



Michaela Potančoková, Jason Sadler, Jakub Bijak, Miguel González-Leonardo and Akira Soto-Nishimura

Documentation of QuantMig Web Tools for Data and Simulations

Deliverable 10.4



QuantMig has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 870299 Website: www.quantmig.eu

History of changes

Version	Date	Changes
1.0	20.07.2023	Issued for Consortium Review
1.1	31.07.2023	First version submitted as official deliverable to the EC

Suggested citation

Potaňoková, M., Sadler, J., Bijak, J., González-Leonardo, M., Soto-Nishimura, A. (2023) Documentation of QuantMig Web tools for Data and Simulations. QuantMig Project Deliverable D10.4. Laxenburg: International Institute for Applied Systems Analysis.

Technical coordination: Daniela Vono de Vilhena, Max-Planck-Gesellschaft

Dissemination level

PU Public

Acknowledgments

We are grateful for the comments and recommendations made by colleagues from the QuantMig team who kindly tested the tools. We are also grateful to developers from the Southampton GeoData Institute team for their dedicated work and smooth collaboration. This document reflects the authors' view and the Research Executive Agency of the European Commission are not responsible for any use that may be made of the information it contains.

Cover photo: [iStockphoto.com/Guenter Guni](https://www.istockphoto.com/GuenterGuni).

1 Table of Contents

	Acknowledgments	i
2	Introduction	3
3	Design and documentation for the QuantMig Migration Estimates Explorer web tool.....	4
3.1	Description of the tool	4
3.2	Technical Implementation	8
4	Design and documentation for the QuantMig Migration Scenarios Explorer web tool.....	9
4.1	Page 1: Introduction and brief description of the tool	9
4.2	Page 2: Migration scenarios	9
4.3	Page 3. Population pyramids.....	15
4.4	Page 4: Indicators	17
4.5	Page 5: Maps	21
4.6	Technical Implementation	23
5	Data Inventories (Migration Data, Migration Drivers, Migration Policies).....	24
5.1	Description of Inventories.....	24
5.1.1	QuantMig Migration Data Inventory	24
5.1.2	QuantMig Migration Driver Inventory	25
5.1.3	DEMIG-QuantMig Migration Policy Database	25
5.2	Technical Implementation	25
	References.....	26
	Appendix I – SQL Database table definitions.....	28
	QuantMig Migration Estimates Explorer Tool.....	28
	QuantMig Migration Scenarios Explorer	30
	Data Inventories (Migration Data, Migration Drivers, Migration Policies).....	32

2 Introduction

This report contains the accompanying documentation to the QuantMig data inventories and QuantMig web tools for user-friendly and interactive presentation of migration estimates and scenarios produced within the project. Two web tools and three data inventories have been developed within the project. This technical report documents their design, functionalities, and technical implementation. All inventories and both web tools are open source. Inventories and web tools were developed in collaboration with and technically implemented by Southampton GeoData Institute (www.geodata.soton.ac.uk) and are accessible in the Data and Estimates section at the QuantMig project web site: https://quantmig.eu/data_and_estimates/

Web tools give access to the data produced by the project. Two separate web tools were developed to facilitate open and user-friendly access to (1) Bayesian estimates of origin-destination migration flows of Aristotelous et al. (2022) and to (2) microsimulations of future population change and population heterogeneity in 31 European countries by alternative migration scenarios (Marois et al. (2023) and Potančoková et al. 2023a). A specific web tool has been designed for each of these outputs.

The QuantMig Migration Estimates Explorer allows the user to explore and download migration flows estimates that integrate the ones previously developed as a part of the Integrated Modelling of European Migration (IMEM) project until 2009 (Raymer et al. 2013), with those produced within QuantMig until 2019. The methodology of the QuantMig migration estimates is available from Aristotelous et al. (2022). Section 3 of this report documents the design of the web tool and the functionalities for exploring immigration, emigration and net migration flows by origin and destination. The tool can be accessed using the following link: http://quantmig.eu/data_and_estimates/estimates_explorer/

The QuantMig Migration Scenarios Explorer allows the user to explore results of simulation of population heterogeneity in 31 European countries from 2020 until 2060 in 15 alternative migration scenarios: a baseline scenario envisaging continuation of historical trends, seven scenarios to evaluate the demographic impacts of short high-immigration events from seven specific regions of origin, and seven high and persistent immigration events from the same seven regions of origin. The methodology of the microsimulations is available from Marois et al. (2023). The web toll gives access to granular results from 15 out of 29 simulated scenarios of migration events into the EU+. Selected scenarios model high-migration events corresponding to the twice-in-a-century frequency of occurrence and illustrate well the diversity of demographic futures in EU+ countries. They show more pronounced demographic and labour force impacts than the 14 scenarios of migration events with once-in-a-decade frequency of occurrence which are not implemented in the web tool and serve as sensitivity analysis. Results for all scenarios and the sensitivity analysis are published in a discussion paper from Potančoková et al. (2023a). The input and output data for all 29 scenarios are available at the Zenodo repository as of August 2023, DOI: 10.5281/zenodo.7728049 (Potančoková et al. 2023b). The tool can be accessed using the following link: http://quantmig.eu/data_and_estimates/scenarios_explorer/

The QuantMig Data Inventories catalogue the plethora of migration data and provide meta-information relevant for the user with respect to scope and variables of interest. Three Data inventories were developed by the QuantMig team: Migration Data Inventory, Migration Driver Inventory and DEMIG-QuantMig Database. Section 5 details on the scope and technical implementation of the Inventories.

The purpose of the Data Inventory is not only to catalogue existing datasets on migration to, from, and within the EU but also to provide key meta-information on each dataset that may be relevant for the user. The inventory is accessible at https://quantmig.eu/data_and_estimates/data_inventory/

The Migration Driver Inventory contains information on the drivers of migration that have been considered and operationalized from 176 academic articles published between 2000 and 2019. This inventory is accessible at https://quantmig.eu/data_and_estimates/driver_inventory/

The DEMIG-QuantMig database provides an update and extension to the DEMIG POLICY database following the same template for methodology, and specifically for assessing policy changes, as in de Haas et al. (2015) and is accessible at https://quantmig.eu/data_and_estimates/policy_database/

3 Design and documentation for the QuantMig Migration Estimates Explorer web tool

3.1 Description of the tool

The QuantMig Migration Estimates Explorer is an online front-end for querying the database of migration estimates by origin, destination, age, gender, and region of birth for 2009–2019, developed within the QuantMig project (Aristotelous et al. 2022). The estimates have been prepared for flows between the 32 countries of the EU+ migration system (EU-27, EFTA, and the United Kingdom, which during the time covered by the estimates was still a part of the EU), as well as for flows into and out of North Macedonia, and the aggregated Rest of the World. Age is available in five-year groups, and the region of birth is binary: within or outside the EU+ system. In addition, for a subset of countries and characteristics (so excluding estimates for Croatia and North Macedonia, with the region of birth disaggregation not available), the estimates for 2002–2008 are also available, resulting from the IMEM project (Raymer et al. 2013). As these estimates were obtained by using a slightly different methodology, the results of the queries involving both sets of years (2002–2008 and 2009–2019) include a note about a possible discontinuity in trends.

The tool enables querying the database and generating summary graphs and tables for three main indicators: immigration, emigration and net migration, as well as preparing files for download in CSV (tables) or PNG (graphs) formats. The queries return summary statistics based on between 1,000 and 5,000 posterior samples for each given estimate, with the underlying statistical details provided in Aristotelous et al. (2022). Standard outputs include posterior medians of the estimates, as well as their 10th, 25th, 75th and 90th percentiles, so as to provide some pre-set measures of the uncertainty of estimation. From the database, either marginal distributions can be obtained (by setting Grouping Type = None), or specific distributions for individual dimensions (with Grouping set to the breakdown of interest, such as Age Group, Gender, Country, or their combinations).

The default query setup is shown in Figure 1, with some options changing depending on whether the user toggles the Region of Birth variable (which is only available at the country level, not disaggregated by age or gender), and on whether they want to include the 2002–2008 period in the query as well. In the latter case, some countries and breakdowns will become unavailable: besides the estimates for Croatia and North Macedonia not being available for the earlier period, the region of birth variable is only included for the 2009–2019 dataset. The data can be visualised

either as time series of selected indicators, or as age profiles, aggregated over a chosen period. Either way, the data can be presented as piecewise-linear or monotonically smoothed function, depending on the option selected by the user. It is worth noting that the query can involve a single country or a group of countries, so that the users can generate different combinations of migration flows: one country to/from another country, one country to/from a group (or all) other countries, and a group of (or all) countries to/from another group of (or all) countries. One important consideration is that if a user wants to generate aggregates for several countries, flows, age groups, etc., this should be done through the interface, and not by adding any values generated for individual countries, flows or groups: as all the presented outcomes are expressed in terms of posterior quantiles (medians, percentiles), it has to be borne in mind that such quantities are not additive. The queries are executed by pressing the 'Generate Chart' button, visible at the bottom of Figure 1, just underneath all options that can be selected by users to tailor the query to their specific needs.

