

A.1 INTRODUCCIÓN

El INEI, desde 1986 a la fecha, viene realizando cuatro Encuestas Demográficas y de Salud Familiar de periodicidad quinquenal: ENDES I, en 1986; ENDES II, en 1991-92; ENDES III, en 1996 y ENDES IV, en el 2000; y, desde octubre del 2003, cuatro operativos de periodicidad anual como ENDES Continua para atender la necesidad de información que apoye el monitoreo y evaluación a nivel nacional de los indicadores de productos y resultados intermedios de los programas y proyectos de población y salud; y, por agregación de muestra de al menos tres años de estimaciones a nivel departamental para indicadores de producto y resultados intermedios.

A fin de proveer información que permita estimar de manera oportuna y confiable los indicadores identificados en los Programas Estratégicos para el periodo 2012 al 2014, se ha diseñado la muestra para la Encuesta Demográfica y de Salud Familiar ENDES Continua 2012 – 2014, ante la culminación de la Muestra Maestra 2009-2011. Esta es una muestra maestra probabilística que fue seleccionada en dos etapas: la selección de conglomerados en la primera etapa y la selección de viviendas en la segunda etapa. La muestra maestra está constituida por 2 852 conglomerados, distribuidos aleatoriamente en cuatro submuestras de conglomerados para ejecutarse en tres años consecutivos. Cada año de estudio tiene 1 426 conglomerados, distribuidos proporcionalmente en el área urbana y rural de cada departamento.

Para la ENDES 2012 el número de conglomerados por departamento fue alrededor de 53, con la excepción de Lima en donde se seleccionaron 200 conglomerados. La muestra total permite sólidas estimaciones de indicadores para los hogares y para las mujeres entre 15-49 años de edad para un número de variables importantes a nivel anual, en cada uno de los 24 departamentos y región natural del país (Costa, Sierra, Selva y Lima Metropolitana).

Para la muestra total, el número de viviendas fue 28 376, de las cuales solo 27 488 fueron entrevistadas. En los hogares entrevistados, se encontraron un total de 24 552 mujeres elegibles y de las cuales se obtuvo un total de 23 888 entrevistas completas. La muestra seleccionada representa la totalidad de la población total del país.

A.2 DISEÑO Y SELECCIÓN DE LA MUESTRA

Marco Muestral

Dado que el diseño para la Encuesta ENDES Continua 2012 - 2014 es bietápica, se utiliza un marco muestral en cada una de las etapas de la selección de las unidades de muestreo.

Para la primera etapa: selección de UPM (conglomerados), se utiliza la información del Censo de Población y Vivienda del 2007. Para la segunda etapa: selección de USM (viviendas), se utiliza el marco muestral proveniente de la actualización cartográfica y registro de edificios y viviendas realizada previamente a las entrevistas, con el objetivo de identificar y registrar cambios en las áreas seleccionadas. Producto de este proceso se tiene un registro actualizado de viviendas que en ese momento existían en el área, el cual constituirá el marco de muestreo para la selección de viviendas.

