ILOSTAT MICRODATA PROCESSING QUICK GUIDE

Principles and methods underlying the ILO's processing of anonymized household survey microdata

International Labour Organization

Department of Statistics

ILOSTAT Microdata Processing Quick Guide

Principles and methods underlying the ILO's processing of anonymized household survey microdata

Information on ILO publications and digital products can be found at: www.ilo.org/publns.

Printed in Switzerland

Contents

Acknowledgements

1. Introduction

The ILO Department of Statistics' mandate is to provide users with relevant, timely, and reliable labour statistics at national, regional and international levels; to develop international standards and guidelines for measuring all crucial aspects of labour markets while ensuring comparability across countries; and to offer technical assistance, cooperation and training to member countries. The department is the focal point for labour statistics within the United Nations System, as well as the focal point for all statistics-related issues within the ILO.

The Data Production and Analysis Unit of the Department of Statistics is in charge of the compilation and dissemination of labour statistics and related information produced by national statistical systems. The unit collects short-term and annual indicators on a wide variety of labour-related topics, and publishes derived indicators on ILOSTAT, the ILO's central statistical repository.[1]

This Quick Guide presents the anonymized microdata processing undertaken by DPAU. It describes why and how the unit carries out this activity, as well as the potential expansion of this work. It also mentions considerations and limitations to take into account by data users.

2. Why does the ILO process microdata?

In 2016, the ILO began systematically processing microdata from household surveys (HHS), mainly labour force surveys (LFS).[2] The overarching aim of this new activity was to increase the coverage and improve the quality and international comparability of data published on ILOSTAT. This exercise was initially conceived to fill data gaps by using estimates derived from anonymized and openly shared microdata sets. The significant improvements achieved in ILOSTAT in terms of data quality (consistency and comparability) and quantity (increased country coverage, and more indicators and disaggregations available) led to further investments in microdata collection and processing. Microdata processing has now become a major activity of DPAU and a major source of labour market statistics published by the ILO.

The ILO's microdata processing is expected to bring further benefits in the short and long run, not only to the ILO Department of Statistics, but also to national statistical systems, researchers, policy-makers and more generally to all data users (internal and external) who are searching for internationally-comparable and up-to-date labour market information.

The use of anonymized datasets for the production of labour market indicators represents a new type of collaboration between national statistical systems and the ILO. This builds on the traditional means of data compilation (detailed labour statistics reported by countries to the ILO through an annual questionnaire of labour statistics[3] or fetched directly by DPAU from national reporting platforms and other official statistical sources).

---

[1] Available at www.ilo.org/ilostat

[2] In particular, Labour Force Surveys (LFS) are used to analyse national labour markets. For further information on the main sources of statistical information please refer to the "Quick guide on sources and uses of Labour Statistics" http://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/publication/wcms_590092.pdf).

[3] Each year, the ILO Department of Statistics sends out an extensive questionnaire on labour statistics to the relevant agencies of each member State, including the national statistical office and the labour ministry. This is an electronic instrument that collects labour market indicators and the related metadata, following international standards and recommendations to the extent possible (so as to ensure cross-country comparability). This questionnaire is a crucial tool allowing us to (1) reach countries that do not openly disseminate data online or for which data is not widely available, (2) obtain more detailed and disaggregated information than what is widely available, and (3) consolidate in one database all the labour market statistics produced by different agencies participating in the national statistical system. The questionnaire covers 17 labour-related topics, and includes around 100 indicators.

More specifically, the LFS microdata processing system that DPAU has put in place is intended to benefit national statistical systems by:

- Reducing the burden of responding to the extensive annual ILOSTAT questionnaire. Having access to microdata files, the ILO is able to derive the necessary labour market indicators directly in compliance with international statistical standards and definitions as well as with national and international rules and guidelines on microdata handling, distribution and confidentiality. Countries sharing microdata with the ILO receive a reduced version of the ILOSTAT questionnaire, focussed on collecting indicators most typically derived from sources other than HHS/LFS.

- Providing the coding used to process the relevant microdata files and obtain the necessary labour market indicators according to standard concepts and definitions, so they can use these coding routines and build on this work.

- Ensuring a higher level of compliance with internationally-agreed concepts, standards, definitions and classifications, favouring the harmonization and comparability of data across countries and over time.

- Bringing to light areas where technical assistance may be necessary, particularly allowing the ILO to provide valuable feedback on:
  - how the relevant survey questionnaire could be improved to better meet international statistical standards; and
  - what computations and calculations done at the national level should be modified to better meet international statistical standards.

- Making accurate, reliable and comparable statistics available in a timely manner, which facilitates monitoring and reporting under internationally-agreed goals such as the Sustainable Development Goals (SDGs).

