# RECOMMENDED SAMPLE DESIGN FOR 1999 *ENQUÊTE PERMANENTE AUPRÈS DES MÉNAGES*

DAVID J. MEGILL
U.S. BUREAU OF THE CENSUS

JANUARY 1999

**RECOMMENDED SAMPLE DESIGN FOR 1999 *ENQUÊTE PERMANENTE AUPRÈS DES MÉNAGES***

TABLE OF CONTENTS

LIST OF TABLES

## 1. Background

The *Direction des Statistiques de Ménages* (DSM) of the *Institut National de la Statistique* (INSTAT) previously conducted the *Enquête Permanente auprès des Ménages* (EPM) in 1994 and 1997. This national household survey, sponsored by the World Bank, collects data on many socioeconomic characteristics including household income and expenditures. The survey methodology is consistent with the World Bank's multinational Priority Survey. The DSM is planning to conduct the data collection for the 1999 EPM from May to July.

The sample design for the 1999 EPM will be based on the master sampling frame developed by the *Direction de la Démographie et Statistiques Sociales* (DDSS) of INSTAT from the data and cartographic materials from the 1993 Madagascar Census of Population and Housing. The methodology for the master sample design is described in the report on «Preliminary Recommendations for Designing the Master Sample for the Madagascar Household Survey Program" (Megill, April 1996), which is also available in a French version, «*Recommandations Préliminaires pour la Conception de l'Échantillon-Maitre pour le Programme d'Enquêtes auprès des Ménages a Madagascar*.» The master sampling frame was used to select the sample segments for the 1997 EPM and well as those for the 1997 *Enquête Démographique et de Santé* (DHS). The tabulation of measures of precision (standard errors, coefficients of variation and design effects) from these surveys will be useful in determining an approximately optimal sample design for the 1999 EPM.

The sample design for the 1999 EPM will be based on the survey objectives as well as resource constraints. The purpose of this report is to document the proposed sampling methodology and estimation procedures for this survey. These recommendations were developed jointly with Jean Razafindravonona, Director of the DSM, and his staff.

The questionnaire for the 1999 EPM will include a few additional questions on family planning and child immunization. Considerations for the precision of the survey results for these characteristics are also discussed in the report.

It is important that the Statisticians at DSM staff receive more training in sampling and estimation methods in order to build their capability to implement appropriate weighting procedures and tabulate measures of precision for the survey results, as well as design efficient samples for future household surveys. Therefore it is recommended that two Statisticians from the DSM attend the 4-week course on «Sampling and Statistical Methods» to be held at U.S. Census Bureau from June 7 to July 2, 1999.

## 2. Survey Objectives and Resource Considerations

The survey objectives are generally defined in terms of the characteristics being measured, the domains of analysis (that is, the level of disaggregation in the survey tables), and the level of accuracy required for the survey results. The sample size for a particular survey is determined by

the level of precision required for the survey estimates for each domain, as well as by the resource and operational constraints.

The questionnaire for the 1999 EPM will be similar to that for the 1997 survey, with socioeconomic characteristics related to health, education, employment, migration, housing, agriculture, household income and expenditures. A few family planning and child immunization questions will be added, and other sections may be modified slightly. The geographic domains defined for the survey tabulations are national, Capital (Antananarivo), other large cities (*grands centres urbains*, or GCUs), other urban and rural, and the six individual *faritany* (Antananarivo, Fianarantsoa, Toamasina, Mahajanga, Toliary and Antsiranana).

The 1999 EPM results will be compared to those from the 1997 and 1994 surveys, in order to examine trends in the socioeconomic characteristics over time. This also has implications for the 1997 EPM sample design, since it is ideal to have some overlap in the sample areas to improve the estimates of trends.

For the family planning and child immunization questions, it is also ideal to have 1999 EPM results comparable to the corresponding estimates from the 1997 DHS. In comparing the geographic domains for these two surveys, the only difference is that for the EPM the other urban domain is divided into separate strata for large cities (GCU) and secondary urban centers (*centres urbaines sécondaires*, or CUS). These can be combined into the DHS categories for the survey tables on family planning and health characteristics which will be compared to the DHS results. However, this also has implications for the sample allocation, which is discussed later in this report.

In terms of resource constraints, the DSM has indicated that the current survey budget will limit the sample size to close to 5,000 households, given that the questionnaire for the 1999 EPM will be more comprehensive than that for the 1997 survey, and different persons in each sample household will have to be interviewed. As a result, it is important that the sample allocation be designed optimally to meet the most important survey objectives. The tabulation of standard errors from the 1997 EPM data were used in determining the most effective sample allocation and to estimate the approximate level of precision which can be expected for the different domains of analysis based on the proposed sample design for the 1999 EPM. At the same time, it is important to concentrate on the quality and operational control for the data collection and other survey operations in order to reduce the nonsampling errors in the survey data.

## 3. Sample Design for the 1997 EPM

Given that objectives of the 1999 EPM are similar to those for the 1997 EPM, it is important first to review the sample design for the 1997 survey. By examining measures of precision from the 1997 EPM data, it is possible to determine whether it is necessary to modify some aspects of the previous sample design to improve the sampling efficiency for the 1999 EPM.
The preliminary sample design for the 1997 EPM was recommended by Dr. Chris Scott, in his

report on «*Enquête Prioritaire a Madagascar: Plan de Sondage Proposé*» (November 1996). A stratified multistage sample was used for the 1997 EPM, based on the master sampling frame. The master sample is stratified by *faritany* and *milieu* (GCU, other urban and rural). For the purposes of the 1993 census, Madagascar was divided into *zones de denombrement* (ZDs) which were covered by individual enumerators. The urban ZDs contain an average of 196 households, and the rural ZDs have an average of 160 households. The primary sampling unit (PSU) for the master sampling frame is defined as an individual ZD or combination of ZDs within a *zone d'equipe* (ZE), with a minimum of 300 households. Within each GCU, the PSUs were classified into four socioeconomic groups, based on data from the 1993 census. The PSUs were ordered by socioeconomic group and geographically within each group, in order to provide a corresponding implicit stratification. For the other urban and rural strata, the PSUs were ordered geographically. The PSUs within each stratum were selected systematically with probability proportional to size (PPS), where the measure of size was based on the number of households in the 1993 census frame. A total of 720 PSUs were selected for the master sample, systematically divided into four nationally-representative subsamples (replicates) of 190 sample PSUs each. Within each sample PSU, one ZD was selected with PPS. The sample ZDs for the 1997 EPM were selected from Subsamples 1, 2 and 3 of the master sample. In order to maintain the PPS nature of the sample areas, the subsample of ZDs for the 1997 EPM within each stratum were selected from the master sample with equal probability. A total of 274 sample ZDs were selected for the 1997 EPM. A new listing of households was conducted in these sample ZDs. In the case of the GCU stratum, a sample of 20 households was selected within each sample ZD; for the other urban and rural strata, 25 households were selected per ZD. Table 1 shows the distribution of the sample ZDs and households by stratum for the 1997 EPM.

Table 1. Distribution of Sample ZDs and Households for 1997 EPM by *Faritany* and *Milieu* (Zone of Residence)

| *Faritany* | GCU | | Other Urban | | Rural | | Total | |
|---|---|---|---|---|---|---|---|---|
| | Sample ZDs | Sample Hhs. | Sample ZDs | Sample Hhs. | Sample ZDs | Sample Hhs. | Sample ZDs | Sample Hhs. |
| Antananarivo | 62 | 1240 | 4 | 100 | 37 | 925 | 103 | 2265 |
| Fianarantsoa | 7 | 140 | 4 | 100 | 30 | 750 | 41 | 990 |
| Toamasina | 11 | 220 | 5 | 125 | 24 | 600 | 40 | 945 |
| Mahajanga | 9 | 180 | 3 | 75 | 20 | 500 | 32 | 755 |
| Toliary | 6 | 120 | 4 | 100 | 21 | 525 | 31 | 745 |
| Antsiranana | 5 | 100 | 3 | 75 | 19 | 475 | 27 | 650 |
| Madagascar | 100 | 2000 | 23 | 575 | 151 | 3775 | 274 | 6350 |

## 4. Review of Measures of Precision and Sampling Efficiency for 1997 EPM

It is important to calculate measures of precision for the most important estimates from the 1997 EPM data, in order to determine whether the reliability is satisfactory for each domain, and to evaluate the efficiency of the sample design. The software package CENVAR (Census Variance Calculation System), a component of IMPS (Integrated Microcomputer Processing System), was used to tabulate the standard error, coefficient of variation (CV), 95 percent confidence interval and design effect (DEFF) for key estimates from the 1997 EPM data. The estimation methodology used by CENVAR is described in Section 12. The design effect is defined as the ratio of the variance for a particular survey estimate based on the actual sample design to the corresponding variance based on a simple random sample of the same size. The DEFF is therefore a measure of the relative sampling efficiency of the sample design, and takes into account the effects of stratification, clustering, and differential sampling rates. It should be pointed out that the DEFT statistic appearing in the CLUSTERS tabulation of standard errors from the 1994 EPM data is defined in terms of the ratio of the standard errors, and is therefore the square root of DEFF. The standard error tables in Annex B of the 1997 Madagascar DHS final report also present the DEFT, which needs to be squared in order to obtain the corresponding DEFF.

The CENVAR analysis was carried out for the 1997 EPM estimates of average household income, average total expenditures, average expenditure for food and non-food items, by geographic domain (*faritany* and *milieu*), and the proportion of households by major economic activity of the head of household. These CENVAR results are presented in Annex I.

At the *faritany* level, the CVs for the 1997 EPM estimates of average household income and expenditures are all below 10 percent except for Toliary, where the CV is 13.6 percent for income and 13.9 percent for expenditures. One reason for the higher CV in Toliary is the very high DEFF (7.76 for income and 8.13 for expenditures), which is about twice the DEFF at the national level. This issue is discussed later in this section. At the national level, the CV for average household income is 3.9 percent, and that for average expenditure is 3.8 percent, indicating a high level of precision.

By studying the design effects from the 1997 EPM data, it is possible to determine how the sample design can be improved for the 1999 survey. The design effect for each domain indicates the relative sampling efficiency in the corresponding strata. It is interesting that the DEFF for the estimate of average household income for the GCU at the national level was only 0.78, indicating a very high level of sampling efficiency, probably due to the implicit socioeconomic stratification which offset the effects of clustering. On the other hand, the DEFF for rural is very high, 4.68, indicating a large between-cluster variability and homogeneity within clusters in the rural areas. For average household expenditures, the DEFF for the rural estimate is even higher, 7.03, which is somewhat surprising, given that at the national level the DEFF is lower for expenditures (2.72) than for income (3.89). This could be effected by differential nonsampling errors in the data

collection for the rural areas, where consumption data are more problematic. The DEFF is high for the rural average food expenditures (6.74), which indicates that one source of the variability could be differences in the quality of the autoconsumption data by area; this requires further study. The DEFF for the other urban stratum appears to be reasonable: 2.94 for income and 1.83 for expenditures.

Given that the DEFFs are somewhat high at the national level (for example, 3.89 for average household income), it is important to examine the potential source of these design effects. The main factors which affect the DEFF are clustering effects due to intraclass correlation (that is, homogeneity within clusters) and differential weights (given the different sampling rates by

$$DEFF_C = 1 + d x [\bar{n} - 1] ,$$

stratum). The DEFF due to the clustering effects ($DEFF_C$) can be expressed as follows:

where:

$d$ = coefficient of intraclass correlation, or measure of homogeneity within the cluster, for a particular characteristic (such as household income)

$\bar{n}$ = average number of households selected per cluster; in the case of the 1997 EPM, $\bar{n}$ = 20 for GCU and $\bar{n}$ = 25 for other urban and rural

It can be seen from this formula that the DEFF is affected by the intraclass correlation and the number of households selected per cluster. In order to reduce the DEFF, it would be necessary to select a smaller number of households per cluster, and include more clusters in the sample.

The other factor which has increased the DEFF for the 1997 EPM estimates at the national and *faritany* levels is the different sampling rates used for the individual strata. Table 2 shows the average weight for the sample households in each stratum (*faritany* by *milieu*) for the 1997 EPM. In this table it can be seen that the weights are much lower for the GCU strata than for the other urban and rural strata. These weights vary by a factor as high as 5.26 (736.00/139.89) in the case of Toliary GCU and other urban. As a result, in this particular case the DEFF for the combined urban estimate of average household income for Toliary is very high (8.73). At the national level, the ratio between the weights for sample households in the GCU and other urban stratum is 3.8. One reason for these differential sampling rates is that in the original sample design for the 1997 EPM proposed in Scott's report, the other urban stratum was combined with the rural stratum in determining the sample allocation. However, most of the survey tables from the 1997 EPM data had the GCU and other urban combined into one category; sometimes the Capital (Antananarivo), other GCU and other urban were tabulated separately.