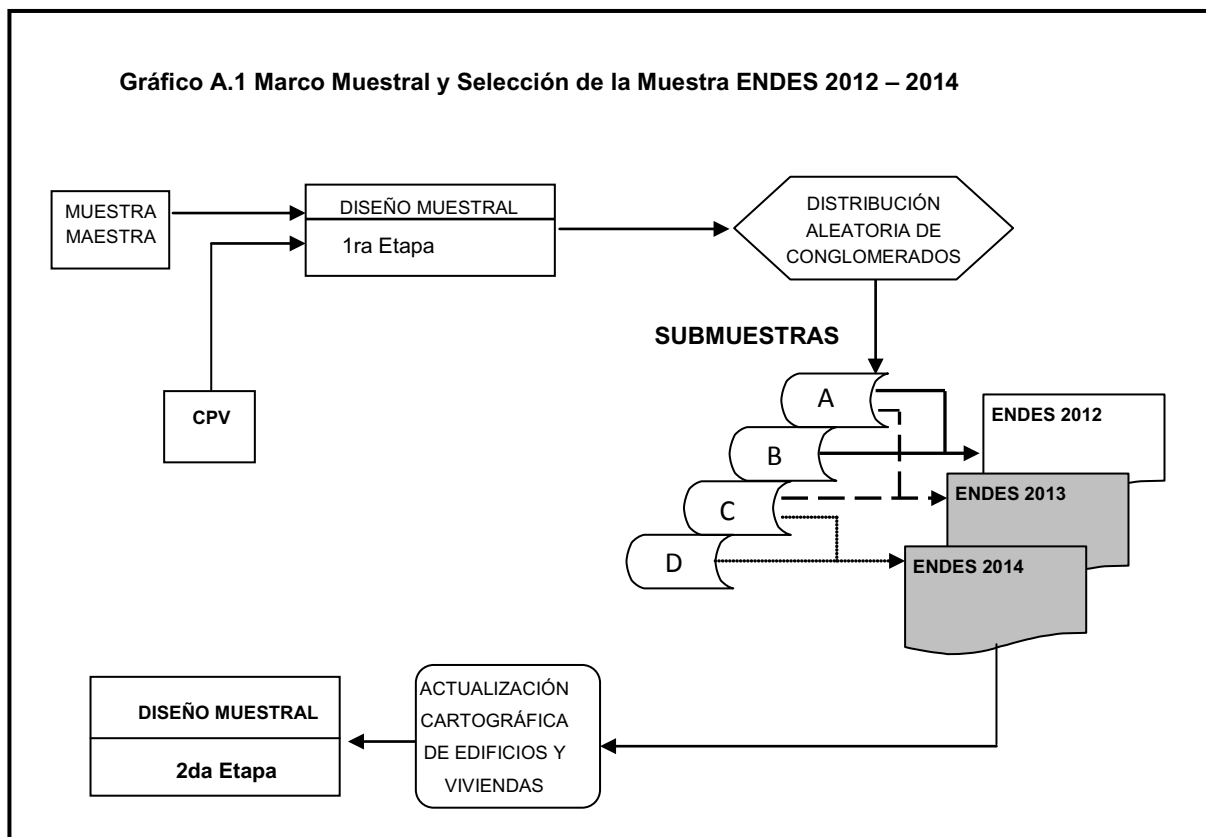
En la práctica, se debe considerar periódicamente un proceso de actualización debido a que con el transcurso del tiempo el número de viviendas contenidas en cada una de las áreas puede cambiar. Este sistema de actualización permite incorporar periódicamente al marco de muestreo en la segunda etapa las nuevas viviendas que aparecen en el proceso de expansión urbana y asignarle una probabilidad no nula de pertenecer a la muestra.

Muestra Panel de conglomerados

La Muestra Panel para la ENDES Continua 2012 - 2014 es de **conglomerados**, a diferencia de otros estudios de encuestas llamadas paneles clásicos, que tienen una muestra permanente de personas que responden a un cuestionario en periodos de tiempos fijos, lo que permite seguir la evolución de los indicadores en el tiempo. Entre las ventajas de esta aplicación se puede enumerar las siguientes:

- Posibilidad de estimar cambios (entre los conglomerados paneles de la encuesta) con mayor precisión, utilizando la correlación entre muestras paneles.
- Facilita el trabajo de campo, ya sea para la ubicación del conglomerado, o búsqueda de hogares.
- Reducción de esfuerzos en relación a los trabajos de actualización cartográfica y registro de viviendas.
- Mayor tiempo destinado a la supervisión del recojo de la información.

En el Gráfico A.1 se describe el uso de los marcos muestrales en las diferentes etapas de muestreo y la distribución de la muestra maestra, en la cual se observa la muestra ENDES 2012, que viene representada por la submuestra A (panel de conglomerados I semestre) y la submuestra C (II semestre):



Unidades de Muestreo

Área Urbana, las unidades de muestreo son:

- La Unidad Primaria de Muestreo (UPM), representada por el conglomerado el cual es el área geográfica conformado por una o varias manzanas que en conjunto tienen 120 viviendas particulares, en promedio.
- La Unidad Secundaria de Muestreo (USM), es la vivienda particular que integra la UPM seleccionada.

Área Rural, las unidades de muestreo son:

- La Unidad Primaria de Muestreo (UPM), es de dos tipos:
 - o El conglomerado el cual esta conformado por una o varias manzanas que en conjunto tienen 120 viviendas particulares, en promedio,
 - o El AER (área de empadronamiento rural) el cual esta conformado por uno o varios centros poblados rurales que en conjunto tienen 120 viviendas particulares (por lo general dispersas en el terreno).
- La Unidad Secundaria de Muestreo (USM), es la vivienda particular que integra la UPM (AER seleccionado).

