Por ello, resulta útil cuando se desea conocer el error promedio típico del modelo. En esta investigación, el MAE permite estimar cuánto se equivoca en promedio el modelo al predecir el índice de presión hospitalaria (Hyndman & Koehler, 2006; Willmott & Matsuura, 2005).

2.8.2. RMSE

La Raíz del Error Cuadrático Medio, o RMSE, calcula la raíz cuadrada del promedio de los errores elevados al cuadrado. Su fórmula general es: $RMSE = \sqrt{[(1/n) \Sigma(y_i - \hat{y}_i)^2]}$. Al elevar los errores al cuadrado, esta métrica penaliza con mayor intensidad los errores grandes, por lo que es sensible a valores extremos (Hyndman & Koehler, 2006).

El RMSE es útil cuando los errores grandes tienen consecuencias importantes. En un contexto hospitalario, subestimar de manera considerable la presión asistencial podría afectar la planificación de camas, personal y recursos. Por ello, esta métrica complementa al MAE al mostrar si el modelo comete errores extremos que podrían ser operativamente relevantes (Willmott & Matsuura, 2005).

2.8.3. WMAPE

El Error Porcentual Absoluto Medio Ponderado, conocido como WMAPE, mide el error absoluto total en relación con el volumen total observado. Una forma común de expresarlo es: $WMAPE = \Sigma|y_i - \hat{y}_i| / \Sigma|y_i|$. Esta métrica permite interpretar el error en términos relativos y

resulta especialmente útil cuando se comparan series con diferentes escalas o magnitudes (Hyndman & Koehler, 2006).

En predicción hospitalaria, el WMAPE es valioso porque permite expresar el error como una proporción del nivel real de presión o demanda. A diferencia del MAPE tradicional, suele ser más estable cuando existen valores pequeños o cercanos a cero, ya que pondera el error por el volumen total observado. En esta investigación, el WMAPE complementa al MAE y al RMSE al ofrecer una interpretación porcentual del desempeño predictivo (Hyndman & Koehler, 2006).

El uso conjunto de MAE, RMSE y WMAPE permite evaluar los modelos desde tres perspectivas: error promedio absoluto, sensibilidad a errores grandes e interpretación relativa del error. Esta combinación ofrece una visión más completa del desempeño y facilita la comparación entre modelos estadísticos, de aprendizaje automático y redes neuronales (Hyndman & Koehler, 2006; Willmott & Matsuura, 2005).

Tabla 2-4

Métricas de evaluación de modelos predictivos

Métrica	**Qué mide**	**Interpretación en el proyecto**
MAE	Error promedio absoluto	Cuánto se equivoca en promedio el modelo al predecir el IPH.
RMSE	Error cuadrático promedio con penalización de errores grandes	Permite detectar modelos que fallan mucho en ciertos periodos.
WMAPE	Error absoluto ponderado respecto al total observado	Expresa el error relativo del modelo frente al nivel real de presión.

CAPÍTULO 3 . DESARROLLO.

El marco metodológico de esta investigación se organiza bajo la metodología CRISP-DM, complementada con un enfoque técnico de preparación de datos por capas Bronze-Silver-Gold. Esta combinación permite estructurar el estudio desde la comprensión del problema hasta la generación de resultados predictivos útiles para la planificación hospitalaria.

El análisis exploratorio de datos (EDA) se desarrolla de forma secuencial. En primer lugar, se diagnostica la integración entre las fuentes de camas y egresos, revisando la cobertura de llaves, el relacionamiento entre registros y la calidad de los datos de origen. Luego, tras el preprocesamiento, se construye la tabla analítica integrada o capa Gold, sobre la cual se realiza un EDA de negocio para validar cobertura, coherencia y comportamiento general del fenómeno hospitalario. Finalmente, se genera la tabla de entrenamiento mediante ingeniería de variables, incorporando información histórica de periodos anteriores, promedios móviles, variables de calendario y la definición de la variable objetivo del modelo.

Figura 3-1

Metodología CRISP-DM

En la fase de comprensión del negocio (3.1), se define el problema desde la planificación hospitalaria: estimar la demanda y relacionarla con la disponibilidad de camas para identificar grupos de establecimientos con posible presión asistencial. En la comprensión de los datos (3.2), se analizan las fuentes de egresos y camas, considerando su granularidad, cobertura temporal, calidad y utilidad para el estudio. Posteriormente, en la preparación de datos (3.3), se aplica el enfoque Bronze-Silver-Gold, donde los datos originales se conservan, depuran, normalizan e integran hasta construir la tabla analítica final. En esta misma fase se realiza la ingeniería de características, la definición del target, la prevención de fuga de información y la partición temporal para entrenamiento, validación y prueba.

En la fase de modelado (3.4), se define el diseño experimental y se entrenan modelos basales, como Naïve y Seasonal Naïve, junto con modelos de Machine Learning como Random Forest y XGBoost. El ajuste de hiperparámetros se realiza mediante RandomizedSearchCV con validación cruzada temporal. Luego, en la evaluación (3.5), los modelos se comparan mediante métricas como MAE, RMSE y WMAPE. Finalmente, la fase de despliegue analítico (3.6) se orienta al apoyo de decisiones mediante tablas, gráficos, predicciones e indicadores que permiten interpretar la demanda esperada frente a la capacidad disponible.

En síntesis, CRISP-DM proporciona la estructura metodológica general, mientras que el enfoque Bronze-Silver-Gold organiza la integración y preparación de los datos. La combinación de ambos enfoques permite desarrollar un proceso ordenado, reproducible y coherente, desde la validación de las fuentes hasta la generación de evidencia predictiva útil para la planificación hospitalaria.

3.1 Fase 1: Comprensión del negocio

La planificación hospitalaria requiere información oportuna y confiable sobre la demanda de atención y la capacidad disponible de los establecimientos de salud. En este contexto, los

egresos hospitalarios permiten aproximar el volumen de atención brindada, mientras que la disponibilidad de camas representa una dimensión clave de la capacidad hospitalaria.

Anticipar la demanda resulta fundamental para apoyar decisiones relacionadas con la asignación de recursos, la organización de servicios y la priorización de capacidad. Cuando la demanda esperada se aproxima o supera la capacidad disponible, pueden generarse situaciones de alta presión operativa que afectan la respuesta institucional.

En Ecuador existen fuentes de información hospitalaria sobre egresos y camas; sin embargo, estas no siempre se encuentran integradas en una estructura analítica común. Esta separación limita el análisis conjunto entre demanda observada y capacidad disponible, así como la construcción de modelos predictivos útiles para la toma de decisiones.

Por ello, esta investigación integra información de egresos hospitalarios y camas correspondientes al período 2022–2024. A partir de estas fuentes, se construye un dataset analítico para estimar la demanda hospitalaria futura en grupos de establecimientos definidos por variables geográficas e institucionales comunes. Posteriormente, dicha demanda se relaciona con la disponibilidad de camas para apoyar la planificación de recursos hospitalarios.

3.1.1 Contexto de planificación hospitalaria

La planificación hospitalaria suele apoyarse en información histórica, reportes administrativos, disponibilidad declarada de recursos y criterios de gestión institucional. Aunque estos elementos son necesarios, pueden resultar insuficientes cuando se requiere anticipar la demanda futura o identificar oportunamente escenarios de alta presión sobre la capacidad disponible.

En muchos casos, la información de demanda y capacidad se analiza por separado. Los egresos hospitalarios permiten conocer el volumen de atención brindada, su comportamiento temporal, características de los pacientes y causas de egreso. Por su parte, los registros de

camas muestran la capacidad reportada por los establecimientos, incluyendo dotación normal, camas disponibles y disponibilidad por áreas o servicios. Sin embargo, si estas fuentes no se integran adecuadamente, se dificulta evaluar la relación entre la demanda esperada y la capacidad disponible.

La integración de datos hospitalarios también presenta desafíos metodológicos. En las bases disponibles para esta investigación, no se identificó un código único común que permita enlazar directamente cada egreso con un hospital individual y su respectiva información de camas. Por esta razón, se validaron previamente las variables comunes de ubicación e institucionalidad, con el fin de definir una unidad analítica consistente para el estudio.

Como resultado, la investigación no trabaja a nivel de hospital individual, sino a nivel de grupos de establecimientos de salud definidos por atributos geográficos e institucionales comunes. Esta decisión permite integrar las fuentes de datos de forma coherente con la granularidad real de la información disponible.

En este escenario, el uso de modelos predictivos representa una oportunidad para fortalecer la planificación hospitalaria. La estimación de la demanda futura de egresos, combinada con la disponibilidad de camas, puede generar indicadores útiles para identificar grupos de establecimientos, provincias o áreas de atención donde la demanda esperada sea alta en relación con la capacidad disponible.

3.1.2 Problema analítico del estudio

Las fuentes de egresos hospitalarios y camas disponibles contienen información relevante para analizar la demanda y la capacidad del sistema hospitalario. Sin embargo, originalmente se encuentran separadas y presentan diferencias de granularidad: los egresos corresponden a eventos hospitalarios individuales, mientras que la información de camas se presenta de forma agregada según características geográficas e institucionales.

Esta diferencia impide construir una base integrada a nivel de hospital individual, especialmente porque no existe un identificador único común que permita enlazar directamente cada egreso con su respectivo establecimiento y registro de camas. Por ello, el estudio define una unidad analítica basada en grupos de establecimientos con atributos geográficos e institucionales comunes, evitando asumir una granularidad que los datos no respaldan.

A esta dificultad se suma el reto de construir modelos predictivos sin fuga de información. Algunas variables derivadas de los egresos, como causas del mismo período, porcentajes por sexo, días de estadía o indicadores calculados con el valor objetivo, pueden ser útiles para el análisis descriptivo, pero no necesariamente válidas como predictores si incorporan información que no estaría disponible antes del período a predecir.

En este sentido, el problema analítico consiste en integrar adecuadamente las fuentes de egresos y camas, construir un dataset de modelado consistente y evaluar modelos predictivos capaces de estimar la demanda hospitalaria futura. Posteriormente, la demanda estimada se relaciona con la disponibilidad de camas para generar un indicador demanda-capacidad que apoye la planificación de recursos hospitalarios.

En síntesis, la investigación busca contar con un enfoque analítico y predictivo que permita estimar la demanda hospitalaria futura e identificar grupos de establecimientos donde dicha demanda pueda ser alta en comparación con la capacidad disponible de camas, utilizando de forma integrada la información histórica de egresos y camas hospitalarias.

3.1.3 Objetivo analítico del modelo

El objetivo analítico del modelo es estimar la demanda hospitalaria futura en grupos de establecimientos de salud del Ecuador, definidos por variables geográficas e institucionales comunes, utilizando información histórica de egresos hospitalarios, disponibilidad de camas y causas de egreso del período 2022–2024. Posteriormente, la demanda estimada se relaciona

con la capacidad disponible de camas, con el propósito de generar un indicador útil para apoyar la planificación de recursos hospitalarios.

Para alcanzar este objetivo, el estudio integra las bases de egresos hospitalarios y camas mediante una llave natural construida a partir de variables comunes. Esta integración permite consolidar una fuente analítica coherente con la granularidad real de los datos disponibles. Además, se evalúa la calidad, cobertura y consistencia de las fuentes originales, considerando aspectos como duplicidad, correspondencia entre bases, disponibilidad de camas y comportamiento de los egresos. Este análisis permite definir adecuadamente la unidad analítica del estudio y evitar interpretaciones a nivel de hospital individual cuando los datos no lo respaldan.

A partir de la tabla integrada, se construye un dataset de modelado predictivo con variables temporales, geográficas, institucionales, disponibilidad de camas, causas de egreso e información histórica de la demanda. Durante este proceso se prioriza la prevención de fuga de información, evitando utilizar variables que no estarían disponibles antes del período a predecir.

Finalmente, se entrenan y comparan modelos de Machine Learning, principalmente Random Forest y XGBoost, y se incorpora un modelo MLP ligero como contraste complementario de Deep Learning. El desempeño se evalúa mediante métricas de regresión como MAE, RMSE y WMAPE, con el fin de identificar el modelo con mejor capacidad predictiva y utilizar sus resultados para estimar la relación demanda-capacidad.

3.1.4 Definición de variables del estudio

La definición de variables del estudio se realizó considerando el objetivo de estimar la demanda hospitalaria futura y relacionarla con la capacidad disponible de camas. Para ello, se distinguieron tres componentes principales: la variable objetivo, las variables explicativas y los indicadores derivados para el análisis operativo.

Esta definición se presenta antes de la construcción de la tabla Gold y del dataset de modelado, ya que orienta qué información debe integrarse, qué variables deben conservarse, cuáles requieren transformación y cuáles no deben utilizarse como predictores para evitar fuga de información.

3.1.4.1. Variable objetivo: demanda hospitalaria

La variable objetivo corresponde a la demanda hospitalaria futura, medida a través del número de egresos hospitalarios estimados para un grupo de establecimientos en un período determinado. En este estudio, los egresos se utilizan como una aproximación de la demanda atendida, debido a que representan eventos hospitalarios efectivamente registrados y permiten cuantificar el volumen de atención finalizada.

A partir de la fuente transaccional de egresos, los registros individuales se agregan por unidad analítica y período temporal para obtener la demanda observada. Esta información histórica se utiliza posteriormente para entrenar modelos predictivos orientados a estimar la demanda futura.

Para el proceso predictivo, la demanda estimada por el modelo se representa como:

$$\widehat{D}_{i,t}$$

Donde:

Símbolo	Significado
$\widehat{D}_{i,t}$	Demanda hospitalaria predicha, medida como egresos hospitalarios estimados
(i)	Grupo de establecimientos de salud
(t)	Período de análisis

Por ejemplo, si para un grupo de establecimientos el modelo estima 80 egresos para un período determinado, ese valor representa la demanda hospitalaria esperada para dicho grupo

y período. Esta predicción no se analiza de forma aislada, sino que posteriormente se compara con la disponibilidad de camas para estimar la presión relativa sobre la capacidad hospitalaria.

3.1.4.2. Variables explicativas disponibles

Las variables explicativas corresponden a los atributos utilizados por los modelos para estimar la demanda hospitalaria. Estas provienen de las fuentes de egresos hospitalarios y camas, así como de transformaciones realizadas durante la construcción del dataset de modelado.

Las variables se organizan en los siguientes grupos:

Tabla 3-1

Variables explicativas para estimar la demanda hospitalaria

Variables	Descripción
Geográficas e institucionales	Incluyen provincia, cantón, parroquia, área de ubicación, clase del establecimiento, tipo de establecimiento, entidad y sector. Estas variables permiten capturar diferencias territoriales e institucionales entre grupos de establecimientos.
De capacidad hospitalaria.	Incluyen la disponibilidad de camas reportada y, cuando corresponde, camas disponibles por áreas o servicios específicos. La variable principal de capacidad es el total de camas disponibles, mientras que la dotación normal se conserva como referencia auxiliar para control de calidad.
Temporales.	Incluyen año, mes, día y otras variables derivadas del calendario. Estas variables permiten capturar patrones temporales, variaciones estacionales y cambios en el comportamiento de los egresos durante el período de estudio.
Históricas de demanda.	Corresponden a variables construidas a partir de egresos observados en períodos anteriores, como egresos previos y promedios móviles. Estas variables permiten que el modelo utilice información histórica disponible antes

	del período que se desea predecir. Por ejemplo, para estimar la demanda de una semana determinada, podrían utilizarse los egresos registrados en semanas anteriores, pero no los egresos de la misma semana que se intenta predecir.
Relacionadas con causas de egreso.	Incluyen agrupaciones de diagnósticos o causas de egreso, como capítulos CIE-10. Estas variables permiten caracterizar el perfil histórico de la demanda hospitalaria. Para evitar fuga de información, su uso como predictor debe limitarse a información histórica o agregada de períodos anteriores, y no a causas observadas en el mismo período que se desea predecir.

En la construcción del dataset de modelado se garantiza que las variables predictoras estén disponibles antes del período que se desea estimar. Por tanto, variables calculadas con información del mismo período objetivo, como egresos por causa del período actual, porcentajes de egresos por sexo del período actual o indicadores derivados directamente del valor objetivo, no se utilizan como predictores directos.

3.1.4.3. Indicador de relación demanda-capacidad (RDC)

Además de predecir la demanda hospitalaria, el estudio calcula un indicador derivado denominado relación demanda-capacidad (RDC). Este indicador compara la demanda hospitalaria estimada con la disponibilidad de camas reportada para cada grupo de establecimientos. Para su cálculo se utiliza la variable total_camas_disponibles, definida como la medida principal y estandarizada de capacidad.

La relación demanda-capacidad se define de la siguiente manera:

$$RDC_{i,t} = \frac{\widehat{D}_{i,t}}{C_{i,t}}$$

Donde:

Símbolo	Significado
$RDC_{i,t}$	Relación demanda-capacidad del grupo de establecimientos (i) en el período (t)
$\widehat{D}_{i,t}$	Demanda hospitalaria predicha, medida como egresos hospitalarios estimados
$C_{i,t}$	Camas disponibles reportadas para el grupo de establecimientos (i) en el período (t)

Ejemplo de interpretación de la RDC:

Si un grupo de establecimientos tiene una demanda predicha de 120 egresos y cuenta con 100 camas disponibles, la relación demanda-capacidad sería:

RDC = 120 / 100 = 1.20

Esto indica que la demanda estimada equivale al 120 % de las camas disponibles. En términos operativos, una RDC superior a 1 sugiere una presión relativa alta, ya que la demanda esperada supera la capacidad reportada. Por el contrario, si la demanda predicha fuera de 70 egresos y el grupo tuviera 100 camas disponibles, la RDC sería 0.70, lo que indicaría una presión relativa menor.

En términos generales, una RDC baja refleja menor presión sobre la capacidad disponible. A medida que la RDC aumenta, la presión relativa también crece. Cuando el indicador se acerca o supera el valor de 1, la demanda estimada iguala o excede la capacidad disponible, lo que puede requerir acciones de monitoreo o planificación preventiva.

Este indicador no estima ocupación hospitalaria efectiva, sino presión asistencial relativa sobre la capacidad reportada. Por ello, no incorpora directamente la dinámica real de uso de camas, la estancia media ni la concurrencia de pacientes. Su utilidad principal es servir como una métrica comparativa para priorizar grupos de establecimientos que podrían requerir mayor atención operativa.

3.1.4.4. Lift de alerta operativa

Para complementar el análisis de priorización, se incorpora el cálculo del lift de alerta operativa. Esta medida permite comparar qué combinaciones de provincia y causa de egreso presentan una asociación mayor o menor con niveles de prioridad alta o media, en relación con el comportamiento promedio del conjunto analizado.

En este estudio, el lift se define como la razón entre la tasa de alerta de una causa específica y la tasa de alerta promedio del conjunto analizado:

$$\text{lift} = \frac{\text{tasa de alerta de la causa}}{\text{tasa de alerta global}}$$

La tasa de alerta corresponde a la proporción de registros clasificados en prioridad alta o media.

Su interpretación práctica es la siguiente:

lift = 1: la causa tiene una asociación con alerta similar al promedio general.

lift > 1: la causa tiene una asociación con alerta mayor que el promedio.

lift < 1: la causa tiene una asociación con alerta menor que el promedio.

Ejemplo de interpretación del lift:

Si se tuviese en un conjunto de registros, el 20 % presenta prioridad alta o media. Si para una combinación específica de provincia y causa de egreso la tasa de alerta es del 40 %, entonces el lift sería:

lift = 40 % / 20 % = 2.0

Esto significa que esa combinación presenta alertas con una frecuencia dos veces mayor que el promedio general. En cambio, si la tasa de alerta de otra combinación fuera del 10 %, el lift sería 0.5, lo que indicaría una asociación menor con niveles de alerta.

En síntesis, el lift permite identificar focos territoriales y clínicos con mayor asociación operativa a niveles de alerta. Sin embargo, debe interpretarse como una medida de priorización y no como evidencia de causalidad clínica.

3.2 Fase 2: Comprensión de los datos

En la fase de comprensión de los datos se analizaron las fuentes de egresos hospitalarios y camas disponibles, con el propósito de evaluar su cobertura, granularidad, calidad y posibilidad de integración. Para ello, el pipeline de preprocesamiento contempló dos procedimientos principales: el análisis de cobertura de llaves y el relacionamiento entre tablas. Los archivos generados, las evidencias gráficas y el detalle técnico completo del proceso se presentan en el Anexo 1.

3.2.1. Fuentes de datos

El estudio utilizó una fuente de egresos hospitalarios compuesta por 3,434,083 registros correspondientes al período 2022–2024. Esta fuente representa la demanda observada de atención hospitalaria, medida mediante eventos de ingreso y egreso registrados en establecimientos de salud del Ecuador.

La distribución de registros de egresos por año fue la siguiente:

Tabla 3-2

Distribución de registros de egresos por año

Año	Registros de egresos
2022	1,130,603
2023	1,170,813
2024	1,132,667

Además, se incorporó información de camas hospitalarias, utilizada como aproximación de la capacidad disponible. Esta fuente contiene datos de dotación normal y disponibilidad de camas por especialidad y período.

La distribución de registros de camas por año fue la siguiente:

Tabla 3-3

Distribución de registros de camas por año

Año	Registros de camas
2022	633
2023	633
2024	627

En conjunto, estas fuentes permiten representar dos dimensiones complementarias del sistema hospitalario: la demanda observada, a partir de los egresos hospitalarios, y la capacidad disponible, a partir de la información de camas.

3.2.2. Variables comunes, llave natural y lógica de relacionamiento entre fuentes

La fuente de egresos hospitalarios contiene registros asociados a eventos de salida hospitalaria. Incluye variables relacionadas con la ubicación del establecimiento, características del paciente, fechas de ingreso y egreso, días de estadía, condición de egreso, especialidad y causa de egreso. Su naturaleza es transaccional, ya que cada registro representa un egreso hospitalario individual.

La fuente de camas hospitalarias contiene información sobre dotación normal y disponibilidad de camas. Incluye variables de ubicación, clase y tipo de establecimiento, entidad, sector, camas disponibles, camas de dotación normal y camas por servicios o áreas de atención. Esta fuente permite aproximar la capacidad hospitalaria disponible en cada período analizado.

Para integrar ambas fuentes se utilizaron variables comunes de ubicación y características institucionales. La llave natural evaluada estuvo conformada por ocho variables:

Tabla 3-4

Variables comunes de ubicación y características institucionales

Variable	Descripción general
Provincia de ubicación	Provincia donde se ubica el establecimiento de salud
Cantón de ubicación	Cantón donde se ubica el establecimiento de salud
Parroquia de ubicación	Parroquia donde se ubica el establecimiento de salud
Área de ubicación	Área de ubicación del establecimiento
Clase del establecimiento	Clase del establecimiento de salud
Tipo de establecimiento	Tipo de establecimiento de salud
Entidad	Entidad a la que pertenece el establecimiento
Sector	Sector institucional del establecimiento

Estas variables permitieron construir una llave natural para el emparejamiento entre camas y egresos. Sin embargo, esta llave no identifica necesariamente a un hospital individual, sino a grupos de establecimientos que comparten atributos geográficos e institucionales comunes. Desde el punto de vista técnico, la llave natural fue definida en el pipeline implementado con lenguaje python mediante ocho columnas comunes entre ambas fuentes:

```
LLAVE = [
    "prov_ubi", "cant_ubi", "parr_ubi", "area_ubi",
    "clase", "tipo", "entidad", "sector"
]
```

En el primer análisis se evaluó la cobertura inicial de llaves entre ambas fuentes. El objetivo fue identificar cuántas combinaciones coincidían exactamente y cuántas permanecían solo en

una de las bases. Los resultados mostraron una cobertura inicial alta en los tres años analizados.

Tabla 3-5

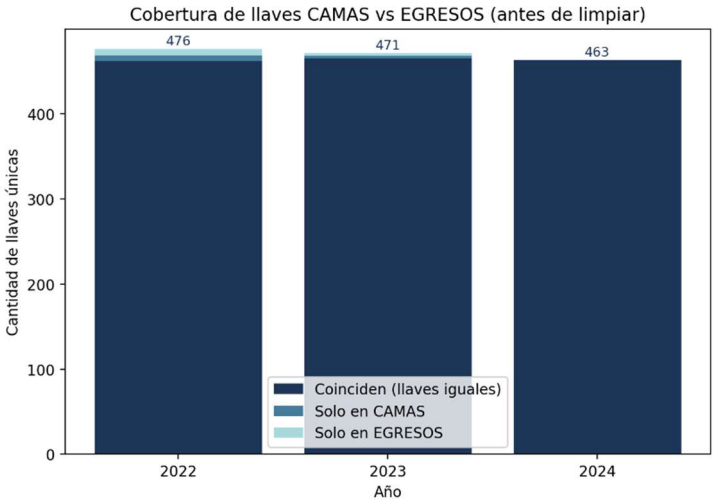
Cobertura inicial alta en los tres años analizados

Año	Llaves únicas CAMAS	Llaves únicas EGRESOS	Llaves que coinciden	Solo CAMAS	Solo EGRESOS	Cobertura CAMAS (%)	Cobertura EGRESOS (%)
2022	469	469	462	7	7	98.51	98.51
2023	468	468	465	3	3	99.36	99.36
2024	463	463	463	0	0	100.00	100.00

Estos resultados evidencian que, antes de la limpieza textual, la mayoría de combinaciones entre ambas fuentes ya presentaba correspondencia directa. Las diferencias identificadas fueron reducidas y se concentraron principalmente en los años 2022 y 2023.

Figura 3-2

Cobertura de llaves CAMAS vs. EGRESOS antes de la limpieza

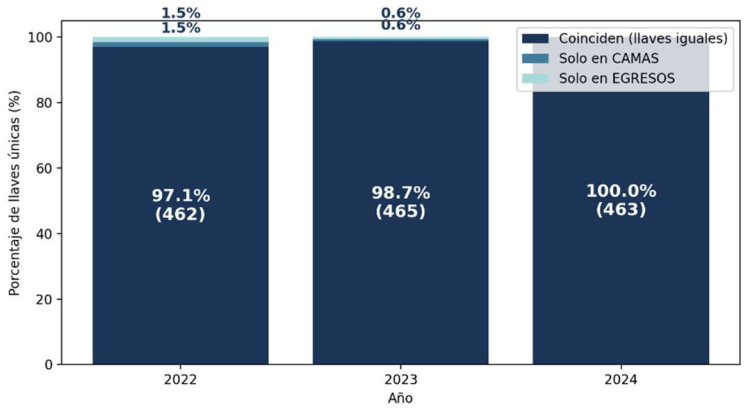


La figura evidencia que, antes de la limpieza, una proporción relevante de llaves de cada fuente no encontraba correspondencia en la otra, lo que confirmó la existencia de diferencias

de escritura y registro entre CAMAS y EGRESOS. Este diagnóstico justificó la etapa de relacionamiento textual descrita a continuación, orientada a recuperar cruces válidos sin forzar coincidencias de alto riesgo.

Figura 3-3

Cobertura porcentual de llaves CAMAS vs. EGRESOS antes de la limpieza



Una vez identificadas las llaves no coincidentes, se aplicó un proceso de normalización y relacionamiento textual. Este procedimiento permitió revisar si las diferencias correspondían a errores ortográficos, variaciones menores de escritura o expresiones equivalentes entre ambas fuentes.

El proceso incluyó la normalización de textos mediante conversión a mayúsculas, eliminación de tildes, unificación de espacios y depuración de caracteres especiales. Posteriormente, se compararon las llaves exclusivas de CAMAS y EGRESOS mediante criterios de similitud por columna y similitud global de la llave completa.

De forma resumida, la lógica de decisión combinó la similitud promedio entre columnas y la similitud global de la llave. Solo se conservaron como candidatos los pares con alta similitud y con diferencias limitadas entre columnas:

```
score_final = round((0.65 * score_cols) + (0.35 * score_key), 4)
if score_final < 0.82 or cols_iguales < (len(LLAVE) - 2):
    aceptar_candidato = True
```

Como resultado del relacionamiento, se identificaron 14 candidatos de emparejamiento entre 2022 y 2023. De ellos, 10 fueron clasificados como coincidencias sugeridas principales, 4 se consolidaron como mapeos formales de equivalencia textual, 4 casos adicionales se resolvieron automáticamente mediante reglas derivadas de esos mapeos y no quedaron casos pendientes de revisión manual.

Tabla 3-6

Resultado del proceso de normalización y relacionamiento textual

Indicador	**Valor**
Candidatos leídos	14
Matches sugeridos en detalle	10
Mapeos generados	4
Casos auto-resueltos por reglas	4
Casos pendientes de revisión manual	0

Los principales mapeos detectados correspondieron a diferencias ortográficas en cantón y parroquia. En particular, se identificaron equivalencias entre “Azoques” y “Azogues”, y entre “San Crstóbal” y “San Cristóbal”. Estas variantes explicaban las discrepancias observadas en 2022 y 2023.

Tabla 3-7

Variantes de las discrepancias observadas en 2022 y 2023

Columna	**Valor observado**	**Valor canónico relacionado**	**Años detectados**	**Número de coincidencias**	**Score promedio**
cant_ubi	Azoques	Azogues	2022, 2023	7	0.9856
parr_ubi	San Crstóbal	San Cristóbal	2022	3	0.9956

Después de aplicar las equivalencias textuales, la cobertura potencial entre CAMAS y EGRESOS alcanzó el 100 % en todos los años analizados. Esto significa que todas las llaves observadas en una fuente encontraron correspondencia en la otra, ya sea por coincidencia exacta o por equivalencia textual validada.

Tabla 3-8

Cobertura ajustada de llaves después del relacionamiento por caracteres

Año	Llaves CAMAS	Llaves EGRESOS	Coincidencias exactas iniciales	Matches sugeridos por caracteres	Coincidencias potenciales	Cobertura potencial CAMAS (%)	Cobertura potencial EGRESOS (%)
2022	469	469	462	7	469	100.00	100.00
2023	468	468	465	3	468	100.00	100.00
2024	463	463	463	0	463	100.00	100.00

En términos metodológicos, esta fase permitió concluir que la llave natural construida a partir de variables geográficas e institucionales es adecuada para integrar ambas fuentes a nivel de combinaciones observadas. No obstante, también confirmó que dicha llave no representa un identificador único de hospital individual, sino una agrupación de establecimientos con atributos comunes. Esta precisión es importante para mantener coherencia entre la estructura de los datos, la integración de fuentes y la definición posterior de la unidad analítica del estudio.

3.2.3 Definición de la unidad analítica

Con base en el análisis exploratorio previo y en la identificación de variables comunes entre fuentes, se definió como unidad analítica el grupo de establecimientos de salud. Esta unidad se representa mediante la combinación de atributos geográficos e institucionales: provincia, cantón, parroquia, área, clase, tipo, entidad y sector.

Esta decisión fue necesaria porque las bases originales no disponen de un identificador único común que permita vincular de forma inequívoca cada egreso hospitalario con un establecimiento individual y su respectiva información de camas. La llave natural construida permite relacionar ambas fuentes a nivel de combinaciones observadas, pero no garantiza identificación hospitalaria individual.

Por ello, el análisis se plantea a nivel de grupos de establecimientos y no a nivel de hospital individual. Esta definición evita asumir una granularidad que los datos no respaldan y mantiene la coherencia metodológica para las etapas posteriores de integración, transformación y modelado predictivo.

A partir de esta unidad analítica, se agregan los egresos hospitalarios, la disponibilidad de camas y las demás variables explicativas necesarias para la construcción del dataset de modelado.

3.2.4. Control de calidad de entrada

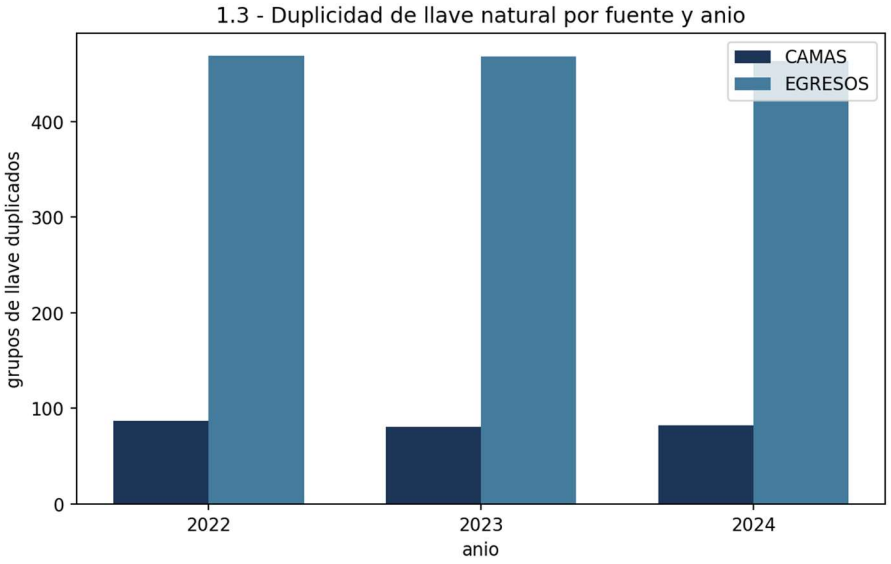
El control de calidad de entrada tuvo como propósito realizar una validación inicial sobre las fuentes CAMAS y EGRESOS antes de la integración final y de la construcción del dataset de modelado. En este punto aún no se ejecuta la unión analítica definitiva; se revisan condiciones de consistencia, completitud, duplicidad, validez temporal y calidad textual para asegurar que los insumos sean adecuados para las siguientes fases.

El primer aspecto evaluado fue la duplicidad de llaves por fuente y año. En CAMAS se observaron entre 81 y 87 grupos de llave duplicados por año, con máximos de repetición

entre 9 y 13 registros por grupo. En EGRESOS, en cambio, todos los grupos de llave aparecen repetidos en los tres años analizados, con máximos de repetición mucho más altos. Este comportamiento es esperable, ya que EGRESOS tiene naturaleza transaccional y registra múltiples eventos hospitalarios para una misma combinación de variables.

Figura 3-4

Duplicidad de llave natural por fuente y año

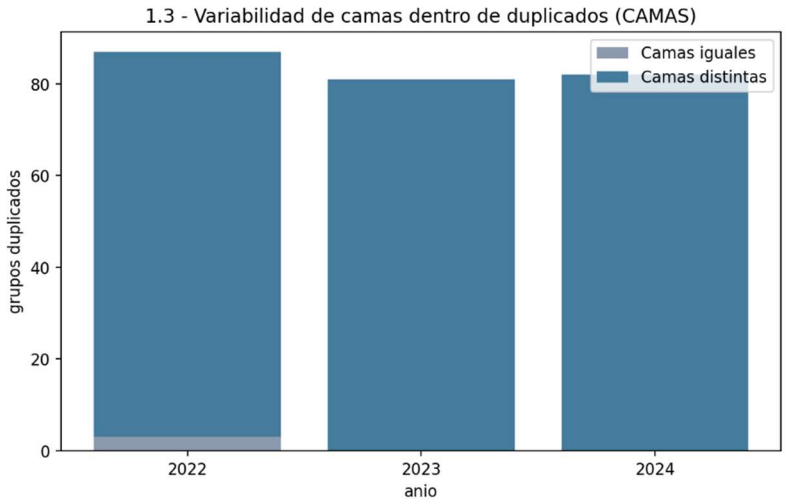


Este resultado confirma que el análisis posterior no debe realizarse mediante un emparejamiento fila a fila entre CAMAS y EGRESOS. Por el contrario, se requiere trabajar mediante agregación por unidad analítica, respetando la naturaleza de cada fuente.

En segundo lugar, se revisó la variabilidad dentro de los duplicados de CAMAS. El análisis mostró que algunos registros duplicados no corresponden únicamente a repeticiones simples, sino que pueden presentar diferencias en las camas reportadas dentro de un mismo grupo de llave. Este hallazgo justifica la aplicación de reglas de consolidación antes de integrar la información en la tabla analítica.

Figura 3-5

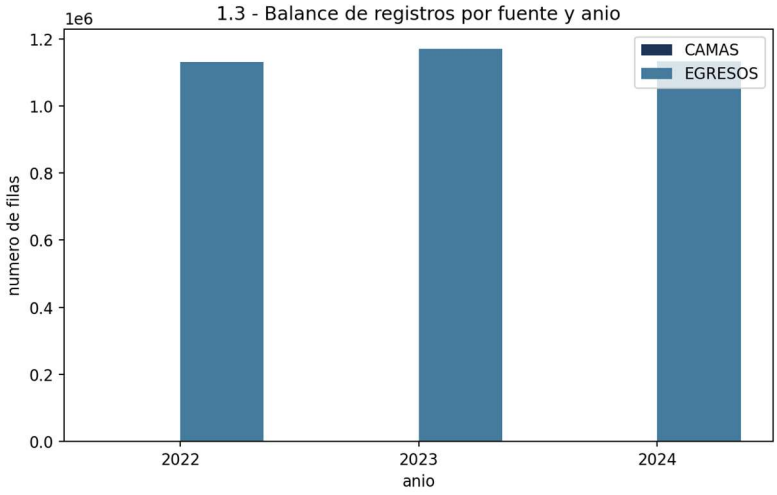
Variabilidad de camas dentro de duplicados



También se confirmó un desbalance natural de escala entre fuentes. EGRESOS concentra más de 1,13 millones de registros por año, mientras que CAMAS trabaja con una cantidad mucho menor de registros. Esta diferencia respalda que el análisis posterior deba realizarse mediante agregaciones y no mediante cruces directos a nivel de registro individual.

Figura 3-6

Balance de registros por fuente y año



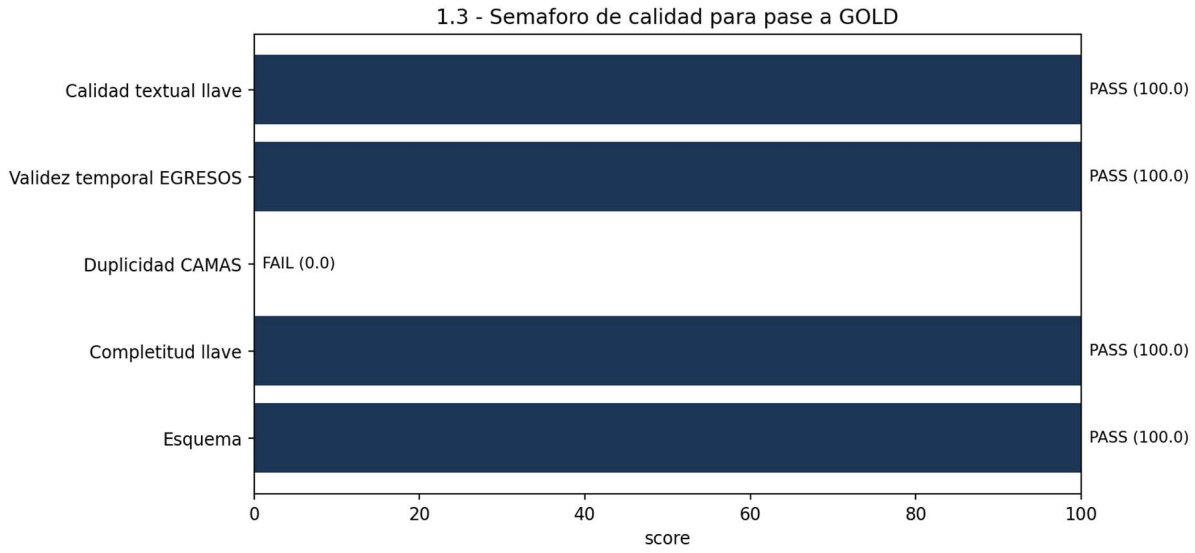
En cuanto a completitud, validez temporal y calidad textual, los resultados fueron satisfactorios. Las columnas críticas evaluadas alcanzaron 100 % de valores no nulos en

CAMAS y EGRESOS para el período 2022–2024. Asimismo, las variables año, mes y día de egreso presentaron 100 % de validez, y no se identificaron problemas relevantes en los textos de las columnas utilizadas para la llave natural, como valores en blanco, tokens nulos, caracteres de control o marcas de formato.

Finalmente, la semaforización global marcó resultado PASS en esquema, completitud de llave, validez temporal y calidad textual. El único resultado marcado como FAIL correspondió a la duplicidad en CAMAS. Sin embargo, esta duplicidad no invalida la fuente, ya que es esperada por la forma en que se registra la información de camas, donde pueden existir varios registros asociados a un mismo grupo, período o categoría de cama.

Figura 3-7

Semáforo de calidad para pase a GOLD



En síntesis, el principal hallazgo de esta etapa se concentra en la necesidad de consolidar registros duplicados en la fuente CAMAS antes de la integración. No se identificaron problemas relevantes de estructura, completitud, fechas ni calidad textual. Por tanto, las fuentes se consideran adecuadas para continuar con el proceso, siempre que se apliquen criterios de consolidación y depuración antes de construir la tabla analítica final.

3.3 Fase 3: Preparación de los datos

La preparación de datos se estructuró bajo un enfoque por capas inspirado en la arquitectura Bronze-Silver-Gold, también conocida como arquitectura medallón. Este patrón organiza el tratamiento de la información en capas progresivas: la capa Bronze conserva los datos originales obtenidos desde las fuentes primarias; la capa Silver contiene datos depurados, estandarizados e integrados; y la capa Gold reúne información consolidada y lista para análisis y modelado (Databricks, s. f.).

En esta investigación, el enfoque Bronze-Silver-Gold se aplica para transformar las fuentes de egresos hospitalarios y camas disponibles en una tabla analítica integrada. Esta organización permite mantener trazabilidad entre los datos originales, las reglas de limpieza y la construcción final del dataset de modelado.

3.3.1. Construcción de la tabla Gold bajo enfoque Bronze-Silver-Gold

3.3.1.1. Lectura de archivos, limpieza y construcción de llave natural

La construcción de la tabla Gold inicia con la lectura de los archivos originales, la limpieza básica de las fuentes y la construcción de la llave natural que permite relacionar egresos hospitalarios y camas disponibles.

En primer lugar, se leen los archivos CSV originales de camas hospitalarias y egresos hospitalarios correspondientes al período 2022–2024.

```
df_camas = pd.read_csv("data/raw/camas_hospitalarias_2023.csv", sep=";")

df_egresos = pd.read_csv("data/raw/egresos_hospitalarios_2023.csv", sep=";")
```

Posteriormente, se homogeneizan los nombres de columnas, se limpian los textos relevantes y se reemplazan valores nulos o equivalentes, como “NAN” o “SIN DATO”, por un valor estándar. Esta limpieza permite reducir diferencias de formato que podrían afectar la integración posterior.

```
df_camas.columns = [col.strip().upper() for col in df_camas.columns]
df_camas = df_camas.applymap(lambda x: str(x).strip().upper() if isinstance(x, str) else x)
df_camas.replace(NULL_LIKE_TOKENS, "SIN_DATO", inplace=True)
```

Finalmente, se construye la llave natural mediante la combinación de variables geográficas e institucionales comunes entre ambas fuentes. Esta llave permite emparejar registros de camas y egresos a nivel de grupos de establecimientos.

```
df_camas["llave_natural"] = (
    df_camas["PROV_UBI"] + "|" +
    df_camas["CANT_UBI"] + "|" +
    df_camas["PARR_UBI"] + "|" +
    df_camas["AREA_UBI"] + "|" +
    df_camas["CLASE"] + "|" +
    df_camas["TIPO"] + "|" +
    df_camas["ENTIDAD"] + "|" +
    df_camas["SECTOR"]
)
```

3.3.1.2. Creación de dimensiones

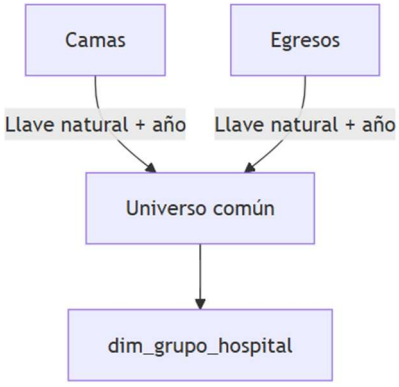
Como parte de la preparación de los datos, se construyeron dimensiones analíticas que permiten organizar la información de forma más clara y reducir la complejidad de las fuentes originales. Estas dimensiones facilitan la integración, agregación e interpretación posterior de los datos.

