



Análisis de Trasplantes Hepáticos

Fundación Santa Fe de Bogotá

Universidad de Los Andes
MINE 4101 – Ciencia de datos
2025-20

- Lina María Bejarano Rivera
- Juan Diego Enríquez Ramos
- Juan Manuel Rivera López
- Johana Alejandra Rátiva Mora



Comprensión del Caso de Uso

El Desafío

La Fundación Santa Fe consolidó datos de pacientes con trasplante hepático.

La información estaba dispersa, dificultando análisis clínico y toma de decisiones.

La Solución

Desarrollar un tablero de control que facilite agilice la consulta de los datos y analizar los datos para encontrar insights para el negocio

Recolección de los Datos

Base de datos recibida el **25 de septiembre de 2025** con información clínica, quirúrgica y epidemiológica de pacientes sometidos a trasplante hepático.

736

Registros de Pacientes

Total de casos de trasplante hepático consolidados

286

Variables Iniciales

Datos operativos, clínicos y quirúrgicos

121

Variables Priorizadas

Seleccionadas tras estrategia de depuración

01

Selección Inicial

Variables relevantes según pregunta de negocio

02

Evaluación de Completitud

Eliminación de variables vacías o sin registros útiles

03

Medición de Variabilidad

Descarte de variables con baja variabilidad o completitud

04

Validación Clínica

Revisión con stakeholders para asegurar pertinencia



Preparación de Datos

Resultado de la estrategia de priorización: de 286 variables iniciales a 121 variables seleccionadas para análisis estadístico.



Limpieza de Datos

Script en Python para depuración avanzada y tratamiento de valores nulos.



Análisis Exploratorio

EDA univariado y bivariado documentado en cuadernillos adjuntos.



Diccionario de Datos

Documentación detallada de variables priorizadas para análisis.



Responsable: Ética y Cumplimiento

1

Clasificación Sin Riesgo

Según Resolución 8430 de 1993 del Ministerio de Salud de Colombia, usando datos retrospectivos anonimizados.

2

Acuerdo de Confidencialidad

Suscrito con el Servicio de Trasplantes, garantizando uso ético y seguro de la información.

3

Ley 1581 de 2012

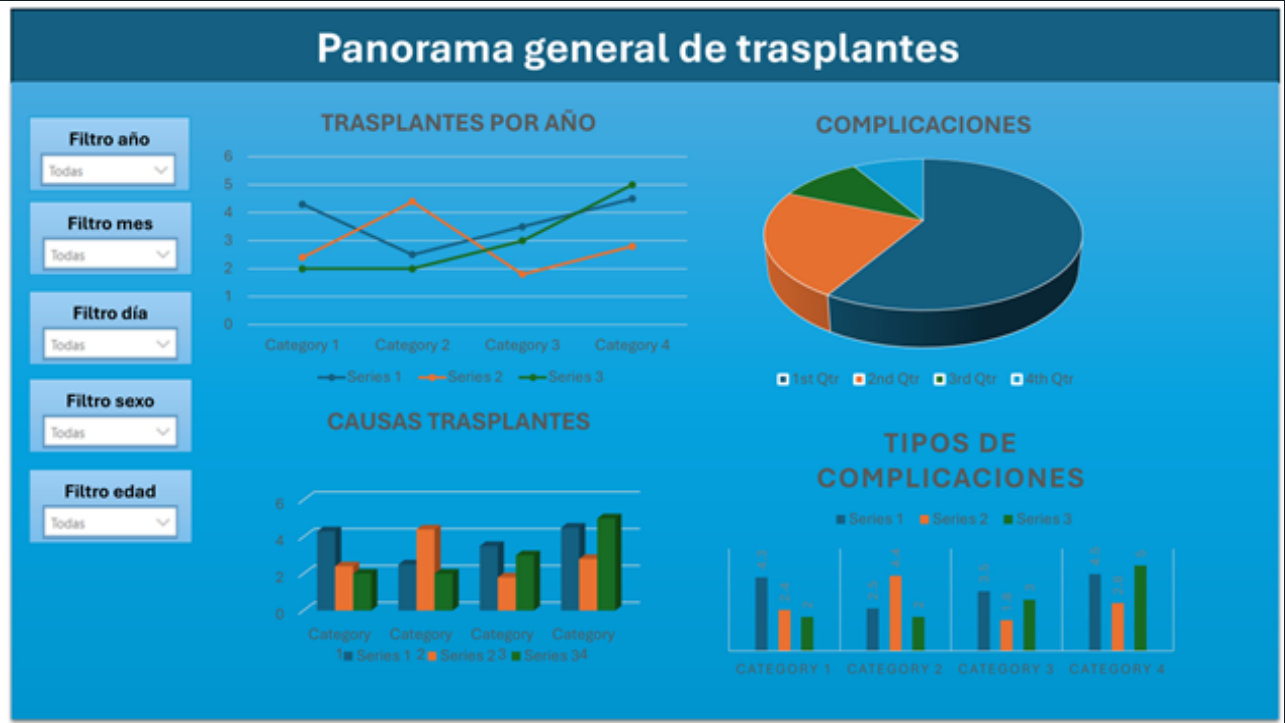
Cumplimiento de protección de datos personales, principios de confidencialidad, legalidad y finalidad.