Composición de la Muestra

El objetivo inicial fue obtener aproximadamente 24 200 entrevistas completas de mujeres elegibles, se utilizó información de la ENDES 2000 para estimar ciertos parámetros del diseño, especialmente las tasas de respuesta. Por ello, se ha estimado un total de 28 376 viviendas a ser seleccionadas para obtener el número antes mencionado de entrevistas completas. Ver cuadro A.1.

Selección de la Muestra

La muestra resultante de la ENDES 2012 se obtuvo en dos etapas:

La primera etapa, consiste en la selección sistemática de las UPM (conglomerados) con probabilidad proporcional a su tamaño. Cada UPM tiene un tamaño promedio de 120 viviendas particulares. Inmediatamente después de la primera etapa de selección, se implementó la actualización y registro exhaustivo de viviendas en las áreas seleccionadas.

Cuadro A.1: Número de conglomerados urbanos y rurales, según departamento, 2012.

Departamento	Número de viviendas seleccionadas	Número de conglomerados seleccionados			
		Total	Sede	Resto Urbano	Rural
Total	28 376	1 426	718	258	450
Amazonas	1 060	46	7	10	29
Áncash	1 333	66	28	10	28
Apurímac	1 121	49	10	7	32
Arequipa	1 164	63	47	8	8
Ayacucho	1 239	55	16	9	30
Cajamarca	1 081	46	7	7	32
Cusco	1 121	52	16	12	24
Huancavelica	994	42	8	2	32
Huánuco	1 032	45	10	7	28
Ica	1 163	58	27	23	8
Junín	1 084	50	19	12	19
La Libertad	1 159	61	41	8	12
Lambayeque	968	55	39	8	8
Lima	2 959	200	180	12	8
Loreto	922	50	25	9	16
Madre De Dios	1 062	49	24	9	16
Moquegua	1 125	53	20	21	12
Pasco	1 175	54	20	14	20
Piura	1 160	50	14	23	13
Puno	1 478	55	9	16	30
San Martín	1 028	46	8	18	20
Tacna	963	62	51	2	9
Tumbes	967	58	46	8	4
Ucayali	1 018	61	46	3	12

Proceso de selección de los conglomerados:

- El número de conglomerados a ser seleccionados en cada departamento (Lima incluye la Provincia Constitucional del Callao) fue calculado como la división del tamaño de muestra esperado en el departamento entre el tamaño promedio de muestra esperado dentro del conglomerado.
- El total de muestra estimado en cada departamento fue distribuido de manera proporcional entre sus subestratos (urbano, semirural y rural) de acuerdo a la población censal, y también dentro de cada subestrato, los conglomerados fueron ordenados de acuerdo a la continuidad geográfica en serpentin.
- Para cada subestrato, se elaboró una lista de conglomerados con sus correspondientes totales de población (viviendas) y también de sus respectivos acumulados parciales de población.
- En cada subestrato, se realizó la selección (sistemática y con probabilidad proporcional a su tamaño poblacional) del número de conglomerados requeridos para la ENDES Continua, digamos , en el total de ellos de dicho subestrato.

- La probabilidad de selección del i-ésimo conglomerado para la ENDES Continua es:

$$P_{li} = \frac{a * M_i}{\sum M_i}$$

Donde:

a : Número de conglomerados a ser seleccionados en el subestrato

M_i : Total de viviendas del i-ésimo conglomerado de acuerdo al censo 2007

$\sum M_i$: Total de viviendas en el subestrato según el censo 2007

La selección de las UPMs, mediante muestreo sistemático con probabilidad proporcional al tamaño, se realizó de la manera siguiente:

- Se listaron las UPMs con su respectivo tamaño (m_i) de viviendas.
- Se acumularon los valores de m_i uno a uno. Al final se verificó que el acumulado total sea igual a la suma de los valores m_i .
- Dado que a es el número de UPM a ser seleccionada, se calcula el intervalo de selección (K), eligiendo al azar un número entre 1 y K . Este número se denomina arranque aleatorio (aa).

$$K = \frac{\text{Total de viviendas en el subestrato}}{\text{Nº de UPMs a elegirse en el subestrato}} = \frac{\sum M_i}{a}$$

El arranque aleatorio (aa) es el número que determina la primera UPM a seleccionar. La segunda UPM a seleccionar sería: (aa) + K , la tercera UPM: (aa) + $2K$, y así sucesivamente hasta completar las UPMs requeridas para la muestra.