Table 2. Average Weight for Sample Households by Stratum (*Faritany* and *Milieu*) for 1997 EPM

| *Faritany* | GCU | Other Urban | Rural |
|---|---|---|---|
| Antananarivo<br>Capital<br>Antsirabe | 151.20<br>150.44<br>155.81 | 540.00 | 567.39 |
| Fianarantsoa | 162.00 | 690.73 | 581.11 |
| Toamasina | 146.00 | 536.00 | 599.63 |
| Mahajanga | 147.89 | 555.00 | 509.45 |
| Toliary | 139.89 | 736.00 | 648.43 |
| Antsiranana | 167.00 | 360.00 | 418.21 |
| Madagascar | 151.25 | 576.05 | 559.99 |

The weights for the sample households in the 1997 EPM data should be based on the inverse of the probability of selection, which takes into account each sampling stage. The basic weight is specified in Section 12 on Estimation Procedures. It can be seen in that formula that although the weights are approximately self-weighting within each stratum, the weights may vary somewhat by ZD within a stratum based on the difference between the number of households in the new listing for the ZD and the corresponding number from the 1993 census frame. However, in examining the final weights assigned to the 1997 EPM data file, it was found that the weights for most of the sample ZDs within each stratum are the same. Apparently some averaging and rounding was carried out in calculating the weights for the sample ZDs in each stratum. One way to evaluate the weights is to compare the weighted number of households by stratum from the survey data with the corresponding total of number of households in the 1993 census frame for each stratum. This comparison is shown in Table 3, which also shows the percent difference between the two figures for each stratum.

The results in Table 3 indicate that the weights appear to be reasonable. At the national level, there was a 7.4 percent increase in the number of households, which reflects the growth in the population in the period between the 1993 census and the 1997 EPM. There were only two small strata where the weighted number of households from the survey data were slightly less than the frame: Toliary GCU and Antsiranana other urban. The overall conclusion from this review is that the final weights are reasonable, and any bias from averaging and rounding the weights within each stratum should be minimal.

Table 3. Weighted Total Number of Households from 1997 EPM Data with Corresponding Number of Households from the 1993 Census Frame, by Stratum

| FARITANY/ Stratum | Weighted Total Hhs. 1997 EMP | No. Hhs. 1993 Census Frame | % Diff. |
|---|---|---|---|
| ANTANANARIVO | 735549 | 705177 | 4.3% |
| Capital | 157960 | 145591 | 8.5% |
| Antsirabe | 26955 | 25146 | 7.2% |
| Other Urban | 50760 | 48675 | 4.3% |
| Rural | 499874 | 485765 | 2.9% |
| FIANARANTSOA | 517084 | 478634 | 8.0% |
| GCU | 22680 | 20439 | 11.0% |
| Other Urban | 61475 | 60510 | 1.6% |
| Rural | 432929 | 397685 | 8.9% |
| TOAMASINA | 457823 | 415218 | 10.3% |
| GCU | 32120 | 29014 | 10.7% |
| Other Urban | 65928 | 62429 | 5.6% |
| Rural | 359775 | 323775 | 11.1% |
| MAHAJANGA | 320422 | 289108 | 10.8% |
| GCU | 26620 | 24222 | 9.9% |
| Other Urban | 41625 | 37554 | 10.8% |
| Rural | 252177 | 227332 | 10.9% |
| TOLIARY | 419003 | 388646 | 7.8% |
| GCU | 15528 | 15571 | -0.3% |
| Other Urban | 72128 | 66332 | 8.7% |
| Rural | 331347 | 306743 | 8.0% |
| ANTSIRANANA | 238644 | 226076 | 5.6% |
| GCU | 16700 | 15055 | 10.9% |
| Other Urban | 26640 | 28466 | -6.4% |
| Rural | 195304 | 182555 | 7.0% |
| MADAGASCAR | 2688525 | 2502859 | 7.4% |
| Capital | 157960 | 145591 | 8.5% |
| Other GCU | 140603 | 129447 | 8.6% |
| Other Urban | 318556 | 303966 | 4.8% |
| Rural | 2071406 | 1923855 | 7.7% |

5. Modification of Sample Design for 1999 EPM

The accuracy of the survey results depends on both the sampling error, which can be measured through variance estimation, and the nonsampling error, which can only partially be measured through expensive reinterview or validation studies. The sampling error is inversely proportional to the sample size. On the other hand, the nonsampling error may increase with the sample size, since it is more difficult to control the quality of a larger operation. It is therefore important that the overall sample size be manageable for quality and operational control purposes.

One conclusion from the review of the CENVAR results from the 1997 EPM data is that it is possible to improve the efficiency of the sample design for the 1999 EPM in order to obtain the same or a higher level of precision with a smaller sample size. Because of resource constraints and the expected longer household interviews for the 1999 EPM, the DSM would like to limit the total sample size to approximately 5,000 households.

The findings from the CENVAR analysis of standard errors and design effects for estimates from the 1997 EPM data can be used to determine a more efficient allocation of the sample ZDs and households by stratum. In order to decrease the design effects, it is recommended to decrease the number of sample households selected per ZD to 16 for the GCU and other urban strata, and 18 for the rural strata. The slightly higher number of sample households per ZD recommended for the rural strata takes into account the higher cost of traveling between sample ZDs in the rural areas. At the same time, the total number of sample ZDs should be increased slightly to 300 (from 274 for the 1997 EPM). Although this will slightly increase the cost of the listing operation and transportation between sample ZDs in the rural areas, it will significantly increase the sampling efficiency, as described later in this section. The design effects for the urban estimates can also be decreased by using a similar overall sampling rate for the GCU and other urban strata within each *faritany*. Table 4 shows the proposed number of sample ZDs and households by stratum for the 1999 EPM.

Table 4. Proposed Number of Sample ZDs and Households for 1999 EPM by *Faritany* and Zone of Residence (*Milieu*)

| *Faritany* | GCU | | Other Urban | | Rural | | Total | |
|---|---|---|---|---|---|---|---|---|
| | Sample ZDs | Sample Hhs. | Sample ZDs | Sample Hhs. | Sample ZDs | Sample Hhs. | Sample ZDs | Sample Hhs. |
| Antananarivo<br>Capital<br>Antsirabe | 42<br>36<br>6 | 672<br>576<br>96 | 12 | 192 | 38 | 684 | 92 | 1548 |
| Fianarantsoa | 4 | 64 | 12 | 192 | 34 | 612 | 50 | 868 |
| Toamasina | 6 | 96 | 12 | 192 | 26 | 468 | 44 | 756 |
| Mahajanga | 8 | 128 | 8 | 128 | 18 | 324 | 34 | 580 |
| Toliary | 4 | 64 | 16 | 256 | 26 | 468 | 46 | 788 |
| Antsiranana | 6 | 96 | 10 | 160 | 18 | 324 | 34 | 580 |
| Madagascar | 70 | 1120 | 70 | 1120 | 160 | 2880 | 300 | 5120 |

In order to determine the level of precision which can be expected from the proposed sample design for the 1999 EPM, the approximate standard errors for average household income were estimated using the CENVAR results from the 1997 data. Within each stratum, the design effect only measures the clustering effect, since the sample for the 1997 EPM is approximately self-weighting within each stratum. Therefore the design effect for each stratum based on the sample

$$DEFF_{97h} = 1 + d_h \, x[\bar{n}_{97h} - 1] \, ,$$

design for the 1997 EPM can be defined follows:

where:

$DEFF_{97h}$ = design effect for estimate of average household income in stratum h based on sample design for the 1997 EPM

$d_h$ = coefficient of intraclass correlation or measure of homogeneity within the ZDs in stratum h for a particular characteristic

$\bar{n}_{97h}$ = average number of households selected per sample ZD within stratum h for the 1997 EPM; $\bar{n}_{97h}$ = 20 for GCU and $\bar{n}$ = 25 for other urban and rural

This expression can then be used to estimate the intraclass correlation coefficient for stratum h,

$$d_h = \frac{DEFF_{97h} - 1}{\bar{n}_{97h} - 1}$$

$d_h$, as follows:
The design effect based on the proposed sample design for the 1999 EPM can then estimated

$$DEFF_{99h} = 1 + (\bar{n}_{99h} - 1) x \left[ \frac{(DEFF_{97h} - 1)}{(\bar{n}_{97h} - 1)} \right],$$

from the CENVAR results based on the 1997 EPM data, as follows:
where:

$DEFF_{99h}$ = design effect for estimate of average household income for stratum h based on proposed sample design for the 1999 EPM

$\bar{n}_{99h}$ = average number of households selected per sample ZD within stratum h in the proposed sample design for the 1999 EPM; $\bar{n}_{99h}$ = 16 for GCU and other urban, and $\bar{n}_{99h}$ = 18 for rural

The ratio between the variance (square of the standard error) for the survey estimate of average household income for stratum h based on the proposed sample design for the 1999 EPM and that based on the 1997 EPM design can be expressed as follows:

$$\frac{var_{99}(\bar{x}_h)}{var_{97}(\bar{x}_h)} = \frac{\frac{s_{xh}^2}{n_{99h}} x\, DEFF_{99h}}{\frac{s_{xh}^2}{n_{97h}} x\, DEFF_{97h}},$$

where:

$var_{99}(\bar{x}_h)$ = variance (square of standard error) for estimate of average household income in stratum h based on proposed sample design for 1999 EPM

$var_{97}(\bar{x}_h)$ = variance for estimate of average household income in stratum h based on

actual sample design for 1997 EPM

$s_{xh}^2$ = population variance for household income in stratum h

$n_{99h}$ = number of sample households for the 1999 EPM in stratum h

$n_{97h}$ = number of sample households for the 1997 EPM in stratum h

From this ratio the variance for the estimate of average household income based on the proposed

$$var_{99}(\bar{x}_h) = var_{97}(\bar{x}_h) x \frac{n_{97h}}{n_{99h}} x \frac{DEFF_{99h}}{DEFF_{97h}},$$

sample design for the 1999 EPM can be expressed as follows:
where $DEFF_{99h}$ is calculated using the formula specified previously.

This formula was used to calculate the approximate variance for the average household income by stratum which would result from the proposed sample design for the 1999 EPM. The variance for the estimate of average household income for each geographic domain was derived as a weighted combination of the variances for the strata within the domain, as follows:

$$var_{99}(\bar{x}_d) = \sum_{hed} \left( \frac{N_h}{N_d} \right)^2 x\, var_{99}(\bar{x}_h),$$

where:

$var_{99}(\bar{x}_d)$ = variance for estimate of average household income in geographic domain d based on proposed sample design for 1999 EPM

$N_h$ = total number of households in the frame for stratum h, based on the 1993 census

$N_d = \sum_{hed} N_h$ = total number of households in the frame for domain d, based on the 1993 census

The sum in this expression is across all strata which belong to geographic domain d. The standard error of the estimate for domain d is equal to the square root of this variance.

These procedures were used to estimate the approximate CVs for the estimates of average household income by domain based on the proposed sample design for the 1999 EPM. These results are shown in Table 5, together with the corresponding CVs from the 1997 EPM data, and

the difference between the CVs from the two surveys.

It can be seen in Table 5 that even though the total number of sample households in the 1997 EPM was 24 percent higher than that proposed for the 1999 EPM, the more efficient design for the latter survey results in slightly lower CVs at the national level and for most of the geographic domains. This illustrates the gain in sampling efficiency from the new sample allocation, the reduction in the number of households selected per sample ZD, and a slight increase in the total number of sample ZDs. In the case of the estimate of average household income for the GCU, the CV based on the 1999 EPM design is slightly higher than that from the previous survey, given the large concentration of the sample in this stratum in the previous design; however, it is still within 5 percent, which is very good. The CVs also increased slightly for two of the *faritany*, Antananarivo and Antsiranana, but the precision of these estimates is still very reasonable. Although the CV for Toliary decreases based on the new sample design, it is still somewhat high (12.15 percent) because of the large between-ZD variability in the sample data for this *faritany*. This may be partly due to a higher level of nonsampling error in the data for Toliary, given that the DEFF is also high for the estimate of average food expenditures in this *faritany*. Such nonsampling error can be reduced for the 1999 EPM by improving the quality control and supervision, in order to improve the accuracy of the results for Toliary.

