

2025

Fundamentos de la Ciencia de Datos

Trabajo Práctico Especial

Horse Colic



Grupo 11

Rivera, Tomas ; Ferraro, Bautista

1. Introduccion

El dataset que se nos asignó es el conjunto de datos de Cólico Equino creado por Mary McLeish y Matt Cecile en la Universidad de Guelph y donado al repositorio de UCI en 1989. Contiene 368 casos y 28 atributos . El núcleo del problema abordado por el dataset no es una enfermedad específica, sino un síndrome de dolor abdominal en caballos denominado **cólico**. Los caballos ingresan a urgencias por la presencia de estos cólicos y es allí donde comienza el desafío para los veterinarios : **Determinar el origen del problema.**

Aquí comenzó nuestro primer desafío : Comenzar a indagar en la problemática de nuestro dataset. Gracias al uso de la IA y las capacidades de la misma logramos crear un informe (Adjunto en las referencias) en donde se explica cuál es el sentido de cada variable y que análisis debemos tener en cuenta para poder expresar al máximo la información presente en el dataset, todo analizado desde el punto de vista veterinario y técnico lo que nos permite cambiar nuestro enfoque sobre el dataset y abrir nuestras mentes para realizar el mejor trabajo posible a pesar de estar en un ambiente no habitual.

No obstante, es aquí donde comenzamos a enfrentarnos a las problemáticas que presenta el conjunto de datos: La gran cantidad de valores faltantes (cerca del 30%) , el poco contexto que tenían las variables y en algunos casos la definición errónea de las mismas.

A continuación se presenta el informe que recopila todo el trabajo realizado , dividido en las 6 distintas hipótesis que planteamos y que consideramos más relevantes e importantes para el correcto análisis del dataset como también para demostrar mediante nuestras herramientas cuales son las evidencias que nos dan las variables que nos pueden ayudar a tomar decisiones que sean adecuadas.

2. Análisis exploratorio de los datos

El Análisis Exploratorio de Datos (EDA) se centró en comprender la naturaleza, la distribución y las interrelaciones de las variables, lo que guió tanto las transformaciones de datos como el planteo de las hipótesis.

Para contextualizar el informe, se presenta un resumen de las variables contenidas en el dataset, clasificadas según su tipo:

- **surgery:** Variable cualitativa nominal que indica si el caballo recibió tratamiento quirúrgico (1) o no (0).
- **Age:** Variable cualitativa nominal que clasifica al animal como adulto (1) o joven <6 meses (0).
- **Hospital Number:** Identificador del caso, variable cualitativa nominal.
- **rectal temperature:** Variable cuantitativa continua que mide la temperatura en °C; se eleva en infecciones o disminuye en shock.
- **pulse:** Variable cuantitativa continua que registra la frecuencia cardíaca; aumenta con dolor o shock.
- **respiratory rate:** Variable cuantitativa continua; muy fluctuante, normal entre 8–10.
- **temperature of extremities:** Variable cualitativa ordinal que evalúa la circulación periférica (normal → fría).

- **peripheral pulse:** Variable cualitativa ordinal que mide la perfusión (normal → ausente).
- **mucous membranes:** Variable cualitativa ordinal que describe el color de mucosas, indicador del estado circulatorio.
- **capillary refill time:** Variable cualitativa ordinal que clasifica el relleno capilar como $<3s$ o $\geq 3s$.
- **pain:** Variable cualitativa nominal que evalúa el nivel de dolor; no debe tratarse como ordinal.
- **peristalsis:** Variable cualitativa ordinal que refleja la actividad intestinal.
- **abdominal distension:** Variable cualitativa ordinal vinculada a gravedad y probabilidad de cirugía.
- **nasogastric tube:** Variable cualitativa ordinal que indica presencia de gas por sonda.
- **nasogastric reflux:** Variable cualitativa ordinal que mide la cantidad de reflujo.
- **nasogastric reflux PH:** Variable cuantitativa continua que registra el pH del reflujo.
- **rectal examination - feces:** Variable cualitativa ordinal que evalúa la presencia de heces.
- **abdomen:** Variable cualitativa nominal que describe hallazgos abdominales.

- **packed cell volume:** Variable cuantitativa continua que mide el hematocrito.
- **total protein:** Variable cuantitativa continua que registra proteína total en sangre.
- **abdominocentesis appearance:** Variable cualitativa nominal que describe el aspecto del líquido abdominal.
- **abdomcentesis total protein:** Variable cuantitativa continua que mide proteína en líquido abdominal.
- **outcome:** Variable cualitativa nominal que indica si el caballo vivió, murió o fue eutanasiado. (1 = vivio ; 2= murio ; 3 =eutanasiado)
- **surgical lesion:** Variable cualitativa nominal que determina si la lesión fue quirúrgica.
- **type of lesion (25–27):** Conjunto de variables cualitativas nominales que describen sitio, tipo y subtipo de la lesión.
- **cp_data:** Variable cualitativa nominal que indica si existen datos patológicos (no relevantes en este dataset).

2.1. El Desafío de los Datos Faltantes (NaN)

El primer hallazgo significativo fue la cantidad de valores faltantes. Aproximadamente el **30% de los datos son nulos**. Se tomó la decisión de no completar aquellos valores debido a que implicaba una gran inversión de tiempo en metodologías para completarlos que no nos iban a garantizar mejores resultados y que, en este contexto, Este patrón de "ausencia informativa" tiene implicaciones directas para el análisis de datos. La imputación simple (reemplazar los valores faltantes con la media, mediana o moda) podría ser perjudicial, ya que destruiría esta valiosa información contextual. Por ejemplo, si los valores faltantes en las variables de abdominocentesis se reemplazan con el valor "normal", el modelo podría interpretar erróneamente un caso grave como uno leve. Imputar la media sería un error metodológico grave, y eliminar o ignorar aquellas filas sería la decisión más segura para no inventar datos.

Es así como trabajamos, dependiendo del enfoque, con aquellas muestras que si nos eran representativas para el contexto que necesitábamos.

```
surgery          2
age              0
hospital_number  0
rectal_temperature 69
pulse            26
respiratory_rate 71
temperature_of_extremities 65
peripheral_pulse 83
mucous_membranes 48
capillary_refill_time 38
pain             63
peristalsis      52
abdominal_distension 65
nasogastric_tube 131
nasogastric_reflux 133
nasogastric_reflux_ph 299
rectal_examination_feces 128
abdomen          143
packed_cell_volume 37
total_protein    43
abdominocentesis_appearance 194
abdominocentesis_total_protein 235
outcome          2
surgical_lesion  0
lesion_1         0
lesion_2         0
lesion_3         0
cp_data          0
dtype: int64
```

Fig. 1: Resultado de la ejecución de la línea de código “df_completo.isnull().sum()” en nuestro dataset donde se aprecia la suma de nulls presentes por variable.

2.2 El porqué y el manejo de los datos faltantes:

Consideramos importante tratar de entender el porqué de esta falta de datos y luego de una investigación de cómo es que se recopilan los mismos concluimos:

La alta tasa de valores faltantes (30%) en este conjunto de datos no debe ser vista como un defecto o un error de recopilación, sino como una característica intrínseca y representativa de la medicina de emergencia equina. Los datos faltantes no son aleatorios; su ausencia a menudo está correlacionada con el estado del paciente y las decisiones clínicas tomadas bajo presión. Así también como una consecuencia de la falta de disponibilidad de equipos o de la prioridad absoluta del veterinario en estabilizar al animal.

Para la validación de nuestras hipótesis , analizamos las variables implicadas en cada una y el % de datos faltantes para corroborar el peso estadístico de los mismos en cada contexto. Cabe destacar que a la hora de plantear las hipótesis, intentamos usar aquellas variables cuyo porcentaje de datos faltantes sea menor al 20% del total de las muestras , descartando aquellas que poseen datos faltantes.

2.3. Encoding de variables:

Eliminación de hospital_number :

Es un identificador numérico que no aporta poder predictivo y puede no ser único, por lo que fue excluido.

Eliminación de cp_data:

Fue eliminada ya que la documentación original del dataset establece explícitamente que esta variable "no tiene ninguna significancia".