En la segunda etapa, un número de viviendas (USM) en cada UPM de la muestra es elegido usando la selección sistemática para cada conglomerado seleccionado.

Antes de la **segunda etapa** de selección previamente se realiza la Actualización Cartográfica y Registro de Viviendas de cada conglomerado. Para el proceso de selección de las USM o viviendas se siguieron los siguientes pasos:

- En cada UPM, se procede a seleccionar en forma sistemática simple aleatoria el número de viviendas requeridas para la ENDES 2012.
- La probabilidad final del j-ésima vivienda está dada por:

$$P_{li} = \frac{a * M_i}{\sum M_i} \left(\frac{n_i}{N_i} \right)$$

Donde:

n_i : Número de viviendas (USMs) seleccionadas en el i-ésimo conglomerado.

N_i : Número total de viviendas (USMs) en el i-ésimo conglomerado

- Una vez establecido el número de viviendas a ser seleccionadas para cada conglomerado, el procedimiento de selección dependerá del tipo de conglomerado, es decir, si es urbano o rural.

Conglomerado urbano, se procedió a seleccionar automáticamente las unidades secundarias de muestreo por selección sistemática, utilizando el siguiente algoritmo:

- Se listan las viviendas particulares de la UPM seleccionada.
- Se enumeran las viviendas en la lista desde 1 hasta M_{ij} .
- Se determina el intervalo de selección (K):

$$K = \frac{M_{ij}}{m_j}$$

Donde:

Total de viviendas ocupadas en el i -ésimo conglomerado

Total de viviendas a seleccionar en el i -ésimo conglomerado

- Se elige al azar un número como arranque aleatorio (aa). Este número muestral debe estar entre 1 y K .
- El arranque aleatorio (aa) es el número muestral que corresponde a la primera vivienda seleccionada. La segunda vivienda seleccionada será el arranque aleatorio (aa) + K . La tercera vivienda seleccionada será el arranque aleatorio (aa) + $2K$, y así sucesivamente hasta completar el número requerido de viviendas a seleccionar dentro del conglomerado.

Conglomerado Rural (AER), la selección de viviendas se hizo en grupos compactos de ocho viviendas contiguas sobre el listado de viviendas de acuerdo con el procedimiento siguiente:

- Se enumeran las viviendas en forma consecutiva en el listado del conglomerado rural.
- Después de cada múltiplo de ocho (es decir, 8, 16, 24, etc.), se traza una línea horizontal (con un marcador de color). Cada grupo está formado por las viviendas entre dos líneas horizontales consecutivas, y el primer grupo está constituido por la primera vivienda en la lista hasta el anterior inmediato a la primera línea horizontal; en el caso de existir más de cuatro viviendas después de la última línea horizontal, el último grupo estará constituido por esas viviendas.
- Si hubieran viviendas (menos de cuatro) después de la última línea horizontal, ellas serán distribuidas uniformemente entre los grupos compactos anteriores en forma sistemática y con inicio al azar.
- Todos los grupos así formados, se denominan grupos compactos, y G_{ij} es el total de grupos compactos formados en el conglomerado.
- Se calcula el número de grupos compactos a seleccionar:

$$g_{ij} = \frac{m_j}{8}$$

- Se selecciona entonces g_{ij} grupos compactos del total G_{ij} , en forma sistemática simple con arranque aleatorio.
- Todas las viviendas dentro de los grupos compactos seleccionados en el conglomerado rural son considerados como la muestra de viviendas de dicho conglomerado.

Metodología de Estimación

Para que las estimaciones derivadas de la ENDES 2012 sean representativas de la población, es necesario multiplicar los datos de cada hogar muestral contenido en los archivos por el peso o factor de ponderación calculada según el diseño muestral.

El factor final para cada hogar tiene dos componentes: el factor básico de muestreo y factor de ajuste por la no respuesta. El factor básico de expansión es determinado por el diseño de la muestra y equivale al inverso de la probabilidad final de selección de cada hogar. Esto quiere decir que los factores de básicos de expansión varían según el departamento de estudio.

Finalmente, el factor de ajuste por no respuesta es utilizado para ajustar los factores básicos de muestreo con el objetivo de tomar en cuenta los hogares no entrevistados en el conglomerado. Este ajuste a los factores básicos es realizado a nivel de departamento.

