Australian Government | Jobs and Skills Australia

Strong and Responsive VET Pathways

2019-20 graduate outcomes from the VET National Data Asset (VNDA)

Technical Paper

13 November 2024

Contents

Context

Purpose

This technical report serves as a companion document to the main 2019-20 VET National Data Asset (VNDA) descriptive statistics report. It provides in-depth information on the analysis methodology, data sources used and variable construction in the descriptive statistics analysis. This supplementary information aims to facilitate a deeper understanding of the results and limitations of the analysis.

The 2019-20 VNDA descriptive statistics release utilises extracted student outcomes data from VNDA for students that completed a nationally recognised VET qualification in 2019-20. It is a follow-up and more comprehensive release compared to the 2018-19 'VET Student Outcomes - Top 100 courses' publication.

Background

VNDA was initiated in 2021 under the National Skills Commission and transitioned to Jobs and Skills Australia following the passing of the Jobs and Skills Australia Act 2022. It is a collaboration between Jobs and Skills Australia and the Australian Bureau of Statistics (ABS).

VNDA is a secure data asset that links unit records from the Total VET Activity (TVA) data collected by the National Centre for Vocational Education Research (NCVER) with government administrative data from the Australian Taxation Office, Department of Social Services (DSS), Department of Education (DoE) and other sources within an ABS secure environment. VNDA leverages the ABS' existing Person Level Integrated Data Asset (PLIDA) and Business Longitudinal Analysis Data Environment (BLADE) data assets.

VNDA enables examination of a broad range of short-term and long-term outcomes following the completion of training. For example, it contains information on employee income, employment status, reliance on and exits from income support, and progression to further education and training. By using administrative data collected over multiple years, a range of measures can be observed in the years following training. It is also possible to compare values before the student's engagement in training. This means VNDA can greatly expand on the insights currently available through the NCVER Student Outcomes Survey. Figure 1 provides a broad indication of the data and timelines tracked through VNDA for this analysis.

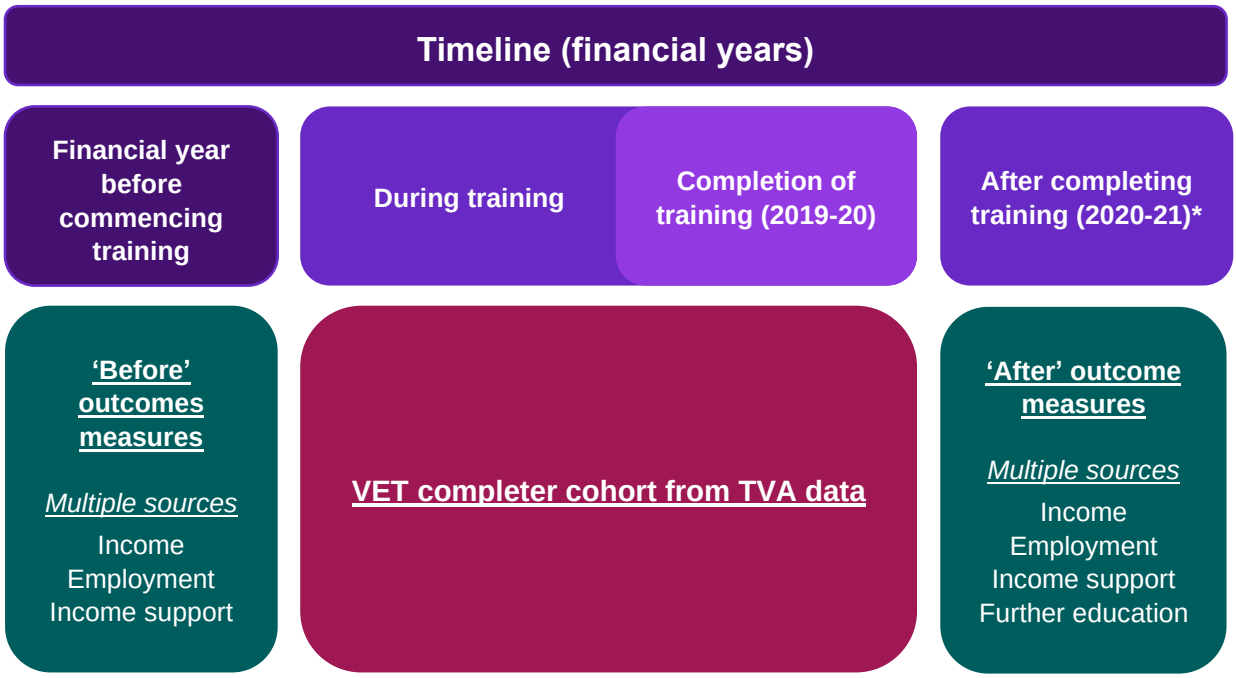

Figure 1. VNDA design for 2019-20. *Please note that income support uses a different post-training period, namely 2021-22.

