

FUNDACIÓN UNIVERSITARIA LOS LIBERTADORES
FACULTAD DE INGENIERÍA Y CIENCIAS BÁSICAS
ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA [METODOLOGÍA VIRTUAL]
TALLER 2 Análisis de Regresión
Modelo de Regresión Lineal Múltiple

PROFESOR: SÉBASTIEN LOZANO FORERO
slozanof@libertadores.edu.co

INSTRUCCIONES

- Resuelva el taller en R (o en Jupyter Notebooks). El entregable deberá ser el archivo pdf con código .R (o el archivo .ipynb)
- Interprete las salidas oportunamente. Asuma poco y no deje detalles sin resolver.
- Plazo Máximo de entrega: Domingo 28 de Noviembre de 2021 (23h59).

Considere el la base de datos `housing.csv`, adjunta. La misma es descrita en detalle en **Boston house prices dataset** del repositorio que se encuentra [aquí](#). A partir de la misma resuelva lo siguiente:

1. Realice una corta presentación del conjunto de datos asegurándose de entender cada variable.
2. Presente los principales resultados de un Análisis Descriptivo y Exploratorio de los datos. Haga énfasis en la relación que pueda existir entre las variables.
3. Proponga el mejor modelo de regresión lineal múltiple para la última columna de la base de datos (valor medio comercial de casas de uso residencial en miles de dólares).
4. Estime el modelo propuesto en el ítem anterior. Verifique sea estadísticamente significativo. Reporte los valores R^2 , AIC , BIC de forma explícita.
5. Verifique los supuestos del modelo. Implemente análisis de Residuos (figuras), calcule y dibuje los residuos estudentizados.
6. Reporte los p-valores de las pruebas de hipótesis que fueron usadas, además del planteamiento de tales pruebas.
7. Existen datos atípicos, influyentes o que llamen su atención?
8. Interprete su modelo en términos de la variable que incluye.

Investigación A partir del modelo, suponga algún valor para la variable respuesta y reporte un Intervalo de Confianza del 95% para el pronóstico de la variable respuesta.