Ideación del Producto

Dashboard en Power BI con tres vistas principales diseñadas para médicos rurales y equipo clínico.

Panorama General

Reportes operativos y descripción poblacional con filtros por edad, sexo, año y variables clínicas.



Perfil del Paciente

Información individual de complicaciones, riesgos y variables relevantes del análisis exploratorio.

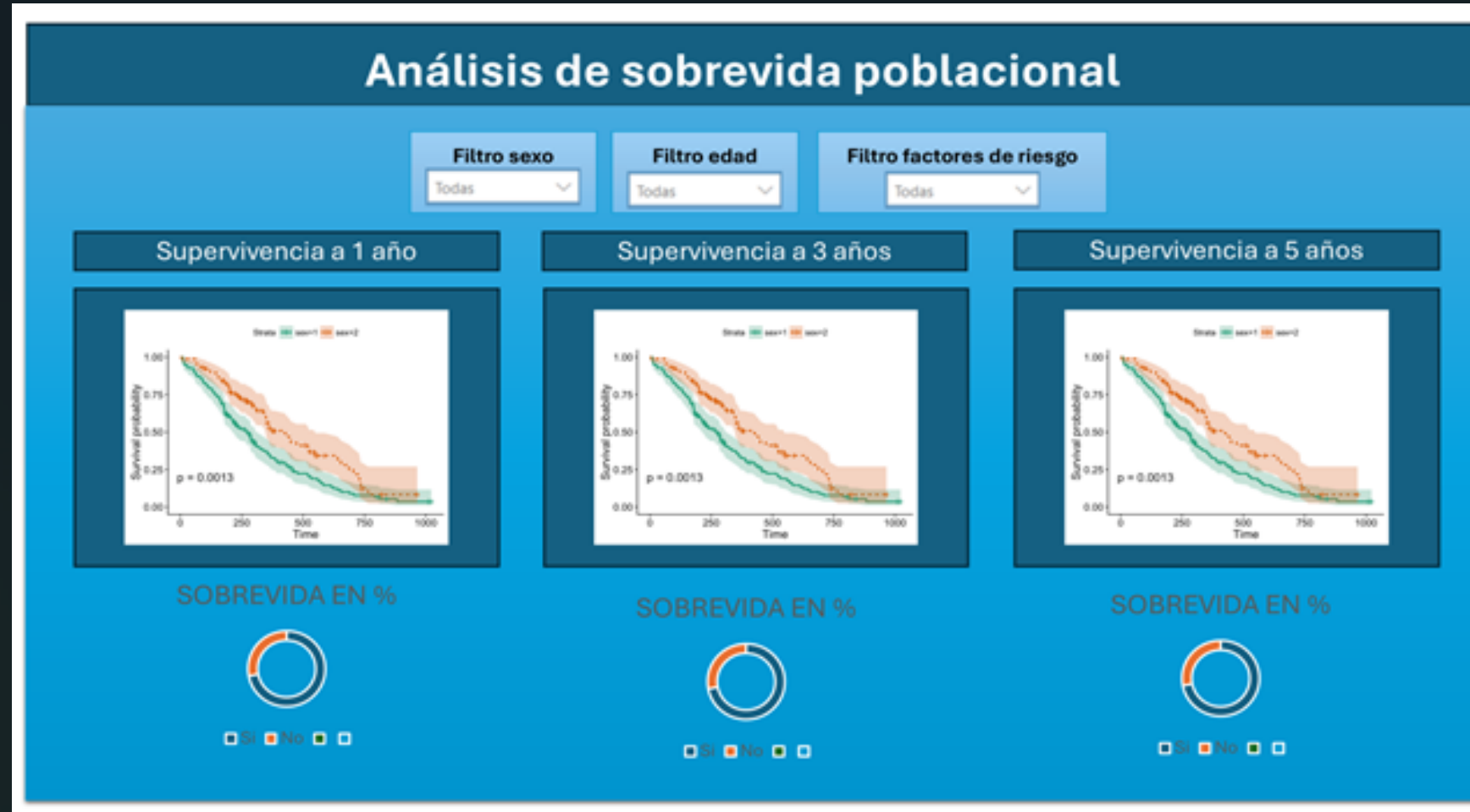


Ideación del Producto

Dashboard en Power BI con tres vistas principales diseñadas para médicos rurales y equipo clínico.

Análisis de Sobrevida

Tasas de supervivencia por grupos y variables que más inciden en resultados.



Ideación del Producto

API para consulta de datos

/paciente/{{código_anonimización}}

Consulta de historia clínica del paciente

/complicaciones/

Predicción de si un paciente tendrá complicaciones dada una serie de variables

Enfoque Analítico

Objetivo del Modelo

- El objetivo de este modelo analítico es explorar y modelar los factores asociados a complicaciones postoperatorias.