As detailed in figure 1, the focus is on short-term outcomes post-training in this descriptive statistics release. This is due to the short analysis window available, which begins 1 January 2015 with the introduction of the Universal Student identifier (USI). As the analysis window widens, future publications could focus on longer term outcomes, such as three years after training, which is the timeline used for higher education graduates in the Graduate Outcomes Survey – Longitudinal (GOS-L).

It should be noted that the majority of the measured outcomes have one financial year after course completion as the post-training period. However, for income support exit rates, two financial years after training were used for the post-training period (so 2021-22 instead of 2020-21) to minimise the impact of increases in social security payments brought on by the COVID pandemic.

Scope

The analysis for this release is limited to domestic, non-school, VET students who completed a nationally recognised VET qualification in financial year 2019-20.

The data were then aggregated as follows:

- National and state and territory totals
- Australian Qualification Framework (AQF) level (national and state/territory)
- Field of Education (FoE) for each AQF level (national and state/territory)
- Qualification (national)
- Selected student characteristics, such as gender, First Nations, disability, age and location (national).

To be included in the 2019-20 release, qualifications had to meet the following criteria:

- Have more than a minimum number of completions
- Have a defined FoE.

Programs were also excluded from the analysis if the release of program-level data may have enabled any Registered Training Organisations (RTOs) that provide the program to be uniquely identified (this is described in the following confidentiality section).

For students that completed multiple qualifications in 2019-20, outcomes are recorded against the qualification with the highest AQF level. In the rare case where a student has completed multiple qualifications at the same highest AQF level, only their first recorded instance in the dataset will be retained.

Data Preparation and Treatment (Methodology)

Data sources and initial linkage

JSA (formerly National Skills Commission) received approval to develop VNDA from the ABS and data custodians in early 2021. Data used in the development of VNDA were sourced from PLIDA. PLIDA is a detailed microdata product available in the ABS DataLab and provides an integrated longitudinal microdata resource, featuring a wide range of social, health, welfare, education, and economic information on Australian residents.

VNDA integrates a subset of person-level data products available in PLIDA, as listed below in Table 1. To achieve person-level data integration between datasets, the ABS provides a *Person Linkage Spine*—or simply “Spine”—that uniquely represents an individual.

The spine facilitates linkages between government data assets which in turn enable measurements of student outcomes before and after training. For example, TVA training data are used to source training incidents for students, which can then be linked to ATO data to measure post-training income of VET graduates. Person-level data are then aggregated to the levels described above to create student summary statistics.

Table 1: Data sources in the VET National Data Asset

Custodian	Dataset name	Acronym	Reference Period used in this report
ATO	Personal Income Tax	PIT	FY 2013-14 to FY 2021-22
ATO	Payment Summary	PS	FY 2013-14 to FY 2021-22
ABS	Census of Population and Housing	Census	2016
Multiple	Combined Locations	CombLoc	2006 – 2021
Multiple	Combined Demographics	CombDem	2006 – 2021
DoE	Higher Education	HE	2005 – 2021
DEWR	Training.gov.au	TGA	2023
DSS	Data Over Multiple Individual Occurrences	DOMINO	2006 – 2023
NCVER	Training activity (NAT00120) file	TVA	2015 – 2021
NCVER	Program completed (NAT00130) file	TVA	2015 - 2021

Overview of linkage between VET completions and enrolments

The analysis described in this paper is limited to VET student completions recorded in the 2019-20 financial year that have an associated TVA enrolment record. To determine the enrolment financial year, TVA completion records were matched with corresponding program enrolment summary records. Matching was achieved through a combination of spine (as described previously), a current program identifier (see characteristics of interest for more details), and an anonymised RTO identifier. Filters were applied to the resulting linked data as detailed in the scoping section.