Table 5. Approximate Standard Errors and CVs for Average Household Income Based on 1999 EPM Sample Design

| DOMAIN | Estimate of Average Household Income (1997 EPM) | Approximate Standard Error, 1999 EPM | Approximate CV, 1999 EPM | CV, 1997 EPM | Difference, 1997 CV - 1999 CV |
|---|---|---|---|---|---|
| Madagascar | 2330089 | 87108.9 | 3.74% | 3.85% | 0.11% |
| Urban | 3154890 | 110723.5 | 3.51% | 4.03% | 0.52% |
| GCU | 3646812 | 180379.9 | 4.95% | 3.66% | -1.29% |
| Other Urban | 2693842 | 133582.3 | 4.96% | 7.93% | 2.97% |
| Rural | 2084361 | 108315.1 | 5.20% | 5.28% | 0.08% |
| FARITANY | | | | | |
| Antananarivo | 2374197 | 177659.3 | 7.48% | 7.02% | -0.46% |
| Finanarantsoa | 1304205 | 93110.4 | 7.14% | 7.60% | 0.46% |
| Toamasina | 3153052 | 230144.0 | 7.30% | 7.67% | 0.37% |
| Mahajanga | 1787341 | 134694.0 | 7.54% | 7.98% | 0.44% |
| Toliary | 2695735 | 327444.4 | 12.15% | 13.58% | 1.43% |
| Antsiranana | 2924924 | 242621.7 | 8.29% | 7.43% | -0.86% |

6. Estimated Precision for Family Planning and Child Immunization Characteristics

Given that the 1999 EPM questionnaire will include questions on family planning and child immunization, it is important to examine the approximate level of precision which can be expected for the corresponding estimates by geographic domain. The final report from the 1997 Madagascar DHS includes Annex B with tabulated measures of precision, which can be used to study the variability for these types of characteristics. Two types of estimates included in the DHS report were selected for this study: the proportion of currently married (*en union*) women age 15 to 49 using modern methods of contraception, and the proportion of children from 12 to 23 months who are fully immunized.

For each characteristic, Annex B in the 1997 DHS report includes estimates of the standard error and the DEFT, which is the square root of the design effect (DEFF) described previously. It is possible to use these estimates to determine the approximate level of precision which can be expected from the proposed sample design for the 1999 EPM. The variance for the survey estimate of a proportion ($Var(\hat{p})$) such as the contraceptive prevalence rate (CPR) can be defined as follows:

$$Var(\hat{p}) = \frac{\hat{p}x\hat{q}}{n-1} xDEFF \ ,$$

where:

$\hat{p}$ = proportion estimated from the survey data

$\hat{q}$ = (1- $\hat{p}$ )

*DEFF* = design effect for characteristic based on sample design

*n* = total number of sample cases

As in the case of the design effect for the variance of average household income, the DEFF depends on the intraclass correlation and the number of observations per cluster, although it is also affected by the stratification. The estimates of standard errors and DEFT in Annex B of the 1997 DHS report are not available by stratum. However, the stratification of the sample for the 1999 EPM is similar to that for the 1997 DHS, which was also based on the master sampling frame. The approximate DEFFs for the contraceptive prevalence and child immunization rates based on the 1999 EPM sample design were estimated from the corresponding estimates of DEFT from the 1997 DHS, taking into account the different number of observations per ZD in the two surveys. This approach was similar to that described in the previous section for estimating the DEFF for average household income based on the new design. For the 1997 DHS an average of about 29 households were selected per sample ZD, resulting in an average of 16 completed interviews of eligible women (*en union*, age 15 to 49) per ZD. Assuming a similar ratio of completed interviews for eligible women in the 1999 EPM, it is estimated that the average number of eligible women interviews per ZD would be about 10 in the urban strata and 11 in the rural strata. Although the total number of sample households for the 1999 EPM is less than that used for the 1997 DHS, the total number of sample ZDs increases about 10 percent, from 269 to 300.

In the case of the immunization rate for children 12 to 23 months, the number of observations per sample ZD is much smaller. The average number of sample children 12 to 23 months per ZD in the 1997 DHS was about 4.2, representing a ratio of about 0.16 per sample household. Applying this same ratio to the sample for the 1999 EPM, the average number of sample children 12 to 23 months per ZD would be about 2.6 for the urban strata and 2.9 for the rural strata.

Table 6 presents the 1997 DHS estimates of the proportion of currently married women using modern methods of contraception with the corresponding standard errors by geographic domain (*faritany* and *milieu*). This table also shows the approximate DEFF and standard errors which can be expected for each domain based on the sample design for the 1999 EPM. Table 7 shows similar results for the child immunization rates.

Table 6. Approximate Standard Errors for Estimates of Current Contraceptive Prevalence Rate for Modern Methods by Domain Based on Proposed Sample Design for 1999 EPM

| DOMAIN | CPR Modern Methods, 1997 DHS | Standard Error, 1997 DHS | Approximate DEFF, 1999 EPM | Approximate Number of Sample CM Women | Approximate Standard Error, 1999 EPM |
|---|---|---|---|---|---|
| Madagascar | 0.097 | 0.008 | 2.43 | 3174 | 0.008 |
| Urban | 0.176 | 0.016 | 2.09 | 1389 | 0.015 |
| Rural | 0.071 | 0.010 | 3.14 | 1786 | 0.011 |
| FARITANY | | | | | |
| Antananarivo | 0.159 | 0.017 | 2.31 | 960 | 0.018 |
| Finanarantsoa | 0.051 | 0.016 | 2.73 | 538 | 0.016 |
| Toamasina | 0.094 | 0.023 | 2.45 | 469 | 0.021 |
| Mahajanga | 0.053 | 0.012 | 1.55 | 360 | 0.015 |
| Toliary | 0.054 | 0.018 | 2.55 | 489 | 0.016 |
| Antsiranana | 0.108 | 0.014 | 1.05 | 360 | 0.017 |

Table 7. Approximate Standard Errors for Estimates of Proportion of Fully Immunized Children Age 12 to 23 Months by Domain Based on Proposed Sample Design for 1999 EPM

| DOMAIN | Proportion Immunized Children, 1997 DHS | Standard Error, 1997 DHS | Approximate DEFF, 1999 EPM | Approximate Number of Sample Children 12-23 M | Approximate Standard Error, 1999 EPM |
|---|---|---|---|---|---|
| Madagascar | 0.362 | 0.019 | 1.49 | 819 | 0.020 |
| Urban | 0.464 | 0.053 | 2.18 | 358 | 0.039 |
| Rural | 0.335 | 0.019 | 1.25 | 461 | 0.025 |

| | | | | | |
|---|---|---|---|---|---|
| FARITANY | | | | | |
| Antananarivo | 0.632 | 0.027 | 1.16 | 248 | 0.033 |
| Finanarantsoa | 0.194 | 0.037 | 1.42 | 139 | 0.040 |
| Toamasina | 0.420 | 0.051 | 1.32 | 121 | 0.052 |
| Mahajanga | 0.190 | 0.053 | 2.18 | 93 | 0.060 |
| Toliary | 0.150 | 0.046 | 1.68 | 126 | 0.041 |
| Antsiranana | 0.156 | 0.041 | 1.35 | 93 | 0.044 |

It can be seen from Tables 6 and 7 that the level of precision from the 1999 EPM for the contraceptive prevalence and child immunization rates will be similar to that from the 1997 DHS, even though the total number of sample households will decrease by about 30 percent. One reason for this is that the sample allocation for the 1999 EPM provides a higher level of sampling efficiency. In general, the intraclass correlation for demographic characteristics is smaller than that for socioeconomic characteristics, so a larger number of households per cluster is normally selected for demographic surveys than for income and expenditure surveys such as the EPM. However, the design effects for both contraceptive prevalence and child immunization from the 1997 Madagascar DHS were somewhat high, indicating a higher clustering effect for these characteristics. This is not surprising, given that contraceptive prevalence and immunization rates are both correlated with socioeconomic characteristics.

As in the case of the 1997 DHS results, the standard errors for the estimates of the proportion of children 12 to 23 months who are fully immunized by *faritany* based on the 1999 EPM design are fairly high, given the relative small number of children in this age group included in the sample. As a result, the corresponding confidence intervals for some *faritanys* are very wide. The precision of the immunization rates is reasonable at the national level by *milieu*, and perhaps some users can group the *faritanys* into regions. In general, surveys specializing in child immunization need a sample of about 250 children (age 12 to 23 months) per domain, which would require a much larger number of sample households. One sampling approach which can be considered for such surveys as the next Multiple Indicators Cluster Survey (MICS) sponsored by UNICEF, is to use a special screening question during the listing operation to identify the households with children in this age group. Then the listed households can then be stratified into two groups: those with eligible children and those without. If the survey covers all households, a higher sampling rate can be used for selecting the households with eligible children, and the data can be weighted accordingly.

## 7. Considerations for Overlapping Sample with 1997 EPM

One of the analytical objectives of the 1999 EPM is to compare the results to those from the 1997 EPM in order to determine trends in the socioeconomic indicators over time. In order to improve the precision of estimates of differences in income, expenditures and other characteristics between 1997 and 1999, it is recommended that part of the 1997 sample ZDs be maintained in the sample

for the 1999 EPM. The resulting correlation between the estimates from the two surveys will improve the reliability of the results for trends and comparisons. Another advantage of keeping a large proportion of the 1997 EPM sample ZDs in the new sample is that updated sketch maps have already been prepared for these areas from the previous survey, which will improve the quality and cost-effectiveness of the new listing of households.

In some countries which conduct continuous labor force surveys on a quarterly or annual basis, a sample rotation scheme is used in which a portion of the selected households is maintained in the sample from one survey to the next, while another portion is rotated. This increases the correlation between the samples to improve the estimates of trend. However, the Madagascar EPM is only conducted once every two years, and a new listing of households is carried out each time. Since there will be many changes in the households within sample ZDs, it would complicate the field and processing procedures if an attempt is made to select some of the same households in the previous sample. For example, it would be necessary to match the previous sample of households in each sample ZD to the new listing, with potential matching errors when different names are used. For this reason it is recommended to limit the overlap to sample areas, and to select a new sample of households from the current listing. Although some of the previous sample households may be selected again, most of the sample households will be new. However, if most of the sample ZDs from the 1997 EPM are maintained in the sample, the correlation between the households in these areas will improve the reliability for the estimates of trends over time. In this case the sample design can take advantage of the intraclass correlation. The CENVAR results from the 1997 EPM data indicate that this intraclass correlation is generally high for socioeconomic characteristics.

8. Selection of Sample ZDs for 1999 EPM

In comparing the distribution of the sample ZDs for the 1997 EPM shown in Table 1 with the corresponding distribution proposed for the 1999 EPM in Table 4, it can be seen that for some strata the number of sample ZDs will be increased, and for other strata the sample size will be reduced. In order to maximize the number of sample ZDs which overlap between the two surveys, it is recommended to use the following procedures for selecting the sample ZDs for the 1999 EPM:

(1) In the case of strata where the number of sample ZDs in the 1997 EPM is higher than the corresponding number proposed for the 1999 EPM, all of the sample ZDs for new survey should be selected as a subsample of the 1997 sample ZDs, systematically with equal probability.

(2) For strata where the number of sample ZDs in the 1997 EPM is lower than the corresponding number proposed for the 1999 EPM, all of the 1997 sample ZDs should be maintained in the sample. The additional sample ZDs should be selected systematically with equal probability from the ZDs identified in Subsamples 1, 2 and 3 from the master sample, excluding those in the 1997 EPM sample.

Given that the sample ZDs for the 1997 EPM were selected from Subsamples 1, 2 and 3 of the master sample, it is recommended that the new sample ZDs also be selected from these subsamples. Each nationally-representative subsample has 190 sample ZDs, so the three subsamples with a total of 570 sample ZDs should be large enough for selecting the new sample ZDs.

In order to examine the resulting overlap between the 1997 and 1999 sample ZDs, Table 6 shows the distribution of the 1999 EPM sample ZDs to be selected from the 1997 sample, and the number of additional sample ZDs which will be selected, by stratum (*faritany* and *milieu*). It can be seen in this table that at the national level a total of 80 percent of sample ZDs for the 1999 EPM will be kept from the 1997 sample, which should be very efficient for estimating trends between the two surveys. Given the differences in the sample allocation between the two surveys, the percentage of overlap varies by stratum. In the case of the GCU, all of the sample ZDs for the 1999 EPM except for one in Antsiranana will be selected from the 1997 sample; given the higher variability for economic characteristics in the large cities, the high level of overlap for this stratum is important. On the other hand, only about a third of the sample ZDs for the other urban strata will be come from the previous sample, given the lower sampling rate for these strata in the 1997 design. In the case of the rural strata, the overlap in the sample ZDs will be slightly over 92 percent.

