


COLEGIO BRASILIA BOSA
INSTITUCIÓN EDUCATIVA DISTRITAL

	AREA ACADEMICO	VERSION 01
	"FORMACIÓN INTEGRAL HACIA LA EXCELENCIA HUMANA Y LABORAL"	
	GUIA APOYO ESCOLAR	CBB-GA01

IDENTIFICACIÓN

ÁREA: MEDIA INTEGRAL DÉCIMO-Logica y Fundamentos de Matemáticas-TALLER 4

DOCENTE: Milton Gamboa Ortiz

Taller Diagnóstico

Preliminares: fórmulas y herramientas de apoyo

Antes de iniciar el análisis estadístico, es importante recordar algunas fórmulas básicas y comprender su utilidad. Estas medidas nos permitirán describir y analizar de forma clara el comportamiento de los datos relacionados con el rendimiento de los estudiantes. Además, se incluye un bloque de código en R para quienes deseen usar herramientas digitales para automatizar los cálculos y construir gráficos.

Tabla de fórmulas y explicaciones

Medida	Fórmula / Método	Explicación breve
Media	$\bar{x} = \frac{\sum x_i}{n}$	Suma de todos los datos dividida entre el total de valores.
Mediana	Si n impar: $x_{\frac{n+1}{2}}$ Si n par: $\frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}$	Valor central cuando los datos están ordenados.
Moda	Valor que más se repite	Se identifica observando el valor más frecuente.
Rango	máx – mín	Diferencia entre el mayor y menor valor.
Cuartiles	Q1, Q2 (mediana), Q3	Dividen el conjunto en cuatro partes iguales.
Rango inter-cuartílico	$RIQ = Q_3 - Q_1$	Mide la dispersión de los datos centrales (50%).
Desviación estándar (muestral)	$s = \sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2}$	Mide qué tanto se alejan los datos del promedio.

Código en R para quienes deseen automatizar los cálculos

```
# Ingresar los datos (reemplazar los valores)
tiempos <- c(ingresa los datos)

# Crear intervalos
breaks <- seq(0, 100, by = 10)
intervalos <- cut(tiempos, breaks = breaks, right = FALSE)

# Tabla de frecuencias
tabla <- table(intervalos)
tabla_df <- as.data.frame(tabla)
colnames(tabla_df) <- c("Intervalo", "Frecuencia")

# Agregar centro de clase
centros <- sapply(1:(length(breaks)-1), function(i) mean(c(breaks[i], breaks[i+1])))
tabla_df$Centro <- centros

# Graficar histograma con ggplot2
library(ggplot2)
ggplot(tabla_df, aes(x = factor(Centro), y = Frecuencia)) +
  geom_bar(stat = "identity", fill = "skyblue", color = "black") +
  labs(title = "Histograma de tiempos (por intervalos)",
       x = "Centro del intervalo (minutos)",
       y = "Frecuencia") +
  theme_minimal()

# Medidas estadísticas
media <- mean(tiempos)
mediana <- median(tiempos)
moda <- as.numeric(names(sort(table(tiempos), decreasing = TRUE)[1]))
rango <- max(tiempos) - min(tiempos)
cuartiles <- quantile(tiempos, probs = c(0.25, 0.5, 0.75))
RIQ <- IQR(tiempos)
desviacion <- sd(tiempos)

# Mostrar resultados
cat("Media:", media, "\n")
cat("Mediana:", mediana, "\n")
cat("Moda:", moda, "\n")
cat("Rango:", rango, "\n")
cat("Cuartiles:", cuartiles, "\n")
cat("RIQ:", RIQ, "\n")
cat("Desviación estándar:", desviacion, "\n")
```

Taller: Análisis estadístico de una carrera deportiva

Situación contextual

Durante los Juegos Intercolegiados del colegio **Brasilia Bosa**, se realizó una carrera de resistencia de 5 kilómetros con la participación de 100 estudiantes de distintos grados. El profesor de Educación Física registró el *tiempo que cada estudiante tardó en completar la carrera*, expresado en minutos.

Ahora, el docente desea analizar estos datos para comprender mejor el rendimiento general del grupo. Para ello, te ha pedido que realices un análisis estadístico utilizando técnicas como la tabla de frecuencias, medidas de tendencia central, medidas de dispersión e interpretación gráfica.

3, 5, 7, 9, 10, 12, 14, 15, 16, 18, 19, 20, 21, 22, 23, 24, 25, 26, 26, 27, 28, 28, 29, 29, 30, 30, 30, 30, 31, 31, 31, 32, 32, 33, 33, 34, 34, 34, 35, 35, 36, 36, 36, 36, 37, 37, 38, 38, 39, 39, 40, 40, 40, 40, 41, 41, 41, 42, 42, 43, 43, 44, 44, 45, 45, 46, 46, 47, 47, 48, 48, 49, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 62, 64, 65, 66, 68, 70, 72, 74, 75, 77, 78, 80, 82, 84, 86, 90

Instrucciones

1. Construye una **tabla de frecuencias** agrupando los datos en intervalos.
2. Dibuja un **histograma** usando los datos agrupados.
3. Calcula las siguientes **medidas de tendencia central**:
 - Media
 - Mediana
 - Moda
4. Calcula las siguientes **medidas de dispersión**:
 - Rango
 - Cuartiles (Q1, Q2, Q3)
 - Rango intercuartílico (RIQ)
 - Desviación estándar
5. Responde las siguientes preguntas de análisis.

Preguntas de análisis con apoyo teórico

1. ¿Por qué crees que la media es diferente a la mediana en este conjunto de datos?

$$\text{Media: } \bar{x} = \frac{\sum x_i}{n}$$

Mediana: valor central de los datos ordenados.

Reflexiona sobre cómo los valores extremos afectan la media en comparación con la mediana.

2. ¿Qué información te da la moda sobre el comportamiento de los estudiantes durante la carrera?

Moda: valor que más se repite.

Comenta si este tiempo representa un grupo significativo de estudiantes y por qué.

3. **¿Qué grupo de tiempos concentra la mayor cantidad de estudiantes? ¿Qué puede significar esto para la preparación física?**

Revisa tu tabla de frecuencias e identifica el intervalo más alto. Relaciona ese grupo con el nivel físico promedio de los estudiantes.

4. **¿Qué ventajas tiene usar el rango intercuartílico en lugar del rango total?**

Rango: máx - mín

RIQ: $Q_3 - Q_1$

Explica por qué es útil centrarse en los datos intermedios en lugar de los extremos.

5. **Si la desviación estándar es alta, ¿qué implicaciones tiene para la organización de futuras competencias?**

Desviación estándar: $s = \sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2}$

Explica si los estudiantes tuvieron un rendimiento homogéneo o muy variable, y qué ajustes se podrían hacer.