Principles of Sample Generation and Estimation of Indices of Sample Survey of Population (Households) regarding Their Economic Activity

For the purpose of conducting a sample survey of population (households) regarding their economic activity (hereinafter - SEAP) a national territorial probability sample of non-institutional households has been built. The validity period of the geographical sample for the survey is as follows:
- in urban settlements – January, 2009 – December, 2013;
- in rural settlements – May, 2008 – December, 2013.

The monthly sampling frame in 2010 averaged 16 600 households. All members of the selected households aged 15 - 70 years old inclusive were subject to interviewing. The results of the survey by means of statistical weighting cover the entire population of Ukraine of the specified age.

1. The procedure of sample generation

A sample is generated on the basis of the stratified multistage sampling procedure.
The general diagram of generating a household sampling frame is shown in Figure 1.

The sample generation procedure encompasses the following stages:
1) exclusion of areas that cannot be surveyed;
2) exclusion of population which is not subject to survey;
3) stratification of general population;
4) selection of first-stage geographical units;
5) selection of second-stage geographical units;
6) selection of households.

In the process of generating a geographical sample rural settlements (village councils) located in the alienation zone (zone I) and the zone of unconditional (mandatory) evacuation (zone II) of radiation pollution resulting from the Chernobyl nuclear power disaster are excluded from the total territory of Ukraine. Accordingly, the population living in this area is also excluded from the population of Ukraine and related areas.

The institutional population - conscripts, prisoners, persons permanently residing in nursing homes, senior centres etc., are excluded from the population.

Stratification of the general population is carried out for the purpose of adequate reflection of main features of the administrative territorial division of Ukraine and for ensuring selection from more homogeneous groups of households by their main characteristics. To this end, sample generation involves detection of strata which, within each region (27 regions: the Autonomous Republic of Crimea, 24 regions and the cities of Kyiv and Sevastopol) correspond to cities and city councils with a population of 100,000 people and more (hereinafter - the “cities” of the region strata), towns, city councils and town councils with population of less than 100,000 people (hereinafter - the “towns” of the region strata) and districts (except urban districts), i.e. administrative districts in rural areas (hereinafter - the “districts” of the region strata).
Thus, 52 strata have been formed in urban settlements (according to the number of regions and with allowance for lack of a “towns” stratum in Kyiv and Sevastopol municipalities), and 490 strata have been formed in rural areas (according to the number of rural areas). The size of the sample is distributed between the strata according to their size (population of a stratum) and the need to ensure a certain level of reliability of key indicators[1].

[1] Methods of generating sampling frames for conducting sample surveys of population (households) in 2009-2013: living conditions of households, economic activity of population and agricultural activities of the population in rural areas. - K: State Statistics Committee, 2009, in Ukrainian.

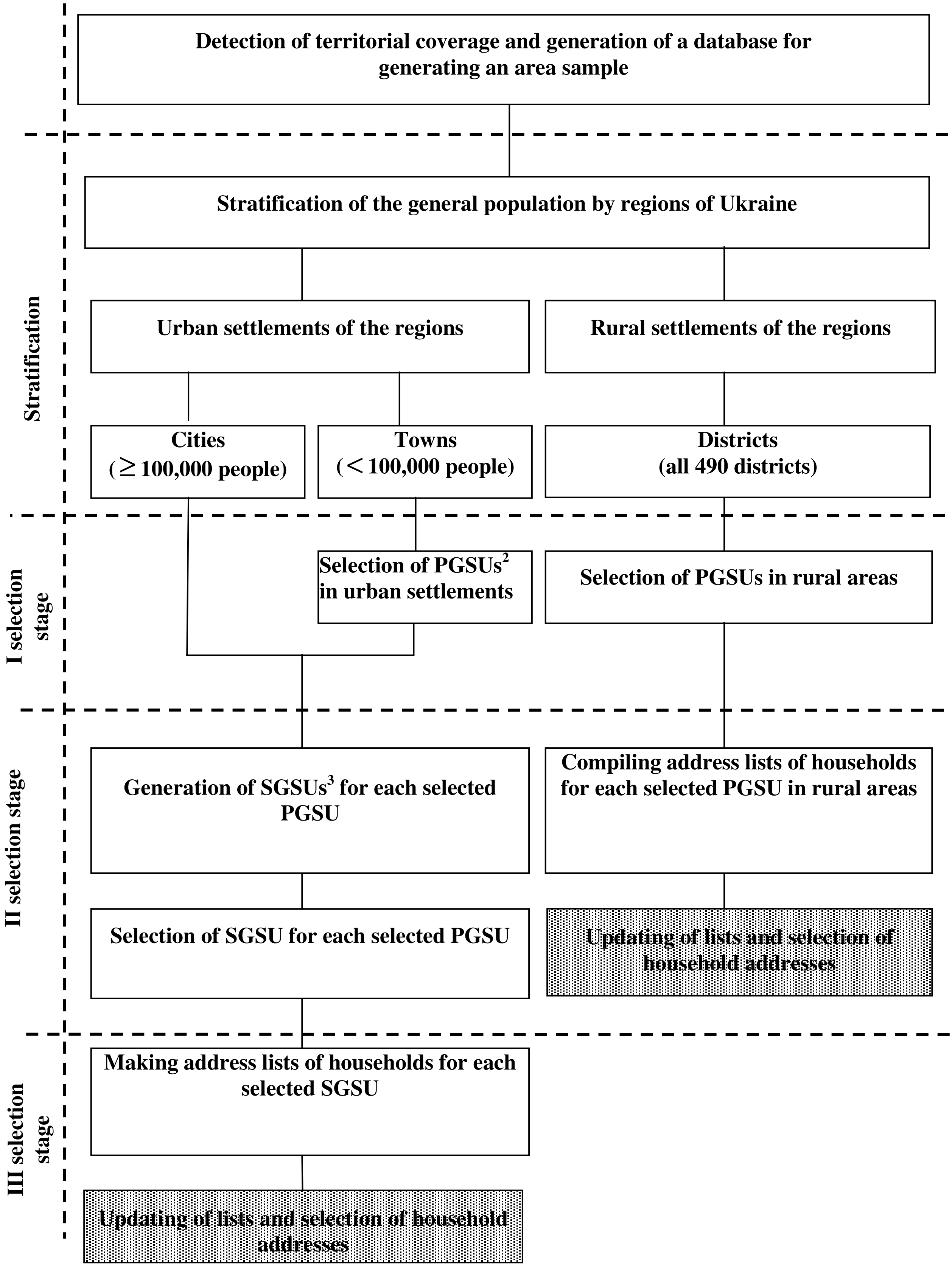

Figure 1. A sample generation diagram for base sample surveys of population (households) in 2009-2013.

