

# Laboratorio 3

## Instrucciones

El laboratorio 3 deberá ser contestado en equipos de 2 o 3 personas (max.) y el archivo con las respuestas lo debes enviar por email a mi correo electrónico ([sergio.bejar@cide.edu](mailto:sergio.bejar@cide.edu)) a más tardar el viernes 26 de abril, 2023 a las 12:00pm.

La base de datos que vamos a utilizar se encuentra en el folder “Files” del repositorio Github de la clase y se llama (“river.csv”)

## Ejercicio 1

Para este ejercicio vamos a suponer que queremos saber si existe relación (causal) entre el ingreso de un individuo y su nivel educativo. Específicamente, nuestra hipótesis es que a mayor número de años de educación, mayor será el ingreso anual en promedio.

1. Elabora un gráfico de dispersión (scatterplot) de nuestras variables de interés (ingreso (income) y educación (educ)). (5 pts.)
2. ¿Qué clase de relación existe entre ingreso y educación? (2 pts.)
3. Estima una regresión lineal en donde la variable dependiente sea ingreso anual (income) y la variable independiente sea educación (educ). (5 pts.)
4. Interpreta el coeficiente del intercepto. (2 pts.)
5. Interpreta el efecto de educación en la variable dependiente (ingreso). (5 pts.)
6. ¿Cuál es el valor esperado de  $y$  ( $\hat{y}$ ) cuando educación es igual a 20? (1 pt.)

## Ejercicio 2

Ahora asumamos que estamos interesados en entender la relación que existe entre la variable categorica ‘género’ (gender) y el ‘ingreso anual’ (income).

1. Recodifica la variable categórica “gender” de tal forma que “mujer” tome el valor de 0 y “hombre” sea igual a 1. (2 pts.)

2. Grafica la relación que existe entre “género” e ingreso. (5 pts.)
3. ¿Qué podemos inferir del gráfico que obtuviste en el inciso anterior? (2 pts.)
4. Estima una regresión lineal en donde la variable dependiente sea “ingreso anual” y la variable independiente sea “gender”. (5 pts.)
5. Interpreta los resultados de la regresión. (6 pts.)