Table 8. Proposed Number of Sample ZDs for 1999 EPM Overlapping with 1997 Sample, and Newly Selected ZDs, by *Faritany* and Zone of Residence (*Milieu*)

| *Faritany* | GCU | | Other Urban | | Rural | | Total | |
|---|---|---|---|---|---|---|---|---|
| | 1997 Sample ZDs | New Sample ZDs | 1997 Sample ZDs | New Sample ZDs | 1997 Sample ZDs | New Sample ZDs | 1997 Sample ZDs | New Sample ZDs |
| Antananarivo | 42 | - | 4 | 8 | 37 | 1 | 83 | 9 |
| Capital | 36 | - | | | | | | |
| Antsirabe | 6 | - | | | | | | |
| Fianarantsoa | 4 | - | 4 | 8 | 30 | 4 | 38 | 12 |
| Toamasina | 6 | - | 5 | 7 | 24 | 2 | 35 | 9 |
| Mahajanga | 8 | - | 3 | 5 | 18 | - | 29 | 5 |
| Toliary | 4 | - | 4 | 12 | 21 | 5 | 29 | 17 |
| Antsiranana | 5 | 1 | 3 | 7 | 18 | - | 26 | 8 |
| Madagascar | 69 | 1 | 23 | 47 | 148 | 12 | 240 | 60 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | | |

When the master sample design was implemented, a database file identifying the master sample of 760 sample ZDs was generated. This file contains information on all the geographical codes and sampling frame information for the sample ZDs, including the number of households in the frame for each sample PSU and ZD, and a subsample code from 1 to 4. This file will be used for selecting the 1999 EPM sample, and it is important to electronically maintain all of the geographic and sampling information in the computer files generated for the new sample. This information will also be used in calculating the weights, as described in the Section 12 on Estimation Procedures.

In the database file with the master sample of 760 ZDs a new «flag» field was generated to identify the 274 ZDs in the sample for the 1997 EPM. Eric Ribaira created this field by matching a file with the 1997 EPM sample ZDs to the original master sample file. In order to carry out the two steps specified for selecting the sample ZDs for the 1999 EPM, the 1997 EPM flag field was used to generate two subfiles with all the information from the frame: one file with the 274 ZDs selected in the 1997 sample, and a second file with all of the ZDs from Subsamples 1, 2 and 3 of the master sample which were not included in the 1997 sample. The first file is used for selecting the 1999 EPM sample ZDs from the previous sample (step 1), and the second file will be used for selecting the additional sample ZDs for the 1999 EPM (step 2). Based on the distribution of the sample ZDs shown in Table 6, a total of 240 sample ZDs will be selected from the first file, and 60 from the second file. The two sample files will then be joined to obtain the database file with the final sample of 300 sample ZDs for the 1999 EPM. This file will contain the information needed to calculate the weights for the sample households later.

The database programs which were used for selecting the PSUs by stratum with PPS for the master sample were modified for selecting the sample ZDs from the two database files. In order to select the the sample ZDs with equal probability within each stratum, the measures of size were all replaced with the value 1.

## 9. Segmenting Large Sample ZDs

In the case of the 1997 DHS, the DDSS staff divided some of the large sample ZDs into smaller segments with well defined boundaries, and one segment was selected to be listed for the survey. Since it is difficult to divide the ZD sketch map into segments with an equal number of households, it is recommended that the sample segment be selected with PPS. In this case the weight for the sample households in the segment would have an additional factor equal to the inverse of the proportion of the ZD households in the selected segment. The procedures for segmenting the large sample ZDs can be based on the experience from the 1997 DHS. Annex II presents an example of an Excel spreadsheet which can be used for selecting a segment with PPS within a large sample ZD. The DSM has a copy of this file.

## 10. Listing and Selection of Sample Households

A new listing of households will be conducted in the 300 sample ZDs for the 1999 EPM. Within each sample ZD in the GCU and other urban strata a sample of 16 households will be selected systematically from the new listing, and 18 households will be selected within each sample rural ZD.

It is ideal if the listing can be conducted prior to the data collection for the survey, so that the listing results can be reviewed, and the selection of sample households can be controlled in the office. Sometimes it is found that when the selection of households is carried out in the field, the interviewers can bias the selection to make the work easier. For example, studies in some countries have found that households selected by the interviewers have a smaller average number of persons than the overall average household size; less accessible households would also have a smaller chance of being selected. For this reason the supervision and control of the sample selection is very important.

In the case of the 1997 EPM the selection of households was conducted in the field immediately following the listing, in order to reduce the costs of the fieldwork. In order to reduce the potential selection bias, the supervisors reviewed the listing results and carried out the selection of sample households. Given the budget constraints, the DSM plans to use a similar procedure for the 1999 survey. In order to facilitate the selection of sample households and avoid sampling bias, it is recommended to use an Excel spreadsheet to determine the sampling interval, random start and selection numbers.

11. Considerations for Noninterview Sample Households

In examining the expected level of precision from the proposed sample for the 1999 EPM, it is assumed that the effective sample size will be about 5,120 completed household questionnaires. There are two alternative procedures to compensate for noninterview households in order to assure a minimum effective sample size:

(1) Increase the number of households selected in each ZD based on the expected overall noninterview rate. For example, if it is expected that the noninterview rate will be 10 percent, a sample of 18 households can be selected within each sample urban ZD, and 20 households can be selected within each sample rural ZD. The advantage of this alternative is that it avoids having replacement procedures which may be affected by the interviewers. The disadvantage is that it makes it more difficult to control the overall sample size and the workload for individual interviewers. For example, the workload per sample urban ZD could vary from 12 to 18 household interviews, but the average should be at least 16.

(2) The second option is to substitute all of the noninterview sample households with randomly selected replacement households within the same sample ZD. This alternative is only recommended if the replacement procedures can be carefully controlled by the supervisor. It is also important to obtain the reason for each

noninterview, since the weight adjustment procedures will be different for out-of-scope households (such as demolished or vacant housing units). At the time the original sample of households is selected for each sample ZD, an additional number of households can be selected as possible replacements. The list of possible replacement households would be kept by the supervisor. The interviewer should first make a very strong effort to interview the original sample of households, with multiple call-backs. Only the supervisor should make the decision to replace an original sample household, after a final attempt to complete the interview, and assign the new household to be substituted from the list of replacements. The advantage of this procedure is that it ensures an exact number of completed interviews, and makes it easier to control the workload for each interviewer.

Given that the individual household interviews for the 1999 EPM will be fairly lengthy, it is important to control the interviewer workload. Therefore it is recommended to use the second procedure, and replace the noninterview households. For example, in order to obtain 16 completed household interviews for each sample urban ZD, a total of 24 sample households can be selected systematically within the PSU from the updated listing sheet. Then 8 of these households can be selected systematically to be used as possible replacements, and the remaining 16 will be treated as the original sample households for that ZD. In the case of the sample rural ZDs, a sample of 26 sample households can be selected systematically: 18 in the original sample and 8 as possible replacements. Annex III presents an example of an Excel spreadsheet which can be used for the systematic selection of households from the listing for each sample ZD, including the set of replacement households. The DSM has a copy of this file.

It should be pointed out that under either option the survey results will be affected by noninterview bias, although this bias should be fairly small compared to the standard error. The noninterview bias is due to the fact that the characteristics of interviewed households may be slightly different from those of households which cannot be interviewed. For example, the wealthier households may have a higher noninterview rate, resulting in a potential bias for average household income and expenditures. The only effect of replacing the noninterview households is to maintain the effective sample size.

## 12. Estimation Procedures

### 12.1. Weighting Procedures

In order for the sample estimates from a particular survey to be representative of the population, it is necessary to multiply the data by a sampling weight, or expansion factor. The basic weight for each sample household would be equal to the inverse of its probability of selection (calculated by multiplying the probabilities at each sampling stage). The master sample design will be approximately self-weighting within stratum. Since all survey data will be processed by computer, it should be easy to attach a weight to each sample household record in the computer files, and the tabulation programs can weight the data automatically. Using the database files generated

from the master sampling frame, the sampling probabilities at each stage of selection can be maintained in a microcomputer file so that the overall probability and corresponding weight can be calculated for each sample segment. The weights will probably vary somewhat by sample segment, since in many cases the measure of size in the sampling frame may be slightly different from the actual number of households listed.

The PSUs within each stratum were selected with PPS. In the case of sample PSUs with more than one ZD, an individual ZD was selected with PPS within the sample PSU. Since this procedure is equivalent to selecting the ZDs with PPS at the first stage, the ZDs are treated as PSUs for the estimation purposes. Based on the master sample design, the probability of selection for the sample households in a sample segment would be calculated as follows:

$$p_{hi} = \frac{m_h \, x \, N_{hi}}{N_h} \, x \, p_{Shi} \, x \, \frac{n_{hi}}{N_{h'i}} \, ,$$

where:

$p_{hi}$ = probability of selection for the sample households in the i-th sample ZD in stratum h

$m_h$ = number of sample ZDs selected in stratum h for the EPM, specified in Table 6

$N_h$ = total number of households in the frame for stratum h

$N_{hi}$ = total number of households in the frame for the i-th sample ZD in stratum h

$p_{Shi}$ = probability of selection of sample segment within the i-th sample ZD in stratum h

$n_{hi}$ = number of sample households selected in the i-th sample ZD in stratum h; for the 1999 EPM, $n_{hi}$ = 16 for GCU and other urban strata, and $n_{hi}$ = 18 for the rural strata

$N'_{hi}$ = total number of households listed in the i-th sample ZD (or segment) in stratum h

The three components of this probability of selection correspond to the individual sampling stages. The term $p_{Shi}$ is included to allow for the possibility of segmenting large ZDs for the 1999 EPM; in the case of sample ZDs which are not segmented, $p_{Shi}$ = 1. The basic sampling weight, or expansion factor, is calculated as the inverse of this probability of selection. In this case, the weight can be simplified as follows:

$$W_{hi} = \frac{N_h \, x \, N_{h'i}}{m_h \, x \, N_{hi} \, x \, p_{Shi} \, x \, n_{hi}} ,$$

where:

$W_{hi}$ = basic weight for the sample households in the i-th sample ZD in stratum h

It can be seen that if $n_{hi}$ is constant for each stratum (for example, 16 sample households per urban ZD and 18 per rural ZD for the EPM) and $N'_{hi} = N_{hi}$ (that is, the number of households listed in the sample ZD is equal to the corresponding number in the frame), the sample would be self-weighting within the stratum. The weights will actually vary slightly within a stratum based on the difference between $N'_{hi}$ and $N_{hi}$.

It is also important to adjust the weights to take into account the noninterview rate for each survey. Since the weights will be calculated at the level of the sample segment, it would be advantageous to adjust the weights at this level. Therefore the final weight ($W'_{hi}$) for the sample households in the i-th sample ZD in stratum h can be expressed as follows:

$$W_{h'i} = W_{hi} \, x \, \frac{n_{h'i}''}{n_{hi}} ,$$

where:

$n'_{hi}$ = total number of valid (occupied) sample households selected in the i-th sample ZD in stratum h (that is, the number of interviews plus the number of noninterviews in the sample segment)

$n''_{hi}$ = total number of interviewed sample households selected in the i-th sample ZD in stratum h, including replacement households

If the survey data will later be used to tabulate the distribution of the total population by different characteristics, it will also be possible to adjust the final weights using population projections based on demographic analysis.

In order to calculate the appropriate weights for the 1999 EPM, a spreadsheet was developed for the 300 sample ZDs which includes all of the geographical identification and measures of size from the master sampling frame for each sample ZD. After the listing operation and data collection for the survey, it will be necessary to enter the number of households listed in each sample ZD and the number of completed interviews in order to calculate the final adjusted weights. This spreadsheet can also be used to compare the number of households listed in each sample ZD to the corresponding number in the frame. Whenever there is a large difference between the number of households in the listing and the frame for a sample ZD or segment (for

example, greater than 50 percent), the listing and sketch maps should be verified to ensure that the correct boundaries were used. If a problem is found, it may be necessary to adjust the weight accordingly. For example, if a small *fokontany* with 25 percent of the households in the sample ZD is missing from the listing, the weight for the households in that ZD can be adjusted by the

inverse of the proportion listed (that is, 1/0.75 = 1.33); in this case, the value 0.75 can be entered for $p_{Shi}$ in the spreadsheet, and the weight would be adjusted automatically.

