

# 2018-19 WGEA Data Quality Report



This report communicates the processes that were followed to assess the quality of the 2018-19 employee census data. For the 2018-19 reporting period:

- The cut-off date for inclusion in the dataset was on 17 September 2019
- The dataset and gender equality indicator (GEI) scorecard were released on 19 November 2019

**Data quality is defined as** is a measure of how *fit for purpose* a given data set is, and that there exists many aspects of data quality known as dimensions which contribute to how fit for purpose data is. In line with the literature and national data agencies and departments, the Agency uses the six data quality dimensions: relevance, accuracy, timeliness, accessibility, interpretability, and coherence.

## Timeliness

There were 4,841 reporting organisations\* (covering 9,388 businesses) included the 2018 – 19 dataset. This represents organisations who have submitted valid data to the Agency as at the cut-off date (90.7% of all relevant reporting organisations for the 2018 – 19 reporting period). Table 1 below shows that reporting timeliness by industry ranges from 98% (Electricity, gas, water and waste services sector) to 84% (Accommodation and food services sector).

**Table 1: The 2018-19 WGEA dataset – timeliness by industry**

Industry (ANZSIC Division) <sup>1</sup>	Number of reporting organisations	Number of reporting organisations	Timeliness rate (%)
Accommodation and food services	306	257	84.0
Administrative and support services	296	267	90.2
Agriculture, forestry and fishing	55	52	94.5
Arts and recreation services	116	107	92.2
Construction	236	209	88.6
Education and training	566	534	94.3
Electricity, gas, water and waste services	50	49	98.0
Financial and insurance services	269	254	94.4
Health care and social assistance	729	668	91.6
Information media and telecommunications	155	150	96.8
Manufacturing	665	598	89.9
Mining	163	150	92.0
Other services	159	145	91.2
Professional, scientific and technical services	600	550	91.7
Public administration and safety	32	30	93.8
Rental, hiring and real estate services	90	83	92.2
Retail trade	349	296	84.8
Transport, postal and warehousing	224	192	85.7
Wholesale trade	273	250	91.6
No Industry specified	3	-	-
<b>Total</b>	<b>5,336</b>	<b>4,841</b>	<b>90.7</b>

<sup>1</sup> Based on ANZSIC code provided by the reporting organisation

\* Relevant employers that submit reports to the Agency, sometimes on behalf of other subsidiary entities within their corporate structure.

## Coverage/Relevance

The 2018 – 19 dataset covers 4,341,295 employee positions in the non-public sector. This is 41.4% of the estimated overall Australian workforce as at May 2018. Table 2 below shows that the WGEA dataset covers over 60% of the workforce in the administrative and support services, financial and insurance services, information media and communications, and retail trade sectors. The dataset has lower coverage in public administration and safety (where the public sector is a dominant employer), other services (where small businesses dominate) and agriculture, forestry and fishing and construction (where small to medium businesses are common).

**Table 2: The 2018-19 WGEA dataset - coverage of all Australian employees**

Industry (ANZSIC Division)	Employees in the workforce - (ABS Labour Force Survey)	Employees in the dataset	The WGEA coverage of total employees in Australia 2017-18
	('000)	('000)	(%)
Accommodation and food services	812.9	226.6	27.9
Administrative and support services <sup>1</sup>	295.3	309.2	104.7
Agriculture, forestry and fishing	161.6	23.6	14.6
Arts and recreation services	197.0	91.8	46.6
Construction	754.2	130.0	17.2
Education and training	951.1	441.6	46.4
Electricity, gas, water and waste services	141.2	50.3	35.6
Financial and insurance services	399.6	274.6	68.7
Health care and social assistance	1,511.0	682.5	45.2
Information media and telecommunications	202.7	122.5	60.4
Manufacturing	838.8	352.8	42.1
Mining	228.3	161.9	70.9
Other services	342.3	58.9	17.2
Professional, scientific and technical services	784.1	301.8	38.5
Public administration and safety	772.6	37.1	4.8
Rental, hiring and real estate services	166.2	47.2	28.4
Retail trade	1,129.1	694.2	61.5
Transport, postal and warehousing	494.7	201.9	40.8
Wholesale trade	302.7	132.9	43.9
<b>Total</b>	<b>10,485.6</b>	<b>4,341.3</b>	<b>41.4</b>

<sup>1</sup> Note: WGEA reporting guidelines require organisations involved in labour hire functions to capture their contracted staff within their WGEA report – this can lead to inconsistency in the capture of these employees between the WGEA and ABS.

