

## **DIOC 2015/16**

### **VARIABLES, COVERAGE, SOURCES**

The Database on Immigrants in OECD Countries (DIOC) 2015/16 is an update of DIOC 2000/01, DIOC 2005/06 and DIOC 2010/11. This note provides detailed information on the variables in the data files, the coverage of the data, the data sources, and the methods used to ensure reliability and protect confidentiality. DIOC 2015/16 (release 1.0) contains four separate files covering a specific theme: age and nationality, duration of stay, labour force status and occupation. As shown in the tables below, each file includes a number of core variables (detailed country of birth, education and sex), which makes it possible to generate a great variety of cross-tabulations on the characteristics of the foreign-born populations in OECD countries by country of birth. However, each table is independent and cross-tabulations between tables are not possible.

## 1.1. Classifications and variables in data files

### 1.1.1. Variables by data file

#### File A - Reference population: all persons

variable name	explanation	categories / classification
country	country of residence	ISO 3166-1 (alpha 3)
coub	country of birth	ISO 3166-1 (alpha 3)
regionb	region of birth	AFRI = Africa ASIA = Asia EURO = Europe NOAM = Northern America OCEA = Oceania SCAC = South and Central America and the Caribbean UNK = Unknown
sex		1 = Male 2 = Female
age	age groups	1 = 0-14 2 = 15-24 3 = 25-34 4 = 35-44 5 = 45-54 6 = 55-64 7 = 65+ 26 = 15-64 99 = Unknown
age_ifs	age groups (broad)	0 = 0-14 1 = 15-24 2 = 25-64 3 = 65+ 12 = 15-64 99 = Unknown
edu_detailed	educational attainment (detailed) ISCED-08	1 = ISCED 01/02      Pre-primary education 2 = ISCED 1          Primary education or first stage of basic education 3 = ISCED 2          Lower secondary education 4 = ISCED 3          Upper secondary education 5 = ISCED 4          Post-secondary non-tertiary education 6 = ISCED 5/6/7      First stage of tertiary education (Bachelor and Master) 7 = ISCED 8          Second stage of tertiary education (PhD) 12 = ISCED 01/02/1    Pre-primary/primary education 123 = ISCED 01/02/1/2 Pre-primary/primary/lower secondary education 56 = ISCED 4/5/6/7    Post-secondary non-tertiary or first stage of tertiary (Bachelor and Master) education 67 = ISCED 5/6/7/8    First stage (Bachelor and Master) or second stage (PhD) of tertiary education 99 = Unknown
edu_ifs	educational attainment (broad)	1 = ISCED 0/1/2      low 2 = ISCED 3/4        medium 3 = ISCED 5/6/7/8    high 99 = Unknown
nationality	nationality (detailed)	1 = National at birth 2 = National by acquisition (foreigner at birth) 3 = Foreigner with the country of birth as the country of nationality 4 = Foreigner with other nationality 34 = Foreigner (exact nationality not specified) 99 = Unknown
national	indicates whether national	0 = Foreigner 1 = National 99 = Unknown
fborn	indicates whether foreign-born	0 = Native-born 1 = Foreign-born 99 = Unknown
oeed	indicates whether country of residence is an OECD member	0 = not OECD member 1 = OECD member
oeedb	indicates whether country of birth is an OECD member	0 = not OECD member 1 = OECD member

**File B - Reference population: persons aged 15 and above**

<b>variable name</b>	<b>explanation</b>	<b>categories / classification</b>
<b>country</b>	country of residence	ISO 3166-1 (alpha 3)
<b>coub</b>	country of birth	ISO 3166-1 (alpha 3)
<b>regionb</b>	region of birth	AFRI = Africa ASIA = Asia EURO = Europe NOAM = Northern America OCEA = Oceania SCAC = South and Central America and the Caribbean UNK = Unknown
<b>sex</b>		1 = Male 2 = Female
<b>edu_cen</b>	educational attainment ISCED 2011	1 = ISCED 01/02/1 Pre-primary/Primary education 2 = ISCED 2 Lower secondary education 3 = ISCED 3/4 Upper secondary/Post-secondary non-tertiary education 4 = ISCED 5/6/7 First stage of tertiary education (Bachelor and Master) 5 = ISCED 8 Second stage of tertiary education (PhD) 12 = ISCED 01/02/1/2 Pre-primary/primary/lower secondary education 45 = ISCED 5/6/7/8 First stage (Bachelor and Master) or second stage (PhD) of tertiary education 99 = Unknown
<b>edu_ifs</b>	educational attainment (broad)	1 = ISCED 0/1/2 low 2 = ISCED 3/4 medium 3 = ISCED 5/6/7/8 high 99 = Unknown
<b>dos_cen</b>	duration of stay (detailed)	0 = native-born 1 = one year or less 2 = one to five years 3 = five to ten years 4 = ten to twenty years 5 = more than twenty years 12 = zero to five years 345 = five years or more 45 = ten years or more 99 = unknown
<b>dos_ifs</b>	duration of stay (broad)	0 = native-born 1 = five years or less 2 = five to ten years 3 = more than ten years 23 = more than 5 years 99 = unknown
<b>fborn</b>	indicates whether foreign-born	0 = Native-born 1 = Foreign-born 99 = Unknown
<b>oecd</b>	indicates whether country of residence is an OECD member	0 = not OECD member 1 = OECD member
<b>oecdb</b>	indicates whether country of birth is an OECD member	0 = not OECD member 1 = OECD member