12.2. Survey Estimates

The most common survey estimates to be calculated from the household surveys are in the form of totals and ratios. The survey estimate of a total can be expressed as follows:

$$\hat{Y} = \sum_{h=1}^{L} \sum_{i=1}^{m_h} \sum_{j=1}^{n_{hi}}{}'' W_{h'i}\, y_{hij} \ ,$$

where:

L = number of strata

$y_{hij}$ = value of variable y for the j-th sample household in the i-th sample ZD in stratum h

The survey estimate of a ratio is defined as follows:

$\hat{R} = \frac{\hat{Y}}{\hat{X}}$, where $\hat{Y}$ and $\hat{X}$ are estimates of totals for variables y and x, respectively, calculated as specified previously.

In the case of a stratified multistage sample design such as the EPM, survey estimates of means and proportions are special types of ratios. In the case of the mean, the variable X, in the denominator of the ratio, is defined to be equal to 1 for each element so that the denominator is equal to the sum of the weights. In the case of a proportion, the variable X in the denominator is also defined to equal 1 for all elements; the variable Y in the numerator is binomial and is defined to equal either 0 or 1, depending on the absence or presence, respectively, of the specified characteristic.

12.3. Variance Estimation Procedures

In the publication of the results from the EPM it is important to include a statement on the accuracy of the survey data. In addition to presenting tables with calculated sampling errors for the most important survey estimates, the different sources of nonsampling error should be described.

The standard error, or square root of the variance, is used to measure the sampling error, although it may also include a small part of the nonsampling error. The variance estimator should take into account the different aspects of the sample design, such as the stratification and clustering. In order to avoid the time and effort it would require to develop custom variance programs, it would be ideal to use an available software package to tabulate the measures of precision. One such program available for calculating the standard errors for survey data from stratified multistage sample designs such as the EPM is CENVAR, a component of the Integrated Microcomputer Processing System (IMPS). CENVAR is menu-driven and user-friendly. It uses the data dictionary defined in the DATADICT component of IMPS. It can be used to calculate the standard errors of totals, means, proportions and other ratios. It produces subpopulation estimates for each category of a classification variable, and these variables can be cross-classified. For each estimate, CENVAR calculates the standard error, coefficient of variation (CV), 95 percent confidence interval and the design effect (DEFF). This software package uses an ultimate cluster variance estimator. CENVAR was used for calculating the standard errors from the 1997 EPM data. The DSM has a copy of this software and the corresponding manuals.

In order to tabulate estimates of standard errors using CENVAR, it is generally necessary to produce a new data input file from the original survey data. Since the CENVAR package will only accept one type of record, it is necessary to generate one record for each unit of analysis in the CENVAR data input file. For example, in the case of household estimates, such as average household income, it would be necessary to generate one record for each sample household. For estimates by person, such as the employment rate, the CENVAR input file should have one record for each in-scope sample person. Each record in the CENVAR data input file should include fields for the stratum, cluster and weight, in addition to the classification and analysis variables which are required for the particular CENVAR analysis. The classification variables are used to produce subpopulation estimates for all their respective categories. The analysis variables are generally continuous variables, such as income and expediture data, or count variables, which are equal to 1 if the unit has a certain characteristic and 0 otherwise. CENVAR automatically creates a count variable named INTERCEPT, which is equal to 1 for each record. The INTERCEPT variable can be used to obtain the estimate of the weighted total number of units (for example, the total number of persons or households), or it can be used in the denominator of a ratio in order to obtain a mean or proportion. It can also be used as a classification variable with one category in order to produce estimates at the national level.

CENVAR does not accept any blanks in the file. In the case of classification variables, any record with a blank should be imputed with a special code to identify "missing" or "not applicable." The CENVAR output will include estimates for these categories, which can then be deleted from the tabulations which will be published. For analysis variables, CENVAR assumes that any missing

values are imputed. Once the file is zero-filled, CENVAR will treat any missing value as 0, thus introducing a downward bias in the estimates of means when there are missing values.

The ultimate cluster variance estimator for a total used by CENVAR can be expressed as follows:

Variance Estimator of a Total

$$V(\hat{Y}) = \sum_{h=1}^{L} \left[ \frac{m_h}{m_h - 1} \sum_{i=1}^{m_h} \left( \hat{Y}_{hi} - \frac{\hat{Y}_h}{n_h} \right)^2 \right],$$

$$\hat{Y}_{hi} = \sum_{j=1}^{n_{hi}}{}'' W_{h'i}\, y_{hij}$$

$$\hat{Y}_h = \sum_{i=1}^{n_h} \hat{Y}_{hi}$$

where:

The variance estimator of a ratio used by CENVAR can be expressed as follows:

Variance Estimator of a Ratio

$$V(\hat{R}) = \frac{1}{\hat{X}^2} \left[ V(\hat{Y}) + \hat{R}^2 V(\hat{X}) - 2\, \hat{R}\, COV(\hat{X},\hat{Y}) \right],$$

where:

$$COV(\hat{X},\hat{Y}) = \sum_{h=1}^{L} \left[ \frac{m_h}{m_h - 1} \sum_{i=1}^{m_h} \left( \hat{X}_{hi} - \frac{\hat{X}_h}{m_h} \right) \left( \hat{Y}_{hi} - \frac{\hat{Y}_h}{m_h} \right) \right]$$

$V(\hat{Y})$ and $V(\hat{X})$ are calculated according to formula for the variance of a total.

**ANNEX I**

**TABLEAU DES ERREURS DE SONDAGE POUR *L'ENQUÊTE PRIORITAIRE* 1997**

1. Revenu moyen par ménage selon le domaine géographique (en FMG) (*)

| Categorie | Valeur | Erreur Type | C.V. (%) | Intervalle de confiance 95% Lim. Infer. | Intervalle de confiance 95% Lim. Super. | Effet de Sondage | Nombre de Ménages |
|---|---|---|---|---|---|---|---|
| MADAGASCAR | 2,330,089 | 89,796 | 3.85 | 2,154,088 | 2,506,089 | 3.89 | 6,226 |
| FARITANY | | | | | | | |
| ANTANANARIVO | 2,374,197 | 166,586 | 7.02 | 2,047,689 | 2,700,706 | 2.44 | 2,198 |
| FIANARANTSOA | 1,304,205 | 99,172 | 7.60 | 1,109,827 | 1,498,583 | 4.73 | 974 |
| TOAMASINA | 3,153,052 | 241,825 | 7.67 | 2,679,075 | 3,627,029 | 5.51 | 943 |
| MAHAJANGA | 1,787,341 | 142,619 | 7.98 | 1,507,808 | 2,066,874 | 3.87 | 750 |
| TOLIARY | 2,695,735 | 366,046 | 13.58 | 1,978,284 | 3,413,185 | 7.76 | 720 |
| ANTSIRANANA | 2,924,924 | 217,233 | 7.43 | 2,499,148 | 3,350,700 | 1.32 | 641 |
| MILIEU | | | | | | | |
| URBAIN | 3,154,890 | 127,264 | 4.03 | 2,905,452 | 3,404,328 | 1.68 | 2,527 |
| RURAL | 2,084,361 | 109,993 | 5.28 | 1,868,775 | 2,299,948 | 4.68 | 3,699 |
| MILIEU2 | | | | | | | |
| GCU | 3,646,812 | 133,311 | 3.66 | 3,385,522 | 3,908,102 | 0.78 | 1,974 |
| CUS | 2,693,842 | 213,512 | 7.93 | 2,275,357 | 3,112,326 | 2.94 | 553 |
| RURAL | 2,084,361 | 109,993 | 5.28 | 1,868,775 | 2,299,948 | 4.68 | 3,699 |
| MILIEU3 | | | | | | | |
| CAPITAL | 3,670,963 | 195,706 | 5.33 | 3,287,379 | 4,054,547 | 0.73 | 1,050 |
| GCU | 3,619,679 | 177,749 | 4.91 | 3,271,292 | 3,968,067 | 0.86 | 924 |
| CUS | 2,693,842 | 213,512 | 7.93 | 2,275,357 | 3,112,326 | 2.94 | 553 |
| RURAL | 2,084,361 | 109,993 | 5.28 | 1,868,775 | 2,299,948 | 4.68 | 3,699 |

(*) Calcul excluant les valeurs aberrantes.

1. Revenu moyen par ménage selon le domaine géographique (en FMG) (*) - Continuation

| Categorie | | Valeur | Erreur Type | C.V. (%) | Intervalle de confiance 95% Lim. Infer. | Lim. Super. | Effet de Sondage | Nombre de Ménages |
|---|---|---|---|---|---|---|---|---|
| FARITANY PAR MILIEU | | | | | | | | |
| ANTANANARIVO | URBAIN | 3,463,033 | 158,079 | 4.56 | 3,153,198 | 3,772,869 | 0.80 | 1,317 |
| ANTANANARIVO | RURAL | 1,860,845 | 232,773 | 12.51 | 1,404,610 | 2,317,079 | 3.20 | 881 |
| FIANARANTSOA | URBAIN | 1,929,555 | 276,484 | 14.33 | 1,387,645 | 2,471,464 | 2.95 | 229 |
| FIANARANTSOA | RURAL | 1,182,646 | 104,264 | 8.82 | 978,288 | 1,387,005 | 5.73 | 745 |
| TOAMASINA | URBAIN | 3,769,605 | 430,770 | 11.43 | 2,925,295 | 4,613,914 | 1.98 | 343 |
| TOAMASINA | RURAL | 2,985,025 | 284,584 | 9.53 | 2,427,240 | 3,542,810 | 8.06 | 600 |
| MAHAJANGA | URBAIN | 3,025,566 | 452,492 | 14.96 | 2,138,683 | 3,912,450 | 5.68 | 255 |
| MAHAJANGA | RURAL | 1,452,248 | 134,225 | 9.24 | 1,189,168 | 1,715,329 | 3.65 | 495 |
| TOLIARY | URBAIN | 2,011,202 | 431,415 | 21.45 | 1,165,629 | 2,856,775 | 8.73 | 209 |
| TOLIARY | RURAL | 2,876,824 | 448,831 | 15.60 | 1,997,116 | 3,756,532 | 7.77 | 511 |
| ANTSIRANANA | URBAIN | 4,984,645 | 291,880 | 5.86 | 4,412,560 | 5,556,730 | 0.50 | 174 |
| ANTSIRANANA | RURAL | 2,467,850 | 257,848 | 10.45 | 1,962,469 | 2,973,232 | 1.57 | 467 |
| FARITANY PAR MILIEU2 (STRATE) | | | | | | | | |
| ANTANANARIVO | GCU | 3,584,398 | 179,028 | 4.99 | 3,233,503 | 3,935,293 | 0.77 | 1,223 |
| ANTANANARIVO | CUS | 3,020,910 | 347,060 | 11.49 | 2,340,673 | 3,701,147 | 1.01 | 94 |
| ANTANANARIVO | RURAL | 1,860,845 | 232,773 | 12.51 | 1,404,610 | 2,317,079 | 3.20 | 881 |
| FIANARANTSOA | GCU | 2,065,854 | 547,096 | 26.48 | 993,546 | 3,138,161 | 2.30 | 140 |
| FIANARANTSOA | CUS | 1,879,270 | 321,789 | 17.12 | 1,248,563 | 2,509,977 | 3.36 | 89 |
| FIANARANTSOA | RURAL | 1,182,646 | 104,264 | 8.82 | 978,288 | 1,387,005 | 5.73 | 745 |
| TOAMASINA | GCU | 4,620,661 | 392,407 | 8.49 | 3,851,543 | 5,389,779 | 0.66 | 220 |
| TOAMASINA | CUS | 3,354,972 | 609,629 | 18.17 | 2,160,099 | 4,549,844 | 2.50 | 123 |
| TOAMASINA | RURAL | 2,985,025 | 284,584 | 9.53 | 2,427,240 | 3,542,810 | 8.06 | 600 |
| MAHAJANGA | GCU | 3,718,588 | 381,302 | 10.25 | 2,971,237 | 4,465,939 | 1.00 | 180 |
| MAHAJANGA | CUS | 2,582,366 | 701,019 | 27.15 | 1,208,369 | 3,956,362 | 15.14 | 75 |
| MAHAJANGA | RURAL | 1,452,248 | 134,225 | 9.24 | 1,189,168 | 1,715,329 | 3.65 | 495 |
| TOLIARY | GCU | 2,984,981 | 253,961 | 8.51 | 2,487,217 | 3,482,745 | 0.36 | 111 |
| TOLIARY | CUS | 1,801,563 | 521,556 | 28.95 | 779,312 | 2,823,813 | 12.55 | 98 |
| TOLIARY | RURAL | 2,876,824 | 448,831 | 15.60 | 1,997,116 | 3,756,532 | 7.77 | 511 |
| ANTSIRANANA | GCU | 5,112,889 | 459,848 | 8.99 | 4,211,587 | 6,014,192 | 0.60 | 100 |
| ANTSIRANANA | CUS | 4,904,252 | 378,161 | 7.71 | 4,163,056 | 5,645,447 | 0.46 | 74 |
| ANTSIRANANA | RURAL | 2,467,850 | 257,848 | 10.45 | 1,962,469 | 2,973,232 | 1.57 | 467 |

(*) Calcul excluant les valeurs aberrantes.

