Introduction

This document provides the technical details about the calculations used for the graduate outcomes data on the ComparED website, derived from the Graduate Outcomes Survey (GOS). It is intended for an audience with some technical background who wish to understand the statistical details of the calculations.

Data sources, variables and coverage

Graduate Outcomes Survey (GOS)

The GOS is an annual Australian survey of higher education institution graduates' employment and further study outcomes, as well as other information related to skills utilisation and course experience. Graduates are surveyed approximately four to six months after completing their courses and asked about their employment status, the type of work gained and any further study undertaken.

The GOS uses labour force indicator definitions informed by the Standards for Labour Force Statistics used by the Australian Bureau of Statistics (ABS)[1].

For the purposes of calculating results for the ComparED website, a graduate's labour force status is determined according to the definitions as detailed in this paper.

Indicators

The following graduate outcomes indicators are used on the ComparED website:

- *full-time employment*: the number of Australian graduates in full-time employment, expressed as a proportion of all Australian graduates available for full-time work;
- *overall employment*: the number of Australian graduates in any employment (including full-time, part-time or casual work), expressed as a proportion of all Australian graduates available for employment;
- *full-time study*: the proportion of Australian graduates proceeding to further full-time study; and
- *median salary*: the median salary of Australian graduates in full-time employment.

Full-time employment relates to graduates who were usually or actually in paid employment for at least 35 hours per week in the week before the survey.

Overall employment relates to graduates employed for one or more hours, as a proportion of those available for employment.

All graduate outcomes indicators are calculated from three years of pooled data. This incorporates the most recent year of published data and the two immediately preceding years. For example, indicators released in association with the 2025 GOS were based on results from the 2025, 2024 and 2023 surveys. In this paper these years are notated as Y1, Y2 and Y3, where Y1 is the most recent year of published data.

Indicators are calculated separately for undergraduate and postgraduate coursework graduates.

[1] Concepts and sources | Australian Bureau of Statistics

Coverage

The variables that were used to filter the data can be found in **Table 1**.

The coverage for each variable is applied before the calculation of the indicators and the SAS code used is provided in brackets after each variable in the table. The full code to create the indicators is available from the Social Research Centre (SRC) on request.

Table 1: Data coverage for the graduate employment outcomes indicators

	Graduate outcomes indicators			
Variables (coverage)	**Full-Time Employment**	**Overall employment**	**Further Full-Time Study**	**Median Salary**
Undergraduate level				
Undergraduate level graduates only **(if E310 in (8,9,10,13,20,21,22,23))**	**X**	**X**	**X**	**X**
Domestic graduates **(if E358 in (1,2,3,8))**	**X**	**X**	**X**	**X**
For Institution level analysis: **(if ANALYSIS = 1)** For study area analysis: **(if ANALYSIS in (1,2))**	**X**	**X**	**X**	**X**
Online responses only **(if SURVEY = 1)**	**X**	**X**	**X**	**X**
Available for full-time employment **(if AVAILFT = 1)**	**X**			
Available for employment **(if AVAILEMP = 1)**		**X**		
Question regarding further study answered **(if FURSTUD in (1,2,5))**			**X**	
In full-time employment **(if FULLEMP=1)**				**X**
Lower end and top 1% of salary responses deleted **(salary = salarya*trimsal)**				**X**
After data are pooled:				
Total minimum sample size of 25 across the pooled data years **(if n ≥ 25)**	**X**	**X**	**X**	**X**
Postgraduate coursework level				
Post-graduate coursework graduates only **(if E310 in (4,5,6,7,11,12,14))**	**X**	**X**	**X**	**X**
Domestic graduates **(if E358 in (1,2,3,8))**	**X**	**X**	**X**	**X**
For Institution level analysis: **(if ANALYSIS = 1)** For study area analysis: **(if ANALYSIS in (1,2))**	**X**	**X**	**X**	**X**
Online responses only **(if SURVEY = 1)**	**X**	**X**	**X**	**X**
Available for full-time employment **(if AVAILFT = 1)**	**X**			
Available for employment **(if AVAILEMP = 1)**		**X**		
Question regarding further study answered **(if FURSTUD in (1,2,5))**			**X**	
In full-time employment **(if FULLEMP=1)**				**X**
Lower end and top 1% of salary responses deleted **(salary = salarya*trimsal)**				**X**
After data are pooled:				
Total minimum sample size of 25 across the pooled data years **(if n ≥ 25)**	**X**	**X**	**X**	**X**

X Indicates that the restriction is applied to the data before a particular indicator is calculated.

Standard errors and confidence intervals

Standard errors, and therefore the 90% confidence intervals, were calculated using the Finite Population Correction (FPC) to account for the relatively large samples when compared with the population. The FPC is generally used in cases where the sampling fraction, the proportion of the population sampled, exceeds 5%.

