

Safe Work Australia

National Dataset for Compensation- based Statistics

Explanatory notes

FEBRUARY 2023

CONTENTS

National Data Set for Compensation-Based Statistics	3
Scope and coverage.....	3
Serious claims	3
Age of employee.....	4
Time lost from work	4
Compensation paid.....	4
Industry classification	4
Occupation classification	4
Details of injuries and diseases.....	5
Insufficiently coded data	5
Confidentiality	5
Time-series analyses.....	5
Denominators used to calculate rates	5
Reliability of data	6

Explanatory notes: National Dataset for Compensation-based Statistics

Safe Work Australia compiles information on work-related injuries, diseases and fatalities in two primary datasets: the National Dataset for Compensation-based Statistics and the Traumatic Injury Fatalities database. This document provides explanatory notes for the National Dataset for Compensation-based Statistics.

National Data Set for Compensation-Based Statistics

Safe Work Australia compiles national workers' compensation statistics using data obtained from workers' compensation authorities in each state, territory and the Commonwealth Government. These data are collated into the National Dataset for Compensation-based Statistics (NDS), which is Safe Work Australia's primary source of information on work-related injuries and diseases.

Scope and coverage

The NDS is comprised of accepted workers' compensation claims, which are presented by the financial year of lodgement. A financial year begins on 1 July and ends on 30 June.

The NDS does not cover all cases of work-related injuries and diseases for the reasons below:

- While state, territory and Commonwealth Government workers' compensation legislations provide coverage for the majority of employees, some specific groups of workers are covered under separate legislations. Claims lodged by police in Western Australia and military personnel of the Australian Defence Forces are not included.
- Work-related injuries and diseases of self-employed workers are under-represented because workers' compensation schemes do not generally cover self-employed workers. Around 10% of Australian workers are self-employed. Denominators used to calculate rates only include the jobs and hours of employees who are eligible for workers' compensation.
- Mesothelioma claims are under-represented because many mesothelioma cases, which are commonly linked to work-related exposure to asbestos, are compensated through mechanisms other than workers' compensation.
- Diseases are under-represented because many diseases result from long-term exposure to agents or have long latency periods, which makes the link between the work-related disease and the workplace difficult to establish.

Serious claims

A serious claim is an accepted workers' compensation claim that involves one or more weeks away from work and excludes all fatalities, and all injuries experienced while travelling to or from work or while on a break away from the workplace.

Age of employee

The age of an employee is derived from their date of birth and the date on which the injury occurred or the disease was first reported to the employer. Data related to the open-ended age group of 65 years and above should be used with caution as claims can be made by people who are no longer in the workforce, particularly for work-related diseases that arise after employees retire. The employee estimates only count people who are currently working. It is possible that rates for this age group overstate the actual rates.

Time lost from work

Time lost from work is measured in working weeks and excludes estimates of future absences. Time lost from work comprises the total period of time for which compensation was paid. The time lost is not necessarily continuous and may occur over a number of separate periods. Where an employee returns to work on a part-time basis, they may continue to receive pro-rata payments and the total number of hours for which compensation has been paid is included in calculating the time lost.

The median is used as the measure of central tendency because long-term claims that involve lengthy periods of time lost from work would significantly skew an arithmetic mean.

Data from the preliminary year are excluded when reporting time lost from work because claims from the preliminary year are likely to be open and claimants may accrue more time off work in subsequent years.

Compensation paid

The median is used as the measure of central tendency because long-term claims that involve lengthy periods of time lost from work would significantly skew an arithmetic mean. The calculation of median compensation payments includes payments for all serious claims for each financial year. Amounts of compensation paid are rounded to the nearest \$100.

Data from the preliminary year are excluded when reporting compensation paid because claims from the preliminary year are likely to be open and claimants may accrue more compensation payments in subsequent years.

Industry classification

Information about the industry of the claimant is coded using the Australian and New Zealand Standard Industrial Classification, 2006.

The industry of the claimant is based on the industry of the business that employs the claimant. As industry is defined using the industry of the employer, a claim made by a person employed under labour hire arrangements is coded to the labour supply services industry class, which is in the administrative and support services industry. Industry of employer will be different to industry of workplace for some workers.

Occupation classification

Information about the occupation of the claimant is coded using the Australian and New Zealand Standard Classification of Occupations, First Edition.

Details of injuries and diseases

Information about injuries and diseases of claimants is coded using the Type of Occurrence Classification System, Third Edition, Revision 1. The classification system is used to code the:

- nature of injury or disease
- bodily location of injury or disease
- mechanism of injury or disease
- breakdown agency of injury or disease, and
- agency of injury or disease.

Insufficiently coded data

Some claims are not fully coded due to insufficient information being provided at the time of the claim. Where there is a significant number of incompletely coded claims, they may be separately identified in residual categories like 'other and unspecified' or 'not elsewhere classified'. These claims are included when totals are calculated.

Confidentiality

Claim numbers are rounded to the nearest 5 to help protect confidential information about employers and employees. Due to rounding, differences may appear between the reported totals and the sums of rows or columns. Rates and percentages are calculated using unrounded numbers.

Time-series analyses

Comparison of preliminary and non-preliminary data should be done with caution. Non-preliminary data are more likely to have been finalised and are generally more accurate than preliminary data.

When analysing trends over time, consideration needs to be given to legislative changes that may influence trends in workers' compensation data. Information on workers' compensation arrangements can be found in Safe Work Australia's Comparison of Workers' Compensation Arrangements in Australia and New Zealand.

Denominators used to calculate rates

Estimates of the number of employees and hours worked for each Australian workers' compensation jurisdiction are supplied annually by the ABS. The data are primarily derived from the Labour Force Survey, which are adjusted to account for differences in scope between the Labour Force Survey and workers' compensation coverage.

The ABS provides two sets of estimates for each jurisdiction—one is split by sex, age and industry and the second is split by occupation. This restricts presentation of rates to the categories supported by the ABS data. Therefore, it is not possible to calculate rates for occupational groups within specific industries.

Reliability of data

Data are subject to two types of errors—non-sampling errors and sampling errors.

Non-sampling error

Non-sampling errors may occur in any statistical collection due to:

- incorrect inclusion or exclusion of respondents or cases.
- non-response of respondents.
- inaccurate information from respondents.
- inaccurate recording of information by data collectors.
- deficiencies in data collection materials and processes.
- errors that occur during the entry, coding and editing of data.

Non-sampling errors may affect the numerator and denominator data. It is difficult to quantify non-sampling error.

Sampling error

Sampling error is a measure of the variation that occurs when a sample, rather than an entire population, is surveyed. Standard error indicates the extent to which an estimate of a sample varies from the estimate of the population from which it is drawn. When a standard error is expressed as a percentage of the estimate to which it relates, it is known as a relative standard error (RSE).

The denominator data used to calculate rates of serious claims are subject to sampling error. Rates that are calculated using denominators with high RSEs are annotated with one asterisk (*) to indicate that the RSE of the denominator is 25% or higher and two asterisks (**) if the RSE is 50% or higher.