

# Econometría I

## Reto 01

Carlos A. Yanes Guerra

Universidad del Norte | Departamento de Economía

### Contenido

Antes de empezar . . . . .	1
Brief de investigación . . . . .	1
Objetivo . . . . .	1
Base simulada . . . . .	2
<b>Preguntas del reto</b>	<b>2</b>
<b>Criterios de evaluación (rúbrica breve)</b>	<b>3</b>

### Antes de empezar

Recuerde que los retos se desarrollan en clase en 1 h 50 min. Puede trabajar solo o en pareja (máx. 2 personas). No se permite el uso de paquetes, comandos, funciones o código distintos a los utilizados en la guía del curso. Puede consultarla o utilizarla como acompañamiento de su grupo.

Recuerde que, una vez cumplido el tiempo, el parcial comienza a calificarse con penalización: por cada 20 minutos de retraso, la nota se reduce en 10%.

Por ejemplo: si entrega a las 10:01, la calificación se realiza sobre 90%; entre 10:20 y 10:30, sobre 80%; entre 10:31 y 10:40, sobre 70%; y si entrega después de las 10:41, sobre 60%.

Los trabajos no entregados tienen calificación 0.0. No hay reposición.

### Brief de investigación

Usted hace parte de un equipo de análisis económico que estudia la relación entre ingresos, educación y empleo en un conjunto de hogares urbanos. Para este ejercicio, usted trabajará con una base de datos simulada que representa información individual de personas en edad de trabajar.

### Objetivo

Realizar un análisis completo de la base de datos usando la **guía base del curso**.

- id: identificador del individuo
- edad: edad en años
- educacion: años de educación formal
- ingreso: ingreso mensual (en pesos)
- empleado: condición laboral (1 = empleado, 0 = no empleado)
- sexo: categoría nominal (Hombre / Mujer)

## Base simulada

Corra lo siguiente en su colab o archivo .rmd (R Markdown)

```
# Fijar semilla para reproducibilidad
set.seed(2026)

# Número de individuos
n <- 30
id <- 1:n
edad <- sample(22:60, n, replace = TRUE)
educacion <- sample(5:20, n, replace = TRUE)
ingreso <- round(rnorm(
  n = n,
  mean = 1500000,
  sd = 400000
), -3)

# es hora de introducir algunos valores que no se recopilaron

ingreso[sample(1:n, 4)] <- NA
empleado <- sample(c(0, 1), n, replace = TRUE, prob = c(0.35, 0.65))
sexo <- sample(c("Hombre", "Mujer"), n, replace = TRUE)

hogares <- data.frame(
  id = id,
  edad = edad,
  educacion = educacion,
  ingreso = ingreso,
  empleado = empleado,
  sexo = sexo,
  stringsAsFactors = FALSE
)
```

## Preguntas del reto

1.) Recree la base de datos en R utilizando los vectores proporcionados por el docente y conviértala en un data frame.

Luego:

- Identifique cuántas filas y columnas tiene la base.
- Clasifique cada variable según su tipo de dato (binaria, continua, nominal, índice, etc.).
- Explique por qué esta base corresponde a un corte transversal y no a una serie de tiempo o panel.

2.) Utilizando el data frame ejecutado en código:

- Seleccione únicamente la variable ingreso usando el método apropiado en R.
- Calcule el ingreso promedio ignorando valores faltantes.
- Identifique cuántos individuos no reportan ingreso y explique por qué es importante detectarlos antes de un análisis econométrico.

3.) ree las siguientes variables nuevas en la base de datos:

- ingreso\_anual: ingreso mensual multiplicado por 12.
- ingreso\_educacion: ingreso mensual dividido entre años de educación.
- grupo\_edad: variable categórica que clasifique a los individuos en:
  - Jóvenes ( $\leq 30$  años)

- Adultos (31–50 años)
- Mayores (> 50 años)

Explique brevemente qué interpretación económica puede tener cada nueva variable.

4.) Ahora trabaje únicamente con los individuos que están **empleados**.

- Cree un subconjunto del data frame que cumpla esta condición.
- Calcule el ingreso promedio por sexo dentro de este grupo.
- Comente si observa alguna diferencia relevante y qué limitaciones tiene este resultado.

5.) Con base en las variables disponibles y las creadas:

- Identifique una variable dependiente razonable para un modelo econométrico.
- Proponga al menos dos variables que podrían explicar esa variable dependiente y justifique su elección.

---

## Criterios de evaluación (rúbrica breve)

- **Correctitud numérica (10%):** Son
- **Evidencia (25%):** resultados claros
- **Interpretación econométrica (25%)**
- **Presentación consultiva (40%):** Trabajo con títulos, unidades, etc.