Recodificación Binaria

Las variables binarias (surgery, surgical_lesion, capillary_refill_time) se codificaron para que el valor 2 pasará a ser 0. Esto asegura que la clase positiva sea 1 y la negativa sea 0, estandarizando el formato para los modelos.

Ajuste de age: El código 9 (para caballo joven en el set de datos) se unificó con la clase 0 (Joven), dejando 1 para el caballo Adulto.

2.4. Manejo de columnas lesion1, lesion2, lesión 3:

Al explorar esta columna notamos que toda la información acerca de las 3 posibles lesiones que podía presentar un caballo se encontraba unificada en una única columna (lesión 1) y el resto contenía todos ceros.

Optamos por crear nuevas columnas que descomponen la información de las lesiones en cuatro categorías: sitio, tipo , subtipo y lesion_especifica para un mejor manejo de la variable y una más fácil comprensión .

2.5. Las Variables de Resultado: Outcome y

Surgical_lesion:

Estas variables describen el resultado final del caso y el diagnóstico definitivo. La elección de cuál de ellas usar como objetivo (target) para un modelo de aprendizaje automático determinará la pregunta que el modelo está diseñado para responder.

Aquí es donde plantamos las bases de nuestro análisis : Seleccionar como variables objetivo a **surgical_lesion** y **outcome** y luego “tirar” de ellas para analizarlos:

2.5.1. Outcome :

Esta es la variable de resultado final a largo plazo. Si bien parece un objetivo de predicción obvio, está significativamente influenciado por factores no biológicos. La decisión de eutanasiar a un caballo (3=was euthanized) no siempre se basa en la inevitabilidad médica. A menudo, es una decisión tomada por el propietario debido a un pronóstico reservado, los altos costos asociados con la cirugía y los cuidados intensivos postoperatorios, o una combinación de ambos. Por lo tanto, un modelo entrenado para predecir outcome podría aprender a identificar no solo los casos médicamente desesperados, sino también aquellos que son "caros de tratar" o que tienen un pronóstico incierto, introduciendo un sesgo socioeconómico en la predicción.

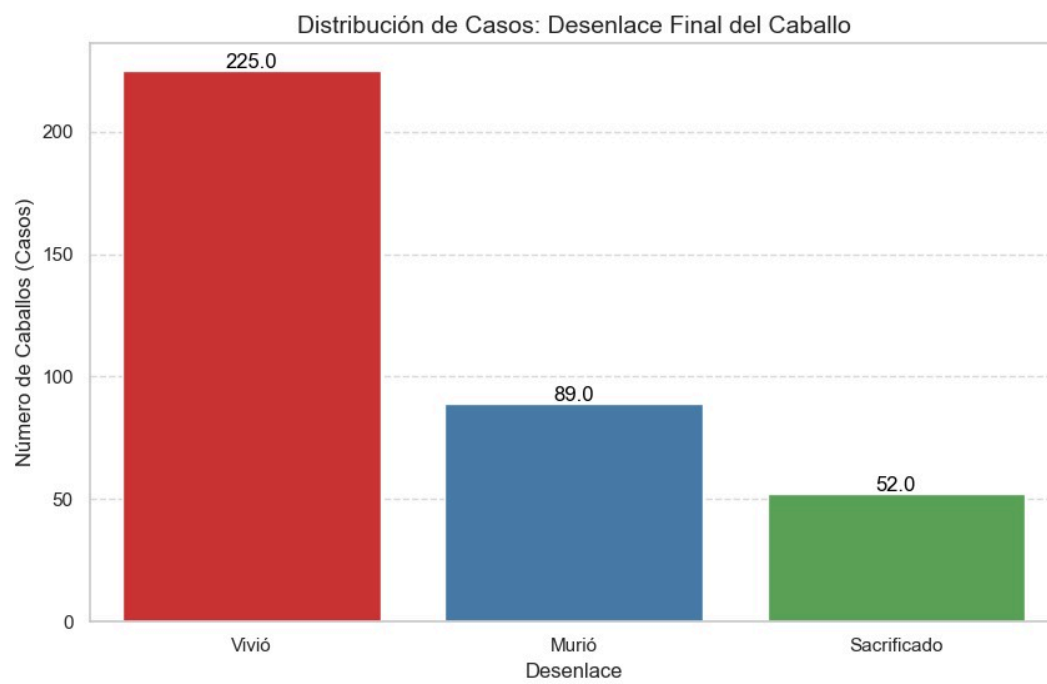


Fig. 2: Distribucion de la variable outcome.

Nótese la gran cantidad de desenlaces fatales (141).

2.5.2. Surgical_Lesion:

Esta variable es, desde una perspectiva clínica, el objetivo de clasificación **más valioso y limpio**. Indica si el problema subyacente del caballo era una lesión que requería intervención quirúrgica. Es crucial entender que este valor se determinó retrospectivamente, con certeza, ya sea durante la cirugía o en la autopsia. Esto la convierte en una etiqueta o "verdad fundamental", libre de la incertidumbre y las conjeturas del diagnóstico inicial. Un modelo entrenado para predecir surgical_lesion a partir de las variables del examen inicial está, en efecto, intentando predecir el diagnóstico final y definitivo. Su propósito es emular y apoyar la decisión más crítica del veterinario en el campo: ¿necesita este caballo ser referido para cirugía?

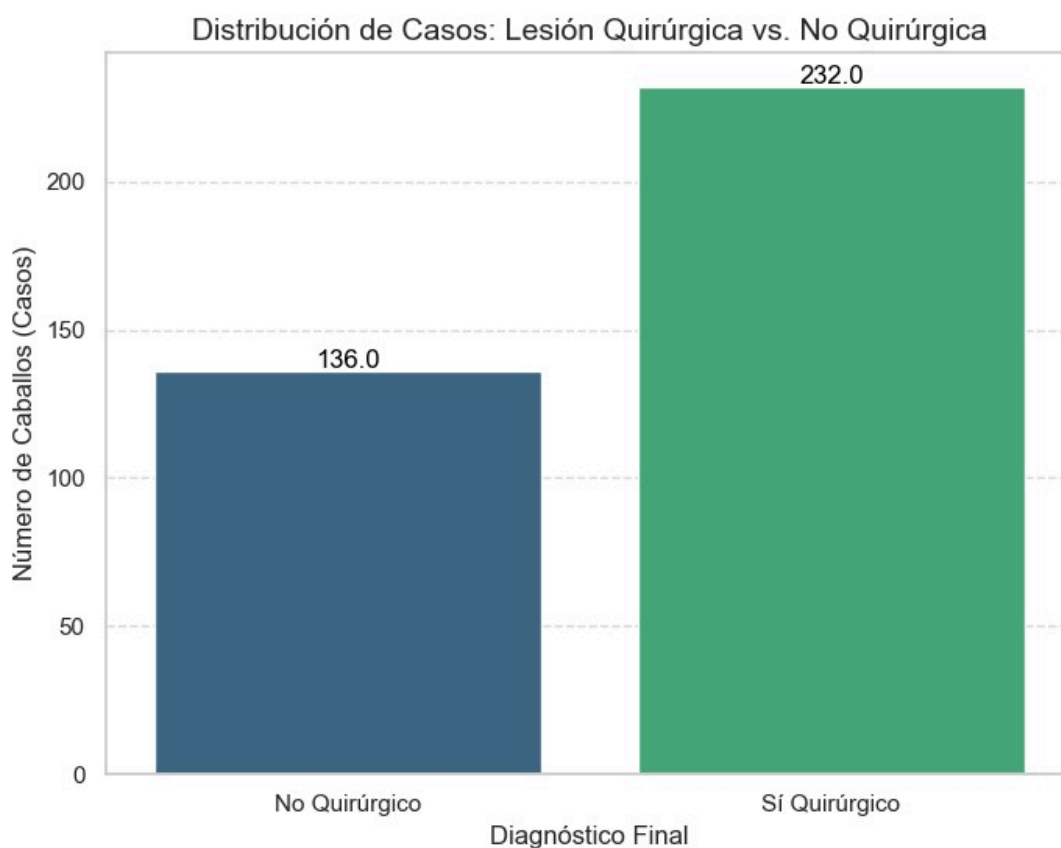


Fig. 3: Distribucion de la variable surgical_lesion.