Moreover, the microdata processing initiative has led to a number of data improvements and new ILO data offerings, including:

- Increased data and metadata coverage and improved data and metadata quality on ILOSTAT. Data entry errors have been reduced and many data gaps have been filled without demanding extra resources from national data providers.

- Improvements in ILO's international data reporting (e.g. reporting of data and metadata on SDG indicators under ILO custodianship) by enabling the production of more comparable labour statistics in a timely, transparent and efficient manner.

- The ability to respond to requests from other specialized agencies, research institutions and ILO departments and field offices for specific tabulations and/or cross-tabulations (not collected via the annual ILOSTAT questionnaire).

3. How do we do it?

The microdata processing begins with a research phase to identify the primary source of labour statistics at the national level. Commonly, this is the labour force survey (or a similar type of

household survey)[4], designed to measure employment and unemployment (and other labour-related items) following statistically robust and reliable methods.

After identifying the relevant survey, we proceed to retrieve or request the corresponding datasets and supporting documentation (guidelines to collect, process and disseminate information, questionnaires, data dictionaries, methodological documents and reports, etc.). Depending on the national rules for data handling and dissemination, this step can be accomplished by:

- Downloading all files needed for the analysis directly from the national microdata catalogue (usually accessible from the national statistical office's website); or
- Requesting the anonymized data and related metadata files from the national statistical office or the national focal point.

The actual processing begins once all the necessary files are available. The microdata processing workflow comprises several steps, which require the use of statistical software (namely Stata and R), as detailed below. Figure 1 outlines the steps in the microdata processing workflow to derive the desired labour market indicators.

- The first step in the actual processing is the microdata preparation. Given that Stata is the statistical software used for the definition of variables, if needed, datasets are converted to Stata format (*.dta*). It is also possible that data from one survey is divided into different files (for instance, information may be split by questionnaire module or by geographical coverage); if this is the case, the corresponding files are merged (or appended).

- Once a microdata file is ready to be used, it is read and handled in Stata. Using this software and following international concepts and definitions to the greatest extent possible, a set of variables is derived using the relevant survey questions and/or response categories. Box 1 provides an example of how variables are defined, by showing how the labour force status variable is determined. The classifications used in the microdata processing for the various breakdowns available are the relevant international standard classifications (where these exist), such as the ISCED[5] for the level of education, the ICSE[6] for status in employment, the ISIC[7] for economic activity and the ISCO[8] for occupation. This requires mapping the national classification items to the international standard classification items, which we do based on the specifications stated in the survey technical documentation.[9]

- The newly created variables (named with the prefix *ilo*) are saved separately in a new *.dta* file with no traceability with respect to the original dataset. All the derived indicators are carefully compared to those reported by the national statistical office or the national focal point. Wherever differences exist, these are tracked and documented. Two additional files are created during this step: (1) a read-me file with information on the

---

[4] Some countries draw their official labour market statistics from other types of household surveys than labour force surveys, such as income and expenditure surveys and living conditions surveys. Following national specifications, the available databases and related materials are also downloaded and processed to produce labour-related indicators. In other cases, usually where there is no regular household survey in place, establishment surveys and/or administrative records are the main source of labour market information.

[5] International Standard Classification of Education

[6] International Classification of Status in Employment

[7] International Standard Industrial Classification of All Economic Activities

[8] International Standard Classification of Occupations

[9] This step is undertaken in order to be in line with internationally agreed concepts. When information is available, the level defined for ISIC and ISCO variables' is the 2-digit level. In the absence of national or regional correspondence tables or references, only the more aggregated version of the classifications is defined (if possible).

definitions and concepts used, and (2) a framework file containing all the necessary codes to include notes to indicators, classifications and values to tabulations or cross-tabulations (for instance, the number of criteria used to define unemployment or possible deviations from standard classifications such as ISIC or ISCO).[10]

- After generating the standardized dataset and documents described above, these are used as input files in R for the production of pre-coded indicators (based on the Horvitz-Thompson estimator of totals and proportions[11]) and their related quality measures. This process also checks for data consistency and completeness across indicators.[12] If at least one of the indicators under evaluation does not pass the verification tests, the process is re-started from the pre-processing stages, that is, we go back to the definition of standard variables trying to identify issues and correct them.

- Finally, the decision of whether the derived indicators are disseminated on ILOSTAT is based on a reliability compliance test, which verifies that the sample size used for deriving estimates of totals and proportions in each tabulation and cross-tabulation meets our requirements. The rules for dissemination are set as follows: a value is (1) not published if the sample size used for producing it is less than 5, (2) published with a warning if the sample size lies between 5 and 14 (inclusive), and (3) published with no remark associated with the value if the sample size is greater or equal to 15.[13]

---

[10] For a complete list of the created variables, please refer to the Annex.
[11] Särndal, C. E., Swensson, B., & Wretman, J. (2003). *Model assisted survey sampling*. Springer Science & Business Media.
[12] For instance, that employment, unemployment and population outside the labour force of working-age sum up to the working-age population.
[13] These rules apply equally for monthly, quarterly and annual indicators.