Análisis Exploratorio

Identificar patrones, distribuciones y relaciones entre variables clínicas y quirúrgicas.



Regresión Logística

Estimar probabilidad de complicación según edad, estadio tumoral y marcadores bioquímicos.



Evaluación

Métricas de exactitud, precisión y recall para medir capacidad predictiva.

KPIs de Impacto

1

Compleitud de Datos

Porcentaje de mejora en calidad, incentivando captura más limpia con diccionarios de datos.

2

Tiempo de Depuración

Reducción en duración del proceso de limpieza y estandarización de datos.

3

Comprensión del Paciente

Tiempo promedio para entender estado del paciente mediante perfil dinámico.

4

Investigación Derivada

Abstracts, manuscritos y nuevas líneas de investigación clínica generadas.

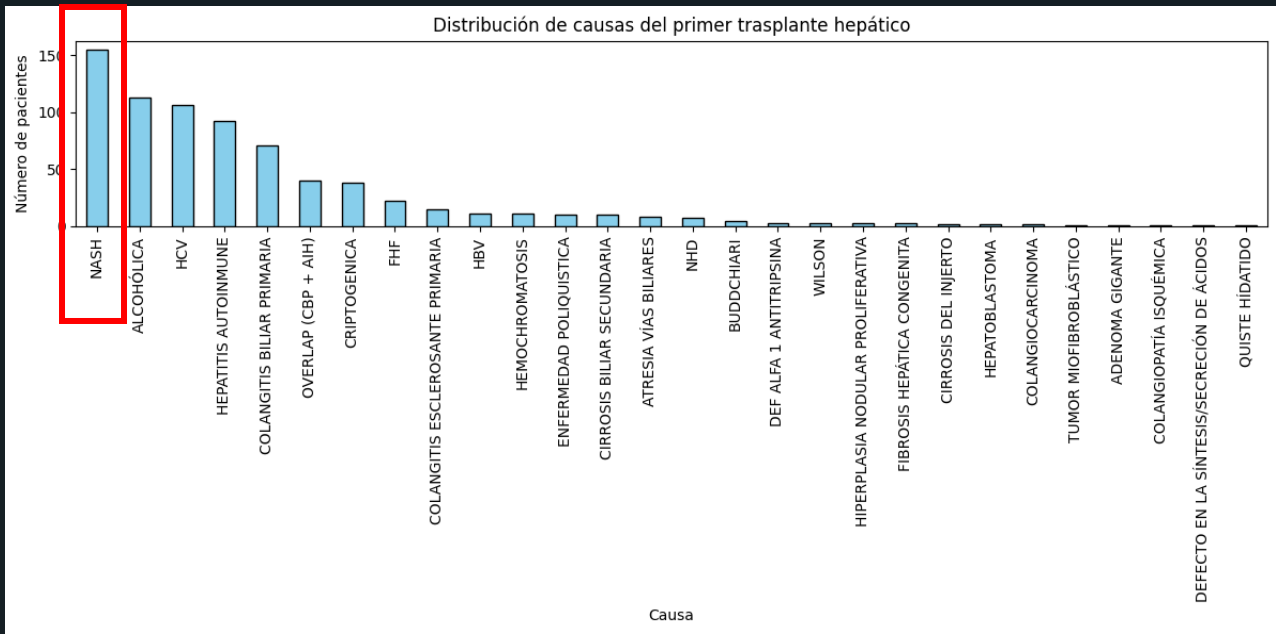
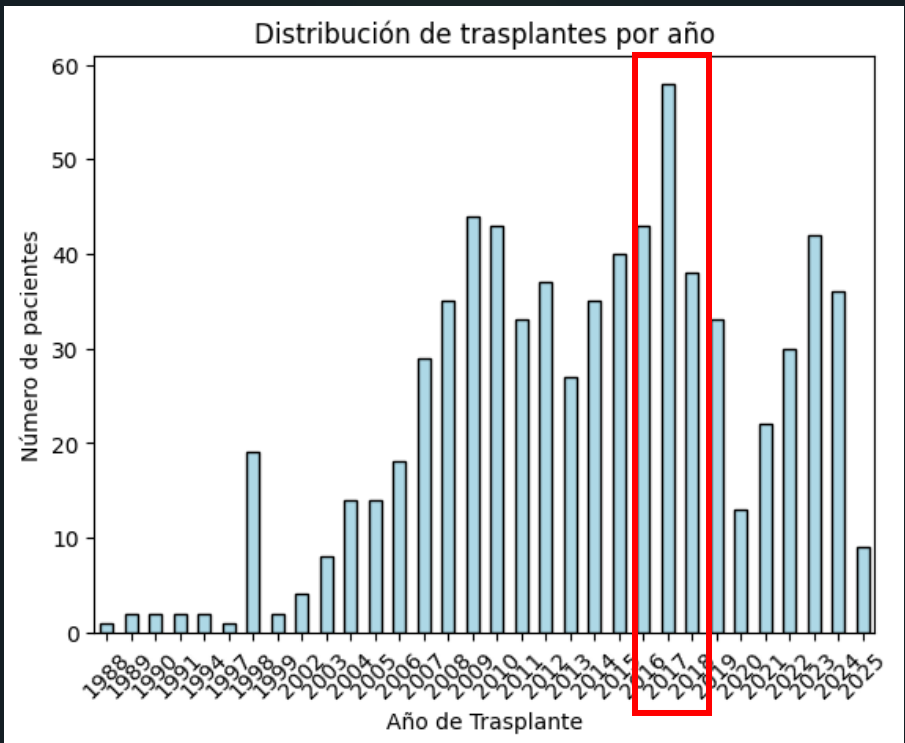
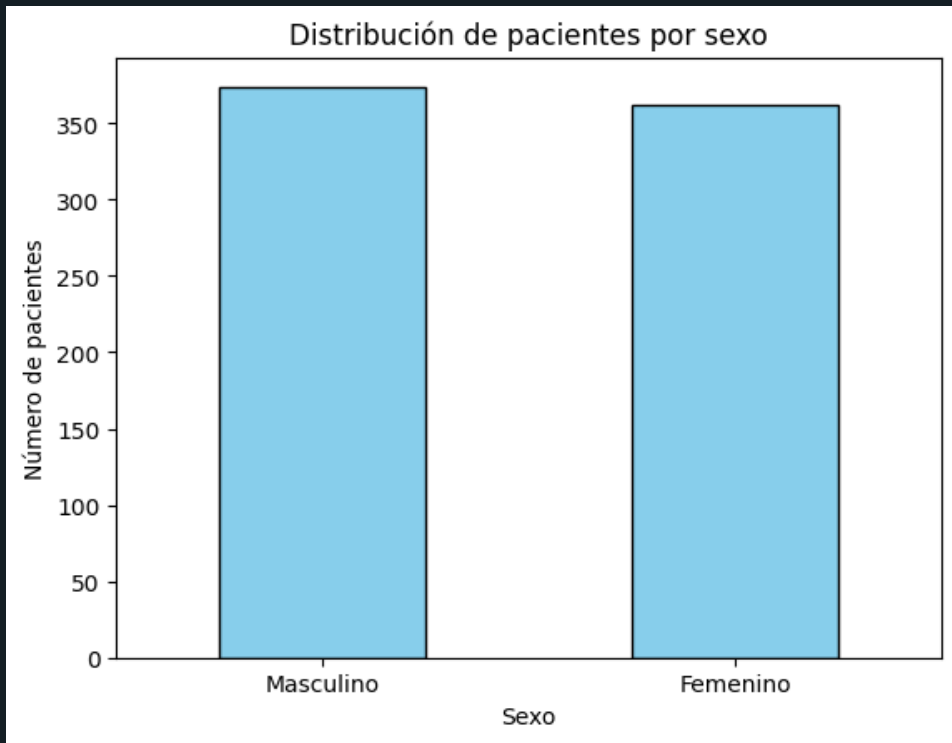
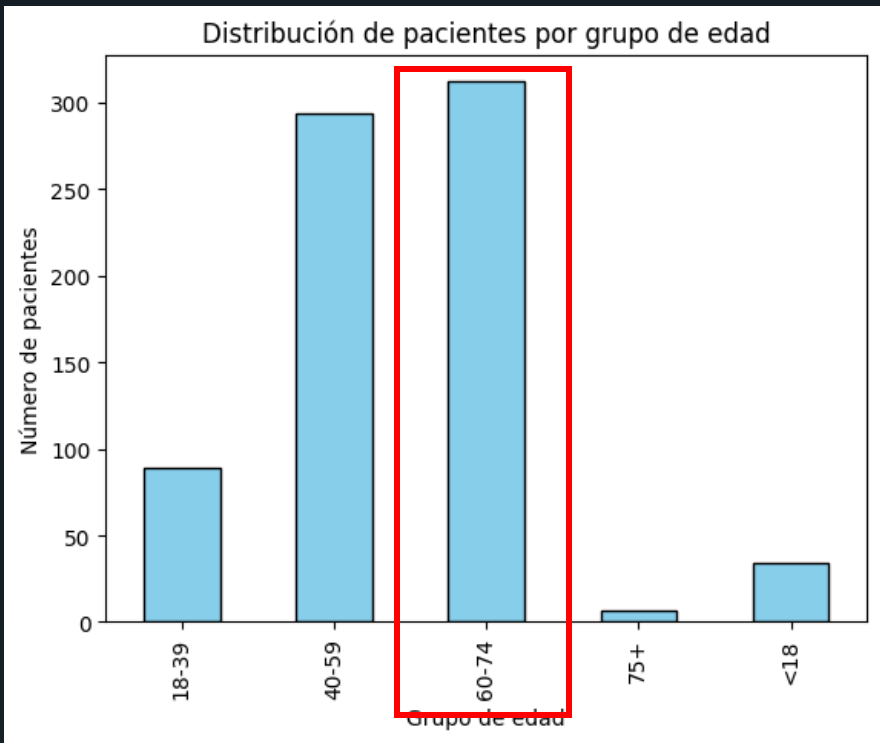
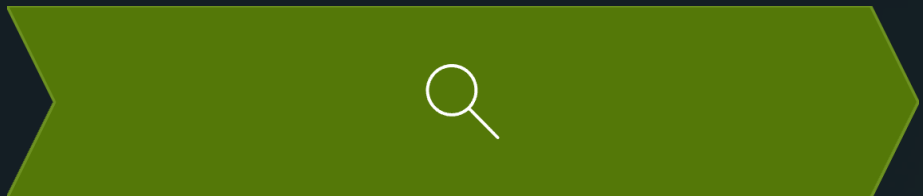
5