Consultation on analytical approach

To facilitate expert input into the development of student outcome measures, JSA established the VNDA Technical Advisory Group (TAG) in March 2021, comprising all VNDA data custodians.

During 2022, custodian involvement and support provided endorsement of outcome measures to ensure VNDA outputs are fit for purpose and appropriately utilised.

Collaboration with data custodians assisted to develop:

- the definition of key outcomes, including VET graduate employment income and employment status
- an approach, in partnership with the ABS, to assess missing data and identify and resolve outliers
- an overall data quality approach, in partnership with the ABS

- a methodology to produce indicators of course performance, and
- the selection of models, variables, cohorts and data sources.

Characteristics of interest

Table 2 provides details of the input variables used to segment the outcome measures for the descriptive statistics, with more complete definitions provided in Appendix A. Wherever possible, VNDA definitions adhere closely to the rules used by NCVER in its reporting of VET data. In some instances, however, it has been necessary to adopt a different approach, for example:

- NCVER definitions are at a subject level, whereas VNDA seeks to represent data at the qualification (program) level. Such definitions can change over time, either because of a genuine change in student or training characteristics or because of data reporting issues.
- VNDA definitions need to consider disclosure risk, leading JSA to adopt a single definition for state of training delivery when NCVER may use different rules for different collections and reports.

JSA consulted with NCVER and other VNDA data custodians on the definitions for characteristics of interest through the VNDA TAG.

Table 2: Student and training characteristics

Student or training characteristic	Brief definition
Current program identifier	A program that uniquely identifies the training package qualification or accredited qualification, agglomerating superseded versions into the same program.
Pre-training period	The financial year preceding the commencement of the training in the program completed by the student.
Post-training period	The financial year following the financial year in which the student completed the program, that is 2020-21. Please note that a different post-training period was used for income support exits. Please refer to 'receipt of income support' in Appendix A for more information.
First Nations people status	Indicates whether a student has identified as a First Nations person.
Apprentice and trainee status	Indicates whether a student is an apprentice or trainee in TVA enrolments.
Gender	Gender is self-reported by the student and indicates whether a student identifies as female, male, or other.
State of training	The state or territory of a student's training at the time of course completion.
Disability status	Indicates whether a student identified as living with a disability.
Secondary school status	Students who are undertaking 'VET in School' program while also completing secondary school studies. Note these students have been scoped out of the analysis.
international student status	Students enrolled with a national funding source in categories 'international client', 'international onshore client',

	or 'international offshore client'. Please note that international students have been scoped out of the analysis.
Location	Indicates student location (in terms of remoteness area) at time of course completion. Location currently has three distinct categories namely: Major City, Regional and Remote.
Completion time	The number of days recorded between a student's first activity start date and the completion date for their qualification.
Completion age	The student age (years) at the completion date.

Outcome measures

Table 3 provides brief definitions for outcome measures used to compute descriptive statistics; Appendix B contains complete definitions. Outcome measure definitions were also discussed at the VNDA TAG, with the objective of aligning outcome measures as closely as possible with other existing definitions (particularly regarding treatment of income by the ATO).

Table 3: Student outcome measures

Outcome measure	Brief definition
Income	Gross employee income (salary and wage data only, no business income), as per completed tax return OR information from payment summary if no tax return was finalised. In the event of missing tax return or payment summary data, it can be supplemented with income data from DSS.
Change in income	Employee income in year following completion (2020-21) minus employee income in year prior to commencement (adjusted to 2020-21 dollars to ensure comparability).
Employment (post)	Employment status in the year following completion, as indicated by presence of Income.
Further education	An indicator variable representing whether a student enrolled in a higher-level VET or higher education (university) course, in the same year as the completion of the VET course, or the year following completion of the VET course.
Income support exit rate	Percentage of students that were on income support in the month of June in the year prior to enrolment and were no longer on income support in the month of June, 2 years after completion.

Limitations of employee income

The current income and associated employment measures in VNDA are solely based on employee income. This means business income has not been included, and courses that commonly lead to self-employment may be under-reporting their income and employment

outcomes. JSA is exploring ways to incorporate business income into future iterations of VNDA.

Wage Price Index

Employee income statistics were determined at two specific time intervals: the financial year prior to student enrolment, and the financial year post completion (i.e. 2020-21). Given the enrolment year varies between students, income statistics were adjusted using the ABS Wage Price Index (WPI) to ensure a relative comparison. The WPI measures changes in the price of labour in the Australian labour market. Applying the index to income recorded in a financial year adjusts wage values relative to a reference year. All income statistics in this release use 2020-21 as the reference year.

