2020 censo de Población y Vivienda

Design of the census sample

Census of population and household

INEGI

Instituto Nacional de Estadística y Geografía

Census of population and household 2020

Design of the census sample

Complementary works published by INEGI on the subject:
Design of the census sample 2010.

Catalogación en la fuente INEGI:

304.601072 Censo de Población y Vivienda (2020).
Censo de Población y Vivienda 2020 : diseño de la muestra censal / Instituto Nacional de Estadística y Geografía.-- México : INEGI, c2021.

vii, 14 p.

1. México - Población - Censos, 2020 - Metodología. 2. México - Población - Censos, 2020 - Metodología. I. Instituto Nacional de Estadística y Geografía (México).

Conociendo México
800 111 4634
www.inegi.org.mx atencion.usuarios@inegi.org.mx

INEGI Informa @INEGI_INFORMA

Registration pending

2021, **Instituto Nacional de Estadística y Geografía**
Edificio Sede
Avenida Héroe de Nacozari Sur 2301
Fraccionamiento Jardines del Parque,
20276 Aguascalientes, Aguascalientes,
Aguascalientes, entre la calle INEGI,
Avenida del Lago y Avenida Paseo de
las Garzas.

PRESENTATION

The **Instituto Nacional de Estadística y Geografía (INEGI)** in the exercise of the attributions that the Law of the National System of Statistical and Geographical Information confers on it, carried out the 2020 Population and Household Census.

As part of the methodological documentation that accompanies the final results of this statistical project, the Design of the 2020 census sample is presented in order to make information available to users that allows them to have an overview of the methodological bases based on which the design of the census sample was carried out.

In this way, the Institute, as part of the National System of Statistical and Geographic Information, fulfills the function of providing society and the State with quality, pertinent, truthful, and timely information to contribute to the development of Mexico.

INDEX

INTRODUCTION

As of the XII General Population and Household Census 2000, the information was collected using two types of questionnaires: basic and extended. In the same way, during the 2020 Population and Household Census, these two types of questionnaires were applied. With the Expanded Questionnaire, around 4 million households were registered in the national territory.

The Expanded Questionnaire was applied in selected households using probabilistic criteria and includes the questions of the Basic Questionnaire and additional questions that delve into some aspects of the characteristics of the households and their inhabitants, so that all households were surveyed with the Basic Questionnaire and some, those selected, answered additional questions.

This document describes all the technical aspects related to the 2020 census sample, the sampling frame, the sample size, the sampling scheme, as well as the obtaining of estimators.

It is vitally important that the users of the information in the census sample consider the design of the sample to obtain the estimators they require.

The sample design allows generating estimators with acceptable precision and confidence for the vast majority of the additional variables included in the Extended Questionnaire and for municipal disaggregation levels, however, whenever an estimator is obtained, precision and confidence must be considered, before making a decision based on it.

I. OBJETIVES

The 2020 census sample has the general objective of providing information on the characteristics of the households and their inhabitants, with measurable precision and confidence for rates, averages and proportions, at the national level, by federative entity, by municipality or territorial demarcation of Mexico City and for each of the localities with 50 thousand or more inhabitants.

In particular, the sample seeks to provide information for the following levels of geographic breakdown:

- National.
- For each of the 32 states.
- Each of the 2 469 municipalities and territorial demarcations in the country.
- Each of the localities with 50,000 or more inhabitants.

Additionally, the sample allows the generation of indicators at the minimum level of geographic disaggregation allowed by the confidentiality criterion for the municipalities that meet the following criteria, based on the results of the 2015 Intercensal Survey:

- Municipalities with less than 1,300 inhabited private households.
- Municipalities with a very high degree of social backwardness[1] and with 10,000 or less inhabited private households.
- Municipalities with a lower Human Development Index[2] (HDI less than or equal to 0.55) and with 10,000 or less inhabited private households.
- Municipalities with 40% or more of their population aged 3 years and over who speak an indigenous language and who do not speak Spanish.
- Municipalities with 40% or more of their population considered Afro-descendant.

[1] https://www.coneval.org.mx/Medicion/IRS/Paginas/Indice_Rezago_Social_2015.aspx
[2] https://www.mx.undp.org/content/mexico/es/household/library/poverty/informe-de-desarrollo-humano-municipal- 2010-2015--transformando-.html

II. SAMPLE DESIGN

The design of the 2020 census sample is stratified by conglomerates and was selected in a single stage, that is, complete areas, blocks or localities were selected, in which the Expanded Questionnaire was applied in all inhabited private households.

