



EUROPEAN COMMISSION
EUROSTAT

Directorate F: Social Statistics
Unit F-4: Income and living conditions; Quality of life

HOUSEHOLD BUDGET SURVEY 2020

ANONYMISATION CRITERIA

GENERAL RULES:

- Due to the presence of two different (non-nested) classifications related to the geographical detail of the household (the classical NUTS classification and the degree of urbanisation), for many MSs there were several unique cases in the size of municipalities at NUTS2 level. For this reason the variable HA08 (Region of residence) is released at NUTS1 level.
- Variables HA06 (Stratum) and HA07 (Primary sampling unit) are dropped.
- The variables MB01A (Country of birth), MB011A (Country of main citizenship) and MB012A (Country of residence) are released according to the following classification: National, Non-National but EU-National, Non-National and non-EU-National.
- The “sin-consumption” variables are dropped from the SUF. The “sin-consumption” variables are the following: HE023 (Narcotics), HE0230; HE122 (Prostitution), HE1220, HE12200; HE0943 (Games of chance), HE09430.
- MB03 (Age (in completed years) of household member) is recoded to 5 classes.
- MB04 (Marital status of household member): the categories of widowed and divorced, are aggregated.
- ME0908 (Occupation in Main Job) – codes 'Z1','Z2','Z3' are aggregated to "AF" i.e. “Armed forces” and codes '11','12','13' , '14' are aggregated to "LM" i.e. “Legislators & managers”.

COUNTRY SPECIFIC RULES:

In the case of **Slovenia** additional anonymisation rules had to be applied:

- (1) HH012 (Income in kind from employment
HH095 (Monetary net income (total monetary income from all sources minus income taxes))):

MF099 (Total Income from all sources (net amount) corresponding to each single member of the family):

- Top coding 20 (replacement of the highest 20 unweighted values with their weighted average).
- Rounding to the nearest 10 EUR.

(2) HH023 (Income in kind from non-salaried activities):

HH032 (Imputed rent):

- Rounding to the nearest 10 EUR.
- HH099 (Net income (total income from all sources including non-monetary components minus income taxes)):
- Recomputation of variables: $HH099 = HH012 + HH023 + HH032 + HH095$.

(3) MB01A (Country of birth)

MB011A (Country of main citizenship)

MB012A (Country of residence).

- Recoding into two categories (national and other):
 - National: 10 → 10
 - Other: 21, 22 → 20
 - Not Specified: 99 → 99

(4) MC01A (Highest educational attainment level):

MC02B (Level of the current/most recent formal education or training activity)

- Recoding into three categories (primary, secondary and tertiary):
 - Primary: 10, Z1, Z2 → 2
 - Secondary: Z3, Z4 → 3
 - Tertiary: Z5, Z6, Z7, Z8 → 5
 - Not Applicable: 88 → 88
 - Not Stated: 99 → 99

(5) ME04 (Economic activity of the local unit (main job) (NACE Rev. 2)):

- - Recoding into the following categories:
 - A → 'A'
 - B, C, D, E → 'B-E'
 - F → 'F'
 - G → 'G'
 - H → 'H'
 - I → 'I'
 - J → 'J'
 - K → 'K'
 - L, M, N → 'L-N'

- O → 'O'
- P → 'P'
- Q → 'Q'
- R, S, T, U → 'R-U'
- Not Specified: Z → 'Z'

(6) ME0908 (Occupation in main job (ISCO-08)):

- Grouping according to the first 'digit':
 - Z1, Z2, Z3 → 00
 - 11, 12, 13, 14 → 10
 - 21, 22, 23, 24, 25, 26 → 20
 - 31, 32, 33, 34, 35 → 30
 - 41, 42, 43, 44 → 40
 - 51, 52, 53, 54 → 50
 - 61, 62, 63 → 60
 - 71, 72, 73, 74, 75 → 70
 - 81, 82, 83 → 80
 - 91, 92, 93, 94, 95, 96 → 90
 - Not Applicable: 98 → 98
 - Not Specified: 99 → 99

In the case of **Malta** additional anonymisation rules had to be applied:

Categorical variables:

- (1) HB05: Top-coded at 6+
- (2) HB051: Top-coded at 2+
- (3) HB052: Top-coded at 2+
- (4) HB053: Top-coded at 2+
- (5) HB054: Top-coded at 2+
- (6) HB055: Top-coded at 2+
- (7) HB056: Top-coded at 3+
- (8) HB057: Top-coded at 2+
- (9) HB0761: Top-coded at 3+
- (10) HB0762: Top-coded at 3+
- (11) HC23: Merged as follows:
 - Z1, Z2, Z3, Z4, Z5 and Z6 → 1_6
 - Z8, Z9, 10 and 11 → 8_11
 - Z7 → 7
- (12) HD20: Top-coded at 3+

- (13) MB01: Aggregated into:
 - 10 - National
 - 21_22 - Non-national.
- (14) MB01.1: Aggregated into:
 - 10 - National
 - 21_22 - Non-national.
- (15) ME01: Merged as follows:
 - 3, 4, 5, 6 and 7 → 3_7
- (16) ME0908 Merged as follows:
 - 92 and 95 → 92_95
- (17) ME12: Merged as follows:
 - 1 and 2 → 1_2
 - 3, 4 and 5 → “3_5”.

Income and Expenditure variables:

- (18) MF099:
 - Suppressed negative values (replaced with ‘.’).
- (19) Values for income and expenditure variables at household level are suppressed if both of the following two rules are satisfied:
 - a) For the household reference person only, there is a record with a unique combination of the following demographic characteristics: HA09 (Degree of urbanisation), MB02 (Sex) and MB03_10years (Age re-grouped in ten-year classes).
 - b) For the household reference person, the ‘pro-capite’ value (i.e. value per number of person in the household) of these variables is greater than the upper risk threshold and smaller than the lower risk threshold, as represented hereunder.
 - Lower Risk Threshold < ‘pro capite’ value < Upper Risk Threshold
 - Lower Risk Threshold = $Q1 - 3*IQR$
 - Upper Risk Threshold = $Q3 + 3*IQR$
 - Where:
 - Q1 = Quartile 1 (i.e. the 25th percentile) calculated from the ‘pro-capite’ value
 - Q3 = Quartile 3 (i.e. the 75th percentile) calculated from the ‘pro-capite’ value
 - IQR = $Q3 - Q1$ (i.e. the inter-quartile range).