

**Diseño de muestra
Encuesta de Hogares
Encuesta de Niñas, Niños y Adolescentes que
realizan una actividad laboral o trabajan
2016**

Febrero 2017

UNIDAD DE MUESTREO Y LOGISTICA

Contenido

1. INTRODUCCIÓN.....	3
ASPECTOS METODOLÓGICOS DE LA ENCUESTA DE HOGARES 2016 Y ENNA.	3
2. OBJETIVOS DEL DISEÑO DE LA MUESTRA.	3
OBJETIVO GENERAL.	3
OBJETIVOS ESPECÍFICOS.....	3
3. ÁMBITO GEOGRÁFICO.....	4
4. UNIVERSO DE ESTUDIO.	4
5. UNIDADES DE OBSERVACIÓN, ANÁLISIS.....	4
6. UNIDADES DE MUESTREO.	5
7. PERIODO DE LEVANTAMIENTO DE DATOS.....	5
8. MARCO MUESTRAL.	5
CRITERIOS DE CONGLOMERACIÓN Y ESTRATIFICACIÓN DEL MARCO MUESTRAL.....	6
VARIABLE DE CONGLOMERACIÓN.	6
VARIABLE DE ESTRATIFICACIÓN.....	7
9. CARACTERÍSTICAS GENERALES DEL DISEÑO DE MUESTRA.....	7
10. NIVEL DE DESAGREGACIÓN DE LA INFORMACIÓN.	8
11. SELECCIÓN DE LA MUESTRA.	8
12. CÁLCULO DEL TAMAÑO DE LA MUESTRA.	9
13. DISTRIBUCIÓN DE LA MUESTRA.....	9
14. FACTORES DE EXPANSIÓN.	10
15. AJUSTES A LOS FACTORES DE EXPANSIÓN.....	11
POR PROYECCIONES DE POBLACIÓN AL 2016.	11
16. ESTIMADORES Y ERRORES DE MUESTREO.	12
17. INCIDENCIA DE CAMPO.	12

Contenido de Tabla

TABLA 1: ESTRUCTURA DEL MARCO MUESTRAL	6
TABLA 2: DISTRIBUCIÓN DE UPM POR DEPARTAMENTO	7
TABLA 3: ESTRATO GEOGRÁFICO.....	7
TABLA 4: REAGRUPACIÓN GEOGRÁFICA	8
TABLA 5: COEFICIENTE DE VARIACIÓN (CV) DEL TAMAÑO DE MUESTRA PLANIFICADA	10
TABLA 6: DISTRIBUCIÓN DE LA MUESTRA DE VIVIENDAS Y UPMS POR DEPARTAMENTO Y ÁREA GEOGRÁFICA	11
TABLA 7: COEFICIENTE DE VARIACIÓN (CV) DEL TAMAÑO DE MUESTRA EJECUTADA.....	13
TABLA 8: INCIDENCIA DE CAMPO POR DEPARTAMENTO	14

1. Introducción.

Las encuestas de hogares se desarrollaron a partir de 1980 con las encuestas permanentes de hogares, las cuales se remplazaron por las encuestas integradas de hogares EIH, a partir de 1999 a 2002 se desarrollaron encuestas puntuales del programa para el mejoramiento de las encuestas y la Medición de Condiciones de Vida MECOVI. Con la culminación del programa MECOVI, se inició la encuesta continua de hogares durante 2003 y 2004, incluyendo temas generales del programa MECOVI así como el de presupuestos familiares y un módulo sobre gestión de gobierno y democracia, propuesto en las reuniones de trabajo de la comunidad Andina de Naciones. Desde el 2005 a 2009 se retoma las encuestas puntuales de corte transversal, manteniendo como temática de exploración "las condiciones de vida de la población" donde los resultados se muestran a nivel nacional, nacional urbano, nacional rural.

A partir del 2011 se realiza la desagregación a nivel nacional, nacional urbano, nacional rural y departamental, donde los tamaños de muestra son calculados de manera independiente por departamento excepto para el departamento de Beni y Pando, donde las estimaciones deben de considerarse de manera conjunta.

Es en la gestión 2016, en cumplimiento de la Ley Nro. 548 y la política de fortalecimiento y actualización de las estadísticas e indicadores sobre trabajo de Niñas, Niños y Adolescentes, el Instituto Nacional de Estadística, adiciona a la encuesta de hogares la Encuesta de Niñas, Niños y Adolescentes que realizan una Actividad Laboral o Trabajan "EH-ENNA 2016" en coordinación con el Ministerio de Trabajo, Empleo y Prevención Social y el Ministerio de planificación y desarrollo.