1. POPULATION UNDER STUDY

The population under study is made up of inhabited private households and their habitual residents within the national territory.

2. STUDY DOMAINS

The following study domains were defined:

- National.
- State.
- State with breakdown into four sizes of locality:
  - Less than 2,500 inhabitants.
  - From 2,500 to 14,999 inhabitants.
  - From 15,000 to 49,999 inhabitants.
  - or 50,000 and more inhabitants.
- Each of the 2 453 municipalities of the 31 states and each of the 16 territorial demarcations of Mexico City.
- The localities of 50 thousand or more inhabitants[3] of the country existing in 2019.

It is important to note that during 2020 the formation of four municipalities was consolidated (Honduras de la Sierra, Chiapas; Seybaplaya, Campeche; Hueyapan, Morelos and San Quintín, Baja California), so they were not considered in the design of the sample. Likewise, in the municipality of La Magdalena Tlaltelulco, Tlaxcala, it was not possible to carry out the collection of information. Therefore, only the estimate of household and population is presented for the municipalities of La Magdalena Tlaltelulco, Tlaxcala; Honduras de la Sierra, Chiapas; Seybaplaya, Campeche. For the municipalities of Hueyapan, Morelos and San Quintín, Baja California, the results should be carefully reviewed before being used, considering the standard error and the confidence intervals. Next, the characteristics of the design are presented considering the 2,465 municipalities included in the sampling frame used for the design of the census sample.

[3] After the 2020 Census, 18 localities were identified that exceed 50 thousand inhabitants and that were not considered in the design because before the census event they did not exceed that population. The results for these municipalities should be carefully reviewed before being used, considering the standard error and the confidence intervals.

The 2,465 municipalities existing in 2019 were grouped as follows:

GROUPING OF MUNICIPALITIES	
Municipalities that are counted with Extended Questionnaire (certainty)	Municipalities with less than 1,300 private households.
	Municipalities with populations of interest (HDI, social backwardness, indigenous language, and Afro-descendants).
Rest of the municipalities	Without localities of more than 50 thousand inhabitants.
	With locations of more than 50 thousand inhabitants.

3. SAMPLE FRAME

The sampling frame is made up of the AGEB, town and block catalogs of the Geostatistical Framework 2019, as well as the statistics associated with each area.

The framework for the 2020 Population and Household Census was based on the information obtained from the 2010 Population and Housing Census, the Intercensal Survey 2015, the Single Registry of Housing 2012 (RUV 2012) and the Economic Censuses 2014 and 2019.

Some areas of the cartographic catalog were not considered within the sampling frame because they did not include the population under study:

- Localities without a habitual resident population in the last three events (Censuses and Intercensal Survey 2015).
- Areas in which there are only collective households.
- Areas by type (under bridge, median, economic, roundabout, park or garden, etc.).
- Areas that only have services (schools, buildings, bodies of water, etc.).

4. SAMPLING SCHEME

The sampling scheme for the census sample is stratified, by conglomerates and in a single selection stage, the primary sampling units correspond to blocks or complete localities, that is, in each of the selected areas all inhabited private households are registered using the Extended Questionnaire. For the selection of the areas, a simple random sampling without replacement was used for each defined stratum.

Because more details are required from some sectors of the population, it was determined to survey all municipalities that met at least one of the following criteria with the Expanded Questionnaire:

MUNICIPALITIES TO BE CENSUSED WITH EXPANDED QUESTIONNAIRE[4]		786
Estimation from the Intercensal Survey 2015	With less than 1 300 inhabited private households.	683
	Municipalities with 40% or more of their population aged 3 years and over who speak an indigenous language and who do not speak Spanish.	22
	Municipalities with 40% or more of its population considered Afro-descendant.	12
Municipalities with a very high degree of social backwardness in 2015 with 10,000 or less inhabited private households.		169
Municipalities with a Human Development Index (HDI) less than 0.55 and with 10,000 or less inhabited private households.		131

4.1 Sampling units

The primary sampling units are entire geographic areas, whether they are blocks or towns. In this sense, the sampling units to be selected (blocks or localities) vary according to the sample design applied in each one of them.