Note. Marked in grey are stages of sample generation implemented annually; marked in white are stages implemented once for the entire validity period of an area sample, i.e. from 2009 to 2013 inclusive.

[2] PGSUs are primary geographical sampling units: in urban settlements - cities and urban-type settlements, in rural areas - village councils.

[3] SGSUs are secondary geographical sampling units: in urban areas - instructor stations of the 30 September, 2007 early parliamentary elections.

1.1 Generation and selection of primary geographical sampling units

This stage of sample generation involves selection of city and town councils in urban settlements and village councils in rural areas. Selected first-stage geographical units are primary geographical sampling units.

In urban settlements, generation of a sampling frame and selection of PGSUs goes separately within each "city" and "town" stratum of the region. Cities, urban-type settlements, city and village councils are accepted as geographical units. PGSUs within the "city" strata are selected with probability 1, i.e. the sample includes all cities and city councils without selection, whereas within the “town” strata they are selected based on a random sampling mechanism with probability proportional to the size – amount of a PTSU population. Geographical units in the “towns” strata are arranged on the basis of their size. This increases the representativity of an area sample in terms of representation of towns of different sizes.

A number of interviewers and a sample size are identified along with PGSUs selection.

The terms of conducting surveys stipulate that every PGSU of a “town” stratum shall be examined by one interviewer.

In rural areas, generation of a sampling frame and selection of PGSUs is carried out within each “district” stratum generated in the rural areas of the regions. Separate village councils or groups of village councils formed with due regard for exceeding a set minimum total number of households are accepted as PGSUs.

PGSUs in the rural areas are selected within each stratum based on a random sampling mechanism with probability proportional to the size – a number of households in a PGSU. The PGSUs for conducting SEAP are selected among the PGSUs selected for a survey of agricultural activity of the population (SAAP) living in rural areas and arranged according to a “geographical serpentine” pattern.

Approximately one third of the PGSUs covered by the 2004-2008 SEAP was preserved in rural areas in the process of generating an area sample for conducting SEAP in 2009-2013 in order to improve the reliability of estimation of changes and the possibility to mitigate the transition in terms of the continuity of time series.

1.2 Generation and selection of secondary geographical sampling units

Generation and selection of secondary geographical sampling units is the second stage of the procedure of generation of a household sampling frame in urban settlements. SGSUs are selected within each PGSU during generation of an area sample once for its entire validity period.

The basis for SGSUs generation in urban settlements is the Central Election Commission of Ukraine information and analytical data about location and characteristics of the polling stations in certain Ukrainian city (village) councils based on the materials of the 30 September, 2007 early parliamentary elections. The choice of this source ensues from the fact that the 2001 census data which were used in the previous five-year area sample cycle are no longer up to date.

Selection of SGSUs is based on a random sampling mechanism with probability proportional to the size - the number of the electorate). All units of the second stage of selection within each selected city or village council are arranged according to a “geographical serpentine” pattern. Each interviewer surveys four SGSUs.

The distribution of the total number of SGSUs in urban settlements and of PGSUs in rural areas for conducting SEAP in 2009-2013 is shown in Table 1.

Table 1

Distribution of the total number of selected geographical units for conducting SEAP by the regions in 2009-2013

	Total SGSUs in urban settlements	including in		Total PGSUs in rural areas
		cities	towns	
Ukraine	1544	828	716	584
Autonomous Republic of Crimea	64	36	28	17
Vinnytsia	52	20	32	29
Volyn	44	16	28	19
Dnipropetrovsk	120	88	32	24
Donetsk	140	88	52	24
Zhytomyr	36	12	24	24
Transcarpathean	36	8	28	21
Zaporizhia	68	48	20	27
Ivano-Frankivsk	36	16	20	21
Kyiv	60	12	48	26
Kirovohrad	48	20	28	22
Luhansk	76	40	36	23
Lviv	68	28	40	31
Mykolaiv	48	24	24	21
Odesa	72	40	32	28
Poltava	40	20	20	27
Rivne	28	12	16	18
Sumy	40	12	28	20
Ternopil	32	12	20	21
Kharkiv	88	48	40	29
Kherson	52	24	28	29
Khmelnytsk	40	12	28	22
Cherkasy	44	16	28	21
Chernivtsi	32	16	16	17
Chernihiv	36	16	20	23
Kyiv (city council)	112	112	x	x
Sevastopol (city council)	32	32	x	x

1.2 Selection of households

The last stage of sample generation provides for selection of households. For this purpose a sampling frame – a specially arranged full list of household addresses (updated lists of residents or lists according to House Registers) is created for each SGSU in urban settlements and on each PGSU in rural areas. Households are selected from the sampling frame through a systematic sampling procedure, and at the same time rotation groups are created in accordance with a household rotation pattern in the sample.

2. Rotation of the households in a sample

The use of a household rotation pattern in the sample makes it possible to improve the reliability of comparisons of indicators in movement, which ensures updating of a sampling frame for each observation period and at the same time makes it possible to keep a certain part of it unchanged.

According to the existing rotation pattern (Fig. 2) each selected household is surveyed for a total of 6 periods: during three consecutive months followed by a nine month break, and then it is interviewed for another three months. Thus the household stays in the sample for 15 months but is interviewed only 6 times.

