

Distr.: General  
26 October 2017

English

---

## **United Nations Economic Commission for Europe**

Conference of European Statisticians

### **Work Session on Migration Statistics**

Geneva, Switzerland

30-31 October 2017

Item 3 of the provisional agenda

### **Data integration and administrative data**

## **Questionnaire on data integration for measuring migration**

**Note by the UNECE Task Force on Data Integration for Measuring Migration**

### *Summary*

This document is an addendum to the draft report on 'Data Integration for Measuring Migration', prepared by the UNECE Task Force on 'Data Integration for Measuring Migration' and includes a questionnaire for a survey conducted by the Task Force in the preparation of the report

---

## Questionnaire on data integration for measuring migration

21 September 2016

### Introduction

The UNECE Task Force on Data Integration for Measuring Migration is currently investigating the use of data integration for measuring international migration. The main outcomes of the work by the Task Force are to establish guidance, from use of data integration, on the compilation of migration statistics from a range of data sources and on the presentation of comparisons of immigration/ emigration by country of origin/ destination.

'Data integration' is often mentioned as one of most effective approaches for enriching and improving statistical information. This questionnaire intends to collect information on current practices by National Statistical Offices in the use of combined or integrated data sources for the measurement of immigration and emigration flows, and for generating statistics on the migrant population and the foreign-born population.

*For the purposes of this questionnaire, data integration refers to the production of international migration statistics outputs (including both statistical reports and data) by:*

- *matching data between multiple data sources at the unit record level ("micro level") and/or*
- *combining data sources at the "macro level" using statistical modelling*

If you have any questions or concerns regarding the concept of data integration, the scope of this survey or the contents of this questionnaire you may contact us at [Social.Stats@unece.org](mailto:Social.Stats@unece.org).

### CONTACT INFORMATION

Name: \_\_\_\_\_

Position/title: \_\_\_\_\_

Email Address (for possible follow-up):

\_\_\_\_\_

Institution: \_\_\_\_\_

Country: \_\_\_\_\_

### A. INTERNATIONAL MIGRATION FLOWS

1. Does your country produce statistics or collect data on international migration flows or has it done so in the past?

\_ YES

\_ NO => GO TO 9

2. Are/were these statistical outputs based on:

\_ A single source: => GO TO 9

\_ More than one data source

Describe in a few lines the name and main features of these source(s).

**3. How are/were these data sources integrated?**

- Integrated at the individual record level (matching on individual data)
- Integrated at a macro-level (using statistical modelling or other methods to combine aggregated data)
- Integrated at both the individual record and macro-level
- Sources not combined to produce a statistical output => **GO TO 9**
- Other (please specify) \_\_\_\_\_

**4. In the context of migration flow statistics, observation units are migration events, individuals, households or some combination of the above. When integrating data from different sources to generate these statistical outputs, did the data sources have:**

- identical observation units
- mutually exclusive observation units (combination of two distinct population sub-groups)
- partially overlapping observation units

Comments \_\_\_\_\_

**5. Many different demographic or other variables are often included in migration flow statistics. When integrating data from different source to generate these statistical outputs, did the data sources have:**

- identical variables
- mutually exclusive variables (no overlap other than ID variables used for matching)
- partially overlapping variables

Comments \_\_\_\_\_

**6. Are/were data integrated as part of:**

- Regularly published statistical outputs => **GO TO 8**
- Past pilot studies
- Ongoing pilot studies
- Both regularly published statistics and past or ongoing pilot studies
- Other (including ad hoc outputs) => **GO TO 8**

**7. What was the result of the pilot study?**

- Will implement techniques in future regular data publication
- No implementation yet, further study required
- No implementation, techniques not suitable
- Other (please specify)

8. Describe briefly the sources and techniques used as well as any other sources and techniques used in the past for the purpose of data integration. If a link to a methodological document is available please provide it.

**B. STOCKS OF FOREIGN POPULATION /POPULATION BORN ABROAD**

9. Does your country produce statistics or collect data on foreign population or population born abroad or has it done so in the past?

YES  
 NO => GO TO 17

10. Are/were these statistical outputs based on:

A single source: => GO TO 17  
 More than one data source

Describe in a few lines the name and main features of these source(s).

