



---

# INSTITUTO POLITÉCNICO NACIONAL

ESCUELA SUPERIOR DE CÓMPUTO

*Ingeniería en Sistemas Computacionales*

## Statistical Tools For Data Analytics

### **Profesor:**

Gonzales Cisneros Alejandro

### **Tarea 1:**

Estadística descriptiva

### **Alumno:**

Flores Osorio Adolfo Angel

Grupo: 6CV2



24 de febrero de 2025

Calcular el coeficiente de correlación muestral para los siguientes conjuntos de datos apareados.

Ejercicio1. Tabla de la correlación de las puntuaciones de IQ y el salario en miles de dólares de 30 trabajadores.

Trabajador i	Puntuación IQ X <sub>i</sub>	Salario anual y <sub>i</sub> (en miles de dólares)
1	110	68
2	107	30
3	83	13
4	87	24
5	117	40
6	104	22
7	110	25
8	118	62
9	116	45
10	94	70
11	93	15
12	101	22
13	93	18
14	76	20
15	91	14
16	84	19
17	83	16
18	112	52
19	80	11
20	91	13
21	113	29
22	124	71
23	79	19
24	116	43
25	113	44
26	94	17
27	95	15
28	104	30
29	115	63
30	90	16

promedio x <sub>i</sub> :	99.7666667
promedio y <sub>i</sub> :	31.5333333
n:	30

$$\frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} = 0.71866693$$

El coeficiente de correlación muestral es -0.718, lo que indica una correlación positiva fuerte

## Ejercicio2. Consumo de cigarros y radicales libres

Persona	Numero de cigarrillos consumidos (x_i)	Radicales libres (y_i)
1	18	202
2	32	644
3	25	411
4	60	755
5	12	144
6	25	302
7	50	512
8	15	223
9	22	183
10	30	375

promedio x_i:	28.9
promedio y_i:	375.1
n:	10

$$\frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} = 0.8759639$$

El coeficiente de correlación muestral es 0.875, lo que indica una correlación positiva fuerte

## Ejercicio3. Pulsaciones por minuto y años de escolarización completos

	Persona									
	1	2	3	4	5	6	7	8	9	10
Años de escolarización	12	16	13	18	19	12	18	19	12	14
pulsaciones	73	67	74	63	73	84	60	62	76	71

promedio x_i:	15.3
promedio y_i:	70.3
n:	10

$$\frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} = -0.7638033$$

El coeficiente de correlación muestral es -0.763, es una correlación negativa, lo que indica que, a mayor escolarización, tiende a haber menos pulsaciones por minuto