



Technical Paper for Research Publications

The Queensland Sentencing Advisory Council produce a range of publications that contain research and statistical information. This technical paper explains some of the terminology, data sources, counting rules and methodologies that are common across most of these publications.

Data sources

The data used in Council publications is sourced from several different agencies.

The most common source of data used in most of the Council's research publications is the Courts Database which contains information collected from administrative information systems used by court staff. Courts Services Queensland use the Queensland Wide Interlinked Courts (QWIC) case management system to record information about all criminal cases that are heard in Queensland, including sentencing outcomes. Data extracted from QWIC is provided to the Queensland Government Statistician's Office (QGSO), Queensland Treasury which maintains the Courts Database. QGSO provides a copy of the Courts Database to the Queensland Sentencing Advisory Council on an annual basis.

The Courts Database is continually updated as more information is entered into QWIC, extracted, and provided to QGSO. The data presented in each research publication will be clearly marked in terms of the valid timeframe from which the data has been derived. The information presented in research publications may vary from data published elsewhere by other agencies due to differences in the dates that data were extracted.

Two other commonly used sources of data are data extracted by Queensland Corrective Services from the Integrated Offender Management System (IOMS) and by the Queensland Police Service from the Queensland Police Records and Information Management Exchange (QPRIME). The datasets obtained from these and other agencies are requested on an ad-hoc basis and vary depending on the types of research being conducted. Details on ad-hoc data requests, and the methodologies and counting rules are explained in detail in each Council publication that uses these data sources.

On occasion, additional information about cases, including court outcomes, is sourced from sentencing remarks obtained from the Queensland Sentencing Information Service (QGIS).

A note on terminology

The word *data* has two constructions: in academic literature it is often treated as a plural noun (*data are...*); whereas in modern non-scientific English it is treated as a mass noun which takes singular verbs (*data is...*).

This difference stems from differing definitions. Academic journals tend to consider *data* to be the plural of *datum* which is defined as a single observation or fact. In this sense, the word *data* is plural and refers to many facts or observations. However, in modern English, the word *data* has a meaning synonymous with that of the word *information* and has a life of its own quite independent of the now defunct *datum*.

As the community is one of the Council's primary audiences, it has decided to use language that is accessible to the public. Consistent with this, research publications published from February 2021 will use singular verbs, singular modifiers and the singular pronoun when referring to data.

Limitations

The data presented in the Council's research publications is a simplified representation of Queensland's complex criminal justice system and subject to a range of limitations. Caution should therefore be used when interpreting the data, particularly due to the following:

- Data is derived from administrative systems that are designed for operational, rather than research purposes. The accuracy of information presented in research publications reflects how administrative information is structured, entered, maintained and extracted from these administrative systems.
- Information on sentencing outcomes includes information about the court that imposed the sentence, and not the court of original lodgement, as this is how this information is recorded administratively.
- Sentencing details are provided in relation to the original, or 'first instance' judgments relating to the offences dealt with. Information relating to any appeals and their outcomes are not included in the administrative datasets.
- An offender may be sentenced for multiple offences at the same time. The sentencing outcomes presented in research publications generally relate to the most serious offence (MSO) for which the offender was sentenced, though additional details of other offences and sentencing outcomes are also presented where relevant.

Counting rules and methodologies

Children

Offenders are classified as adults or children based on the penalty type imposed. If the penalty imposed is made pursuant to the *Youth Justice Act 1992* (Qld), the offender is categorised as a child. If the penalty is imposed under the *Penalties and Sentences Act 1992* (Qld), the offender is categorised as an adult.

Determining offender characteristics

Due to inconsistencies in the way administrative data is recorded, some characteristics of a person, such as a person's gender, age or Aboriginal and Torres Strait Islander status may be recorded differently each time a person interacts with the criminal justice system. To improve the integrity of analysis in this report, the most commonly occurring characteristics for each person have been applied across all data points.

