Migrant Employment Data

2021 Residence Visa Output

Data dictionary for September 2025 release

This data dictionary details the various outputs produced by the MED process, defines the variables observed in the data, and provides a comprehensive list of all categories encountered in the data.

The September 2025 release of the Migrant Employment Data (MED) 2021 Residence Visa (2021RV) Output contains finalised data from January 2022 through May 2025 inclusive, and provisional data from June 2025 through September 2025 inclusive. The data set includes figures on the head count number of 2021 Residence Visa residents employed by area of residence and industry of employment. Note that this is a subset of the wider Migrant Employment Data, and that other outputs provide figures on all migrants, not just those who were approved for the 2021 Residence Visa.

The MED is produced from the Integrated Data Infrastructure (IDI), managed by Stats NZ. The results are not official statistics. They have been created for research purposes from the IDI. The opinions, findings, recommendations, and conclusions expressed in this report are those of the authors, not Stats NZ.

Access to the anonymised data used in this study was provided by Stats NZ under the security and confidentiality provisions of the Statistics Act 1975. Only people authorised by the Statistics Act 1975 are allowed to see data about a particular person, household, business, or organisation, and the results in this report have been confidentialised to protect these groups from identification and to keep their data safe.

Careful consideration has been given to the privacy, security and confidentiality issues associated with using administrative and survey data in the IDI. Further detail can be found in the Privacy impact assessment for the IDI available from www.stats.govt.nz.

Contents

Background and context

The MED 2021RV Output provides a measure of the employment outcomes of residents who achieved residence through the 2021 Residence Visa. This dataset measures the headcount number of these people who are in waged and salaried employment, per month, per category, for a number of different categories. We categorise people based on their area of residence and industry of employment. Note that this data *only* includes those people who achieved residence through the 2021 Residence Visa.

The MED is produced through the Stats NZ Integrated Data Infrastructure (IDI). The IDI is a powerful database of unit-level data on New Zealand and New Zealanders. By coupling this data with Immigration New Zealand's administrative dataset on people's visa status, we are able to produce an incredibly detailed base dataset that we then aggregate to form the MED.

The MED includes people in the labour force who are in New Zealand for less than 12 months, such as some seasonal workers and working holiday-makers. These people are not included in official migration statistics published by Stats NZ (see Stats NZ's webpage on International Migration[1]), and will be generally classified as visitors to New Zealand (irrespective of their visa type on arrival).

This dataset is built from administrative data, which may be subject to change. While we make every effort to ensure that this data is accurate, you may still see revisions to data between releases. For more information, please see our section **Caveats and assumptions**.

NOTES FOR THIS RELEASE

The release notes below apply only to this most recent release of the MED 2021RV Output. For historical release notes, please **contact us**.

No specific notes for this release.

USAGE NOTES

The following disclaimer applies to this data. If you publish data from the MED, we ask that you publish this disclaimer alongside the data, in compliance with Stats NZ's policies:

---

[1]https://www.stats.govt.nz/publications/?filters=International+migration%2CInformation+releases

HOW TO ACCESS THE MED

We provide the MED in two forms:

- as a series of zipped collections of comma-separated value (CSV) files, provided through MBIE's website
- through MBIE's `WIN7207` SQL server.

Please note that the CSV files, even zipped, are relatively large (on the order of 10MB). This makes them impractical to email to our stakeholders. By offering them through MBIE's website we are able to provide you with a link to download the files.

Data on the `WIN7207` SQL server is only accessible to MBIE employees. If you wish to access this data through SQL, please get in touch so we can set up your access permissions. If you have access to the data in the `WIN7207` SQL server, please see the section **Accessing data through SQL** at the end of this dictionary for more information on the table structure.

We will notify you via email when we update the MED. This email will include a link to download the CSV files off of our SFT server. This link will expire in three days, but if you need to access it outside of this three-day period you can email us and we can create a new link for you.

DATA ANONYMISATION

We take privacy seriously when producing the MED, and rigorously follow Stats NZ's microdata output guide to ensure the privacy of New Zealanders. We apply the following rules to all counts released as part of the MED:

For outputs which include a breakdown by industry or sector:

- Any counts where the number of employers is less than 3 are suppressed.
- Any counts below six (including zero counts) are suppressed.
- All counts which are not suppressed are rounded following the table below:

Count	Rounded to base
<18	3
19	2
20-99	5
100-999	10
1000+	100

For outputs which do not include a breakdown by industry or sector:

- Any counts below six (including zero counts) are suppressed.
- All counts which are not suppressed are random rounded to base three. This means:
  - If a count is a multiple of three, it is not altered.
  - If a count is not a multiple of three, it is randomly rounded either to the nearest multiple of three (two thirds chance) or to the next-nearest multiple of three (one third chance).