Figure 1. Flowchart of actions taken in microdata processing to derive labour market indicators
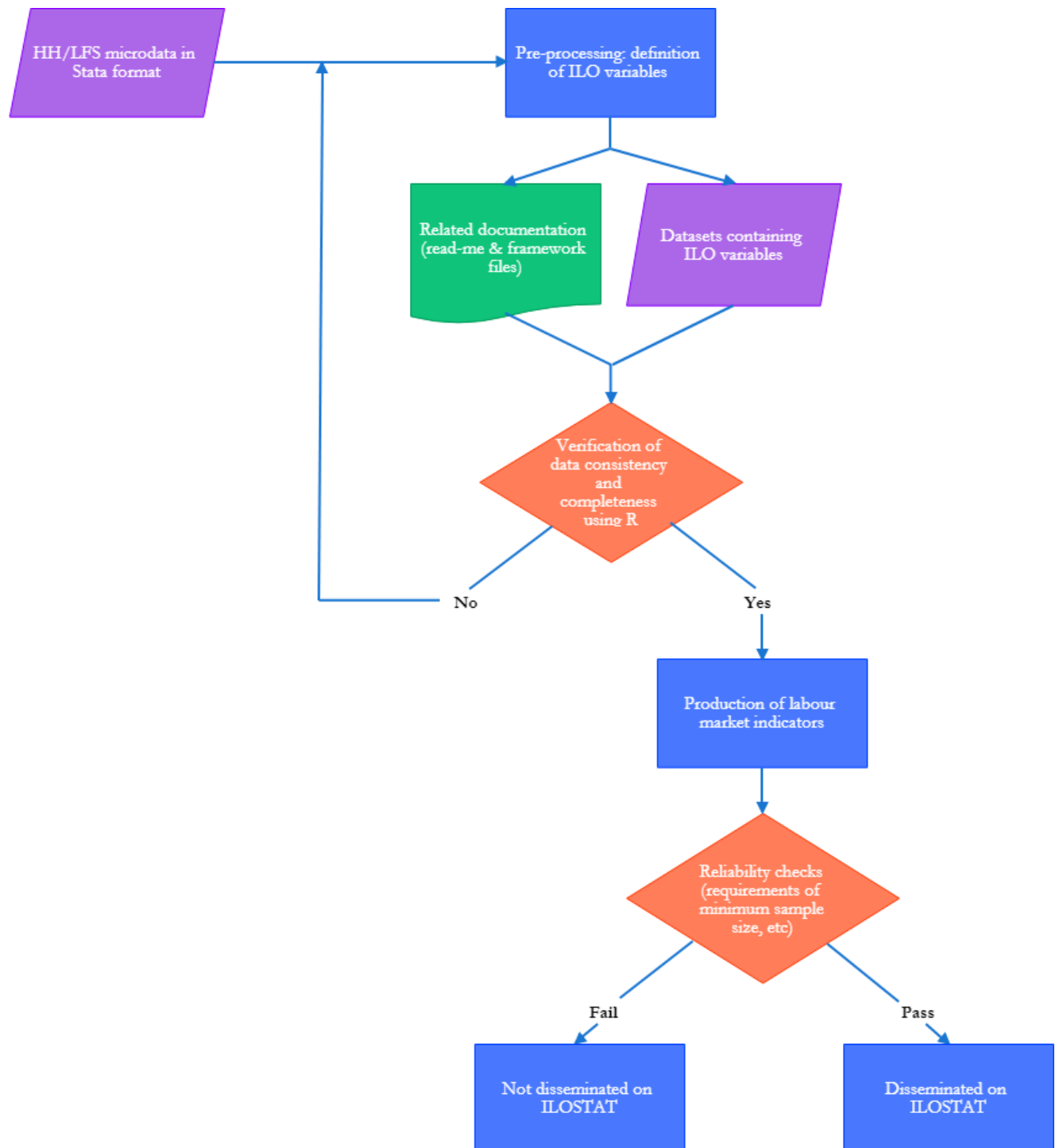
Box 1. Concepts and definitions (example)