	Months											
	01	*02*	*03*	*04*	*05*	*06*	*07*	*08*	*09*	*10*	*11*	*12*
Rotation groups	33											
	34	34										
	35	35	35									
		36	36	36								
			37	37	37							
				38	38	38						
					39	39	39					
						40	40	40				
							41	41	41			
								42	42	42		
									43	43	43	
										44	44	44
											45	45
												46
	45											
	46	46										
	47	47	47									
		48	48	48								
			49	49	49							
				50	50	50						
					51	51	51					
						52	52	52				
							53	53	53			
								54	54	54		
									55	55	55	
										56	56	56
											57	57
												58

Figure 2. Household rotation pattern of the 2011 sampling frame

3. Application of survey results to general population

Application of selective data obtained from the answers of respondents aged 15-70 to the general population - the entire population of Ukraine of specified age, is the basis for calculation of the indicators of economic activity, employment and unemployment which are obtained from sample surveys reports. For this purpose, based on the provisions of "The Methods of Calculation of Statistical Weights For Application of Results of Sample Survey of Population (Households) Regarding Their Economic Activity to the General Population", approved by Order No. 520 of the State Statistics Committee of Ukraine of 14.11.2006, we develop a system of statistical weights which are designed primarily to account for the following:

- household selection probabilities;
- actual levels of refusals of households and individuals from participating in the survey;
- harmonization of survey results with the demographic statistics data regarding the number and the gender and age structure of the population etc.

The final weight for the *r-th* person is calculated as the product of the base weight of a household and the corresponding weighting coefficients (reweighting coefficients) by the formula:

$$w_r = w_{B\,r} \cdot k_{1_r} \cdot \ldots \cdot k_{n_r}, \qquad (1)$$

where w_r is the ultimate weight of a person and $w_{B\,r}$ is the base weight of a household a person is a member of;

$k_{1_r},..,k_{n_r}$ are weighting coefficients; n is a number of stages of adjusting statistical weights.

The system of weighting coefficients has to be as detailed as possible in order to reflect effects of various factors on the ultimate weights of persons.

3.1 Finding a base weight

In a household probability sample each household i of the general population has a certain general probability p_i to be included to the sample. In other words, this household represents $1/p_i$ of the general population households and, hence, it has to be calculated $1/p_i$ times. The value $1/p_i$ is a base weight of a household for applying the data to the general population:

$$w_{Bi} = w_{B\,r} = 1/\,p_i\,. \qquad (2)$$

The sample design implemented in the survey stipulates the use of the following components for designing base weights:

– PGSU selection probability – P_{1i};

– probability of selecting SGSUs among urban settlements or selecting households in a rural area – P_{2i};

– probability of selecting households within a GSSU in urban settlements – P_{3i}.

The reciprocal variable of the product of these three probabilities for an urban area or of two probabilities for a rural area will be the base weight of a household.

In a general case of a three-stage sample this condition can be expressed by the following formula:

$$w_{Bi} = w_{B\,r} = \frac{1}{(P_{1i} \cdot P_{2i} \cdot P_{3i})}, \qquad (3)$$

In the process of generating a sampling frame of households for monthly surveys in 2011 the average share of household selection was 1/888, i.e. one household of the sample represented an average of 888 households of the general population.

3.2 Accounting for household non-response

A special system of cells is created to account for household non-responses. The system of cells is a system of attributes to be used for grouping the units of general population and (or) a sampling frame. The main tasks which are fulfilled by creation of cells are accounting for peculiarities of the geographical coverage of a sample in the process of adjusting the statistical weights of households and securing a certain number of households in each cell.

Creation of a system of cells in SEAP involves the use of the following classification variables:

- region - 27 discrete values (the Autonomous Republic of Crimea, 24 oblasts of Ukraine, cities of Kyiv and Sevastopol);
- type of area - 2 discrete values (urban areas, rural areas).

Besides, it accounts for the degree of territorial proximity groups of households which are included in a particular cell.

The above information is used for construction of cells by merging separate geographical units according to a special pattern: SGSUs in urban areas corresponding to PGSUs in rural areas.

The established minimum number of households which were surveyed in each cell is equal to 50. A cell with a number of surveyed households smaller than the established value is merged with another (adjacent) cell belonging to the same administrative unit. Thus, an individual system of cells is created for each survey in order to adjust the statistical weights of households.

The total number of households in cell n_l is determined based on the results of selection of households in geographical units. Non-residential premises, locked rooms and nonexistent addresses are excluded from the total number of households in the cell.

The non-response factor is calculated for each l cell by the formula:

$$k_l = k_{1r} = \frac{\sum_{i=1}^{n} w_{Bi} \cdot \lambda_{li}}{\sum_{i=1}^{n} w_{Bi} \cdot \lambda_{li} \cdot \eta_{li}}, \qquad (4)$$

where $\lambda_{li} = \begin{cases} 1, \text{ якщо } i \in M_l; \\ 0, \text{ якщо } i \notin M_l;. \end{cases}$

M_l – a household set of l cell;

$\eta_{li} = l \begin{cases} 1, & \textit{if a household took part in a survey} \\ 0 & \textit{if a household refused to take part in a survey} \end{cases};$

The refusal factor values k_{2r} are the same within each cell.

The weights of households with account for household non-response w_{1i} are calculated by the following formula:

$$w_{1i} = w_{Bi} \cdot k_{1r}. \qquad (5)$$

3.3 Accounting for non-response of persons

A special system of cells is created based on the aggregate of households surveyed to account for refusals of persons to be surveyed.

The cell code is defined similarly to the cell code for accounting for household non-responses to be surveyed.

The total number of persons in cell n_p is determined based on the results of the interviews for the current month. The non-response factor k_{2r} is calculated for each cell p by the formula:

$$k_p = k_{2r} = \frac{\sum_{i=1}^{n} w_{1i} \cdot \lambda_{pi}}{\sum_{i=1}^{n} w_{1i} \cdot \lambda_{pi} \cdot \eta_{pi}}, \quad (6)$$

where

$$\lambda_{pr} = \begin{cases} 1, & if \quad p \in M_p; \\ 0, & if \quad p \notin M_p; \end{cases}$$

$$\eta_{pi} = \begin{cases} 1, & \textit{if a person took part in the survey} \\ 0, & \textit{If a person refused to take part in the survey} \end{cases};$$

M_p is a set of persons of cell p;

The non-response factor values k_{2r} are the same within each cell.