Aspectos Metodológicos de la Encuesta de Hogares 2016 y ENNA.

2. Objetivos del diseño de la muestra.

Objetivo General.

La Encuesta de Hogares 2016 tiene el siguiente objetivo:

- El objetivo principal del diseño muestral de la EH-ENNA 2016, es determinar el tamaño de muestra, para recopilar información de los hogares, que permitan obtener resultados confiables.

Objetivos Específicos.

La Encuesta de Hogares 2016 tiene los siguientes objetivos específicos:

- Generar indicadores relacionados con el empleo, desempleo y otras características económicas confiables, que permitan conocer la evolución de la pobreza, el bienestar y la tendencia de las condiciones de vida de los hogares

del país, con resultados a nivel nacional, con desagregación nacional urbano y nacional rural y a nivel departamental, excepto Beni y Pando que será de manera conjunta.

- Generar indicadores de desocupación para las ciudades capitales de los departamentos del eje central (La Paz, Cochabamba, Santa Cruz).
- Generar la cuantificación de niñas, niños y adolescentes de 5 a 17 años de edad confiables, que realizan actividad laboral y las determinantes que inciden en la actividad laboral o trabajo en el marco de la Ley 548, mediante el módulo de Niñas, Niños y Adolescentes entre 5 y 17 años de edad, que realizan actividad laboral o trabajo 2016 aplicado en la Encuesta de Hogares.
- Identificar las condiciones demográficas y socio económicas de los NNA con actividad laboral.

3. Ámbito Geográfico.

El ámbito geográfico de la "Encuesta de Hogares-Encuesta de Niñas, Niños y Adolescentes que realizan una actividad laboral o trabajan 2016 son el área urbana y rural de todos los departamentos del país: Chuquisaca, La Paz, Cochabamba, Oruro, Potosí, Tarija, Santa Cruz, Beni y Pando.

4. Universo de estudio.

La investigación está dirigida al conjunto de hogares establecidos en viviendas particulares ocupadas dentro del territorio de Bolivia, en el periodo definido para la ejecución de la encuesta en la gestión 2016.

Se excluye de la investigación a la población residente en viviendas colectivas tales como hospitales, hoteles, cárceles, recintos militares, hogares de ancianos, unidades educativas, etc.; se incluyen a hogares que, forman un grupo independiente y además residen dentro de estos establecimientos, ejemplo: porteros, cuidadores, etc.

5. Unidades de observación, análisis.

De acuerdo a los objetivos de la investigación, la unidad de observación es el Hogar. Las unidades de análisis son: el hogar como unidad de consumo, los miembros del hogar en cuanto a sus características sociodemográficas, ocupacionales y de ingreso; y, las niñas, niños y adolescentes entre 5 y 17 años de edad.

6. Unidades de muestreo.

La unidad de muestreo, en su última etapa, es la vivienda particular ocupada, ésta tiene permanencia fija en el tiempo y espacio, característica que la habilita para ser utilizada como unidad de selección en el diseño muestral. La unidad de muestreo, de acuerdo a las áreas de estudio, se conforma de la siguiente manera:

- UPM - La unidad primaria de muestreo es el conglomerado de viviendas, en las ciudades amanzanadas puede estar compuesta por una o más manzanas y en área dispersa está compuesta por la comunidad o parte de ella.
- USM - La unidad secundaria de muestreo conforman las viviendas particulares dentro de las UPM.

7. Periodo de levantamiento de datos.

El levantamiento de información primaria corresponde al periodo del 19 de Octubre al 20 de Diciembre de 2016.

8. Marco Muestral.

Para la Encuesta de Hogares - Encuesta de Niñas, Niños y Adolescentes que realizan una actividad laboral o trabajan 2016, se utilizó el Marco Muestral - 2012.

La información del Marco Muestral - 2012 (MM), se encuentra basada en el Censo de Población y Vivienda de 2012 (CNPV-2012) y el Censo Nacional Agropecuario de 2013 (CNA-2013).