This means that all counts within the MED have an inherent introduced uncertainty. When two counts are added or subtracted, their uncertainty is also added together (so if two counts both have an uncertainty of ±2, the sum of the counts has an uncertainty of ±4). This process—referred to as 'compounding'—means that totals derived from random-rounded figures tend to have a high uncertainty. We recommend that rather than adding together values within the MED yourself, you instead use the data provided by our team at the more aggregated level.

OUTPUT FREQUENCY

We release the MED every three months. Until our December 2021 release (released in February 2022), we released the MED two months after the end of the quarter, ie:

Data to the end of...	...will be released at the end of
March	May
June	August
September	November
December	February

From the March 2022 release (to be released in June 2022) onwards, we anticipate releasing the MED three months after the end of the quarter, ie:

Data to the end of...	...will be released at the end of
March	June
June	September
September	December
December	March

List of outputs

The MED 2021RV Output contains information on the number of people who were approved for the 2021 Residence Visa, employed per industry (ANZSIC division and subdivision), regional council area per month. The dataset contains one- and two-way breakdowns of the data, providing a total of five discrete datasets.

As each of these sets is randomly-rounded, we provide data at different aggregations so that you can access the figures you need without calculating them.

CATEGORY AND COLUMN DEFINITIONS

Migrant Employment Data counts are broken down by a number of different categories. The categories used in creating the MED 2021RV Output, and their definitions, are as follows. Descriptions of possible values (for category variables) are defined following this section.

Date

> **Relevant column:** Date

The date for which this data applies. Data is provided on a monthly basis, and each date is given as the last day of the month. Each value is a snapshot as at that month.

This release contains finalised data from January 2012 through May 2025 inclusive, and provisional data from June 2025 through September 2025 inclusive.

Region

> **Relevant column:** Region

We generally aggregate MED by Regional Council area, or region. A person's region is determined based on their current residential address at the time of the measurement, which is in turn calculated based on administrative data (including health, tax, and other data which is linked to a person's residential address). If a person does not have a recognised residential address, their area is assigned based on the area of their employer.

The following values of region occur in the MED:

- Auckland Region
- Bay of Plenty Region
- Canterbury Region
- Gisborne Region
- Hawke's Bay Region
- Manawatu-Wanganui Region
- Marlborough Region
- Nelson Region
- Northland Region
- Otago Region
- Southland Region

- Taranaki Region
- Tasman Region
- Waikato Region
- Wellington Region
- West Coast Region
- Area Outside Region
- Unknown

Counts of people who have an address, but whose address cannot be linked to a region, are given the region *Area Outside Region*. Counts of people who cannot be linked to an address are given the region *Unknown*.

ANZSIC Division and Subdivision

> **Relevant columns:** ANZSIC1_Code, ANZSIC1_Description, ANZSIC2_Code, ANZSIC2_Description

Each person is allocated to an industry based on the enterprise that pays their wages. Almost all people are allocated at an Enterprise level, rather than a Business Unit (eg site or plant) level. It should be noted that many Enterprises operate in multiple industries: in this case the Enterprise (and the people it pays) will be allocated to the industry that most closely matches the Enterprise's main activity. The Australian and New Zealand Standard Industrial Classification (ANZSIC) is a classification system for industries and industry groupings. We provide our data grouped by **ANZSIC division** and **ANZSIC subdivision**, and provide both the ANZSIC code and the text description of each ANZSIC division and subdivision code. A person is allocated to a given ANZSIC division and subdivision based the Enterprise they are paid by.

ANZSIC divisions represent the most aggregated categorisation of industry levels, and can be represented either by a a one-letter code (eg 'A') or by a broad description (eg 'Agriculture, Forestry and Fishing').

ANZSIC subdivisions represent a more granular layer of detail than ANZSIC divisions, and are represented either by a three-symbol code (eg 'A01') or a more detailed description (eg 'Agriculture').

Our data is broken down into a total of 20 ANZSIC divisions and 87 ANZSIC subdivisions. For reasons of space, we do not list all of these divisions and subdivisions here. The full list can be found on Stats NZ's Ariā[2] page.