**External sources:** ABS (2019) 6291.0.55.003 - *Labour Force, Australia, Detailed - Electronic Delivery, June 2019*. Result cited is for May 2018, available in data cube EQ05

## Data quality checks

Each submission undergoes a series of automated quality checks. Appendix I lists the data quality checks that were applied for the 2018 – 19 reporting period. Organisations with potential errors were sent an email with a request to correct the data and resubmit their reports or contact the Agency. The Agency accepts anomalies if the employer provides a legitimate reason. Examples of legitimate reasons are listed on Appendix I.

## Excluded reports

There were 21 organisations with legitimate anomalies that were excluded from the dataset to prevent the distortion of benchmark results. These organisations tended to have less than 10 employees and/or had no managers in their workforce profile.

**Table 3: Industry breakdown of 21 reports excluded from the 2018 - 19 dataset**

Division	Number of reports excluded from dataset	Number of reports included in dataset	Proportion of reports excluded from dataset
Accommodation and Food Services	1	257	0.39%
Administrative and Support Services	2	267	0.74%
Agriculture, Forestry and Fishing	1	52	1.89%
Arts and Recreation Services	0	107	0.00%
Construction	1	209	0.48%
Education and Training	0	534	0.00%
Electricity, Gas, Water and Waste Services	0	49	0.00%
Financial and Insurance Services	3	254	1.17%
Health Care and Social Assistance	2	668	0.30%
Information Media and Telecommunications	0	150	0.00%
Manufacturing	2	598	0.33%
Mining	0	150	0.00%
Other Services	1	145	0.68%
Professional, Scientific and Technical Services	6	550	1.08%
Public Administration and Safety	0	30	0.00%
Rental, Hiring and Real Estate Services	0	83	0.00%
Retail Trade	1	296	0.34%
Transport, Postal and Warehousing	0	192	0.00%
Wholesale Trade	1	250	0.40%
<b>Total</b>	<b>21</b>	<b>4,841</b>	<b>0.43%</b>

### Data quality checks - Questionnaire data

The most common anomalies that were accepted by the Agency as legitimate are included in the dataset relate to the governing bodies data. Table 4 shows that these anomalies affect less than 2% of organisations in the dataset. This is an insignificant impact.

**Table 3: Common questionnaire anomalies – 2018 -19 dataset**

Anomaly	Number of organisations affected	% of organisations in the dataset (4, 841)
Target (%) is less than or equal to percentage of women on governing body	87	1.8
Organisation has no chair on its governing body	46	1.0
Too many chairs for orgs	32	0.7
Too many board members for orgs	30	0.6

### Data quality - remuneration data

Organisations are able to provide remuneration data in unit level or aggregated format. Unit level data enables a richer analysis of remuneration data.

In the 2018 – 19 reporting year, the unit level dataset accounted for 2,297,561 records (53% of the 4,341,295 employee records).

Known limitations of the benchmarks remuneration data provided by employers in general are summarised below:

- Approximately 0.7% of employee salaries are below \$13,000, which is the minimum wage for 15 year olds. Most of these salaries are legitimate as some employees are under 15 years of age or are on a disability scheme payment in this dataset. There are legitimate cases where an employee has no salary (for example, in some religious organisations; and when an employee works on commission only).
- The data for casual employees includes a ‘casual loading’ and cannot be compared to non-casual employee remuneration data.
- Some non-executive board directors have been incorrectly inputted as key management personnel in the workplace profile, which means that some salaries are particularly low for this category.
- Approximately 1.6% of employers reported the same base salary and total remuneration amounts for some employees (noting that this situation can be legitimate under certain circumstances – for example, employees who are under 18 and work less than 30 hours a week, or employees that earn less than \$450 a month).
- It is possible that salaries of some part-time or casual employees have not been annualised and/or converted to full-time equivalent amounts, which could lead to more variance in the salary data.
- Table 4 shows that there is more variance in the salary and remuneration unit level data submitted for Full-time employees compared to Part-time and Casual employees. Median absolute deviation (MAD) is used as a measure of variance as it is robust against outlier records.

**Table 4: Median absolute deviation (MAD) of annualised salaries 2018 – 19 unit level dataset**

Employment status	Annualised salary MAD	Annualised remuneration MAD
Full-time	\$ 33,418	\$ 42,715
Part-time	\$ 11,623	\$ 14,417
Casual	\$ 14,052	\$ 15,951

## WGEA census data - changes over time in reporting organisations

The overall size of the comparison group may have changed from last year. Although the most recent year-on-year change was minimal, with an overall increase of 23 reporting organisations, the composition of the comparison group may be affected (Table 5).