La estructura del Marco Muestral está definida por las siguientes variables:

Tabla 1: Estructura del Marco Muestral

Tipo de Variable	Nombre de la variable	Variable
Variables de División Política Administrativa	• I02_DEPTO	Departamento: código y descripción
	• I03_PROV • PROV	Provincia: código y descripción
	• I04_SECC • MUN	Sección municipal: código y descripción
Variables de organización municipal administrativa rural	• I04_SECC • MUN	Cantón: código y descripción
	• I06_CIUCOM • NOMBRE_COMUNIDAD	Comunidad: código y descripción
Variables de Organización Municipal administrativa urbana	• I04_SECC • MUN	Sección municipal: código y descripción
	• ID_UNICO	Manzana

VARIABLES PARA EL DISEÑO DE MUESTRA	<ul style="list-style-type: none"> • ID_UNICO_UPM • upm • NUM_UPM • ADAREA 	UPM dado por la unidad de cartografía UPM normalizado UPM dado por la unidad de Muestreo Manzana o comunidad en el Área
	<ul style="list-style-type: none"> • AREA_CATEG6 • URBRUR 	Estrato geográfico estadístico Área Urbana o Rural
	<ul style="list-style-type: none"> • nvivienda • vivOcup • vivDesoc 	Número de viviendas Totales Número de viviendas Ocupadas Número de viviendas Desocupadas
	<ul style="list-style-type: none"> • TOTPERS_VIV • TOT_HOMBRES_VIV • TOT_MUJERES_VIV 	Total de persona Total de personas hombres Total de personas mujeres

Fuente: Instituto Nacional de Estadística

Los planos y mapas cartográficos constituyen parte del marco muestral para la encuesta, mediante el cual se identifican las áreas correspondientes a las unidades primarias de muestreo seleccionadas y sirven de guía para ubicación en el terreno a las brigadas de encuestadores y supervisores.

Criterios de Conglomeración y Estratificación del Marco Muestral.

Variable de Conglomeración.

El Marco Muestral 2012 está constituido por 102.593 UPM, para la conformación de los conglomerados en cada uno de los departamentos se realizó la agrupación, división de las manzanas o comunidades considerando el número de viviendas en el conglomerado.

A continuación se observa el número de UPM a Nivel Nacional por Departamento.

Tabla 2: Distribución de UPM por Departamento

Departamento	Frecuencia	Porcentaje
Bolivia	102'593	100.0
Chuquisaca	5'362	5.2
La Paz	29'610	28.9
Cochabamba	18'041	17.6
Oruro	7'384	7.2
Potosí	8'519	8.3
Tarija	4'578	4.5
Santa Cruz	23'690	23.1
Beni	4'350	4.2
Pando	1'059	1.0

Fuente: Instituto Nacional de Estadística

Variable de Estratificación.

Con el propósito de obtener precisión en los estimadores en el sistema de encuestas a Hogares se realizó la estratificación en base a características socioeconómicas de manera que los conglomerados sean homogéneos dentro de los grupos y heterogéneos entre ellos.

El marco muestral 2012 posee una variables de estratificación que combina la estratificación geografía y la socioeconómica.

A continuación se describe del Estrato geográfico y estrato socioeconómico:

Tabla 3: Estrato geográfico

Estrato Geográfico	Descripción	Área geográfica
1	Ciudades Capitales más la ciudad de El Alto	1. Urbano
2	Ciudades Intermedias(10.000 y más habitantes)	
3	Centros Poblados (2.000 a 9.999 habitantes)	
6	Ciudad de El Alto	
4	Localidades con características amanzanadas (250 a 1.999 habitantes)	2. Rural
5	Áreas Dispersas (menos de 250 habitantes)	
7	Área Dispersa (mayor a 2000 habitantes)	

Fuente: Instituto Nacional de Estadística

Tabla 4: Estrato geográfico

Estrato Estadístico Socioeconómico	Descripción
1	Bajo
2	Medio Bajo
3	Medio Alto
4	Alto

Fuente: Instituto Nacional de Estadística

9. Características generales del diseño de muestra.

El diseño de la muestra para la EH-ENNA 2016, posee las características de un muestreo probabilístico, estratificado y bietápico.

- Probabilístico, porque la probabilidad de selección es conocida y diferente de cero para cada elemento objeto de estudio, lo cual permite realizar inferencia estadística al total de la población objetivo.
- Por conglomerados, debido a que las unidades objeto de estudio (viviendas/hogares) se encuentran agrupadas en áreas conformadas, con fines

de muestreo, en base a la información y cartografía del Censo de Población y Vivienda 2012, denominadas conglomerados.

- c) Estratificado, porque previo a la selección de conglomerados, estos, son clasificados con el propósito de mejorar la precisión de los estimadores y permite un mejor control en la distribución de la muestra.
- d) Bi-etapico, porque los elementos pertenecientes a la muestra se seleccionan en dos etapas:
 - a. Etapa 1: Selección de conglomerados (UPM), la selección se la realiza proporcional al tamaño de las viviendas.
 - b. Etapa 2: selección de viviendas dentro la UPM elegida, la selección de esta segunda etapa es de manera sistemática con arranque aleatorio.

Realizar desagregaciones mayores de los resultados estará sujeto a la valoración de los errores muestrales.