Medians

Median statistics were used for continuous variables and provide an indication of central tendency. Medians have been used for:

- Income (in 2020-21 dollars)
- Age at completion (in years)
- Duration of training (in days)

For the two income measures, zero values were excluded when calculating medians, as follows:

- *Income post completion*—where no income was reported for a student in the 2020-21 financial year, these results were excluded when calculating the median income.
- *Change in income*—where a student had zero income for either the pre-training or post-completion income, this individual's change in income was not used in calculating the median change in income.

Percentages

Percentage statistics are used for all other outcomes and student characteristics that are binary or categorical (e.g. percentage of students employed after training and percentage of female students). The form of the percentage is provided below.

If y is the number of student completers with a particular characteristic post training and z is the total number of student completers, then the percentage x is:

$$x\% = \frac{y}{z} \times 100\%$$

Data confidentiality

JSA analyses and reports on course-level outcomes for identified VET courses. However, only statistics that meet output rules (refer to Input and output clearance | Australian Bureau of Statistics (abs.gov.au) for more information) can be released from the secure ABS DataLab environment. These rules safeguard against the risk of releasing granular-level data that could potentially identify either an individual or an RTO.

Primary units (students)

After a review early this year, JSA decided to adopt the method of perturbation to protect student level data. Perturbation is a technique which involves making small, random adjustments to data, and it is considered to be a satisfactory technique for avoiding the release of identifiable data, while maximising the range of information that can be released. These adjustments generally have a negligible impact on the underlying pattern of the statistics. Adoption of this method has also led to a faster turnaround for releasing data from VNDA while enabling the release of data at additional levels of aggregation (levels are detailed in the 'scoping' section above).

For percentage statistics, perturbation involves applying jitter to both the numerator and denominator. Essentially, applying jitter safeguards against users unpicking the statistics and disclosing person-level data. Very occasionally, perturbation may produce statistics that are non-sensical. For example, the numerator being greater than the denominator resulting in percentages greater than 100%. In this case, the resultant statistics were adjusted to equal 100%. Perturbation was also applied to median calculations to safeguard student data.

While perturbation generally has a negligible impact on the accuracy of the statistics released, users of the student outcomes data described in this report should be aware that perturbation potentially has a greater impact on accuracy for statistics which summarise smaller groups of students and those groups with greater variation. Statistics summarising very small groups of students continue to be suppressed, in accordance with ABS DataLab output clearance requirements.

Secondary units (RTOs)

Where there is a risk of disclosure of an RTO, either by few RTOs offering a qualification, FOE within an AQF level (AQF-FOE) or a large concentration of students studying at one particular RTO, suppression is then applied. In this instance, all statistics for that level can't be released from the ABS secure environment and are excluded from the analysis.

Rounding

Given the apparent confidentiality treatments and adjustments described above, JSA wanted to avoid conveying unwarranted precision with the results, and therefore rounded all percentages to the nearest whole number and rounded median dollar amounts to the nearest hundred.

Appendix A – Student and training characteristic definitions

Current Program Identifier

JSA Definition	Uniquely identifies a qualification, course, or skill set.
Data Source	NCVER Training activity, Program completed and Training.gov.au
Treatment	Aggregated cluster grouping of superseded and current program identifier.

Pre-training Period

JSA Definition	'Pre-training' is defined as the financial year prior to the earliest available date of enrolment for each student.
Data Source	NCVER Training activity
Treatment	Aggregated cluster grouping of superseded and current program identifier.

Post-training Period

JSA Definition	'Post-training' is defined as the financial year after the completion of a course, with the exception of the income support exit measure, which defines it as a 2-year period after course completion.
Data Source	NCVER Program completed
Treatment	Financial year derived from completion date for student program.

First-Nations status

JSA Definition	Indicates whether a student has identified as being First Nations.
Data Source	NCVER Training activity, Program completed, and ABS Combined Demographics
Treatment	The minimum value reported in the first subject and completion record by each student will be taken as their First Nations status. This approach prioritises students who have reported as Aboriginal or Torres Strait Islander (1, 2 or 3) in the TVA dataset, before non-Indigenous (4 or @/missing). For the remaining students that have missing values (4 or @/missing), if possible, we combined (coalesced) with the three ABS Combined Demographics binary variables (ever Indigenous, ever Aboriginal, ever Torres Strait Islander).