## File C - Reference population: persons aged 15 and above

variable name	explanation	categories / classification
<b>country</b>	country of residence	ISO 3166-1 (alpha 3)
<b>coub</b>	country of birth	ISO 3166-1 (alpha 3)
<b>regionb</b>	region of birth	AFRI = Africa ASIA = Asia EURO = Europe NOAM = Northern America OCEA = Oceania SCAC = South and Central America and the Caribbean UNK = Unknown
<b>sex</b>		1 = Male 2 = Female
<b>age</b>	age groups (broad)	1 = 15-24 2 = 25-64 3 = 65+ 1564 = 15-64 99 = Unknown
<b>edu_cen</b>	educational attainment ISCED 2011	1 = ISCED 01/02/1 Pre-primary/Primary education 2 = ISCED 2 Lower secondary education 3 = ISCED 3/4 Upper secondary/Post-secondary non-tertiary education 4 = ISCED 5/6/7 First stage of tertiary education (Bachelor and Master) 5 = ISCED 8 Second stage of tertiary education (PhD) 12 = ISCED 01/02/1/2 Pre-primary/primary/lower secondary education 45 = ISCED 5/6/7/8 First stage (Bachelor and Master) or second stage (PhD) of tertiary education 99 = Unknown
<b>edu_lfs</b>	educational attainment (broad)	1 = ISCED 0/1/2 low 2 = ISCED 3/4 medium 3 = ISCED 5/6/7/8 high 99 = Unknown
<b>lfs</b>	labour force status	1 = employed 2 = unemployed 3 = inactive 99 = Unknown
<b>fborn</b>	indicates whether foreign-born	0 = Native-born 1 = Foreign-born 99 = Unknown
<b>oecd</b>	indicates whether country of residence is an OECD member	0 = not OECD member 1 = OECD member
<b>oecdb</b>	indicates whether country of birth is an OECD member	0 = not OECD member 1 = OECD member

**File D - Reference population: employed persons aged 15 and above**

<b>variable name</b>	<b>explanation</b>	<b>categories / classification</b>
<b>country</b>	country of residence	ISO 3166-1 (alpha 3)
<b>coub</b>	country of birth	ISO 3166-1 (alpha 3)
<b>regionb</b>	region of birth	AFRI = Africa ASIA = Asia EURO = Europe NOAM = Northern America OCEA = Oceania SCAC = South and Central America and the Caribbean UNK = Unknown
<b>sex</b>		1 = Male 2 = Female
<b>edu_cen</b>	educational attainment ISCED 2011	1 = ISCED 01/02/1    Pre-primary/Primary education 2 = ISCED 2            Lower secondary education 3 = ISCED 3/4        Upper secondary/Post-secondary non-tertiary education 4 = ISCED 5/6/7     First stage of tertiary education (Bachelor and Master) 5 = ISCED 8            Second stage of tertiary education (PhD) 12 = ISCED 01/02/1/2    Pre-primary/primary/lower secondary education 45 = ISCED 5/6/7/8    First stage (Bachelor and Master) or second stage (PhD) of tertiary education
<b>edu_ifs</b>	educational attainment (broad)	1 = ISCED 0/1/2        low 2 = ISCED 3/4         medium 3 = ISCED 5/6/7/8    high 99 = Unknown
<b>occ_1d</b>	ISCO 2008 occupations (broad)	ISCO 1-digit occupation codes
<b>occ_08</b>	ISCO 2008 occupations	ISCO 2-digit occupation codes
<b>occ_oth</b>	other occupations	other occupation codes which cannot be matched with ISCO occupations
<b>fborn</b>	indicates whether foreign-born	0 = Native-born 1 = Foreign-born 99 = Unknown
<b>oecd</b>	indicates whether country of residence is an OECD member	0 = not OECD member 1 = OECD member
<b>oecdb</b>	indicates whether country of birth is an OECD member	0 = not OECD member 1 = OECD member
<b>skill_occ</b>	skill level of occupation	1 = ISCO skill level 1 2 = ISCO skill level 2 3 = ISCO skill level 3 99 = Unknown
<b>overqualified</b>	indicates whether worker is overqualified (highly educated only)	0 = not overqualified 1 = overqualified 99 = Unknown