Dimensión dim_grupo_hospital: La dimensión dim_grupo_hospital se construye a partir de la llave natural definida previamente. Su propósito es asignar un identificador único, denominado grupo_hospital_id, a cada combinación de atributos geográficos e institucionales comunes. Esta dimensión permite trabajar con la unidad analítica definida para el estudio: grupos de establecimientos de salud.

La construcción de esta dimensión se realiza cruzando las llaves naturales de CAMAS y EGRESOS por año. Solo se conservan las combinaciones que existen en ambas fuentes, lo que garantiza que el análisis posterior se base en información comparable.

Figura 3-8

Creación de la dimensión dim_grupo_hospital



El diagrama resume la lógica de construcción de la dimensión de grupos hospitalarios: ambas fuentes se proyectan sobre la llave natural más el año, se conserva el universo común y de él se deriva dim_grupo_hospital con un identificador propio. Esta dimensión es la columna vertebral del modelo de datos, pues garantiza que egresos y camas se refieran siempre al mismo grupo de establecimientos.

Desde el punto de vista técnico, la lógica general fue la siguiente:

```
# Cruce de llaves naturales y filtrado de universo común
llaves = ["anio", "llave_natural"]
df_camas_matched = df_camas.merge(df_egresos[llaves], on=llaves, how="inner")
df_egresos_matched = df_egresos.merge(df_camas[llaves], on=llaves, how="inner")
```

```
# Construcción de la dimensión de grupo hospitalario
dim_grupo_hospital =
df_camas_matched[["llave_natural"]].drop_duplicates().reset_index(drop=True)
dim_grupo_hospital["grupo_hospital_id"] = dim_grupo_hospital.index + 1
```

Ejemplo de código Python de esta fase

Tabla 3-9

Ejemplo de la Dimensión de grupo hospitalario

grupo_hospital_id	natural_key	prov_ubi	cant_ubi	parr_ubi	area_ubi	clase	tipo	entidad	sector
1	GUAYAS||GU	GUAYAS	GUAYAQUIL	AYACUCHO	URBANA	GINECO-OBST	AGUDO	PRIVADOS SIN FIN	PRIVADO SIN FINES DE LUCRO
2	SANTO DOMI	SANTO DOMI	SANTO DOMI	CHIGUILPE	URBANA	HOSPITAL BÃ	SIN TIPO HOS	PRIVADOS CON FI	PRIVADO CON FINES DE LUCRO
3	PICHINCHA||	PICHINCHA	QUITO	LA CONCEPCI	URBANA	HOSPITAL BÃ	SIN TIPO HOS	PRIVADOS CON FI	PRIVADO CON FINES DE LUCRO
4	PICHINCHA||	PICHINCHA	QUITO	RUMIPAMBA	URBANA	HOSPITAL GE	AGUDO	PRIVADOS CON FI	PRIVADO CON FINES DE LUCRO
5	GUAYAS||PE	GUAYAS	PEDRO CARB	SABANILLA	RURAL	ESTABLECIMI	AGUDO	PRIVADOS SIN FIN	PRIVADO SIN FINES DE LUCRO
6	GUAYAS||DA	GUAYAS	DAULE	BANIFE	URBANA	CLÃNICA GEN	CLÃNICAS GE	PRIVADOS SIN FIN	PRIVADO SIN FINES DE LUCRO
7	MORONA SAN	MORONA SAN	MORONA	MACAS	URBANA	HOSPITAL BÃ	SIN TIPO HOS	INSTITUTO ECUAT	PÃšBLICO

Esta dimensión es clave porque evita trabajar a nivel de hospital individual cuando las fuentes no permiten identificarlo de forma inequívoca. Además, facilita la agregación de egresos, camas y variables explicativas para el modelado predictivo.

Dimensión dim_area_cama: La dimensión dim_area_cama permite agrupar y describir las áreas funcionales de las camas hospitalarias. Su objetivo es reducir la cantidad de columnas originales asociadas a especialidades o servicios y convertirlas en grupos funcionales más interpretables.

En los datos originales existen 44 códigos individuales de áreas de cama, representados por columnas específicas, como dotpedi, dotneon o dotcardipe. Durante la preparación, estos códigos se consolidan en 16 grupos funcionales, como pediátrico, quirúrgico, rehabilitación, medicina o salud mental.

Tabla 3-10

Ejemplo de agrupación por grupo funcional

Código original	Grupo funcional
pedi	pediatrico
neon	pediatrico
cardipe	pediatrico
ciruped	pediatrico

Esta agrupación reduce la dimensionalidad de la información y facilita el análisis posterior, ya que el modelo trabaja con variables agregadas y más estables en lugar de múltiples columnas específicas de baja trazabilidad.

De forma resumida, la definición de grupos funcionales se implementó mediante un diccionario de equivalencias:

```
AREA_GROUPS_DEF_RAW = {
    "ginecologia_obstetricia": ["gino", "obste"],
    "pediatrico": ["pedi", "neon", "cardipe", "ciruped"],
```

Ejemplo de código Python de definición de grupos de áreas

Posteriormente, cada grupo de área recibe un identificador único, un código y un nombre descriptivo.

```
dim_area_cama = pd.DataFrame([
    {
        "area_cama_id": i + 1,
        "codigo_area": group_name,
        "nombre_area": group_name.replace("_", " ").title(),
    }
```

```
    for i, group_name in enumerate(AREA_GROUPS_DEF.keys())
])
```

Ejemplo de código Python de creación de la dimensión dim_area_cama

Tabla 3-11

Ejemplo de la dimensión dim_area_cama

area_cama_id	codigo_area	nombre_area
1	medicina	Medicina
2	cardiologia	Cardiologia
3	ginecologia_obstetricia	Ginecologia Obstetricia
4	pediatrico	Pediatrico
5	salud_mental	Salud Mental

Dimensión dim_causa

La dimensión dim_causa permite agrupar y estandarizar las causas de egreso hospitalario según los grandes capítulos de la Clasificación Internacional de Enfermedades CIE-10. Esta dimensión facilita el análisis agregado de causas de egreso y permite incorporar información clínica de forma más ordenada en el proceso analítico.

Para su construcción, se extraen las causas originales reportadas en los egresos hospitalarios, se limpian los valores y se clasifican en grupos funcionales basados en capítulos CIE-10. En este caso, se pasó de 221 códigos individuales a 22 grupos funcionales, lo que simplifica el análisis y mejora la interpretación de los resultados.

Tabla 3-12

Ejemplo de agrupación por grupo funcional (grupo_causa)

Código original	**Grupo funcional (grupo_causa)**
A15	a00_b99 (Enfermedades infecciosas y parasitarias)
C34	c00_d48 (Neoplasias)
J18	j00_j99 (Enfermedades del sistema respiratorio)

La lógica general de construcción fue la siguiente:

```
    dim_causa = (
        df_egresos_silver[["cau221rx"]]
        .dropna(subset=["cau221rx"])
        .drop_duplicates()
        .copy()
    )
    dim_causa["grupo_causa"] = dim_causa["cau221rx"].apply(clasificar_grupo_causa_cie10)
    dim_causa = dim_causa[dim_causa["grupo_causa"].notna()].reset_index(drop=True)
else:
    dim_causa = pd.DataFrame(columns=["cau221rx", "grupo_causa"])
```

Esta dimensión permite reducir la complejidad de las causas originales y trabajar con categorías clínicas más interpretables para el análisis descriptivo, la priorización operativa y la construcción de variables relacionadas con el perfil histórico de la demanda.

3.3.1.3 Construcción de la capa Silver mediante tablas de hechos

Las tablas de hechos se construyen como estructuras intermedias de la capa Silver, con el propósito de resumir y organizar los eventos registrados en las fuentes originales. Estas tablas capturan medidas cuantitativas asociadas a un nivel específico de análisis, como egresos hospitalarios, días de estadía o disponibilidad de camas, y permiten consolidar la información antes de integrarla en la tabla Gold.

En este estudio se construyeron tres tablas de hechos principales:

fact_egresos_hospital_mes_dia.

Contiene el conteo de egresos hospitalarios por grupo hospitalario, año, mes y día. Además,

incorpora agregados como días de estadía, distribución por sexo, edad promedio y otros indicadores demográficos relevantes.

```
fact_egresos_hospital_mes_dia =
df_egresos_validos.groupby(group_cols).size().reset_index(name="egresos_totales")
```

fact_camas_hospital_anio_resumen.

Resume la dotación y disponibilidad de camas por grupo hospitalario, año y área funcional. Esta tabla permite consolidar la capacidad hospitalaria que posteriormente se relaciona con los egresos.

```
fact_camas_hospital_anio_resumen = df_camas_silver.copy()
fact_camas_hospital_anio_resumen["anio"] =
fact_camas_hospital_anio_resumen["anio_archivo"]
# ... agregación y suma de columnas de camas ...
```

fact_egresos_hospital_mes_causa.

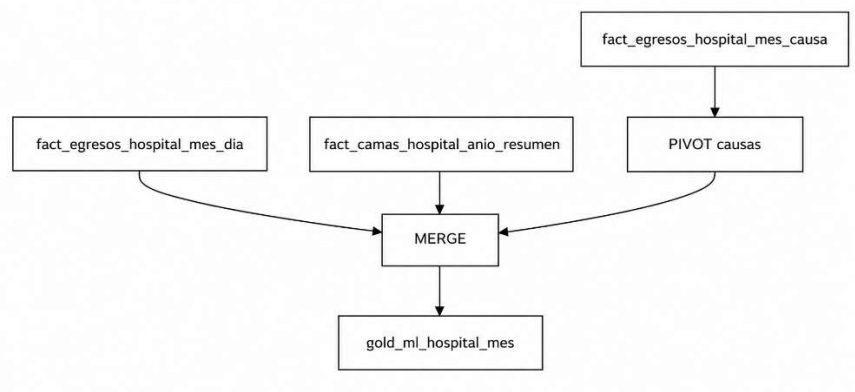
Contiene el conteo de egresos por grupo hospitalario, período y grupo de causa CIE-10. Esta información se transforma posteriormente mediante pivoteo para incorporar las causas como columnas agregadas en la tabla Gold.

```
fact_egresos_hospital_mes_causa = (
    df_egresos_mes_con_causa
    .groupby(group_cols + ["grupo_causa"], as_index=False)
    .size()
    .rename(columns={"size": "egresos_causa"})
)
```

A partir de estas tablas, se integra la información de actividad hospitalaria, capacidad disponible y causas de egreso. Primero, se combinan los egresos diarios con el resumen anual de camas. Luego, los egresos por causa se transforman para que cada grupo CIE-10 quede representado como una variable agregada. Finalmente, todos los insumos se consolidan en la tabla Gold, denominada gold_ml_hospital_mes_dia, que constituye la base final para el análisis y el modelado predictivo.

Figura 3-9

Flujo de integración para construir la tabla Gold



El flujo muestra cómo se integra la tabla Gold: los hechos de egresos diarios, el resumen anual de camas y la matriz de causas pivotada convergen en una única operación de unión que produce gold_ml_hospital_mes. Esta consolidación permite que las fases posteriores trabajen sobre una sola tabla trazable en lugar de múltiples fuentes intermedias.

3.3.1.4. Reducción y agrupación de columnas de disponibilidad

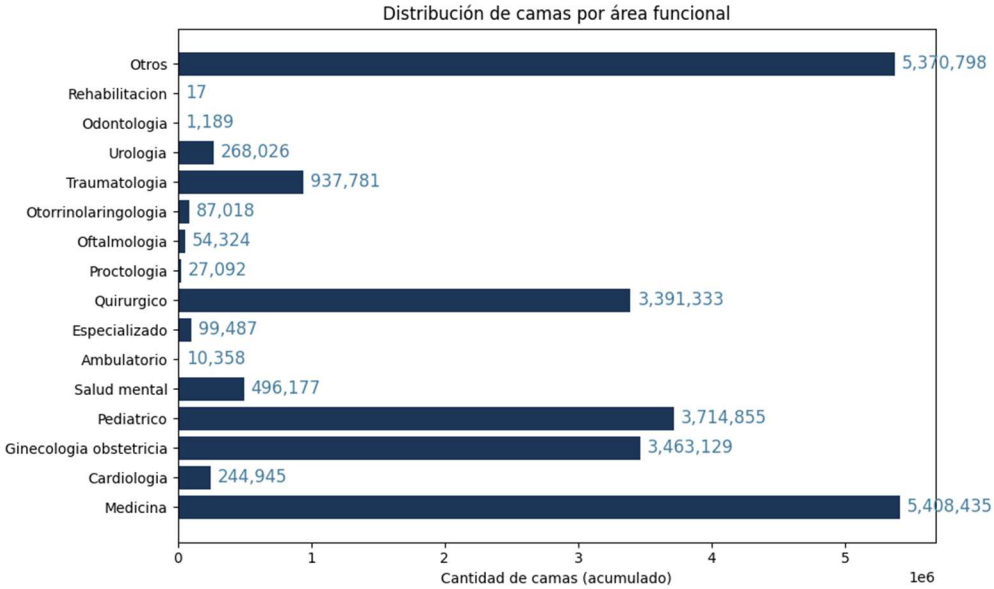
En los datos originales existen múltiples columnas específicas para cada área funcional de cama: 53 columnas de disponibilidad, identificadas con el prefijo dis*, y 63 columnas de dotación, identificadas con el prefijo dot*. Estas columnas están asociadas a 44 códigos individuales de áreas de cama.

Para simplificar el análisis, estos códigos se consolidaron en 16 grupos funcionales definidos en la dimensión dim_area_cama, como pediátrico, quirúrgico, rehabilitación, medicina y

salud mental. De esta forma, el modelo pasa de manejar más de 100 columnas individuales a un conjunto reducido de variables agregadas por área funcional, como camas_disponibles_*. Esta reducción mejora la trazabilidad, facilita la interpretación de la capacidad hospitalaria y evita trabajar con columnas demasiado específicas que podrían dificultar el análisis y el modelado.

Figura 3-10

Distribución de camas por área functional



La distribución revela que la capacidad instalada se concentra en pocas áreas funcionales: medicina, el grupo de otras áreas, pediatría, ginecología-obstetricia y el área quirúrgica acumulan la gran mayoría de camas, mientras especialidades como rehabilitación u odontología son marginales. Esta concentración respalda la agrupación funcional utilizada, que reduce dimensionalidad sin perder las áreas operativamente dominantes.

3.3.1.5. Agrupación de causas CIE-10

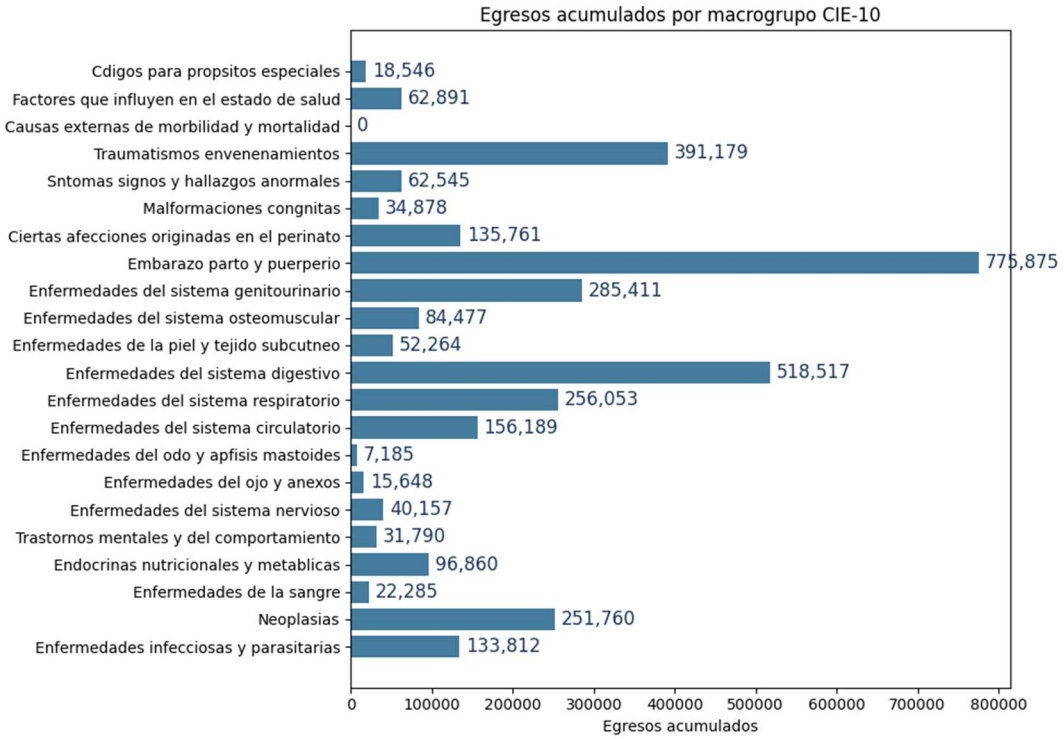
En los datos originales se registran 221 causas individuales de egreso hospitalario. Para facilitar su uso analítico, estas causas se agruparon en 22 macrogrupos funcionales definidos por los capítulos de la CIE-10, según la dimensión dim_causa.

Este proceso permite transformar una variable clínica muy detallada en categorías más estables e interpretables, como enfermedades infecciosas, neoplasias o enfermedades del sistema respiratorio. Además, se generan columnas agregadas de egresos por causa y columnas porcentuales que representan la participación de cada macrogrupo dentro del total de egresos.

De esta manera, el dataset final mantiene información clínica relevante, pero en una estructura más simple y comparable entre grupos hospitalarios y períodos.

Figura 3-11

Agrupación y consolidación de causas hospitalarias CIE-10



La consolidación por macrogrupos CIE-10 muestra que embarazo, parto y puerperio es la categoría con más egresos acumulados, seguida por enfermedades del sistema digestivo y traumatismos y envenenamientos. La figura valida la reducción de más de doscientas causas específicas a veintidós macrogrupos estándar, manteniendo la interpretación clínica de los principales generadores de demanda.

3.3.1.6. Agregación de camas

La agregación de camas se realizó por grupo hospitalario, año y área funcional. En este proceso se consolidaron tanto las camas de dotación como las camas disponibles para cada uno de los 16 grupos funcionales definidos previamente.

Además, se preservaron variables base y se generaron campos derivados para control de consistencia y trazabilidad:

- total_camas_disponibles_original: valor original reportado de camas disponibles.
- total_camas_disponibles_ajustada: valor ajustado para evitar que las camas disponibles excedan la dotación registrada.
- flag_camas_disp_mayor_dotacion: indicador que identifica casos donde las camas disponibles superan la dotación.
- total_camas_disponibles_incluye_especiales: suma de camas base y camas especiales.

Esta estructura permite mantener control sobre los valores originales y ajustados, facilitando auditorías y análisis posteriores.

3.3.1.7. Agregación de egresos

Los egresos hospitalarios se normalizaron y agregaron por grupo hospitalario, año, mes y día. Durante este proceso se calcularon indicadores como total de egresos, días de estadía, distribución por sexo, edad promedio y proporciones demográficas relevantes.

Adicionalmente, las causas de egreso se agruparon en 22 macrogrupos CIE-10, generando columnas de egresos por causa y columnas porcentuales asociadas. Estos agregados constituyen insumos directos para la tabla Gold y para la posterior construcción del dataset de modelado.

3.3.1.8. Integración final y almacenamiento en Parquet

En esta etapa se integraron las tablas de hechos de egresos, camas y causas mediante uniones por grupo_hospital_id, año y, cuando corresponde, área funcional. Además, se incorporaron las dimensiones de grupo hospitalario, área de cama y causa de egreso para enriquecer el dataset con descripciones y clasificaciones útiles para el análisis.

El resultado final se ordenó y se almacenó en formato Parquet, debido a su eficiencia en almacenamiento, compresión y lectura para procesos analíticos. También se generó una versión en CSV para facilitar la revisión externa y el consumo por otras herramientas.

3.3.1.9. Reglas de consistencia y trazabilidad

Durante la construcción de la tabla Gold se mantuvo como unidad analítica central el grupo hospitalario, evitando asumir una identificación individual de hospitales que no está respaldada por las fuentes originales.

Asimismo, se preservaron tanto los valores originales como las versiones ajustadas de variables críticas, especialmente las relacionadas con camas disponibles. Esto permite mantener trazabilidad sobre las transformaciones aplicadas y facilita la auditoría del proceso.

Las columnas relacionadas con causas de egreso, tanto en conteo como en porcentaje, permanecen disponibles en el dataset final. Sin embargo, su uso como variables predictoras debe considerar la prevención de fuga de información, utilizando únicamente información disponible antes del período que se desea predecir.

3.3.1.10. Resultados finales de la tabla Gold

El proceso de integración produjo una tabla Gold consolidada con los siguientes resultados:

Filas Gold: 381.864 registros consolidados.
Columnas Gold: 103 variables finales, incluyendo métricas, indicadores y dimensiones relevantes.
Grupos hospitalarios únicos: 591 unidades analíticas distintas.

Estos resultados reflejan la consolidación estructurada de la información hospitalaria y dejan disponible una base analítica para el análisis descriptivo, la construcción del dataset de modelado y la evaluación predictiva.

3.3.2 Análisis exploratorio (EDA) de datos sobre la tabla Gold

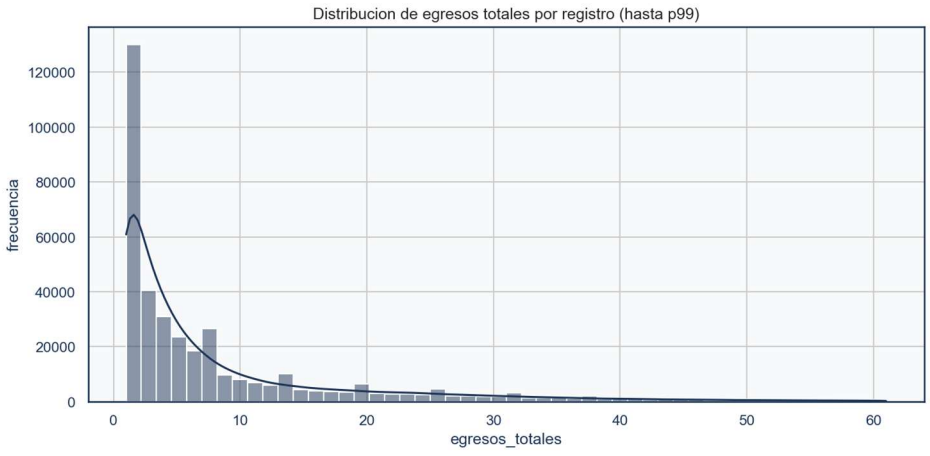
El análisis exploratorio de datos (EDA) sobre la tabla Gold permitió validar la calidad de la integración, identificar patrones relevantes y justificar decisiones posteriores de modelado. En esta etapa se analizaron la distribución de egresos, la evolución temporal, la concentración por grupos de establecimientos, la relación entre demanda y capacidad, la reducción de variables, las causas de egreso, las variables demográficas y la presencia de valores atípicos.

3.3.2.1. Distribución y volumen de egresos hospitalarios

Se analizó la distribución total de egresos hospitalarios para identificar la dispersión de la demanda y posibles concentraciones extremas en determinados grupos de establecimientos.

Figura 3-12

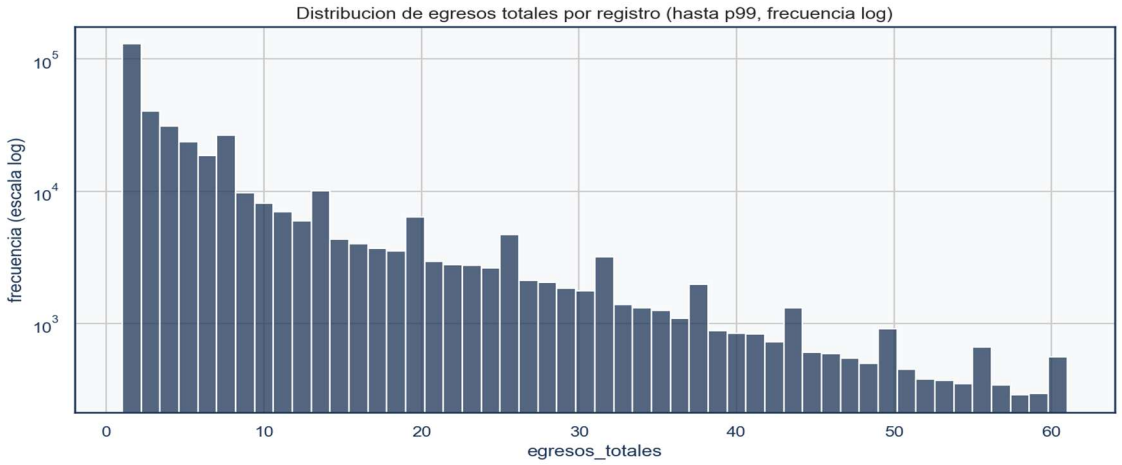
Distribución de egresos hospitalarios.



La distribución de egresos es marcadamente asimétrica: la mayoría de grupos hospitalarios registra volúmenes bajos y una cola larga concentra los volúmenes altos. Este comportamiento anticipa la necesidad de escalas transformadas en el análisis y de escenarios con y sin valores extremos en el modelado.

Figura 3-13

Distribución de egresos hospitalarios en escala logarítmica



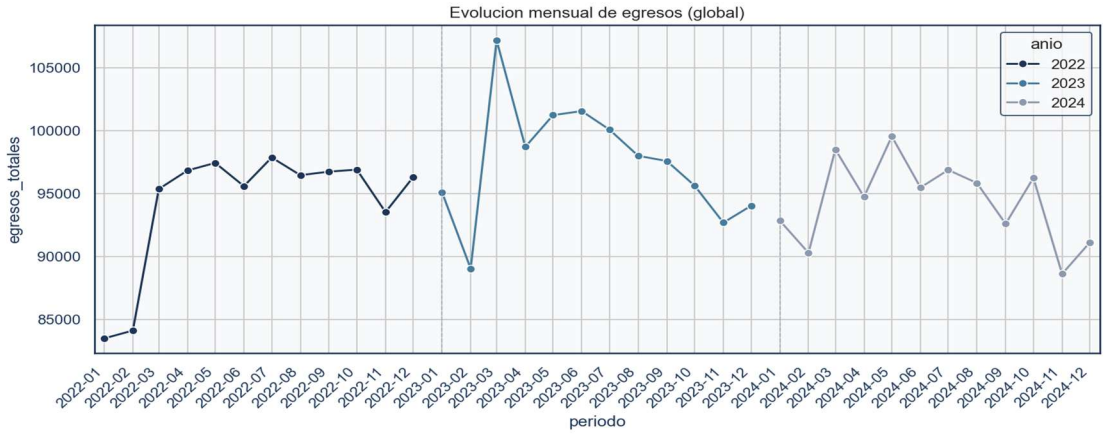
Los histogramas muestran una alta concentración de grupos con bajo volumen de egresos y una cola larga de grupos con mayor actividad. Este comportamiento evidencia una distribución asimétrica, con presencia de valores extremos. Por ello, el uso de escalas logarítmicas resulta útil para visualizar mejor la variabilidad de la demanda hospitalaria.

3.3.2.2. Evolución temporal y estacionalidad

La evolución mensual global de los egresos permite observar tendencias, posibles patrones estacionales y cambios en el comportamiento de la demanda durante el período 2022–2024, como se observa en la siguiente figura.

Figura 3-14

Evolución mensual global de egresos



La serie mensual global muestra un nivel de actividad relativamente estable con oscilaciones estacionales reconocibles y sin quiebres abruptos en el período 2022-2024. Esta regularidad respalda el uso de rezagos semanales y ventanas móviles como señales predictivas.

Figura 3-15

Heatmap de egresos por año y mes

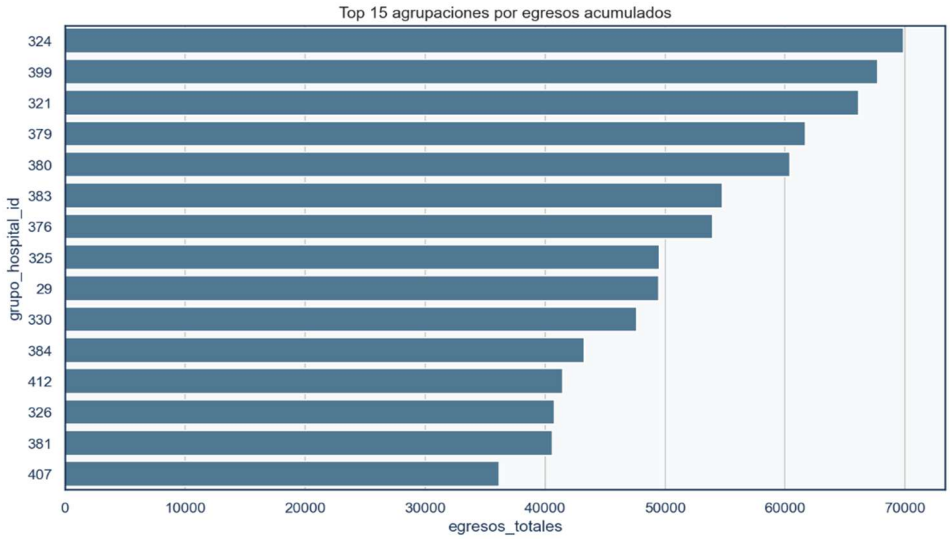
En el mapa de calor de egresos por año y mes, los resultados muestran variaciones mensuales en la demanda hospitalaria. Se observan meses con mayor concentración de egresos y otros con menor actividad. En particular, mayo aparece como uno de los meses con mayor recurrencia de egresos en los años analizados, mientras que enero y febrero presentan menores volúmenes relativos. Estos patrones justifican la incorporación de variables temporales en el dataset de modelado.

3.3.2.3. Cobertura y concentración por grupo de establecimientos

Se identificaron los grupos de establecimientos con mayor volumen de egresos y su comportamiento anual. Este análisis permite reconocer la concentración de la demanda y la heterogeneidad existente entre unidades analíticas. En la siguiente imagen se muestra las agrupaciones de los hospitales con su volumen de egresos acumulado.

Figura 3-16

Principales grupos de establecimientos por volumen de egresos



Un número reducido de grupos de establecimientos concentra una parte sustancial de los egresos nacionales, mientras la mayoría opera con volúmenes menores. Esta heterogeneidad justifica el uso de una métrica ponderada por volumen (WMAPE) y de una priorización relativa en lugar de umbrales únicos.

Una visualización de la información de los diferentes grupos a nivel anual, permite en primer lugar comprender las diferencias entre egresos por año, y adicionalmente ayuda a determinar la consistencia entre los diferentes grupos de hospitales, para determinar el nivel de homogeneidad y las posibles diferencias que pueden existir para tomar decisiones en como tratar los datos. De esta forma en la figura a continuación se muestra para los tres años de estudio 2022, 2023 y 2024, el top 10 de agrupaciones de hospitales con la cantidad de egresos.

Figura 3-17

Principales grupos de establecimientos por año

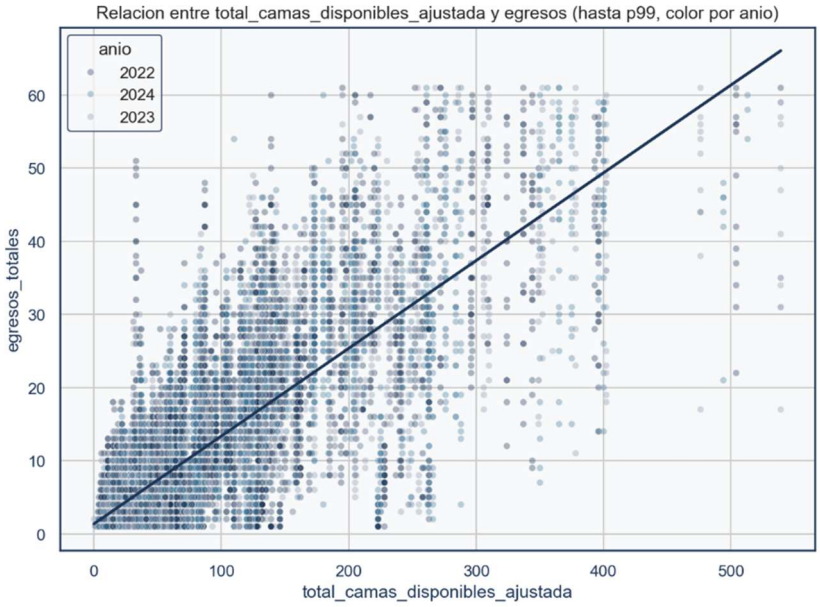
Los resultados de los tres gráficos, muestran que un conjunto reducido de grupos concentra una parte importante de los egresos, mientras que la mayoría presenta volúmenes menores. Esta concentración confirma que la red hospitalaria no es homogénea y que el modelado debe considerar diferencias estructurales entre grupos de establecimientos.

3.3.2.4. Relación entre egresos y camas disponibles

Se analizó la relación entre el total de camas disponibles ajustadas y el volumen de egresos hospitalarios mediante un gráfico de dispersión diferenciado por año. Para facilitar la interpretación y reducir el efecto de valores extremos, la visualización se limitó hasta el percentil 99 de ambas variables.

Figura 3-18

Dispersión de egresos frente a camas disponibles ajustadas

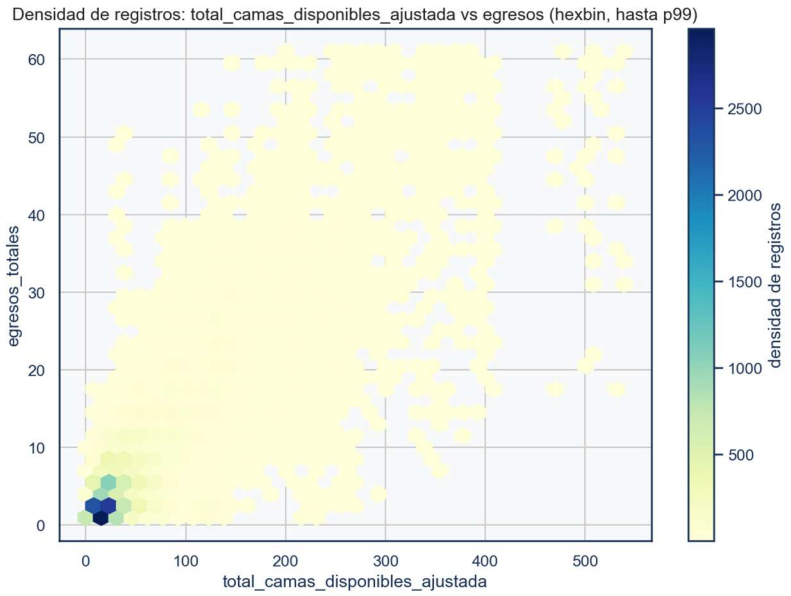


Se observa una relación positiva entre la disponibilidad de camas y los egresos; sin embargo, la amplia dispersión indica que grupos con capacidades similares pueden registrar niveles de demanda muy diferentes. Este comportamiento, consistente entre los años analizados, sustenta el uso de un indicador relativo que vincule la demanda hospitalaria con la capacidad disponible.

A continuación, se analizó la relación existente entre egresos versus camas disponibles, teniendo en cuenta que la información viene de dos archivos diferentes, los cuales se unificaron mediante la llave natural definida. De la misma forma, se visualizó este comportamiento utilizando una escala logarítmica. Lo cual se observa en los siguientes gráficos.

Figura 3-19

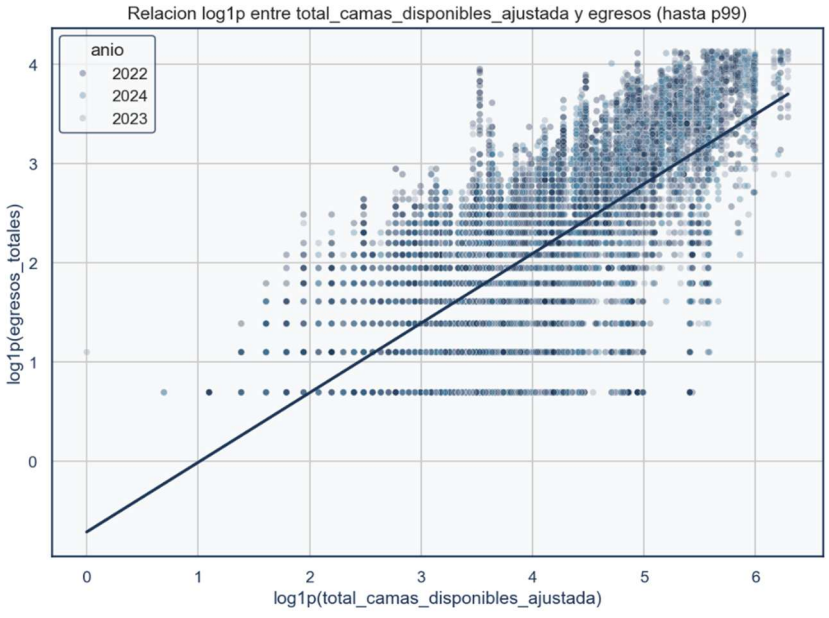
Hexbin de egresos frente a camas disponibles



La visualización hexbin confirma que la densidad de observaciones se concentra en combinaciones de baja capacidad y bajo volumen, con casos progresivamente más dispersos hacia valores altos. El patrón sugiere que el comportamiento típico del sistema se define en los grupos pequeños y medianos.

Figura 3-20

Dispersión logarítmica de egresos frente a camas disponibles



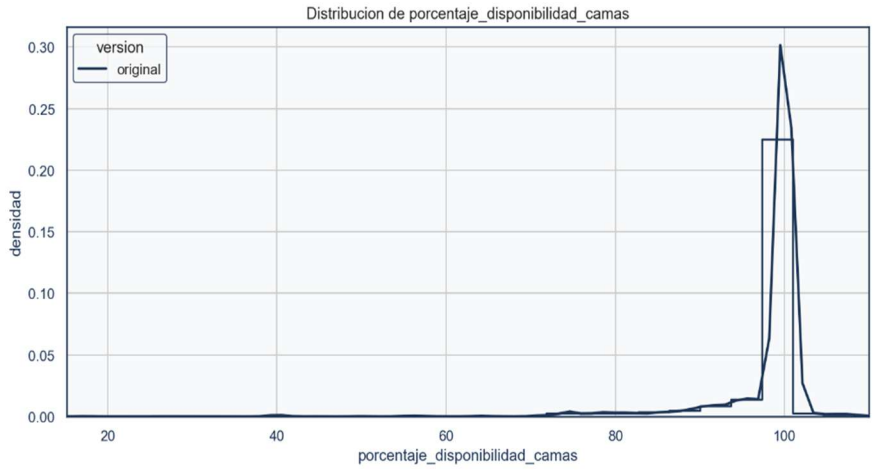
La curva acumulada evidencia una fuerte concentración de registros en torno al 100 % de disponibilidad de camas, mientras que una proporción menor presenta valores inferiores o superiores a este nivel. Este comportamiento permite identificar la concentración de la variable y posibles valores que requieren revisión o tratamiento durante la preparación de los datos.

3.3.2.5. Calidad y reducción de variables de camas

Se evaluó la consistencia de las variables asociadas a la capacidad hospitalaria mediante la comparación entre la dotación total de camas y la disponibilidad reportada. El propósito fue corregir registros inconsistentes, reducir la duplicidad de información y seleccionar la variable que representara mejor la capacidad operativa de cada grupo hospitalario.

Figura 3-21

Comparación entre disponibilidad original y ajustada



La comparación permitió identificar registros en los que las camas disponibles superaban la dotación total, situación que no es posible en la práctica, pues un establecimiento no puede disponer de más camas de las que posee. Para corregir estos casos, la disponibilidad ajustada se limitó al valor máximo de la dotación reportada.

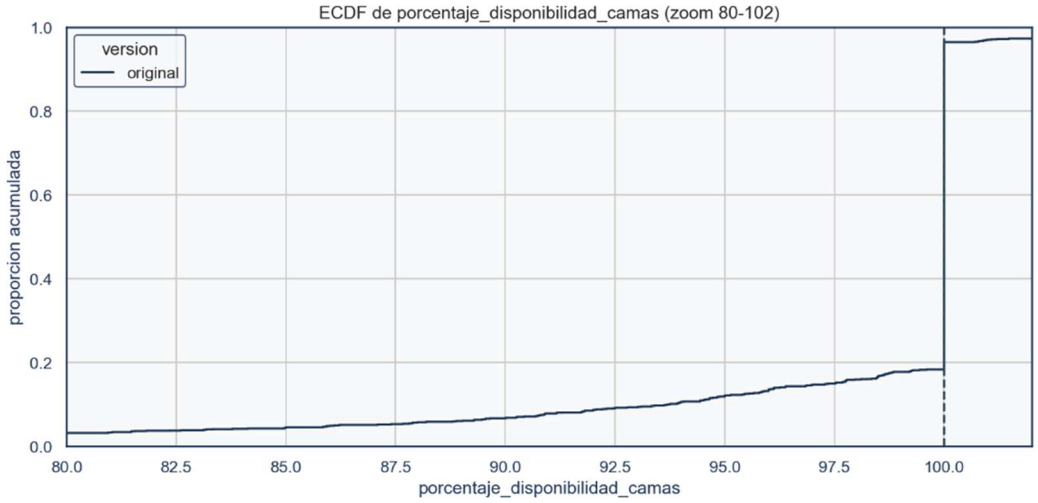
Para los análisis posteriores se seleccionó la disponibilidad ajustada como principal medida de capacidad operativa, ya que representa las camas que efectivamente podían utilizarse. La dotación se mantuvo como referencia para validar los datos y calcular el porcentaje de

disponibilidad, pero no se incorporó simultáneamente como una variable principal para evitar información redundante y posiblemente inconsistente.

Una vez seleccionada la disponibilidad como medida de la capacidad operativa, se analizó el porcentaje que representa con respecto al total de camas registrado. Para ello se utilizó una función de distribución acumulada empírica (ECDF), que permite conocer la proporción de registros con un porcentaje de disponibilidad igual o inferior a cada valor observado.

Figura 3-22

Distribución acumulada de camas disponibles



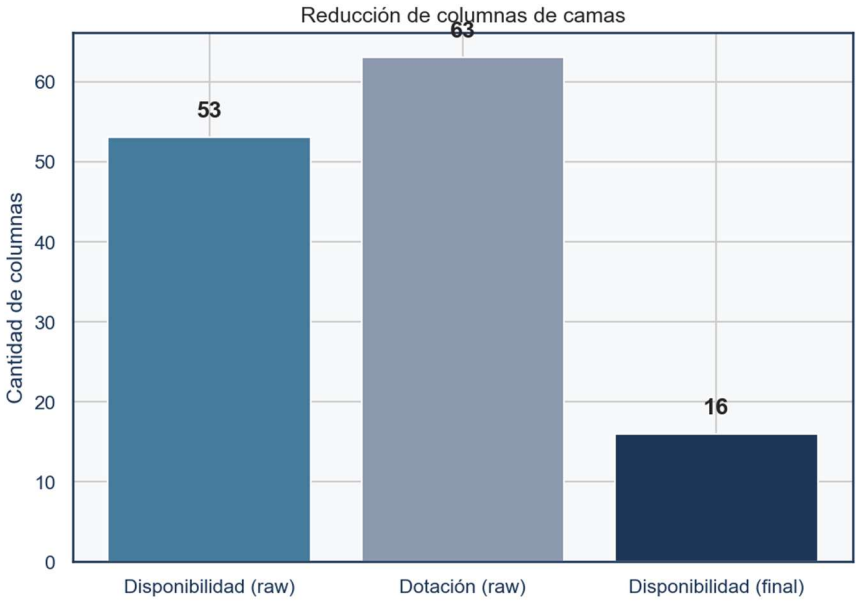
La gráfica muestra que la mayor parte de los registros se concentra en el 100 % de disponibilidad. En términos sencillos, si se consideraran 100 registros, aproximadamente 18 tendrían menos del 100 % de camas disponibles, cerca de 79 reportarían exactamente el 100 % y alrededor de 3 superarían ese valor. Esto significa que, en la mayoría de los casos, se informó que todas las camas registradas estaban disponibles, mientras que fueron pocos los casos con una disponibilidad menor o superior al total esperado.

Después de revisar las variables de dotación y disponibilidad, se simplificó el conjunto de columnas relacionadas con camas. El propósito fue conservar las variables con mayor

utilidad para el análisis y reducir aquellas que contenían información repetida o demasiado específica.