\_\_\_\_\_

11. How are/were these data sources integrated?

Integrated at the individual record level (matching on individual data)  
 Integrated at a macro-level (using statistical modelling or other methods to combine aggregated data)  
 Integrated at both the individual record and macro levels  
 Sources not combined to produce a statistical output => GO TO 17  
 Other (please specify) \_\_\_\_\_

12. In the context of migration stock statistics, observation units are individuals, households or some combination of the above. When integrating data from different sources to generate these statistical outputs, did the data sources have:

identical observation units  
 mutually exclusive observation units (combination of two distinct population sub-groups)  
 partially overlapping observation units

Comments \_\_\_\_\_

13. Many different demographic or other variables are often included in migration stock statistics. When integrating data from different source to generate these statistical outputs, did the data sources have:

identical variables  
 mutually exclusive variables (no overlap other than ID variables used for matching)  
 partially overlapping variables

Comments \_\_\_\_\_

**14. Are/were data integrated as part of:**

- Regularly published statistics => **GO TO 16**
- Past pilot studies
- Ongoing pilot studies
- Both regularly published statistics and past or ongoing pilot studies
- Other (including ad hoc outputs) => **GO TO 16**

**15. What was the result of the pilot study?**

- Will implement techniques in future regular data publication
- No implementation yet, further study required
- No implementation, techniques not suitable
- Other (please specify)

**16. Describe briefly the sources and techniques used as well as any other sources and techniques used in the past for the purpose of data integration.**

**If a link to a methodological document is available, please provide it.**

**C. OTHER RELEVANT INFORMATION ON MIGRANTS OR FOREIGN POPULATION**

**17. Does your country integrate data from different sources to generate other statistical outputs *not already mentioned* in this questionnaire on migrants or foreign population (residence or job permits, regulated migration, short-term migration, asylum seekers, socio-economic characteristics of migrants, etc) or has it done so in the past?**

- YES
- NO => **GO TO 24**

Describe the content of the most relevant output (most successful implementation of data integration)

**18. Referring to the most relevant statistical output from the previous question, how are/were different sources used to generate this output?**

- Different sources integrated at the individual record level (matching on individual data)
- Different sources integrated at a macro-level (using statistical modelling or other methods to combine aggregated data)
- Different sources both integrated at the individual record and macro levels
- Different sources not combined to produce a statistical output => **GO TO 24**

Describe in a few lines the name and main features of these source(s).

**19. In the context of statistics on migrants or foreign population, observation units are migration events, individuals, households or some combination of the above. Referring to the most relevant statistical output described previously, did the data sources have:**

- identical observation units*

- mutually exclusive observation units (combination of two distinct population sub-groups)*
- partially overlapping observation units*

Comments \_\_\_\_\_

**20. Many different demographic or other variables are often included in statistics of migrants or foreign population. When integrating data from different source to generate these statistical outputs, did the data sources have:**

- identical variables*
- mutually exclusive variables (no overlap other than ID variables used for matching)*
- partially overlapping variables*

Comments \_\_\_\_\_

**21. Referring to the most relevant statistical output described previously, are/were these data integrated as part of:**

- Regularly published statistics => **GO TO 23**
- Past pilot studies
- Ongoing pilot studies
- Both regularly published statistics and past or ongoing pilot studies
- Other (including ad hoc outputs) => **GO TO 23**

**22. What was the result of the pilot study?**

- Will implement techniques in future regular data publication
- No implementation yet, further study required
- No implementation, techniques not suitable
- Other (please specify)

**23. Referring to the most relevant statistical output described previously, describe briefly the sources and techniques used as well as any other sources and techniques used in the past for the purpose of data integration. If a link to a methodological document is available, please provide it.**

**24. Any other comments / suggestions (max 10 lines)**

Thank you very much for your cooperation!

Any additional methodological documents that you would like to share can be sent to [Social.Stats@unece.org](mailto:Social.Stats@unece.org).

Results of this survey will be made available as part of the final report of the UNECE Task Force on Data Integration for Measuring Migration (expected in 2018).