### 1.1.2. Country of birth

The database provides information on countries of birth using the ISO 3166-1 alpha-3 classification (see <https://www.iso.org/iso-3166-country-codes.html> for more information). To produce a consistent list of countries of birth across receiving OECD countries, some minor adjustments were made, especially with respect to small islands and overseas territories. This recoding may explain small differences that might exist with national estimates for foreign-born and native-born populations. The following recoding was carried out:

Australia	Denmark	France	United Kingdom	Portugal	United States
- Heard & McDonald Islands	- Faroe Islands - Greenland	- French southern - Tromelin Island - Guadeloupe - Martinique - Réunion - Juan De Nova - Guyane - Mayotte - Wallis and Fortuna - Saint Martin - Saint Barthélemy - New Caledonia - Saint-Pierre-et-Miquelon	- Channel Islands - Isle of Man - Isle of Sark - South Georgia and - Jersey - British Indian Territory	- Madeira Islands - Azores Islands	- US minor islands - Christmas Island - Wake Island - Palmyra Atoll - Navassa Island - Midway Islands - Johnston Atoll - Howland Island - Baker Island
Netherlands	Finland	Norway			
- Bonaire, Saint Eustasius, and Saba - Sint Maarten	- Aland Island	- Svalbard and Jan Mayen - Bouvet Island			

With regard to split, recomposed or newly constituted countries, there was little choice but to respect the coding in the national data collection, which varies from one country to another. In the United States, for example, most people born in the Korean peninsula indicated they were born in Korea and not in South or North Korea. Similarly, in the Japanese census data, it is not possible to identify in which part of the Korean peninsula a person was born. For this reason, the place of birth of people born on the current territories of the Republic of Korea or the Democratic People’s Republic of Korea is therefore noted KOREA-NS. For some people in some OECD countries, the Czech Republic and the Slovak Republic are aggregated under the name of the former Czechoslovakia (CSFR) as country of birth. The same applies to the former USSR and the former Yugoslavia (FYUG).

Due to confidentiality and reliability issues or imprecise information, the place of birth is sometimes recorded at the continental level such as Africa, Asia, Europe, etc. instead of the country level (see additional details in the section on “Reliability thresholds and confidentiality issues” below). Regarding imprecise or missing information on the place of birth, for foreign-born people whose country of birth is unknown or is too imprecise to fall into one of the continental categories, the country of birth is coded as Other (OTH). Since the definition of the category Other is specific to each OECD reporting country, the overall Other category does not have any particular meaning in terms of country or region of origin.

For people whose birth status (native or foreign-born) is unknown, the country of birth is coded as Unknown (UNK). The detailed list of the countries and regions of birth represented in the database is provided in Table A.1.

#### ► *Special cases*

**Austria, Belgium, Czech Republic, Estonia, Iceland and Poland** –These countries were unable to provide data for this update of DIOC. As a result, the DIOC 2010/11 data were

adjusted via raking to be consistent with existing and publicly available data for 2015/16. The European Labour Force Survey (EU LFS) provided updated distributions by aggregated regions of birth for sex, age, education, nationality, duration of stay, labour force status and occupation. Eurostat's Unified Demographic Data Collection (UNIDEMO, [https://ec.europa.eu/eurostat/cache/metadata/en/demo\\_pop\\_esms.htm](https://ec.europa.eu/eurostat/cache/metadata/en/demo_pop_esms.htm)) provided updated totals by detailed country of birth, sex, and age. DIOC 2010/11 data were first iteratively raked to the control distributions from the EU LFS. The resulting intermediate totals were then iteratively raked to the UNIDEMO totals.

In the case of Poland, no detailed data on country of birth exists from the Eurostat Unified Demographic Data Collection, and EU LFS data were deemed insufficient for the adjustment methodology. The UNIDEMO does, however, include aggregated regions of birth for Poland. The DIOC 2010/11 totals for Poland were thus directly raked to the grouped UNIDEMO totals.

**Germany** – Missing country of birth information of ethnic Germans was imputed using the distribution of ethnic Germans for whom such information is available. Imputations were made randomly based on available data categorised by gender, age, education, duration of stay and labour force status.

**Israel** – Country of birth is not available for the 0-14 age group.

**Japan, Korea** – Since data based on the country of birth are not available, a citizenship-based definition of migration is used. Immigrants are defined to be individuals living in the host country and not holding the nationality of that country. This definition has obvious shortcomings. First, while there is an overlap between foreigners and the foreign-born, there is generally a significant difference between the two population figures. Second, this difference between the foreigners and the foreign-born applies not only to the absolute numbers, but also to the distribution of demographic and labour market characteristics. This implies that the Japanese and Korean data are not directly comparable to those of other countries.

**New Zealand** – Data were provided on the total number of foreign-born for 64 countries of birth, but data on demographic and labour-market characteristics were supplied only for the top 10 countries of birth (for only the top six countries in the case of labour-force status) and for the native-born. Demographic and labour market characteristics for all other countries of birth were imputed by applying the distributions by country of birth for those variables in DIOC 2010/11 to the total for each country of birth.