3. Hipótesis planteadas y resolución

3.1. Hipótesis 1:

"Se postula que la mediana de la frecuencia cardíaca en la población estudiada es superior al límite fisiológico de referencia (44 lpm)."

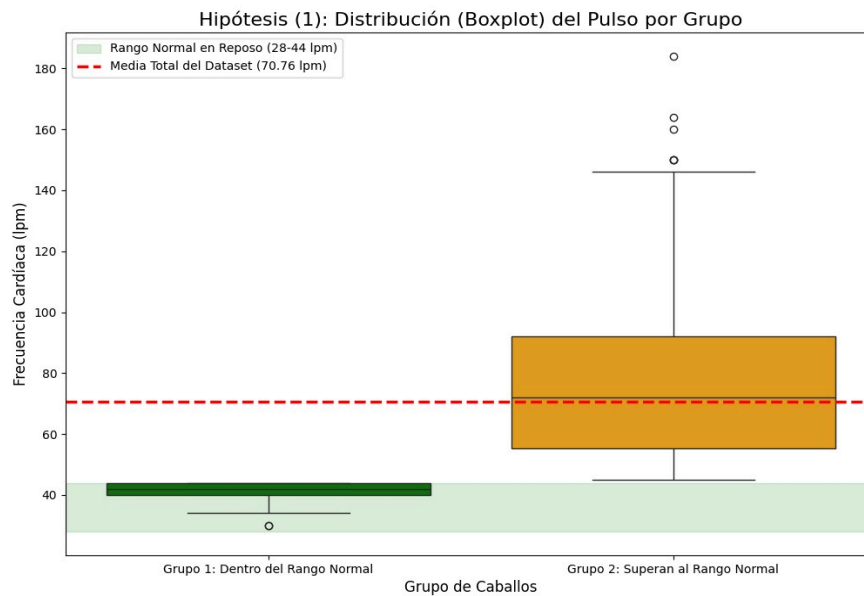


Fig. 4: Distribución (Boxplot) del Pulso por grupo

Esta hipótesis la formulamos en base al análisis de los datos. Observamos que los valores que presentaban las muestras eran notablemente más altos que lo estipulado como "Normal" en la descripción del dataset por lo que decidimos demostrarlo gráficamente con un boxplot.

Para validar esta hipótesis , primero debemos analizar la distribución de los datos:

Normalidad: Grafico QQplot:

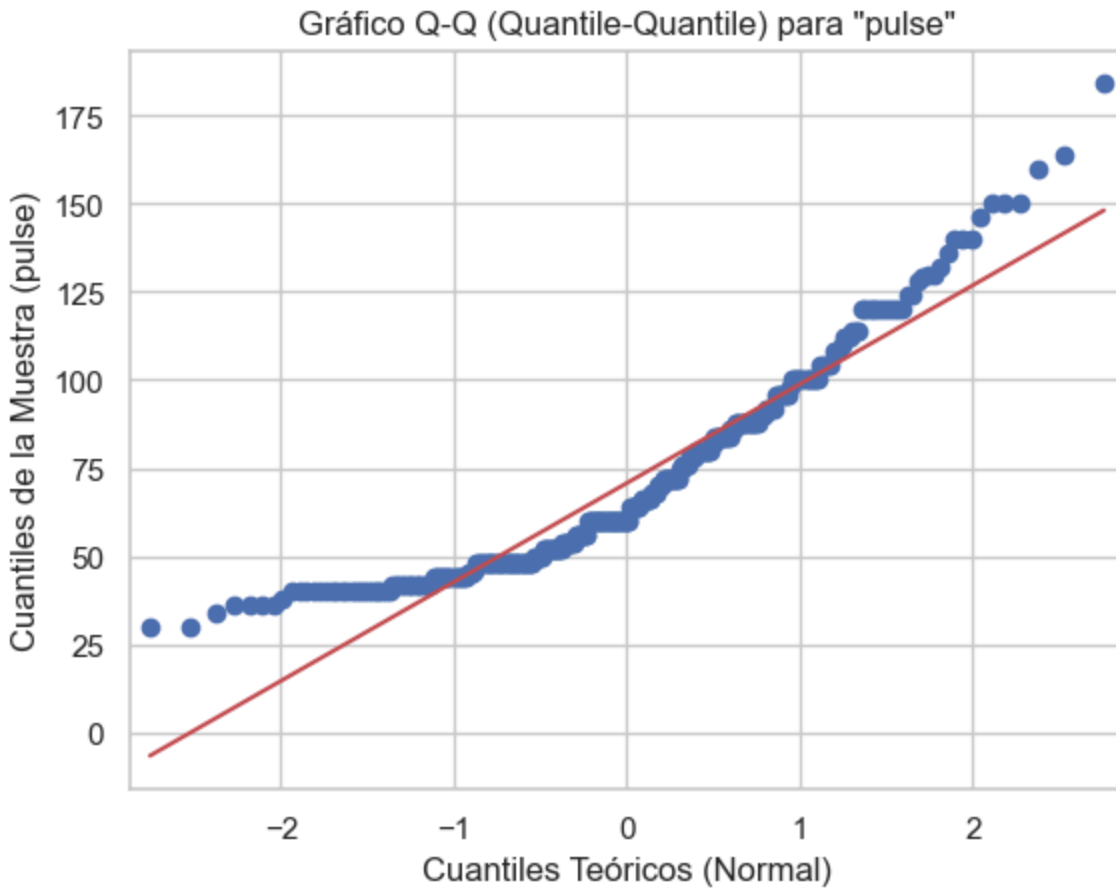


Fig. 5: Gráfico QQ plot para "pulse"

Como se observa en el gráfico, la variable pulse claramente no sigue una distribución normal por lo que no podemos usar un test paramétrico para la validación de la hipótesis por lo que optamos por el test no paramétrico : Wilcoxon.

```
--- Test No Paramétrico: Wilcoxon de Rangos Signados ---  
H0: La MEDIANA de 'pulse' <= 44  
Ha: La MEDIANA de 'pulse' > 44  
  
Mediana real de la muestra: 60.0  
Cantidad de muestras válidas (N): 342  
  
--- Resultados del Test ---  
Estadístico W (W-statistic): 51285.5  
Valor P (p-value): 8.419104712797174e-48  
  
Nivel de significancia (alpha): 0.05  
Resultado: p-value (8.42e-48) < alpha (0.05)  
Conclusión: Se RECHAZA la Hipótesis Nula (H0).  
La mediana del pulso es significativamente MAYOR que 44 lpm.
```

Fig. 6: Resultados del test Wilcoxon.

Como el $p\text{-value} < \alpha$ (0,05) **rechazamos** la hipótesis nula.

El análisis arrojó un resultado claro : La media total del dataset (70,76 lpm) se escapa totalmente del rango normal (Franja verde) estipulado por el mismo. Esto es algo totalmente lógico si nos enfocamos en el estado de shock (dolor, deshidratación,etc..) el que entran los caballos a las clínicas. Además, al demostrar que la población está sesgada hacia la taquicardia, podemos definir al pulso como variable de interés.

A futuro, se podría **categorizar** la variable pulso en nuevos rangos (ej: Taquicardia Leve, Moderada, Severa) y ver si esas categorías se relacionan con el tipo de lesión.

3.2. Hipótesis 2:

“La probabilidad de una `surgical_lesion` (lesión quirúrgica) aumentará a medida que se incremente el nivel de `abdominal_distension` (distensión abdominal)”.