**Table 5: Industry breakdown of reporting organisations in WGEA's benchmark dataset over time**

Division	2013-14	2014-15	2015-16	2016-17	2017-18	2018-19
Accommodation and food services	248	258	260	233	236	257
Administrative and support services	227	239	253	253	254	267
Agriculture, forestry and fishing	42	46	47	47	49	52
Arts and recreation services	98	98	106	100	102	107
Construction	192	203	195	202	202	209
Education and training	491	520	526	512	514	534
Electricity, gas, water and waste services	51	53	52	47	46	49
Financial and insurance services	225	238	232	238	254	254
Health care and social assistance	539	613	652	652	648	668
Information media and telecommunications	119	125	134	132	136	150
Manufacturing	633	663	636	613	583	598
Mining	162	169	154	135	140	150
Other services	130	149	142	142	140	145
Professional, scientific and technical services	433	472	488	513	514	550
Public administration and safety	19	19	22	17	21	30
Rental, hiring and real estate services	63	72	80	76	82	83
Retail trade	293	315	303	294	305	296
Transport, postal and warehousing	181	196	190	186	187	192
Wholesale trade	208	222	225	229	231	250
<b>Total</b>	<b>4,354</b>	<b>4,670</b>	<b>4,697</b>	<b>4,621</b>	<b>4,644</b>	<b>4,841</b>

Organisations may have changed size due to restructure or downsizing (Table 6).

- Organisations may have modified their ownership structure.
- Organisations may have chosen to report in a different way this year (e.g. as a collective last year and as separate subsidiaries this year).

**Table 6: Organisation size breakdown of reporting organisations in WGEA’s benchmark dataset over time**

Organisation size category (number of employees)	2013-14	2014-15	2015-16	2016-17	2017-18	2018-19
0-249	1,962	2,181	2,163	2,096	2,045	2,137
250-499	1,032	1,091	1,125	1,125	1,129	1,158
500-999	644	656	648	648	655	713
1000-4999	716	742	632	615	688	692
5000+	131	127	129	129	127	141
<b>Total</b>	<b>4,354</b>	<b>4,670</b>	<b>4,697</b>	<b>4,621</b>	<b>4,644</b>	<b>4,841</b>



## WGEA census – changes over time in employee numbers

The longitudinal census dataset shows a growth in employee numbers.

→ Table 7 and 8 show that employee coverage has been consistent since 2013 – 14 and have grown steadily across industries

**Table 7: Industry breakdown of employee records in WGEA's benchmark dataset over time**

Division	2013-14	2014-15	2015-16	2016-17	2017-18	2018-19
Accommodation and food services	173,653	177,140	190,167	202,871	203,434	226,641
Administrative and support services	196,917	211,735	237,001	276,728	305,937	309,210
Agriculture, forestry and fishing	22,379	25,082	27,480	27,716	21,424	23,599
Arts and recreation services	95,105	93,460	95,579	87,645	89,102	91,770
Construction	143,259	132,805	117,004	121,141	124,862	129,979
Education and training	381,484	396,159	413,532	408,027	420,626	441,565
Electricity, gas, water and waste services	45,454	47,646	44,226	42,387	43,279	50,321
Financial and insurance services	267,363	275,319	273,307	272,757	273,038	274,570
Health care and social assistance	515,176	559,088	593,819	627,746	655,949	682,519
Information media and telecommunications	131,697	131,798	131,647	128,702	120,508	122,453
Manufacturing	371,937	366,111	345,539	338,569	344,270	352,754
Mining	190,171	177,639	148,724	136,545	142,411	161,870
Other services	50,627	59,080	55,940	62,048	53,079	58,865
Professional, scientific and technical services	288,272	291,561	289,332	276,852	283,413	301,848
Public administration and safety	27,405	25,247	29,569	22,721	34,475	37,115
Rental, hiring and real estate services	34,337	36,450	40,934	41,775	43,844	47,165
Retail trade	648,558	653,173	681,384	666,328	682,834	694,211
Transport, postal and warehousing	207,845	208,998	199,019	195,557	192,749	201,892
Wholesale trade	100,261	106,331	111,101	115,990	121,106	132,948
<b>Total</b>	<b>3,891,900</b>	<b>3,974,822</b>	<b>4,025,304</b>	<b>4,052,105</b>	<b>4,156,340</b>	<b>4,341,295</b>