For example, if an offender is recorded as having a different date of birth across different cases, the date of birth that is most frequently recorded across the highest number of cases is the date of birth that is used for all analyses pertaining to that offender.

If half of an offender's cases recorded the person as being male, and half as being female, the person's gender is classified as unknown. If half of an offender's cases recorded the person as being Aboriginal and Torres Strait Islander, the person is classified as being Aboriginal and Torres Strait Islander.

Multiple penalties

In some instances, it is possible for a single offence to result in multiple penalties. For example, a court may sentence an offender to both a probation order and a community service order.¹ Courts also have the power to impose a fine in addition to any other sentence imposed.² Where multiple penalties are imposed for a single offence, most research publications will include details only for the most serious penalty imposed. The seriousness of penalties is classified according to the sentence type classification produced by the Australian Bureau of Statistics as an appendix to its *Criminal Courts, Australia* publication.

Offence categories

Where data is analysed by offence categories, the categories listed reflect the classifications assigned by the Australian and New Zealand Standard Offence Classification (ANZSOC) scheme. This classification scheme aims to provide a uniform national statistical framework for classifying offences used by criminal justice.³ In some instances, the Queensland Extension to the Australian Standard Offence Classification (QASOC) may be used.⁴ The

¹ *Penalties and Sentences Act 1992* (Qld) s 109.

² *Penalties and Sentences Act 1992* (Qld) s 45(2).

³ For more information, please see Australian Bureau of Statistics, *Australian and New Zealand Standard Offence Classification (ANZSOC), 2011* (02 June 2011).

⁴ For more information, please see Office of Economic and Statistical Research, *Australian Standard Offence Classification (Queensland Extension)*, (2008).

QASOC extension adds a fourth level to the three-level national classification which provides Queensland-specific distinctions which are considered important by Queensland agencies.

Offender rates

Rates are calculated using the Australian Bureau of Statistics (ABS) estimated resident population (ERP) and expressed as the number of offenders per 1,000 or 100,000 population⁵ aged 10 years and over. Where rates are presented by gender, age, Aboriginal and Torres Strait Islander status or geographical location, the rates are calculated using the population of the relevant group. For example, the statement '88.7 offenders per 1,000 people' should be read as '88.7 offenders per 1,000 people in the relevant group aged 10 years and over'.

Remoteness areas

In the Courts Database, the extent of geographical data provided to the Queensland Sentencing Advisory Council includes the suburb and postcode of a defendant's residential address, and the suburb and postcode in which an offence was committed. For analytical purposes, the Council maps suburb and postcode data to standardised geographical areas using correspondence files provided by the ABS and the QGSO, Queensland Treasury.

For example, data on remoteness areas is provided as part of the ABS's Australian Statistical Geographic Standard (ASGS).⁶ The ASGS was developed to collect and disseminate geographically classified statistics. The ASGS remoteness structure divides Australia into five classes of remoteness based on relative access to services. The ASGS provides correspondence files which enable the Council to map combinations of postcodes and suburbs to relevant remoteness areas.

Unique offenders

Throughout the Council's research publications, offenders sentenced at various court events are 'linked' to identify where the same individual has been sentenced multiple times at different court events. This linkage process allows the analysis of 'unique offenders' and provides a history of prior and subsequent court events for each person, allowing for analysis of recidivism.

Prior to 2018, the Council undertook inhouse linking, using an offender's name and date of birth to identify unique offenders using a data linking software application known as LinkKing. From 2019, the QGSO undertakes a computational methodology each year to identify unique offenders in the Courts Database. This process involves deterministic and probabilistic matching techniques to account for anomalies in relation to common administrative data errors, such as the use of nicknames and incorrectly transposed details.

Expected release date

The expected release date is the date at which it is anticipated that an offender will be released from custody using the administrative data that is collected at a sentencing event. As such, the expected release date is an approximation, and the actual release date may be later than expected (for example, if a prisoner's parole application is denied, they might spend more time in custody than estimated).