The weights of persons with account for their non-response to be surveyed w_{2r} are calculated by the following formula:

$$w_{2r} = w_{1i} \cdot k_{2r} \quad (7)$$

3.4 Elimination of post-stratification effects

To eliminate the post-stratification effects, a system of cells can be created based on the following classification variables:
- region - 27 discrete values (the Autonomous Republic of Crimea, 24 regions, cities of Kyiv and Sevastopol);
- type of area - 2 discrete values (urban areas, rural areas);
- respondent gender - 2 discrete values (female, male);
- respondent age group - 11 discrete values (15-19 years old, 20-24 years old, 25-29 years old, 30-34 years old, 35-39 years old, 40-44 years old, 45-49 years old, 50 - 54 years old, 55-59 years old, 60-64 years old, 65-70 years old).

for this purpose the cell code is calculated by the following formula:

c_q = a region attribute * 1000 + an area type attribute * 100 + a gender attribute * 10 + an age group attribute.
The total number of persons in cell n_q is determined by the results of the survey for the current month.
The total number of persons of the general population cell N_q is determined once a year according to the demographic statistics as of January 1 of the current year.

The factors for elimination of post-stratification effects k_{3r} are calculated for each cell q by the following formula:

$$k_q = k_{3r} = \frac{N_q}{\sum_r w_{2r} \cdot \lambda_{qr}}, \quad (8)$$

where $\lambda_{qr} = \begin{cases} 1, & if \quad q \in M_q; \\ 0 & if \quad q \notin M_q; \end{cases}$

M_q is a set of persons of cell q.

The non-response factor values k_{3r} are the same within each cell. The weights of persons after elimination of post-stratification effects 3_{3r} are calculated by the following formula:

$$w_{3r} = w_{2r} \cdot k_{3r} \qquad (9)$$

The ultimate weights of persons w_r equal w_{3r}.

4. Assessment of sample quality

Since stratification of the sample was based on the type of area, assessment of the sample quality involved the use of the data about the sex and age structure of the population aged 15-70. The comparative data are presented in Table 2.

Table 2

Comparison of gender and age structure of persons aged 15-70 according to the demographic statistics as of 01.01.2011 and the SEAP results (unweighted)

Age groups	Both sexes			Women			Men		
	Survey results	Demographic statistics	Bias (percentage points)	Survey results	Demographic statistics	Bias (percentage points)	Survey results	Demographic statistics	Bias (percentage points)
Population aged 15–70 including by age groups:	100.0	100.0	0.0	100.0	100.0	0.0	100.0	100.0	0.0
15-19	6.8	7.8	-1.0	6.0	7.2	-1.2	7.6	8.5	-0.9
20-24	7.5	10.4	-2.9	6.7	9.6	-2.9	8.4	11.2	-2.8
25-29	8.1	11.0	-2.9	7.6	10.2	-2.6	8.8	11.8	-3.0
30-34	8.1	9.9	-1.8	7.9	9.4	-1.5	8.3	10.5	-2.2
35-39	9.2	9.5	-0.3	9.1	9.2	-0.1	9.3	9.9	-0.6
40-44	9.4	8.9	0.5	9.4	8.7	0.7	9.5	9.1	0.4
45-49	10.5	9.7	0.8	10.2	9.8	0.4	10.8	9.6	1.2
50-54	11.5	10.2	1.3	11.4	10.5	0.9	11.4	9.8	1.6
55-59	10.6	8.9	1.7	11.5	9.5	2.0	9.7	8.1	1.6
60-64	9.9	7.4	2.5	10.5	8.3	2.2	9.2	6.5	2.7
65-70	8.4	6.3	2.1	9.7	7.6	2.1	7.0	5.0	2.0

The data in Table 2 indicate that the sample reflects adequately the gender-age structure of the population. The largest deviation is 3.0 percentage points for men aged 25-29.

5. Methodological clarifications on data quality assessment

For assessment of the quality of the indicators based on the data of a sample survey, a sample error value indicator is used. It defines the limits of confidence intervals of possible values of the indicator according to the survey results. The estimate of an index for which a sample error value is considerable as compared with the value of the estimate per se, cannot be used in the analysis of survey results.
A sampling error (*SE*) is calculated as a standard error according to the following formula:

$$SE=\sqrt{\frac{\sigma^2}{n}}, \qquad (10)$$

where σ^2 is a variance which describes the indicator values variation on sample units;
n – sample size.

Value σ^2 is calculated by the following complex design formula for a sample:

$$\sigma^2 = deff \times \frac{\sum_{i=1}^{i=1} (y_i - \bar{y})^2}{n}, \qquad (11)$$

where y_i is an attribute value for a household *i*;

$\bar{y}$ is a mean attribute value for a sampling frame;

deff is a parameter reflecting the effect of a sample design on the attribute variance value (a design effect).

To determine the variances, standard errors and design effect values for samples with a complex design special calculation methods are used.

The quality of sample surveys data is characterized by the error limit of the sample *LSE* and the coefficient of variation *CV*.

The value of an error limit of the sample LSE determines the confidence interval limits for the indicator estimate and is calculated by the following formula:

$$LSE = t \cdot SE, \qquad (12)$$

where *t* is a confidence number which determines the ratio between a limit and a standard error at a given probability (*p* is a probability that the sampling error for assessing the value shall not exceed *LSE*).

Error limit of the sample is used to construct confidence limits estimates (bounds of confidence intervals). For example,

$$\bar{y}_L = \bar{y} - LSE;\ \bar{y}_L = \bar{y} + LSE, \qquad (13)$$

where y_L is the lower confidence limit of the interval estimation of a mean value;

y_R is the upper confidence limit of the interval estimation of a mean value.

The coefficient of variation *CV* is calculated by the following formula:

$$CV = \frac{SE}{\bar{y}} x100\ (\%), \qquad (14)$$

The coefficient of variation is used as an indicator of data suitability for analysis. If *CV*≤5%, the estimate is regarded reliable; if 5%<*CV*≤10%, the estimate is suitable for a quantitative analysis but its reliability is not high enough; if 10%<*CV*≤25% , the estimate is suitable only for a qualitative analysis and should be used with discretion, if 25%<*CV*≤50%, the estimate is regarded unreliable, if *CV* is more than 50%, the information cannot be published due to its unsuitability for use.

