

NOTAS · 1

Pensada para 90 a 120 minutos. Es la clase con más fórmulas de la unidad: conviene hacer las cuentas de los cinco hogares en el pizarrón, a la par de las diapositivas.



NOTAS · 2

Las dos respuestas son correctas y la diferencia es grande. Preguntar cuál pondrían en un titular, y por qué creen que la media queda más arriba.

Se responde en la parte 2. Adelanto: unos pocos ingresos muy altos tiran de la media, y a la mediana no la mueven.



NOTAS · 3

Posición, dispersión y forma son las tres preguntas que se le hacen a una variable cuantitativa: dónde está, cuánto varía y si es simétrica.



SEMANA 3 · LA SUMATORIA Y LA MEDIA

Cinco hogares

Personas que viven en cada uno de los cinco primeros hogares de la base de la EPH.

Hogar, i	1	2	3	4	5
Personas, x_i	6	2	4	1	11

n = 5: la cantidad de datos. x_i es el dato del hogar i: $x_1 = 6$, $x_5 = 11$.

5 / 30

NOTAS · 5

La notación parece un detalle, pero es lo que permite escribir una cuenta una sola vez para cualquier cantidad de datos. Escribirla en el pizarrón y dejarla toda la clase.

SEMANA 3 · LA SUMATORIA Y LA MEDIA

La sumatoria y la media

$\Sigma\, x_i$ se lee «la suma de los x_i, para i de 1 a n».

La suma

$\Sigma\, x_i = 6 + 2 + 4 + 1 + 11 = 24$

La media

$\bar{x} = \Sigma\, x_i / n = 24 / 5 = 4{,}8$ personas

Los desvíos

$\Sigma\, (x_i - \bar{x}) = 0$

Los desvíos respecto de la media siempre suman cero: la media es el punto de equilibrio.

6 / 30

NOTAS · 6

La media de la muestra se escribe $\bar{x}$; la de la población, μ. Es la distinción de parámetro y estadístico de la semana 1.

Que verifiquen la tercera tarjeta con los cinco hogares: 1,2 + −2,8 + −0,8 + −3,8 + 6,2 = 0.



NOTAS · 7

Ordenados: 1, 2, 4, 6, 11. La mediana es el tercero, 4. Sin el hogar de 11, quedan cuatro y la mediana es el promedio de los dos del medio.

La media bajó 1,55 personas y la mediana, 1: la media usa el valor de cada dato, la mediana solo su orden.



SEMANA 3 · MEDIDAS DE POSICIÓN

Tres medidas del centro

	Qué es	Para tener en cuenta
Media	El total repartido en partes iguales	Usa todos los valores: la mueve un dato extremo
Mediana	El valor del medio de los datos ordenados	Resistente: no la mueven los extremos
Modo	El valor más frecuente	El único que sirve para una nominal; puede haber varios

9 / 30

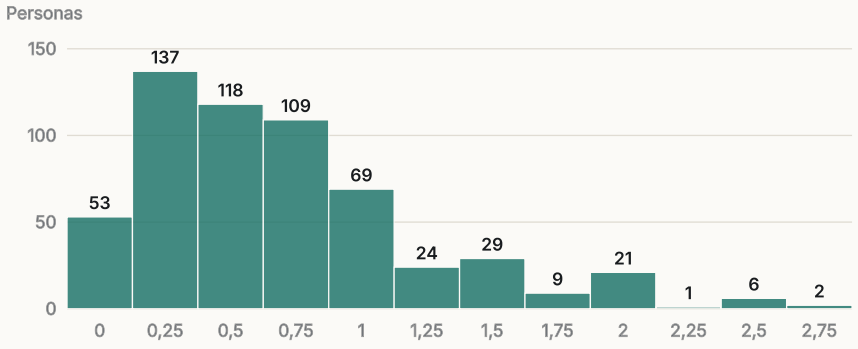
NOTAS · 9

Resistente quiere decir que cambiar unos pocos datos, aunque sea mucho, la cambia poco. La mediana lo es; la media, no.