**Table 8: Organisation size breakdown of employee records in WGEA's benchmark dataset over time**

Organisation size category (number of employees)	2013-14	2014-15	2015-16	2016-17	2017-18	2018-19
0-249	299,257	328,561	330,555	323,755	312,720	327,714
250-499	360,654	380,351	394,934	387,497	393,243	405,576
500-999	450,894	455,943	448,712	452,048	457,310	500,763
1000-4999	1,223,448	1,242,618	1,264,610	1,311,176	1,430,303	1,428,474
5000+	1,557,647	1,567,349	1,586,493	1,577,629	1,562,764	1,678,768
<b>Total</b>	<b>3,891,900</b>	<b>3,974,822</b>	<b>4,025,304</b>	<b>4,052,105</b>	<b>4,156,340</b>	<b>4,341,295</b>

## Appendix I – 2019 Data Quality Checks

Data quality check	Explanation of data quality anomaly provided to reporting organisations	Example of a legitimate reason for an anomaly
<b>Total remuneration is lower than expected given the base salary entered and superannuation requirements</b>	The data entered into the 'Total remuneration' column of your aggregate workplace profile is lower than expected when taking into account the base salary entered and superannuation requirements. You may have calculated additional payments (beyond base salary) incorrectly.	If an employee earns less than \$450 per month, superannuation is not legally required to be paid. In this case, base salary and total remuneration may be the same.  There may be other circumstances where superannuation is not legally required.
<b>Base salary is low</b>	The data entered into the 'Base salary' column in your aggregate workplace profile suggests that those salaries may not have been converted to annualised and full-time equivalent amounts. This has been flagged as a possible error because base salaries cannot fall below \$35,000, the minimum adult full-time equivalent base salary, or \$13,000, the minimum non-adult full-time equivalent base wage.	There are trainees or apprentices, employees on a disability scheme and/or employees under 21 years of age in this occupation.
<b>Total remuneration is high</b>	The data entered into the 'Total remuneration' column of your aggregate workplace profile is considerably higher than other employees' remuneration in this occupational category. You may have calculated additional payments incorrectly, or provided the TOTAL remuneration of all employees in a category instead of the AVERAGE remuneration paid to employees in that category.	Employees in this occupation are paid above market rates and/or receive large bonuses or other discretionary payments.
<b>Part-time base salary remuneration is less than full-time remuneration</b>	The data entered into the 'Base salary' column for part-time employees in the aggregate workplace profile suggests that those salaries have not been converted to annualised and full-time equivalent amounts, as they are considerably less than the average full-time base salary for that occupation.	Part-time employees are trainees or apprentices, on a disability scheme and/or under 21 years of age.
<b>Part-time total remuneration is less than full-time remuneration</b>	The data entered into the 'Total remuneration' column for part-time employees in the aggregate workplace profile suggests that those salaries have not been converted to annualised and full-time equivalent amounts. The current figures are considerably less than the average full-time remuneration for that occupation.	Part-time employees are trainees or apprentices, on a disability scheme and/or under 21 years of age.

Data quality check	Explanation of data quality anomaly provided to reporting organisations	Example of a legitimate reason for an anomaly
<b>Casual base salary remuneration is less than full-time remuneration</b>	The data entered into the 'Base salary' column for casual employees in the aggregate workplace profile suggests that those salaries have not been converted to annualised and full-time equivalent amounts, as they are considerably less than the average full-time base salary for that occupation.	Casual employees are trainees or apprentices, on a disability scheme and/or under 21 years of age.
<b>Casual total remuneration is less than full-time remuneration</b>	The data entered into the 'Total remuneration' column for casual employees in the aggregate workplace profile suggests that those salaries have not been converted to annualised and full-time equivalent amounts. The current figures are considerably less than the average full-time remuneration for that occupation.	Casual employees are trainees or apprentices, on a disability scheme and/or under 21 years of age.
<b>Reporting levels from CEO are not consecutive</b>	Your workplace profile contains gaps in the number of reporting levels to the CEO. The reporting levels should be consecutive, e.g. -1,-2,-3, NOT e.g. -1,-3,-4 where -2 is missing. Please see page 25 of the Reference Guide for more detail.	Equity partners usually occupy an entire reporting level. For example, they represent the entire '-1' level as they report directly to the CEO/head of business.
<b>All manager categories report to someone higher than CEO</b>	All of your managers have been allocated a reporting level to the CEO of "+1". This reporting level is only to be used for those individuals who are based overseas and who are more senior than the CEO. This usually applies to only a few individuals in global organisations.	It is a multi-national organisation which employs managers in Australia who:  are more senior in the hierarchy than the Australian CEO/head of business, and report to someone overseas.
<b>Multiple CEOs</b>	Your workplace profile has multiple CEOs. This may be because other categories of managers have been incorrectly classified as CEOs. Managers need to be grouped into the five standardised management categories according to definitions provided in the Reference Guide. Please see page 21 of the Reference Guide for more detail.	The report covers multiple organisations with registered ABNs, and/or the CEO/head of business role is shared.
<b>100 percent of workforce is one gender</b>	Your workplace profile shows that 100 percent of your workforce is one gender.	Only females/males were employed at the time of reporting.
<b>One or more CEO and no other managers</b>	Your workplace profile contains no data for manager categories other than CEO. This means your company has one or more CEOs but no managers. If this is an error remember managers need to be grouped into the five standardised management categories according to the definitions provided in the Reference Guide, then their reporting level to the CEO indicated.	The organisation only has a CEO/head of business and no other managers.