**Turkey** – No data were provided for this update of DIOC. As a result, the DIOC 2010/11 data were adjusted via raking to be consistent with existing and publicly available data for 2015/16. The Turkish Labour Force Survey (LFS) provided updated distributions by aggregated regions of birth for sex, age, education, nationality, duration of stay, labour force status and occupation. Turkstat's Address Based Population Registration System Statistics (ABPRSS, [http://web.turkstat.gov.tr/PreIstatistikTablo.do?istab\\_id=2428](http://web.turkstat.gov.tr/PreIstatistikTablo.do?istab_id=2428)) provided updated totals by detailed country of birth, sex, and age. DIOC 2010/11 data were first iteratively raked to the control distributions from the LFS. The resulting intermediate totals were then iteratively raked to the ABPRSS totals.

**United States** – It is not possible to release data on countries of origin with populations in the United States under 15 000 due to the regulations of the American Community Survey. However, the total number of foreigners is provided in the total number by region of origin, and missing country of birth information was imputed using the data in DIOC 2010/11. The assumption on which this imputation is based is that foreigners in the United States

(aged 15+) originating from those suppressed countries have the same sex, age and education distribution as in 2010/11. Note that there is no detailed information on nationality or education (instead there is the broad nationality variable and the broad educational attainment variable) in File A; and no detailed information on education in Files B, C or D: There is no age information in File C for those whose country of birth was imputed.

### ***1.1.3. Age***

File A generally includes all individuals of all ages, although for some countries the 0-14 age group may be omitted. Files B, C and D include only the population aged 15 years and over.

#### **► *Special cases***

**Austria** – In File D, no detailed age information is available, and these individuals are recoded as 1564.

**New Zealand** – Detailed information on age is not available and no information is available for the population aged 0-14 years. The variable `age_lfs` provides a broader breakdown of age.

**Turkey** – Detailed age between 15 and 64 is not available for some combinations of country of birth and sex for File A. These individuals are recoded as 23 for the age variable, and as 12 for the `age_lfs` variable. In File D, no detailed age information is available, and these individuals are recoded as 1564.

### ***1.1.4. Education***

The International Standard Classification of Education (ISCED 2011; <http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf>) was used as a baseline. Due to the differences in national systems and the resulting imperfect mapping, there is no observation in category ISCED 4 for Australia, Austria, Denmark, France, Israel, Korea, Slovenia, Switzerland, the United Kingdom and the United States. For Canada, persons with ISCED 4 level are not separable from persons with ISCED 5/6/7 levels.

To guarantee comparability of data across countries, a broader three-category classification (low, medium, high) has been created for all countries under the variable named `edu_lfs`. This broader classification distinguishes between low levels of education (ISCED 0 to 2), medium levels of education (ISCED 3 and 4) and high levels of education (ISCED 5 to 8).

#### **► *Special cases***

**Austria, Poland, New Zealand and Turkey** – No detailed information on education is available. Educational attainment is coded according to the broader classification (`edu_lfs`) in all Files.

**Belgium, Czech Republic, Estonia, Greece, Hungary, Iceland, Slovak Republic, Slovenia** – Information on educational attainment for these countries is based on information in DIOC 2010/11 derived from ISCED 1997 instead of ISCED 2011, and as such the pre-primary level (ISCED 2011 01/02) cannot be separated from the primary/first stage of basic education level (ISCED 2011 1). Individuals with this level of education are coded as 12 in the education variable in all files. No detailed information on education is available for the Czech Republic in File B.



**Canada** – Individuals with primary education and lower secondary education cannot be separated from each other and they are recoded as 123 in File A. Additionally, individuals with post-secondary non-tertiary education (ISCED 4) cannot be separated from individuals with first stage of tertiary education (Bachelor and Master, ISCED 5/6/7) and they are recoded as 56 in File A.

For Files B, C, and D, individuals with primary education and lower secondary education cannot be separated from each other and they are recoded as 12. Individuals with post-secondary non-tertiary education (ISCED 4) are included with individuals with first stage of tertiary education (Bachelor and Master, ISCED 5/6/7) instead of with individuals with upper secondary education (ISCED 3) as this was the coding in the supplied data.

**Chile** – Individuals with an ISCED 5/6/7 level of education cannot be separated from those with an ISCED 8 level of education and are recoded as 67 in File A.

**Finland** – Education data are based on Statistics Finland’s Register of Completed Education and Degrees, which refer to the post-comprehensive school educational qualifications and degrees attained (ISCED 3/4/5/6/7/8). It is not possible to separate persons with low education (ISCED 01/02/1/2) from those whose education is unknown. Both individuals with low education and unknown education are coded as 123 in File A and 12 in Files B, C and D. Since the Register of Completed Education and Degrees does not have total coverage on the educational attainment of recent immigrants, low-educated are overrepresented among them.