SEMANA 3 · MEDIDAS DE POSICIÓN

Los ingresos: una cola larga a la derecha

Personas ocupadas por ingreso de la ocupación principal, en intervalos de 250.000 pesos. Debajo de cada barra, dónde empieza, en millones.



Media: $ 881.992. Mediana: $ 700.000. Otras 16 personas ganan 3 millones o más y no entran en el gráfico. Fuente: INDEC, Encuesta Permanente de Hogares, primer trimestre de 2026. Gran Catamarca.

10 / 30

NOTAS · 10

La mayoría gana menos de un millón y unos pocos se estiran hacia la derecha. Esos pocos suben la media, pero a la mediana no la mueven: para la mediana da igual que el que más gana cobre 3 millones o 14.

Por eso los ingresos se informan con la mediana. La media sigue sirviendo para otra pregunta: cuánto suman entre todos.

PREGUNTA PARA LA CLASE

¿Media, mediana o modo?

a) El ingreso típico de los hogares de una ciudad

b) El consumo de agua de los hogares de un barrio, para dimensionar la red

c) El tiempo de espera típico en una guardia

d) El talle de zapatillas que más conviene tener en stock

e) La nota de un parcial con notas parejas, sin extremos

11 / 30

NOTAS · 11

Mediana: los ingresos son asimétricos. Media: lo que importa es el total, y el total es la media por n. Mediana: los tiempos tienen cola a la derecha. Modo: interesa el más frecuente, y el talle se usa como categoría. Media o mediana: con datos simétricos dan casi lo mismo, y la media aprovecha toda la información.

SEMANA 3 · MEDIDAS DE POSICIÓN

El modo de una variable continua

$ 800.000

es el ingreso más repetido: lo declararon 47 personas

$ 500.000

lo declararon 45: por 2 casos no es el modo

Con datos continuos el modo depende de cómo se redondea al responder. Es la medida natural de las nominales: el modo de la categoría ocupacional es «Obrero o empleado».

12 / 30

NOTAS · 12

La gente responde el ingreso en números redondos: por eso hay valores repetidos. Si todos dijeran el monto exacto, ningún valor se repetiría y no habría modo.

En una tabla agrupada, la clase modal (la del intervalo con más datos) es más estable que el modo de los datos sueltos.

SEMANA 3 · MEDIDAS DE POSICIÓN

La media ponderada

Cada persona encuestada representa a una cantidad distinta de habitantes: su ponderador, PONDERA.

36,2 años

edad media de las personas encuestadas: cada una pesa lo mismo

33,7 años

edad media ponderada: estima la de Gran Catamarca

$\bar{x}_p = \Sigma\, w_i\, x_i \,/\, \Sigma\, w_i$, donde w_i es el peso de cada dato. Si todos los pesos son iguales, es la media de siempre.

13 / 30

NOTAS · 13

La diferencia, 2,4 años, dice que en la muestra hay, en proporción, más personas mayores que en la población. El ponderador lo corrige.

El mismo cálculo sirve para promediar promedios: el promedio de tres comisiones se pondera por la cantidad de estudiantes de cada una. Está en el práctico.

SEMANA 3 · MEDIDAS DE POSICIÓN

Los cuantiles

Cortan los datos ordenados en partes con la misma cantidad de casos.

Cuantil	Ingreso	Qué dice
Primer cuartil, Q_1	$ 400.000	El 25 % gana esto o menos
Mediana, el segundo cuartil	$ 700.000	La mitad gana esto o menos
Tercer cuartil, Q_3	$ 1.000.000	El 75 % gana esto o menos
Percentil 90	$ 1.600.000	Solo el 10 % gana más

Personas ocupadas que declararon el ingreso de su ocupación principal: 594. Cuartiles por el método de R y de Excel; SPSS usa otra convención y puede dar algo distinto.

14 / 30

NOTAS · 14

El cuantil p deja una proporción p de los datos por debajo. Cuartiles: p = 0,25, 0,50, 0,75. Deciles: de a 0,10. Percentiles: de a 0,01.