Number of people

> **Relevant column:** Number_of_People

For each row in the data, we provide the head-count number of people who fall into the specified categories for the given month. In particular:

- A person is considered to be on a given visa type for the month if they spent any time during the month on that visa type.

---

[2]http://aria.stats.govt.nz/aria/#ClassificationView:uri=http://stats.govt.nz/cms/ClassificationVersion/CARS5587

- A person is considered to be in a given region for the month if their residential address was listed in the region at any point in the month.
- A person is considered to be working in a given industry for the month if:

  - they receive any wages or salaries during the month from a business which is considered to be in that industry

- they receive wages or salaries from the same business for both the previous and the following months, and that business is considered to be in that industry. This allows us to bridge small gaps in a person's pay, which will generally be due to extended leave or administrative issues and will not actually reflect the person's leaving and re-entering employment.

Note that this means a person may be double-counted in the data. For example, if a person switches from a Student visa to a Post-Study Work visa during the month, they will appear under both the *Student* visa category and the *Work - PSW* visa category for that month. Similarly, someone shifting regions or industries may be double-counted in the data.

In addition, this is merely a measure of the number of people working in each visa/region/industry, and is not a measure of full-time equivalents.

Status

Each value is given a status, which may be *Provisional* or *Final*. *Provisional* values are produced based on Inland Revenue's ad-hoc weekly tax filings, while *Final* values are produced based on Inland Revenue's cleaned Employer Monthly Scheduling (EMS) data.

For more information on the difference between provisional and final data, please see the **Caveats and assumptions** section.

Caveats and assumptions

These results are not official statistics. They have been created for research purposes from the Stats NZ Integrated Data Infrastructure (IDI), which is carefully managed by Stats NZ. For more information about the IDI, please visit Stats NZ's website[3].

COVERAGE OF EMPLOYEES

The figures provided in this data only include people who are aged 15 or over, are employed, and receive a wage or salary. This means that the data excludes:

- people aged 14 and under
- most people who are self-employed
- people who are unemployed
- people who are not in the labour force.

It is worth noting that some people who are self-employed are also waged or salaried. This may occur if the person operates a business as a sole trader, and the business then pays the person a wage or salary. In this case, the person *is* included in the MED.

PROVISIONAL AND FINAL DATA

We attempt to make the MED as timely a series as possible. We do this by using both cleaned and ad-hoc tax data provided to the datalab by Inland Revenue (IR).

As IR's ad-hoc tax data is added to their cleaned data series, we may see some revisions. This means you may see changes to the data from release to release. When data is derived from IR's ad-hoc tax data, we have marked it as *Provisional*, which means it may change between this release and then next (when it will become *Final*). We expect the last (approximately) half-year of data in each release to be provisional: as IR's cleaned tax data becomes available we will replace previous provisional data with the new final data.

While provisional data should give a good idea of recent trends, we may still see revisions to this data as it is reconciled with final data. We may also see an increase in the number of people listed against the *Unknown* region and/or industry in our provisional data.

You may also see revisions to final data between releases. This may be because of revisions IR has made to its tax data, revisions we have made to our migration data, or improvements we have made to our methodology. We expect these changes, when they do occur, to be minor. Where possible, we will advise you of these changes through our release notes.

FURTHER INFORMATION ON DATA PROCESSING

If you have any concerns about the quality of your data, or further questions about the methodology and assumptions made in assembling the MED series, we encourage you to get in touch.

---

[3] https://www.stats.govt.nz/integrated-data/

Contact us

The MED series is produced by the Migration, Workforce and Workplace team, Data Service Delivery. For information about the series, and to be notified when the data is updated, please get in touch with the team through the following email:

migration_evidence_insights@mbie.govt.nz.

Appendix A: Accessing data through SQL

When you are approved for SQL access, you will be granted `SELECT` access to the `Migration.med` schema on the WIN7207 SQL Server. This database consists of:

- A series of data warehousing tables which store historical and current MED data.
- A series of views which allow you to access the current MED data.

We recommend accessing the MED through the provided SQL views, which are prefixed with `vw_`. For example, you could fetch MED for the Auckland region for 2020 as follows:

```
SELECT *
  FROM [Migration].[med].[vw_base_anzsic2_region_visa]
WHERE
[Region] = 'Auckland Region'
AND
YEAR([Date]) = 2020;
```

For more information on SQL access, please contact us.