ESTADISTICA.AR

Posición, dispersión y forma

Estadística · Semana 3 · Unidad 2: Estadística descriptiva

¿Cuánto gana una persona ocupada en Gran Catamarca?

$ 881.992

si se reparte el total en partes iguales: la **media**

$ 700.000

gana la persona del medio, con la mitad por debajo y la mitad por encima: la **mediana**

Personas ocupadas que declararon el ingreso de su ocupación principal: 594. Fuente: INDEC, Encuesta Permanente de Hogares, primer trimestre de 2026. Gran Catamarca.

Hoy

1

La sumatoria y la media

Resumir muchos datos en un número

2

Medidas de posición

Dónde está el centro, y algo más

3

Medidas de dispersión

Cuánto se alejan los datos del centro

4

Posición relativa y forma

Dónde está un dato, y cómo es la distribución

5

El diagrama de caja

Cinco números y los atípicos

1

La sumatoria y la media

Resumir muchos datos en un número

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: $x_1 = 6$, $x_5 = 11$.

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.

Un hogar grande mueve la media

4,8 y 4

la media y la mediana de los cinco hogares

3,25 y 3

la media y la mediana sin el hogar de 11 personas

La mediana es el valor del medio de los datos ordenados; con una cantidad par, el promedio de los dos del medio.

2

Medidas de posición

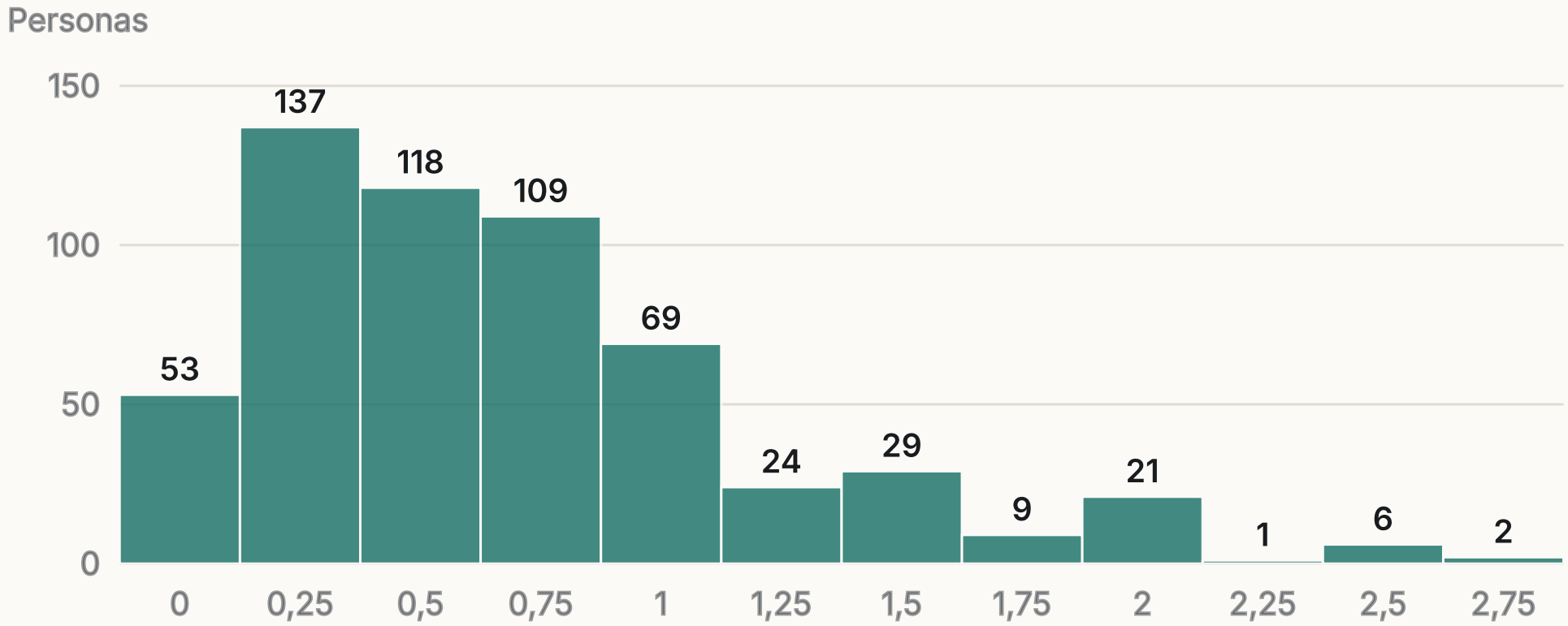
Dónde está el centro, y algo más

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

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.

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

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».

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.

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.

3

Medidas de dispersión

Cuánto se alejan los datos del centro

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.

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.

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.

El coeficiente de variación

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

53,5 %

personas por hogar

57,8 %

edad de las personas encuestadas

106,1 %

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.

4

Posición relativa y forma

Dónde está un dato, y cómo es la distribución

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.