10. Nivel de desagregación de la información.

Se garantiza generar estimaciones de indicadores de pobreza, con errores muestrales admisibles, para los siguientes niveles de desagregación:

- Nacional
- Nacional Urbano
- Nacional Rural
- Departamental (En el caso de Beni y Pando se obtendrán estimaciones conjuntas)

Presentar indicadores de ocupación para las siguientes ciudades capitales: La Paz, Cochabamba y Santa Cruz.

Realizar desagregaciones mayores a los planteados en el diseño, estará sujeto a la valoración de los errores muestrales.

11. Selección de la Muestra.

Para la selección de la muestra en las diferentes etapas fue de la siguiente manera:

- En la primera etapa se utiliza la Selección Sistemática con probabilidad proporcional al tamaño (PPT) de viviendas.
- En la última etapa la selección de viviendas es sistemática con arranque aleatorio.

12. Cálculo del tamaño de la muestra.

Las variables que han sido, determinantes para el cálculo del tamaño de muestra apropiados para el estudio fueron: la **tasa de desempleo, pobreza y la condición de ocupación del niño o adolescente**. Los datos de referencia para el cálculo del tamaño de muestra fueron en base a indicadores de pobreza y empleo obtenidos con la información de la Encuesta de Hogares 2014 y la condición de ocupación del niño o adolescente de la Encuesta de Trabajo Infantil 2008.

Para garantizar la precisión en los estimadores, se procedió de manera independiente en cada uno de los dominios de estudio. En el caso de las ciudades capitales de los departamentos de La Paz, Cochabamba y Santa Cruz se empleó el coeficiente de variación de la tasa de desempleo; en el resto se consideró el indicador de pobreza.

Tabla 5: Coeficiente de variación (CV) del tamaño de muestra planificada

Departamento/Ciudad		Número de viviendas (muestra planificada 2016)	Coeficiente de Variación esperado (CV) en base a la Gestión 2014		
			CV de desempleo	CV de pobreza	CV de pobreza extrema
Nacional	Nacional	11.280	---	1,37	2,07
Chuquisaca	Chuquisaca	804		3,95	5,57
La Paz	La Paz (Ciudad)	1.020	7,5	5,76	9,17
	La Paz (Resto)	804		3,79	4,88
	La Paz (El alto)	1.020		5,33	9,00
Cochabamba	Cochabamba (Ciudad)	1.356	7,3	4,50	10,17
	Cochabamba (Resto)	804		4,62	6,96
Oruro	Oruro	792		6,47	10,08
Potosí	Potosí	732		3,85	4,68
Tarija	Tarija	1.032		5,42	10,07
Santa Cruz	Santa Cruz (Ciudad)	1.128	7,9	5,30	10,08
	Santa Cruz (Resto)	792		5,75	9,64
Beni	Beni	540		6,45	11,45
Pando	Pando	456		6,52	9,74

Fuente: Instituto Nacional de Estadística

13. Distribución de la muestra.

El cuadro siguiente presenta la distribución de la muestra por departamento y área geográfica, de la Encuesta de Hogares - Encuesta de Niñas, Niños y Adolescentes que realizan una actividad laboral o trabajan 2016:

Tabla 6: Distribución de la muestra de viviendas y UPM por departamento y área geográfica

	Número de viviendas de la Muestra					Numero de UPM de la Muestra				
	Total	Capital/ Conurbación	Ciudad Intermedia	Centro Poblado	Rural	Total	Capital/ Conurbación	Ciudad Intermedia	Centro Poblado	Rural
NACIONAL	11.280	6.156	1.752	648	2.724	940	513	146	54	227
Chuquisaca	804	360	0	96	348	67	30	0	8	29
La Paz	2.844	2.040	348	60	396	237	170	29	5	33
Cochabamba	2.160	1.356	336	108	360	180	113	28	9	30
Oruro	792	372	48	36	336	66	31	4	3	28
Potosí	732	144	120	72	396	61	12	10	6	33
Tarija	1.032	408	300	36	288	86	34	25	3	24
Santa Cruz	1.920	1.128	384	144	264	160	94	32	12	22
Beni	540	120	216	60	144	45	10	18	5	12
Pando	456	228	0	36	192	38	19	0	3	16

Fuente: Instituto Nacional de Estadística

14. Factores de expansión.

El factor de expansión es el inverso de la probabilidad de selección de la vivienda. El factor de expansión final, tiene el ajuste de la no-respuesta, considerando los resultados de incidencias de campo y el total poblacional proyectado para ese año, por esta razón los factores de expansión deben ser calculados una vez realizada la encuesta.