- Localities with 250 or more inhabited private households. The Primary Sampling Unit (PSU) is generally made up of blocks.
- The localities with less than 250 households, the PSUs are the localities themselves.

[4] Some municipalities meet more than one inclusion criterion, so the sum of them does not correspond to the total number of municipalities surveyed with an expanded questionnaire.

4.2 Stratification

The 2,453 municipalities of the 31 states and the 16 territorial demarcations of Mexico City were classified into two groups: “municipalities selected with certainty” and the rest of the municipalities. In the case of the selected municipalities with certainty, all their areas were included in the sample with probability one, therefore, all the households in these areas were surveyed applying the Expanded Questionnaire.

In the rest of the municipalities, complete areas were selected (blocks or localities with less than 250 inhabited private households). Within each sampled municipality, strata were defined by locality size as follows.

- Less than 50,000 inhabitants:
  - Less than 2,500 inhabitants.
  - 2,500 to 4,999 inhabitants.
  - 5,000 to 14,999 inhabitants.
  - 15,000 to 49,999 inhabitants.
- For localities with 50 thousand and more inhabitants, each of them formed a stratum.

5. SAMPLE SIZE

To guarantee that the estimates obtained from the 2020 Population and Household Census sample have acceptable precision and confidence, it is necessary that the sample size defined is sufficient for each parameter of interest.

In general, to obtain the sample size, a confidence level of 90%, a relative error of 0.2, a design effect of 1.44 and a response rate of 90%, as the case may be, were considered. With these parameters, the sample size distributed as follows:

- Municipalities with 1,300 households or less or with certainty were surveyed.
- In municipalities with more than 1,300 households, at least 2,057 households and if the municipality includes towns with more than 200,000 inhabitants, at least 4,470 households.

The sample sizes were obtained by the following expression:

$$n = \frac{Z^2}{r^2} x \frac{q}{p} x \frac{DEFF}{1 - TNR}$$

Where:
p = Proportion to be estimated
q = 1 - pp
r = Maximum relative error
z = Confidence level at 90%
TNR = Non-response rate
$DEFT$ = Design effect

Additionally, a finite population adjustment was made to the required sample size.
The expression is the following:

$$n' = \frac{n}{1 + \frac{n}{N}}$$

Where:
nn ′ = Sample size adjusted for finite population
nn = Estimated sample size
NN = Total households

Fixed sample sizes guarantee municipal estimates with acceptable precision and confidence (according to fixed parameters) for close proportions to 0.01 or greater.

5.1 Affixation of the sample within the strata

In the case of the selected municipalities with certainty, it is not necessary to carry out some type of allocation of the sample, but for the rest of the municipalities the allocation of the sample size was carried out in proportion to the total number of private households inhabited in the municipality.

$$n_h = \frac{N_h}{N_{mun}} n_{mun}$$

Where:
N_h = Total households in stratum h
N_{mun} = Total households in the municipality
n_{mun} = Sample size by municipality
n_h = Sample size to be selected by stratum

5.2 Expansion factors

According to the sampling scheme, the probability of selection of each of the PSU (blocks or localities) within each stratum is given by the following formula:

$$P_{hj} = \frac{n_h}{N_h}$$

n_h = Number of PSUs in the sample in the *h - th* stratum
N_h = Total PSU in the *h - th* stratum

In this way, the expansion factor for each of the households (and each of its residents) of the *j-th* PSU of the *h-th* stratum is given by:

$$F_{hj} = \frac{1}{P_{hj}} = \frac{N_h}{n_h}$$

For each PSU, the expansion factors were multiplied by an additional factor called the non-response adjustment.

The sample design is not self-weighted, that is, each sampled household represents a different number of households. The expansion factors are different between areas and the non-response rate is also differential, and therefore, any indicator or tabulation based on the census sample requires the use of expansion factors.

5.3 Calculation of estimators

The estimator is a statistic obtained from the data collected from only a part of the population (sample), it is used as an approximation of an unknown parameter of the population.

For the 2020 census sample, the estimator of the total was calculated using the method proposed by Horvitz-Thompson, from which the estimators of the averages, proportions and ratios are derived.

There are various statistical packages with which the estimators can be obtained, for the case of the 2020 census sample, the SAS 9.4 software was used.

