

Laboratorio # 2 de analítica de datos

1. Utilizando la base de datos de MPG (Millas por galón) obtener un modelo de regresión con un r^2 superior al 88% utilizando una entrada o varias siguiendo los siguientes requerimientos.
 - a. El entrenamiento y prueba tiene que estar dividido 60% y 40% respectivamente.
 - b. Tienen que realizar un análisis de las variables para conocer si existen datos atípicos o no. En caso de confirmar su existencia deben procesarlos.
 - c. Deben identificar el lugar exacto de estos atípicos y guardar sus índices en una variable externa.
 - d. Cuantificar la cantidad de registros que existen por vehículo.
2. Seleccionando 5 vehículos a elección obtener una regresión por cada uno de ellos, buscando el mejor rendimiento posible. Además:
 - a. Comparar visualmente las respuestas de todas las regresiones usando un gráfico de línea o dispersión.