The expected release date is used for calculating recidivism (that is, a person's ability to reoffend is limited during the time they are incarcerated – recidivism is usually calculated from the point at which a person is released into the community and has the opportunity to offend), and is also used to provide greater context for analysis pertaining to custodial outcomes (especially in comparing the seriousness of different types of custodial penalties).

A person's expected release date is calculated as:

- For non-custodial penalties and wholly suspended sentences, the expected release date is the date the penalty was given.
- For partially suspended sentences, the expected release date is the date of sentence, plus any days of actual imprisonment to be served, less any days of declared pre-sentence custody.
- For sentences of imprisonment, the expected release date is the date an offender becomes eligible for parole. If no parole date is specified at sentencing, parole eligibility is estimated at 50 per cent of the sentence (less any pre-sentence custody), or 80 per cent for cases where a serious violent offence declaration is made.

⁵ The scale is chosen based on the number of offenders.

⁶ Australian Bureau of Statistics, Australian Statistical Geography Standard (ASGS): Volume 5 - Remoteness Structure, July 2016, (16 March 2018).

Recidivism methodology

There are considerable challenges in measuring recidivism. The methodology used in recidivism analysis varies depending on factors such as the quality and availability of data, and the specific research question that is being addressed. In most of research publications, the Council generally operationalises recidivism as any sentencing event that is followed by another sentencing event within two years of an offender's expected release from custody. Each Council publication may use different methodologies or counting rules to assess recidivism in specific situations. As such, each publication that contains recidivism analysis will provide details on the exact methodology used, and this technical paper will not go into this level of detail.

Generally, traffic offences are excluded from recidivism analysis.

Statistical analysis

The research publications produced by the Council contain both descriptive and inferential statistics.

Descriptive statistics

Descriptive statistics help to describe and summarise data in a meaningful way. These include numbers, rates, proportions and measures of central tendency (such as the average or median). Descriptive statistics are often presented in charts and tables. The following statement is an example of descriptive statistics:

In Queensland, the average age of sentenced offenders is 31.2 years, and 76.4 per cent are men.

Inferential statistics and statistical significance

Inferential statistics test whether differences or changes observed in descriptive statistics are 'statistically significant'. A change or difference is deemed to be 'statistically significant' if the probability of it occurring is unlikely to have occurred by chance. Another way of looking at this is to consider that there are always variations in the number of cases sentenced each year, and there is always some variation in the percentages and numbers of the various attributes which are recorded. Statistically significant changes are those that are of such a magnitude that they exceed the level of change that could be expected due to usual variation alone.

In the Council's research publications, a finding is considered to be 'statistically significant' if the probability of the result occurring by chance is less than 5 per cent.

The following statement is an example of inferential statistics:

Aboriginal and Torres Strait Islander peoples sentenced in Queensland were younger than their non-Indigenous counterparts and this difference was statistically significant.

A note on terminology

In the Council's research publications, the word 'significant' is only used when referring to statistical significance. If a publication states that a change or difference was 'significant', it means that statistical tests were performed, and the results were found to be statistically significant.

Types of statistical analysis

A range of statistical tests are used throughout the Council's research publications, depending on the specific research questions being addressed and the types of data being analysed.

Univariate analysis refers to any statistical analysis which involves one outcome variable, and can be contrasted with multivariable analysis which involves more than one output variable. The phrase bivariate analysis is sometimes used to refer to the statistical analysis of two variables to determine whether there is a relationship between them.

A common statistical test used in many research publications is the *two-sample independent groups t-test* which compares the average of two datasets to determine whether there is a statistically significant difference in the average value of each group. The following statement is an example of the findings from a t-test where the average value (age) of two groups (men and women) is analysed:

On average, male offenders were younger than female offenders, and this difference was statistically significant.