Figura 3-23

Reducción de columnas de camas



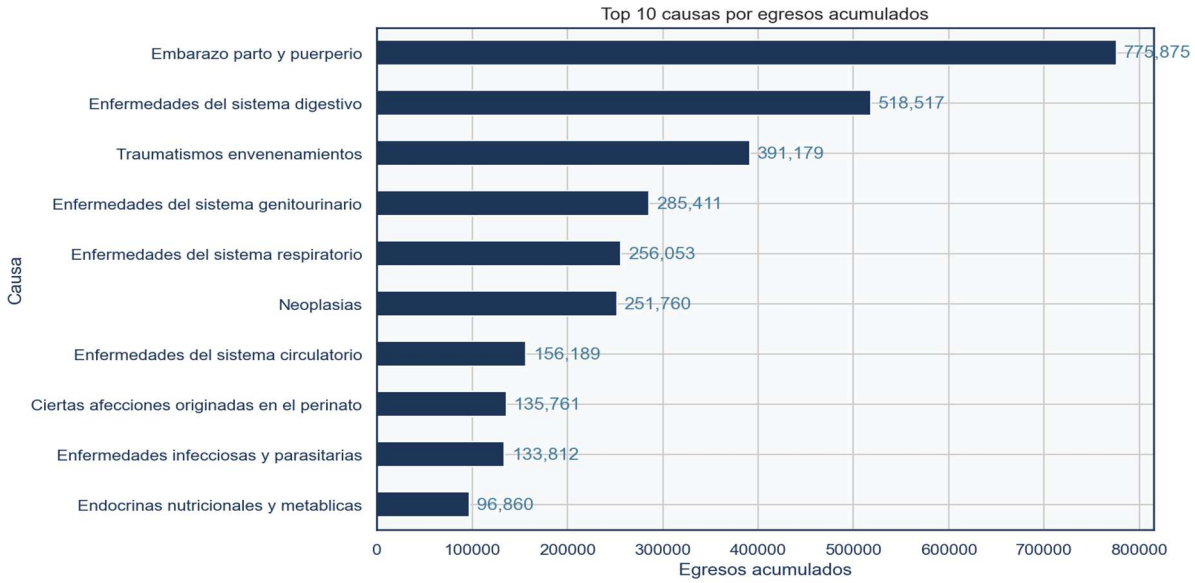
La reducción permitió pasar de múltiples columnas originales a un conjunto más claro y manejable de variables. Se priorizó la disponibilidad de camas por representar mejor la capacidad operativa, mientras que las columnas redundantes se eliminaron o agruparon en categorías funcionales, facilitando la interpretación y el procesamiento posterior de los datos.

3.3.2.6. Causas de egreso y agrupación CIE-10

Se analizaron las principales causas de egreso hospitalario y su agrupación en macrogrupos definidos según capítulos CIE-10. Este análisis permite simplificar la información clínica sin perder capacidad interpretativa.

Figura 3-24

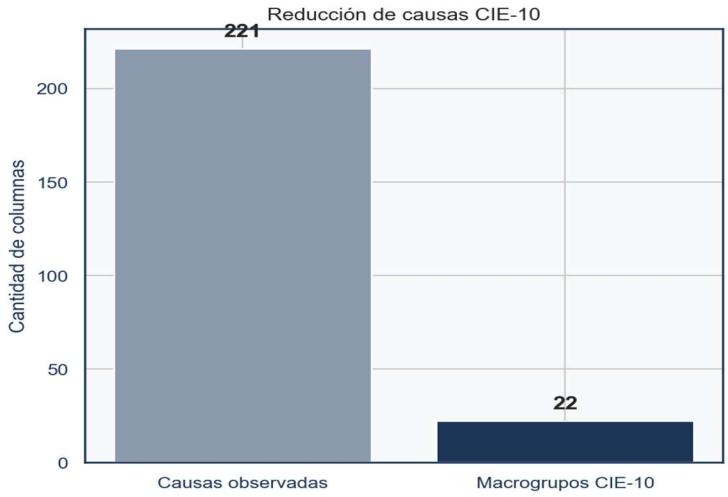
Principales causas de egreso



Las causas más frecuentes de egreso corresponden a embarazo y parto, afecciones digestivas y respiratorias, consistentes con el perfil hospitalario nacional. Como se verá en el análisis operativo, la frecuencia de una causa no coincide necesariamente con su asociación a alertas de presión.

Figura 3-25

Reducción de causas a macrogrupos CIE-10



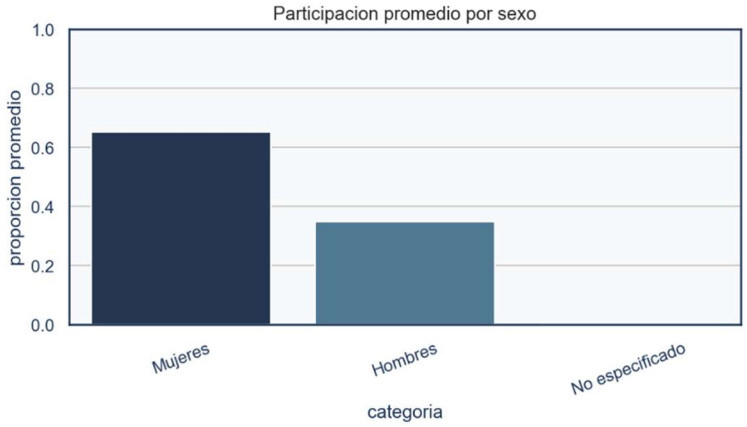
Los resultados muestran que algunas causas concentran una proporción importante de los egresos. La agrupación en macrogrupos CIE-10 permite trabajar con categorías más estables, comparables e interpretables, facilitando su uso en análisis descriptivo y modelado.

3.3.2.7. Variables demográficas y días de estadía

Se analizaron variables demográficas y operativas complementarias, como participación por sexo, edad, días promedio de estadía y correlación entre variables numéricas clave.

Figura 3-26

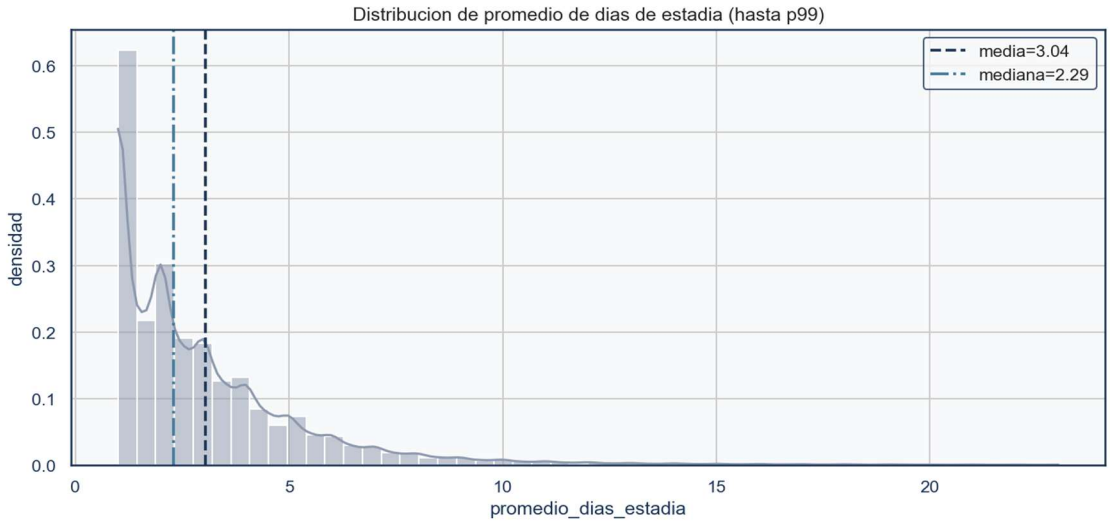
Participación promedio por sexo



La participación promedio por sexo muestra un predominio de mujeres en los egresos, explicado en gran medida por el peso del macrogrupo de embarazo, parto y puerperio. La variable se incorporó al dataset como parte del bloque demográfico.

Figura 3-27

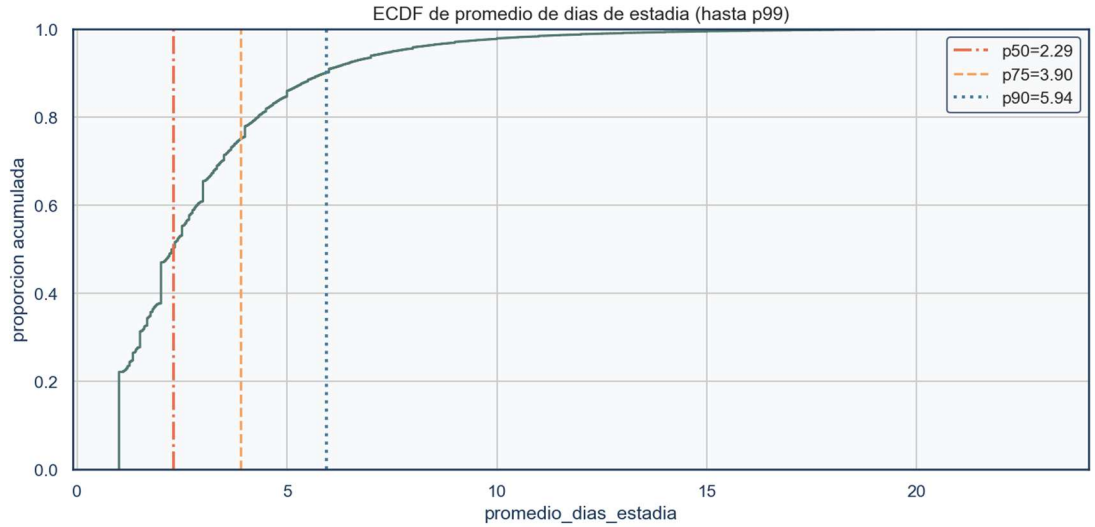
Distribución del promedio de días de estadía



El promedio de días de estadía se concentra en valores bajos, con media cercana a 3.0 días y mediana de 2.3, y una cola larga de estadías prolongadas. La estadía aporta señal sobre la rotación de camas y por ello integra el conjunto de variables del modelado.

Figura 3-28

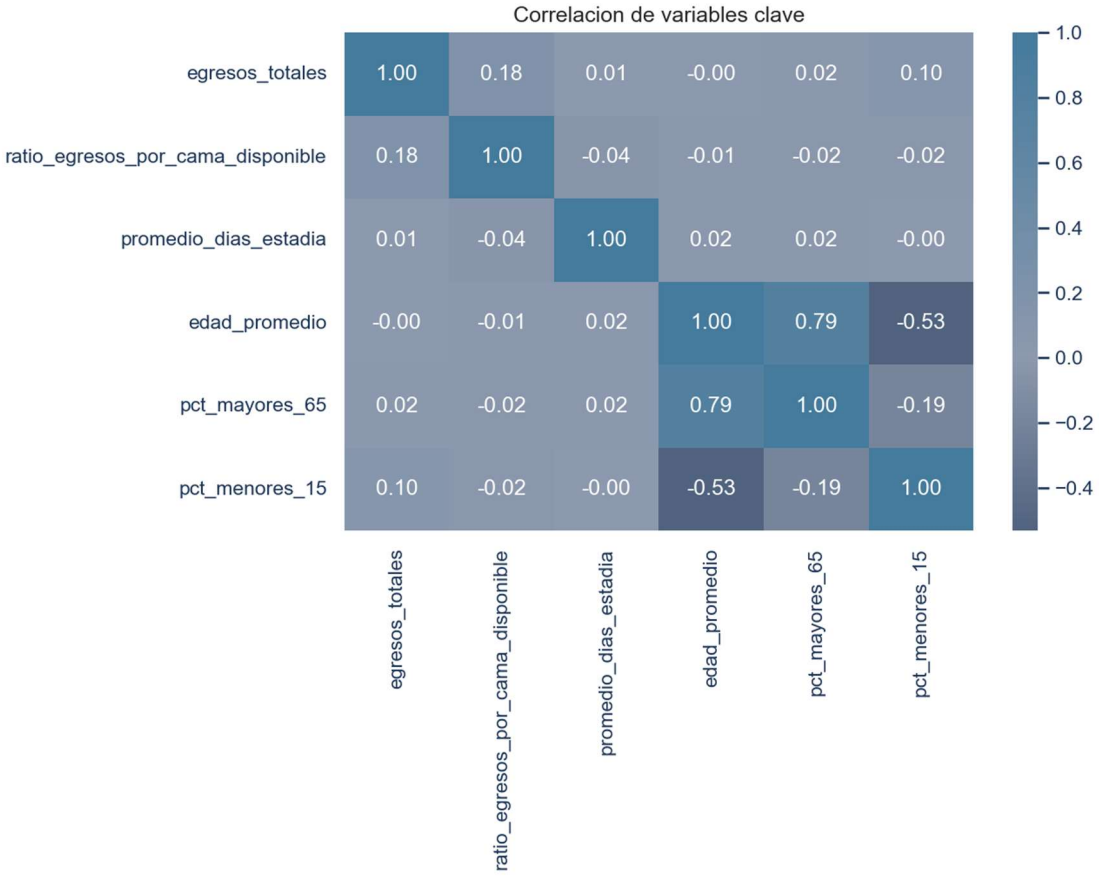
Distribución acumulada del promedio de días de estadía



La curva acumulada precisa los cuantiles de la estadía promedio: la mitad de los registros no supera los 2.3 días y el noventa por ciento se mantiene por debajo de aproximadamente cinco días. Estos umbrales delimitan el comportamiento típico frente a los casos de larga estancia.

Figura 3-29

Heatmap de correlación de variables clave



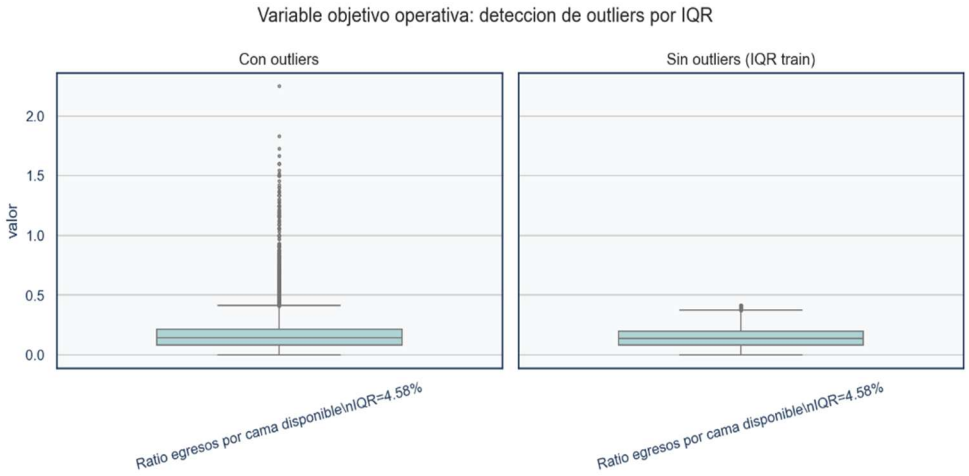
La proporción por sexo se mantiene relativamente equilibrada entre grupos, mientras que la estadía promedio presenta mayor variabilidad. El heatmap de correlación se utiliza como herramienta descriptiva para observar relaciones lineales entre variables numéricas, pero no como criterio automático de eliminación de variables. La selección final de variables se evalúa posteriormente mediante escenarios de modelado y métricas de desempeño como MAE, RMSE y WMAPE.

3.3.2.8. Detección de outliers y calidad de datos

Se aplicó detección de valores atípicos mediante el criterio IQR. Los umbrales se estimaron en el conjunto de entrenamiento y se aplicaron sin recalibración a validación y prueba, con el fin de evitar fuga de información y mantener consistencia temporal.

Figura 3-30

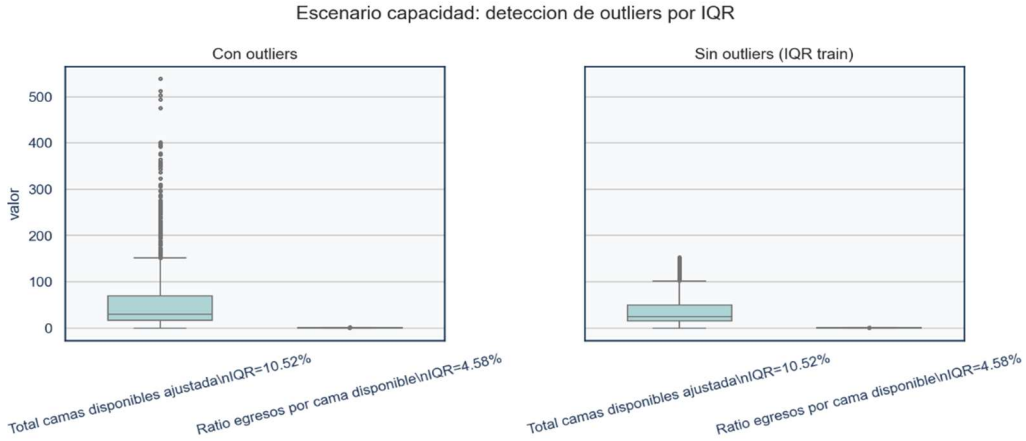
Comparación de variable objetivo operativa con y sin outliers



En la variable objetivo operativa, definida como el ratio de egresos por cama disponible, la proporción de valores fuera de rango fue de 4,58 %. La comparación visual muestra que la eliminación de outliers reduce la cola alta y la dispersión extrema, mientras mantiene relativamente estable la tendencia central.

Figura 3-31

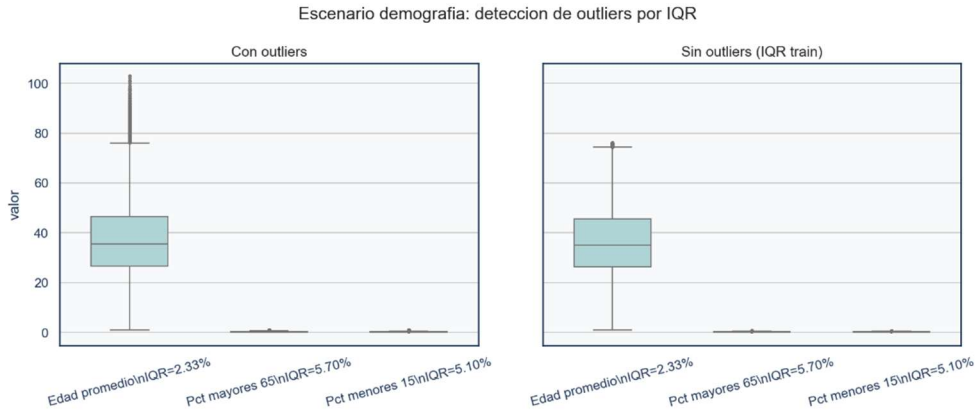
Outliers en variables de capacidad



En el bloque de capacidad, se identificó 10,52 % de valores atípicos en camas disponibles y 4,58 % en el ratio operativo. El filtrado reduce valores extremos de gran magnitud y mejora la comparabilidad entre unidades analíticas.

Figura 3-32

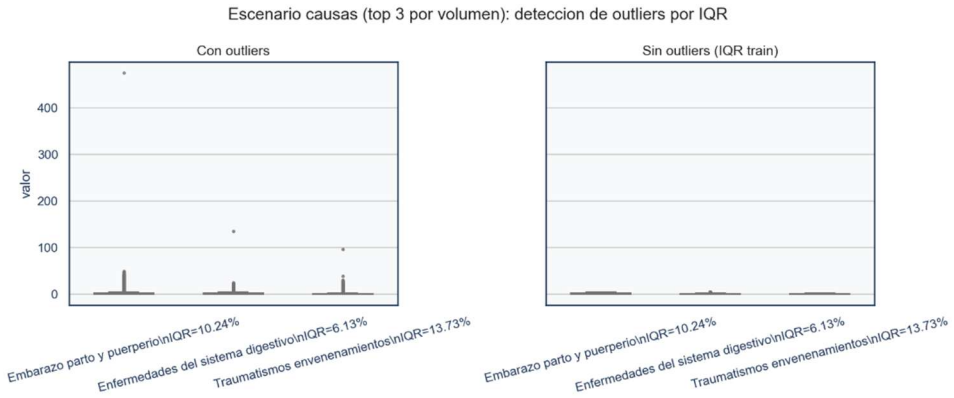
Outliers en variables demográficas



En el bloque demográfico, los porcentajes de valores atípicos fueron moderados: 2,33 % para edad promedio, 5,70 % para porcentaje de mayores de 65 años y 5,10 % para porcentaje de menores de 15 años. Estos resultados sugieren una depuración de extremos puntuales sin cambios sustantivos en la estructura central del bloque demográfico.

Figura 3-33

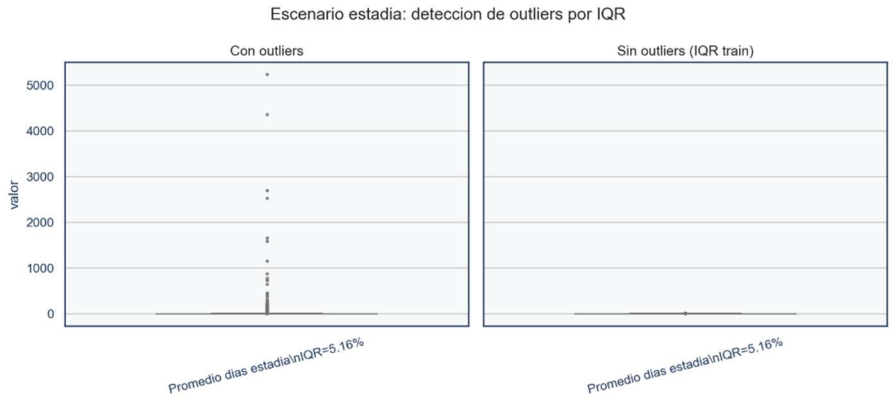
Outliers en variables de causas



En el bloque de causas, los porcentajes IQR fueron mayores para algunas categorías: 10,24 % para embarazo, parto y puerperio; 6,13 % para enfermedades del sistema digestivo; y 13,73 % para traumatismos y envenenamientos. La limpieza reduce colas largas y facilita la lectura de la variabilidad típica, manteniendo la jerarquía relativa entre causas.

Figura 3-34

Outliers en promedio de días de estadía



En el bloque de estadía, el promedio de días de estadía presentó 5,16 % de valores atípicos. La comparación visual muestra reducción de extremos y conservación del patrón central, lo que respalda su uso como predictor complementario.

El filtrado de outliers se aplicó sobre 40 columnas relacionadas con demanda, ratio operativo, capacidad y causas con variables históricas. La regla utilizada fue por fila: si una observación resultaba atípica en al menos una de las columnas evaluadas, se eliminaba del escenario sin outliers. Por esta razón, el porcentaje total de registros excluidos fue alto: 79,42 % en entrenamiento, 80,56 % en validación, 80,47 % en prueba y 79,78 % del total.

Este resultado debe interpretarse con cautela. La eliminación de outliers no se considera una depuración definitiva del dataset, sino un escenario experimental para comparar el comportamiento de los modelos con y sin valores extremos. Por ello, el estudio conserva ambos escenarios de evaluación.

Tabla 3-13

Columnas consideradas en el filtrado IQR por fila

Grupo	Criterio de columnas	Número de columnas	Ejemplos
Demanda total	egresos_totales y derivados temporales	7	egresos_totales, egresos_totales_lag_7, egresos_totales_roll_mean_28
Ratio operativo	ratio_egresos_por_cama_disponible y derivados temporales	7	ratio_egresos_por_cama_disponible, ratio_egresos_por_cama_disponible_lag_7, ratio_egresos_por_cama_disponible_roll_std_28
Capacidad	total_camas_disponibles_original y derivados temporales	7	total_camas_disponibles_original, total_camas_disponibles_original_lag_7, total_camas_disponibles_original_roll_mean_28
Causas (porcentaje)	pct_egresos_causa_ y lag 7	30	pct_egresos_causa_enfermedades_del_sistema_respiratorio, pct_egresos_causa_neoplasias_lag_7
Total	Suma de todas las columnas evaluadas por IQR	51	-

3.3.2.9. Síntesis y hallazgos clave

El EDA sobre la tabla Gold permitió confirmar que la integración de datos produjo una base analítica adecuada para el modelado predictivo. Los principales hallazgos fueron la alta concentración de egresos en ciertos grupos de establecimientos, la presencia de patrones temporales, la relación positiva pero dispersa entre egresos y camas disponibles, y la necesidad de reducir columnas de camas y agrupar causas CIE-10 para mejorar la interpretabilidad.

Asimismo, el análisis de outliers justificó la construcción de escenarios con y sin valores extremos. Dado que el filtrado por fila elimina una proporción elevada de registros, ambos escenarios se mantienen para comparar su impacto en el desempeño predictivo. En conjunto, estos resultados respaldan las transformaciones aplicadas antes de la construcción final del dataset de modelado.

3.3.3 Construcción del dataset de modelado e ingeniería de características

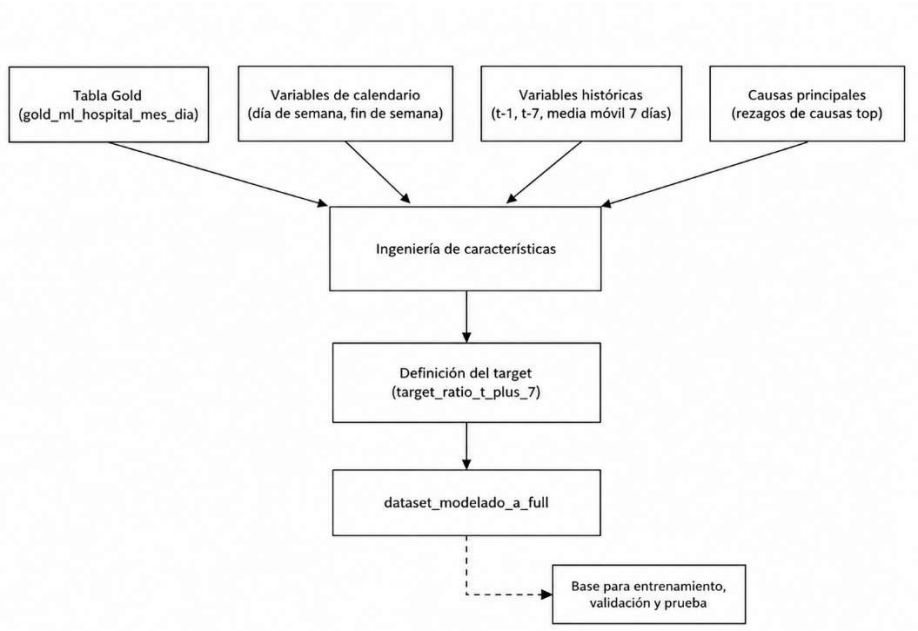
A partir de la tabla Gold se construyó el dataset de modelado denominado dataset_modelado_a_full, almacenado en formato Parquet. Esta tabla constituye el insumo central para las etapas de entrenamiento, validación y evaluación de los modelos predictivos, manteniendo trazabilidad con las fases previas de preparación de datos.

La tabla resultante contiene 373.569 registros y amplía la estructura original de Gold mediante variables derivadas de ingeniería de características. Estas variables buscan capturar patrones temporales, comportamiento histórico de la demanda, disponibilidad de camas, composición por causas y señales de calendario. La granularidad del dataset es diaria por grupo hospitalario, con clave grupo_hospital_id-fecha y sin duplicados en dicha combinación.

El período analizado abarca desde el 8 de enero de 2022 hasta el 24 de diciembre de 2024 e incluye 591 grupos hospitalarios únicos. Además, la tabla integra variables numéricas y no numéricas que combinan señales de demanda, capacidad y contexto operativo.

Figura 3-35

Modelado e ingeniería de características



3.3.3.1. Creación de la tabla dataset_modelado

El dataset_modelado_a_full se construyó a partir de la tabla Gold mediante la incorporación de variables temporales, variables de calendario, información histórica de demanda y capacidad, estadísticas móviles y variables relacionadas con causas de egreso.

La variable objetivo se incorporó en esta tabla como target_ratio_t_plus_7, asociada al horizonte predictivo de siete días. Esta variable no presenta valores nulos dentro del dataset final de modelado.

La síntesis de las principales variables creadas se presenta en la siguiente tabla, y el total de columnas en el siguiente gráfico:

Figura 3-36

Evolución de dimensionalidad: Gold vs Modelado

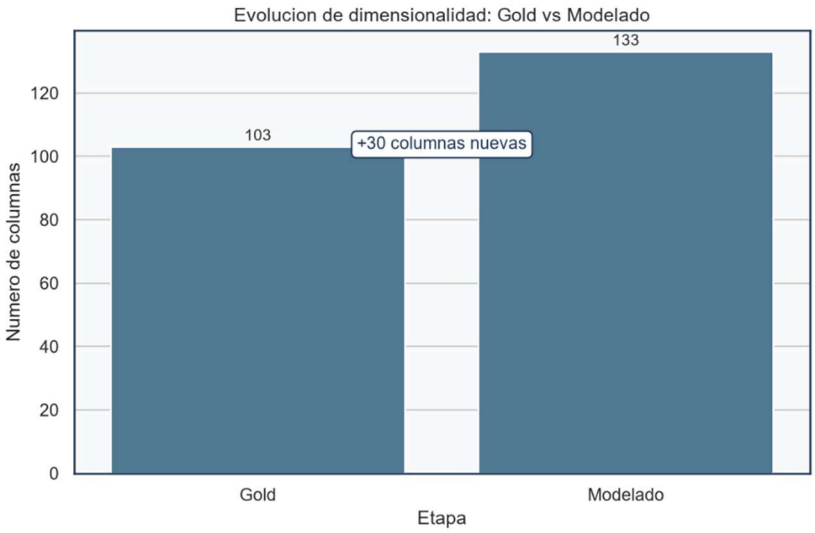


Tabla 3-14

Resumen actualizado de columnas en dataset_modelado_a_full

Grupo	Variables	Cantidad	Observación
Identificación y tiempo	grupo_hospital_id, fecha	2	Definen la unidad analítica grupo hospitalario-día y permiten ordenar cronológicamente la serie para partición temporal y trazabilidad del modelado.
Variable objetivo	target_ratio_t_plus_7	1	Representa la señal a predecir: el ratio egresos/cama con 7 días de anticipación,

Grupo	Variables	Cantidad	Observación
			alineado con el horizonte operativo del estudio.
Calendario operativo	dow, is_weekend	2	Incorporan estacionalidad semanal interpretable, distinguiendo patrón por día de semana y efecto fin de semana en la dinámica de demanda-capacidad.
Rezagos de variables base	ratio_egresos_por_cama_disponible_lag_1, ratio_egresos_por_cama_disponible_lag_7, egresos_totales_lag_1, egresos_totales_lag_7, total_camas_disponibles_original_lag_1, total_camas_disponibles_original_lag_7	6	Capturan memoria temporal reciente de uso, demanda y oferta de camas, aportando contexto autoregresivo de corto plazo al pronóstico.
Rolling de variables base	roll_mean_7, roll_std_7, roll_mean_28, roll_std_28 para cada variable base (ratio, egresos_totales, total_camas_disponibles_original)	12	Resumen nivel y variabilidad de corto y mediano plazo; estabilizan ruido diario y mejoran la lectura de

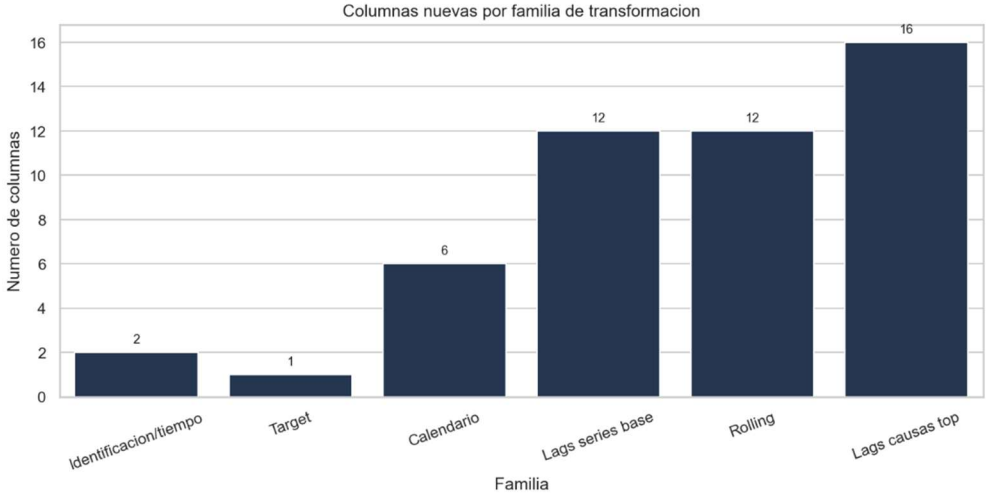
Grupo	Variables	Cantidad	Observación
			tendencia y volatilidad operativa.
Rezagos de composición por causa top	pct_egresos_causa_*_lag_7 (8 causas)	8	Añaden contexto epidemiológico reciente del mix de causas, útil para explicar cambios en la intensidad de egresos entre grupos y periodos.
Total columnas de ingeniería		29	Conjunto de variables explicativas diseñado para maximizar capacidad predictiva y mantener interpretabilidad del modelo.
Total adicional incluyendo estructurales + target		31	Cobertura completa para modelado supervisado: variables explicativas, estructura temporal y variable objetivo.

3.3.3.2. Selección y transformación de variables

La selección de variables se realizó con el objetivo de mantener información relevante para la predicción y reducir el riesgo de incorporar variables redundantes o no disponibles antes del período a estimar. Para ello, las variables se organizaron en grupos: identificación y tiempo, calendario, variables históricas, estadísticas móviles y causas principales.

Figura 3-37

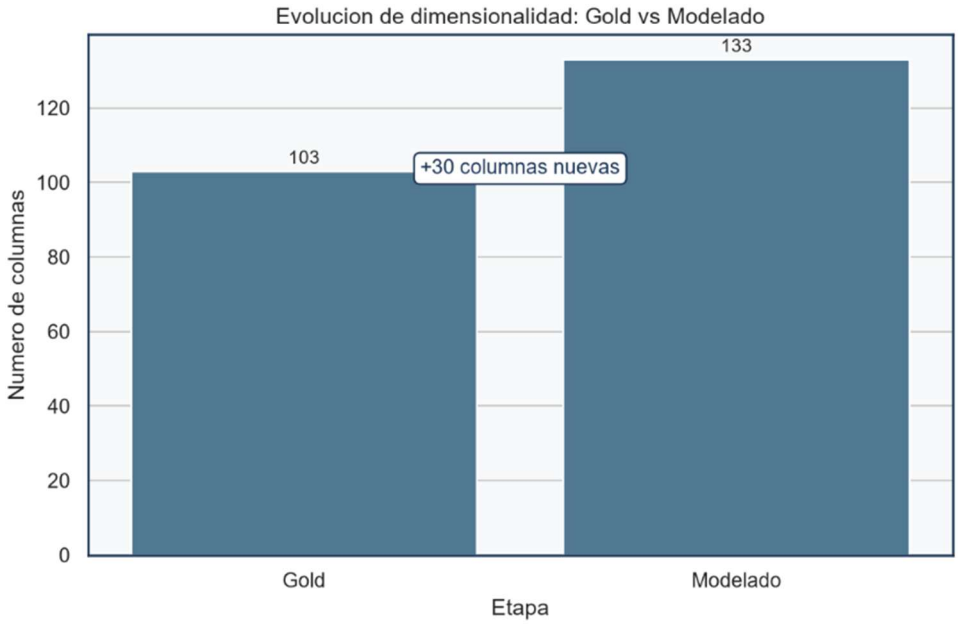
Agrupación de variables para modelado



Posteriormente, se verificó el cambio de dimensionalidad entre la tabla Gold y el dataset de modelado. Este incremento responde a la incorporación de variables derivadas que permiten capturar memoria histórica, tendencia reciente y comportamiento operativo de los grupos hospitalarios.

Figura 3-38

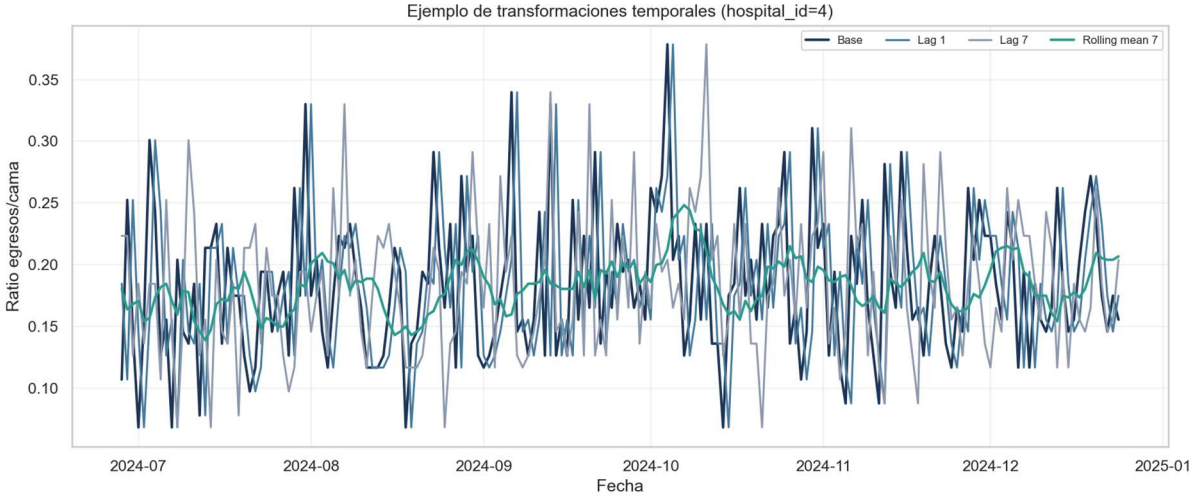
Comparación de dimensionalidad entre Gold y dataset de modelado



Las transformaciones temporales incluyeron valores históricos de períodos anteriores, como t-1 y t-7, y promedios móviles con ventana de 7 días. Estas variables permiten que el modelo utilice información pasada de cada grupo hospitalario para estimar el comportamiento futuro.

Figura 3-39

Variables históricas y estadísticas móviles



La selección de causas principales se realizó con criterio temporal, utilizando únicamente información disponible en el período de entrenamiento. De esta forma, se evita que la selección de variables incorpore información futura proveniente de validación o prueba.

3.3.3.3. Definición del target

La variable objetivo del estudio se define como el ratio de egresos hospitalarios por cama disponible en un horizonte de siete días. Este indicador permite representar la presión asistencial esperada, al relacionar la demanda observada con la capacidad disponible.

La formulación general del ratio en el tiempo t es:

$$\text{Target}_{h,t} = \frac{\text{egresos}_{h,t}}{\text{camas}_{h,t}}$$

Donde h representa el grupo hospitalario y t el día de referencia.

Para fines predictivos, se utiliza un horizonte de siete días. Por tanto, la variable objetivo final se expresa como:

$$y_{h,t} = \left(\frac{\text{egresos}}{\text{camas}}\right)_{h,\,t+7}$$

Esto significa que, con la información disponible hasta el tiempo t, el modelo estima la presión asistencial esperada una semana después. Por ejemplo, si t corresponde al 10 de mayo de 2024, el modelo utiliza únicamente la información disponible hasta esa fecha para estimar el ratio egresos/camas del 17 de mayo de 2024.

Bajo esta definición, valores altos del target indican mayor presión relativa de demanda frente a la capacidad disponible, mientras que valores bajos reflejan menor presión asistencial. Esta formulación se mantiene constante en todos los escenarios experimentales; lo que cambia entre escenarios es el conjunto de variables utilizadas y el tratamiento de outliers, no la variable objetivo.

3.3.4. Prevención de fuga de información

Para asegurar que los modelos aprendan patrones reales y no información futura, se aplicaron controles de fuga sobre dataset_modelado_a_full antes del entrenamiento. Estos controles fueron especialmente importantes por tratarse de un problema con estructura temporal.

En primer lugar, cada serie se ordenó cronológicamente por grupo hospitalario y fecha. Esto garantiza que las transformaciones temporales se calculen respetando la secuencia de cada grupo.

```
# 1) Orden temporal por hospital
df.sort_values(["hospital_id", "fecha"]).reset_index(drop=True)
```

La variable objetivo se definió con horizonte de siete días mediante desplazamiento hacia adelante dentro de cada grupo hospitalario. Con ello, cada fila en t busca predecir el valor de presión asistencial en t+7.

```
# 2) Definición del target (horizonte t+7)
df["target_ratio_t_plus_7"] = (
    df.groupby("hospital_id")["ratio_egresos_por_cama_disponible"].shift(-
TARGET_HORIZON_DAYS)
)
```

Las variables históricas se generaron mediante rezagos y estadísticas móviles calculadas únicamente con información previa al día de referencia. Para las medias y desviaciones móviles se aplicó desplazamiento previo, evitando que el valor actual o futuro contamine la predicción.

```
# 4) Rolling sin acceso a futuro (con shift(1))
def _add_rolling_features(df, cols, windows):
```

```
    g = df.groupby("hospital_id", group_keys=False)
    for col in cols:
        shifted = g[col].shift(1)
        for w in windows:
            df[f"{col}_roll_mean_{w}"] = shifted.rolling(w, min_periods=3).mean()
            df[f"{col}_roll_std_{w}"] = shifted.rolling(w, min_periods=3).std()
    return df
```

Las variables de calendario, como día de semana, fin de semana y codificaciones cíclicas, se incluyeron porque son conocidas antes del período de predicción. Por tanto, aportan información estacional sin introducir fuga.

La selección de causas principales también se realizó con datos del conjunto de entrenamiento, evitando utilizar información de validación o prueba para decidir qué causas conservar.

```
•   # 5) Causas top seleccionadas con entrenamiento
•   df_top = df[df["fecha"] <= TRAIN_END].copy()
if df_top.empty:
    df_top = df


top_causas = (
    df_top[causa_cols]
    .apply(pd.to_numeric, errors="coerce")
    .mean(skipna=True)
    .sort_values(ascending=False)
```

La separación entre entrenamiento, validación y prueba se realizó de forma estrictamente temporal, sin mezclar observaciones futuras con datos usados para entrenar.

```
# 6) Split temporal estricto
def split_df(df):
    train = df[df["fecha"] <= TRAIN_END].copy()
    valid = df[(df["fecha"] > TRAIN_END) & (df["fecha"] <= VALID_END)].copy()
    test = df[df["fecha"] > VALID_END].copy()
    return train, valid, test
```

En el escenario sin outliers, los límites IQR se calcularon únicamente con el conjunto de entrenamiento y luego se aplicaron a validación y prueba sin recalibración. Además, la variable objetivo y los identificadores fueron excluidos de la regla de filtrado para evitar que el resultado condicionara la definición del escenario.

Bajo estas reglas se definieron dos escenarios experimentales. En el escenario con outliers se conserva el conjunto completo después de la ingeniería de características. En el escenario sin outliers se excluyen observaciones fuera de los límites IQR aprendidos en entrenamiento y aplicados de forma consistente a validación y prueba.

En síntesis, el dataset de modelado fue construido respetando el orden temporal, evitando el uso de información futura y manteniendo una separación cronológica entre entrenamiento, validación y prueba. Estos controles permiten que la evaluación de los modelos sea más realista y coherente con el uso operativo esperado.

3.4 Fase 4. Modelado

La cuarta fase corresponde al modelado predictivo. En esta etapa se utilizan las particiones temporales del dataset de modelado para entrenar, validar y comparar distintos enfoques bajo una matriz experimental definida por escenarios de datos, conjuntos de variables y modelos. El objetivo de esta fase es identificar los modelos con mejor capacidad para estimar la presión hospitalaria futura, manteniendo criterios de reproducibilidad, comparabilidad y prevención de fuga de información. Para ello, se aplicó una estrategia de búsqueda aleatoria de hiperparámetros en dos fases, con validación cruzada temporal, respetando el orden cronológico de los datos.

3.4.1. División temporal de datos: entrenamiento, validación y prueba

Una vez construida la tabla de modelado y definida la variable objetivo, el conjunto de datos se dividió temporalmente en tres subconjuntos: entrenamiento, validación y prueba. Esta partición respetó el orden cronológico de las observaciones, con el fin de simular un entorno real de predicción y evitar fuga de información entre períodos.