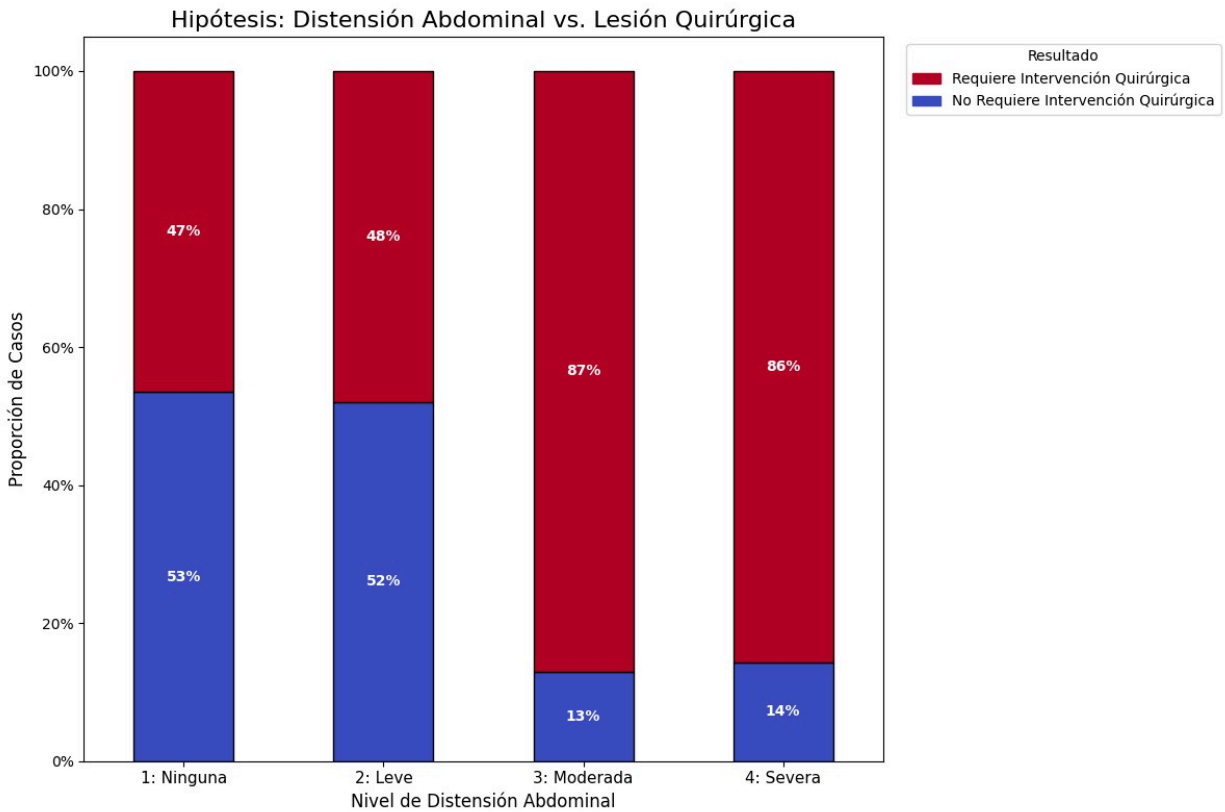


Fig. 7: Gráfico de barras apiladas que muestra la proporción de casos que requieren (rojo) o no requieren (azul) intervención quirúrgica, según el nivel de distensión abdominal.

El gráfico muestra una clara correlación positiva. La probabilidad de requerir una intervención quirúrgica aumenta drásticamente con la gravedad de la distensión, pasando de un 47% en el Nivel 1 (Ninguna) a un 86% en el Nivel 4 (Severa). Esto confirma

que la distensión severa es un signo clínico de alta prioridad, tal como lo describe la documentación del dataset.

Para esta hipótesis, elegimos el test de Chi-cuadrado (no-paramétrico) porque es la herramienta correcta para comparar conteos en diferentes categorías:

Hipótesis Nula (H0): No hay asociación. La abdominal_distension y la surgical_lesion son independientes.

Hipótesis Alternativa (Ha): Sí hay asociación. Las variables son dependientes (qué es lo que queremos probar).

El test arrojó un **Valor P (p-value): 9.15e-11** lo que es muy inferior al nivel de significancia de alfa = 0.05, lo que nos lleva a **rechazar** la hipótesis nula de que no hay asociación. La abdominal_distension y la surgical_lesion son dependientes.

La distensión abdominal es un signo que el veterinario ve "desde afuera" o con un examen físico rápido, sin necesitar equipos de laboratorio complejos. Esto es de gran ayuda ya que permite acortar los tiempos de diagnóstico que podrían marcar la diferencia entre salvar un caballo o no.

En conclusión, el gráfico y el test confirma que la **abdominal_distension** es un predictor importante cuando está presente, pero su ausencia no es, bajo ningún concepto, un indicador fiable para descartar una lesión quirúrgica.

3.3. Hipótesis :

"Sugerimos que la evaluación conjunta de los signos de shock (pulso, temperatura y mucosas) es un predictor de mortalidad más robusto y fiable que el análisis del pulso por sí solo, ya que reduce el ruido de factores aislados como el estrés."

Comparación de Predictores de Resultado Fatal

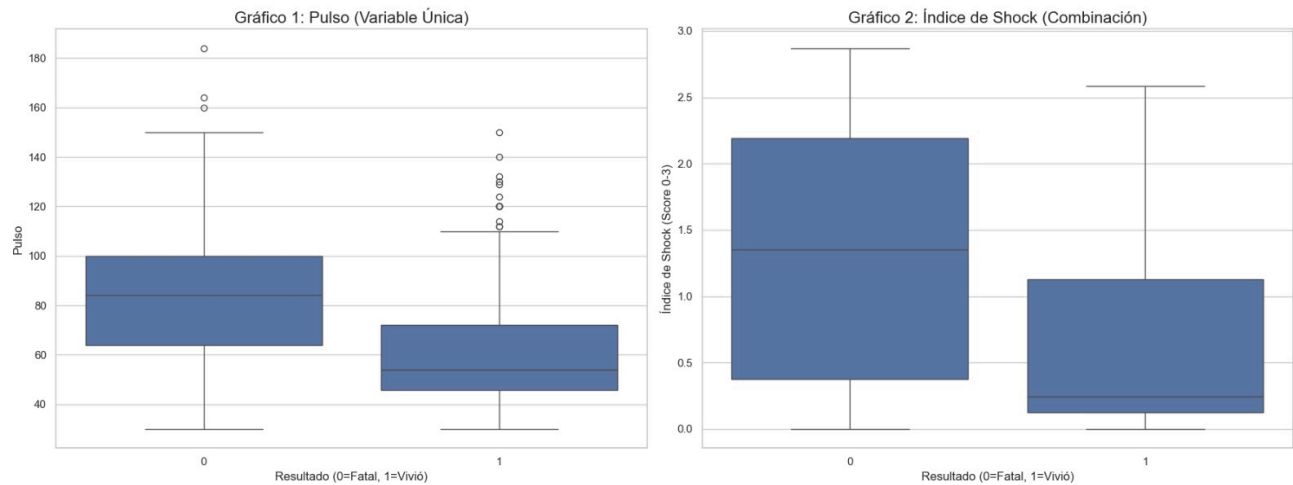


Fig. 8: Comparación de 'Pulso' (izquierda) e 'Índice de Shock' (derecha) como predictores del resultado final (0=Fatal, 1=Vivió).

En este contexto, nos encontramos nuevamente con la presencia de valores faltantes (7% en el pulso, 17,66% en temperature_of_extremities y un 13% en mucous_membranes) y se opta por mantener la metodología de trabajo planteada en las hipótesis anteriores.

Para esta hipótesis nos planteamos **combinar** algunas variables que el dataset indica como generadoras de shock para crear un índice y poder compararlo con variables como el pulso por sí solo y ver si este índice puede ser un predictor de un resultado fatal más fuerte.

Esto lo hacemos para poder eliminar el “ruido” que puede tener la variable por sí sola al combinarla con otras variables generadoras de shock.

Inicialmente nuestra idea era utilizar un test-t para comparar medias de dos grupos independientes pero al verificar los supuestos nos encontramos con que el Test de normalidad de Shapiro-Wilk arrojó lo siguiente:

```
--- 1. Test de Normalidad (Shapiro-Wilk) ---  
Grupo 'Fatal' (shock_index): p-value = 3.97e-07  
-> Conclusión: Los datos del grupo 'Fatal' NO son normales.  
Grupo 'Vivió' (shock_index): p-value = 2.76e-16  
-> Conclusión: Los datos del grupo 'Vivió' NO son normales.
```

Fig. 9: Resultado del test Shapiro-Wilk.

Como podemos observar, ambos p_valores son $< 0,05$ por lo que se rechaza la hipótesis nula de que los datos se distribuyen de manera normal lo que nos limitaría a usar tests no paramétricos.

Como complemento, buscamos testear la Homogeneidad de las varianzas usando Levene, obteniendo los siguientes resultados:

```
--- 2. Test de Homogeneidad de Varianzas (Levene) ---  
Resultado del Test de Levene: p-value = 9.36e-03  
-> Conclusión: Las varianzas NO son homogéneas (son diferentes).
```

Fig. 10: Resultado del test Levene.