In order to calculate the standard errors for the survey estimates, no non-response bias was assumed and the Agresti-Coull method[2] for confidence intervals for proportions was used.

The general formula used for confidence intervals for proportions was:

$$CI\ bound(\hat{p}) = \tilde{p} \pm z_{\frac{\propto}{2}} \times FPC\ \times SE(\tilde{p}) = \tilde{p} \pm z_{0.05} \times \sqrt{1 - \frac{n}{N}} \sqrt{\frac{\tilde{p}(1 - \tilde{p})}{\tilde{n}}}$$

$$\tilde{p} = \frac{\tilde{y}}{\tilde{n}}$$

$$\tilde{y} = y + \frac{z^2{}_{\frac{\alpha}{2}}}{2} = y + \frac{z^2{}_{0.05}}{2}$$

$$\tilde{n} = n + z^2{}_{\frac{\alpha}{2}} = n + z^2{}_{0.05}$$

Where:

$\hat{p} = \frac{y}{n}$ denotes the estimated proportion from the survey data

$\tilde{p}$ denotes the adjusted proportion used in the Agresti–Coull confidence interval

FPC denotes the finite population correction which adjusts the estimated survey standard error based on how large the sample was compared with the population

$z_{0.05}$ or $z_{a/2}$ denotes the 95th quantile of the standard Normal distribution, i.e., $Z \sim N\ (0, 1)$

n denotes the sample size

N denotes the corresponding population size

y denotes the number of sampled units exhibiting the characteristic of interest.

As mentioned above, the FPC was used to calculate the 90% confidence intervals and the standard errors, therefore an estimate of the population was required.

The Student Data Collection Completions data files were used for the GOS population estimates. The majors of the courses were used to determine the study area. For any year of GOS data, the corresponding population is drawn from Completions data file from the immediately preceding year. For example, for the 2025 GOS the population is drawn from the 2024 Completions data file.

The variables that were used to filter the population data can be found in **Table 2**.

The coverage for each variable is applied before the calculation of the population and the SAS code used is provided in brackets after each variable in the table.

[2] Agresti, A., & Coull, B. A. (1998). Approximate is Better than "Exact" for Interval Estimation of Binomial Proportions. *The American Statistician*, *52*(2), 119–126. https://doi.org/10.1080/00031305.1998.10480550

Table 2: Data coverage for the graduate employment outcomes indicator populations

Variables (coverage)	Graduate outcomes indicators: Full-Time Employment	Overall employment	Further Full-Time Study	Median Salary
Undergraduate level				
Undergraduate level graduates only **(if E310 in (8,9,10,13,20,21,22,23))**	X	X	X	X
Australian residents **(if E358 in (1,2,3,8))**	X	X	X	X
Postgraduate coursework level				
Post-graduate coursework graduates only **(if E310 in (4,5,6,7,11,12,14))**	X	X	X	X
Australian residents **(if E358 in (1,2,3,8))**	X	X	X	X

However, the population estimates will not match perfectly with the survey data, as GOS data uses major as listed by the student, major as listed by the institution or course field of education based on availability in that priority order while the Completions file lists the majors as determined by the graduate's institution.

To determine the population numbers for the indicators, all records for graduates by major were used to allow for a mismatch in order of major on the Student Data Collection when compared with the order of major selected by the combination used in the GOS. This means that the population estimates will potentially be an over-estimate in some study areas, which will mean that some standard errors will be over-estimated.

In addition, the discrepancies between the two files can lead to a survey sample size greater than or equal to the population estimate in an institution by study area level. There are also cases where there are no population figures on the completions file that match an institution by study area group, i.e. the population estimate is missing as the student(s) has written a major not coded by the graduate's institution in the Completions file or an institution was registered too late to be included in the completions file.

In these cases, the FPC was set to 1 as equivalent to not applying FPC.

Calculation of indicators and confidence intervals

Full-Time Employment

This indicator is defined as the three-year combined proportion of Australian graduates in full-time employment as a percentage of graduates available for full-time work. This indicator can be expressed as 'the proportion of Australian graduates available for full-time work who were in full-time employment at the time of the survey'. It is calculated separately for undergraduate and postgraduate coursework level graduates.

Graduates are considered full-time employed if they were usually or actually in paid employment for at least 35 hours per week in the week before the survey.