Select Country: (Select Multiple Countries)

Austria

Query Type:
 Immigration Emigration Net Migration

Horizontal Axis:
 Year Age Group

Grouping Type:
 None

Cubic Interpolation Mode:
 Default Monotone

Advanced Options

Select Partner Country or Countries:

<input type="checkbox"/> Austria	<input checked="" type="checkbox"/> Belgium	<input checked="" type="checkbox"/> Bulgaria	<input checked="" type="checkbox"/> Croatia
<input checked="" type="checkbox"/> Cyprus	<input checked="" type="checkbox"/> Czechia	<input checked="" type="checkbox"/> Denmark	<input checked="" type="checkbox"/> Estonia
<input checked="" type="checkbox"/> Finland	<input checked="" type="checkbox"/> France	<input checked="" type="checkbox"/> Germany	<input checked="" type="checkbox"/> Greece
<input checked="" type="checkbox"/> Hungary	<input checked="" type="checkbox"/> Iceland	<input checked="" type="checkbox"/> Ireland	<input checked="" type="checkbox"/> Italy
<input checked="" type="checkbox"/> Latvia	<input checked="" type="checkbox"/> Liechtenstein	<input checked="" type="checkbox"/> Lithuania	<input checked="" type="checkbox"/> Luxembourg
<input checked="" type="checkbox"/> Malta	<input checked="" type="checkbox"/> Netherlands	<input checked="" type="checkbox"/> North Macedonia	<input checked="" type="checkbox"/> Norway
<input checked="" type="checkbox"/> Poland	<input checked="" type="checkbox"/> Portugal	<input checked="" type="checkbox"/> Rest World	<input checked="" type="checkbox"/> Romania
<input checked="" type="checkbox"/> Slovakia	<input checked="" type="checkbox"/> Slovenia	<input checked="" type="checkbox"/> Spain	<input checked="" type="checkbox"/> Sweden
<input checked="" type="checkbox"/> Switzerland	<input checked="" type="checkbox"/> United Kingdom		

Select All **Select EU15** **Invert Selection** **Unselect All**

Toggle Region Of Birth: (this will remove certain axis/options and groupings)

Query original IMEM 2002-8 data: (this will remove certain axis/options and groupings)

Select Age Groups:

<input checked="" type="checkbox"/> 0-4	<input checked="" type="checkbox"/> 5-9	<input checked="" type="checkbox"/> 10-14	<input checked="" type="checkbox"/> 15-19
<input checked="" type="checkbox"/> 20-24	<input checked="" type="checkbox"/> 25-29	<input checked="" type="checkbox"/> 30-34	<input checked="" type="checkbox"/> 35-39
<input checked="" type="checkbox"/> 40-44	<input checked="" type="checkbox"/> 45-49	<input checked="" type="checkbox"/> 50-54	<input checked="" type="checkbox"/> 55-59
<input checked="" type="checkbox"/> 60-64	<input checked="" type="checkbox"/> 65-69	<input checked="" type="checkbox"/> 70-74	<input checked="" type="checkbox"/> 75-79
<input checked="" type="checkbox"/> 80-84	<input checked="" type="checkbox"/> 85+		

Select All **Invert Selection** **Unselect All**

Select Genders:
 Female Male

Select Measures Of Uncertainty:
 10th Percentile 1st Quartile Median 3rd Quartile 90th Percentile

Select Years:

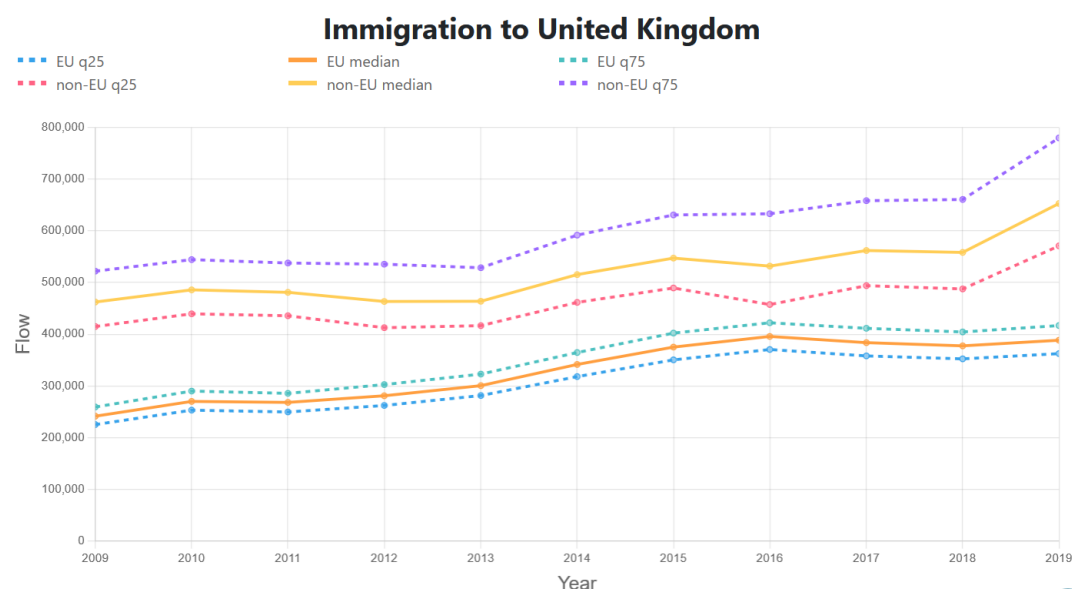
<input checked="" type="checkbox"/> 2009	<input checked="" type="checkbox"/> 2010	<input checked="" type="checkbox"/> 2011	<input checked="" type="checkbox"/> 2012
<input checked="" type="checkbox"/> 2013	<input checked="" type="checkbox"/> 2014	<input checked="" type="checkbox"/> 2015	<input checked="" type="checkbox"/> 2016
<input checked="" type="checkbox"/> 2017	<input checked="" type="checkbox"/> 2018	<input checked="" type="checkbox"/> 2019	

Select All **Invert Selection** **Unselect All**

Generate Chart

Figure 1. Basic interface of the QuantMig Migration Estimates Explorer

An example of the results of a query are shown in Figure 2, which presents immigration to the UK in the period 2009–2019 only, but grouped by region of birth (EU born versus non-EU born). The selected indicators (the medians, as well as lower and upper quartiles) are shown on the graph and provided in a tabular form underneath. As mentioned above, both the table and the graph are downloadable for further analysis and processing. In addition, the graph can be zoomed in or out along the horizontal axis by stretching out or pinching in two fingers on the touchpad, screen or a similar device (where available).



[Download as CSV](#)
[Download as PNG](#)

year	country	q25	median	q75
2009	EU	225124.69	241310.49	259145.81
2010	EU	253007.45	269877.92	289914.33
2011	EU	249485.63	267988.53	285547.76
2012	EU	262051.07	280830.22	302478.00
2013	EU	281284.84	300392.60	322774.69
2014	EU	317937.10	341413.99	364515.53
2015	EU	350340.26	375022.86	402138.62
2016	EU	370373.75	395637.26	422123.41
2017	EU	357980.02	383615.08	411402.02
2018	EU	352278.83	377418.56	404325.21
2019	EU	362423.03	388290.84	416726.99
2009	non-EU	414874.94	462249.45	521978.25

Figure 2. Sample outputs of the QuantMig Migration Estimates Explorer

Whenever the 2002–08 dataset is toggled, the outcomes of the analysis are additionally equipped with the following caveat, displayed above the graph: *“There may be discontinuity between the estimates and their uncertainty assessments between 2002–08 and 2009–19, due to different levels of data availability and varying extent of data harmonisation in those periods, and to slight changes to the estimation model. Please bear that in mind when using these estimates for modelling, especially in a time series framework”.*

3.2 Technical Implementation

The QuantMig website is delivered via a content management framework, developed by the Southampton GeoData Institute team, based on open source technologies (PHP, Apache, MySQL, PostgreSQL) and open standards for web development (XHTML, CSS, JavaScript), supplemented with widely used open source JavaScript libraries (jQuery, Bootstrap, Font Awesome). Together this delivers a clean, accessible site, with “responsive design” to support a full range of standard browsers and devices.

The QuantMig Migration Estimates Explorer tool is embedded within the main QuantMig website, making use of the JavaScript tools to deliver an intuitive user interface, extended with the open source Chart.js JavaScript graphing library to deliver interactive graphs to visualise the migration estimates.

The data for the QuantMig Migration Estimates Explorer – posterior samples of the estimates of individual flows – was imported to a PostgreSQL (v14) database, hosted within the server infrastructure used to deliver the QuantMig website, managed by the GeoData team at the University of Southampton. The database provides a suitable tool to store the large number of data values, with indexes to allow efficient query and statistical functions (Ordered-Set Aggregate Functions – discrete percentile) to deliver the user functionality required. Estimates data for 2002 to 2008 from the previous IMEM project (Raymer et al. 2013) was also incorporated into the database, so that they could be accessed seamlessly by website users alongside the 2009-2019 estimates developed within the current project.

An Asynchronous JavaScript (AJAX) method is used to pass data between the back-end server website and database tools and the front end, delivering an interactive experience for users, without the need to reload the entire page when choosing estimate variables to explore. A database query caching mechanism is used to save the results of any previous request, so that resource intensive and time-consuming database queries for a set of variable selections are run only once, then available instantly for all website users.

The SQL database table definitions are included in Appendix I.

4 Design and documentation for the QuantMig Migration Scenarios Explorer web tool

The QuantMig Migration Scenarios Explorer web tool allows the user to query the inputs necessary to understand the differences between the migration scenarios (assumptions on immigration into the EU+ countries from other world regions) and outputs of the QuantMig-Mic population projection simulations for 31 European countries for 2020-2060 in a user-friendly and visually attractive manner. The web tool consists of five Pages and the design is inspired by the Wittgenstein Centre Human Capital Data Explorer (Wittgenstein Centre for Demography and Global Human Capital 2018)¹ but adapted to the specific needs of QuantMig.