5.4 Variance estimators

The variance of the linear estimators presented in the 2020 census sample tabulations was obtained using the Taylor series method. As for the 2010 Census and the 2015 Intercensal Survey, this method was selected in such a way that users who exploit the census sample obtain the same results using some of the most widely used statistical packages.

Likewise, it was decided that the estimation of the variance of the non-linear indicators should be approximated by resampling techniques. Both for the 2010 census sample, the 2015 Intercensal Survey and for the 2020 census sample, it was decided to use the technique known as Jackknife.

5.5 Standard error

The standard error is calculated with the sample data and is defined as the square root of the statistical variance of the estimator. It measures the degree of precision with which the estimator approximates the result that would have been obtained if the entire population under study had been interviewed under the same conditions. It is also known as sampling error or sampling error. By itself it is difficult to interpret, so the coefficient of variation and the confidence interval are used.

5.6 Coefficient of Variation (CV)

For the purposes of presenting the results of the 2020 census sample, the coefficient of variation is defined as the percentage variation of the standard error to the central estimate. It is the quotient between the standard error and the estimator multiplied by 100. It is also known as the relative error. It is a uniform criterion for determining the precision of an estimator.

5.7 Confidence interval

The confidence interval is one of the most effective forms of statistical inference. From this it is possible to know how correct the estimate is made, since it indicates that the true value is between the lower limit and the upper limit of the interval, with a predetermined probability or confidence level α. For the census sample, 90% confidence was used.

5.8 Total estimators

The estimator of the total population for a given study domain d is the sum of the estimators of the total of the strata:

$$\hat{Y}_d = \sum_{h=1}^{H} \hat{Y}_h$$

$$\hat{Y}_d = \sum_{h=1}^{H} \sum_{j=1}^{n_h} F_{hj} \sum_{k=1}^{M_{hj}} y_{hjk}$$

Where:

Y_d= Estimated total of the characteristic of interest for domain *d*

Y_h= Estimated total of the characteristic of interest in stratum *h* of domain *d*

F_{hj}= Expansion factor of the *j - th* PSU in the *h - th* stratum

y_{hjk}= Value of the characteristic of interest in the *k - th* house of the *j - th* PSU in the *h - th* stratum

H= Number of strata in domain *d*

n_h= Number of PSUs selected in the *h - th* stratum

M_{hj}= Number of units in sample within the *j - th* PSU in the *h - th* stratum

By having independent samples per stratum, the variance of the estimator of the total it is equal to the sum of the variances of the estimators of the total strata.

$$var(\hat{Y}_d) = \sum_{h=1}^{H} var(\hat{Y}_h) = \sum_{h=1}^{H} N_h^2 \left(1 - \frac{n_h}{N_h}\right) \frac{S_h^2}{n_h}$$

Where $\hat{S}_h^2 = \sum_{j=1}^{n_h} \frac{(y_{hj} - \hat{\bar{y}}_h)^2}{n_h - 1}$ is the variance of stratum *h*.

And the estimator of the variance of the estimator of the total, for a given domain, is given by:

$$\widehat{var}(\hat{Y}_d) = \sum_{h=1}^{H} \widehat{var}(\hat{Y}_h) = \sum_{h=1}^{H} N_h^2 \left(1 - \frac{n_h}{N_h}\right) \frac{\hat{S}_h^2}{n_h}$$

Where:

$$\hat{S}_h^2 = \sum_{j=1}^{n_h} \frac{\left(y_{hj} - \hat{\bar{y}}_h\right)^2}{n_h - 1}$$

Note that $\hat{S}_h^2$ is the same expression as S_h^2 , although the first refers to the values of the sample and the second to the population value of each stratum.

Variance $var(\hat{Y}_d)$ measures the variability of the estimators in the domain and $\widehat{var}(\hat{Y}_h)$ in the different strata.

The factor $(1 - \frac{n_h}{N_h})$ called Finite Population Correction Factor (FCP) in general tends to 1 for the various study domains of the 2020 census sample, so it was not considered for the calculation of the variance estimators; Furthermore, not involving the FCP simplified the calculations.

Then the standard error associated with $\hat{Y}_d$ is given by:

$$EE_{\hat{Y}_d} = \sqrt{\widehat{var}(\hat{Y}_d)}$$

And the estimate of its coefficient of variation (CV) is:

$$CV = \frac{EE_{\hat{Y}_d}}{\hat{Y}_d} \times 100$$

This coefficient is also known as the relative error.