The full-time employment indicator is calculated as follows:

$$FTEMP_{pooled} = \frac{Number\ of\ graduates\ in\ full-time\ employment_{\mathrm{Y3-Y1}}}{Number\ of\ graduates\ available\ for\ full-time\ work_{\mathrm{Y3-Y1}}}$$

Where:

Number of graduates in full-time employment$_{Y3–Y1}$ is the total number of Australian graduates, in the relevant three years combined, in full-time employment at the time of the survey after filters are applied (**FULLEMP = 1**)

Number of graduates available for full-time work$_{Y3–Y1}$ is the total number of Australian graduates, in the relevant three years combined, who were in full-time employment or actively seeking full-time employment but not in full-time employment at the time of the survey after filters are applied (**AVAILFT = 1**)

The 90% confidence interval for the full-time employment indicator is calculated as follows:

$$90\%CI_{FTEMP} = \widetilde{FTEMP}_{pooled} \pm z_{\frac{\propto}{2}} \times FPC \times SE_{\widetilde{FTEMP}}$$

$$= \widetilde{FTEMP}_{pooled} \pm z_{0.05} \times \sqrt{1 - \frac{n}{N}} \times \sqrt{\frac{\widetilde{FTEMP}_{pooled} \times \left(1 - \widetilde{FTEMP}_{pooled}\right)}{\tilde{n}}}$$

$$\widetilde{FTEMP}_{pooled} = \frac{\tilde{y}}{\tilde{n}}$$

$$\tilde{y} = Number\ of\ graduates\ in\ full-time\ employment_{\mathrm{Y3-Y1}} + \frac{z^2{}_{0.05}}{2}$$
$$= Number\ of\ graduates\ in\ full-time\ employment_{\mathrm{Y3-Y1}} + \frac{1.645^2}{2}$$

$$\tilde{n} = Number\ of\ graduates\ available\ for\ full-time\ work_{\mathrm{Y3-Y1}} + z^2{}_{0.05}$$
$$= Number\ of\ graduates\ available\ for\ full-time\ work_{\mathrm{Y3-Y1}} + 1.645^2$$

Where:

$\widetilde{FTEMP}_{pooled}$ is an adjusted estimated proportion used only in confidence interval calculations

FPC is the finite population correction which adjusts the estimated survey standard error based on how large the sample was compared with the population

N is the estimated population in the institution study area or institution as relevant, over the three years combined

n is the number of graduates available for full-time work in the institution study area or institution as relevant, over the three years combined

$z_{0.05}$ is the 95th quantile of the standard Normal distribution ~ N(0,1)

Overall employment

This indicator is defined as the three-year combined proportion of Australian graduates in employment as a percentage of graduates available for work. This indicator can be expressed as 'the proportion of Australian graduates available for any kind of work (including full-time, part-time or casual work) who were in employment at the time of the survey'. It is calculated separately for undergraduate and postgraduate coursework level graduates.

Graduates are considered employed if they were usually or actually in paid employment for one or more hours in the week before the survey.

The overall employment indicator is calculated as follows:

$$OVEMP_{pooled} = \frac{Number\ of\ graduates\ in\ employment_{\mathrm{Y3-Y1}}}{Number\ of\ graduates\ available\ for\ work_{\mathrm{Y3-Y1}}}$$

Where:

Number of graduates in employment$_{Y3–Y1}$ is the total number of Australian graduates, in the relevant three years combined, in any form of employment at the time of the survey after filters are applied (**GENEMP = 1**)

Number of graduates available for work$_{Y3–Y1}$ is the total number of Australian graduates, in the relevant three years combined, who were in employment or actively seeking employment but not in employment at the time of the survey after filters are applied (**AVAILEMP = 1**)

The 90% confidence interval for the overall employment indicator is calculated as follows:

$$90\%CI_{OVEMP} = \widetilde{OVEMP}_{pooled} \pm z_{\frac{\propto}{2}} \times FPC \times SE_{\widetilde{OVEMP}}$$

$$= \widetilde{OVEMP}_{pooled} \pm z_{0.05} \times \sqrt{1 - \frac{n}{N}} \times \sqrt{\frac{\widetilde{OVEMP}_{pooled} \times \left(1 - \widetilde{OVEMP}_{pooled}\right)}{\tilde{n}}}$$

$$\widetilde{OVEMP}_{pooled} = \frac{\tilde{y}}{\tilde{n}}$$

$$\tilde{y} = Number\ of\ graduates\ in\ employment_{\mathrm{Y3-Y1}} + \frac{z^2{}_{0.05}}{2}$$

$$= Number\ of\ graduates\ in\ employment_{\mathrm{Y3-Y1}} + \frac{1.645^2}{2}$$

$$\tilde{n} = Number\ of\ graduates\ available\ for\ work_{\mathrm{Y3-Y1}} + z^2{}_{0.05}$$

$$= Number\ of\ graduates\ available\ for\ work_{\mathrm{Y3-Y1}} + 1.645^2$$

Where:

$\widetilde{OVEMP}_{pooled}$ is an adjusted estimated proportion used only in confidence interval calculations

FPC is the finite population correction which adjusts the estimated survey standard error based on how large the sample was compared with the population

N is the estimated population in the institution study area or institution as relevant, over the three years combined

n is the number of graduates available for any kind of work in the institution study area or institution as relevant, over the three years combined

$z_{0.05}$ is the 95th quantile of the standard Normal distribution ~ N(0,1)

The restrictions for this indicator can be found in Tables 1 and 2.