Las fórmulas con las que se calculará las probabilidades, vienen dadas por:

$$P(viv_{ijh}) = \left(\frac{A_h N_{jh}}{N_h} \right) \cdot \left(\frac{VE_{jh}}{VOE_{jh}} \right) = \left(\frac{A_h N_{jh}}{N_h} \right) \cdot \left(\frac{c}{VOE_{jh}} \right)$$

Donde:

$P(viv_{ijh})$: Probabilidad de seleccionar la i-ésima vivienda de la j-ésima UPM, del estrato h de una UPM del área amanzanada.
A_h	: Número de UPM's seleccionadas del estrato h.
N_h	: Número de viviendas del estrato h.
N_{jh}	: Número de viviendas en la j-ésima UPM del estrato h.
VE_{jh}	: Total de viviendas elegidas (viviendas elegidas ocupadas) en la j-ésima UPM, del estrato h, es decir, $VE_{jh} = c$
VOE_{jh}	: Total de viviendas objeto de estudio (viviendas ocupadas) listadas en la

	j-ésima UPM, del estrato h.
c	: Número fijo de viviendas seleccionadas en la última etapa, $c = 12$.

El procedimiento habitual indica que a los factores de expansión se aplique la corrección de las proyecciones de población, para que los mismos expresen el total de la población boliviana excepto las personas que habitan en viviendas colectivas que no son objeto de estudio en la encuesta.

15. Ajustes a los Factores de Expansión.

Por Incidencias Finales.

Los factores de expansión se corrigen por incidencias de campo, es decir, que se utiliza los valores de la muestra efectiva por UPM.

El ajuste se realiza multiplicando el factor básico por la relación: $\frac{c}{V_{jh}}$

V_{jh} : es la muestra efectiva, número de viviendas con encuestas válidas en la j-ésima UPM del estrato h.

c : es el número de viviendas planificadas para ser encuestadas = 12

Por Proyecciones de Población al 2016.

Para estimar el Total de hogares y el Total de población, se corrige el factor de expansión con las proyecciones de población que genere el INE para la gestión 2016, esta información por cada dominio de estudio.

La corrección se realiza multiplicando por: $\frac{P_h}{\hat{P}_h}$

Donde:

P_h	: Población en el h-ésimo dominio, según la proyección.
\hat{P}_h	: Población en el h-ésimo dominio, a la que expande la encuesta.

Los factores de expansión finales estarán dados por:

$$F_{jh} = \frac{1}{P(Viv_{ijh})} \times \frac{c}{V_{jh}} \times \frac{P_h}{\hat{P}_h}$$

16. Estimadores y Errores de Muestreo.

Se ha señalado que el diseño empleado es complejo, por lo cual se requiere utilizar fórmulas especiales que consideran los efectos de la estratificación y la conglomeración al momento de obtener los resultados de la encuesta.

Para calcular los estimadores, es fundamental que la base de datos de la encuesta esté expandida, vale decir, que se empleen los factores de expansión. Para el cálculo de los intervalos de confianza, coeficiente de variación, error estándar, efecto de diseño, entre otros, es necesario aplicar el diseño de la muestra antes de utilizar la base de datos en la generación de indicadores.

Se recomienda valerse de algún software estadístico que permita realizar el cálculo de los errores muestrales aplicando el diseño de muestra empleado; entre los programas estadísticos más utilizados para realizar este procedimiento están el SPSS.

17. Incidencia de campo.

Una vez finalizado el operativo de campo se procede a calcular la incidencia de campo en los diferentes departamentos.

A continuación se presenta la tabla que refleja las incidencias de campo de la encuesta por departamento y área (urbana y rural):

Tabla 8: Incidencia de campo por departamento

Departamento / Ciudad	Muestra planificada			Muestra Ejecutada			Incidencia de campo (%)		
	Total	Urbano	Rural	Total	Urbano	Rural	Total	Urbano	Rural
NACIONAL	11280	8568	2712	11062	8373	2689	98.1	97.7	99.2
Chuquisaca	804	468	336	797	463	334	99.1	98.9	99.4
La Paz	2844	2448	396	2759	2364	395	97	96.6	99.7
Cochabamba	2160	1800	360	2121	1761	360	98.2	97.8	100
Oruro	792	456	336	778	444	334	98.2	97.4	99.4
Potosí	732	336	396	730	334	396	99.7	99.4	100
Tarija	1032	744	288	1010	723	287	97.9	97.2	99.7
Santa Cruz	1920	1656	264	1887	1630	257	98.3	98.4	97.3
Beni	540	396	144	537	394	143	99.4	99.5	99.3
Pando	456	264	192	443	260	183	97.1	98.5	95.3