**Ireland** – Individuals with an ISCED 01/02 level of education (pre-primary) cannot be separated from those with an ISCED 1 level of education (primary) and are recoded as 12 in File A.

**Japan** – No information is available on educational attainment.

**Korea** – No information is available on educational attainment in Files C or D.

**Latvia** – Information on educational attainment is available for persons aged 15-74 years.

**United Kingdom** – Education has been recoded by the OECD based on the correspondence table provided in the fourth volume of the 2016 LFS user guide (pages 101 to 108): [http://doc.ukdataservice.ac.uk/doc/8367/mrdoc/pdf/lfs\\_user\\_guide\\_vol4\\_derivedvariables\\_2016.pdf](http://doc.ukdataservice.ac.uk/doc/8367/mrdoc/pdf/lfs_user_guide_vol4_derivedvariables_2016.pdf). Education is only recorded for those aged 16 and above.

### *1.1.5. Nationality*

For some countries, detailed nationality information is recorded, whereas for some others only a broader breakdown was possible. To guarantee comparability of data across countries, the broader (two-category) classification has been extended to all countries under the variable named national.

#### ► *Special cases*

Australia, Chile, the Czech Republic, Denmark, Estonia, Greece, Iceland, Ireland, Israel, Italy, Korea, Latvia, Mexico, Netherlands, Slovenia, Spain and the United Kingdom only provide the broader breakdown between nationals and foreigners.

There is no information on nationality for Austria, Japan, New Zealand, Poland and Turkey.

### 1.1.6. Duration of Stay

Duration of stay is only recorded for the foreign-born population aged 15 and over. For most of the countries, duration of stay is recorded in five categories, whereas for some others only a broader breakdown was possible. To guarantee comparability of data across countries, the broader (three-category) classification has been extended to all countries, under the variable `dos_lfs`.

#### ► *Special cases*

**Austria, Czech Republic, Iceland and Turkey** – Information on duration of stay is only available with the broader breakdown.

**Mexico** – Individuals who arrived more than five years ago cannot be separated from those who arrived more than 10 years ago and they all are recorded as 345 in File B. Individuals who arrived less than one year ago cannot be separated from those who arrived less than five years ago and are coded as 12 in File B.

**Japan, Korea** – Information on duration of stay is not available.

**Poland** – Individuals who arrived more than 20 years ago cannot be separated from those who arrived more than 10 years ago and they are recorded as 45 in File B.

### 1.1.7. Labour force status

For most countries, labour force status is recorded according to guidelines issued by the International Labour Organisation (see the Resolution of the 13th International Conference of Labour Statisticians at [https://www.ilo.org/global/statistics-and-databases/meetings-and-events/international-conference-of-labour-statisticians/WCMS\\_221500/lang--en/index.htm](https://www.ilo.org/global/statistics-and-databases/meetings-and-events/international-conference-of-labour-statisticians/WCMS_221500/lang--en/index.htm)).

### 1.1.8. Occupations (only for employed people aged 15+)

Occupations are recorded according to the International Labour Organization's 2008 revision of the International standard classification of occupations (ISCO-08, see <https://www.ilo.org/public/english/bureau/stat/isco/index.htm>), at the sub-major group (two digit) level. Some countries have only been able to provide data at the major group (one digit) level. The list of occupations in ISCO (one-digit level) is detailed in Table A.3. The complete list of ISCO codes is available at <http://www.ilo.org/public/english/bureau/stat/isco/index.htm>.

Countries do not necessarily record activities in censuses or registers according to ISCO and may use national classifications instead. Most national classifications are close enough to ISCO to allow a correct mapping at the two-digit level, but this is not always the case. When the internal logic of the national classification is too distant from that of ISCO, the mapping is at best imperfect.

#### ► *Special cases*

**Japan, United States** – These countries provided data using national classifications that are not easily mapped to ISCO categories. The categories of their respective national classifications are recorded in the variable `occ_other`. See Table A.4 for occupation codes used in the Population Census of Japan. The occupation codes for the United States are the 23 major groups of the 2010 Standard Occupational Classification (SOC). See Table A.5 for the 2010 SOC occupational codes and groups, and see <https://www.bls.gov/soc/2010/home.htm> for more information on the 2010 SOC.

**Iceland, Korea, Mexico and Portugal** – Information on occupations is only available for the ISCO major-group (one digit) level.

**New Zealand, Poland and Turkey** – No information on occupations is available.

### *1.1.9. Skill level and Overqualification*

A person is considered overqualified if he/she has tertiary education (ISCED 5 to 8) and a skill level of occupation low or medium (ISCO skill level 4 to 9). For more information on skill level, please see: [http://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/--publ/documents/publication/wcms\\_172572.pdf](http://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/--publ/documents/publication/wcms_172572.pdf) (p. 12). For more information on occupation, please see: <http://www.ilo.org/public/english/bureau/stat/isco/index.htm>.