1. Revenu moyen par ménage selon le groupe socioéconomique (en FMG) (*) - Continuation

| Categorie | Valeur | Erreur Type | C.V. (%) | Intervalle de confiance 95% Lim. Infer. | Lim. Super. | Effet de Sondage | Nombre de Ménages |
|---|---|---|---|---|---|---|---|
| GROUPE SOCIOÉCONOMIQUE | | | | | | | |
| GRAND EXPL.AGRIC | 3,512,508 | 395,036 | 11.25 | 2,738,237 | 4,286,780 | 1.78 | 343 |
| MOYEN EXPL.AGRIC | 1,974,779 | 139,313 | 7.05 | 1,701,725 | 2,247,833 | 4.58 | 1,007 |
| PETIT EXPL.AGRIC | 1,521,320 | 90,731 | 5.96 | 1,343,487 | 1,699,153 | 5.59 | 2,002 |
| ELEVEUR | 1,959,570 | 709,527 | 36.21 | 568,897 | 3,350,244 | 1.39 | 26 |
| PECHEUR | 4,332,774 | 1,469,041 | 33.91 | 1,453,453 | 7,212,095 | 6.62 | 56 |
| CHASSEURS | 1,271,918 | 507,421 | 39.89 | 277,374 | 2,266,463 | 1.37 | 2 |
| PETIT EXP.N.AGR. | 3,693,264 | 459,014 | 12.43 | 2,793,596 | 4,592,931 | 1.61 | 158 |
| COMMERCANTS | 4,497,421 | 317,440 | 7.06 | 3,875,238 | 5,119,604 | 1.16 | 341 |
| PETIT ENT. SERV. | 5,862,404 | 1,006,811 | 17.17 | 3,889,055 | 7,835,754 | 1.65 | 86 |
| GR.MOY.ENT.SERV. | 15,999,066 | 11,408,519 | 71.31 | -6,361,631 | 38,359,763 | 1.26 | 17 |
| CADRE SALARIE | 4,413,064 | 200,821 | 4.55 | 4,019,454 | 4,806,673 | 0.88 | 426 |
| EMPLOYE, OUVRIER | 2,853,468 | 137,539 | 4.82 | 2,583,892 | 3,123,045 | 1.05 | 754 |
| M.D'OUVRE S.QUAL | 2,284,334 | 188,323 | 8.24 | 1,915,222 | 2,653,447 | 0.88 | 223 |
| AIDES FAMILIAUX | 4,420,464 | 2,186,255 | 49.46 | 135,405 | 8,705,523 | 1.40 | 7 |
| FEM./HOM.AU FOYE | 1,165,847 | 194,355 | 16.67 | 784,910 | 1,546,783 | 1.66 | 31 |
| CLERGE | 3,149,329 | 807,221 | 25.63 | 1,567,175 | 4,731,483 | 1.29 | 16 |
| APPRENTI | 1,785,575 | 823,766 | 46.13 | 170,994 | 3,400,156 | 0.57 | 3 |
| AUTRES | 2,201,696 | 149,127 | 6.77 | 1,909,407 | 2,493,985 | 1.22 | 728 |

(*) Calcul excluant les valeurs aberrantes.

2. Dépenses totales par ménage selon le domaine géographique (en FMG)

| Categorie | Valeur | Erreur Type | C.V. (%) | Intervalle de confiance 95% Lim. Infer. | Lim. Super. | Effet de Sondage | Nombre de Ménages |
|---|---|---|---|---|---|---|---|
| MADAGASCAR | 1,923,079 | 72,286 | 3.76 | 1,781,399 | 2,064,759 | 2.72 | 6,226 |
| FARITANY | | | | | | | |
| ANTANANARIVO | 2,976,277 | 182,358 | 6.13 | 2,618,856 | 3,333,698 | 2.08 | 2,198 |
| FIANARANTSOA | 1,150,914 | 90,201 | 7.84 | 974,119 | 1,327,709 | 3.58 | 974 |
| TOAMASINA | 1,494,625 | 102,563 | 6.86 | 1,293,603 | 1,695,648 | 3.02 | 943 |
| MAHAJANGA | 1,212,502 | 107,862 | 8.90 | 1,001,092 | 1,423,912 | 2.19 | 750 |
| TOLIARY | 2,000,738 | 277,518 | 13.87 | 1,456,802 | 2,544,674 | 8.13 | 720 |
| ANTSIRANANA | 1,989,689 | 100,100 | 5.03 | 1,793,492 | 2,185,886 | 0.54 | 641 |
| MILIEU | | | | | | | |
| URBAIN | 3,945,586 | 175,669 | 4.45 | 3,601,275 | 4,289,897 | 1.33 | 2,527 |
| RURAL | 1,320,528 | 77,039 | 5.83 | 1,169,532 | 1,471,524 | 7.03 | 3,699 |
| MILIEU2 | | | | | | | |
| GCU | 4,994,851 | 221,923 | 4.44 | 4,559,882 | 5,429,820 | 0.98 | 1,974 |
| CUS | 2,962,175 | 271,474 | 9.16 | 2,430,085 | 3,494,264 | 1.83 | 553 |
| RURAL | 1,320,528 | 77,039 | 5.83 | 1,169,532 | 1,471,524 | 7.03 | 3,699 |
| MILIEU3 | | | | | | | |
| CAPITAL | 5,686,915 | 338,468 | 5.95 | 5,023,517 | 6,350,313 | 0.84 | 1,050 |
| GCU | 4,217,353 | 229,835 | 5.45 | 3,766,876 | 4,667,830 | 1.04 | 924 |
| CUS | 2,962,175 | 271,474 | 9.16 | 2,430,085 | 3,494,264 | 1.83 | 553 |
| RURAL | 1,320,528 | 77,039 | 5.83 | 1,169,532 | 1,471,524 | 7.03 | 3,699 |

2. Dépenses totales par ménage selon le domaine géographique (en FMG) - Continuation

| Categorie | | Valeur | Erreur Type | C.V. (%) | Intervalle de confiance 95% Lim. Infer. | Lim. Super. | Effet de Sondage | Nombre de Ménages |
|---|---|---|---|---|---|---|---|---|
| FARITANY PAR MILIEU | | | | | | | | |
| ANTANANARIVO | URBAIN | 5,488,859 | 368,050 | 6.71 | 4,767,480 | 6,210,237 | 1.16 | 1,317 |
| ANTANANARIVO | RURAL | 1,791,673 | 198,727 | 11.09 | 1,402,167 | 2,181,178 | 7.98 | 881 |
| FIANARANTSOA | URBAIN | 2,296,468 | 225,079 | 9.80 | 1,855,314 | 2,737,623 | 1.75 | 229 |
| FIANARANTSOA | RURAL | 928,235 | 97,065 | 10.46 | 737,987 | 1,118,483 | 5.11 | 745 |
| TOAMASINA | URBAIN | 3,400,473 | 389,063 | 11.44 | 2,637,910 | 4,163,036 | 3.25 | 343 |
| TOAMASINA | RURAL | 975,233 | 78,178 | 8.02 | 822,005 | 1,128,461 | 9.04 | 600 |
| MAHAJANGA | URBAIN | 3,030,720 | 392,601 | 12.95 | 2,261,221 | 3,800,218 | 1.96 | 255 |
| MAHAJANGA | RURAL | 720,450 | 85,131 | 11.82 | 553,593 | 887,307 | 8.55 | 495 |
| TOLIARY | URBAIN | 1,860,556 | 406,398 | 21.84 | 1,064,016 | 2,657,097 | 8.67 | 209 |
| TOLIARY | RURAL | 2,037,822 | 334,227 | 16.40 | 1,382,737 | 2,692,908 | 8.09 | 511 |
| ANTSIRANANA | URBAIN | 5,646,517 | 370,935 | 6.57 | 4,919,484 | 6,373,550 | 0.44 | 174 |
| ANTSIRANANA | RURAL | 1,178,201 | 95,130 | 8.07 | 991,747 | 1,364,655 | 2.01 | 467 |
| FARITANY PAR MILIEU2 (STRATE) | | | | | | | | |
| ANTANANARIVO | GCU | 5,239,866 | 326,353 | 6.23 | 4,600,214 | 5,879,517 | 1.02 | 1,223 |
| ANTANANARIVO | CUS | 6,395,923 | 1,208,610 | 18.90 | 4,027,048 | 8,764,798 | 1.31 | 94 |
| ANTANANARIVO | RURAL | 1,791,673 | 198,727 | 11.09 | 1,402,167 | 2,181,178 | 7.98 | 881 |
| FIANARANTSOA | GCU | 2,977,151 | 685,748 | 23.03 | 1,633,085 | 4,321,217 | 1.73 | 140 |
| FIANARANTSOA | CUS | 2,045,343 | 181,947 | 8.90 | 1,688,727 | 2,401,960 | 2.13 | 89 |
| FIANARANTSOA | RURAL | 928,235 | 97,065 | 10.46 | 737,987 | 1,118,483 | 5.11 | 745 |
| TOAMASINA | GCU | 5,050,489 | 342,763 | 6.79 | 4,378,673 | 5,722,304 | 0.76 | 220 |
| TOAMASINA | CUS | 2,596,589 | 550,119 | 21.19 | 1,518,356 | 3,674,821 | 5.66 | 123 |
| TOAMASINA | RURAL | 975,233 | 78,178 | 8.02 | 822,005 | 1,128,461 | 9.04 | 600 |
| MAHAJANGA | GCU | 4,364,489 | 573,367 | 13.14 | 3,240,689 | 5,488,288 | 0.82 | 180 |
| MAHAJANGA | CUS | 2,177,748 | 528,545 | 24.27 | 1,141,799 | 3,213,697 | 10.11 | 75 |
| MAHAJANGA | RURAL | 720,450 | 85,131 | 11.82 | 553,593 | 887,307 | 8.55 | 495 |
| TOLIARY | GCU | 3,930,687 | 197,454 | 5.02 | 3,543,677 | 4,317,696 | 0.24 | 111 |
| TOLIARY | CUS | 1,414,890 | 493,304 | 34.87 | 448,015 | 2,381,766 | 17.33 | 98 |
| TOLIARY | RURAL | 2,037,822 | 334,227 | 16.40 | 1,382,737 | 2,692,908 | 8.09 | 511 |
| ANTSIRANANA | GCU | 6,909,338 | 710,149 | 10.28 | 5,517,446 | 8,301,231 | 0.77 | 100 |
| ANTSIRANANA | CUS | 4,854,883 | 396,829 | 8.17 | 4,077,099 | 5,632,667 | 0.29 | 74 |
| ANTSIRANANA | RURAL | 1,178,201 | 95,130 | 8.07 | 991,747 | 1,364,655 | 2.01 | 467 |