Apprentice and Trainee (A&T)

JSA Definition	A student with an apprenticeship/traineeship training contract.
Data Source	NCVER Training activity
Treatment	The methodology used to develop the A&T indicator is based on the presence of an A&T flag available in TVA enrolments. The following rules were applied to identify the A&T cohort: 1. If an A&T flag exists for a student for the first subject as part of a program, then that student is categorised as an A&T for that whole program. 2. Otherwise, the student is categorised as a non-A&T student.

Gender

JSA Definition	The gender variable indicates whether a student identifies as female, male, or other. Gender is self-reported by the student and must not be determined by the training organisation. JSA uses the TVA collection as the primary source, supplemented with data from other sources in some instances, to identify the gender of students per program enrolment.
Data Source	NCVER Training activity and Program completed
Treatment	1. Where a student has completed a program, JSA will take the most recently reported gender value from the TVA completions dataset to determine a student's overall gender for that program enrolment. JSA assumes this to be the last data collection point for a student and therefore their most current gender identity. 2. Where the reported gender at completion of a program is unspecified (@), JSA will take the value reported at commencement of the program (the gender value reported at the first activity start date in TVA enrolments). 3. JSA expects there to be instances where students 'have not reported gender at both time points'. Therefore, where a student has reported as unspecified at both completion and commencement of their program, JSA will convert all remaining unspecified values to missing and coalesce the TVA gender variable with the Combined Demographics module. This helps reduce the number of unspecified values, by filling in the TVA gender variable with a known value reported in the Combined Demographics module (where available). 4. Additionally, in the rare scenario where a student has reported multiple gender values at completion of their program and JSA cannot determine whether the student has reported F, M or X (non-binary), then JSA will take their gender from Combined Demographics. For example, where a student has reported an M and an F at program completion, then JSA will resolve the inconsistency by using the gender from Combined Demographics module.

State of Training

JSA Definition	The state or territory of a student's training for a program enrolment.
Data Source	NCVER Training activity and Program completed
Treatment	JSA approach to deriving a state of training variable for each student per program enrolment is as follows: 1. In the first instance, the JSA will take the value reported under state of student residence at completion of a program, where the state value is 01, 02, 03, 04, 05, 06, 07, or 08.

2. Where the state of student residence at completion of a program is not a code of 01-08, then the JSA will apply the following steps to allocate a state of training:
    a. Where the state of student residence value at completion is 09, 99, or @@ (see Table 7 for code descriptions), or where a student has duplicate values (i.e. more than one 01-08 state code reported at this time point), then the JSA will take the value reported under state of delivery location at commencement of the program.
    b. Where the state of delivery location value at commencement is 09, 99, or @@, or where a student has duplicate values (i.e., more than one 01-08 state code), then JSA considers the state of training to be unknown.

Students Living with Disability

JSA Definition	The Disability status indicates whether students consider themselves to have a disability, impairment, or long-term condition.
Data Source	NCVER Training activity and Program completed
Treatment	JSA has used the disability status for the first subject in a student's program.

Secondary School Status

JSA Definition	Students who are undertaking vocational education and training (VET) in school program while also completing secondary school studies.
Data Source	NCVER Training activity
Treatment	Where a student has received a VET in Schools Flag value of 'Y' as the last program subject enrolment, then that student will be considered a secondary school student.

Student location

JSA Definition	The location of a student's training for a program enrolment.
Data Source	ABS Combined locations
Treatment	To ensure consistency with state of training, JSA has used student location at 2019-20, or financial year of course completion. The following steps are used to allocate students to a location in a FY: • allocation to statistical area (SA) - assign each student to a SA, based on the SA with the highest cumulative duration during that financial year. In the unlikely event of ties, the smallest SA code out of all tied SAs was used. • Using the statistical area, remoteness/location codes were merged from the Accessibility/Remoteness Index of Australia Plus (ARIA+) (see: Remoteness Areas | Australian Bureau of Statistics (abs.gov.au). These remoteness codes serve as the student location variable. Note that in order to output more granular data, inner and outer regional areas were merged into one group (namely regional). In addition, remote and very remote areas have been merged into one group (namely remote).

