

Laboratorio 1: Empezar con R y RStudio

Trabajar con datos en R

Curso: IF108

Semestre: Agosto–Diciembre 2025

Profesor: Dr. Marco A. González Tagle

Alumno(a): _____

Fecha de entrega: 01/09/2025

Respuestas completas (Parte I y Parte II)

Parte I – R y RStudio

1) Primer contacto con la consola de R

A continuación se muestran las asignaciones de variables para los gastos mensuales de la estudiante, el cálculo del total mensual, así como los totales para un semestre (5 meses) y para el año escolar (10 meses).

```
# Suma directa (gastos de la estudiante)
300 + 240 + 1527 + 400 + 1500 + 1833
# Salida esperada: 5800

# Asignaciones (objetos)
celular      <- 300
transporte   <- 240
comestibles  <- 1527
gimnasio     <- 400
alquiler     <- 1500
otros        <- 1833

# Total mensual como objeto
total <- celular + transporte + comestibles + gimnasio + alquiler + otros
total # 5800

# Gasto en 1 semestre (5 meses)
total_semestre <- total * 5
total_semestre # 29000

# Gasto en 1 año escolar (10 meses)
total_anual <- total * 10
total_anual # 58000
```

2) Nombres de objetos válidos e inválidos

Los nombres de objetos no deben iniciar con un dígito o guion bajo, ni contener comas o espacios. Ejemplos válidos e inválidos:

```
# Válidos
yo_uso_guion_bajo <- 1
otras.personas.usan.punto <- 2
OtrosUsanMayusculas <- 3
this_is_a_really_long_name <- 3.5

# Inválidos (producen error si se intentan ejecutar)
# 5variable <- 5      # inicia con dígito
# _invalid <- 10      # inicia con guion bajo
# my,variable <- 3     # contiene coma
# my variable <- 1     # contiene espacio
```

3) Funciones básicas

```
abs(10) # 10
abs(-4) # 4
sqrt(9) # 3
log(2)  # ~0.6931
```

4) Comentarios en R

```
# R ignora todo lo que esté a la derecha de '#'
2 * 9 # 18
4 + 5 # 9
```

5) Sensibilidad a mayúsculas/minúsculas

```
celular <- 300
Celular <- -300
CELULAR <- 8000

celular + Celular    # 0
CELULAR - celular    # 7700
```

6) Obtener ayuda

```
# Si conoces el nombre de la función
help(abs)
?mean

# Búsqueda por palabra clave
help.search("absolute")
??absolute
```

7) Autoevaluación: vectores y gráficos

Se construye un vector con los gastos, se grafica el diagrama de barras, y luego se ordena el vector en forma decreciente para graficarlo nuevamente. Se añaden etiquetas a las barras.

```
# Vector con los gastos
gastos <- c(celular, transporte, comestibles, gimnasio, alquiler, otros)

# Gráfico de barras (panel de gráficos en RStudio)
barplot(gastos, main = "Gastos mensuales")

# Ordenar y graficar en orden decreciente
gastos_orden <- sort(gastos, decreasing = TRUE)
barplot(gastos_orden, main = "Gastos (orden decreciente)")

# Colocar nombres a las barras
nombres <- c("Celular", "Transporte", "Comestibles", "Gimnasio", "Alquiler", "Otros")
barplot(gastos, names.arg = nombres, main = "Gastos con etiquetas")
```

Parte II – Variables

Problema 1: Tipo de variable en encuesta

- 1 Nombre de estudiante → Cualitativa nominal.
- 2 Fecha de nacimiento → Cuantitativa (tiempo/fecha; se maneja como tipo Date).
- 3 Edad (años) → Cuantitativa discreta.
- 4 Dirección de casa → Cualitativa nominal.
- 5 Número de teléfono → Cualitativa nominal (identificador; no aritmética).
- 6 Área principal de estudio → Cualitativa nominal.
- 7 Grado universitario (primero, segundo, tercero, último) → Cualitativa ordinal.
- 8 Puntaje en mitad de período (0–100) → Cuantitativa continua.
- 9 Calificación general (A, B, C, D, F) → Cualitativa ordinal.
- 10 Tiempo para finalizar examen (min) → Cuantitativa continua.
- 11 Número de hermanos → Cuantitativa discreta.

Problema 2: 14 variables (objeto = Países)

Cuantitativas (7):

- Población total.
- PIB anual.
- Superficie (km²).
- Densidad poblacional.
- Esperanza de vida.
- Número de aeropuertos.
- Temperatura media anual.

Cualitativas (7):

- Continente.
- Idioma oficial.
- Sistema de gobierno.
- Colores/símbolos de la bandera.
- Religión predominante.
- Grupo étnico principal.
- Membresía en la ONU.

Problema 3: Variable numérica como etiquetas

Para los valores 1 = Twitter, 2 = correo electrónico, 3 = mensaje de texto, 4 = Facebook, 5 = blog, la variable es cualitativa nominal. Los números son códigos de categorías; no tienen significado aritmético ni un orden inherente.

Problema 4: Identificar individuos, variables y tipo

¿Cantidad promedio de horas que los estudiantes de universidades públicas trabajan cada semana?

- Individuos: estudiantes de universidades públicas.
- Variable: horas trabajadas por semana.
- Tipo: cuantitativa continua.

¿Qué proporción de todos los estudiantes universitarios de México están inscritos en una universidad pública?

- Individuos: estudiantes universitarios de México.
- Variable: tipo de universidad (pública/privada).
- Tipo: cualitativa nominal (la proporción se calcula sobre esta categoría).

En las universidades públicas, ¿las estudiantes femeninas tienen un promedio de CENEVAL más alto que los estudiantes varones?

- Individuos: estudiantes en universidades públicas.
- Variables: sexo (cualitativa nominal), puntaje CENEVAL (cuantitativa continua).
- Comparación de medias por grupo.

¿Es más probable que los atletas universitarios reciban asesoramiento académico que los atletas no universitarios?

- Individuos: estudiantes universitarios.
- Variables: condición de atleta (sí/no, cualitativa nominal), asesoría académica (sí/no, cualitativa nominal).
- Comparación de proporciones.

Si reuniéramos datos para responder a las preguntas anteriores, ¿qué datos podrían analizarse con un histograma? ¿Cómo lo sabes?

- Variables cuantitativas (por ejemplo: horas trabajadas, puntaje CENEVAL, tiempo en minutos).
- Porque el histograma representa distribuciones numéricas (continuas o discretas).