Following the **_Resolution concerning statistics of work, employment and labour underutilization_** adopted by the 19th International Conference of Labour Statisticians (para 11), "[…] the concept **_labour force_** refers to the current supply of labour for the production of goods and services in exchange for pay or profit […]". Thus, (para 15) "[…] people may be classified in a short reference period […], according to their **_labour force status_** as being: in **employment**, in **unemployment** or **outside the labour force** […]".
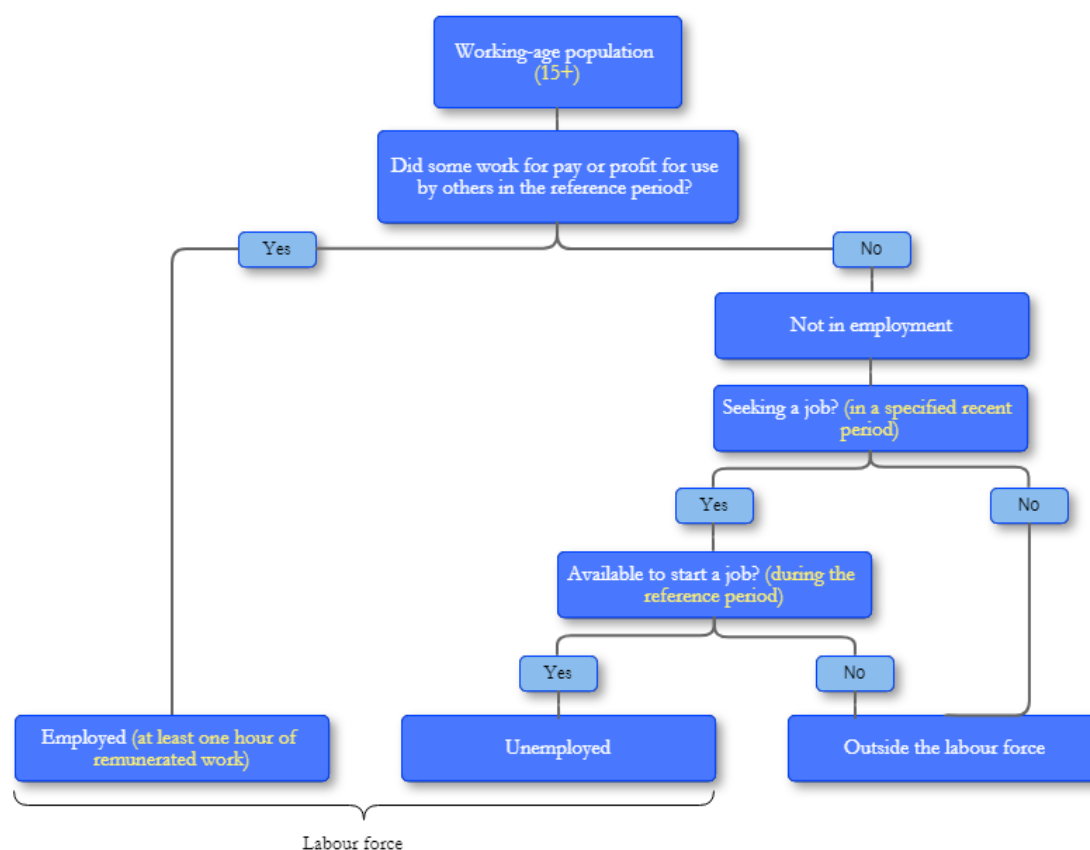
Where:

- "**_Employment_** comprises all persons of working age who during a specified brief period, such as one week or one day, were in the following categories (a) paid employment (whether at work or with a job but not at work); or b) self-employment (whether at work or with an enterprise but not at work)" (para 27), and

- "the **_Unemployed_** comprise all persons of working age who were: (a) without work during the reference period, i.e. were not in paid employment or self-employment; (b) currently available for work, i.e. were available for paid employment or self-employment during the reference period; and (c) seeking work, i.e. had taken specific steps in a specified recent period to seek paid employment. Future starters,

that is, persons who did not look for work but have a future labour market stake (made arrangements for a future job start) are also counted as unemployed, as well as participants in skills training or retraining schemes within employment promotion programmes, who on that basis, were "not in employment", not "currently available" and did not "seek employment" because they had a job offer to start within a short subsequent period generally not greater than three months and persons "not in employment" who carried out activities to migrate abroad in order to work for pay or profit but who were still waiting for the opportunity to leave" (para. 47)

The following piece of code shows a generic example to produce the labour force status variable:

```
* --------------------------------------------------------------------
* --------------------------------------------------------------------
*                    Labour Force Status ('ilo_lfs')                 *
* --------------------------------------------------------------------
* --------------------------------------------------------------------

gen ilo_lfs=.
    replace ilo_lfs=1 if (condition_1==1 | ...  | condition_n==1) & ilo_wap==1
    replace ilo_lfs=2 if ilo_lfs!=1 & (seeking==1 & available==1) & ilo_wap==1
    replace ilo_lfs=3 if !inlist(ilo_lfs,1,2) & ilo_wap==1
            label define label_ilo_lfs 1 "1 - Employed" 2 "2 - Unemployed" 3 "3 - Outside Labour Force
            label value ilo_lfs label_ilo_lfs
            label var ilo_lfs "Labour Force Status"
```

Where:

Category 1 refers to persons classified as employed given that they satisfy the necessary conditions (`condition_1 OR … OR condition_n`) and they are part of the working-age population; category 2 comprises all persons classified as unemployed (`ilo_lfs!=1 AND seeking AND available`); and category 3 refers to outside the labour force.