User enters the web tool through Page 1 which introduces the information necessary to comprehend QuantMig migration scenarios. Page 2 is designed to allow the users to visualise immigration assumptions used in the simulated migration scenarios of high migration events. The user can explore the results of the simulations through pages 3, 4 and 5. On Page 3, users can produce their own population pyramids and investigate population heterogeneity of populations and the labour force by age, sex, place of birth, educational attainment and labour force status. User can create two charts to facilitate comparisons across scenarios, countries and over time. Population pyramids give access to the granular microsimulation outputs. Web tool also goes beyond the raw model outputs by giving the user an access to indicators derived from raw model output data and presented at pages 4 and 5. The temporal change in the indicators across scenarios can be explored on Pages 4 through line charts. The same indicators are visualised on Page 5 in choropleth maps in order to allow comparisons across all 31 simulated countries at a glance. The selection of the indicators for the web tool allows the user to analyse the effect of the different migration scenarios on population size, population ageing, changes in population diversity with respect to human capital (educational attainments) and by place of birth. Users can also download the visuals they co-created as well as the underlying data. We next elaborate the documentation of the tool for each page.

4.1 Page 1: Introduction and brief description of the tool

The user enters the QuantMig Migration Scenarios Explorer through Page 1 which features introductory text that provides in a nutshell information about how to understand and interpret QuantMig migration scenarios. The intro page also contains links to the embedded user guide for each Page of the web tool, a short video which walks the user through the tool and the Frequently Asked Questions (FAQs).

4.2 Page 2: Migration scenarios

The second page allows the user to query and visualise the projected immigration flows (i.e. immigration assumptions used in the scenarios) from East Asia, Latin America, North Africa, Other Europe, Sub-Saharan Africa, South and South-East Asia and West Asia to the 31 European countries in the model (Figure 1). This page allows users to view and understand the migration assumptions used in the model, and become familiar with the scenarios' migration assumptions before exploring the micro simulation results in the following Pages. Flows from specific origins

¹ Version 2, available at: <http://dataexplorer.wittgensteincentre.org/wcde-v2>

can be combined in a single plot for the baseline, short or persistent high-migration event scenarios. In other words, users are allowed select specific immigration flows by origin for the three above-mentioned scenarios, and chose individual countries of destination: Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Germany, Denmark, Estonia, Spain, Finland, France, Greece, Croatia, Hungary, Ireland, Italy, Lithuania, Luxembourg, Latvia, Malta, Netherlands, Poland, Portugal, Romania, Sweden, Slovenia and Slovakia (EU27), the United Kingdom, Switzerland, Norway and Iceland (UK+EFTA) countries). Summary outputs are available as well for the EU27, UK+EFTA and EU+ (alle 31 countries included in the model).

For instance, let's suppose that the user wants to visualise immigration flows from Latin America to Spain in three scenarios. In this case, the user would select Spain from the "Destination country" drop-down menu and tick the scenarios Baseline from Latin America, Persistent-high immigration event from Latin America and Short-high immigration event from Latin America. The user executes the query using the "Generate chart" button and can also download the data from the query in CSV format and download the chart as PGN file using the buttons located below the chart.

Figure 3 illustrates the example above and provides the visual of the Page 2.

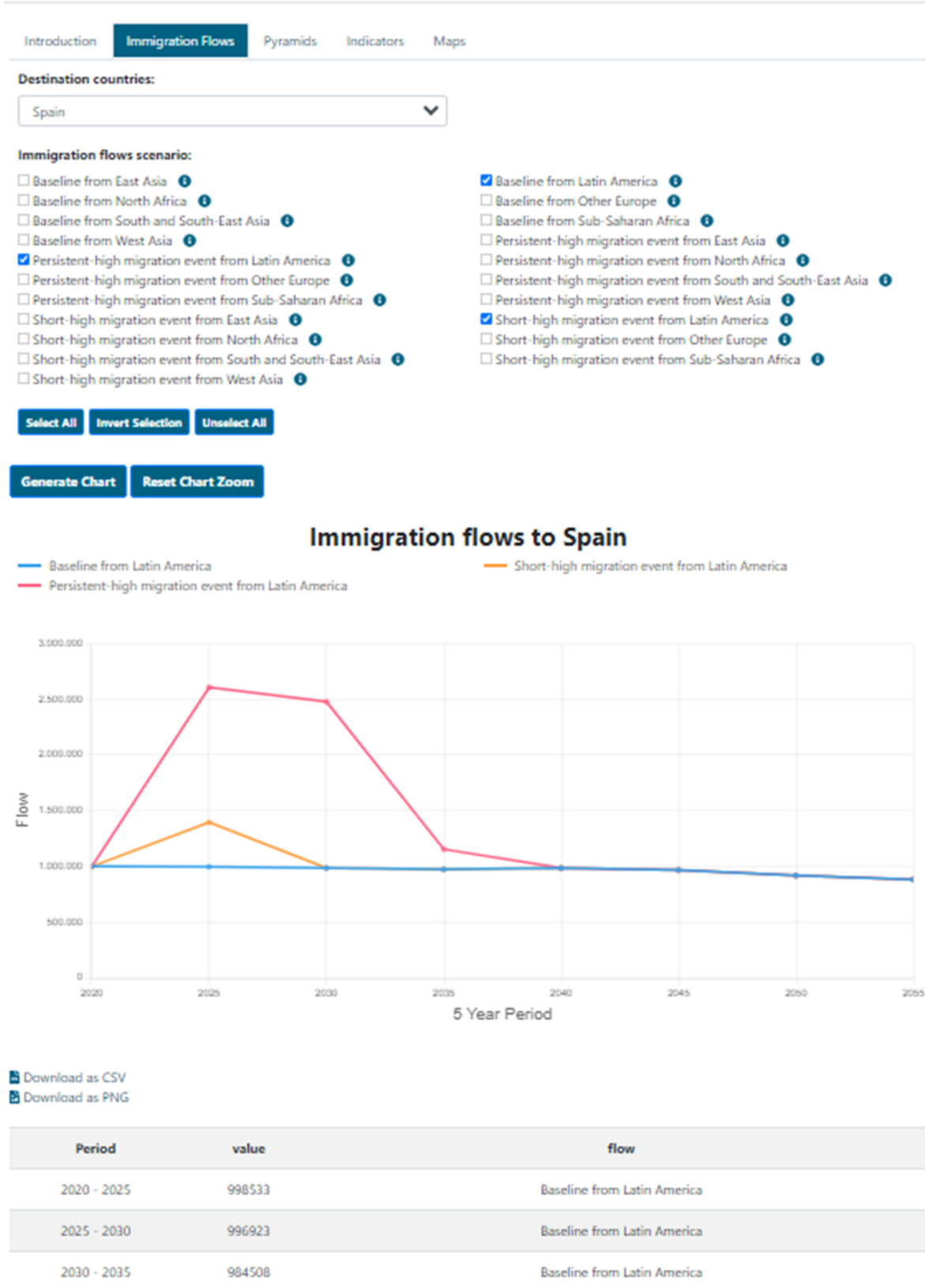


Figure 3. Page 2: Immigration flows.

Narratives for each scenario can be displayed using the “i” symbol at each selection option and also show when mousing over the query field. The descriptions of the selected scenarior are also displayed below the chart and the data table view at the bottom of the Page 2.

The scenario descriptions used in the tool are listed below.

Baseline migration scenario

Immigration from the rest of the world regions into the EU+ continues with the same intensity as in the second decade of 21st century and immigrants from specific world regions will be attracted mainly towards those EU+ countries where compatriots from that given region have migrated to in the past and where the existing migration networks can support them. The actual immigration volumes from the world regions into EU+ countries are derived by applying the average emigration rate from a region A into the EU+ during 2011-2019 onto projected young population in each world region. We assume that the war against Ukraine will continue beyond 2023 and the combat will cease by 2025, resulting in return of 60% of the refugees back to Ukraine. Emigration of native-born, EU+born and persons born outside the EU+ varies with respect to the immigration scenario and is modelled through average emigration rates by country of residence and place of birth. For more details see (Potančoková et al. 2023).

Short-high migration event from Other Europe

An immigration event from Other Europe into the EU+ countries with the frequency of occurrence twice-in-a-century (taking the modelled immigration corresponding to 98th quantile of Pareto distribution, Bijak 2023), and this event takes place for one calendar year within the 5-year period 2025-2029. Immigration from all other world regions follows the baseline scenario. Before and after the extreme event immigration from Other Europe returns to the levels of the baseline scenario. Such an event can be an outcome of humanitarian or natural disasters with temporary migration and high probabilities of return which can be resulting from a speedy policy reaction to the crisis that provoked high immigration. For more details see Potančoková et al 2023.

Short-high migration event from North Africa

An immigration event from North Africa into the EU+ countries with the frequency of occurrence twice-in-a-century (taking the modelled immigration corresponding to 98th quantile of Pareto distribution, Bijak 2023), and this event takes place for one calendar year within the 5-year period 2025-2029. Immigration from all other world regions follows the baseline scenario. Before and after the extreme event immigration from North Africa returns to the levels of the baseline scenario. Such an event can be an outcome of humanitarian or natural disasters with temporary migration and high probabilities of return which can be resulting from a speedy policy reaction to the crisis that provoked high immigration. For more details see Potančoková et al 2023.

Short-high migration event from Sub-Saharan Africa

An immigration event from Sub-Saharan Africa into the EU+ countries with the frequency of occurrence twice-in-a-century (taking the modelled immigration corresponding to 98th quantile of Pareto distribution, Bijak 2023), and this event takes place for one calendar year within the 5-year period 2025-2029. Immigration from all other world regions follows the baseline scenario. Before and after the extreme event immigration from Sub-Saharan Africa returns to the levels of the baseline scenario. Such an event can be an outcome of humanitarian or natural disasters with temporary migration and high probabilities of return which can be resulting from a speedy policy reaction to the crisis that provoked high immigration. For more details see Potančoková et al 2023.

Short-high migration event from Latin America

An immigration event from Latin America into the EU+ countries with the frequency of occurrence twice-in-a-century (taking the modelled immigration corresponding to 98th quantile of Pareto distribution, Bijak 2023), and this event takes place for one calendar year within the 5-year period 2025-2029. Immigration from all other world regions follows the baseline scenario. Before and after the extreme event immigration from Latin America returns to the levels of the baseline scenario.