Hay varias formas de calcularlos cuando el corte cae entre dos datos. Con muchos datos dan casi lo mismo; con pocos, no. El práctico compara las dos más usadas.

3

Medidas de dispersión

Cuánto se alejan los datos del centro

15 / 30

SEMANA 3 · MEDIDAS DE DISPERSIÓN

Los desvíos de los cinco hogares

Hogar	x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1	6	1,2	1,44
2	2	−2,8	7,84
3	4	−0,8	0,64
4	1	−3,8	14,44
5	11	6,2	38,44
Suma	24	0	62,80

$\bar{x}$ = 4,8. La columna de los desvíos suma cero: por eso se elevan al cuadrado antes de sumar.

16 / 30

NOTAS · 16

El cuadrado hace que todos los desvíos cuenten como positivos y que los grandes pesen mucho más que los chicos.

SEMANA 3 · MEDIDAS DE DISPERSIÓN

Varianza y desvío estándar

Varianza

$s^2 = \Sigma (x_i - \bar{x})^2 / (n - 1) = 62{,}80 / 4 = 15{,}70$

En personas al cuadrado: no se interpreta directamente.

Desvío estándar

$s = \sqrt{15{,}70} = 3{,}96$ personas

En la unidad de los datos: lo típico que un dato se aleja de la media.

¿Por qué n − 1?

Los desvíos suman cero: sabiendo n − 1, el último está determinado.

Con n, la varianza de la muestra subestima la de la población.

17 / 30

NOTAS · 17

Para la población entera se divide por N y se escribe σ^2. Con n − 1 la varianza de la muestra es un buen estimador de la de la población: se ve en la unidad de estimación.

R, Excel (DESVEST.M), SPSS e InfoStat dividen por n − 1.

SEMANA 3 · MEDIDAS DE DISPERSIÓN

Cuatro medidas de la dispersión del ingreso

Medida	Valor	Qué mira
Rango	$ 13.950.000	Solo los dos extremos: del menor al mayor
Rango intercuartílico	$ 600.000	La mitad central: $Q_3 - Q_1$. Resistente
Desvío estándar	$ 935.780	Todos los datos: lo mueven los extremos
Coeficiente de variación	106,1 %	El desvío como proporción de la media, sin unidad

Personas ocupadas que declararon el ingreso de su ocupación principal: 594.

18 / 30

NOTAS · 18

El rango va de $ 50.000 a $ 14.000.000: depende de dos personas. El RIC dice que la mitad central gana entre $ 400.000 y $ 1.000.000.

La regla es la misma que con el centro: la mediana va con el RIC, y la media con el desvío estándar.

SEMANA 3 · MEDIDAS DE DISPERSIÓN

El coeficiente de variación

CV = s / x̄. Para comparar la variabilidad de variables con distinta unidad o con medias muy distintas.

53,5 %	57,8 %	106,1 %
personas por hogar	edad de las personas encuestadas	ingreso de las personas ocupadas

Años, pesos y personas no se comparan; los porcentajes, sí. Por encima de 30 % se suele hablar de una dispersión alta.

19 / 30

NOTAS · 19

El desvío del ingreso es de casi un millón de pesos y el de la edad, de unos 20 años: compararlos no tiene sentido. El CV los pone en la misma escala.

Solo tiene sentido en variables de razón, con un cero verdadero: el CV de una temperatura en grados Celsius cambia si se pasa a Fahrenheit.

4

Posición relativa y forma

Dónde está un dato, y cómo es la distribución

20 / 30

SEMANA 3 · POSICIÓN RELATIVA Y FORMA

El puntaje z

$z = (x - \bar{x}) / s$: a cuántos desvíos estándar de la media está un dato.

z = 0
El dato está justo en la media.

z positivo
Por encima de la media; negativo, por debajo.

z más allá de ±2
Poco habitual, si la distribución es más o menos simétrica.

21 / 30

NOTAS · 21

El z no tiene unidad: por eso sirve para comparar posiciones en variables distintas, o en grupos con medias y desvíos distintos.