Como podemos observar, el **p_valor** es $< 0,05$ por lo que se rechaza la hipótesis nula de la homogeneidad de varianzas.

En su lugar, se seleccionó correctamente la alternativa no paramétrica: el Test de **Mann-Whitney U**. Este test compara las medianas (o más precisamente, las distribuciones) de dos muestras independientes sin asumir una distribución específica. Las hipótesis planteadas fueron las siguientes:

H0: La distribución de 'pulse' es la misma para ambos grupos.

Ha: La distribución es diferente (uno es mayor que el otro).

El test arrojó los siguientes resultados:

```
--- Conclusión Comparativa (Test Mann-Whitney U) ---  
P-value de 'pulse' solo:          2.08e-14  
P-value de 'shock_index':         2.05e-16  
  
Estadístico U1 de 'pulse':        20194.0000  
Estadístico U1 de 'shock_index':   23871.5000  
  
Validación: El 'shock_index' (p-value más bajo y U-stat más alto)  
es un predictor no paramétrico más significativo que el pulso por sí solo.
```

Fig. 11: Resultado del test Mann-Whitney U.

El rechazo de la hipótesis nula nos permite concluir que : Se confirma la hipótesis multivariada con un alto grado de certeza estadística. El Test de Mann-Whitney U, valida que el shock_index (el índice combinado) es un predictor de mortalidad significativamente más potente que el pulse (pulso) por sí solo.

3.4. Hipótesis 4:

“ Los caballos que no sobreviven (mueren o son sacrificados) poseen una media de pulso más alto que los caballos que sí lo hacen.”

Esta hipótesis válida el pulso como **herramienta de decisión**. Si confirmamos que el pulso alto equivale a muerte, le damos al profesional el respaldo numérico para tomar decisiones acertadas en menor tiempo.

Volvemos a hacer énfasis en el pulso ya que el informe es clarísimo en esto: La frecuencia cardíaca es posiblemente el signo vital más importante y un pulso elevado se asocia con un compromiso cardiovascular severo y un pronóstico reservado.

Además , nuestra hipótesis 1 ya validada , nos sirve como complemento para el análisis de esta nueva hipótesis que busca si existe una relación entre el pulso alto y un resultado fatal.

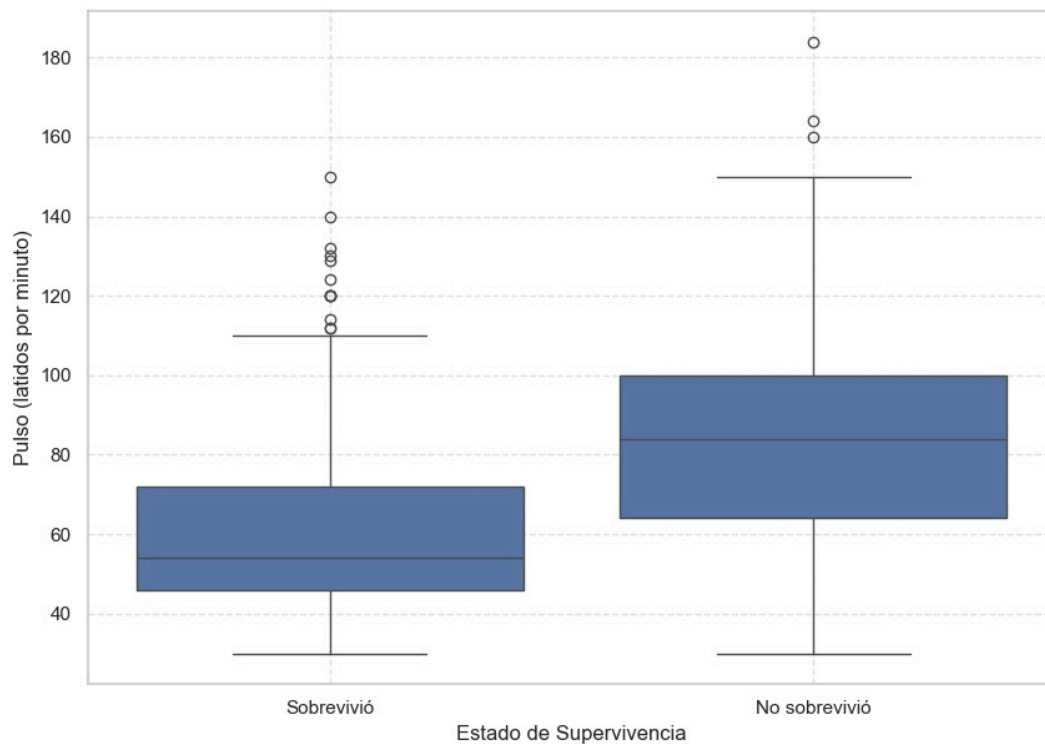


Fig. 12: Boxplot que relaciona el pulso dado el estado de supervivencia.

El boxplot arrojó un resultado claro: Observamos como la media del grupo perteneciente a No sobrevivió (Outcome= 2 o Outcome =3) es claramente superior a la del grupo que Sobrevivió.

Para abordar el testeo de esta hipótesis , primero se buscan probar los supuestos y se obtiene que los datos no son Normales (mediante Shapiro) y las varianzas no son Homogéneas (Levene) por lo que se optó por validar la hipótesis con Mann-Whitney U:

H0 : El pulso de ‘Sobrevivio’ es significativamente mayor a la del grupo ‘No Sobrevivió’.

H1 : El pulso de ‘No sobrevivió’ es significativamente mayor a la del grupo ‘Sobrevivió’.

```
--- Preparación del Test ---  
Número de casos 'Sobrevivió' con datos de pulso: 212  
Número de casos 'No sobrevivió' con datos de pulso: 128  
  
--- Resultado del Test de Mann-Whitney U ---  
Estadístico U: 6942.0  
Valor p (p-value): 2.083956044297287e-14  
Interpretación: ( $p < 0.05$ ) Rechazamos la hipótesis nula.  
Conclusión: ¡Hipótesis validada! Existe una diferencia estadísticamente significativa.  
El pulso del grupo 'No sobrevivió' es significativamente MAYOR que el del grupo 'Sobrevivió'.
```

Fig. 13 : Resultado del test Mann-Whitney U.

Concluimos que : **Existe una fuerte evidencia estadística** que demuestra que el pulso (frecuencia cardíaca) es un indicador crítico del desenlace del caballo. Los caballos que no sobreviven (ya sea por muerte o eutanasia) tienen un ritmo cardíaco significativamente más alto que los que logran sobrevivir.

3.5. Hipótesis 5:

" Los signos vitales iniciales (pulso, temperatura rectal y frecuencia respiratoria) son los factores preponderantes que definen el pronóstico vital . La interacción de estas variables es suficiente para explicar la mayoría de los desenlaces fatales, por encima de otros signos clínicos."

En la atención de emergencias equinas, el tiempo y los recursos suelen ser limitados. Mientras que en un hospital de alta complejidad se pueden medir variables avanzadas (como proteínas totales o pH abdominal), el veterinario de campo a menudo debe tomar decisiones de vida o muerte (cirugía o eutanasia) contando solo con herramientas básicas: un termómetro y un estetoscopio.