Such an event can be an outcome of humanitarian or natural disasters with temporary migration and high probabilities of return which can be resulting from a speedy policy reaction to the crisis that provoked high immigration. For more details see Potančoková et al 2023.

Short-high migration event from West Asia

An immigration event from West Asia into the EU+ countries with the frequency of occurrence twice-in-a-century (taking the modelled immigration corresponding to 98th quantile of Pareto distribution, Bijak 2023), and this event takes place for one calendar year within the 5-year period 2025-2029. Immigration from all other world regions follows the baseline scenario. Before and after the extreme event immigration from West Asia returns to the levels of the baseline scenario. Such an event can be an outcome of humanitarian or natural disasters with temporary migration and high probabilities of return which can be resulting from a speedy policy reaction to the crisis that provoked high immigration. For more details see Potančoková et al 2023.

Short-high migration event from South & South-East Asia

An immigration event from South & South-East Asia into the EU+ countries with the frequency of occurrence twice-in-a-century (taking the modelled immigration corresponding to 98th quantile of Pareto distribution, Bijak 2023), and this event takes place for one calendar year within the 5-year period 2025-2029. Immigration from all other world regions follows the baseline scenario. Before and after the extreme event immigration from South & South-East Asia returns to the levels of the baseline scenario. Such an event can be an outcome of humanitarian or natural disasters with temporary migration and high probabilities of return which can be resulting from a speedy policy reaction to the crisis that provoked high immigration. For more details see Potančoková et al 2023.

Short-high migration event from East Asia

An immigration event from East Asia into the EU+ countries with the frequency of occurrence twice-in-a-century (taking the modelled immigration corresponding to 98th quantile of Pareto distribution, Bijak 2023), and this event takes place for one calendar year within the 5-year period 2025-2029. Immigration from all other world regions follows the baseline scenario. Before and after the extreme event immigration from East Asia returns to the levels of the baseline scenario. Such an event can be an outcome of humanitarian or natural disasters with temporary migration and high probabilities of return which can be resulting from a speedy policy reaction to the crisis that provoked high immigration. For more details see Potančoková et al 2023.

Persistent-high migration event from Other Europe

The initial short-high migration event from Other Europe – which corresponds to the frequency of occurrence twice-in-a-century – is followed by a gradually declining immigration from that region for another decade. After a decade immigration from Other Europe continues as in the baseline scenario. The persistence in migration is envisaged because of the initial event's effect on establishment of migration networks, subsequent family reunifications and chain migration, as well as due to persistence of migration drivers stimulating emigration from the origin countries. Elevated migration flows thus take place between 2027 and 2036. Immigration from all other world regions follows the baseline scenario. For more details see Potančoková et al 2023.

Persistent-high migration event from North Africa

The initial short-high migration event from North Africa – which corresponds to the frequency of occurrence twice-in-a-century – is followed by a gradually declining immigration from that region for another decade. After a decade immigration from North Africa continues as in the baseline scenario. The persistence in migration is envisaged because of the initial event's effect on establishment of migration networks, subsequent family reunifications and chain migration, as

well as due to persistence of migration drivers stimulating emigration from the origin countries. Elevated migration flows thus take place between 2027 and 2036. Immigration from all other world regions follows the baseline scenario. For more details see Potančoková et al 2023.

Persistent-high migration event from Sub-Saharan Africa

The initial short-high migration event from Sub-Saharan Africa – which corresponds to the frequency of occurrence twice-in-a-century – is followed by a gradually declining immigration from that region for another decade. After a decade immigration from Sub-Saharan Africa continues as in the baseline scenario. The persistence in migration is envisaged because of the initial event's effect on establishment of migration networks, subsequent family reunifications and chain migration, as well as due to persistence of migration drivers stimulating emigration from the origin countries. Elevated migration flows thus take place between 2027 and 2036. Immigration from all other world regions follows the baseline scenario. For more details see Potančoková et al 2023.

Persistent-high migration event from Latin America

The initial short-high migration event from Latin America – which corresponds to the frequency of occurrence twice-in-a-century – is followed by a gradually declining immigration from that region for another decade. After a decade immigration from Latin America continues as in the baseline scenario. The persistence in migration is envisaged because of the initial event's effect on establishment of migration networks, subsequent family reunifications and chain migration, as well as due to persistence of migration drivers stimulating emigration from the origin countries. Elevated migration flows thus take place between 2027 and 2036. Immigration from all other world regions follows the baseline scenario. For more details see Potančoková et al 2023.

Persistent-high migration event from West Asia

The initial short-high migration event from West Asia – which corresponds to the frequency of occurrence twice-in-a-century – is followed by a gradually declining immigration from that region for another decade. After a decade immigration from West Asia continues as in the baseline scenario. The persistence in migration is envisaged because of the initial event's effect on establishment of migration networks, subsequent family reunifications and chain migration, as well as due to persistence of migration drivers stimulating emigration from the origin countries. Elevated migration flows thus take place between 2027 and 2036. Immigration from all other world regions follows the baseline scenario. For more details see Potančoková et al 2023.

Persistent-high migration event from South & South-East Asia

The initial short-high migration event from South & South-East Asia – which corresponds to the frequency of occurrence twice-in-a-century – is followed by a gradually declining immigration from that region for another decade. After a decade immigration from South & South-East Asia continues as in the baseline scenario. The persistence in migration is envisaged because of the initial event's effect on establishment of migration networks, subsequent family reunifications and chain migration, as well as due to persistence of migration drivers stimulating emigration from the origin countries. Elevated migration flows thus take place between 2027 and 2036. Immigration from all other world regions follows the baseline scenario. For more details see Potančoková et al 2023.

Persistent-high migration event from East Asia

The initial short-high migration event from East Asia – which corresponds to the frequency of occurrence twice-in-a-century – is followed by a gradually declining immigration from that region for another decade. After a decade immigration from East Asia continues as in the baseline

scenario. The persistence in migration is envisaged because of the initial event's effect on establishment of migration networks, subsequent family reunifications and chain migration, as well as due to persistence of migration drivers stimulating emigration from the origin countries. Elevated migration flows thus take place between 2027 and 2036. Immigration from all other world regions follows the baseline scenario. For more details see Potančoková et al 2023.

4.3 Page 3. Population pyramids

In the third page, the user can query granular microsimulation outputs and explore the projected demographic diversity of European populations for the period 2020-2060 according to the different migration scenarios. For that purpose, we use population pyramids which allow users to visualise population stocks by multiple age, sex and up to three additional characteristics. This is achieved by allowing the user to choose the main, grouping variable and filtering additional two variables in Advanced options (Figure 4).

Users can select individual destination countries within EU+ states, as well as EU+, EU27 and UK+EFTA from the drop-down menu and customise immigration flows from the different place of birth regions. Using Grouping variable users can visualise population pyramid by no additional variable (None option in the drop-down menu, returns simple pyramid by age and sex), by education (below secondary, secondary and post-secondary), labour force status (active or inactive in the labour force) or place of birth (native-born, foreign-born, EU+ born and born in each of the seven world regions or origin outside EU+ as described in the previous section + North America and Australia). Generate chart button generates the plot for the query, the legend can be seen after ticking the box with that option.

Advanced options allow the user to add additional variables and construct population pyramids for specific population groups by filtering the variables. The three dimensions – education, labour status and place of birth can be used as a filters. Therefore, users can filter/select a specific group of education or region of birth in the charts.

To allow comparisons between the countries, scenarios or over time, two panels with population pyramids are included (left and right). To facilitate comparison between both panels, we included an option to harmonise the X-axis in both plots. Users can download the generated plots in PNG format and can also download the underlying data in a CSV file using the option displayed below the charts. The underlying data are also visible in the table view placed below the charts.

Let's explain the functioning of this page using an example. Let's suppose that a user wants to analyse the impact of a high and persistent immigration event from East Asia on the population composition by place of birth for individuals with post-secondary education in France and compared this output with the baseline immigration scenario in the same country, based on observed migration trend from 2011-2019. In this case, the user would France and the destination country, the baseline scenario within immigration flows scenario, select place of birth as a variable, move educational attainment as a filter, and filter populations with post-secondary education. Then the user would use the option "Copy input fields across" to replicate the same population pyramid at right but selecting the "High and persistent immigration event from East Asia" within immigration flows scenarios in the population pyramid at right, see Figure 4 below.