The confidence interval presented in the tables was constructed with a confidence of 90% and is given by:

$$P[\hat{Y}_d - 1.645\sqrt{\widehat{var}(\hat{Y}_d)} < Y_d < \hat{Y}_d + 1.645\sqrt{\widehat{var}(\hat{Y}_d)}] \approx 0.90$$

5.9 Average estimators

$$\hat{\bar{Y}}_d = \frac{\hat{Y}_d}{\hat{M}_d} = \frac{\hat{Y}_d}{\sum_{h=1}^{H}\sum_{j=1}^{nh} F_{hj}}$$

$\hat{M}_d$ = Estimation of the total in domain *d*.

And the variance of $\hat{\bar{Y}}_d$ is:

$$var\left(\hat{\bar{Y}}_d\right) = \sum_{h=1}^{H} var\left(\frac{\hat{Y}_d}{\hat{M}_d}\right) = \sum_{h=1}^{H} \frac{1}{M_d^2} var\left(\hat{Y}_d\right) = \sum_{h=1}^{H} \frac{N_h^2}{M_d^2}\left(1 - \frac{n_h}{N_h}\right)\frac{S_h^2}{n_h}$$

If what is required is to estimate the variance of the sample mean of a particular characteristic, the formula is used:

$$\widehat{var}\left(\hat{\bar{Y}}_d\right) = \sum_{h=1}^{H} \widehat{var}\left(\frac{\hat{Y}_d}{\hat{M}_d}\right) = \sum_{h=1}^{H} \frac{1}{\hat{M}_d^2} \widehat{var}\left(\hat{Y}_d\right) = \sum_{h=1}^{H} \frac{N_h^2}{\hat{M}_d^2}\left(1 - \frac{n_h}{N_h}\right)\frac{\hat{S}_h^2}{n_h}$$

Where the standard error for the mean is given by:

$$EE_{\hat{\bar{Y}}_d} = \sqrt{\widehat{var}\left(\hat{\bar{Y}}_d\right)}$$

And the estimate of the coefficient of variation (CV) is:

$$CV = \frac{EE_{\hat{\bar{Y}}_d}}{\hat{\bar{Y}}_d} \times 100$$

Similarly, to estimate the mean in domain d, the approximate 90 percent confidence interval can be constructed:

$$P[\hat{\bar{Y}}_d - 1.645\sqrt{\widehat{var}\left(\hat{\bar{Y}}_d\right)} < \bar{Y}_d < \hat{\bar{Y}}_d + 1.645\sqrt{\widehat{var}\left(\hat{\bar{Y}}_d\right)}] \approx 0.90$$

5.10 Proportion estimators

If what is required to estimate is the proportion P of elements of the population that have a certain characteristic, the equivalences given by:

$$\bar{Y}_d = P_d, \bar{y}_d = p_d$$

So the estimator for the proportion is:

$$\hat{P}_d = \frac{\hat{Y}_d}{\hat{M}_d} = \frac{\hat{Y}_d}{\sum_{h=1}^{H}\sum_{j=1}^{nh} F_{hj}}$$

With an estimated variance given by the following expression:

$$\widehat{var}\left(\hat{\bar{Y}}_d\right) = \widehat{var}\left(\hat{P}_d\right) = \sum_{h=1}^{H} \frac{N_h^2}{\hat{M}_d^2}\left(1 - \frac{n_h}{N_h}\right)\frac{p_d(1-p_d)}{n_h}$$

Where the standard error of the proportion is obtained with:

$$EE_{\hat{P}_d} = \sqrt{\widehat{var}\left(\hat{P}_d\right)}$$

And the estimate of its coefficient of variation (CV) is given by:

$$CV = \frac{EE_{\hat{P}_d}}{\hat{P}_d} \times 100$$

Unlike the previous ones, *logit* confidence intervals are used to calculate the confidence interval for a percentage or proportion.