2. Dépenses totales par ménage selon le groupe socioéconomique (en FMG) - Continuation

| Categorie | Valeur | Erreur Type | C.V. (%) | Intervalle de confiance 95% Lim. Infer. | Lim. Super. | Effet de Sondage | Nombre de Ménages |
|---|---|---|---|---|---|---|---|
| GROUPE SOCIOÉCONOMIQUE | | | | | | | |
| GRAND EXPL.AGRIC | 1,543,491 | 168,798 | 10.94 | 1,212,646 | 1,874,336 | 2.33 | 343 |
| MOYEN EXPL.AGRIC | 1,106,826 | 66,413 | 6.00 | 976,657 | 1,236,995 | 1.36 | 1,007 |
| PETIT EXPL.AGRIC | 1,003,084 | 47,855 | 4.77 | 909,287 | 1,096,880 | 5.54 | 2,002 |
| ELEVEUR | 1,699,855 | 343,178 | 20.19 | 1,027,227 | 2,372,483 | 1.10 | 26 |
| PECHEUR | 1,697,858 | 401,390 | 23.64 | 911,132 | 2,484,583 | 5.41 | 56 |
| CHASSEURS | 437,050 | 112,937 | 25.84 | 215,694 | 658,407 | 1.37 | 2 |
| PETIT EXP.N.AGR. | 3,306,198 | 406,316 | 12.29 | 2,509,818 | 4,102,578 | 1.04 | 158 |
| COMMERCANTS | 4,193,314 | 272,094 | 6.49 | 3,660,010 | 4,726,619 | 0.64 | 341 |
| PETIT ENT. SERV. | 8,055,188 | 3,159,441 | 39.22 | 1,862,684 | 14,247,693 | 1.83 | 86 |
| GR.MOY.ENT.SERV. | 19,131,272 | 5,392,465 | 28.19 | 8,562,040 | 29,700,503 | 1.23 | 17 |
| CADRE SALARIE | 5,111,201 | 303,964 | 5.95 | 4,515,432 | 5,706,970 | 1.11 | 426 |
| EMPLOYE, OUVRIER | 3,470,829 | 140,098 | 4.04 | 3,196,236 | 3,745,421 | 0.87 | 754 |
| M.D'OUVRE S.QUAL | 2,160,284 | 155,871 | 7.22 | 1,854,777 | 2,465,792 | 1.39 | 223 |
| AIDES FAMILIAUX | 2,200,945 | 637,960 | 28.99 | 950,542 | 3,451,347 | 0.44 | 7 |
| FEM./HOM.AU FOYE | 1,346,855 | 324,030 | 24.06 | 711,756 | 1,981,955 | 1.07 | 31 |
| CLERGE | 2,548,661 | 711,902 | 27.93 | 1,153,334 | 3,943,989 | 1.09 | 16 |
| APPRENTI | 1,722,270 | 738,719 | 42.89 | 274,382 | 3,170,158 | 0.67 | 3 |
| AUTRES | 3,142,433 | 210,429 | 6.70 | 2,729,992 | 3,554,874 | 0.98 | 728 |

3. Dépenses alimentaires par ménage selon le domaine géographique (en FMG)

| Categorie | Valeur | Erreur Type | C.V. (%) | Intervalle de confiance 95% Lim. Infer. | Lim. Super. | Effet de Sondage | Nombre de Ménages |
|---|---|---|---|---|---|---|---|
| MADAGASCAR | 1,309,884 | 46,086 | 3.52 | 1,219,555 | 1,400,213 | 3.27 | 6,226 |
| FARITANY | | | | | | | |
| ANTANANARIVO | 1,912,120 | 98,078 | 5.13 | 1,719,887 | 2,104,353 | 2.34 | 2,198 |
| FIANARANTSOA | 809,951 | 57,405 | 7.09 | 697,437 | 922,465 | 4.72 | 974 |
| TOAMASINA | 1,050,857 | 75,261 | 7.16 | 903,346 | 1,198,368 | 3.10 | 943 |
| MAHAJANGA | 799,427 | 73,053 | 9.14 | 656,243 | 942,612 | 2.77 | 750 |
| TOLIARY | 1,504,366 | 201,284 | 13.38 | 1,109,849 | 1,898,883 | 7.92 | 720 |
| ANTSIRANANA | 1,377,746 | 82,791 | 6.01 | 1,215,476 | 1,540,016 | 0.66 | 641 |
| MILIEU | | | | | | | |
| URBAIN | 2,575,799 | 89,411 | 3.47 | 2,400,554 | 2,751,045 | 1.30 | 2,527 |
| RURAL | 932,739 | 53,310 | 5.72 | 828,252 | 1,037,226 | 6.74 | 3,699 |
| MILIEU2 | | | | | | | |
| GCU | 3,248,913 | 106,966 | 3.29 | 3,039,260 | 3,458,566 | 0.69 | 1,974 |
| CUS | 1,944,932 | 141,266 | 7.26 | 1,668,051 | 2,221,812 | 2.73 | 553 |
| RURAL | 932,739 | 53,310 | 5.72 | 828,252 | 1,037,226 | 6.74 | 3,699 |
| MILIEU3 | | | | | | | |
| CAPITAL | 3,562,125 | 157,331 | 4.42 | 3,253,757 | 3,870,493 | 0.55 | 1,050 |
| GCU | 2,897,036 | 120,368 | 4.15 | 2,661,115 | 3,132,956 | 0.85 | 924 |
| CUS | 1,944,932 | 141,266 | 7.26 | 1,668,051 | 2,221,812 | 2.73 | 553 |
| RURAL | 932,739 | 53,310 | 5.72 | 828,252 | 1,037,226 | 6.74 | 3,699 |

3. Dépenses alimentaires par ménage selon le domaine géographique (en FMG) - Continuation

| Categorie | | Valeur | Erreur Type | C.V. (%) | Intervalle de confiance 95% Lim. Infer. | Lim. Super. | Effet de Sondage | Nombre de Ménages |
|---|---|---|---|---|---|---|---|---|
| FARITANY PAR MILIEU | | | | | | | | |
| ANTANANARIVO | URBAIN | 3,298,228 | 130,149 | 3.95 | 3,043,137 | 3,553,320 | 0.71 | 1,317 |
| ANTANANARIVO | RURAL | 1,258,613 | 129,660 | 10.30 | 1,004,479 | 1,512,747 | 6.86 | 881 |
| FIANARANTSOA | URBAIN | 1,652,458 | 133,250 | 8.06 | 1,391,287 | 1,913,628 | 2.20 | 229 |
| FIANARANTSOA | RURAL | 646,180 | 62,551 | 9.68 | 523,580 | 768,781 | 7.43 | 745 |
| TOAMASINA | URBAIN | 2,382,552 | 273,821 | 11.49 | 1,845,863 | 2,919,241 | 3.18 | 343 |
| TOAMASINA | RURAL | 687,936 | 61,293 | 8.91 | 567,801 | 808,071 | 8.00 | 600 |
| MAHAJANGA | URBAIN | 2,036,270 | 228,538 | 11.22 | 1,588,337 | 2,484,204 | 2.31 | 255 |
| MAHAJANGA | RURAL | 464,708 | 68,302 | 14.70 | 330,836 | 598,581 | 8.33 | 495 |
| TOLIARY | URBAIN | 1,426,993 | 337,076 | 23.62 | 766,323 | 2,087,663 | 8.57 | 209 |
| TOLIARY | RURAL | 1,524,835 | 238,502 | 15.64 | 1,057,371 | 1,992,298 | 7.85 | 511 |
| ANTSIRANANA | URBAIN | 4,050,485 | 294,285 | 7.27 | 3,473,687 | 4,627,283 | 0.47 | 174 |
| ANTSIRANANA | RURAL | 784,638 | 81,265 | 10.36 | 625,358 | 943,918 | 3.08 | 467 |
| FARITANY PAR MILIEU2 (STRATE) | | | | | | | | |
| ANTANANARIVO | GCU | 3,334,199 | 157,965 | 4.74 | 3,024,588 | 3,643,810 | 0.70 | 1,223 |
| ANTANANARIVO | CUS | 3,167,189 | 185,687 | 5.86 | 2,803,243 | 3,531,135 | 0.82 | 94 |
| ANTANANARIVO | RURAL | 1,258,613 | 129,660 | 10.30 | 1,004,479 | 1,512,747 | 6.86 | 881 |
| FIANARANTSOA | GCU | 1,961,372 | 308,061 | 15.71 | 1,357,574 | 2,565,171 | 2.82 | 140 |
| FIANARANTSOA | CUS | 1,538,490 | 144,441 | 9.39 | 1,255,386 | 1,821,594 | 2.05 | 89 |
| FIANARANTSOA | RURAL | 646,180 | 62,551 | 9.68 | 523,580 | 768,781 | 7.43 | 745 |
| TOAMASINA | GCU | 3,517,154 | 159,758 | 4.54 | 3,204,028 | 3,830,280 | 0.33 | 220 |
| TOAMASINA | CUS | 1,829,776 | 397,172 | 21.71 | 1,051,320 | 2,608,232 | 5.69 | 123 |
| TOAMASINA | RURAL | 687,936 | 61,293 | 8.91 | 567,801 | 808,071 | 8.00 | 600 |
| MAHAJANGA | GCU | 2,769,771 | 282,822 | 10.21 | 2,215,440 | 3,324,102 | 0.78 | 180 |
| MAHAJANGA | CUS | 1,567,183 | 328,076 | 20.93 | 924,154 | 2,210,211 | 8.51 | 75 |
| MAHAJANGA | RURAL | 464,708 | 68,302 | 14.70 | 330,836 | 598,581 | 8.33 | 495 |
| TOLIARY | GCU | 3,000,839 | 150,625 | 5.02 | 2,705,614 | 3,296,065 | 0.22 | 111 |
| TOLIARY | CUS | 1,088,169 | 409,308 | 37.61 | 285,925 | 1,890,413 | 15.35 | 98 |
| TOLIARY | RURAL | 1,524,835 | 238,502 | 15.64 | 1,057,371 | 1,992,298 | 7.85 | 511 |
| ANTSIRANANA | GCU | 4,531,645 | 322,761 | 7.12 | 3,899,033 | 5,164,257 | 0.66 | 100 |
| ANTSIRANANA | CUS | 3,748,857 | 429,847 | 11.47 | 2,906,357 | 4,591,356 | 0.44 | 74 |
| ANTSIRANANA | RURAL | 784,638 | 81,265 | 10.36 | 625,358 | 943,918 | 3.08 | 467 |

3. Dépenses alimentaires par ménage selon le groupe socioéconomique (en FMG) - Continuation

| Categorie | Valeur | Erreur Type | C.V. (%) | Intervalle de confiance 95% Lim. Infer. | Lim. Super. | Effet de Sondage | Nombre de Ménages |
|---|---|---|---|---|---|---|---|
| GROUPE SOCIOÉCONOMIQUE | | | | | | | |
| GRAND EXPL.AGRIC | 1,094,718 | 136,195 | 12.44 | 827,776 | 1,361,661 | 2.00 | 343 |
| MOYEN EXPL.AGRIC | 787,424 | 56,473 | 7.17 | 676,736 | 898,111 | 1.47 | 1,007 |
| PETIT EXPL.AGRIC | 732,877 | 38,745 | 5.29 | 656,937 | 808,817 | 5.39 | 2,002 |
| ELEVEUR | 1,251,698 | 295,528 | 23.61 | 672,463 | 1,830,934 | 1.45 | 26 |
| PECHEUR | 1,269,677 | 335,733 | 26.44 | 611,641 | 1,927,713 | 5.50 | 56 |
| CHASSEURS | 293,399 | 105,655 | 36.01 | 86,315 | 500,483 | 1.37 | 2 |
| PETIT EXP.N.AGR. | 2,370,920 | 312,915 | 13.20 | 1,757,607 | 2,984,233 | 1.01 | 158 |
| COMMERCANTS | 2,811,768 | 196,072 | 6.97 | 2,427,466 | 3,196,070 | 0.63 | 341 |
| PETIT ENT. SERV. | 3,201,852 | 383,448 | 11.98 | 2,450,295 | 3,953,410 | 1.29 | 86 |
| GR.MOY.ENT.SERV. | 7,677,787 | 1,508,275 | 19.64 | 4,721,567 | 10,634,007 | 0.70 | 17 |
| CADRE SALARIE | 3,188,255 | 157,145 | 4.93 | 2,880,252 | 3,496,259 | 1.23 | 426 |
| EMPLOYE, OUVRIER | 2,515,553 | 116,202 | 4.62 | 2,287,797 | 2,743,309 | 0.95 | 754 |
| M.D'OUVRE S.QUAL | 1,717,901 | 131,789 | 7.67 | 1,459,596 | 1,976,207 | 1.58 | 223 |
| AIDES FAMILIAUX | 1,636,398 | 518,573 | 31.69 | 619,995 | 2,652,801 | 0.43 | 7 |
| FEM./HOM.AU FOYE | 1,112,675 | 328,705 | 29.54 | 468,413 | 1,756,938 | 1.13 | 31 |
| CLERGE | 1,605,540 | 345,844 | 21.54 | 927,687 | 2,283,393 | 0.90 | 16 |
| APPRENTI | 1,203,542 | 529,225 | 43.97 | 166,260 | 2,240,824 | 0.67 | 3 |
| AUTRES | 2,034,459 | 104,817 | 5.15 | 1,829,018 | 2,239,900 | 1.29 | 728 |