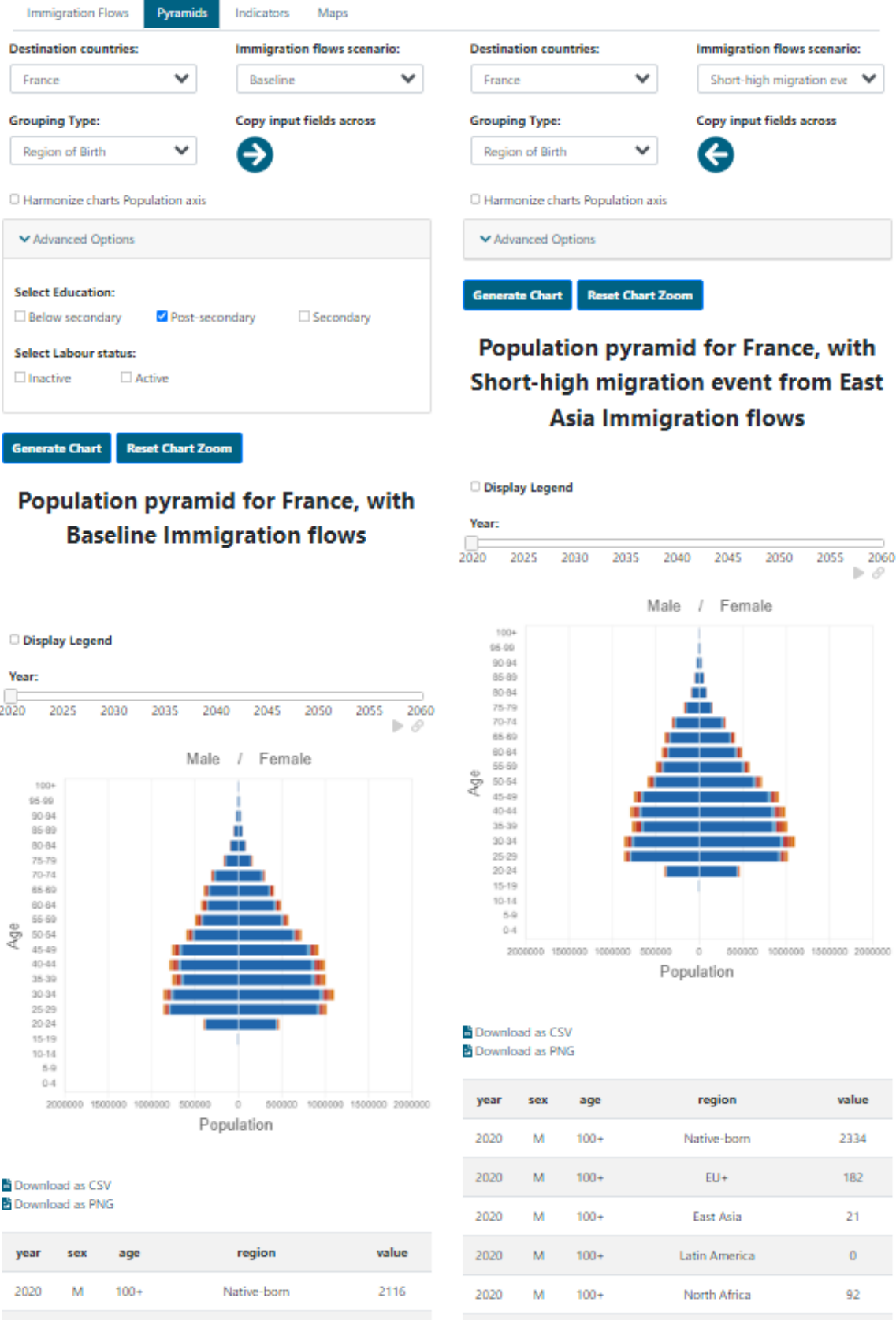


Figure 4. Page 3: Population pyramids.

4.4 Page 4: Indicators

Besides the results from the model outputs, the web tool gives the users the option to explore analytical indicators derived from the raw results. These are visualised in line charts on Page 4 and as choropleth maps on Page 5. The indicators were selected for the assessment of the demographic and labour force impacts of different migration scenarios over time. Users can visualise effects on population sizes and composition and labour force and also plot the total immigration, emigration and net migration flows (i.e. not only assumptions as for immigration into the EU+ shown in the Page 1 but model outputs that include migration between the 31 EU+ countries).

Indicators are available for total population of the destination country and in addition to that also for population of all Foreign-born and for population of persons born outside the EU+ (first generation immigrants). The indicators are available for each of 31 destination countries and for EU+, EU27 and UK+EFTA (Figure 5). The categories included in the chart, which are represented using lines, correspond to the different immigration flows scenarios. Users can select and deselect as many scenarios as they want to visualise. This means to add or remove scenarios/lines in the chart.

Let's briefly explain the functioning of this page using an example. Let's think that a user wants to visualise the effect of different immigration scenarios on the foreign-born working age population (aged 15 to 64) size in 2060 in Austria and Denmark. In this case, the user would select the foreign-born in the variable place of birth, the indicator Working-age population (15-64) and the specific immigration flows scenarios he is interested in. The user would use the "Copy input fields across" bottom and select Austria in the left chart and Denmark in the right one.

Immigration Flows
Pyramids
Indicators
Maps

Place of birth:
Foreign-born

Indicator:
Working age population

Destination countries:
Austria

Copy input fields across
➔

Immigration flows scenario:

- Baseline
- Persistent-high migration event from East Asia
- Persistent-high migration event from Latin America
- Persistent-high migration event from North Africa
- Persistent-high migration event from other Europe
- Persistent-high migration event from South and South-East Asia
- Persistent-high migration event from Sub-Saharan Africa
- Persistent-high migration event from West Asia
- Short-high migration event from East Asia
- Short-high migration event from Latin America
- Short-high migration event from North Africa
- Short-high migration event from other Europe
- Short-high migration event from South and South-East Asia
- Short-high migration event from Sub-Saharan Africa
- Short-high migration event from West Asia

Select All Invert Selection Unselect All

Generate Chart Reset Chart Zoom

Working age population (15-64), Foreign-born, Austria

Display Legend

- Baseline
- Persistent-high migration event from East Asia
- Persistent-high migration event from other Europe

- Persistent-high migration event from South and South-East Asia

Download as CSV
Download as PNG

year	value	flow
2020	1673841	Baseline
2025	1963154	Baseline

Place of birth:
Foreign-born

Indicator:
Working age population

Destination countries:
Denmark

Copy input fields across
➔

Immigration flows scenario:

- Baseline
- Persistent-high migration event from East Asia
- Persistent-high migration event from Latin America
- Persistent-high migration event from North Africa
- Persistent-high migration event from other Europe
- Persistent-high migration event from South and South-East Asia
- Persistent-high migration event from Sub-Saharan Africa
- Persistent-high migration event from West Asia
- Short-high migration event from East Asia
- Short-high migration event from Latin America
- Short-high migration event from North Africa
- Short-high migration event from other Europe
- Short-high migration event from South and South-East Asia
- Short-high migration event from Sub-Saharan Africa
- Short-high migration event from West Asia

Select All Invert Selection Unselect All

Generate Chart Reset Chart Zoom

Working age population (15-64), Foreign-born, Denmark

Display Legend

- Baseline
- Persistent-high migration event from East Asia
- Persistent-high migration event from other Europe

- Persistent-high migration event from South and South-East Asia

Download as CSV
Download as PNG

year	value	flow
2020	668076	Baseline
2025	795859	Baseline

Figure 5. Design for Page 4: Indicators.

Table 1: Listing of indicators by type of indicator and place of birth

Type	Variable	Place of birth		
		Total	Foreign-born	Born outside EU+
Population composition	Population	X	X	X
	Native-born population	X		
	% native-born	X		
	% foreign-born	X		
	% born outside EU+	X		
	Mean age	X	X	X
	Age dependency ratio	X		
Labour force	Working age population (15-64)	X	X	X
	Native-born working age population (15-64)	X		
	Labour force	X	X	X
	% of active population	X	X	X
	% foreign-born within the labour force	X		
	% born outside EU+ within the labour force	X		
	Inactive population	X	X	X
	% inactive at working age population	X	X	X
	% foreign-born within inactive population	X		
	% born outside EU+ within inactive population	X		
	Labour force dependency ratio	X		
Education	% below secondary educated (age 25-39)	X	X	X
	% postsecondary educated population (age 25-39)	X	X	X
Gender	Female inactive population	X	X	X
	% inactive female at working age population	X	X	X
	% female within inactive population	X		
	% female foreign-born within inactive population	X		
	% female born outside EU+ within inactive population	X		
Migration	Immigration outside EU+	X		
	Immigration within EU+	X		
	Emigration outside EU+	X		
	Emigration within EU+	X		
	Net migration outside EU+	X		
	Net migration within EU+	X		

Some indicators are calculated for the total population and others for specific groups (i.e., the foreign-born and individuals born outside EU+). The listing below shows which indicators are available within each type of indicator group and specific populations/place of birth. Below the chart, the page includes a table with all values, the definition of selected indicators and an option to download a CSV file with the data and a PNG file with the chart.

Pages 4 and 5 include the following thematically grouped indicators:

Population composition

Population: Number of inhabitants.

Native-born population: Inhabitants who were born in in the country of residence.

% native-born: Percentage of inhabitants who were born in the country of residence.

% foreign-born: Percentage of inhabitants who were born abroad.

% born outside EU+: Percentage of inhabitants who were born outside the EU+ countries.

Mean age: Average age of the population calculated as the arithmetic mean.

Age dependency ratio: Population below 15 years old and over 64, divided by individuals from 15 to 64 years old.

Labour force

Working age population (15-64): Population from 15 to 64-year-old.

Native-born working age population (15-64): Population from 15 to 64-year-old who were born in the country of residence.

Labour force: Active population (employed and unemployed persons).

% of active population: Percentage of active population (employed and unemployed persons).

% foreign-born within the labour force: Percentage of foreign-born population within the total labour force.

% born outside EU+ within the labour force: Percentage of individuals born outside EU+ within the total labour force.

Inactive population: Individuals who declare themselves to not be in paid work.

% inactive at working age population: Percentage of inactive population from 15 to 64-year-old.

% foreign-born within inactive population: Percentage foreign-born within the total inactive population.

% born outside EU+ within inactive population: Percentage individuals born outside EU+ within the total inactive population.

Labour force dependency ratio: Inactive population divided by active population.

Education

% below secondary educated (age 25-39): Percentage of population from 25 to 39-years-old with less than secondary education.

% postsecondary educated population (age 25-39): Percentage of population from 25 to 39-years-old with more than secondary education.

Gender

Female inactive population: Women who declare themselves to not be in paid work.

% inactive female at working age population: Percentage of inactive women from 15 to 64-years-old.

% female within inactive population: Percentage of women within the total inactive population.

% female foreign-born within inactive population: Percentage of foreign-born women within the total inactive population.