Frecuencia de Uso

Aumento del acceso y aprovechamiento de información antes subutilizada.

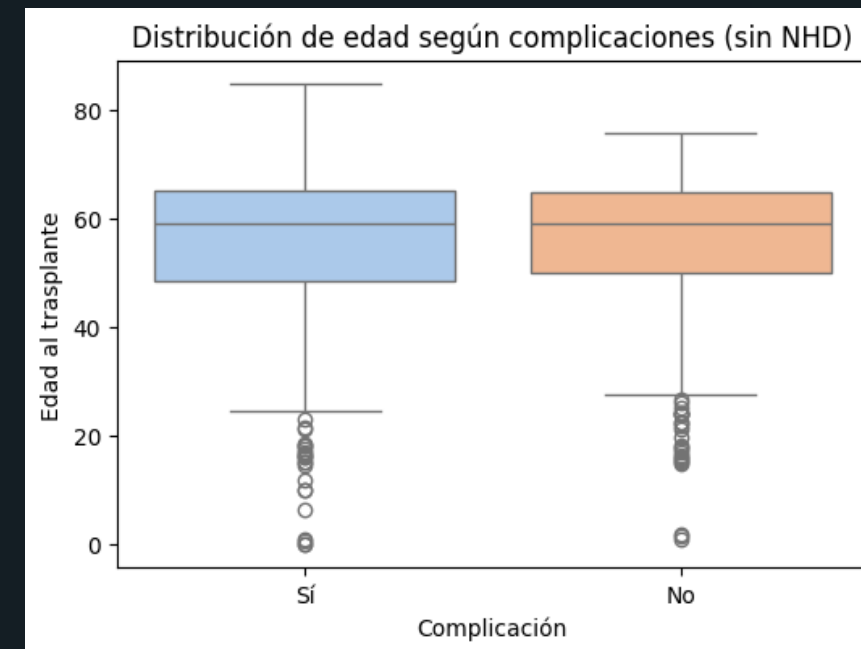