#### ► *Special cases*

**New Zealand, Poland and Turkey** – As no information on occupations is available, it is not possible to determine skill level or overqualification of individuals in these countries.

### *1.1.10. Unknowns*

**Country of birth** – For most countries, the share of the population for whom the country of birth is completely undetermined is very small. For the OECD area as a whole, the share of people aged 15+<sup>1</sup> with an unknown place of birth is about 0.3%. The share of those with unknown country of birth is highest in Australia (7%). In nine countries, this share is below 1%, and for the remaining 25 countries it is negligible (less than 0.05%).

**Age** – The share of persons with unknown age is zero or negligible in all countries, except for New Zealand (0.07%), Mexico (0.08%) and Japan (1.14%).

**Educational attainment** – Most countries do not collect education information for those aged under 15, and thus, education of those aged under 15 is coded unknown. The share of persons that are 15 or older with unknown education is on average 2% in the OECD, but varies across countries. This share is highest in Australia (13%), Latvia (12%) and Luxembourg (11%). In 13 other countries it is higher than 1% and in 18 countries it is negligible.

**Duration of stay** – The share of migrants with unknown duration of stay in the destination country is 3% on average in the OECD but varies across countries. In 18 countries it remains below 1% but in some countries, namely Canada, Chile, Denmark, France, Hungary, Iceland, Ireland and Latvia, the share of migrants with unknown duration of residence is over 5%.

**Labour Force Status** – The share of persons with unknown labour force status is 1.2% across OECD countries in the database. This share is over 5% in Australia, the Czech Republic, Japan, Latvia, Norway and the Slovak Republic.

**Occupation** – The share of persons with unknown occupation is 1.2% in the OECD. This share is over 5% in the Czech Republic, Denmark, Ireland, Israel, Norway, the Slovak Republic, and Sweden.

---

<sup>1</sup> Country of birth is not recorded for individuals under 15 years of age in Israel. Including them would increase the overall share with unknown country of birth to 0.5%.

## 1.2. Coverage

The database contains information on several demographic and labour market characteristics of the population by country of birth in all 35 OECD countries around 2015/16.<sup>2</sup> The OECD countries included are Australia, Austria, Belgium, Canada, Chile, the Czech Republic, Denmark, Estonia, Finland, France (includes Overseas France), Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States.

All of the thematic files of the database include three core variables: the country of residence, the country of birth and educational attainment. Other variables available in the database include age, gender, citizenship, duration of stay, labour force status and occupation.

File A covers all persons, whereas File B and File C cover all individuals aged 15 and above. File D covers all employed individuals aged 15 and above.

### ► *Special cases*

**File A** – The reference population for Switzerland, Luxembourg and New Zealand is 15+. There is no information on those who are under 15. Caution should be used while making country comparisons.

**File B** –

**File C** – Labour force status is recorded only for those aged 16 and above for Spain and the United States. Information on labour force status is available for persons aged 15-74 years for Latvia, and for persons aged 15-64 years for Austria and Turkey. For Finland, if minors are not classified as unemployed, they are classified as inactive. Since 2005, the register data used to produce these tables has not covered minors' employment (persons younger than 18 years).

**File D** – For Latvia, information on occupation is available for persons aged 15-74 years.

## 1.3. Sources

The sources for this database are census data from the 2015/16 round of censuses, population registers for the 2015/16 period and representative surveys for the 2015/16 period. For countries that were unable to provide data, multiple sources were used (see the Special Cases in the “Country of birth” section). Table A.2 lists the sources for each country.

---

<sup>2</sup> Lithuania became the 36th country to join the OECD in 2018. Because it was not part of the organisation in 2015/16, it is not included in DIOC 2015/16.

---

## 1.4. Methods

### *1.4.1. Reliability thresholds and confidentiality issues*

Due to concerns about confidentiality and reliability, the data should meet certain standards to be published, in particular when the data source is a survey. To this end, for each File (A, B, C, D), two different tables are published:

- One table where the country of birth variable is aggregated at the regional level, but which includes the country of destination (Files A\_1, B\_1, C\_1 and D\_1)
- One table where the country of destination is omitted, but which includes the country of birth (Files A\_2, B\_2, C\_2 and D\_2); these tables therefore provide figures for the OECD as a whole as a destination.

In addition, only the less detailed versions of the following variables are included in the public files: age, education, nationality, duration of stay and occupation.

Some countries that provided census or register data have applied a random rounding procedure intended to prevent the disclosure of individual information. This procedure consists of randomly rounding each data cell to the closest upper or lower multiple of 3 or 5.

