Ejercicio 1

Determine si las siguientes afirmaciones sobre los valores atípicos son verdaderas o falsas:

- El análisis de boxplot permite detectar outliers multivariados.
- Isolation Forest es un método basado en densidad que permite detectar tanto outliers univariados como multivariados.
- La única forma de tratar valores atípicos es eliminar los valores más altos y más bajos de una variable. _____
- Si se calcula el z-score para un conjunto de datos y se quieren eliminar los outliers deben eliminarse los valores dentro del rango -3 y 3._____

Ejercicio 2

Determinar si las siguientes afirmaciones son verdaderas o falsas:

- El coeficiente de correlación de pearson mide cuán relacionadas están dos variables sea o no de forma lineal.
- El método de regresión logística se utiliza para resolver problemas de clasificación binaria.
- La regresión lineal simple consiste en predecir un valor numérico en rango continuo en base a más de una variable predictora. _____

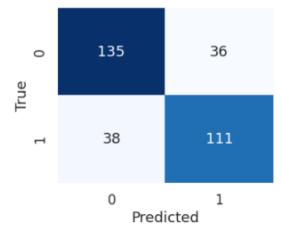
Ejercicio 3

Determinar cuáles de las siguientes opciones son características que comparten los métodos de Bagging y Boosting. Indicar verdadero o falso:

- Se basan en entrenar varios modelos, y luego tomar una única decisión.
- Cada modelo se entrena en paralelo.
- Cada modelo es influenciado por la performance de los otros.
- Utilizan datasets de entrenamiento generados mediante Boostrap.

Ejercicio 4

Se entrenó un modelo de clasificación (MODELO A) para detectar personas enfermas (clase 1) y sanas (clase 0). Luego se evaluó el modelo en los datos de test y se obtuvo la siguiente matriz de confusión:



Luego se entrenó un segundo modelo de clasificación (MODELO B), se evaluó sobre los mismo	<u>)S</u>
<u>datos de test</u> y se obtuvieron los siguientes resultados:	

True Positive: 49 True Negative: 130 False Positive: 41 False Negative: 100

¿Qué modelo (A o B) elegiría en términos de Accuracy?: _	
Accuracy Modelo A	
Accuracy Modelo B	