Rendimiento de la Muestra

En los Cuadros A.2 y A.3. se presentan las tasas de respuesta a nivel departamental, para la encuesta de hogares y para la encuesta de mujeres. De un total de 1426 conglomerados seleccionados, se logró implementar la encuesta en todos ellos.

De los 28 376 hogares que fueron seleccionados, se logró entrevistar a 27 488 hogares, con una tasa de respuesta de hogares de 99,0 por ciento, tasa que varía ligeramente de acuerdo al departamento. Las tasas más bajas se encontraron en Lima, la Provincia Constitucional del Callao y La Libertad (95,6, 96,8 y 97,8 por ciento respectivamente) y la más altas en los departamentos de Ayacucho y San Martín (100,0 por ciento respectivamente).

De un total de 24 552 mujeres elegibles, se logró entrevistar a 23 888, obteniéndose una tasa de respuesta individual de 97,3 por ciento. Las mayores tasas de respuesta se lograron en los departamentos de San Martín (99,1 por ciento), Huánuco (98,9 por ciento), Amazonas (98,8 por ciento) y las más bajas en Lima, la Provincia Constitucional del Callao, Moquegua y Tacna en proporciones que varían entre 92,3 y 96,7 por ciento.

Cuadro A.2 Implementación de la muestra: Hogares

Distribución porcentual de la población total en los hogares por resultados de la entrevista de hogar y tasa de respuesta de hogares, según departamento, Perú 2012.

Departamento	Resultado									Total	Hogares	Tasa de respuesta (TRHH)
	Entrevistas completas (HC)	Hogar sin persona competente (HP)	Aplazada (HA)	Rechazada (HR)	Vivienda no encontrada (VNE)	Hogar ausente (HA)	Vivienda desocupada/ no es vivienda (VD)	Vivienda destruida (VDD)	Otro (HO)			
Amazonas	96,5	0,0	0,0	0,2	0,0	1,2	0,8	0,0	1,2	100,0	1 060	99,8
Áncash	95,7	0,1	0,2	0,2	0,0	0,8	1,1	0,0	1,8	100,0	1 333	99,5
Apurímac	95,3	0,1	0,0	0,3	0,0	0,8	0,8	0,1	2,7	100,0	1 121	99,6
Arequipa	97,1	0,1	0,1	0,6	0,0	1,1	0,9	0,1	0,1	100,0	1 164	99,2
Ayacucho	96,3	0,0	0,0	0,0	0,0	1,1	1,5	0,0	1,1	100,0	1 239	100,0
Cajamarca	95,7	0,0	0,0	0,5	0,0	1,6	0,7	0,2	1,3	100,0	1 081	99,5
Callao	95,4	0,0	0,0	3,2	0,0	1,4	0,0	0,0	0,0	100,0	281	96,8
Cusco	93,6	0,2	0,0	0,4	0,0	2,2	1,5	0,0	2,1	100,0	1 121	99,4
Huancavelica	96,5	0,0	0,0	0,1	0,0	0,5	0,6	0,0	2,3	100,0	994	99,9
Huánuco	98,1	0,0	0,0	0,1	0,0	0,0	0,6	0,0	1,3	100,0	1 032	99,9
Ica	98,4	0,1	0,2	0,3	0,1	0,7	0,3	0,0	0,0	100,0	1 163	99,4
Junín	96,3	0,1	0,1	0,6	0,1	0,8	0,6	0,0	1,4	100,0	1 084	99,1
La Libertad	96,1	0,0	0,1	2,1	0,0	0,9	0,2	0,0	0,6	100,0	1 159	97,8
Lambayeque	96,9	0,0	0,0	0,8	0,0	1,0	0,3	0,0	0,9	100,0	968	99,2
Lima	93,0	0,2	0,5	3,6	0,0	2,4	0,1	0,0	0,2	100,0	2 678	95,6
Loreto	97,0	0,0	0,0	0,5	0,0	0,4	1,3	0,1	0,7	100,0	922	99,4
Madre de Dios	95,9	0,0	0,1	0,6	0,0	1,9	1,2	0,2	0,2	100,0	1 062	99,3
Moquegua	94,3	0,0	0,0	1,1	0,3	2,0	1,4	0,0	1,0	100,0	1 125	98,6
Pasco	95,1	0,2	0,0	0,3	0,0	1,8	0,9	0,0	1,8	100,0	1 175	99,6
Piura	97,2	0,2	0,0	0,3	0,0	0,4	1,3	0,1	0,5	100,0	1 160	99,5
Puno	93,8	0,0	0,1	0,1	0,0	2,0	0,8	0,2	3,0	100,0	1 478	99,8
San Martín	96,7	0,0	0,0	0,0	0,0	0,6	1,4	0,3	1,1	100,0	1 028	100,0
Tacna	97,0	0,0	0,1	1,1	0,0	0,5	0,7	0,0	0,5	100,0	963	98,7
Tumbes	98,1	0,0	0,0	0,4	0,0	0,7	0,5	0,0	0,2	100,0	967	99,6
Ucayali	97,6	0,0	0,0	0,3	0,0	0,3	0,7	0,3	0,8	100,0	1 018	99,7
Total	95,9	0,1	0,1	0,8	0,0	1,2	0,8	0,1	1,1	100,0	28 376	99,0