PREGUNTA PARA LA CLASE

Una persona ocupada de 60 años gana $ 2.000.000 por mes. ¿Se aleja más del resto por su edad o por su ingreso?

22 / 30

NOTAS · 22

Entre las personas ocupadas, la edad media es 41,2 años con un desvío de 12,3: z = (60 − 41,2) / 12,3 = 1,52.

El ingreso medio es $ 881.992 con un desvío de $ 935.780: z = 1,19. Se aleja más por la edad, aunque 2 millones suene a mucho.

Con una distribución tan asimétrica como la del ingreso, el z hay que tomarlo con cuidado: el desvío está inflado por la cola. El percentil es una alternativa más robusta.

SEMANA 3 · POSICIÓN RELATIVA Y FORMA

La forma y el orden de las medidas

Simétrica

La media y la mediana casi coinciden.

La estatura de adultos de un mismo sexo.

Asimétrica a la derecha

La media queda **por encima** de la mediana.

Los ingresos, los tiempos de espera.

Asimétrica a la izquierda

La media queda **por debajo** de la mediana.

La edad al morir, la nota en un examen fácil.

23 / 30

NOTAS · 23

La media se corre hacia la cola porque la cola tira de ella; la mediana se queda con el dato del medio.

SEMANA 3 · POSICIÓN RELATIVA Y FORMA

El coeficiente de asimetría

Cero si la distribución es simétrica; positivo con cola a la derecha; negativo con cola a la izquierda.

0,38

edad: casi simétrica

1,00

personas por hogar: cola a la derecha

6,43

ingreso: muy asimétrico

Calculado como Excel (COEFICIENTE.ASIMETRIA) y SPSS, con la corrección por tamaño de muestra.

24 / 30

NOTAS · 24

Más allá de ±1 se suele hablar de una asimetría marcada. El número sirve para comparar; para ver la forma, el histograma o la caja dicen más.

5

El diagrama de caja

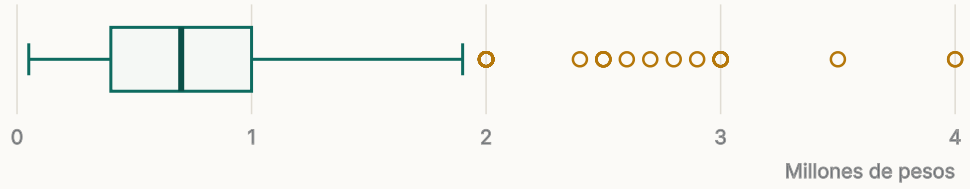
Cinco números y los atípicos

25 / 30

SEMANA 3 · EL DIAGRAMA DE CAJA

El diagrama de caja

El ingreso de las personas ocupadas de Gran Catamarca.



La caja va de Q_1 a Q_3, con la mediana adentro. Los círculos son atípicos. Otros 6 atípicos, de más de 4 millones, no entran en el gráfico.

26 / 30

NOTAS · 26

La caja contiene a la mitad central. Los bigotes llegan hasta el dato más extremo que no es atípico.

Se ve la asimetría sin hacer cuentas: el bigote de la derecha es mucho más largo que el de la izquierda, y todos los atípicos están del lado de los ingresos altos.

SEMANA 3 · EL DIAGRAMA DE CAJA

Cómo se arma

	Ingreso de los ocupados
Q_1, mediana y Q_3	$ 400.000, $ 700.000 y $ 1.000.000
Rango intercuartílico, RIC	$ 600.000
Límite superior: Q_3 + 1,5 RIC	$ 1.900.000
Límite inferior: Q_1 − 1,5 RIC	$ −500.000
Bigotes: el menor y el mayor dato dentro de los límites	$ 50.000 y $ 1.900.000
Atípicos: fuera de los límites	46 personas

La regla de 1,5 RIC es de Tukey (1977). Los límites no se dibujan: marcan dónde termina el bigote.