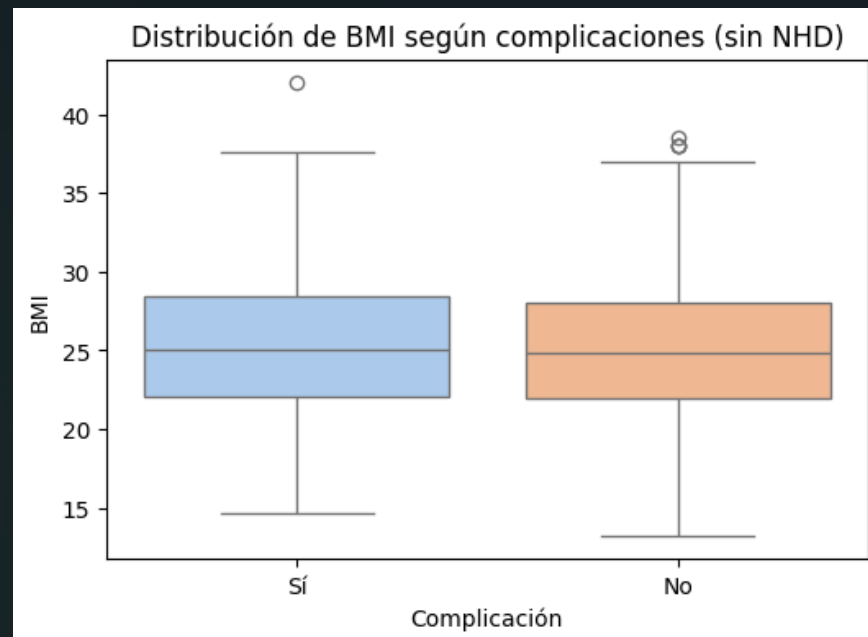
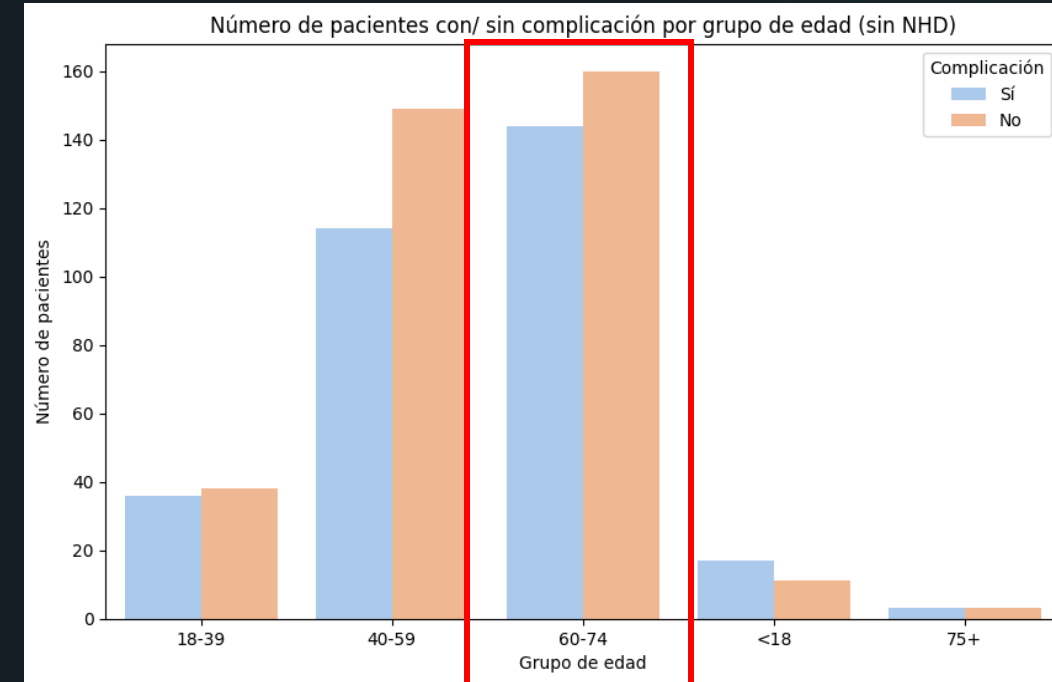
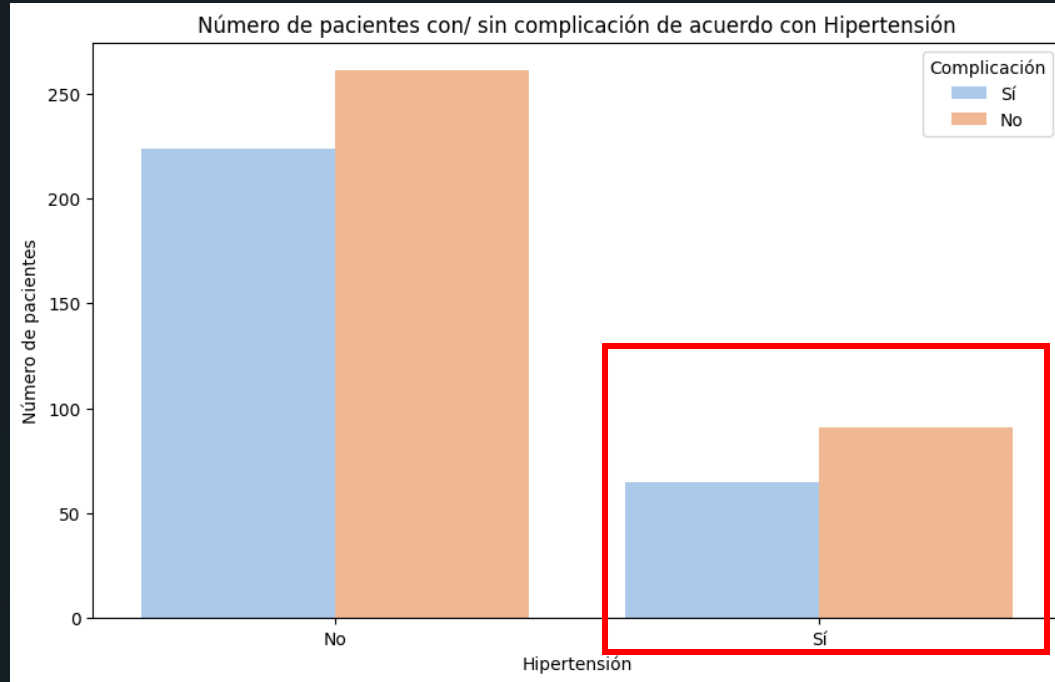
Enfoque Analítico