Las observaciones con fecha menor o igual al 31 de diciembre de 2023 fueron asignadas al conjunto de entrenamiento. Las observaciones posteriores a esa fecha y hasta el 30 de junio de 2024 fueron destinadas al conjunto de validación. Finalmente, las observaciones posteriores al 30 de junio de 2024 fueron reservadas para el conjunto de prueba.

```
TRAIN_END = pd.Timestamp("2023-12-31")
VALID_END = pd.Timestamp("2024-06-30")
def split_df(df):
    train = df[df["fecha"] <= TRAIN_END].copy()
    valid = df[(df["fecha"] > TRAIN_END) & (df["fecha"] <= VALID_END)].copy()
    test = df[df["fecha"] > VALID_END].copy()
    return train, valid, test
```

La partición temporal se aplicó de forma consistente en los dos escenarios experimentales definidos: con outliers y sin outliers. Esto garantiza que la comparación de desempeño entre escenarios responda al tratamiento de valores atípicos y no a diferencias en los períodos evaluados.

```
train_con, valid_con, test_con = split_df(model_df_con_outliers)
train_sin, valid_sin, test_sin = split_df(model_df_sin_outliers)
```

En términos operativos, esta división permitió entrenar los modelos con datos históricos, ajustar hiperparámetros con un período intermedio de validación y evaluar el desempeño final sobre un conjunto futuro no utilizado durante el entrenamiento.

3.4.2. Diseño experimental y búsqueda de hiperparámetros en dos fases

El proceso de benchmarking se organizó mediante una matriz experimental que combina escenarios de datos, conjuntos de variables, modelos y particiones de evaluación. Esta estructura permitió comparar los modelos de forma homogénea, manteniendo constante la definición del target y la lógica de partición temporal.

3.4.2.1. Escenario con outliers

El escenario con outliers conserva la totalidad de las observaciones del dataset de modelado. En este caso no se aplica filtrado de valores atípicos, por lo que la evaluación se realiza sobre la distribución original de los datos. Este escenario funciona como línea base para analizar si los valores extremos aportan señal predictiva o introducen ruido en el ajuste.

```
odel_df_con_outliers = model_df.copy()
train_con, valid_con, test_con = split_df(model_df_con_outliers)
```

Tabla 3-15

Filas resultantes del escenario con outliers

Escenario	Filas totales	Entrenamiento	Validación	Prueba
con outliers	373.569	250.507	63.073	59.989

En síntesis, este escenario preserva la distribución completa de los datos y sirve como referencia para comparar el efecto del filtrado de valores extremos.

3.4.2.2. Escenario sin outliers

El escenario sin outliers mantiene la misma estructura temporal y la misma ingeniería de características, pero elimina observaciones atípicas antes del entrenamiento. Su objetivo es evaluar si la depuración de valores extremos mejora la estabilidad del modelo y su capacidad predictiva frente al escenario base.

El filtrado se realizó mediante el criterio IQR. Los límites se calcularon únicamente con el conjunto de entrenamiento y luego se aplicaron sin recalibración sobre validación y prueba. De esta forma se evita fuga de información y se preserva la comparabilidad temporal.

```
bounds = {}


for col in features_outlier:
    s_train = pd.to_numeric(train_con[col], errors="coerce")
    q1 = s_train.quantile(0.25)
    q3 = s_train.quantile(0.75)
    iqr = q3 - q1
    if pd.isna(iqr) or iqr == 0:
        continue
    bounds[col] = (q1 - 1.5 * iqr, q3 + 1.5 * iqr)
```

```
def filtrar_outliers_con_bounds(df_split):
    if not bounds:
        return df_split.copy()
    mask_outlier = np.zeros(len(df_split), dtype=bool)
    for col, (low, high) in bounds.items():
        s = pd.to_numeric(df_split[col], errors="coerce")
        mask_outlier |= (s < low) | (s > high)
    return df_split.loc[~mask_outlier].copy()

train_sin = filtrar_outliers_con_bounds(train_con)
valid_sin = filtrar_outliers_con_bounds(valid_con)
test_sin = filtrar_outliers_con_bounds(test_con)
```

Tabla 3-16

Filas resultantes del escenario sin outliers

Escenario	Filas totales	Entrenamiento	Validación	Prueba
sin outliers	71.771	17.810	12.263	16.872

El escenario con_outliers mantiene una mayor cobertura de observaciones y conserva una representación más amplia de los patrones históricos, incluyendo valores extremos. En cambio, el escenario sin_outliers reduce el número de registros debido al filtrado aplicado sobre valores atípicos. En ambos casos, las particiones mantienen una estructura temporal consistente entre entrenamiento, validación y prueba, lo que permite comparar el efecto del tratamiento de outliers bajo una lógica cronológica y sin fuga de información.

3.4.2.3. Matriz de combinaciones

La etapa de modelado se organizó mediante una matriz experimental que cruza cuatro dimensiones: escenario de datos, conjunto de variables, modelo y partición de evaluación. En total, se evaluaron 2 × 6 × 4 × 2 = 96 combinaciones experimentales.

Tabla 3-17

Matriz experimental de modelado

Factor	Niveles	Cantidad
Escenario de datos	con outliers, sin outliers	2
Conjunto de variables	all, capacidad, demografia, causas, capacidad+demografia	6
Modelo	Naive_t, SeasonalNaive_lag7, RandomForest, XGBoost	4
Partición de evaluación	valid, test	2

Esta estructura permite comparar el desempeño bajo una misma base temporal, evaluar la contribución de cada grupo de variables y medir el impacto del tratamiento de outliers. Así, cualquier diferencia observada en las métricas responde al modelo, al conjunto de variables o al escenario de datos, y no a cambios en el criterio experimental.

La implementación se realizó mediante diccionarios de configuración que definen los escenarios, conjuntos de variables y modelos disponibles. El benchmark recorre sistemáticamente estas combinaciones y genera resultados para validación y prueba.

```
DATA_SCENARIOS = {
    "con_outliers": MODEL_DIR / "con_outliers",
    "sin_outliers": MODEL_DIR / "sin_outliers",
}
MODEL_NAMES = ("Naive_t", "SeasonalNaive_lag7", "RandomForest", "XGBoost")
```

3.4.2.4. Estrategia de búsqueda de hiperparámetros en dos fases

El ajuste de hiperparámetros para Random Forest y XGBoost se realizó mediante RandomizedSearchCV con validación cruzada temporal. Para equilibrar desempeño predictivo y costo computacional, se utilizó una estrategia de búsqueda en dos fases.

En la primera fase se ejecutó una búsqueda inicial con pocas iteraciones sobre las combinaciones de escenario de datos y conjunto de variables. Esta etapa permitió identificar las configuraciones más prometedoras según el WMAPE obtenido en validación.

Posteriormente, en la segunda fase, se realizó una búsqueda más amplia sobre las 10 combinaciones finalistas, concentrando el esfuerzo computacional en las alternativas con mejor desempeño inicial.

En ambas fases se utilizó TimeSeriesSplit con tres pliegues, garantizando que cada bloque de validación fuera posterior al bloque de entrenamiento. Esta decisión preserva el orden temporal de los datos y evita que la búsqueda de hiperparámetros utilice información futura.

```
cv_obj = TimeSeriesSplit(n_splits=3)
search = RandomizedSearchCV(
    estimator=model_builder(),
    param_distributions=param_distributions,
    n_iter=rs_n_iter,
    cv=cv_obj,
    scoring="neg_mean_absolute_error",
    random_state=42,
    n_jobs=-1,
    refit=False,
)
search.fit(x_search, y_search)
```

```
best_params = search.best_params_
model = model_builder(**best_params)
model.fit(x_train, y_train)
```

En total, el proceso ejecutó 1.532 ajustes internos de modelo antes de generar las 96 evaluaciones finales de la matriz experimental. Esta cantidad incluye las búsquedas de validación cruzada y los reentrenamientos finales de las configuraciones seleccionadas.

Los espacios de búsqueda para cada algoritmo fueron definidos con rangos moderados de complejidad. En Random Forest se consideraron parámetros como número de árboles, profundidad máxima, tamaño mínimo de hoja y número de variables evaluadas por división. En XGBoost se exploraron parámetros como número de estimadores, tasa de aprendizaje, profundidad máxima, submuestreo, proporción de columnas y regularización. El detalle completo de los rangos evaluados se presenta en el anexo correspondiente.

Esta estrategia de dos fases permitió reducir el costo computacional frente a una búsqueda exhaustiva, sin perder capacidad para identificar configuraciones con buen desempeño. Los resultados finales de la matriz experimental quedaron registrados en el archivo de benchmark 20_benchmark_modelos_a_escenarios.csv.

3.4.3. Modelos de Machine Learning

Esta sección presenta los modelos utilizados en el benchmark principal de Machine Learning y su rol dentro del sistema de evaluación. Se incluyeron dos modelos base, utilizados como referencia mínima, y dos modelos de aprendizaje automático sobre datos tabulares: Random Forest y XGBoost.

Tabla 3-18

Modelos de Machine Learning

Modelo	Enfoque	Rol en el estudio	Fortalezas principales
Random Forest	Bagging	Referencia no lineal robusta	Estabilidad y reducción de varianza
XGBoost	Boosting	Modelo de alta capacidad predictiva	Mayor capacidad de ajuste con control de regularización

3.4.3.1. Modelos base

Los modelos base se utilizaron para establecer una referencia inicial de comparación. Su propósito no es maximizar la precisión, sino verificar si los modelos más complejos aportan una mejora real frente a reglas simples de pronóstico.

Naive_t proyecta el valor reciente de la serie, mientras que SeasonalNaive_lag7 utiliza el valor observado siete días antes, capturando un posible patrón semanal. Estos modelos permiten interpretar con mayor claridad la ganancia obtenida por Random Forest y XGBoost.

3.4.3.2. Random Forest

Random Forest se utilizó como modelo de ensamble basado en múltiples árboles de decisión. Su enfoque de bagging permite reducir la varianza al promediar árboles entrenados con aleatoriedad en filas y variables. Esto lo hace adecuado para datos tabulares con variables heterogéneas, relaciones no lineales e interacciones entre predictores.

La configuración base utilizada fue la siguiente:

```
m = RandomForestRegressor(
    n_estimators=RF_N_ESTIMATORS,
    max_depth=12,
    min_samples_leaf=8,
```

```
    random_state=42,
    n_jobs=-1,
)
```

En este estudio, Random Forest aporta una referencia sólida y estable para evaluar el comportamiento de modelos no lineales bajo la misma estructura temporal y los mismos conjuntos de variables.

3.4.3.3. XGBoost

XGBoost se incorporó como modelo de boosting secuencial, orientado a capturar patrones complejos mediante la corrección progresiva de errores. Su capacidad de regularización lo hace útil cuando existen relaciones no lineales entre variables históricas, capacidad instalada y composición de egresos.

La configuración base utilizada fue la siguiente:

```
m = XGBRegressor(
    n_estimators=XGB_N_ESTIMATORS,
    learning_rate=0.06,
    max_depth=6,
    subsample=0.85,
    colsample_bytree=0.85,
    objective="reg:squarederror",
    random_state=42,
    n_jobs=-1,
)
```

En este estudio, XGBoost representa el modelo de mayor flexibilidad dentro del benchmark principal de Machine Learning, complementando la estabilidad de Random Forest con una mayor capacidad de ajuste.

3.4.3.4. Preparación de entrada para los modelos

Para garantizar comparabilidad entre algoritmos, todos los modelos utilizaron el mismo esquema de preparación de entrada en cada combinación evaluada. Se excluyeron identificadores y la variable objetivo; posteriormente, los valores faltantes fueron imputados con la mediana calculada únicamente sobre el conjunto de entrenamiento. Esa misma mediana se aplicó de forma consistente a validación y prueba.

```
exclude = {TARGET_COL, "fecha", "hospital_id", "grupo_hospital_id"}
all_numeric = [
    c for c in train_df.columns
    if c not in exclude and pd.api.types.is_numeric_dtype(train_df[c])
]
```

```
x_train = train_df[feature_cols].copy()
x_valid = valid_df[feature_cols].copy()
x_test = test_df[feature_cols].copy()


med = x_train.median(numeric_only=True)
x_train = x_train.fillna(med)
x_valid = x_valid.fillna(med)
x_test = x_test.fillna(med)
```

Este procedimiento evita tratamientos diferentes entre particiones y mantiene una base homogénea para comparar el desempeño de los modelos mediante MAE, RMSE y WMAPE.

3.5 Fase 5. Evaluación

En esta fase se presentan y comparan los resultados obtenidos por los modelos predictivos bajo los escenarios de datos, conjuntos de variables y particiones temporales definidos previamente. La evaluación se realiza mediante métricas de regresión y tablas de síntesis, con el propósito de identificar el desempeño relativo de los modelos base y de los modelos de Machine Learning en un marco homogéneo, trazable y reproducible.

3.5.1. Métricas de evaluación

Para evaluar el desempeño de los modelos se utilizaron tres métricas de regresión: MAE, RMSE y WMAPE. Estas métricas permiten analizar el error desde perspectivas complementarias: error absoluto promedio, sensibilidad a errores grandes y error relativo ponderado respecto al volumen real observado.

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

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

$$WMAPE = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{\sum_{i=1}^{n}|y_i|}$$

Donde y_i representa el valor real, $\hat{y}_i$ la predicción del modelo y n el número de observaciones evaluadas.

Desde el punto de vista técnico, las métricas se calcularon de forma consistente para los conjuntos de validación y prueba:

```
def wmape(y_true: np.ndarray, y_pred: np.ndarray) -> float:
    denom = np.abs(y_true).sum()
    if denom == 0:
        return float("nan")
    return float(np.abs(y_true - y_pred).sum() / denom)


def metricas(y_true: pd.Series, y_pred: np.ndarray) -> dict[str, float]:
    yt = np.asarray(y_true, dtype=float)
    yp = np.asarray(y_pred, dtype=float)

    mae = float(mean_absolute_error(yt, yp))
    rmse = float(np.sqrt(mean_squared_error(yt, yp)))

    return {
        "MAE": mae,
        "RMSE": rmse,
        "WMAPE": wmape(yt, yp),
    }
```

En conjunto, MAE permite interpretar el error medio, RMSE enfatiza errores grandes y WMAPE facilita una lectura proporcional del error entre escenarios con distintas escalas de demanda.

3.5.2. Evaluación de modelos de Machine Learning

En esta sección se comparan los resultados obtenidos por los modelos evaluados en la matriz experimental. La comparación considera los escenarios con y sin outliers, los distintos conjuntos de variables y las particiones de validación y prueba. El propósito es identificar qué modelos y configuraciones ofrecen mejor desempeño predictivo, y verificar si los modelos de Machine Learning superan a las líneas base definidas para el estudio.

3.5.2.1. Criterios de comparación

La comparación incluyó cuatro modelos: Naive_t, SeasonalNaive_lag7, RandomForest y XGBoost. Los dos primeros se utilizaron como modelos base, mientras que RandomForest y XGBoost representaron los modelos principales de Machine Learning. Esta distinción permite evaluar si los modelos más complejos aportan una mejora real frente a estrategias simples de pronóstico.

La evaluación se aplicó sobre las particiones de validación y prueba, manteniendo constante la separación temporal. La matriz experimental consideró dos escenarios de datos, con outliers y sin outliers, y seis conjuntos de variables: all, capacidad, demografía, causas, estadía y capacidad+demografía. Para la comparación global se priorizó WMAPE como métrica principal, complementada por MAE y RMSE.

3.5.2.2. Resultados comparativos

Para facilitar la interpretación, los resultados se presentan en tablas de síntesis agregadas sobre el conjunto de prueba.

Tabla 3-19

Promedio global por modelo (test)

Modelo	WMAPE promedio	MAE promedio	RMSE promedio	N configuraciones
RandomForest	0.4429	0.0683	0.0940	12
XGBoost	0.4430	0.0683	0.0941	12
Naive_t	0.4689	0.0770	0.1203	12
SeasonalNaive_lag7	0.4793	0.0738	0.1106	12

En promedio global, RandomForest y XGBoost presentan un desempeño prácticamente equivalente y superior al de los modelos base. Esto indica que los modelos de Machine Learning aportan una mejora frente a reglas simples de persistencia temporal.

Tabla 3-20

Promedio por escenario de datos en prueba

Escenario	WMAPE promedio	MAE promedio	RMSE promedio	N configuraciones
sin_outliers	0.4396	0.0632	0.0846	24
con_outliers	0.4463	0.0733	0.1036	24

El escenario sin outliers muestra un desempeño promedio ligeramente mejor que el escenario con outliers. Sin embargo, la diferencia en WMAPE es moderada, por lo que ambos escenarios se consideran útiles para el análisis comparativo.

Tabla 3-21

Promedio por conjunto de variables en prueba

Feature set	WMAPE promedio	MAE promedio	RMSE promedio	N configuraciones
all	0.3763	0.0579	0.0819	**4**
capacidad+demografia	0.3838	0.0590	0.0829	**4**
capacidad	0.3839	0.0591	0.0829	**4**
causas	0.5006	0.0773	0.1051	**4**
estadia	0.5038	0.0777	0.1055	**4**
demografia	0.5093	0.0786	0.1061	**4**

Los mejores resultados se concentran en los conjuntos con mayor información disponible, especialmente all, capacidad+demografía y capacidad. En cambio, los subconjuntos aislados de causas, estadía y demografía presentan menor capacidad predictiva cuando se evalúan de forma independiente.

Tabla 3-22

Mejores configuraciones individuales en prueba

Ranking	Modelo	Escenario	Feature set	WMAPE	MAE	RMSE	N features
1	XGBoost	con_outliers	all	0.3691	0.0606	0.0888	122
2	RandomForest	con_outliers	all	0.3701	0.0608	0.0886	122
3	RandomForest	con_outliers	capacidad	0.3757	0.0617	0.0893	45

Ranking	Modelo	Escenario	Feature set	WMAPE	MAE	RMSE	N features
4	RandomForest	con_outliers	capacidad+demografia	0.3759	0.0617	0.0893	82
5	XGBoost	con_outliers	capacidad+demografia	0.3766	0.0618	0.0896	82
6	XGBoost	con_outliers	capacidad	0.3766	0.0619	0.0896	45

El mejor resultado puntual se obtuvo con XGBoost en el escenario con outliers y el conjunto completo de variables, con un WMAPE de 0.3691. RandomForest presentó un resultado muy cercano bajo la misma configuración. Esto confirma que ambos modelos tienen desempeños competitivos, aunque XGBoost alcanza el mejor valor individual.

Tabla 3-23
Comparación de WMAPE por modelo

Modelo	WMAPE promedio
RandomForest	0.4429
XGBoost	0.4430
Naive_t	0.4689
SeasonalNaive_lag7	0.4793

Tabla 3-24

Tabla Comparación de WMAPE por conjunto de variables

Feature set	WMAPE promedio
all	0.3763
capacidad+demografía	0.3838
capacidad	0.3839
causas	0.5006
estadía	0.5038
demografía	0.5093

3.5.3. Síntesis de hallazgos de la fase

La evaluación permitió identificar cuatro hallazgos principales. En primer lugar, los modelos de Machine Learning superan en promedio a los modelos base, lo que evidencia una ganancia frente a reglas simples de pronóstico. En segundo lugar, RandomForest y XGBoost presentan desempeños muy similares, con una ligera ventaja puntual de XGBoost en la mejor configuración individual. En tercer lugar, el escenario sin outliers mejora levemente el desempeño promedio global, aunque el mejor resultado puntual se obtiene en el escenario con outliers. Finalmente, los conjuntos de variables más completos concentran la mayor capacidad predictiva, especialmente all, capacidad+demografía y capacidad.

3.5.4. Análisis complementario de errores

Como complemento de la comparación principal, se revisó el comportamiento del error en el conjunto de prueba considerando dos niveles de lectura: el desempeño puntual del feature set all y el desempeño promedio por escenario.

Tabla 3-25

Comparación por modelo en prueba para el feature set all

Modelo	WMAPE con_outliers	WMAPE sin_outliers	Delta (sin - con)
RandomForest	0.369245	0.377669	0.008424
XGBoost	0.369217	0.376252	0.007035

En el conjunto completo de variables, ambos modelos presentan mejor desempeño puntual en el escenario con outliers. Esta diferencia no contradice el resultado promedio global, sino que muestra que los valores extremos pueden aportar señal predictiva en configuraciones con mayor cobertura de variables.

Tabla 3-26

Promedio global por escenario en prueba, considerando modelos de Machine Learning

Escenario	MAE promedio	RMSE promedio	WMAPE promedio	N configuraciones
con_outliers	0.073044	0.103198	0.444724	12
sin_outliers	0.060126	0.081056	0.431333	12

A nivel global, el escenario sin outliers reduce el error promedio en las tres métricas. Esto sugiere mayor estabilidad general del escenario filtrado, aunque no necesariamente produce el mejor resultado puntual en todas las configuraciones.

Tabla 3-27

Ranking de feature sets por escenario en prueba

Escenario	Feature set	WMAPE promedio
con_outliers	all	0.369231
con_outliers	capacidad+demografia	0.374487
con_outliers	capacidad	0.375133
con_outliers	causas	0.513132
con_outliers	estadia	0.515637
con_outliers	demografia	0.520725
sin_outliers	all	0.376960
sin_outliers	capacidad	0.383399
sin_outliers	capacidad+demografia	0.383494
sin_outliers	causas	0.471938
sin_outliers	estadia	0.482346
sin_outliers	demografia	0.489861

El patrón es consistente en ambos escenarios: all y capacidad concentran los menores errores, mientras que demografía y estadía presentan los niveles más altos cuando se evalúan de forma aislada.

En síntesis, el análisis de errores muestra dos resultados complementarios. En configuraciones puntuales de alta señal, como el feature set all, el escenario con outliers puede rendir mejor. En cambio, cuando se observa el promedio global del benchmark, el escenario sin outliers presenta menor error agregado. Por ello, el estudio reporta de forma conjunta el mejor resultado puntual y el desempeño promedio por escenario.

3.6 FASE 6. Despliegue

En esta fase, los resultados del modelo seleccionado se transforman en salidas analíticas útiles para la interpretación y la planificación hospitalaria. A diferencia de la fase de benchmark, cuyo objetivo fue comparar múltiples combinaciones de escenarios, conjuntos de variables y modelos, en esta etapa se operacionaliza una configuración final para generar predicciones trazables por grupo hospitalario y fecha.

La lógica de esta fase mantiene coherencia con el diseño experimental previo. La configuración final se selecciona con base en el desempeño observado en validación, utilizando WMAPE como métrica principal. Posteriormente, el conjunto de prueba se conserva para reportar el desempeño final fuera de muestra, evitando que los datos de prueba influyan en la selección del modelo.

3.6.1. Generación de predicciones

En esta subsección se describe la generación de predicciones del modelo final seleccionado. El objetivo es producir una salida detallada e interpretable por grupo hospitalario y fecha, asociada al indicador de presión asistencial definido como la razón egresos/camas en el horizonte de siete días.

La selección del modelo final se realizó a partir de los resultados de validación, comparando escenarios de datos, conjuntos de variables y modelos. Luego, la configuración seleccionada fue evaluada en prueba para documentar su capacidad de generalización sobre datos no utilizados durante el ajuste.

Para evitar ambigüedades en la interpretación de resultados, se distinguen tres niveles de lectura:

Tabla 3-28

Niveles de lectura

Criterio de decisión	Qué compara	Mejor resultado observado	Valor
Mejor escenario promedio	Promedio del benchmark por escenario	sin_outliers	0.431333
Mejor modelo promedio global	Promedio del benchmark por modelo	XGBoost	0.437925
Mejor configuración puntual en prueba	Una combinación específica de escenario + feature set + modelo	XGBoost + con_outliers + all	0.369217

Estos resultados no son contradictorios. Un escenario puede presentar mejor desempeño promedio, un modelo puede liderar el promedio global y una combinación puntual puede alcanzar el menor error individual. Por ello, la interpretación considera tanto el rendimiento agregado como el mejor resultado específico.

Los principales productos generados en esta etapa fueron los siguientes:

Tabla 3-29

Principales productos generados en esta etapa

Producto	Contenido	Archivo
Predicción detallada	Predicción por observación: escenario, feature set, modelo, split, hospital, fecha, y_real, y_pred, error y métricas	22_prediccion_detallada_modelos_valid_test.csv
Resumen textual	Síntesis de resultados y métricas de predicción	22_prediccion_detallada_modelos_resumen.md

Producto	Contenido	Archivo
Gráficos de apoyo	Visualizaciones comparativas y diagnósticas	resumen_escenarios

El archivo de predicción detallada contiene 3.529.032 registros correspondientes a observaciones de validación y prueba. Este volumen permite mantener trazabilidad de los resultados por grupo hospitalario, fecha, modelo, escenario y conjunto de variables.

Tabla 3-30

Vista general del archivo de predicción detallada

A	B	C	D	E	F	G	H	I	J	K	L	M	N
data_scenario	feature_set	modelo	split	hospital_id	fecha	y_real	y_pred	n_features	MAE	RMSE	WMAPE	grupo_hospit	area_ubi
con_outliers	all	Naive_t	test	4	1/7/2024	0.203883495	0.067961165	140	0.077023720	0.120279184	0.468955113	4	URBANA
con_outliers	all	Naive_t	test	9	1/7/2024	0.1	0.05	140	0.077023720	0.120279184	0.468955113	9	URBANA
con_outliers	all	Naive_t	test	13	1/7/2024	0.16	0.12	140	0.077023720	0.120279184	0.468955113	13	URBANA
con_outliers	all	Naive_t	test	16	1/7/2024	0.020833333	0.041666666	140	0.077023720	0.120279184	0.468955113	16	URBANA
con_outliers	all	Naive_t	test	17	1/7/2024	0.076923076	0.128205128	140	0.077023720	0.120279184	0.468955113	17	URBANA

Entre las columnas principales del archivo se encuentran: data_scenario, feature_set, modelo, split, hospital_id o grupo_hospital_id, fecha, y_real, y_pred, error, abs_error y ape. Estas variables permiten calcular métricas posteriores, identificar errores extremos y comparar el desempeño entre configuraciones.

La lógica general de ejecución consistió en recorrer los escenarios seleccionados, cargar las particiones de entrenamiento, validación y prueba, preparar las variables explicativas, entrenar los modelos y generar predicciones para validación y prueba. El siguiente fragmento resume el proceso:

```
for data_scenario, data_dir in selected_scenarios.items():
    train_df = limitar_filas(pd.read_parquet(train_path), max_train_rows)
    valid_df = limitar_filas(pd.read_parquet(valid_path), max_valid_rows)
    test_df = limitar_filas(pd.read_parquet(test_path), max_test_rows)

    y_train = pd.to_numeric(train_df[TARGET_COL], errors="coerce")
```

```
    y_valid = pd.to_numeric(valid_df[TARGET_COL], errors="coerce")
    y_test = pd.to_numeric(test_df[TARGET_COL], errors="coerce")

    context_cols, top_causas = columnas_contexto(train_df)

    for feature_set, selector in selected_feature_sets.items():
        x_train, x_valid, x_test, feature_cols = preparar_features(
            train_df, valid_df, test_df, selector
        )

        for nombre, (pred_valid, pred_test) in predicciones_baseline(valid_df, test_df).items():
            detalles.append(armar_detalle(...))

        for nombre, model in modelos_ml().items():
            model.fit(x_train, y_train)
            pred_valid = model.predict(x_valid)
            pred_test = model.predict(x_test)
            detalles.append(armar_detalle(...))
```

Posteriormente, cada predicción se organizó en una tabla detallada por observación. Esta tabla incluye los valores reales, los valores predichos y los errores asociados.

```
out["error"] = out["y_pred"] - out["y_real"]
    out["abs_error"] = np.abs(out["error"])
    out["ape"] = np.abs(out["error"].values) / denom
    return out
```

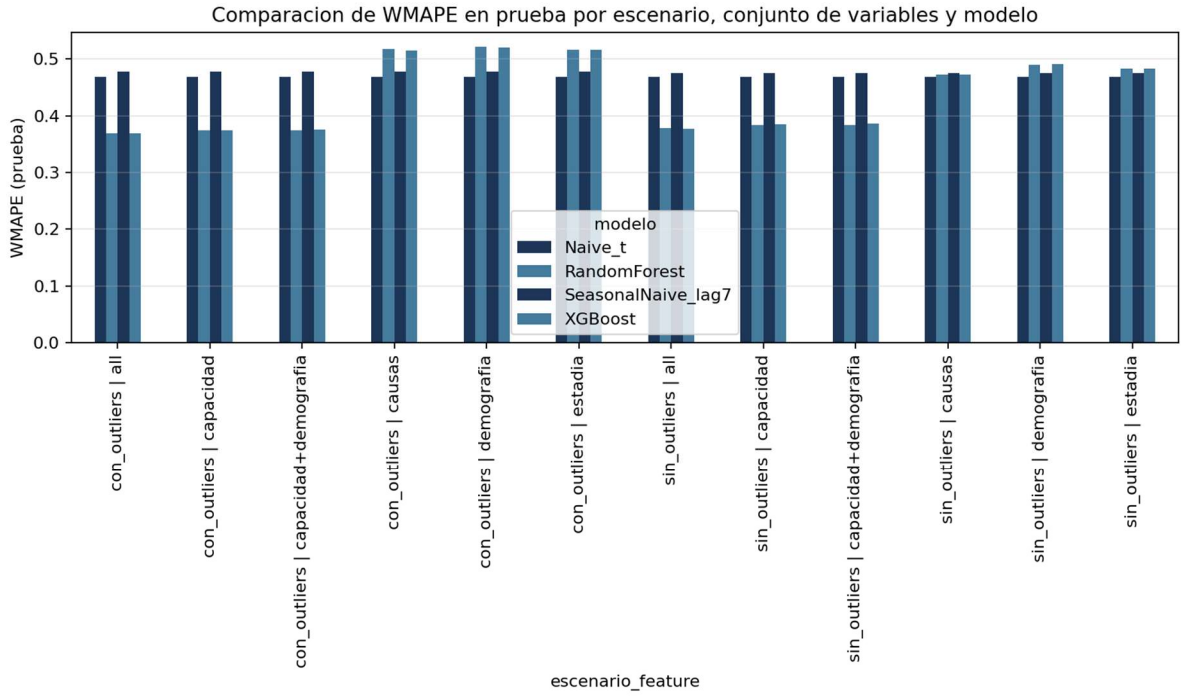
También se incluyeron predicciones de modelos base dentro del mismo flujo, con el fin de conservar comparabilidad entre modelos simples y modelos de Machine Learning. Naive_t utiliza el valor contemporáneo disponible de la serie, mientras que SeasonalNaive_lag7 utiliza el valor rezagado siete días.

3.6.1.1. Evidencia gráfica asociada

La comparación global del error en prueba se resume mediante una visualización del WMAPE por combinación de escenario, conjunto de variables y modelo.

Figura 3-40

resumen_wmape_test

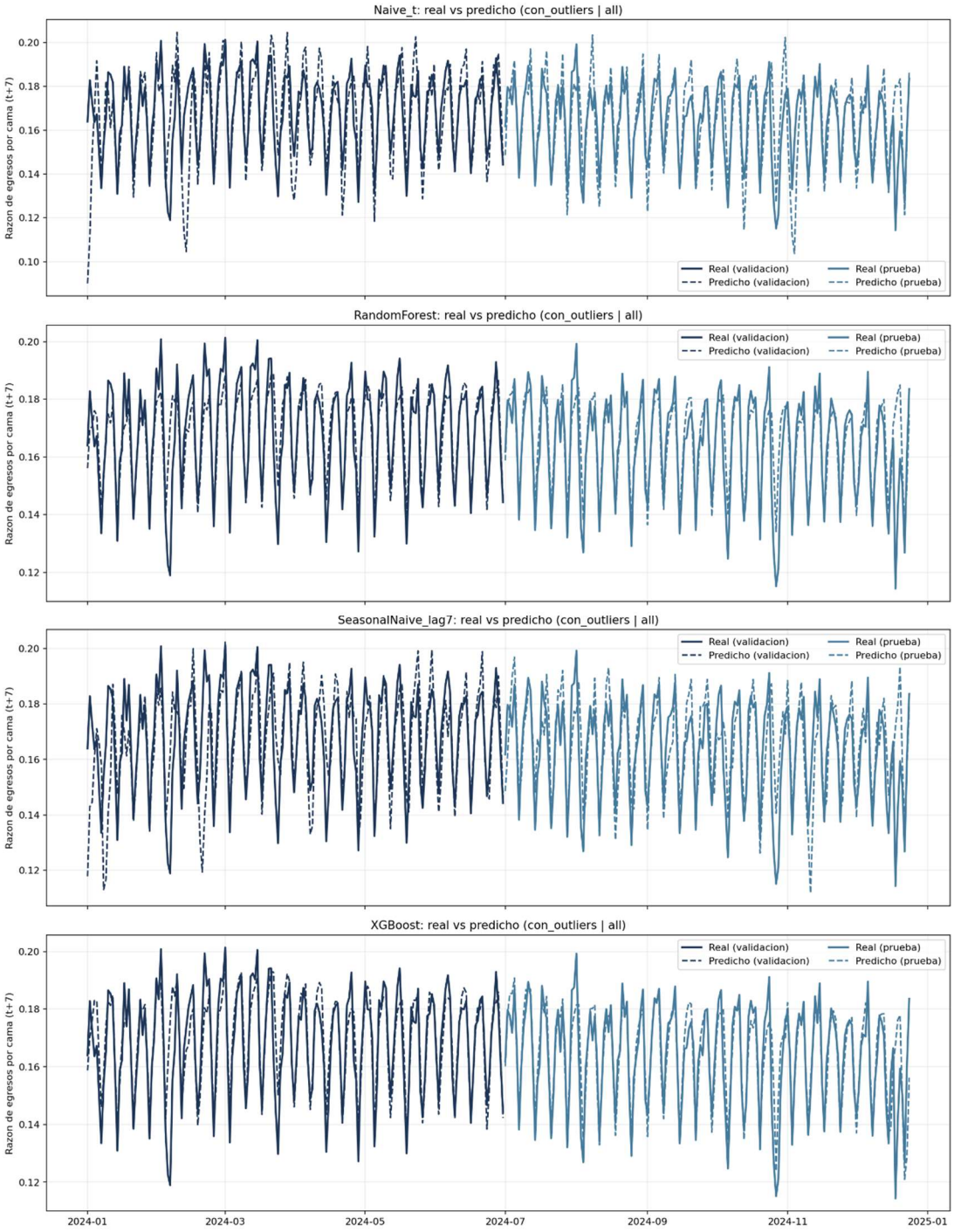


Esta figura permite identificar de forma directa las configuraciones con menor error. La lectura principal confirma que los mejores desempeños se concentran en combinaciones con el conjunto de variables all, y que la diferencia entre XGBoost y RandomForest es reducida en las configuraciones líderes.

Adicionalmente, se generaron series de predicción para comparar visualmente el comportamiento entre valores reales y predichos.

Figura 3-41

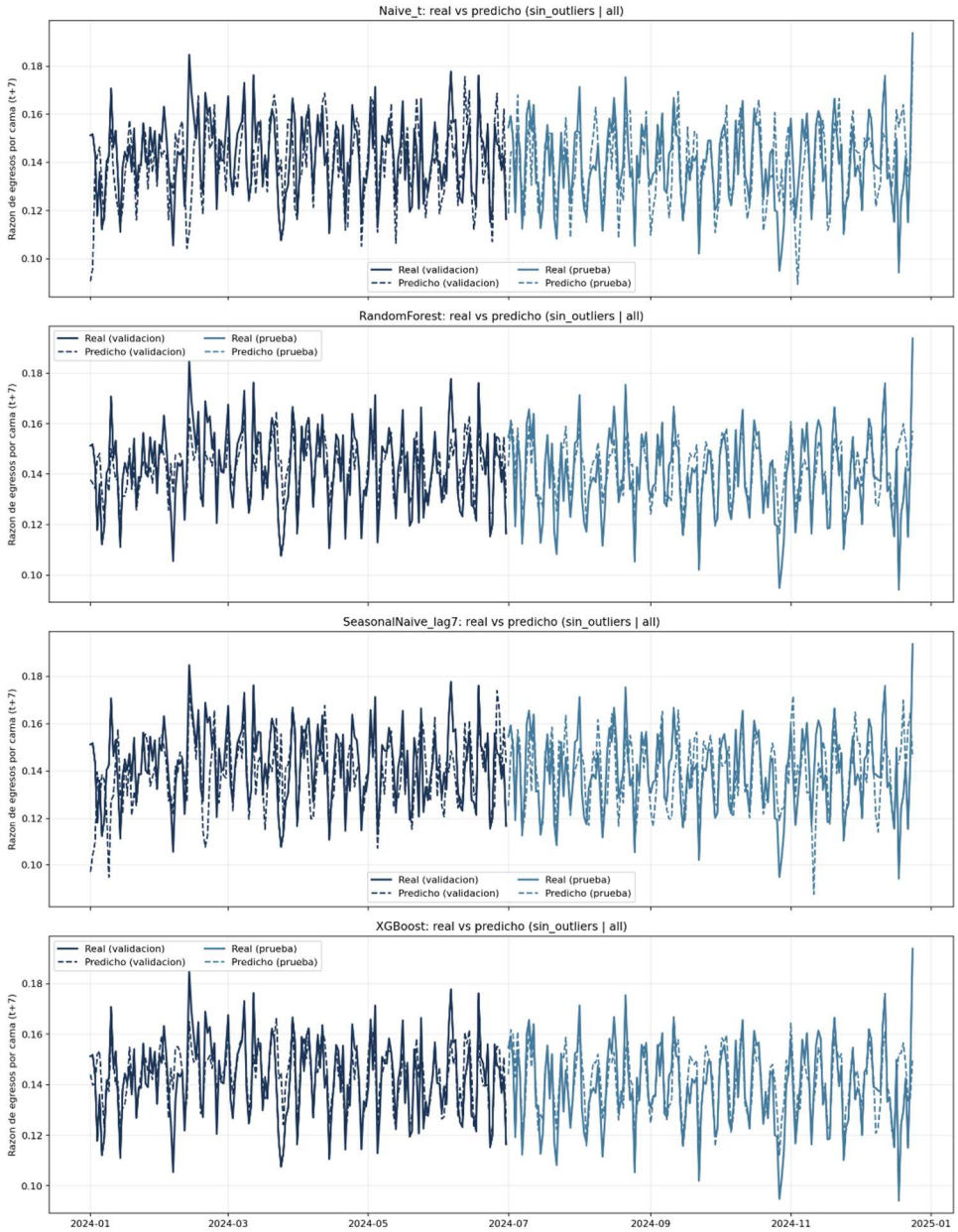
Predicciones_lineas_con_outliers_all



En el escenario con outliers y con el conjunto completo de variables, la serie predicha reproduce la dinámica general de la serie real y conserva sensibilidad frente a variaciones temporales. Esta visualización respalda que el conjunto all concentra suficiente señal predictiva incluso en presencia de valores extremos.

Figura 3-42

Predicciones_lineas_sin_outliers



En el escenario sin outliers y con el conjunto completo de variables, la serie predicha mantiene coherencia temporal y reduce el efecto de valores extremos. Esta lectura es consistente con el mejor desempeño promedio global del escenario sin outliers en las métricas agregadas.

En síntesis, la generación de predicciones permite pasar del benchmark comparativo a una salida analítica trazable, interpretable y útil para revisar presión hospitalaria esperada por grupo hospitalario y fecha.

3.6.2 Indicador de relación demanda-capacidad

El indicador operativo de esta fase es la presión asistencial esperada, expresada como relación demanda-capacidad. Dado que la variable objetivo del modelo corresponde a la razón entre egresos hospitalarios y camas disponibles, la predicción generada por el modelo puede interpretarse directamente como un indicador de presión hospitalaria futura.

En la salida final, la variable demanda_capacidad equivale al valor predicho por el modelo, representado por y_pred. Por tanto, no se introduce una nueva variable independiente, sino una lectura operativa y gerencial de la predicción.

```
detalle["demanda_capacidad"] = detalle["y_pred"]
detalle["error"] = detalle["y_pred"] - detalle["y_real"]
detalle["abs_error"] = np.abs(detalle["error"])
```

Las variables principales asociadas al indicador son las siguientes:

Tabla 3-31

Principales variables asociadas al indicador

Variable	Descripción
y_pred	Razón egresos/camas estimada por el modelo
y_real	Razón observada
demanda_capacidad	Indicador operativo de presión asistencial esperada
error	Diferencia entre predicción y valor real
abs_error	Error absoluto por observación
ape	Error porcentual absoluto

Además del valor continuo del indicador, se generó una clasificación operativa por niveles de presión. Esta clasificación permite traducir la predicción numérica en categorías más fáciles de interpretar para análisis y priorización.

La clasificación se realizó mediante cuantiles calculados sobre el conjunto de prueba. Los valores por debajo de la mediana se clasificaron como presión baja; los valores entre la mediana y el percentil 90 como presión media; y los valores iguales o superiores al percentil 90 como presión alta.

```
p50 = detalle_test["demanda_capacidad"].quantile(0.50)
p90 = detalle_test["demanda_capacidad"].quantile(0.90)


detalle_test["nivel_presion"] = np.select(
    [
        detalle_test["demanda_capacidad"] < p50,
```

```
        (detalle_test["demanda_capacidad"] >= p50) & (detalle_test["demanda_capacidad"] <
p90),
        detalle_test["demanda_capacidad"] >= p90,
    ],
    ["Baja", "Media", "Alta"],
    default="Sin clasificar",
)
```

Bajo esta lógica, un valor alto de demanda_capacidad indica mayor presión esperada de egresos frente a la capacidad disponible de camas. En cambio, un valor bajo refleja una presión relativa menor. Esta clasificación no debe interpretarse como ocupación hospitalaria real, sino como una señal comparativa para identificar grupos hospitalarios y fechas que podrían requerir mayor atención operativa.

3.6.3 Interpretación para la planificación hospitalaria

La presente sección desarrolla la interpretación operativa de los resultados del modelo seleccionado, con el propósito de transformar la predicción de demanda hospitalaria en insumos útiles para la planificación y la toma de decisiones. Para ello, se analiza primero el desempeño global del modelo XGBoost mediante métricas de error, residuales y calibración; posteriormente, se examinan diferencias territoriales por área geográfica y se identifican grupos hospitalarios sensibles según presión y error acumulado. A partir de esta base, se construye un esquema de priorización relativa mensual mediante semáforos operativos, el cual permite clasificar los grupos hospitalarios en niveles de prioridad alta, media y base dentro de cada mes de análisis. Finalmente, la interpretación se amplía hacia una lectura territorial y clínica, incorporando rankings por provincia, tendencias mensuales, causas predominantes de egreso, focos provincia-causa y desagregaciones por grupo hospitalario. De

esta manera, el capítulo no solo evalúa el comportamiento estadístico del modelo, sino que traduce sus resultados en criterios accionables para monitoreo, alerta temprana y planificación preventiva de recursos hospitalarios.

Para facilitar la lectura operativa de los resultados, se empleó un sistema de semaforización relativa mensual. Este sistema no busca clasificar el riesgo hospitalario en términos absolutos, sino establecer prioridades comparativas dentro de cada mes, permitiendo identificar qué grupos requieren atención preferente frente al resto de grupos evaluados en el mismo período.

Los niveles de prioridad se interpretan así:

- ROJO_REL: prioridad alta del mes. El grupo se ubica entre los casos con mayor presión comparativa del periodo y requiere intervención inmediata.
- AMARILLO_REL: prioridad media del mes. El grupo requiere vigilancia reforzada y preparación preventiva.
- VERDE_REL: prioridad base del mes. El grupo no está entre los más urgentes del corte, pero mantiene seguimiento rutinario.

El término relativo significa que la comparación se hace dentro del mismo mes. En otras palabras, un grupo queda en rojo porque está entre los más presionados de ese mes, no porque tenga necesariamente el mismo valor absoluto que un rojo de otro mes.

3.6.3.1 Lectura global del desempeño

El modelo de referencia seleccionado para la etapa operativa fue XGBoost, bajo el escenario con_outliers y con el conjunto completo de variables.