% female born outside EU+ within inactive population: Percentage of women born outside the EU+ countries within the total inactive population.

Migration

Immigration outside EU+: Number of immigrants from countries outside EU+.

Immigration within EU+: Number of immigrants from EU+ countries.

Emigration outside EU+: Number of emigrants from countries outside EU+.

Emigration within EU+: Number of emigrants from EU+ countries.

Net-migration outside EU+: Difference between the flow of immigrants and emigrants from and to countries outside EU+.

Net-migration within EU+: Difference between the flow of immigrants and emigrants from and to countries within EU+.

4.5 Page 5: Maps

Page 5 visualises the same indicators as Page 4 in choropleth maps. Users can query the database using the same dimensions as in Indicators Page 4: place of birth, immigration scenarios and destination countries (Figure 6). The only difference is that Page 5 shows the results using two thematic maps (left and right) which allows cross-country comparisons between different immigration assumptions. Users can control the indicators selection and migration scenario selection through filters.

For instance, let's suppose that a user wants to compare the percentage of foreign-born population in 2060 between the baseline scenario and a High-persistent immigration event from North Africa. In this case, the user would select the total population within Place of birth, the indicator '% of foreign-born' as indicator, the baseline scenario in the left panel and the year 2060. Then, the user would press the bottom Copy input across and change the immigration flow scenario to High-persistent immigration event from North Africa in the right chart.

Indicator definitions are displayed under "Indicator Information" at the bottom of the page below the chart and the data table view, see also Figure 6.

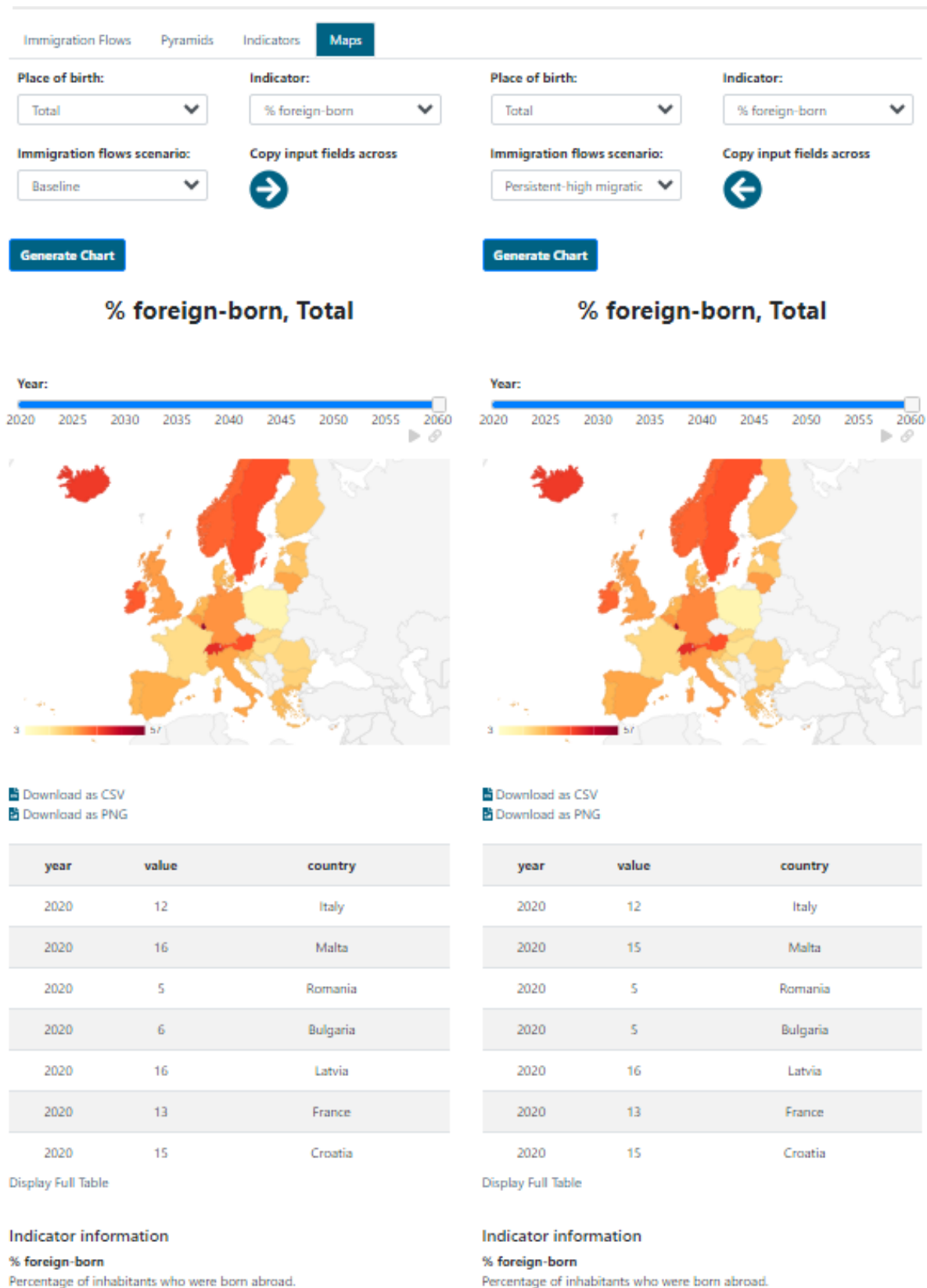


Figure 6. Design for Page 5: Maps.

4.6 Technical Implementation

The QuantMig website is delivered via a content management framework, developed by the Southampton GeoData team, based on open source technologies (PHP, Apache, MySQL, PostgreSQL) and open standards for web development (XHTML, CSS, JavaScript), supplemented with widely used open source JavaScript libraries (jQuery, Bootstrap, Font Awesome). Together this delivers a clean, accessible site, with “responsive design” to support a full range of standard browsers and devices.

The QuantMig migration Scenarios Explorer is embedded within the main QuantMig website, making use of the JavaScript tools to deliver an intuitive user interface, extended with the open source Chart.js JavaScript graphing library to deliver interactive graphs to visualise the Scenario data. For the choropleth maps used on Page 5 of the explorer, the freely available Google Geochart service and JavaScript library was used to visualise country data interactively through the same interface web page framework as the other graphs.

The Scenario data was imported to a PostgreSQL (v14) database, hosted within the server infrastructure used to deliver the QuantMig website, managed by the GeoData team at the University of Southampton. The database provides a suitable tool to store the data values, with indexes and supporting reference table structures to allow efficient query and to deliver the user functionality required.

An Asynchronous JavaScript (AJAX) method is used to pass data between the back-end server website and database tools and the front end, delivering an interactive experience for users, without the need to reload the entire page when choosing estimate variables to explore. A database query caching mechanism is used to save the results of any previous request, so that resource intensive and time-consuming database queries for a set of variable selections are run only once, then available instantly for all website users. On pages where website users are able to scroll through all pyramid charts or maps of all available data years, these are generated on user selection, allowing users to animate through all visualisations smoothly.

The SQL database table definitions are included in Appendix I.

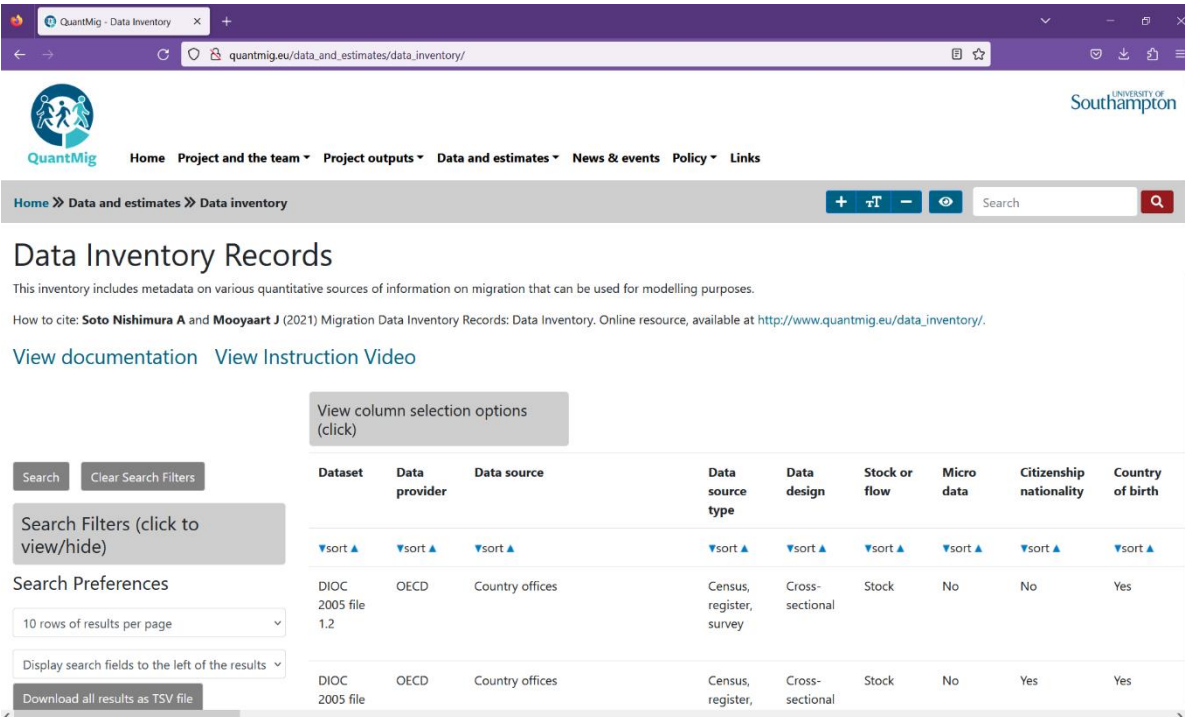
5 Data Inventories (Migration Data, Migration Drivers, Migration Policies)

5.1 Description of Inventories

5.1.1 QuantMig Migration Data Inventory

The QuantMig Migration Data Inventory (Soto Nishimura and Mooyaart 2021) is available from the QuantMig website: http://www.quantmig.eu/data_and_estimates/data_inventory/, where it can be accessed via the **Data and Estimates** section. The underlying open data deposit is also available from Zenodo, with DOI [10.5281/zenodo.7689165](https://doi.org/10.5281/zenodo.7689165). The purpose of the Data Inventory is not only to catalogue existing datasets on migration to, from, and within the EU but also to provide key meta-information on each dataset that may be relevant for the user.

