**Maestría en Ciencia y Análisis de Datos- Universidad Mayor de San Andrés**

**Modelos lineales y modelos lineales generalizados**

NOMBRE:

INFORMACION DE CONTACTO:

Correo electrónico:

Número de teléfono celular (opcional):

**PRACTICA 2**

**(cada pregunta tiene un valor de 3 puntos,**

**Límite de entrega: Domingo 28 de Julio a medianoche.**

**Entrega en la plataforma Moodle. Trabajar con un nivel de significancia de 5%)**

1. Utilizando los datos de salarios\_posgrado.csv (link GitHub debajo), estime los valores puntuales de los parámetros , , , *,*  en la ecuación:

<https://raw.githubusercontent.com/rogon666/UMSA/main/MLMLG/datos/salarios_posgrado.csv>

Donde “salarios” son los salarios de las i-personas en unidades monetarias (dólares), “educación” son los anos de educación, “edad” es la edad de la persona en años, “antigüedad” es la antigüedad de trabajo en la empresa (en años) y vve es el ratio de las ventas de las empresa dividido entre su valor de cotización en bolsa.

1. Evalué el supuesto de linealidad del modelo con la prueba de linealidad Harvey-Collier (1.5 puntos) y evalúe la normalidad de los residuos de la regresión anterior con la prueba Jarque-Bera (1.5 puntos). Interprete los resultados.
2. Evalúe la presencia de heteroscedasticidad con la prueba de homocedasticidad Breusch-Pagan (1.5 puntos) y evalúe la presencia de autocorrelación de órdenes 1, 2, y 3 con la prueba de correlación serial Breusch-Godfrey (1.5 puntos). Interprete los resultados.
3. Evalúe la presencia de multicolinealidad con los factores de inflación de varianza (VIF). Interprete los resultados. Reduzca el modelo si encuentra evidencia de multicolinealidad **severa**.
4. Obtenga resultados del **modelo final reducido** (si es que decide reducir el modelo) con estimaciones de errores estándar obtenidas con una matriz sándwich robusta a heteroscedasticidad (si es que solamente existe heteroscedasticidad y no existe evidencia de autocorrelación) o una matriz robusta a heteroscedasticidad y autocorrelación, si existencia evidencia de heteroscedasticidad y autocorrelación (1.5 puntos). Interprete los resultados: interprete el ajuste del modelo, los resultados de los coeficientes en base a la estimación puntual, y la significancia estadística (prueba de hipótesis) basada en los valores p (p-values) asociados al estadígrafo t de los estimadores (1.5 puntos).
5. Copie debajo los códigos R que utilizó para obtener las estimaciones.