The whole process, from obtaining the files until the publication of estimates, is marked by a strong collaboration between the ILO Department of Statistics, national focal points and national statistical offices. The materials developed by DPAU during the processing are available to national statistical offices upon request. The unit can also provide feedback on how to improve the survey questionnaire and on how to define and compute labour market indicators at the national level to better satisfy internationally-agreed standards.

The results derived from the anonymized microdata processing are in line with the recommendations on data and metadata flows, reporting and validation made by the United Nations Statistical Commission.[14] They are not intended to replace data or metadata officially transmitted by the national data provider to the ILO already following internationally-agreed standards, but rather to improve the quality of both statistical reporting systems.

---

[14] The Inter-agency and Expert Group on Sustainable Development Goal Indicators (IAEG-SDGs). "Global Consultation: Guidelines on Data Flows and Global Data Reporting for Sustainable Development Goals". January 2018.

4. Possibility of obtaining additional data not included in ILOSTAT

A maximum set of 74 variables can be derived through our current microdata processing workflow.[15] This includes the generation of variables by breakdowns using defined and aggregated categories: for instance, in the case of the classification of employment by occupation in main job, only one variable is needed for defining the ISCO classification at the 2-digit level (selected categories), 1-digit level and skill level (when available).[16] The following table shows the aggregations made for the ISCO-08 when passing from 1-digit level to the skill level:

Broad Skill Levels	ISCO-08
Skill levels 3 and 4 (high)	1. Managers
	2. Professionals
	3. Technicians and associate professionals
Skill level 2 (medium)	4. Clerical support workers
	5. Service and sales workers
	6. Skilled agricultural, forestry and fishery workers
	7. Craft and related trades workers
	8. Plant and machine operators, and assemblers
Skill level 1 (low)	9. Elementary occupations
Armed forces	0. Armed forces occupations
Not elsewhere classified	X. Not elsewhere classified

All the information is updated on a weekly basis and made available to the public in the ILO's online database, ILOSTAT. Users can download data using the customized download functionality, which allows to select specific queries (countries, indicators, time range and format). Entire datasets can also be downloaded using the bulk download facility.[17]

If the statistics sought by the data user are not part of the set of indicators disseminated on ILOSTAT, if feasible, ad-hoc tabulations can be produced by DPAU on the basis of processed microdata. Whether the inquiry is feasible will depend on the availability of the desired tabulation (or variables needed for its computation), the satisfaction of the reliability thresholds and compliance with national and regional data confidentiality regulations.

---

[15] For a complete list of the created variables, please refer to the Annex.

[16] For detailed information on the various ISCO classifications, see: https://www.ilo.org/public/english/bureau/stat/isco/

[17] For further information on the bulk download facility, please refer to: "Bulk download in ILOSTAT: Instructions and Guidelines", available at: http://www.ilo.org/ilostat-files/Documents/ILOSTAT_BulkDownload_Guidelines.pdf

5. Considerations and limitations

The ILO Department of Statistics is not responsible for survey design, data collection, cleaning or dissemination, and therefore does not have any type of legal ownership of the microdata files (and related files) mentioned in this document. The rights obtained by the ILO for using the anonymized datasets in the production of the statistics described in this Quick Guide are not transmissible to a third party.

Estimates produced and disseminated by the ILO may differ from those derived and published at the national level; these figures will be replaced when data produced by national statistical systems that are consistent with the international standards become available. ILOSTAT data users are able to clearly identify figures published as a result of ILO microdata processing by checking the related notes and metadata in the corresponding ILOSTAT table.

The thresholds established by the ILO in order to evaluate the reliability of estimates (please refer to section 3) might be different from those that are set at the national level. In many cases, those used by the Department of Statistics are stricter than the national ones, and therefore the dissemination of information is more limited.

In some cases, published annual figures are the result of annual averages of monthly or quarterly data points. Thus, the final figure related to a given year may differ from the annual figure disseminated at the national level where the country uses a different reference period for the annual data (e.g. one country may use one specific month or quarter as representative of the year rather than the annual average). For information on the data reference period used please refer to the notes on the values and tables on ILOSTAT.