6. Characteristics of reliability of key indicators of economic activity of the population obtained on the basis of sample surveys in 2011

Annexes 1.1-3.5 show the values of reliability characteristics (standard sample error, error limit of the sample, and coefficient of variation) for estimates of absolute and relative average annual and quarterly indicators of economic activity of the population, employment and unemployment in relation to gender, type of area, age groups, education both for Ukraine as a whole and for each region.

These data indicate that the estimates of the indicators are fairly reliable both for Ukraine as a whole and section-wise. Thus, the coefficient of variation for estimating the average annual number of economically active population for Ukraine is 0.9%, which indicates a high reliability of the estimate (Annex 1.1). The coefficient of variation for estimating of the number of economically active population has the largest value for the rural population of Ukraine - 2.6%. The quarterly data are also fairly reliable and can be used for both quantitative and qualitative analysis.

The error limit of the sample for an average number of economically active population is 397,700 people. This means that with a probability value of 0.95 the average number of economically active population ranges within 22,056,900 + / - 397,700 persons, i.e. within an interval of 21,659,200 ÷ 22,454,600.

Estimates of the indicator of the number of employed persons have equally high level of reliability (Annex 2.1), with the maximum value of the coefficient of variation not exceeding 2.9% for the rural population in Q4.

For the estimates of the number of the unemployed population in Ukraine as a whole (Annex 3.1), the maximum value of the coefficient of variation is 10.1% in rural areas in Q3. Thus, the reliability of the estimates of the unemployed population is somewhat lower than the estimates of the economically active and employed population.

On the whole, the findings are similar regarding the degree of data reliability for the estimates of relative indicators of the level of economic activity of the population (Annex 1.2), the level of employment (Annex 2.2) and unemployment (Annex 3.2). For the first two indicators the coefficients of variation are lower than those for estimating the respective absolute indicators, which is natural and accounts for the properties of the indicators.

The reliability of estimating the level of unemployment is almost the same as for estimating the number of the unemployed. This is driven by the need to account for the reliability of estimating the number of economically active population in the sections under consideration. The maximum value of the coefficient of variation here is 10.0% for rural areas in Q3. Hence, the data presented can be used for conducting a quantitative and a qualitative analysis.

The estimates of comparative indices - the level of economic activity of the population (Annex 1.3) and the level of employment (Annex 2.3) in relation to age groups have a fairly high level of reliability for Ukraine as a whole. The value of the coefficient of variation does not exceed 4.2% in all age groups for levels of economic activity and employment of the population. The largest coefficients of variation are observed in the extreme age groups - 15-19 and 60-70. With regard to unemployment (Annex 3.3), the data are less reliable: the coefficient of variation does not exceed 4% only for the 20-24 age group, whereas for all other groups it ranges from 5 to 12%, and for persons above active working age (60-64) the coefficient of variation is 48%. Given the value of the standard error of the sample, it is lower than in the other age groups. These values of variation coefficients for the said age groups are explained by the low level of unemployment.

The estimates of the level of economic activity (Annex 1.4) and the level of employment (Annex 2.4) in relation to education levels have also a fairly high level of reliability for Ukraine as a whole. As regards the unemployment rates (Annex 3.4), the coefficient of variation does not exceed 5% for the index of the levels of complete higher and comprehensive secondary education. For persons with incomplete education the coefficient of variation is 5.5%. This indicates sufficient reliability of the index and its suitability for a quantitative and a qualitative analysis. For persons with basic higher and basic comprehensive secondary education the coefficients of variation are 12.7% and 9.9% respectively. The reliability of the unemployment rate index among persons with elementary general education or persons with no education, is low, therefore the information is not released to the public (the coefficient of variation exceeded 50%).

As regards the regional level, the estimates of the indicators here are generally less reliable than for the nationwide level. This results, primarily, from a considerably smaller sample size for each particular region.

For the level of economic activity and the level of employment the estimates precision obtained for the regions is satisfactory (Annexes 1.5, 2.5). The coefficient of variation for almost all the regions does not exceed 5%.

Particular attention should be paid to the unemployment rate by regions of Ukraine (Annex 3.5). According to the findings of the reliability analysis, the variation coefficient values for seven regions do not exceed 10%, for eleven regions they range within 10% - 15%, and for nine regions they exceed 15%.

Annex 1.1

Reliability estimates of the index of economically active population aged 15-70 by gender and location in 2011

	Economically active population aged 15-70, thousands of persons	Standard error of sample (SE), thousands of persons	Error limit of sample, (LSE), thousands of persons	Coefficient of variation (CV), %
Q1				
Total	22033,1	260,2	509,9	1,2
including				
women	10625,6	134,8	264,3	1,3
men	11407,5	155,7	305,1	1,4
urban settlements	15163,6	164,7	322,8	1,1
rural areas	6869,5	202,1	396,0	2,9
Q2				
Total	22083,5	221,2	433,6	1,0
including				
women	10663,7	127,5	249,9	1,2
men	11419,8	133,2	261,0	1,2
urban settlements	15089,7	134,7	264,1	0,9
rural areas	6993,8	184,5	361,6	2,6
Q3				
Total	22313,7	213,0	417,5	1,0
including				
women	10724,4	131,1	256,9	1,2
men	11589,3	135,3	265,2	1,2
urban settlements	15168,4	123,1	241,4	0,8
Rural areas	7145,3	183,0	358,7	2,6
Q4				
Total	21797,2	289,3	567,0	1,3
including				
women	10414,4	149,0	292,0	1,4
men	11382,8	167,6	328,4	1,5
urban settlements	14918,2	199,2	390,4	1,3
rural areas	6879,0	207,2	406,1	3,0
Average per year				
Total	22056,9	202,9	397,7	0,9
including				
women	10607,0	112,9	221,2	1,1
men	11449,9	114,4	224,3	1,0
Urban settlements	15085,0	98,2	192,6	0,7
Rural areas	6971,9	182,0	356,6	2,6

Annex 1.2

Reliability estimates of the index of economic activity rate of population aged 15-70 by gender and location in 2011