The *logit* confidence limits P_L and P_U are calculated as:

$$Limite\ inferior\ P_L = \frac{e^{\lambda_L}}{(1+e^{\lambda_L})}$$

$$Limite\ superior\ P_U = \frac{e^{\lambda_U}}{(1+e^{\lambda_U})}$$

Where:

$$\lambda_L = \log\left(\frac{\hat{p}}{1-\hat{p}}\right) - z_{df,\frac{\alpha}{2}} \times \frac{EE_{\hat{p}}}{\left(\frac{\hat{p}}{1-\hat{p}}\right)}$$

$$\lambda_U = \log\left(\frac{\hat{p}}{1-\hat{p}}\right) + z_{df,\frac{\alpha}{2}} \times \frac{EE_{\hat{p}}}{\left(\frac{\hat{p}}{1-\hat{p}}\right)}$$

5.11 Ratio estimators

Suppose that we want to calculate the estimator of the ratio of the characteristics Y and X, with $R = Y/X$, which is given in the following expression:

$$\hat{R}_d = \frac{\hat{Y}_d}{\hat{X}_d} = \frac{\sum_{h=1}^{H}\sum_{j=1}^{n_h} F_{hj} \sum_{k=1}^{M_{hj}} y_{hjk}}{\sum_{h=1}^{H}\sum_{j=1}^{n_h} F_{hj} \sum_{k=1}^{M_{hj}} x_{hjk}}$$

Where the estimate of its variance is obtained:

$$\widehat{var}(\hat{R}_d) = \sum_{h=1}^{H} \widehat{var_h}(\hat{R}_d)$$

Therefore, the standard error of the ratio is estimated by:

$$EE_{\hat{R}_d} = \sqrt{\widehat{var}(\hat{R}_d)}$$

And the estimate of its coefficient of variation (CV) is calculated:

$$CV = \frac{EE_{\hat{R}_d}}{\hat{R}_d} \times 100$$

For the ratio estimator in domain *d*, the approximate 90% confidence interval can be constructed with the following expression:

$$P[\hat{R}_d - 1.645\sqrt{\widehat{var}(\hat{R}_d)} < R_d < \hat{R}_d + 1.645\sqrt{\widehat{var}(\hat{R}_d)}] \approx 0.90$$

With a design effect (DEFF) associated with the proportion, it will be given by:

$$DEFF_{\hat{R}_d} = \frac{\widehat{var}(\hat{R}_d)}{\hat{V}_{MAS}(\hat{R}_d)}$$

Where $\hat{V}_{MAS}(\hat{R}_d)$ is the estimated variance of the estimator of the ratio considering a simple random sampling (MAS).

5.12 Estimation of variance by the Jackknife method

The Jackknife estimation method is part of the resampling techniques, which are used mainly when it is required to estimate the variance of non-linear estimators. The estimate calculated with this method for the presentation of the results of the 2020 census sample was the Global Fertility Rate.

Said method is based on making estimates eliminating, each time, a different PSU from the complete sample to create replicas; therefore, we have the same number of estimates as the number of PSUs in the sample. The Jackknife variance estimator for the estimator $\hat{\theta}$ is calculated by the expression:

$$\hat{V}_{JK}(\hat{\theta}) = \sum_{r=1}^{R} \alpha_r (\hat{\theta} - \hat{\theta}_r)^2$$

Where:

θ = Estimator calculated from the complete sample

θ_r= Estimator calculated in the *r - th* replication using the replicated $F_h^{(r)}$, given by the expression:

$$F_h^{(r)} \begin{cases} F_h & \text{If the PSU eliminated in the } r \text{ - } th \text{ replication does not belong to the donor stratum} \\ \dfrac{F_h}{\alpha_r} & \text{If the PSU eliminated in the } r \text{ - } th \text{ replica belongs to the donor stratum} \end{cases}$$

α_r = Jackknife coefficients, $\alpha_r = \dfrac{n_{hr}-1}{n_{hr}}$

h_r= Donor stratum or stratum from which the PSU is eliminated in the *r - th* replication

n_{hr}= Number of PSU in donor stratum

La estimación de la varianza Jackknife tiene $R - H$ grados de libertad.

5.13 Consideration to the Jackknife method

The method assumes that there is more than one PSU per stratum in the sample, otherwise a special treatment must be given. This consists of not considering for the elimination of PSU the strata that have only one PSU (the case of surveyed areas, for example); however, they should be considered for the calculation of the $\hat{\theta}_r$ as well as for $\hat{\theta}$. If the census areas were not considered in the calculations, to avoid the problem of strata with a single PSU in the sample, biased estimates would be obtained.

Conociendo
México

800 111 46 34
www.inegi.org.mx
atencion.usuarios@inegi.org.mx

INEGI Informa

Translation by Sistemas de Información Geográfica SA de CV, Esri oficial distributor for México