## Annex A. Appendix

**Table A.1. List of countries and regions of birth represented in the database**

Regional groups		Individual countries (cont.)		Individual countries (cont.)	
AFRI	Africa	CSFR	Former Czechoslovakia	IRQ	Iraq
ASIA	Asia	CSFR-CZE	Czech Republic	ISL	Iceland
EURO	Europe	CSFR-SVK	Slovak Republic	ISR	Israel
NOAM	North America	CUB	Cuba	ITA	Italy
OCEA	Oceania	CYM	Cayman Islands	JAM	Jamaica
SCAC	South and Central America and the Caribbean	CYP	Cyprus <sup>1</sup>	JOR	Jordan
		DEU	Germany	JPN	Japan
		DJI	Djibouti	KEN	Kenya
		DMA	Dominica	KHM	Cambodia
		DNK	Denmark	KIR	Kiribati
		DOM	Dominican Republic	KNA	Saint Kitts and Nevis
		DZA	Algeria	KOREA-NS	Korea
		ECU	Ecuador	KWT	Kuwait
		EGY	Egypt	LAO	Lao People's Democratic Republic
		ERI	Eritrea	LBN	Lebanon
		ESH	Western Sahara	LBR	Liberia
		ESP	Spain	LBY	Libya
		ETH	Ethiopia	LCA	Saint Lucia
		FIN	Finland	LIE	Liechtenstein
		FJI	Fiji	LKA	Sri Lanka
		FLK	Falkland Islands	LSO	Lesotho
		FRA	France	LUX	Luxembourg
		FSM	Micronesia	MAR	Morocco
		FYUG	Former Yugoslavia	MCO	Monaco
		FYUG-BIH	Bosnia and Herzegovina	MDG	Madagascar
		FYUG-HRV	Croatia	MDV	Maldives
		FYUG-MKD	North Macedonia	MEX	Mexico
		FYUG-MNE	Montenegro	MHL	Marshall Islands
		FYUG-SRB	Serbia	MLI	Mali
		FYUG-SVN	Slovenia	MLT	Malta
		FYUG-XXK	Kosovo <sup>2</sup>	MMR	Myanmar
		GAB	Gabon	MNG	Mongolia
		GBR	United Kingdom	MNP	Northern Mariana Islands
		GHA	Ghana	MOZ	Mozambique
		GIB	Gibraltar	MRT	Mauritania
		GIN	Guinea	MSR	Montserrat
		GMB	Gambia	MUS	Mauritius
		GNB	Guinea-Bissau	MWI	Malawi
		GNQ	Equatorial Guinea	MYS	Malaysia
		GRC	Greece	NAM	Namibia
		GRD	Grenada	NER	Niger
		GTM	Guatemala	NGA	Nigeria
		GUM	Guam	NIC	Nicaragua
		GUY	Guyana	NIU	Niue
		HKG	Hong Kong (China)	NLD	Netherlands
		HND	Honduras	NOR	Norway
		HTI	Haiti	NPL	Nepal
		HUN	Hungary	NZL	New Zealand
		IDN	Indonesia	OMN	Oman
		IND	India	PAK	Pakistan
		IRL	Ireland	PAN	Panama
		IRN	Iran	PER	Peru



2: The designation of Kosovo is without prejudice to positions on status, and is in line with United Nations Security Council Resolution 1244/99 and the Advisory Opinion of the International Court of Justice on Kosovo's declaration of independence.