	Rate of economic activity of the population aged 15-70, %	Standard error of sample (SE), %	Error limit of sample (LSE), %	Coefficient of variation (CV), %
Q1				
Total	64,2	0,4	0,7	0,6
including				
women	58,6	0,4	0,9	0,8
men	70,4	0,5	0,9	0,7
urban settlements	63,0	0,3	0,6	0,5
rural areas	67,1	1,0	1,9	1,4
Q2				
Total	64,3	0,5	0,9	0,7
including				
women	58,9	0,6	1,2	1,0
men	70,5	0,5	1,1	0,8
urban settlements	62,7	0,4	0,8	0,6
rural areas	68,3	1,3	2,5	1,9
Q3				
Total	65,0	0,4	0,8	0,6
including				
women	59,2	0,6	1,1	1,0
men	71,5	0,4	0,8	0,6
urban settlements	63,0	0,3	0,6	0,5
Rural areas	69,8	1,2	2,3	1,7
Q4				
Total	63,5	0,3	0,7	0,5
including				
women	57,5	0,5	0,9	0,8
men	70,2	0,4	0,8	0,6
urban settlements	61,9	0,2	0,5	0,4
rural areas	67,2	1,0	2,0	1,5
Average per year				
Total	64,3	0,3	0,6	0,5
including				
women	58,5	0,4	0,9	0,7
men	70,7	0,3	0,6	0,5
Urban settlements	62,6	0,2	0,3	0,3
Rural areas	68,1	1,0	2,0	1,5

Annex 1.3

Reliability estimates of the index of economic activity rate of population aged 15-70 by age groups in 2011

	Rate of economic activity of the population aged 15-70, %	Standard error of sample (SE), %	Error limit of sample (LSE), %	Coefficient of variation (CV), %
Total population aged 15-70	64,3	0,3	0,6	0,5
including				
15-19	14,7	0,5	1,0	3,6
20-24	62,0	0,9	1,8	1,5
25-29	79,7	0,5	0,9	0,6
30-34	83,0	0,5	1,0	0,6
35-39	85,4	0,4	0,9	0,5
40-44	86,2	0,4	0,7	0,4
45-49	84,4	0,5	0,9	0,5
50-54	77,6	0,6	1,1	0,7
55-59	51,3	0,7	1,4	1,3
60-64	28,1	0,7	1,3	2,4
65-70	19,5	0,8	1,6	4,2
Persons of working age	72,7	0,3	0,6	0,4
Persons above working age	27,9	0,7	1,3	2,4

Annex 1.4

Reliability estimates of the index of economic activity rate of population aged 15-70 by level of education in 2011

	Rate of economic activity of the population aged 15-70, %	Standard error of sample (SE), %	Error limit of sample (LSE), %	Coefficient of variation (CV), %
Total population aged 15-70	64,3	0,3	0,6	0,5
Including persons with education:				
Complete higher education	78,7	0,3	0,6	0,4
Basic higher education	51,4	1,4	2,8	2,8
Incomplete higher education	71,1	0,4	0,8	0,6
Complete comprehensive secondary education	62,6	0,4	0,8	0,7
Basic comprehensive secondary education	38,2	1,1	2,1	2,8
Elementary general education or no education at all	13,5	1,1	2,1	7,8

Annex 1.5

Reliability estimates of the index of economic activity rate of population aged 15-70 by regions in 2011

	Rate of economic activity of the population aged 15-70, %	Standard error of sample (SE), %	Error limit of sample (LSE), %	Coefficient of variation (CV), %
Ukraine	**64,3**	**0,3**	**0,6**	**0,5**
Autonomous Republic of Crimea	65,6	1,1	2,1	1,6
regions				
Vinnytsia	64,7	2,4	4,8	3,7
Volyn	64,3	2,0	3,8	3,0
Dnipropetrovsk	65,2	1,0	2,0	1,6
Donetsk	64,8	0,7	1,3	1,0
Zhytomyr	65,8	2,2	4,2	3,3
Transcarpathean	62,6	3,2	6,2	5,0
Zaporizhia	65,1	1,1	2,1	1,7
Ivano-Frankivsk	57,2	1,7	3,3	3,0
Kyiv	63,1	1,2	2,4	1,9
Kirovohrad	63,3	1,8	3,6	2,9
Luhansk	61,3	1,0	2,0	1,7
Lviv	63,1	1,3	2,6	2,1
Mykolaiv	65,1	2,1	4,1	3,2
Odesa	61,8	1,4	2,7	2,3
Poltava	64,7	1,6	3,2	2,5
Rivne	65,6	1,5	2,9	2,3
Sumy	65,1	1,7	3,4	2,6
Ternopil	61,0	3,1	6,0	5,0
Kharkiv	65,1	1,2	2,4	1,9
Kherson	64,3	1,6	3,2	2,5
Khmelnytsk	64,4	3,2	6,2	4,9
Cherkasy	65,3	2,1	4,1	3,2
Chernivtsi	62,9	1,9	3,7	3,0
Chernihiv	66,4	1,5	2,9	2,3
cities				
Kyiv	68,2	0,9	1,8	1,3
Sevastopol	66,0	1,3	2,6	2,0

Annex 2.1

Reliability estimates of the index of number of occupied population aged 15-70 by gender and location in 2011

	Number of occupied population aged 15-70, thousands of persons	Standard error of sample (SE), thousands of persons	Error limit of sample (LSE), thousands of persons	Coefficient of variation (CV), %
Q1				
Total	20108,2	232,0	454,8	1,2
including				
women	9884,5	127,0	249,0	1,3
men	10223,7	135,5	265,5	1,3
urban settlements	13859,3	152,2	298,3	1,1
rural areas	6248,9	176,2	345,3	2,8
Q2				
Total	20387,3	240,8	471,9	1,2
including				
women	9942,2	128,0	250,8	1,3
men	10445,1	138,2	270,8	1,3
urban settlements	13860,4	166,7	326,7	1,2
rural areas	6526,9	184,3	361,3	2,8
Q3				
Total	20783,0	216,0	423,3	1,0
including				
women	10076,1	131,5	257,8	1,3
men	10706,9	125,0	244,9	1,2
urban settlements	14111,7	131,4	257,6	0,9
rural areas	6671,3	182,4	357,5	2,7
Q4				
Total	20018,4	275,7	540,3	1,4
including				
women	9622,5	133,0	260,7	1,4
men	10395,9	163,8	321,0	1,6
urban settlements	13664,3	204,0	399,7	1,5
rural areas	6354,1	184,4	361,5	2,9
Average per year				
Total	20324,2	198,7	389,4	1,0
including				
women	9881,3	111,5	218,6	1,1
men	10442,9	107,4	210,5	1,0
urban settlements	13873,9	109,4	214,3	0,8
rural areas	6450,3	171,6	336,3	2,7