Se plantea el uso de Regresión Logística para analizar si es posible predecir una surgical lesion en base a combinar los signos vitales que ya hemos visto en hipótesis anteriores que influyen en los resultados de surgical lesion y outcome. Para esto, primero buscamos respuestas en análisis de algunos gráficos:

Análisis de Predictores vs. Lesión Quirúrgica

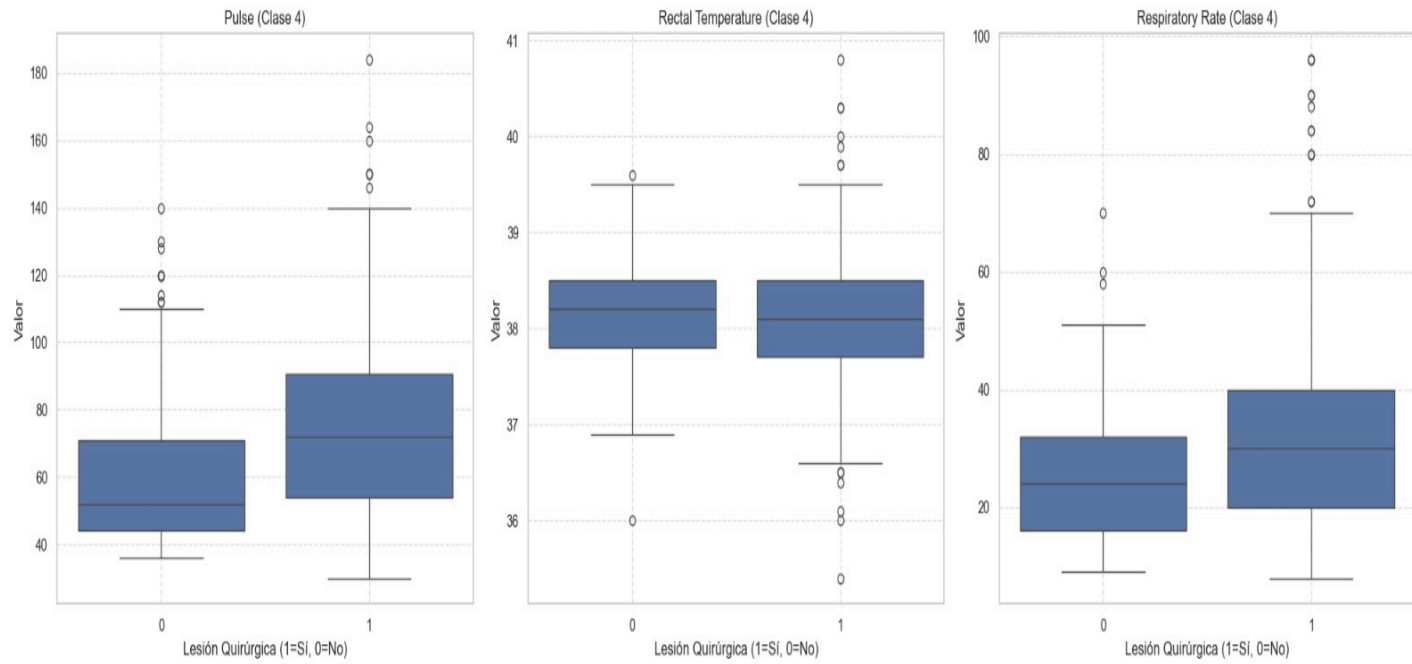


Fig. 14 : Boxplots de variables : Pulso, Rectal_temperature , respiratory_rate ; Eje (x) : Lesion Quirurgica (1= Si ; 2=No)

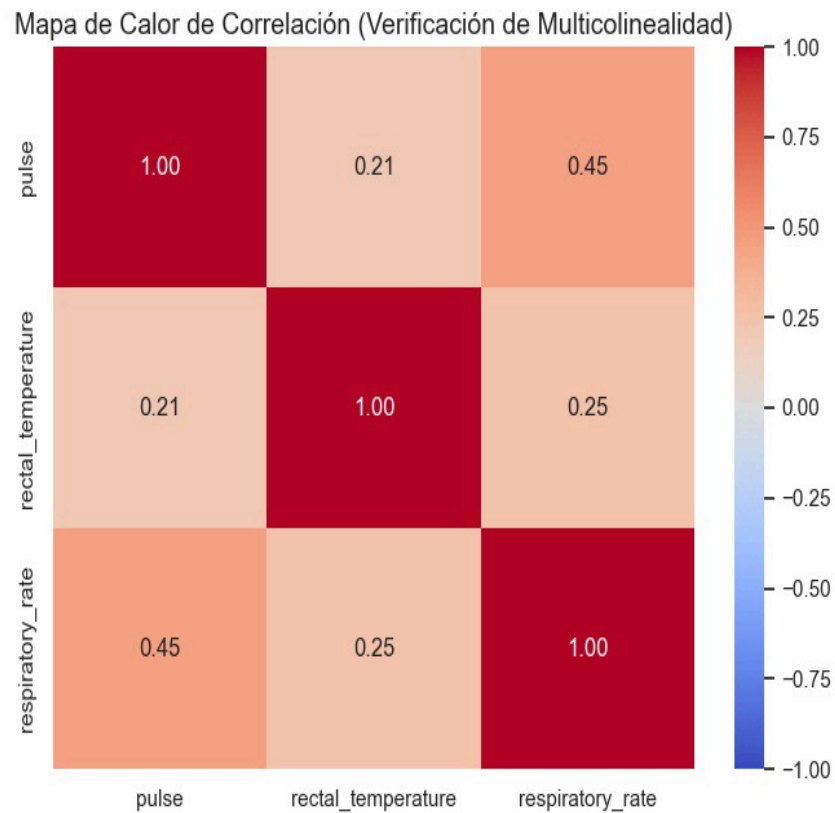


Fig. 15: Matriz de Correlación.

Viendo los Boxplots observamos como el pulso parece ser el predictor individual más fuerte. La temperatura parece ser más débil. Nuestra hipótesis es que, al combinarlos en un modelo de Regresión Logística, el conjunto será más fuerte que las partes.

Por otro lado, el supuesto principal que debemos verificar para una Regresión Logística es la ausencia de multicolinealidad .

Esto significa que nuestras variables predictoras (pulse, rectal_temperature, respiratory_rate) no deben estar fuertemente correlacionadas entre sí. Si lo estuvieran (ej. correlación > 0.7), el modelo no podría saber cuál de ellas es la que realmente importa. Viendo la matriz , observamos como todas las correlaciones entre nuestras

variables predictoras son débiles. El valor más alto es 0.45, que no es lo suficientemente grande como para causar problemas de multicolinealidad.

Esto nos da luz verde para usar estas tres variables juntas en el modelo de Regresión Logística, ya que podemos estar seguros de que cada una aporta información independiente:

```
--- Bloque 3: Test (Modelo de Regresión Logística) ---  
Filas originales en df_completo: 368  
Filas que se usarán para el modelo (sin NaNs): 256  
  
--- Resultados de la Evaluación del Modelo ---  
Matriz de Confusión:  
[[19 13]  
 [11 34]]  
  
Reporte de Clasificación:  
              precision    recall  f1-score   support  
  
0: No Quirúrgico      0.63      0.59      0.61        32  
1: Quirúrgico         0.72      0.76      0.74        45  
  
   accuracy              0.69        77  
  macro avg              0.68        77  
 weighted avg            0.69        77  
  
Accuracy (Exactitud) del modelo: 0.6883  
Error (1 - Accuracy): 0.3117  
Precision (Precisión) para 'Quirúrgico': 0.7234  
Recall (Sensibilidad) para 'Quirúrgico': 0.7556
```

Fig. 16. : Resultados modelo de regresión .