Tabla 3-32

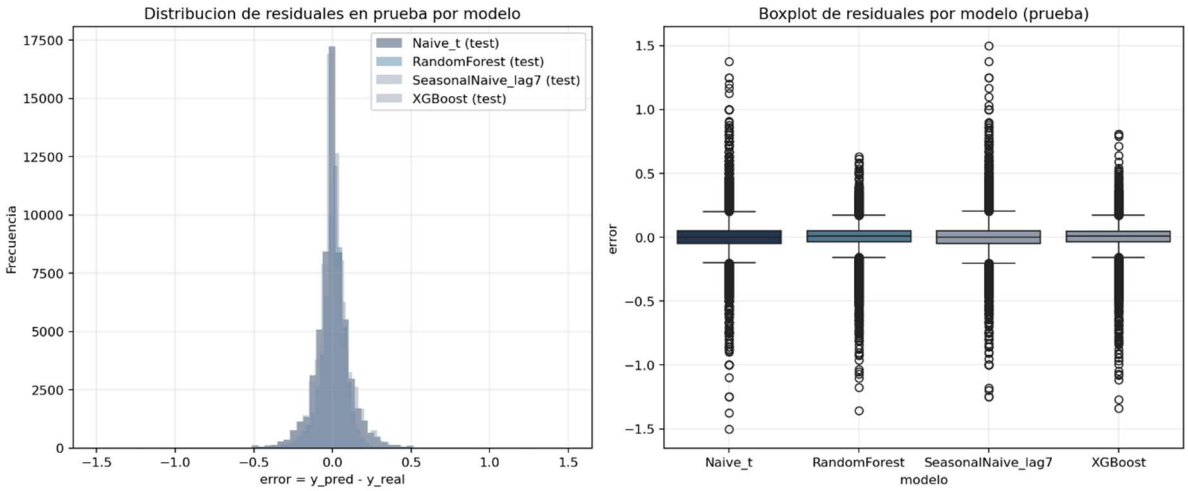
Modelo de Referencia seleccionado

Escenario	Split	Modelo	Feature set	WMAPE	RMSE	MAE	n_features	n_test
con_outliers	test	XGBoost	all	0.369118	0.088753	0.060626	122	59,989

La tabla anterior la distribución de los errores de predicción (residuales) en los modelos evaluados. Estos errores representan la diferencia entre la demanda estimada por el modelo y la demanda real observada. En la parte izquierda se observa que la mayoría de errores se concentra cerca de cero, lo que indica que las predicciones suelen mantenerse próximas a los valores reales. En la parte derecha, el boxplot compara los errores de los cuatro modelos: Naive_t, RandomForest, SeasonalNaive_lag7 y XGBoost. XGBoost presenta una mediana cercana a cero y una dispersión moderada, lo que respalda su estabilidad como modelo seleccionado. Aunque existen valores extremos en todos los modelos, estos reflejan casos puntuales de alta variabilidad hospitalaria. De esta forma el estudio realiza un análisis con la opción de casos especiales (outliers) y sin estos.

Figura 3-43

imagen de residuales

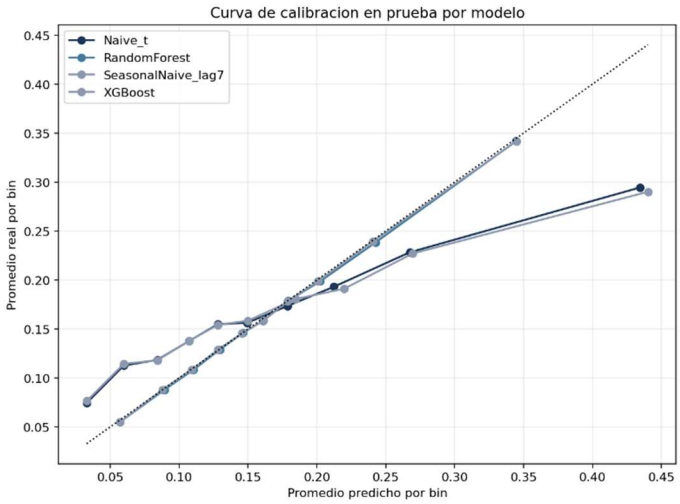


La Figura anterior presenta la curva de calibración de los modelos evaluados, comparando el promedio predicho con el promedio real observado en distintos rangos. La línea punteada representa una calibración ideal, donde la predicción coincide con el valor real. Se observa que XGBoost mantiene una trayectoria cercana a esta línea en varios puntos, especialmente en los rangos bajos y medios, lo que respalda su utilidad para estimar niveles relativos de

presión hospitalaria. En los valores más altos se aprecian ciertas desviaciones, por lo que estas predicciones deben interpretarse como señales de alerta operativa y no como valores exactos.

Figura 3-44

Imagen de calibración



La curva de calibración compara el promedio predicho con el promedio real por tramo: RandomForest y XGBoost se mantienen próximos a la diagonal ideal en la mayor parte del rango, mientras las referencias naive se desvían en los tramos altos. Esto indica que lo que los modelos de aprendizaje señalan como presión alta tiende a corresponder a presión alta observada.

Sección de Código relacionada:

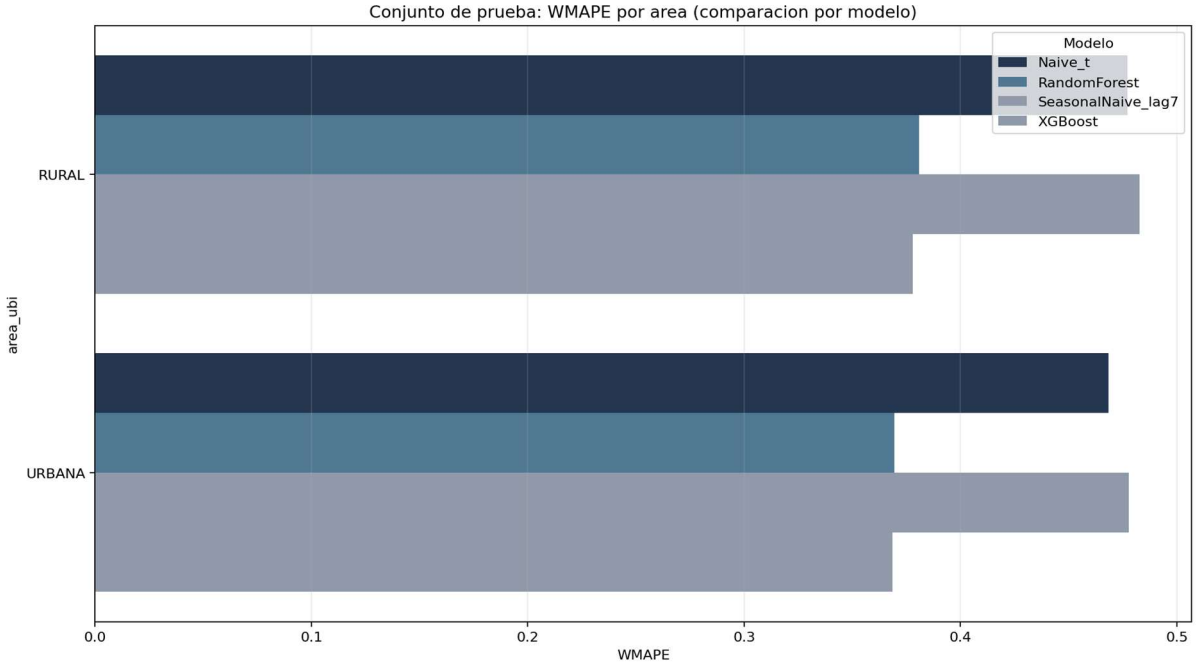
```
filtro = (
    (bench["split"] == "test")
    & (bench["data_scenario"] == "con_outliers")
    & (bench["feature_set"] == "all")
    & (bench["modelo"] == "XGBoost")
)
resumen_modelo = bench.loc[filtro, ["wmape", "rmse", "mae", "n_features", "n_test"]]
print(resumen_modelo)
```

3.6.3.2 Diferencias territoriales por área geográfica

La Figura 3.6.3-3 evidencia que RandomForest y XGBoost presentan mejores resultados de WMAPE frente a los modelos base Naive_t y SeasonalNaive_lag7, tanto en el área urbana como rural. Aunque las diferencias entre áreas no son extremas, el gráfico confirma que el desempeño del modelo puede variar según el contexto territorial. Por ello, la evaluación por área permite identificar dónde las predicciones requieren mayor seguimiento operativo.

Figura 3-45

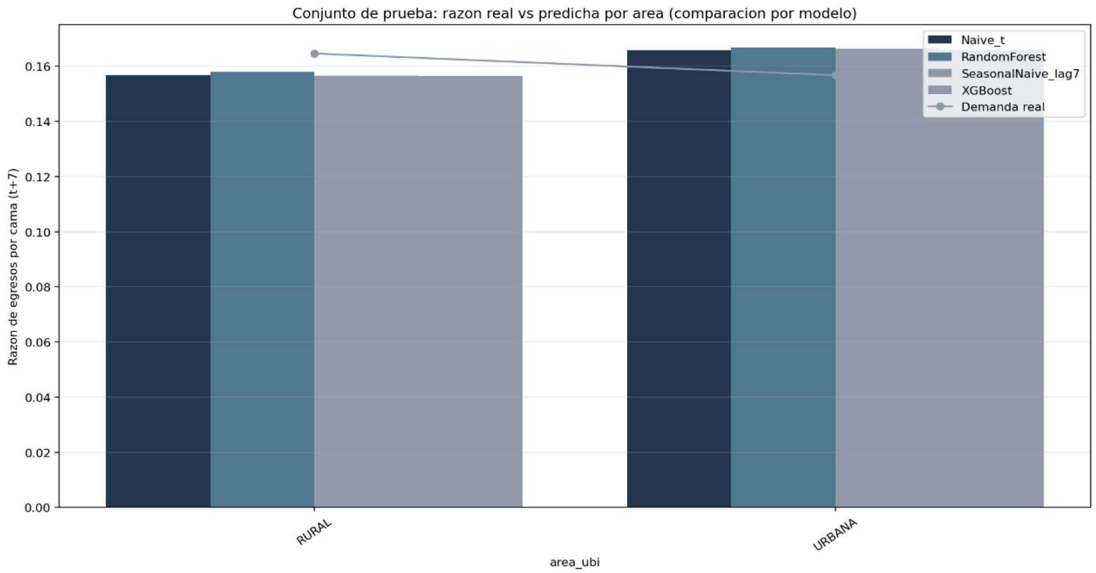
WMAPE por área



La Figura anterior compara la razón real y la razón predicha de egresos por cama en las áreas rural y urbana. Se observa que los modelos generan valores cercanos a la demanda real, especialmente en el área urbana, donde las predicciones se aproximan mejor al valor observado. En el área rural existe una ligera diferencia entre la razón real y la predicha, lo que sugiere mayor variabilidad o menor estabilidad en ese segmento. En general, el gráfico muestra que el modelo mantiene una aproximación razonable por área geográfica, aunque las predicciones deben interpretarse como una referencia de apoyo para la planificación y no como valores exactos.

Figura 3-46

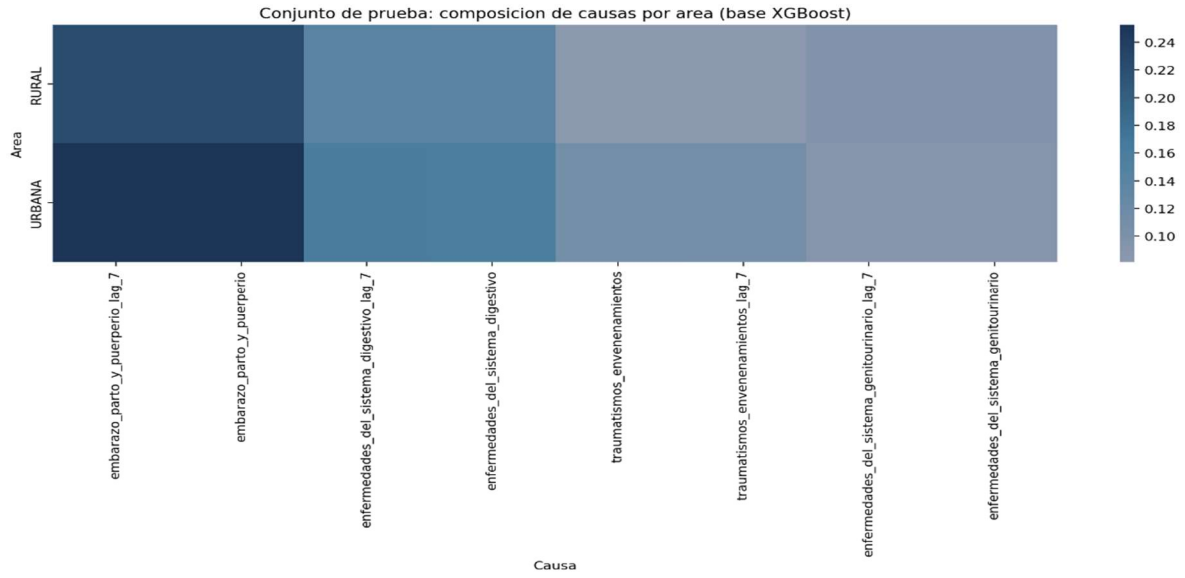
Demanda real vs predicha por área



El mapa de calor evidencia heterogeneidad en la composición de causas entre áreas, lo que aporta contexto para explicar variaciones de desempeño por segmento territorial. Esta lectura sugiere que la calidad predictiva no depende únicamente del modelo, sino también de diferencias estructurales en el perfil de egresos que conviene considerar en decisiones de planificación. (ver siguiente figura)

Figura 3-47

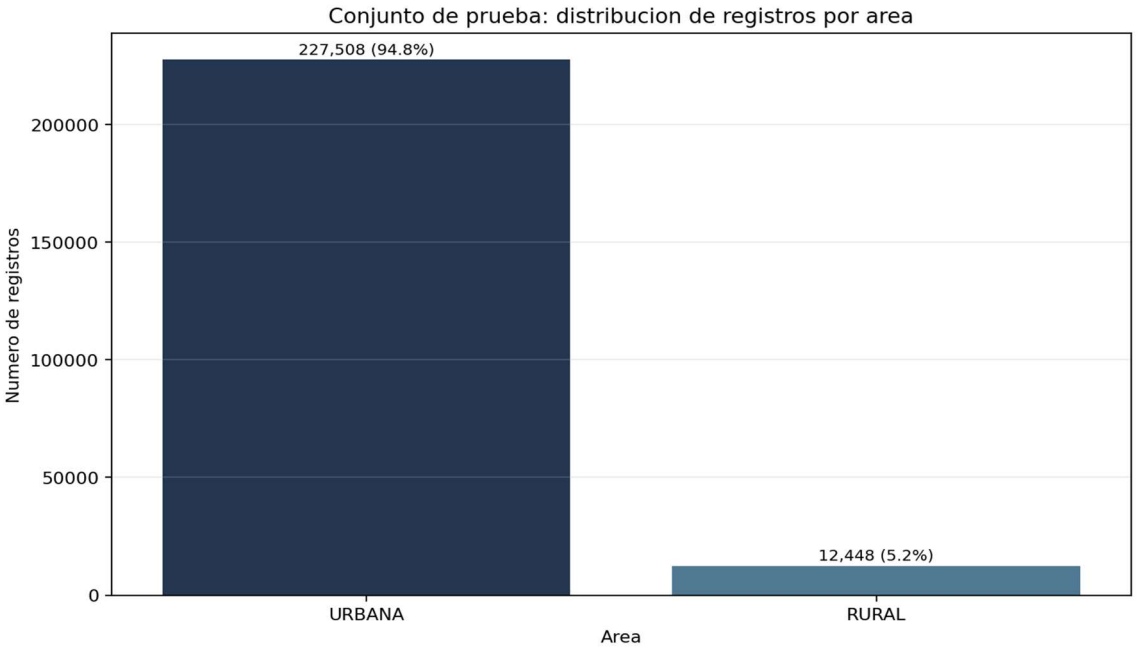
Mapa de calor de causas por área



La distribución de registros por área contextualiza la lectura de métricas, mostrando la concentración del volumen en ciertos segmentos territoriales. Esto justifica que la priorización de recursos de monitoreo se apoye no solo en error relativo, sino también en peso operativo por cantidad de casos.

Figura 3-48

Distribución de registros por área



Interpretación:

- La presión hospitalaria no se distribuye de manera homogénea entre territorios.
- El impacto práctico del error depende del nivel de actividad asistencial de cada zona.
- Esta diferencia justifica complementar la evaluación estadística con lectura territorial y operativa. Preferiblemente a nivel de provincia para tener más visibilidad.

Sección de Código relacionada:

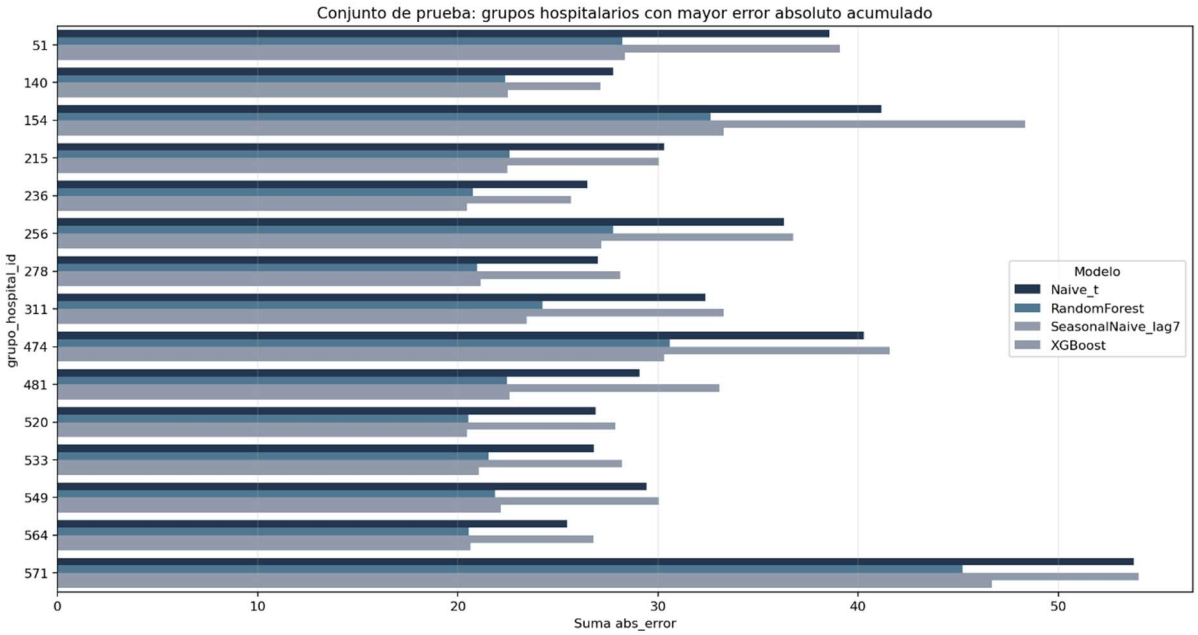
```
resumen_area = (
    det.groupby("area_ubi", as_index=False)
    .agg(wmape_area=("ape", "mean"), n_registros=("y_true", "count"))
    .sort_values("wmape_area", ascending=False)
)
```

3.6.3.3 Priorización territorial de riesgo por grupo hospitalario

La Figura siguiente permite identificar los grupos hospitalarios que concentran mayor error absoluto acumulado en el conjunto de prueba. Se observa que algunos grupos, como el 571, 154, 474 y 51, presentan errores más altos que el resto, lo que indica que en estos casos la predicción fue más difícil para los modelos evaluados. Esta situación puede estar asociada a mayor variabilidad en la demanda, comportamiento atípico o diferencias en la capacidad hospitalaria registrada. Por ello, estos grupos deben considerarse puntos sensibles para la planificación, ya que requieren mayor seguimiento, validación de datos y análisis operativo antes de tomar decisiones basadas únicamente en la predicción del modelo.

Figura 3-49

Top hospitales por error

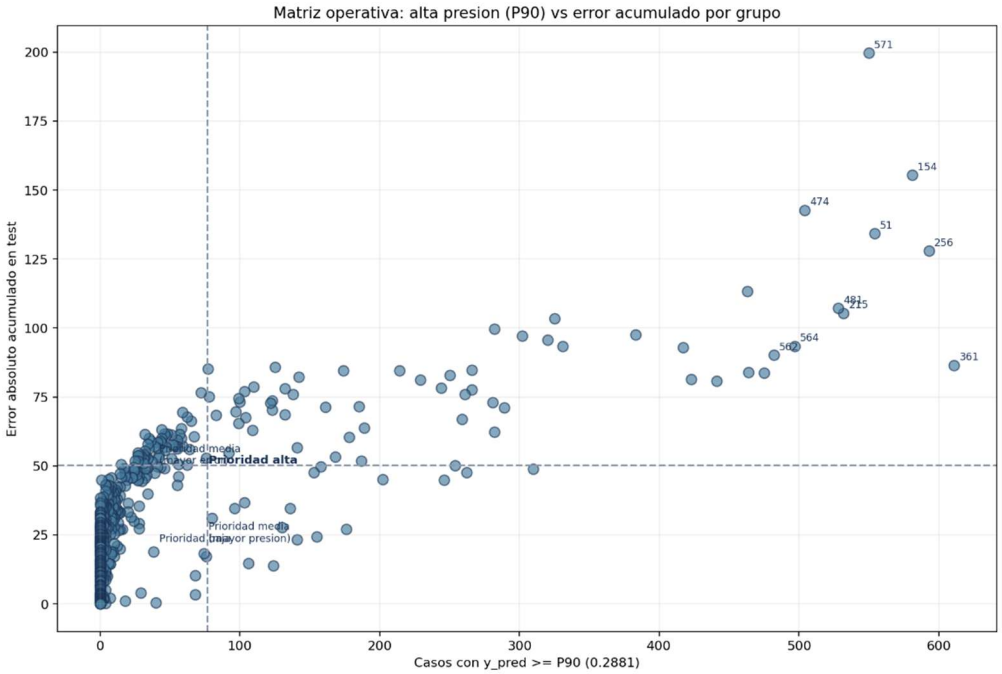


La Figura anterior presenta una matriz operativa que relaciona los casos de alta presión proyectada con el error absoluto acumulado por grupo hospitalario. Los grupos ubicados en la parte superior derecha, como 571, 154, 474, 51 y 256, combinan un alto número de casos de presión elevada con errores acumulados importantes, por lo que deben considerarse puntos críticos para seguimiento operativo. Estos grupos no solo concentran mayor demanda

proyectada, sino que además presentan mayor incertidumbre predictiva, lo que justifica una revisión más detallada de su capacidad, registros históricos y comportamiento de egresos. En cambio, los grupos cercanos al origen reflejan menor presión y menor error, por lo que requieren un monitoreo rutinario.

Figura 3-50

Matriz de priorización operativa



La matriz cruza el número de casos de alta presión proyectada (sobre el percentil 90) con el error absoluto acumulado por grupo, delimitando cuadrantes de prioridad. Los grupos del cuadrante superior derecho —alta presión y mayor error— son los candidatos naturales a monitoreo reforzado y validación de datos.

Sección de Código relacionada:

```
g = det.groupby("grupo_hospital_id", as_index=False).agg(
    casos_alta_presion=("is_alta", "sum"),
    abs_error_acumulado=("abs_error", "sum")
)
p75_alta = g["casos_alta_presion"].quantile(0.75)
p75_error = g["abs_error_acumulado"].quantile(0.75)
```

3.6.3.4 Priorización relativa mensual como criterio operativo principal

En la etapa de validación operativa se aplicó al conjunto de prueba del modelo una regla de clasificación basada en límites fijos sobre el indicador de relación demanda-capacidad. Para cada uno de los 461 grupos hospitalarios evaluados en el corte final, se calculó el valor proyectado del indicador y se lo clasificó con los umbrales definidos: verde para valores menores a 0.95, amarillo para valores entre 0.95 y 1.10, y rojo para valores iguales o superiores a 1.10. Esta regla se incorporó como referencia inicial porque ofrecía una interpretación directa y sencilla para la gestión hospitalaria, al comparar cada grupo contra un estándar único e independiente del mes analizado.

Sin embargo, al revisar los resultados se observó que la clasificación no aportaba capacidad discriminativa: los 461 grupos hospitalarios evaluados quedaron ubicados en nivel verde. Esto indicó que los límites fijos eran demasiado altos para la distribución observada en el corte final y que, en consecuencia, la regla no permitía distinguir con claridad qué grupos requerían mayor atención operativa. Por esa razón, el criterio absoluto no resultó útil para la priorización.

A partir de esa evidencia, se adoptó un criterio relativo mensual basado en percentiles (P85 y P95) calculados dentro de cada mes de referencia. A diferencia de la regla anterior, este enfoque no compara a todos los grupos contra un mismo valor fijo, sino que identifica cuáles presentan mayor presión dentro de su propio período. De este modo, la clasificación se ajusta mejor a la variación observada mes a mes y permite resaltar los casos que realmente requieren seguimiento o intervención.

La Tabla 3-31 muestra que la proporción de registros clasificados como prioridad alta y media se mantuvo cercana al 15 % durante todo el período evaluado. Esto confirma que el criterio relativo permite seleccionar de manera estable un subconjunto manejable de casos prioritarios por mes, independientemente del volumen mensual de actividad hospitalaria.

Tabla 3-33

Distribución mensual de prioridad relativa

Mes	Total registros evaluados	Alta	Media	Base	% prioridad alta + media
2024-07	10,760	538	1,076	9,146	15.000%
2024-08	10,737	537	1,074	9,126	15.004%
2024-09	10,364	519	1,036	8,809	15.004%
2024-10	10,668	534	1,067	9,067	15.008%
2024-11	10,102	506	1,010	8,586	15.007%
2024-12	7,358	368	736	6,254	15.004%

Interpretación: La tabla confirma que, mes a mes, el criterio relativo siempre marca aproximadamente al 15 % de los registros como prioritarios. Esto es útil porque la cantidad de casos señalados se mantiene estable aunque cambie el número total de registros en cada mes. En otras palabras, el método no se basa en un umbral fijo, sino en la comparación de cada caso con los demás registros del mismo período.

La sección de Código implementada es la siguiente:

```
filt["indice_presion_t_plus_7"] = pd.to_numeric(filt["y_pred"], errors="coerce")
filt["p85"] = filt["mes_referencia"].map(
    filt.groupby("mes_referencia")["indice_presion_t_plus_7"].quantile(0.85).to_di
ct()
)
filt["p95"] = filt["mes_referencia"].map(
    filt.groupby("mes_referencia")["indice_presion_t_plus_7"].quantile(0.95).to_di
ct()
)
filt["semaforo_relativo"] = filt.apply(
    lambda r: _semaforo_relativo(r["indice_presion_t_plus_7"], r["p85"],
r["p95"]), axis=1
)
```

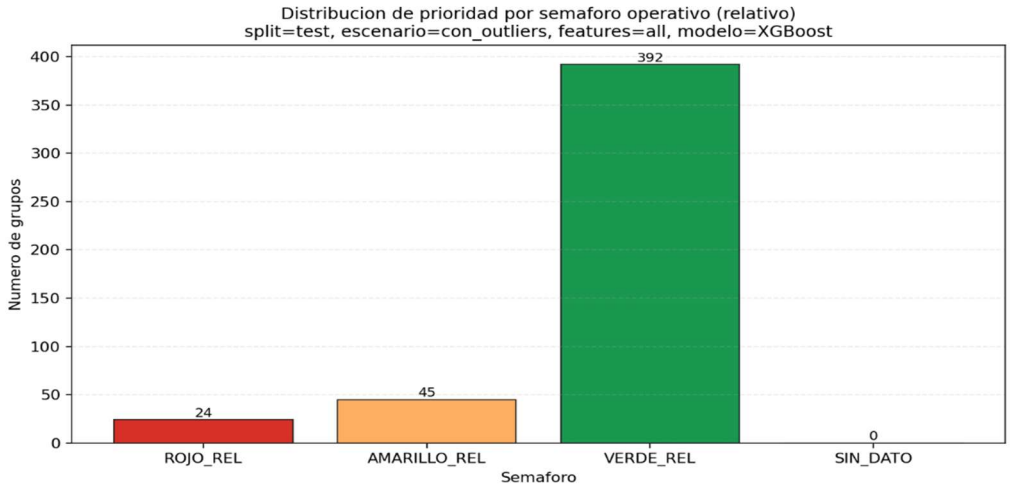
Se trabajó sobre un subconjunto filtrado de las predicciones del modelo, correspondiente al split de prueba y a la configuración final del experimento, que incluye el escenario con

outliers, el conjunto de variables completo y el modelo XGBoost. Este filtro se aplicó para conservar únicamente los registros válidos del corte evaluado, descartando valores faltantes y evitando que la clasificación relativa se calcule sobre información que no forma parte de la validación final. Sobre ese subconjunto se aplicó una regla de clasificación fila por fila, en la que cada valor de presión proyectada se comparó con los percentiles mensuales para asignar la categoría correspondiente.

3.6.3.4.1 Distribución global del semáforo relativo

Una vez definido el semáforo relativo mensual, se analizó su distribución general para identificar cuántos grupos hospitalarios quedaron clasificados en cada nivel de prioridad operativa. Esta visualización permite verificar si el criterio de percentiles separa adecuadamente los casos de mayor atención sin generar un exceso de alertas. Los grupos se clasifican en tres categorías: ROJO_REL, que representa prioridad alta; AMARILLO_REL, que representa prioridad media; y VERDE_REL, que corresponde a prioridad base. Estas categorías se calculan de forma relativa dentro de cada mes de referencia, por lo que la prioridad asignada depende de la comparación entre grupos evaluados en el mismo período. La distribución resultante se presenta en la siguiente figura.

Figura 3-51
Distribución del semáforo



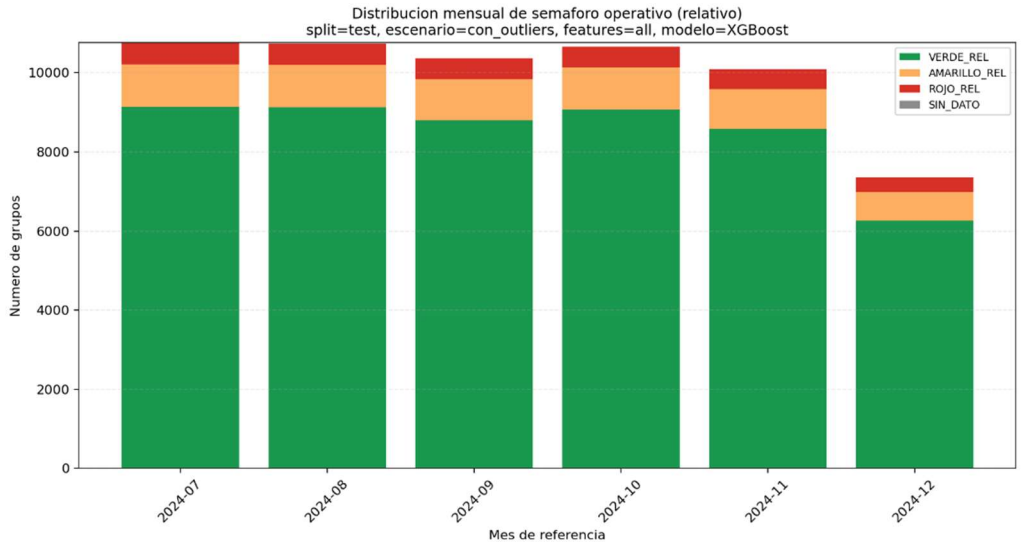
Interpretación: De los 461 grupos hospitalarios únicos evaluados, 24 grupos (5.2 %) alcanzaron prioridad alta, 45 grupos (9.8 %) prioridad media y 392 grupos (85 %) prioridad base. Esto evidencia que el semáforo relativo permite concentrar la atención en un subconjunto reducido de grupos con mayor presión comparativa. En términos operativos, aproximadamente el 15 % de los grupos requiere seguimiento prioritario, mientras que la mayoría permanece en monitoreo rutinario. Esta distribución evita saturar la gestión hospitalaria y facilita focalizar recursos en los casos más relevantes.

3.6.3.4.2 Consistencia temporal de la priorización relativa

La figura siguiente muestra la distribución mensual del semáforo relativo durante el período de prueba, comprendido entre julio y diciembre de 2024. Cada barra representa los registros evaluados en un mes y los segmenta según su nivel de prioridad operativa: alta, media o base. Esta visualización permite comprobar si la regla de priorización mantiene un comportamiento estable a lo largo del tiempo, aun cuando el volumen mensual de registros evaluados varía entre períodos.

Figura 3-52

Semáforo por mes



Interpretación:

La distribución mensual evidencia que la proporción de registros en prioridad alta y media se mantiene cercana al 15 % durante todos los meses analizados. Esto confirma que el semáforo relativo permite seleccionar de manera consistente un subconjunto manejable de casos prioritarios para seguimiento operativo. Aunque diciembre presenta un menor volumen de registros evaluados frente a los meses anteriores, la proporción de alertas se mantiene estable, lo que indica que la regla de priorización no depende del volumen total del mes, sino de la comparación relativa entre los registros de cada período.

Sección de Código relacionada:

```
det["mes_ref"] = pd.to_datetime(det["fecha"]).dt.to_period("M").astype(str)
p85 = det.groupby("mes_ref")["y_pred"].quantile(0.85)
p95 = det.groupby("mes_ref")["y_pred"].quantile(0.95)

det = det.merge(p85.rename("p85"), on="mes_ref").merge(p95.rename("p95"), on="mes_ref")
det["semaforo"] = np.where(
    det["y_pred"] >= det["p95"], "ROJO_REL",
    np.where(det["y_pred"] >= det["p85"], "AMARILLO_REL", "VERDE_REL")
)
```

3.6.3.5 Priorización por provincia y mes

La priorización se desagregó por provincia y mes con el fin de identificar en qué territorios se concentra con mayor frecuencia la prioridad operativa alta y media. Esta lectura permite pasar de una visión general del semáforo relativo a una interpretación territorial más útil para la planificación hospitalaria, ya que facilita reconocer provincias que requieren mayor monitoreo, preparación preventiva o revisión de capacidad durante el período analizado. La

siguiente tabla presenta el ranking completo de provincias según su porcentaje promedio de prioridad.

Tabla 3-34

Ranking completo de provincias por prioridad promedio

Rank	Provincia	% prioridad promedio
1	ORELLANA	46.41%
2	SUCUMBÍOS	44.71%
3	SANTA ELENA	25.68%
4	AZUAY	25.43%
5	PICHINCHA	24.55%
6	GUAYAS	16.32%
7	MANABÍ	14.50%
8	EL ORO	14.13%
9	SANTO DOMINGO DE LOS TSÁCHILAS	11.04%
10	ESMERALDAS	10.26%
11	CARCHI	10.26%
12	TUNGURAHUA	9.84%
13	IMBABURA	8.95%
14	LOS RÍOS	7.69%
15	LOJA	5.78%
16	CAÑAR	4.85%
17	COTOPAXI	4.65%
18	GALÁPAGOS	4.37%
19	MORONA SANTIAGO	1.55%
20	CHIMBORAZO	0.43%
21	NAPO	0.35%
22	BOLÍVAR	0.24%
23	PASTAZA	0.00%
24	ZAMORA CHINCHIPE	0.00%

Interpretación: El ranking muestra que Orellana y Sucumbíos presentan los mayores porcentajes promedio de prioridad, con valores superiores al 40 %, lo que indica una concentración recurrente de registros en niveles de alerta alta o media. En un segundo grupo aparecen Santa Elena, Azuay y Pichincha, con porcentajes cercanos al 25 %,

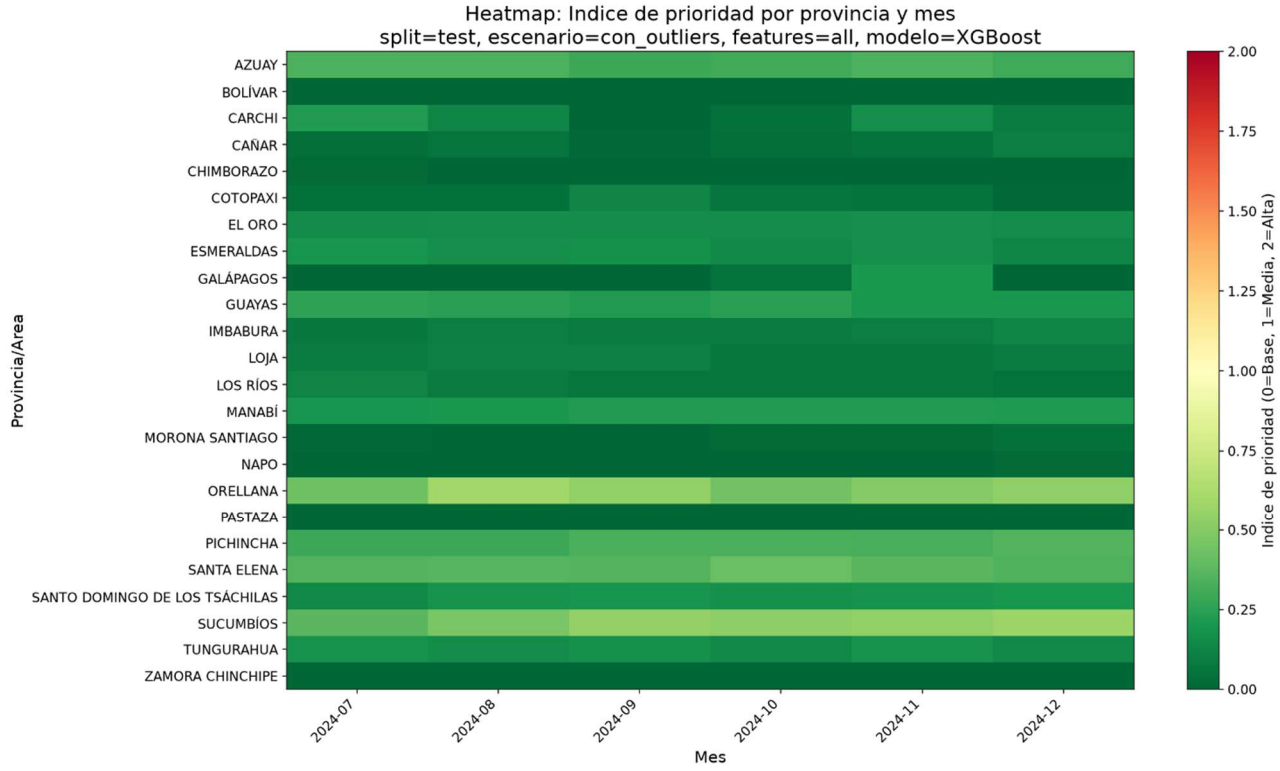
mientras que la mayoría de provincias se ubica por debajo del 15 %. Es importante señalar que un valor de 0.00 % no implica ausencia total de riesgo hospitalario, sino que, bajo el criterio relativo aplicado en el período de prueba, esas provincias no registraron casos clasificados en prioridad alta o media.

3.6.3.5.1 Distribución provincia-mes de la prioridad operativa

La Figura siguiente presenta un mapa de calor que cruza las 24 provincias con los seis meses del período de prueba. Cada celda muestra el índice promedio de prioridad operativa, donde 0 representa prioridad base, 1 prioridad media y 2 prioridad alta. Esta visualización permite identificar de forma rápida en qué provincias y meses se concentra con mayor intensidad la prioridad operativa.

Figura 3-53

Heatmap provincia-mes



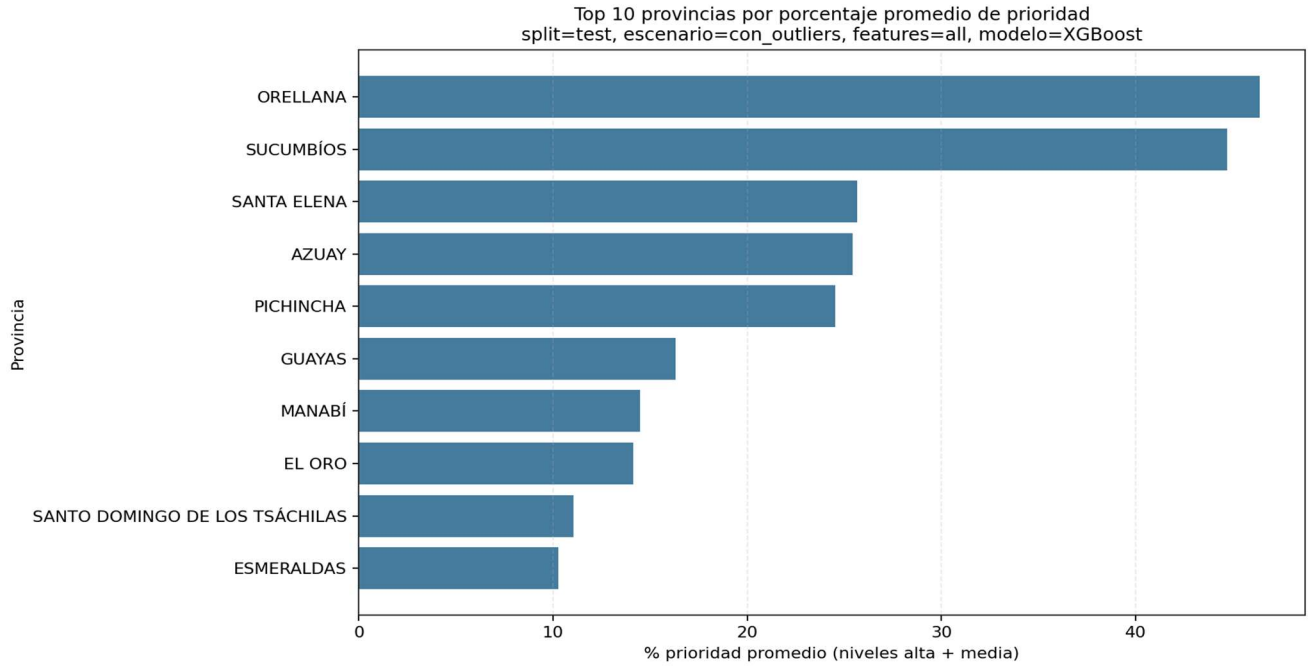
Interpretación: El mapa evidencia que Orellana presenta los índices de prioridad más altos durante la mayor parte del período, especialmente en agosto y septiembre. En contraste, la mayoría de provincias mantiene valores cercanos a cero, lo que sugiere que la prioridad operativa se concentra en pocos territorios. Pastaza y Zamora Chinchipe registran índice cero en todos los meses analizados, lo cual no implica ausencia total de demanda hospitalaria, sino que sus registros no alcanzaron niveles de prioridad alta o media bajo el criterio relativo aplicado.

3.6.3.5.2 Provincias con mayor prioridad promedio

La Figura siguiente presenta las diez provincias con mayor porcentaje promedio de registros clasificados en prioridad alta o media durante el período de prueba. Esta visualización permite identificar los territorios que concentran con mayor frecuencia alertas operativas y facilita establecer una primera lectura comparativa para la planificación hospitalaria.

Figura 3-54

Top provincias por prioridad promedio



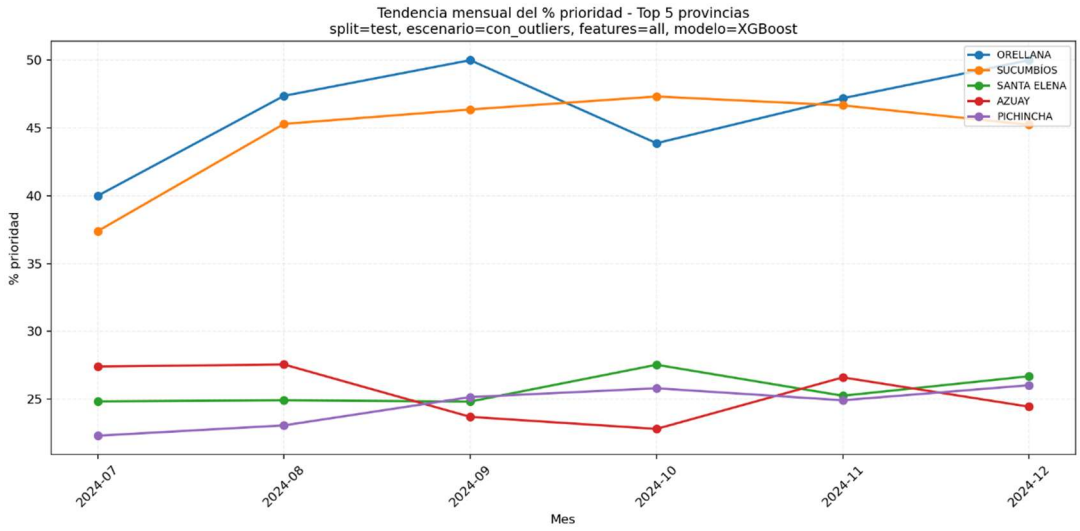
Interpretación: Orellana y Sucumbíos registran los porcentajes promedio más altos, con valores superiores al 40 %, lo que las ubica como las provincias de mayor recurrencia de prioridad operativa. A partir del tercer lugar, representado por Santa Elena, el porcentaje desciende a aproximadamente 25 %, evidenciando una brecha relevante frente a las dos primeras provincias. Esta diferencia permite distinguir un primer grupo territorial que requiere seguimiento prioritario y un segundo grupo con presión moderada, pero todavía relevante para la gestión hospitalaria.