Annex 2.2

Reliability estimates of the index of employment rate of population aged 15-70 by gender and location in 2011

	Employment rate of population aged 15-70, %	Standard error of sample (SE), %	Error limit of sample (LSE), %	Coefficient of variation (CV), %
Q1				
Total	58,6	0,4	0,7	0,6
including				
women	54,6	0,4	0,8	0,8
men	63,1	0,5	0,9	0,8
urban settlements	57,5	0,3	0,6	0,5
rural areas	61,0	1,0	1,9	1,6
Q2				
Total	59,4	0,5	1,0	0,9
including				
women	54,9	0,6	1,2	1,1
men	64,5	0,6	1,2	1,0
urban settlements	57,6	0,5	1,0	0,9
rural areas	63,7	1,4	2,7	2,2
Q3				
Total	60,6	0,4	0,9	0,7
including				
women	55,6	0,6	1,1	1,0
men	66,1	0,5	1,0	0,7
urban settlements	58,6	0,4	0,7	0,6
rural areas	65,2	1,2	2,4	1,9
Q4				
Total	58,3	0,3	0,7	0,6
including				
women	53,1	0,4	0,8	0,8
men	64,2	0,5	0,9	0,7
urban settlements	56,7	0,3	0,6	0,5
rural areas	62,1	0,9	1,8	1,5
Average per year				
Total	59,2	0,3	0,7	0,6
including				
women	54,5	0,4	0,9	0,8
men	64,4	0,4	0,7	0,6
urban settlements	57,6	0,2	0,4	0,4
rural areas	63,0	1,1	2,1	1,7

Annex 2.3

Reliability estimates of the index of employment rate of population aged 15-70 by age groups in 2011

	Employment rate of population aged 15-70, %	Standard error of sample (SE), %	Error limit of sample (LSE), %	Coefficient of variation (CV), %
All population aged 15-70	59,2	0,3	0,7	0,6
including				
15-19	11,2	0,4	0,9	4,0
20-24	51,0	0,9	1,7	1,7
25-29	72,4	0,6	1,2	0,8
30-34	76,9	0,6	1,3	0,8
35-39	79,3	0,5	1,0	0,6
40-44	80,4	0,6	1,2	0,8
45-49	79,3	0,6	1,2	0,7
50-54	72,9	0,5	1,0	0,7
55-59	49,5	0,7	1,4	1,4
60-64	28,1	0,7	1,3	2,4
65-70	19,5	0,8	1,6	4,2
Persons of working age	66,5	0,3	0,6	0,5
Persons above working age	27,9	0,7	1,3	2,4

Annex 2.4

Reliability estimates of the index of employment rate of population aged 15-70 by level of education in 2011

	Employment rate of population aged 15-70, %	Standard error of sample (SE), %	Error limit of sample (LSE), %	Coefficient of variation (CV), %
All population aged 15-70	59,2	0,3	0,7	0,6
Including persons with education				
complete higher	73,4	0,4	0,8	0,5
basic higher	41,1	1,6	3,0	3,8
incomplete higher	66,1	0,5	0,9	0,7
complete comprehensive secondary	57,1	0,5	1,0	0,9
basic comprehensive secondary	35,6	1,0	2,1	2,9
elementary general or no education at all	13,5	1,0	2,0	7,7

Annex 2.5

Reliability estimates of the index of employment rate of population aged 15-70 by regions in 2011

	Employment rate of population aged 15-70, %	Standard error of sample (SE), %	Error limit of sample (LSE), %	Coefficient of variation (CV), %
Ukraine	**59,2**	**0,3**	**0,7**	**0,6**
Autonomous Republic of Crimea	61,7	1,3	2,6	2,2
regions				
Vinnytsia	58,4	1,7	3,4	2,9
Volyn	59,0	1,5	2,9	2,5
Dnipropetrovsk	60,7	1,0	2,0	1,7
Donetsk	59,5	0,7	1,3	1,1
Zhytomyr	59,2	1,8	3,5	3,0
Transcarpathean	56,6	3,5	6,9	6,2
Zaporizhia	60,4	1,2	2,4	2,0
Ivano-Frankivsk	52,2	1,8	3,5	3,4
Kyiv	58,8	1,5	2,9	2,5
Kirovohrad	57,8	1,9	3,8	3,3
Luhansk	57,3	1,0	2,0	1,8
Lviv	58,3	1,4	2,7	2,4
Mykolaiv	59,8	1,9	3,7	3,1
Odesa	58,1	1,4	2,7	2,3
Poltava	58,8	1,7	3,3	2,9
Rivne	58,8	1,5	2,9	2,5
Sumy	59,2	1,6	3,1	2,7
Ternopil	54,6	3,5	6,9	6,4
Kharkiv	60,6	1,3	2,6	2,2
Kherson	58,5	1,6	3,2	2,8
Khmelnytsk	58,7	2,9	5,7	4,9
Cherkasy	59,3	2,3	4,5	3,9
Chernivtsi	57,7	2,2	4,2	3,7
Chernihiv	59,5	1,2	2,3	2,0
cities				
Kyiv	64,4	0,9	1,8	1,5
Sevastopol	61,9	1,2	2,3	1,9

Annex 3.1

Reliability estimates of the index of unemployed population aged 15-70 by gender and location in 2011

