

The availability of statistical data in labor and skills domains covering South Mediterranean countries¹ for the production of a data visualization platform - “YEM Skills Panorama”

This summary provides a high-level overview of the process of identifying potential international depositories of (comparable) statistical data for representing the labour market and skills situation in South Mediterranean countries (Lebanon, Jordan, Israel, Palestine, Morocco, Libya, Algeria and Tunisia). It is to be noted that some statistical data is also likely to be available from national statistical sources; however, in these cases the data would need to undergo evaluation and possibly transformation to make it (where possible) comparable to data from other countries.

Statistical data “value chain” and international data depositories

Three main types of actors can be identified in relation to sourcing and presenting statistical data at the international level. Firstly, there are the data producing entities – national statistical offices and/or institutes who have the mandate for collecting and compiling statistical data at national level. These entities possess the most detailed statistical data and, depending on their capacity, can produce fine-grained extractions across multiple dimensions as are covered in national surveys. However, sourcing cross-national data from national institutions is burdensome, given the fact that the data normally is hosted in the language of the country and is structured and defined in line with national practices. To ensure cross-country data comparability, such data needs to be translated as well as transformed (to the extent possible) in line with international standards to make it sufficiently comparable across countries.

The task of compiling comparable cross-national data is carried out by international organisations within the domain of their expertise. This in particular concerns the International Labour Organisation (ILO) in relation to labour market statistical data and the UNESCO Institute of Statistics (UIS) in relation to education statistical data. In a more limited scope, the Organisation for Economic Cooperation and Development (OECD) also compiles cross-country comparable statistical data, however this is largely limited to covering the OECD Member countries. In addition, the World-Bank, primarily covering economic statistics, in its public data depository also seem to host some labour-market and education-related data. All the above-mentioned institutions thus host data depositories (databases) covering a large number of country-level **aggregated indicators**, sources and compiled from national statistical offices and/or institutes.

In addition, some of the above-mentioned institutions also pursue cross-national surveys, which enables constructing internationally **comparable micro-data sets**. The most notable of such international surveys in the labour market, education and skills domains include:

- Programme for International Assessment of Student Competences (PISA) by the OECD
- Programme for International Assessment of Adult Competences (PIAAC) by the OECD

¹ Algeria, Israel, Jordan, Lebanon, Libya, Morocco, Palestine and Tunisia.

- STEP Skills Measurement Programme by the World Bank
- School-to-work transition surveys by the ILO

Furthermore, other international organisations may also in their data collection exercises include certain labour market or education variables, but such individual variables are likely to provide only very limited level of detailed and breakdowns in relation to labour market/education. Thus, a decision was made to focus on data sources which cover statistics specifically dedicated to monitoring labour market, education and skills domain.

Finally, there are several initiatives to increase the accessibility of data related to skills at the international level. Notably, the OECD have developed two data depositories focusing on skills:

- **World Indicators of Skills for Employment**, developed around 2014/2015, compiling a large number of indicators covering around 200 countries and providing, where available, time-series data. Unfortunately, the data has not been updated since that time and the most recent year covered by the database is 2014.
- **Skills for Jobs database (S4J)**, developed around 2017/2018, covering OECD and EU countries but also several other countries (Argentina, Brazil, Malaysia, Peru and South Africa). At this stage only a single year is provided in the public database (2015 for labour market indicators and 2016 for mismatch indicators). In addition, for easier exploration of skills for job database, OECD also developed a visualisation platform for navigating the data.

The example of CEDEFOP's "Skills Panorama"

CEDEFOP's Skills Panorama seems to be one-of-a-kind data visualisation platform covering around 66 indicator domains related to skills and labour market for the 27 EU Member States. Each of the indicator domain includes one or several specific charts presenting the data through different breakdowns. CEDEFOP uses multiple pan-EU data sources for building the visualisation platform:

- EU Labour Force Survey (source - Eurostat)
- EU ICT Usage in Households and by Individuals (source – Eurostat)
- EU Income and Living Conditions Survey (source – Eurostat)
- National Accounts (source – Eurostat)
- European Skills and Job Survey (source - CEDEFOP)
- CEDEFOP Skills Forecast (source – CEDEFOP)
- European Skills Index (source – CEDEFOP)
- Skills OVATE (source – CEDEFOP)
- PIAAC (source – OECD)
- PISA (source – OECD)
- European database of tasks indices (JRC)

Data sources and their availability

Statistical data for labour market and skills therefore can be sourced from these main types of sources:

- General population surveys, most notably national labour force surveys which are, to a large extent, implemented in most of the world countries. National labour force surveys are aggregated into a single database by the ILO. At least some statistical data in the labour market

domain is available for most of the YEM countries, however regular time-series data is available only for less than half of YEM countries (also depending on the specific indicator).

- Specialised surveys, most notably the PISA and PIAAC surveys implemented by the OECD as well as the STEP survey implemented by the World Bank. In terms of YEM countries, PIAAC covers only a single country (Israel) while PISA covers several of the countries depending on the wave of the survey. Specialised European surveys do not cover any of the YEM countries.
- Administrative education data. Such data from national data sources is compiled by UNESCO Institute of Statistics via the UOE data collection. At least some statistical data in the education domain is available for most of the YEM countries, however regular time-series data is available only for a few YEM countries (also depending on the specific indicator).
- Aggregated Economic data (i.e., GDP) is largely available for all the YEM countries.

In addition, ILO produces modelled estimates for core labour market indicators which covers universally all the countries in the world. Given such universal coverage and presence of key skills-related indicators, it is recommended for the YEM Skills Panorama to focus on representing data from ILO modelled estimates database. This data could be supplemented by data coming from PISA (available for a significant number of YEM countries) as well as economic (GDP) data, covering all YEM countries.

Availability of data cross-tabulations

The final consideration as regards data availability for the data visualisation platform is the availability of multiple breakdowns within a single data source/table. In some of the indicator domains in CEDEFOP's skills panorama, the data is disaggregated across multiple dimensions in a single data. One example of such desegregation is the distribution of employment across occupational and sectoral categories. This, for example, allows to show what how many individuals are working in which occupations in one (or several) selected economic sectors. Conversely, the analyst could select one specific occupational group and see how the jobs belonging to such occupational groups is distributed across sectors of the economy. Unfortunately, such level of data-desegregation is currently not included in the ILO modelled estimates database; whereas in the general ILO tables the coverage of YEM countries is rather poor.

Final list of indicator domains selected for data visualisation platform

Taking into consideration the results of data mapping as reported above, it is proposed for data visualisation platform to include only domains which have universal or very good coverage of YEM countries. Such indicator domains include:

- Population data
- Working-age population data (such data includes, as a rule of thumb, individuals between 15 and 64 years of age)
- Labour force data (labour force includes individuals active in the labour market, i.e. employed and unemployed persons)
- Employment data
- Unemployment data
- Self-employment data
- Data about young adults not in employment, education or training (NEET)

- Distribution of jobs in terms of skills requirements
- Distribution of jobs in terms of occupations
- Distribution of jobs in terms of economic sectors
- PISA data (while this data does not cover all YEM countries, given its thematic relevance it is proposed to be included in the data visualisation platform)
- GDP data