Cuadro A.3 Implementación de la muestra: mujeres

Distribución porcentual de mujeres elegibles en la muestra por resultados de las entrevistas de mujeres y tasa de respuesta, según departamento, Perú 2012.

Departamento	Resultado								Total	Mujeres	Tasa de respuesta (TRM)
	Entrevistas completas (HC)	Completa (TRM)	Ausente	Aplazada	Rechazada	Parcialmente completa	Incapacitada	Otra			
Amazonas	98,6	98,8	0,7	0,0	0,1	0,0	0,5	0,0	100,0	881	98,8
Áncash	97,4	97,9	0,4	0,3	0,6	0,0	0,8	0,0	100,0	1 142	97,9
Apurímac	98,3	98,6	0,1	0,1	0,1	0,0	1,0	0,0	100,0	802	98,6
Arequipa	97,3	98,1	1,2	0,0	0,4	0,0	0,3	0,0	100,0	933	98,1
Ayacucho	98,7	98,7	0,6	0,0	0,2	0,0	0,3	0,1	100,0	954	98,7
Cajamarca	97,1	97,5	0,3	0,1	0,3	0,0	1,7	0,0	100,0	930	97,5
Callao	90,5	93,5	4,8	0,3	0,7	0,0	0,7	0,0	100,0	292	93,5
Cusco	97,4	97,9	0,7	0,0	0,6	0,0	0,7	0,0	100,0	828	97,9
Huancavelica	98,3	98,4	0,1	0,0	0,5	0,1	0,8	0,0	100,0	773	98,4
Huánuco	98,8	98,9	0,4	0,0	0,0	0,0	0,7	0,0	100,0	911	98,9
Ica	97,3	97,9	1,7	0,1	0,1	0,0	0,2	0,0	100,0	1 041	97,9
Junín	97,4	98,3	1,0	0,0	0,2	0,0	0,5	0,0	100,0	935	98,3
La Libertad	94,7	96,8	1,2	0,0	1,3	0,0	0,8	0,0	100,0	1 027	96,8
Lambayeque	97,9	98,7	0,2	0,1	0,6	0,0	0,3	0,1	100,0	993	98,7
Lima	88,2	92,3	4,7	0,6	1,5	0,2	0,7	0,1	100,0	2 655	92,3
Loreto	98,1	98,7	0,3	0,0	0,4	0,1	0,5	0,0	100,0	986	98,7
Madre de Dios	97,2	97,9	0,9	0,0	0,6	0,2	0,4	0,0	100,0	939	97,9
Moquegua	94,3	95,6	2,7	0,4	0,6	0,1	0,4	0,1	100,0	771	95,6
Pasco	97,2	97,7	1,4	0,0	0,4	0,0	0,5	0,0	100,0	989	97,7
Piura	97,9	98,4	0,6	0,0	0,0	0,1	0,9	0,0	100,0	1 199	98,4
Puno	96,7	96,9	1,8	0,5	0,4	0,0	0,3	0,1	100,0	971	96,9
San Martín	99,1	99,1	0,4	0,0	0,1	0,0	0,3	0,0	100,0	894	99,1
Tacna	95,5	96,7	1,0	0,1	1,8	0,1	0,2	0,0	100,0	819	96,7
Tumbes	96,6	97,0	0,7	0,0	1,1	0,0	1,2	0,0	100,0	946	97,0
Ucayali	98,0	98,3	1,4	0,0	0,0	0,0	0,2	0,1	100,0	941	98,3
Total	96,3	97,3	1,3	0,1	0,6	0,0	0,6	0,0	100,0	24 552	97,3