27 / 30

NOTAS · 27

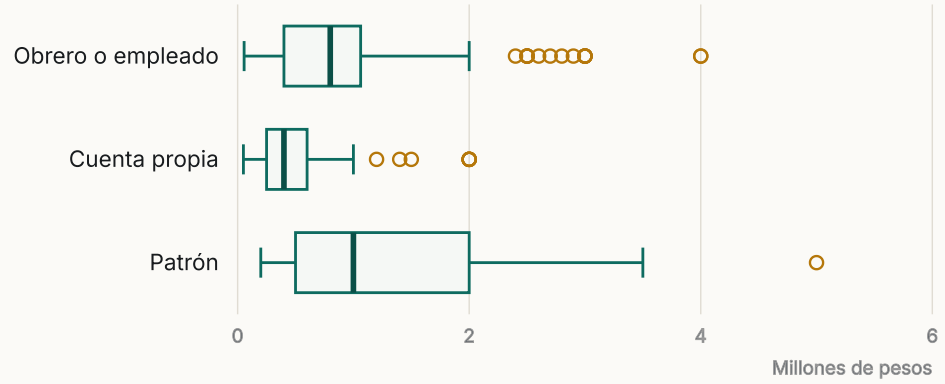
El límite inferior es negativo, así que ningún ingreso queda por debajo: los 46 atípicos están todos arriba. Es lo que pasa en una distribución con cola a la derecha.

Que los marque como atípicos no quiere decir que sean errores: la regla está pensada para distribuciones más o menos simétricas.

SEMANA 3 · EL DIAGRAMA DE CAJA

Comparar grupos

El ingreso de las personas ocupadas, por categoría ocupacional.



Obrero o empleado: 444. Cuenta propia: 109. Patrón: 41. Otros 5 atípicos, de más de 5 millones, no entran. Fuente: INDEC, Encuesta Permanente de Hogares, primer trimestre de 2026. Gran Catamarca.

28 / 30

NOTAS · 28

Pedir que comparen: ¿quiénes ganan más en el centro? ¿Quiénes son más dispares? La caja de los patrones es la más ancha: el RIC más grande. La de cuenta propia, la más baja.

Las medianas: obrero o empleado $ 800.000; cuenta propia $ 400.000; patrón $ 1.000.000. Para comparar grupos, las cajas lado a lado son el mejor gráfico.

SEMANA 3 · EL DIAGRAMA DE CAJA

Un atípico: ¿lo saco o lo dejo?

Se corrige o se saca

- Si es un **error**: una edad de 230 años. Se corrige o queda como dato faltante.
- Si **no pertenece** a la población estudiada, con un criterio escrito antes de mirar los datos.

Se deja

- Si es un **valor real**: un ingreso de varios millones existe.
- Se eligen medidas resistentes, o se informa con y sin ese caso.

La regla de 1,5 RIC encuentra los casos que hay que mirar; no decide cuáles se borran.

29 / 30

NOTAS · 29

El práctico tiene un ejercicio para discutir tres casos de este tipo. Sacar un dato real porque molesta es describir una población que no existe.

SEMANA 3 · POSICIÓN, DISPERSIÓN Y FORMA

Para llevarse

- La media usa el valor de cada dato y la mueven los extremos; la mediana usa el orden y es resistente.
- Con asimetría a la derecha, la media queda por encima de la mediana: los ingresos se informan con la mediana.
- Media con desvío estándar; mediana con rango intercuartílico.
- El coeficiente de variación y el puntaje z no tienen unidad: sirven para comparar.
- Un atípico se mira antes de decidir: si es real, se queda.

Guía de trabajo práctico 3
estadistica.ar/docentes/trabajos-practicos/semana-3

Calculadora de descriptivos
estadistica.ar/herramientas/calc-descriptivos

Cuál medida de posición usar
estadistica.ar/conceptos/eleccion-de-medida-de-posicion

Diagrama de caja
estadistica.ar/conceptos/diagrama-de-caja

30 / 30

NOTAS · 30

Para el práctico: calculadora científica, y la base de la EPH para el ejercicio con software.