6. Annex

Variable	Name of the variable in Stata: description
1. Identifier	ilo_key: identifier
2. Sample weight	ilo_weight: is used to give a certain weight to each observation in the sample in order for the sample to represent the overall population covered. It should stay at the unit level.
3. Time period	ilo_time: corresponds to the Gregorian calendar. It can take 3 forms: annual (2017); quarterly (2017Q1) or monthly (2017M01).
4. Geographical coverage	ilo_geo: follows the national definition of geographical areas and therefore the national boundary between urban and rural areas.
5. Sex	ilo_sex: follows directly what is in the microdata. However, to make sure that it follows the same standard across countries, we might have to recode 1 for male and 2 for female.
6. Age	ilo_age: reported age.
7. Age: 5 year bands	ilo_age_5yrbands: divides the population into 5 year age-bands until 64 and then a category includes everyone above 65.
8. Age: 10 year bands	ilo_age_10yrbands: divides the population into 10 year age-bands until 64 and then a category includes everyone above 65.
9. Age: aggregate categories	ilo_age_aggregate: 5 categories: children (<15); youth (15-24); adults (25-54) and older population: 55-64 and 65+.
10. Level of education	ilo_edu_isced97/isced11: maps the highest level of education of the respondent with ISCED 11. However some microdatasets still use a link to ISCED 97 (second best option).
11. Level of education: aggregate categories	ilo_edu_aggregate: 5 categories: less than basic, basic, intermediate, advanced and level not stated.
12. Educational attendance	ilo_edu_attendance: refers to the attendance in education of the respondent.
13. Marital status	ilo_mrts_details: divides the population into 6 categories: single, married, union/cohabiting, widowed, divorced/separated and not elsewhere classified.

14. Marital status: aggregate categories	ilo_mrts_aggregate: 3 categories: "Single/ Widowed/Divorced","Married/Union/Co-habiting" and "Not elsewhere classified".
15. Disability status	ilo_dsb_details: follow the Washington Group recommendation, at least 4 core questions (difficulties in seeing, hearing, walking and remembering) are asked to all the respondents with details as defined: "No - no difficulty", "Yes - some difficulty", "Yes - a lot of difficulty" or "Cannot do at all".
16. Disability status: aggregate categories	ilo_dsb_aggregate: as a second best option, an aggregated category can be created with just 2 options, either the person is with or without disabilities.
17. Working-age population	ilo_wap: persons aged 15 and above.
18. Labour force status	ilo_lfs: follows the "Resolution concerning statistics of work, employment and labour underutilization" adopted by the 19th International Conference of Labour Statisticians (October 2013). 3 categories: employed, unemployed, outside the labour force. Priority is given to employment over the other two categories, and to unemployment over outside the labour force. The three categories of labour force status are mutually exclusive and exhaustive. The sum of persons in employment and in unemployment equals the labour force.
19. Multiple job holder	ilo_mjh: refers directly to the answer of the respondent and whether he/she assessed to have only one job or more than one.
20. Status in employment in main job	ilo_job1_ste_icse93: maps the status in employment with the International Classification of Status in Employment (ICSE) 93.[18] If this is not possible we use aggregated categories (second best option). The ICSE-93 consists of the following groups: Employees, employers, own-account workers, members of producers' cooperatives,

[18] The 20th International Conference of Labour Statisticians held in October of 2018 in Geneva, has adopted the following resolution concerning statistics on work relationships, in substitution for the resolution of status in employment of 1993: https://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/meetingdocument/wcms_647343.pdf. No country has implemented it in their labour force survey yet and therefore no data has been produced using it.

	contributing family workers and workers not classifiable by status.
21. Status in employment in main job: aggregate categories	ilo_job1_ste_aggregate: as a second best option, an aggregated category can be created with just 2 options, either the person is an employee or self-employed.
22. Status in employment in second job	ilo_job2_ste_icse93: idem as for main job.
23. Status in employment in second job: aggregate categories	ilo_job2_ste_aggregate: idem as for main job.
24. Economic activity in main job: 2-digit level	ilo_job1_eco_isic4(_isic3)_2digits: maps the economic activity of the respondent with ISIC Revision 4 or Revision 3.1(second best option) at the 2-digit level.
25. Economic activity in main job: 1-digit level	ilo_job1_eco_isic4(_isic3): maps the economic activity of the respondent only at the first digit.
26. Economic activity in main job: aggregate level	ilo_job1_eco_aggregate: maps the economic activity of the respondent at the aggregated level of classification.
27. Economic activity in second job: 2-digit level	ilo_job2_eco_isic4(_isic3)_2digits: idem as for main job.
28. Economic activity in second job: 1-digit level	ilo_job2_eco_isic4(_isic3): idem as for main job.
29. Economic activity in main job: aggregate level	ilo_job2_eco_aggregate: idem as for main job.
30. Occupation in main job: 2-digit level	ilo_job1_ocu_isco08(_isco88)_2digits: maps the occupation of the respondent with ISCO-08 or ISCO-88 (second best option) at the 2-digit level.
31. Occupation in main job: 1-digit level	ilo_job1_ocu_isco08(_isco88): maps the occupation of the respondent only at the first digit.
32. Occupation in main job: aggregate categories	ilo_job1_ocu_aggregate: maps the occupation of the respondent at the aggregated level of classification.
33. Occupation in main job: skill level	ilo_job1_ocu_skill: maps the occupation of the respondent at the skill level of classification.
34. Occupation in second job: 2-digit level	ilo_job2_ocu_isco08(_isco88)_2digits: idem as for main job.
35. Occupation in second job: 1-digit level	ilo_job2_ocu_isco08(_isco88): idem as for main job.
36. Occupation in second job: aggregate categories	ilo_job2_ocu_aggregate: idem as for main job.