3.6.3.5.3 Tendencia mensual de las provincias críticas

La Figura siguiente muestra la evolución mensual del porcentaje de registros clasificados en prioridad alta o media para las cinco provincias con mayor promedio de alerta durante el período de prueba. Esta visualización permite identificar si la prioridad operativa se mantiene estable, aumenta o disminuye en el tiempo, y ayuda a diferenciar territorios con alertas recurrentes frente a aquellos con variaciones puntuales.

Figura 3-55

Tendencia top 5 provincias



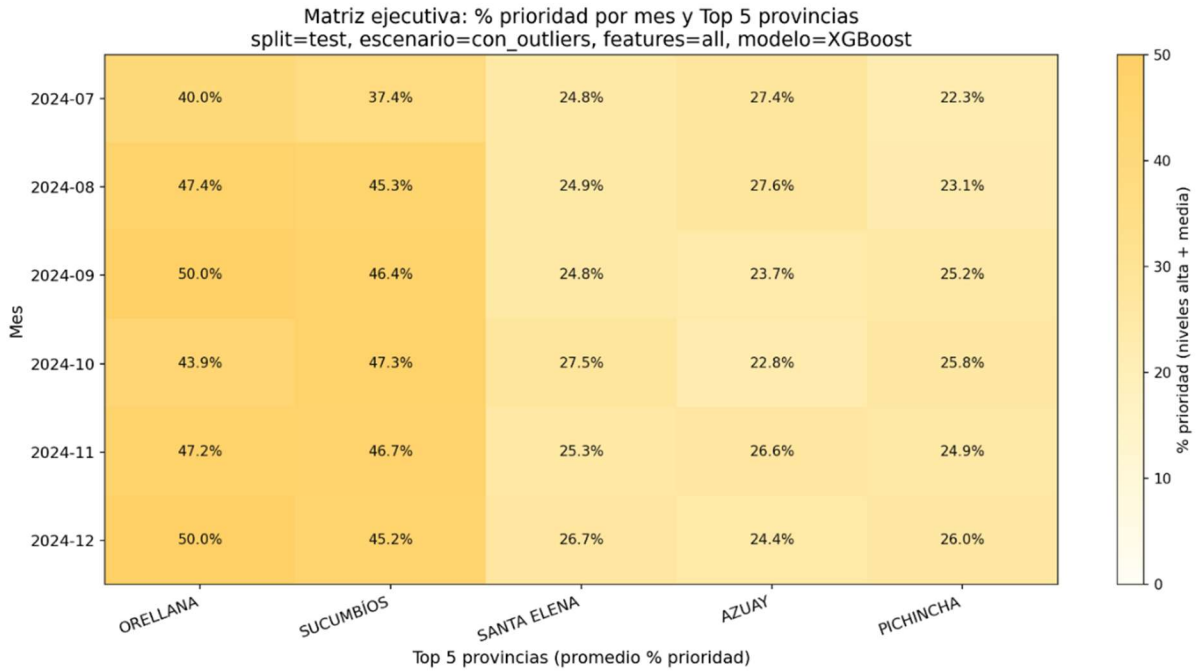
Interpretación: Orellana y Sucumbíos mantienen porcentajes elevados durante todo el semestre, con valores entre aproximadamente 37 % y 50 %, sin evidenciar una reducción sostenida. Orellana alcanza sus valores más altos en septiembre y diciembre, lo que indica meses de mayor prioridad relativa dentro del período analizado. Santa Elena, Azuay y Pichincha se mantienen alrededor del 25 %, con variaciones moderadas, lo que sugiere una prioridad territorial estable, aunque menor que la observada en Orellana y Sucumbíos.

3.6.3.5.4 Matriz ejecutiva de provincias prioritarias

La Figura siguiente presenta un mapa de calor con las cinco provincias de mayor prioridad promedio durante el período de prueba. Cada celda muestra el porcentaje mensual de registros clasificados en prioridad alta o media, acompañado de una codificación de color que facilita la comparación visual entre provincias y meses. Esta matriz funciona como una vista ejecutiva, ya que permite identificar rápidamente los territorios y períodos que requieren mayor seguimiento operativo.

Figura 3-56

Matriz ejecutiva top 5



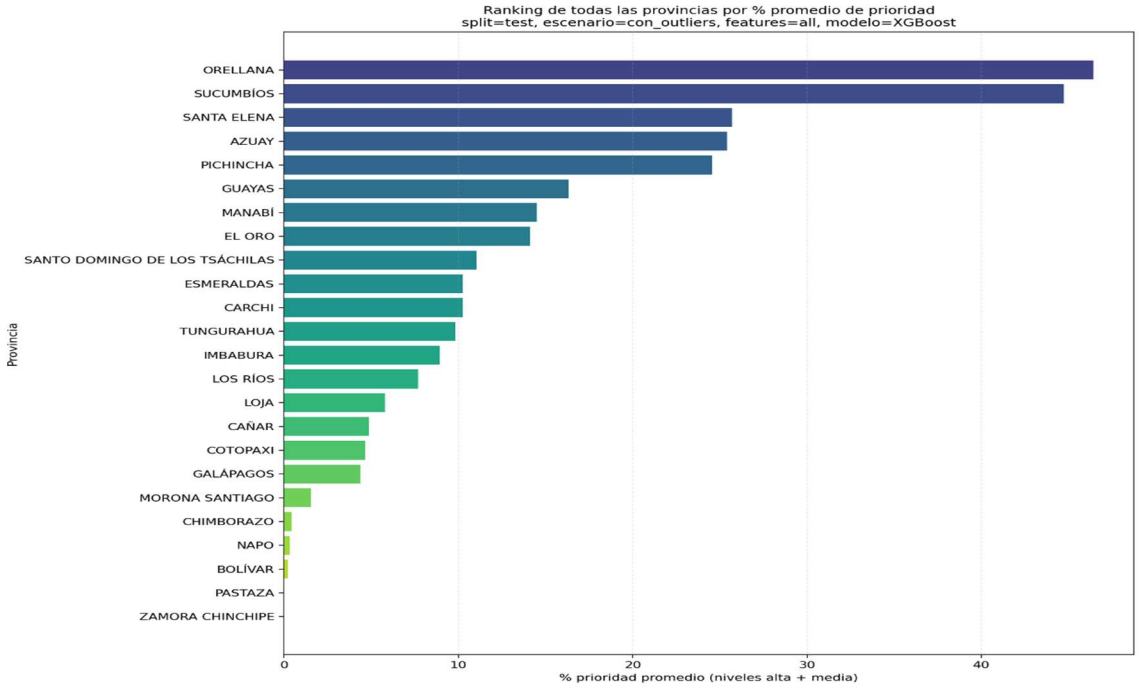
Interpretación: Orellana presenta los valores más altos de la matriz, alcanzando 50.0 % en septiembre y diciembre, lo que significa que en esos meses la mitad de sus registros evaluados se clasificaron en prioridad alta o media. Sucumbíos mantiene también porcentajes elevados durante todo el período, con valores entre 37 % y 47 %. En comparación, Santa Elena, Azuay y Pichincha se ubican alrededor del 25 %, mostrando una prioridad relevante pero menor. Esta diferencia evidencia que Orellana y Sucumbíos mantienen un patrón de prioridad más alto y persistente frente al resto de provincias analizadas.

3.6.3.5.5 Ranking nacional de prioridad por provincia

La Figura siguiente presenta el ranking completo de las 24 provincias, ordenadas de mayor a menor porcentaje promedio de registros clasificados en prioridad alta o media. Esta visualización permite comparar la distribución territorial de la prioridad operativa y ubicar tanto las provincias con mayor concentración de alertas como aquellas con menor presencia relativa durante el período de prueba.

Figura 3-57

Ranking de todas las provincias



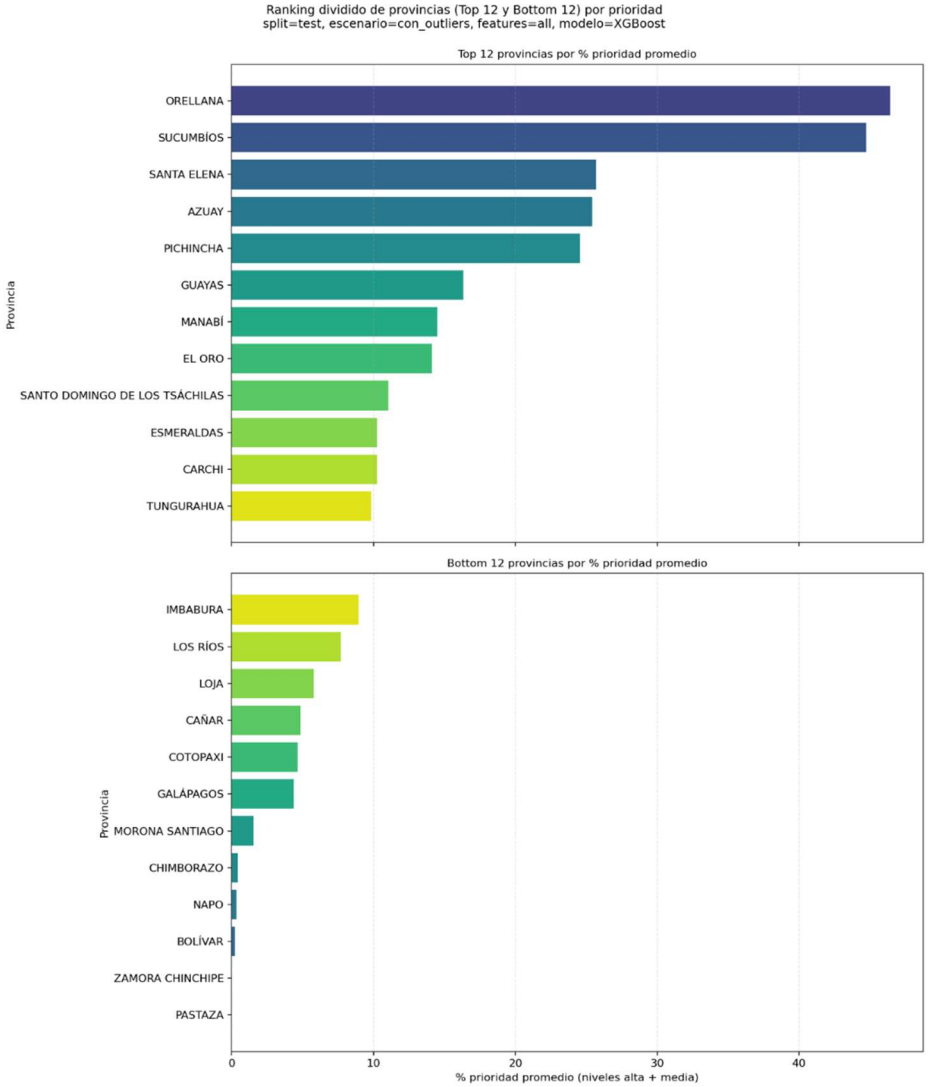
Interpretación: El ranking evidencia una diferencia marcada entre las dos primeras provincias y el resto del país. Orellana y Sucumbíos presentan los porcentajes más altos, con aproximadamente 46 % y 44 %, respectivamente, casi duplicando el valor observado en Santa Elena, que se ubica alrededor del 26 %. En el extremo inferior, provincias como Pastaza, Zamora Chinchipe, Bolívar, Napo, Chimborazo y Morona Santiago registran porcentajes inferiores al 2 %. Estos valores no implican ausencia total de demanda hospitalaria, sino una baja presencia de registros clasificados en prioridad alta o media bajo el criterio relativo aplicado en el período analizado.

3.6.3.5.6 Comparación entre provincias de alta y baja prioridad

La Figura siguiente divide el ranking provincial en dos paneles: las doce provincias con mayor prioridad promedio y las doce con menor prioridad promedio. Esta separación facilita la comparación visual entre territorios, evitando que las provincias con valores más altos oculten las diferencias existentes entre aquellas con menor porcentaje de registros priorizados. Su utilidad práctica es distinguir los territorios que requieren mayor seguimiento operativo de aquellos que, durante el período analizado, presentaron menor frecuencia de alertas relativas.

Figura 3-58

Ranking top 12 y bottom 12



Interpretación: El panel superior muestra que Orellana y Sucumbíos son las únicas provincias que superan el 40 % de registros clasificados en prioridad alta o media. Las demás provincias del grupo superior se ubican aproximadamente entre el 10 % y el 26 %, lo que evidencia un segundo nivel de prioridad territorial. En el panel inferior se observa un conjunto de provincias con porcentajes inferiores al 2 %, entre ellas Pastaza, Zamora Chinchipe, Bolívar, Napo, Chimborazo y Morona Santiago. Estos resultados no implican ausencia de demanda hospitalaria, sino una menor presencia de registros clasificados como prioritarios bajo el criterio relativo aplicado en el período de prueba.

Sección de Código relacionada:

```
det["is_alerta"] = det["semaforo"].isin(["ROJO_REL", "AMARILLO_REL"]).astype(int)


prov_mes = det.groupby(["provincia", "mes_ref"], as_index=False).agg(
    total=("is_alerta", "count"),
    alertas=("is_alerta", "sum")
)
prov_mes["pct_alerta"] = 100 * prov_mes["alertas"] / prov_mes["total"]
```

3.6.3.6 Priorización accionable por provincia, mes y grupo

La salida operativa final permitió identificar 720 registros accionables a nivel provincia-mes-grupo hospitalario. Estos registros corresponden a combinaciones específicas que requieren seguimiento diferenciado según su nivel de prioridad relativa. La Tabla 3.6.3-4 resume la distribución de estos casos por clase operativa, diferenciando prioridad alta, media y base. Esta clasificación permite transformar la predicción del modelo en una lista concreta de casos que pueden ser utilizados para monitoreo, revisión de capacidad y planificación preventiva.

Tabla 3-35

Distribución de casos accionables por clase operativa

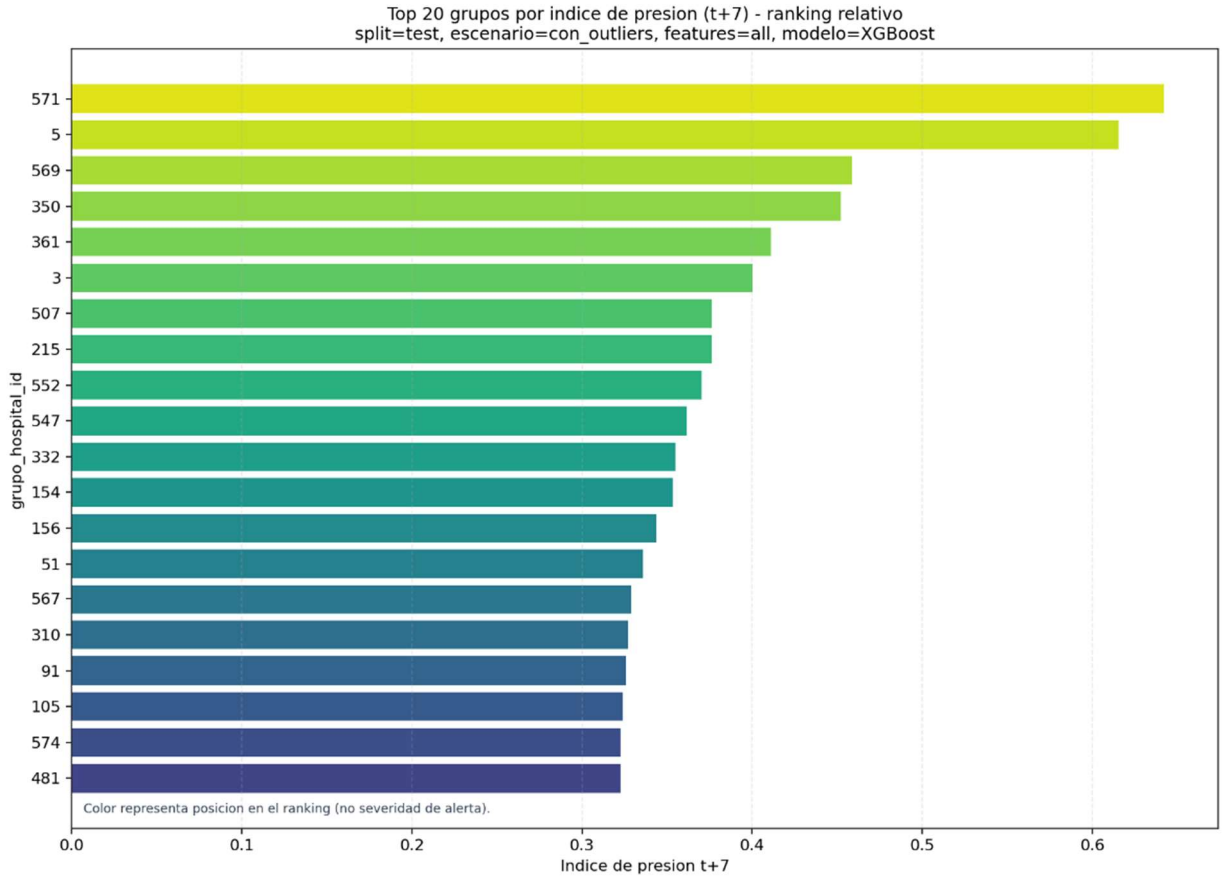
Clase operativa	Conteo
ROJO_REL	358
AMARILLO_REL	174
VERDE_REL	188

3.6.3.6.1 Grupos hospitalarios con mayor presión proyectada

La Figura siguiente presenta los 20 grupos hospitalarios con mayor índice de presión proyectada para el horizonte de predicción t+7. Este índice relaciona la demanda estimada por el modelo con la capacidad residual disponible, por lo que permite identificar los grupos donde la presión esperada podría ser más relevante en el siguiente período mensual. La visualización facilita priorizar aquellos grupos que requieren mayor atención operativa antes de que la demanda se materialice.

Figura 3-59

Top 20 por índice de presión



Interpretación: El grupo 571 presenta el índice de presión proyectada más alto, con un valor superior a 0.62, seguido por el grupo 5, con un valor cercano a 0.61. También se observan otros grupos con índices superiores a 0.40, lo que indica una mayor relación entre

demanda esperada y capacidad residual. Estos grupos deben considerarse prioritarios para seguimiento, validación de información y planificación preventiva de camas, personal y recursos. En términos operativos, el gráfico permite pasar de una alerta general a una lista concreta de grupos hospitalarios que requieren atención anticipada en el corto plazo.

Sección de Código relacionada:

```
top_accionable = (
    det.sort_values("y_pred", ascending=False)
    .loc[:, ["provincia", "mes_ref", "grupo_hospital_id", "y_pred", "semaforo"]]
    .head(720)
)
```

3.6.3.7 Análisis complementario de causas de egreso asociadas a prioridad operativa

Como complemento al análisis territorial, se incorporó una lectura de causas predominantes de egreso con el fin de identificar qué perfiles clínicos aparecen con mayor frecuencia asociados a registros clasificados en prioridad alta o media. Este análisis permite enriquecer la interpretación operativa, ya que no solo muestra dónde se concentra la prioridad, sino también con qué tipo de demanda hospitalaria se relaciona. Es importante señalar que esta revisión no busca establecer causalidad clínica, sino reconocer patrones de asociación que pueden apoyar el monitoreo, la planificación preventiva y la revisión de capacidad en determinados servicios o territorios.

La Tabla 3.6.3-5 presenta las diez causas predominantes con mayor asociación a prioridad operativa a nivel nacional. El indicador de lift compara la frecuencia de alerta de cada causa frente a la tasa promedio del conjunto: valores superiores a 1 indican una asociación mayor al promedio, mientras que valores cercanos o inferiores a 1 reflejan una asociación similar o menor.

Tabla 3-36

Top 10 causas predominantes asociadas a prioridad alta o media a nivel nacional

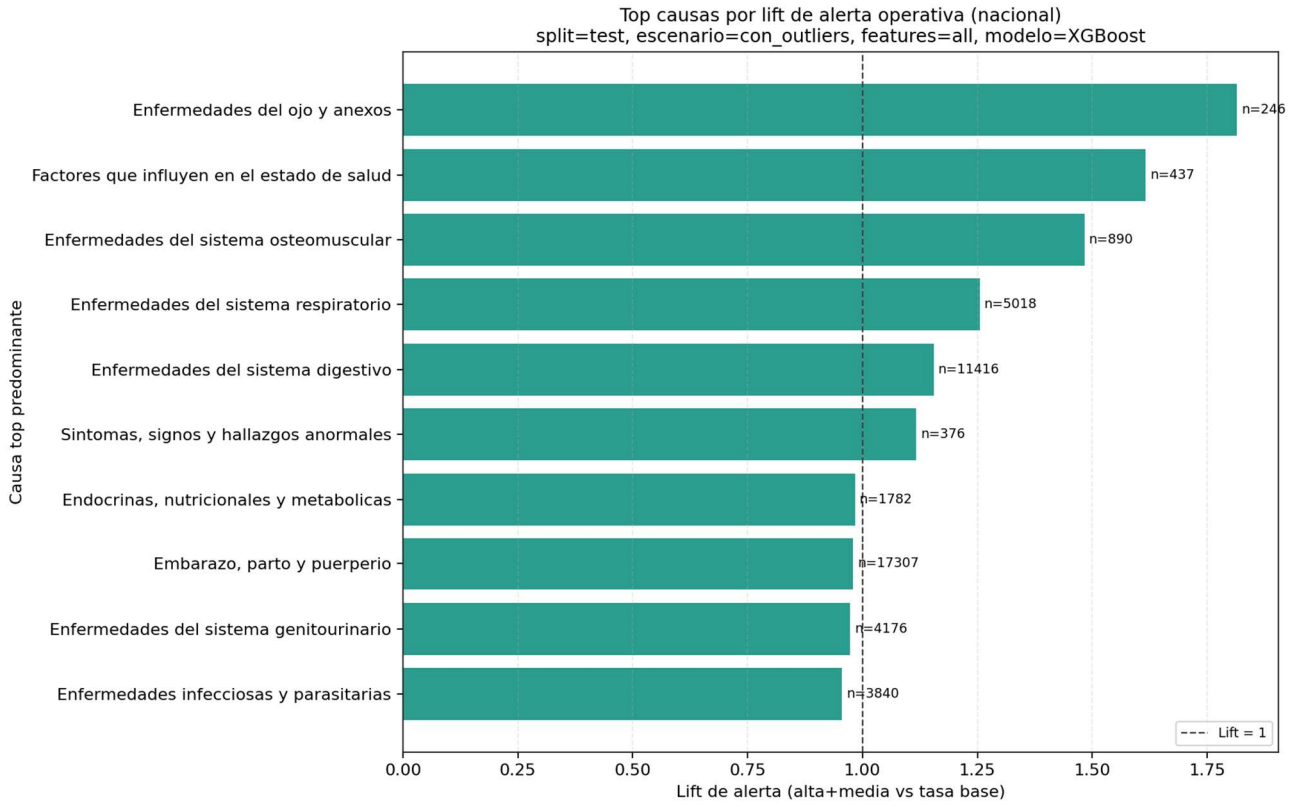
Rank	Causa predominante	Registros	% alerta	% alta	% media	Lift
1	Enfermedades del ojo y anexos	246	27.24%	15.45%	11.79%	1.82
2	Factores que influyen en el estado de salud	437	24.26%	10.07%	14.19%	1.62
3	Enfermedades del sistema osteomuscular	890	22.25%	7.30%	14.94%	1.48
4	Enfermedades del sistema respiratorio	5,018	18.83%	7.89%	10.94%	1.26
5	Enfermedades del sistema digestivo	11,416	17.33%	5.62%	11.70%	1.15
6	Síntomas, signos y hallazgos anormales	376	16.76%	5.85%	10.90%	1.12
7	Endocrinas, nutricionales y metabólicas	1,782	14.76%	6.34%	8.42%	0.98
8	Embarazo, parto y puerperio	17,307	14.70%	5.67%	9.03%	0.98
9	Enfermedades del sistema genitourinario	4,176	14.61%	3.95%	10.66%	0.97
10	Enfermedades infecciosas y parasitarias	3,840	14.32%	3.78%	10.55%	0.95

3.6.3.7.1 Causas clínicas con mayor lift de alerta

La Figura siguiente presenta las diez categorías diagnósticas CIE-10 con mayor lift de alerta a nivel nacional. El lift permite comparar la proporción de registros en alerta de cada causa frente al promedio general del conjunto. Un valor superior a 1 indica que esa causa se asocia con una frecuencia de alerta mayor al promedio, mientras que un valor cercano a 1 refleja un comportamiento similar al promedio. Esta visualización es útil porque permite identificar causas que, aunque no necesariamente sean las más frecuentes, aparecen más vinculadas a escenarios de prioridad operativa.

Figura 3-60

Gráfico nacional de lift por causa



Interpretación: “Enfermedades del ojo y anexos” presenta el lift más alto, con un valor aproximado de 1.82, lo que indica una asociación superior al promedio con registros clasificados en alerta. También destacan “Factores que influyen en el estado de salud” y “Enfermedades del sistema osteomuscular”, con valores mayores a 1. En contraste, causas de

alto volumen como “Embarazo, parto y puerperio” muestran un lift cercano a 1, lo que significa que su comportamiento de alerta es similar al promedio nacional. Esto confirma que las causas más frecuentes no necesariamente son las que presentan mayor asociación relativa con prioridad operativa.

Sección de Código relacionada:

```
base_alerta = det["is_alerta"].mean()


causas = det.groupby("causa_top", as_index=False).agg(
    registros=("is_alerta", "count"),
    tasa_alerta=("is_alerta", "mean")
)
causas["lift"] = causas["tasa_alerta"] / base_alerta
causas_top10 = causas.sort_values("lift", ascending=False).head(10)
```

3.6.3.8 Causas y focos territoriales de presión operativa

La lectura nacional de causas se complementó con un análisis territorial para identificar combinaciones específicas entre provincia y causa predominante que presentan mayor asociación con prioridad operativa alta o media. Este enfoque permite reconocer que una misma causa de egreso puede comportarse de manera distinta según el territorio, debido a diferencias en volumen de atención, capacidad disponible, perfil de servicios o concentración de determinados tipos de demanda. La Tabla 3.6.3-6 presenta ejemplos de combinaciones provincia-causa con mayores niveles de alerta y lift durante el período analizado.

Tabla 3-37

Ejemplos territoriales de asociación entre causa predominante y prioridad operativa

Provincia	Causa predominante	Registros	% alerta	% alta	% media	Lift
SUCUMBÍOS	Enfermedades del sistema genitourinario	36	77.78%	11.11%	66.67%	5.18
ORELLANA	Embarazo, parto y puerperio	184	64.67%	1.63%	63.04%	4.31
MANABÍ	Factores que influyen en el estado de salud	56	60.71%	41.07%	19.64%	4.05
ORELLANA	Traumatismos y envenenamientos	54	59.26%	3.70%	55.56%	3.95
GUAYAS	Enfermedades del ojo y anexos	44	54.55%	34.09%	20.45%	3.64
AZUAY	Enfermedades del ojo y anexos	35	54.29%	42.86%	11.43%	3.62
PICHINCHA	Endocrinas, nutricionales y metabólicas	338	44.38%	26.92%	17.46%	2.96
SUCUMBÍOS	Embarazo, parto y puerperio	314	39.49%	3.18%	36.31%	2.63
AZUAY	Enfermedades del sistema genitourinario	339	38.64%	13.27%	25.37%	2.58
PICHINCHA	Enfermedades infecciosas y parasitarias	466	38.84%	12.88%	25.97%	2.59

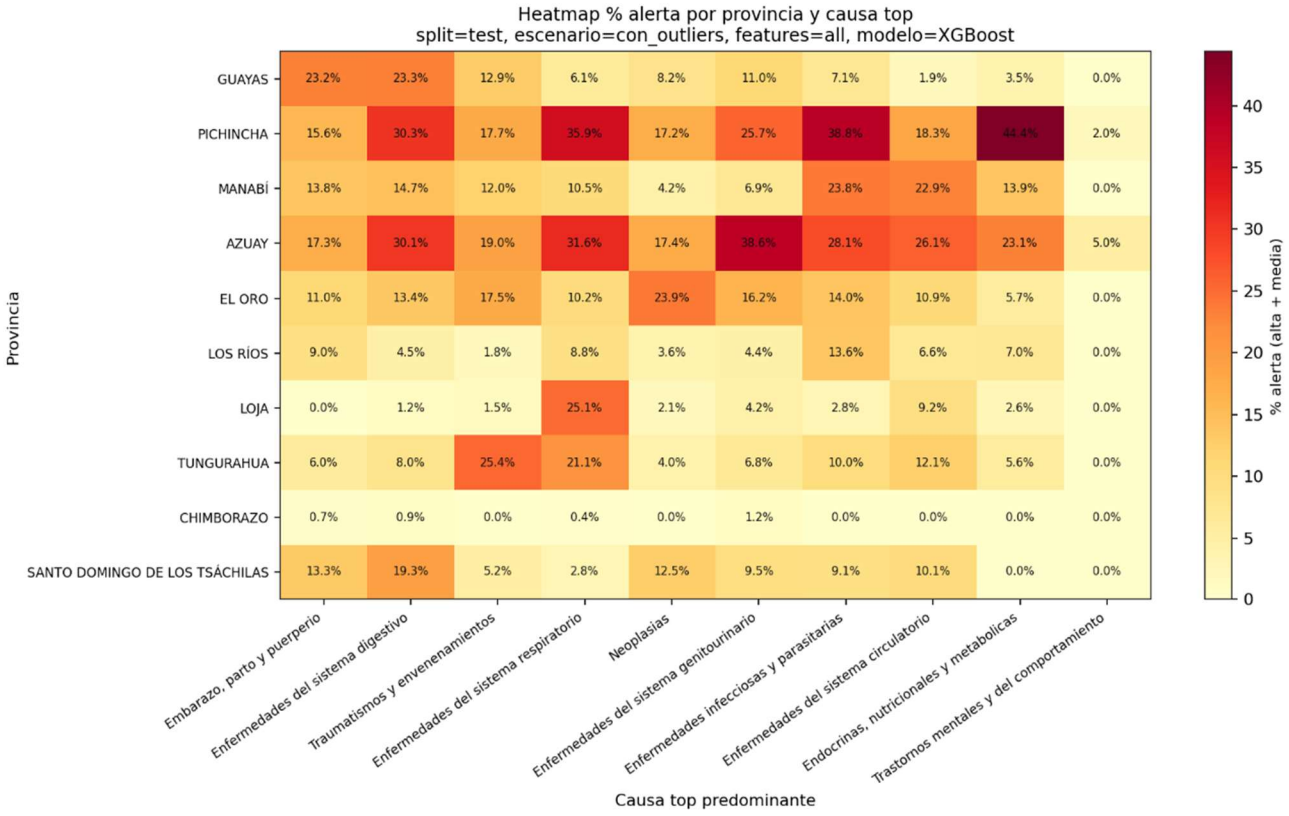
Interpretación: La tabla evidencia que la prioridad operativa no depende únicamente de la causa clínica, sino también del territorio donde se presenta. Por ejemplo, Sucumbíos con enfermedades del sistema genitourinario alcanza un 77.78 % de registros en alerta y un lift de 5.18, mientras que Orellana con embarazo, parto y puerperio registra 64.67 % de alerta y un lift de 4.31. Estos resultados muestran focos territoriales específicos donde la combinación entre provincia y tipo de demanda hospitalaria requiere mayor seguimiento. Sin embargo, los valores deben interpretarse como asociaciones operativas y no como relaciones causales clínicas.

3.6.3.8.1 Cruce provincia-causa en provincias de mayor volumen

La Figura siguiente presenta un mapa de calor que cruza las diez provincias con mayor volumen de egresos y las diez causas diagnósticas más frecuentes. Cada celda muestra el porcentaje de registros clasificados en prioridad alta o media para una combinación específica provincia-causa. Esta visualización permite identificar focos donde coinciden un territorio de alto volumen y una causa clínica con mayor presencia relativa de alertas operativas.

Figura 3-61

Heatmap provincia-causa de alerta



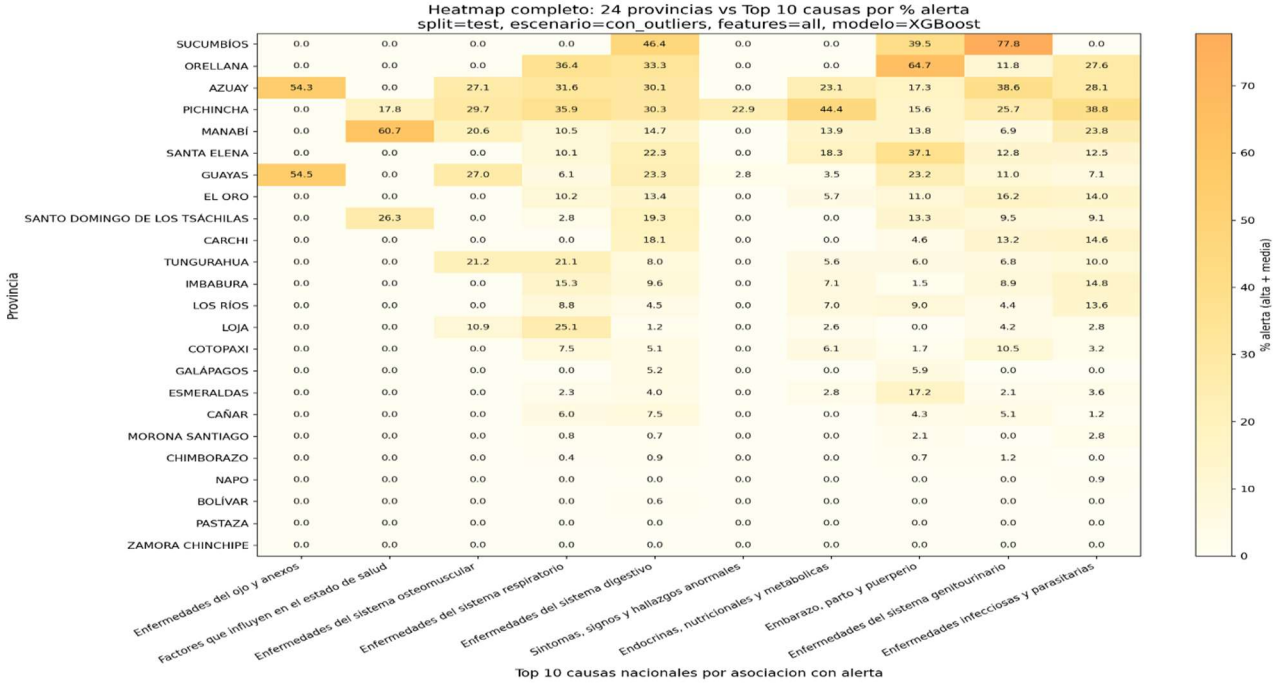
Interpretación: El mapa muestra que las combinaciones con mayor porcentaje de alerta se concentran principalmente en Pichincha. La celda más alta corresponde a Pichincha con enfermedades endocrinas, nutricionales y metabólicas, con 44.4 %, seguida de Pichincha con enfermedades infecciosas y parasitarias, con 38.8 %. En contraste, Chimborazo presenta valores cercanos a cero en las causas analizadas, lo que indica una baja presencia de registros clasificados en prioridad alta o media bajo el criterio relativo aplicado durante el período de prueba.

3.6.3.8.2 Mapa completo de alerta por provincia y causa

La Figura siguiente presenta un mapa de calor que cruza las 24 provincias del país con las diez causas diagnósticas de mayor asociación con alerta a nivel nacional. Cada celda muestra el porcentaje de registros clasificados en prioridad alta o media para una combinación específica provincia-causa. Esta visualización permite ampliar el análisis territorial y detectar focos de alerta que podrían no observarse únicamente en el ranking nacional de causas.

Figura 3-62

Heatmap completo 24 provincias vs top 10 causas por porcentaje de alerta



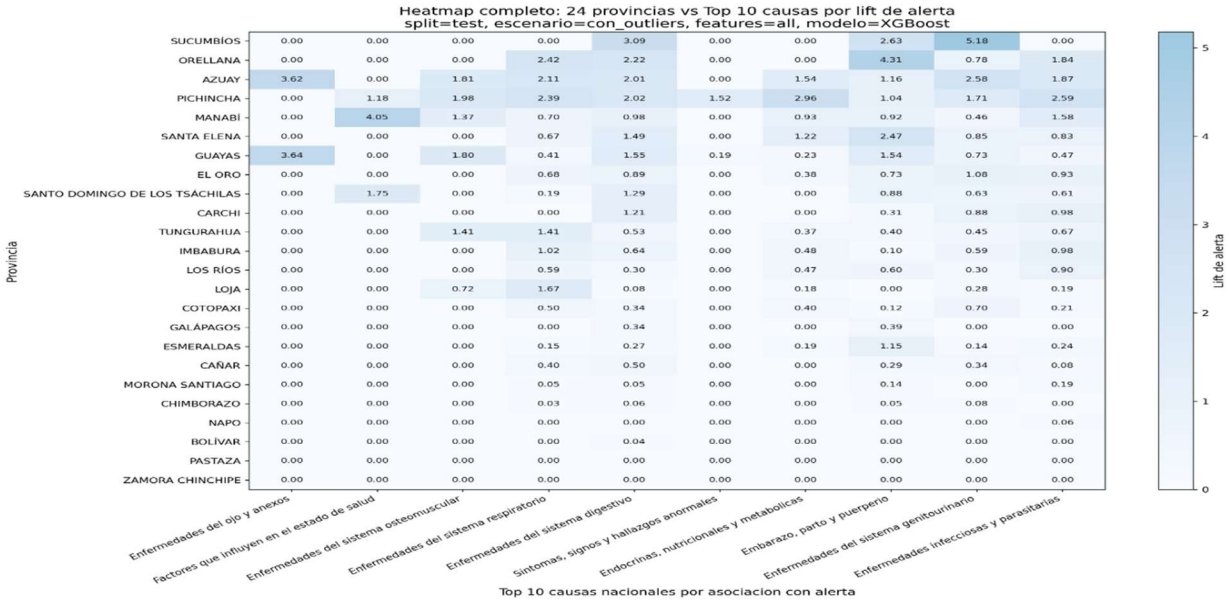
Interpretación: El mayor porcentaje de alerta se observa en la combinación Sucumbíos con enfermedades del sistema genitourinario, con 77.8 %, lo que significa que casi 8 de cada 10 registros de esa causa en la provincia fueron clasificados en prioridad alta o media. También destaca Orellana con embarazo, parto y puerperio, con 64.7 %, evidenciando una presencia relevante de alertas asociadas a esta causa en el período analizado. En contraste, provincias como Napo, Bolívar, Pastaza y Zamora Chinchipe presentan valores cercanos a cero en la mayoría de causas, lo que indica una baja presencia de registros priorizados bajo el criterio relativo aplicado.

3.6.3.8.3 Riesgo relativo por provincia y causa mediante lift

La Figura Siguiente presenta el cruce entre las 24 provincias y las diez causas diagnósticas analizadas, utilizando el lift de alerta en lugar del porcentaje absoluto. A diferencia del mapa anterior, este indicador compara cada combinación provincia-causa frente a la tasa promedio de alerta del conjunto. De esta manera, permite identificar combinaciones que presentan una asociación relativa mayor con prioridad operativa, incluso cuando su volumen de registros no es el más alto.

Figura 3-63

Mapa de calor completo 24 provincias vs top 10 causas por lift



Interpretación: El mapa de lift permite detectar focos que podrían no destacar únicamente por porcentaje absoluto o volumen de casos. Las combinaciones con valores superiores a 1 indican una asociación mayor al promedio con registros clasificados en prioridad alta o media, mientras que valores cercanos a 1 reflejan un comportamiento similar al promedio general. Esta lectura complementa el mapa de porcentajes, ya que ayuda a identificar provincias y causas cuya alerta relativa es proporcionalmente más alta. No obstante, los

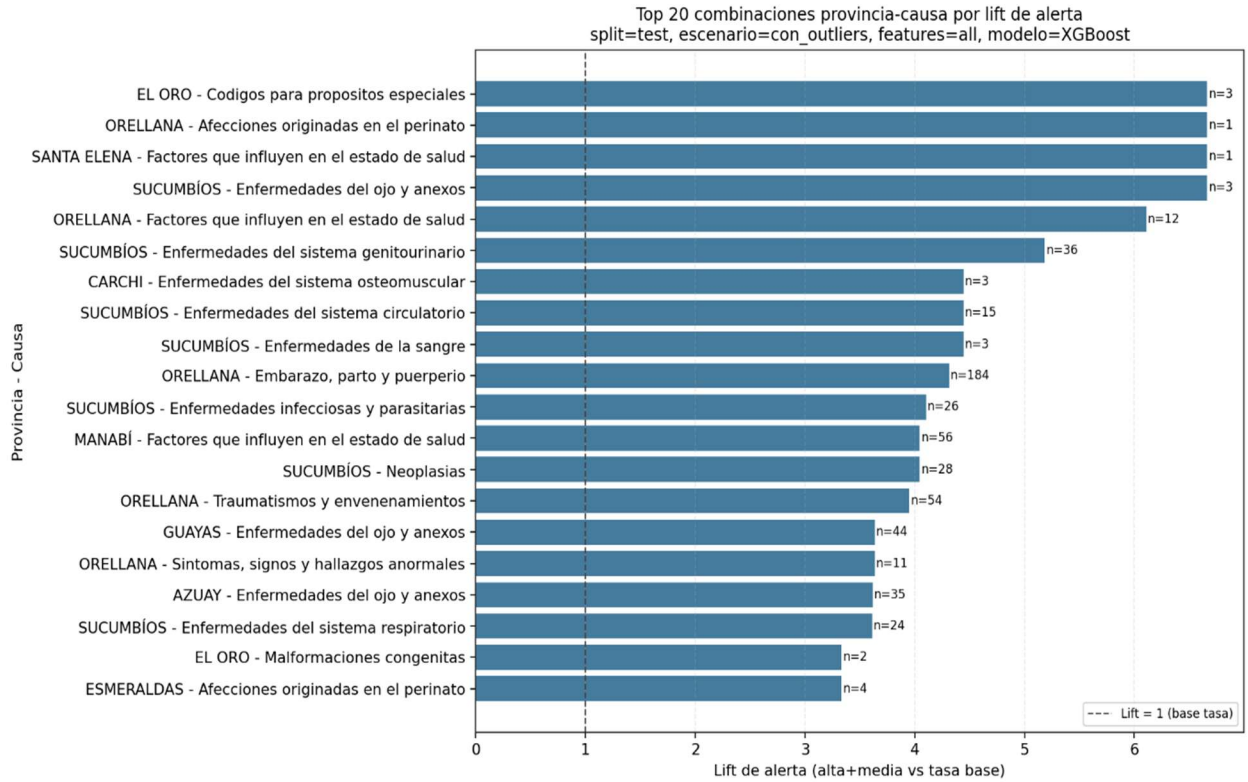
valores de lift deben interpretarse con cautela cuando el número de registros es bajo, debido a que muestras pequeñas pueden amplificar artificialmente la asociación observada.

3.6.3.8.4 Combinaciones provincia-causa con mayor prioridad operativa

La Figura siguiente presenta las 20 combinaciones provincia-causa con mayor lift de alerta, integrando la dimensión territorial y la composición clínica de los egresos. Cada barra muestra el valor de lift y se acompaña del tamaño de muestra (n), lo que permite evaluar si la asociación observada se apoya en un volumen suficiente de registros o si debe interpretarse con mayor cautela.

