

# Práctica #2 de la unidad 1

Juan Pablo Salinas Muñoz

2025-08-26

**Objetivo:** El alumno obtendrá medidas y gráficos resumen mediante el uso de software estadístico.

**Software a emplear:** RStudio

**Instrucciones:** El reporte con las respuestas a los siguientes ejercicios debe ser elaborado en RMarkdown y entregado en la plataforma Aula Virtual dentro del apartado correspondiente con formato PDF.

**Ejercicio 1** Realiza lo indicado en cada uno de los siguientes incisos y contesta correctamente.

**(a) [Valor 20%]** Ejecuta el siguiente fragmento de código (code chunk) para abrir el archivo de datos referente a los resultados del concurso de programación del ICPC 2024 en México. Responde, ¿Cuántas instancias (renglones, equipos) tiene el conjunto de datos? ¿Cuántas variables (columnas) tiene el conjunto de datos?

```
library(readxl)
datos <- read_excel(file.choose())
dim(datos)

## [1] 482 17
```

\*\*Respuesta del inciso (a):482, 17

**(b) [Valor 20%]** Ejecuta el siguiente fragmento de código (code chunk) y responde, ¿Cuántos estados de la república participaron en el concurso? ¿Cuál estado de la república mexicana tuvo mayor presencia de equipos en el concurso?

```
with(datos, table(ESTADO))

## ESTADO
## Aguascalientes Baja California Chiapas Chihuahua
Ciudad México
## 16 2 2 15
69
## Coahuila Durango Estado México Guanajuato
Hidalgo
## 28 19 25 31
13
## Jalisco Michoacán Morelos Nuevo León
Oaxaca
## 77 5 2 60
9
```

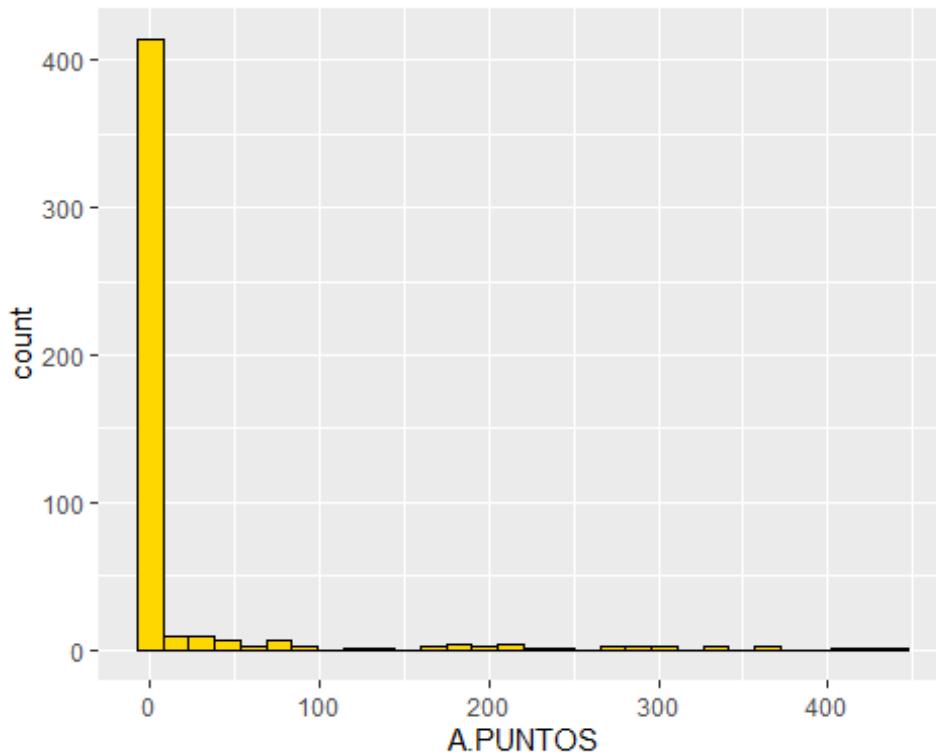
	Puebla	Querétaro	Quintana Roo	San Luis Potosí
Sinaloa	24	6	10	11
16				
Tamaulipas	5	1	16	Zacatecas
##				20

\*\*Respuesta del inciso (b):24, Jalisco

(c) [Valor 20%] Ejecuta el siguiente fragmento de código (code chunk) e interpreta el gráfico.

```
library(ggplot2)
p <- ggplot(data=datos, aes(x=A.PUNTOS)) +
  geom_histogram(color="black", fill="gold")
plot(p)

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



\*\*Respuesta

del inciso (c):La mayoría de equipos obtuvo una puntuación mínima

(d) [Valor 20%] Calcula y reporta la media y la desviación estándar de la cantidad de problemas resueltos en la primera fecha. Has lo mismo para la segunda y la tercera fecha. Compara los resultados entre las fechas e interpreta.

```
# media
mean(datos$F1.PROBLEMAS)
```

```
## [1] 1.695021

# desviación estandar
sd(datos$F1.PROBLEMAS)

## [1] 1.245506

# media
mean(datos$F2.PROBLEMAS)

## [1] 1.46888

# desviación estandar
sd(datos$F2.PROBLEMAS)

## [1] 1.667865

# media
mean(datos$F3.PROBLEMAS)

## [1] 2.278008

# desviación estandar
sd(datos$F3.PROBLEMAS)

## [1] 1.927343
```

\*\*Respuesta del inciso (d):*Hubieron más problemas resueltos en la 3ra fecha, pero también tuvo mayor dispersión.*

**Ejercicio 2. [Valor 20%]** Comenta los aprendizajes obtenidos al realizar la presente práctica, así como, las dificultades o retos enfrentados.

\*\*Respuesta del ejercicio 2:*Manejar los datos de un excel con R Markdown*