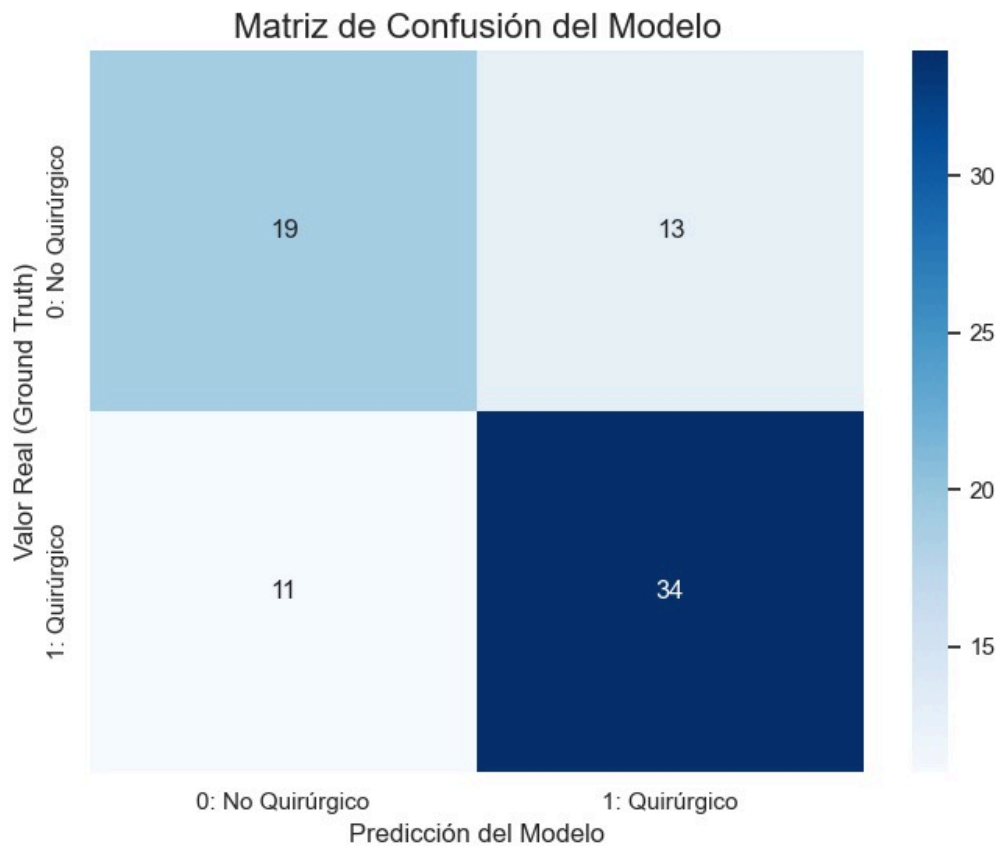


Fig. 17 : Resultados modelo de regresión .

Conclusión :

Para validar la hipótesis multivariada, construimos un modelo de Regresión Logística para predecir si un caso requería cirugía (surgical_lesion) usando la combinación de pulse, rectal_temperature y respiratory_rate. Tras confirmar con una matriz de correlación que no había multicolinealidad problemática entre los predictores , preparamos los datos eliminando filas con valores nulos, lo que resultó en 256 casos completos. Al dividir los datos (70% entrenamiento, 30% prueba) y entrenar el modelo, los resultados validan la hipótesis: el modelo alcanzó una Exactitud (Accuracy) general del 69%, demostrando ser significativamente mejor que una predicción al azar. Más importante aún, el reporte de clasificación mostró un Recall (Sensibilidad) del 76% para

la clase "Quirúrgico", lo que significa que el modelo identificó correctamente a 3 de cada 4 caballos que sí necesitaban la operación. La Matriz de Confusión confirma este poder predictivo, aunque también revela 11 "Falsos Negativos", indicando que, si bien la hipótesis es correcta, un modelo clínicamente fiable requeriría predictores adicionales para lograr una accuracy mucho mayor y ser fiable.

3.6. Hipótesis 6:

“La parálisis intestinal afecta la frecuencia respiratoria”

El caballo tendrá una frecuencia respiratoria (respiratory_rate) significativamente más alta cuando su movimiento intestinal (peristalsis) es anormal o ausente, en comparación con cuando es normal. Entender esto ayuda al veterinario a no confundirse: si ve un caballo respirando rápido, puede que no sea un problema pulmonar, sino una consecuencia directa de que el intestino está "parado" e hinchado.

La conexión no es directa, es mecánica. Los informes clínicos nos dicen que la frecuencia respiratoria aumenta por la distensión abdominal. La distensión se produce cuando el intestino se paraliza (peristalsis ausente) y acumula gas. Estamos probando si el mal funcionamiento digestivo se ve reflejado de forma estadística en el sistema respiratorio.

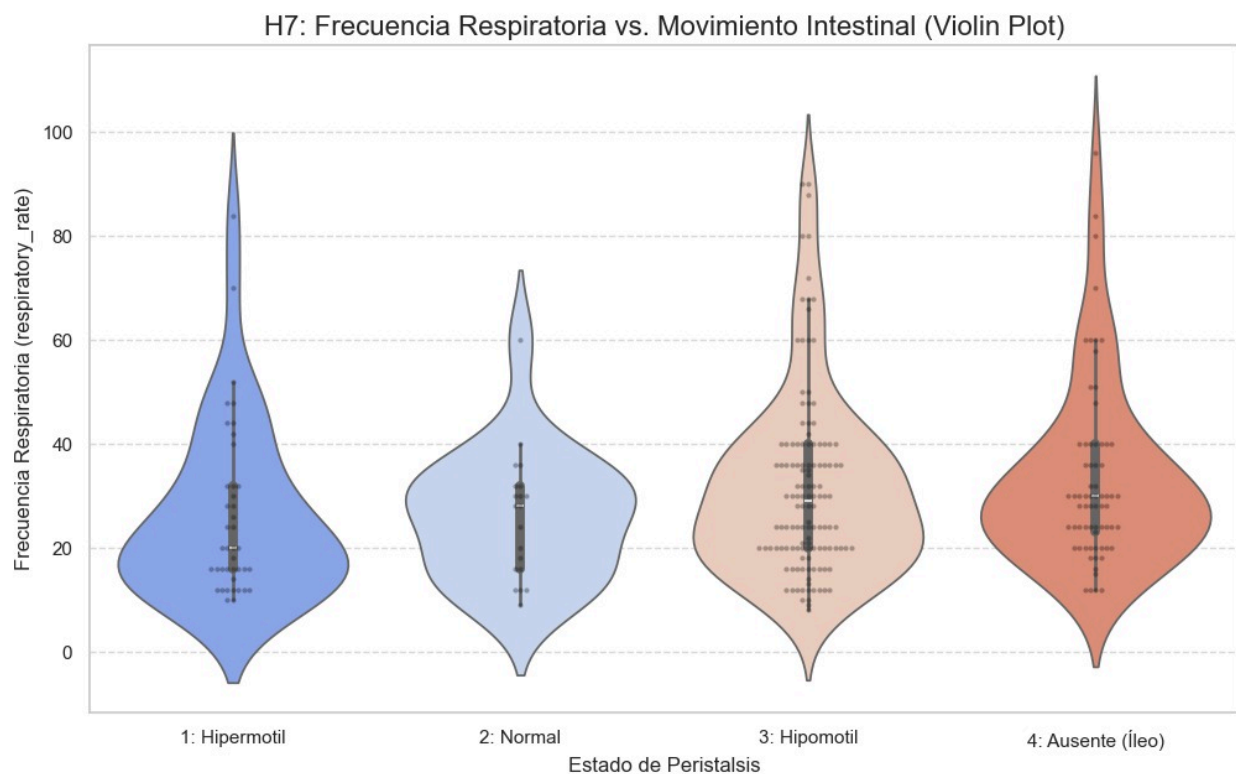


Fig. 18. Gráfico de Violín: Muestra la distribución (KDE) y el boxplot interno.

Planteamos las siguientes hipótesis a testear con Kruskal-Wallis :

H0 (Hipótesis Nula): Las medianas son iguales en todos los grupos.

H1 (Hipótesis Alternativa): Al menos una mediana es diferente.

Debido a que estamos comparando tres o más grupos independientes, nuestra variable (**peristalsis**) es categórica y define 4 grupos y **respiratory_rate** es una variable dependiente continua , y conociendo que los supuestos no se cumplen para un test paramétrico , optamos por Kruskal-Wallis y obtuvimos los siguientes resultados:

```
--- Resultados del Test de Kruskal-Wallis (H7) ---  
Estadístico H: 9.0319  
Valor p (p-value): 0.028870150282033488
```

```
Interpretación (p < 0.05):  
Resultado ESTADÍSTICAMENTE SIGNIFICATIVO.  
Rechazamos la hipótesis nula (H0).  
Conclusión: Existe una diferencia significativa en la frecuencia respiratoria entre al menos dos de los grupos de peristalsis.
```

Fig. 19. : Resultado del test Kruskal-Wallis.

Como el **p valor = 0.028870.. < 0.05** , podemos concluir que existe una diferencia significativa en la frecuencia respiratoria entre al menos dos de los grupos de peristalsis lo que nos lleva a afirmar que la parálisis intestinal está asociada con un aumento real y medible en la frecuencia respiratoria.

4. Conclusiones

Hemos llegado al final de este análisis, logrando transformar un dataset intimidante y complejo en una serie de conclusiones claras y accionables. El objetivo de este trabajo no era solo aplicar tests estadísticos, sino entender por qué los aplicábamos, usando el contexto clínico como nuestra guía principal.