37. Occupation in second job: skill level	ilo_job2_ocu_skill: idem as for main job.
38. Institutional sector of economic activities	ilo_job1_ins_sector: refers directly to the answer of the respondent and whether he/she assessed to have a job in the public or private sector. If the question leads to more answers (public and private sectors are divided into sub-categories), we map them based on the national definitions to the broad categories of public and private sectors.
39. Usual hours of work in main job	ilo_job1_how_usual: follows the "Resolution concerning the measurement of working time Adopted by the 18th International Conference of Labour Statisticians (November-December 2008)". The variable produced is weekly hours usually worked in main job.
40. Usual hours of work in main job: bands	ilo_job1_how_usual_bands: an aggregated category can be created with 8 categories: no hours usually worked (0), between 1 and 14, between 15 and 29, between 30 and 34, between 35 and 39, between 40 and 48, 49 or more, and not elsewhere classified.
41. Actual hours of work in main job	ilo_job1_how_actual: follows the "Resolution concerning the measurement of working time Adopted by the 18th International Conference of Labour Statisticians (November-December 2008)". The variable produced is weekly hours actually worked in main job.
42. Actual hours of work in main job: bands	ilo_job1_how_actual_bands: an aggregated category can be created with 8 categories: no hours actually worked (0), between 1 and 14, between 15 and 29, between 30 and 34, between 35 and 39, between 40 and 48, 49 or more and not elsewhere classified.
43. Usual hours of work in second job	ilo_job2_how_usual: idem as for main job.
44. Usual hours of work in second job: bands	ilo_job2_how_usual_bands: idem as for main job.
45. Actual hours of work in second job	ilo_job2_how_actual: idem as for main job.
46. Actual hours of work in second job: bands	ilo_job2_how_actual_bands: idem as for main job.
47. Usual hours of work in all jobs	ilo_joball_how_usual: idem as for main job.
48. Usual hours of work in all jobs: bands	ilo_joball_how_usual_bands: idem as for main job.
49. Actual hours of work in all jobs	ilo_joball_how_actual: follows the "Resolution concerning the measurement of working time adopted by the 18th International Conference of Labour Statisticians

	(November-December 2008)". Data on weekly hours of work are presented, whenever possible, on the basis of the mean number of hours of work per week, and with reference to hours worked in all jobs of employed persons and in all types of working time arrangements (e.g. full-time and part-time).
50. Actual hours of work in all jobs: bands	ilo_joball_how_actual_bands: idem as for main job.
51. Working time arrangement	ilo_job1_job_time: it either follows the answer of the respondent based on a self-assessment question or based on a defined national hour threshold.
52. Type of contract	ilo_job1_job_contract: it classifies as permanent those persons having a contract without limit of time. Any contract with a specific duration will be classified as temporary.
53. Unit of production: formal/informal sector	ilo_job1_ife_prod: maps persons in informal or formal sector based on the 15th ICLS, the 17th ICLS and the Document on Measuring informality: a statistical manual on the informal sector and informal employment, published in 2013.[19] We define whether the person works in the formal sector (all workers in incorporated enterprises), or the informal sector (all workers in unincorporated enterprises that produce at least partly for the market and are not registered). The key questions from a labour force survey used are: institutional sector; destination of production; bookkeeping; registration of the unit; location of workplace; size and social security coverage.
54. Nature of job: formal/informal employment	ilo_job1_ife_nature: maps persons in informal or formal sector based on the 15th ICLS, the 17th ICLS and the Document on Measuring informality: a statistical manual on the informal sector and informal employment, published in 2013. We define whether the person's main job is formal or informal. If the person is an employee, this is defined based on the attachment to a national labour legislation or the entitlement to certain employment benefits (paid vacation, paid sick leave and contribution to pension funds). If the person is self-employed, it depends on the unit of production as defined in "ilo_job1_ife_prod".

