

Universidad Nacional del Altiplano Facultad de Ingeniera Estadística y Informática  
**Docente:** Fred Torres Cruz  
**Autor:** Jhoan Jeremy Chavez Lima

Trabajo Encargado - N° 001

## Análisis del coeficiente de variación en un conjunto de datos de no menos de 100 registros

### Introducción

Este código en R genera un conjunto de datos ficticios con las calificaciones de 100 estudiantes en 6 materias y calcula, para cada asignatura, su media, desviación estándar y el coeficiente de variación (CV). El objetivo es medir la dispersión relativa de las calificaciones y comparar en qué materias existe mayor variabilidad en el rendimiento de los estudiantes. Al final, se muestra un resumen ordenado de los coeficientes de variación de mayor a menor.

### Datos

Se generó un conjunto de datos ficticios que contiene las calificaciones de **100 estudiantes** en las siguientes materias:

- **Matemáticas:** 85, 90, 78, 92, 88, 76, 94, 82, 87, 91, 79, 86, 93, 84, 89, 77, 95, 83, 88, 90, ...
- **Física:** 72, 68, 85, 79, 91, 63, 88, 75, 82, 69, 87, 74, 90, 71, 84, 66, 89, 77, 83, 70, ...
- **Inglés:** 88, 87, 89, 86, 90, 85, 91, 88, 87, 89, 86, 90, 88, 87, 89, 85, 91, 88, 87, 90, ...
- **Historia:** 95, 68, 82, 91, 73, 86, 79, 94, 71, 88, 76, 93, 69, 85, 78, 96, 70, 87, 80, 92, ...
- **Química:** 45, 92, 78, 34, 88, 67, 41, 95, 73, 38, 89, 64, 47, 91, 76, 52, 85, 69, 43, 94, ...
- **Educación Física:** 98, 95, 97, 99, 96, 94, 98, 97, 95, 99, 96, 98, 94, 97, 99, 95, 98, 96, 97, 94, ...

Cada lista contiene los registros de los 100 estudiantes, representando sus calificaciones en cada asignatura.

## Métodos

Se realizaron los siguientes pasos en el análisis:

- Generación de un conjunto de datos ficticios con 100 estudiantes y 6 materias.
- Cálculo de medidas estadísticas descriptivas: media, desviación estándar y coeficiente de variación (CV).
- Comparación de la variabilidad relativa entre materias mediante el coeficiente de variación.
- Elaboración de un resumen ordenado de los coeficientes de variación de mayor a menor.

```
1 # Creamos datos los cuales son (100 estudiantes, 6 materias)
2 datos <- data.frame(
3   matematicas = c(85, 90, 78, 92, 88, 76, 94, 82, 87, 91, 79, 86, 93, 84,
4     89,
5     ... ), # <- se acort para brevedad
6   fisica = c(72, 68, 85, 79, 91, 63, 88, 75, 82, 69, ...),
7   ingles = c(88, 87, 89, 86, 90, 85, 91, 88, 87, 89, ...),
8   historia = c(95, 68, 82, 91, 73, 86, 79, 94, 71, 88, ...),
9   quimica = c(45, 92, 78, 34, 88, 67, 41, 95, 73, 38, ...),
10  educacion_fisica = c(98, 95, 97, 99, 96, 94, 98, 97, 95, 99, ...)
11 )
12 # Ver n mero de estudiantes y materias
13 cat("Tenemos", nrow(datos), "estudiantes y", ncol(datos), "materias\n\n")
14
15 # Calcular coeficiente de variaci n por materia
16 media_mat <- mean(datos$matematicas)
17 desv_mat <- sd(datos$matematicas)
18 cv_mat <- (desv_mat / media_mat) * 100
19
20 # ... (repetido para cada materia)
21
22 # Resumen ordenado de CV
23 materias <- c("Matem ticas", "F sica", "Ingl s", "Historia", "Qu mica",
24   "Ed. F sica")
25 cv_valores <- c(cv_mat, cv_fis, cv_ing, cv_hist, cv_quim, cv_ef)
26 orden <- order(cv_valores, decreasing = TRUE)
27 for(i in 1:length(orden)) {
28   cat(i, ".", materias[orden[i]], ":", round(cv_valores[orden[i]], 2), "%\n")
29 }
```

Listing 1: Código R utilizado

## Resultados

- La materia con mayor variabilidad relativa fue **Química** ( $CV = 26,89\%$ ).

Cuadro 1: Medidas descriptivas y coeficientes de variación por materia

Materia	Media (M)	Desviación Est. (DE)	CV (%)
Matemáticas	87.00	5.14	5.91
Física	78.13	7.70	9.86
Inglés	88.10	1.61	1.83
Historia	83.47	8.49	10.17
Química	67.65	18.20	26.89
Educación Física	96.76	1.62	1.67

- La materia con menor variabilidad fue **Educación Física** ( $CV = 1,67\%$ ).
- El orden de dispersión de mayor a menor fue: Química > Historia > Física > Matemáticas > Inglés > Educación Física.

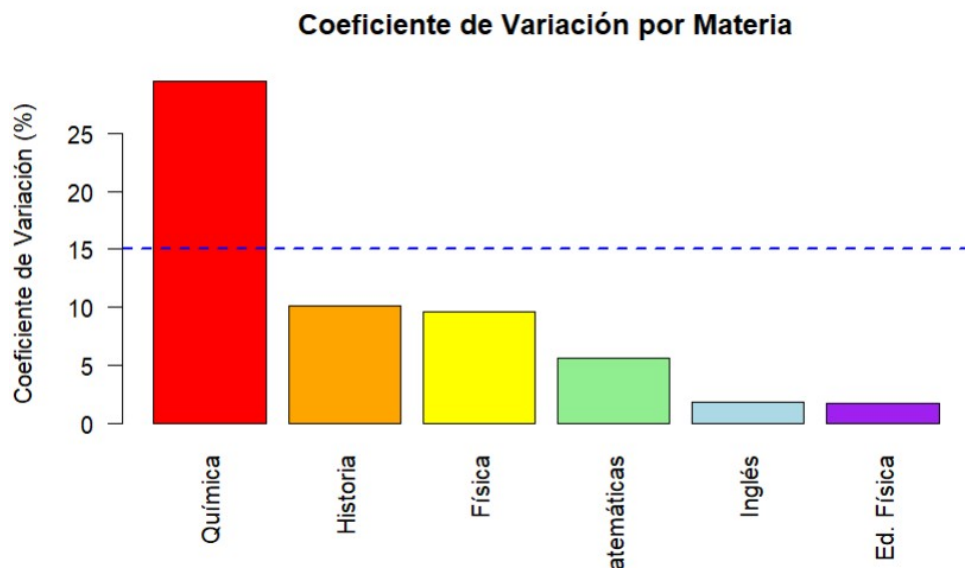


Figura 1: Coeficiente de variación comparado entre materias

## Conclusión

El análisis descriptivo permitió identificar diferencias en la variabilidad de las calificaciones entre materias. La asignatura de **Química** presentó la mayor dispersión relativa en el rendimiento de los estudiantes ( $CV = 26,89\%$ ), mientras que **Educación Física** mostró la menor variabilidad ( $CV = 1,67\%$ ), lo que sugiere un desempeño más homogéneo en esta última. En general, el coeficiente de variación resultó ser una medida útil para comparar la estabilidad del rendimiento académico entre distintas áreas.

## Repositorio

Todo el código y datos disponibles en:  
<https://github.com/JJereChavez/Estad-stica-Computacional.git>