Table A.2. Variables included in the database and detailed sources by country

File reference	A	B	C	D
Main theme	Nationality and Age	Duration of stay	Labour force status	Occupation
<b>Variables included</b>	- Country of residence - Country of birth - Education - Sex - Age - Nationality	- Country of residence - Country of birth - Education - Sex - Duration of stay	- Country of residence - Country of birth - Education - Sex - Age - Labour force status	- Country of residence - Country of birth - Education - Sex - Occupation - Skill level - Overqualification
<b>Reference population</b>	0+	15+	15+	15+ employed
<b>Australia</b>	Census 2016	Census 2016	Census 2016	Census 2016
<b>Austria</b>	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)
<b>Belgium</b>	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)
<b>Canada</b>	Census 2016	Census 2016	Census 2016	Census 2016
<b>Switzerland</b>	Structural Survey 2016	Structural Survey 2016	Structural Survey 2016	Structural Survey 2016
<b>Chile</b>	Labour Force Survey 2015/16	Labour Force Survey 2015/16	Labour Force Survey 2015/16	Labour Force Survey 2015/16
<b>Czech Republic</b>	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)
<b>Germany</b>	Microcensus 2016	Microcensus 2016	Microcensus 2016	Microcensus 2016
<b>Denmark</b>	Population register 2016	Population register 2016	Population register 2016	Population register 2016
<b>Spain</b>	Labour Force Survey 2016	Labour Force Survey 2016	Labour Force Survey 2016	Labour Force Survey 2016
<b>Estonia</b>	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)
<b>Finland</b>	Population registers 2015	Population registers 2015	Population registers 2015	Population registers 2015
<b>France</b>	Census 2015	Census 2015	Census 2015	Census 2015
<b>United Kingdom</b>	Labour Force Survey 2016	Labour Force Survey 2016	Labour Force Survey 2016	Labour Force Survey 2016
<b>Greece</b>	DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)
<b>Hungary</b>	DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)
<b>Ireland</b>	Census 2016	Census 2016	Census 2016	Census 2016
<b>Iceland</b>	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)
<b>Israel</b>	Labour Force Survey 2016	Labour Force Survey 2016	Labour Force Survey 2016	Labour Force Survey 2016
<b>Italy</b>	Labour Force Survey 2015/16	Labour Force Survey 2015/16	Labour Force Survey 2015/16	Labour Force Survey 2015/16
<b>Japan</b>	Population Census 2015	Population Census 2015	Population Census 2015	Population Census 2015
<b>Korea</b>	2015 Population and Housing Census (20% Sample survey)	2015 Population and Housing Census (20% Sample survey)	2015 Population and Housing Census (20% Sample survey)	2015 Population and Housing Census (20% Sample survey)
<b>Luxembourg</b>	Labour Force Survey 2016	Labour Force Survey 2016	Labour Force Survey 2016	Labour Force Survey 2016
<b>Latvia</b>	Labour Force Survey 2016	Labour Force Survey 2016	Labour Force Survey 2016	Labour Force Survey 2016
<b>Mexico</b>	Intercensal Survey 2015	Intercensal Survey 2015	Intercensal Survey 2015	Intercensal Survey 2015
<b>Netherlands</b>	Labour Force Survey 2016	Labour Force Survey 2016	Labour Force Survey 2016	Labour Force Survey 2016
<b>Norway</b>	Population register 2016	Population register 2016	Population register 2016	Population register 2016
<b>New Zealand</b>	Household Labour Force Survey 2016	Household Labour Force Survey 2016	Household Labour Force Survey 2016	Household Labour Force Survey 2016
<b>Poland</b>	DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)
<b>Portugal</b>	Labour Force Survey 2016	Labour Force Survey 2016	Labour Force Survey 2016	Labour Force Survey 2016
<b>Slovak Republic</b>	DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)
<b>Slovenia</b>	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)	European Labour Force Survey 2016 (EU LFS), DIOC 2010/11 and Eurostat's Unified Demographic Data Collection 2015-2016 (UNIDEMO)
<b>Sweden</b>	Population register 2016	Population register 2016	Population register 2016	Population register 2016
<b>Turkey</b>	Labour Force Survey 2015, DIOC 2010/11 and Turkstat's Address Based Population Registration System Statistics 2015 (ABPRSS)	Labour Force Survey 2015, DIOC 2010/11 and Turkstat's Address Based Population Registration System Statistics 2015 (ABPRSS)	Labour Force Survey 2015, DIOC 2010/11 and Turkstat's Address Based Population Registration System Statistics 2015 (ABPRSS)	Labour Force Survey 2015, DIOC 2010/11 and Turkstat's Address Based Population Registration System Statistics 2015 (ABPRSS)
<b>United States</b>	American Community Survey 2016	American Community Survey 2016	American Community Survey 2016	American Community Survey 2016

**Table A.3. ISCO-08 Codes (major-group level)**

Major group	Occupation
1	Managers
2	Professionals
3	Technicians and Associate Professionals
4	Clerical Support Workers
5	Services and Sales Workers
6	Skilled Agricultural, Forestry and Fishery Workers
7	Craft and Related Trades Workers
8	Plant and Machine Operators and Assemblers
9	Elementary Occupations
0	Armed Forces Occupations

**Table A.4. Occupation codes in Population Census of Japan**

Category code	Occupation
A	Administrative And Managerial Workers
B	Professional And Engineering Workers
C	Clerical Workers
D	Sales Workers
E	Service Workers
F	Security Workers
G	Agriculture, Forestry And Fishery Workers
H	Manufacturing Process Workers
I	Transport And Machine Operation Workers
J	Construction And Mining Workers
K	Carrying, Cleaning, Packaging, And Related Workers
L	Workers Not Classifiable By Occupation

**Table A.5. Major groups of the Standard Occupational Classification (SOC) System, United States**

<b>Code</b>	<b>Major Group</b>
11	Management Occupations
13	Business and Financial Operations Occupations
15	Computer and mathematical occupations
17	Architecture and Engineering Occupations
19	Life, Physical, and Social Science Occupations
21	Community and Social Service Occupations
23	Legal Occupations
25	Education, Training, and Library Occupations
27	Arts, Design, Entertainment, Sports, and Media Occupations
29	Healthcare Practitioners and Technical Occupations
31	Healthcare Support Occupations
33	Protective Service Occupations
35	Food Preparation and Serving Related Occupations
37	Building and Grounds Cleaning and Maintenance Occupations
39	Personal Care and Service Occupations
41	Sales and Related Occupations
43	Office and Administrative Support Occupations
45	Farming, Fishing, and Forestry Occupations
47	Construction and Extraction Occupations
49	Installation, Maintenance, and Repair Occupations
51	Production Occupations
53	Transportation and Material Moving Occupations
55	Military Specific Occupations