- Adjunto notebook limpieza de datos



Enfoque Analítico

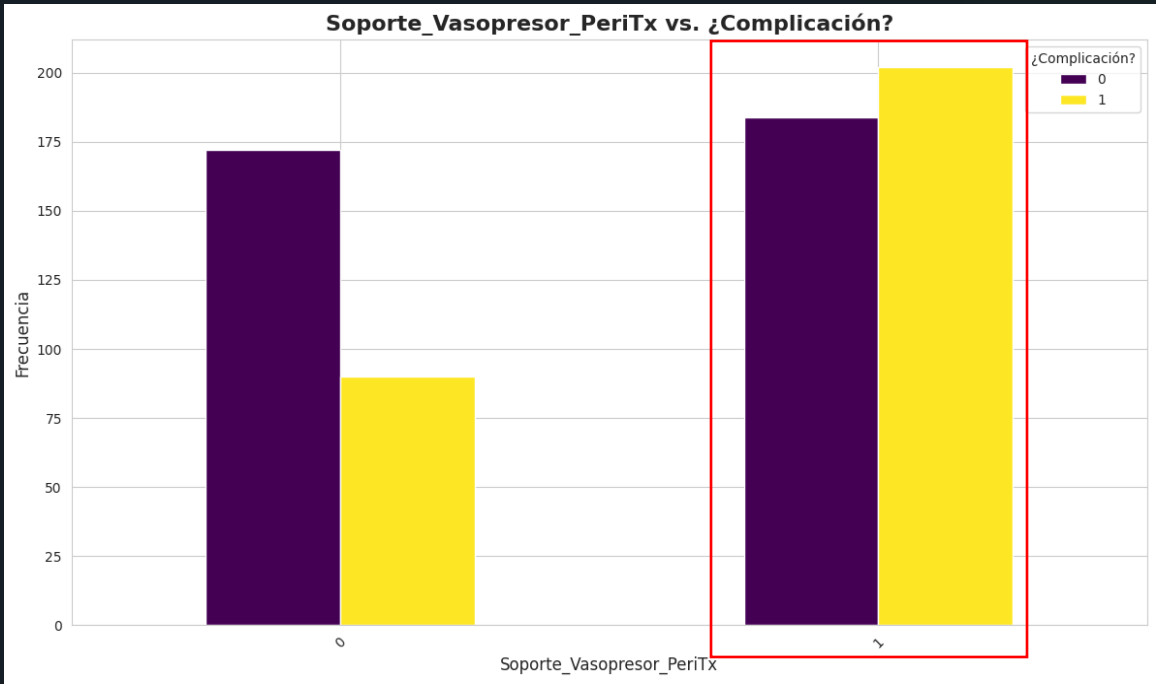
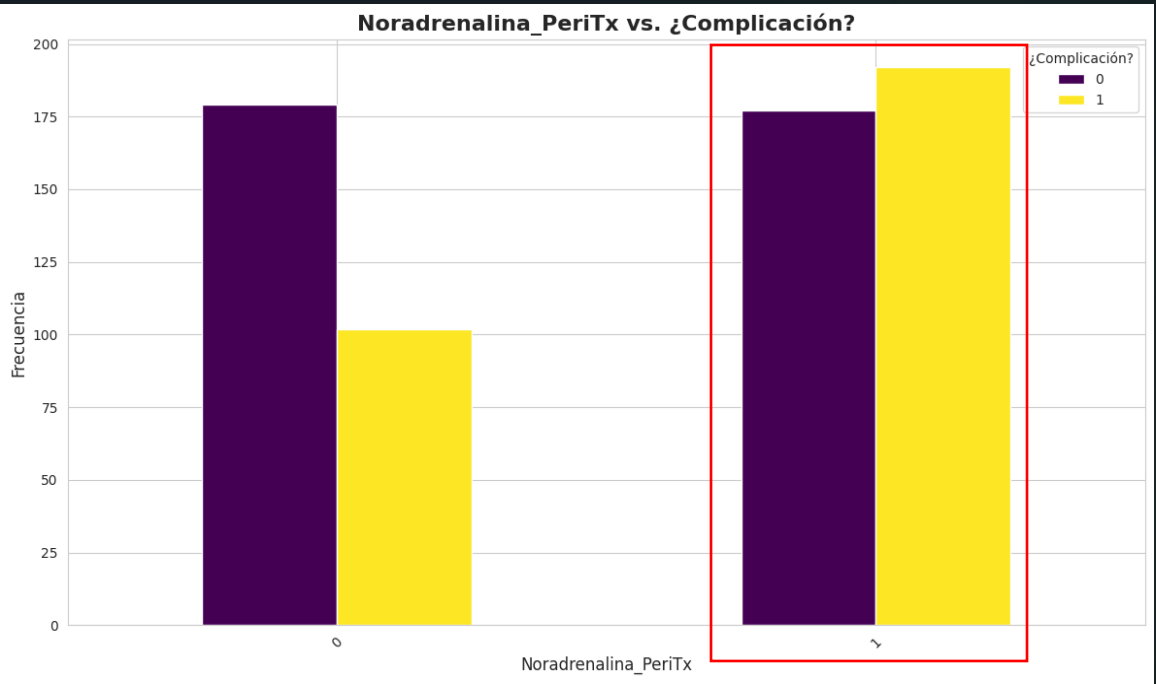
- Adjunto notebook limpieza de datos