Appendix B – Outcome measure definitions

Employee Income

JSA Definition	Income earned as an employee (excluding any income earned through self-employment or from the investment of existing capital)
Business Definition	Capture employee income from employment to understand outcomes achieved through VET training.
Data Source	ATO Personal Income Tax Return, ATO Payment Summary, Department of Social Services (DSS) Data Over Multiple Individual Occurrences
Treatment	Employee income is sourced from three locations: 1. ATO Personal Income Tax Return (PIT) 2. ATO Payment Summary (PS) 3. DSS Data Over Multiple Individual Occurrences (DOMINO) The following priority method is used: • Where a PIT record exists (for a given individual in a given year), then PIT is used. • If no PIT record exists, then PS is used. • If no PIT record, and no PS record, then DOMINO is used. PIT employee income is defined by the adding up the following data items: • Salary or wages • Allowances, earnings, tips, directors' fees etc. • Attributed personal services income • Reportable employer superannuation contributions • Total reportable fringe benefits PS employee income is defined by adding up the following data items: • Gross payment amount • Total allowances • Reportable employer superannuation contributions • Reportable fringe benefits DOMINO employee income is defined by adding up the following data items: • Continuous employee income • Variable employee income

Employment Status

JSA Definition	Employment status is defined in Australia with reference to categories used in the Census of Population and Housing (Census) and the Labour Force Survey (LFS). That is, those employed are distinguished from those who are not employed and seeking paid employment ('unemployment'), with those not working and not seeking paid employment regarded as 'not in the labour force'.
Business Definition	Infer employment status from the employee income data to understand outcomes achieved through VET training.
Data Source	ATO Personal Income Tax Return, ATO Payment Summary, DSS DOMINO
Treatment	Employment status is not consistently recorded in any single administrative database for the Australian population in a way that aligns with the requirements for the VNDA project. Consequently, JSA defines employment status in a given financial year as a binary flag based on whether that individual received any employee income in that year.

Engagement with Further Education

JSA Definition	A student is considered to have progressed to further education where they have commenced in a qualification at a higher level of education than they completed in either VET or Higher Education in the same financial year as completing training or in the subsequent financial year.
Business Definition	Understand student educational pathways based on whether students' progress to further education after completing a VET qualification.
Data Source	NCVER Training activity, Program completed and DoE Higher Education
Treatment	TVA data uses the ABS Australian Standard Classification of Education (ASCED) to classify the level of education for VET qualifications, whereas Higher Education Information Management System (HEIMS) uses its own classification. To align with ASCED, enable comparison with TVA data and de-duplicate multiple commencements, JSA uses the following processes: 1. Reclassifiyng the levels of education categories to align with the ASCED codes 2. Where a student commenced in more than one level of education in the same financial year, the lowest ASCED value (corresponding to the highest level of study) is taken to capture the student's highest level of education commencement. The process is done for both TVA and higher education data, when the highest level of enrolment for VET and the highest level for higher education is obtained separately. In the instance where a VET completing student has progressed to a TVA level of education that is unassigned (for example, 999, 991 or 992), then that student will not be considered to have progressed to further study for that subsequent program commencement.

Receipt of Income Support

JSA Definition	The income support payments used in this analysis are generally income tested allowances primarily paid to those of working age who are expected to be looking for paid work or undertaking training and education to improve their employment prospects. These payments would be expected to reduce with the skills gained in a qualification. The presence of at least one payment at a specific time and financial year would mean the student is an income support recipient for that financial year.

Income Support Exit rate	Based on a cohort analysis considering only graduates who were on income support in the month of June prior to enrolment, this indicator measures the percentage of graduates that were no longer receiving income support payments in the month of June two years after completing their qualification Exit rate=(No. of students who are on income support prior but off income support post)/(No. of students on income support prior)
Business Definition	Explore the role the VET system is playing in increasing a student's employability and enabling them to transition off social benefit payments.
Data Source	DSS - Data over Multiple Individual Occurrences (DOMINO)
Treatment	Define receipt of income support in a given financial year as a binary flag on whether an individual received payment(s) during the final reference month of the financial year (June). Using the final reference month of the financial year seems to capture an appropriate segment of the population on transitional and long-term payments. Given the reference period for completion was 2019-20 for this report, the year after training should have been June 2021. Instead, the report uses two years post-training, June 2022, to avoid the impact of the COVID outbreak.