The information that can be found in the inventory includes the country and year coverage of the dataset, availability, whether the dataset contains demographic information such as age, sex, labor status, and education, and dataset design (e.g., longitudinal). The inventory also includes information on different quality criteria, which can help users identify the appropriate dataset. More information on the methodology of the QuantMig Migration Data Inventory can be found in Mooyaart et al. (2020), and on the website, which also includes a brief video tutorial. A screenshot of a sample database view is enclosed in Figure 7 below.



The screenshot shows the main view of the QuantMig Migration Data Inventory. The page title is "Data Inventory Records". Below the title, there is a search bar and a table of records. The table has the following columns: Dataset, Data provider, Data source, Data source type, Data design, Stock or flow, Micro data, Citizenship nationality, and Country of birth. The table contains two rows of data.

Dataset	Data provider	Data source	Data source type	Data design	Stock or flow	Micro data	Citizenship nationality	Country of birth
DIOC 2005 file 1.2	OECD	Country offices	Census, register, survey	Cross-sectional	Stock	No	No	Yes
DIOC 2005 file	OECD	Country offices	Census, register,	Cross-sectional	Stock	No	Yes	Yes

Figure 7. Sample screenshot of the main view of the QuantMig Migration Data Inventory

5.1.2 QuantMig Migration Driver Inventory

The QuantMig Migration Driver Inventory (Soto Nishimura 2022) is also available from the same section of the QuantMig website, via: http://quantmig.eu/data_and_estimates/driver_inventory/. The underlying open data deposit is also available from Zenodo, DOI: [10.5281/zenodo.7689155](https://doi.org/10.5281/zenodo.7689155). The Migration Driver Inventory contains information on the drivers of migration that have been considered and operationalized from 176 academic articles published between 2000 and 2019. The inventory shows which migration drivers each article analyzed, the original data source used for the migration driver, the spatio-temporal coverage, in terms of country and year information available in the data source, the availability of the data, and the aggregation level of the data (individual, subnational, country). The database also contains some information on the results of the articles. Where appropriate, the effect direction and the level of statistical significance of the relationships with specific variables were recorded as well. More information on the methodology of the Migration driver database can be found in Soto Nishimura and Czaika (2022), and on the website, which also includes a video tutorial.

5.1.3 DEMIG-QuantMig Migration Policy Database

The DEMIG-QuantMig Migration Policy Database builds on the work of the DEMIG project (DEMIG 2015), specifically the DEMIG POLICY database, assembled according to the methodology proposed by de Haas et al. (2015), and available at <https://www.migrationinstitute.org/data/demig-data/demig-policy>. The DEMIG-QuantMig database provides an update and extension to the DEMIG POLICY data, for the sake of comparability and compatibility following the same template for methodology, and specifically for assessing policy changes, as in de Haas et al. (2015). At the time of writing, the database is undergoing final checks, and is expected to be released in summer 2023 via the QuantMig website, at http://quantmig.eu/data_and_estimates/policy_database. The underlying open data deposit is also available from Zenodo, DOI [10.5281/zenodo.8169110](https://doi.org/10.5281/zenodo.8169110).

On the whole, the DEMIG-QuantMig Migration Policy Database tracks more than 7,600 migration policy changes enacted by 31 European (EU and non-EU) countries for the period 1990 to 2020. The policy measures are coded according to the policy area and migrant group targeted (in terms of pathway and countries of origin), as well as the change in restrictiveness they introduce in the existing legal system. The database also includes detailed information about the source of individual policy changes, the magnitude of those changes assessed on a four-point scale (fine-tuning, minor, mid-level and major changes), the level of relevant legislation, and the policy tools used. The database allows for both quantitative and qualitative research on the long-term evolution and effectiveness of migration policies. At the time of writing (July 2023), a similar tutorial video as for the Migration Data and Driver Inventories is also being planned for the DEMIG-QuantMig Migration Policy Database.

5.2 Technical Implementation

The QuantMig website (www.quantmig.eu) is delivered via a content management framework developed by the University of Southampton GeoData team (www.geodata.soton.ac.uk), based on open-source technologies (PHP, Apache, MySQL, PostgreSQL) and open standards for web development (XHTML, CSS, JavaScript), supplemented with widely used open-source JavaScript libraries (jQuery, Bootstrap, Font Awesome). Together this delivers a clean, accessible site, with

“responsive design” to support a full range of standard browsers and devices.

The QuantMig Data and Driver Inventory and Policy Database tools are embedded within the main QuantMig website, making use of basic JavaScript to deliver a flexible and intuitive user interface. The user interface initially presents a simplified view of the imported spreadsheet data, with the options to add additional columns, filter and sort the data, or export any view (or the whole database) to a tab-separated (TSV) file via the links provided in Section 5.1, enabling further analysis in popular tools, such as MS Excel, as well as more advanced analytical environments, such as R. In the online versions of the databases, each column is filterable, sortable, and columns can be selected not to appear on the screen, so that the users can customise both the view of the inventory and the information provided. In this way, all three collections – the Migration Data and Driver Inventories, as well as the Migration Policy Database – can both serve as online reference collections, as well as source information for further, bespoke analysis.

The inputs to the QuantMig Data and Driver Inventory and Policy Database, in each case, were provided in Excel Spreadsheet format and imported into a single table within the website MySQL database, with columns to match the spreadsheet data. When website users populate data filters via the presented form options, the values are added to the MySQL database query that retrieves rows from the database, meeting the filter criteria, then the rows are displayed in a table on the webpage. The SQL database table definitions are included in Appendix I.

References

- Aristotelous, G., Smith, P. W. F., and Bijak, J. (2022) Technical Report: Estimation Methodology. QuantMig Project Deliverable D6.3. Southampton: University of Southampton.
- DEMIG (2015) *DEMIG POLICY, version 1.3, Online Edition*. Oxford: International Migration Institute, University of Oxford. www.migrationdeterminants.eu
- de Haas, H., Natter, K. and Vezzoli, S. (2015) Conceptualizing and measuring migration policy change. *Comparative Migration Studies* 3: 15.
- Mooyaart, J., Soto-Nishimura, A., Czaika, M. and de Valk, H. (2020) QuantMig Migration Data Inventory: Documentation. QuantMig Project Deliverable D4.1. The Hague and Krems: Netherlands Interdisciplinary Demographic Institute and Danube University Krems.
- Potančoková, M., Marois, G., and González-Leonardo, M. (2023a) Discussion paper: Demographic and labour force implications of high immigration events scenarios. IIASA Report. International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria.
- Potančoková, M., Marois, G., and González-Leonardo, M. (2023b) QuantMig microsimulation population projection model and migration scenarios for 31 European countries. Zenodo. <https://doi.org/10.5281/zenodo.7728049>
- Raymer, J., Wiśniowski, A., Forster, J. J., Smith, P. W. F., and Bijak, J. (2013) Integrated Modeling of European Migration. *Journal of the American Statistical Association*, 108(503), 801–819.
- Soto Nishimura, A., and Czaika, M. (2022) Migration Drivers Database, QuantMig Deliverable 5.6, Danube University Krems.
- Soto Nishimura, A., and Mooyaart, J. (2021) Migration Data Inventory Records: Data Inventory. Online resource, available at http://quantmig.eu/data_and_estimates/data_inventory/
- Soto Nishimura A (2022) Migration Drivers Data Inventory Records, 2000-2019: Data Inventory.

Online resource, available at http://quantmig.eu/data_and_estimates/driver_inventory/
Wittgenstein Centre for Demography and Global Human Capital (2018). *Wittgenstein Centre Data Explorer Version 2.0*. Available at: <http://www.wittgensteincentre.org/dataexplorer>