Enfoque Analítico

- Adjunto notebook limpieza de datos

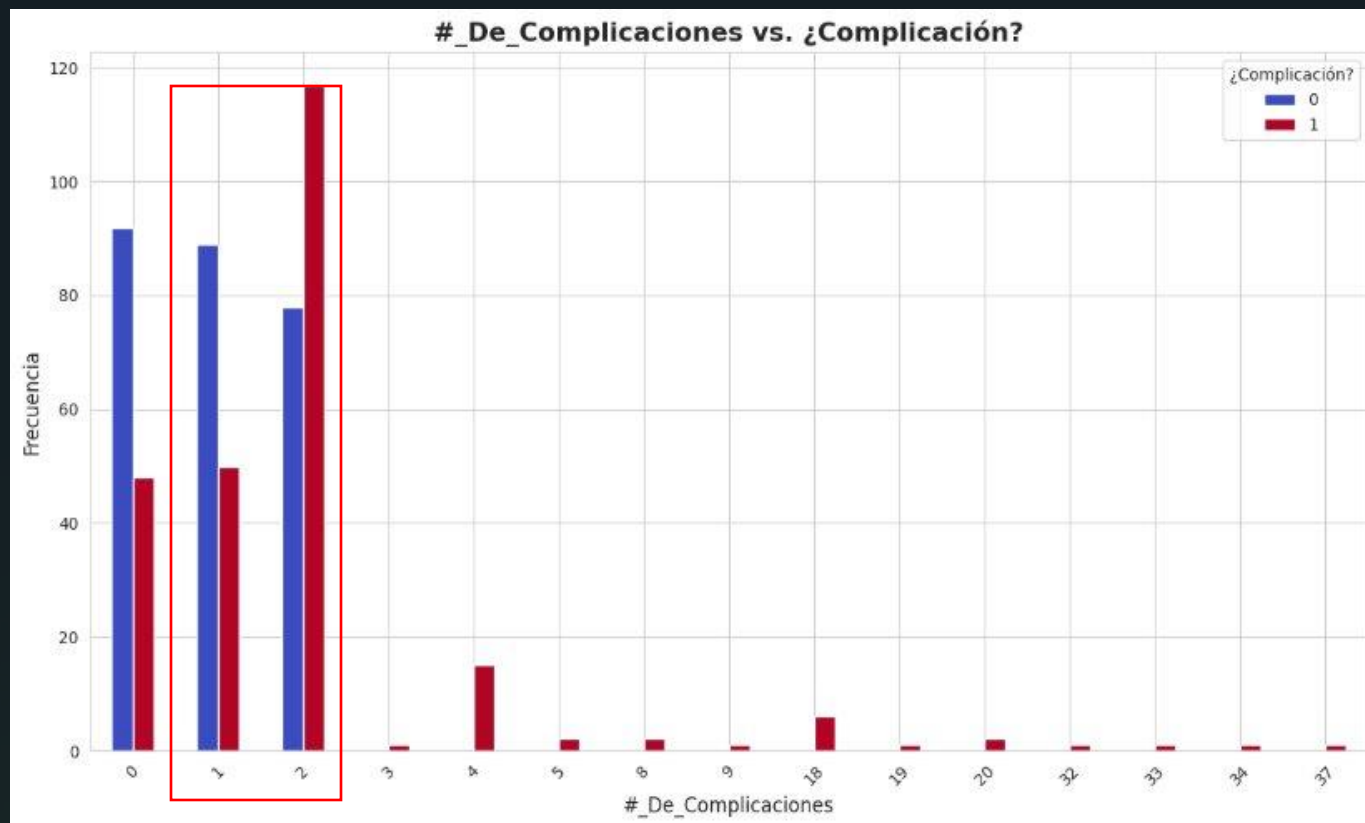
El uso de algunas hormonas usadas durante la cirugía (vasopresina y noradrenalina) tienen relación estadística con la ocurrencia de complicaciones



Enfoque Analítico



- Adjunto notebook limpieza de datos



- “¿Complicación?” presenta problemas de consistencia con otras variables complementarias:
 - Complejidad_asociada
 - #_de_complicaciones
 - arritmia_pop
 - falla_cardiaca
 - Infección
 - rechazo_agudo
 - rechazo_cronico
 - Infarto



Conclusiones y Próximos Pasos

- “¿Complicación?” necesita ser depurada y homologada antes de avanzar en la modelación predictiva.
- Si el paciente necesita algún tipo de soporte durante la cirugía es posible que luego presente complicaciones
- Hay variables que tienen una tendencia de presentar complicaciones, pero no llega a ser significativa. Sin embargo, estas variables tienen sentido clínico y no deben ser excluidas
- Dada la complejidad de caso, es necesario realizar **análisis multivariados** para identificar variables que estén relacionadas a las complicaciones