El desafío más grande fue, sin duda, la naturaleza de los datos: un 30% de valores faltantes. Nuestra decisión metodológica de no imputar (rellenar) estos datos se basó en un hallazgo clave: la "ausencia informativa". Entendimos que imputar un valor "normal" a una prueba que no se hizo *precisamente por la gravedad* del caballo habría sido un error metodológico que falsearía los resultados. Esto nos obligó a ser más robustos en nuestro enfoque, usando `dropna()` de forma selectiva para cada hipótesis y eligiendo tests no paramétricos (como Mann-Whitney U y Kruskal-Wallis) que no dependen de una distribución normal que, como demostramos, nuestros datos no tenían.

Nuestros hallazgos validaron fuertemente la intuición veterinaria. Demostramos con una confianza estadística altísima (un p-value de $2.08e-14$) que el pulso es un indicador abrumador de un desenlace fatal, siendo la "alarma" más clara del sistema. Sin embargo, el hallazgo más potente provino de nuestro modelo multivariado: combinando solo tres signos vitales, nuestra Regresión Logística fue capaz de predecir la necesidad de cirugía (`surgical_lesion`) con un Recall (Sensibilidad) del 76%. Esto confirma nuestra hipótesis de que, si bien el pulso es importante, es la *combinación* de factores lo que realmente permite construir un modelo predictivo útil.

Este trabajo no solo cumplió con los objetivos del práctico, sino que demuestra el poder de la ciencia de datos: logramos encontrar patrones estadísticos claros en medio del "caos" de un escenario de emergencia real. Estos resultados, en manos de un experto, son la base para crear herramientas que pueden ayudar a un veterinario a tomar decisiones más rápidas y, en última instancia, salvar vidas.

Referencias

A continuación se adjuntan todos los links a las páginas web utilizadas para la recopilación de información técnica para el entendimiento del contexto:

Dataset

[Horse Colic - UCI Machine Learning Repository](#)

Contexto Clínico del Cólico Equino

1. American College of Veterinary Surgeons. (s.f.). *Colic in Horses*.
<https://www.acvs.org/large-animal/colic-in-horses/> [14]
2. Merck Veterinary Manual. (s.f.). *Overview of Colic in Horses*.
<https://www.msdsvetmanual.com/digestive-system/colic-in-horses/overview-of-colic-in-horses> [8]
3. Merck Veterinary Manual. (s.f.). *Colic in Horses - Horse Owners*.
<https://www.merckvetmanual.com/horse-owners/digestive-disorders-of-horses/colic-in-horses> [7]
4. Royal Veterinary College. (s.f.). *Colic Fact File*.
<https://www.rvc.ac.uk/equine-vet/information-and-advice/fact-files/colic> [1]
5. Cornell eCommons. (s.f.). *Equine colic*.
<https://ecommons.cornell.edu/items/311b57a7-51a3-459e-b6f8-12a451d58710> [2]
6. University of Minnesota Extension. (s.f.). *Colic in your horse*.
<https://extension.umn.edu/horse-health/colic-your-horse> [15]
7. UK Ag Equine Programs. (s.f.). *Colic in Horses: General Review*.
<https://equine.mgcafe.uky.edu/news-story/colic-horses-general-review> [10]

8. Veterinary Extension, University of Florida. (s.f.). *A Horse Owner's Guide to Colic*.
<https://extension.vetmed.ufl.edu/files/2011/10/Horse-Owners-Guide-to-Colic.pdf> [17]
9. Wikipedia. (s.f.). *Horse colic*. https://en.wikipedia.org/wiki/Horse_colic [16]
10. Morris Animal Foundation. (s.f.). *Colic – an Emergency for Horse Owners & a Top Research Focus for the Foundation*.
<https://www.morrisanimalfoundation.org/article/colic-2022-studies> [11]

Proceso de Diagnóstico y Variables

11. DVM360. (s.f.). *Evaluating equine colic: Surgeon's perspective (Proceedings)*.
<https://www.dvm360.com/view/evaluating-equine-colic-surgeons-perspective-proceedings> [9]
12. Vet Times. (s.f.). *Equine colic: management and decision-making in the field*.
<https://www.vettimes.com/clinical/equine/equine-colic-management-and-decision-making-in-the-field> [18]
13. Merck Veterinary Manual. (s.f.). *The Veterinarian's Examination of a Horse with Colic*.
<https://www.merckvetmanual.com/multimedia/table/the-veterinarians-examination-of-a-horse-with-colic> [19]
14. WVS Academy. (s.f.). *Clinical parameters*.
<https://wvs.academy/quick-reference/horse/clinical-parameters/> [24]
15. Texas A&M. (s.f.). *Understanding Vital Life Signs in Horses*.
<https://texashelp.tamu.edu/wp-content/uploads/2016/02/understanding-vital-life-signs-in-horses.pdf> [25]
16. B&W Equine Vets. (s.f.). *Your Horse's Normal Vital Signs*.
<https://www.bwequinevets.co.uk/article/4/horses-normal-vital-signs/> [31]
17. Central Lakes Equine. (s.f.). *Reflux, what is this?*.
<https://www.centrallakesequine.co.nz/new-blog-1/2025/9/21/reflux-what-is-this> [12]
18. B&W Equine Vets. (s.f.). *Nasogastric intubation*.
https://www.bwequinevets.co.uk/uploads/Nasogastric_intubation_2018.pdf [13]
19. Horse Side Vet Guide. (s.f.). *Nasogastric Intubation*.
<https://horsesidevetguide.com/drv/Diagnostic/82/nasogastric-intubation/> [32]

20. Merck Veterinary Manual. (s.f.). *Emergency Procedures in Horses*.
<https://www.merckvetmanual.com/emergency-medicine-and-critical-care/emergency-medicine-in-horses/emergency-procedures-in-horses> [33]
21. EquiManagement. (s.f.). *Abdominal Fluid Interpretation in Colic*.
<https://equimanagement.com/research-medical/abdominal-fluid-interpretation-in-colic/> [27]
22. PubMed. (s.f.). *Interpreting abdominal fluid in colic horses: Understanding and...*
<https://pubmed.ncbi.nlm.nih.gov/35044063/> [37]
23. AVMA Journals. (s.f.). *How to perform abdominocentesis and interpret abdominal fluid in equine patients*.
<https://avmajournals.avma.org/view/journals/javma/263/8/javma.25.02.0123.xml> [26]
24. University of Minnesota. (s.f.). *Abdominocentesis and Peritoneal Fluid Analysis*.
https://open.lib.umn.edu/app/uploads/sites/208/2019/04/Practical_Guide_To_Equine_Colic_Abdominocentesis.pdf [35]
25. Ivis.org. (s.f.). *Abdominocentesis (Peritoneal Tap)*.
<https://www.ivis.org/library/manual-of-equine-diagnostic-procedures/abdominocentesis-peritoneal-tap> [36]
26. ResearchGate. (s.f.). *Morbidity associated with abdominocentesis — a prospective study*.
https://www.researchgate.net/publication/230088552_Morbidity_associated_with_abdominocentesis-_a_prospective_study [40]
27. MAG Online Library. (s.f.). *Abdominal paracentesis in the field setting for horses with acute abdominal pain*.
<https://www.magonlinelibrary.com/doi/pdf/10.12968/ukve.2024.8.4.184?download=true> [41]
28. Vet Times. (s.f.). *Abdominocentesis in horses*.
<https://www.vettimes.com/clinical/equine/abdominocentesis-in-horses-cpdequine> [42]
29. PubMed Central. (s.f.). *Early recognition of pain: improving colic outcomes in horses in Senegal*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC11410767/> [28]
30. AWS. (s.f.). *Managing Pain Associated with Colic*.
http://vetfolio-vetstreet.s3.amazonaws.com/mmah/e7/ad1a65f5d44b28b70b77b5d250426e/filePVE_03_06_294.pdf [30]

31. EquiManagement. (s.f.). *Equine Pain Assessment Tools and Techniques*.
<https://equimanagement.com/research-medical/sports-medicine/equine-pain-assessment-tools-and-techniques/> [43]
32. PubMed. (s.f.). *Equine colic: A comprehensive overview of the sonographic evaluation....*
<https://pubmed.ncbi.nlm.nih.gov/40276205/> [3]