Appendix I – SQL Database table definitions

QuantMig Migration Estimates Explorer Tool

The following SQL commands were used in the creation of the QuantMig Migration Estimates Explorer database structure within the QuantMig PostgreSQL v14 instance.

```
# A table to import the raw Country of Birth data

create table ODcobT_import (
  id          integer,
  iteration   smallint,
  orig        varchar(17),
  dest        varchar(17),
  cob         varchar(17),
  year        smallint,
  val         numeric(9,2)
);

# Reference table for countries, populated from the import table

create table countries (
  id          serial primary key,
  country     varchar(17),
  code        char(3)
);

# Country of Birth table created from the import
# 58080000 rows, 3.3GB on disk

create table ODcobT as
select iteration, o.id as orig, d.id as dest, c.id as cob, year, val
from ODcobT_import
left join countries o on orig = o.country
left join countries d on dest = d.country
left join countries c on cob = c.country;

# Indexes for query performance

create index ODcobTa_it on ODcobT (iteration);
create index ODcobTa_or on ODcobT (orig);
create index ODcobTa_de on ODcobT (dest);
create index ODcobTa_co on ODcobT (cob);
create index ODcobTa_yr on ODcobT (year);
create index ODcobTa_va on ODcobT (year, val);

# A table to import the raw Age and Gender data

create table ODAST_import (
  id          integer,
  iteration   smallint,
  orig        varchar(17),
  dest        varchar(17),
  age         varchar(10),
  sex         char(1),
  year        smallint,
  val         numeric(9,2)
);

# Reference table for age groups, populated from the import table
```

```
create table ages (
  id          serial primary key,
  age         varchar(10),
  age_name    varchar(10)
);

# Main data table created from the import
# 444312000 rows, 25 GB on disk

create table ODAST as
select iteration, o.id as orig, d.id as dest, a.id as age, sex, year, val
from ODAST_import
left join countries o on orig = o.country
left join countries d on dest = d.country
left join ages a on ODAST_import.age = a.age
;

# Indexes for query performance

create index ODASTa_it on ODAST (iteration);
create index ODASTa_or on ODAST (orig);
create index ODASTa_de on ODAST (dest);
create index ODASTa_co on ODAST (age);
create index ODASTa_se on ODAST (sex);
create index ODASTa_yr on ODAST (year);

# A table to import the 2002-2008 IMEM data

create table IMEM2002_import (
  iteration  smallint,
  orig       char(3),
  dest       char(3),
  ageid      smallint,
  age        char(5),
  sex        char(1),
  year       smallint,
  val        numeric(9,2)
);

# Main 2002-2008 data table created from imort
# 444312000 rows, 25 GB on disk

create table IMEM2002 as
select iteration, o.id as orig, d.id as dest, a.id as age, sex, year, val
from IMEM2002_import
left join countries o on orig = o.code
left join countries d on dest = d.code
left join ages a on IMEM2002_import.ageid = a.id
;

# Indexes for query performance

create index IMEM2002a_it on IMEM2002 (iteration);
create index IMEM2002a_or on IMEM2002 (orig);
create index IMEM2002a_de on IMEM2002 (dest);
create index IMEM2002a_co on IMEM2002 (age);
create index IMEM2002a_se on IMEM2002 (sex);
create index IMEM2002a_yr on IMEM2002 (year);
create index IMEM2002a_va on IMEM2002 (year, val);

# A large table with indexes to aggregate all ODAST data across all years

create table imem_odast as
select * from IMEM2002
```

```

union
select * from odast
;

create index imem_odast_it on imem_odast (iteration);
create index imem_odast_or on imem_odast (orig);
create index imem_odast_de on imem_odast (dest);
create index imem_odast_co on imem_odast (age);
create index imem_odast_se on imem_odast (sex);
create index imem_odast_yr on imem_odast (year);
create index imem_odast_va on imem_odast (year, val);

```

QuantMig Migration Scenarios Explorer

The following SQL commands were used in the creation of the QuantMig Migration Scenarios Explorer database structure within the QuantMig PostgreSQL v14 instance.

```

# Tables to import the raw data

create table Flows_1_import (
  dest_country varchar(20),
  flows        varchar(80),
  year         smallint,
  val          numeric(9,1),
  description  text
);

create table Flows_2_import (
  year         smallint,
  age          char(5),
  sex          char(1),
  dest_country varchar(20),
  education    varchar(20),
  region_birth varchar(30),
  labour_status varchar(1),
  val          numeric(9,1),
  flows        varchar(80)
);

create table Flows_34_import (
  dest_country varchar(20),
  year         smallint,
  val          numeric(10,1),
  indicator_type varchar(25),
  indicator     varchar(80),
  place_birth  varchar(20),
  flows        varchar(80),
  description  text
);

# Reference tables for all variables, populated from the import data

create table flow_countries as table countries;

create table regions (
  id          serial primary key,
  region     varchar(30)
);

```

```

create table flows (
  id          serial primary key,
  flow       varchar(80)
);

create table flows1 (
  id          serial primary key,
  flow       varchar(80),
  description text
);

create table place_births (
  id          serial primary key,
  flow       varchar(80)
);

create table flow_ages (
  id          serial primary key,
  age        varchar(10)
);

create table educations (
  id          serial primary key,
  education  varchar(30)
);

create table indicators (
  id          serial primary key,
  indicator_type varchar(25),
  indicator   varchar(80),
  description text
);

# Main data tables with indexes, created from the import tables
# 5712, 11566800, 240720 rows

create table Flows_1 as
select d.id as dest_country, f.id as flows, year, val
from Flows_1_import
left join flow_countries d on dest_country = d.country
left join flows1 f on flows = f.flow;

create index flows1_cou on flows_1(dest_country);
create index flows1_flo on flows_1(flows);
create index flows1_yea on flows_1(year);

create table Flows_2 as
select a.id as age, sex, d.id as dest_country, e.id as education, r.id as
region_birth, labour_status, f.id as flows, year, val
from Flows_2_import
left join flow_ages a on Flows_2_import.age = a.age
left join flow_countries d on dest_country = d.country
left join educations e on Flows_2_import.education = e.education
left join regions r on region_birth = r.region
left join flows f on flows = f.flow
;

create index flows2_age on flows_2(age);
create index flows2_sex on flows_2(sex);
create index flows2_cou on flows_2(dest_country);
create index flows2_edu on flows_2(education);
create index flows2_reg on flows_2(region_birth);
create index flows2_lab on flows_2(labour_status);
create index flows2_flo on flows_2(flows);

```

```

create index flows2_yea on flows_2(year);

create table Flows_34 as
select d.id as dest_country, i.id as indicator, p.id as place_birth, f.id as
flows, year, val
from Flows_34_import
left join flow_countries d on dest_country = d.country
left join indicators i on Flows_34_import.indicator = i.indicator and
Flows_34_import.indicator_type = i.indicator_type
left join place_births p using(place_birth)
left join flows f on flows = f.flow
;

create index flows34_cou on flows_34(dest_country);
create index flows34_ind on flows_34(indicator);
create index flows34_bir on flows_34(place_birth);
create index flows34_flo on flows_34(flows);
create index flows34_yea on flows_34(year);

```

Data Inventories (Migration Data, Migration Drivers, Migration Policies)

The following SQL commands were used in the creation of the Data Inventories database structure within the QuantMig MySQL v5.7 instance.

```

# Data Inventory (143 rows)

CREATE TABLE data_inventory (
  dataset varchar(1028) DEFAULT NULL,
  data_provider varchar(1028) DEFAULT NULL,
  data_source varchar(1028) DEFAULT NULL,
  data_source_type text,
  data_design varchar(1028) DEFAULT NULL,
  stock_or_flow varchar(1028) DEFAULT NULL,
  micro_data varchar(1028) DEFAULT NULL,
  citizenship_nationality varchar(1028) DEFAULT NULL,
  country_of_birth varchar(1028) DEFAULT NULL,
  previous_residence varchar(1028) DEFAULT NULL,
  migration_type text,
  migration_definition_detail varchar(1028) DEFAULT NULL,
  migration_definition_short varchar(1028) DEFAULT NULL,
  migration_time_definition varchar(1028) DEFAULT NULL,
  origin_countries text,
  destination_countries text,
  intra_eu varchar(1028) DEFAULT NULL,
  aggregation_level varchar(1028) DEFAULT NULL,
  years text,
  update_status varchar(1028) DEFAULT NULL,
  sex varchar(1028) DEFAULT NULL,
  age varchar(1028) DEFAULT NULL,
  labor varchar(1028) DEFAULT NULL,
  education varchar(1028) DEFAULT NULL,
  q1_geographical_coverage varchar(1028) DEFAULT NULL,
  q2_migrant_characteristics varchar(1028) DEFAULT NULL,
  q3_completeness varchar(1028) DEFAULT NULL,
  q4_target_type varchar(1028) DEFAULT NULL,
  q5_timeliness varchar(1028) DEFAULT NULL,
  missings varchar(1028) DEFAULT NULL,
  accessiblity varchar(1028) DEFAULT NULL,
  link_dataset varchar(1028) DEFAULT NULL,

```

```

link_documentation varchar(1028) DEFAULT NULL,
link_report varchar(1028) DEFAULT NULL,
notes varchar(1028) DEFAULT NULL,
id int(11) NOT NULL AUTO_INCREMENT,
TS timestamp NOT NULL DEFAULT CURRENT_TIMESTAMP ON UPDATE CURRENT_TIMESTAMP,
origin_countries_new text,
origin_countries_orig text,
destination_countries_new text,
destination_countries_orig text,
PRIMARY KEY (id)
) ENGINE=InnoDB DEFAULT CHARSET=latin1;

```

Migration Drivers Data Inventory (2258 rows)

```

CREATE TABLE migration_driver_inventory (
  id int(11) NOT NULL AUTO_INCREMENT,
  TS timestamp NOT NULL DEFAULT CURRENT_TIMESTAMP ON UPDATE CURRENT_TIMESTAMP,
  driver_dimension text,
  driver_factor text,
  specific_driver text,
  driver_id text,
  variable text,
  variable_perspective text,
  type_of_migration text,
  data_scale_type text,
  data_aggregation_level text,
  data_type text,
  dataset_name text,
  data_provider text,
  data_source text,
  iso3 text,
  year text,
  country_coverage text,
  availability text,
  update_status text,
  article_year text,
  article_names text,
  link_to_articles text,
  dataset_link text,
  documentation_link text,
  notes text,
  effect text,
  PRIMARY KEY (id)
) ENGINE=InnoDB DEFAULT CHARSET=latin1;

```

Migration Policy database (7643 rows)

```

create table policy_database (
  id int(11) NOT NULL AUTO_INCREMENT,
  Country varchar(20),
  Summary text,
  Year varchar(10),
  Description text,
  Source text,
  Magnitude text,
  Level_of_legislation text,
  Policy_area text,
  Policy_tool text,
  Target_group text,
  Target_origin text,
  Specific_nationalities text,
  Restrictiveness text,

```

Comments text,

PRIMARY KEY (id),
KEY (country),
KEY (Year)

) DEFAULT CHARSET=Latin1;