Further Full-Time Study

This indicator is defined as the three-year combined proportion of Australian graduates proceeding to full-time study. This indicator can be expressed as 'the proportion of Australian graduates who were in full-time study at the time of the survey'. It is calculated separately for undergraduate and postgraduate coursework level graduates.

The further full-time study indicator is calculated as follows:

$$STUDY_{pooled} = \frac{Number\ of\ graduates\ in\ full-time\ study_{\mathrm{Y3-Y1}}}{Number\ of\ graduates\ who\ responded_{\mathrm{Y3-Y1}}}$$

Where:

Number of graduates in full-time study$_{Y3–Y1}$ is the total number of Australian graduates, in the relevant three years combined, who were in full-time study at the time of the survey after filters are applied (**FURSTUD = 1**)

Number of graduates who responded$_{Y3–Y1}$ is the total number of Australian graduates, in the relevant three years combined, who responded to the study question after filters are applied (**FURSTUD in (1,2,5)**)

The 90% confidence interval for the further full-time study indicator is calculated as follows:

$$90\%CI_{STUDY} = \widetilde{STUDY}_{pooled} \pm z_{\frac{\propto}{2}} \times FPC \times SE_{\widetilde{STUDY}}$$

$$= \widetilde{STUDY}_{pooled} \pm z_{0.05} \times \sqrt{1 - \frac{n}{N}} \times \sqrt{\frac{\widetilde{STUDY}_{pooled} \times \left(1 - \widetilde{STUDY}_{pooled}\right)}{\tilde{n}}}$$

$$\widetilde{STUDY}_{pooled} = \frac{\tilde{y}}{\tilde{n}}$$

$$\tilde{y} = Number\ of\ graduates\ in\ full-time\ study_{\mathrm{Y3-Y1}} + \frac{z^2{}_{0.05}}{2}$$

$$= Number\ of\ graduates\ in\ full-time\ study_{\mathrm{Y3-Y1}} + \frac{1.645^2}{2}$$

$$\tilde{n} = Number\ of\ graduates\ who\ responded_{\mathrm{Y3-Y1}} + z^2{}_{0.05}$$
$$= Number\ of\ graduates\ who\ responded_{\mathrm{Y3-Y1}} + 1.645^2$$

Where:

$\widetilde{STUDY}_{pooled}$ is an adjusted estimated proportion used only in confidence interval calculations

FPC is the finite population correction which adjusts the estimated survey standard error based on how large the sample was compared with the population

N is the estimated population in the institution study area or institution as relevant, over the three years combined

n is the number of graduates who had a valid response in the institution study area or institution as relevant, over the three years combined

$z_{0.05}$ is the 95th quantile of the standard Normal distribution ~ N(0,1)

Median salary

This indicator is defined as the three-year combined median salary of Australian graduates in full-time employment. This indicator can be expressed as 'the median salary of Australian graduates who were in full-time employment at the time of the survey'. It is calculated separately for undergraduate and postgraduate coursework level graduates.

The median salary indicator is calculated as follows:

$$SALARY_{pooled} = Round(Median\ Salary_{\mathrm{Y3-Y1}}, 100)$$

Where:

Median Salary$_{Y3–Y1}$ is the 50th percentile (median) of the salary of Australian graduates, over the combined three years, who were in full-time employment at the time of the survey after filters are applied

Round(,100) means round to the nearest $100

Unlike proportions, there are no formulae available to calculate the standard error of a median – particularly of one that is not from a normal distribution.

In order to estimate the standard error of the survey medians, 1,000 simple random samples with replacement (SRSWR) were drawn within the institution study area or institution strata, as relevant. The median was calculated for each bootstrap sample, and the standard deviation of the resulting 1,000 medians was used as the bootstrap estimate of the standard error. This estimate was then used to construct 90% confidence intervals.

The 90% confidence interval for the median salary indicator is calculated as follows:

$$90\%CI_{SALARY} = SALARY_{pooled} \pm \mathrm{t}_{\frac{\alpha}{2}(\mathrm{n\text{-}1})} \times \hat{SE}_{SALARY}$$

$$= SALARY_{pooled} \pm \mathrm{t}_{0.05(\mathrm{n\text{-}1})} \times Stddev\left(Salary_{1000\,SRSWR}\right)$$

Where:

n is the number of graduates who had a valid response in the institution study area or institution as relevant, over the three years combined

$t_{0.05(n-1)}$ is the 95th quantile from the Student's t distribution with (*n*-1) degrees of freedom

$Stddev\left(Salary_{1000\,SRSWR}\right)$ is the standard deviation of the medians from the 1,000 SRSWR samples within the relevant strata