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?

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.

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.

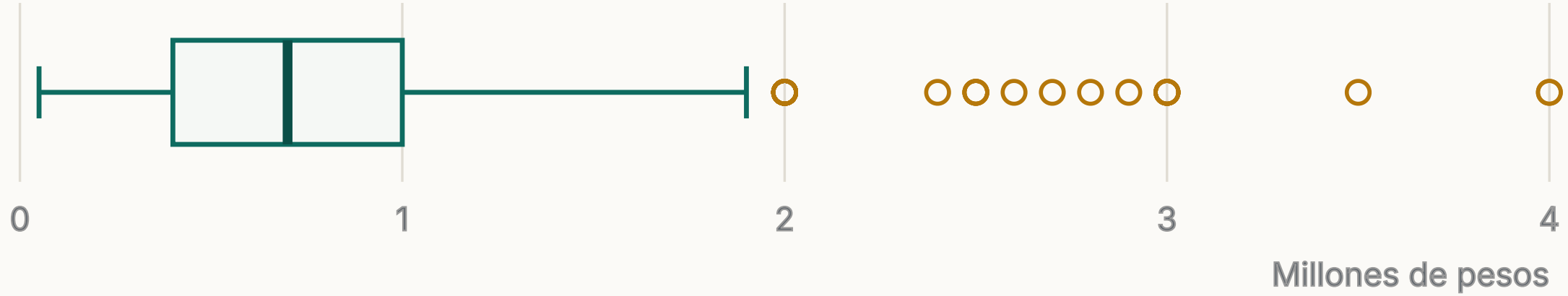
5

El diagrama de caja

Cinco números y los atípicos

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.

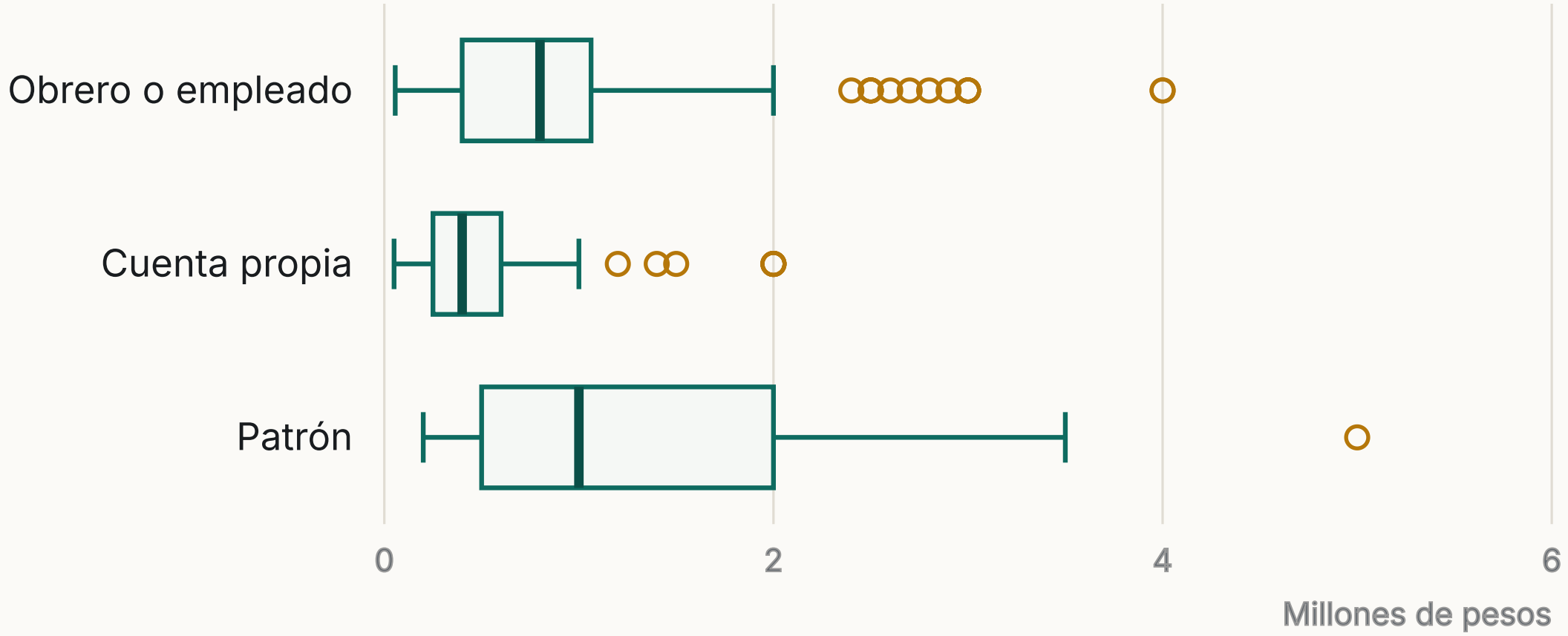
Cómo se arma

	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.

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.

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.

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

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

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

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