Figura 3-64

Ranking de las 10 causas principales por provincia según asociación con alerta operativa



Interpretación: Las combinaciones con tamaños de muestra muy pequeños pueden presentar valores de lift elevados por efecto de pocos registros, por lo que deben interpretarse con cautela. Entre las combinaciones con mayor soporte de datos destaca Sucumbíos con enfermedades del sistema genitourinario, con 36 registros y un lift aproximado de 5.18, lo

que indica una asociación con alerta más de cinco veces superior al promedio general. También resalta Orellana con embarazo, parto y puerperio, con 184 registros y un lift cercano a 4.3, combinando un volumen relevante con una alta asociación relativa. En conjunto, este análisis permite identificar focos provincia-causa que requieren revisión prioritaria, seguimiento operativo y validación complementaria antes de definir acciones de planificación hospitalaria.

Sección de Código relacionada:

```
focos = det.groupby(["provincia", "causa_top"], as_index=False).agg(
    registros=("is_alerta", "count"),
    tasa_alerta=("is_alerta", "mean")
)
focos["lift"] = focos["tasa_alerta"] / det["is_alerta"].mean()
```

3.6.3.9 Focos por grupo hospitalario en combinaciones provincia-causa

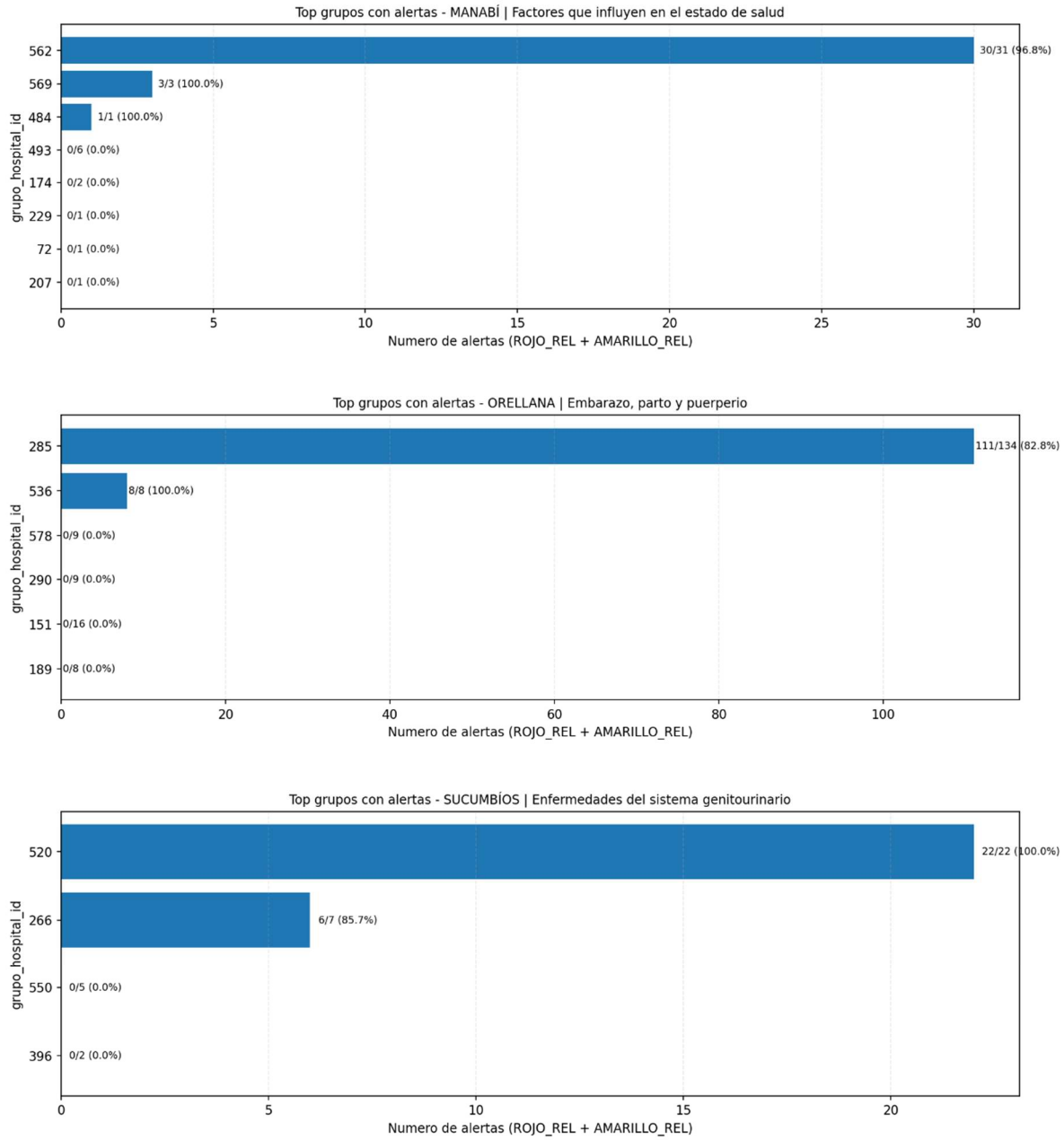
Después de identificar las combinaciones provincia-causa con mayor asociación a prioridad operativa, el análisis se desagregó al nivel de grupo hospitalario. Esta lectura permite precisar qué grupos concentran la mayor cantidad de registros priorizados dentro de cada foco territorial y clínico, fortaleciendo la utilidad operativa del modelo para seguimiento y planificación preventiva

3.6.3.9.1 Desagregación de focos por grupo hospitalario

La Figura siguiente presenta una serie de subgráficos, uno por cada foco activo identificado. En cada caso se comparan los registros totales y los registros clasificados en alerta por grupo hospitalario. Esta visualización permite identificar qué grupos tienen mayor participación dentro de cada combinación provincia-causa y, por tanto, cuáles requieren una revisión más específica en términos de capacidad, demanda y seguimiento operativo.

Figura 3-65

Grupos hospitalarios con más alertas en focos priorizados



Interpretación: En el foco “MANABÍ | Factores que influyen en el estado de salud”, el grupo 562 concentra 31 registros, de los cuales 30 fueron clasificados en alerta, alcanzando una tasa de 96.8 %. Esto lo convierte en el principal grupo a revisar dentro de ese foco. En el caso de “ORELLANA | Embarazo, parto y puerperio”, el grupo 285 presenta 134 registros y 111 alertas, con una tasa de 82.8 %, lo que evidencia una concentración importante de prioridad operativa. Por otra parte, el grupo 536 registra 100 % de alerta, pero sobre una base de solo 8

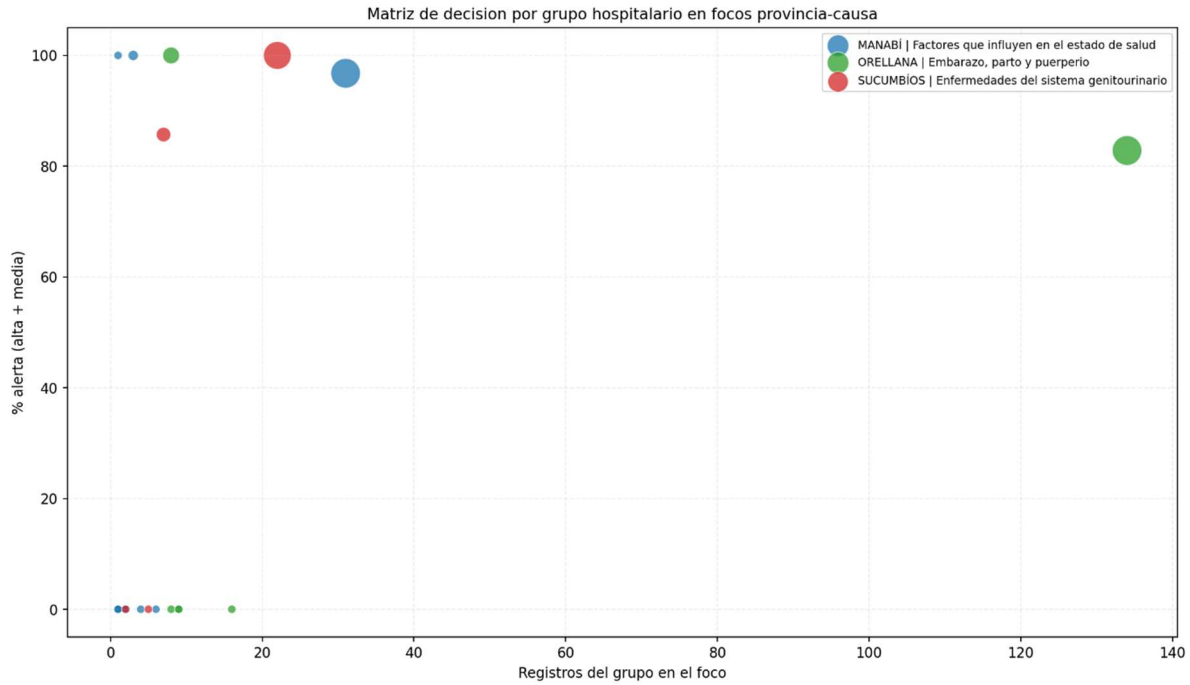
registros, por lo que este resultado debe interpretarse con cautela y validarse con información complementaria.

3.6.3.9.2 Priorización de grupos según volumen y tasa de alerta

La Figura siguiente presenta un gráfico de dispersión en el que cada punto representa un grupo hospitalario dentro de un foco provincia-causa. El eje X muestra el número de registros del grupo en el foco, mientras que el eje Y representa el porcentaje de registros clasificados en alerta. El tamaño del punto refleja el volumen de registros y el color identifica el foco al que pertenece. Esta visualización permite distinguir los grupos que combinan mayor volumen con mayor tasa de alerta, facilitando una priorización más precisa para el seguimiento operativo.

Figura 3-66

Matriz de decisión por grupo hospitalario en focos priorizados



Interpretación: Los grupos de mayor relevancia operativa son aquellos ubicados hacia la parte superior derecha del gráfico, ya que combinan alto volumen de registros con una alta proporción de alertas. En este sentido, el grupo asociado al foco “ORELLANA | Embarazo, parto y puerperio”, con aproximadamente 134 registros y 83 % de alerta, representa uno de los casos con mayor peso conjunto en el análisis. En cambio, los puntos con 100 % de alerta pero con muy pocos registros deben interpretarse con cautela, ya que pueden corresponder a casos puntuales o muestras pequeñas que requieren validación adicional antes de definir acciones de intervención.

3.6.3.10 Hallazgos para planificación hospitalaria

A partir de la interpretación operativa del modelo, se identifican los siguientes hallazgos para la planificación hospitalaria:

• El modelo seleccionado permite apoyar decisiones operativas al identificar escenarios comparativos de mayor presión hospitalaria.

• El semáforo relativo mensual transformó la predicción en una clasificación útil de prioridad alta, media y base.

• La prioridad operativa presenta una distribución territorial heterogénea, con mayor concentración en determinadas provincias y meses.

• La lectura provincia-mes permitió identificar territorios con mayor recurrencia de registros priorizados, facilitando el monitoreo focalizado.

• El análisis de causas de egreso aportó una lectura complementaria sobre los perfiles clínicos más asociados a prioridad operativa.

• Las causas más frecuentes no necesariamente fueron las más asociadas a alerta, por lo que el uso del lift permitió enriquecer la interpretación.

• La combinación de provincia, mes, causa predominante y grupo hospitalario permitió pasar de una predicción general a focos específicos de seguimiento.

• La desagregación por grupo hospitalario fortaleció el valor práctico del modelo para planificación preventiva de camas, personal y recursos.

Síntesis:

La predicción hospitalaria adquiere mayor utilidad cuando se traduce en una lógica operativa concreta. El enfoque aplicado permitió convertir los resultados del modelo en criterios de priorización por mes, territorio, causa predominante y grupo hospitalario, aportando una base práctica para el monitoreo focalizado, la alerta temprana y la planificación preventiva de recursos hospitalarios.

3.6.4 Conclusiones de los modelos de Machine Learning

1. **El benchmark mostró resultados competitivos entre escenarios, pero la mejor configuración puntual fue XGBoost con outliers y todas las variables.**

   El promedio de desempeño entre escenarios fue cercano, con WMAPE de 0.458198 en *sin_outliers* y 0.459889 en *con_outliers*. Sin embargo, la mejor configuración puntual correspondió a XGBoost bajo el escenario *con_outliers* y *feature set all*, con WMAPE = 0.369118, RMSE = 0.088753 y MAE = 0.060626. El mejor RandomForest quedó muy próximo, con WMAPE = 0.370068, lo que evidencia alta competitividad entre ambos modelos.

2. **El conjunto completo de variables aportó mayor valor predictivo para la decisión operativa.**

   Las mejores configuraciones se concentraron en el conjunto *all*, lo que respalda la utilidad de integrar señales de capacidad, demanda histórica, demografía, causas de egreso y estadía para representar mejor la dinámica hospitalaria.

3. **El manejo de outliers debe definirse según el objetivo de uso del modelo.**

   Aunque el escenario *sin_outliers* presentó una leve ventaja promedio, el mejor resultado puntual se obtuvo en *con_outliers*. Esto indica que la decisión de conservar

o excluir valores extremos no debe basarse únicamente en la métrica agregada, sino en el propósito operativo: estabilidad general o sensibilidad frente a escenarios de mayor presión.

4. **El modelo final mostró aplicabilidad sobre volumen real de datos.**

   La configuración seleccionada se aplicó sobre 59,989 registros de prueba y mantuvo un desempeño robusto, con WMAPE = 0.369118 y MAE = 0.060626. Esto respalda su uso como herramienta institucional para monitoreo, anticipación y priorización operativa.

5. **El semáforo absoluto no fue útil para discriminar prioridades, mientras que el semáforo relativo mensual generó una señal accionable.**

   En el corte final, el semáforo absoluto clasificó el 100 % de los 461 grupos evaluados en verde, por lo que no permitió diferenciar niveles de intervención. En cambio, el semáforo relativo mensual separó 392 grupos en prioridad base, 45 en prioridad media y 24 en prioridad alta, manteniendo una proporción prioritaria cercana al 15 % mensual.

6. **La lectura territorial permitió focalizar mejor la planificación hospitalaria.**

   El análisis por provincia evidenció una concentración relevante de prioridad en Orellana y Sucumbíos, con porcentajes promedio de 46.41 % y 44.71 %, respectivamente. Esta desagregación permite asignar esfuerzos de monitoreo y preparación preventiva con mayor precisión territorial.

7. **La salida provincia-mes-grupo convirtió la predicción en una herramienta de decisión concreta.**

   Se generaron 720 registros dentro del corte operativo, distribuidos en 358 ROJO_REL, 174 AMARILLO_REL y 188 VERDE_REL. Esta salida permite

diferenciar casos de intervención prioritaria, monitoreo reforzado y seguimiento rutinario.

8. **El análisis por causas y combinaciones provincia-causa aportó explicabilidad operativa.**

   A nivel nacional, algunas causas presentaron mayor asociación relativa con alerta, como enfermedades del ojo y anexos, factores que influyen en el estado de salud y enfermedades del sistema osteomuscular. A nivel territorial, combinaciones como Sucumbíos con enfermedades del sistema genitourinario y Orellana con embarazo, parto y puerperio mostraron asociaciones elevadas. Esto confirma que la frecuencia de egresos y la asociación con prioridad operativa no son equivalentes.

9. **La comparación con Deep Learning refuerza que la selección final debe ser criterio-dependiente.**

   Aunque MLP_light alcanzó un mejor WMAPE puntual de 0.362742, la selección final no debe depender únicamente de la familia algorítmica, sino de la combinación entre desempeño, trazabilidad, estabilidad, interpretabilidad y utilidad operativa.

3.7 Herramientas utilizadas

El desarrollo, procesamiento de datos y experimentación de este trabajo se realizó con herramientas consolidadas en ciencia de datos y aprendizaje automático, priorizando reproducibilidad, trazabilidad y eficiencia en el manejo de información hospitalaria.

3.7.1 Lenguaje de programación

Se utilizó Python 3.11 como lenguaje principal por su ecosistema maduro para análisis de datos y modelado predictivo.

Además, se trabajó con un entorno virtual (venv) para aislar dependencias y garantizar la reproducibilidad de los resultados en distintas ejecuciones.

3.7.2 Bibliotecas para procesamiento y análisis de datos

Para el tratamiento y preparación de la información se emplearon principalmente:

- pandas: manipulación, limpieza, transformación y análisis de datos tabulares.
- numpy: operaciones numéricas vectorizadas y soporte para estructuras multidimensionales.

3.7.3 Bibliotecas para Machine Learning

La etapa de modelado y evaluación se implementó con:

- scikit-learn: entrenamiento y evaluación de modelos, incluyendo RandomForestRegressor y métricas de desempeño.
- xgboost: implementación de XGBRegressor para boosting de gradiente.
- joblib: serialización y persistencia de modelos entrenados (cuando aplica en el flujo experimental).

3.7.4 Bibliotecas para Deep Learning

En la fase presentada en este capítulo no se implementaron modelos de Deep Learning en producción analítica principal.

No obstante, la estructura del proyecto permite incorporar en trabajos futuros bibliotecas especializadas para redes neuronales, según los requerimientos metodológicos y de capacidad computacional.

3.7.5 Almacenamiento y formatos de datos

Para el almacenamiento e intercambio de datos se utilizaron:

- Parquet: formato columnar eficiente para lectura/escritura en flujos analíticos y modelado.
- CSV: formato interoperable para revisión, exportación y validación manual.
- pathlib: gestión robusta y portable de rutas y archivos dentro del proyecto.

3.7.6 Entorno de ejecución y organización del proyecto

El desarrollo del proyecto se realizó en un entorno de ejecución local basado en Python, utilizando una estructura organizada de carpetas para separar datos originales, datos procesados, scripts, resultados, reportes y visualizaciones. Esta organización permitió mantener trazabilidad entre las etapas de preprocesamiento, construcción de la tabla Gold, generación del dataset de modelado, entrenamiento de modelos, evaluación y despliegue analítico.

La ejecución se apoyó en scripts modulares, cada uno asociado a una fase específica del pipeline. Esta separación facilitó la reproducción del proceso, la revisión de resultados intermedios y la actualización controlada de los artefactos generados.

En síntesis, se utilizó una organización modular de scripts y datos para separar las etapas de preprocesamiento, modelado, evaluación y generación de resultados.

3.8. Experimento complementario de Deep Learning (MLP)

3.8.1. Propósito del experimento complementario

Con el fin de fortalecer la validez comparativa del estudio, se incorporó un experimento complementario de Deep Learning mediante una red neuronal densa ligera, denominada MLP_light. Este experimento no sustituye ni redefine la metodología principal basada en modelos de Machine Learning, sino que actúa como contraste técnico para evaluar si una alternativa de mayor complejidad aporta mejoras relevantes frente a Random Forest y XGBoost.

3.8.2. Principio de no alteración del pipeline principal

El experimento se diseñó bajo el criterio de no intervención del pipeline oficial. En consecuencia, no se realizaron cambios en los orquestadores ni en el flujo principal de preprocesamiento, modelado y visualización. La ejecución del componente adicional se

implementó como un script independiente, benchmark_mlp_adicional.py, manteniendo intacta la trazabilidad del proceso original.

3.8.3. Estrategia de particionamiento temporal

Los datos modelados se dividieron en tres conjuntos temporales: entrenamiento, validación y prueba. Esta partición respetó el orden cronológico de las observaciones, evitando que información futura contaminara el entrenamiento y manteniendo coherencia con el horizonte predictivo de siete días.

Tabla 3-38

Partición temporal utilizada en el experimento MLP

Escenario	n_train	n_valid	n_test	Total	% Train	% Valid	% Test
con_outliers	250,507	63,073	59,989	373,569	67.06%	16.88%	16.06%
sin_outliers	71,771	17,810	16,872	106,453	67.42%	16.73%	15.85%

Esta estrategia permite entrenar el modelo con información histórica, ajustar el entrenamiento mediante validación y evaluar el desempeño final sobre datos posteriores no utilizados durante el ajuste.

3.8.4. Especificación técnica del MLP ligero

El modelo de Deep Learning se configuró como una red neuronal densa de baja complejidad, con dos capas ocultas de 64 y 32 neuronas, regularización mediante dropout y una capa de salida lineal para regresión.

Las variables de entrada fueron estandarizadas ajustando el escalador únicamente con el conjunto de entrenamiento y aplicando la misma transformación a validación y prueba. Esta decisión evita fuga de información y mantiene consistencia con el tratamiento aplicado en los modelos de Machine Learning.

El modelo se denomina MLP_light porque su propósito fue funcionar como benchmark complementario bajo condiciones controladas, priorizando comparabilidad, trazabilidad y costo computacional moderado, sin buscar una arquitectura profunda de alta complejidad.

3.8.5. Parámetros de entrenamiento y control de sobreajuste

El entrenamiento utilizó el optimizador Adam y una función de pérdida basada en el error absoluto medio. La configuración se mantuvo acotada, con 30 épocas máximas, batch size de 256 y early stopping con paciencia de 5 épocas. Además, se utilizó dropout de 0,15 como mecanismo adicional de regularización.

Estas decisiones permitieron controlar el riesgo de sobreajuste y mantener el experimento dentro de un costo computacional razonable. El objetivo no fue optimizar exhaustivamente una arquitectura neuronal, sino evaluar si un MLP ligero podía competir con los modelos principales de Machine Learning.

3.8.6. Métricas de evaluación

El desempeño del MLP_light se evaluó con las mismas métricas utilizadas en el benchmark principal: MAE, RMSE y WMAPE. Esta decisión permite comparar sus resultados de forma consistente con Random Forest, XGBoost y los modelos base.

MAE permite medir el error absoluto promedio, RMSE penaliza con mayor fuerza los errores grandes y WMAPE facilita la comparación proporcional del error respecto al volumen real observado. Para mantener coherencia metodológica, la evaluación se realizó sobre los mismos escenarios de datos y particiones temporales definidos previamente.

3.8.7. Resultados del experimento MLP_light

Los resultados del experimento complementario muestran que el MLP_light alcanzó un desempeño competitivo frente a los modelos de Machine Learning. En particular, obtuvo el

mejor WMAPE puntual del estudio bajo una configuración específica, lo que evidencia que una red neuronal densa ligera puede capturar patrones relevantes en el dataset construido. Sin embargo, este resultado debe interpretarse con cautela. El MLP_light fue incorporado como experimento complementario y no como reemplazo del benchmark principal. Su evaluación permite ampliar la comparación metodológica, pero la tesis mantiene como eje central los modelos Random Forest y XGBoost, debido a su mayor interpretabilidad, estabilidad y alineación con datos tabulares.

3.8.8. Interpretación del experimento complementario

El experimento con MLP_light aporta evidencia adicional sobre la viabilidad de utilizar técnicas de Deep Learning en el problema de predicción de presión hospitalaria. Su desempeño competitivo sugiere que existen patrones no lineales que pueden ser aprovechados por arquitecturas neuronales simples.

No obstante, el uso de Deep Learning requiere mayor cuidado en la configuración, escalamiento, control de sobreajuste e interpretación de resultados. Por ello, el MLP_light se considera una referencia complementaria que fortalece la comparación, pero no modifica las conclusiones centrales del estudio.

En síntesis, el experimento confirma que una red neuronal densa ligera puede ofrecer resultados competitivos bajo ciertas configuraciones, aunque los modelos de Machine Learning siguen constituyendo la línea principal del estudio por su equilibrio entre desempeño, interpretabilidad y trazabilidad metodológica.

CAPÍTULO 4 ANÁLISIS DE RESULTADOS

4.1. Introducción del capítulo

En este capítulo se presentan y analizan los principales resultados obtenidos a partir del enfoque predictivo desarrollado para estimar la presión hospitalaria con siete días de anticipación. Mientras el capítulo anterior describe las fases de preparación, modelado, evaluación y despliegue, incluyendo en la sección 3.6.3 una interpretación detallada de los resultados operativos, este capítulo sintetiza los hallazgos finales más relevantes y discute su utilidad para la planificación hospitalaria.

El análisis se organiza en cinco fases. Primero, se resume la consistencia de los datos utilizados en la evaluación. Segundo, se presentan los hallazgos principales del benchmark de Machine Learning y del experimento complementario con MLP_light. Tercero, se analiza la lectura de errores y el modelo de referencia seleccionado. Cuarto, se presentan los resultados operativos derivados de las predicciones, incluyendo priorización, provincias con mayor presión, causas asociadas y lift de alerta operativa. Finalmente, se discuten los resultados desde una perspectiva de apoyo a la toma de decisiones.

Las métricas utilizadas fueron MAE, RMSE y WMAPE. MAE mide el error absoluto medio; RMSE penaliza con mayor fuerza los errores grandes; y WMAPE expresa el error porcentual absoluto ponderado respecto al valor real observado. Para la interpretación principal se priorizó WMAPE, debido a que permite evaluar el error relativo ponderado en escenarios con diferencias de volumen entre grupos hospitalarios.

4.2. Consistencia de datos y particiones de evaluación

La evaluación se realizó sobre dos escenarios de modelado: con_outliers y sin_outliers. Como se observa en la siguiente tabla, el escenario con_outliers conserva la distribución completa de los datos, incluyendo observaciones extremas. El escenario sin_outliers reduce el número de registros debido al filtrado aplicado sobre valores atípicos, con el propósito de evaluar si la reducción de extremos mejora la estabilidad de los modelos.

Tabla 4-1

Tamaño de particiones por escenario

Escenario	n_train	n_valid	n_test
con_outliers	250.507	63.073	59.989
sin_outliers	71.771	17.810	16.872

Esta separación temporal es importante porque el objetivo del estudio no es explicar únicamente el comportamiento histórico, sino evaluar si los modelos son capaces de anticipar la presión hospitalaria en períodos posteriores a los utilizados para el entrenamiento.

4.3. Resultados predictivos de los modelos

El benchmark principal evaluó modelos de Machine Learning, específicamente RandomForest y XGBoost, bajo distintos conjuntos de variables. Adicionalmente, se ejecutó un experimento complementario de Deep Learning con una red neuronal densa ligera denominada MLP_light. La comparación permitió observar no solo qué modelo obtuvo menor error, sino también qué tipo de variables aportaron mayor señal predictiva.

4.3.1. Mejores resultados del benchmark de Machine Learning

La siguiente tabla presenta las mejores configuraciones obtenidas por los modelos de Machine Learning en el conjunto de prueba, ordenadas por WMAPE.

Tabla 4-2

Mejores resultados en prueba para Machine Learning

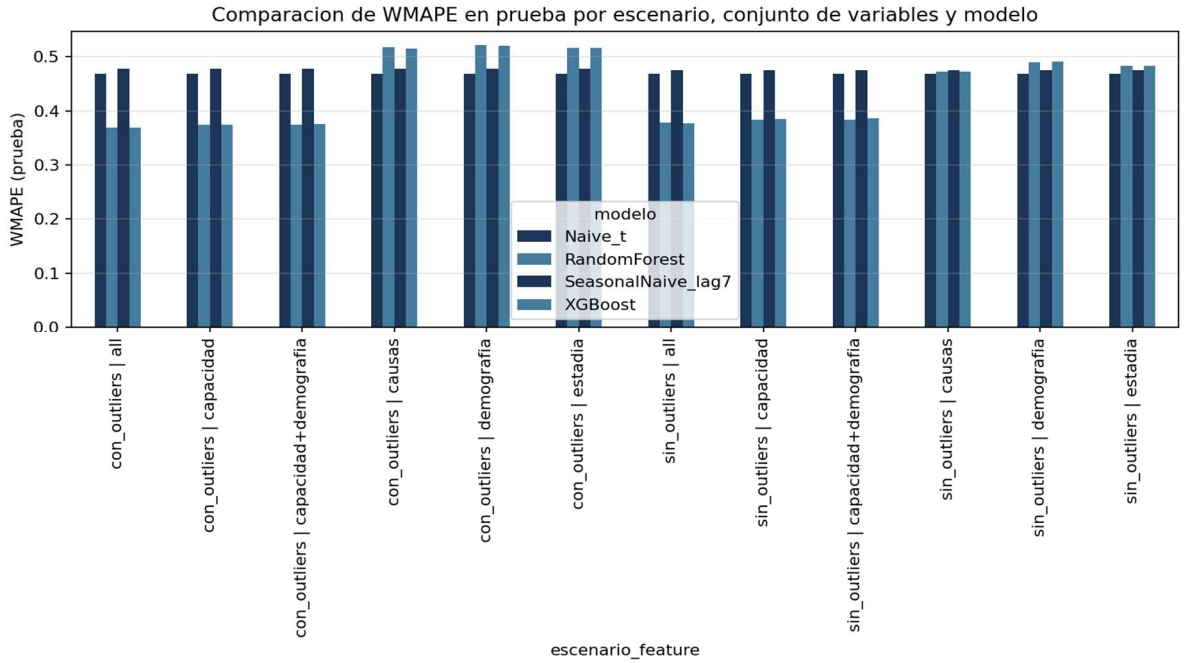
Escenario	Feature set	Modelo	MAE	RMSE	WMAPE
con_outliers	All	XGBoost	0.060626	0.088753	0.369118
con_outliers	All	RandomForest	0.060782	0.088605	0.370068
con_outliers	Capacidad	RandomForest	0.061715	0.089298	0.375746
con_outliers	capacidad+demografia	RandomForest	0.061744	0.089338	0.375925
con_outliers	capacidad+demografia	XGBoost	0.061847	0.089583	0.376555
con_outliers	Capacidad	XGBoost	0.061863	0.089556	0.376647
sin_outliers	All	XGBoost	0.055012	0.075095	0.382482
sin_outliers	All	RandomForest	0.055146	0.075149	0.383410
sin_outliers	capacidad+demografia	RandomForest	0.056134	0.076220	0.390281
sin_outliers	Capacidad	RandomForest	0.056183	0.076267	0.390621

El mejor resultado puntual del benchmark de Machine Learning corresponde a XGBoost en el escenario con_outliers y con el conjunto completo de variables, con un WMAPE de 0.369118. RandomForest obtuvo un resultado muy cercano bajo la misma configuración, con un WMAPE de 0.370068. Esta diferencia mínima evidencia que ambos modelos capturan de forma similar la estructura predictiva del problema.

Un hallazgo relevante es que las mejores configuraciones se concentran en el conjunto all y en los conjuntos relacionados con capacidad. Esto sugiere que la presión hospitalaria no depende de una sola dimensión, sino de la combinación entre demanda histórica, capacidad disponible, variables temporales, causas de egreso, estadía y características agregadas de los grupos hospitalarios.

Figura 4-1

Comparación de WMAPE en el conjunto de prueba



La comparación global confirma dos patrones: los modelos de aprendizaje superan a las referencias naive en los escenarios con valores atípicos, y ninguna configuración domina en todos los cortes. El mejor resultado corresponde a XGBoost con el conjunto completo de variables en el escenario con outliers, base de la selección del modelo de referencia.

4.3.2. Comparación global entre modelos

La siguiente tabla resume el desempeño promedio de los modelos evaluados en prueba. Esta tabla integra los resultados del benchmark de Machine Learning y del experimento complementario con MLP_light.

Tabla 4-3

Promedio en prueba por modelo

Modelo	MAE promedio	RMSE promedio	WMAPE promedio
MLP_light	0.064437	0.094714	0.424328
RandomForest	0.068259	0.094049	0.442896
XGBoost	0.068264	0.094090	0.442956

MLP_light obtuvo el mejor WMAPE promedio entre los modelos comparados. Sin embargo, RandomForest y XGBoost presentaron valores ligeramente mejores de RMSE promedio. Esto muestra que la interpretación del desempeño depende de la métrica considerada: un modelo puede ser mejor en error relativo ponderado, pero no necesariamente en la penalización de errores grandes.

El mejor resultado puntual de todo el estudio también corresponde a MLP_light, con un WMAPE de 0.362742 en el escenario con_outliers y con el conjunto all. Este resultado muestra que una red neuronal densa ligera puede capturar patrones útiles en datos hospitalarios tabulares cuando dispone de una representación amplia de variables. No obstante, su ventaja disminuye cuando se utilizan subconjuntos más reducidos como causas, demografía o estadía.

Por esta razón, el MLP_light se interpreta como un experimento complementario con desempeño competitivo, pero no como sustituto automático de los modelos de Machine Learning. RandomForest y XGBoost mantienen ventajas prácticas especialmente en facilidad de mantenimiento y adopción.

[Aquí va una nueva figura que debe mostrar: comparación de WMAPE entre MLP_light, RandomForest y XGBoost, considerando el promedio en prueba de cada modelo.]

4.3.3. Lectura de errores por escenario

Al comparar los escenarios con_outliers y sin_outliers, se observa que el escenario sin_outliers obtiene mejores métricas promedio, lo que sugiere mayor estabilidad cuando se reducen valores extremos. Sin embargo, el mejor resultado puntual aparece en el escenario con_outliers, específicamente con XGBoost y el conjunto all.

Esta diferencia es importante porque muestra que el tratamiento de outliers no debe evaluarse de forma aislada. Si el objetivo es obtener estabilidad promedio, el escenario sin_outliers resulta favorable. En cambio, si el objetivo es conservar información asociada a situaciones de mayor variabilidad hospitalaria, el escenario con_outliers puede aportar señales relevantes para la planificación operativa.

En términos de interpretación, los errores observados no invalidan la utilidad del modelo. El WMAPE obtenido en las mejores configuraciones indica que el modelo no debe entenderse como una predicción exacta de cada valor futuro, sino como una herramienta para anticipar presión relativa, priorizar casos y orientar el monitoreo preventivo.

4.4. Modelo de referencia para la interpretación operativa

Aunque MLP_light obtuvo el mejor WMAPE puntual, el análisis operativo se construyó tomando como referencia el modelo XGBoost en el escenario con_outliers y con el conjunto completo de variables. Esta decisión responde a tres razones principales: fue la mejor configuración del benchmark principal de Machine Learning, mantiene trazabilidad dentro del pipeline principal y ofrece una interpretación más directa para un contexto institucional.

Tabla 4-4

Modelo de referencia para la interpretación operativa

Elemento	Valor
Modelo	XGBoost
Escenario	con_outliers
Feature set	all
MAE	0.060626
RMSE	0.088753
WMAPE	0.369118
Uso principal	Interpretación operativa y priorización hospitalaria

La selección de este modelo como referencia no implica descartar el aporte del MLP_light. Más bien, permite diferenciar entre el mejor desempeño experimental y el modelo utilizado para construir una lectura operativa trazable. En una aplicación institucional, este criterio es relevante porque la utilidad del modelo no depende únicamente de una métrica, sino también de su facilidad de explicación, reproducción y mantenimiento.

4.5. Priorización operativa a partir de las predicciones

El principal aporte aplicado del estudio consiste en transformar las predicciones en niveles de prioridad para apoyar la planificación hospitalaria. La variable predicha representa la relación esperada entre demanda hospitalaria y capacidad disponible a siete días. Por tanto, valores más altos reflejan mayor presión asistencial esperada.

Para facilitar la interpretación, las predicciones fueron clasificadas mediante un semáforo relativo mensual. Esta clasificación no representa un diagnóstico clínico ni una medición

absoluta de congestión, sino una priorización comparativa entre grupos hospitalarios evaluados dentro del mismo período.

Tabla 4-5

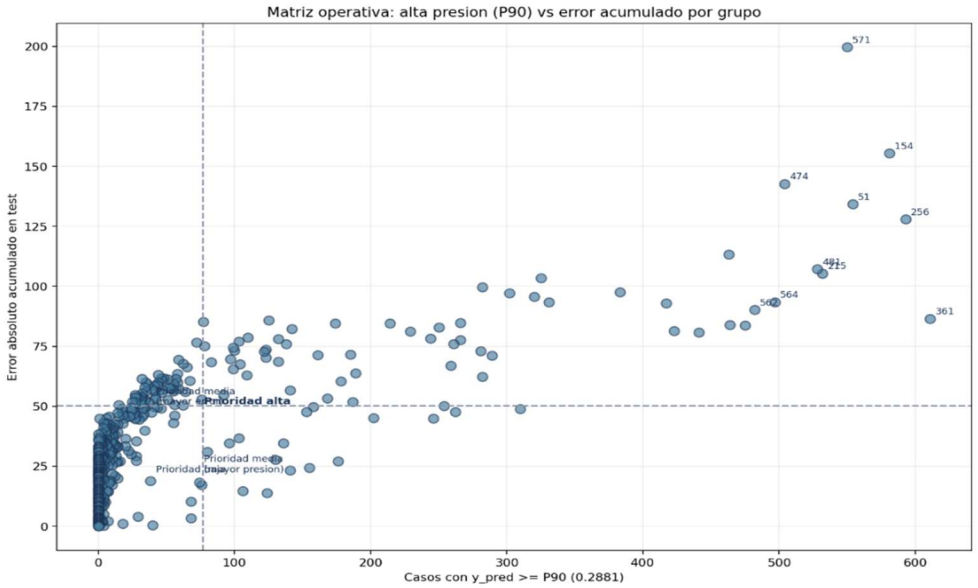
Niveles de prioridad operativa

Nivel	Interpretación
Rojo relativo	Prioridad alta
Amarillo relativo	Prioridad media
Verde relativo	Prioridad baja o base

El uso de un semáforo relativo permite focalizar la atención en los grupos hospitalarios con mayor presión esperada, sin asumir que todos los casos clasificados como prioridad alta representan necesariamente saturación real. Esta lectura es útil para monitoreo preventivo, revisión de capacidad, análisis territorial y planificación anticipada de recursos.

Figura 4-2

Matriz de priorización operativa a partir del modelo de referencia



Aplicada sobre el modelo de referencia, la matriz operativa traduce el desempeño en una regla de gestión: los grupos con mayor presión proyectada y mayor error acumulado se asignan a seguimiento diario, y los restantes a monitoreo semanal o quincenal. De este modo el error del modelo no solo se mide, sino que modula la intensidad del control.

4.6. Resultados territoriales de prioridad

El análisis por provincia permitió identificar territorios con mayor recurrencia de prioridad alta o media durante el período de prueba. Esta lectura es relevante porque permite pasar de una evaluación global del modelo a una interpretación territorial útil para la planificación hospitalaria.

Tabla 4-6

Provincias con mayor prioridad promedio

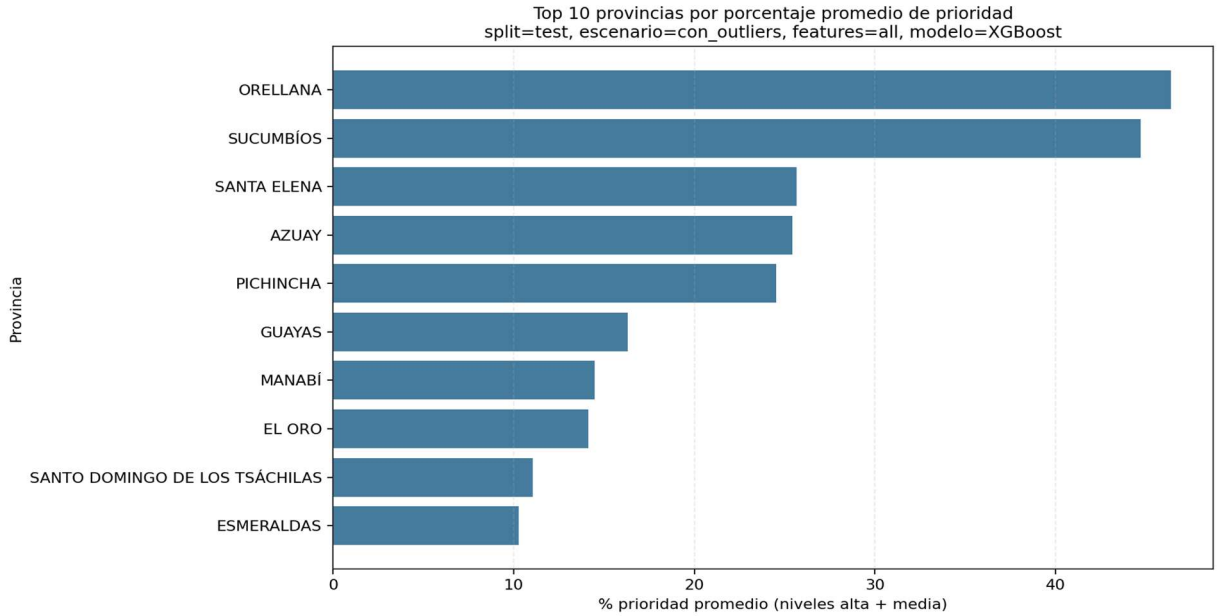
Ranking	Provincia	Prioridad promedio
1	Orellana	46.41 %
2	Sucumbíos	44.71 %
3	Santa Elena	25.68 %
4	Azuay	25.43 %
5	Pichincha	24.55 %
6	Guayas	16.32 %
7	Manabí	14.50 %
8	El Oro	14.13 %
9	Santo Domingo de los Tsáchilas	11.04 %
10	Esmeraldas	10.26 %

Los resultados muestran que Orellana y Sucumbíos presentan los mayores porcentajes de prioridad promedio. Ambas provincias mantienen una recurrencia elevada de registros clasificados con prioridad alta o media, lo que sugiere una presión relativa más persistente durante el período evaluado.

Santa Elena, Azuay y Pichincha también muestran niveles relevantes de prioridad, aunque menores que los observados en Orellana y Sucumbíos. Esta diferencia evidencia que la presión hospitalaria esperada no se distribuye de manera homogénea en el territorio, por lo que la planificación debería considerar diferencias provinciales y no únicamente indicadores agregados nacionales.

Figura 4-3

Provincias con mayor prioridad promedio



La lectura territorial permite identificar provincias que podrían requerir monitoreo más frecuente, revisión de disponibilidad de camas y análisis complementario de recursos. Sin embargo, estos resultados deben interpretarse como señales de priorización y no como evidencia definitiva de saturación, debido a que el indicador modelado representa presión relativa y no ocupación hospitalaria real.

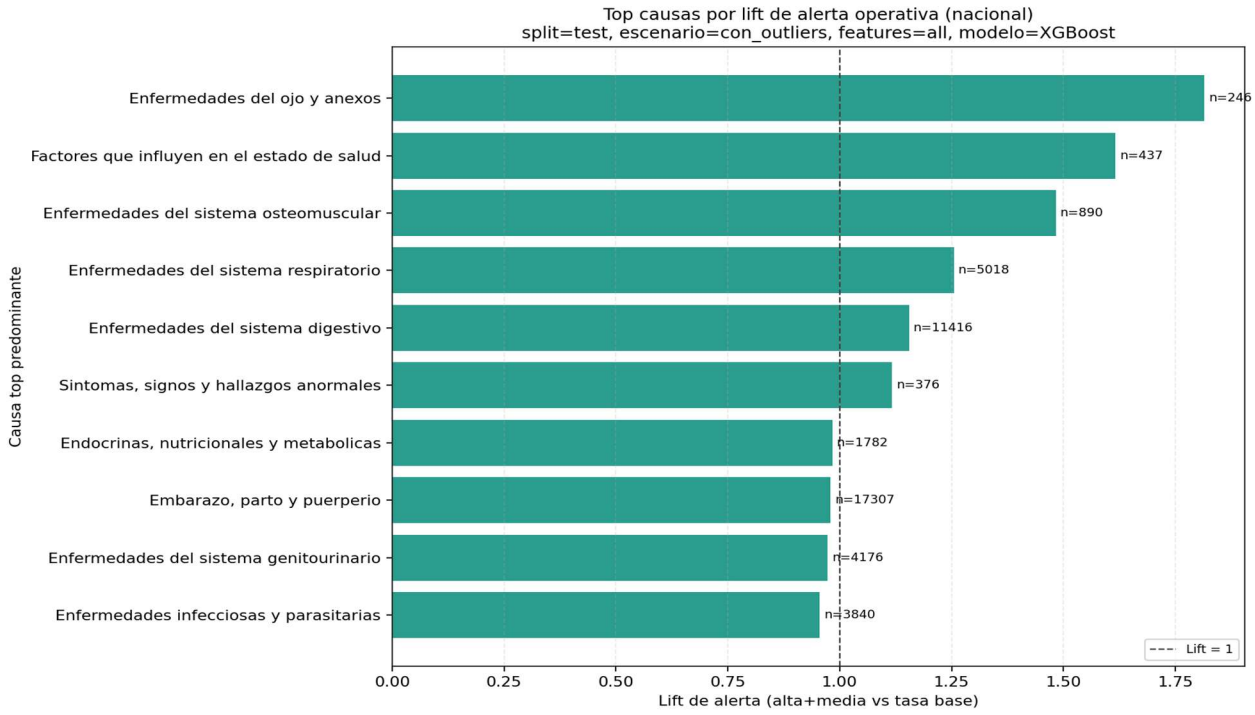
4.7. Causas predominantes y prioridad operativa

Además del análisis territorial, el estudio incorporó una lectura por causas predominantes de egreso hospitalario. Esta dimensión permite explorar si ciertas categorías clínicas aparecen con mayor frecuencia en registros clasificados con prioridad alta o media.