Data quality check	Explanation of data quality anomaly provided to reporting organisations	Example of a legitimate reason for an anomaly
<b>More than 20 percent of non-managers are classified in the 'other' occupational category</b>	20% or more non-managers have been categorised as being employed in the occupational category of "Other". This appears to be an error as it means a significant portion of your employees do not fit into the standard occupation categories for non-managers.	Employee works in a highly specialised or unique role that genuinely cannot be categorised in any other non-manager category.
<b>More than 20% of non-managers are classified as graduates</b>	You have categorised 20 percent or more non-managers as being employed as 'graduates'. This appears to be an error as it means at least one in five non-manager employees are on a formal graduate program.	There was a high intake of employees in the formal graduate program and/or the graduate program spans multiple years.
<b>Workforce does not contain any managers</b>	Your workplace profile contains no managers. Managers need to be grouped into the five standardised management categories according to definitions provided in the Reference Guide, then the reporting level to the CEO indicated.	It is a global organisation and does not have any managers employed in Australia.
<b>Organisation does not have a CEO/Head of Business</b>	Your workplace profile does not include a CEO. You do not need to include salary data, however you must include the gender and employment status of your CEO.	The organisation has been placed into administration and there is no CEO/head of business.
<b>No directors on the governing body/board</b>	Your response to Q2.1 of the reporting questionnaire shows that you have no board chairs or members for at least one of your governing bodies.	There are multiple owners and they are all recorded as chairs
<b>Too many board members for orgs</b>	Your response to Q2.1 of the questionnaire shows the number of board members is more than 20 for one or more of your governing bodies.	The organisation is a union, religious or other organisation with a high number of boards.
<b>Too many chairs for orgs</b>	Your response to Q2.1 of the questionnaire shows that at least one of your governing bodies has more than two chairs.	The organisation is a union, religious or other organisation with a high number of boards.
<b>The target set for the governing body is less than or equal to the representation of women on the board. The target has already been reached.</b>	Your response to Q2.1 in the reporting questionnaire shows that the target set for one or more of your governing bodies is less than or equal to the representation of women on your board.	The target date falls in the reporting period or later.
<b>Primary Carers Leave - high number of weeks</b>	Your response to Q5.1 of the questionnaire shows the number of weeks offered for primary carer's leave is more than one year (52 weeks). That may be an error as it is an unusually large amount of leave.	Industry standard

Data quality check	Explanation of data quality anomaly provided to reporting organisations	Example of a legitimate reason for an anomaly
<b>Secondary Carers Leave - high number of days</b>	Your response to Q6.1 of the questionnaire shows the number of days offered for secondary carer's leave is more than six months (approximately 180 days). That may be an error as it is an unusually large amount of leave.	Industry standard
<b>0 weeks of paid primary carer's leave has been entered, despite the organisation indicating that it does offer primary carer's leave.</b>	Your response to Q5.1 of the reporting questionnaire shows that you offer 0 weeks of primary carer's leave, despite your organisation indicating it does offer paid primary carer's leave.	None
<b>0 days of paid secondary carer's leave has been entered, despite the organisation indicating that it does offer secondary carer's leave.</b>	Your response to Q6.1 of the reporting questionnaire shows that you offer 0 days of secondary carer's leave, despite your organisation indicating it does offer paid secondary carer's leave.	None
<b>Number of resignations while on parental leave is greater than the number of employees on parental leave</b>	Your responses to Q7 and Q8 of the reporting questionnaire show the number of female manager resignations while on parental leave is considerably larger than the total number of female managers who utilised parental leave	None
<b>Number of promotions is greater than the number of appointments.</b>	Your response to Q1.11 in the reporting questionnaire shows the number of promotions is larger than the total number of appointed, however the number of appointments should also include the number of promotions.	None