Another common test is the Pearson's Chi-square test, which is commonly used to evaluate the independence of two categorical variables. That is, it can be used to determine whether a relationship exists between two categorical variables. The following statement is an example of the findings from a chi-square test where one categorical variable (gender) is tested to see if it is related to a second variable (aggravating circumstances):

The proportion of men sentenced for unlawful use of a motor vehicle with aggravating circumstances was significantly higher than for women.

Definitions

Aboriginal and Torres Strait Islander status	<p>A defendant's Aboriginal and Torres Strait Islander status is based on information recorded by police and transferred to the administrative systems used by Queensland Courts. The information is based on self-identification by the offender and the quality of data is dependent on the accurate recording of data by police and whether the data was able to be transferred to the courts administrative systems. As such, cases which are not brought by police (for example, traffic offences) may be missing information on Aboriginal and Torres Strait Islander status.</p> <p>A defendant's Aboriginal and Torres Strait Islander status is based on one of the following four options:</p> <ol style="list-style-type: none">1. neither Aboriginal nor Torres Strait Islander (non-Indigenous),2. Aboriginal,3. Torres Strait Islander, and4. both Aboriginal and Torres Strait Islander.
Average (or mean)	<p>The average is a measure used to describe the central position of a dataset.</p> <p>The average is calculated by adding up all the values in a dataset and dividing the sum by the total number of values.</p> <p>The average is affected by outliers – extreme values at either end of the distribution can cause the average to shift significantly. When the sample size is large and does not include extreme (or outlier) values, the mean usually provides the preferred measure of central tendency.</p>
Case (or sentencing event)	<p>A case (also referred to as a 'sentencing event') refers to the collection of charges for a single offender that are finalised on the same day at the same court level and court location.</p> <p>Where there are multiple offenders dealt with jointly during a court event, each offender is recorded as a separate case.</p>
Charge	<p>A formal accusation of an offence. Each sentencing event relates to at least one charge. One charge per event is flagged as the most serious offence (MSO).</p>
Children	<p>Offenders are classified as adults or children based on the penalty imposed. If the penalty is imposed pursuant to the <i>Youth Justice Act 1992</i> (Qld), the offender is categorised as a child. If the penalty is imposed under the <i>Penalties and Sentences Act 1992</i> (Qld), the offender is categorised as an adult.</p> <p>For the purposes of research publications, the term 'child' is used to refer to a person aged under 18 years.⁷ Any child under the age of 10 years at the time of the offence is not considered to be legally responsible for any offending behaviour. A child under 14 can only be criminally responsible if the prosecution proves the child had the capacity to understand they should not do the act or make the omission at the time of doing it.</p>

⁷ Until 12 February 2018, children aged between 10 and 16 were sentenced as children. From 12 February 2018, 17-year olds are also treated as children.

Court level

The type of court in which the offender was sentenced. Research publications will usually refer to three tiers of courts:

1. Magistrates Courts
2. District Court
3. Supreme Court
4. The District Court and Supreme Court are often referred to collectively as the 'higher courts'. The Magistrates Courts are sometimes referred to as the 'lower courts'.

For offenders sentenced as children, outcomes will usually be reported at two court levels:

1. Childrens Court (Magistrates Court level)
2. Higher courts (for sentences imposed by the Childrens Court of Queensland (as constituted by a Childrens Court judge or a District Court judge) or the Supreme Court).