Un hallazgo importante es que las causas más frecuentes no siempre son las más asociadas a mayor prioridad operativa. Por ello, el análisis no se limitó al conteo de causas, sino que incorporó el lift de alerta operativa. Este indicador compara la tasa de alerta de una causa frente a la tasa promedio del conjunto analizado. Un lift mayor que 1 indica una asociación superior al promedio; un lift cercano a 1 refleja un comportamiento similar al promedio; y un lift menor que 1 indica una asociación inferior al promedio.

Figura 4-4

Lift nacional de alerta operativa por causa predominante



El lift aporta una lectura complementaria a la frecuencia. Mientras la frecuencia indica volumen, el lift indica asociación relativa con alerta operativa. Esto permite identificar causas

que, aunque no sean las más numerosas, pueden estar más vinculadas con escenarios de presión hospitalaria esperada.

Desde la perspectiva de planificación, esta distinción es relevante. Una causa frecuente puede requerir capacidad estructural permanente, mientras que una causa con lift elevado puede indicar un foco específico de monitoreo o preparación preventiva.

4.8. Focos provincia-causa de alerta operativa

El análisis nacional de causas se complementó con una lectura territorial por combinación provincia-causa. Esta aproximación permitió identificar focos donde una causa predominante presenta una asociación superior con prioridad alta o media dentro de una provincia específica.

Tabla 4-7

Ejemplos de focos provincia-causa con mayor asociación operativa

Provincia	Causa predominante	Registros	% alerta	% alta	% media	Lift
Sucumbíos	Enfermedades del sistema genitourinario	36	77.78 %	11.11 %	66.67 %	5.18
Orellana	Embarazo, parto y puerperio	184	64.67 %	1.63 %	63.04 %	4.31
Manabí	Factores que influyen en el estado de salud	56	60.71 %	41.07 %	19.64 %	4.05
Orellana	Traumatismos y envenenamientos	54	59.26 %	3.70 %	55.56 %	3.95
Guayas	Enfermedades del ojo y anexos	44	54.55 %	34.09 %	20.45 %	3.64
Azuay	Enfermedades del ojo y anexos	35	54.29 %	42.86 %	11.43 %	3.62
Pichincha	Endocrinas, nutricionales y metabólicas	338	44.38 %	26.92 %	17.46 %	2.96
Sucumbíos	Embarazo, parto y puerperio	314	39.49 %	3.18 %	36.31 %	2.63

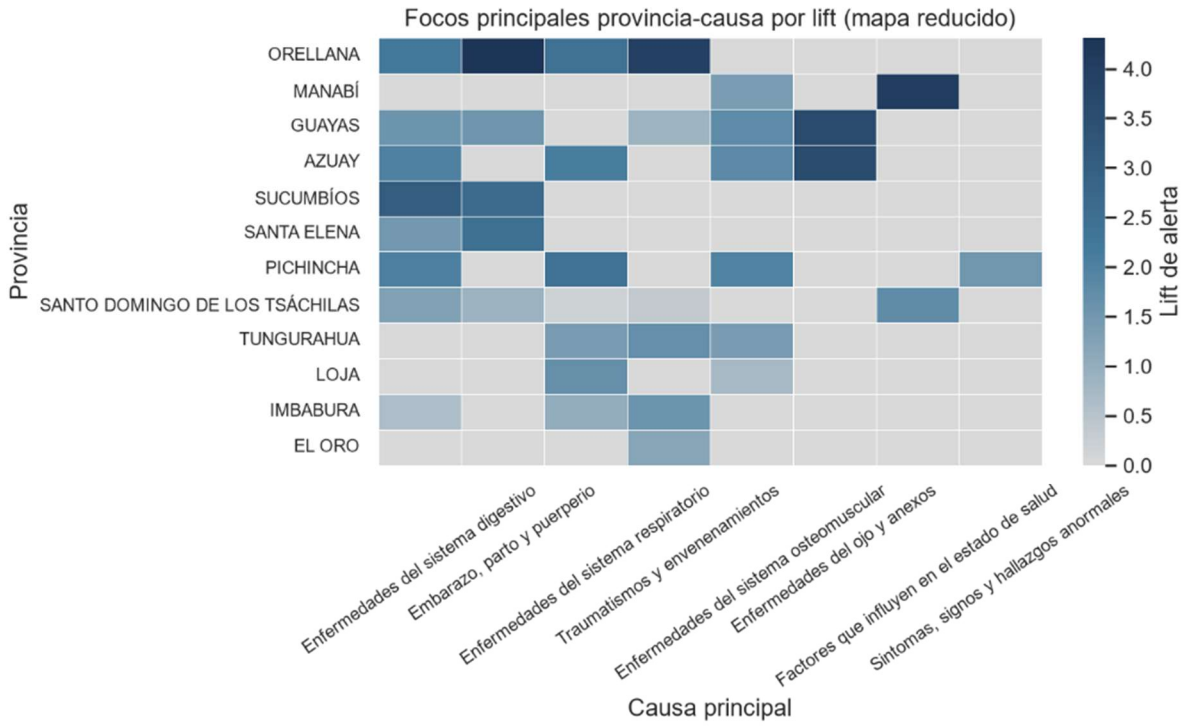
Entre los focos identificados destaca Sucumbíos con enfermedades del sistema genitourinario, con un lift de 5.18. Esto indica que dicha combinación provincia-causa presenta una asociación con alerta operativa más de cinco veces superior al promedio general. También resalta Orellana con embarazo, parto y puerperio, con 184 registros y un lift de 4.31, lo que combina volumen relevante con una asociación elevada a prioridad.

Estos resultados no deben interpretarse como causalidad clínica. Su propósito es identificar señales operativas que podrían orientar revisiones adicionales, análisis territorial, validación con equipos de gestión hospitalaria y preparación preventiva.

Por ejemplo, en la siguiente figura se evidencia que los focos de mayor asociación operativa no se distribuyen de manera uniforme entre provincias ni entre causas predominantes. Se observan valores altos de lift en Orellana, especialmente en embarazo, parto y puerperio, con un valor cercano a 4.3, y en enfermedades del sistema respiratorio y traumatismos, con valores alrededor de 3.8 a 4.0. También destaca Manabí en factores que influyen en el estado de salud, con un lift aproximado de 4.0, y Sucumbíos en enfermedades del sistema digestivo y embarazo, parto y puerperio, con valores superiores a 2.5. Estos resultados sugieren que ciertos territorios presentan combinaciones específicas de causa y presión esperada, por lo que el análisis provincia-causa aporta una lectura más precisa que un ranking nacional agregado.

Figura 4-5

Focos principales provincia por causa



El mapa de calor identifica los focos provincia-causa con mayor lift de alerta: Orellana destaca en embarazo-parto y afecciones respiratorias, Manabí en enfermedades del ojo y Guayas y Azuay en el sistema osteomuscular. Esta lectura orienta la preparación preventiva hacia combinaciones específicas de territorio y perfil clínico, más allá del volumen bruto de egresos.

4.9. Interpretación integrada de los hallazgos

Los resultados permiten identificar varios hallazgos relevantes para la planificación hospitalaria.

En primer lugar, los modelos predictivos evaluados lograron generar estimaciones útiles de presión hospitalaria esperada a siete días. XGBoost y RandomForest mostraron desempeños muy cercanos, mientras que MLP_light alcanzó el mejor WMAPE puntual y promedio. Esto evidencia que tanto los modelos de árboles como una red neuronal densa ligera pueden capturar patrones relevantes en los datos hospitalarios integrados.

En segundo lugar, el conjunto completo de variables fue el más informativo. Las mejores configuraciones se obtuvieron cuando se utilizaron variables de capacidad, demanda histórica, temporalidad, causas de egreso, demografía y estadía. Esto confirma que la presión hospitalaria es un fenómeno multidimensional y que su predicción mejora cuando se integran distintas fuentes de información.

En tercer lugar, la comparación de escenarios muestra una tensión entre estabilidad y cobertura. El escenario sin_outliers mejora las métricas promedio, mientras que el escenario con_outliers conserva mayor variabilidad y produce el mejor resultado puntual en el benchmark de Machine Learning. Esta diferencia es importante porque los valores extremos pueden representar situaciones operativas relevantes en el contexto hospitalario.

En cuarto lugar, el aporte principal del estudio no se limita a reducir el error de predicción. La transformación de las predicciones en niveles de prioridad, rankings territoriales, análisis de causas y lift de alerta operativa permite convertir el modelo en una herramienta de apoyo a la gestión. Esta lectura operativa es fundamental porque conecta el desempeño estadístico con decisiones concretas de monitoreo y planificación.

En quinto lugar, los resultados territoriales muestran que la presión esperada no se distribuye de forma homogénea. Provincias como Orellana y Sucumbíos presentan mayor recurrencia de prioridad alta o media, mientras que ciertas combinaciones provincia-causa muestran asociaciones especialmente elevadas con alerta operativa. Esto sugiere que la planificación hospitalaria puede beneficiarse de enfoques diferenciados por territorio y perfil de demanda.

4.10. Alcance de la interpretación operativa

Los resultados deben interpretarse con cautela. El modelo no predice ocupación real de camas ni reemplaza la evaluación clínica o administrativa de los establecimientos de salud. La variable modelada representa una relación esperada entre demanda y capacidad disponible, construida a partir de egresos hospitalarios y camas disponibles.

Por tanto, las alertas generadas deben entenderse como señales de priorización relativa. Su valor radica en orientar el monitoreo, anticipar posibles presiones y focalizar la revisión de recursos. Antes de implementar decisiones operativas, los resultados deberían contrastarse con información institucional adicional, como disponibilidad real de personal, ocupación diaria, derivaciones, cartera de servicios y eventos coyunturales no capturados por los datos históricos.

Desde esta perspectiva, el modelo funciona como una herramienta de apoyo analítico. Su utilidad está en transformar datos históricos en evidencia cuantitativa para fortalecer la planificación preventiva, no en sustituir el juicio experto de los responsables de la gestión hospitalaria.

4.11. Cierre del capítulo

El análisis de resultados evidencia que el enfoque predictivo desarrollado permite estimar la presión hospitalaria esperada y transformarla en criterios operativos de priorización. Los modelos de Machine Learning, especialmente XGBoost y RandomForest, mostraron resultados robustos y trazables. El experimento complementario con MLP_light evidenció mejoras en WMAPE bajo configuraciones favorables, especialmente con el conjunto completo de variables.

Más allá de la comparación de modelos, el principal aporte del estudio está en la interpretación operativa de las predicciones. La clasificación por niveles de prioridad, el análisis por provincia, la revisión de causas predominantes y el uso del lift de alerta operativa permiten identificar focos que podrían requerir mayor monitoreo y preparación preventiva.

En conclusión, el modelo no debe entenderse como una herramienta de decisión automática, sino como un insumo analítico para anticipar presión asistencial, orientar la revisión de capacidad y fortalecer la planificación hospitalaria basada en datos.

CAPÍTULO 5 CONCLUSIONES Y RECOMENDACIONES

5.1 Síntesis final de la investigación

La investigación cumplió su propósito de desarrollar y validar un enfoque predictivo para apoyar la planificación hospitalaria a partir de datos de egresos y disponibilidad de camas del período 2022-2024. El trabajo integró fuentes heterogéneas, construyó un flujo reproducible de preparación y modelado, y transformó los resultados en una salida operativa orientada a la priorización.

La contribución principal de la tesis no se limita a comparar algoritmos, sino a conectar la cadena completa de valor analítica: integración de datos, control de calidad, construcción del dataset de modelado, prevención de fuga de información, evaluación comparativa y traducción de resultados a criterios de gestión hospitalaria. En este sentido, el estudio demuestra que el valor del modelo depende tanto de su desempeño estadístico como de su capacidad para convertirse en una herramienta útil para el monitoreo, la alerta temprana y la planificación preventiva.

Los resultados evidencian que no existe una configuración universalmente superior. El desempeño varía según el tratamiento de outliers, el conjunto de variables y la familia de modelos utilizada. Por ello, la selección final debe responder al propósito operativo del análisis, considerando no solo la métrica de error, sino también la trazabilidad, estabilidad, interpretabilidad y utilidad para la toma de decisiones.

5.2 Conclusiones por objetivo de investigación

En relación con el objetivo general, se concluye que el enfoque propuesto es viable para estimar la presión asistencial esperada mediante la relación demanda-capacidad y apoyar la planificación hospitalaria con criterio preventivo. La investigación permitió integrar datos,

construir un dataset de modelado, evaluar modelos predictivos y traducir los resultados en una salida operativa útil para la toma de decisiones.

Objetivo específico 1: Integrar bases de egresos y camas del período 2022-2024.

Este objetivo se cumplió mediante la consolidación de las fuentes de egresos hospitalarios y disponibilidad de camas en una base integrada, trazable y utilizable para analítica predictiva. Esta integración permitió construir una visión conjunta entre demanda observada y capacidad hospitalaria.

Objetivo específico 2: Evaluar la calidad, cobertura y granularidad de los datos para definir la unidad analítica.

Este objetivo se cumplió a través de la revisión, depuración y estandarización de los datos, así como de la identificación de inconsistencias, valores extremos y diferencias de cobertura entre fuentes. Como resultado, se definió una unidad analítica consistente para el modelado y se construyeron escenarios comparables con y sin tratamiento de outliers.

Objetivo específico 3: Construir un dataset de modelado sin fuga de información.

Este objetivo se cumplió mediante un diseño temporal estricto, separando los conjuntos de entrenamiento, validación y prueba. Además, las reglas de tratamiento de datos y generación de variables fueron aplicadas de forma controlada para evitar que información futura influyera en el entrenamiento de los modelos.

Objetivo específico 4: Entrenar modelos de Machine Learning y contrastarlos con un experimento complementario de Deep Learning.

Este objetivo se cumplió mediante el entrenamiento y evaluación de modelos como RandomForest, XGBoost y una red neuronal densa ligera o MLP_light. Todos los modelos fueron comparados bajo un protocolo homogéneo, utilizando los mismos escenarios, particiones y conjuntos de variables.

Objetivo específico 5: Comparar el desempeño de los modelos mediante MAE, RMSE y WMAPE.

Este objetivo se cumplió al evaluar los modelos con métricas complementarias de error absoluto, error cuadrático y error relativo ponderado. Los resultados permitieron diferenciar entre robustez promedio y mejor configuración puntual, mostrando que la selección del modelo depende del escenario, las variables y el propósito operativo.

Objetivo específico 6: Estimar la presión asistencial esperada para identificar prioridades territoriales, causas asociadas y grupos hospitalarios con mayor presión relativa.

Este objetivo se cumplió mediante la construcción de una salida operativa basada en semáforo relativo mensual, rankings por provincia y mes, análisis por causas de egreso, lift de alerta operativa y priorización por grupo hospitalario. Esta salida permitió traducir las predicciones en insumos útiles para monitoreo, alerta temprana y planificación preventiva.

En conjunto, los objetivos fueron alcanzados y el estudio aporta evidencia metodológica y aplicada para apoyar decisiones hospitalarias más anticipativas, focalizadas y sustentadas en datos.

5.3 Aportes de la tesis

Los aportes de la tesis se sintetizan en cuatro niveles principales.

En primer lugar, el aporte metodológico consiste en el diseño de un flujo integral y reproducible para datos hospitalarios, desde la integración de fuentes hasta la generación de una salida operativa. Este flujo permite ordenar el proceso analítico y reducir riesgos metodológicos asociados a inconsistencias, fuga de información o falta de trazabilidad.

En segundo lugar, el aporte técnico se refleja en la comparación sistemática de modelos, escenarios y conjuntos de variables. La evaluación permitió contrastar alternativas de Machine Learning y Deep Learning bajo condiciones homogéneas, mostrando que la

selección del modelo debe sustentarse en evidencia empírica y no únicamente en la complejidad algorítmica.

En tercer lugar, el aporte aplicado se encuentra en la traducción de las predicciones a una lógica de priorización operativa. El semáforo relativo mensual, la lectura territorial por provincia y mes, y la desagregación por grupo hospitalario permiten convertir los resultados del modelo en insumos concretos para monitoreo y planificación preventiva.

Finalmente, el aporte académico se relaciona con la generación de evidencia sobre el uso de modelos predictivos en datos tabulares hospitalarios. La investigación muestra que el desempeño depende del escenario, del tratamiento de valores atípicos, del conjunto de variables y del criterio de evaluación, por lo que no existe una solución única aplicable a todos los contextos.

5.4 Implicaciones para la planificación hospitalaria

Desde la perspectiva de gestión, el estudio muestra que la predicción puede funcionar como un mecanismo de alerta temprana para anticipar escenarios de presión asistencial. Su utilidad no está únicamente en estimar un valor futuro, sino en ordenar prioridades y orientar la atención hacia los grupos hospitalarios, territorios y períodos que requieren mayor seguimiento.

El análisis evidencia que el semáforo absoluto puede no ser suficiente para discriminar prioridades en determinados cortes. En cambio, el semáforo relativo mensual permite comparar los registros dentro de cada período y seleccionar un subconjunto manejable de casos prioritarios. Esto facilita una gestión más focalizada y evita saturar el proceso de seguimiento con alertas poco diferenciadas.

La desagregación territorial por provincia y mes aporta una lectura útil para asignar esfuerzos operativos de manera diferenciada. Asimismo, la incorporación de causas predominantes de

egreso permite complementar la lectura territorial con una visión clínica-operativa, sin asumir causalidad, pero identificando patrones asociados a mayor prioridad.

Adicionalmente, el uso del lift de alerta operativa permitió diferenciar entre causas frecuentes y causas con mayor asociación relativa a prioridad. Esta lectura es relevante porque una causa con alto volumen no necesariamente representa mayor presión operativa. En cambio, el análisis provincia-causa permitió identificar focos específicos de seguimiento, útiles para orientar revisiones territoriales, validar patrones con equipos de gestión y planificar acciones preventivas de manera más focalizada.

En conjunto, la combinación de grupo hospitalario, territorio, tiempo y causa predominante transforma el modelo en una herramienta de apoyo a la decisión. El modelo no reemplaza el criterio institucional ni la validación de los equipos de gestión, pero sí aporta evidencia comparable y oportuna para fortalecer la planificación de camas, personal y recursos.

5.5 Limitaciones del estudio

El estudio presenta limitaciones que deben ser consideradas al interpretar sus resultados.

La primera limitación se relaciona con la dependencia de registros administrativos. La calidad, consistencia y completitud de los datos de egresos y camas condicionan el desempeño del modelo y la confiabilidad de las salidas operativas.

La segunda limitación corresponde al tratamiento de outliers. Los resultados muestran que la exclusión o conservación de valores extremos puede modificar el desempeño según el escenario y el objetivo de uso. Por ello, no debe asumirse una regla única para todos los contextos.

La tercera limitación se vincula con la heterogeneidad entre unidades analíticas. Los grupos hospitalarios pueden presentar diferencias importantes en volumen, especialización, capacidad instalada y dinámica de demanda, lo que influye en la estabilidad de las predicciones.

La cuarta limitación está asociada al alcance temporal del estudio. El período 2022-2024 permite construir una base histórica relevante, pero los patrones de demanda hospitalaria pueden cambiar con el tiempo, por lo que el modelo requiere recalibración y monitoreo continuo para mantener su utilidad.

Finalmente, el componente de Deep Learning se mantuvo como experimento complementario. Aunque aportó evidencia relevante, su desarrollo no reemplaza la necesidad de futuras pruebas con arquitecturas adicionales, mayor ajuste de hiperparámetros y validaciones más amplias.

Otra limitación del estudio es que los resultados del modelo deben entenderse como señales de apoyo para la planificación, y no como una medición directa de ocupación real o de congestión hospitalaria. De igual manera, el análisis por causas permite identificar patrones asociados a mayor prioridad, pero no demuestra que esas causas sean el origen directo de la presión hospitalaria observada.

Estas limitaciones no invalidan los hallazgos obtenidos, pero delimitan su alcance y refuerzan la necesidad de actualización, validación y mejora continua si el enfoque se implementa en un entorno institucional.

5.6 Recomendaciones

A partir de los resultados obtenidos, se recomienda mantener una evaluación periódica de los modelos y escenarios de predicción. La evidencia muestra que el desempeño puede variar según el tratamiento de outliers, el conjunto de variables y el objetivo operativo, por lo que no conviene establecer una única configuración fija sin revisión continua.

También se recomienda conservar una base amplia de variables cuando el propósito sea mejorar la capacidad predictiva y capturar mejor la dinámica hospitalaria. Las variables de capacidad, demanda histórica, demografía, causas de egreso y estadía aportan señales complementarias que fortalecen el modelado.

Desde el punto de vista operativo, se recomienda utilizar el semáforo relativo mensual como mecanismo base de priorización. Este enfoque permite identificar casos prioritarios dentro de cada período y facilita un seguimiento comparable entre meses.

Asimismo, se recomienda institucionalizar una lectura territorial periódica por provincia y mes, de modo que los equipos de gestión puedan identificar territorios con mayor recurrencia de registros priorizados y asignar recursos de forma diferenciada.

Se recomienda también mantener trazabilidad completa del pipeline, incluyendo reglas de limpieza, tratamiento de outliers, generación de variables, particiones temporales, entrenamiento de modelos y producción de salidas operativas. Esto es fundamental para reproducibilidad, auditoría y mejora continua.

Finalmente, se recomienda establecer mecanismos de recalibración programada y monitoreo de deriva del modelo. Si cambian los patrones de demanda, la disponibilidad de camas o la calidad de los registros, el desempeño predictivo puede modificarse y requerir ajustes.

5.7 Líneas de trabajo futuro

Como continuidad de esta investigación, se propone ampliar el horizonte temporal de análisis para evaluar la estabilidad interanual de los modelos y verificar si los patrones identificados se mantienen en períodos posteriores.

También se recomienda profundizar en estrategias de tratamiento de valores atípicos y calibración por subgrupos territoriales, considerando que los outliers pueden representar tanto ruido estadístico como eventos relevantes para la gestión hospitalaria.

Otra línea de trabajo futuro consiste en incorporar variables exógenas, siempre que exista disponibilidad y calidad suficiente. Variables epidemiológicas, demográficas, climáticas, socioeconómicas o de calendario podrían enriquecer la capacidad del modelo para anticipar variaciones en la demanda.

Asimismo, se sugiere explorar arquitecturas adicionales de Deep Learning y compararlas bajo protocolos homogéneos, manteniendo como criterio central no solo el desempeño numérico, sino también la estabilidad, interpretabilidad y utilidad operativa.

Finalmente, se recomienda avanzar hacia una capa visual ejecutiva orientada a usuarios de gestión, que permita consultar semáforos, rankings territoriales, grupos priorizados y evolución temporal de manera clara y oportuna.

5.8 Conclusión final

La tesis demuestra que es posible construir un sistema analítico robusto para apoyar la planificación hospitalaria mediante la integración de datos, el modelado predictivo comparativo y la traducción operativa de resultados. El valor del enfoque no depende únicamente de seleccionar el algoritmo con mejor desempeño puntual, sino de articular adecuadamente calidad de datos, diseño experimental, criterios de evaluación y propósito de uso.

La conclusión central es que la predicción hospitalaria adquiere mayor utilidad cuando se convierte en una herramienta de priorización. Bajo este enfoque, los resultados permiten identificar escenarios de mayor presión relativa por mes, territorio, causa predominante y grupo hospitalario, aportando una base práctica para decisiones más anticipativas, focalizadas y sustentadas en evidencia.

REFERENCIAS BIBLIOGRÁFICAS.

Amirsmaili, M., Mehrolhassani, M. H., & Iranmanesh, M. (2026). The role of external factors in shaping the supply and demand of general practitioners in Iran: A qualitative study. *BMC Primary Care, 27*(1), 2.

Andersen, R. M. (1995). Revisiting the behavioral model and access to medical care: Does it matter? *Journal of Health and Social Behavior, 36*(1), 1–10.

Babitsch, B., Gohl, D., & von Lengerke, T. (2012). Re-revisiting Andersen's Behavioral Model of Health Services Use: A systematic review of studies from 1998–2011. *GMS Psycho-Social-Medicine, 9*, Doc11.

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research, 13*, 281–305.

Blanco, J., Ferreras, M., & Cosido, O. (2026). Predictive modeling of hospital emergency department demand using artificial intelligence: A systematic review. *International Journal of Medical Informatics, 207*, 106215.

Boyle, J., Zeitz, K., Hoffman, R., Khanna, S., & Beltrame, J. (2014). Probability of severe adverse events as a function of hospital occupancy. *IEEE Journal of Biomedical and Health Informatics, 18*(1), 15–20.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). Wiley.

Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5–32.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.

Databricks. (s. f.). *What is medallion architecture?* Recuperado el 18 de junio de 2026, de https://www.databricks.com/blog/what-is-medallion-architecture

Fallahnezhad, M., Langarizadeh, M., & Vahabzadeh, A. (2024). Key performance indicators of hospital supply chain: A systematic review. *BMC Health Services Research, 24*, 1610.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

Griffiths, J. D., Knight, V., & Komenda, I. (2013). Bed management in a critical care unit. *IMA Journal of Management Mathematics, 24*(2), 137–153.

Hadian, S. A., Rezayatmand, R., & Pourghaderi, A. R. (2024). Hospital performance evaluation indicators: A scoping review. *BMC Health Services Research*.

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting, 22*(4), 679–688.

Işli, D., & Aydın, L. (2026). Enhancing hospital drug and medical supply request processes through a clinically prioritized and demand-variability-aware adaptive decision support framework. International Journal of Medical Informatics.

Khalifa, M., Khalid, P., & Al-Juma, S. (2015). Developing strategic health care key performance indicators: A case study on a tertiary care hospital. *Procedia Computer Science, 63*, 459–466.

McRae, S. (2021). Long-term forecasting of regional demand for hospital services. Operations Research for Health Care.

Núñez, A., Neriz, L., & Mateo, R. (2018). Emergency departments key performance indicators: A unified framework and its practice. *International Journal of Health Planning and Management*.

Orhan, F., & Kurutkan, M. N. (2025). Predicting total healthcare demand using machine learning: Separate and combined analysis of predisposing, enabling, and need factors. *BMC Health Services Research*.

Pan, P., & Yan, R. (2026). Forecast analysis of Shanghai medical service demand based on ARIMA model. *Scientific Reports, 16*, 16747. https://doi.org/10.1038/s41598-026-47940-6

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825–2830.

Rodoreda-Pallàs, B., Morín-Fraile, V., & Lumillo-Gutierrez, I. (2026). Healthcare utilization and cost analysis by social determinants of health: A prospective cohort study in Catalonia, Spain. *BMC Health Services Research*.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature, 323*, 533–536.

Sasikala, K., Karpagalakshmi, S., Rajeswari, M., Abinaya, C., & Janani, E. (2026). Hybrid machine learning and queuing theory approach for real-time hospital resource allocation. En *IEEE International Conference on Electronic Systems and Intelligent Computing (ICESIC) 2026 Proceedings*.

Shahidian, N., Abreu, P., & Barbosa-Póvoa, A. (2025). Short-term forecasting of emergency medical services demand exploring machine learning. Computers & Industrial Engineering.

Wang, Y., Han, Y., Wang, S., Guan, Z., & Chang, C. (2026). Feasibility of short-term hospital mask demand forecasting using a backpropagation neural network under data scarcity. *Scientific Reports, 16*(1), 14904. https://doi.org/10.1038/s41598-026-44754-4

Wiler, J. L., Welch, S., Pines, J., Schuur, J., Jouriles, N., & Stone-Griffith, S. (2015). Emergency department performance measures updates: Proceedings of the 2014 Emergency

Department Benchmarking Alliance Consensus Summit. *Academic Emergency Medicine, 22*(5), 542–553.

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error over the root mean square error in assessing average model performance. *Climate Research, 30*, 79–82.

Zhao, Y. (2026). Dynamic demand forecasting and capacity assessment for emergency care system: A hybrid time series and machine learning approach with evidence from Beijing. *BMC Health Services Research, 26*(1), 568. https://doi.org/10.1186/s12913-026-14356-x

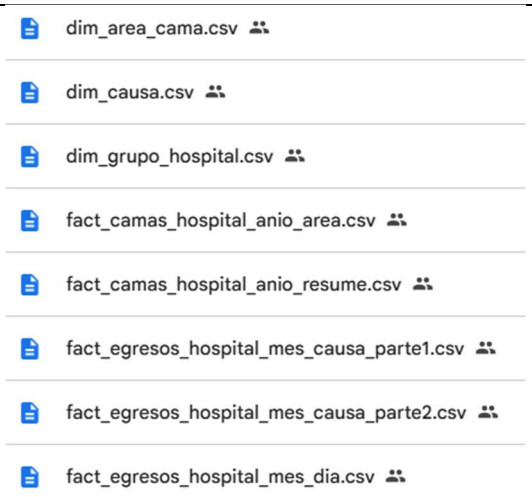
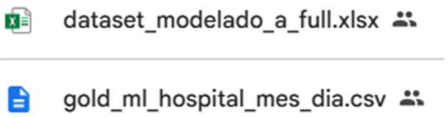
APÉNDICES:

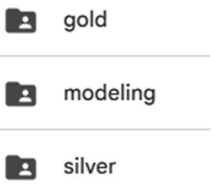
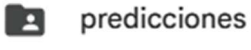
Apéndice 1. Enlace a código Generado:

https://drive.google.com/drive/folders/1UI5SOK7d5pF528tUh6aVtA19AjksAC6h?usp=drive_link

Dentro del proyecto se tiene la siguiente estructura:

data	En esta carpeta se encuentran todos los datos antes y despúes de procesarlos.
csv	
bronze (original)	
Datos_Abierto_ECH_2024 Datos_Abierto_EEH_2024 Datos_abiertos_ECH_2022 Datos_Abiertos_ECH_2023 Datos_abiertos_EEH_2022 Datos_Abiertos_EEH_2023	Datos originales de egresos y camas hospitalarias por separado. (años 2022, 2023, 2024).
silver	
	Tablas de dimensiones y tablas de hechos para visualización. Los scripts utilizan las tablas parquet con el mismo nombre y funcionan como temporales para armar agrupaciones y es la base para la tabla final Gold.

	
	
	- Archivo final para visualización con columnas adicionales para modelos - Tabla GOLD agrupada por día y mes de las transacciones de los años 2022, 2023 y 2024

	
	Archivos generados para almacenar toda la información del proceso en formato parquet en las capas silver, gold y diferentes escenarios con outliers, sin outliers, y separación en: TRAIN, TEST y VALIDACION
	

22_prediccion_detallada_modelos_valid_test.csv	Archivo que contiene todas las predicciones para los modelos evaluados naive, naive_t, XGBoost, y Random Forest

Orden de ejecución	src	
	utilidades	
	paleta_colores.py	Mantiene arreglo de colores para los gráficos.
	orquestacion	
1	regenerar_fase_pre_tesis.py	Orquestador de la fase previa al modelado denominada **preprocesamiento**. Ejecuta scripts de diagnóstico, relacionamiento, EDA previo y finaliza con la construcción de la tabla GOLD.
2	regenerar_fase_tesis.py	Orquestador de la fase denominada **modelado**. Ejecuta EDA GOLD, modelado, benchmark, predicciones y entregables finales.

	preprocesamiento	
1.1	analisis_cobertura_llaves_spark.py	Cálculos de cobertura de llaves entre fuentes con Spark. (camas y egresos).
1.2	relacionamiento_llaves.py	Resuelve/normaliza el relacionamiento de llaves entre tablas. (camas y egresos).
1.3	eda_previo_preprocesamiento.py	Genera EDA inicial de datos antes del preprocesamiento. Permite comprender inicialmente como se encuentran los datos.
1.4	prediccionHospitalariaPreprocesamiento.py	Ejecuta el preprocesamiento principal y construye datasets intermedios y GOLD.
1.5	diccionario_datos_gold.py	Genera el diccionario de datos del dataset GOLD.
	modelado	
2.1	eda_gold_hospital.py	Genera visualizaciones EDA del GOLD para análisis y documentación.

2.2	construir_dataset_modelado_a.py	Construye dataset de modelado con target, lags, rolling y particiones train/valid/test.
2.3	generar_diccionario_modelado_a.py	Genera el diccionario del dataset de modelado.
2.4	benchmark_modelos_a.py	Ejecuta benchmark de modelos y escenarios (incluye búsqueda aleatoria configurada por variables de entorno). Para modelos de machine learning.
2.5	benchmark_mlp_adicional.py	Ejecuta benchmark para modelo de Deep Learning.
2.6	generar_prediccion_detallada_modelos.py	Genera predicciones detalladas para validación y prueba.
2.7	graficar_predicciones_modelos.py	Genera gráficos resumen de desempeño y análisis de predicciones.
2.8	resumen_mejor_modelo_escenario.py	Resume mejor modelo por escenario/split con métricas.

2.9	generar_entregable_priorizacion_grupos.py	Construye el entregable operativo final (tablas y figuras de priorización/semaforización).

Apéndice 2. Resumen de Diccionario de Datos de tabla GOLD

Grupo de columnas	**Criterio de agrupación**	**Total de columnas**	**Tipo**	**Ejemplos**	**Descripción**
Identificación y tiempo	grupo_hospital_id, anio, mes, dia	4	int64	grupo_hospital_id, anio, mes	Identificación del grupo hospitalario y referencia temporal del registro.
Ubicación e institución	prov_ubi, cant_ubi, parr_ubi, area_ubi, clase, tipo, entidad, sector	8	object	prov_ubi, area_ubi, entidad	Contexto geográfico y administrativo del establecimiento.

Grupo de columnas	Criterio de agrupación	Total de columnas	Tipo	Ejemplos	Descripción
Egresos y demografía	egresos_totales, estadía, sexo, edad y proporciones poblacionales	14	int64 / float64	egresos_totales, promedio_dias_estadia, pct_mayores_65	Magnitud de egresos, composición demográfica y severidad operativa.
Capacidad total de camas	totales de dotación/disponibilidad y bandera de consistencia	5	int64 / float64	total_camas_dotacion, total_camas_disponibles_original, flag_camas_disp_mayor_dotacion	Estado global de capacidad instalada y validación de calidad de dato.
Ratios operativos	indicadores derivados de egresos y camas	2	float64	ratio_egresos_por_cama_disponible, porcentaje_disponibilidad_camas	Indicadores sintéticos de presión y disponibilidad.

Grupo de columnas	Criterio de agrupación	Total de columnas	Tipo	Ejemplos	Descripción
Camas por servicio	columnas que inician con camas_ (excepto camas_disponibles_)	10	int64	camas_emergencia, camas_intensivos_adultos, camas_recuperacion	Distribución de camas por tipo de servicio clínico.
Camas disponibles por área funcional	columnas que inician con camas_disponibles_	16	int64	camas_disponibles_medicina, camas_disponibles_quirurgico, camas_disponibles_otros	Capacidad disponible desagregada por áreas funcionales.
Egresos por causa CIE-10 (conteo)	columnas que inician con egresos_causa_	22	int64	egresos_causa_neoplasias, egresos_causa_enfermedades_del_sistema_respiratorio	Conteo de egresos por macrogrupo diagnóstico CIE-10.
Egresos por causa CIE-10 (proporción)	columnas que inician con pct_egresos_causa_	22	float64	pct_egresos_causa_neoplasias, pct_egresos_causa_enfermedades_del_sistema_digestivo	Proporción de egresos por macrogrupo

Grupo de columnas	**Criterio de agrupación**	**Total de columnas**	**Tipo**	**Ejemplos**	**Descripción**
					diagnóstico CIE-10.
Total	**Suma de todos los grupos anteriores**	**103**			

Apéndice 3. Diccionario de columnas creadas en la fase de Ingeniería de Características para la tabla de modelado

nueva_columna	tipo	descripción
fecha	datetime64[us]	Fecha completa del registro (año-mes-día).
target_ratio_t_plus_7	float64	Variable objetivo: ratio de egresos por cama disponible desplazado a t+7 días.
dow	int32	Día de semana numérico de la fecha (0=lunes, 6=domingo).
is_weekend	int32	Indicador de fin de semana (1 si sábado/domingo, 0 caso contrario).
ratio_egresos_por_cama_disponible_lag_1	float64	Rezago de 1 día del ratio de egresos por cama disponible.
ratio_egresos_por_cama_disponible_lag_7	float64	Rezago de 7 días del ratio de egresos por cama disponible.

nueva_columna	tipo	descripción
egresos_totales_lag_1	float64	Rezago de 1 día de egresos totales.
egresos_totales_lag_7	float64	Rezago de 7 días de egresos totales.
total_camas_disponibles_original_lag_1	float64	Rezago de 1 día de camas disponibles originales.
total_camas_disponibles_original_lag_7	float64	Rezago de 7 días de camas disponibles originales.
ratio_egresos_por_cama_disponible_roll_mean_7	float64	Media móvil histórica (sin fuga) de 7 días del ratio de egresos por cama disponible.
ratio_egresos_por_cama_disponible_roll_std_7	float64	Desviación estándar móvil histórica (sin fuga) de 7 días del ratio de egresos por cama disponible.
ratio_egresos_por_cama_disponible_roll_mean_28	float64	Media móvil histórica (sin fuga) de 28 días del ratio de egresos por cama disponible.
ratio_egresos_por_cama_disponible_roll_std_28	float64	Desviación estándar móvil histórica (sin fuga) de 28 días del ratio de egresos por cama disponible.
egresos_totales_roll_mean_7	float64	Media móvil histórica (sin fuga) de 7 días de egresos totales.
egresos_totales_roll_std_7	float64	Desviación estándar móvil histórica (sin fuga) de 7 días de egresos totales.
egresos_totales_roll_mean_28	float64	Media móvil histórica (sin fuga) de 28 días de egresos totales.
egresos_totales_roll_std_28	float64	Desviación estándar móvil histórica (sin fuga) de 28 días de egresos totales.

nueva_columna	tipo	descripción
total_camas_disponibles_original_roll_mean_7	float64	Media móvil histórica (sin fuga) de 7 días de camas disponibles originales.
total_camas_disponibles_original_roll_std_7	float64	Desviación estándar móvil histórica (sin fuga) de 7 días de camas disponibles originales.
total_camas_disponibles_original_roll_mean_28	float64	Media móvil histórica (sin fuga) de 28 días de camas disponibles originales.
total_camas_disponibles_original_roll_std_28	float64	Desviación estándar móvil histórica (sin fuga) de 28 días de camas disponibles originales.
pct_egresos_causa_embarazo_parto_y_puerperio_lag_7	float64	Rezago de 7 días del porcentaje de egresos por causa embarazo, parto y puerperio.
pct_egresos_causa_enfermedades_del_sistema_digestivo_lag_7	float64	Rezago de 7 días del porcentaje de egresos por causa enfermedades del sistema digestivo.
pct_egresos_causa_traumatismos_envenenamientos_lag_7	float64	Rezago de 7 días del porcentaje de egresos por causa traumatismos y envenenamientos.
pct_egresos_causa_enfermedades_del_sistema_genitourinario_lag_7	float64	Rezago de 7 días del porcentaje de egresos por causa enfermedades del sistema genitourinario.
pct_egresos_causa_enfermedades_del_sistema_respiratorio_lag_7	float64	Rezago de 7 días del porcentaje de egresos por causa enfermedades del sistema respiratorio.
pct_egresos_causa_neoplasias_lag_7	float64	Rezago de 7 días del porcentaje de egresos por causa neoplasias.

nueva_columna	tipo	descripción
pct_egresos_causa_enfermedades_infecciosas_y_parasitarias_lag_7	float64	Rezago de 7 días del porcentaje de egresos por causa enfermedades infecciosas y parasitarias.
pct_egresos_causa_enfermedades_del_sistema_circulatorio_lag_7	float64	Rezago de 7 días del porcentaje de egresos por causa enfermedades del sistema circulatorio.

Apéndice 4. Muestra de 10 registros de predicción detallada

Figura A1

Muestra de registros de predicción detallada por grupo hospitalario y fecha

	A	B	C	D	E	F	G	H	I	J	K	L
1	data_scena	feature_se	modelo	split	fecha	y_real	y_pred	n_features	MAE	RMSE	WMAPE	grupo_hos
2	con_outliers	all	Naive_t	test	1/7/2024	0.203883495	0.067961165	140	0.077023720	0.120279184	0.468955113	4
3	con_outliers	all	Naive_t	test	1/7/2024	0.1	0.05	140	0.077023720	0.120279184	0.468955113	9
4	con_outliers	all	Naive_t	test	1/7/2024	0.16	0.12	140	0.077023720	0.120279184	0.468955113	13
5	con_outliers	all	Naive_t	test	1/7/2024	0.020833333	0.041666666	140	0.077023720	0.120279184	0.468955113	16
6	con_outliers	all	Naive_t	test	1/7/2024	0.076923076	0.128205128	140	0.077023720	0.120279184	0.468955113	17
7	con_outliers	all	Naive_t	test	1/7/2024	0.083333333	0.208333333	140	0.077023720	0.120279184	0.468955113	27
8	con_outliers	all	Naive_t	test	1/7/2024	0.183333333	0.166666666	140	0.077023720	0.120279184	0.468955113	28
9	con_outliers	all	Naive_t	test	1/7/2024	0.211111111	0.138888888	140	0.077023720	0.120279184	0.468955113	29
10	con_outliers	all	Naive_t	test	1/7/2024	0.090090090	0.108108108	140	0.077023720	0.120279184	0.468955113	33

	N	O	P	Q	R	S	T	U	V	W	X	Y
1	egresos_to	total_cama	total_cama	total_cama	total_cama	ratio_egre	total_cama	pct_egresc	pct_egresc	pct_egresc	pct_egresc	pct_egresc
2	7	103	103	114	103	0.067961165	103	0.0	0.0	0.0	0.1053	0.1429
3	1	20	20	38	20	0.05	20	0.6667	1.0	1.0	0.0	0.0
4	3	25	25	25	25	0.12	25	1.0	0.6667	0.0	0.0	0.5
5	2	48	48	59	48	0.041666666	48	1.0	0.0	0.0	0.0	0.0
6	5	39	39	85	39	0.128205128	39	0.125	0.0	0.0	0.0	0.25
7	5	24	24	41	25	0.208333333	24	0.0	0.2	0.25	0.0	0.0
8	10	60	60	94	60	0.166666666	60	0.0	0.0	0.0	0.0	0.0
9	25	180	180	229	181	0.138888888	180	0.0294	0.04	0.0	0.0882	0.0
10	12	111	111	199	111	0.108108108	111	0.0	0.0	0.1667	0.0	0.5

	Z	AA	AB	AC	AD	AE	AF	AG
1	pct_egresc	pct_egresc	pct_egresc	causa_top	error	abs_error	ape	pct_egresc
2	0.4286	0.0	0.1429	pct_egresos_	-0.135922330	0.135922330	0.6666666666666666	
3	0.0	0.0	0.0	pct_egresos_	-0.05	0.05	0.5	
4	0.0	0.3333	0.5	pct_egresos_	-0.040000000	0.040000000	0.25000000000000006	
5	0.5	0.0	0.0	pct_egresos_	0.020833333	0.020833333	1.0	
6	0.2	0.4	0.375	pct_egresos_	0.051282051	0.051282051	0.6666666666666664	
7	0.2	0.0	0.25	pct_egresos_	0.125	0.125	1.5	
8	0.0	0.5	1.0	pct_egresos_	-0.016666666	0.016666666	0.0909090909090909	
9	0.08	0.24	0.5	pct_egresos_	-0.072222222	0.072222222	0.3421052631578947	
10	0.0	0.5833	0.1667	pct_egresos_	0.018018018	0.018018018	0.20000000000000012	