4. Dépenses non-alimentaires par ménage selon le domaine géographique (en FMG)

| Categorie | Valeur | Erreur Type | C.V. (%) | Intervalle de confiance 95% Lim. Infer. | Lim. Super. | Effet de Sondage | Nombre de Ménages |
|---|---|---|---|---|---|---|---|
| MADAGASCAR | 613,195 | 34,549 | 5.63 | 545,479 | 680,911 | 1.56 | 6,226 |
| FARITANY | | | | | | | |
| ANTANANARIVO | 1,064,157 | 108,623 | 10.21 | 851,257 | 1,277,057 | 1.45 | 2,198 |
| FIANARANTSOA | 340,963 | 39,863 | 11.69 | 262,832 | 419,093 | 1.59 | 974 |
| TOAMASINA | 443,768 | 33,877 | 7.63 | 377,369 | 510,167 | 2.30 | 943 |
| MAHAJANGA | 413,075 | 44,537 | 10.78 | 325,782 | 500,368 | 1.63 | 750 |
| TOLIARY | 496,372 | 84,984 | 17.12 | 329,804 | 662,939 | 3.24 | 720 |
| ANTSIRANANA | 611,943 | 38,874 | 6.35 | 535,749 | 688,136 | 0.56 | 641 |
| MILIEU | | | | | | | |
| URBAIN | 1,369,786 | 115,822 | 8.46 | 1,142,775 | 1,596,798 | 1.15 | 2,527 |
| RURAL | 387,789 | 28,070 | 7.24 | 332,771 | 442,807 | 3.90 | 3,699 |
| MILIEU2 | | | | | | | |
| GCU | 1,745,938 | 127,647 | 7.31 | 1,495,749 | 1,996,126 | 0.79 | 1,974 |
| CUS | 1,017,243 | 190,893 | 18.77 | 643,092 | 1,391,394 | 1.44 | 553 |
| RURAL | 387,789 | 28,070 | 7.24 | 332,771 | 442,807 | 3.90 | 3,699 |
| MILIEU3 | | | | | | | |
| CAPITAL | 2,124,790 | 202,841 | 9.55 | 1,727,221 | 2,522,359 | 0.68 | 1,050 |
| GCU | 1,320,318 | 123,015 | 9.32 | 1,079,208 | 1,561,428 | 0.92 | 924 |
| CUS | 1,017,243 | 190,893 | 18.77 | 643,092 | 1,391,394 | 1.44 | 553 |
| RURAL | 387,789 | 28,070 | 7.24 | 332,771 | 442,807 | 3.90 | 3,699 |

4. Dépenses non-alimentaires par ménage selon le domaine géographique (en FMG) - Continuation

| Categorie | | Valeur | Erreur Type | C.V. (%) | Intervalle de confiance 95% Lim. Infer. | Lim. Super. | Effet de Sondage | Nombre de Ménages |
|---|---|---|---|---|---|---|---|---|
| FARITANY PAR MILIEU | | | | | | | | |
| ANTANANARIVO | URBAIN | 2,190,630 | 283,983 | 12.96 | 1,634,024 | 2,747,236 | 1.14 | 1,317 |
| ANTANANARIVO | RURAL | 533,060 | 82,152 | 15.41 | 372,042 | 694,079 | 6.56 | 881 |
| FIANARANTSOA | URBAIN | 644,010 | 111,676 | 17.34 | 425,125 | 862,896 | 0.92 | 229 |
| FIANARANTSOA | RURAL | 282,055 | 42,025 | 14.90 | 199,685 | 364,425 | 1.97 | 745 |
| TOAMASINA | URBAIN | 1,017,921 | 137,314 | 13.49 | 748,786 | 1,287,056 | 2.38 | 343 |
| TOAMASINA | RURAL | 287,297 | 22,100 | 7.69 | 243,981 | 330,612 | 6.00 | 600 |
| MAHAJANGA | URBAIN | 994,449 | 187,961 | 18.90 | 626,046 | 1,362,852 | 1.56 | 255 |
| MAHAJANGA | RURAL | 255,742 | 24,393 | 9.54 | 207,930 | 303,553 | 5.66 | 495 |
| TOLIARY | URBAIN | 433,563 | 73,633 | 16.98 | 289,243 | 577,884 | 5.08 | 209 |
| TOLIARY | RURAL | 512,987 | 105,738 | 20.61 | 305,741 | 720,234 | 3.21 | 511 |
| ANTSIRANANA | URBAIN | 1,596,032 | 162,769 | 10.20 | 1,277,004 | 1,915,060 | 0.49 | 174 |
| ANTSIRANANA | RURAL | 393,563 | 30,832 | 7.83 | 333,133 | 453,993 | 1.18 | 467 |
| FARITANY PAR MILIEU2 (STRATE) | | | | | | | | |
| ANTANANARIVO | GCU | 1,905,666 | 187,040 | 9.81 | 1,539,067 | 2,272,266 | 0.78 | 1,223 |
| ANTANANARIVO | CUS | 3,228,733 | 1,107,006 | 34.29 | 1,059,001 | 5,398,466 | 1.33 | 94 |
| ANTANANARIVO | RURAL | 533,060 | 82,152 | 15.41 | 372,042 | 694,079 | 6.56 | 881 |
| FIANARANTSOA | GCU | 1,015,778 | 397,990 | 39.18 | 235,719 | 1,795,838 | 0.96 | 140 |
| FIANARANTSOA | CUS | 506,854 | 48,888 | 9.65 | 411,034 | 602,674 | 0.94 | 89 |
| FIANARANTSOA | RURAL | 282,055 | 42,025 | 14.90 | 199,685 | 364,425 | 1.97 | 745 |
| TOAMASINA | GCU | 1,533,335 | 223,637 | 14.58 | 1,095,007 | 1,971,662 | 1.37 | 220 |
| TOAMASINA | CUS | 766,812 | 171,517 | 22.37 | 430,640 | 1,102,985 | 3.86 | 123 |
| TOAMASINA | RURAL | 287,297 | 22,100 | 7.69 | 243,981 | 330,612 | 6.00 | 600 |
| MAHAJANGA | GCU | 1,594,718 | 318,269 | 19.96 | 970,911 | 2,218,525 | 0.81 | 180 |
| MAHAJANGA | CUS | 610,565 | 230,971 | 37.83 | 157,862 | 1,063,269 | 9.75 | 75 |
| MAHAJANGA | RURAL | 255,742 | 24,393 | 9.54 | 207,930 | 303,553 | 5.66 | 495 |
| TOLIARY | GCU | 929,847 | 76,147 | 8.19 | 780,600 | 1,079,095 | 0.41 | 111 |
| TOLIARY | CUS | 326,721 | 88,252 | 27.01 | 153,747 | 499,695 | 14.52 | 98 |
| TOLIARY | RURAL | 512,987 | 105,738 | 20.61 | 305,741 | 720,234 | 3.21 | 511 |
| ANTSIRANANA | GCU | 2,377,693 | 418,212 | 17.59 | 1,557,999 | 3,197,388 | 0.58 | 100 |
| ANTSIRANANA | CUS | 1,106,026 | 42,670 | 3.86 | 1,022,393 | 1,189,660 | 0.13 | 74 |
| ANTSIRANANA | RURAL | 393,563 | 30,832 | 7.83 | 333,133 | 453,993 | 1.18 | 467 |

4. Dépenses non-alimentaires par ménage selon le groupe socioéconomique (en FMG) - Continuation

| Categorie | Valeur | Erreur Type | C.V. (%) | Intervalle de confiance 95% Lim. Infer. | Lim. Super. | Effet de Sondage | Nombre de Ménages |
|---|---|---|---|---|---|---|---|
| GROUPE SOCIOÉCONOMIQUE | | | | | | | |
| GRAND EXPL.AGRIC | 448,772 | 48,851 | 10.89 | 353,025 | 544,520 | 2.78 | 343 |
| MOYEN EXPL.AGRIC | 319,402 | 18,213 | 5.70 | 283,704 | 355,100 | 1.20 | 1,007 |
| PETIT EXPL.AGRIC | 270,206 | 11,048 | 4.09 | 248,553 | 291,859 | 3.75 | 2,002 |
| ELEVEUR | 448,157 | 81,069 | 18.09 | 289,261 | 607,053 | 0.52 | 26 |
| PECHEUR | 428,180 | 75,548 | 17.64 | 280,106 | 576,255 | 2.01 | 56 |
| CHASSEURS | 143,652 | 7,282 | 5.07 | 129,380 | 157,924 | 1.37 | 2 |
| PETIT EXP.N.AGR. | 935,278 | 151,549 | 16.20 | 638,242 | 1,232,314 | 0.72 | 158 |
| COMMERCANTS | 1,381,546 | 133,123 | 9.64 | 1,120,625 | 1,642,467 | 0.65 | 341 |
| PETIT ENT. SERV. | 4,853,336 | 2,826,534 | 58.24 | -686,670 | 10,393,342 | 1.76 | 86 |
| GR.MOY.ENT.SERV. | 11,453,485 | 4,942,065 | 43.15 | 1,767,038 | 21,139,931 | 1.27 | 17 |
| CADRE SALARIE | 1,922,946 | 197,543 | 10.27 | 1,535,762 | 2,310,131 | 0.93 | 426 |
| EMPLOYE, OUVRIER | 955,275 | 40,074 | 4.20 | 876,730 | 1,033,821 | 0.63 | 754 |
| M.D'OUVRE S.QUAL | 442,383 | 34,696 | 7.84 | 374,378 | 510,388 | 0.70 | 223 |
| AIDES FAMILIAUX | 564,546 | 121,125 | 21.46 | 327,142 | 801,950 | 0.48 | 7 |
| FEM./HOM.AU FOYE | 234,180 | 26,809 | 11.45 | 181,635 | 286,725 | 0.67 | 31 |
| CLERGE | 943,121 | 387,613 | 41.10 | 183,400 | 1,702,843 | 1.24 | 16 |
| APPRENTI | 518,728 | 209,862 | 40.46 | 107,398 | 930,058 | 0.67 | 3 |
| AUTRES | 1,107,974 | 130,042 | 11.74 | 853,092 | 1,362,856 | 0.68 | 728 |

5. Distribution des ménages selon l’activité principale du chef (en pourcentage)

| Categorie | Valeur (%) | Erreur Type | C.V. (%) | Intervalle de confiance 95% Lim. Infer. | Lim. Super. | Effet de Sondage | Nombre de Ménages |
|---|---|---|---|---|---|---|---|
| GROUPE SOCIOÉCONOMIQUE | | | | | | | |
| GRAND EXPL.AGRIC | 7.2 | 0.9 | 12.72 | 5.4 | 9.0 | 7.81 | 343 |
| MOYEN EXPL.AGRIC | 20.1 | 1.3 | 6.54 | 17.5 | 22.7 | 6.70 | 1,007 |
| PETIT EXPL.AGRIC | 40.6 | 1.9 | 4.62 | 36.9 | 44.3 | 9.08 | 2,002 |
| ELEVEUR | 0.5 | 0.1 | 27.79 | 0.2 | 0.7 | 2.30 | 26 |
| PECHEUR | 1.0 | 0.5 | 45.65 | 0.1 | 1.9 | 13.02 | 56 |
| CHASSEURS | 0.0 | 0.0 | 70.70 | -0.0 | 0.1 | 1.37 | 2 |
| PETIT EXP.N.AGR. | 2.2 | 0.4 | 18.63 | 1.4 | 2.9 | 4.77 | 158 |
| COMMERCANTS | 4.1 | 0.4 | 10.06 | 3.3 | 4.9 | 2.67 | 341 |
| PETIT ENT. SERV. | 0.7 | 0.1 | 13.66 | 0.5 | 0.9 | 0.87 | 86 |
| GR.MOY.ENT.SERV. | 0.2 | 0.0 | 24.29 | 0.1 | 0.3 | 0.69 | 17 |
| CADRE SALARIE | 5.1 | 0.4 | 8.68 | 4.2 | 5.9 | 2.50 | 426 |
| EMPLOYE, OUVRIER | 6.9 | 0.5 | 7.48 | 5.9 | 7.9 | 2.59 | 754 |
| M.D'OUVRE S.QUAL | 2.6 | 0.3 | 11.39 | 2.0 | 3.2 | 2.18 | 223 |
| AIDES FAMILIAUX | 0.1 | 0.0 | 46.36 | 0.0 | 0.2 | 1.15 | 7 |
| FEM./HOM.AU FOYE | 0.6 | 0.2 | 34.12 | 0.2 | 0.9 | 4.10 | 31 |
| CLERGE | 0.3 | 0.1 | 27.73 | 0.1 | 0.5 | 1.45 | 16 |
| APPRENTI | 0.0 | 0.0 | 68.67 | -0.0 | 0.1 | 0.92 | 3 |
| AUTRES | 7.9 | 0.4 | 5.69 | 7.0 | 8.7 | 1.72 | 728 |