Interquartile range (IQR)	A measure of dispersion among values. It represents the middle 50 per cent range of values (that is, between the 25th and 75th percentiles). A smaller IQR represents less variability in the variable being measured.																						
Mean	See 'average'.																						
Median	<p>The median is a measure used to describe the central value of a dataset.</p> <p>The median is the middle value (or the half-way point) of an ordered dataset. Half of the values lie above the median, and half below.</p> <p>The advantage of using the median is that, compared to the mean, it is relatively unaffected by extreme scores at either end of the distribution.</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td>1</td><td>1</td><td>4</td><td>5</td><td>8</td><td>9</td><td>11</td><td>11</td><td>13</td><td>14</td><td>87</td> </tr> <tr> <td colspan="5"></td> <td style="text-align: center;">Median</td> <td colspan="5"></td> </tr> </table>	1	1	4	5	8	9	11	11	13	14	87						Median					
1	1	4	5	8	9	11	11	13	14	87													
					Median																		
Mode	The mode refers to the value that occurs most often in a dataset. s																						
Most serious offence (MSO)	Defined as the offence that received the most serious sentence, as ranked by the classification scheme used by the ABS. ⁸ For example, immediate imprisonment exceeds a partially suspended sentence even if the other penalty is longer (and therefore arguably, more severe).																						
Offender	<p>An offender in the context of the Council's research publications refers to a person who has been sentenced for one or more criminal offences.</p> <p>A case may involve multiple offenders, and an offender may appear in multiple cases. Where there are multiple offenders in an individual case, these offenders are counted separately. Where an offender appears across multiple cases, each case is counted separately.</p>																						

⁸ See 'Appendix 3 (Sentence Type Classification)' to the explanatory notes for Australian Bureau of Statistics, *Criminal Courts, Australia, 2018-19* (27 February 2020).

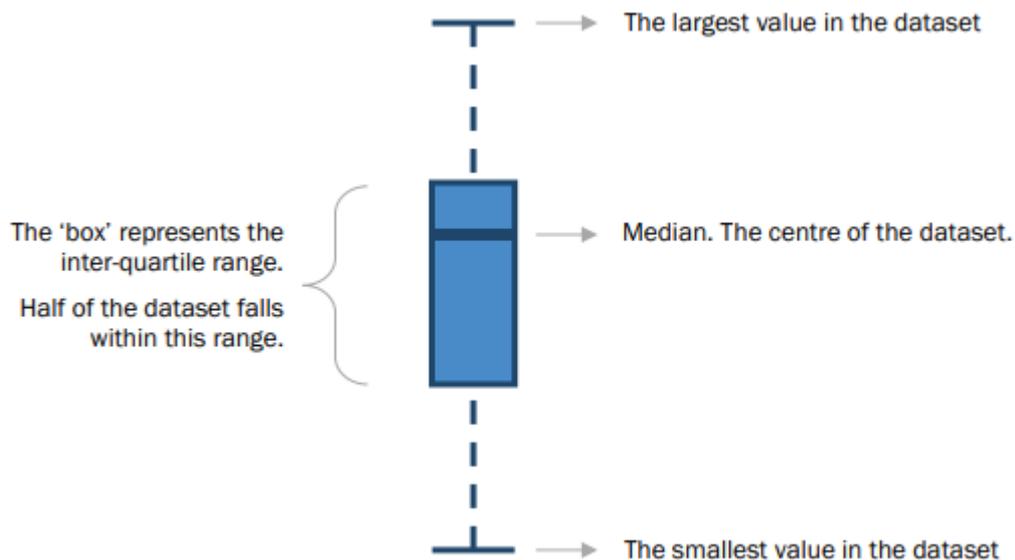
Interpreting a boxplot

The largest and smallest values are shown at the ends of each 'whisker' that attaches to either end of the boxplot.

The middle 50 per cent of all of the data points lie within the box.

A quarter of the data falls above and below the box.

The median is the line inside the middle of the box.



Interpreting a violin plot

A violin plot helps to visualise the full distribution of the dataset. Where a boxplot only shows certain key values (ie, the maximum, minimum, median and inter-quartile range), the violin plot provides a visual depiction of the entire dataset. The thinner parts of the chart represent fewer items, and the thicker parts of the chart represent a higher number of items.

Outliers are sometimes displayed separately to preserve scale → ● n=12

The largest value in the dataset →

The smallest value in the dataset →

The thickest part of the plot represents the most common value (also known as the 'mode')