	Unemployed population aged 15-70, thousands of persons	Standard error of sample (SE), %	Error limit of sample (LSE), %	Coefficient of variation (CV), %
Q1				
Total	1924,9	78,7	154,3	4,1
including				
women	741,1	45,7	89,5	6,2
men	1183,8	54,3	106,4	4,6
urban settlements	1304,3	54,9	107,7	4,2
rural areas	620,6	56,5	110,7	9,1
Q2				
Total	1696,2	71,5	140,2	4,2
including				
women	721,5	42,1	82,4	5,8
men	974,7	56,2	110,1	5,8
urban settlements	1229,3	58,4	114,4	4,7
rural areas	466,9	38,2	74,9	8,2
Q3				
Total	1530,7	65,5	128,3	4,3
including				
women	648,3	28,9	56,6	4,5
men	882,4	59,0	115,7	6,7
urban settlements	1056,7	45,7	89,7	4,3
rural areas	474,0	47,7	93,4	10,1
Q4				
Total	1778,8	82,6	161,9	4,6
including				
women	791,9	42,1	82,6	5,3
men	986,9	65,9	129,2	6,7
urban settlements	1253,9	66,4	130,1	5,3
rural areas	524,9	49,8	97,7	9,5
Average per year				
Total	1732,7	51,0	100,0	2,9
including				
women	725,7	23,8	46,6	3,3
men	1007,0	39,8	78,0	4,0
urban settlements	1211,1	33,0	64,7	2,7
rural areas	521,6	38,7	75,9	7,4

Annex 3.2

Reliability estimates of the index of unemployment rate of population aged 15-70 by gender and location in 2011

	Unemployment rate of population aged 15-70, %	Standard error of sample (SE), %	Error limit of sample (LSE), %	Coefficient of variation (CV), %
Q1				
Total	8,7	0,3	0,6	3,7
including				
women	7,0	0,4	0,8	5,9
men	10,4	0,4	0,8	4,0
urban settlements	8,6	0,3	0,7	3,9
rural areas	9,0	0,7	1,4	7,8
Q2				
Total	7,7	0,3	0,7	4,4
including				
women	6,8	0,4	0,8	5,7
men	8,5	0,5	1,0	5,8
urban settlements	8,1	0,4	0,8	5,2
rural areas	6,7	0,6	1,1	8,3
Q3				
Total	6,9	0,3	0,6	4,2
including				
women	6,0	0,3	0,5	4,5
men	7,6	0,5	0,9	6,3
urban settlements	7,0	0,3	0,6	4,4
rural areas	6,6	0,7	1,3	10,0
Q4				
Total	8,2	0,4	0,7	4,4
including				
women	7,6	0,4	0,7	4,8
men	8,7	0,6	1,1	6,4
urban settlements	8,4	0,4	0,9	5,3
rural areas	7,6	0,6	1,2	8,2
Average per year				
Total	7,9	0,2	0,5	3,1
including				
women	6,8	0,2	0,4	3,3
men	8,8	0,3	0,6	3,7
urban settlements	8,0	0,2	0,5	2,9
rural areas	7,5	0,5	1,0	6,9

Annex 3.3

Reliability estimates of the index of unemployment rate of population aged 15-70 by age groups in 2011

	Unemployment rate of population aged 15-70, %	Standard error of sample (SE), %	Error limit of sample (LSE), %	Coefficient of variation (CV), %
All population aged 15-70	7,9	0,2	0,4	2,9
including				
15-19	23,7	2,2	4,2	9,1
20-24	17,7	0,7	1,4	4,0
25-29	9,2	0,6	1,1	6,2
30-34	7,4	0,5	1,0	6,7
35-39	7,2	0,4	0,9	6,2
40-44	6,7	0,6	1,2	9,3
45-49	6,1	0,5	0,9	7,6
50-54	6,0	0,3	0,6	5,1
55-59	3,5	0,4	0,8	12,1
60-64	0,1	0,0	0,1	48,0
65-70	-	-	-	-
Persons of working age	8,6	0,2	0,5	2,8
Persons above working age	0,1	0,0	0,1	37,0

Annex 3.4

Reliability estimates of the index of unemployment rate of population aged 15-70 By level of education in 2011

	Unemployment rate of population aged 15-70, %	Standard error of sample (SE), %	Error limit of sample (LSE), %	Coefficient of variation (CV), %
All population aged 15-70	7,9	0,2	0,4	2,9
Including persons with education				
complete higher	6,7	0,3	0,5	4,0
basic higher	20,1	2,6	5,0	12,7
incomplete higher	7,1	0,4	0,8	5,5
complete comprehensive secondary	8,8	0,4	0,7	4,3
basic comprehensive secondary	6,9	0,7	1,3	9,9
elementary general or no education at all	:	:	:	78,8

Annex 3.5

Reliability estimates of the index of unemployment rate of population aged 15-70 by regions in 2011

	Unemployment rate of population aged 15-70, %	Standard error of sample (SE), %	Error limit of sample (LSE), %	Coefficient of variation (CV), %
Ukraine	**7,9**	**0,2**	**0,4**	**2,9**
Autonomous Republic of Crimea	6,1	1,3	2,6	21,6
regions				
Vinnytsia	9,7	1,3	2,5	13,1
Volyn	8,3	1,2	2,4	14,7
Dnipropetrovsk	6,8	0,9	1,7	13,0
Donetsk	8,2	0,5	1,0	6,3
Zhytomyr	10,0	1,5	2,9	14,6
Transcarpathean	9,6	1,8	3,5	18,9
Zaporizhia	7,2	0,8	1,5	10,8
Ivano-Frankivsk	8,7	0,9	1,7	10,1
Kyiv	6,7	1,2	2,3	17,2
Kirovohrad	8,6	0,8	1,5	8,9
Luhansk	6,6	0,7	1,3	10,3
Lviv	7,7	0,5	1,0	6,5
Mykolaiv	8,1	1,3	2,5	15,9
Odesa	6,0	0,8	1,6	13,4
Poltava	9,2	0,7	1,4	8,1
Rivne	10,4	1,5	2,9	14,4
Sumy	9,1	1,5	3,0	16,8
Ternopil	10,4	1,8	3,5	17,3
Kharkiv	7,0	0,9	1,8	13,0
Kherson	9,0	0,8	1,6	9,1
Khmelnytsk	8,8	1,6	3,2	18,4
Cherkasy	9,2	1,6	3,1	17,4
Chernivtsi	8,2	2,0	3,9	24,5
Chernihiv	10,4	1,3	2,6	12,8
cities				
Kyiv	5,6	0,5	1,0	9,5
Sevastopol	6,2	0,6	1,1	9,0