---

[19] Available at : https://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_222979.pdf

	Finally, all contributing family workers are classified as holding informal jobs.
55. Monthly labour-related income in main job: employees	ilo_job1_lri_ees: follows the "Resolution concerning the measurement of employment-related income adopted by the 16ᵗʰ International Conference of Labour Statisticians (October 1998)".
56. Monthly labour-related income in main job: self-employed	ilo_job1_lri_slf: follows the "Resolution concerning the measurement of employment-related income adopted by the 16ᵗʰ International Conference of Labour Statisticians (October 1998)".
57. Time-related underemployed	ilo_joball_tru: follows the "Resolution concerning statistics of work, employment and labour underutilization adopted by the 19ᵗʰ International Conference of Labour Statisticians (October 2013)". Persons in time-related underemployment comprise all persons in employment, who satisfy the following three criteria during the reference period: want to work additional hours, currently available to work additional hours and worked less than a threshold relating to working time. The hour threshold is chosen according to national circumstances. In the absence of nationally-defined threshold, the most widely used practice of 35 hours per week is applied.
58. Cases of non-fatal occupational injury	ilo_joball_oi_case: follows the "Resolution concerning statistics of occupational injuries (resulting from occupational accidents), adopted by the 16ᵗʰ International Conference of Labour Statisticians (October 1998)".
59. Days lost due to cases of occupational injury	ilo_joball_oi_day: follows the "Resolution concerning statistics of occupational injuries (resulting from occupational accidents), adopted by the 16ᵗʰ International Conference of Labour Statisticians (October 1998)".
60. Category of unemployment	ilo_cat_une: If there is a direct question in the national questionnaire, we follow the answer given by the respondent. Otherwise, we check if a previous situation is defined somewhere else in the questionnaire.
61. Duration of unemployment	ilo_dur_details: collects information on the duration of the search for employment. It starts when the unemployed person began carrying out activities to "seek employment" or at the end of the last job. In case both are

	defined, we consider the shortest of the two time periods.
62. Duration of unemployment: aggregate categories	ilo_dur_aggregate: an aggregated category can be created with 4 categories: less than 6 months, 6 to less than 12 months, 12 months or more and not elsewhere classified.
63. Previous economic activity: 2-digit level	ilo_preveco_isic3(_isic4)_2digits: maps the previous economic activity of the respondent with ISIC Revision 4 or Revision 3 (second best option) at the 2-digit level.
64. Previous economic activity: 1-digit level	ilo_preveco_isic3(_isic4): maps the previous economic activity of the respondent at the first digit.
65. Previous economic activity: aggregate level	ilo_preveco_aggregate: maps the previous economic activity of the respondent at the aggregated level of classification.
66. Previous occupation: 2-digit level	ilo_prevocu_isco08(_isco88)_2digits: maps the previous occupation of the respondent with ISCO-08 or ISCO-88 (second best option) at the 2-digit level.
67. Previous occupation: 1-digit level	ilo_prevocu_isco08(_isco88): maps the previous occupation of the respondent at the first digit.
68. Previous occupation: aggregate categories	ilo_prevocu_aggregate: maps the previous occupation of the respondent at the aggregated level of classification.
69. Previous occupation: skill level	ilo_prevocu_skill: maps the previous occupation of the respondent by skill level.
70. Degree of labour market attachment	ilo_olf_dlma: follows the "Resolution concerning statistics of work, employment and labour underutilization adopted by the 19th International Conference of Labour Statisticians (October 2013)". The degrees of labour market attachment are classified as follows: seeking, not available (unavailable jobseekers); not seeking, available (available potential jobseekers); not seeking, not available, willing (willing non-jobseekers); not seeking, not available, not willing and not elsewhere classified.
71. Reasons for not seeking a job	ilo_olf_reason: follows the "Resolution concerning statistics of work, employment and labour underutilization Adopted by the 19th International Conference of Labour Statisticians (October 2013)". The following categories are defined: labour market (past failure to find a suitable job, lack of experience,

	qualifications or jobs matching the person's skills, lack of jobs in the area, considered too young or too old by prospective employers, does not know how/where to find a job, waiting for an answer after an application, seasonal break, bad weather); personal/family-related (own illness, disability, studies, social exclusion, pregnancy, presence of small children, refusal by family); does not need/want to work (retired, other sources of income: pensions, rents); and not elsewhere classified (other, lack of infrastructure).
72. Discouraged job-seekers	ilo_seekers: follows the "Resolution concerning statistics of work, employment and labour underutilization Adopted by the 19th International Conference of Labour Statisticians (October 2013)". Discouraged job-seekers refer to all persons of working age who, during the short reference period, were neither in employment nor in unemployment, were currently available, and did not seek employment in the recent past period for labour market related reasons (as listed under labour market reasons defined in "ilo_olf_reason").
73. Subsistence farming	ilo_sub: measurement of subsistence farming following the "Resolution concerning statistics of work, employment and labour underutilization Adopted by the 19th International Conference of Labour Statisticians (October 2013)".
74. Youth not in education, employment or training (NEET)	ilo_neet: Youth (defined as persons aged 15 - 24 years) who were, during a specified reference period (e.g., one week): not employed, not enrolled in school and not enrolled in a formal training programme (e.g. vocational training).